ON THE ASSEMBLY HISTORY OF STELLAR COMPONENTS IN MASSIVE GALAXIES

JAEHYUN LEE AND SUKYOUNG K. YI

Research School of Astronomy and Astrophysics, Australia National University, Canberra 2611, Australia; yki@anu.edu.au

Received 2012 August 24; accepted 2013 February 4; published 2013 March 5

ABSTRACT

Matsuoka & Kawara showed that the number density of the most massive galaxies (log \( M/M_\odot = 11.5-12.0 \)) increases faster than that of the next massive group (log \( M/M_\odot = 11.0-11.5 \)) during \( 0 < z < 1 \). This appears to be in contradiction to the apparent “downsizing effect.” We attempt to understand the two observational findings in the context of the hierarchical merger paradigm using semi-analytic techniques. Our models closely reproduce the result of Matsuoka & Kawara. Downsizing can also be understood as larger galaxies have, on average, smaller assembly ages but larger stellar ages. Our fiducial models further reveal details of the history of the stellar mass growth of massive galaxies. The most massive galaxies (log \( M/M_\odot = 11.5-12.0 \) at \( z = 0 \)), which are mostly the brightest cluster galaxies, obtain roughly 70% of their stellar components via merger accretion. The role of merger accretion monotonically declines with galaxy mass: 40% for log \( M/M_\odot = 11.0-11.5 \) and 20% for log \( M/M_\odot = 10.5-11.0 \) at \( z = 0 \). The specific accreted stellar mass rates via galaxy mergers decline very slowly during the whole redshift range, while specific star formation rates sharply decrease with time. In the case of the most massive galaxies, merger accretion becomes the most important channel for the stellar mass growth at \( z \sim 2 \). On the other hand, in situ star formation is always the dominant channel in \( L_\ast \) galaxies.

Key words: galaxies: elliptical and lenticular, cD – galaxies: evolution – galaxies: formation – galaxies: stellar content

1. INTRODUCTION

Dynamical realizations based on the concordance \( \Lambda \)CDM cosmology (e.g., Spergel et al. 2007) have been remarkably successful at reproducing large-scale structures in the universe (e.g., Springel et al. 2006). In the common perception of this paradigm, large galaxies grow hierarchically through numerous galaxy mergers that follow mergers between dark halos.

Supporting this view, galaxies with disturbed features have frequently been witnessed (e.g., Arp 1966; Schweizer & Seitzer 1988). With the advent of deep and wide-field surveys, the arguments could be investigated in greater detail. Recent studies based on ultra-deep imaging data are particularly noteworthy. From the \( \mu_\ast = 28 \) mag arcsec\(^{-2} \) deep images, van Dokkum (2005) found that about 50% of red bulge-dominant galaxies in field environments show tidal debris. Merger galaxy fraction has been found to be almost as high in cluster environments (Sheen et al. 2012). Kaviraj et al. (2007) and Kaviraj (2010) claimed that residual star formation found in a large fraction (\( \sim 30\% \)) of local massive early-type galaxies is related to mergers or interactions. As observation techniques reach deeper and hidden nature of galaxies, “peculiar” is no longer a synonym of “rare.”

High-redshift surveys allow us to directly investigate of the role of galaxy mergers and interactions on galaxy evolution. Using the \( K \)-band Hubble diagram, Aragon-Salamanca et al. (1998) found that brightest cluster galaxies (BCGs) have increased their mass by a factor of two to four with no or negative luminosity evolution during \( 0 < z < 1 \). From the GOODS fields, Bundy et al. (2009) found that the pair fraction of massive (log \( M/M_\odot > 11.0 \)) red spheroidal galaxies is higher than that of less massive systems (log \( M/M_\odot \sim 10.0 \)). In addition, the pair fraction unaccompanied by star formation increases with time. They concluded that massive galaxies grow primarily through dry or minor mergers, at least at \( z \lesssim 1 \).

Almost simultaneously, from the UKIDSS and the SDSS II Supernova Survey, Matsuoka & Kawara (2010, hereafter MK10) found that during \( z < 1 \) the number density of the most massive galaxies (log \( M/M_\odot = 11.5–12.0 \)) increased more rapidly than that of the next massive group (log \( M/M_\odot = 11.0–11.5 \)). They also showed that more massive systems have a lower blue galaxy fraction than less massive systems, and that the fraction decreases with time. All these observations seem to imply that during \( z \lesssim 1 \) massive galaxies are mainly brought up via mergers.

Not all observations naively support the hierarchical merger picture. The tight color–magnitude relation found among early-type galaxies is more simply, albeit not exclusively, explained by monolithic formation scenarios (e.g., Bower et al. 1992; Kodama & Arimoto 1997). Furthermore, the “downsizing” effect, where larger galaxies appear to be older and are thus suspected to have formed earlier, seems to be inconsistent with the new paradigm (Cowie et al. 1996; Glazebrook et al. 2004; Cimatti et al. 2004). On the face of it, the inconsistency seems a counterevidence of the hierarchical picture; however, some studies have pointed out that downsizing can be understood by the hierarchical paradigm reasonably enough. Based on semi-analytic models, De Lucia et al. (2006) showed that the star formation rates of all the progenitors of more massive galaxies peak earlier and decrease faster than those of less massive galaxies. In fact, downsizing could be a natural result of the hierarchical clusterings of dark halos. Neistein et al. (2006) demonstrated that downsizing appears in the total (combined) mass evolution of all the progenitors of a dark matter halo. If the stellar mass of a galaxy correlates with the mass of its dark matter halo (Moster et al. 2010), downsizing may very well originate from the bottom-up assembly of dark matter halos. Using semi-analytic approaches, many studies have shown that the assembly and formation history of stellar components, especially in massive, red and luminous, or elliptical galaxies, can be different in the hierarchical universe (Kauffmann 1999; Baugh et al. 1996; De Lucia et al. 2006; De Lucia & Blaizot 2007; Almeida et al. 2008). More direct hydrodynamic simulations performed for
smaller volumes have yielded consistent results (Oser et al. 2010; Lackner et al. 2012).

