Mon. Not. R. Astron. Soc. **366,** 858–864 (2006) doi:10.1111/j.1365-2966.2005.09883.x

Pure luminosity evolution models: too few massive galaxies at intermediate and high redshift

M. G. Kitzbichler★ and S. D. M. White
Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching bei München, Germany

Accepted 2005 November 17. Received 2005 November 16; in original form 2005 September 20

ABSTRACT

We compare recent galaxy data at low and high redshift to pure luminosity evolution (PLE) models, which assume that massive galaxies were assembled and formed most of their stars at high redshift ($z > 3$) and have evolved without merging or substantial dust obscuration since then. Previous studies, typically comparing to only one or a few different PLE parametrizations, painted a diverse picture of whether or not the evolution of bright early-type galaxies is consistent with such models. Here we attempt to gain further insight by exploring a wider parameter space. Our models span the full range of plausible metallicities, initial mass functions (IMFs) and star formation histories. We require them to reproduce the abundance of galaxies by colour and luminosity in the Sloan Digital Sky Survey, and we investigate whether they can simultaneously fit (i) the observed galaxy counts as a function of redshift in magnitude-limited surveys with $K < 20$, and (ii) the colour and M/L ratio evolution of red-sequence galaxies in clusters. All models that are consistent with (ii) predict galaxy counts at $1.5 < z < 3$ that lie above the observations. This finding does not change with the incorporation of moderate dust extinction, confirming previous studies, which concluded that, for an IMF slope similar to the Salpeter value, such models lie far above the data. The progenitors of most present-day massive galaxies must be much more heavily extincted than currently known galaxies at $z \geqslant 1.5$ to match the observed counts at these redshifts. Alternatively, the majority of massive galaxies may have assembled at later redshifts as suggested by some hierarchical formation models.

Key words: galaxies: elliptical and lenticular, cD – galaxies: evolution – galaxies: formation – galaxies: general – galaxies: luminosity function, mass function.

1 INTRODUCTION

A very wide range of evolutionary histories appear consistent with the observed properties of the present-day population of galaxies. The simplest and most conservative assumption may be that most galaxies were assembled at some early time and their differing stellar populations reflect differing subsequent star formation histories. Massive galaxies – big ellipticals, S0 galaxies and early-type spirals – appear to be dominated by old stellar populations, so their star formation rates (SFRs) must have been high at early times and must thereafter have declined steeply. Many less massive galaxies – late-type spirals and irregulars – show evidence for substantial recent star formation, so their SFRs may have varied much less. The light of some is clearly dominated by stars from a recent burst.

The recent evolution of the galaxy population in such a scenario can be modelled by adopting the backwards-in-time technique first introduced by Tinsley (see Tinsley 1980). This requires three main ingredients: the present-day luminosity function (LF) of galaxies divided by morphological type (or better by colour); a parametrization of the mean star formation history (SFH) for each type (or colour class); and a global cosmological model to relate times, distances and redshifts. The SFH is fed into stellar population synthesis models, which determine how the luminosities and colours of each type evolve with time. These can then be combined with the cosmological model to predict counts of galaxies as a function of apparent magnitude, observed colour and redshift.

Kauffmann & Charlot (1998, hereafter KC98) compared available data to the redshift distribution predicted for complete K-band-limited galaxy samples by such pure luminosity evolution (PLE) models assuming an Einstein–de Sitter cosmology. They found the models to overpredict counts at redshifts $z > 1$ by a large factor. Since then, a number of similar studies have updated the cosmological model to the current concordance cosmology and have presented new observational samples that cover wider areas or go significantly deeper. While the improved observations have reduced the statistical uncertainties, they have not substantially changed the redshift distributions from those used by KC98. The change to ΛCDM

★E-mail: mgk@mpa-garching.mpg.de (MGK); swhite@mpa-garching.mpg.de (SDMW)

significantly reduced the discrepancy, however, by bringing down the number of high-redshift objects predicted at a given K magnitude.

Fontana et al. (1999) published a study based on photometric redshifts for a $K \leqslant 21$ sample of 319 galaxies in several small fields. Despite using a ΛCDM model, their conclusion agreed with that of KC98: the observed redshift distribution disagreed with their PLE model. Rudnick et al. (2001) found the same result when comparing a range of published PLE models with their photometric redshifts for 95 galaxies with $K_{s,AB} \leqslant 22$ in the *Hubble Deep Field* South. In part II of a series of papers on the Las Campanas Infrared (LCIR) Survey, Firth et al. (2002) present photometric redshifts for 3177 galaxies down to $H \leqslant 20$. They compare these to a number of different PLE models, and again find the abundance of high-redshift objects to be overpredicted. All these studies echoed the conclusion of KC98 that the data suggest that many present-day massive galaxies were assembled at relatively low redshift.

Other recent work based on similar data disagrees with this conclusion. Kashikawa et al. (2003) and Cimatti et al. (2002a) both compare to a modified 'PLE' model by Totani et al. (2001). This incorporates a metallicity-dependent dust treatment and high-z selection effects, as well as a simplified parametrization of mergers (so it is not strictly a PLE model). The galactic-wind model adopted for the formation of ellipticals (Arimoto & Yoshii 1987) assumes an initial starburst phase with a dust optical depth of $\tau = 10$ in B band. This corresponds to an extinction of 5 mag at 1.1 μm (observer-frame K band at $z = 1$) even after 1 Gyr (see fig. 1 in Totani & Yoshii 2000). This model is able to fit the observed redshift distributions because its large assumed dust extinction hides most massive galaxies at redshifts beyond around $z = 1.5$, even in observed K band.

