

# Sub-millimetre galaxies in cosmological hydrodynamic simulations: Source number counts and the spatial clustering

Ikkoh Shimizu,<sup>1,2\*</sup> Naoki Yoshida<sup>2,3</sup> †, Takashi Okamoto<sup>4</sup> ‡

<sup>1</sup>College of General Education, Osaka Sangyo University, 3-1-1 Nakagaito, Daito, Osaka 574-8530, Japan

<sup>2</sup>Kavli Institute for the Physics and Mathematics of the Universe, TODIAS,

The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan

<sup>3</sup>Department of Physics, The University of Tokyo, 7-3-1 Hongo, Tokyo 113-0033, Japan

<sup>4</sup>Center for Computational Sciences, University of Tsukuba, Tsukuba 305-8577, Japan

In original form 2011 November 11

## ABSTRACT

We use large cosmological Smoothed-Particle-Hydrodynamics simulations to study the formation and evolution of sub-millimetre galaxies (SMGs). In our previous work, we studied the statistical properties of ultra-violet selected star-forming galaxies at high redshifts. We populate the same cosmological simulations with SMGs by calculating the reprocess of stellar light by dust grains into far-infrared to millimetre wavebands in a self-consistent manner. We generate light-cone outputs to compare directly the statistical properties of the simulated SMGs with available observations. Our model reproduces the submm source number counts and the clustering amplitude. We show that bright SMGs with flux  $S > 1$  mJy reside in halos with mass of  $\sim 10^{13} M_{\odot}$  and have stellar masses greater than  $10^{11} M_{\odot}$ . The angular cross-correlation between the SMGs and Lyman- $\alpha$  emitters is significantly weaker than that between the SMGs and Lyman-break galaxies. The cross-correlation is also weaker than the auto-correlation of the SMGs. The redshift distribution of the SMGs shows a broad peak at  $z \sim 2$ , where Bright SMGs contribute significantly to the global cosmic star formation rate density. Our model predicts that there are hundreds of SMGs with  $S > 0.1$  mJy at  $z > 5$  per 1 square degree field. Such SMGs can be detected by ALMA.

**Key words:** Galaxies – sub-millimetre galaxy; Galaxies – Formation; Galaxies – correlation function

## 1 INTRODUCTION

An array of recent millimetre and sub-millimetre (submm) observations revealed the physical properties of sub-millimetre bright galaxies (SMGs) with luminosities  $10^{12} \sim 10^{13} L_{\odot}$  (Swinbank et al. 2004; Chapman et al. 2005; Capak et al. 2008; Coppin et al. 2009; Daddi et al. 2009; Tamura et al. 2009; Knudsen et al. 2010; Riechers et al. 2010; Hatsukade et al. 2011). The power sources of such luminous SMGs are thought to be very active star formation with  $100 - 1000 M_{\odot} \text{yr}^{-1}$  and/or active galactic nuclei (AGNs). Ultra-violet photons from massive stars or AGNs are converted to photons in submm to infrared bands by dust grains via their thermal emission. Alexander et al. (2003) find that, in the Chandra Deep Field North survey, X-rays

are detected from more than one-third of bright SMGs, which indicates a significant contribution from AGNs.

SMGs provide important information on the star formation rate, the stellar initial mass function, and the chemical evolution of the galaxies at high redshifts. Because SMGs are highly obscured by dust, they often appear dark in UV and optical wavebands. Thus measuring their redshifts by conventional techniques is very difficult. Moreover, identifying the optical counterpart is not easy because the spatial resolution of the currently operating submm telescopes is much worse than that of large optical telescopes.

So far, most of the SMGs with measured redshifts are at  $z < 3$  (Swinbank et al. 2004; Chapman et al. 2005), but a few SMGs have been found at higher redshifts (Capak et al. 2008; Coppin et al. 2009; Daddi et al. 2009; Knudsen et al. 2010). SMGs generally show strong clustering (Webb et al. 2003; Blain et al. 2004; Scott et al. 2006; Weiβ et al. 2009; Cooray et al. 2010; Hickox et al. 2012). Maddox et al. (2010) measured the angular correlation of

\* E-mail:ikk.shimizu@ipmu.jp

† E-mail:naoki.yoshida@ipmu.jp

‡ E-mail:tokamoto@ccs.tsukuba.ac.jp

## 2 Shimizu, Yoshida and Okamoto

the 350  $\mu\text{m}$  and 500  $\mu\text{m}$  selected SMGs in the Herschel-ATLAS survey. The SMGs are strongly clustered while the 250  $\mu\text{m}$  selected sample showed weak or no clustering signals. It is thought that SMGs are very massive systems with large gas reservoir ( $\sim 10^{11} \text{ M}_\odot$ ) and with large stellar masses ( $\sim 10^{11} \text{ M}_\odot$ ) (Greve et al. 2005; Tacconi et al. 2006; Wardlow et al. 2011). Also, SMGs are likely to be ancestors of massive elliptical galaxies in the local universe (Lilly et al. 1996; Smail et al. 2004).

Chapman et al. (2005) report that the redshift distribution of SMGs shows a peak at  $z \sim 2$ , similarly to that of AGNs, although the size of the sample is small. Several SMGs have counterparts of star-forming galaxies such as Lyman break galaxies (LBGs) and Lyman  $\alpha$  emitters (LAEs) (Geach et al. 2005, 2007; Beelen et al. 2008; Daddi et al. 2009). However, many of high-redshift galaxies are not seen as SMGs probably because their submm fluxes are below typical observational flux limits ( $> 1 \text{ mJy}$ ). Because observations so far detect very bright SMGs, it remains unclear if SMGs are generally associated with LBGs/LAEs. Matsuda et al. (2007) and Tamura et al. (2010) failed to detect submm sources at the location of LAEs in a protocluster region even though the target galaxies are very bright in Ly $\alpha$ .

Dayal et al. (2010) use cosmological simulations and found that the submm fluxes of high- $z$  LAEs would be less than 0.1 mJy. They argue, however, that many of such galaxies can be detected by ALMA. Yajima, Umemura, & Mori (2011) study the evolution of submm flux of a simulated galaxy by using detailed three-dimensional radiative transfer calculations. They show that the simulated galaxy in a very early LAE phase can be detected by ALMA. Observing galaxies in the early formation phase is important to understand the early chemical evolution.

One of the most important observational quantities is the source number count of SMGs (Greve et al. 2004; Laurent et al. 2005; Perera et al. 2008; Austermann et al. 2009, 2010; Scott et al. 2010; Hatsukade et al. 2011). Unfortunately, there remain substantial uncertainties in theoretical models. Baugh et al. (2005) calculate the source number count of SMGs using a semi-analytic galaxy formation model. Interestingly, in order to reproduce the observed source number count, they make a crucial assumption that the stellar initial mass function (IMF) is ‘top-heavy’ when starburst occurs (see also Lacey et al. (2010)). More recently, Fontanot et al. (2007) reproduced the source number count of SMGs without assuming top-heavy IMF. They argue that the major difference between Baugh et al. (2005) and their study may be in the gas cooling model. In Fontanot et al. (2007), the hot gas in a galaxy is assumed to cool more efficiently than in Baugh et al. (2005). However, because of the efficient cooling, the model of Fontanot et al. (2007) overproduces massive galaxies at lower redshifts. It appears that some efficient star-formation recipe is needed in the semi-analytic models, in order to reproduce the source number count of SMGs. It is interesting to see if the same is true for cosmological hydrodynamic simulations of galaxy formation. Davé et al. (2010) study the physical properties of SMGs using cosmological simulations. They claim that smooth infalling gas or/and gas-rich satellites are important to produce large star formation rates. Although their

finding seems reasonable, they employ a simple model that galaxies with very large star-formation rates are identified as SMGs. They do not calculate dust absorption and the resulting submm flux. It is important and timely to study in detail the statistical properties of SMGs over a wide range of redshift using cosmological hydrodynamic simulations by considering dust absorption and re-emission consistently.

