

# Chapter 2

## STARPY: Bayesian inference of a galaxy's star formation history

*The work in the following chapter has been published in Smethurst et al. (2015).*

### 2.1 Star Formation History Models

The quenched star formation history (SFH) of a galaxy can be simply modelled as an exponentially declining star formation rate (SFR) across cosmic time ( $0 \leq t$  [Gyr]  $\leq 13.8$ ) as:

$$SFR = \begin{cases} I_{sfr}(t_q) & \text{if } t < t_q \\ I_{sfr}(t_q) \times \exp\left(\frac{-(t-t_q)}{\tau}\right) & \text{if } t > t_q \end{cases} \quad (2.1)$$

where  $t_q$  is the onset time of quenching,  $\tau$  is the timescale over which the quenching occurs and  $I_{sfr}$  is an initial constant star formation rate dependent on  $t_q$ . A smaller  $\tau$  value corresponds to a rapid quench, whereas a larger  $\tau$  value corresponds to a slower quench.

Here I assume that all galaxies formed at a time  $t = 0$  Gyr with an initial burst of star formation. The mass of this initial burst is controlled by the value of the  $I_{sfr}$  which is set as the average specific SFR (sSFR) at the time of quenching  $t_q$ . Peng et al. (2010) defined a relation (their equation 1) between the average sSFR and redshift (cosmic time,  $t$ ) by fitting to measurements of the mean sSFR of blue star

forming galaxies from SDSS, zCOSMOS and literature values at increasing redshifts (Elbaz et al., 2007; Daddi et al., 2007):

$$sSFR(m, t) = 2.5 \left( \frac{m}{10^{10} M_{\odot}} \right)^{-0.1} \left( \frac{t}{3.5 \text{ Gyr}} \right)^{-2.2} \text{Gyr}^{-1}. \quad (2.2)$$

Beyond  $z \sim 2$  the characteristic SFR flattens and is roughly constant back to  $z \sim 6$ . The cause for this change is not well understood but can be seen across similar observational data (Peng et al., 2010; González et al., 2010; Béthermin et al., 2012). Motivated by these observations, the relation defined in Peng et al. (2010) is taken up to a cosmic time of  $t = 3$  Gyr ( $z \sim 2.3$ ) and prior to this a constant average SFR is assumed (see middle panel of Figure 2.1). At the point of quenching,  $t_q$ , the SFH models are defined to have an  $I_{sfr}$  which lies on this relationship for the sSFR, for a galaxy with mass,  $m = 10^{10.27} M_{\odot}$  (the mean mass of the GZ2-GALEX sample; see left panel of Figure 2.1).

Under these assumptions the average SFR of these models will result in a lower value than the relation defined in Peng et al. (2010) at all cosmic times as each galaxy only resides on the ‘main sequence’ at the point of quenching. However galaxies cannot remain on the ‘main sequence’ from early to late times throughout their entire lifetimes given the unphysical stellar masses and SFRs this would result in at the current epoch in the local Universe (Béthermin et al., 2012; Heinis et al., 2014). If prescriptions for starbursts, mergers, AGN etc. were included in this model, the reproduction of the average SFR across cosmic time would improve; however I have chosen to first focus on the simplest possible model.

Once this evolutionary SFR is obtained, it is convolved with the Bruzual & Charlot (2003) population synthesis models to generate a model SED at each time step. The observed features of galaxy spectra can be modelled using simple stellar population techniques which sum the contributions of individual, coeval, equal-metallicity stars. The accuracy of these predictions depends on the completeness of the input stellar physics. Comprehensive knowledge is therefore required of (i) stellar evolutionary tracks and (ii) the initial mass function (IMF) to synthesise a stellar population accurately.

These stellar population synthesis (SPS) models are an extremely well explored (and often debated) area of astrophysics (Maraston, 2005; Eminian et al., 2008; Conroy et al., 2009; Falkenberg et al., 2009; Chen et al., 2010; Kriek et al., 2010; Miner et al., 2011; Melbourne et al., 2012). In this work I have chosen to utilise the Bruzual

