Bayesian inferences of galaxy formation from the \(K\)-band luminosity and \(\text{H}\,\text{I}\) mass functions of galaxies: constraining star formation and feedback

Yu Lu,\(^1\)\* H. J. Mo,\(^2\) Zhanhui Lu,\(^2\) Neal Katz\(^2\) and Martin D. Weinberg\(^2\)

\(^1\)Physics Department, Kavli Institute for Particle Astrophysics and Cosmology, and SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94305, USA
\(^2\)Department of Astronomy, University of Massachusetts, Amherst, MA 01003-9305, USA

Accepted 2014 June 12. Received 2014 May 29; in original form 2013 October 31

ABSTRACT

We infer mechanisms of galaxy formation for a broad family of semi-analytic models (SAMs) constrained by the \(K\)-band luminosity function and \(\text{H}\,\text{I}\) mass function of local galaxies using tools of Bayesian analysis. Even with a broad search in parameter space the whole model family fails to match to constraining data. In the best-fitting models, the star formation and feedback parameters in low-mass haloes are tightly constrained by the two data sets, and the analysis reveals several generic failures of models that similarly apply to other existing SAMs. First, based on the assumption that baryon accretion follows the dark matter accretion, large mass-loading factors are required for haloes with circular velocities lower than 200 km s\(^{-1}\), and most of the wind mass must be expelled from the haloes. Second, assuming that the feedback is powered by Type II supernovae with a Chabrier initial mass function, the outflow requires more than 25 per cent of the available supernova kinetic energy. Finally, the posterior predictive distributions for the star formation history are dramatically inconsistent with observations for masses similar to or smaller than the Milky Way mass. The inferences suggest that the current model family is still missing some key physical processes that regulate the gas accretion and star formation in galaxies with masses below that of the Milky Way.

Key words: methods: numerical – methods: statistical – galaxies: evolution – galaxies: formation – galaxies: luminosity function, mass function.

1 INTRODUCTION

The formation and evolution of galaxies is regulated by mass accretion, star formation, and associated physical processes, such as feedback from supernova (SN) explosions. Unravelling this physical complexity is one of the great challenges in modern astrophysics (see Mo, van den Bosch & White 2010, for an overview). An important hurdle for any theoretical model of galaxy formation is a natural explanation for the differing shapes of the dark matter halo mass function and the galaxy luminosity and stellar mass function. The mass function of dark matter haloes, \(n(M)\), scales with halo mass as \(n(M) \propto M^{-2}\) at the low-mass end, in contrast to the observed luminosity function of galaxies at low \(z\), \(\Phi(L)\), which has a shallow faint-end slope: \(\Phi(L) \propto L^{-1}\). This suggests that star formation in low-mass haloes must be inefficient, but the physical cause remains unexplained. Similarly, the dependence of the baryonic mass fraction on halo mass requires explanation. One might expect that each halo has a baryonic mass fraction that is close to the universal value of 17 per cent. In reality, the total inventory of baryons in Milky Way-sized haloes is only 5–10 per cent of the halo mass, and the baryon mass fraction decreases rapidly with decreasing halo mass at the low-mass end (e.g. Yang, Mo & van den Bosch 2003; Papastergis et al. 2012). Therefore, some processes must keep a large fraction of baryons outside of galaxies.

The process most often considered to suppress star formation in low-mass haloes is SN feedback; the total amount of energy released by SNe can be larger than the binding energy of the gas in low-mass haloes (Dekel & Silk 1986; White & Zaritsky 1992). Therefore, it is energetically possible to expel large amounts of baryonic matter from low-mass haloes and, thereby, reduce the efficiency of star formation. Feedback may do this in multiple ways. For example, the feedback energy may heat a fraction of the cold interstellar medium (ISM) in the disc, either causing it to mix with hot halo gas or ejecting it directly without mixing (Somerville & Primack 1999; Benson et al. 2003). The feedback energy may also interact with the halo gas, heat it, and then eject it from the halo (Croton et al. 2006). Alternatively, galactic winds may be driven by radiation pressure associated with massive stars and SN explosions (e.g. Murray, Quataert & Thompson 2005). The total amount of

\*E-mail: luyu@stanford.edu
energy in radiation is two orders of magnitude larger than the available kinetic energy in SN explosions, and so the energy budget can be increased significantly. Some simulations of isolated galaxies demonstrate that radiation pressure-driven (also called momentum-driven) winds are effective at ejecting cold gas from galaxies (e.g. Hopkins, Quataert & Murray 2011; Hopkins et al. 2013), but other simulations question their efficiency (e.g. Krumholz & Thompson 2012). Furthermore, cosmological simulations that employ simplified versions of the momentum-driven model have not successfully reproduced the observed galaxy mass function (e.g. Oppenheimer & Dave 2006; Puchwein et al. 2012), although in the latest simulations the discrepancies are predominantly at high masses where SN winds may not be the dominant feedback mechanism (Davé et al. 2013). In summary (1) none of the feedback processes is understood in detail; and (2) the effect of feedback on the star formation efficiency is poorly determined; the efficiency parameter adopted in the models ranges from a few per cent to about 100 per cent (Bower, Benson & Crain 2012; Mutch, Poole & Croton 2013), with a large range of uncertainty.

Observationally, significant outflows are observed only for starburst galaxies, and there is no clear evidence for massive winds from normal spiral galaxies (e.g. Heckman 2003). The mass loading and velocity of galactic winds are poorly constrained (Martin 2005; Weiner et al. 2009; Chen et al. 2010). Martin (2006) estimated that mass outflow rates (OFRs) in the cold component of the wind is on the order of 10 per cent of the star formation rate (SFR) in the starburst, while other observations suggest that the rate may be as high as or even higher than the SFR of the galaxy (Martin, Kobulnicky & Heckman 2002).

The feedback processes described above assume that baryons accrete into a halo at the cosmic baryon fraction and that a large fraction of the cooled baryons are subsequently ejected by feedback. Most semi-analytic models (SAMs) have focused on gas ejection through winds and we will follow suit here. Owing to the uncertainty in the physical details, we use the SAM developed by Lu et al. (2011) that both incorporates sufficiently general parametrizations of galaxy formation mechanisms so that its results will apply to a large class of SAMs and uses advanced Markov chain Monte Carlo (MCMC) techniques to effectively explore the high-dimensional parameter space. Similar approaches have been adopted by Kampakoglou, Trotta & Silk (2008), Henriques et al. (2009, 2013), and Mutch et al. (2013). In addition, Gaussian process model emulators have also been adopted to explore the parameter space of SAMs by Bower et al. (2010) and Gómez et al. (2012). More generally, since the Bayesian inference approach places the model on a probabilistic footing, comparisons between models, probabilities within models, and the consistency of the models given the data (i.e. goodness of fit) can all be assessed quantitatively.

In Lu et al. (2012), we applied the Bayesian SAM to make model inferences conditional on the observed K-band luminosity function of local galaxies. We found that star formation and SN feedback are degenerate when one only uses the luminosity function to constrain the model. The model implied two modes for low-mass haloes: one where the star formation efficiency is increasingly lower for lower halo masses and one where the SN feedback is increasingly effective in reheating the cold gas. The dominating low star formation mode hides a large fraction of baryonic mass in the cold gas phase, thereby vastly overpredicting the H I mass function of galaxies. In this paper, we use both the K-band luminosity function and the H I mass function of galaxies to constrain a model family based on Lu et al. (2012), with additional feedback processes for ejecting baryons out of haloes, to see if there are interesting regions in the extended parameter space that can explain both data sets simultaneously. We choose the K-band luminosity function because it is less affected by dust extinction than are optical bands. The cold gas component, although a relatively small fraction of the total baryon gas (e.g. Bell et al. 2003b; Keres, Yun & Young 2003; Giovanelli et al. 2005; Zwaan et al. 2005), depends crucially on gas cooling and accretion, star formation, and feedback (Lu et al. 2012; Wang, Weinmann & Neistein 2012) and thereby complements the information from the stellar component. We first use our SAM with 17 free parameters to obtain the posterior distribution from the constraining data. Our philosophy in this paper is to take a broad prior range for the parameters, a range that is sometimes broader than a conventional interpretation of observations and theory would allow, but one that is broad enough to contain values assumed in published SAMs. We find that using these very broad ranges for the priors of the model parameters, our posterior distribution recovers many parameter combinations of previously published SAMs, but the model is still unable to match both of the constraining data sets simultaneously. If we restrict our prior ranges to those suggested by other observations (and theory), then the mismatch with the constraining observation becomes much worse, as we discuss later. We then analyze the posterior distribution of the key parameters that govern star formation and feedback in an attempt to gain insight into the underlying physical processes and to investigate the shortcomings of the model. Finally, we use the constrained posterior to make predictions and compare them with existing data to further test the models.

The paper is organized as follows. Section 2 describes the extended model based on Lu et al. (2011, 2012). We present the data, its error model, and the definition of the likelihood function for the Bayesian inference in Section 3. Section 4 presents the results of our model inference using the constrained posterior distribution of model parameters (Section 4.1) and model predictions (Section 4.2). Finally, we discuss our results in Section 5. Throughout the paper, we assume a Λ cold dark matter (ΛCDM) cosmology with ΩM,0 = 0.26, Ωb,0 = 0.044, h = 0.71, n = 0.96, and σ8 = 0.79, which are consistent with the 5-year Wilkinson Microwave Anisotropy Probe (WMAP5) data (Dunkley et al. 2009; Komatsu et al. 2009).

2 THE SEMI-ANALYTIC MODEL

We adopt the SAM developed by Lu et al. (2011; see also Lu et al. 2012), which includes star formation, SN feedback, galaxy mergers, and active galactic nuclei (AGN) feedback. The processes are parametrized with standard analytic functions so that the model encompasses a large range of previously used prescriptions. We only describe the parametrizations relevant to SN feedback and star formation in this section; see Lu et al. (2011) for a comprehensive description of the other details.

