Monolithic or hierarchical star formation? A new statistical analysis

Marios Kampakoglou,[★] Roberto Trotta[★] and Joseph Silk[★]

Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH

ABSTRACT

We consider an analytic model of cosmic star formation which incorporates supernova feedback, gas accretion and enriched outflows, reproducing the history of cosmic star formation, metallicity, Type II supernova rates and the fraction of baryons allocated to structures. We present a new statistical treatment of the available observational data on the star formation rate and metallicity that accounts for the presence of possible systematics. We then employ a Bayesian Markov Chain Monte Carlo method to compare the predictions of our model with observations and derive constraints on the seven free parameters of the model. We find that the dust-correction scheme one chooses to adopt for the star formation data is critical in determining which scenario is favoured between a hierarchical star formation model, where star formation is prolonged by accretion, infall and merging, and a monolithic scenario, where star formation is rapid and efficient. We distinguish between these modes by defining a characteristic minimum mass, $M \gtrsim 10^{11} \, \mathrm{M_\odot}$, in our fiducial model, for early-type galaxies where star formation occurs efficiently. Our results indicate that the hierarchical star formation model can achieve better agreement with the data, but that this requires a high efficiency of supernova-driven outflows. In a monolithic model, our analysis points to the need for a mechanism that drives metal-poor winds, perhaps in the form of supermassive black hole induced outflows. Furthermore, the relative absence of star formation beyond $z \sim 5$ in the monolithic scenario requires an alternative mechanism to dwarf galaxies for re-ionizing the universe at $z \sim 11$, as required by observations of the microwave background. While the monolithic scenario is less favoured in terms of its quality-of-fit, it cannot yet be excluded.

Key words: Galaxy: formation – galaxies: formation.

1 INTRODUCTION

Massive galactic spheroids form either by hierarchical build-up or monolithically. The former scenario is favoured by the theory of, and evidence for, cold dark matter (CDM); the latter is favoured by *some* observations, most notably the evidence for downsizing in stellar mass and on chemical evolution time-scales from the [α/Fe] enhancement at high spheroid masses (Worthey, Faber & Gonzalez 1992). The conventional view of hierarchical build-up fits in with semi-analytical galaxy formation simulations, provided that suitable feedback models are prescribed (Bower et al. 2006; Croton et al. 2006). Gas cooling drives star formation below a critical halo mass and active galactic nucleus (AGN) quenching of star formation occurs at higher masses where the cooling is inefficient. This model reproduces the shape of the galaxy luminosity function at the high-mass end, and also produces old, red massive galaxies. Not yet addressed in these studies is the issue of whether these massive galaxies can be formed sufficiently rapidly, the resolution of which

may lend support to a monolithic formation scenario. Major mergers are a plausible ingredient of a monolithic model, for triggering rapid star formation.

Not all studies agree on the role of major mergers. Not all luminous starbursts are triggered by such mergers. Indeed, in the case of disc galaxy formation, their role is likely to be small. There is evidence that major mergers are not exclusively responsible for the dominant star formation episodes in starbursts forming stars at up to 200 $\mathrm{M_\odot} \, \mathrm{yr^{-1}}$ where thick discs are seen (e.g. Hammer et al. 2005). Such high star formation rates (SFRs), also found in discs at $z \sim 2$ by Forster-Schreiber et al. (2006), may favour an interpretation in terms of bulge formation since the discs are already present. The observed major merger rate is small out to $z \sim 1.2$ despite the strong increase in the comoving SFR (Lotz et al. 2006). Of course, ultraluminous infrared galaxies, with SFRs of $\gtrsim 500 \, \mathrm{M_\odot} \, \mathrm{yr^{-1}}$, are examples of major merger induced star formation. However, for the bulk of star formation in the early universe, there is little evidence to suggest that major mergers play a significant role. This motivates a hierarchical picture of prolonged minor mergers to build up mass.

One can ask more generally whether minor mergers or gas cloud accretion is responsible for the gas supply. With regard to minor mergers, the answer seems to be negative, because the time-scale

★E-mail: mariosk@astro.ox.ac.uk (MK); rxt@astro.ox.ac.uk (RT); silk@astro.ox.ac.uk (JS)

for supplying the gas would be many dynamical time-scales. Yet there are two persuasive arguments for a short star formation time-scale in massive galaxies. Spectral energy distribution (SED) modelling suggests downsizing the most-massive galaxies forming first and with star formation time-scales of the order of a dynamical time. The [α/Fe]-enhancement with increasing spheroid mass independently reinforces this conclusion, since most of the star formation must have occurred before dilution of the star-forming gas by Type Ia supernova (SNIa) ejecta occurred. This result is not, or at least not yet, found in semi-analytical modelling of massive galaxy formation. It provides a strong argument for monolithic formation of massive galaxies, although one cannot of course exclude the possibility that hierarchical formation models will reproduce a similar result once more complex star formation rules are introduced. Indeed, simple spherically symmetric accretion models have been proposed (cf. Birnboim, Dekel & Neistein 2007) that allow the gas to accumulate in an isothermal halo atmosphere prior to an accretion-triggered burst of star formation. In essence, this approach introduces a monolithic formation model in combination with hierarchical gas accumulation. Unfortunately, there is little evidence today for such gas-rich haloes, suggesting that this process, if important, would only have been influential at very early epochs.

To save hierarchical models, one needs an efficient way of converting gas into stars. The typical gas-to-star e-folding time in nearby well-studied sites of star formation, namely spiral galaxies, is several Gyr. By contrast, at a redshift of 2–3, the star formation time-scale is as short as 0.2 Gyr from both spectrophotometric SED filling (Maraston et al. 2006) and [α/Fe] analyses (Thomas et al. 2005). The AGN phenomenon is invoked for stopping star formation by quenching the gas supply. This may keep massive galaxies red, but may not be enough to produce massive galaxies sufficiently early. One resolution may be to invoke positive feedback from AGN-driven outflows that overpressure protogalactic clouds and trigger star formation on a rapid time-scale (Silk 2005). Whether or not this particular solution turns out to resolve the dilemma (if indeed the problem persists in the perspective of improved star formation modelling) is not the issue we focus on in this paper. Rather we revisit the case for hierarchical versus monolithic galaxy formation in terms of accounting for the data on SFR and chemical evolution.

In this paper, we apply a novel statistical approach that rigorously treats the key parameters in semi-analytical galaxy formation theory, with emphasis on reproducing the cosmic star formation and chemical evolution histories. Numerical modelling via semi-analytical techniques of a large box of the universe takes up so much computer time and memory that it is impossible to test the robustness of the results. Here, we focus on probing the key parameter space by means of analytical techniques combined with appropriate binning of the full data sets. We find that the characteristic minimum mass for the building blocks of massive galaxies plays a central role. Specifically, the dust-correction scheme one chooses to adopt for the cosmic star formation history data is one of the most-critical factors in determining the balance of evidence in support of a hierarchical star formation model as opposed to a monolithic scenario, where star formation happens predominantly in massive spheroids. Our results indicate that the hierarchical star formation model can achieve better agreement with the data, but that this requires a high efficiency of SN-driven outflows. In a monolithic model, our analysis points to the need for a mechanism that drives metal-poor winds, perhaps in the form of supermassive black hole (SMBH) induced outflows. While the monolithic scenario is less favoured in terms of its quality-of-fit, it cannot yet be excluded.

This paper is organized as follows. Our star formation model is introduced in Section 2, and the dependence of the observable quantities on the free parameters of the model explored in Section 3. We then describe the data employed and our statistical procedure in Section 4. Our results are presented in Section 5 and our conclusions discussed in Section 6. Appendix A gives details of our binning procedure for SFR and metallicity data which accounts for undetected systematics.

2 STAR FORMATION MODEL

In this section, we present a physical model of the cosmic star formation incorporating SN feedback, gas accretion and enriched outflows. Our model builds on the model described in Daigne et al. (2006), and more specifically their 'Model 0'. The main difference at the model level is our choice of using a different chemical evolution model. In fact, we adopt the instantaneous recycling approximation, whereas in Daigne et al. (2006) the metallicity is computed under the delayed enrichment approximation. There are several differences in terms of the statistical treatment of the data and the fitting procedure, that in this work are significantly more sophisticated, as explained in Section 4 and Appendix A.

2.1 Governing equations

The description of baryons in the Universe and the processes that define the evolution of the baryonic mass are of fundamental importance to our model. Following Daigne et al. (2006), we employ three baryon reservoirs in the model, encompassing the interstellar medium (ISM) (gas), the mass in stars and the intergalactic medium (IGM). We denote by M_{gas} the mass of gas, by $M_\star$ the mass in stars and by $M_{struct} = M_{gas} + M_\star$ the total mass in collapsed structures. The IGM and the structures exchange mass through accretion and outflow, while the interaction between stars and gas is governed by star formation and ejection of enriched gas. In the instantaneous recycling approximation adopted here, the accretion rate of the mass in stars is simply equal to the SFR, $\Psi(t)$, that is,

$$\frac{dM_\star}{dt} = \Psi(t). \tag{1}$$

We then have the following set of differential equations governing the evolution of the mass in the three reservoirs:

$$\frac{dM_{IGM}}{dt} = -\frac{dM_{struct}}{dt}, \tag{2}$$

$$\frac{dM_{struct}}{dt} = a_b(t) - o(t), \tag{3}$$

$$\frac{dM_{gas}}{dt} = \frac{dM_{struct}}{dt} - \frac{dM_\star}{dt}. \tag{4}$$

In the above equations, $a_b(t)$ is the rate of baryon accretion, while $o(t)$ is the rate of baryon outflow. The latter includes two terms, accounting for winds powered by stellar explosions and SN ejecta. We neglect SN ejecta since this effect was found to be subdominant (Daigne et al. 2004). The relation between physical time t and redshift z is given by

$$\frac{dt}{dz} = \frac{9.78\,h^{-1}\,\text{Gyr}}{(1+z)\sqrt{\Omega_\Lambda + \Omega_m(1+z)^3}}, \tag{5}$$

where we have assumed a flat Universe. In the following, we will use equation (5) with parameters fixed to the values of the ΛCDM concordance model, that is, a matter density parameter $\Omega_m = 0.27$

and a cosmological constant energy density $\Omega_\Lambda = 0.73$ (both in units of the critical energy density of the Universe), and we will take for the Hubble constant $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}} = 71$.

