Deep infrared studies of massive high redshift galaxies

Labbé, I.

CHAPTER

FIVE

Field Galaxies at $1 < z < 3$ The Color Magnitude Distribution of the evolution and modeling of the blue sequence

ABSTRACT

We use deep near-infrared VLT/ISAAC imaging to study the rest-frame color-magnitude distribution of infrared selected galaxies in the redshift range $1 < z < 3$. We find a well-defined blue peak of star-forming galaxies at all redshifts. The blue galaxies populate a color-magnitude relation (CMR), such that more luminous galaxies in the rest-frame V-band tend to have redder ultraviolet-to-optical colors. The slope of the CMR does not evolve with time, and is similar to the slope of blue, late-type galaxies in the local universe. Analysis of spectra of nearby late-type galaxies suggests that the steepness of the slope can be fully explained by the observed correlation of dust content with optical luminosity. The zeropoint of the blue CMR at a given magnitude reddens smoothly from $z = 3$ to $z = 0$, likely reflecting an increase of the mean stellar age and an increase in the mean dust opacity of blue-sequence galaxies. A distinct feature is that the color scatter around the $z \sim 3$ CMR is asymmetric, with a blue "ridge" and a skew towards red colors. We have explored which types of star formation histories can reproduce the scatter and the skewed shape of the color distribution. These included models with constant star formation rates and sudden cutoffs, exponentially declining star formation rates, burst models, and models with episodic star formation. The episodic models reproduced the color distribution best, with quiescent periods lasting 30-50% of the length of an active period, and duration of the duty cycle between 150 Myr to 1 Gyr. The episodic star formation in these models rejuvenate the galaxies during each episode, making it significantly bluer than a galaxy with constant star formation of the same age. This could be a solution of the enigmatic observation that $z = 3$ galaxies are much bluer than expected if they were as old as the universe. Finally, the color distribution has a strong tail of very red galaxies. The relative number of red galaxies increases sharply from $z \sim 3$ to $z \sim 1$. The rest-frame V-band luminosity density in luminous blue-sequence galaxies is constant, or decreases, whereas that in red galaxies rises with time. We are viewing the progressive formation of red, passively evolving galaxies.

Ivo Labbé, Marijn Franx, Gregory Rudnick, Natascha M. Förster Schreiber, Emanuele Daddi, Pieter G. van Dokkum, Konrad Kuijken, Alan Moorwood, Hans-Walter Rix, Huub Röttgering, Ignacio Trujillo, Arjen van der Wel, Paul van der Werf, & Lottie van Starkenburg

1 Introduction

The formation and evolution of galaxies is a complex process, involving the growth of dark-matter structure from the gradual hierarchical merging of smaller fragments (e.g., White & Frenk 1991; Kauffmann & White 1993), the accretion of gas, the formation of stars, and feedback from supernovae and black holes. Presently, the theories describing the growth of large-scale dark-matter structure are thought to be well-constrained (Freedman et al. 2001; Efstathiou et al. 2002; Spergel et al. 2003), but the formation history of the stars inside the dark-matter halos is still poorly understood. Neither hydrodynamical simulations (e.g., Katz & Gunn 1991; Springel et.al. 2001; Steinmetz & Navarro 2002) nor state-of-the-art semianalytic models (e.g., Kauffmann, White, & Guiderdoni 1993; Somerville & Primack 1999; Cole, Lacey, Baugh, & Frenk 2000) provide unique predictions for the formation of the stars in galaxies, given the large parameter space available to these models. Direct observations are critical to constrain them.

Specifically the observations of massive galaxies provide strong tests for pictures of galaxy formation, as their build-up can be directly observed from high redshift to the present epoch. The largest samples of high-redshift galaxies to date have been selected by their rest-frame UV light, through the Lyman Break technique (LBGs; Steidel et al. 1996a,b, 2003). Unfortunately, the rest-frame UV light is highly susceptible to dust and not a good measure of the number of intermediate and low mass stars, which may dominate the stellar mass. In fact, rest-frame optical observations have shown LBGs to be relatively low mass ($M \sim 10^{10}\mathrm{M}_{\odot}$), unobscured, star-forming galaxies (e.g., Papovich, Dickinson, & Ferguson 2001; Shapley et al. 2001).

The rest-frame optical light is already much less sensitive to the effects of dust obscuration and on-going star formation than the UV, and it is expected to be a better tracer of stellar mass. Recent advances in near-infrared (NIR) capabilities on large telescopes are now making it possible to select statistically meaningful samples of galaxies by their rest-frame optical light out to $z \sim 3$. In this context we started the Faint Infrared Extragalactic Survey (FIRES; Franx et al. 2000), a deep optical-to-infrared multicolor survey of NIR-selected galaxies. The rest-frame optical observations are also useful as much of our knowledge in the local universe is based on studies at these wavelengths.

A particular diagnostic of galaxy formation is the shape of the galaxy luminosity and mass functions as a function of the spectral type. In the local universe these distributions have been determined in great detail (e.g., Strateva et al. 2001; Norberg et al. 2002; Blanton et al. 2003; Kauffmann et al 2003). While a precise characterisation of the local distributions can strongly constrain the models, powerful tests can also be made using their evolution as a function of redshift. For example, Bell et al. (2004a,2004b) show that the V-band luminosity density of photometrically selected early-type galaxies is nearly constant in the range $0.2 < z < 1.1$, suggesting an increase in stellar mass of passively evolving galaxies

over this range. This is not expected from the simplest collapse models, where all galaxies form at high redshift and evolve passively to the present-day.

Different constraints on galaxy formation come from correlations between integrated galaxy properties. For example, morphologically early-type galaxies in the local universe and distant clusters populate a well-defined color-magnitude relation (CMR), which is thought to reflect a sequence of increasing metallicity with stellar mass (e.g., Bower, Lucey, & Ellis 1992; Schweizer & Seitzer 1992; Kodama & Arimoto 1997). The small scatter of this relation implies that early type galaxies formed most of their stars at high redshift (e.g., van Dokkum et al. 1998). Blue, star forming, late-type galaxies also display a color-magnitude relation (Chester & Roberts 1964; Visvanathan 1981; Tully, Mould, & Aaronson 1982), but the scatter is larger (Griersmith 1980), and its origin is probably more complex; it has been interpreted as a sequence in the mean stellar age (Peletier & de Grijs 1998), dust attenuation (Tully et al. 1998), and/or metallicity (Zaritsky, Kennicutt, & Huchra 1994, Bell & De Jong 2000).

Interestingly, Papovich, Dickinson, & Ferguson (2001) found a color-magnitude relation for blue, star-forming galaxies at $z \sim 3$, from a sample of NIR-selected galaxies in the Hubble Deep Field North (HDFN). The trend, which is also seen in the similar field of the Hubble Deep Field South (Labbé et al. 2003), is such that galaxies more luminous in the rest-frame $V-$band, tend to have redder ultraviolet-to-optical colors. A distinct aspect of the high-redshift CMR, is that the galaxies are asymmetrically distributed around the relation, with a well-defined blue envelope (Papovich et al. 2004).

Studies of colors and magnitudes of star-forming galaxies at $z > 2$ also raised questions. A particular puzzle was presented by modeling of their stellar populations, which implied average luminosity weighted ages of a few 100 Myr, much younger than the age of the universe at these redshifts (e.g., Papovich, Dickinson, & Ferguson 2001; Shapley et al. 2001). From this, and from the relative absence of candidates for red, non star-forming galaxies in the HDFN it was suggested that star formation in LBGs occurs with short duty cycles and a timescale between star formation events of $\lesssim 1$ Gyr.

In this paper, we investigate the evolution of the rest-frame colors of galaxies as a function of redshift in the range $1 \lesssim z \lesssim 3$. The deep optical-to-NIR imaging and the homogeneous photometry of the FIRES project is excellently suited for such studies, and we use it to analyze rest-frame ultraviolet-to-optical colors and magnitudes of a sample of 1475 K_s-band selected galaxies.

We focus on the galaxies that populate the blue peak of galaxies at low and high redshift. We wish to understand the nature of the blue color-magnitude relation, the evolution of the colors towards high redshift, and the origin of the conspicious skewed color distribution around the CMR at $z \sim 3$. For the latter we explore models with different types of star formation histories that might produce such distributions. Finally, we chart the evolution of the relative number of red

galaxies over the redshift range $1 \lesssim z \lesssim 3$.

This paper is organized as follows. We present the data in §2, describe the color magnitude distribution of FIRES galaxies in §3, analyze the blue color-magnitude relation in §4, and model the scatter of galaxies around the blue CMR in §5. Finally, §6 presents the evolution of the red galaxy fraction. Where necessary, we adopt an $\Omega_M = 0.3, \Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1}Mpc^{-1} cosmology. We use magnitudes calibrated to models for Vega throughout.

2 The Data

2.1 The Observations and Sample Selection

The observations were obtained as part of the public Faint Infrared Extragalactic Survey (FIRES; Franx et al. 2000) the deepest groundbased NIR survey to date. We cover two fields with existing deep optical WFPC2 imaging from the *Hubble Space Telescope* (HST): the WPFC2-field of HDFS, and the field around the $z = 0.83$ cluster MS1054-03. The observations, data reduction, and assembly of the catalog source catalogs are presented in detail by Labbé et al. (2003) for the HDFS and Förster Schreiber et al. (2004a) for the MS1054 field.

Briefly, we observed in the NIR J_s, H, and K_s bands with the Infrared Spectrometer and Array Camera (ISAAC; Moorwood 1997) at the *Very Large Telescope* (VLT). In the HDFS, a total of 101.5 hours was invested in a single $2.5' \times 2.5'$ pointing, resulting in the deepest groundbased NIR imaging, and the deepest $K-$band to date, even from space. We complemented the existing deep optical HST WFPC2 imaging in the $U_{300}, B_{450}, V_{606}, I_{814}$ bands (Casertano et al. 2000). A further 77 hours of NIR imaging was spent on a mosaic of four ISAAC pointings centered on the $z = 0.83$ foreground cluster MS1054-03, reaching somewhat shallower depths. We complemented the data with WFPC2 mosaics in the V_{606} and I_{814} bands (van Dokkum et al. 2000), and collected additional imaging with the VLT FORS1 instrument in the U, B, and V bands (Förster Schreiber et al. 2004a). In both surveyed fields the effective seeing in the final NIR images was $\approx 0\farcs45 - 0\farcs55$ FWHM.

We detect objects in the K_s-band using version 2.2.2 of the SExtractor software ((Bertin & Arnouts 1996). For consistent photometry accross all bands, all images were aligned, and accurately PSF-matched to the filter in which the image quality was worst. Stellar curve of growth analysis indicates that the fraction of enclosed flux agrees to better than 3% for the apertures relevant to our color measurements. The color measurements were done in a customized isophotal aperture defined from the K_s-image. The estimate of total flux in the K_s-band was computed using SExtractors *AUTO* aperture for isolated sources, and in an adaptive circular aperture for blended sources. In both cases, a minimal aperture correction for the light lost by a point source was applied. Photometric uncertainties were derived empirically from the flux distribution in apertures placed on empty parts of the

maps. For details concerning all aspects of the photometric measurements, see Labbé et al. 2003. The total $5-\sigma$ limiting depth for point sources are $K_s^{tots} = 23.8$ for the HDFS, and 23.1 for the MS1054-field.

2.2 Photometric Redshifts and Rest-Frame Colors

Photometric redshifts were estimated by fitting a linear combination of redshifted empirical galaxy spectra, and a 10 Myr old simple stellar population model (1999 version of Bruzual A. & Charlot 1993) to the observed flux points. The algorithm is described in detail by Rudnick et al. (2001, 2003). We adopt a minimux flux error of 5% for all bands to account for zeropoint uncertainties and for mismatches between the observations and the photo-z template set.

Monte-Carlo simulations were used to estimate the errors $\delta z_{ph,MC}$ on the photometric redshifts. These errors reflect the photometric uncertainties, template mismatch, and the possibility of secondary solutions. We determined the accuray of the technique from comparisons to the available spectroscopy in each field. We find $\delta z =< |z_{spec} - z_{phot}|/(1 + z_{spec}) >= 0.07$, and $\delta z = 0.05$ for sources at $z \geq 2$. The errors calculated from simulations $\delta z_{ph,MC}$ are consistent with this. We identified and removed stars using the method described in Rudnick et al. (2003).

