

# Galaxy Zoo: Investigating the Star Formation History of the Green Valley

R. J. Smethurst,<sup>1</sup> C. J. Lintott,<sup>1,2</sup> B. D. Simmons,<sup>1</sup>

<sup>1</sup> Oxford Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK

<sup>2</sup> Adler Planetarium, 1300 S Lake Shore Drive, Chicago, IL, 60605, USA

9 April 2014

## ABSTRACT

Does galactic evolution proceed through the Green Valley via multiple pathways or as a single population? Motivated by recent results which used a toy model to highlight radically different evolutionary pathways between early- and late-type galaxies, we present results from an advanced Bayesian approach to this problem wherein we model the star formation history of a galaxy and compare the predicted and observed optical and near-ultraviolet colours in the model parameter space. We investigate the most probable values for these parameters for both disc-like and elliptical-like populations of galaxies, incorporating the morphological vote fractions from Galaxy Zoo<sup>\*</sup> into our analysis. We will discuss the implications of our results on the understanding of the simultaneous morphological-colour evolution of galaxies, particularly of those residing in the Green Valley.

## 1 INTRODUCTION

The Green Valley galaxies occupy a low density area of the colour-magnitude diagram between the *red sequence* and *blue cloud*. How they came to be there and why there are relatively so few of them has been a source of debate for many years (*cite lots of people here*).

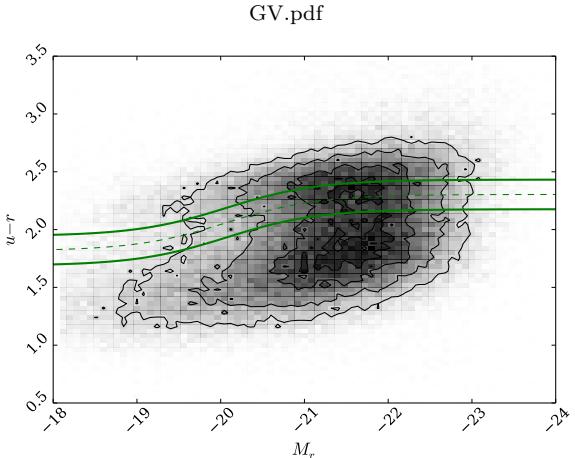
A previous Galaxy Zoo investigation by Schawinski et al. (2014), found two contrastingly different evolutionary pathways between morphological types across the Green Valley using a toy model; this study will expand on this by providing a full Bayesian analysis of the problem first outlined in Schawinski et al. (2014).

## 2 DATA

See Galaxy Zoo 2 data release Willett et al. (2013).

### 2.1 Defining the Green Valley

We use the definition from Baldry et al. (2004) to define which of the GZ2 galaxies are in the Green Valley, which is shown in figure 1 by the dashed line. Anything within  $\pm 1\sigma$  of this relationship was considered a Green Valley galaxy, shown by the solid lines in figure 1. We decided upon a conservative Green Valley in order to remove any influence from tail galaxies in either the red sequence or blue cloud.



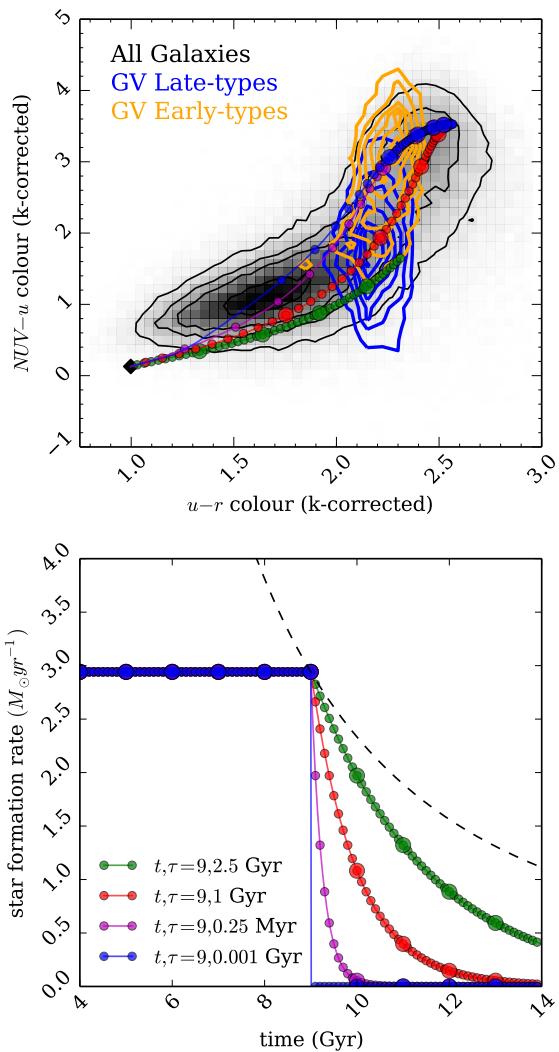
**Figure 1.** Colour-magnitude diagram for the GZ2 population showing the definition between the Blue Cloud and the Red Sequence from Baldry et al. (2004) with the dashed line. The solid lines show  $\pm 1\sigma$  either side of this definition; any galaxy residing in this area of the colour-magnitude diagram is considered a Green Valley galaxy.

## 3 A BAYESIAN ANALYSIS

### 3.1 Stellar Population Synthesis Models

Using an assumed IMF, SPS models calculate the total flux given off by a population, as described in Bruzual & Charlot (2003)

\* This investigation has been made possible by the participation of more than 250,000 volunteers in the Galaxy Zoo project. Their contributions are individually acknowledged at <http://www.galaxyzoo.org/volunteers.aspx>



**Figure 2.** Exponentially declining star formation history models with  $t_{quench} = 9 \text{ Gyr}$  and various  $\tau$  values are shown in the bottom panel, with the dashed line showing the average SFR with lookback time from Peng et al. (2010) which is used to set the  $c_{SFR}$  at  $t_{quench}$  (see section 3.3). The top panel shows the evolution with time of the models across the optical-NUV colour-colour diagram, in comparison to the distribution of the GZ2 galaxies across this colour space.

### 3.2 Modelling the Star Formation History

The star formation history of a population is modelled with an exponentially declining star formation rate (SFR) at time steps from  $t = 0 \text{ Gyr}$  to  $t = 13.7 \text{ Gyr}$  as:

$$SFR = \begin{cases} c_{SFR} & \text{if } t < t_q \\ c_{SFR} \times \exp\left(\frac{-(t-t_q)}{\tau}\right) & \text{if } t > t_q \end{cases}$$

where  $t_q$  is the onset time of quenching,  $\tau$  is the timescale over which the quenching occurs and  $c_{SFR}$  is a constant star formation rate (see section 3.3). A smaller  $\tau$  value corresponds to a rapid quench (almost instantaneous suppression of star formation), whereas a large value of  $\tau$  corresponds to a slower quench (slower than the dynamical

timescale of a galaxy). This model assumes that all of the populations form an initial burst of stars at  $t = 0$ .