2.2 Accretion

We adopt the hierarchical scenario of structure formation, where small structures are formed first. At redshift z, the comoving density of dark matter haloes in the mass range $[M, M + \mathrm{d}M]$ is $f_\mathrm{ps}(M, z)\mathrm{d}M$, normalized in such a way that

$$\int_0^\infty \mathrm{d}M\, M f_\mathrm{ps}(M,\, z) = \rho_\mathrm{DM}, \qquad (6)$$

where ρ_DM is the comoving dark matter density. The distribution function of haloes $f_\mathrm{ps}(M, z)$ is computed using the method described in Jenkins et al. (2001) using code provided by A. Jenkins. It follows the standard theory (Press & Schechter 1974) including the modification of Sheth & Tormen (1999). We assume a primordial power spectrum of fluctuation with a power-law index $n_\mathrm{S} = 1$ and the fitting formula to the exact transfer function for non-baryonic CDM given by Bond & Efstathiou (1999). For the rms amplitude, we adopt a value $\sigma_8 = 0.9$ for mass density fluctuations in a sphere of radius $8h^{-1}$ Mpc.

Using the above expressions for the distribution function of dark matter haloes, we can calculate the fraction of baryons at redshift z that is allocated to structures, by assuming that the baryon density is proportional to the dark matter density, with a proportionality factor given by the ratio of visible to dark matter density – in other words, we assume that light traces matter with no bias. The fraction of baryons in star-forming structures at redshift z is then given by

$$f_\mathrm{bar}(z) = \frac{\int_{M_\mathrm{min}}^\infty \mathrm{d}M\, M f_\mathrm{ps}(M, z)}{\int_0^\infty \mathrm{d}M\, M f_\mathrm{ps}(M,\, z)}, \qquad (7)$$

where M_min is a free parameter controlling the minimum mass (in units of solar masses) of the collapsed structures where star formation can occur.

The accretion rate is then given by (Daigne et al. 2006):

$$a_\mathrm{b}(t) = \Omega_\mathrm{b} \frac{3H_0^2}{8\pi G} \left(\frac{\mathrm{d}t}{\mathrm{d}z}\right)^{-1} \left|\frac{\mathrm{d}f_\mathrm{bar}}{\mathrm{d}z}\right|. \qquad (8)$$

Given a value of M_min (that we adopt as a free parameter, see below), we fix the redshift at which star formation begins, z_init by the requirement that $f_\mathrm{bar}(z_\mathrm{init}) = 0.01$. In other words, the first stars form in collapsed haloes of mass larger than M_min when the fraction of baryons allocated to such structures is more than 1 per cent. We adopt a fixed baryonic density parameter of $\Omega_\mathrm{b} = 0.044$ (from the posterior mean of *WMAP* three-year data combined with all other data sets, Spergel et al. 2007).

2.3 Outflow

The adopted stellar initial mass function (IMF) is of the form

$$\Phi(m) = B \left(\frac{m}{M_\odot}\right)^{-(1+x)} \qquad \text{for } m_\mathrm{l} < m < m_\mathrm{u}, \qquad (9)$$

where the normalization constant B is fixed by the requirement that

$$\int_{m_\mathrm{l}}^{m_\mathrm{u}} m\Phi(m)\mathrm{d}m = M_\star. \qquad (10)$$

For the limits of integration, we fix $m_\mathrm{l} = 0.1\,\mathrm{M_\odot}$ and $m_\mathrm{u} = 100\,\mathrm{M_\odot}$ (Pagel 1997). Therefore, the only parameter needed to define the

IMF is its power-law index, x. The quantity x is used as a free parameter in this model.

We model the outflow powered by stellar explosions as follows:

$$o(t) = \frac{2\epsilon}{v_\mathrm{esc}^2(z)} \int_{m_0}^{100\,\mathrm{M_\odot}} \mathrm{d}m\,\Phi(m)\Psi(t - \tau_\mathrm{s}(m))E_\mathrm{kin}(m), \qquad (11)$$

where

$$m_0 = \max(8\,\mathrm{M_\odot}, m_\mathrm{d}(t)) \qquad (12)$$

and $\Phi(m)$ is the IMF defined above, $\tau_\mathrm{s}(m)$ is the lifetime of a star of mass m and $m_\mathrm{d}(t)$ is the mass of stars that die at age t. Furthermore, $E_\mathrm{kin}(m)$ is the kinetic energy released by the explosion of a star of mass m, that we take to be a fixed constant independent of mass, $E_\mathrm{kin}(m) = 10^{51}$ erg (a mass dependence could easily be taken into account). The free parameter ϵ controls the fraction of the kinetic energy of SNe that is available to power the winds, and $v_\mathrm{esc}^2(z)$ is the mean square of the escape velocity of structures at redshift z.

In order to compute the stellar lifetime $\tau_\mathrm{s}(m)$, we assume it to be equal to the time that a star of mass m spends on the main sequence. Thus, the age of a star of mass m is given by

$$\tau_\mathrm{s}(m) = (m/\mathrm{M_\odot})^{-2.5}\, t_\odot, \qquad (13)$$

where $t_\odot$ is the total time that a star of mass $M = \mathrm{M_\odot}$ will spend on the main sequence and we adopt a value $t_\odot = 9$ Gyr. To compute $m_\mathrm{d}(t)$ in equation (12), we solve equation (13) for m, thereby obtaining the mass of stars $m_\mathrm{d}(t)$ that die at age t.

The escape velocity is obtained by assuming virialized haloes and averaging over the distribution function, thus obtaining:[1]

$$v_\mathrm{esc}^2(z) = \frac{\int_{M_\mathrm{min}}^\infty \mathrm{d}M\, M f_\mathrm{ps}(M, z)(2GM/R(M))}{\int_{M_\mathrm{min}}^\infty \mathrm{d}M\, M f_\mathrm{ps}(M, z)}, \qquad (14)$$

where $R(M)$ is the radius of a dark matter halo of mass M given by the following expression:

$$R(M) = \left\{\frac{3M}{178\rho_\mathrm{c}[\Omega_\mathrm{m}(1+z)^3 + \Omega_\Lambda]4\pi}\right\}^{1/3}, \qquad (15)$$

where ρ_c is the critical density of the universe today. The factor 178 is the overdensity (relative to the critical density) at virialization for an Einstein–de Sitter model (Coles & Lucchin 1995).

2.4 Star formation and supernova rate

Following Daigne et al. (2006), we adopt an exponentially decreasing SFR:

$$\Psi(t) = \nu M_\mathrm{struct}(t) \exp\left[\frac{-(t - t_\mathrm{init})}{\tau}\right], \qquad (16)$$

where t_init is the time corresponding to the redshift z_init when star formation starts (as defined above), τ is a characteristic time-scale that we take as a free parameter and ν is a normalization parameter (with dimensions of inverse time).

The SN rate is strongly linked to the SFR because of the short lifetime of massive progenitors with $M > 8\,\mathrm{M_\odot}$. We can therefore

---

[1] This is the escape velocity from R to infinity, not the escape velocity from the sites of star formation that are deeper in the potential well. This approximation does not affect the results for the massive spheroids since even the shallower potential at R is deep enough to prevent winds from being effective, see the discussion in Section 5. For less-massive systems, we expect this approximation to result in slightly smaller values for the parameter ϵ than one would otherwise obtain.

Table 1. Upper part: free model parameters and priors used in the analysis. Top-hat (flat) priors have been adopted on the parameter ranges indicated. Lower part: model parameters that have been fixed.