We combine the observed SEDs and photometric redshifts to derive rest-frame luminosties L_λ^{rest}. We used a method of estimating L_λ^{rest} that interpolates directly between the observed fluxes, using the templates as a guide. The rest-frame phometric system, and details on estimating L_λ^{rest} are described extensively by (Rudnick et al. 2003). Throughout we will use the rest-frame UX, B, and V filters Beers et al. (1990) and the HST/FOC $F140W, F170W$, and $F220W$ filters, which we will call $1400, 1700$, and 2200 throughout. We adopted the photometric system of Bessell (1990) for the optical filters, which was calibrated to the Dreiling and Bell (1980) model spectrum for Vega. The HST/FOC UV zeropoints were calibrated to the Kurucz (1992) model for Vega.

The rest-frame luminosities are sensitive to the uncertainties in the photometric redshifts. Therefore we only analyze the sample of galaxies with $\delta z_{ph,MC}/(1 + z_{ph}) < 0.2$, keeping 1354 out of 1475 galaxies. The median $\delta z_{ph,MC}/(1 + z_{ph})$ for the remaining sample is 0.05. We checked that the color distribution of the rejected galaxies was consistent with that of the galaxies we kept.

The reduced images, photometric catalog, redshifts, and rest-frame luminosities are all available on-line through the FIRES homepage[1].

[1] http://www.strw.leidenuniv.nl/~fires

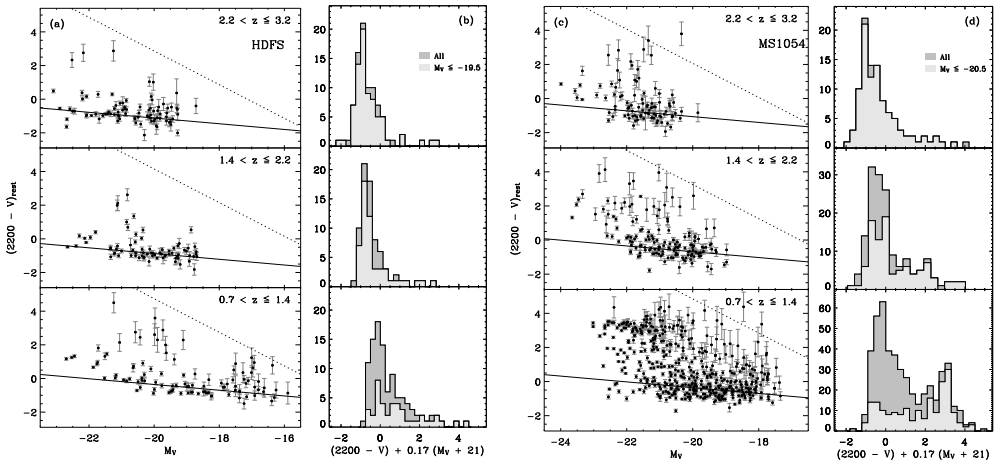

Figure 1 — The rest-frame $2200 - V$ colors against absolute magnitude in the $V-$band for galaxies in the field of the HDFS (a) and in the field of MS1054 (c). The two samples are split into three redshift bins. The errorbars represent the 1σ uncertainties on the rest-frame colors. The dotted line roughly marks a conservative color limit, corresponding to the 2σ flux uncertainty in the observed filter that is closest to 2200-band at the maximum redshift of the bin. Galaxies redward of this line have uncertain rest-frame colors, but can be observed. The solid line shows a fit of a linear relation with a fixed slope of -0.17 to the galaxies in the blue peak of the color-magnitude distribution. Panels b (HDFS) and d (MS1054) show histograms of the colors after the slope is subtracted. The peak of this distribution is normalized to the intercept of the CMR at $M_V = -21$. We show the color distributions of all detected galaxies in each bin (*dark gray histograms*) and those to a limiting magnitude of $M_V \leq -19.5$ and $M_V \leq -20.5$ in the field of the HDFS and MS1054 respectively (*light gray histograms*).

3 The rest-frame Color-Magnitude Distribution of Galaxies from $z \sim 1$ to $z \sim 3$

A well-studied diagnostic in the local universe is the optical color-magnitude distribution (e.g., Strateva et al. 2001; Hogg et al. 2002; Baldry et al. 2004). Quantifying this distribution out to higher redshifts, will provide strong constraints on models of galaxy formation, which must reproduce these observations.

To quantify the evolution of galaxy colors and magnitude with redshift, we compare the distribution at the rest-frame $2200-V$ color versus absolute $V-$band magnitude. We present evolution in terms of the $2200 - V$ color in panels a and c of Figure 1. These colors are sensitive to variations in stellar population age, as they span the Balmer and 4000 Å breaks, and attenuation by dust, by virtue of the long wavelength baseline extending into the UV. We prefer the rest-frame 2200-band to the more commonly used U-band, because at $z = 2 - 3$ the 2200-band is probed by our deepest optical observations. In addition, the uncertainties in our photometric redshifts tend to increase the scatter more in the U and B colors than in the 2200 colors as the former bands directly straddle the Balmer

and 4000 Å breaks.

We divide the galaxies in three redshift bins centered at z=1.0, 1.8, and z=2.7. The redshift range is defined so that the rest-frame 2200 and V-band always lies within the range of our observed filters. The number of redshift bins is a compromise between keeping a statistically meaningful sample at each redshift, while reducing the effects of evolution over the time span in each bin. The limits of the bins are defined so that galaxies in the smaller field of the HDFS are equally divided in number. The redshift bins of the larger MS1054 field are matched to those of the HDFS.

Clearly, in both fields and in all redshift bins there is a well-defined ridge of blue galaxies. Furthermore, the ridge is tilted: the more luminous galaxies along the ridge tend to have redder $2200 - V$ colors. This ridge is what we define as the color magnitude relation of blue galaxies.

There is no reason to think that this color-magnitude relation is due to some selection effect; we can easily detect galaxies that are bright in M_V and blue in $2200 - V$, and our photometry is sufficiently deep to ensure that the trend is not caused by a lack of faint red galaxies. Only at at low redshifts ($z \sim 1$) and faint magnitudes may we miss some of the bluest galaxies, as the sample was selected the observed K_s-filter, which is significantly redder than the redshifted $V-$band at $z \sim 1$.

It is striking that the blue color magnitude relation has a very well- defined boundary to the blue, and a much more extended tail to the red. This red tail extends up to 4 magnitudes, and defines a clear red color magnitude relation in the lowest redshift bin in the MS1054 field. This is mainly caused by the color-magnitude relation of the cluster galaxies at $z = 0.83$, and is also observed in the field up to $z = 1$ (e.g., Bell et al. 2004a)

We now proceed to analyze the properties of the blue color magnitude in sections 4 and 5, and will turn to the red galaxies in section 6.

4 The Color-Magnitude Relation of Blue Field Galaxies

Two key features of the blue CMR in Fig. 1 are that the peak of the distribution, and hence the zeropoint of the blue sequence, reddens from $z \sim 3$ to $z \sim 1$, while on first sight the slope does not seem to vary appreciably. We proceed by quantifying the slope and zeropoint of the relation and their evolution with redshift.

4.1 The Slope and its Evolution

We assume that the distribution of colors and magnitudes can be characterized by a linear relation. We obtain the best fitting slope as follows: for each value of the slope we obtain a distribution of color residuals. We select the slope that maximizes the peak of the histogram of the residuals. This so-called “mode re-

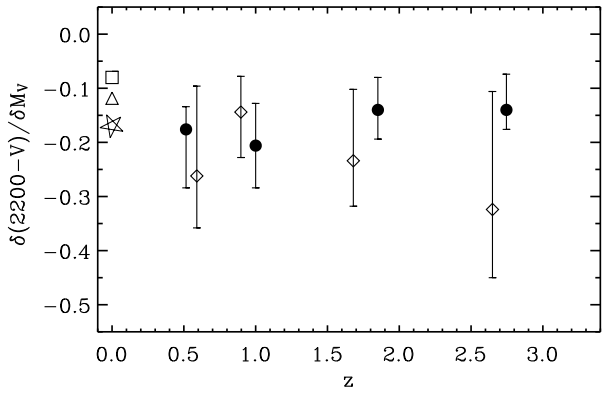

Figure 2 — The evolution of the slope of the blue color-magnitude relation from linear fits to the galaxy distributions in the HDFS (*filled circles*) and MS1054 (*diamonds*). The results are plotted at the median redshift of the galaxies in the bin. The uncertainties correspond the 68% confidence interval obtained with bootstrap resampling. The data are consistent with a constant slope with redshift. Also drawn are values for blue late type galaxies from the Nearby Field Galaxy Survey (Jansen, Franx, & Fabricant 2000a), where we assume the CMR is caused by a systematic trend with dust reddening (*star*), or stellar age (*triangle*). Finally, we show the value for the metallicity-luminosity relation of local early type galaxies (*square*; Bower, Lucey, & Ellis 1992).

gression" technique is very insensitive to outliers. The histogram is not calculated in discrete bins, but using a kernel density esimator with a small gaussian kernel of $\sigma = 0.2$, comparable to the photometric uncertainty in the colors of individual galaxies. The results depend slightly on the smoothing parameter used but these effects are small compared to the intrinsic uncertainty caused by the limited number of galaxies in the sample.

These uncertainties were calculated by bootstrap resampling the color-magnitude distributions 200 times, and repeating the fitting procedures. We selected the central 68% of the best-fit slopes as our confidence interval.

Figure 2 shows the values of the slope $\delta(2200 - V)/\delta M_V$ versus redshift in fields of the HDFS and MS1054. We have included an additional low-redshift bin centered at $z \sim 0.5$, also determined from our data. To compare to observations at $z = 0$, we derived the $2200 - V$ slope from the $U - V$ slope under three different assumptions:

In the first case we assume the slope to result from a metallicity-luminosity sequence identical to that of nearby early-type galaxies, where more luminous galaxies are more metal-rich. We used observations of Coma cluster galaxies (Bower, Lucey, & Ellis 1992) and converted the observed slope in $U - V$ colors to $2200 - V$ colors with the help of Bruzual & Charlot (2003) models. These models were constructed to match observed colors of the cluster galaxies on the red CM-relation,

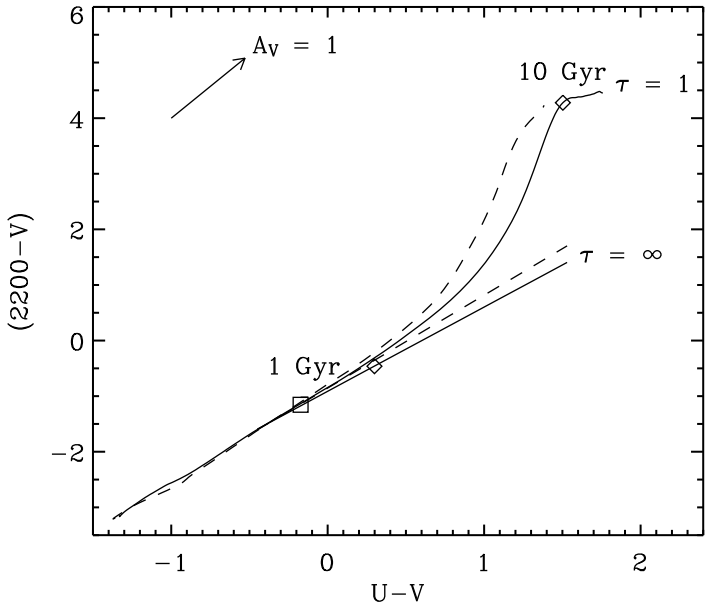

Figure 3 — The relation between $2200 - V$ and $U - V$ colors for Bruzual & Charlot(2003) stellar population models with declining star formation rates and a range of timescales τ. Solar metallicity (*solid lines*) and 1/3 solar models (*dashed lines*) are shown. The square indicates stellar ages of ∼1 Gyr, the diamonds indicate ages of ∼10 Gyr. Over a range of stellar population ages and metallicities the relationship is tight for blue galaxies, allowing a fairly accurate transformation of their colors. Also drawn is a Calzetti et al. (2000) attenuation vector of $\mathrm{A}_V = 1$. In the presence of dust, the transformation of $U - V$ to $2200 - V$ colors is done with stellar tracks that already include reddening.

assuming the expected star formation history, i.e., high formation redshift and passive evolution.