In this study, we look further into the details of the assembly history of stellar components in massive galaxies. We use semi-analytic approaches because these are more effective for constructing models of a large volume of the universe. Motivated mainly by the empirical results of MK10, we investigate the factors that drive the growth of galaxy stellar mass as a function of time and mass.

2. MODEL

We have developed our own semi-analytic model for galaxy formation and evolution. In this section, we briefly introduce physical ingredients, with more focus on the prescriptions that play important roles in this study.

2.1. Dark Matter Halo Merger Trees and Galaxy Mergers

In the context of the two-step galaxy formation theory (White & Rees 1978), the first step involves the construction of dark halo merger trees. We ran dark matter N-body volume simulations using the GADGET-2 code (Springel 2005). The simulation was performed using the standard ΛCDM cosmology parameters derived from the Wilkinson Microwave Anisotropy Probe seven-year observations, Ω_m = 0.266, Ω_L = 0.734, σ_8 = 0.801, Ω_b = 0.0449, and H_0 = 71 km s^{-1} Mpc^{-1} (Jarosik et al. 2011). The periodic cube size of the simulation was 70 h^{-1} Mpc on the side with 512^3 collisionless particles, mass of a particle was 1.9 × 10^8 h^{-1} M_{⊙}, and mass resolution of a halo was ~10^{11} h^{-1} M_{⊙}. We identify halo structures in the simulation box, using a halo finding code developed by Tweed et al. (2009), which is based on the AdaptaHOP (Aubert et al. 2004). Halo merger trees were generated by backtracking the infall histories of subhalos with increasing redshift. We calculate galaxy merger timescales, using the positional information of subhalos extracted from our N-body simulation. In this study, we assume that the mass distribution of dark matter halos basically follows the Navarro–Frank–White (NFW) profile (Navarro et al. 1996).

Subhalos can disappear before arriving at central regions if they are heavily embedded in their host halo density profiles or their mass decreases below the resolution limit of our numerical simulation. We take these numerical artifacts into consideration. Thus, once a halo enters into a more massive halo, we additionally calculate its merger timescale, t_{merge}, using the following fitting formula suggested by Jiang et al. (2008):

\[
t_{\text{merge}}(\text{Gyr}) = \frac{0.94 + 0.60 M_{\text{host}}}{2C} \frac{1}{M_{\text{sat}} \ln[1 + (M_{\text{host}}/M_{\text{sat}})]} R_{\text{vir}},
\]

where \( \epsilon \) is the circularity of the orbit of a satellite halo, \( C \) is a constant, approximately equal to 0.43, \( M_{\text{host}} \) is the mass of a host halo, \( M_{\text{sat}} \) is the mass of a satellite halo, \( R_{\text{vir}} \) is the virial radius of a host halo, and \( V_{\epsilon} \) is the circular velocity of a host halo at \( R_{\text{vir}} \). If a subhalo disappears within 0.1 \( R_{\text{vir}} \) of its host halo, a galaxy in the subhalo is regarded as having merged with its central galaxy. On the other hand, if it disappears outside 0.1 \( R_{\text{vir}} \), we assume that the galaxy in the subhalo merges with its central galaxy at \( t_{\text{merge}} \) given by Equation (1) after the subhalo becomes a satellite of its host halo. In that case, we consider dynamical friction to analytically compute the positions and velocities of subhalos. We adopt the dynamical friction prescription introduced by Binney & Tremaine (2008):

\[
dv{\Delta t_{\text{dyne}}} = -\frac{G M_{\text{sat}}(t)}{r^2} \ln \Lambda \left( \frac{V_c}{v} \right)^2 \times \left\{ \text{erf} \left( \frac{v}{V_c} \right) - \frac{\sqrt{\pi}}{2} \left( \frac{v}{V_c} \right) \exp \left[ -\left( \frac{v}{V_c} \right)^2 \right] \right\} e^{-v},
\]

where \( M_{\text{sat}} \) is the mass of a subhalo, which is initially defined as the mass at a previous time step after which the subhalo cannot be resolved in the N-body simulation anymore, \( r \) is the distance between the subhalo and the center of its host halo, \( \ln \Lambda \) is a Coulomb logarithm with \( \Lambda = 1 + M_{\text{host}}/M_{\text{sat}} \) adopted by Springel et al. (2001), \( V_c \) is the circular velocity of the host halo at the virial radius, and \( v \) is the orbital velocity of the subhalo.

During the orbital motion of a satellite halo in its host halo, the dark matter of the satellite halo is stripped due to dynamical friction. If the satellite is resolved in an N-body simulation, dark matter of the satellite would be naturally stripped. On the other hand, if the halo is not resolved in the simulation but considered to orbit around its host halo, stripping due to dynamical friction should be computed analytically. We evaluate the amount of dark matter stripped by dynamical friction by adopting the concept of the sphere of influence (\( r_{\text{soi}} \)) within which dark matter particles are bound to the satellite halos, using the following formula (Battin 1987):

\[
r_{\text{soi}} \sim r \left[ \left( \frac{M_{\text{sat,tot}}}{M_{\text{host}}(<r)} \right)^{-0.4} (1 + 3 \cos^2 \theta)^{0.1} + 0.4 \cos \theta \left( \frac{1 + 6 \cos^2 \theta}{1 + 3 \cos^2 \theta} \right) \right]^{-1},
\]

where \( r \) is the distance between the centers of the satellite and its host halos, \( M_{\text{sat,tot}} \) is the total (baryon+dark matter) mass of the satellite halo, \( M_{\text{host}}(<r) \) is the total mass of the central halo within \( r \), and \( \theta \) is the angle between the line connecting the particle to the center of the satellite halo and the line connecting the centers of the satellite and the host halos. During a time step, \( \Delta t \), we assume that a satellite halo loses \( \delta M_{\text{sat}} = M_{\text{sat}}(r > r_{\text{soi}}) \delta t / \Delta t_{\text{dyne}} \), where \( \Delta t_{\text{dyne}} \) is the dynamical timescale of the satellite halo, and \( M_{\text{sat}}(r > r_{\text{soi}}) \) is the mass of dark matter outside \( r_{\text{soi}} \) of the satellite halo.