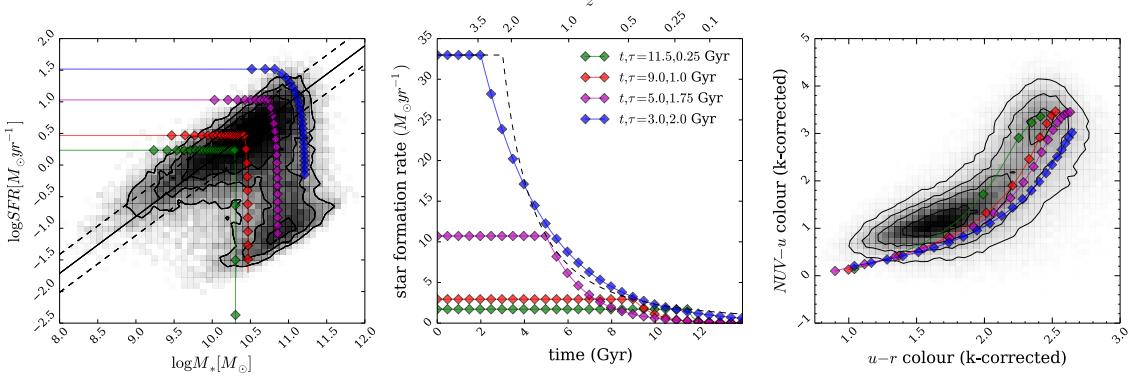


Figure 2.1: Left panel: SFR-stellar mass plane for all 126,316 galaxies in the GZ2-GALEX sample (shaded contours), with model galaxy trajectories shown by the coloured lines, with each point representing a time step of 0.5 Gyr. The ‘main sequence’ of star formation as defined by Peng et al. (2010) is shown by the solid line with  $\pm 1\sigma$  (dashed lines). Middle panel: The SFHs of the models are shown, where the SFR is initially constant before quenching at time  $t_q$  and thereafter exponentially declining with a characteristic timescale  $\tau$ . The SFR at the point of quenching is set to be consistent with the typical SFR of a star-forming galaxy at the quenching time,  $t_q$  (dashed curve; Peng et al. 2010). Right panel: The full range of models can reproduce the observed colour-colour properties of the sample; for clarity the figures show only 4 of the possible models explored in this study. Note that some of the model tracks produce colours redder than the apparent peak of the red sequence in the GZ2 subsample; however this is not the *true* peak of the red sequence due to the necessity for NUV colours from GALEX.

& Charlot (2003) *GALExEV* SPS models, along with a Chabrier IMF (Chabrier, 2003), across a large wavelength range ( $0.0091 < \lambda [\mu\text{m}] < 160$ ) with solar metallicity (m62 in the Bruzual & Charlot (2003) models; hereafter BC03), to allow a direct comparison with Schawinski et al. (2014).

Fluxes from stars younger than 3 Myr in the SPS model are suppressed to mimic the large optical depth of protostars embedded in dusty formation clouds (as in Schawinski et al. 2014). Filter transmission curves are then applied to the fluxes to obtain AB magnitudes and ultimately colours. For a particular galaxy at an observed redshift,  $z$ , I calculate the observed time,  $t^{obs}$  for that galaxy using the standard cosmological conversion between redshift and time provided in the ASTROPY *Python* module (Astropy Collaboration et al., 2013). The predicted colours of the SFH models at the observed redshift of each individual galaxy can then be compared to the observed colours directly.

Figure 2.2 shows these predicted optical and NUV colours at a time of  $t^{obs} = 12.8$  Gyr (the average observed time of the GZ2-GALEX sample,  $z \sim 0.076$ ) for the exponential SFH model. These predicted colours will be referred to as  $d_{c,p}(t_q, \tau, t^{obs})$ , where  $c=\{\text{opt,NUV}\}$  and  $p = \text{predicted}$ . The SFR at a time of  $t^{obs} = 12.8$  Gyr is also shown in Figure 2.2 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 the value of  $\tau$  and so is ideal for tracing any recent star formation in a population. 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  $t_q$ . Together the two colours are ideal for tracing the effects of  $t_q$  and  $\tau$  in a population.

This model is not a fully hydrodynamical simulation, it is a simple model built in order to test our understanding of the evolution of galaxy populations. These models are therefore not expected to accurately determine the SFH of every galaxy in the GZ2-GALEX sample, in particular galaxies which have not undergone any quenching. In this case the models described above can only attribute a constant star formation rate to these unquenched galaxies. In reality, there are many possible forms of SFH that a galaxy can take, a few of which have been investigated in previous literature; starbursts (Canalizo & Stockton, 2001), a power law (Glazebrook et al., 2003), single stellar populations (Trager et al., 2000; Sánchez-Blázquez et al., 2006; Vazdekis et al., 2010), log-normal distributions (Abramson et al., 2016) and metallicity enrichment