The other two are derived from a linear fit to the $U - V$ colors of nearby blue-sequence galaxies from the Nearby Field Galaxy Survey (NGFS; Jansen et al. 2000a). Here we synthesized the $2200 - V$ slope under two assumptions. First, we assumed the CMR to reflect a systematic age trend, where higher-luminosity galaxies have higher stellar ages, and we used Bruzual & Charlot (2003) models to transform $U - V$ colors to $2200 - V$ colors (see Figure 3) For galaxies with blue colors, the models give $\Delta(2200-V) \approx 1.5\Delta(U-V)$ for a range of metallicities and stellar ages. Second, we assume the slope is the result of increasing dust opacity with luminosity. Adopting the Calzetti et al. (2000) dust law yields $\Delta(2200-V) \approx 2.3\Delta(U-V)$. An SMC extinction law (Gordon et al. 2003) would yield similar values.

Fig. 2 shows that the slope of the CM relation does not depend on redshift.

Only if we use the $z = 0$ slope derived for very red early-type galaxies in Coma is there any hint of evolution. This model is rather extreme, however, and will not be considered further. The measured evolution of the slope is $0.01z \pm 0.03$, consistent with zero. The error-weigthed mean value of all FIRES measurements is

$$\delta(2200 - V)/\delta M_V = -0.17 \pm 0.021$$

If the slope were due to an age gradient as a function of magnitude, one might expect the slope to steepen with redshift, but we see no significant effect in the data presented here. We analyze the cause of the relation later in §4.3.

We note that Barmby et al. (2004) found a linear CMR for faint galaxies, and an upturn of the CMR at bright magnitudes. Figure 1 shows some evidence for an upturn at the bright end, but obviously our sample is too small to describe this properly, and we hence focus entirely on the linear part of the relation.

4.2 The Zeropoint and its Evolution

We determine the zeropoint by assuming that the slope of the CMR does not evolve and remains at the mean value of $\delta(2200 - V)/\delta M_V = -0.17$. We subtract the slope so that all galaxy colors are normalized to the color at $M_V = -21$. The histogram of normalized colors was determined as described before, and the zeropoint was determined from the location of the peak.

The evolution of the CMR zeropoint at $M_V = -21$ with redshift is shown in figure 4. The errorbars reflect the central 68% confidence interval of the zeropoint, obtained with the bootstrapping technique. We also present the CMR zeropoint of a local galaxy sample from the Nearby Field Galaxy Survey (Jansen, Fabricant, Franx, & Caldwell 2000). This value was derived from a fit to the $U - V$ colors of blue-sequence galaxies. We used tracks in Figure 3 to transform from $U - V$ colors to $2200 - V$, applying a small correction for a dust reddening of $E(B - V) = 0.15$. The color excess was obtained from the spectra of the galaxies (see §4.3), and adopting a Calzetti et al. (2000) dust law.

Clearly, the color of the blue CMR at fixed absolute magnitude reddens monotonically from $z \sim 3$ to $z \sim 0.5$. The galaxies become redder in $2200 - V$ by ≈ 1.2 mag from $z \sim 2.7$ to $z \sim 0.5$ (from the FIRES data alone), or by ≈ 1.45 from $z \sim 2.7$ to $z = 0$ Surprisingly a straight line describes the points very well, and we find

$$2200 - V = 0.24 \pm 0.06 - (0.46 \pm 0.03)z$$

The colors in the independent fields of the HDFS and MS1054 agree very well, and if extrapolated to $z = 0$, are consistent with the $z = 0$ point from the NFGS.

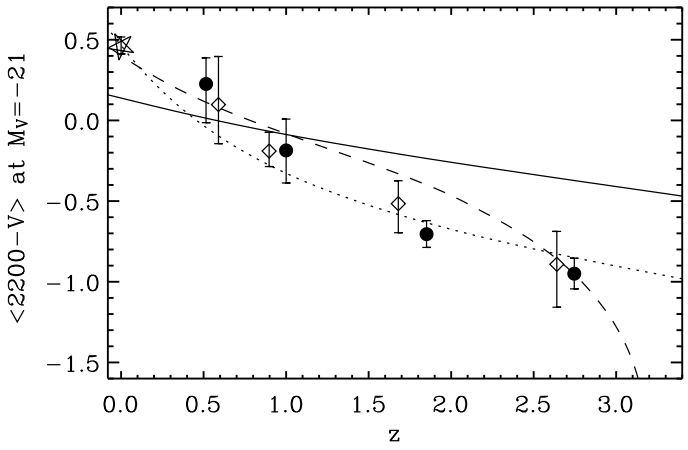

Figure 4 — The intercepts of fits to the blue CMR at fixed $M_V = -21$, marking the color evolution of the blue sequence as a function of redshift. We show the results in the field of the HDFS field (*filled circles*) and in the field of MS1054 (*diamonds*). The errorbars correspond to the 68% confidence interval derived from bootstrap resampling. The star indicates the $z = 0$ relation from the NFGS (Jansen, Franx, & Fabricant 2000a). The lines represent tracks of Bruzual & Charlot (2003) stellar population models. We show a model with formation redshift $z_f = 3.2$, a star formation timescale $\tau = 10$ Gyr, and fixed reddening of $E(B-V) = 0.15$ (*dashed line*); one with $z_f = 10$, constant star formation, and fixed $E(B-V) = 0.15$ (*solid line*); a model with $z_f = 10$, $\tau = 10$ Gyr, and $E(B-V)$ evolving linearly in time from 0 at $z = 10$ to 0.15 at $z = 0$ (*dotted line*).

This is encouraging considering that absolute calibration between such different surveys is difficult, and that the NFGS data has been transformed to $2200 - V$ colors from other passbands.

Next we compare the observed color evolution to predictions from simple stellar populations models. We assume that the galaxies remain on the ridge of the CMR throughout their life. This assumption may very well be wrong, but more extensive modeling is beyond the scope of this paper. Obviously for such simple models, the galaxies have all the same color, and these follow directly from the colors of the stellar population model, depending on star formation history, and the dust absorption only.

We calculate rest-frame colors using Bruzual & Charlot (2003) stellar population synthesis models with solar metallicity and exponentially declining star formation rates. We added dust reddening to the model colors using a Calzetti et al. (2000) dust law. To generate colors at constant absolute magnitude we apply a small color correction to account for $V-$band luminosity evolution of the stellar population. It reflects the fact that galaxies that were brighter in the past populated a different, redder part of the CMR. Similarly, when we allow varying levels

of dust attenuation, we apply a correction to account for dimming of the V-band light. We use the measured slope to apply the correction, the total amplitude of the effect is less than $\Delta(2200 - V) \lesssim 0.1$ for most models.

The simple model tracks are shown in Figure 4. A model with constant star formation and formation redshift $z_f = 10$ fits remarkably bad: the evolution is much slower than the observed evolution. When we explore models with exponentially declining star formation rates, we find that a decline time scale $\tau = 10Gyr$ and formation redshift $z_f = 3.2$ fits best. We note that we restricted the fit to $zf \geq 3.2$ as the color-magnitude relation is already in place at redshifts lower than that.

Naturally, the models can be made to fit perfectly by allowing variable reddening by dust. As an example we show a $\tau = 10$ Gyr model with formation redshift $z_f = 10$ with reddening evolving linearly with time from $E(B-V) = 0$ at $z_f = 10$ to $E(B-V) = 0.15$ at $z = 0$.

We conclude that the relatively strong color evolution in the interval $0 < z < 3$ is likely caused by both aging of the stellar population and increasing levels of dust attenuation with time.

4.3 The Origin of the Blue Sequence in the Local Universe

It is impossible to determine the cause of the color-magnitude relation of blue galaxies from broad band photometry alone. Fits of models to the photometry produce age and dust estimates which are very uncertain (e.g., Shapley et al. 2001, Papovich et al. 2001, Förster-Schreiber et al. 2004). Spectroscopy is needed for more direct estimates of the reddening and ages. Unfortunately, the restframe optical spectroscopy of distant galaxies is generally not deep enough for the detection of Hβ, and the balmer decrement cannot be determined to high enough accuracy. Hence we can only analyze the relation for nearby galaxies which have spectroscopy with high signal-to-noise ratio.

We use the local sample of galaxies in the Nearby Field Galaxy Survey (Jansen, Fabricant, Franx, & Caldwell 2000; Jansen, Franx, Fabricant, & Caldwell 2000), a spectrophotometric survey of a subsample of 196 galaxies in the CfA redshift survey (Huchra et al 1993). It was carefully selected from $\sim$ 2400 galaxies to closely match the distributions of morphology and magnitude of the nearby galaxy population, and covers a large range of magnitudes $-15 < M_B < -23$. The strength of the NFGS is the simultaneous availability of integrated broadband photometry, and integrated spectrophotometry of all galaxies, including line fluxes and equivalent widths of $[OII]$ H$_\alpha$, and H$_\beta$. Integrated spectra and photometry are essential to enable a fair comparison to the integrated photometry of our high-redsfhift galaxy sample.

The trends of $U - V$ colors, dust absorption, metallicity, and Hα equivalent width for normal nearby galaxies are illustrated in Figure 5. The distribution

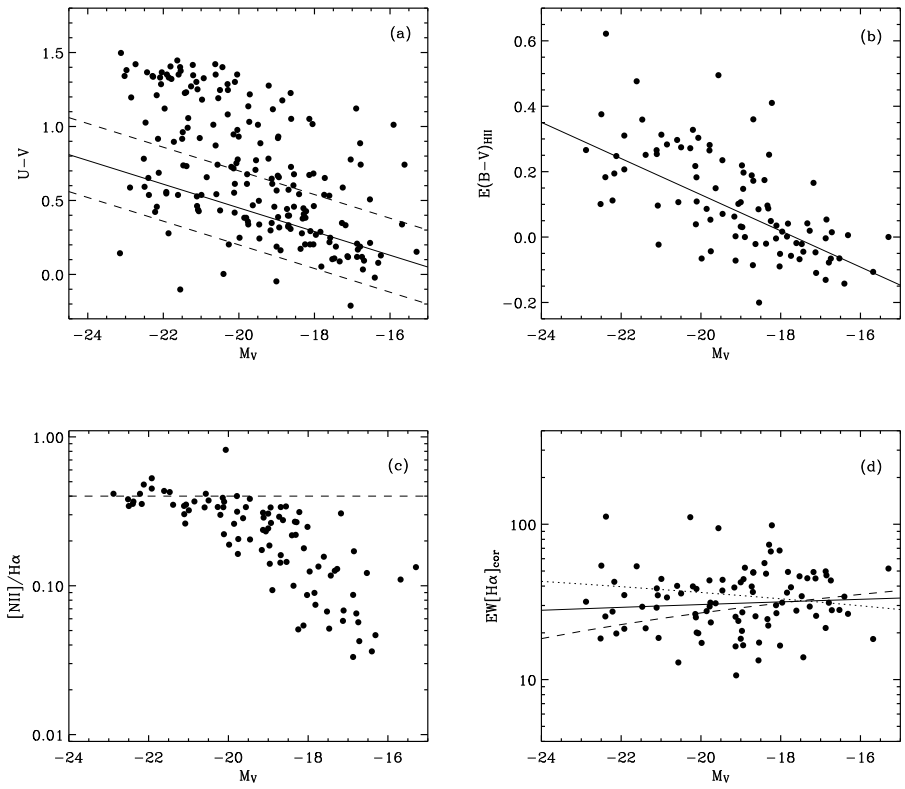

Figure 5 — $U-V$ colors and nebular emission lines properties versus absolute V-band magnitude of nearby normal galaxies from the Nearby Field Galaxy Survey (Jansen et al. 2000a). (*a*) The $U-V$ color versus absolute V magnitude. A linear fit to the blue sequence is shown (*solid line*). In the other panels we plot only the galaxies that are within 0.25 mag of the blue color-magnitude rellation (*dashed lines*). (b) The $E(B-V)_{HII}$ color excess versus absolute V-band magnitude, derived from the observed ratio of integrated fluxes of Hα and Hβ. The fluxes are corrected for Balmer absorption and Galactic reddening. The line shows a linear fit to the data (*d*). The metallicity sensitive [N II]λ6584/Hα ratio. The dashed line indicates Solar metallicity. (c) The Hα equivalent width EW[Hα], corrected for Balmer absorption and attenuation by dust. The correction for dust absorption was obtained from the measured $E(B-V)_{HII}$ and assuming an absorption of the stellar continuum $E(B-V)_{cont} = rE(B-V)_{HII}$, where $r = 0.7$. The solid line is a linear fit to the data. Also shown are linear fits to the data in the case that $r = 1$ (*dashed line*) and $r = 0.2$ (*dotted line*).