In this paper we are primarily concerned with traditional PLE models in which mergers are neglected and extinction is assumed weak, in particular for massive galaxies after their initial burst of star formation is complete. We will, however, comment briefly on the effects of dust in later sections.

One of the most recent studies comparing PLE predictions to the redshift distributions of K-selected samples is that of Somerville et al. (2004), who found that, although such models overproduce the counts at high redshift, the discrepancy is quite modest. They took advantage of the newly acquired K20 and GOODS survey data, which we also use here, together with other recent high-quality survey data, for comparison to our own PLE models. As we will see, our conclusions do not agree with those of Somerville et al. (2004) even for similar models.

More involved studies of number density evolution as a function of galaxy type yield similarly controversial results. Im et al. (2002) found that DEEP survey observations in the Groth strip are consistent with PLE and also with a minor merger scenario out to $z = 1$, as long as major star-forming bursts in this redshift interval are excluded. Using LCIR data, Chen et al. (2003) performed a study of the number density evolution of galaxies by comparing the LF in R band at four different redshifts. They give estimates for the evolution of the comoving luminosity density l_R in the interval $0.3 < z < 1.5$ of at the most $\times 3$ for L^* galaxies and $\times 6$ for $1.6\,L^*$ galaxies. Pozzetti et al. (2003) based their work on the K20 survey, which is also used in this paper, finding that, out to intermediate redshifts of about 1–1.5, PLE models are consistent with the observations, and that in this redshift range the number density of E/S0 galaxy types decreases at the most by 30 per cent. In the most comprehensive number density evolution study performed recently, Bell et al. (2004) took advantage of the wide area covered by the COMBO-17 survey. They conclude that the colour of red galaxies at a given rest-frame magnitude becomes bluer with redshift, consistent with passive ageing of stellar populations, but that the stellar mass on the red sequence has increased at least by a factor of 2 since $z = 1$. This appears consistent with a hierarchical buildup of stellar mass by mergers in a ΛCDM universe. Most recently, Saracco et al. (2005) identified seven bright massive galaxies in the MUNICS survey at redshifts beyond 1.2 which look already evolved. This is consistent with no evolution of the number density of E/S0 galaxies out to a redshift of $z = 1.7$, at the same time putting the formation of these galaxies to redshifts well beyond $z_{\rm f} = 2$. Despite their relatively poor statistics, these authors conclude that massive ellipticals did not form recently, and argue that this finding contradicts the hierarchical model of galaxy formation. Other papers testing this hypothesis are Cimatti et al. (2004) and Glazebrook et al. (2004), the latter finding that only one-third of present-day massive galaxies were present at $z = 1.8$. Although many hierarchical models of galaxy formation predict even fewer galaxies at this redshift, this is not an intrinsic problem of hierarchical growth, but rather a reflection of the specific star formation recipes adopted.

In this paper we investigate a number of traditional PLE models spanning the full plausible range of metallicity (Z), initial mass function (IMF) and star formation history (SFH). Section 2 describes how our models are set up to reproduce the present-day LFs as a function of colour in the Sloan Digital Sky Survey (SDSS) (Section 2.1) and how various different SFRs and metallicities are assigned to the different colour classes (Section 2.1) in order to follow their luminosity evolution backwards in time. We establish the range of allowed parameters and present five models to illustrate the resulting range of evolutionary predictions. We check that our models reproduce the local K-band LF, as observed by the 2MASS survey (Section 2.2) as well as the passive evolution of colour and M/L ratio observed for cluster elliptical galaxies. In Section 3 we compare the predictions of these models with counts as a function of redshift in recent deep K-selected surveys.

Finally, in the concluding section we discuss possible interpretations of our primary result, that there are fewer luminous galaxies observed at $z \gtrsim 1.2$ than are expected on the basis of traditional PLE models. One possibility is that much more dust obscures the majority of massive high-redshift galaxies than is present in the galaxies that have so far been observed. Alternatively, many present-day massive galaxies simply were not yet assembled by $z \sim 2$.

2 THE MODELS

As mentioned above, traditional PLE models require knowledge of the present-day LFs of galaxies as a function of their colour. For each colour class, an SFH model is assumed that reproduces its $z = 0$ colour, and this SFH is then used to predict the LF and the spectral energy distribution (SED) of galaxies of this class at all earlier times. Combining the different classes, galaxy counts can then be predicted as a function of observed magnitude, colour and redshift in any observed photometric band for any assumed cosmological model. In the following we adopt the cosmological parameters of the present standard concordance cosmology: $\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\ {\rm km\,s^{-1}\,Mpc^{-1}}$.