Our model assumes that quiescent star formation occurs in a high-density, exponentially distributed cold gas disc. The characteristic radius for the exponential disc, rdisc, is related to the virial radius and the spin parameter of the host halo as follows:

\[ r_{\text{disc}} = \lambda r_{\text{vir}} / \sqrt{2} \]

where \( \lambda \) is the spin parameter of the halo (Mo, Mao & White 1998). We choose a single value \( \lambda = 0.035 \) for all haloes; this value (e.g. Kennicutt 1998; Kennicutt et al. 2007; Bigiel et al. 2008). For an exponential disc, the cold
gas mass with a surface density higher than the threshold surface density $\Sigma_{SF}$ is

$$m_{sd} = m_{\text{cold}} \left[ 1 - \left( 1 + \ln \frac{m_{\text{cold}}}{2\pi f_{\text{SF}} r_{\text{disc},0}^2} \right) \frac{2\pi f_{\text{SF}} r_{\text{disc},0}^2}{m_{\text{cold}}} \right],$$  

(1)

where $m_{\text{cold}}$ is the total cold gas mass in the galaxy, and $f_{\text{SF}}$ is a model parameter defined as

$$f_{\text{SF}} = \left( \frac{\lambda}{0.035} \right)^2 \Sigma_{SF}.$$  

(2)

The model parameter $f_{\text{SF}}$ describes both the size of the disc and the threshold surface density of the cold gas. Thus, the parameters $\lambda$ and $\Sigma_{SF}$ are intrinsically degenerate through the combined parameter $f_{\text{SF}}$. If $f_{\text{SF}}$ is significantly lower than $10 M_\odot$ pc$^{-2}$, then either the cold gas disc is more compact than the fiducial size, or the threshold surface density is required to be lower than $10 M_\odot$ pc$^{-2}$. We assign a generous prior for the parameter $f_{\text{SF}}$ to accommodate the broad uncertainties in the implementation of star formation in the model (see Table 1), and we test the sensitivity of our results on the choice of the prior, which will be discussed in Section 4.1.

We assume that the SFR, $\phi$, is proportional to the mass of star-forming gas and inversely proportional to the dynamical time-scale of the disc, $t_{\text{disc}} = r_{\text{disc}}/V_{\text{vir}}$:

$$\phi = \epsilon_{sd} \frac{m_{sd}}{t_{\text{disc}}}.$$  

(3)

where $\epsilon_{sd}$ is an overall star formation efficiency factor. To further generalize the dependence of the SFR on halo properties, the model further assumes that $\epsilon_{sd}$ has a broken power-law dependence on the circular velocity of the host halo:

$$\epsilon_{sd} = \left\{ \begin{array}{ll}
\alpha_{SF} \quad & V_{\text{vir}} \geq V_{\text{SF}}, \\
\alpha_{SF} \left( \frac{V_{\text{vir}}}{V_{\text{SF}}} \right)^{\beta_{SF}} & V_{\text{vir}} < V_{\text{SF}}, 
\end{array} \right.$$  

(4)

where $\alpha_{SF}$, $\beta_{SF}$, and $V_{\text{SF}}$ characterize the star formation efficiency. For haloes with a circular velocity higher than the critical velocity scale, $V_{\text{SF}}$, the star formation efficiency is a constant, $\alpha_{SF}$; for haloes with a circular velocity lower than $V_{\text{SF}}$, the star formation efficiency follows a power-law function of halo circular velocity defined by $\alpha_{SF}$ and $\beta_{SF}$. The pivot point $V_{\text{SF}}$ in the parametrization is needed to avoid unrealistically high star formation efficiencies in high $V_{\text{vir}}$ haloes when $\beta_{SF}$ is positive. However, we note that it introduces an intrinsic degeneracy into the parametrization. For example, when $\beta_{SF} = 0$, $V_{\text{SF}}$ does not have any impact on the model. This affects our study, as we discuss in Section 4.1. We note that allowing $\beta_{SF}$ to have non-zero values conflicts with some observational results, which find that the SFR is largely determined by the available supply of cold dense gas and not on the mass (or circular velocity) of the host halo (Schmidt 1959; Kennicutt 1989).

Star formation in our SAM can both consume cold gas by forming stars and affect the cold gas supply for subsequent star formation through feedback. In general, star formation feedback can influence the cold gas in three different ways: first, feedback can heat up the cold gas in the disc so that it joins the hot halo; second, feedback can directly eject cold gas out of the gravitational potential of the host halo; and third, feedback can expel some of the hot halo gas out of the host halo. Our model attempts to capture all these processes. In our previous studies (Lu et al. 2011), we did not explore the dependence on cold gas ejection owing to its strong degeneracy with other mechanisms when the luminosity function or stellar mass function alone is used as a constraint. The addition of the H i mass function breaks this degeneracy, and we will see that cold gas ejection is required by the H i data. Our previous papers enforced conservation of the mechanical energy produced by SN explosions for the chosen stellar initial mass function (IMF) when modelling the mass loading in galaxy winds. In this paper, we allow the possibility of other energy sources, e.g. radiatively driven winds.

We parametrize the mass-loss rate of the cold gas to be proportional to the SFR, $\phi$, as

$$m_{out} = \alpha_{ld} \phi,$$  

(5)

where the coefficient $\alpha_{ld}$ is the loading factor of the outflow. Because the quantity of energy or momentum produced during stellar evolution is finite, the mass loading depends on the initial wind velocity and the physics of the coupling with the multiphase ISM. Two types of mass-loading models have been extensively explored in the literature: (1) an energy-driven wind that couples a fixed fraction of the feedback energy with the wind, which motivates $\alpha_{ld} \propto v_c^{-2}$ (e.g. Okamoto et al. 2010); (2) a momentum-driven wind that couples a fixed fraction of momentum (e.g. in a radiation field) with the

---

Table 1. Model parameters.

| # | Parameter | Meaning | Prior |
|---|-----------|---------|-------|
| 1 | $\log M_{CC}$ (M$_\odot$) | Cooling cut-off halo mass | [1.5, 4.5] |
| 2 | $\log \alpha_{SF}$ | Star formation efficiency power-law amplitude | [−3, 0] |
| 3 | $\beta_{SF}$ | Star formation efficiency power-law index | [−1, 12] |
| 4 | $\log V_{\text{SF}}$ (km s$^{-1}$) | Star formation law turn-over halo circular velocity | [1.5, 3.0] |
| 5 | $\log f_{\text{SF}}$ | Star formation threshold gas surface density | [0, 1.2] |
| 6 | $\log \alpha_{SN}$ | SN feedback energy fraction | [−3, 1] |
| 7 | $\log \alpha_{LD}$ | SN feedback reheating power-law amplitude | [−3, 2] |
| 8 | $\beta_{LD}$ | SN feedback reheating power-law index | [0, 14] |
| 9 | $\log V_{\text{EJ}}$ | SN feedback ejection pivot halo circular velocity | [1.6, 3.0] |
| 10 | $\log \epsilon_{\text{EJ}}$ | SN feedback ejection power-law index | [0, 10] |
| 11 | $\epsilon_{\text{EJ}}$ | SN feedback ejection escaping velocity factor | [1, 5] |
| 12 | $\log f_{\text{W}}$ | Fraction of surplus SN feedback energy used for powering wind | [−3, 0] |
| 13 | $\log f_{\text{RI}}$ | Fraction of re-infall ejected hot gas | [−2, 0] |
| 14 | $\log f_{\text{IN}}$ | Merging time-scale in dynamical friction time-scale | [0, 2] |
| 15 | $\log \alpha_{SB}$ | Merger triggered starburst efficiency power-law amplitude | [−2, 0] |
| 16 | $\beta_{SB}$ | Merger triggered starburst efficiency power-law index | [0, 2] |
| 17 | \(\arctan(\sigma_{\text{IN}})\) | K-band luminosity function faint-end incompleteness | [0, 0.177] |
wind, which motivates $\alpha_{\rm LD} \propto v^{-1}$ (e.g. Oppenheimer & Dave 2006; Oppenheimer et al. 2010; Hopkins et al. 2013). Attempts to observe $v_\sigma$ and infer its dependence on galaxy properties (e.g. Martin 1999) have been inconclusive. Some theoretical models assume a constant wind velocity (e.g. Croton et al. 2006), while others assume a halo-mass-dependent velocity (e.g. Puchwein et al. 2012). To parametrize the various possibilities, we assume the loading factor to be a power-law function of halo circular velocity, $V_{\rm vir}$,

$$\alpha_{\rm LD} = \alpha_{\rm LD} \left( \frac{V_{\rm vir}}{V_0} \right)^{-\beta_{\rm LD}},$$

where the power index $\beta_{\rm LD}$ and the normalization $\alpha_{\rm LD}$ are model parameters, and $V_0$ is an arbitrary velocity scale that we fix at $220 \, \text{km} \, \text{s}^{-1}$.

In this model, a given value of $\beta_{\rm LD}$ does not determine a particular wind launching mechanism (energy driven or momentum driven). For example, $\beta_{\rm LD} = 0$ implies independence of $V_{\rm vir}$, but this does not affect the $\alpha_{\rm LD}$-$v_\sigma$ relation predicted by a particular wind driving mechanism, because the $v_\sigma$-$V_{\rm vir}$ relation is undetermined. Only in the special case when $v_\sigma \propto V_{\rm vir}$ is the value of $\beta_{\rm LD}$ related to the wind driving mechanism: $\beta_{\rm LD} = 2$ for an energy-driven wind and $\beta_{\rm LD} = 1$ for a momentum-driven wind. Because of all the uncertainties, we assign a large prior range for $\beta_{\rm LD}$, allowing $\beta_{\rm LD}$ to take any value between 0 and 8 in our inference.