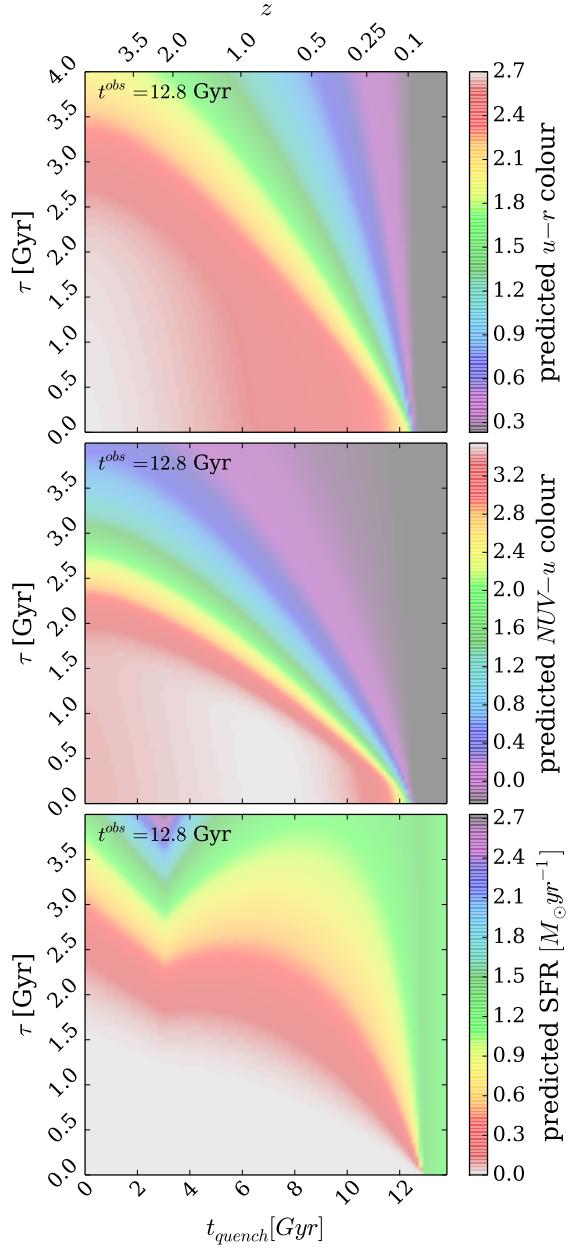


Figure 2.2: Quenching timescale  $\tau$  versus quenching onset time  $t_q$  in all three panels for the quenched SFH models used in STARPY. Colour shadings show model predictions of the  $u - r$  optical colour (top panel),  $NUV - u$  colour (middle panel), and star formation rate (lower panel), at  $t^{obs} = 12.8$  Gyr, the mean observed redshift of the GZ2 sample (see Section 2.1). The combination of optical and NUV colours is a sensitive measure of the  $\theta = [t_q, \tau]$  parameter space. Note that all models with  $t > 12.8$  Gyr are effectively un-quenched. The ‘kink’ in the bottom panel is due to the assumption that the sSFR is constant prior to  $t \sim 3$  Gyr ( $z \sim 2.2$ ).

(De Lucia et al., 2014). Incorporating these different SFHs along with prescriptions for mergers and a possible reinvigoration of star formation post quench (e.g. see recent work by Pontzen et al. 2016) into the SFH models is a possible future extension to this work once the results of this study are well enough understood to permit additional complexity to be added.

## 2.2 Probabilistic Fitting Methods

In order to achieve robust conclusions I conducted a Bayesian analysis (Sivia & Skilling, 2006; Mackay, 2003) of the predicted colours from the SFH models in comparison to the observed colours of the GZ2-GALEX sample. This approach requires consideration of all possible combinations of  $\theta \equiv (t_q, \tau)$ . Assuming that all galaxies formed at  $t = 0$  Gyr with an initial burst of star formation, we can assume that the ‘age’ of each galaxy in the GZ2 sample is equivalent to an observed time,  $t_k^{obs}$ . I then used this ‘age’ to calculate the predicted model colours at this cosmic time for a given combination of  $\theta$ :  $d_{c,p}(\theta_k, t_k^{obs})$  for both optical and NUV ( $c = opt, NUV$ ) colours. The predicted model colours can now directly be compared with the observed GZ2-GALEX sample colours, so that for a single galaxy  $k$  with optical ( $u - r$ ) colour,  $d_{opt,k}$  and NUV ( $NUV - u$ ) colour,  $d_{NUV,k}$ , the likelihood of a given model  $P(d_k|\theta_k, t_k^{obs})$  is:

$$P(d_k|\theta_k, t_k^{obs}) = \frac{1}{\sqrt{2\pi\sigma_{opt,k}^2}} \frac{1}{\sqrt{2\pi\sigma_{NUV,k}^2}} \exp \left[ -\frac{(d_{opt,k} - d_{opt,p}(\theta_k, t_k^{obs}))^2}{\sigma_{opt,k}^2} \right] \\ \exp \left[ -\frac{(d_{NUV,k} - d_{NUV,p}(\theta_k, t_k^{obs}))^2}{\sigma_{NUV,k}^2} \right]. \quad (2.3)$$

Here I have assumed that  $P(d_{opt}|\theta_k, t_k^{obs})$  and  $P(d_{NUV}|\theta_k, t_k^{obs})$  are independent of each other and that the errors on the observed colours are also independent. To obtain the probability of a combination of  $\theta$  values given the GZ2 data:  $P(\theta_k|d_k, t^{obs})$ , i.e. how likely is a single SFH model given the observed colours of a single GZ2 galaxy, I utilise Bayes’ theorem:

$$P(\theta_k|d_k, t^{obs}) = \frac{P(d_k|\theta_k, t^{obs})P(\theta_k)}{\int P(d_k|\theta_k, t^{obs})P(\theta_k)d\theta_k}. \quad (2.4)$$

I assume a flat prior on the model parameters so that:

$$P(\theta_k) = \begin{cases} 1 & \text{if } 0 \leq t_q \text{ [Gyr]} \leq 13.8 \text{ and } 0 \leq \tau \text{ [Gyr]} \leq 4 \\ 0 & \text{otherwise.} \end{cases} \quad (2.5)$$

As the denominator of Equation 2.4 is a normalisation factor, comparison between likelihoods for two different SFH models (i.e., two different combinations of  $\theta_k = [t_q, \tau]$ ) is equivalent to a comparison of the numerators. Markov Chain Monte Carlo (MCMC; Mackay 2003; Foreman-Mackey et al. 2013; Goodman & Weare 2010) provides a robust comparison of the likelihoods between  $\theta$  values; here I choose *emcee*,<sup>1</sup> a Python implementation of an affine invariant ensemble sampler by Foreman-Mackey et al. (2013).