Once this evolutionary SFR is obtained, we convolve them with the BC03 population synthesis models to generate a model SED at each time step. We suppress the fluxes from stars younger than 3 *Myr* to mimic the effect of birth clouds, then apply filter transmission curves to obtain AB magnitudes and therefore colours. How these colours evolve with each time step in the optical-NUV colour space can be seen in the top panel of Figure 2.

We can ‘observe’ these model stellar populations at a given time in their history; this correlates to the lookback time at which we observe them and therefore, if they were real populations, a measurable redshift (assuming all galaxies are formed at a time  $t = 0$ ). Therefore, the lookback time was calculated from the redshift (using the *cosmology* package provided in the Python module *astroPy*) of each galaxy in the GZ2 sample, in order to compare the observed colours to the predicted colours of the model at that time. This lookback time  $t^{lb}$  can be thought of as a galaxy’s age when we assume that all galaxies formed with an initial burst of star formation at  $t = 0$ .

Figure 3 shows the predicted optical and NUV colours at a time of  $t = 12.8 \text{ Gyr}$  (the average look back time of the Galaxy Zoo 2 sample) provided by the exponential SFH model. These predicted colours will be referred to as  $d_p(t_q, \tau, t^{lb})$ . The SFR at a time of  $t = 12.8 \text{ Gyr}$  is also shown in Figure 3 to compare how this correlates with the predicted colours. The  $u - r$  predicted colour shows an immediate correlation with the SFR, however the  $NUV - u$  colour is more sensitive to  $\tau$  and so is a good colour to investigate any recent star formation. At small  $\tau$  (rapid quenching timescales) the  $NUV - u$  colour is insensitive to  $t_q$ , whereas at large  $\tau$  (slow quenching timescales) the colour is very sensitive to  $t_q$ . This therefore makes it an ideal colour for tracing both  $t_q$  and  $\tau$  in a population.

### 3.3 Constraining the Models

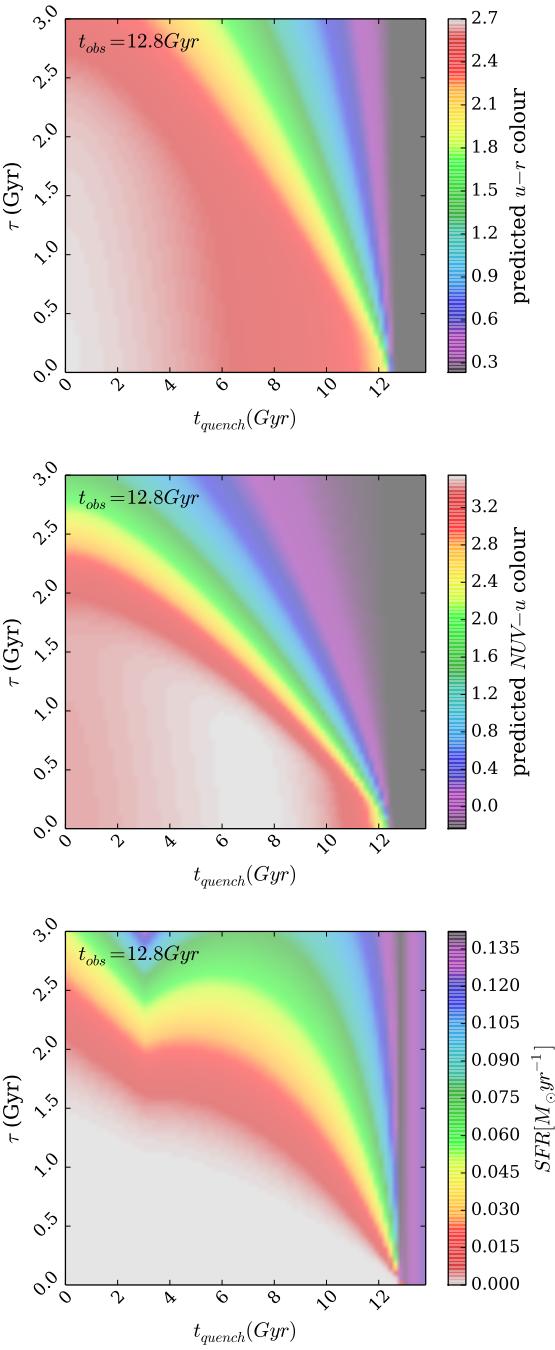
The relation defined in equation 1 and shown in Figure 3 of Peng et al. (2010) was used to constrain the initial constant SFR; at the time  $t_{quench}$ , the models have a SFR which lies on this relationship (see right hand panel of figure 4). The left hand plot in figure 4 show how these models reproduce the observed relationship between SFR and mass of a galaxy, including how at the time of quenching they reside on the “main sequence” of star formation, shown by the solid black line.

### 3.4 Bayesian Analysis

To conduct a fully Bayesian analysis of the problem outlined above, we must consider all possible combinations of  $(t_q, \tau) = \theta$  which will be distributed with a mean and standard deviation, so that:

$$w = (\mu_\theta, \sigma_\theta) = (\mu_{t_q}, \sigma_{t_q}, \mu_\tau, \sigma_\tau)$$

We can then define the Bayesian probability  $P(\theta|w) = P(t_q, \tau|w) = P(t_q|w)P(\tau|w)$  assuming that  $P(t_q|w)$  and



**Figure 3.** Model predictions of the  $u - r$  and  $NUV - u$  colours in the left and middle panels respectively when the population is observed at  $12.8 \text{ Gyr}$  (the average look back time of the GZ2 sample, calculated from the observed redshift). In the right hand panel, the predicted SFR by a time of  $t \sim 12.8 \text{ Gyr}$  from the models described in sections 3.2 and 3.3.

$P(\tau|w)$  are independent of each other:

$$P(t_q, \tau|w) = \frac{1}{\sqrt{4\pi^2 \sigma_{t_q}^2 \sigma_\tau^2}} \exp\left(-\frac{(t_q - \mu_{t_q})^2}{2\sigma_{t_q}^2}\right) \exp\left(-\frac{(\tau - \mu_\tau)^2}{2\sigma_\tau^2}\right).$$