2.1 From the local LF to the models

Our PLE models are normalized to the LFs at redshift $z_{\rm LF} = 0.1$ recently obtained by Blanton et al. (2003) from the data of the SDSS survey (York et al. 2000). For practical reasons, they give the LF in blueshifted SDSS magnitudes corresponding to the filter wavebands at redshift $z = 0.1$, denoted ${}^{0.1}u$, ${}^{0.1}g$, etc. For our purposes the great advantages of these data are their high quality, their superb statistical precision and the fact that they are given in colour–luminosity space (see Fig. 1). We separate the data distribution into five colour ranges and calculate the parameters (see Table 1) for a Schechter function fit to the LF of each colour bin independently. These parametrized LFs are shown in Fig. 2.

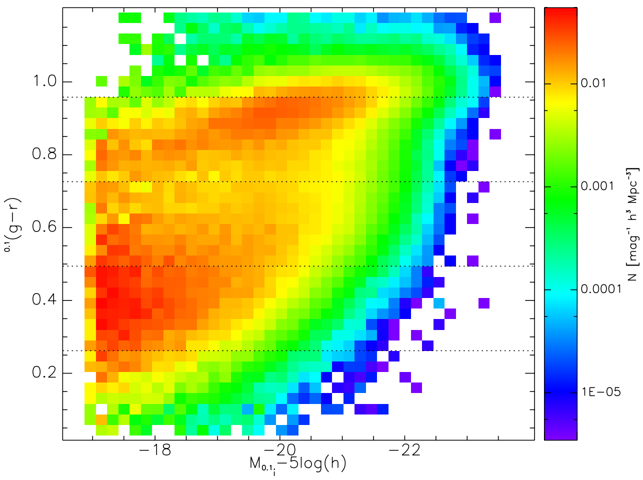

Figure 1. Two-dimensional luminosity functions by colour and absolute magnitude (using data from Blanton et al. 2003). The horizontal black dotted lines indicate the colours separating our different colour classes.

Table 1. Definition of the different galaxy types according to their colour. The parameters of the Schechter function fits to the respective LFs are also given here.

Type	Colour ${}^{0.1}(g-r)$		LF – Schechter fit		
	Mean	Range	Φ (Mpc $h^{-1})^{-3}$	α	${}_{0.1i}M_*$
1	1.01	0.96–1.19	2.377×10^{-3}	−0.11	−20.96
2	0.87	0.73–0.96	8.406×10^{-3}	−0.60	−20.61
3	0.61	0.49–0.73	5.169×10^{-3}	−0.89	−20.49
4	0.40	0.26–0.49	4.382×10^{-3}	−1.29	−19.84
5	0.20	0.03–0.26	9.596×10^{-4}	−1.51	−19.11

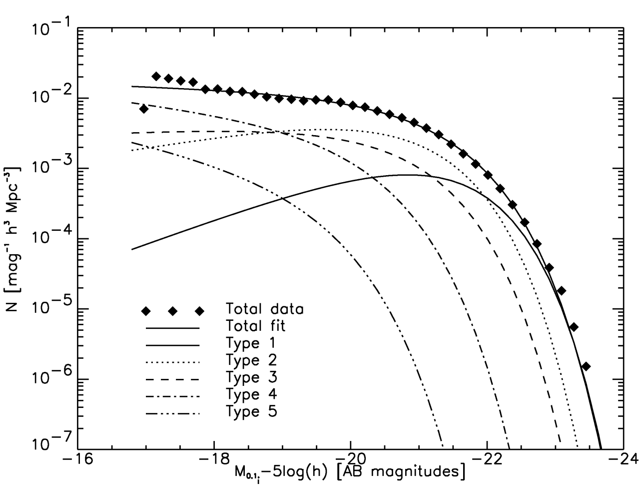

Figure 2. Schechter function fits to the luminosity functions of SDSS galaxies in our five different colour classes (see also Table 1).

We use the fits of Fig. 2 to construct PLE models as described in Gardner (1998) – except for the slight complication that $z_{\mathrm{LF}} = 0.1$. The five colour classes are identified with five SFHs that reproduce their broad-band colours according to the stellar population synthesis models of Bruzual & Charlot (2003). For each galaxy type the spectrum and the LF can then be evolved backwards in time in order to predict the properties of the galaxy population at earlier redshifts.

The assignment of SFH to present-day colour is far from unique, so we construct a variety of possible models differing in their IMF, metallicity, formation redshift z_{f} (defined as the redshift when stars start to form) and e-folding time-scale τ for an assumed exponentially declining SFR. We assume all colour classes to have the same z_{f}, except for the bluest one, which often cannot be fitted by any exponentially declining SFR. This is a particular problem for models with a steep IMF. In such cases we assume an SFH with constant SFR seen at a fixed age, implying *no* evolution with redshift. This is the standard fix for this problem, which is, in any case, irrelevant for the questions we study here.