In this paper, we study the statistical properties of SMGs such as the stellar mass function, the source number count and the angular correlation function. To this end, we perform large cosmological hydrodynamic simulations based on the standard  $\Lambda\text{CDM}$  cosmology. Our simulations follow star formation, supernova feedback, and metal enrichment self-consistently. For the galaxies identified in our cosmological simulations, we calculate the spectral evolution and dust extinction to estimate the FIR luminosity. In our earlier work (Shimizu, Yoshida & Okamoto 2011), we used the same set of simulations to study the statistical properties of LAEs at  $z = 3.1$ . Our model successfully reproduced all the available observational data at  $z = 3.1$  such as the Lyman- $\alpha$  luminosity function, the angular correlation functions, and the Lyman- $\alpha$  equivalent width distribution. Our aim in the present paper is to build a consistent theoretical model of galaxy formation that reproduces the observed properties of both ultra-violet selected galaxies and SMGs.

Throughout the present paper, we adopt the  $\Lambda\text{CDM}$  cosmology with the matter density  $\Omega_M = 0.27$ , the cosmological constant  $\Omega_\Lambda = 0.73$ , the Hubble constant  $h = 0.7$  in units of  $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and the baryon density  $\Omega_B = 0.046$ . The matter density fluctuations are normalised by setting  $\sigma_8 = 0.81$  (Spergel et al. 2003). All magnitudes are expressed in the AB system, and the Ly $\alpha$  EW<sub>Ly $\alpha$</sub>  values in this paper are in the rest frame.

## 2 THEORETICAL MODEL

Our simulation code is based on an updated version of the Tree-PM smoothed particle hydrodynamics (SPH) code GADGET-3 which is a successor of Tree-PM SPH code GADGET-2 (Springel 2005). We employ  $N = 2 \times 640^3$  particles in a comoving volume of  $100 h^{-1} \text{ Mpc}$  on a side. The mass of a dark matter particle is  $2.41 \times 10^8 h^{-1} \text{ M}_\odot$  and that of a gas particle is  $4.95 \times 10^7 h^{-1} \text{ M}_\odot$ , respectively. We implement relevant physical processes such as star formation, supernova feedback and chemical enrichment following Okamoto, Nemmen & Bower (2008), Okamoto, & Frenk (2009) and Okamoto et al. (2010). In Okamoto et al. (2010), two wind models for supernova (SN) feedback are examined. One model assumes that the initial wind speed of gas around stellar particles that trigger SN is proportional to the local velocity dispersion of dark matter ( $\sigma$ ). The model is motivated by recent observations of Martin (2005). The other model assumes a constant initial wind speed, as has been used in many cosmological hydrodynamic simulations (Springel & Hernquist 2003a; Nagamine, Springel & Hernquist 2004). The major difference between the two models is the wind mass-loading per star formation rate. In the former model, the mass-loading is proportional to  $\sigma^{-2}$  whereas it is constant for all galaxies in the latter model. Okamoto et al. (2010) show that the former model reproduce the observed satellite luminos-

ity function and the luminosity-metallicity relation in the Local Group satellites. We set the initial wind speed of gas to be five times the local velocity dispersion, following Okamoto et al. (2010).

In our previous paper (Shimizu, Yoshida & Okamoto 2011), we studied the statistical properties of UV-selected galaxies. We implemented wavelength dependent dust absorption to calculate the UV and Lyman- $\alpha$  luminosities of the galaxies. In the present paper, in order to calculate the FIR luminosity, we assume that the absorbed UV photons are re-emitted in FIR by thermal dust grains.

## 2.1 Dust absorption

The treatment of dust absorption is described in Shimizu, Yoshida & Okamoto (2011). Briefly, we calculate the optical depth  $\tau_d(\lambda)$  for UV continuum photons on the assumption that dust is distributed in a sphere with a certain radius around a galaxy. Then we use only one length scale, the effective radius of dust distribution  $r_d$ , to evaluate the optical depth.

$$\tau_d(\lambda) = \frac{3\Sigma_d Q(\lambda)}{4a_d s}, \quad (1)$$

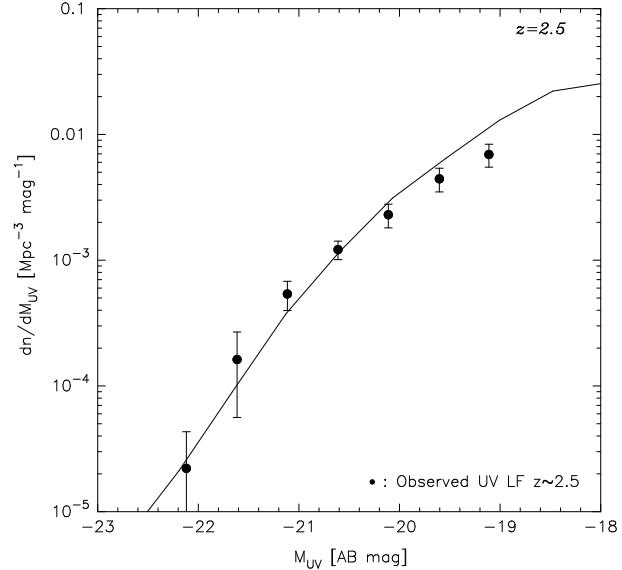
where  $a_d$  and  $s$  are the dust grain size and the material density of dust grains, respectively, and  $\Sigma_d = M_d/(\pi r_d^2)$  is the dust surface mass density. For simplicity, we assume a single-sized dust component. We use the dust optical constant  $Q(\lambda)$  as a function of wavelength given in Draine & Lee (1984), so that we obtain realistic SEDs for individual galaxies. We set  $a_d = 0.1 \mu\text{m}$  and  $s = 2.5 \text{ g cm}^{-3}$  for dust grains that are produced by supernovae (Todini & Ferrara 2001; Nozawa et al. 2003). In this case, the resulting extinction curve is flatter than the SMC extinction curve and the Calzetti extinction curve, because  $Q(\lambda)$  for large grains is nearly independent of wavelength (Hirashita et al. 2008, 2010). Intriguingly, recent observations of ultraluminous infrared galaxies at  $z \sim 1$  suggest that the extinction curve is flatter than the Calzetti extinction curve (Shimizu et al. 2011). Furthermore, the extinction curve for the SN-produced dust size distribution reproduces the observation well (Hirashita et al. 2008, 2010).

Although our single-sized dust model might seem crude, the resulting extinction curve does not significantly affect our main results presented in the paper. In section 4, we explicitly show that adopting the steep SMC extinction curve yields virtually the same statistical quantities for the SMG population.

We calculate the dust surface mass density as  $\Sigma_d = M_d/(\pi r_d^2)$  where  $M_d$  and  $r_d$  are the dust mass and the effective dust distribution radius of each simulated galaxy, respectively. The dust mass is calculated by

$$M_d = 0.01 m_g \left( \frac{Z}{Z_\odot} \right), \quad (2)$$

where  $m_g$ ,  $Z$  and  $Z_\odot$  denote the gas mass, the gas metallicity and the solar metallicity, respectively (Draine et al. 2007). We adopt  $Z_\odot = 0.02$ , and then the above equation gives  $M_d = 0.01 m_g$  for  $Z = Z_\odot$ , which means that one percent (in mass) of the gas in a simulated galaxy with  $Z = Z_\odot$  is locked in dust grains. Hence the dust mass scales with the gas metallicity. We assume that the effective dust distribution radius of a galaxy  $r_d$  is a fraction of its virial radius,



**Figure 1.** UV (1500 Å at rest-frame) luminosity function of the simulated galaxies at  $z = 2.5$ . The solid line shows our simulation result, which we compare with the observational data (points with error bars) of Oesch et al. (2010).

$$r_d = e \times r_{\text{vir}}, \quad (3)$$

where  $e$  is a global parameter to be kept constant for all the galaxies. The numerical value of  $e$  is determined by matching the UV luminosity function at  $z = 2.5$  (see Fig. 1). The best fit value at  $z = 2.5$  is  $e = 0.09$ .

We calculate the escape fraction of the UV continuum photons  $f_c$  using the following equation:

$$f_c(\lambda) = \frac{1 - e^{-\tau_d(\lambda)}}{\tau_d(\lambda)}. \quad (4)$$

The FIR luminosity  $L_{\text{FIR}}$  is then set equal to the total energy of UV photons absorbed by dust,

$$L_{\text{FIR}} = \int [L^{\text{int}}(\nu) - L^{\text{real}}(\nu)] d\nu, \quad (5)$$

where  $L^{\text{int}}(\nu)$  is the intrinsic luminosity (per frequency) of the galaxy, and  $L^{\text{real}}(\nu)$  is the luminosity after dust absorption. The above integration is evaluated over from UV to optical wavelengths. In the next subsection, we describe the calculation of the FIR flux.