$U-V$ colors versus absolute V magnitudes is shown in (a). A blue sequence and a red sequence of galaxies are visible. A linear fit to the blue sequence using the same technique as described in §4.1 gives a slope of $\delta(U-V)/\delta M_V = -0.08$. We isolate the blue galaxies and only consider the spectral properties of galaxies with colors $\Delta(U-V) < 0.25$ from the linear fit.

In (b) we show the reddening $E(B-V)_{HII}$ towards HII regions versus absolute V-band magnitude. The reddening was computed by Jansen et al. (2000) from the observed Balmer decrement Hα/Hβ, assuming the intrinsic ration of 2.85 (case B recombination). As noted by Jansen et al., the sample shows a clear correlation; more luminous galaxies tend to have higher dust opacities.

The slope of the reddening-luminosity relation is $\delta E(B-V)_{HII}/\delta M_V = 0.06$. This reddening estimate applies to the H II regions in the galaxies, and it is generally thought that the mean reddening towards the stars is lower by a factor r: $E(B-V)_{cont} = rE(B-V)_{HII}$ with $0.5 \leq r \leq 1$ (Kennicutt 1998; Calzetti, Kinney, & Storchi-Bergmann 1996; Erb et al. 2003). For a Calzetti et al. (2000) dust law, the implied range of $U-V$ slope is 0.06 - 0.11, close to the observed value. Other forms of the extinction law, appropriate for the Milky Way (MW; Allen 1976) or the Small Magellanic Cloud (SMC; Gordon et al. 2003), yield similar values.

In (b) we show the relation between metallicity and absolute magnitude, as traced by nebular emission from H II regions. The correlation of the [N II]λ6584/Hα line ratio with integrated M_V shows that brighter galaxies are more metal-rich. To estimate the implications of this relation on the broadband colors, we explore two possibilities. Firstly, we use BC03 models to calculate the expected $U-V$ color variations for stellar populations of a wide variety of metallicities. We explored a grid of models with a range of exponentially declining star formation rates ($\tau = 0 - \infty$) and ages ($t = 0 - 13$ Gyr). We find that, regardless of age, models with blue colors $U-V < 0.5$ show variations with metallicities of $\Delta(U-V) \lesssim 0.1$, very small compared to the observed trend.

Finally, we show in (d) the relation between the Hα equivalent width and absolute magnitude, after correction for Balmer absorption and dust absorption. The EW[Hα] measures the instantaneous star formation rate per unit optical luminosity, and can be interpreted as the ratio of present to past averaged star formation rate or the time since the onset of star formation.

The correlation of EW[Hα] with M_V depends on the value of r used for the correction for dust extinction. It is straightforward to see that the extinction correction for log EW[Hα] is proportional to $(1-r)$: if $r = 1$, the H II regions and continuum are equally extincted, whereas for lower values of r, the $H\alpha$ is more extincted than the continuum.

We illustrate the effect in (d). The filled circles and the solid line show the data and a linear fit after correction with $r = 0.7$. There is no relation with absolute magnitude, implying that all the galaxies have the same age. The dashed line shows a linear fit to the data in the case of equal absorption $r = 1$. Now more luminous galaxies have slightly higher mean stellar ages. The dotted line shows $r = 0.2$, leading to a decreasing ages with increasing luminosity.

We can use the derived slopes to calculate the dependence of age and reddening as a function of magnitude. If metallicity variations are ignored, an exact solution

can be derived: The reddening of the H II regions follows immediately from the data, r and the slope of log age versus magnitude follow directly from the observed slopes of $U-V$ versus magnitude and EW[Hα] versus magnitude. The solution is

$$r = 0.78$$

and

$$\delta \log(age)/\delta M_V = 0.027$$

This age gradient is very small, and produces a very shallow color magnitude slope of 0.01, opposite to the observed slope. The main cause of the color magnitude relation is the variation of dust reddening with magnitude. These results does not change significantly if metallicity contributes -0.01 to the color magnitude relation.

We conclude that the data on local galaxies imply that the CMR for blue galaxies is mostly caused by variations in reddening, with only small contributions from age and metallicity variations.

4.4 Comparison to $z \sim 3$ Galaxies

We turn to our broadband observations at $z \sim 3$ and compare the situation in the local universe to that at high redshift. We select the 186 galaxies in the FIRES sample at $2.2 < z < 3.2$, and inspect the properties of the galaxies that lie within $\Delta(2200-V) < 1$ mag of the blue CMR.

4.4.1 The Models

Following previous studies of high redshift galaxies (Papovich, Dickinson, & Ferguson 2001; Shapley et al. 2001; Förster Schreiber et al. 2004), we fit stellar population models to the broadband SEDs of individual galaxies, and interpret the distribution of best-fit ages and extinctions. We use the publicly availably HYPERZ fitting code (Bolzonella, Miralles, & Pello 2000), updated with the synthetic template spectra from the latest version of the Bruzual & Charlot (2003) stellar population synthesis code. We use the Basel 3.1 library (Westera et al. 2002) of theoretical stellar spectra, selected the Padova1994 stellar evolutionary tracks, and adopted a Salpeter IMF with upper and lower mass cut-offs of 0.1 and 100 M$\odot$.

For simplicity, we only discuss the results for models with constant star formation rates and solar metallicity. Metallicities of blue, star-forming galaxies $z \sim 2-3$ are not well constrained,and may be solar (Shapley et al. 2004) or somewhat lower (Pettini et al. 2001). We prefer solar metallicity models as they have been directly calibrated against empirical stellar spectra. We refer to earlier studies for detailed discussions how the model assumptions affect the distribution of best-fit parameters Papovich, Dickinson, & Ferguson (2001); Shapley et al. (2001). Finally, we

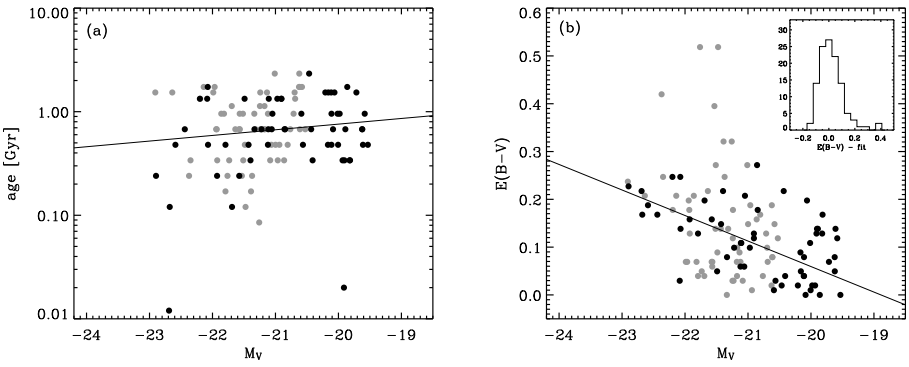

Figure 6 — The result of Bruzual & Charlot stellar population fits to the broadband SEDs FIRES fields at $2.2 < z3.2$. We show the distribution of best-fit stellar ages and extinctions as a function of absolute $V-$band magnitude for solar metallicity, constant star forming models, reddened with according to a Calzetti et al.(2000) dust law. Only the results for the galaxies within 1 mag in 2200-V of the blue CMR are drawn. (*a*) Best-fit ages and (*b*) best fit color excesses E(B-V) versus absolute magnitude. The solid lines shows linear fits to the data. and the inset shows the distribution of residuals around the best fit relation for the color excess.

include reddening by dust by adopting the Calzetti et al. (2000) starburst attenuation law. Hence, the models are characterized by 3 free parameters: age (time since onset of star formation), star formation rate (SFR), and the level of extinction.

4.4.2 Results

We kept 152 out of a total 186 galaxies for which the models could find an acceptable fit, based on the value of χ^2 per degree of freedom. The galaxies that did not fit well, could have had SEDs that are not well-described by our (incomplete) template set, or suffer from contamination by emission lines. The distribution of colors and magnitudes of the rejected galaxies is similar to that of the galaxies with good fits, and their exclusion is not likely to affect the results.

In Fig 6(a) we show the between best-fit age against M_v. A weak trend is indicated, in the sense that the faint galaxies are slightly older than the bright ones. A least squares fit indicates a slope of log(age) versus M_V 0.05 $\pm$ 0.06, significant at the 0.8 σ level.

In contrast, the reddening $E(B-V)$ shows a much stronger correlation with M_V in (c). We fitted a linear relation to the data finding a $\delta E(B-V)/\delta M_V$=0.05 $\pm$ 0.01. Remarkably, this is almost the same as the slope implied for the continuum reddening of nearby galaxies $\delta E(B-V)/\delta M_V$=0.047 (see §4.3).

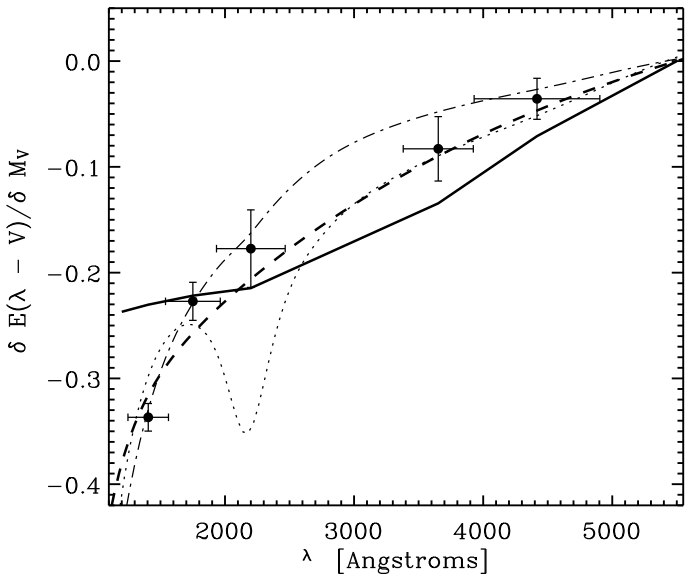

Figure 7 — The steepness of the CMR slope versus wavelength in the field of the HDFS. The data show the slope in the rest-frame $\lambda - V$ color versus M_V as a function of the filter λ. Overplotted are expectations for three extinction laws: the Calzetti et al. (2000) dust law (*dashed line*), the MW extinction law (Allen 1976; *dotted*), and the SMC extinction law (Gordon et al. (2003); *dash-dot line*). The normalization is different for each law: $\delta E(B-V)/\delta M_V$=0.04 for the Calzetti, 0.05 for the MW, and 0.02 for the SMC. The thick solid line represents the color-dependence of the CMR slope in the case that stellar population age correlates with M_V. We show the track for a solar-metallicity, constant star-forming model. The normalization is $\log(\text{age}) \propto -0.30\ M_V$.

4.4.3 The Blue CMR in Various Rest-Frame Colors

As another way to show why the data imply dust as the main cause of the CMR slope, we derived the slope of the CMR for many different restframe filters, and show the results in Figure 7. We described our rest-frame luminosities and colors earlier in §2.2. We overplot the expected relation as a function of wavelength for several extinction curves. These curves were scaled with an arbitrary constant to provide the best fit. As we can see, the Calzetti curve provides the best fit, whereas the SMC and MW curve fit progressively worse. We also show the expected dependence if the CMR is caused by age variations. Again, the amplitude of this curve is fitted to the data points. It does not fit at all to the bluest point, which is the CMR slope in $1400 - V$ color versus M_V.

We conclude that, models fits to SEDs at $z = 3$ are in agreement with the picture in local universe. Reddening by dust explaines the slope of the CMR consistently.

5 Constraints of the Color-Magnitude Relation on the Star Formation Histories of Blue Galaxies at $z \sim 3$

In this section we explore models to produce the narrow and asymmetric color distribution of galaxies along the CMR. We focus on the redshift range $2.2 < z < 3.2$, where the fraction of red galaxies is the lowest, and where perhaps the color distribution may be described by simpler models than at lower redshift. The distribution (Figure 1) is characterized by a blue peak and a skewness to much redder colors. The color distribution has a sharp cutoff at the blue side of the peak, and the scatter of the colors around the peak is quite low: most of the high redshift galaxies occupy a narrow locus in color space.

Here, our basic assumption is that the color scatter around the CMR is caused by age variations. We explore 4 different scenarios. In each scenario, we generate the complete star formation history of model galaxies, typically characterized by 2 or 3 free parameters. We then compile a large library of Monte-Carlo realizations, and generate the expected galaxy color distributions, taking into account observational errors and biases. We compare the model distributions to the observed color distribution to identify the best fitting parameterizations of the star formation histories. Kauffmann et al. (2003) used a similar method to constrain the star formation histories of local galaxies from the Sloan Digital Sky Survey.