Which is equivalent to:

$$P(\theta|w) = \frac{1}{Z_\theta} \exp\left(-\frac{\chi_\theta^2}{2}\right).$$

Therefore if we work in logarithmic probabilities:

$$\log[P(\theta|w)] = -\log(Z_\theta) - \frac{\chi_\theta^2}{2}.$$

We must then find the probability of the data given these values of theta,  $P(\underline{d}|\theta, t_k^{lb})$ :

$$P(\underline{d}|\theta, t_k^{lb}) = \prod_k P(d_k|\theta, t_k^{lb}),$$

where  $d_k$  is a single data point (optical and NUV colours of one galaxy). We calculate  $P(d_k|\theta, t_k^{lb})$  using the predicted values for the optical ( $c = opt$ ) and NUV ( $c = NUV$ ) colours,  $d_{c,p}(\theta, t_k^{lb})$ , for a given combination of  $\theta = (t_q, \tau)$  and a calculated galaxy age  $t^{lb}$  (look back time, calculated from a galaxy's redshift, equivalent to the age of the galaxy assuming that all galaxies formed at  $t = 0 \text{ Gyr}$ ):

$$P(d_k|\theta, t^{lb}) = \frac{1}{\sqrt{2\pi\sigma_{opt,k}^2}} \frac{1}{\sqrt{2\pi\sigma_{NUV,k}^2}} \exp\left(-\frac{(d_{opt,k} - d_{opt,p}(\theta, t_k^{lb}))^2}{\sigma_{opt,k}^2}\right) \exp\left(-\frac{(d_{NUV,k} - d_{NUV,p}(\theta, t_k^{lb}))^2}{\sigma_{NUV,k}^2}\right),$$

where for one combination of  $\theta = (t_q, \tau)$ ,

$$\chi_{c,k}^2 = \frac{(d_{c,k} - d_{c,p}(\theta, t_k^{lb}))^2}{\sigma_{c,k}^2}$$

and

$$Z_k = \sqrt{2\pi\sigma_{c,k}^2}.$$

Again working in logarithmic probabilities:

$$\log(P(\underline{d}|\theta, t^{lb})) = \sum_{c,k} \log(P(d_{c,k}|\theta, t_k^{lb}))$$

$$\log(P(\underline{d}|\theta, t^{lb})) = K - \sum_{c,k} \frac{\chi_{c,k}^2}{2},$$

where  $K$  is a constant:

$$K = - \sum_{c,k} \log Z_{c,k}.$$

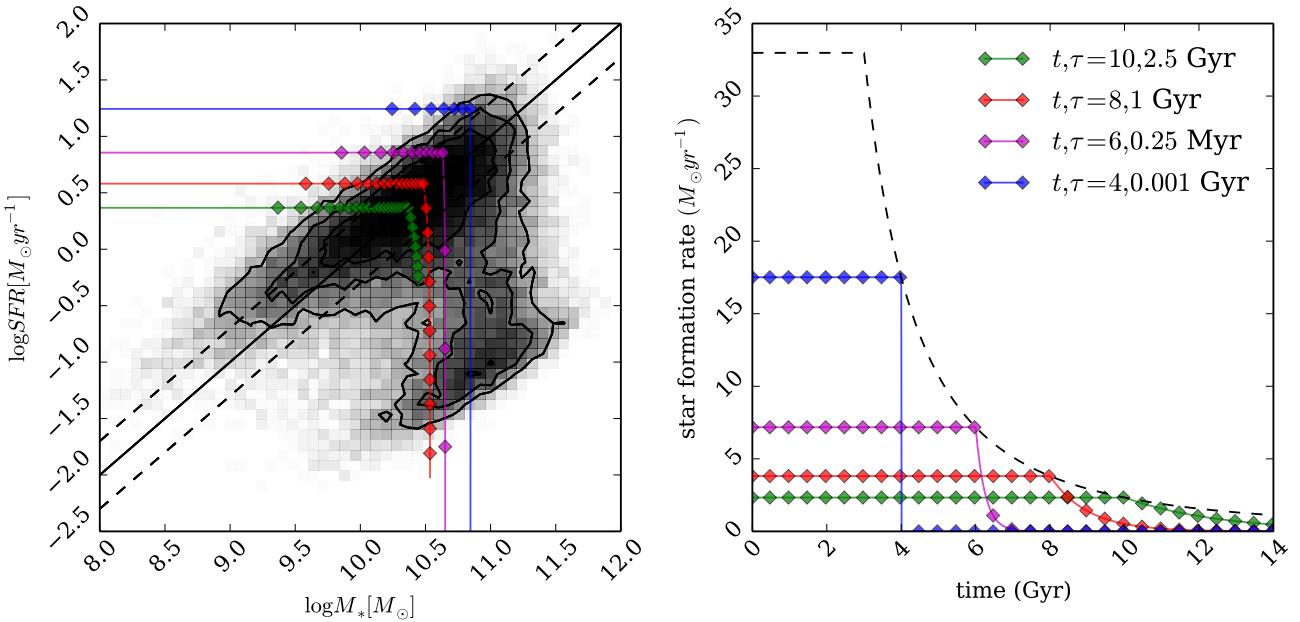
What we need however is the probability of each combination of  $\theta$  given the GZ2 data,  $P(\theta|\underline{d})$ , which we can find by:

$$P(\theta|\underline{d}) = \frac{P(\underline{d}|\theta, t^{lb})P(\theta)}{\int P(\underline{d}|\theta, t^{lb})P(\theta)d\theta},$$

where,

$$P(\underline{d}|\theta, t^{lb})P(\theta) = \exp\left[\log[P(\underline{d}|\theta, t^{lb})] + \log[P(\theta)]\right],$$

and the denominator  $\int P(\underline{d}|\theta, t^{lb})P(\theta)d\theta$  is given by the sum across all the elements of  $P(\underline{d}|\theta, t^{lb})P(\theta)$ . This denominator is a mere normalisation factor, therefore when comparing the likelihoods between two different combinations of  $\theta = (t_q, \tau)$



**Figure 4.** The right hand panel shows the relationship between the star formation rate and stellar mass for the galaxies in the GZ2 sample (black contours), which trace the "main sequence" of star formation (solid line from Peng et al. (2010) with  $\pm 0.3$  dex scatter shown by the dotted lines). The SFRs and masses of the model galaxies are also shown with corresponding star formation histories shown in the left hand panel. The dashed line in the left hand panel shows the relationship between the SFR and time as defined in equation 1 of Peng et al. (2010).

we need only compare the numerator and can also remain in logarithmic probability space. So given the data from the GZ2 sample, we can calculate  $\log[P(\theta|d, t^{lb})]$  for all possible  $\theta$  values and compare these to determine the most likely values for  $\theta$  given the GZ2 data. In order to do this robustly, we performed a Markov Chain Monte Carlo (MCMC) sampling method to cycle through the defined parameter space using a Python implementation of an affine invariant ensemble sampler Foreman-Mackey et al. (2013); *emcee*.

In addition to the colours, the GZ2 data is unique in that it provides information on a galaxy's morphology. Vote fractions from GZ2 users are available for each galaxy, for example if 80 of 100 people thought a galaxy was disc shaped, whereas 20 out of 100 people thought the same galaxy was smooth in shape (i.e. elliptical), that galaxy would have  $p_s = 0.2$  and  $p_d = 0.8$ . In this example this galaxy would be included in the '*clean*' disc sample according to Willett et al. (2013), however we lose the information that a certain majority of people classified this galaxy as *smooth* (i.e. elliptical). If we just run the above sampling method on galaxies in either *clean* sample, we lose any information about any intermediate galaxies and how these contribute to the likelihood of  $P(\theta|d, t^{lb})$ . It is the intermediate galaxies which are thought to be a key to the morphological changes between disc and elliptical galaxies; if this change is associated with the Green Valley in any way then these vote fractions are information we wish to keep in our sampling.