## 2.2 Calculation of the FIR Flux

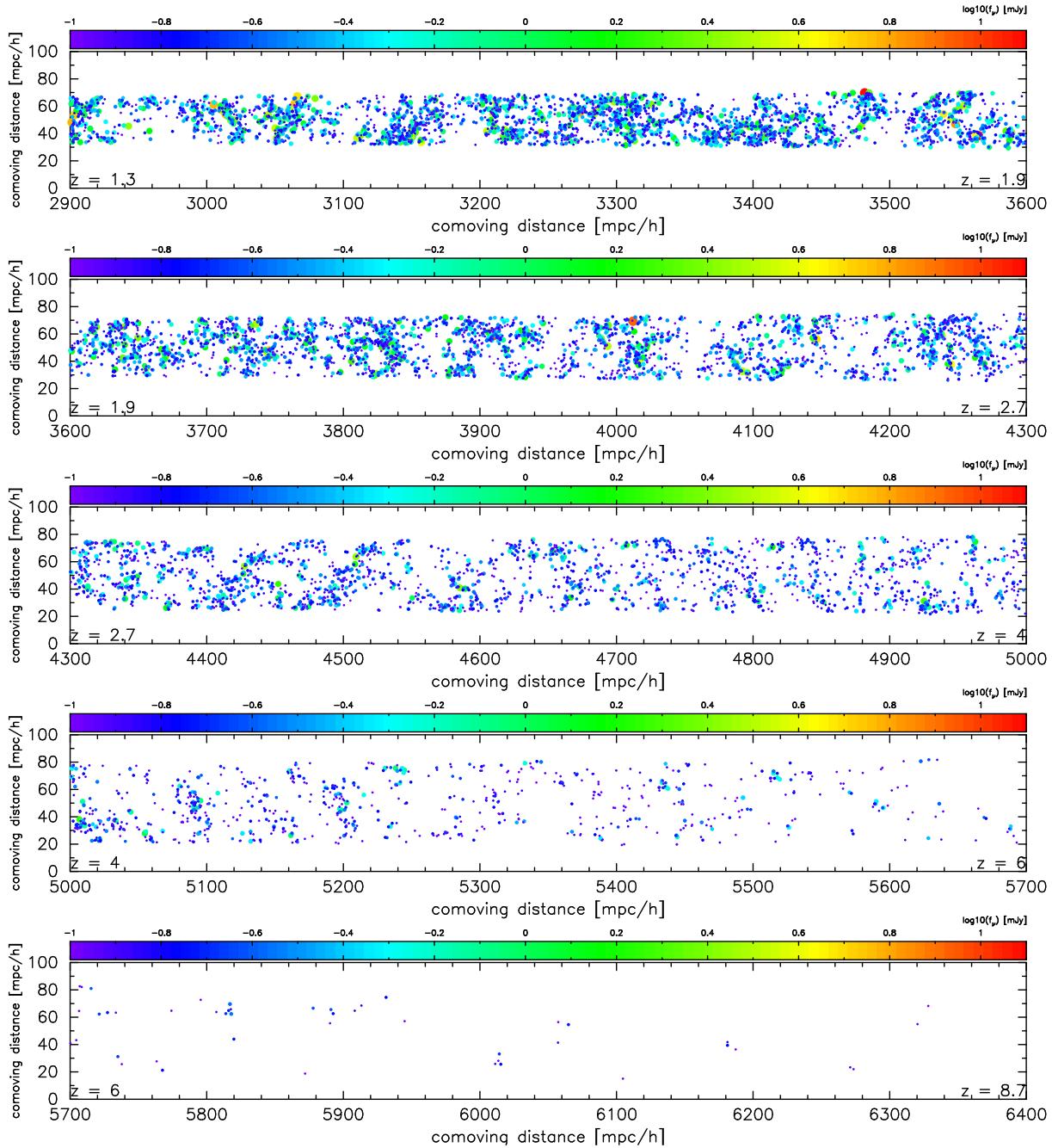
We calculate the FIR flux following Hirashita & Ferrara (2002). For the single-sized dust model, the monochromatic luminosity at frequency  $\nu$  is written as

$$L_\nu^{\text{dust}} = 4\pi M_d \kappa_\nu B_\nu(T_d), \quad (6)$$

where  $\kappa_\nu$ ,  $B_\nu$  and  $T_d$  are the absorption coefficient, the Planck function and the dust grain temperature, respectively. The absorption coefficient  $\kappa_\nu$  in FIR is well described by a power-law  $\kappa_\nu = \kappa_0(\nu/\nu_0)^\beta$  with  $\beta = 1 \sim 2$ .

Then the total FIR luminosity is

$$L_{\text{FIR}} = \int_0^\infty L_\nu^{\text{dust}} d\nu$$



**Figure 2.** The spatial distribution of our simulated SMGs on the past light cone of an observer. We plot SMGs with submm flux greater than 0.1 mJy at 270 GHz. The redshift range is from  $z = 1.3$  to  $z = 8.5$ . The field of view of the light cone is 40 arcmin on a side. The colour indicates the submm flux; blue for faint SMGs and yellow to red for bright SMGs. Also the point size scales with the submm flux. The smallest points and the largest points correspond to 0.1 mJy and for 30 mJy, respectively.

$$\begin{aligned}
 &= 4\pi M_d \kappa_{\nu_0} \nu_0^{-\beta} \left( \frac{kT_d}{h_p} \right)^{4+\beta} \left( \frac{2h_p}{c^2} \right) \\
 &\times \int_0^\infty \frac{x^{3+\beta}}{e^x - 1} dx,
 \end{aligned} \tag{7}$$

where  $k$ ,  $h_p$  and  $c$  are the Boltzmann constant, the Planck constant, and the speed of light, respectively. We denote  $x = h_p \nu / (kT_d)$ . We set a normalization parameter  $\nu_0 = 3.0 \times 10^{12} \text{ Hz}$  which corresponds to the wavelength of  $100 \mu\text{m}$ . The absorption coefficient  $\kappa_0$  is given by  $\kappa_0 = 3Q_{\nu_0}(4as)^{-1}$ ,

where  $Q_{\nu_0}$  is the dust optical constant at  $\nu_0$ . The value of  $\kappa_0$  is  $46.95 \text{ g}^{-1}\text{cm}^2$ . Note that  $Q_{\nu_0}/a$ , and hence  $\kappa_0$ , are independent of dust size.

Using equation (7), we derive the dust temperature  $T_d$  as

$$T_d = 7.64 \left( \frac{L_{\text{FIR}}/L_\odot}{M_d/M_\odot} \right)^{1/6} [\text{K}], \tag{8}$$

where we set  $\beta = 2$  assuming graphite and silicate dust grains (Draine & Lee 1984). Finally, the observed submm flux of a galaxy is obtained from

$$f_\nu^{\text{obs}} = \frac{(1+z)L_{\nu(1+z)}}{4\pi d_L^2}, \quad (9)$$

where  $d_L$  is the luminosity distance to the galaxy. The dust temperature for our sample galaxies is typically  $T_d = 33$  K. To directly compare our model with the observational results of the AzTEC observation, we consider the filter response function of the AzTEC detector when we calculate the submm flux at the observed frame 1.1 mm.

### 3 RESULT

#### 3.1 UV luminosity function

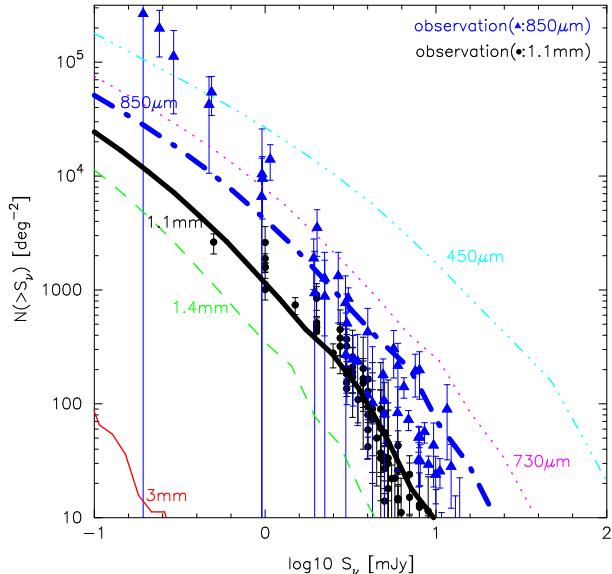
We first examine whether our model reproduces the UV luminosity function. This is an important check because the total FIR luminosity is equal to, by definition in our model, the total UV luminosity absorbed by dust grains. Fig. 1 shows the UV (1500 Å at rest-frame) luminosity function of our model at  $z = 2.5$ . Clearly, our model reproduces the observational data very well. Given the equally good agreement between our model predictions for  $z = 3.1$  LAEs (Shimizu, Yoshida & Okamoto 2011) and other available observational data of UV-selected galaxies (Shimizu et al., in preparation), we are confident that our cosmological simulations can be used to build a self-consistent model for UV-selected galaxies and SMGs at high redshifts.