We limit the range of allowed parameters in our PLE models by requiring consistency with the observed, apparently passive evolution of bright early-type galaxies in clusters. We require the *B*-band mass-to-light ratio of our reddest colour classes to evolve similarly to the measurements of van Dokkum & Stanford (2003). As the left three panels in Fig. 3 show, this mainly constrains the slope of the IMF, given that one has considerable freedom in the choice of the formation redshift z_{f}. IMFs with a power-law exponent of $x = 2.0$ (where the Salpeter exponent is $x = 1.35$) are excluded, except possibly for the lowest formation redshifts. We nevertheless adopt this slope for model 4 below in order to study its implications. We

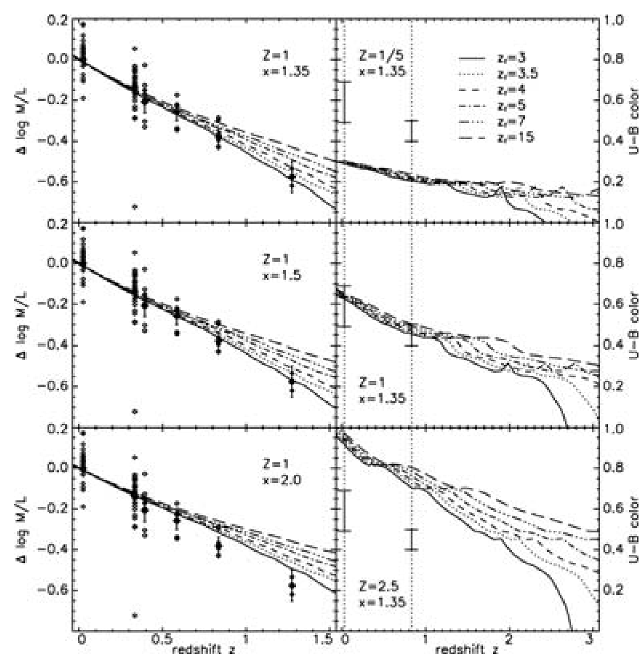

Figure 3. Left: Evolution of the mass-to-light ratio of cluster ellipticals in the *B* band (using data given by van Dokkum & Stanford 2003). The smaller open symbols denote individual galaxies, while the larger filled symbols stand for data averaged over a number of massive galaxies in a cluster. The model predictions are shown for different z_{f} and IMF slopes ranging from $x = 1.35$ at the top to $x = 2$ at the bottom. Right: Rest-frame $U - B$ evolution of model early-type galaxies compared to the rest-frame $U - B$ colours of cluster ellipticals at $z = 0.87$ (MS 1054−03) and at $z = 0.023$ (Coma). Model predictions are shown for different z_{f} and for three metallicities, 0.2, 1.0 and 2.5 $Z_{\odot}$ from top to bottom.

Table 2. Definition of the different models. The parameters given are: x, slope of the IMF; $z_{\rm form}$, formation redshift; and τ, exponential fall-off time of the SFR.[a]

Model	0	1	2	3	4
x	1.35	1.5	1.35	1.5	2.0
$z_{\rm form}$	15	15	3.5	3.5	3.5
τ_1	1.5	2.0	1.5	1.5	1.5
τ_2	3.0	3.0	2.5	2.5	3.0
τ_3	6.0	10.0	5.0	7.0	30.0
τ_4	∞	$^*\infty$	∞	$^*\infty$	$^*\infty$
τ_5	$^*\infty$	$^*\infty$	$^*\infty$	$^*\infty$	$^*\infty$

[a]Here ∞ means constant star formation rate; and the * denotes galaxy types without evolution.

note that recent work on IMFs at high redshift have tended to argue for $x < 1.35$ ('top-heavy IMFs') in order to explain the high luminosities of submillimetre luminous galaxies and the apparently high aggregate metal yields of early generations of stars (see Nagashima et al. 2005).

We also require the rest-frame $U - B$ colours of the reddest colour class to match those of bright ellipticals in two clusters, the Coma cluster at $z = 0.023$ and MS 1054−03 at $z = 0.87$ (Gavazzi, Boselli & Kennicutt 1991; van Dokkum et al. 1999). This allows only a narrow range of metallicities for these bright early types, namely approximately solar, as can be seen from the three right-hand panels in Fig. 3, which show the evolution in rest-frame colour for stellar populations of given metallicity formed with a Salpeter IMF in a single burst at a variety of redshifts. IMF variations have very little effect on this colour since it is dominated by main-sequence turnoff stars (as explained by Bruzual & Charlot 2003).

We present results for five representative models that are at least marginally consistent with all these constraints. Their parameters are summarized in Table 2 and were selected to cover the whole range of permitted values.

2.2 The *K*-band LF as a consistency test

The LFs used here were measured in the rest-frame $^{0.1}i$ band. We can check the reliability of our stellar population models for the five colour classes by using them to predict the K-band (2.2 μm) LF of local galaxies. This is of particular interest because near-infrared light is a relatively good tracer of stellar mass, depending only weakly on dust content and SFH. We therefore compare the present-day K-band LF produced by our models to the observed function as given by Kochanek et al. (2001). As can be seen in Fig. 4, models and data agree reasonably well apart from a slight magnitude offset, perhaps ~0.15 mag, at the bright end. This is likely to be due to the rather bright isophotal magnitudes used by Kochanek et al., in contrast to the surface-brightness independent Petrosian magnitudes of the SDSS survey. The difference is most pronounced for elliptical galaxies with de Vaucouleurs-type surface brightness profiles. These dominate the bright end of the LF (see also Andreon 2002).