After the gas is removed from the galaxy, it can be trapped within the potential well of the dark matter halo or ejected from the halo into the intergalactic medium (IGM). We assume that the fraction of removed gas that is ejected from the halo is

$$f_{\text{ej}} = \left[ 1 + \left( \frac{V_{\text{vir}}}{V_{\text{ej}}} \right)^{\beta_{\text{ej}}} \right]^{-1},$$

where $V_{\text{ej}}$ and $\beta_{\text{ej}}$ are free parameters. Thus, for a halo with a circular velocity lower than $V_{\text{ej}}$, most of the outflow gas is ejected from the halo, while for haloes with circular velocities much larger than $V_{\text{ej}}$, the ejected fraction follows a power-law function of the halo circular velocity. Our model also includes an outflow of hot halo gas. The strength of the effect is controlled by the parameter $\alpha_{\text{SN}}$, which describes the fraction of the total kinetic energy released by Type II SN that is needed to power all the feedback processes in the model. The excess energy over the amount used for reheating and ejecting the cold gas is used to power an outflow of hot halo gas. The amount of hot gas mass expelled is written as

$$\Delta m_{\text{wind}} = \epsilon_W \left\{ \alpha_{\text{SN}} \frac{V_{\text{SN}}}{V_{\text{esc}}} \right\} \rho_{\text{ld}} \left[ \left( \frac{V_{\text{vir}}}{V_{\text{esc}}} \right)^2 + f_{\text{ej}} \right] \phi \Delta t,$$

where $V_{\text{esc}}$ is the escape velocity of the halo. For a Navarro–Frenk–White (NFW) halo with a concentration $c$ (Navarro, Frenk & White 1996), $V_{\text{esc}} = V_{\text{vir}} \sqrt{\frac{c+1}{c+15}}$ (Puchwein & Springel 2013).

The model assumes that the ejected gas can re-collapse into the halo at later times as hot halo gas. As in Springel et al. (2001) and De Lucia & Blaizot (2007), the rate of re-infall of ejected gas is given by

$$m_{\text{e-re-infall}} = f_{\text{ej}} M_{\text{ej}} \left( \frac{M_{\text{ej}}}{\tau_{\text{dyn}}} \right).$$

Here, $f_{\text{ej}}$ is a free parameter, $M_{\text{ej}}$ is the total mass of gas in the ‘ejected’ reservoir, including both ejected gas from the cold gas disc and expelled gas from the hot halo, and $\tau_{\text{dyn}} = r_{\text{vir}}/V_{\text{vir}}$ is the dynamical time of the halo.

To convert the stellar mass to stellar light, we apply the stellar population synthesis (SPS) model of Bruzual (2007, BC07), which includes an improved treatment of thermally pulsing asymptotic giant branch (AGB) stars. The SPS model is used to predict the $K$-band luminosity of a galaxy from its star formation history. To implement the SPS model, we use a Chabrier IMF (Chabrier 2003), and create a look-up table for the luminosities on a grid with 220 values of age evenly spaced from 0 to 15 Gyr, and with $Z = 0.0001, 0.0004, 0.0008, 0.02$, and 0.05.

Finally, our SAM uses Monte-Carlo-generated halo merger trees tuned to match the conditional mass functions found in $N$-body simulations following Parkinson, Cole & Helly (2008) with final masses in the range $10^9$ to $10^{15} h^{-1} M_{\odot}$ (see Lu et al. 2011, for additional details). Since haloes and their merger trees are randomly sampled from the halo mass function, a model prediction based on a finite merger tree sample suffers from sampling variance. We, therefore, choose the number of halo merger trees in each mass bin such that the standard deviation in the predictions, namely the $K$-band luminosity function and the $H\alpha$ mass function, are at least two times smaller than the error in the observational data in all the bins. Specifically, we use 1000 merger trees for haloes with present masses in the range $10^{11} - 10^{13} h^{-1} M_{\odot}$, 1500 in $10^{12.5} - 10^{13.5} h^{-1} M_{\odot}$, 400 in $10^9 - 10^{10} h^{-1} M_{\odot}$, 400 in $10^{10} - 10^{11} h^{-1} M_{\odot}$, and 100 in $10^{13.5} - 10^{14} h^{-1} M_{\odot}$. The logarithmic halo masses are evenly distributed in each range. Since massive haloes are rare in the assumed cosmology, their contribution to scatter in the stellar mass function is negligible. We choose a mass resolution for the merger trees that varies with the final halo mass. For haloes with final masses smaller than $10^{10} h^{-1} M_{\odot}$, the mass resolution is $10^{0.5} h^{-1} M_{\odot}$; for haloes with final mass between $10^{10}$ and $10^{12} h^{-1} M_{\odot}$, the mass resolution is $10^{0.2} h^{-1} M_{\odot}$; for haloes with final mass between $10^{12}$ and $10^{14} h^{-1} M_{\odot}$, the mass resolution is $10^{0.1} h^{-1} M_{\odot}$; and for haloes with final masses larger than $10^{14} h^{-1} M_{\odot}$, it is $10^{0.0} h^{-1} M_{\odot}$. All merger trees are sampled at 100 redshifts equally spaced in log$(1 + z)$ from $z = 7$ to 0. To make predictions for the total galaxy population we weight each predicted galaxy by the halo mass function of Sheth, Mo & Tormen (2001). The merger tree set we adopt allows us to run each model rapidly with sufficient accuracy so that we can explore the parameter space within a reasonable amount of time. After we obtained the posterior, we checked our model predictions using a larger number of trees and higher time and mass resolutions. The deviations of the results from those obtained with the smaller tree sample and lower resolutions described above were found to be well within the $1\sigma$ range of the posterior predictive distributions, demonstrating that the number of trees and mass resolution that we adopt are adequate for our purposes.

### 3 DATA AND LIKELIHOOD

Our SAM inference is conditional on both the $K$-band luminosity function and $H\alpha$ mass function of galaxies to constrain models. We use the $K$-band luminosity function from Bell et al. (2003a); the details of the data and how it is used to constrain the model can be found in Lu et al. (2012). Different from the luminosity function, whose errors can be safely approximated by Poisson errors, the $H\alpha$ mass function contains errors that are correlated across different bins. Here, we give a detailed description for our use of the $H\alpha$ mass function from Zwaan et al. (2005), derived from the HI Parkes All Sky Survey (HIPASS) Bright Galaxy Catalogue (hereafter HIPASS BGC); see Koribalski et al. (2004), and show how we derive a full covariance matrix for the likelihood function utilized in our Bayesian analysis.
3.1 Uncertainties and covariance in the observed H\(_i\) mass function

A galaxy’s H\(_i\) mass is estimated from its 21-cm flux density and distance as
\[ M_{\text{HI}} = 2.36 \times 10^5 d^2 F_{\text{HI}} M_\odot, \]
where \( d \) is the distance of the source in Mpc, and \( F_{\text{HI}} \) is the integrated 21-cm flux density in Janskys \( \text{km}^{-1} \text{s}^{-1} \). The value of \( F_{\text{HI}} \) is provided for each galaxy in the catalogue, along with its uncertainty, \( \sigma(F_{\text{HI}}) \). We use the Hubble distance,
\[ d = v_{\text{LG}} / H_0, \]
where \( H_0 \) is the Hubble constant in units of \( \text{km} \text{s}^{-1} \text{Mpc}^{-1} \), \( v_{\text{LG}} \) is the recession velocity of the source, and \( v_{\text{sys}} \), corrected to the Local Group rest frame:
\[ v_{\text{LG}} = v_{\text{sys}} + 300 \sin l \cos b, \]
where \( l \) and \( b \) are the Galactic longitude and latitude of the source, respectively. The observational error in \( v_{\text{sys}} \) is provided in the source catalogue and is used in our analysis.

The detectability of a H\(_i\) galaxy is affected by both its peak flux density, \( S_p \), and its linewidth, \( v_{20} \), defined to be the wavelength difference between the two points where the flux density is 20 per cent of \( S_p \). Following Koribalski et al. (2004), we estimate the uncertainty in the peak flux density as
\[ \sigma(S_p) = \text{rms}^2 + (0.05 S_p)^2, \]
where the rms = 13 mJy. The uncertainty in \( v_{20} \) is estimated as \( \sigma(v_{20}) = 3 \sigma(v_{\text{sys}}) \).

Using these values of \( v_{\text{LG}}, F_{\text{HI}}, S_p, \) and \( v_{20} \) and their standard deviations, we randomly generate 1000 replicas containing 1000 galaxies, each assuming independent Gaussian distributions. For each of the replicas, we adopt the same procedure used by Zwaan et al. (2005) to obtain the H\(_i\) mass function. Specifically, we apply the two-dimensional, stepwise maximum likelihood method proposed by Zwaan et al. (2003) to find the maximum likelihood estimator of the mass function. These ensembles yield the average mass function and the corresponding covariance matrix, \( \Sigma_{\text{obs}} \). This covariance matrix includes both H\(_i\)-measurement variance and the sampling variance owing to the finite survey volume. We will assume that both types of variance are independent and use this covariance matrix to specify the likelihood function in Section 3.3.