#### 3.2 The light cone output

Most of the recent observations do not provide redshifts for SMGs. In order to directly compare our model with the observations, we generate a light cone output which extends from  $z = 1.3$  to  $z = 8.5$  using a number of simulation outputs. We coordinate the output redshifts to fill the volume of a lightcone from  $z = 1.3$  to  $z = 8.5$  without gaps. We then randomly shift each simulation box so that the same objects (at different epochs) do not appear multiple times on a single line-of-sight. Fig. 2 shows the distribution of simulated SMGs on the light cone from  $z = 1.3$  to  $z = 8.7$ . For this figure, we set the SMG detection limit to be 0.1 mJy at 270 GHz. Most of the SMGs are located at  $z = 1 - 4$ , but there are dozens of SMGs even at  $z > 6$ . Such faint SMGs can be detected by ALMA with less than one hour integration using 16 antennas at the 350 GHz band. The high redshift SMGs will provide invaluable information on star formation activities in the early universe.

#### 3.3 The source number counts

We use the generated light cone output to conduct mock observations of the SMGs. Fig. 3 shows the number counts of SMGs. The filled circles and the filled triangles are the observational results at the observed frame 1.1 mm and 850 μm, respectively. The lines shows our model predictions for various bands. To directly compare our model with the observation, we use the filter response function of the AzTEC



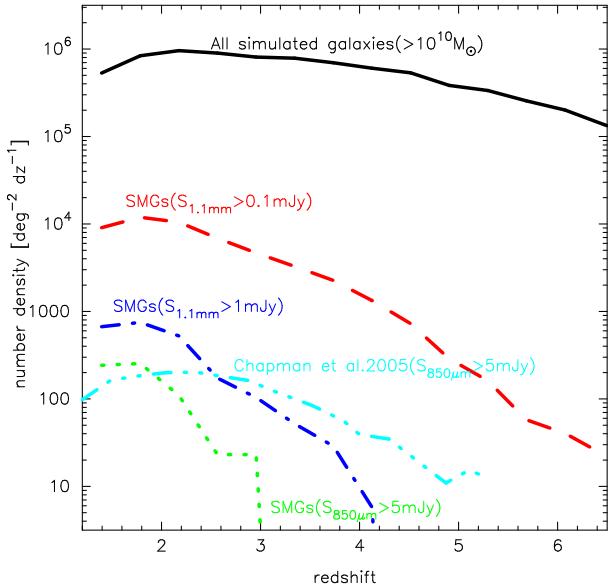
**Figure 3.** The source number counts of SMGs for six observed frame fluxes. The thick solid line shows the number count at the observed frame 1.1 mm (AzTEC detector band). The thick dash-dotted line is for the observed frame 850 μm, and the other lines are for the observed frames 3.0 mm, 1.4 mm, 730 μm and 450 μm which correspond to the five ALMA bands, respectively. The filled circles with error bars show the observed number counts at 1.1 mm by AzTEC (Greve et al. 2004; Laurent et al. 2005; Perera et al. 2008; Austermann et al. 2009, 2010; Scott et al. 2010; Hatsukade et al. 2011). The filled triangles with error bars show the observed number counts at 850 μm (Chapman et al. 2005).

detector when we calculate the submm flux at the observed frame 1.1 mm.

Our model reproduces the observational results remarkably well. The 850 μm source count is somewhat smaller than the observation at  $S < 1$  mJy, but the error bars are also large at the faint end. The thin solid, dashes, dotted and double-dotted lines are the predicted SMG number counts at observed frames 3.0 mm, 1.4 mm, 730 μm and 450 μm, which correspond to the ALMA detector bands. The number of SMGs increases with shorter wavelength, suggesting that SMGs are detected easier at shorter wavelengths; this is another manifestation of the negative k-correction for SMGs.

#### 3.4 The SMG number density evolution

In this subsection, we study the redshift evolution of the SMG population in more detail. Fig. 4 shows the evolution of the number density of SMGs. The dashed and dash-dotted lines represent the number counts for the detection limit of 0.1 mJy and 1 mJy at the observed frame 1.1 mm. We also plot our model number count at 850 μm with the detection limit 5 mJy. The triple-dotted line shows the observational result at 850 μm from Chapman et al. (2005). Note that the sample of Chapman et al. (2005) consists of SMGs in multiple fields with various areas. Thus we normalize the observed redshift distribution arbitrarily to make the comparison easy. We also plot the number of all the simulated galaxies (the solid line) with halo mass more than  $10^{10} M_\odot$ . Regardless of the selection, the SMG number counts peak



**Figure 4.** The evolution of the number density of SMGs as a function of redshift. The dashed and dash-dotted lines represent our model prediction for the detection limit of 0.1 mJy and 1 mJy at the observed frame 1.1 mm, respectively. The dotted line is also our model prediction for a high detection limit of 5 mJy at the observed frame 850  $\mu$ m, to be compared with the observation of Chapman et al. (2005). For reference, we plot the number density of all simulated galaxies (solid line) with halo masses greater than  $10^{10} M_\odot$ .

at  $z \sim 2 - 3$ . Our model lacks very bright SMGs at high redshifts. This is likely due to the small simulation box size (cosmic variance). Note also that, in general, determining the redshift distribution of the SMGs are not trivial if only submm observations are used.

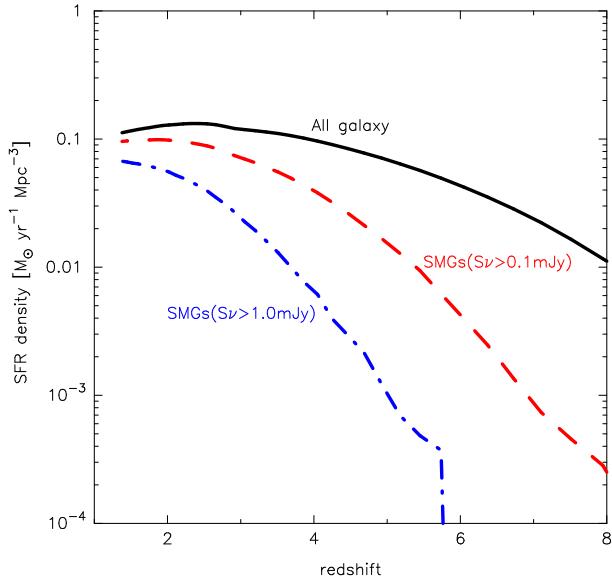
In our model, approximately 90% of SMGs with  $> 0.1$  mJy at 1.1 mm are at  $z > 2$ . This implies that most of the observed SMGs are high- $z$  galaxies, and the contribution from low- $z$  galaxies is small.

We calculate the contribution from SMGs to the global cosmic star formation rate (SFR) evolution. Fig. 5 shows the sum of the SFRs of all the simulated galaxies (solid line), that of SMGs with  $> 0.1$  mJy at observed frame 1.1 mm (dashed line) and that of SMGs with  $> 1$  mJy at observed frame 1.1 mm (dot-dashed line).

The contribution of SMGs to the cosmic SFR is small at  $z > 6$ . However the contribution rapidly increases with decreasing redshift. At  $z \sim 2$ , Bright SMGs with  $S > 1$  mJy at 1.1 mm contributes roughly one third to the total SFR. Clearly, it is important to include SMGs to estimate the true global star-formation rate at  $z \sim 1 - 4$ .