This method allows for a more efficient exploration of the parameter space by avoiding those areas with low likelihood. A large number of ‘walkers’ are started at an initial position where the likelihood is calculated; from there they individually ‘jump’ to a new area of parameter space. If the likelihood in this new area is greater (less) than the original position then the ‘walkers’ accept (reject) this change in position. Any new position then influences the direction of the ‘jumps’ of other walkers. This is repeated for the defined number of steps after an initial ‘burn-in’ phase. *emcee* returns the positions of these ‘walkers’, which are analogous to the regions of high probability in the model parameter space.

The routine outlined above has been coded using the *Python* programming language into a package named STARPY which has been made freely available to download<sup>2</sup>. An example output from this module for a single galaxy from the GZ2-GALEX sample in the red sequence is shown in Figure 2.2.

## 2.3 Testing STARPY

In order to test that STARPY can find the correct quenching model for a given observed colour, 25 synthesised galaxies were created with known SFHs (i.e. known values of  $\theta = [t_q, \tau]$ ) from which optical and NUV colours were generated using the BC03 SPS models. These were input into STARPY to test whether the known values of  $\theta$  were

---

<sup>1</sup>[emcee13.iel.fm/emcee/](http://emcee13.iel.fm/emcee/)

<sup>2</sup>[github.com/zooniverse/starp](https://github.com/zooniverse/starp)

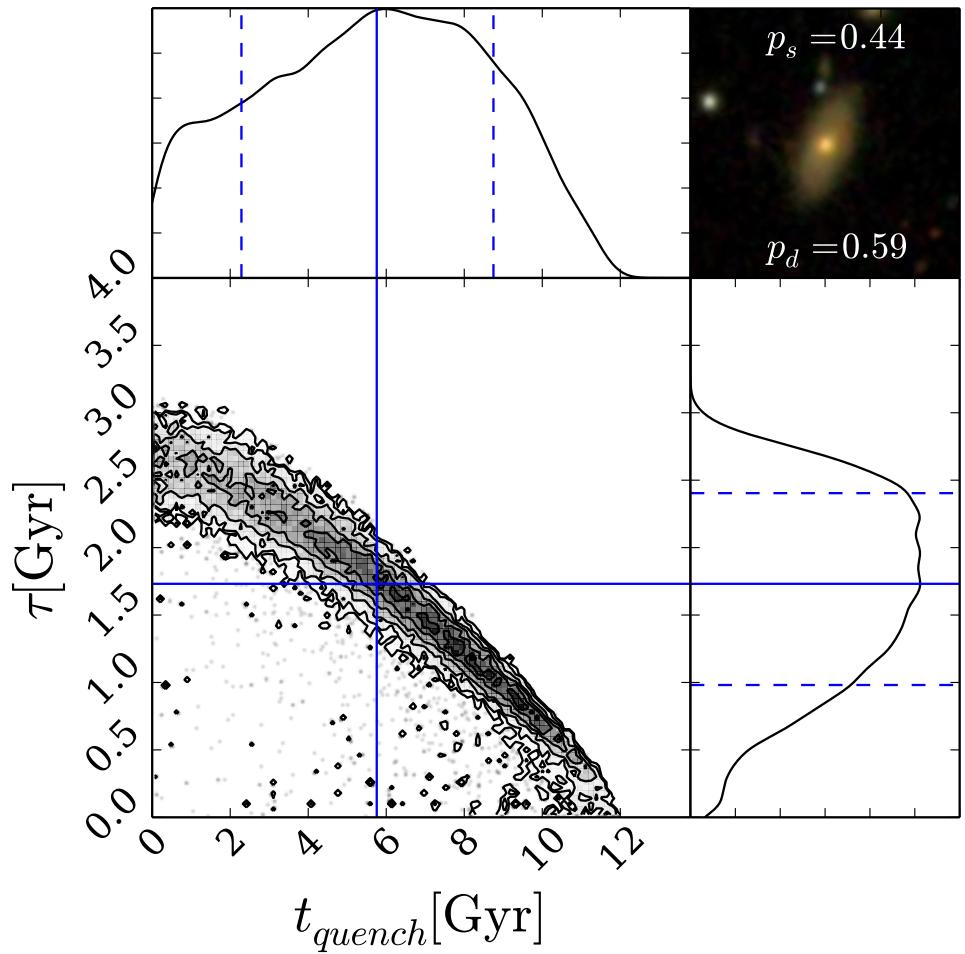


Figure 2.3: Example output from STARPY for a galaxy within the red sequence. The contours show the positions of the ‘walkers’ in the Markov Chain (which are analogous to the areas of high probability) for the quenching models described by  $\theta = [t_q, \tau]$ . The histograms show the 1D projection along each axis. Solid (dashed) blue lines show the best fit parameters (with  $\pm 1\sigma$ ) to the data. The postage stamp image from SDSS is shown in the top right along with the debiased vote fractions for smooth ( $p_s$ ) and disc ( $p_d$ ) from Galaxy Zoo 2.