3.2 Corrections for molecular fraction and incompleteness

Our SAM predicts a galaxy’s total cold mass, \( M_{\text{cold}} \), which includes atomic H\(_i\) gas, molecular H\(_2\) gas, as well as heavier elements. We make the following assumptions to predict the H\(_i\) mass function. To start, we write
\[ M_{\text{cold}} = M_{\text{HI}} + M_{\text{H_2}}, \]
where \( \beta = 0.74 \) is used to correct for the mass in helium (He) and in the small fraction of heavier elements in a cosmic gas. Given the molecular to atomic hydrogen mass ratio, \( \eta = M_{\text{H_2}} / M_{\text{HI}} \), we have
\[ M_{\text{HI}} = \frac{\beta M_{\text{cold}}}{1 + \eta}. \]

The value of \( \eta \) depends on galaxy properties. Following Obreschkow et al. (2009), we assume that \( \eta \) depends only on the total mass of cold gas:
\[ \log \eta = q + k \log \left( \frac{M_{\text{cold}}}{10^9 h^{-2} M_\odot} \right) + \sigma, \]
where \( q = -0.51^{+0.03}_{-0.04}, k = -0.24^{+0.05}_{-0.05} \), and \( \sigma = 0.30 \). The H\(_i\) mass derived from the observed 21-cm flux is only a fraction of the total H\(_i\) mass owing to self-absorption. Although the details of self-absorption depend on inclination, the average correction expected for a sample of randomly oriented galaxies is about 15 per cent (Zwaan et al. 2003). This gives a relation between the predicted, observed H\(_i\) mass, \( M_{\text{HI}} \), and the real H\(_i\) mass \( M_{\text{HI}} \):
\[ M_{\text{HI}} = (1 - f_{\text{HI}}) M_{\text{HI}} = \frac{1 - f_{\text{HI}}}{1 + \eta} M_{\text{cold}}. \]

We include this effect in our model predictions for the H\(_i\) mass function as \( \Phi^p(M_{\text{HI}}) = \Phi(M_{\text{HI}}) (1 - \zeta) \).

3.3 The likelihood function

For the K-band luminosity function, we use the same likelihood function as described in Lu et al. (2012). For the H\(_i\) mass function, we construct a likelihood function that independently represents the H\(_i\)-measurement error and the sample variance. To start, we neglect the covariance between measurement error and sampling variance. Such covariance may be important for small sample sizes but is unimportant for the current sample where the variance is measurement dominated (Papastergis, private communication). The uncertainties in the process leading to the H\(_i\) mass estimate for each
galaxy (Sections 3.1 and 3.2) induce correlations between the H I mass bins. Ideally, the error model describing this covariance would be used to predict the observed data. However, our galaxy formation model is not capable of predicting all the necessary observables (e.g. $w_{20}$ and $S_{20}$). Instead, we proceed by making two additional assumptions: (1) the likelihood function is a multidimensional normal distribution; and (2) the sampling is a pure Poisson process. Then, the errors in the simulated data presented in Section 3.1 can be modelled as follows. We define the equivalent volume, $V_{\text{eq}}$, implied by the H I measurement process including all the selection effects described in Section 3.2. This equivalent volume can be estimated from the data as $V_{\text{eq}} = N_{\text{obs}}/\Phi_{\text{obs}}$. The Poisson variance in each modelled bin is then $\Phi_{\text{mod}}/V_{\text{eq}}$ and only affects the diagonal terms of the covariance matrix. We, therefore, write the diagonal terms as

$$\sigma_{\text{mod}}^2 = \sigma_{\text{inst}}^2 + \left(\frac{\Phi_{\text{mod}}}{V_{\text{eq}}}\right)^2,$$ (22)

where $\sigma_{\text{inst}}^2$ is some unknown measurement variance. Under the same assumptions, the diagonal elements of the simulated covariance matrix from Section 3.1 are

$$\sigma_{\text{obs}}^2 = \sigma_{\text{inst}}^2 + \left(\frac{\Phi_{\text{obs}}}{V_{\text{eq}}}\right)^2.$$ (23)

We now use equations (22) and (23) to eliminate $\sigma_{\text{inst}}^2$:

$$\sigma_{\text{mod}}^2 = \sigma_{\text{obs}}^2 + \frac{\Phi_{\text{mod}} - \Phi_{\text{obs}}}{V_{\text{eq}}}.$$ (24)

This replaces the variance in the diagonal terms of the simulated covariance matrix to obtain the predicted covariance $\Sigma_{\text{pred}}$. Thus, the likelihood function is written as

$$L(\Phi_{\text{obs}}|\Theta) = \frac{L_{\text{eq}}}{(2\pi)^{L_{\text{eq}}/2}|\det(\Sigma_{\text{pred}})|^{1/2}} \times \exp \left[-\frac{1}{2}(\Phi_{\text{obs}} - \Phi_{\text{mod}})^T \Sigma_{\text{pred}}^{-1} (\Phi_{\text{obs}} - \Phi_{\text{mod}})\right].$$ (25)

where $\Theta$ denotes the model parameter vector, $\Phi_{\text{obs}}$ and $\Phi_{\text{mod}}$ are, respectively, the vectors of the observed and predicted H I mass functions over the H I mass bins, and $I$ is the rank of the covariance matrix. We assume that the K-band luminosity function and the H I mass function are statistically independent and, therefore, the total likelihood is simply the product of the two:

$$L(D|\Theta) = L(D_K|\Theta) \times L(D_{\text{H I}}|\Theta),$$ (26)

where $D_K$ and $D_{\text{H I}}$ denote the observed data of the K-band luminosity function and the H I mass function, respectively.

### 4 MODEL INFERENCE

#### 4.1 Inference from the posterior distribution

We used a tempered version of the differential evolution algorithm (ter Braak 2006) implemented in the Bayesian Inference Engine (BIE; see Weinberg 2013) to sample the posterior density distribution with 256 chains. All simulations run for more than 7000 iterations and the convergence is monitored with the Gelman–Rubin $R$ test (Gelman & Rubin 1992). We declare convergence when $R \lesssim 1.2$. All simulations typically converge after 3000 iterations, and the converged states are used to sample the posterior distribution. After removing 12 outlier chains, we obtain $\sim10^6$ chain states. As the autocorrelation length of the chains is typically $\sim10$, there are about $10^5$ independent chain states to sample the posterior distribution.

Fig. 1 shows the posterior predictions for the K-band luminosity function and the H I mass function of local galaxies. We account for potential incompleteness in the faint end of the K-band luminosity function by marginalizing over the incompleteness parameter $\alpha_{\text{IN}}$ (see Lu et al. 2012, for details). The orange and yellow bands in each panel encompass 66 and 95 per cent credible ranges of the predictions, respectively, while the red lines are the median values.
of the observation should be likely under the distribution of the discrepancy statistics of the posterior models. On the other hand, large discrepancies between the observed data and the posterior predictive distribution indicate that the model performs poorly. Following our previous paper (Lu et al. 2012), we adopt the Bayesian p-value to quantify the goodness of fit, with p defined as

\[ p = \frac{1}{N} \sum_{i=1}^{N} I_{\chi^2_{\text{mod}} \geq \chi^2_{\text{obs}}}, \]

where \( \chi^2_{\text{mod}} \) and \( \chi^2_{\text{obs}} \) are the discrepancy statistics for the model and the data. \( I_q \) is the indication function for the condition \( q \), with \( I_q = 1 \) if \( q \) is true and \( I_q = 0 \) otherwise. In other words, the fraction of posterior model samples with \( \chi^2_{\text{mod}} \geq \chi^2_{\text{obs}} \) is an estimate of \( p \).

A small value of the p reflects the implausibility of the data under the model and, therefore, suggests problems with the model. For the data and posterior distribution shown in Fig. 1, the Bayesian p-value from the PPC for the H1 mass function is \( p_{H1} = 0.25 \), which indicates an adequate fit, but for the K-band luminosity function it is only \( p_K = 0.011 \). This is much lower than the value of \( p_K = 0.66 \) obtained in Lu et al. (2012), which used only the K-band luminosity function as the data constraint. Hence, the model constrained using both the H1 mass function and the K-band luminosity function does not fit the observational data well. In the remainder of this paper, we will explore the causes for this discrepancy and their implication for models of galaxy formation generally.

Fig. 2 shows marginalized posterior distributions for 10 parameters characterizing star formation and feedback and we now describe those with marginal distributions that differ significantly from those in Lu et al. (2012). In the upper left-hand panel of Fig. 2, the vertical axis shows the model parameter \( V_{\text{EJ}} \), the halo circular velocity below which all outflowing mass reheat from the disc must be ejected out of the host halo. The horizontal axis shows the parameter \( M_{\text{CC}} \), the critical halo mass above which radiative cooling of the halo gas is shut off to mimic radio-mode AGN feedback (Bower et al. 2006; Croton et al. 2006). The mass \( M_{\text{CC}} \) is well constrained, with a value about \( 10^{12.5-12.6} M_\odot \). This result also agrees with the value obtained by Lu et al. (2012) using the K-band luminosity function alone as a data constraint, and with other existing SAMs that implement ‘halo quenching’ (e.g. Cattaneo et al. 2006; Somerville et al. 2008). Including the H1 mass function as an additional constraint does not affect the inferred value of \( M_{\text{CC}} \), significantly because \( M_{\text{CC}} \) controls gas cooling and star formation in massive haloes while the H1 mass function mainly constrains low-mass haloes, which contain most of the H1 gas.

The value of \( V_{\text{EJ}} \) is also well constrained, with \( log V_{\text{EJ}} \approx 2.3 \) (\( V_{\text{EJ}} \approx 200 \text{ km s}^{-1} \)), suggesting that all haloes with circular velocities lower than this are required to have efficient mass ejection to fit the data, and the same ejection process can be inefficient in haloes with higher circular velocities. There is also a non-negligible tail to even larger \( V_{\text{EJ}} \); models in this tail require efficient mass ejection at even higher halo circular velocities. This result agrees with both the model of Somerville et al. (2008) and a recent analysis of the Munich model (Henriques et al. 2013): efficient gas ejection from haloes is required to match multiple data sets simultaneously. Without the H1 constraint, the observed K-band luminosity function of galaxies can be reproduced by making the star formation efficiency a strong function of circular velocity, resulting in a massive cold gas component. However, this makes the amplitude of the corresponding predicted H1 mass function an order of magnitude higher than that observed (Lu et al. 2012; Wang et al. 2012). In our model, the strong degeneracy between outflow and star formation efficiency is broken by the constraint provided by the H1 mass function. Consequently, one requires efficient gas ejection from haloes to prevent a large reservoir of cold gas and only requires the star formation efficiency to be a weak function of halo circular velocity (see below). If we fix \( V_{\text{EJ}} = 0 \), we find that the model can no longer match the observational data; the best-fitting model, i.e. in this case the model with the maximum a posteriori (MAP) probability, has a likelihood that is almost 3000 000 times lower than our fiducial MAP. This suggests that efficient gas ejection is unavoidable in our model family.