We can incorporate these GZ2 vote fractions into our sampling by considering them as fractions which that galaxy contributes to the likelihood  $P(d_k|\theta)$ . For example a galaxy which has  $p_s = 0.9$  should carry more weight in the overall likelihood than a galaxy with  $p_s = 0.1$ . Therefore the

likelihood can now be thought of as:

$$P(\underline{d}|\theta, \underline{t}^{lb}) = \prod_k p_k P(d_k|\theta, t_k^{lb}),$$

where  $p_k$  is either  $p_s$  or  $p_d$  for an individual galaxy. We can then feed the code with the colours for all of the GZ2 data along with first the  $p_s$  vote fractions to find the most likely parameters for  $\theta$  for elliptical galaxies and then with the  $p_d$  vote fractions to find the most likely parameters for  $\theta$  for disc galaxies. However, this is costly in computing time, therefore we perform our sampling across four parameters so that  $\theta = (t_s, \tau_s, t_d, \tau_d)$  and our likelihood function is then:

$$P(\underline{d}|\theta, \underline{t}^{lb}) = \prod_k \left[ p_{s,k} P(d_k|\theta_s, t_k^{lb}) + p_{d,k} P(d_k|\theta_d, t_k^{lb}) \right],$$

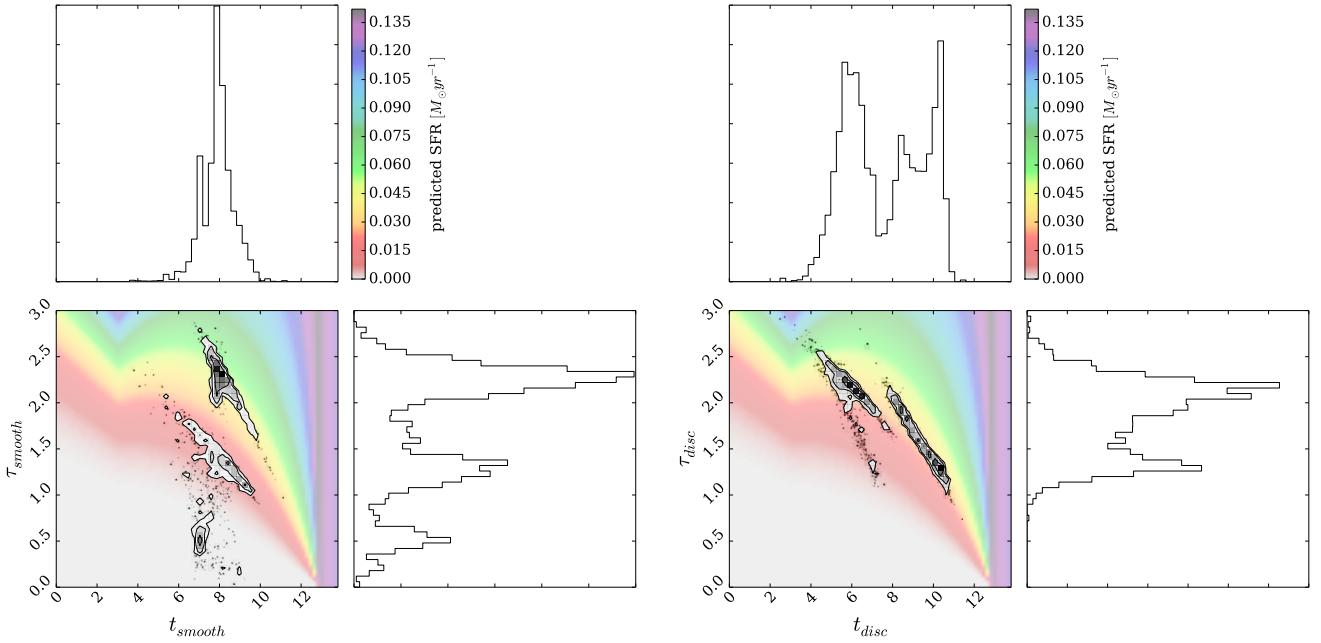
or,

$$\log \left[ P(\underline{d}|\theta, \underline{t}^{lb}) \right] = \sum_k \log \left[ p_{s,k} P(d_k|\theta_s, t_k^{lb}) + p_{d,k} P(d_k|\theta_d, t_k^{lb}) \right].$$

The code searches through the  $\theta$  parameter space to find the region that then maximises  $P(\theta|\underline{d})$  to return four parameter values for  $t_s, \tau_s, t_d$  and  $\tau_d$ . If we find that  $t_s \sim t_d$  and  $\tau_s \sim \tau_d$  for the Green Valley galaxies then we can conclude that this area of the colour-magnitude diagram consists of single population.

## 4 RESULTS

The contour plots that follow show the regions of high likelihood for the various samples of galaxies of the SFH parameters outlined in section 3.2, produced using an MCMC sampling method. It is expected that those subsets of galaxies



**Figure 5.** For all the galaxies in the Galaxy Zoo 2 sample, contour and histogram plots show the regions of greatest likelihood for an exponential model star formation history parameters [ $t_{quench}$  and  $\tau_{quench}$ ] for both smooth-like(left) and disc-like (right) galaxies.  $t_q$  is the time at which quenching occurs (Gyr) and  $\tau_q$  is the time scale on which quenching occurs (Gyr; the larger the  $\tau_q$ , the slower the quenching). Background colours show the star formation rate predicted by this model after a time  $t \sim 12.8$  Gyr, which is the mean look back time of the galaxies in the GZ2 sample. Galaxies contribute to  $[t_q, \tau_q]_{smooth}$  and  $[t_q, \tau_q]_{disc}$  according to their Galaxy Zoo 2 vote fraction (i.e. a galaxy with  $p_{disc} \sim p_{smooth} \sim 0.5$  will contribute equally to each set of parameters).

with large numbers will rapidly converge the ‘walkers’ into the regions of high likelihood without the need to explore the whole parameter space. Conversely the smaller galaxy subsets will allow a full exploration of the parameter space. It is also expected that those subsets of galaxies with narrow ranges of either NUV or optical colours will also cause the ‘walkers’ to converge rapidly into the regions of high likelihood without the need to explore the whole of the parameter space.

This raises the issue of whether the sampling method is capable of finding multiple likelihood peaks, however the figures below show that this is not the case.

## 5 DISCUSSION

In figure 5 we can immediately see that the two populations occupy very different locations in the  $[t, \tau]$  parameter space supporting the conclusion in Schawinski et al. (2014) that early- and late- type galaxies have different quenching timescales. We can also see a significant *correlation* between the two parameters  $[t, \tau]$  (if quenching occurs earlier, then the quenching timescale increases) particularly for the disc-like galaxies.