We assume that stellar components in satellite galaxies merging with their hosts constitute the bulge component of the host. If the mass ratio of baryonic mass (\( m_{\text{cold}} + m_\ast \)) between merging galaxies, \( m_{\text{secondary}} / m_{\text{primary}} \), is greater than 0.25, then it is assumed that all the stellar components of the host galaxy quickly become bulge components of the remnant, as well.

Empirical studies have shown that intracluster light originates from the extended diffuse stellar components of the brightest galaxies in groups or clusters (e.g., Feldmeier et al. 2002; Gonzalez et al. 2005; Zibetti et al. 2005). They suggested that diffuse stellar components are the stars scattered from satellite galaxies during tidal stripping or mergers into central galaxies. It has been suggested that 10%–40% of stellar components in satellite galaxies turn into diffuse stellar components in each galaxy merger (Murante et al. 2004; Monaco et al. 2006).
We adopt the value of 40% in this study because it resulted in the best reproduction of empirical data. The amount of stellar mass that a central galaxy acquires via merger is \((1 - f_{\text{scatte}})M_{*\text{sat}}\), where \(f_{\text{scatte}}\) is the fraction of scattered stellar components and \(M_{*\text{sat}}\) is the stellar mass of a satellite galaxy.

### 2.2. Gas Cooling, Star Formation, and Recycling

We assume that the baryonic fraction in accreted dark matter follows the global baryonic fraction, \(\Omega_b/\Omega_m\). Baryons are accreted onto dark halos and shock-heated to become hot gas components.

Gas accretion onto a galactic disk plane via atomic cooling of hot gas is calculated based on the model proposed by White & Frenk (1991). The cooling timescale at distance \(r\) from the center of a halo is estimated by the following formula:

\[
\tau_{\text{cool}}(r) = \frac{3}{2} \frac{\rho_s(r)kT}{\mu n_e n_i \Lambda(Z, T)},
\]

where \(\rho_s\) is the gas density within radius \(r\) of a dark matter halo, \(T\) is the temperature of gas, assumed to be the virial temperature in the model, \(Z\) is metallicity, \(\mu\) is the mean molecular weight of gas with \(Z\) and \(T\), and \(m_p\) is the mass of a proton, \(n_e\) is the number density of electrons, \(n_i\) is the number density of ions, and \(\Lambda\) is the cooling function with \(Z\) and \(T\). The values of \(\mu\), \(n_e\), \(n_i\), and \(\Lambda\) are determined by referring to Sutherland & Dopita (1993). We do not include self-consistent chemical evolution in our calculation. Instead, we take the metallicity of hot gas components in halos as a constant, \(0.3Z_\odot\), which is comparable to the metallicity of observed clusters with various masses (Arnaud et al. 1992). Because the cooling function, \(\Lambda\), is sensitive to metallicity, our results should not be taken too literally. However, relative analysis, a main tool of this study, is affected little by the details in the treatment of chemical evolution.

It is assumed that the gas density follows a singular isothermal profile truncated at \(R_{\text{vir}}\): \(\rho_s(r) = m_{\text{cool}}/(4\pi R_{\text{vir}}^2 r^2)\). Substituting the formula for \(\rho_s(r)\) in Equation (4) and adopting the dynamical timescale of a halo, \(t_{\text{dyn}}, \tau_{\text{cool}}\), one can derive the cooling radius, \(r_{\text{cool}}\), within which hot gas can cool within \(\tau_{\text{cool}}\): For the case of \(r_{\text{cool}} > R_{\text{vir}}\), the cooling rate is rather restrained by the free-fall rate than the cooling rate, so that \(m_{\text{cool}} = m_{\text{bol}}/(2\tau_{\text{cool}})\). In contrast, if \(r_{\text{cool}} < R_{\text{vir}}\), \(m_{\text{cool}} = m_{\text{cold}}/(2R_{\text{vir}}/c_{\text{cool}})\).

In our model, stars can be formed through a quiescent mode, in which cold gas turns into disk stellar components via gas contraction on a disk, or a burst mode, which is induced by galaxy mergers. Star formation rate in the quiescent mode are delineated by a simple law proposed by Kauffmann et al. (1993) as follows:

\[
m_* = \alpha \frac{m_{\text{cold}}}{t_{\text{dyn}, \text{gal}}},
\]

where \(\alpha\) is the empirically determined star formation efficiency, \(m_{\text{cold}}\) is the amount of cold gas, and \(t_{\text{dyn}, \text{gal}}\) is the dynamical timescale of cold gas disk assumed to be \(0.1t_{\text{dyn}}\).

Observations (e.g., Borne et al. 2000; Woods & Geller 2007) and hydrodynamic simulations of galaxy mergers (e.g., Cox et al. 2008, hereafter C08) have shown that galaxy mergers can give rise to rapid star formation. We follow the conventional treatment: stars formed in the quiescent mode belong to a galactic disk, while stars born in the burst mode become bulge components. We adopt the prescription for merger-induced starbursts described in Somerville et al. (2008, hereafter S08), which formulates the prescription based on C08. S08 defines the burst efficiency, \(\epsilon_{\text{burst}}\), to parameterize the fraction of the cold gas reservoir involved in a merger induced starburst as follows:

\[
\epsilon_{\text{burst}} = \epsilon_{\text{burst}, 0} \mu^{\gamma_{\text{burst}}},
\]

where \(\mu\) is the mass ratio between a host galaxy and its merger counterpart, \(\gamma_{\text{burst}}\) is the bulge-to-total mass ratio \((B/T)\) of the host galaxy, and \(\epsilon_{\text{burst}, 0}\) is the burst efficiency fitted by the following formula:

\[
\epsilon_{\text{burst}, 0} = 0.60(V_{\text{vir}}/(\text{km s}^{-1}))^{0.07}
\times (1 + q_{\text{EOS}})^{-0.17}(1 + f_{T})^{0.07}(1 + z)^{0.04},
\]