4.1.1 Star formation

With the H1 mass function as an additional constraint, the posterior probability of the star formation parameters is now constrained to a fairly narrow range (see Panel 2 of Fig. 2). The posterior of log (\( \alpha_{\text{SF}} \)) peaks sharply at \( \sim 1.8 \), implying that a galaxy converts about \( \approx 1.6 \) per cent of its cold gas into stars in every dynamical time. This is consistent with the observed efficiency in galaxies (Elmegreen 1997; Silk 1997). The posterior for \( \beta_{\text{SF}} \) is now peaked near zero, much lower than the value of \( \beta_{\text{SF}} \approx 10 \) obtained in Lu et al. (2012) constrained using the K-band luminosity function alone. This low value for \( \beta_{\text{SF}} \) means that the SFR is largely determined by the available supply of cold dense gas and does not depend on the mass (or circular velocity) of the host halo. This is consistent with observed star formation laws (Schmidt 1959; Kennicutt 1989), in which the SFR depends only on gas properties. However, there still remains a weaker mode – its posterior odds relative to the main mode is 2.5 times smaller – where \( \alpha_{\text{SF}} \approx 2 \) per cent and \( \beta_{\text{SF}} \approx 2.5 \).

We will show that the parameter \( \beta_{\text{SF}} \) is covariant with feedback parameters in another panel of the figure.

In Lu et al. (2012), the star formation efficiency parameter \( \alpha_{\text{SF}} \) was found to be strongly covariant with the threshold of cold gas surface density for star formation, \( f_{\text{SF}} \); a high star formation efficiency with a high threshold value was equivalent to a low star formation efficiency with a low threshold value. As expected, the addition of the H1 mass function breaks this degeneracy. Panel 3 of Fig. 2 shows the marginalized posterior probability distribution in the \( \log f_{\text{SF}} - \log \alpha_{\text{SF}} \) plane. Remember that in our parametrization, \( f_{\text{SF}} \) is equal to the surface gas density threshold for star formation, \( \Sigma_{\text{SF}} \), under the assumption that \( r_{\text{disc}} = r_{\text{disc}, 0} \) i.e. that the disc size is given by the model of Mo et al. (1998). The combined data sets require that the threshold surface density for star formation is small; we find \( \Sigma_{\text{SF}} \approx 1 M_\odot \text{ pc}^{-2} \). This value is about an order of magnitude lower than the observed threshold surface density (Kennicutt 1998; Bigiel et al. 2008; Leroy et al. 2008) and is also right against the prior bound, so even lower values may be preferred by the model, although such low values would be in even greater conflict with observations. In the model, a high gas surface density threshold would result in too much cold gas in low-mass haloes at large radii and too little star formation to power enough outflows. This problem was highlighted by Mo et al. (2005), who found that the standard threshold surface density for star formation would lead to an H1 mass function that was too high, even if each low-mass halo only contained a cold gas disc at the threshold surface density.

Our SAM requires both powerful, efficient feedback to remove the gas from the halo and a threshold gas surface density that is much lower than the observed star formation threshold (and still fails to match the data). To explore explicitly the sensitivity to a higher surface density threshold, we constrained the prior of \( f_{\text{SF}} \) to
Bayesian inference on star formation feedback

Figure 2. The marginalized posterior distributions of 10 model parameters. Panel 1: AGN feedback shuts off gas cooling for haloes with $M_{\text{vir}} > M_{\text{CC}}$ and winds expel gas from haloes with circular velocities $V_{\text{vir}} > V_{EJ}$. Panel 2: $\alpha_{\text{SF}}$ parametrizes the normalization of star formation efficiency, which is a constant for haloes with a circular velocity higher than parameter $V_{\text{SF}}$, and a power-law function of $V_{\text{vir}}$ with $\beta_{\text{SF}}$ being the power-law index of the relation. Panel 3: $f_{\text{SF}}$ is the cold gas surface density threshold for star formation and $\alpha_{\text{SF}}$ is as above. Panel 4: $\alpha_{\text{LD}}$ parametrizes the loading factor of the SN feedback and $\beta_{\text{LD}}$ is the power index of the power-law dependence of the loading factor on the halo circular velocity. The blue lines denote the kinetic energy limit based on Type II SN for outflows in haloes with a circular velocity of 220 km s$^{-1}$ for two IMFs: Salpeter (solid) and Chabrier (dashed). The red lines show the extreme parameter combination of the two mass-loading parameters for haloes with a circular velocity of 100 km s$^{-1}$. Models in regions to the right of the limiting lines require more than 100 per cent of the total available energy for the cold gas ejection. The magenta dotted line denotes models with a loading factor equal to 4 for haloes with a circular velocity of 70 km s$^{-1}$. Panel 5: the joint marginalized posterior distribution in the $\beta_{\text{LD}}$-$V_{\text{SF}}$ plane. Panel 6: the joint marginalized posterior distribution in the $\beta_{\text{LD}}$-$f_{\text{SF}}$ plane. Panel 7: $\alpha_{\text{SN}}$ characterizes the fraction of SN energy that is required to power the ejection of hot halo gas. Panel 8: the joint marginalized posterior distribution in the $\beta_{\text{LD}}$-log$f_{\text{DF}}$ plane, where $f_{\text{DF}}$ is a coefficient characterizing the satellite dynamical friction time-scale. The grey-scale (colour in on-line version) denotes the confidence levels as shown by the colour bar at the top of the figure. Symbols with different colours in the various panels denote the corresponding parameters adopted in some earlier SAMs, as indicated. They do not appear in every panel for all models because some parameters are absent in the parametrization of those models.

range between 3 and $30 M_\odot$ pc$^{-2}$. We find that this greatly reduces the goodness of fit and that the best fits are always obtained with $f_{\text{SF}}$ at the lower boundary of the prior. The MAP using this restricted prior has a likelihood that is smaller than the MAP in the fiducial run by a factor of over 30 000. If we further restrict $f_{\text{SF}}$ to match the observed value of $10 M_\odot$ pc$^{-2}$, the likelihood decreases by a further factor of over 60 000 000, and the $p$-value for the $H_1$ mass function becomes very small, $p_{H_1} < 0.005$, indicating an unacceptable fit.
Thus, the current model family cannot explain the observed H\textsc{i} mass function using the observed gas density threshold for star formation.

4.1.2 Feedback

Panel 4 of Fig. 2 shows the marginalized posterior probability distribution of the two SN gas ejection parameters: $\alpha_{\text{LD}}$ and $\beta_{\text{LD}}$. For comparison, we also plot the combinations of these parameters used in a number of existing SAMs. These models include Cole et al. (1994), Kauffmann & Charlot (1998), Somerville & Primack (1999), Somerville et al. (2008), De Lucia, Kauffmann & White (2004), Kang et al. (2005), Menci et al. (2005), Cattaneo et al. (2006), and Guo et al. (2011). The parameter choices are converted into our parametrization, where the amplitude of the mass-loading factor is normalized at a halo circular velocity scale of 220 km s$^{-1}$. As expected, because of the similarity in our phenomenological parametrizations, most of the other SAMs are located in the high-probability region of our inferred posterior. These two parameters are covariant and populate three main modes: $\beta_{\text{LD}} \approx 0$ that is equivalent to the Croton model (Croton et al. 2006), $\beta_{\text{LD}} \approx 2$ that is similar to the Somerville model (Somerville et al. 2008) and a number of other models, and a large value of $\beta_{\text{LD}} \approx 6$ (e.g. Cole et al. 1994).

Because we parametrize the outflow loading factor as a power-law function of halo circular velocity, $\alpha_{\text{LD}} = \alpha_{\text{LD}}(V_{\text{vir}}/220 \text{ km s}^{-1})^{-\beta_{\text{LD}}}$, a covariance represented by a straight line in the log $\alpha_{\text{LD}}-\beta_{\text{LD}}$ plane corresponds to a group of models that have a fixed loading factor for a specific circular velocity. The magenta dotted line in Panel 4 of Fig. 2 roughly captures the covariance of the two parameters and corresponds to a loading factor of 4 for haloes with a circular velocity of 70 km s$^{-1}$ or a mass of about $1.5 \times 10^{11} M_\odot$ at the present time. The covariance suggests that these haloes may be critical for fitting the $K$-band luminosity function and the H\textsc{i} mass function of galaxies simultaneously.

Our fiducial model does not limit the total energy available for feedback but, of course, nature does impose limits. The total energy available for outflows powered by Type II SNe is

$$E_{\text{SN,max}} = \phi_{\text{SN}} \eta_{\text{SN}}$$

where $\epsilon_{\text{SN}} \sim 10^{51}$ erg $= 5 \times 10^7 M_\odot$ km$^2$ s$^{-2}$ is the total energy output of a SN, and $\eta_{\text{SN}}$ is the number of SNe expected from the formation of 1 $M_\odot$ of stars. Assuming that Type II SN progenitors are stars with masses $8 < m < 20 M_\odot$, then $\eta_{\text{SN}} = 5.88 \times 10^{-3} M_\odot^{-1}$ for a Salpeter IMF (Salpeter 1955), and $\eta_{\text{SN}} = 9.76 \times 10^{-3} M_\odot^{-1}$ for a Chabrier IMF (Chabrier 2003). The total energy consumed by the outflow depends on the detailed processes and the final state of the outflow material. If the outflow is thermalized and settles into an equilibrium state in the gravitational potential of the host halo, the energy required is $\frac{5}{4} m_\odot V_{\text{esc}}^2$, where $m_\odot$ is the mass of the reheated gas (e.g. Kauffmann & Charlot 1998). Since only a fraction of the energy is expected to heat the gas:

$$\eta_{\text{SN}} \phi_{\text{SN}} \geq \frac{5}{4} \alpha_{\text{LD}} \left( \frac{220 \text{ km s}^{-1}}{V_{\text{vir}}} \right)^{\beta_{\text{LD}}} V_{\text{esc}}^2.$$

Thus, for haloes with $V_{\text{vir}} = 220$ km s$^{-1}$, the maximum normalization of the loading factor, $\alpha_{\text{LD,max}}(V_{\text{vir}} = 220) = 4.13 \times 10^{\eta_{\text{SN}}}$, which is 4.8 for a Salpeter IMF and 8.0 for a Chabrier IMF. These values are plotted in the lower right-hand panel of Fig. 2 as the blue solid line and the blue dashed line, respectively.