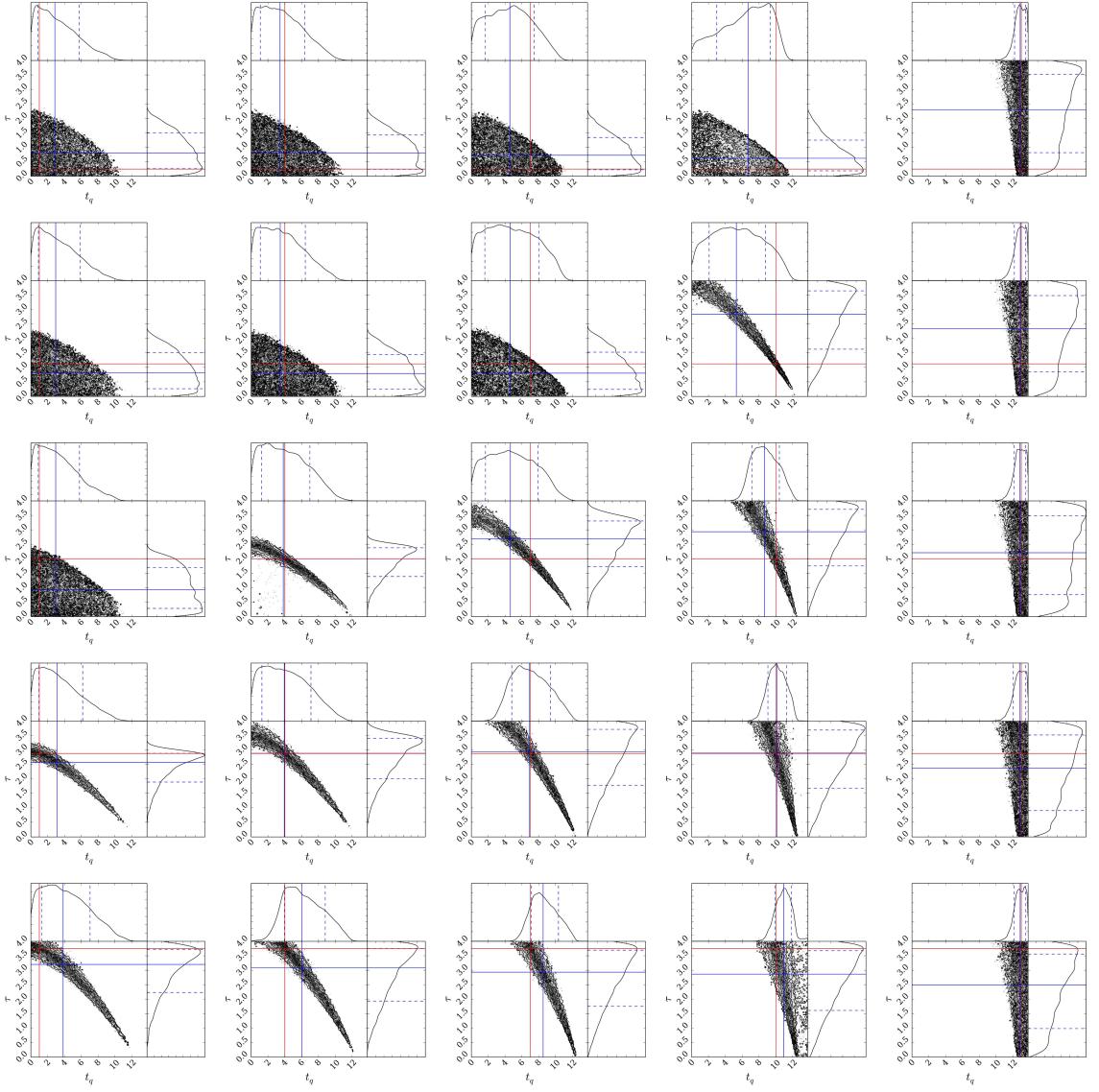


Figure 2.4: Results from STARPY for an array of synthesised galaxies with known, i.e. true,  $t_q$  and  $\tau$  values (marked by the red lines) using the complete function to calculate the predicted colour of a proposed set of  $\theta$  values in each MCMC iteration, assuming an error on the calculated known colours of  $\sigma_{u-r} = 0.124$  and  $\sigma_{NUV-u} = 0.215$  (the average errors on the GZ sample colours). I also assume that each synthesised galaxy has been observed at a redshift of  $z = 0$ . In each case STARPY succeeds (50th percentile best fit parameters are shown by the blue lines) in locating the true parameter values within the degeneracies of the star formation history model.