Quantity	Symbol	Defined	Prior range or value
Minimum mass of collapsed haloes (M_{min} in $M_{\odot}$)	$\log M_{min}$	Section 2.2	$5 \leqslant \log M_{min} \leqslant 13$
SNII energy efficiency factor	ϵ	Equation (11)	$0.01 \leqslant \epsilon \leqslant 0.45$
IMF power-law index	x	Equation (9)	$3 \leqslant x \leqslant 2$
SFR normalization parameter (Gyr^{-1})	ν	Equation (16)	$0.01 \leqslant \nu \leqslant 5$
SFR time-scale (Gyr)	τ	Equation (16)	$1 \leqslant \tau \leqslant 5$
Wind load factor	η	Equation (18)	$0 \leqslant \eta \leqslant 30$
Metal dilution factor	f_{dil}	Equation (19)	$1 \leqslant f_{dil} \leqslant 30$
Baryon density parameter	Ω_b		0.044
Matter density parameter	Ω_m		0.27
Cosmological constant density parameter	Ω_{Λ}		0.70
Hubble constant ($km\,s^{-1}\,Mpc^{-1}$)	H_0		71
rms fluctuation amplitude	σ_8		0.9
Dark matter to baryons bias parameter	b		1.0
Minimum fraction of baryons when star formation begins	$f_{bar}(z_{init})$		0.01
Kinetic energy from stellar explosions	E_{kin}		10^{51} erg
Yield	q		0.02
Metallicity of accreted material	Z_F		0

assume that core-collapse SNe are strongly correlated with instantaneous SFR, and the SNII rate $\Psi_{SNII}(t)$ is given by

$$\Psi_{SNII}(t) = \int_{8\,M_{\odot}}^{m_u} \Phi(m)\Psi(t - \tau_s(m))dm. \qquad (17)$$

2.5 Chemical evolution model

Chemical evolution is included in the model using the instantaneous recycling approximation, that is, we assume that all processes involving stellar evolution, nucleosynthesis and recycling take place instantaneously on the time-scale of galactic evolution. The equation of galactic chemical evolution is (Pagel 1997):

$$\Sigma_g \frac{dZ}{dt} = q\Psi(t) + [Z_F - Z(t)]a_b(t) - (\eta - 1)Z(t)o(t), \qquad (18)$$

where Σ_g is the density of the gas (in units of $M_{\odot}/Mpc^3$), η is a multiple of the nucleosynthetic yield that parametrizes the metallicity Z of the SN ejecta (Dalcanton 2006) (also called the 'load factor', adopted here as a free parameter) and q is the yield. We fix the value of the yield to $q = 0.02$ and assume that the mass accreted to the disc has zero metallicity i.e. we fix $Z_F = 0$). Furthermore, we normalize the metallicities to the solar value for which we adopt $Z_{\odot} = 0.02$. The chemical evolution of the ISM, described by equation (18), contains three terms. The first one represents the chemical enrichment due to the evolution of stars. The second term represents the dilution of metallicity (if $Z_F < Z(t)$) or the chemical enrichment [if $Z_F > Z(t)$] of the ISM due to accreted material. The last term describes the dilution of metallicity (if $\eta > 1$) or the chemical enrichment (if $\eta < 1$) of the ISM due to galactic winds powered by stellar explosions.

In recent theoretical work, what has been dubbed 'the missing metals problem' has received considerable attention (see Prochaska et al. 2003 for an extensive discussion of this problem), namely the fact that the mean metallicity is ~10 times lower than the value expected from the inferred star formation history. This problem may indicate a serious flaw in our understanding of the interplay between star formation and metal-enrichment. Therefore, we have introduced in our chemical evolution model an extra parameter f_{dil}

accounting phenomenologically for these effects. This allows the metallicity predictions of equation (18) to be adjusted to match observational data. Thus, we rescale the metallicity values given by solving equation (18) by a factor f_{dil}, that is,

$$\tilde{Z}(t) = \frac{Z(t)}{f_{dil}}. \qquad (19)$$

2.6 Summary of model parameters

To summarize, our model is characterized by a set of seven free parameters, that we denote by θ:

$$t\theta = (\log M_{min}, \epsilon, x, \nu, \tau, \eta, f_{dil}). \qquad (20)$$

The free parameters of the model (and the ones that we have chosen to fix) are summarized in Table 1, where we also give the prior ranges for our statistical analysis, that is, the ranges within which their values are allowed to vary (see Section 4.2 for more details).

3 INFLUENCE OF MODEL PARAMETERS ON OBSERVABLE QUANTITIES

In this section, we discuss the impact of each of the seven free parameters in our model (given in equation 20 and in the upper part of Table 1) on the physical observables introduced above, namely the SFR, SNII rate, metallicity and baryonic fraction in structures. We also present a physical interpretation of the observed behaviour of these quantities. As a fiducial model, we fix the parameter values to the following values: $\log M_{min} = 8$, $\epsilon = 0.1$, $x = 1.7$, $\tau = 3$, $\nu = 1.4$, $\eta = 10$ and $f_{dil} = 2$. We then proceed to vary one of the parameters at a time to get a feeling for the physical impact of each of them.

Fig. 1 shows the model dependence on the minimum mass of collapsed dark matter haloes, $\log M_{min}$. Smaller values of this parameter describe a scenario where star formation is hierarchical and follows the growth of structures, while higher values of $\log M_{min}$ correspond to star formation occurring in massive spheroids. Correspondingly, for small $\log M_{min}$ star formation begins earlier, as apparent from the top panel of Fig. 1. At small redshift, a smaller $\log M_{min}$ leads to reduced SFR, since the relatively strong winds
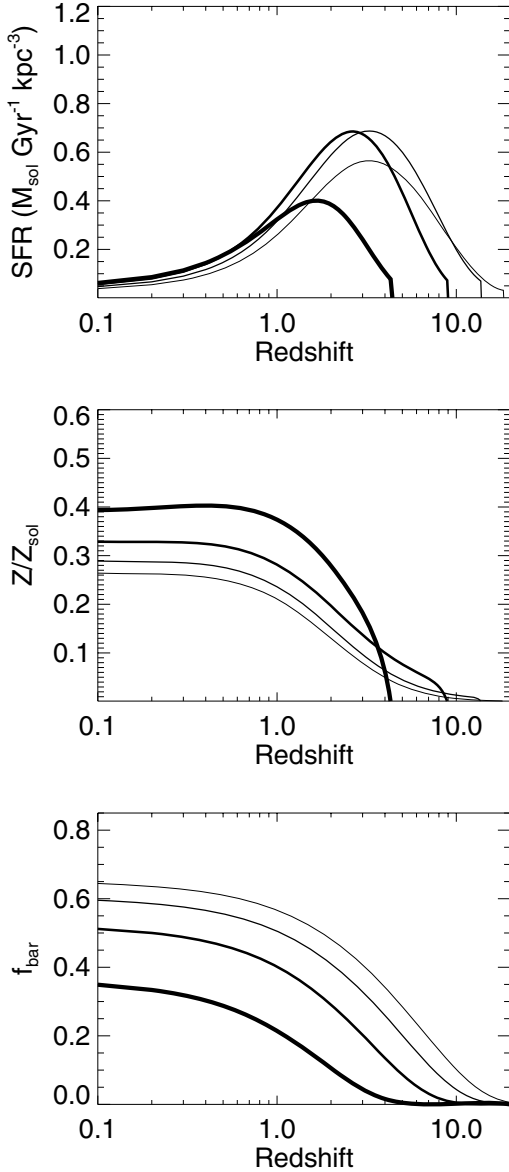
Figure 1. Dependence of the SFR, metallicity and baryonic fraction in structures (panels from top to bottom) on the minimum mass of collapsed dark matter haloes, $\log M_{\min}$. The curves are for $\log M_{\min} = 6, 8, 10$ and 12, from the thin to thick line.

Figure 2. Dependence of the SFR (top panel) and of the metallicity (bottom panel) on the model parameter ϵ, describing the strength of galactic winds. The curves are for $\epsilon = 0.05, 0.15, 0.25$ and 0.35, from the thin to thick line.

($\epsilon = 0.1$ in this example) drive the gas out of the system for shallower potentials. For large $\log M_{\min}$, the build-up of metals is delayed in time but the metallicity can reach larger values, since SN-powered winds are less important in massive systems (middle panel). As it is clear from equation (7), the percentage of baryons allocated to dark matter haloes, $f_{\rm bar}$, increases for decreasing $\log M_{\min}$ (bottom panel). Due to the short lifetime of massive progenitors, the SN rate is essentially identical to the SFR, and we therefore do not display it.

In Fig. 2, we show the model sensitivity to the parameter ϵ, defining the percentage of SN energy that goes to the ISM. This parameter essentially describes the strength of galactic winds. The physical interpretation of high values of ϵ is that strong winds driven by feedback energy are maintained in dark matter haloes. For increasing value of ϵ, galactic winds become stronger and the SFR is
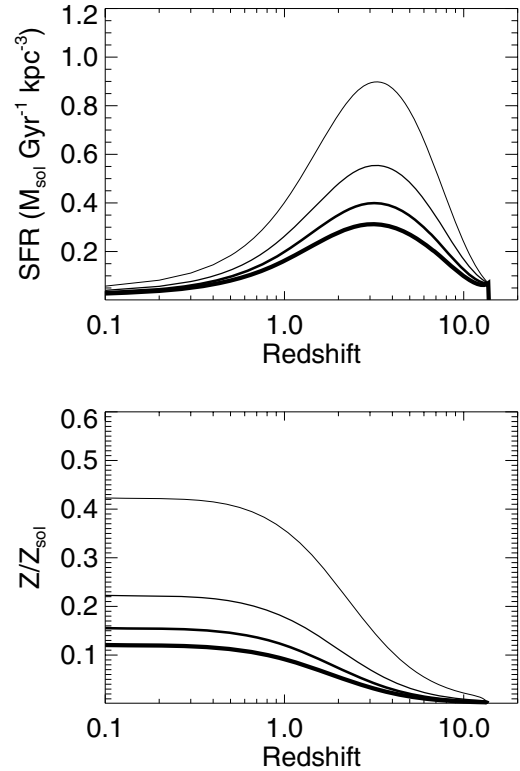
reduced since less gas is available to make stars (top panel). This effect is more important for the shallower gravitational potential of the low-mass haloes, that is, for smaller $\log M_{\min}$ (in this example, $\log M_{\min} = 8$). As already remarked above, higher values of ϵ correspond to winds sweeping out metals from the ISM and thus to lower metallicity (bottom panel). Again, this effect is most important for the low-mass haloes where their gravitational potential is relatively shallow.