### 3.5 The stellar mass of SMGs

Fig. 6 shows the SFR (the left panel) and specific SFR (the right panel) as a function of the stellar mass. We plot all the SMGs in our light cone data from  $z = 1.3$  to  $z = 8.7$ . There are tens of SMGs with the SFR greater than  $1000 M_\odot \text{yr}^{-1}$ , whereas the specific SFR is roughly constant ( $\sim 10^{-9} \text{ yr}^{-1}$ ). Bright SMGs have very large stellar masses in excess of  $10^{12} M_\odot$ . Some of our simulated SMGs with largest stel-



**Figure 5.** The star formation rate density contributed by SMGs. The solid, dashed and dash-dotted lines are for all the galaxies, SMGs with  $S > 0.1$  mJy at observed frame 1.1 mm and SMGs with  $S > 1$  mJy at observed frame 1.1 mm, respectively.

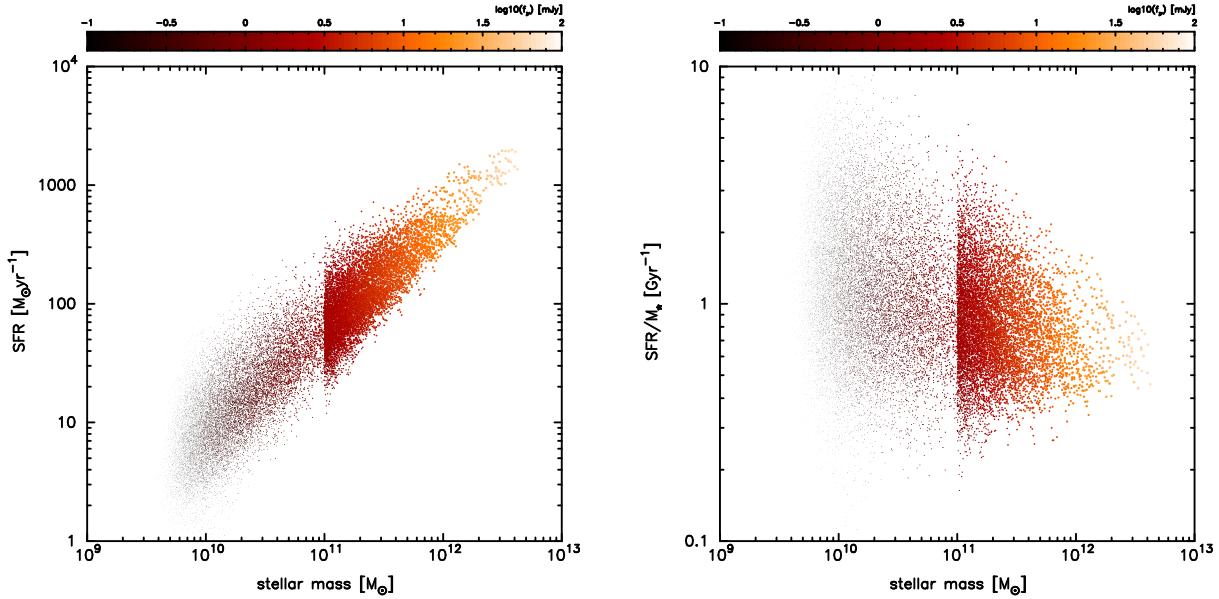
lar masses ( $> 10^{12} M_\odot$ ) consist of multiple clumps within a single host halo. However, we do not treat each such clump of star particles as an individual galaxy. Essentially we regard one halo as a SMG. This crude definition is motivated by the fact that the current observations do not have high enough angular resolution to resolve the internal structure within a halo. See also the discussion in section 3.6 and Fig. 10 which shows the host halo mass of the SMGs.

### 3.6 The distribution of SMGs and the angular correlation functions

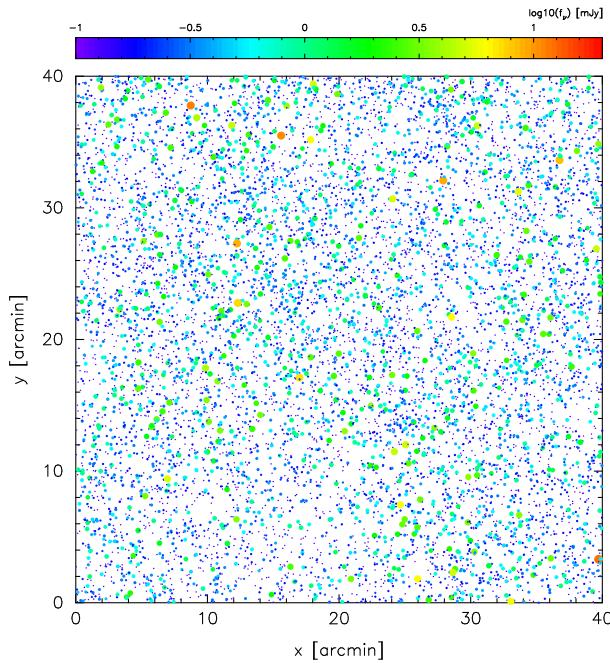
Fig. 7 shows the projected distribution of the SMGs on the lightcone. We plot the SMGs with  $S > 0.1$  mJy at the observed frame 1.1 mm. As is naively expected, brighter SMGs are strongly clustered, whereas faint SMGs appear to be distributed more or less homogeneously over the field of view.

Next, we calculate the angular two-point correlation function (ACF) of the SMGs to quantify the clustering strength. In Fig. 8, the solid line is the correlation function of our simulated SMG sample. For comparison, we have chosen the same detection limit of 1.26 mJy as the observation of Williams et al. (2011). Our simulation result is consistent with the observational data although with substantial scatter for both the data.

We calculate the angular cross correlation between the simulated SMGs and other UV-selected galaxies. We use one simulation box at  $z = 3.1$  for this calculation. We consider two flux limits for our simulated SMGs, 0.1 mJy and 1 mJy at observed frame 1.1 mm. The number of simulated SMGs for the two cases is then 993 and 18, respectively. We identify simulated galaxies which satisfy  $\text{EW}_{\text{Ly}\alpha} > 30 \text{ \AA}$  and  $L_{\text{Ly}\alpha}^{\text{obs}} > 1.0 \times 10^{42} [\text{erg/s}]$  as LAEs (Shimizu, Yoshida & Okamoto 2011). Here  $L_{\text{Ly}\alpha}^{\text{obs}}$  and  $\text{EW}_{\text{Ly}\alpha}$  are the observed Ly $\alpha$  luminosity and the rest frame Ly $\alpha$  equivalent width. We define



**Figure 6.** The star formation rates of the simulated SMGs (left) and the specific star formation rates (right) as a function of the stellar mass. We use bigger size points for brighter SMGs. Also the colour indicates the submm flux. To reduce the number of points, we randomly chose one tenth of the galaxies for stellar masses below  $10^{11} M_{\odot}$ .



**Figure 7.** The projected distribution of SMGs in a field-of-view of  $40 \text{ arcmin} \times 40 \text{ arcmin}$ . The left and right panel show SMGs with  $S > 0.1 \text{ mJy}$ . The point size scales with the submm flux. The smallest points and the largest points are for  $\sim 0.1 \text{ mJy}$  and for  $\sim 30 \text{ mJy}$ , respectively. Also the point colour indicates the submm flux.

LBGs at  $z \sim 3$  by the following selection:

$$23 < z_{AB} \leq 26.0, \quad (10)$$

$$U_{AB} - V_{AB} \geq 0.8, \quad V_{AB} - z_{AB} \leq 2.7, \quad (11)$$

$$U_{AB} - V_{AB} \geq 1.8 \quad (V_{AB} - z_{AB}) + 1.6, \quad (12)$$

where  $U_{AB}$ ,  $V_{AB}$  and  $z_{AB}$  denote AB magnitude of U band,

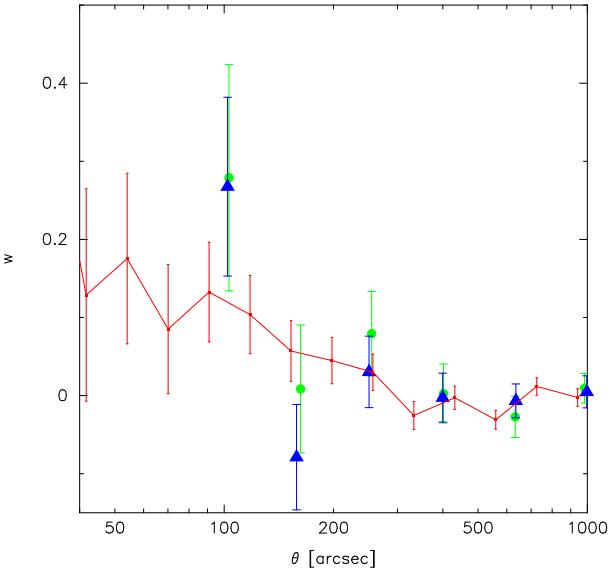
V band and z band, respectively. We select the LBGs in this way consistently with the observation of Yoshida et al. (2008). Finally, we locate 2913 LAEs and 3211 LBGs in the simulation box at  $z \sim 3$ .