A desirable property of this method, as will become apparent below, is that by using the information contained in the galaxy color distribution, we can resolve some of the degeneracies produced by fitting SEDs to the broadband colors of individual galaxies (e.g., Papovich, Dickinson, & Ferguson 2001), in particular the degeneracies in prior star formation history.

With this in mind, we will now discuss the model ingredients, the generation of the library, the fitting methodology, and the results for several different parameterizations of the star formation histories.

5.1 A library of Star Formation Histories

As the basic ingredient, we used the solar metallicy BC03 models as discussed in §4.4, and we fixed the BC03 model parameters except the star formation history: we only explore the effect of aging on the galaxy color distributions. We will compare the predictions of the models directly to our rest-frame luminosities and colors (see §2.2).

We generated a large library of Monte-Carlo realizations of stellar populations with different star formation histories. We classify the formation histories into 4 seperate scenarios, characterized by distinct parameterizations of the star formation rates. They range from simple constant star formation, to complex models including bursts. Details of the parameterizations are given together with the results in §5.3. We do not introduce additional complexity, such as a distribution of metallicities, or distributions in the star formation rate parameters, as we do not

have sufficient observational constraints to justify more freedom in the models.

For each scenario, and for each SFH parameter combination, we generated between 100-200 realizations of individual star formation histories. For every instance we saved a selection of parameters, including colors, magnitudes, star formation rates, and masses. The final libraries of the scenarios contain on the order of $\sim 10^5$ unique star formation histories each.

5.2 Fitting Method

Briefly, we draw samples from the libraries, generate distributions of colors and magnitudes, and compare them to rest-frame luminosity and colors of FIRES galaxies in the redshift range $2.2 < z < 3.2$. We fit the models seperately to the color distributions of the HDFS and MS1054 fields, as these have different depths and areas.

5.2.1 Creating Mock Observations

We need to account for three essential aspects of the data: the observations are magnitude limited, contain photometric errors, and have additional color scatter by variations in dust content.

It is evident that magnitude limits can substantially alter the resulting color distribution if galaxies in the models evolve strongly in luminosity. For example, an otherwise undetected galaxy undergoing a massive burst will temporarily increase in luminosity, and can enter a magnitude-limited sample, changing the color distribution. This effect is enhanced if there are many more galaxies below the magnitude limit than above. Hence, the steepness of the faint end slope of the luminosity function also plays a role.

We adopt a faint-end slope of $\alpha = -1.6$ according to the rest-frame far-UV luminosity function of Steidel et al. (1999). Shapley et al. (2001) found a steeper slope, but it resulted from a positive correlation of observed R and $R - K_s$ photometry, which we do not see in our data. We applied the luminosity function in the following way.

From the models we generate sets of galaxies with the same redshift distribution as the observed distribution. We then draw a luminosity from a luminosity function with $\alpha = -1.6$ and scale the model galaxy to that luminosity. Instead of using the instantaneous luminosity of the mock galaxy at the time of observation to compute the scaling, we use its median luminosity over the redshift range $2.2 < z < 3.2$.

Furthermore, we add photometric errors to the model colors as a function of model luminosity. The standard deviations of the errors were determined from a linear fit to the errors on the rest-frame luminosities and colors as a function of rest-frame M_V. Hence, these include the photometric redshift uncertainties. The fit gives a mean error in the rest-frame $2200 - V$ color of 0.15 in the HDFS and

0.19 for the MS1054.

Finally, we include reddening by dust in the models, adopting the Calzetti et al. (2000) starburst attenuation law. We add a distribution of color excesses

$$E(B-V) = 0.1 \pm 0.05(1\sigma)$$

to the model colors and magnitudes, where $E(B-V)$ is required to be greater than 0. The mean value is appropriate for $M_V = -21$ galaxies from the best-fit models in §4.3, and we used a scatter of 0.05 equal to that found in local galaxies (see §4.2) and similar to the distribution of $E(B-V)$ in §4.3. Using this distribution of extinctions, the scatter added to the model colors is $\sigma(2200-V)_{dust} \sim 0.2$. Adopting an SMC extinction law would result in $\sigma(2200-V)_{dust} \sim 0.15$. We remark that the extinction variations of the FIRES galaxies at $z = 2-3$ are well-constrained by the low scatter in the rest-frame far-UV colors.

Summing up, the average $(2200-V)$ scatter introduced into the models by including both photometric uncertainties and dust variations is 0.25 mag for the HDFS, and 0.28 mag for the MS1054 field. Note that the width of the observed scatter at $2.2 < z < 3.2$, which we characterize by the central 32% of the color distribution, is 0.58 and 0.83 (for the HDFS and MS1054, respectively; see also Table 2).

Finally, to complete the mock color distributions, we imposed a magnitude limit of $M_V = -19.5$ for the HDFS field, and $M_V = -20.5$ for the field of MS1054. We are complete for all SED types down to these magnitude limits.

5.2.2 The Fitting

First, we subtracted the color-magnitude relation $\delta(2200-V)/\delta M_V$ from the observed colors as in §4.2, and we normalize the $2200-V$ color distribution to the color of the CMR intercept at $M_V = -21$. Then, for each scenario and for each field, we found the best-fit parameters by performing the two-sided Kolmogorov-Smirnov (KS) test on the unbinned color distributions of the models and the data. Next, we multiplied the KS-test probabilities of the individual fields, and we selected the parameter combination that yielded the highest probability.

A property of the KS-test is that it tends to be more sensitive around the median values of the distribution than at the extreme ends (Press et al. 1992). Given this behaviour, it is easy to see that our assumptions for the stellar population model, such as IMF and metallicity, may impose strong, yet undesirable, constraints on the fits. For example, a steeper IMF slope less rich in high-mass stars (e.g., Scalo instead of Salpeter) results in inherently redder spectra, whereas lower metallicity of the stellar populations leads to intrinsically bluer spectra. The KS-test would be sensitive to these systematic color changes in color, violation our basic assumption that the variations in color are caused by variations in age alone.

Furthermore, both the FIRES fields contain some very red galaxies (Franx et al. 2003; van Dokkum et al. 2003). These galaxies are thought to have much larger dust extinctions, higher ages, and perhaps different star formation histories than the $z \sim 3$ blue sequence galaxies we model here (see, e.g., van Dokkum et al. 2004; Förster Schreiber et al. 2004a). In principle, we do not expect our simple models to account for them, and while the KS-test is relatively insensitive to a small numbers of outliers, we might consider excluding them from the fit.

For these reasons we perform 3 different KS-tests, which we will refer to as KS1, KS2, and KS3. Firstly, we do a straightforward two-sided KS-test on the model and observed distributions. For the second, we normalize both the model and the observed color distribution to the median, and hence we only test for the shape of the color distribution, and not the absolute value of the color. Lastly, we compare the median normalized distributions after excluding galaxies that are $\Delta(2200 - V) > 1.75$ redder than the mode of the distributions, thereby removing the reddest outliers. The mode is calculated as in §4.1.

In the subsequent presentation of the results we focus on the second test (KS2), which is sensitive to the shape only, but we discuss the outcome of the other tests where appropriate.

We note in advance, that due to inherent differences between the fields the combined KS probabilities will never reach high values. For example, a two-sided KS-test comparing the MS1054 data to the observations of the HDFS (to a magnitude limit of $M_V = -20.5$) gives a probability of 70%. We do not expect any model that fits both data sets to exceed this probability.

5.3 Results

5.3.1 Constant Star Formation

In scenario 1, we assume galaxies start forming stars at random redshifts z_f. We take z_f to be distributed uniformly in time from $z = 2.2$ up to a certain certain maximum redshift z_{max}, where z_{max} is between $3.2 < z_{max} < 10$ in steps of 0.2. The star formation rate of each individual galaxy is constant for a certain time t_{sf} after which star formation ceases. We construct predictions for values of t_{sf} sampled logarithmically in 10 steps from 0.05 to 3 Gyr. Hence, this model is characterized by two parameters: z_{max} and t_{sf}.

In Figure 8(a) we show the color evolution as a function of time for a characteristic star formation history in this model, together with a schematic representation of the star formation rate. We also overplot the color intercept of the observed CMR in the HDFS and MS1054 fields at the median redshift of observation. In (b) we show the corresponding color-magnitude evolution. Figure 8 (c) shows the entire color-magnitude distribution of the best-fit models in the fields of the HDFS and MS1054. The observations are overplotted after subtracting the color-magnitude relation $-0.17(M_V + 21)$ (see also §4.1). In (d) we show the absolute

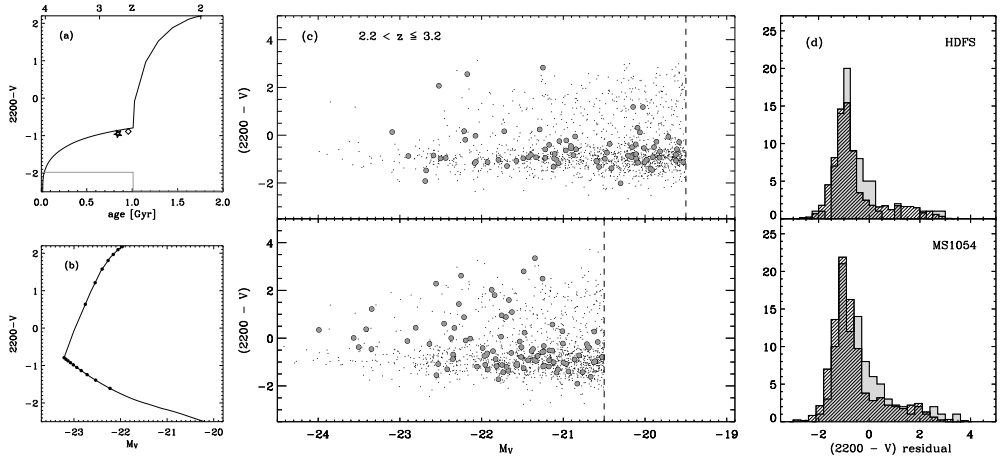

Figure 8 — The results of model 1. Galaxies in this model start off at random redshifts $z_0 < z_{max}$, and form stars at a constant rate for a duration t_{sf} before stopping. The maximum formation redshift z_{max} and t_{sf} are the free parameters. (*a*) The stellar population track of $2200 - V$ color against age for a characteristic galaxy ($z_0 = 4$, $t_{sf} = 1$ Gyr). The star formation rate is illustrated schematically by the gray line. Also drawn is the observed $2200 - V$ color of the blue CMR at fixed $M_V = -21$ in the HDFS (*star*) and MS1054 (*diamond*). A mean dust reddening of $E(B-V) = 0.1 \approx E(2200-V) = 0.4$ (Calzetti et al 2000) is added to the model color. (*b*) The track of $2200 - V$ color against absolute V-band magnitude for the same galaxy in steps of 100 Myr (*filled circles*) (*c*) The full color-magnitude diagram of the best-fit model (*black points*) in the redshift range $2.2 < z < 3.2$. The filled gray circles are the data from the HDFS (*top*) and MS1054 (*bottom*). The color distribution of the model is broadened to account for photometric errors and scatter in the dust properties, where we used $\sigma(2200-V)_{dust} = 0.2$. Only galaxies brighter than the absolute magnitude cut-off (*dashed line*) are included in the fit. (*d*) The histograms of residual $2200 - V$ colors show the data (*gray histograms*) and the best-fit model (*hatched histograms*). The best-fit parameters are $z_{max} = 4.6$ and $t_{sf} = 1$ Gyr.

color histograms of the best-fit model together with the data. We note that the "shape-sensitive" KS2 test was used to find the best fit.

The best-fit model in the KS2 test has a maximum formation redshift $z_{max} = 4.6$ and a constant star formation timescale $t_{sf} = 1$ Gyr. Some characteristics are reproduced, such as the narrow blue peak, but the asymmetric profile in the blue peak is not, although there is a low-level tail to red colors containing passively evolving galaxies. The fit is rather poor; the KS2-test assigns a 0.11 probability to the fit, the KS1 gives a maximum 0.01, while the KS3 gives 0.04.