3 COMPARISON OF *K*-BAND SELECTED REDSHIFT DISTRIBUTIONS

In this paper we compare to the same deep surveys as Somerville et al. (2004), namely GOODS CDF-S covering about 160 arcmin^2 with photometric redshifts obtained by Mobasher et al. (2004),

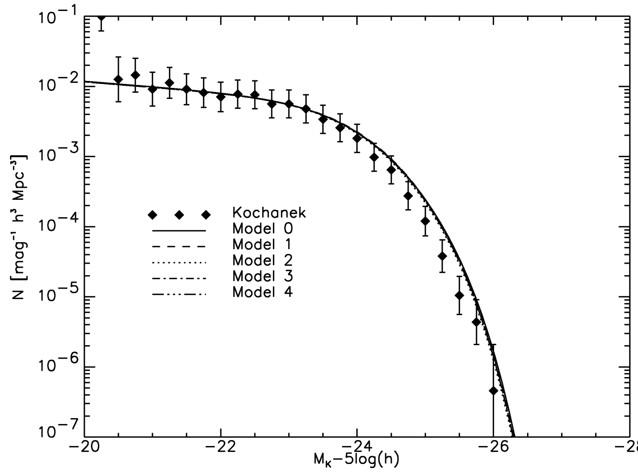

Figure 4. Comparison of our model K-band LFs with the one that Kochanek et al. (2001) derived from 2MASS data. The slight offset at bright magnitudes can be accounted for by differing magnitude definitions in the SDSS and 2MASS surveys.

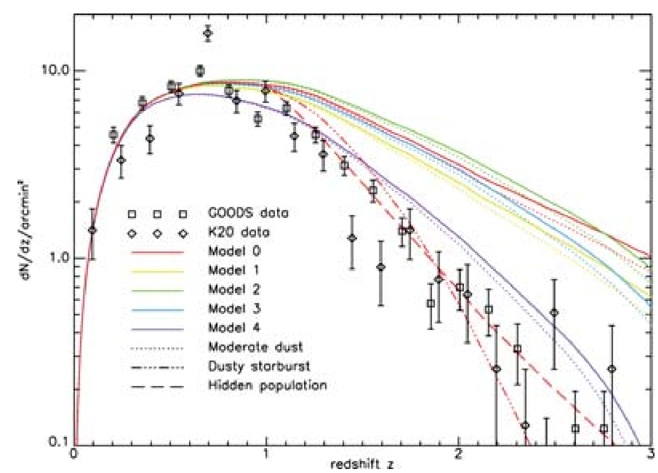

Figure 5. Differential redshift distributions for $K < 20$ galaxies. The errors plotted on the observational data points are approximate Poisson errors. Our five PLE models without dust are shown as full curves of different colours as indicated in the figure. Dotted lines denote the same models with moderate dust extinction. The effects of much more substantial extinction are illustrated for model 0 only, denoted by the dot-dashed curve (dusty starburst) and the dashed curve (hidden population). See text for details.

and K20 carried out in a smaller area of the same field covering 52 arcmin^2 but providing spectroscopic redshifts rather than photometric ones (Cimatti et al. 2002b). The differential distribution of galaxies per arcmin^2 and per unit redshift interval is shown in Fig. 5 for both data sets, binned to $\Delta z = 0.15$ and with Poisson error bars. Clearly there is some substructure in these distributions as a result of the relatively small fields surveyed. In particular, at $z \sim 0.7$ there is a prominent peak in the K20 data. This feature is still visible in Fig. 6, the cumulative redshift distribution of galaxies. In a larger comoving volume, such fluctuations should average out consistent with the smoother curves obtained for the somewhat larger GOODS survey.

Superposed on the observational data in Figs 5 and 6 we show the differential and cumulative redshift distributions predicted by the various models specified in Table 2. In addition, three

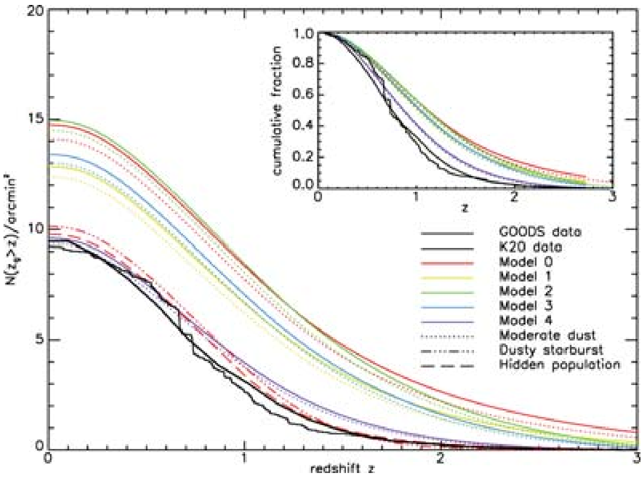

Figure 6. Cumulative redshift distributions corresponding to the differential distributions of Fig. 5. Colours and line styles have the same meaning as in Fig. 5. The inset shows the same distributions normalized to 1.0 at $z = 0$ rather than the absolute counts per arcmin^2.

Table 3. Predicted and observed galaxy counts per arcmin^2 in the redshift range $1 < z < 3$ for a magnitude limit in the K band of $K < 20$.