The sensitivity to the parameter x, giving the slope of the IMF, is shown in Fig. 3. We can see a strong influence of x on the SNII rates (middle panel), a consequence of equation (17). Decreasing the value of x (i.e. making the IMF shallower) corresponds to a larger number of more massive stars, and hence the SNII rate increases. Taking into account that in our model each SN gives a constant percentage of its energy to the ISM, small values of x result in stronger galactic winds and so in smaller SFRs (top panel) and a less-enriched ISM, hence smaller metallicity (bottom panel). For the extreme case that $x = 1$ (very flat IMF), the SNII rate is very large at high redshift causing very strong winds that reduce the SFR quickly. This causes the spike in Fig. 3 (middle panel).

Fig. 4 shows the model sensitivity to the parameter ν, connected with star formation through the proportionality factor that defines the efficiency of star formation, see equation (16). Increasing the value of ν the star formation becomes more efficient and the ISM becomes highly enriched in metals by evolving stars. On the contrary, smaller values of ν lead to a less-efficient star formation.

The influence of the parameter τ, defining the characteristic time-scale of star formation, is displayed in Fig. 5. Decreasing the value of τ leads to the star formation activity ending sooner and to an ISM which is therefore poorer in metals. Larger values of τ result in an
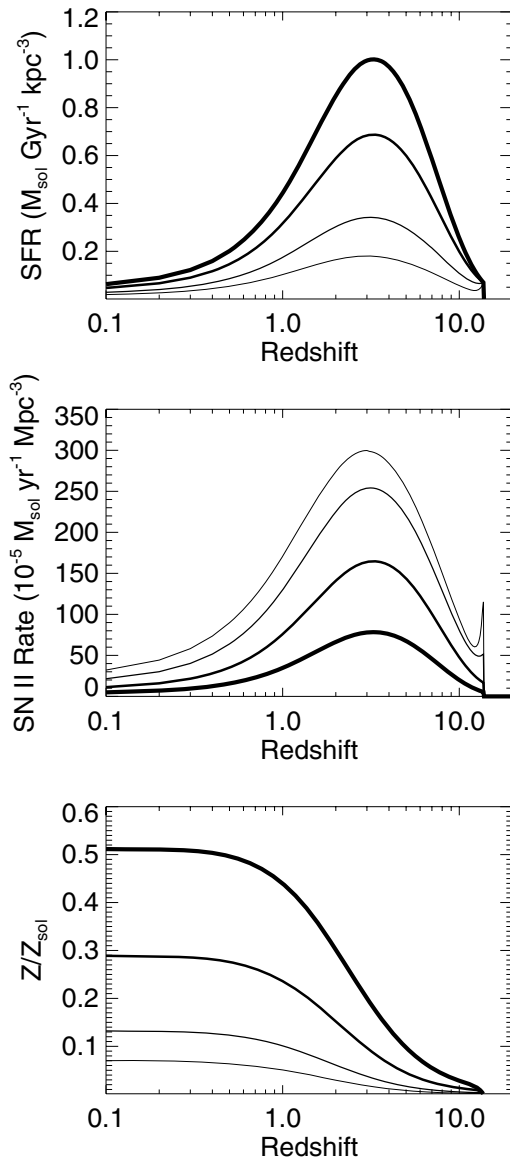
Figure 3. Dependence of the SFR (top panel), SN rate (middle panel) and of the metallicity (bottom panel) on the model parameter x, controlling the slope of the IMF. Curves are for $x = 1.0, 1.4, 1.7$ and 2.0, from the thin to thick line.

enriched ISM since galaxies are active, in terms of star formation, for a longer period.

The influence of the parameter η, controlling the metallicity of the ejecta, is displayed in Fig. 6. A larger value of η leads to a decrease in metallicity of the system, since the metallicity of the winds is increased by a factor of η with respect to the mean metallicity, see equation (18). Finally, we do not display the impact of the dilution factor $f_{\rm dil}$, since its value merely rescales the metallicity by a multiplicative factor, see equation (19).

We now turn to discuss the data employed and the details of our statistical treatment and fitting procedure.

4 DATA AND STATISTICAL ANALYSIS

Combining different types of observations to maximize their constraining power on multidimensional parameter spaces has become
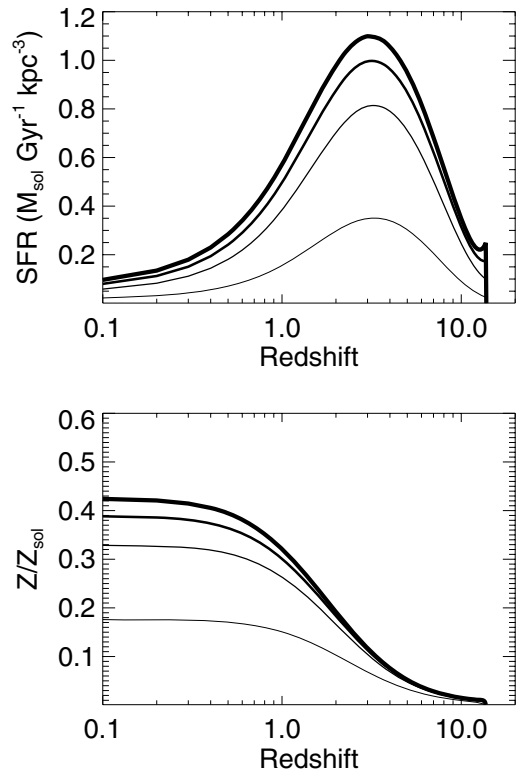

Figure 4. Dependence of the SFR (top panel) and the metallicity (bottom panel) on the parameter ν, controlling the efficiency of star formation. The curves are for $\nu = 0.5, 2, 3.5$ and 5, from the thin to thick line.

a common approach in cosmology. Following an approach similar in spirit, in this work we perform a simultaneous analysis of star formation history, SN rates, metallicity and baryonic fraction data in order to find tight constraints on the parameters of our model, equation (20). One of the aims of this paper is to provide the first complete statistical analysis of existing metallicity, SFR, SNII rate and local collapse baryon fraction data in a realistic model. We first describe the data employed in Section 4.1, and then we outline the Bayesian fitting procedure that vastly improves on usual fixed-grid scans in Section 4.2.

4.1 Observational constraints

The usual compilations of measurements for the SFR and metallicity observations (such as the ones used e.g. in Daigne et al. 2006) are unsuitable for a robust statistical analysis, because of the large systematic differences among measurements at about the same redshift performed over a range of different systems. In fact, when using such a 'raw' data compilation the statistical fit is usually dominated by only a few data points with very small error bars, while the large majority of observations carry almost no statistical weight. This is clearly less than satisfactory. To cure this effect, it becomes important to bin the observations in such a way as to account for possible systematic uncertainties among different measurements at the same redshift. This problem is addressed here for the first time by employing a Bayesian procedure that accounts for possible systematic differences between measurements, based on the treatment given in (Press 1996). The details of the method are given in Appendix A.

We apply the binning procedure described in Appendix A to the data points for the metallicity in the ISM given by Prochaska et al.
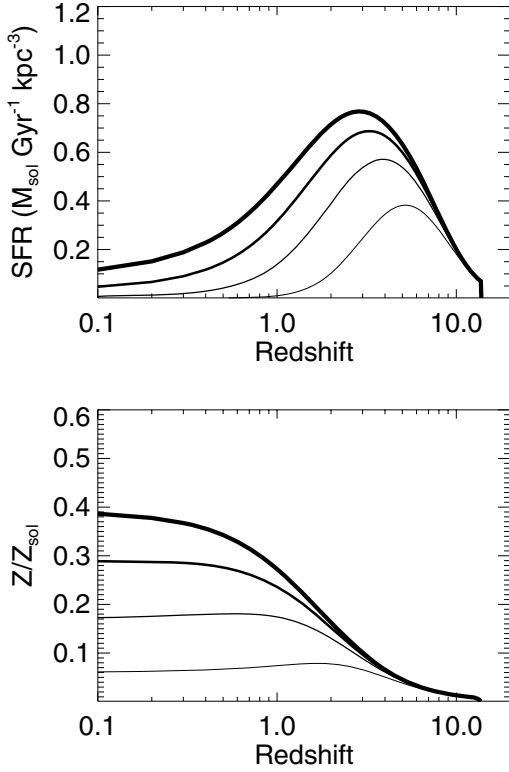
Figure 5. Sensitivity of the model to the parameter τ, giving the characteristic time-scale of star formation. The curves are for $\tau = 1, 2, 3$ and 4, from the thin to thick line.
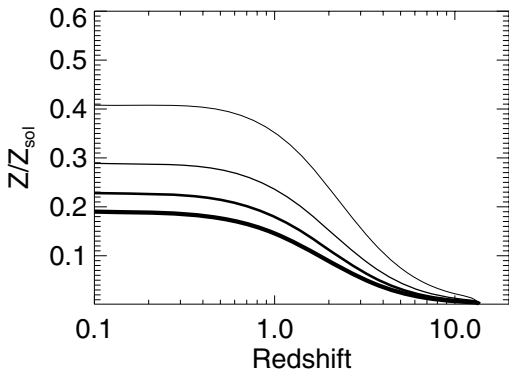

Figure 6. Sensitivity of the metallicity to the parameter η. Curves are for $\eta = 5, 10, 15$ and 20, from the thin to thick line.

(2003). By using equation (A3), we place the 125 measurements in eight bins, ranging in redshift from $z = 0.85$ to 4.45. The bin distribution and spacing have been chosen to obtain a reasonable large number of points in each bin, while simultaneously having a sufficiently small redshift spacing between bins. The measurements of [M/H] number density relative to solar metallicity obtained after the statistical rebinning, are summarized in Table 2.