If the outflow is ejected directly from the centre of the host halo, the energy required to escape is $\frac{3}{2} m_\odot V_{\text{esc}}^2$, where $V_{\text{esc}}$ is the escape velocity. For a NFW halo (Navarro et al. 1996) with a concentration $c = 10$, the escape velocity from the centre is $V_{\text{esc}}^2 \approx 13 V_{\text{vir}}^2$. Thus, to eject the gas from such a halo, the lower limit on the required energy is $E = \frac{3}{2} m_\odot V_{\text{sc}}^2$, where $m_\odot$ is the total gas mass loaded in the wind. The mass-loading factor in such haloes is, therefore, limited by the total energy provided by SN explosions, and the solid red line and the red dashed line in Panel 4 of Fig. 2 show these limits for a Salpeter and a Chabrier IMF, respectively. Models lying above such a limiting line need more energy than is available from haloes with circular velocities equal to or below 100 km s$^{-1}$. The posterior contours for all the modes with $\beta < 4$ are $\sim$0.3–0.6 dex to the left of but not far from the energy limits for haloes with $V_{\text{vir}} \lesssim 100$ km s$^{-1}$, indicating that a large fraction, i.e. 25–50 per cent, of all the SN energy has to be put into the outflow in such galaxies.

The required high loading factor results from the need to simultaneously match the shallow low-end slope of both the $K$-band luminosity function and the H\textsc{i} mass function. The total baryonic mass that can be accreted into a halo is $f_b m_\odot$, where $f_b$ is the universal baryon fraction. Suppose a fraction $f_e$ of the gas can cool and collapse on to the central galaxy. For low-mass haloes with $V_{\text{vir}} \lesssim 150$ km s$^{-1}$, $f_e$ is expected to be close to unity because of the high cooling efficiency (Thoul & Weinberg 1995; Lu & Mo 2007). If the expelled gas can fall back on to the galaxy and go through multiple cycles, the effective $f_e$ can be larger than 1. However, to explain the shallow low-end mass of the $H$ mass function and the faint end slope of the $K$-band luminosity function, the cold gas and stellar mass have to be very low relative to $f_b m_\odot$. Indeed, the stellar mass to halo mass ratio for haloes with masses $\sim 10^{11} M_\odot$ is only about $3 \times 10^{-4}$ (Papastergis et al. 2012; Yang et al. 2012; Behroozi, Wechsler & Conroy 2013), and the cold gas mass to stellar mass ratio is no more than 10 (Kannappan 2004; Papastergis et al. 2012). Thus, the gas associated with low-mass haloes either is never accreted into the galaxy or has been expelled by galactic winds. In the latter case, if we conservatively assume that the cold gas mass is 10 times the stellar mass, then the total cold baryonic mass in a $10^{11} M_\odot$ halo is $\sim 0.033$ of the halo mass. If we make the further conservative assumption that for every unit of stellar mass formed at early times only half of it remains in stars today owing to stellar mass loss, the total mass that was ever involved in star formation is $\sim 0.006$ of the final halo mass. Based on these conservative assumptions, one finds that the mass-loading factor is $(f_b - 0.033)/0.006 = 23$, which is close to the upper limit obtained from the total kinetic energy expected from SN.

The posterior marginals (Fig. 2) imply that 25–50 per cent of the available SN energy, depending on the assumed IMF, is required to power the wind. This required high efficiency of energy loading poses a severe challenge for the standard SN feedback scenario (e.g. Dekel & Silk 1986; White & Zaritsky 1992). Indeed, detailed hydrodynamical simulations by Mac Low & Ferrara (1999) and Strickland & Stevens (2000) demonstrated that SN feedback is very inefficient in expelling mass because the onset of Rayleigh–Taylor instabilities severely limits the mass-loading efficiency of galactic winds. There are a couple of possible solutions to this dilemma. First, a top-heavy or bottom-light IMF would yield a larger $\eta_{\text{SN}}$. Our data constraints are from the local galaxy population, which by themselves do not distinguish different IMFs. Second, the ultraviolet (UV) photons from massive stars could effectively transfer radiation momenta to the gas via dust opacity, thereby producing momentum-driven winds (Murray et al. 2005). The total radiation energy available is about 100 times as high as the kinetic energy of SN explosions, vitiating the mechanical energy limit. The viability of this mechanism is uncertain: while some simulations
show that it is effective (e.g. Hopkins et al. 2011), others find that the development of Rayleigh–Taylor-type instabilities can suppress the formation of outflows (Krumholz & Thompson 2012). Thus, although the required high efficiency may rule out mechanisms based on SN energy-driven winds alone, the viability including momentum-driven winds is still unclear.

The star formation feedback models implemented in SAMs are crude. In our model, the outflow reduces the total cold gas mass of the galaxy and the remaining cold gas is redistributed in an exponential profile at each time step. This is similar to some cosmological hydrodynamical simulations (e.g. Davé et al. 2013) where the effective viscosity of the disc gas, which could include actual gas viscosity and/or gravitational torques, causes the cold gas from the outer disc to move rapidly inwards to replenish the star-forming gas, and so star formation can continue near the disc centre even if the total amount of gas in the disc is reduced. However, since the gas viscosity in such simulations is artificial and the resolution is still limited, it is unclear whether the effect is physical or numerical. For example, if the effective viscosity of the disc gas were unimportant, much of the cold gas would stay in the outer region of the disc without forming stars. Some simulations of individual discs (Forbes, Krumholz & Burkert 2012) do find that star formation feedback is localized to regions where the star formation is active, producing a disc with a reduced cold gas density only in the central part of the disc, leaving the cold gas in the outskirts intact. Others confirm the cosmological simulation results (e.g. Hirschmann et al. 2013).

To explore this in our SAM, we have considered a model in which star formation feedback is assumed to only affect the gas within the radius where star formation can happen, i.e. the cold density is above the threshold surface density. Without a lower limit on $f_{\text{SF}}$, the posterior distribution is nearly unchanged owing to the low surface density threshold for star formation and the high OFR required by the shallow low-mass end of the galaxy luminosity function. However, if the value of $f_{\text{SF}}$ is forced to match the observed star formation threshold surface density of 10 M$_{\odot}$ pc$^{-2}$, then the model fails to fit the data ($p_{\text{SF}} < 0.005$), consistent with the results obtained by Mo et al. (2005).

As we remarked before, we find that there are three major modes in the log $\beta_{\text{LD}}$-$\beta_{\text{PLM}}$ plane located at $\beta_{\text{LD}} \approx 0$, 2, and 6, as one can see in Panel 4 of Fig. 2. The relative posterior odds in each mode are 4:1:2, respectively. Hence, the $\beta = 0$ mode is favoured by the combined data sets in terms of marginalized posterior probability. Since our prior distribution is intentionally very broad, as we describe in Section 2, and this will quantitatively affect these odds. That said, the maximum likelihood is located in the $\beta_{\text{LD}} \approx 6$ mode, and the highest likelihood values in the $\beta_{\text{LD}} \approx 2$ and 0 modes are 30 and 1400 times lower, indicating that those two modes produce a worse fit to the data than the $\beta_{\text{LD}} \approx 6$ mode. The PPC $p$-values calculated for the $K$-band luminosity function predicted by each of the modes also indicate that the $\beta_{\text{LD}} \approx 6$ mode produces a better, but still inadequate, fit to the data, $p_K = 0.016$ for the $\beta_{\text{LD}} \approx 6$ mode, and $p_K < 0.005$ for the other two modes. In Fig. 3, we show the $K$-band luminosity function produced by the three modes. The shaded blue region shows the posterior predictions of models with $\beta_{\text{LD}} < 1.25$, the magenta region shows the posterior predictions of models with $1.25 < \beta_{\text{LD}} < 4$, and the green region shows the posterior predictions of models with $\beta_{\text{LD}} > 4$. Note that the three modes produce indistinguishable predictions for the H$_\text{i}$ mass function, as the predictive distribution of the H$_\text{i}$ mass function produced by the entire posterior shown in Fig. 1 is narrow. The predicted faint-end slope of the $K$-band luminosity function is steeper for a smaller value of $\beta_{\text{LD}}$, and the current data apparently favour the mode with the highest $\beta_{\text{LD}}$. This trend confirms recent hydro-dynamical simulation results of Puchwein & Springel (2013) and Davé et al. (2013). Moreover, this demonstrates that discrimination between these modes could be possible given improved data for the faint end of the luminosity function.

Even though the $\beta_{\text{LD}} \approx 0$ mode has relatively lower likelihood, it has the highest marginal posterior probability. This implies that the modes with high values of $\beta_{\text{LD}}$ are supported by proportionately smaller volumes in parameter space relative to the $\beta_{\text{LD}} \approx 0$ mode. This could be investigated in detail using the distribution of marginal likelihood, but the required resolution to compute the volume in the high-dimensional parameter space is computationally prohibitive. Nonetheless, relying on the posterior marginals shown in Fig. 2, we believe that the complex posterior results from the fine-tuning necessary to simultaneously match the $K$-band luminosity and H$_\text{i}$ mass function. Specifically, Panel 5 indicates that $\beta_{\text{SF}}$ correlates with $\beta_{\text{LD}}$. As $\beta_{\text{SF}}$ increases, the increasingly smaller star formation efficiency in low circular velocity haloes leaves more cold gas in the disc. To simultaneously match the H$_\text{i}$ mass function, the model thus needs to increase the feedback efficiency in lower circular velocity haloes by increasing $\beta_{\text{LD}}$.