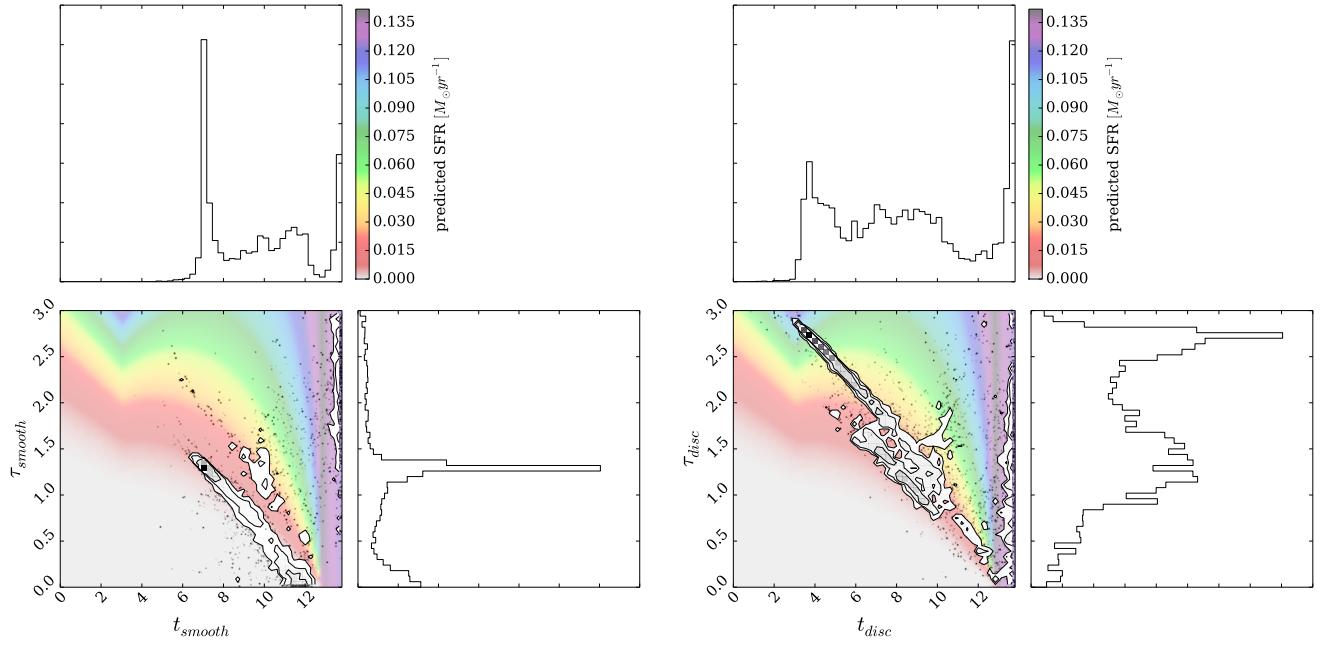
For the disc-like galaxies we have a bimodal likelihood in the SFH parameter space, whereas for the smooth-like galaxies we have a trimodal likelihood. For the disc galaxies we see no likelihood below  $\tau \sim 1$  Gyr (rapid quenching), whereas we do see an area of likelihood for the smooth galaxies - most likely caused by typical massive, red early type galaxies. The likelihood for  $t_{quench}$  for the disc-like galaxies

spans a much larger range than for the smooth-like galaxies and also has a correlation with  $\tau$ ; when quenching occurs at an earlier time, the quenching timescale is much longer, whereas when quenching occurs at a later time, the quenching timescale is quicker. Perhaps this is due to a dependency on the environment which galaxies occupied at earlier times compared to later times.

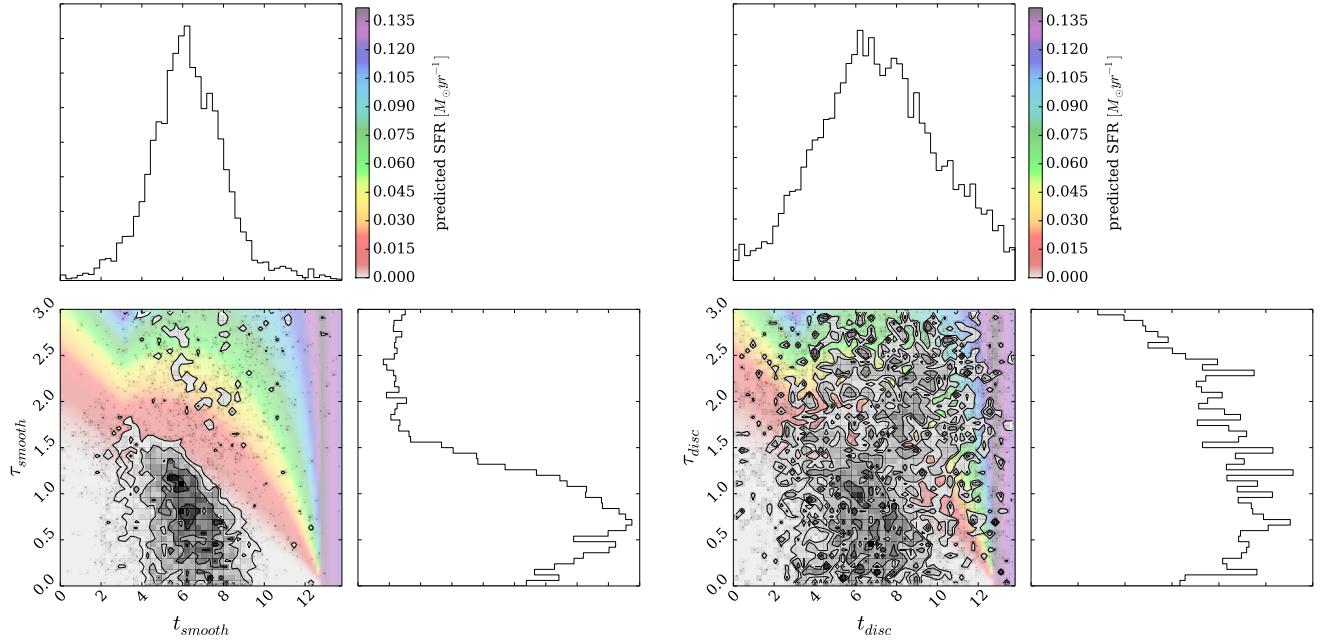
In figure 6 we can see, defined to be in the optical Red Sequence, a prevalence for fast quenching timescales for the smooth-like galaxies and slow quenching timescales for the disc-like galaxies ; again confirming the results found by Schawinski et al. (2014). However there is also a significant likelihood for blue NUV colours in the smooth parameters (slow quenching at late timescales) and for red NUV colours in the disc parameters. Could this be due to:

- (i) a population of S0 galaxies (with GZ2 likelihoods  $p_s \sim p_d \sim 0.5$ ) which contribute to both sets of parameters?
- (ii) bulge dominated disc galaxies impacting on the likely disc parameters?
- (iii) NUV blue smooth galaxies impacting on the likely smooth parameters?