For the case of cosmic SFR data, our statistical rebinning is modified in order to take into account the redshift uncertainty in the raw data. Details are given in Appendix A2. We take the compilation of 'raw' data out to $z \sim 5$ from Hopkins (2004), excluding only one measurement corresponding to the cosmic star formation at $z = 0.005 \pm 0.005$, reported by Condon (1989). Instead, we replace this

Table 2. Binned measurements of [M/H] number density relative to solar values, after the statistical treatment of the data (see Appendix A for details).

Redshift z_b	Metallicity [M/H]/[M/H]$_\odot$	Number of points
0.85	-0.83 ± 0.11	6
1.45	-1.06 ± 0.09	4
1.95	-0.93 ± 0.12	17
2.45	-1.36 ± 0.14	29
2.95	-1.56 ± 0.19	18
3.45	-1.78 ± 0.08	28
3.95	-1.80 ± 0.06	16
4.45	-1.76 ± 0.11	7

Table 3. SFR density data after our statistical binning of the 'raw' SFR data compilation (see Appendix A2 for details). No dust correction has been applied to these values.

Redshift z_b	SFR density $\log(\dot{\rho}_\star)\,(\mathrm{M}_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3})$
0.012	-1.78 ± 0.18
0.135	-1.45 ± 0.07
0.275	-1.45 ± 0.06
0.405	-1.37 ± 0.23
0.580	-1.08 ± 0.08
0.755	-1.03 ± 0.06
0.905	-0.98 ± 0.07
1.150	-0.92 ± 0.09
1.650	-0.63 ± 0.26
2.520	-0.63 ± 0.21
3.770	-0.79 ± 0.10
5.120	-0.88 ± 0.29

point by more recent measurement at the same redshift as reported by the same author (Condon, Cotton & Broderick 2002). Both these measurements use as cosmic star formation estimator counts at 1.4 GHz. From the raw data, we derive binned values in 12 redshift bins, with centres ranging from $z = 0.035$ to 5.12 by using equation (A6). The resulting bins with their errors are summarized in Table 3.

Furthermore, the SFR predictions of our model are corrected to account for dust absorption. There are large uncertainties associated with dust absorption correction, and this is why at low redshift (i.e. for bins with $z_b \leqslant 3$) we employ both a 'normal dust correction' of 1.0 mag and a 'large dust correction' of 1.8 mag. These two choices are made in view of the fact that they seem to bracket the expected values valid over a broad range of systems (Schiminovich et al. 2005). For bins at a higher redshift ($z_b > 3$), we adopt a fixed dust correction of 0.4 mag, following Schiminovich et al. (2005). We will see in the next section that the dust absorption correction scheme one adopts has a crucial impact on the resulting physical scenario.

The present-day fraction of baryons in structures, as estimated by Fukugita & Peebles (2004), is taken to be

$$f_{\mathrm{bar}}(z = 0) = 0.61 \pm 0.11. \tag{21}$$

The data for the core-collapse SNe are taken from the Great Observatories Origins Deep Survey (GOODS, Dahlen et al. 2004). The GOODS core-collapse SN rates have been placed in two bins at $z = 0.3 \pm 0.2$ and 0.7 ± 0.2. For the local rate (at $z = 0$), we adopt

Table 4. Measurements of SNII rate as a function of redshift.

Redshift z	SNII rate $(10^{-5}\,\mathrm{M_\odot\,yr^{-1}\,Mpc^{-3}})$
0.0	6.16 ± 2.92
0.3	$26.20^{+7.83}_{-9.18}$
0.7	$41.32^{+11.06}_{-10.75}$

the value from Cappellaro, Evans & Turatto (1999). We convert the local rate from SN units as described in Dahlen et al. (2004). The three abovementioned data points are summarized in Table 4.

4.2 Bayesian Markov Chain Monte Carlo analysis

After the statistical rebinning of the data described above, the likelihood function $P(d|\theta)$ is the sum of four independent terms, describing the observations of the SFR, the metallicity, the SN rate and the baryonic fraction:

$$P(d|\theta) = \mathcal{L}_{\mathrm{SFR}} + \mathcal{L}_{\mathrm{met}} + \mathcal{L}_{\mathrm{SN}} + \mathcal{L}_{\mathrm{b}}. \qquad (22)$$

We model each of the above four terms as a product of the data points for each observable, taken to be independent and with Gaussian noise

$$\chi^2_{\mathrm{obs}} = -2\ln\mathcal{L}_{\mathrm{obs}} = \sum_{i=1}^{N_{\mathrm{obs}}} \frac{(y_i - d_i)^2}{\sigma_i^2}, \qquad (23)$$

where 'obs' stands for SFR, metallicity, SN or baryon fraction in structures, and the means d_i and standard deviations σ_i of the data points are given in Tables 2–4 and equation (21). The normalization constant does not matter, as we are only interested in the relative posterior probability density, as we now discuss.

From the likelihood function of equation (23), we obtain the posterior probability for the parameters of interest, $P(\theta|d)$, via Bayes' theorem,

$$P(\theta|d) = \frac{P(d|\theta)P(\theta)}{P(d)}, \qquad (24)$$

where $P(\theta)$ is the prior probability distribution ('prior' for short) and $P(d)$ is a normalization constant that does not depend on the parameters and can therefore be neglected in the following (see Trotta 2007a,b for more details on Bayesian parameter inference and model comparison). We adopt flat (i.e. top-hat) priors on our set of parameters θ given in equation (20) in the ranges given in Table 1, which means that the posterior probability distribution function (pdf) is simply proportional to the likelihood.

In order to explore efficiently our seven-dimensional parameter space, we employ a Markov Chain Monte Carlo (MCMC) procedure, with some of the routines adapted from the publicly available COSMOMC package.[2] The great advantages of MCMC methods are that the computational time scales approximately linearly with the number of dimensions of the parameter space, and that the marginalized posterior distribution for the parameters of interest and their correlations can be simply recovered by plotting histograms of the sample list. We follow the procedure outlined in de Austri, Trotta & Roszkowski (2006), to which we refer for further details. Here, we only briefly sketch the main points.

[2] Available from cosmologist.info.

The aim of an MCMC is to produce a series of samples in parameter space (a Markov Chain) with the property that the density of points is proportional to the probability distribution (the target density) one is interested in mapping, in our case the posterior pdf of equation (24). There are several algorithms that can produce a chain with the required properties. Here, we employ the Metropolis–Hastings algorithm (Metropolis et al. 1953; Hastings 1970): the chain is started from a random point in parameter space, θ_0, and a new point θ_1 is proposed with an arbitrarily proposal density distribution $q(\theta_n, \theta_{n+1})$. The transition kernel $T(\theta_n, \theta_{n+1})$ gives the conditional probability for the chain to move from θ_n to θ_{n+1}, and it must satisfy the 'detailed balance' condition

$$P(\theta_{n+1}\,|\,d)T(\theta_{n+1}, \theta_n) = P(\theta_n\,|\,d)T(\theta_n, \theta_{n+1}) \qquad (25)$$

so that the posterior $P(\theta|d)$ is the stationary distribution of the chain. This is achieved by defining the transition kernel as

$$T(\theta_n, \theta_{n+1}) \equiv q(\theta_n, \theta_{n+1})\alpha(\theta_n, \theta_{n+1}), \qquad (26)$$

$$\alpha(\theta_n, \theta_{n+1}) \equiv \min\left\{1, \frac{P(\theta_{n+1}\,|\,d)q(\theta_{n+1}, \theta_n)}{P(\theta_n\,|\,d)q(\theta_n, \theta_{n+1})}\right\}, \qquad (27)$$

where $\alpha\,(\theta_n, \theta_{n+1})$ gives the probability that the new point is accepted. Since $P(\theta|d) \propto P(d|\theta)\,P(\theta)$ and for the usual case of a symmetric proposal density, $q(\theta_n, \theta_{n+1}) = q(\theta_{n+1}, \theta_n)$, the new step is always accepted if it improves on the posterior, otherwise it is accepted with probability $P(d|\theta_{n+1})\,P(\theta_{n+1})/P(d|\theta_n)\,P(\theta_n)$. The result is a sample list from the target distribution, from which all the statistical quantities of interest can readily be evaluated. Further details about MCMC methods can be found, for example, in MacKay (2003).

Our Bayesian MCMC analysis allows us not only to determine efficiently the best-fitting value of the parameters, but also to explore correlations between the model parameters and estimate marginalized high-probability regions, to which we now turn our attention.

5 RESULTS AND DISCUSSION

As mentioned above, we investigate two different dust-correction schemes for SFR data at low ($z < 3$) redshift, one termed 'normal dust correction' and the other 'high dust correction'. This is expected to roughly bracket the range of possible corrections. The outcome of our analysis is strongly dependent on which dust correction one chooses to employ, with the normal dust correction implying hierarchical star formation, while the high dust correction favours the monolithic scenario.

5.1 Best-fitting models and parameter constraints

The values of the best-fitting model parameters for both dust-correction schemes are given in Table 5, and the corresponding SFR, SN rate, metallicity evolution and baryonic fraction in structures are shown in Fig. 7. The one-dimensional posterior probability distributions (with all other parameters marginalized, i.e. integrated over) are plotted in Fig. 8.

We first discuss the case with the normal dust correction applied. In order to fit the (dust-corrected) SFR at both high and small redshifts, the model requires a small minimal mass ($\log M_{\mathrm{min}} \sim 6$) and strong winds ($\epsilon \sim 0.3$). Although the value of the SN energy transfer parameter is quite large, it is not too far away from theoretical predictions, which give an upper limit of $\epsilon = 0.22$ (Larson 1974). An IMF power-law index $x \sim 1.8$, slightly larger than the Scalo IMF, is also preferred, which translates in fewer available SNe. This is

Table 5. Best-fitting parameter values and marginalized 68 and 95 per cent intervals for the normal (1.0 mag for $z \leqslant 3$) and high (1.8 mag for $z \leqslant 3$) dust corrections. For cases where only an upper or lower limit is found within our prior ranges, we give one-tail intervals. We also give the best-fitting χ^2 and the reduced χ^2, where the number of d.o.f. is 17 (for 24 data points and seven free parameters).