Model	Counts (arcmin^{-2})				
	$z \geqslant 1$	$z \geqslant 1.5$	$z \geqslant 2$	$z \geqslant 2.5$	$z \geqslant 3$
0	8.36	4.76	2.65	1.48	0.80
1	6.66	3.46	1.70	0.81	0.35
2	8.47	4.62	2.22	0.88	0.21
3	7.05	3.62	1.62	0.58	0.11
4	3.96	1.56	0.50	0.12	0.01
0 (moderate dust)	7.74	4.21	2.21	1.14	0.56
1 (moderate dust)	6.21	3.08	1.41	0.61	0.23
2 (moderate dust)	8.04	4.27	1.97	0.75	0.16
3 (moderate dust)	6.70	3.34	1.43	0.48	0.08
4 (moderate dust)	3.76	1.42	0.43	0.09	0.00
Dusty starburst	3.77	0.90	0.11	0.01	0.00
Hidden population	3.39	0.95	0.28	0.09	0.03
K20	2.63	0.73	0.17	0.06	0.00
GOODS	3.04	0.93	0.29	0.04	0.00

extensions to these models are presented, incorporating dust extinction or complete obscuration of populations.

The first, which we refer to as a 'moderate dust model', is the treatment advocated by Gardner (1998), whose number-count model we adopted. For more details, see his paper and references therein. Secondly, we insert a redshift-dependent dust optical depth in Gardner's slab model for extinction in massive galaxies (assuming that dust and young stars are intermingled), with τ_B scaling as $(z - 1)^n$ for $z \geqslant 1$, mimicking perhaps the case in which this population was born in a very dusty starburst. Thirdly, as a slight variation, we assume that only a redshift-dependent fraction $f \sim z^{-s}$ for $z \geqslant 1$ of the progenitors of present-day ellipticals is visible, the remainder being completely obscured by dust. These two models are loosely based on the results of Totani et al. (2001), whose modified 'PLE' model seems to be able to match observations (see Section 1).

It should be noted here that all dust treatments have a significant effect only at higher redshifts, in particular in the observed K band, which becomes seriously affected by dust extinction only at redshifts beyond $z \sim$ 1–1.5 when it starts to enter rest-frame optical wavelengths.

Figs 5 and 6 show differential and cumulative counts per arcmin^2. In the inset of Fig. 6, we additionally show cumulative plots normalized to unity, demonstrating that the predicted redshift distributions differ in shape as well as in amplitude.

In order to quantify the obvious discrepancy between observations and models, Table 3 presents expected and measured counts integrated over various redshift ranges. The standard Salpeter model, model 0, overpredicts the observed counts beyond $z = 1$ by a factor of almost 3, beyond $z = 1.5$ by more than a factor of 5, and beyond $z = 2$ by nearly an order of magnitude. The assumption of 'moderate dust extinction' barely affects this problem, since even at $z = 3$ the differential counts are only lowered by about 20 per cent for all models.

By construction, the models with substantial high-z dust extinction can bring the counts down to the right numbers. The required slab model extinction in the first case is $\tau_B = 7(z - 1)^{1/2}$ for $z > 1$, resulting in an equivalent foreground screen extinction in the rest-frame B band of 2.11 mag at $z = 2$. In the second model the dependence of the fraction of visible galaxies is $f = z^{-5/2}$ for $z > 1$, which hides about 80 per cent of the massive galaxy population at $z = 2$. It should be emphasized that these are just toy models, which help to indicate the required magnitude of extinction effects.

4 DISCUSSION AND CONCLUSIONS

The problem we study in this paper is whether the available observational data are consistent with the idea that present-day luminous galaxies assembled the bulk of their stars at high redshift. If so, it should be possible to find a set of parameters such that traditional PLE models can simultaneously reproduce: (i) the present-day luminosity and colour distributions of massive galaxies; (ii) the passive evolution in colour and M/L ratio observed for massive early-type galaxies in clusters; and (iii) the observed galaxy counts as a function of redshift in deep surveys. Near-infrared limited surveys are best suited for this purpose since the observed magnitudes are then a fair indicator of stellar mass and are only weakly affected by dust. We therefore chose K-band data from the K20 and GOODS CDF-S surveys for comparison with our models.

Out to redshift $z \sim 1$ our model predictions are very similar to each other and also fit the data reasonably well, given their error bars. At higher redshifts all models predict too many galaxies. Only model 4, with $x = 2$, comes close to the data. Obviously the assumed IMF slope has the largest impact on the predicted number of galaxies at high redshift; the second and third best models are the two with $x = 1.5$. Changing the formation redshift only mildly influences the shape of the distributions at $z < 2.5$. The more conventional standard model 0, using a Salpeter IMF, and its low $z_{\rm f}$ pendant, model 1, produce the predictions most inconsistent with the data. This may be understood by recalling that the light of old stellar populations is dominated by stars with masses near the main-sequence turnoff. For younger populations this turnoff is at higher masses. Hence a shallower IMF implies brighter galaxies at early times, and so more high-redshift galaxies above any apparent magnitude limit. The B-band M/L ratio evolution of the brightest and reddest galaxies is an important constraint on our models because it is also sensitive to the IMF for the same reasons. As already noted in Section 2.1, models with $x = 2$ are inconsistent with observation, except possibly for very low formation redshifts. Finally, since most models for the light output and metal production of high-redshift galaxies require IMFs with substantially *more* high-mass stars than Salpeter (e.g. Nagashima et al. 2005), an IMF as steep as $x = 2$ appears very unlikely as an explanation of the apparent lack of high-redshift massive galaxies. This is an important result, since many observational publications still compare their data to PLE models with rather steep Scalo IMFs, finding reasonably good agreement (e.g. Cimatti et al. 2002a, and references therein).