where \(V_{\text{vir}}\) is virial velocity, \(q_{\text{EOS}}\) is the effective equation of state of gas, \(f_T \equiv m_{\text{cold}}/(m_{\text{cold}} + m_*)\) is the fraction of cold gas, and \(z\) is the redshift when the disks of progenitor galaxies are constructed. \(q_{\text{EOS}}\) was suggested to parameterize the multiphase nature of interstellar medium (ISM): \(q_{\text{EOS}} = 0\) indicates an isothermal state, and \(q_{\text{EOS}} = 1\) represents the fully pressurized multiphase ISM. In this study, we adopt \(q_{\text{EOS}} = 1\) at which gas is dynamically stable, so that starbursts are more suppressed than the case of \(q_{\text{EOS}} = 0\). The redshift dependency of \(\epsilon_{\text{burst}, 0}\) is very weak, and thus we assume \((1 + z)^{0.04} \sim 1\). The burst timescale, \(\tau_{\text{burst}}\), is also formulated by S08 as follows:

\[
\tau_{\text{burst}} = 191\text{Gyr}[V_{\text{vir}}/(\text{km s}^{-1})]^{-1.88}
\times (1 + q_{\text{EOS}})^{2.58}(1 + f_{T})^{-0.74}(1 + z)^{-0.16}.
\]

The mass ratio, \(\mu\), is the ratio of the total mass of central regions \((m_{\text{DM, core}} + m_* + m_{\text{cold}})\) of a host galaxy to that of its merger counterpart. Following S08, we calculate the core mass of a dark matter halo, \(m_{\text{DM, core}} = m_{\text{DM}}/r < R_{\text{vir}}\), where \(r_s \equiv R_{\text{vir}}/c_{\text{NFW}}\).

The concentration index of the NFW profile, \(c_{\text{NFW}}\), is derived based on the fitting function suggested by Macciò et al. (2007).

The parameter \(\gamma_{\text{burst}}\) is determined by the \(B/T\) of a host galaxy as follows:

\[
\gamma_{\text{burst}} = \begin{cases} 0.61 & \text{if } B/T \leq 0.085 \\ 0.74 & \text{if } 0.085 < B/T \leq 0.25 \\ 1.02 & \text{if } B/T > 0.25. \end{cases}
\]

C08 showed that the burst efficiency of a host galaxy with a high \(B/T\) is lower than that of a galaxy with a lower \(B/T\), because more massive bulges stabilize the galaxies and reduce the burst efficiency more effectively. Because C08 demonstrated that mergers with mass ratios below 1:10 are not associated with starbursts, we assume \(\epsilon_{\text{burst}} = 0\) if \(\mu < 0.1\). With the ingredients, the amount of stars born in burst modes is calculated as \(m_{\text{burst}} = \epsilon_{\text{burst}}m_{\text{cold}}\). We assume that \(m_{\text{burst}}\) turns into stars for \(\tau_{\text{burst}}\) in a uniform rate, \(m_{\text{burst}} = m_{\text{burst}}/\tau_{\text{burst}}\). While merger-induced starbursts occur in a galaxy, the quiescent mode also still goes on in our models.

In our model, the recycling of stellar mass loss is considered in great detail. We compute the mass loss of every single population at each epoch after a new stellar population is born. The mass loss of a single population is calculated as follows. We adopt the Scalo initial mass function (Scalo 1986). The lifetime of a star with mass \(M\), \(\tau_M\), is computed by a broken-power law (Ferreras & Silk 2000), which is obtained from the data of Tinsley (1980) and Schaller et al. (1992):

\[
\tau_M(\text{Gyr}) = \begin{cases} 9.694(M/M_\odot)^{-2.76} & M < 10M_\odot \\ 0.095(M/M_\odot)^{-0.76} & M \geq 10M_\odot. \end{cases}
\]
At the end of its lifetime (after \( \tau_m \) elapses), a star returns most of its mass into space, leaving small remnants such as a white dwarf, a neutron star, or a black hole. The remnant mass of a star with mass \( M \), \( \omega_M \), is suggested by Ferreras & Silk (2000) as follows:

\[
\frac{\omega_M}{M} = \begin{cases} 
0.1(M/M_\odot) + 0.45 & M < 10 M_\odot \\
1.5 & 10 M_\odot \leq M < 25 M_\odot \\
0.61(M/M_\odot) - 13.75 & M \geq 25 M_\odot.
\end{cases}
\]  

In this study, we simply assume that half of the mass loss returns to cold gas components and the rest becomes hot gas.

### 2.3. Environmental Effect

The Chandra X-Ray Observatory revealed that massive satellite galaxies in nearby clusters have hot gas components (Sun et al. 2007; Jeltema et al. 2007), in contradiction to expectations based on an instantaneous hot gas stripping scenario for satellite galaxies in cluster environments (e.g., Kauffmann et al. 1999; Somerville et al. 2008). The old assumption predicted a higher fraction of massive satellites in large halos than observed, known as the satellite overquenching problem (Kimm et al. 2009). Kimm et al. (2011) showed that a gradual, rather than instant, and more realistic stripping of the hot gas reservoir relieves the above-mentioned problem to some degree. Therefore, we implemented the gradual hot gas stripping of satellite galaxies in our model by considering tidal stripping (see Kimm et al. 2011) and ram pressure stripping that uses the prescriptions of McCarthy et al. (2008), which had been modified for semi-analytic models by Font et al. (2008). However, we are still missing a realistic prescription on cold gas stripping. Empirical evidence for cold gas stripping is clear (Vollmer et al. 2008; Chung et al. 2008a, 2008b), and a theoretical study using a semi-analytic approach shows its effect in galaxy evolution (Tecce et al. 2010). Thus, it should be considered in our model in due course.

### 2.4. Feedback Processes

Feedback mechanisms have been introduced into galaxy formation theory to reconcile the discrepancy between galaxy and dark matter halo mass functions. Hydrogen atoms, neutralized at the recombination, may be reionized at later epochs \((z < 10)\) due to background high-energy photons. This mechanism may suppress the growth of small galaxies (Gnedin 2000; Somerville 2002; Benson et al. 2002a, 2002b). Supernova feedback is thought to be effective at disturbing the growth of small galaxies by ejecting cold gas (White & Rees 1978; Dekel & Silk 1986), and active galactic nucleus (AGN) feedback is considered more effective in massive galaxies with a large black hole (Silk & Rees 1998; Schawinski et al. 2006, 2007).