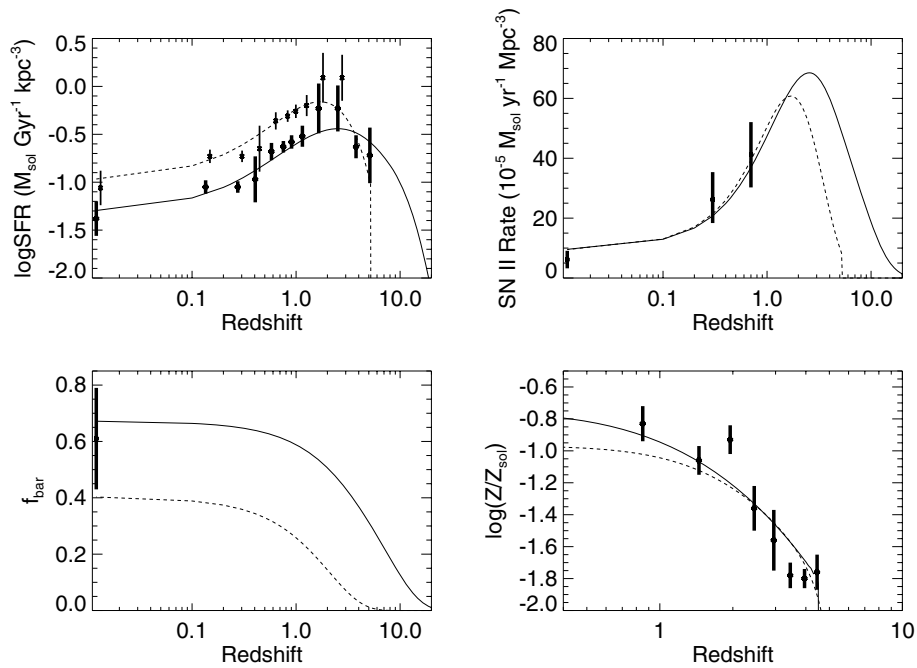| Parameter	Normal dust correction			High dust correction		
	Best fit	68 per cent range	95 per cent range	Best fit	68 per cent range	95 per cent range
$\log M_{min}$	5.17	<6.11	<7.33	11.60	(11.27, 11.61)	(10.97, 11.61)
ϵ	0.32	>0.30	>0.21	0.23	<0.17	0.39
x	1.77	(1.73, 1.82)	(1.68, 1.87)	1.97	>1.95	>1.90
ν (Gyr^{-1})	4.15	>2.20	>1.01	1.81	(1.29, 1.90)	(1.04, 2.25)
τ (Gyr)	3.59	(3.48, 4.22)	(3.23, 4.69)	3.65	(3.45, 4.07)	(3.24, 4.46)
η	8.74	(4.21, 15.70)	(2.24, 24.49)	0.02	<3.80	15.59
f_{dil}	2.72	(1.84, 5.52)	(1.27, 9.34)	20.75	(15.65, 22.34)	(13.06, 26.28)
χ^2	26.60			33.3		
χ^2/d.o.f.	1.6			2.0		

Figure 7. Best-fitting models for the normal (solid line, hierarchical star formation, $\chi^2 = 26.60$) and high (dashed, monolithic scenario, $\chi^2 = 33.3$) dust corrections, with parameters as in Table 5. In the top left-hand panel, showing the SFR, the low-redshift ($z \leqslant 3$) data have been corrected for dust employing a normal dust correction (1.0 mag, lower data points) or a high dust correction (1.8 mag, upper data points). The high dust-correction data have been shifted slightly to the right-hand side for display purposes.

linked to the high value of ϵ, since the energy transfer is so efficient that a large number of SNe are not needed to get the appropriate feedback energy to reproduce the data sets. The metallicity load factor η can be connected with the IMF power-law index x and Dalcanton (2006) gives η values for a variety of IMFs. The value of η for the Scalo (1986) IMF ($x = 1.7$, close to our best-fitting value, $x = 1.77$) is $\eta = 16.8$–18.6, in reasonable agreement with our value, $\eta = 8.74$. This leaves metal-rich outflows as the only viable mechanism for producing the low effective yields observed in gas-rich galaxies. in agreement with suggestions presented in Dalcanton (2006). The dilution factor f_{dil} is of the order of 2, which again is very reasonable, given the complex physics this parameter is supposed to summarize. The value of the χ^2 of the best-fitting model in this case is 26.60 for 17 degrees of freedom (d.o.f.), which suggests that our model captures the essential features of the data. Fig. 7 shows the best-fitting models for normal dust correction (solid line) and high

dust correction (dashed line). Both models provide an acceptable fit to the data, although in the normal dust-correction case the low-redshift metallicity and the present-day baryon fraction in structures appear in better agreement with the data. For a redshift above $z \approx$ 5, the metallicity of the hierarchical model drops very sharply to 0 because of the very significant winds.

Turning now to the high dust-correction case, we note that the preferred values of the parameters in our model are very different from the previous case. Most importantly, a high dust correction at small redshift boosts the value of the SFR for $z \leqslant 3$, and this pushes our model to very large values of $\log M_{min}$, of the order of $\log M_{min} \sim 11$–12. This implies that star formation occurs monolithically in heavy spheroids, as discussed in the Introduction section. We expect dry mergers to play a significant role in the build-up of massive $\log M_{min} \sim 13$ ellipticals in agreement with observations showing that present-day spheroidal galaxies on average have
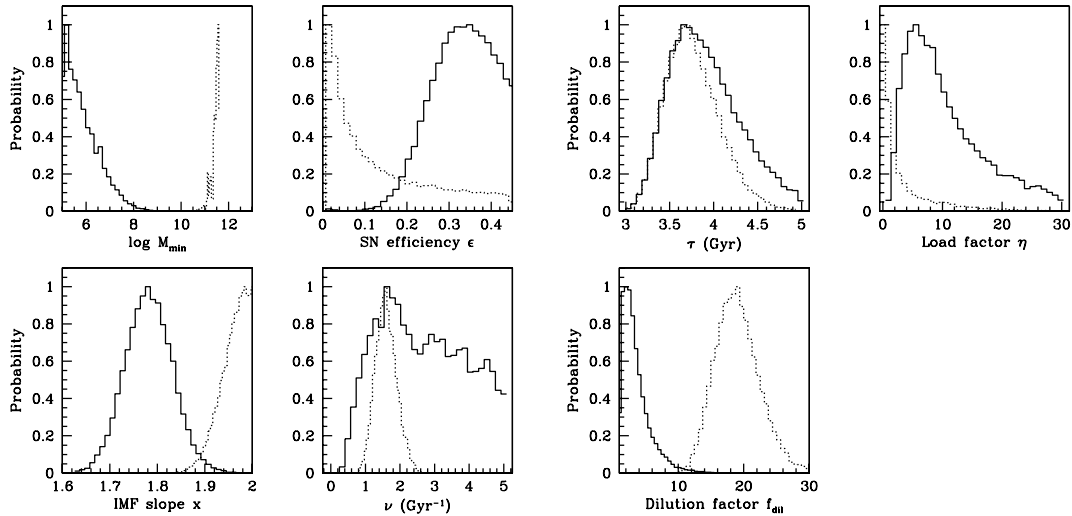
Figure 8. One-dimensional marginalized posterior probability distributions of the model parameters (normalized to their peak values). The solid histograms are for the normal dust-correction case (hierarchical scenario), and the dotted histograms are for the high dust-correction case (monolithic model).

undergone between 0.5 and 2 major dry mergers since $z \sim 0.7$ (Bell et al. 2006). Furthermore, we see from Fig. 7 (dashed curves) that the onset of both the SFR and the metal build-up is significantly delayed in this scenario, until about $z \approx 5$. The SN energy transfer parameter ϵ becomes essentially irrelevant for such large values of the minimum mass, since the potential is deep enough to retain the ejected gas. The peak in the probability distribution for ϵ observed in Fig. 8 is therefore mostly a consequence of a volume effect of our Bayesian MCMC scanning technique. The star formation time-scale $\tau \sim 3.5$ Gyr is in good agreement with theoretical models for Milky Way size disc galaxies (with virial mass $\log M_{vir}$ close to our best-fitting value for $\log M_{min}$). The IMF index is tilted towards extreme values, thus reducing the SN rates but boosting the SFR (cf. Fig. 3). This in turns increases the metallicity, and a large dilution factor, $f_{dil} \sim 20$, is required to bring the predictions in line with observations. We note that this agrees within a factor of 2 with the value already found in previous works on the metallicity of SN ejecta, which was of the order of 10. However, the extremely steep IMF that this model prefers ($x \approx 2$) appears to exclude the possibility that stellar explosions are the main mechanism that drive galactic winds. This is reasonable, since SN-driven gas flows cannot escape from massive galaxies' potential wells. A resolution to this wind dilemma could come from the hypothesis of SMBH-induced out-flows (Silk 2005). In fact, the very low value of the load factor ($\eta = 0.02$) is consistent with this scenario since the SMBH under-goes most of its growth in the gas-rich phase and its outflow expels mostly unprocessed gas. Although our model does not include the physics of SMBHs, it is tempting to say that our best-fitting model suggests that SMBHs should play a key role in the evolution of massive spheroids.