Fig. 9 shows the SMG-LAE and SMG-LBG angular cross-correlation functions. For comparison, we also plot the auto correlation function of SMGs (solid circles with error bars) and the cross-correlation between SMGs and dark halos with mass greater than  $> 10^{10} M_{\odot}$  (stars with error bars). For the SMG flux limit of  $0.1 \text{ mJy}$ , the SMG-LBG cross correlation is significantly stronger than the SMG-LAE cross correlation. The former is indeed very similar in the shape and the amplitude to the auto-correlation of the SMGs, implying that the SMGs are a similar population to the LBGs (see also the SFR shown in Fig. 6). We have checked that many UV bright galaxies in our sample are identified as SMGs.

### 3.7 The host halo mass of SMGs

Fig. 10 shows the submm flux of simulated SMGs as a function of their host halo mass. We use all the SMGs in our light cone data from  $z = 1.3$  to  $z = 8.7$ . The submm flux of simulated SMGs roughly scales with their host halo mass, but with large dispersions for low mass objects. We show the dispersion of host halo mass and submm flux in each mass bin. We also plot dark matter halo mass function at  $z = 2$ . Clearly, bright SMGs with ( $S > 1 \text{ mJy}$ ) reside in rare, large halos with several times  $10^{12} M_{\odot}$ .

It is important to note that, because the angular resolution of the current mm/submm observations is still limited to  $\sim 30 - 60$  arcseconds, multiple galaxies in a galaxy-group size halo are not well resolved. Future observations using ALMA will likely find multiple sources in massive halos. From Fig. 10, we expect that SMGs hosted by smaller halos



**Figure 8.** We plot the angular two-point correlation function of the simulated SMGs (solid line with error bars). The points with error bars are the observational data of Williams et al. (2011). Different symbols refer to different catalogues of SMGs. The circle and triangle points are  $3.5\sigma$  sources and  $3\sigma$  sources, respectively.

with several times  $10^{11} M_{\odot}$  can be detected by observation with a detection limit of 0.1 mJy.

#### 4 CONCLUSIONS AND DISCUSSION

We have studied a variety of statistical properties of SMGs using a large cosmological hydrodynamic simulation. The simulation follows the formation and evolution of star-forming galaxies by employing a new feedback model of Okamoto et al. (2010). Earlier in Shimizu, Yoshida & Okamoto (2011), the same simulation was used to study the properties of LAEs at  $z = 3.1$ . We have shown in the present paper that our SMG model is consistent with a number of observational data currently available. The redshift distribution of the SMGs peaks broadly at  $z \sim 2$ . Most of the SMGs are at  $z > 2$ , and the contribution from low- $z$  ( $z < 1$ ) galaxies to the submm source count is small. The contribution from the SMGs to the global star formation rate density increases with decreasing redshift. Bright SMGs contributes significantly to the star formation rate at  $z \sim 2 - 3$ . We calculate the angular two point correlation to quantify the clustering amplitude of SMGs. Our result is consistent with the observations of Williams et al. (2011). We also study the angular cross correlation functions between SMGs and LBGs and the correlation between SMGs and LAEs. The former is significantly stronger than the latter. Considering the high star-formation rates of the simulated SMGs, we argue that SMGs are similar population to LBGs. Finally, we explicitly showed that the bright SMGs preferentially reside in massive halos ( $> 10^{12} M_{\odot}$ ) and that their typical stellar mass are greater than  $10^{11} M_{\odot}$ .

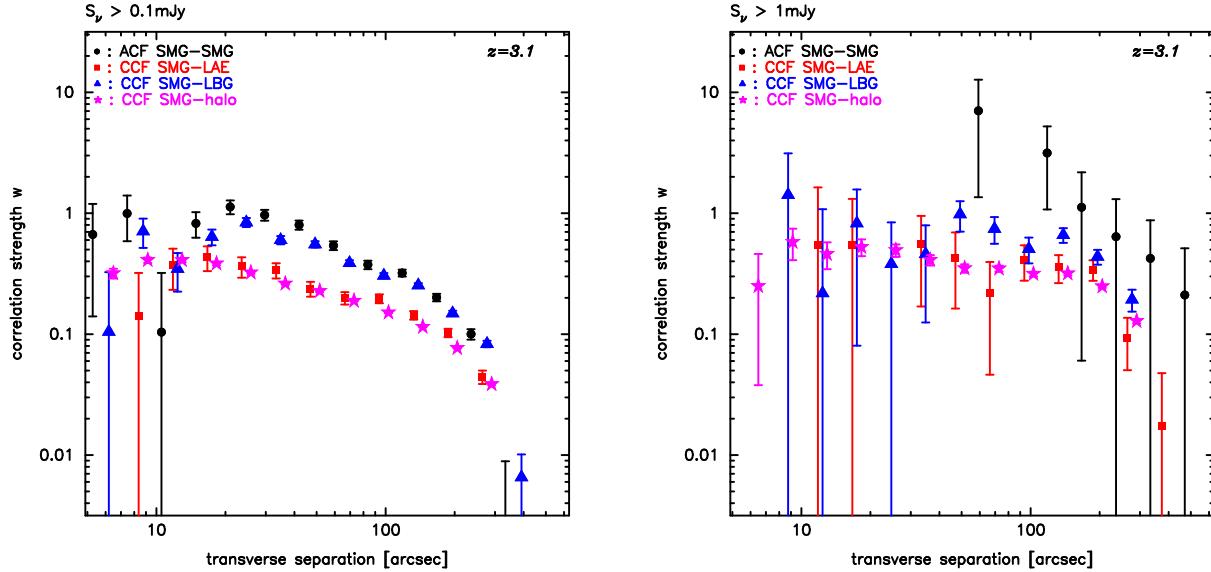
In our simple model of a single-sized, supernovae produced dust, the resulting extinction curve is flatter than the standard ones such the SMC extinction curve or Calzetti extinction curve. This mean that the reprocess of UV photons

to FIR by dust works more efficiently in our model. However, interestingly, we have found that the characteristic feature of the extinction curve does not significantly affect our main results.

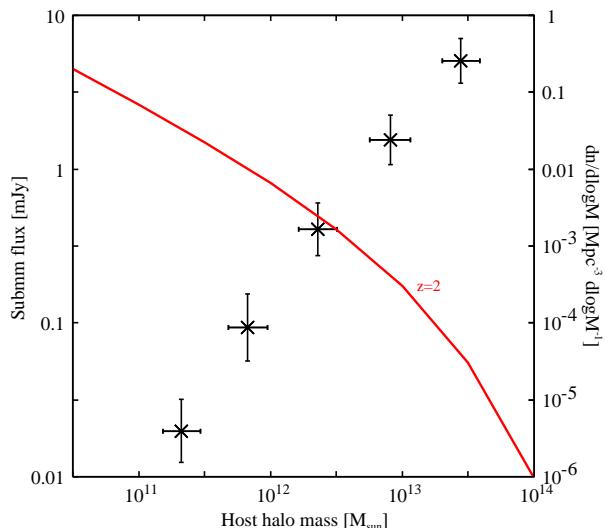
For direct comparison, we calculate the FIR luminosities for our galaxy samples using the SMC extinction curve. We then estimate the SMG number count following otherwise the same procedures as in our fiducial model. Fig. 11 compares the source number count of our model (solid line) and the one calculated using the SMC extinction curve (dashed line). Clearly the difference between our base model and the other is very small. This can be explained by the fact that we normalize the overall luminosity of the galaxies by matching the UV luminosity function at rest-frame 1500 Å to the observed one (see Fig. 1). Even for a different extinction model, the overall level of extinction in UV is always normalized at 1500 Å. Therefore, the difference in the exact shape of the extinction curves near UV does not affect significantly the SMGs number count in our model, as is explicitly shown in Fig. 11. Note, however, that the UV to optical colour of individual galaxies is visibly affected by the extinction curve. Our base model predicts that simulated galaxies appear bluer in UV to optical range than in the case with SMC or Calzetti extinction curve models. Fig. 12 shows, as an illustrative example, the SED of one of our SMG samples from UV to millimetre. The SED for the same galaxy but with SMC extinction curve is also shown. We compare them with the observed SEDs for several SMGs (Michałowski, Hjorth & Watson 2010). Some SMGs have flat (blue) SEDs at UV to optical range whereas others appear redder. The full SEDs are measured only for a limited number of samples. It would be highly interesting to observe SMGs in detail in rest frame optical to UV in order to derive the dust extinction law, and hence the dust size distribution, for star burst galaxies.