We can compare figure 6 with figure 7 which shows the SFH parameters but only for those galaxies defined to be both in the optical Red Sequence and the GZ2 ‘clean’ sample (i.e.  $p_d \geq 0.8$  and  $p_s \geq 0.8$ ). This enables us to disentangle which galaxies are contributing to which areas of high likelihood in figure 6. We can therefore see that the typical clean, smooth, red sequence galaxy has undergone a SFH with a relatively rapid quench at various times, resulting in a very low current SFR. These are therefore the typical ‘red and



**Figure 6.** Same as for Figure 5 but for galaxies defined as optical Red Sequence Baldry et al. (2004).

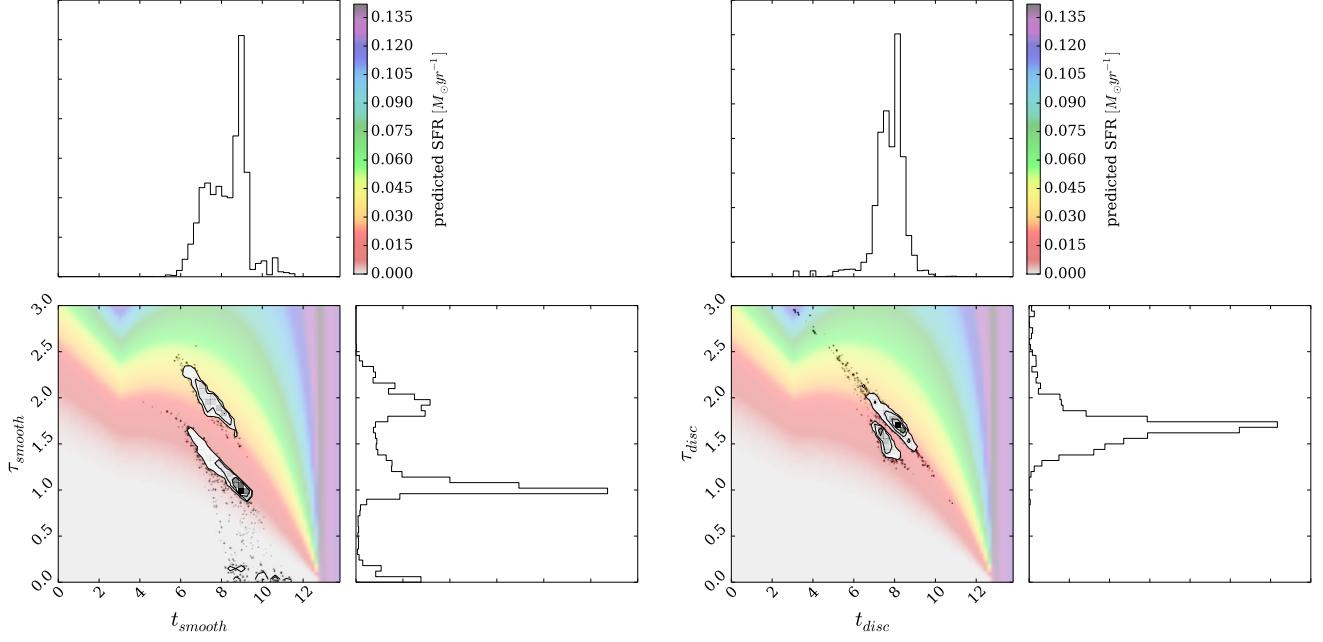


**Figure 7.** Same as for Figure 5 but for galaxies from the clean sample and defined as optical Red Sequence Baldry et al. (2004).

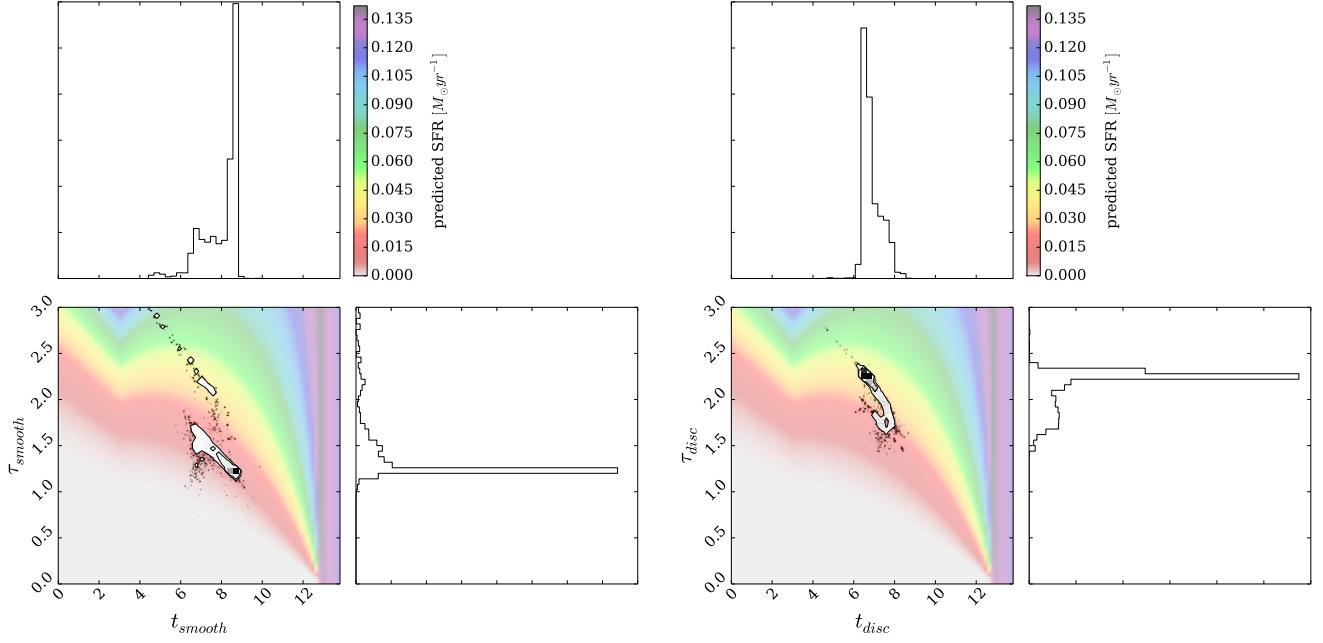
'dead' galaxies expected to be found in the red sequence. As for the clean disc galaxies in the red sequence, presumably red spiral galaxies, there is no region of the parameter space which has a preferred likelihood, suggesting that although these disc galaxies are optically red, they have a wide range of NUV-u colours, giving rise to a wide range of timescales for the most recent episode star formation episode. This suggests that there is no 'typical' red sequence disc galaxy and

that there are many possible routes for a disc galaxy to the red sequence.