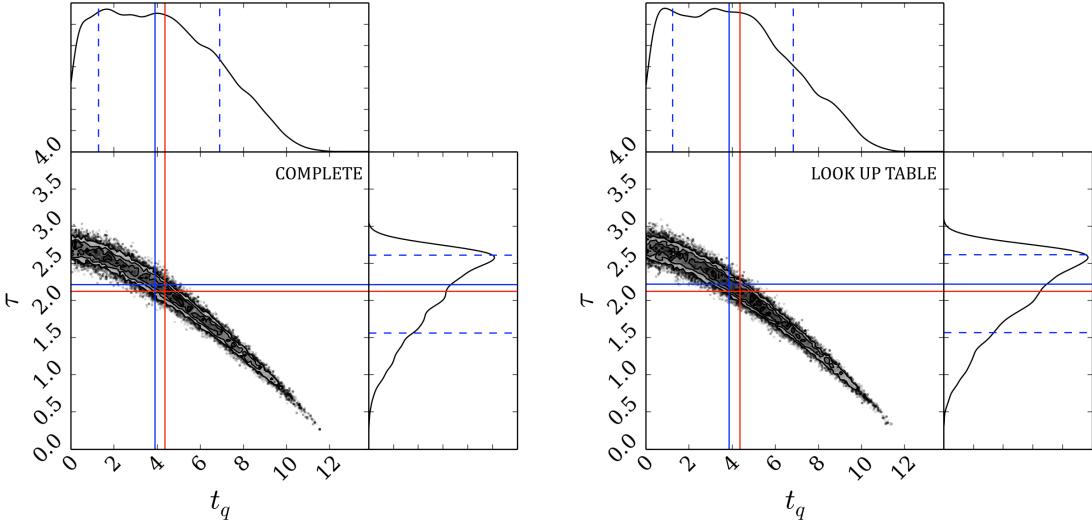


Figure 2.5: Left panel: Results from STARPY for true  $t_q$  and  $\tau$  values (red lines) using the complete function to calculate the predicted colour of a proposed set of  $\theta$  values in each MCMC iteration. The median walker position (the 50th percentile of the Bayesian probability distribution) is shown by the solid blue line with the dashed lines encompassing  $68\%(\pm 1\sigma)$  of the samples (the 16th and 84th percentile positions). The time taken to run for a single galaxy using this method is approximately 2 hours. Right panel: Results from STARPY for true  $t_q$  and  $\tau$  values using a look up table generated from the complete function to calculate the predicted colour of a proposed set of  $\theta$  values in each MCMC iteration. The time taken to run for a single galaxy using this method is approximately 2 minutes.

reproduced, within error, for each of the 25 synthesised galaxies. Figure 2.4 shows the results for each of these synthesised galaxies, with the known values of  $\theta$  shown by the red lines. In some cases this red line does not coincide with the inferred best fit  $\theta$  values shown by the blue lines, however in all cases the intersection of the red lines is within the sample contours; therefore STARPY succeeds in locating the true parameter values within the degeneracies of the SFH model.

## 2.4 Speeding up STARPY

I wish to consider the SFH model parameters for a large populations of galaxies across the colour magnitude diagram, however for each combination of  $\theta$  values which *emcee*

Table 2.1: Median walker positions (the 50th percentile; as shown by the blue solid lines in Figure 2.5) found by STARPY for a single galaxy, using the complete star formation history function and a look up table to speed up the run time. The errors quoted define the region in which 68% of the samples are located, shown by the dashed blue lines in Figure 2.5. The known true values are also quoted, as shown by the red lines in Figure 2.5. All values are quoted to three significant figures.

	$t_q$	$\tau$
True	4.37	2.12
Complete	$3.893 \pm^{3.014}_{2.622}$	$2.215 \pm^{0.395}_{0.652}$
Look up table	$3.850 \pm^{2.988}_{2.619}$	$2.218 \pm^{0.399}_{0.649}$

proposes for a single galaxy, a new SFH must be built, prior to convolving it with the BC03 SPS models at the observed age and then predicted colours calculated from the resultant SED. For a single galaxy this takes up to 2 hours on a typical desktop machine for long Markov Chains. A 3-dimensional look-up table was therefore generated at 50  $t^{obs}$ , 100  $t_{quench}$  and 100  $\tau$  values; this was then interpolated over for a given observed galaxy’s age and proposed  $\theta$  values at each step in the Markov Chain. This ensured that a single galaxy takes approximately 2 minutes to run on a typical desktop machine.

Figure 2.5 shows an example of how using the look up table in place of the full function does not affect the results to a significant level. Table 2.4 quotes the median walker positions (the 50th percentile of the Bayesian probability distribution) along with their  $\pm 1\sigma$  ranges for both methods in comparison to the true values specified to test STARPY. The uncertainties incorporated into the quoted values by using the look up table are therefore minimal with a maximum  $\Delta = 0.043$ .