We utilize the reionization prescription of Benson et al. (2002b). The prescription allows an inflow of baryons into a dark halo via accretion of dark matter when \( V_{\text{vir}} > V_{\text{reionization}} \) throughout the age of the universe, where \( V_{\text{reionization}} \) is the suppression velocity of reionization. If a halo has a lower value of \( V_{\text{vir}} \) than the criterion, the inflow of baryons is allowed only at \( z > z_{\text{reionization}} \), where \( z_{\text{reionization}} \) is the suppression redshift of reionization. In this study, we adopt \( V_{\text{reionization}} = 30 \text{ km s}^{-1} \) and \( z_{\text{reionization}} = 8 \) following Benson et al. (2002b).

We follow the prescriptions of S08 for supernova feedback. These prescriptions take into account not only the amount of reheated gas but also the fraction of reheated gas blown away from halos. The reheating rate of cold gas due to supernova feedback is formulated as follows:

\[
\dot{m}_{\text{r}} = \epsilon_{\text{SN}} \left( \frac{200 \text{ km s}^{-1}}{V_{\text{disk}}} \right) \alpha_{\text{SN}} \dot{m}_*, \tag{12}
\]

where \( \epsilon_{\text{SN}} \) and \( \alpha_{\text{SN}} \) are free parameters, \( V_{\text{disk}} \) is the rotational velocity of a disk, and \( \dot{m}_* \) is the star formation rate. S08 assumes that the rotational velocity of a disk is the same as the maximum rotational velocity of the DM halo. We calculate the fraction of the reheated gas that has enough kinetic energy to escape from the halo as

\[
f_{\text{ej}} = \left[ 1.0 + \left( \frac{V_{\text{vir}}}{V_{\text{ej}}} \right)^2 \right]^{-1}, \tag{13}
\]

where \( \alpha_{\text{ej}} = 6 \) and \( V_{\text{ej}} \sim 100-150 \text{ km s}^{-1} \). In the case of a satellite halo, a fraction of reheated gas by supernova feedback, \( f_{\text{ej}} \dot{m}_{\text{r}} \), is ejected from the satellite and added to the hot gas reservoir of its host halo. On the other hand, it is assumed that the ejected gas from main halos is diffused through intercluster medium.

We take the quasar-mode and radio-mode AGN feedback into account in our model. It has been suggested that the quasar mode is induced by an inflow of cold gas into the central supermassive black hole (SMBH) of the central galaxy during major mergers (Kauffmann & Haehnelt 2000). The increasing mass of the SMBH via the accretion of cold gas can be expressed as

\[
\Delta m_{\text{BH, Q}} = \frac{f_{\text{BH}} m_{\text{cold}}}{1 + (280 \text{ km s}^{-1}/V_{\text{vir}})^2}, \tag{14}
\]

where \( f_{\text{BH}} \) is the efficiency of gas accretion. In this study, we take the modified parameter proposed by Croton et al. (2006): \( f_{\text{BH}} = f_{\text{BH}}(M_{\text{BH}}/M_{\text{host}}) \), where \( f_{\text{BH}} \) is the original form of the parameter introduced by Kauffmann & Haehnelt (2000), \( M_{\text{host}} \) is the mass of a host galaxy, and \( M_{\text{BH}} \) is the mass of the host galaxy’s merger counterpart. In this study, we assume that the lifetime of the quasar mode, \( t_{\text{QSO}} \), is 0.2 Gyr, as Martini & Weinberg (2001) and Martini & Schneider (2003) suggested \( t_{\text{QSO}} < 0.3 \text{ Gyr} \). Thus, we have \( m_{\text{BH, Q}} = \Delta m_{\text{BH, Q}} / t_{\text{QSO}} \). It is generally thought that the quasar mode is caused by a high accretion rate of cold gas, resulting in a rapid growth of an SMBH.

The radio-mode feedback releases low-Eddington-ratio energy through the accretion of hot gas distributed throughout the halo. Although the energy released from the radio-mode AGN is far less than that from the quasar mode, it is regarded that the radio mode supplies enough energy to the surrounding medium to interrupt gas cooling or to blow away (some of) the cold gas. We implement radio-mode feedback into our model following the prescription of Croton et al. (2006):

\[
\dot{m}_{\text{BH, R}} = \kappa_{\text{AGN}} \left( \frac{m_{\text{BH}}}{10^8 M_\odot} \right) \left( \frac{f_{\text{hot}}}{0.1} \right) \left( \frac{V_{\text{vir}}}{200 \text{ km s}^{-1}} \right)^3, \tag{15}
\]

where \( \kappa_{\text{AGN}} \) is a free parameter with units of \( M_\odot \text{ yr}^{-1} \), \( m_{\text{BH}} \) is the black hole mass, \( f_{\text{hot}} \) is the fraction of hot gas with respect to the total halo mass, and \( V_{\text{vir}} \) is the virial velocity.

We assume that the amount of energy generated by the accretion of gas into the SMBH is given as follows:

\[
\dot{L}_{\text{BH}} = \eta \dot{m}_{\text{BH}} c^2 = \eta (\dot{m}_{\text{BH, Q}} + \dot{m}_{\text{BH, R}}) c^2, \tag{16}
\]
where $\eta = 0.1$ is the standard efficiency of the conversion of rest mass to radiation, and $c$ is the speed of light. The reduced cooling rate of gas is computed by

$$m_{\text{cool}}' = m_{\text{cool}} - \frac{L_{\text{BH}}}{0.5V_{\text{vir}}},$$

where the minimum of $m_{\text{cool}}'$ is set to be zero.

### 2.5. Model Calibrations

Our models are based on the conventional techniques and ingredients used in up-to-date semi-analytic models; hence, the output is not particularly noteworthy compared to other successful models. Our models roughly match the global star formation history, galaxy mass functions, black hole mass versus bulge mass relation, etc. Figures 1 and 2 display comparisons of the cosmic star formation history and the galaxy stellar mass functions in the local universe from empirical data (Panter et al. 2007) and our fiducial model. While there still is a large room for improvement, we decide to focus on the mass growth histories of massive galaxies.