In order to understand the behaviour of this model, we explored the z_{max} and t_{sf} parameter space. We found this model cannot produce the distinct skewness towards red colors in the blue peak, even if it cuts off star formation in a substantial fraction of the galaxies around the epoch of observation. This would be only mechanism to produce a red skew in this model, but the reddening of the passively

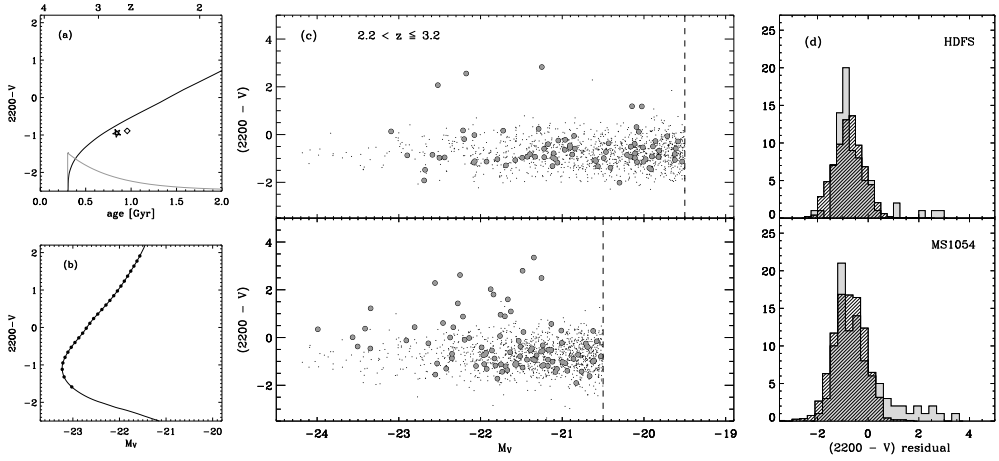

Figure 9 — Same as Figure 8 for model 2. Galaxies start forming stars at random redshifts $z_0 < z_{max}$, and form stars at an exponentially declining rate with e-folding time τ. The maximum formation redshift z_{max} and τ are the free parameters in this model. A characteristic galaxy shown in (*a,b*) has $z_0 = 3.5$ and $\tau = 0.5$ Gyr. The best-fit parameters of the in model (c,d) are $z_{max} = 4.2$ and $\tau = 0.5$ Gyr.

evolving galaxies is so rapid (see Fig. 8a) that instead of a red wing, it produces a prominent second red peak of passively evolving systems.

5.3.2 Exponentially Declining Star Formation

Scenario 2 is almost identical to the first, but now each individual galaxy has a single exponentially declining star formation rate with timescale τ. This model is characterized by two parameters: τ and z_{max}, where τ is sampled logarithmically in 10 steps from 0.05 to 3 Gyr.

Figure 9 shows a characteristic star formation history, and the fitting results of scenario 2. The KS2 test yields a best-fit model with a maximum formation redshift $z_{max} = 4.2$ and a constant star formation timescale $\tau = 0.5$ Gyr. It fits very poorly however, as the distribution is too broad and symmetric around the median. The KS2 test rules out this model at the 99% confidence level. The KS1 and KS3 test give similar answers.

Exploring the z_{max} and τ parameter space, we can understand why the fit is always poor. The only way for this model to produce a red skew in the color distribution, is by having a relatively small value for the star formation timescale τ compared to the mean age of the galaxies at $2.2 < z < 3.2$. In that case, a substantial number of galaxies is entering a "post-starburst" phase, where the

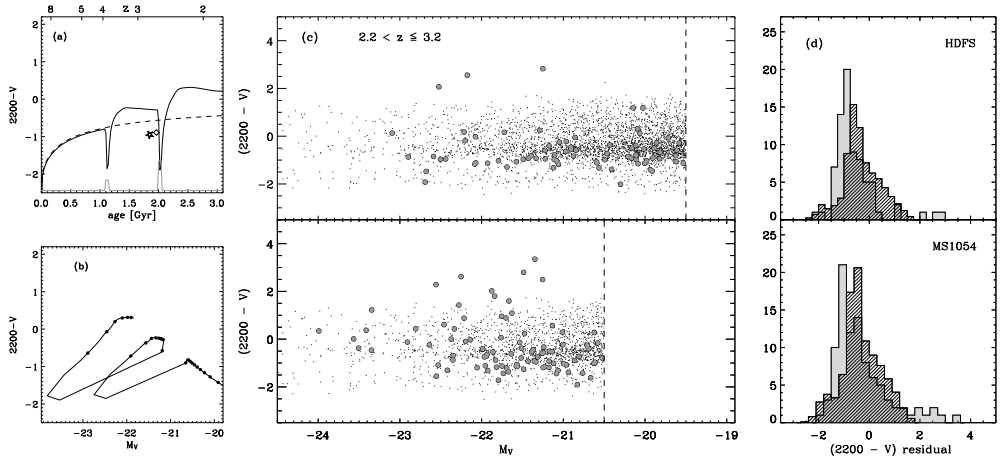

Figure 10 — Same as Figure 8 for model 3. Repeated burst models, with a multicompenent star formation history: an underlying constant star formation rate, and superimposed bursts of a certain strength $r = M_{burst}/M_{tot}$ and frequency n. The formation redshift is fixed to $z = 10$. A characteristic galaxy shown in (a,b) has $n = 1$ and $r = 1$ Gyr^{-1}. The best-fit parameters of the model in (c,d) are $n = 0.3$ Gyr^{-1} and $r = 4$. The dashed line in (a) represents a constant star forming model (Bruzual & Charlot 2003).

mean stellar age is $t > \tau$ and the instantaneous SFR becomes much smaller than the past average. These galaxies gradually move away from the blue CMR, creating a skewness to red colors.

However, $\tau = 0.5$ Gyr models redden more quickly than CSF models at all ages, as can be seen comparing Fig 9a with Fig 8a. This results in somewhat redder mean colors, and a broader spread on the blue side of the peak, created by the newly formed galaxies that are continuously added to the sample.

5.3.3 Repeated Bursts

In scenario 3, all galaxies start forming stars at a fixed $z = 10$. The stars form in two modes: a mode of underlying constant star formation, and superimposed on this, random star bursts. The bursts are distributed uniformly in time with frequency n: the average number of bursts per Gyr. The amplitude of the burst is parameterized as the mass fraction $r = M_{burst}/M_{tot}$ where M_{burst} is the stellar mass formed in the burst and M_{tot} is the total mass formed by the constant star formation and any previous bursts combined. During a burst, stars form at a constant, elevated rate for a fixed time $t_{burst} = 100$ Myr. Thus, there are two free parameters in this model: the burst frequency n, and the burst strength r. We sample n as $n^{1/2}$ from 0.1 to 6 Gyr^{-1}, and r as $r^{1/3}$ from 0.01 to 4.

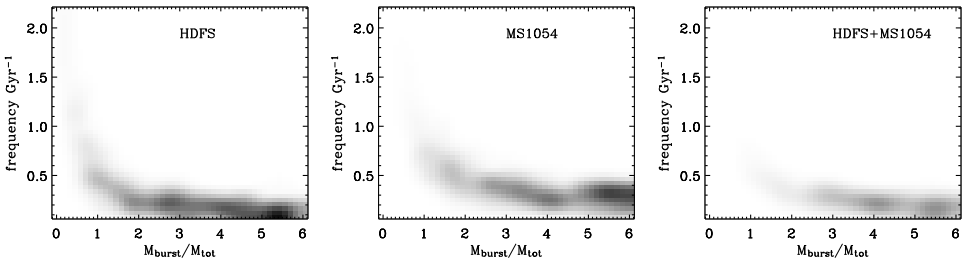

Figure 11 — The map of KS-probabilities versus model parameter repeated burst models of scenario 3. The free parameter were the burst frequency, and the strength $r = M_{burst}/M_{tot}$ of the bursts. The grayscale linearly encodes the probability of the fit. It shows the degeneracies of the model parameters, and regions of parameter space that are clearly excluded. The panels correspond to the model fits to the HDFS data, MS1054 data, and the data sets combined.

Figure 10 shows a characteristic star formation history and the best-fit results of scenario 3. The KS2 test yields a best-fit model with an average burst frequency of 0.3 Gyr^{-1} and a mass fraction formed in each burst of $r = M_{burst}/M_{tot} = 4$ (= 400%): extremely massive, but relatively infrequent bursts. This solution reproduces the correct shape of the color distribution, i.e., the blue cut-off and red skew, but not the absolute colors. The median $2200 - V$ color is 0.4 mag too red, reflecting the $z = 10$ formation redshift. The KS2 test probability for this model is 0.35, with the KS3 test giving a similar value. The KS1 test rejects the model at the 99% confidence level, as a result of the wrong median color.

We explore in Figure 11 the KS2-fit probability over the relevant part of n, r parameter space. In both fields, only models with infrequent massive bursts are allowed, although the exact strength is relatively unconstrained. There is notable difference between the HDFS and MS1054 fields. Specifically, the broader red wing in MS1054 observations compared to the HDFS favors a higher burst frequency.

As illustrated in Figure 10(a), the skewness towards red colors is produced by post-starburst galaxies. These galaxies have just formed large numbers of A-stars in the previous burst that outshine for some time the O- and B-stars formed in the underlying mode of constant star formation, hence biasing the integrated colors to the red. Interestingly, the color of a galaxy after a starburst always remain redder compared to constant star formation (or no bursts). Concluding, recurrent bursts do not "rejuvenate" the ensemble of galaxies. Rather, after a very brief blue period during the burst, the galaxy is redder forever after.

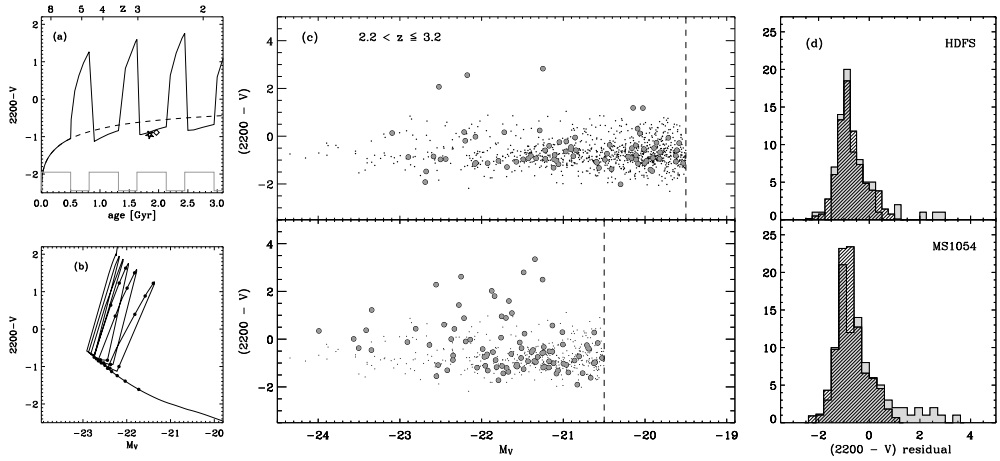

Figure 12 — Same as Figure 8 for model 4. Episodic star formation model with active periods of star formation and times of quiescence. The free parameter are the star formation duration t_{sf}, the fractional duration of quiescence r_q, and the remaining fraction of star formation r_{sfr} in the non-active period. The formation redshift is fixed to $z = 10$. A characteristic galaxy shown in (*a,b*) has $t_{sf} = 500$ Myr, $r_{tq} = 0.6$ Gyr^{-1}), and $r_{sfr} = 0.02$. The best-fit parameters of the model in (c,d) are star formation duration $t_{sfr} = 200$ Myr, $r_{tq} = 0.4$, and $r_{sfr} = 0.02$. Note that the stellar population after resuming star formation is bluer than before it stopped.

5.3.4 Episodic Star Formation: The Duty Cycle

In the final scenario, galaxies again start at a fixed $z = 10$, and subsequently form stars at a constant rate for an "active" period of length t_{sf}. Then star formation suddenly drops to some fraction r_{sfr} of the nominal value, while the galaxy passes through a quiescent period that takes a fraction r_{tq} of the "active" episode $r_{tq} = t_q/t_{sf}$. This constitutes one duty cycle, and these star formation histories are characterized by repeating cycles of fixed length. We randomize only the phases of the cycles. There are three free parameters in this model: t_{sf}, r_{tq}, and r_{sfr}. We sampled t_{sf} logarithmically in 10 steps from 0.05 to 3.0 Gyr, r_{tq} logarithmically in 16 steps from 0.04 to 40, and r_{sfr} logarithmically in 10 steps from 0.01 to 0.8.