In figures 8 and 9 we can make similar comparisons as above but for the Green Valley galaxies. In figure 8 we still see the correlation between the two parameters as for all of the galaxies but the areas of high likelihood for the parameters has changed. We have a very clear bimodal likelihood in the parameter space for smooth galaxies; this is also slightly apparent in the likelihoods for the parameters for the disc-



**Figure 8.** Same as for Figure 5 but for galaxies defined as optical Green Valley Baldry et al. (2004).

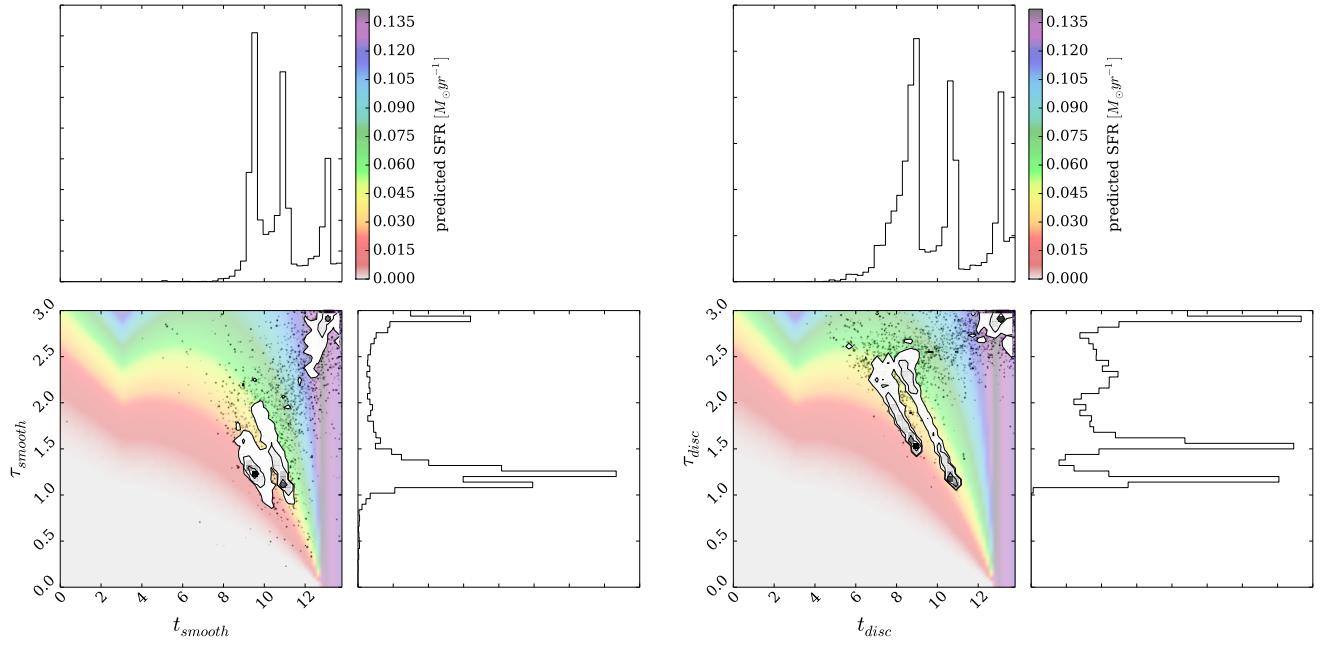


**Figure 9.** Same as for Figure 5 but for galaxies from the clean sample and defined as optical Green Valley Baldry et al. (2004).

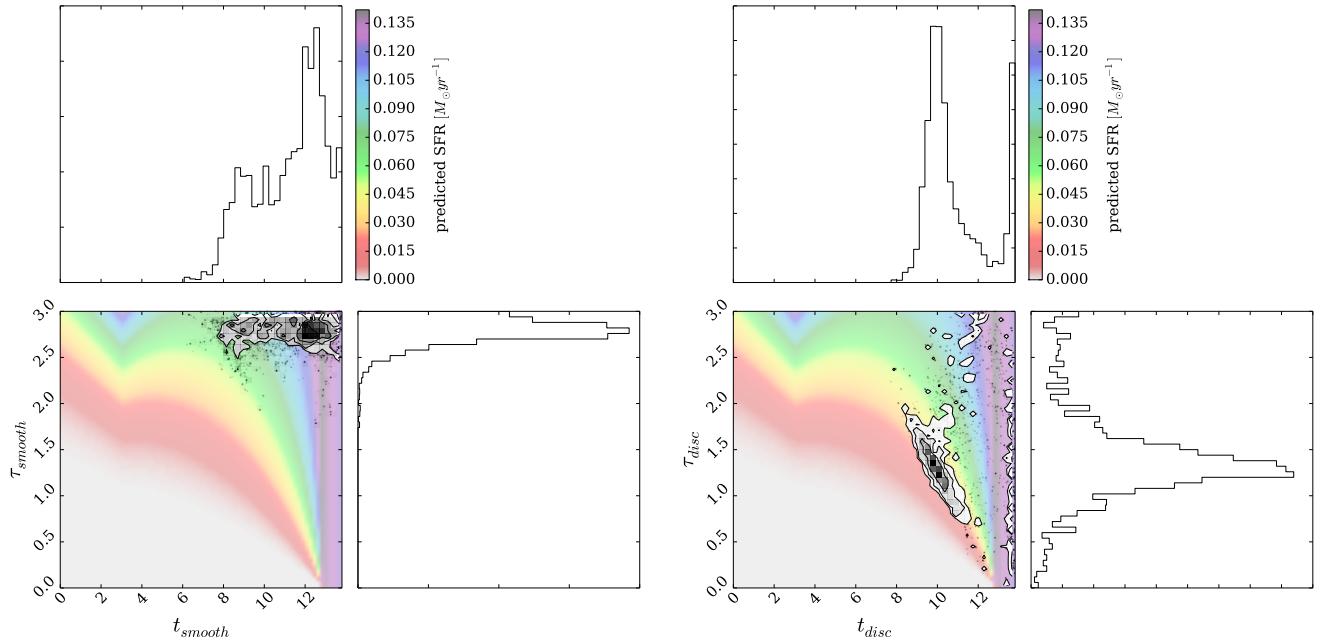
like green valley galaxies, but is not as pronounced. If we compare this to figure 9 we can see that this bimodality has disappeared for both smooth- an disc- like populations. This suggests that there are not just two routes for galaxies through the green valley but *three*: with the smooth, intermediate (e.g. S0 galaxies) and disc galaxies each having different SFHs which cause them to traverse across the Green Valley.