Using this lookup table, each of the 126,316 total galaxies in the GZ2-GALEX sample was run through STARPY on multiple cores of a computer cluster to obtain the Markov Chain positions (analogous to  $P(\theta_k|d_k)$ ) for each galaxy,  $k$  (see Figure 2.2). In each case the Markov Chain consisted of 100 ‘walkers’ which took 400 steps in the ‘burn-in’ phase and 400 steps thereafter, at which point the MCMC acceptance fraction was checked to be within the range  $0.25 < f_{acc} < 0.5$  (which was true in all cases). Due to the Bayesian nature of this method, a statistical test on the results is not possible; the output is probabilistic in nature across the entirety of the parameter space.

## 2.5 POPSTARPY: studying populations of galaxies with STARPY

To study the SFH of a large population of galaxies, the individual galaxy walker positions output by STARPY (analogous to the posterior probability distribution) are combined across  $[t, \tau]$  space. The Markov Chain walker positions are binned and weighted by their corresponding logarithmic posterior probability  $\log[P(\theta_k|d_k)]$ , provided by the *emcee* package, in order to emphasise the features and differences between various populations. This weighting by  $\log[P(\theta_k|d_k)]$  is to minimise the contribution of galaxies poorly fit by this exponentially declining SFH. This is no longer inference but merely a method to visualise the results across a population of galaxies.

I also discard those walker positions with a corresponding normalised posterior probability of  $P(\theta_k|d_k) < 0.2$  in order to exclude galaxies which are not well fit by the quenching model, therefore galaxies in each sample which reside on the main sequence will not contribute to the final population distribution of quenching parameters. This raises the issue of whether I exclude a significant fraction of the GZ2-GALEX sample and whether those galaxies reside in a specific location of the colour-magnitude. The fraction of galaxies which had all or more than half of their walker positions discarded due to low probability are shown in Table 2.5. Using the  $P(\theta_k|d_k) < 0.2$  constraint, 2.4%, 7.0% and 5.4% of green, red and blue galaxies respectively had *all* of their walker positions discarded.

This is not a significant fraction of either population, therefore the STARPY module is effective in fitting the majority of galaxies and this method of discarding walker positions ensures that poorly fit galaxies are removed from the analysis of the results. Figure 2.6 shows that these galaxies with discarded walker positions are also scattered across the optical-NUV colour-colour diagram and therefore STARPY is also effective in fitting galaxies across this entire plane.

Figure 2.4 shows how peaks in the histograms are found across all areas of the parameter space in both dimensions  $[t, \tau]$ , ensuring that any conclusions drawn from combined population distributions are due to a superposition of extended probability distributions, as opposed to a bimodal distribution of probability distributions across all galaxies.

Table 2.2: The number of galaxies in each population which had walker positions discarded due to low posterior probability values in order to exclude those galaxies from the analysis which were poorly fit by the SFH quenching model.

	<b>Red Sequence</b>	<b>Green Valley</b>	<b>Blue Cloud</b>
All walkers discarded	1420 (7.00%)	437 (2.41%)	3109 (5.37%)
More than half walker positions discarded	2010 (9.92%)	779 (4.30%)	6669 (11.52%)

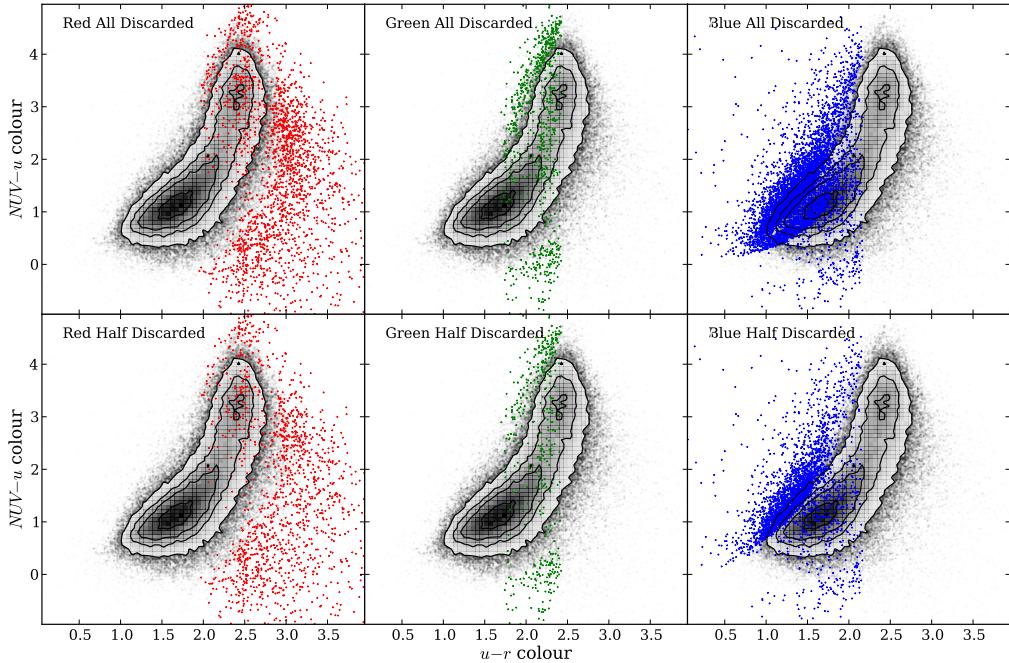


Figure 2.6: Contours show the full GZ2 subsample optical-NUV colour-colour diagram. The points show the positions of the galaxies which had all (top panels) or more than half (bottom panel) of their walker positions discarded due to their low probability for the red sequence (left), green valley (middle) and blue cloud (right).