Next, we discuss the source confusion limit of submm observations. We have seen in previous sections (Fig. 6 and Fig. 10), that some simulated SMGs have large host halo masses with  $> 10^{13} M_{\odot}$  and/or large stellar masses with  $> 10^{12} M_{\odot}$ . Such massive systems often have multiple substructures (main and satellite galaxies). Basically each such satellite galaxy should be treated as an independent galaxy. Note, however, that the angular resolution of the current submm/mm observation is worse than that of optical observation. For example, the resolution of the AzTEC telescope is 30 arcseconds which corresponds 260 kpc (physical scale) at  $z = 2$ . The virial radii of the galactic halos in our simulation are only as large as  $\sim 200$  kpc (physical size) at  $z = 2$ . Thus, our treatment that a halo hosts one SMG as a whole is reasonable, even realistic, if we compare with observations with angular resolution of 0.5-1 arcminutes.

Finally, we discuss the evolution of the stellar mass function. Our galaxy formation model is defined at high redshift, rather than being calibrated against galaxies at the present epoch. We choose a few model parameters such as the overall normalization of dust extinction to reproduce the observational data of star-forming galaxies at high- $z$  ( $z \sim 3$ ) (Shimizu, Yoshida & Okamoto 2011). Contrastingly, previous semi-analytic models (e.g., Baugh et al. (2005) and Lacey et al. (2010)) determined the main physical parameters to match the observations of the present-day galaxies. For example, Baugh et al. (2005) needed to



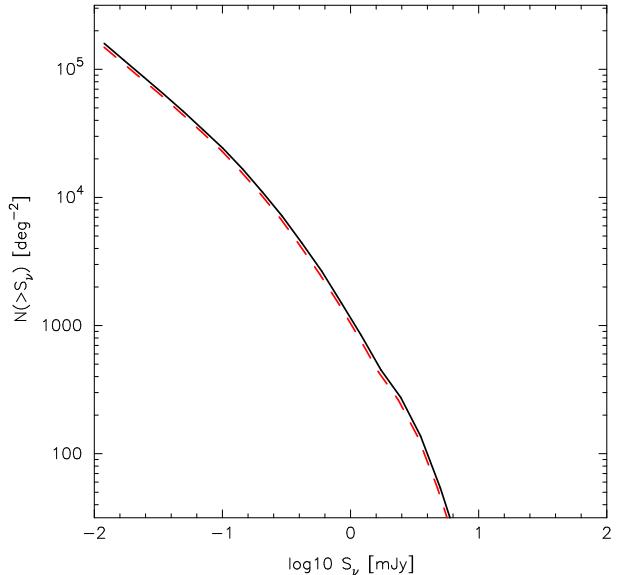
**Figure 9.** We plot the angular cross-correlation with two other populations of galaxies; the SMG-LAE cross-correlation (squares with error bars) and the SMG-LBG cross-correlation (triangles with error bars). For comparison, we also plot the angular auto-correlation of the SMGs (solid circles with error bars) and the angular cross correlation between the SMGs and dark matter halos with  $10^{10} M_\odot$  (stars with error bars). We set two detection limits for the submm flux; 0.1 mJy for the left panel and 1 mJy for the right panel, respectively.



**Figure 10.** The mean submm flux of our simulated SMG samples as a function of their host halo mass. We also plot the halo mass function at  $z = 2$  (the solid line). Bright SMGs with  $S > 1 \text{ mJy}$  are hosted by dark halos with mass greater than  $\sim 10^{13} M_\odot$ .

adopt an extreme IMF (top-heavy IMF) for starburst galaxies at high redshift to reproduce the observed SMG number counts. More recently, Fontanot et al. (2007) reproduced the SMGs number count adopting the Salpeter IMF for all their galaxy samples. However, their model assumes more efficient gas cooling in galactic haloes than in Baugh et al. (2005) and Lacey et al. (2010). Consequently, very massive galaxies with high star-formation rates are formed, which results in disagreement in the stellar mass function at low redshifts ( $z < 1$ ). Clearly it is important to study the stellar mass function of our model.

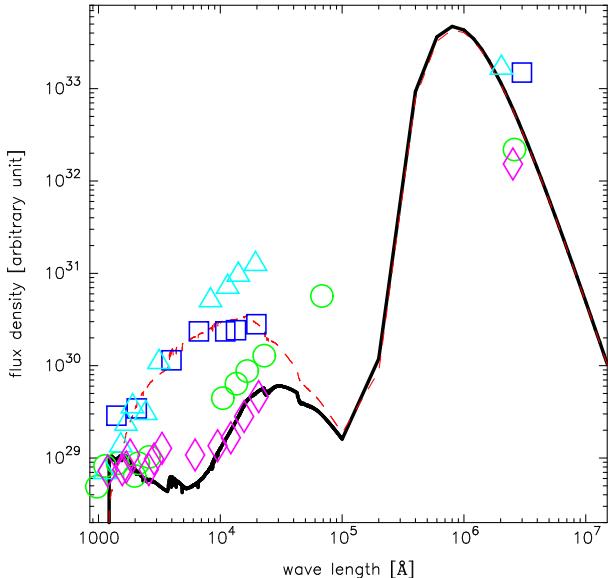
Fig. 13 shows the stellar mass function of our simulated



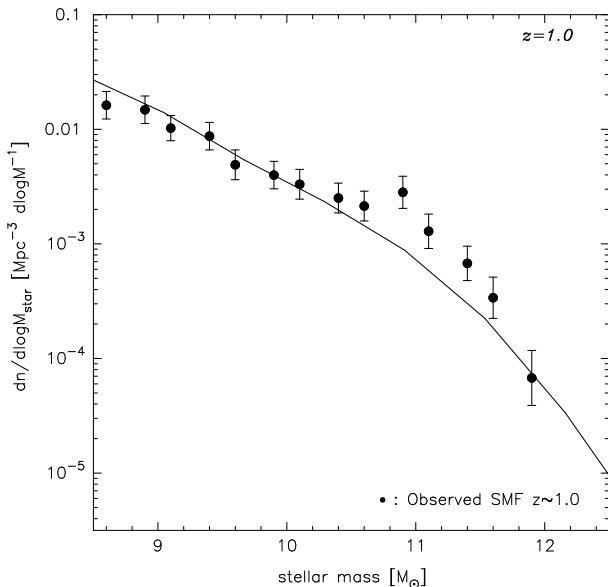
**Figure 11.** SMG number count at 1.1 mm for the two dust extinction models; our model based on a single-sized dust (solid) and the other using the SMC extinction curve (dashed).

galaxies at  $z = 1$ . We also plot the observational data (solid points with error bars) of Mortlock et al. (2011). Our model is in good agreement with the data from the GOODS NICMOS survey as is clearly seen in Fig. 13. This provides yet another support for our galaxy formation model that covers a multiple populations, from LAEs, LBGs to SMGs.

In order to study the present-day stellar mass function, we have run our cosmological simulation down to  $z = 0$ . Our simulation over-predicts the number density of massive galaxies with stellar mass greater than  $10^{11} M_\odot$ . Therefore, although we have successfully constructed a consistent



**Figure 12.** An example SED of one of our simulated SMGs. The thick solid line and thin dashed line represent the dust-extinct SED for our simple dust model and that for the SMC extinction curve, respectively. We also show the observed SEDs for several SMGs (Michałowski, Hjorth & Watson 2010) by open circles, open squares, open triangles and open diamonds. The luminosity of the observed galaxies are arbitrary normalized for this comparison.