In general, we observe that the high dust-correction case seems to stretch our model parameters to extreme values, suggesting either a strong tension between data sets (mostly SFR and metallicity data) or a failure of the model to fully encapsulate all of the relevant physical processes. Even though with a reduced χ^2 per d.o.f. of 2.0 this scenario is less favoured than the hierarchical star formation model discussed above, it appears that the monolothic formation model cannot be dismissed yet. It is interesting that our seven-parameter model is able to describe both cases, and that the SFR dust correction plays a major role in defining which scenario is preferred.

5.2 Correlations among parameters

We now turn to discuss the most-relevant correlations among the model parameters in light of their physical interpretation and of their impact on the observables, as shown in Section 3. Fig. 9 shows a selection of two-dimensional joint posterior probability distribu-tions for $\log M_{min}, \epsilon, x, \nu, \tau$, and f_{dil}, thus giving complementary information to the one-dimensional distributions plotted in Fig. 8. The contours enclose joint two-dimensional 68 and 95 per cent regions, with all other parameters marginalized, for both the 'high dust-correction' case (dashed) and the 'normal dust-correction' case (solid).
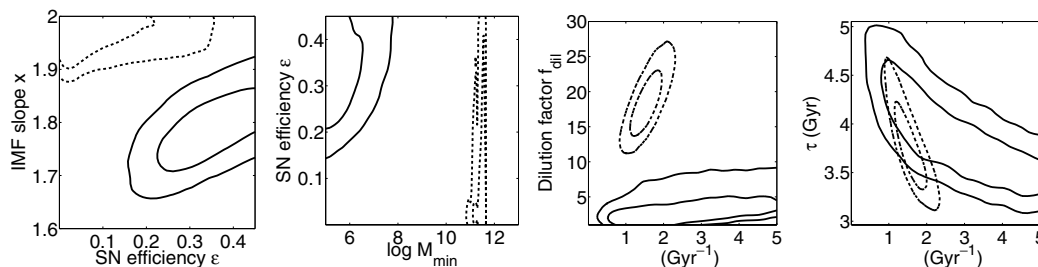

Figure 9. Contours enclosing joint two-dimensional 68 and 95 per cent regions, with all other parameters marginalized, for both the 'high dust-correction' case (dashed, monolithic scenario) and the 'normal dust-correction' case (solid, hierarchical star formation).

In the first panel on the left-hand side of Fig. 9, showing the x–ϵ plane, we observe a positive correlation between the IMF power-law index and the SNII energy efficiency factor. This is expected, since an IMF with a higher power-law index produces less SNe, each of which has to contribute more energy, leading to higher values for the parameter ϵ (cf. Figs 2 and 3, panels showing the SFR and metallicity dependence). For the 'high dust-correction' case (dashed lines), this correlation is weaker, confirming our conclusion that SNeII cannot drive the winds in the massive spheroids. The ϵ– $\log M_{min}$ plane shows that for structures of smaller mass ('normal dust-correction' case, solid lines) the parameter ϵ needs to be large, while for high-mass structures (as preferred in the 'high dust-correction' case, dashed curves), ϵ is essentially unconstrained, indicating that SN feedback is irrelevant for massive spheroids. The different physical processes taking place in small structures and massive spheroids can be further investigated by looking at the correlations in the f_{dil}–ν plane. We expect to find a positive correlation among ν and the dilution factor f_{dil}, as larger ν increases the SFR (cf. Fig. 4), thus leading to a more metal-rich ISM. To bring this back in line with the data, a larger dilution factor is needed. The above line of reasoning explains the strong positive correlation one observes for the high-mass structures (dashed) where winds do not play a strong role and metals cannot escape from the structure. In contrast, metal-rich winds are dominant for smaller structures (solid curves), thus expelling most of the metals produced. This results in almost no correlation between f_{dil} and ν, since the impact of ν on the SFR and metallicity predictions can be mimicked by a different combination of values for ϵ and $\log M_{min}$. Finally, in the extreme right-hand panel of Fig. 9, we display the probability distribution in the τ–ν plane, which exhibits a strong negative correlation. Again, this is expected on the grounds that large values of the parameter ν increase the SFR (cf. Fig. 4) and a smaller time-scale is thus required in order to quench star formation fast enough (see Fig. 5).

6 CONCLUSIONS

We have presented a well-motivated physical model of the cosmic star formation incorporating SN feedback, gas accretion and enriched outflows. We computed the cosmic star formation history and the chemical evolution in the ISM of forming galaxies as a function of redshift, and we presented for the first time a full statistical treatment of the observational data, which accounts for the possibility of systematic errors in the data sets.

We have employed four different observational data – the observed cosmic SFR up to $z \sim 5$, the observed rate of SNeII up to $z \sim 0.7$, the present baryon fraction in structures and the evolution of the metal content in the ISM – to derive constraints on the free parameters of our model. After employing a Bayesian procedure to rebin the SFR and SN rate data, we found that the low-redshift ($z \leqslant 3$) SFR dust correction adopted has a critical impact on the scenario favoured by the data.

For what we have termed 'normal dust correction', the hierarchical star formation model is preferred, where star formation occurs in small structures first and SN winds are important. While the wind load factor remains poorly constrained, we can conclude that larger values are preferred, in agreement with previous work (Dalcanton 2006). Applying a larger dust correction at small redshifts, we found that the data on the contrary favour high values for the minimum mass of a dark halo of the collapsed structures (monolithic star formation scenario). This case requires a large dilution factor, a rather extreme IMF slope and a fairly small wind load factor, as the model parameters are pushed to the boundaries of the available range. We have suggested that this might be interpreted in terms of the presence of outflow from SMBHs, but this possibility will require further investigation. It is worth noting that the monolithic star formation scenario has very little star formation beyond $z \sim 5$. Observations of the E-mode polarization power spectrum of the cosmic microwave background, however, indicate that the Universe was re-ionized around $z \sim 11$ (Spergel et al. 2006). This means that in this scenario the re-ionization mechanism has to be found elsewhere than in massive ultraviolet-emitting stars. Several alternatives have been explored in the literature, for example, re-ionization by decaying particles (Hansen & Haiman 2004), or a high-redshift population of miniquasars that can re-ionize the IGM up to 50 per cent ionization fraction (Dijkstra, Haiman & Loeb 2004).

For both models, the IMF slope is large. Unfortunately, this does not help in distinguishing one model from the other, since observations have so far not yielded convincing results concerning the form of the stellar IMF or its variations in space and time (Scalo 1998). The most important difference among the IMFs is that the fraction of high-mass stars is larger for a shallower IMF. Since only high-mass stars emit significant amount of ultraviolet light, this results in a spectrum which is more shifted towards the ultraviolet for a typical galaxy with a, for example, Salpeter IMF ($x = 1.35$) as compared with a Scalo IMF ($x = 1.7$). In turn, this leads to a different re-ionization history, which can be, in principle, compared with the optical depth to re-ionization as inferred from cosmic microwave background polarization measurements.

While the monolithic scenario is less preferred in terms of quality-of-fit, it is clear that more work is required to be able to draw firm conclusions as to the viability of the two different models. Of particular importance remains the statistical treatment of the data, for which we have here presented a new procedure that we hope will prove useful for future work.

REFERENCES

Bell E. F. et al., 2006, ApJ, 640, 241
Birnboim Y., Dekel A., Neistein E., 2007, MNRAS, 380, 393
Bond J. R., Efstathiou G., 1984, ApJ, 285, L45
Bower R., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Cappellaro E., Evans R., Turatto M., 1999, A&A, 351, 459
Coles P., Lucchin F., 1995, Cosmology. John Wiley & Sons Ltd, New York
Condon J. J., 1989, ApJ, 338, 13
Condon J. J., Cotton W. D., Broderick J. J., 2002, AJ, 124, 675
Croton D. et al., 2006, MNRAS 365, 11
Dahlen T. et al., 2004, ApJ, 613, 189
Daigne F., Olive K. A., Vangioni-Flam E., Silk J., Audouze J., 2004, ApJ, 617, 693
Daigne F., Olive K. A., Silk J., Stoehr F., Vangioni-Flam E., 2006, ApJ, 647, 773
Dalcanton J. J., 2007, ApJ, 658, 941
Dijkstra M., Haiman Z., Loeb A., 2004, ApJ, 613, 646
de Austri R. R., Trotta R., Roszkowski L., 2006, JHEP, 0605, 002
Efstathiou G., 2000, MNRAS, 317, 697

Forster-Schreiber N. et al., 2006, ApJ, 645, 1062

Fukugita M., Peebles P. J. E., 2004, ApJ, 616, 643

Hammer F., Flores H., Elbaz D., Zheng X. Z., Liang Y. C., Cesarsky C., 2005, A&A, 430, 115

Hansen H. H., Haiman Z., 2004, ApJ, 600, 26

Hastings W. K., 1970, Biometrika, 57, 97

Hopkins A. M., 2004, ApJ, 615, 209

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Larson R. B., 1974, MNRAS, 169, 229

Lotz J. et al., 2006, preprint (astro-ph/0602088)

MacKay D., 2003, Information Theory, Inference, and Learning Algorithms. Cambridge Univ. Press, Cambridge

Maraston C., Daddi E., Renzini A., Cimatti A., Dickinson M., Papovich C., Pasquali A., Pirzkal N., 2006, ApJ, 652, 85

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E. J., 1953, Chem. Phys., 21, 1087

Pagel B. E., 1997, Nucleosynthesis and Chemical Evolution of Galaxies. Cambridge Univ. Press, Cambridge