Figure 12 shows a characteristic star formation history and the fitting results of scenario 4. Using the KS2 test, the best-fit model galaxies have active periods of $t_{sf} = 200$ Myrs, and quiescent periods of 50 Myr ($r_{tq} = 0.4$), during which star formation drops to $r_{sfr} = 2\%$ of original rate. Figure 12(d) shows that the key features of the observed color distribution are correctly reproduced: the blue mean colors and the asymmetric shape of the observed distribution. The fit probability of the KS2 test is 0.58, and the KS1 test yields comparable probabilities, indicating that the model also produces the correct blue absolute colors. The KS3 test, which

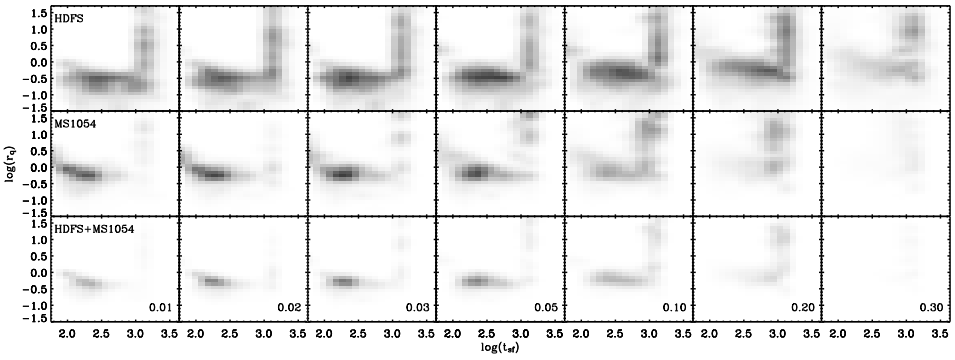

Figure 13 — The map of KS-probabilities versus model parameters in the duty-cycle models of scenario 4. The free parameter were the star formation duration t_{sf}, the fractional duration of quiescence r_q, and the remaining fraction of star formation r_{sfr} in the non-active period. The seven panels in each row correspond to 2-dimensional slices of the 3-dimensional parameter cube. Each slice is taken at a fixed r_{sfr}, which is indicated in the lower-right corner. The three rows correspond to the model fits to the HDFS data, MS1054 data, and the data sets combined. The grayscale encodes linearly the probability of the fit.

excludes the reddest galaxies, yields 0.89, reflecting that the mismatch with the model, but also between the two fields mutually, is mainly the enhanced number of very red galaxies in the field of MS1054 (see also Förster Schreiber et al. 2004). We discussed in §5.2 that our models are not necessarily expected to describe the small number of very red galaxies, and we conclude that this model provides the best description of the shape of the blue peak.

In Figure 13 we show the distribution of the (KS2) fit probabilities over the entire 3-dimensional parameter space. We present 2-dimensional slices of parameter space at steps in fraction of residual star formation rate r_{sfr}. The parameter space allowed by the two datasets are somewhat different. The scatter in the HDFS is intrinsically smaller, leading to smaller values of r_q and allowing longer star formation durations t_{sf}. Also noticable is the "plume" of reasonably high probabilities in t_{sfr}, r_{tq}. This region of parameter space resembles the model discussed in §5.3.1, where star formation is constant and stops exactly at the age of observation. The broader red wing of MS1054 clearly favors longer quiescent periods and short duty-cycles. It can also be seen that the length of the cycles is not well-constrained, and depends on the level of residual star formation rate r_{sfr}. If the intra-burst star formation rates are high, then the quiescent periods can be longer.

A striking aspect of the color evolution of a galaxy in this model is that when star formation resumes after the quiescent period, then the color of the population is at least as blue as, or even bluer than before quiescence (see Fig 12a). The high probabilities of the KS1 test already suggested that, in contrast to the bursting

model, galaxies with cycling star formation histories manage to maintain their extremely blue, colors despite the high formation redshift $z = 10$. The repeating "rejuvenating" of the colors leads to a substantial slower color evolution with time than a constant star forming model, as evidenced from the blue "ridge" of the color evolution track in Fig 12(a).

5.4 Discussion

The results presented here show that if we use the information contained in the observed color distribution of the galaxies, we can effectively constrain the simple models of the star formation history.

The first two scenarios (constant SFR and cut-off, and declining SFRs) were selected for their simplicity, and because these type of star formation histories are often assumed in SED modeling of high redshift galaxies (e.g., Shapley et al. 2001; Papovich, Dickinson, & Ferguson 2001; van Dokkum et al. 2004; Förster Schreiber et al. 2004a).

It is worrying that these scenarios generally fail to reproduce the observed color distribution. In addition, the parameter values for the best fit are not realistic. The maximum formation redshift is generally low $z = 4 - 4.5$, and the timescale of star formation (t_{sf} or τ) is comparable to the mean age of the galaxies at the $2.2 < z < 3.2$ redshift of observation. Hence, this is a special moment in the formation history, with many galaxies switching from active star formation to passive evolution or much lower SFRs. Such models predict profound evolution in the color distribution from $z = 4$ through $z = 2$, in contrast to the modest changes in the observed color distribution (Papovich, Dickinson, & Ferguson 2001; Papovich et al. 2004, this work).

The more complex models we considered are characterized by a fixed, high formation redshift ($z = 10$) and constant star formation, which is modulated by random events, such as starbursts.

The repeated burst model (scenario 3) with underlying constant star formation, does explain the shape of the observed color distribution, but we disfavore it for two reasons. Firstly, the adding of bursts does not rejuvenate the galaxy colors but leads to exceedingly red mean colors, mismatching the observations. We note that part of if can be resolved by tuning model assumptions, e.g., modifying metallicity, IMF, formation redshift, etc.

The second, more suspicious aspect is that the model needs to produce bursts with a high mass fraction, so that the longer-lived, but intrinsically fainter stars produced in the burst, outshine the luminous O- and B-stars created in the underlying mode of constant star formation. High mass fractions imply extreme instant star formation rates. Given a burst duration of 100 Myr, an $r = M_{burst}/M_{tot} \approx 4$ burst at $z = 3$ leads to a $60-$fold increase of the instant star formation rate, transforming an ordinary L^* Lyman break galaxy into a monster with SFRs of 2000

$M_{\odot}yr^{-1}$. Such galaxies are not observed, unless they are also temporarily highly abscured, and show up at sub-mm wavelengths as "SCUBA" sources. Slightly longer burst durations do not alter this picture.

It is evident that the underlying constant star formation artificially introduced the need for massive bursts. Should it temporarily cease, then the bursts would not need to involve such high mass fractions to produce a red wing in the color distribution. In fact, if the underlying star formation temporarily ceased, one does not need bursts at all: hence the episodic model.

The episodic models (scenario 4) reproduced the color distribution best of the scenarios explored here. The best-fit parameters correspond to quiescent periods lasting 30-50% of the time of an active period, and typical duration of the total duty cycle of 150 Myr to 1 Gyr. The relative fraction of the quiescent period is better constrained than the length of the duty cycle, which correlates with the level of residual star formation in the passive periods.

The episodic star formation in these models rejuvenate the galaxies during each episode, making it significantly bluer than a galaxy with constant star formation of the same age. This also demonstrates the dependence of the derived ages on prior star formation history. If the broadband SEDs of our mock galaxies at $z \sim 3$ were fit with constant star forming stellar populations, then the best-fit ages would be ~ 500 Myr, or a "formation redshift" of the galaxies of $z_f \sim 4$, instead of the true $z_f = 10$.

The episodic model could be a solution to the enigma presented by studies of the broadband SEDs of $z \sim 3$ Lyman Break Galaxies (Papovich, Dickinson, & Ferguson 2001; Shapley et al. 2001). When the SEDs of LBGs were fit with constant or declining star forming models, the resulting best-fit ages (100-300 Myr) were much smaller than the cosmic time span of the $2 < z < 3.5$ observation epoch or the age of the universe (1 Gyr, and 2 Gyr, respectively). Papovich, Dickinson, & Ferguson (2001) already cautioned that constant or declining SFH are probably wrong for most galaxies, and suggested that episodic star formation could provide a mechanism to rejuvenate the appearance of the galaxies. Here we have formulated such a model, subjected it to a quantitative test, and constrained its parameters.

Even more constraints can be placed by taking into account the observed evolution of the color distribution as a function of redshift. The models described here may be tested directly by the data at lower redshift. However, one reason why we focused on the redshift range $2.2 < z < 3.2$ is that here the fraction of red galaxies is the lowest, and our simple models are perhaps appropriate. At lower redshift the color distribution is more complex with a prominent blue and red peak (see Figure 1), and such simple models might not apply.

6 The Onset of the Red galaxies

We could already see in Figure 1 that the color distribution in the FIRES fields evolves strongly from $z \sim 3$ to $z \sim 1$. This trend is particularly clear in the lightgray color histograms in (b) and (d). Most notably, at $z \sim 3$ most galaxies are on the blue sequence, and there is no evidence for a well-populated red peak. The onset of a red peak is tentatively observed in the histograms at $z \lesssim 2$, in particular in the field of MS1054. The red peak in the $z \sim 1$ bin of the MS1054 field contains a major contribution of the cluster of galaxies at $z = 0.83$. A prominent red sequence is also observed in photometrically selected samples in the field up to $z = 1$ (e.g., Bell et al. 2004b, Kodama et al. 2004).

We calculate the red galaxy fraction as a function of redshift using a simple color criterion to separate red from blue galaxies. Classically, color-limited fractions are defined relative to the red color-magnitude relation (Butcher & Oemler 1984), but such a definition is unusable here, as the red sequence is virtually absent at $z \sim 3$ in our sample. We therefore define red galaxies as all galaxies more than 1.5 magnitudes redder than the blue color-magnitude relation: $(2200 - V) + 0.17M_V > 1.5$. The threshold was set to be $2 - 3$ times the scatter in $2200 - V$ color around the blue sequence. There is no evidence from the data that the scatter is a function of redshift (see Table 2).

Figure 14(a) shows the evolution of the fraction of red galaxies by number (N_{red}/N_{tot}). We also show the absolute luminosity density contributed by the red and blue galaxies seperately in (b), and the fraction of total luminosity contributed by red galaxies (c). We calculated the fractions to a fixed rest-frame magnitude limit to which we are complete for all SED types at $z \sim 3$. The limits are $M_V = -19.5$ in the field of the HDFS, and $M_V = -20.5$ in the field of MS1054. The luminosity densities were computed by adding the luminosities of the galaxies above the magnitude limit, and dividing it by the cosmic volume of the redshift bin. The uncertainties in all estimates were obtained by bootstrapping the color-magnitude distributions.

In both fields, we find a sharp increase from $z \sim 3$ to $z \sim 1$ in the relative number, the absolute rest-frame $V-$band luminosity density, and the relative $V-$band luminosity density of red galaxies. At the same time V-band luminosity density in luminous blue-sequence galaxies remains constant, or decreases.

The differences between the fractions in the fields are substantial. In the lowest redshift bin ($z \sim 1$), part of it can be explained by the contribution of the cluster. At higher redshift however, small number statistics and the variations in space density of the red galaxies due to large scale structure are the probable cause (see, e.g., Daddi et al. 2003). The FIRES fields are still small; the total surveyed area is less than 30 arcmin^2. We note that effect of the $2200 - V$ color scatter introduced by the limited accuracy of our photometric redshift technique, 0.32 in the median for the red galaxies, is minor. The dominant error is the limited number of galaxies in the sample, which is accounted for in the bootstrap calculation of

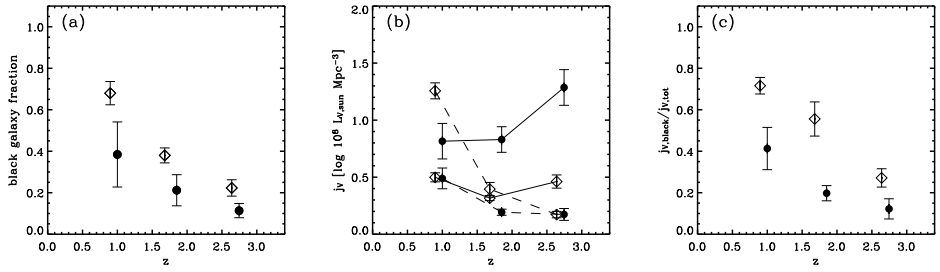

Figure 14 — Evolution of the fraction of red galaxies as a function of redshift. in the HDFS (*filled circles*) and MS1054 (*diamonds*). In (*a*) we show the number fraction of galaxies that are $(2200 - V) > 1.5$ mag redder than the blue sequence. (*b*) The absolute luminosity density in luminous blue (*solid line*) and red galaxies (*dashed line*). (*c*) The relative fraction of the total luminosity density in red galaxies. There is a sharp increase in the luminosity density in red galaxies between z=3 and z=1. We note that the field of MS1054 contains a massive cluster at $z = 0.83$.

the uncertainties.