Our model 0 is very similar to the PLE model used by Somerville et al. (2004) but, whereas we find it to be badly inconsistent with the data, they concluded that any problem is marginal. There are two reasons for this discrepancy. Looking at their fig. 1, there is clearly a problem in going from their differential redshift distribution, which is very similar to our own, to the cumulative distribution, which predicts substantially fewer high-redshift galaxies than does ours. In addition, they compare the cumulative distribution to the data after normalizing both to unity (as in the inset to Fig. 6), which then misses the fact that the total predicted galaxy count at $K < 20$ is substantially larger than observed.

All of our unobscured models with $x \leqslant 1.5$ overpredict the counts at redshifts $z > 1$ by a large factor, as can be seen in Table 3. In the interval $1 < z < 2$ these models all predict more than twice the number of galaxies observed, and in the interval $2 < z < 3$ they are off by factors between 4 and 11. Could cosmic variance or dust account for this? The clustering of galaxies has the greatest effect at low redshift, where the observed volume is comparatively small and clear evidence of large fluctuations is seen in Fig. 5 at $z = 0.7$ in the K20 data. However, in this range the models still agree quite well with the data. Only at higher redshifts do they deviate. Also the model predictions are obviously systematically too high at all z, which is not what one would expect if the effect was due to cosmic variance. Finally, models and data also disagree in the normalized version of the diagram (inset in Fig. 6).

Extinction by dust, on the other hand, might indeed be important. From looking at Figs 5 and 6 as well as Table 3, it becomes clear that the simple dust treatment conventionally applied to PLE models (e.g. Gardner 1998) is not sufficient. A more extreme assumption about the amount of extinction at high redshift like that of Totani et al. (2001) is needed.

To assess how much dust is required to bring our PLE models into agreement with the data, consider placing a foreground screen in front of all galaxies at $z = 1.5$, thereby translating their apparent luminosity function fainter by some fixed amount. We find that to lower the count for model 0 in Fig. 5 by the factor of 2.1 needed to bring it into agreement with the GOODS data at this redshift requires 0.7 mag of extinction at observed K (i.e. at rest-frame z). Carrying out a similar calculation at $z = 2$, we find that 1.0 mag of extinction is required at observed K (now rest-frame r) to reduce the abundance by the required factor of 5.1. These numbers are consistent with the slightly more detailed models of Figs 5 and 6. For comparison, Kauffmann et al. (2003) analysed dust attenuation in a sample of 122 808 low-redshift galaxies drawn from the SDSS, finding a typical (median) attenuation of 0.2–0.3 mag in the z band for massive galaxies. We thus need more dust in high-redshift massive galaxies than is seen in local galaxies to reconcile our PLE models with the data.

Note that these toy dust models substantially *under*-predict the amount of dust needed to get agreement, since they assume the stellar populations of the dusty galaxies to be just as old as in the unobscured models, which are no longer rapidly forming stars in type 1 galaxies at $z < 2$. In the nearby Universe, younger stellar populations are almost always present in dusty galaxies, and the enhanced luminosity due to the young stars cancels almost exactly the attenuation effects of the dust, resulting in a mean apparent M/L ratio for red galaxies that depends weakly on dust content at red optical wavelength (Bell & de Jong 2001; Kauffmann et al. 2003). If high-redshift galaxies behave similarly, then dust will *not* help to reconcile our PLE models with the data.

Recent data do, in fact, indicate such behaviour, as illustrated in Fig. 7. Here we compare the evolution of rest-frame M/L_V predicted by our model 0 and by its 'dusty starburst' variant with observational estimates for two classes of high-redshift galaxy. Fontana et al. (2004) provide estimates for ∼140 K-selected early-type galaxies from the K20 survey. This sample contains both unobscured 'passive' systems and dusty, star-forming galaxies. We plot mean values with their uncertainties for bins centred at $z = 0.85$, 1.25 and 1.75. Smail et al. (2004) provide estimates from optical follow-up of a sample of 96 submillimetre-selected galaxies. These are all highly obscured, strongly star-forming galaxies. We plot mean M/L_V estimates with their quoted uncertainties for the two star formation histories considered by Smail et al., namely a single short burst (the lower value) and a constant SFR (the upper value). The theoretical curves correspond to colour types 1 and 2, which contribute 48 and 47 per cent, respectively, of all galaxies at $z > 1.5$ and $K < 20$ in our model 0.

Fig. 7 demonstrates that the M/L_V values assumed in our standard PLE model with no or moderate obscuration are very similar to those measured in real high-redshift galaxies, even though the observed systems are both dusty and star-forming. On the other hand, when we fit our ad hoc 'dusty starburst' model to the observed redshift counts, it predicts M/L_V values well above those estimated for high-redshift galaxies, even those selected specifically for the strength of their dust emission. Thus PLE models with moderate obscuration match the observed mass-to-light ratios at high redshift but overpredict abundances, while models with sufficient obscuration to fit the observed abundances substantially overpredict high-redshift mass-to-light ratios.