Press W. H., 1996, preprint (astro-ph/9604126)

Press W. H., Schechter P., 1974, ApJ, 187, 425

Prochaska J. X., Gawiser E., Wolfe A. M., Castro S., Djorgovski S. G., 2003, ApJ, 595, L9

Salpeter E. E., 1955, ApJ, 121, 161

Scalo J. M., 1986, Fundam. Cosm. Phys., 11, 1

Scalo J. M., 1998, preprint (astro-ph/9811341)

Schiminovich D. et al., 2005, ApJ, 619, L47

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Silk J., 2005, MNRAS, 364, 1337

Spergel D. N. et al., 2007, ApJS, 170, 377

Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, ApJ, 621, 673

Trotta R., 2007a, MNRAS, 378, 72

Trotta R., 2007b, MNRAS, 378, 819

Worthey G., Faber S. M., Gonzalez J. J., 1992, ApJ, 398, 69

APPENDIX A: BINNING OF DATA ACCOUNTING FOR UNDETECTED SYSTEMATICS

A1 No redshift uncertainty

We wish to define B bins in redshift space. Within each bin b, $1 \leqslant b \leqslant B$, we have a collection of measurements (in our case, metallicity or SFR data), each with its own statistical accuracy and possibly an unspecified systematic error. The fact that systematic differences above the quoted statistical errors dominate the raw data is apparent from a plot of the unbinned metallicity or SFR observations, that show a scatter of up to an order of magnitude for observations at about the same redshift. The origin of the systematic discrepancy can vary from underestimated statistical errors in the observation to intrinsic dispersion in the observed systems to differences in the way the data are collected. In the presence of systematic errors, we cannot simply take the weighted average of the data within each bin. Instead, we model the presence of unknown systematics as follows.

Let us consider the measurement of a quantity y_b in a top-hat bin b, $1 \leqslant b \leqslant B$ – in our case, this represents the metallicity value at the redshift of the bin, z_b, and we assume that we can neglect the redshift uncertainty of the measurements (this issue is addressed in the next section). Each measurement consists of a central value d_i and a statistical error σ_i, $1 \leqslant i \leqslant N_b$, for N_b different measurements within bin b. If the ith datum does not suffer from a systematic error (or where the systematic error, S_i, is negligible compared with the quoted statistical error), the likelihood function is modelled as a Gaussian with the quoted standard deviation σ_i:

$$P_{i,\mathrm{g}}(d_i|y_b) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{d_i - y_b}{\sigma_i}\right)^2\right]. \tag{A1}$$

For the sake of brevity, let us denote such measurements as 'good' measurements, as indicated by the subscript 'g'. If the datum suffers from an undetected systematic, that is, the dominant error is $S_i \gg \sigma_i$, the likelihood is instead given by (neglecting the statistical error with respect to the systematic one):

$$P_{i,\mathrm{s}}(d_i|y_b) = \frac{1}{\sqrt{2\pi}S_i} \exp\left[-\frac{1}{2}\left(\frac{d_i - y_b}{S_i}\right)^2\right], \tag{A2}$$

where the subscript 's' denotes 'systematics', or 'spoiled' measurements, for brevity. Now of course we do not know which measurements suffer from systematic, but this can be determined statistically using the following procedure (adapted from Press 1996).

We denote by p the probability that each of the measurements i in bin b is a 'good' one. Conversely, $1 - p$ is the probability that the datum suffers from systematics. Furthermore, we include a binary vector $\mathcal{V} = (\mathcal{V}_1, \dots, \mathcal{V}_{N_b})$, whose elements $\mathcal{V}_i(1 \leqslant i \leqslant N_b)$ can be either 0 or 1, determining whether the datum i is a good one (for $\mathcal{V}_i = 1$) or a spoiled one (for $\mathcal{V}_i = 0$). We can then compute the posterior probability for the value of the observed quantity y_b in bin b by multiplying the individual contributions of the measurements in the same bin and marginalizing over the unknown quantities p and $\mathcal{V}$ (see equation 16 in Press 1996):

$$P(y_b|d_b) \propto \int \mathrm{d}p \prod_{i=1}^{N_b} [p P_{\mathrm{g},i} + (1 - p)P_{\mathrm{s},i}], \tag{A3}$$

where d_b denotes the collection of measurements in bin d, that is, $d_b = (d_1, \dots, d_{N_b})$. For the prior probability on p, we have assumed a flat prior distribution between $0 \leqslant p \leqslant 1$ and the proportionality factor might be determined by requiring that the likelihood be normalized to unity, but this is not necessary in our application. The precise numerical value of the error associated with systematics, S_i, does not matter, as long as $S_i \gg \sigma_i$. In our case, we take S_i to be unity on a log scale, corresponding to one order of magnitude uncertainty on the observable.

From the posterior distribution (A3), the central value of the bin b is obtained as the peak of the distribution, while the standard deviation is defined as the range enclosing 68.4 per cent (1σ range) of the probability. These values are given in Table 2 for the metallicity data, and are then used for the likelihood function employed in the fit of the model. Of course, one could as well employ the full probability distribution of equation (A3) as the likelihood function, but for simplicity we have summarized it as a Gaussian with mean and standard deviation computed as described above. The collection of raw, unbinned metallicity data and the resulting bins from our statistical treatment are shown in Fig. A1.

A2 Accounting for redshift uncertainty

When the observed quantity suffers from a substantial redshift uncertainty, as in the case of the SFR data, we need to take into account the redshift error in our binning procedure, as this introduces a further uncertainty as to which bin a given datum belongs to. The above procedure is then modified as follows.

The probability that an observation with central redshift z_i and redshift uncertainty τ_i belongs to the bth redshift bin (centred at
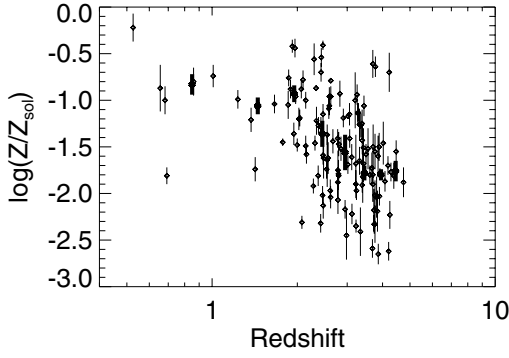
Figure A1. Raw metallicity data and the binned values after the statistical treatment.
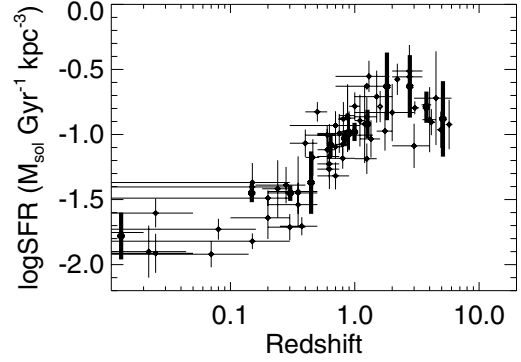

Figure A2. Raw SFR data and the binned values after the statistical treatment including redshift uncertainties. No dust correction has been applied to the data at this stage.

redshift z_b) is modelled as a Gaussian, that is,

$$P(z_i|z_b) = \frac{1}{\sqrt{2\pi}\tau_i} \exp\left[-\frac{1}{2}\left(\frac{z_i - z_b}{\tau_i}\right)^2\right]. \quad (A4)$$

Given the uncertainty on the location of the measurements in redshift, it is now impossible to assign data points to top-hat bins. Instead, one needs to marginalize over all possible assignments of data points among redshift bins, with each point's contribution weighted by the conditional probability of equation (A4). For each bin, let us introduce a new binary variable, $\mathcal{Z} = (\mathcal{Z}_i, \ldots, \mathcal{Z}_N)$, whose elements indicate whether the ith datum ($1 \leqslant i \leqslant N$) belongs to the bin under consideration ($\mathcal{Z}_i = 1$) or not ($\mathcal{Z}_i = 0$). If we knew which datum belongs to which redshift bin, then we could assign an exact binary sequence to $\mathcal{Z}$ (this corresponds to the case considered in the previous section). Instead, we sum (marginalize) over all possibilities, writing for the posterior probability of the SFR value y_b at redshift z_b, given d, the full collection of data points at all redshifts

$$P(y_b|d) = \sum_{\mathcal{Z}} P(y_b, \mathcal{Z}|d) = \sum_{\mathcal{Z}} P(y_b|\mathcal{Z}, d) P(z_i|z_b, \mathcal{Z}_i = 1),$$
$$(A5)$$

where the conditional probability $P(y_b|\mathcal{Z}, d)$ is given by equation (A3), given a specific assignment for $\mathcal{Z}$. The sum over $\mathcal{Z}$ can be

replaced by a product of binomial terms, so that we finally obtain, using equations (A3) and (A4),

$$P(y_b|d) \propto \int \mathrm{d}p \prod_{i=1}^{N} \{[p P_{g,i} + (1-p) P_{s,i}] P(z_i|z_b)\} - 1. \quad (A6)$$

Note that the product is here over all points in the data set, not just over the ones in a bin, as in equation (A3).

Since we include the full data set for each bin, the resulting errors are, in principle, correlated across bins. However, the Gaussian term of equation (A4) ensures that only 'nearby' points give a non-negligible contribution to the value of bin b. We therefore consider it acceptable to ignore the correlation among bins when using the mean and standard deviation of equation (A6) for the likelihood function for the SFR. The results from this procedure are tabulated in Table 3, and are plotted alongside with the raw, unbinned data in Fig. A2.

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