### 3. EVOLUTION OF GALAXY NUMBER DENSITY

MK10 presented a rapid growth of massive galaxies since $z = 1$, using the United Kingdom Infrared Telescope (UKIRT) Infrared Deep Sky Survey (UKIDSS) and the Sloan Digital Sky Survey (SDSS) II Supernova Survey. Figure 3 shows the number density evolution of the most massive ($\log M/M_\odot = 11.5-12.0$) and the next massive ($\log M/M_\odot = 11.0-11.5$) galaxies in the empirical data derived by MK10 and from our fiducial model at each redshift. The empirical data clearly show that the number of the most massive galaxies rapidly increases between $z = 1$ and 0, while the next massive group experiences a milder evolution. The reproduction of the data by our fiducial model looks reasonably good. We also present the number density evolution of the third massive group ($\log M/M_\odot = 10.5-11.0$), which makes our models “super-$L_*$ galaxies.”

We define “relative number density growth rate,” $\Gamma$, as the ratio of the speeds in number density evolution of the two most massive groups of galaxies as a function of observational limit in redshift, as follows:

$$\Gamma = \frac{n_{\text{most}}(z = 0)/n_{\text{most}}(z)}{n_{\text{next}}(z = 0)/n_{\text{next}}(z)},$$

where $n_{\text{most}}(z = 0)$ and $n_{\text{next}}(z = 0)$ are the number densities of the most massive and next massive groups at $z = 0$, and $n_{\text{most}}(z)$ and $n_{\text{next}}(z)$ are the number densities of the two groups at a redshift, respectively. For example, MK10 compared the number density evolution of the two mass groups between redshift 0 and 1, in which case the observational limit is 1 and the relative number density growth rate becomes 3. We present
In the hierarchical paradigm, a galaxy can have more than one progenitor. Progenitors of a galaxy can be divided into “direct” and “collateral” progenitors. A direct progenitor is the galaxy in the largest halo when a merger between halos takes place. While there can be numerous progenitors, there is only one direct progenitor at each epoch. Collateral progenitors are all the other galaxies that contribute to the final galaxy. In this concept, to build the evolutionary history of a galaxy, one should consider not only direct progenitors, but also merger counterparts or collateral progenitors. Figure 5 shows the average mass evolution of the direct progenitors (solid lines) and all (direct and collateral combined) the progenitors (dotted lines) of Rank 1, 2, and 3 galaxies. The mean stellar masses of the three groups at $z = 0$ are $4.97 \times 10^{11} M_\odot$, $1.55 \times 10^{11} M_\odot$, and $5.25 \times 10^{10} M_\odot$, respectively. Most of the Rank 1 galaxies in our volume are brightest cluster galaxies.

As all the progenitors merge with each other, the dotted lines and the solid lines finally meet at $z = 0$. Stellar mass loss and the scattering of stellar components in satellite galaxies into diffuse stellar components, which takes place when mergers occur, lead to a gradual decrease in the total mass during the evolution. For example, one can see a slight decline in the total stellar mass (red dotted line at the top) after $z \sim 0.5$. This effect is not clearly visible in Rank 2 and 3 galaxies in which star formation is more extended and mergers are less frequent than in Rank 1 galaxies. It is useful to have a definition of the formation redshift, $z_f$. We define it as the redshift at which half of the stellar mass at $z = 0$ has been assembled. In the case of Rank 1 galaxies, half of the final mass is achieved at $z_f, D \sim 0.7$ (denoted as $D_1$ in Figure 5) in direct progenitors and at $z_f, A \sim 3.0 (A_1)$ when all progenitors are combined. Our models suggest $(z_f, D, z_f, A) = (0.9, 2.1)$ for Rank 2, and $(z_f, D, z_f, A) = (1.1, 1.6)$ for Rank 3. Models exhibit a monotonic mass dependence of $z_f, D$ and $z_f, A$ in the sense that, with mass, $z_f, D$ decreases while $z_f, A$ increases. In other words, the mass of the direct progenitors of a more massive group grows more slowly, while its total mass of all the progenitors is assembled earlier than that of a less massive group. The evolutionary histories of direct progenitors are opposite to the pattern of cosmic downsizing. As Neistein et al. (2006) and Oser et al. (2010) pointed out, however, if the growth histories of collateral progenitors of galaxies are also considered, downsizing would be a natural outcome of the hierarchical concept of galaxy formation.

The difference in the growth history between the three groups can be understood in depth through Figure 6, which presents the evolutionary histories of star formation rates (SFRs). The mean SFRs contain the star formation histories of both direct and collateral progenitors. We show the best fitting log-normal function to the three SFR curves. The star formation rates of
more massive galaxies peak earlier, as marked by S1, S2, and S3 in the figure, and decreases faster than those of less massive groups. The same features were noted in an earlier study by De Lucia et al. (2006).

Figure 6 also shows the redshift at which half of the total cumulative star formation has occurred in the three groups of galaxies: (C1, C2, C3) = (2.0, 1.5, 1.4) in \( z \). The general trend shown by these three values agrees with that of \( z_{1, A} \) discussed above. In a sense, it is these values, rather than \( z_{1, A} \), that are closer to the general definition of formation redshift. Again, the more massive galaxies are, the earlier they form their stars, consistent with downsizing. In conclusion, the observational finding of downsizing is a result of the hierarchical galaxy formation process, where more massive galaxies have larger stellar ages and smaller assembly ages (see also, e.g., De Lucia & Blaizot 2007; Kaviraj et al. 2009). Both the larger stellar ages and the smaller assembly ages can be understood as a result of large-scale effect; that is, in a deeper potential well, progenitor galaxies and their stars form earlier, and many more galaxies participate in galaxy mergers for a long period of time.

### 4.2. Origin of Stellar Components

In this section, we investigate how stellar components are assembled into massive galaxies. Stellar components in a galaxy originate either from in situ star formation or from “merger accretion.” Two modes of star formation are considered: “quiescent” mode and merger-induced “starburst.” Stars in a galaxy can therefore have four different origins: (1) in situ quiescent star formation, (2) in situ starburst, (3) merger accretion of stars formed in quiescent mode, and (4) merger accretion of stars formed in burst mode.