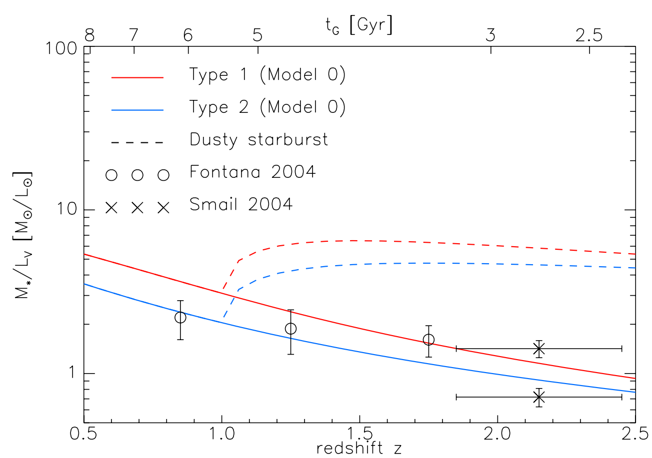

Figure 7. The M_*/L_V ratios of early-type galaxies as a function of redshift taken from the literature compared to colour types 1 and 2 from our model 0. The full lines correspond to the unobscured model, while the dashed lines show the 'dusty starburst' model (see Figs 5 and 6). The data points from Smail et al. (2004) are the mean for a sample of submillimetre galaxies at $\langle z \rangle = 2.2$ for two different assumed SFHs, a single burst (lower value) and a constant SFR (upper value) model. The data from the K20 sample of Fontana et al. (2004) are combined into three redshift bins.

Our main conclusion is thus that 'traditional' PLE models, as originally introduced by Tinsley (1980), cannot reconcile the relatively small number of high-redshift galaxies found in deep K-selected redshift surveys with the abundance of massive galaxies seen in the local Universe. The counterparts of nearby luminous red galaxies just do not seem to be present in sufficient numbers at redshifts of 1.5–2. The areas of current deep surveys are quite small, so there may still be significant uncertainties as a result of cosmic variance. Substantial amounts of dust may also cause many distant massive galaxies to be missed, but only if dust attenuation is not compensated by emission from young stars in the way observed in low-redshift galaxies and if the M/L values estimated for current samples of high-redshift galaxies are atypically small. Observation of the relevant galaxy populations over larger areas and at longer wavelengths will help to get a better understanding of this question.

ACKNOWLEDGMENTS

We would like to thank the anonymous referee for comments that helped to improve this paper. MGK acknowledges a PhD fellowship from the International Max Planck Research School in Astrophysics, and support from a Marie Curie Host Fellowship for Early Stage Research Training.

This work makes use of SDSS Archive data from http://wassup.physics.nyu.edu/manyd/paper/

Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS web site is http:www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, the Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Princeton University, the United States Naval Observatory and the University of Washington.

REFERENCES

Andreon S., 2002, A&A, 382, 495
Bell E. F., de Jong R. S., 2001, ApJ, 550, 212
Bell E. F. et al., 2004, ApJ, 608, 752
Blanton M. R. et al., 2003, ApJ, 594, 186
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Chen H. et al., 2003, ApJ, 586, 745
Cimatti A. et al., 2002a, A&A, 391, L1
Cimatti A. et al., 2002b, A&A, 392, 395
Cimatti A. et al., 2004, Nat, 430, 184
Firth A. E. et al., 2002, MNRAS, 332, 617
Fontana A., Menci N., D'Odorico S., Giallongo E., Poli F., Cristiani S., Moorwood A., Saracco P., 1999, MNRAS, 310, L27
Fontana A. et al., 2004, A&A, 424, 23
Gardner J. P., 1998, PASP, 110, 291
Gavazzi G., Boselli A., Kennicutt R., 1991, AJ, 101, 1207
Glazebrook K. et al., 2004, Nat, 430, 181
Im M. et al., 2002, ApJ, 571, 136
Kashikawa N. et al., 2003, AJ, 125, 53
Kauffmann G., Charlot S., 1998, MNRAS, 297, L23 (KC98)
Kauffmann G. et al., 2003, MNRAS, 341, 33
Kochanek C. S. et al., 2001, ApJ, 560, 566
Mobasher B. et al., 2004, ApJ, 600, L167
Nagashima M., Lacey C. G., Baugh C. M., Frenk C. S., Cole S., 2005, MNRAS, 358, 1247
Pozzetti L. et al., 2003, A&A, 402, 837
Rudnick G. et al., 2001, AJ, 122, 2205
Saracco P. et al., 2005, MNRAS, 357, L40
Smail I., Chapman S. C., Blain A. W., Ivison R. J., 2004, ApJ, 616, 71
Somerville R. S. et al., 2004, ApJ, 600, L135
Tinsley B. M., 1980, ApJ, 241, 41
Totani T., Yoshii Y., 2000, ApJ, 540, 81
Totani T., Yoshii Y., Maihara T., Iwamuro F., Motohara K., 2001, ApJ, 559, 592
van Dokkum P. G., Stanford S. A., 2003, ApJ, 585, 78
van Dokkum P. G., Franx M., Fabricant D., Kelson D. D., Illingworth G. D., 1999, ApJ, 520, L95
York D. G. et al., 2000, AJ, 120, 1579

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