**Figure 13.** The stellar mass function at  $z = 1$ . The solid line is our simulation result. The points with error bars show the observational result from the GOODS NICMOS survey (Mortlock et al. 2011).

model for SMGs and UV-selected galaxies at high redshifts, it appears that our galaxy formation model lacks some physical mechanism(s) that shapes the present-day stellar mass function. As an attempt, we have implemented a model where gas cooling is quenched in large galaxies whose halos'

velocity dispersions are greater than 141 km/sec<sup>1</sup>. Then the resulting stellar mass function closely agrees with the observational data at  $z = 0$  derived by Li & White (2009). We argue that the kind of strong feedback, in terms of star formation efficiency, may be needed for viable galaxy formation models, in order for them to reproduce all the available data from low to high redshifts. Constructing such a perfect model is the ultimate goal of the study of galaxy formation, but it is beyond the scope of the present paper. Further studies on the efficiency of star formation in galaxies are clearly needed using multiple approaches.

Together with our previous study on Lyman- $\alpha$  emitters at  $z = 3.1$  (Shimizu, Yoshida & Okamoto 2011), we now provide a unified galaxy population model within the standard  $\Lambda$  Cold Dark Matter cosmology, which reproduces simultaneously the statistical properties of UV-selected star-forming galaxies and sub-millimetre galaxies at high redshifts.

## ACKNOWLEDGMENTS

The authors are grateful to B. Hatsukade for providing their observational data. IS would like to thank B. Hatsukade, Y. Tamura, T. Nozawa, K. Nagamine and M. Kobayashi for stimulating discussion. Numerical simulations have been performed with the EUP, PRIMO and SGI cluster system installed at the Institute for the Physics and Mathematics of the Universe, University of Tokyo. This work is partially supported by Grant-in-Aid for Young Scientists (S) (20674003) and by the FIRST program Subaru Measurements of Images and Redshifts (SuMIRe) by the Council for Science and Technology Policy. TO acknowledges the financial support of Grant-in-Aid for Young Scientists (B: 24740112) and by MEXT HPCI STRATEGIC PROGRAM.

## REFERENCES

- Alexander, D. M., et al., 2003, AJ, 125, 383
- Arnouts S., et al., 2005, ApJ, 619, L43
- Austermann, J. E., et al. 2009, MNRAS, 393, 1573
- Austermann, J. E., et al. 2010, MNRAS, 401, 160
- Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
- Beelen A., et al. 2008, A&A, 485, 645
- Blain, A. W., Chapman, S. C., Smail, I., Ivison, R. 2004, ApJ, 611, 725
- Capak P. et al., 2008, ApJ, 681, L53
- Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 722
- Cooray, A., et al. 2010, A&A, 518, L22
- Coppin K. E. K. et al., 2009, MNRAS, 395, 1905
- Daddi E. et al., 2009, ApJ, 694, 1517

<sup>1</sup> We have done a few test calculations by varying the threshold velocity dispersion and have found that the run with  $\sigma_{\text{th}} = 141$  km/sec reproduces the break of the observed  $z = 0$  stellar mass function at  $M_* \sim 10^{11} M_\odot$ .

- Davé, R., Finlator, K., Oppenheimer, B. D., Fardal, M., Katz, N., Kereš, D., Weinberg, D. H. 2010, MNRAS, 404, 1355
- Dayal P., Hirashita H., Ferrara A., 2010, MNRAS, 403, 620
- Draine, B. T., Lee, H. M., 1984, ApJ, 285, 89
- Draine B. T. et al., 2007, ApJ, 663, 866
- Fontanot F., Monaco P., Silva L., Grazian A., 2007, MNRAS, 382, 903
- Geach, J. E., et al. 2005, MNRAS, 363, 1398
- Geach, J. E., et al. 2007, ApJ, 655, L9
- Greve et al. 2004, MNRAS, 354, 779
- Greve T. R. et al., 2005, MNRAS, 359, 1165
- Hatsukade, B., et al. 2011, MNRAS, 411, 102
- Hickox, C. R., et al. 2012, MNRAS, 421, 284
- Hirashita, H., Ferrara, A., 2002, MNRAS, 337, 92
- Hirashita H., Nozawa T., Takeuchi T. T., Kozasa T., 2008, MNRAS, 384, 1725
- Hirashita H., Nozawa T., Yan H., Kozasa T., 2010, MNRAS, 404, 1437
- Knudsen K. K., Kneib J.-P., Richard J., Petitpas G., Egami E., 2010, ApJ, 709, 210
- Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., Orsi A., Silva L., Granato G. L., Bressan A., 2010, MNRAS, 405, 2
- Laurent et al. 2005, ApJ, 623, 742
- Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, ApJ, 460, L1
- Maddox, S. J., et al. 2010, A&A, 518, L11
- Martin C. L., 2005, ApJ, 621, 227
- Matsuda Y., et al. 2007, ApJ, 667, 667
- Michałowski, M., Hjorth, J., Watson, D. 2010a, A&A, 514, A67
- Mori M., Umemura M., 2006, Nature, 440, 644
- Mortlock A., Conselice C. J., Bluck A. F. L., Bauer A. E., Gruetzbauch R., Buitrago F., Ownsworth J., 2011, MNRAS, 413, 2845
- Nagamine K., Springel V., Hernquist L., 2004, MNRAS, 348, 421
- Nozawa T., Kozasa T., Umeda H., Maeda K., Nomoto K., 2003, ApJ, 598, 785
- Oesch, P. A. et al., 2010, ApJ, 725, L150
- Okamoto T., Eke V. R., Frenk C. S., Jenkins A., 2005, MNRAS, 363, 1299
- Okamoto, T., Frenk, C. S., Jenkins, A., Theuns, T. 2010, MNRAS, 406, 208
- Okamoto T., Nemmen R. S., Bower R. G., 2008b, MNRAS, 385, 161
- Okamoto T., Frenk C. S., 2009, MNRAS, 399, L174
- Perera et al. 2008, MNRAS, 391, 1227
- Riechers D. A. et al., 2010, ApJ, 720, L131
- Sawicki M., Thompson D., 2006, ApJ, 642, 653
- Scott, S. E., Dunlop, J. S., Serjeant, S. 2006, MNRAS, 370, 1057
- Scott, K. S., et al. 2010, MNRAS, 405, 2260
- Shimizu, I., Umemura, M., Yonehara, A. 2007, MNRAS, 380, 49L
- Shimizu, I., Umemura, M., 2010, MNRAS, 406, 913
- Shimizu I., Yoshida N., Okamoto T., 2011, arXiv:1102.1509
- Shimizu T., Kawara K., Sameshima H., Ienaka N., Nozawa T., Kozasa T., 2011, MNRAS, 418, 625
- Smail, Ian., Chapman, S. C., Blain, A. W., Ivison, R. J., ApJ, 616, 71
- Smail I., Chapman S. C., Blain A. W., Ivison R. J., 2004, ApJ, 616, 71
- Spergel, D. N., et al., 2003, ApJS, 148, 175
- Springel V., Hernquist L., 2003a, MNRAS, 339, 289, MNRAS, 337, 92
- Springel V., 2005, MNRAS, 364, 1105
- Swinbank A. M., Smail I., Chapman S. C., Blain A. W., Ivison R. J., Keel W. C., 2004, ApJ, 617, 64
- Tacconi L. J. et al., 2006, ApJ, 640, 228
- Tamura et al., 2009, Nature, 459, 61
- Tamura Y., et al. 2010, ApJ, 724, 1270
- Todini P., Ferarra A., 2001, MNRAS, 325, 726
- Wardlow, J. L. et al., 2011, MNRAS, 415, 1479
- Webb, T. M., et al. 2003, ApJ, 582, 6
- Weiß, A., et al. 2009, ApJ, 707, 1201
- Williams, C. C., et al. 2011, ApJ, 733, 92
- Yajima H., Umemura M., Mori M., 2011, arXiv:1111.6680
- Yoshida M., Shimasaku K., Ouchi M., Sekiguchi K., Furusawa H., Okamura S. 2008, ApJ, 679, 269
- Yun M. S. et al., 2011, arXiv:1109.6286

This paper has been typeset from a TeX/ L<sup>A</sup>T<sub>E</sub>X file prepared by the author.