In addition, sometimes a particular parametrization can also introduce specific features into the posterior. Panel 6 of Fig. 2 shows an example; the parameter $V_{\text{SF}}$ is not constrained when $\beta_{\text{LD}} \approx 0$, but is strongly constrained when $\beta_{\text{LD}} \approx 6$. When $\beta_{\text{LD}} \approx 0$, $\beta_{\text{SF}}$ is also about 0, meaning a constant star formation efficiency. In this case, as described in Section 2, $V_{\text{SF}}$ no longer has any effect on the model predictions, because the parametrization for the star formation efficiency no longer has a transition scale that $V_{\text{SF}}$ describes. Thus, the likelihood is the same no matter what value $V_{\text{SF}}$ takes when $\beta_{\text{LD}} \approx 0$.

Moreover, the high $\beta_{\text{LD}}$ mode requires two unusual conditions. First, as shown in Panel 7 of Fig. 2, for $\beta_{\text{LD}} \approx 6$, the parameter $\alpha_{\text{SN}}$, which characterizes the fraction of SN kinetic energy used to eject the hot halo gas, is always required to be higher than one, implying that the required energy is higher than the kinetic energy provided by Type II SN. The reason is that $\beta_{\text{LD}} \approx 6$ leads to a
rapid decrease of cold gas ejection with increasing halo circular velocity, and a strong outflow of hot halo gas in relatively massive haloes is needed by the mode to fit the data. This conclusion agrees with the result of Mutch et al. (2013) that one needs more energy than that available from Type II SN to explain the evolution of the galaxy stellar mass function from \( z \approx 0.8 \) to 0. Second, the \( \beta_{LD} \approx 6 \) mode also requires a long time-scale, about 10 times as long as the fiducial dynamical friction time-scale, for satellite galaxies to orbit in the host halo before merging into central galaxies (see Panel 8 of Fig. 2). This can be understood as follows. When \( \beta_{LD} \) is large, the effect of star formation feedback is increasingly weaker for higher mass haloes. In this case, the merging time-scale for satellite galaxies is required to be long to prevent them from merging into the halo centre and hence significantly increasing the mass of the central galaxy. Since a longer merger time also leads to a higher satellite fraction, observational data of satellite fraction can tighten the constraint. We will come back to this topic in a future paper.

### 4.2 Posterior predictions

The model family generally requires efficient outflows from low-mass galaxies for its best fit, but even that is not enough to match both of the constraining observational data sets. We now use the posterior distribution to predict the OFR and other observables to investigate the failure of the model family. Since the marginalized posterior in the \( \alpha_{LD} - \beta_{LD} \) plane is multimodal, we distinguish the modes according to their value of \( \beta_{LD} \) when making model predictions. We marginalize the posterior probability to generate predictions as described in Lu et al. (2012).

Fig. 4 shows the posterior prediction for the cold baryonic mass fraction (cold gas plus stars) of central galaxies normalized by the cosmic baryon fraction, \( f_c \), as a function of halo mass. The red line shows the median of the whole posterior prediction, while the shaded regions with different colours encompass the 95 per cent credible ranges from the modes with \( \beta_{LD} \sim 0, 2, \) and 6, as indicated in the figure. The black bars show the results of abundance matching obtained by Papastergis et al. (2012) using the Arecibo Legacy Fast ALFA (ALFALFA) and Sloan Digital Sky Survey (SDSS) data.

The observed \( K \)-band luminosity function only determines absolute magnitudes up to \( K = \text{Mag}_K - 5 \log h = -20 \), which only constrains the halo virial mass above \( \sim 10^{11} \text{M}_\odot \) (indicated as the vertical line in Fig. 4). The 95 per cent range of the posterior prediction contains the abundance matching result for haloes more massive than a few times \( 10^{11} \text{M}_\odot \), but all three modes overpredict the cold baryon fraction for lower mass haloes. Among the three modes, models with larger \( \beta_{LD} \) produce lower cold baryon mass fractions for low-mass haloes. The mode with \( \beta_{LD} \approx 0 \) vastly overpredicts the cold baryon mass fraction in haloes with masses \( \lesssim 10^{11.5} \text{M}_\odot \).

To explore the physical implications of feedback, we predicted the average SFR and OFR in each mass bin for all central galaxies at \( z = 0 \). Fig. 5 shows the ratio OFR/SFR, which is commonly referred to as the mass-loading factor of the wind. Again the red lines show the medians of the whole posterior prediction, while the shaded regions with different colours encompass the 95 per cent credible range from the three modes defined by their values of \( \beta_{LD} \). The predicted OFR/SFR ratio generally decreases with stellar mass; it is approximately 10 for galaxies with stellar masses \( \lesssim 10^{10} \text{M}_\odot \) and close to 1 for Milky Way mass galaxies. Observationally, there is no strong evidence for the existence of such a strong stellar mass dependence for outflows in normal disc galaxies at the present day (see Rupke, Veilleux & Sanders 2002; Heckman 2003; Veilleux, Cecil & Bland-Hawthorn 2005; Martin 2006). In starburst galaxies, the mass OFRs are roughly comparable to the SFR, so that OFR/SFR \( \approx 1 \), and lower mass galaxies do not show stronger outflows. Thus, the high OFR/SFR ratio predicted for the general dwarf galaxy population does not have any observational support, but then again there is no strong observational evidence against it. Comparing the predictions for the different posterior modes, we see that different choices of \( \beta_{LD} \) produce very different trends of the loading factor with galaxy stellar mass. The \( \beta_{LD} \approx 0 \) mode produces a constant mass-loading factor, OFR/SFR \( \approx 3 \) over a large stellar mass range, while the...
other two modes predict a rapidly decreasing OFR/SFR with increasing stellar mass. The roughly constant OFR/SFR predicted by the $\beta_{\text{LD}} \sim 0$ mode is consistent with the observational results of Martin (2006), but this mode significantly overpredicts the cold baryon mass fraction for low-mass haloes. Clearly, as one needs a large mass-loading factor to achieve better fits to the constraining data, observations of gas outflows in low-mass galaxies, and how they scale with galaxy mass can provide important constraints on the feedback model family considered here.

Fig. 6 shows the predicted cosmic SFR density as a function of redshift using the posterior samples selected from the three modes of $\beta_{\text{LD}}$. The grey shaded region in the figure denotes a compilation of observational estimates of the cosmic SFR density as a function of redshift from Behroozi et al. (2013). Compared to the data, our model family predicts a different shape. The cosmic SFR density in the model reaches its maximum at higher redshifts ($z \approx 3.5$), while data suggest it more sharply peaks around $z \approx 2$, although the observed drop at high redshifts could partly owe to missing low-mass galaxies (Lu et al. 2014). The model also underpredicts the SFR density around $z = 2$, as the median of the posterior predictive distribution is only marginally enclosed by the lower bound of the observational data. The predictions from the $\beta_{\text{LD}} \approx 0$ mode are systematically higher and might be more consistent with the observations. The high $\beta_{\text{LD}}$ mode predicts a SFR density that is too low to match the observational data at $z > 1$. This stems from the powerful SN-driven outflows in low-mass haloes required to fit the local galaxy luminosity function and H I mass function. Because the parametrization for feedback does not have an explicit redshift dependence, such strong outflows work throughout all cosmic time in the model. This feedback suppresses star formation at high redshift where such low-mass haloes dominate.

Fig. 7 shows the SFR as a function of redshift for haloes with masses of $M_{\text{vir}} = 10^{12} M_\odot$ today ($z = 0$) for the three modes in $\beta_{\text{LD}}$. The model predictions are compared with the recent results obtained by Lu et al. (2014) and Behroozi et al. (2013) using empirical models tuned to fit the observed stellar mass functions at different redshifts. The empirical models predict a peak near $z \approx 1$ while our SAM family predicts a much weaker peak at higher $z$. This is consistent with the underprediction of the SFR density at $z \approx 1$ in Fig. 6. The $\beta_{\text{LD}} = 0$ mode also overpredicts the SFR at $z > 2$, although this mode better matches the observed cold gas fraction in haloes with masses of $10^{12} M_\odot$ (see Fig. 4). Overall, the predicted star formation histories for Milky-Way-sized galaxies by our SAM family are inconsistent with the results obtained from the empirical models.

Weinmann et al. (2012) found that SAMs and numerical simulations cannot generally reproduce the observed number density evolution of low-mass galaxies. For comparison, we predict the number densities of galaxies with stellar masses in the range $9.27 \leq \log M_*/M_\odot \leq 9.77$ from $z = 0$ to 6, and compare them with the observational data compiled by Weinmann et al. (2012). Fig. 8 shows that the model family can accommodate the observational data at $z = 0$, but vastly overpredicts the number density of low-mass galaxies at higher redshift. This confirms the results of Weinmann et al. (2012), and suggests that the current model family cannot match the data even if a large parameter space is explored.

Fig. 9 shows the prediction of the comoving mass density of cold gas as a function of redshift. The observational data at high redshifts, estimated using damped Lyman $\alpha$ systems (Péroux et al. 2003; Prochaska & Wolfe 2009), suggest that the cold gas density at high redshifts is higher than in the local universe. Our predicted increase of the cold gas mass density with redshift in the range from $z = 0$ to 2 is consistent with the observations. The model also predicts a decrease of the cold gas density with increasing redshift at $z > 3$. Unfortunately the current data are still too uncertain to provide meaningful constraints on the model.