Figure 7 shows the decomposition of the four channels as a function of time for the three different mass groups. The top and bottom rows show the absolute stellar mass evolution and relative mass fraction evolution, respectively.

Several remarkable features are visible, more easily in the bottom rows. First, in situ quiescent star formation is an important channel for the stellar components in massive galaxies. Its fractional contribution is (30, 60, 80)% in Rank (1, 2, 3) galaxies, respectively. In situ quiescent star formation takes place on a galactic disk, and thus, one may wonder why its contribution is so large in these massive, and probably bulge-dominant galaxies. This is because massive bulge-dominated galaxies at \( z = 0 \) have many late-type progenitors, and that is more pronounced in less massive galaxies (Rank 3) than in Rank 1 galaxies. This is a reflection of the progenitor bias discussed earlier (cf. Franx & van Dokkum 2001; Guo & White 2008; Parry et al. 2009; Parry et al. 2009).

Second, the stars formed in burst mode are an extreme minority (~2%) in these massive galaxies. This result is somewhat unexpected. In the hierarchical universe, larger galaxies are generally a product of numerous mergers between galaxies. Mergers, especially between similar-mass galaxies (major mergers), cause a starburst, and so a naive expectation is that larger galaxies would have a large fraction of stars formed in burst mode. However, it is not that simple. Starbursts are usually a result of major mergers, and major mergers are extremely rare once galaxies become massive, that is, at low redshifts. Major mergers between massive galaxies sometimes occur at low redshifts, but they are most often dry, between early-type galaxies without much cold gas (cf. van Dokkum 2005; Bundy et al. 2007, 2009; Parry et al. 2009). Thus, it has been suggested that the fraction of stars formed in merger-driven burst mode may be less than a few per cent (e.g., Aragon-Salamanca et al. 1998; Rodighiero et al. 2011).

Third, merger accretion is the dominant channel in the most massive galaxies. Roughly 70% of the stellar components in the most massive group (bottom left panel) form outside direct progenitors and get accreted via mergers. This fraction is smaller for Rank 2 (40%) and Rank 3 (20%) galaxies, but is still substantial. It is interesting to note that much of the mass difference between ranks is attributed to the difference in the amount of merger accretion. For example, roughly 80% of the mass difference between Ranks 1 and 2 comes from the differences in merger accretion. The value is ~60% between Ranks 2 and 3. Massive galaxies are so mainly because they have a substantial amount of in situ quiescent star formation, but the most massive galaxies are so because they have acquired a large amount of mass through merger accretion. The result is supported by previous studies. Aragon-Salamanca et al. (1998) confirmed that BCGs have experienced no or negative evolution in luminosity during \( 0 < z < 1 \) from the \( K \)-band Hubble diagram for a sample of BCGs while they have increased their mass by a factor of two to four, depending on cosmological parameters. Thus, it has been understood that merger and accretion may be the most plausible explanation for the evolution. Using a semi-analytic model, De Lucia & Blaizot (2007) showed that most stellar components in model BCGs are formed in the very early age of the universe (80% at \( z \sim 3 \)) in small progenitors and accreted onto BCGs far later (50% after \( z \sim 0.5 \)) via mergers. Oser et al. (2010) presented similar results, using numerical simulations. About 80% of stellar components in simulated massive galaxies (\( M_\star > 2.4 \times 10^{11} M_\odot \) at \( z = 0 \)) are formed outside in the early age (\( z > 3 \)) and brought into the massive galaxies via mergers and accretion. The massive galaxies double their mass after \( z \sim 1 \). They revealed that the fraction gets smaller in less massive galaxies. Parry et al. (2009) also found a similar trend in their investigation on two separate semi-analytic models. They demonstrated that the contribution of mergers to the bulge growth exceeds that of disk instability at \( M_\star > 10^{11.5} M_\odot \). Considering the fact that such massive galaxies (\( M_\star > 10^{11} M_\odot \)) are likely early-type (e.g., Bell et al.
Figure 7. Average mass evolution of stellar components in direct progenitors (upper) and the fraction of each component divided by the mean total stellar mass of direct progenitors at each redshift (bottom). The left, middle, and right panels show the mean evolutionary histories of the direct progenitors of Rank 1, 2, and 3 galaxies, respectively. Each color code represents each stellar component as follows: (1) sky blue: merger accretion of stars formed in quiescent mode, (2) orange: in situ quiescent star formation, (3) blue: in situ starburst, and (4) red: merger accretion of stars formed in burst mode.

Figure 8. Frequency of mergers (per Gyr per galaxy) with a mass ratio $m_2/m_1 \gtrsim 0.1$ experienced by direct progenitors. The solid, dashed, and dotted lines indicate the merger rates of Rank 1, 2, and 3 galaxies, respectively.

2003), it implies that mergers play a more important role in the growth of massive galaxies. It should, however, be noted that disk instability which is more effective to the smaller late-type galaxy evolution may play a role in such progenitor galaxies of present-day massive early types.

Figure 8 shows the merger rate evolution for baryonic mass ratios greater than or equal to 1:10. Although merger rates show stochastic effects, there is a clear decreasing tendency with time, whereas the star formation rates of the three groups drop more sharply, as shown in Figure 6. In general, more massive galaxies are likely to be involved in galaxy mergers more frequently, so that more massive galaxies have many more stellar components born outside and accreted via mergers, as discussed above. During the whole calculation, Rank 1 galaxies undergo about 9.0 mergers with a mass ratio greater than or equal to 1:10 while those in Ranks 2 and 3 experience 4.0 and 1.0 mergers, respectively. During $0 < z < 1$, Rank 1, 2, and 3 galaxies experience 3.5, 1.8, or 0.6 mergers for the same mass ratio criterion. This explains the higher contribution of merger accretion in more massive galaxies, as illustrated in Figure 7.

4.3. Specific Star Formation Rates and Merger Accretion Rates

We found in the previous section that in situ star formation and merger accretion were the two most significant channels for the stellar mass growth of massive galaxies. In this section, we scrutinize their time evolution in greater detail. Specific star formation rates (SSFRs) are normalized growth rates of star formation histories. Likewise, we hereby define the specific stellar accretion rate (SSAR) to evaluate a normalized growth rate via mergers as follows:

$$SSAR = \frac{\Delta M}{M(t)\Delta t}.$$  

where $M(t)$ is the mass of a galaxy at an epoch, and $\Delta M$ is an increment of mass by mergers during a time step $\Delta t$. Because we allow diffuse stellar components due to galaxy mergers, mass increment can be expressed as $\Delta M = (1 - f_{\text{scatter}})M_{*,\text{sat}}$, as described in Section 2.1.