We can also see that quenching for the Green Valley occurs

on longer quenching timescales; there is little to no likelihood for  $\tau$  below  $\sim 1.0$  Gyr for the disc galaxies, however a small likelihood for recent rapid quenching of the smooth galaxies. This may be counter intuitive at first, as one of the main arguments for the lack of galaxies in the green valley is due to the hypothesised rapid movement across it; however, if this is the case then the expected number of galaxies in this region with this type of SFH (rapid quenching) is very small. It is remarkable therefore that we have managed to



**Figure 10.** Same as for Figure 5 but for galaxies defined as optical Blue Cloud Baldry et al. (2004).

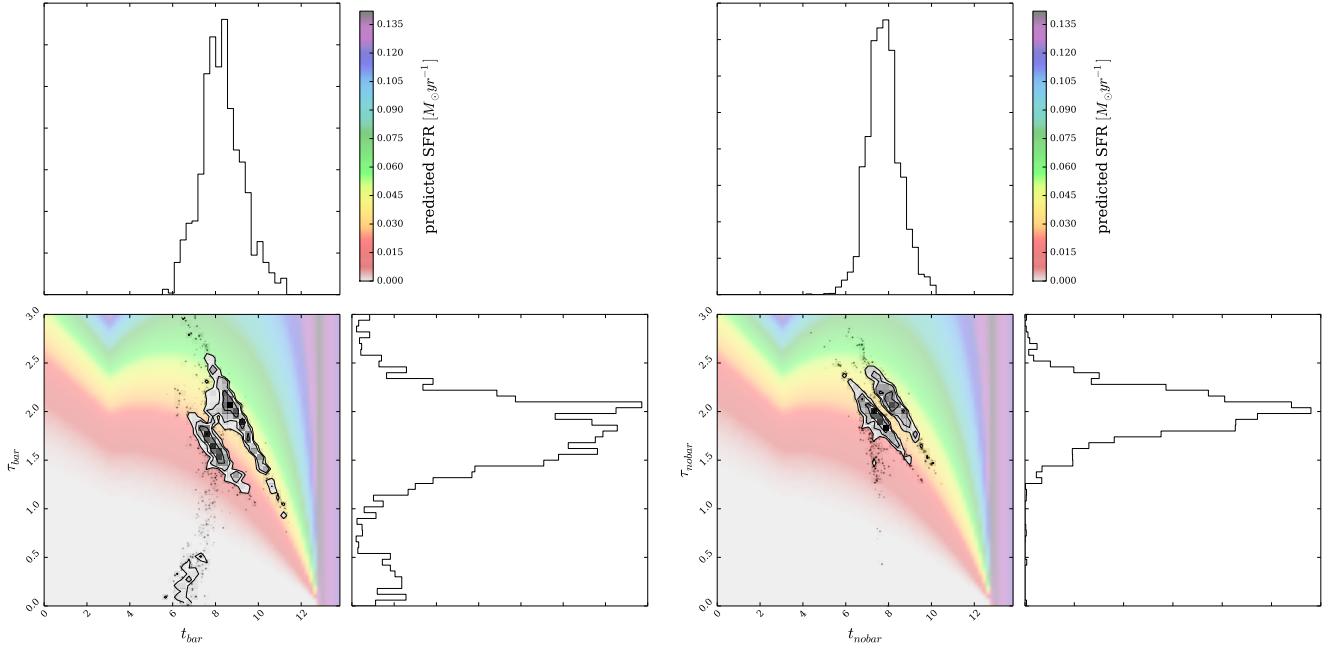


**Figure 11.** Same as for Figure 5 but for galaxies from the clean sample and defined as optical Blue Cloud Baldry et al. (2004).

find some likelihood for this type of SFH in the parameter space (as seen in the left hand panel of figure 8), however only for the smooth-like galaxies, suggesting that it is only early-type galaxies which traverse the green valley in this manner. The disc galaxies come to reside in the green valley after a relatively slow quench causing a transition from the centre to the edge of the blue cloud and eventually into the green valley.

In figure 10 there is a preference for slow quenching

at late times for both smooth- and disc-like galaxies but in this case we see a bimodality (or trimodality) in both populations, this time more strongly in the disc galaxies, compared to figure 5. The disc-like galaxies show a much stronger preference for slow quenching at late times than the smooth-like parameters, typically galaxies in the blue cloud will generally be galaxies whose only SFH to date has been some varying rate of star formation since it's formation with no quenching to date. There is also high likelihood in



**Figure 12.** Same as for Figure 5 but for barred and non barred galaxies.

both samples for middle rate quenching at relatively recent times which could account for those galaxies on the edge of the blue cloud which have begun to quench. The SFH parameter space for the smooth-like and disc-like galaxies are very similar, suggesting that galaxies in the blue cloud are all very similar, regardless of morphology. *I'm still not sure how this compares to the clean sample of blue cloud galaxies in figure 11, leave it with me...*

In figure 12 we can compare the star formation history parameters for barred and unbarred galaxies, selected so that  $N_{count,bar} > 10$  and  $N_{count,nobar} > 10$ . The analysis was then run with  $p_{bar}$  and  $p_{nobar}$  in place of  $p_{smooth}$  and  $p_{disc}$ . We can see that two the populations occupy different areas in the SFH parameter space suggesting that bars have undergone a different formation history to non barred disc galaxies. Barred galaxies simultaneously are more likely to have bluer and redder colours than non barred galaxies, as the bimodality seen in the SFH parameters of the non barred galaxies in the right hand panel of figure 12 has moved outwards. This suggests that the argument for whether bars promote or quench star formation rates in galaxies could have evidence for both possibilities. Perhaps the amount and distribution of gas present in a galaxy prior to the formation of the bar determines what effect the bar will have on the star formation rate. There is also a region of likelihood at very rapid quenching timescales for barred galaxies which is not present in the parameters for the non barred galaxies, supporting the hypothesis that bars turn galaxies red (*cite Maraston papers here*). To explain what we see here we can say either that:

- (i) As a bar forms it can cause a rapid quenching of star formation in it's host galaxy,
- (ii) A rapid quenching of star formation in a galaxy can cause a bar to form.

This analysis however, does not allow for this to be determined.

## 6 CONCLUSION

The three most important points:

- (i) There is a clear correlation between  $t_{quench}$  and  $\tau$ . At earlier times, the quenching timescale is longer, whereas at more recent times, the quenching timescale is shorter on average. Could this be an environmental dependance of quenching with cosmic time?
- (ii) There are a wide range of possible routes to the red sequence for disc galaxies. Some have similar SFHs to the red sequence smooth galaxies, whereas some have similar SFHs to blue cloud disc galaxies. Morphology seems to be irrelevant to the SFH.
- (iii) There are three possible routes through the green valley dependant on morphology.

There is not one specific route to each part of the colour-colour or colour-magnitude diagram, due to the complex interplay between the SFH parameters, however the morphology of a galaxy can have varying impacts and constraints on what these parameters can be.

## References

- Baldry, I. K. et al., 2004, ApJ, 600, 681
- Bruzual, G. & Charlot, S., 2003, MNRAS, 344, 1000
- Foreman-Mackey, D., Hogg, D. W., Lang, D., Goodman, J., 2013, PASP, 125, 306
- Peng, Y. et al., 2010, ApJ, 721, 193
- Schawinski, K. et al., 2014 (arXiv: 1402.4814)
- Willett, K. et al., 2014, MNRAS, 435, 2835