In summary, the model family considered here, which has physical prescriptions similar to most published SAMs, in addition to not fitting one of the constraining data sets, the observed K-band luminosity function, also fails to predict additional observations as follows.

(i) It requires very efficient feedback to eject gas from low-mass haloes. The energy required is probably too high if the outflows are...
The posterior prediction for the evolution of the number density of galaxies with stellar masses in the range \( \log (M_*/M_\odot) = 9.27 - 9.77 \) as a function of redshift. Points with error bars show data compiled by Weinmann et al. (2012). The colour bands and lines are as described in Fig. 4.

![Figure 8](image_url)

**Figure 8.** The posterior prediction for the evolution of the number density of galaxies with stellar masses in the range \( \log (M_*/M_\odot) = 9.27 - 9.77 \) as a function of redshift. Points with error bars show data compiled by Weinmann et al. (2012). The colour bands and lines are as described in Fig. 4.

The primary cause for the failure is that to simultaneously produce the observed stellar component and the observed cold gas fraction at the low-mass end requires the SN feedback to be extremely efficient in expelling gas. This requires a high mass-loading factor that consumes much of the available energy budget. The low \( \beta \) value we found by exhaustively exploring the parameter space is inconsistent with that obtained from empirical models.

(iv) The predicted SFR history of Milky-Way-sized galaxies is inconsistent with that obtained from empirical models.

(v) It overpredicts the number density of low-mass galaxies at high redshift.

These results together suggest that some important physics may still be missing in the current SAMs.

## 5 Discussion and Conclusions

We use the observed low-redshift \( K \)-band luminosity function and \( \text{H}_1 \) gas mass function of galaxies as data constraints to explore star formation and feedback in a SAM of galaxy formation, which uses MCMC-based Bayesian methodology. The combination of these two data sets breaks the degeneracy in the parametrized phenomenology based on using the \( K \)-band luminosity function data constraint alone (Lu et al. 2012). Our model assumes that baryons accrete into the assembling halo where they cool, form stars, and the resulting feedback ejects some fraction of the gas. Our model has many features in common with most published SAMs and our results focus on conclusions generic to a large class of published models. We marginalize over all the uncertainties in the model family that are not constrained by the data sets.

Our broad parameter space covers and extends the ranges of the important parameters for star formation and feedback adopted in many existing SAMs. Our exhaustive parameter-space search results in the robust identification of several compact modes that best match the constraining data. These modes are likelihood dominated, i.e. the existence of these modes does not depend sensitively on the prior distribution as long as the modes are supported. The analyses presented in this paper focus on the behaviour of the model constrained to those modes, namely on the physical implications of the constrained parameters and the predictions of these modes. These conclusions are not expected to depend sensitively on the prior. However, the choice of the prior distribution is generally important in complex high-dimensional problems, especially when the likelihood is not sufficiently informative. In such cases, strong prior belief easily sways weak observational constraints. Conversely, a thorough investigation of the poorly constrained model motivates the acquisition of data that robustly discriminates between competing hypotheses of interest.

Overall, our model fails to reproduce the joint data sets and, therefore, more work will be required to understand the nature of the failure and to propose a remedy before attempting a quantitative Bayesian analysis. To assess the fit, we use the PPC method with a residual sum of squares for each bin, for both the \( K \)-band luminosity function and the \( \text{H}_1 \) gas mass function, as a discrepancy statistic. In this method, the discrepancy with the data is compared with the distribution predicted from the posterior distribution. We find that the data are in the tail of the distribution defined by the posterior, suggesting the implausibility of the data given the model.

The primary cause for the failure is that to simultaneously produce the observed stellar component and the observed cold gas fraction at the low-mass end requires the SN feedback to be extremely efficient in expelling gas. This requires a high mass-loading factor that consumes much of the available energy budget. The low \( \beta \) value we found by exhaustively exploring the parameter space has its roots in the over simplicity of the phenomenological model and suggests that key physical processes are missing or misrepresented.

In addition, we find that the mass-loading factor must depend strongly on halo circular velocity to obtain the shallow faint-end slope of the luminosity function. The model requires that winds from...
haloes with circular velocities lower than $\sim 200 \, \text{km} \, \text{s}^{-1}$ are ejected from the halo, completely. However, in many mass-loading models the mass-loading factor is a weak function of the halo circular velocity. Recent simulations concur: Davé et al. (2013) show that a steeper relation between mass loading and halo circular velocity tends to produce a shallower stellar and H\textsc{i} mass function at the low-mass end. This also suggests that accurate observational data for the faint-end slope of the galaxy luminosity function are crucial to constrain feedback models.

The inferred requirement for efficient feedback is consistent with the recent results of Mutch et al. (2013) and Henriques et al. (2013) who also found that a high efficiency of SN feedback is needed to fit the stellar mass and/or luminosity functions of galaxies at multiple redshifts. However, the high efficiency implied by the posterior distribution obtained here is not supported by hydrodynamical simulations (e.g. Mac Low & Ferrara 1999), which showed that the feedback efficiency of SN kinetic energy is usually quite low. Thus, either some other important energy sources are missing or the way feedback works is not correctly modelled in the current model family. For example, radiation pressure associated with star formation feedback works is not correctly modelled in the current model family. This is a direct consequence of a short radiative cooling time-scale in low-mass haloes combined with the small fraction of baryons in stars and cold gas relative to the cosmic baryon fraction. Moreover, the predicted high mass-loading factor and mass-loss rates are not observed, suggesting that a feedback-generated wind may not be the agent suppressing star formation and cold gas accretion in low-mass haloes, at least at low $z$.

Alternatively, this suggests exploring mechanisms that prevent the gas accretion in the first place. For example, the IGM may be pre-heated by some processes so that low-mass haloes can accrete baryonic matter only at a reduced rate, either by properly including the effects of already considered processes like SN winds or by additional processes like gravitational pancaking or blazar heating (e.g. Mo & Mao 2002; Mo et al. 2005; Lu & Mo 2007; Anderson & Bregman 2010; Zhu, Feng & Fang 2011; Pfrommer, Chang & Broderick 2012). This may also introduce a characteristic time before and after which the star formation and feedback proceed differently, as seems to be required to match the star formation history and number density evolution of low-mass galaxies (see Lu et al. 2014). We will investigate the effects of pre-heating on galaxy formation in a future paper.

**ACKNOWLEDGEMENTS**

This work is partly supported by grants NSF AST-1109354 (HJM and MDW), NSF AST-0908334 (HJM and NK), and NASA ATP NNX10AJ95G (NK). Part of this work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

YL thanks Tom Abel, Peter Behroozi, Eric Bell, Andrew Benson, Darren Croton, Romain Davé, Andrey Kravtsov, Emmanouil Papastergis, Joel Primack, Rachel Somerville, Frank van den Bosch, and Risa Wechsler for useful discussion.

**REFERENCES**

Anderson M. E., Bregman J. N., 2010, ApJ, 714, 320
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003a, ApJ, 585, L117
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003b, ApJS, 149, 289
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Bigiel F., Leroy A., Walter F., Brinks E., De Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, MNRAS, 407, 2017
Bower R. G., Benson A. J., Crain R. A., 2012, MNRAS, 422, 2816
Bruzual G., 2007, in Vallenari A., Tantalo R., Portinari L., Moretti A., eds, ASP Conf. Ser. Vol. 374, From Stars to Galaxies: Building the Pieces to Build Up the Universe. Astron. Soc. Pac., San Francisco, p. 303
Cattaneo A., Dekel A., Devriendt J., Guiderdoni B., Blaizot J., 2006, MNRAS, 370, 1651
Chabrier G., 2003, PASP, 115, 763
Chen Y.-M., Tremonti C. A., Heckman T. M., Kauffmann G., Weiner B. J., Brinchmann J., Wang J., 2010, AJ, 140, 445
Cole S., Aragón-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Croton D. J. et al., 2006, MNRAS, 365, 11 Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645
Dekel A., Silk J., 1986, ApJ, 303, 39
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
De Lucia G., Kauffmann G., White S. D. M., 2004, MNRAS, 349, 1101
Dunkley J. et al., 2009, ApJS, 180, 306
Elmegreen B. G., 1997, ApJ, 486, 944
Forbes J., Krumholz M., Burkert A., 2012, ApJ, 754, 48
Gelman A., Rubin D., 1992, Stat. Sci., 7, 457
Gelman A., Meng X.-L., Stern H., 1996, Stat. Sinica, 6, 735
Gelman A., Carlin J. B., Stern H. S., Rubin D. B., 2003, Bayesian Data Analysis, 2nd edn. CRC Press, Boca Raton, FL
Giovanelli R. et al., 2005, AJ, 130, 2598
Gómez F. A., Coleman-Smith C. E., O’Shea B. W., Tumlinson J., Wolpert R. L., 2012, ApJ, 760, 112
Guo Q. et al., 2011, MNRAS, 413, 101
Heckman T. M., 2003, Rev. Mex. Astron. Astrosf. Conf. Ser., 17, 47
Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, MNRAS, 396, 535
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373
Hirschmann M. et al., 2013, MNRAS, 436, 2929
Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950
Hopkins P. F., Cox T. J., Hernquist L., Narayanan D., Hayward C. C., Murray N., 2013, MNRAS, 430, 1901
Kampakoglou M., Trotta R., Silk J., 2008, MNRAS, 384, 1414
Kang X., Jing Y. P., Mo H. J., Börner G., 2005, ApJ, 631, 21
Kannappan S. J., 2004, ApJ, 611, L89
Kauffmann G., Charlot S., 1998, MNRAS, 294, 705
Kennicutt R. C., Jr, 1989, ApJ, 344, 685
Kennicutt R. C., Jr, 1998, ARA&A, 36, 189
Kennicutt R. C. J. et al., 2007, ApJ, 671, 333
Keres D., Yun M. S., Young J. S., 2003, ApJ, 582, 659

---

**MNFRAS 443, 1252–1266 (2014)**