Figure 9 shows the evolution of the SSFRs (upper) and of the SSARS (bottom) of direct progenitors. In general, more massive galaxies have lower SSFRs, as observations have shown (e.g., Salim et al. 2007; Schiminovich et al. 2007), while they have higher SSARs than less massive galaxies. Star formation rates are decreased by the depression of gas cooling rates due to an increase in the cooling timescale via the growth of halos,
supernova feedback (White & Rees 1978; Dekel & Silk 1986; White & Frenk 1991), and/or AGN feedback (Silk & Rees 1998). Furthermore, the cold gas reservoir of a galaxy could be reduced by feedback processes. Besides, if a galaxy orbits around a more massive galaxy, it becomes redder as it loses its hot gas, which is a source of cold gas, and its cold gas reservoir by tidal and ram pressure stripping (Gunn & Gott 1972; Abadi et al. 1999; Quilis et al. 2000; Chung et al. 2007; Tonnesen & Bryan 2009; Yagi et al. 2010). On the other hand, accretion of stellar components via mergers is determined by gravitational interactions between host and subhalos alone; hence, the accretion rate could remain relatively steady as halos continue to fall into other more massive halos over time.

Figure 10 displays a comparison of the SSFRs and the SSARs of direct progenitors. As illustrated in Figure 7, quiescent star formation dominates the stellar mass growth history in $L_*$ (Rank 3) galaxies (bottom panel). In Rank 2 galaxies, they are comparable to each other most of the time. However, in the most massive (Rank 1) galaxies, merger accretion takes over star formation as the most important channel of stellar mass growth around $z \sim 2$ (top panel). Our result is qualitatively consistent with that of Oser et al. (2010).

### 5. SUMMARY AND DISCUSSION

We have investigated the assembly history of stellar components of massive (super-$L_*$) galaxies, using semi-analytic approaches. Our major results can be summarized as follows.

1. More massive galaxies grow in number faster than less massive galaxies, as a result of the hierarchical nature of galaxy clustering. This result is consistent with the recent observation of MK10.

2. The conflict between the predictions from hierarchical models and downsizing is reconcilable. If we consider only direct progenitors of massive galaxies, our models suggest “upsizing” rather than downsizing; that is, the direct progenitors of more massive galaxies grow more slowly. However, if we consider all the progenitors, direct and collateral, the combined mass suggests downsizing. Our models suggest that more massive galaxies have older stellar ages but younger assembly ages.

3. Merger-induced “bursty” star formation is negligible compared to quiescent disk-mode star formation despite the fact that massive galaxies form through numerous mergers. This is because most of the gas-rich major mergers occur at high redshifts when galaxies are small, and recent major mergers tend to be rare and “dry.”
4. Merger accretion is a growingly more important channel of stellar mass growth in massive galaxies. It accounts for 70% of the final stellar mass in the most massive galaxies in our sample \((\log M/M_\odot = 11.5-12.0)\). It is merger accretion that causes much of the mass difference between massive galaxies. This implies that environments play a central role in the growth of massive galaxies.

5. In the most massive galaxies, which are likely bright cluster galaxies, merger accretion has remained the most important channel of stellar mass growth ever since \(z \sim 2\).

The origin of massive galaxies is a pivotal subject of cosmological paradigms and subsidiary galaxy formation theories. The simplest views based on some pieces of observations, such as downsizing, may favor simplistic formation scenarios, while dynamical models of the universe based on the current cosmology and other statistical aspects of observations, such as galaxy luminosity functions, indicate the other direction. This conflict is at its maximum when it comes to the formation of massive galaxies. The goal of this study is to reconcile the various perspectives and to have an accurate understanding on their formation. In this study, we showed that galaxy models in the hierarchical paradigm provide explanations to seemingly contradicting empirical constraints: the faster growth of the number of more massive galaxies (MK10) and downsizing (Cowie et al. 1996). This is encouraging.

Galaxy formation models, whether hydrodynamic or semianalytic, are still incomplete in many aspects: it is often claimed that much of their incompleteness is caused by our limited knowledge of baryon physics. The success of understanding massive galaxy formation on the other hand seems to be more hinged upon our knowledge of large-scale clusterings, and thus dark matter physics. Massive galaxies achieve their grandeur through mergers; thus, only by a realistic consideration of large-scale clustering information is it possible to accurately reconstruct their formation history.

We thank Taysun Kimm and Sadegh Khochfar for their feedback on the early stage of our code development and Intae Jung for helping us run cosmological volume simulations. We thank the anonymous referee for a number of comments and suggestions that improved the clarity of the paper. We acknowledge the support from the National Research Foundation of Korea through the Center for Galaxy Evolution Research (No. 2010-0027910), Doyon grant (No. 20090078756), and DRC grant. This study is in part supported by the Korea Astronomy and Space Science Institute.

REFERENCES

Abadi, M. G., Moore, B., & Bower, R. G. 1999, MNRAS, 308, 947
Almeida, C., Baugh, C. M., Wake, D. A., et al. 2008, MNRAS, 386, 2145
Aragón-Salamanca, A., Baugh, C. M., & Kauffmann, G. 1998, MNRAS, 297, 427
Arnaud, M., Rothenflug, R., Boulade, O., & Vigroux, L. 1992, A&A, 254, 49
Aznar, H. 1966, ApJS, 14, 1
Aubert, D., Pichon, C., & Colombi, S. 2004, MNRAS, 352, 376
Battin, R. H. 1987, An Introduction to the Mathematics and Methods of Astrodynamics (New York: AIAA)
Baugh, C. M., Cole, S., & Frenk, C. S. 1996, MNRAS, 283, 1361
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289
Benson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2002a, MNRAS, 333, 177
Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., & Frenk, C. S. 2002b, MNRAS, 333, 156