The evolution of the red galaxies at high redshift appears to be different from the evolution between $z = 1$ and $z = 0$. Classic studies of the colors of galaxies at $z < 1$ found no evidence luminosity evolution in red galaxies, although the errorbars were fairly large (e.g., Lilly et al. 1995; Lin et al. 1999; Pozzetti et al. 2003). Recently, Bell et al.(2004b) found a constant luminosity density in photometrically selected red galaxies in the range $0.2 < z < 1.1$, and interpreted this as an increase in stellar mass in the early-type galaxy population, in apparent agreement with the hierarchical models of galaxy formation of Cole, Lacey, Baugh, & Frenk (2000).

We must defer such interpretations for our sample, until we better understand the nature of the red galaxies at $z > 1$. The foremost questions are whether all red galaxies are truly early-types with passively evolving stellar populations, or do other factors, such as reddening by dust play a role. Furthermore, when does the narrow red sequence establish it self?

In the local universe, $\sim 70\%$ of the galaxies in the narrow red sequence are morphologically early-type (Strateva et al. 2001; Hogg et al. 2002). The morphologies of $z \sim 0.7$ red peak galaxies in the GEMS survey (Rix et al. 2004), indicate that 85% of their rest-frame V-band luminosity density comes from visually classified E/S0/Sa (Bell et al. 2004a). In contrast, 70% of the Extremely red objects (EROs) at $1 < z < 2$ are believed to be dust reddened star-forming objects (e.g., Yan & Thompson 2003; Moustakas et al. 2004), while many of the $z \sim 3$ red galaxies in the FIRES sample also show signs of highly reddened star formation (van Dokkum et al. 2004; Förster Schreiber et al. 2004). Recent mid-IR imaging with IRAC on the Spitzer Space Telescope in the HDFS, indicates that only 25% of them are

dominated by passively evolving stellar populations (Labbé et al. 2004). Clearly, future study through spectra and MIR-imaging is necessary to give us a better understanding of the nature of the red colors

The next step in this kind of analysis would be to establish at what redshift the narrow red sequence establishes itself. Unfortunately, our photometric redshifts are too uncertain to allow an determination rest-frame colors with an accuracy better than 0.04 mag, typical of the scatter in the red color-magnitude relation (Bower, Lucey, & Ellis 1992). Hence spectroscopy is needed to establish the onset of the red color-magnitude relation.

7 Summary and Conclusions

We used deep near-infrared VLT/ISAAC imaging to study the rest-frame color-magnitude distribution of infrared selected galaxies in the redshift range $1 < z < 3$. We found a well-defined blue peak of star-forming galaxies at all redshifts. The blue galaxies populate a color-magnitude relation (CMR), such that more luminous galaxies in the rest-frame V-band tend to have redder ultraviolet-to-optical colors.

The slope of the CMR does not evolve with time, and is similar to the slope of blue, late-type galaxies in the local universe. Analysis of the spectra of nearby late-type galaxies from the NFGS suggests that the slope can be fully explained by the observed correlation of dust content with optical luminosity. The zeropoint of the blue peak at a given magnitude reddens smoothly from $z = 3$ to $z = 0$, likely reflecting an increase of the mean stellar age and an increase in the mean dust opacity of blue-sequence galaxies.

A key feature of the blue CMR relation is that the color distribution around it is asymmetric, with a blue "ridge" and a skew towards red colors. Assuming the scatter is caused by variations in mean stellar age, we have constructed models to explore the constraints that these observations place on the star formation history of blue field galaxies at $z = 2-3$. In the best-fitting models, galaxies form stars in short "duty-cycles", characterized by alternating episodes of active star formation and quiescence. In these particular models, the best constrained parameter is the relative duration of the quiescent period, which is 30-50% of the length of an active period. The best-fit total length of the duty-cycle is uncertain, as it correlates with the amount of residual star formation during quiescence; the data allow a range of 150 Myr to 1 Gyr.

The models can explain the color distribution well, and, surprisingly, can also produce the very blue colors of $z = 2 - 3$ galaxies. The colors are bluer than the colors of galaxies with the same age with constant star formation, because the galaxies are rejuvenated by each burst after each quiescent period. These models can therefor solve the paradox that $z \sim 3$ galaxies appear much younger than the age of the universe when fit with constant star formation models (e.g., Papovich, Dickinson, & Ferguson 2001). The solution is simply that the luminosity weighted

ages of the stars derived from model fitting should not be interpreted as the time since the galaxy first started forming stars.

Finally, we find a sharp increase from $z \sim 3$ to $z \sim 1$ in the relative number, the absolute rest-frame $V-$band luminosity density, and the relative luminosity density of luminous red galaxies, while the V-band luminosity density in luminous blue-sequence galaxies is constant, or decreases. Studies at redshifts $z \leq 1$ imply very little evolution, suggesting that the bulk of the evolution takes place round $z \sim 2$. Obviously, our fields are very small, and the variations between the fields are large enough to suggest that the uncertainties in the red galaxy fraction are dominated by large scale structure (see e.g., Daddi et al. 2003.

On the other hand, the scatter and evolution of the blue galaxy sequence is similar between the fields, indicating that field-to-field variations do not play a large role there. Here the challenge is to confirm directly the cause of the relation between color and magnitude, and the cause of the skewness to red colors of the scatter around the relation. Very high signal-to-noise spectroscopy in the Near-IR will be needed for this purpose, to measure the balmer emission lines, in order to estimate ages and reddening. Furthermore, these studies need to be performed at higher redshifts. The advent of multi-object NIR spectrographs on 8-10m class telescopes will make such studies feasible in the near future.

Acknowledgments

We thank the staff at ESO for their dedicated work in taking these data and making them available. This research was supported by grants from the Netherlands Foundation for Research (NWO), the Leids Kerkhoven-Bosscha Fonds, and the Lorentz Center.

References

Allen, D. A. 1976, MNRAS, 174, 29P
Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezić, Ž., Lupton, R. H., Nichol, R. C., & Szalay, A. S. 2004, ApJ, 600, 681
Bell, E. F., et al. 2004, ApJ, 600, L11
Bell, E. F., et al. 2004, ApJ, 608, 752
Bell, E. F. & de Jong, R. S. 2000, MNRAS, 312, 497
Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393
Bessel, M. S. 1990, A&AS, 83, 357
Blanton, M. R., et al. 2003, ApJ, 592, 819
Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601
Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476
Bruzual A., G. & Charlot, S. 1993, ApJ, 405, 538
Bruzual et Charlot et al. 2003 (BC03)
Butcher, H. & Oemler, A. 1984, ApJ, 285, 426
Casertano, S. et al., AJ, 120, pp. 2747–2824, 2000

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1996, ApJ, 458, 132
Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, ApJ, 533, 682
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Chester, C. & Roberts, M. S. 1964, AJ, 69, 635
Daddi, E., et al. 2003, ApJ, 588, 50
Dreiling, L. A. & Bell, R. A. 1980, ApJ, 241, 736
Efstathiou, G., et al. 2002, MNRAS, 330, L29
Erb, D. K., Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., Hunt, M. P., Moorwood, A. F. M., & Cuby, J. 2003, ApJ, 591, 101
Franx, M. et al., *The Messenger* **99**, pp. 20–22, 2000
Franx, M. et al. 2003, ApJ, 587, L79
Forster Schreiber et al. 2004,ApJ, accepted
Forster Schreiber et al. 2004, in preparation
Freedman, W. L., et al. 2001, ApJ, 553, 47
Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, ApJ, 594, 279
Griersmith, D. 1980, AJ, 85, 1295
Hogg, D. W., et al. 2003, ApJ, 585, L5
Jansen, R. A., Fabricant, D., Franx, M., & Caldwell, N. 2000, ApJS, 126, 331
Jansen, R. A., Franx, M., Fabricant, D., & Caldwell, N. 2000, ApJS, 126, 271
Jansen, R. A., Franx, M., & Fabricant, D. 2001, ApJ, 551, 825
Katz, N. & Gunn, J. E. 1991, ApJ, 377, 365
Kauffmann, G. & White, S. D. M. 1993, MNRAS, 261, 921
Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201
Kauffmann, G., Charlot, S., & White, S. D. M. 1996, MNRAS, 283, L117
Kauffmann, G., et al. 2003, MNRAS, 341, 33
Kennicutt, R. C. 1998, ApJ, 498, 541
Kodama, T. & Arimoto, N. 1997, A&A, 320, 41
Labbé, I., et al. 2003, AJ, 125, 1107
Labbé et al 2004, in preparation
Le Borgne, J.-F., et al. 2003, A&A, 402, 433
Moorwood, A. F. 1997, Proc. SPIE, 2871, 1146
Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620
Papovich, C., et al. 2004, ApJ, 600, L111
Peletier, R. F. & de Grijs, R. 1998, MNRAS, 300, L3
Pettini, M., Shapley, A. E., Steidel, C. C., Cuby, J., Dickinson, M., Moorwood, A. F. M., Adelberger, K. L., & Giavalisco, M. 2001, ApJ, 554, 981
Press et al. 1992, *Numerical Recipes in C*
Rix, H., et al. 2004, ApJS, 152, 163
Rudnick, G. et al. 2001, AJ, 122, 2205
Rudnick, G., et al. 2003, ApJ, 599, 847
Schweizer, F. & Seitzer, P. 1992, AJ, 104, 1039
Shapley, A. E., Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., & Pettini, M. 2001, ApJ, 562, 95
Shapley et al., 2004, ApJ, accepted
Spergel, D. N., et al. 2003, ApJS, 148, 175
Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726
Somerville, R. S. & Primack, J. R. 1999, MNRAS, 310, 1087

Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996, AJ, 112, 352
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, ApJ, 592, 72
Steinmetz, M. & Navarro, J. F. 2002, New Astronomy, 7, 155
Tully, R. B., Mould, J. R., & Aaronson, M. 1982, ApJ, 257, 527
Tully, R. B., Pierce, M. J., Huang, J., Saunders, W., Verheijen, M. A. W., & Witchalls, P. L. 1998, AJ, 115, 2264
van Dokkum, P. G., Franx, M., Kelson, D. D., Illingworth, G. D., Fisher, D., & Fabricant, D. 1998, ApJ, 500, 714
van Dokkum, P. G., Franx, M., Fabricant, D., Illingworth, G. D., & Kelson, D. D. 2000, ApJ, 541, 95
van Dokkum, P. G. et al. 2003, ApJ, 587, L83
van Dokkum, P. G., et al. 2004, ApJ, 611, 703
Visvanathan, N. 1981, A&A, 100, L20
Westera, P., Lejeune, T., Buser, R., Cuisinier, F., & Bruzual, G. 2002, A&A, 381, 524
White, S. D. M. & Frenk, C. S. 1991, ApJ, 379, 52
Zaritsky, D., Kennicutt, R. C., & Huchra, J. P. 1994, ApJ, 420, 87

Table 1 — The Blue Color Magnitude Relation

	HDFS		MS1054	
z[1]	a[2]	b[3]	a[2]	b[3]
$0.5 - 0.7$[4]	-0.176	0.226	-0.262	0.098
$0.7 - 1.4$	-0.206	-0.186	-0.144	-0.190
$1.4 - 2.2$	-0.140	-0.705	-0.234	-0.516
$2.2 - 3.2$	-0.139	-0.950	-0.324	-0.892

[1] The redshift range of the subsample
[2] The best-fit slope of the blue CMR.
[3] The intercept of the blue CMR, after fitting *with a slope fixed at* -0.17.
[4] The lower redshift bound was extended to z=0.35 for the HDFS, as the F300W filter is bluer than the classical U-band.

Table 2 — The Scatter around the Blue CM

	HDFS		MS1054	
z[1]	σ^2_{obs}	σ^3_{true}	σ^2_{obs}	σ^3_{true}
$0.5 - 0.7$[4]	0.53	0.46	0.79	0.73
$0.7 - 1.4$	0.66	0.61	0.72	0.65
$1.4 - 2.2$	0.37	0.27	0.73	0.66
$2.2 - 3.2$	0.57	0.51	0.83	0.77

[1] The redshift range of the subsample.
[2] The measured central 32% of the color distribution after rejecting galaxies with $\Delta(2200 - V) + 0.17 M_V > 1.5$.
[3] The intrinsic color distribution after subtracting the observational errors.
[4] The lower redshift bound was extended to z=0.35 for the HDFS, as the F300W filter is bluer than the classical U-band.