The classifications from Galaxy Zoo 2 provide a uniquely powerful continuous measurements of a galaxy’s morphology, therefore I utilise the debiased user vote fractions to obtain separate population density distributions for both smooth and disc galaxies. This is obtained by also weighting by the morphology vote fraction when the binned walker positions are combined. This ensures that the entirety of the population is used, with galaxies with a higher  $p_d$  contributing more to the disc weighted than the smooth weighted population distribution. This negates the need for a threshold on the GZ2 vote fractions (e.g.,  $p_d > 0.8$  as used in Schawinski et al., 2014). These distributions will be referred to as the population densities.

For example, the galaxy shown in Figure 2.2 would contribute almost evenly to both the smooth and disc parameters due to the GZ2 vote fractions. Since galaxies with similar vote fractions contain both a bulge and disc component, this method is effective in incorporating intermediate galaxies which are thought to be crucial to the morphological changes between early- and late-type galaxies. It was the consideration of these intermediate galaxies which was excluded from the investigation by Schawinski et al. (2014).

### 2.5.1 Alternative Hierarchical Bayesian approach

The approach presented above relies upon a visualisation of the SFHs across each population, with no inference involved beyond the use of STARPY to derive the individual galaxy SFHs. An alternative approach to this problem would be to use a hierarchical Bayesian method to determine the ‘hyper-parameters’ that describe the distribution of the parent population  $\theta' = [t'_q, \tau']$  that each individual galaxy’s SFH is drawn from.

The posterior PDF for  $\vec{\theta}'$  to describe such a galaxy population:

$$P(\vec{\theta}'|\vec{d}) = \frac{P(\vec{d}|\vec{\theta}')P(\vec{\theta}')}{P(\vec{d})}, \quad (2.6)$$

where  $\vec{d}$  represents all of the optical and NUV colour data in a population  $\{\vec{d}_k\}$ . For one galaxy,  $k$ , the marginalised likelihood is:

$$P(d_k|\vec{\theta}') = \iint P(d_k|t_k, \tau_k)P(t_k, \tau_k|\vec{\theta}') \, dt_k \, d\tau_k \quad (2.7)$$

and for all galaxies,  $N$ , therefore:

$$P(\vec{d}|\vec{\theta}') = \prod_k^N P(d_k|\vec{\theta}'). \quad (2.8)$$

Using STARPY for an individual galaxy,  $k$  the output is the ‘interim’ posterior  $P(t_k, \tau_k|d_k)$  which I can relate to  $P(d_k|t_k, \tau_k)$  so that:

$$P(d_k|\vec{\theta}') = \iint P(t_k, \tau_k|d_k) \cdot P(d_k) \cdot \frac{P(t_k, \tau_k|\vec{\theta}')}{P(t_k, \tau_k)} dt_k d\tau_k. \quad (2.9)$$

In order to calculate this I draw  $N_s$  random samples,  $r$ , from each interim posterior,  $P(t_k, \tau_k|d_k)$  so that Equation 2.9 can be expressed as a sum over a number of random samples,  $N_s$  (as with the calculation of an expected mean):

$$P(d_k|\vec{\theta}') = \frac{P(d_k)}{N_s} \sum_r^{N_s} \frac{P(t_{k,r}, \tau_{k,r}|\vec{\theta}')}{P(t_k, \tau_k)}, \quad (2.10)$$

for the  $r^{th}$  sample of  $N_s$  total samples taken from one galaxy’s,  $k$ , interim posterior PDF. This fraction is known as the ‘importance weight’,  $w_r$ , in importance sampling.

However, I also have two morphological vote fractions that I can weight by to determine separate hyper-parameters,  $\vec{\theta}' = [\vec{\theta}_d', \vec{\theta}_s']$ , for both disc,  $d$ , and smooth,  $s$ , galaxies. Therefore:

$$w_r = \frac{P(t_{k,r}, \tau_{k,r}|\vec{\theta}')}{P(t_k, \tau_k)} = \frac{p_{d,k}P(t_{k,r}, \tau_{k,r}|\vec{\theta}_d') + p_{s,k}P(t_{k,r}, \tau_{k,r}|\vec{\theta}_s')}{P(t_k, \tau_k)} \quad (2.11)$$

If we substitute equation 2.10 into equation 2.6 we find that the  $P(d_k)$  terms cancel and we are left with:

$$P(\vec{\theta}'|\vec{d}) = P(\vec{\theta}') \prod_k^N \frac{1}{N_{s,k}} \sum_r^{N_s} w_r, \quad (2.12)$$

where  $P(\vec{\theta}')$  is the assumed prior on the hyper-parameters, which is assumed to be uniform.

This approach is heavily dependent on what shape is assumed for the hyper-distribution; a decision which is not trivial. It is often common for this function to take the form of a multi-component Gaussian mixture model (Mackay, 2003;

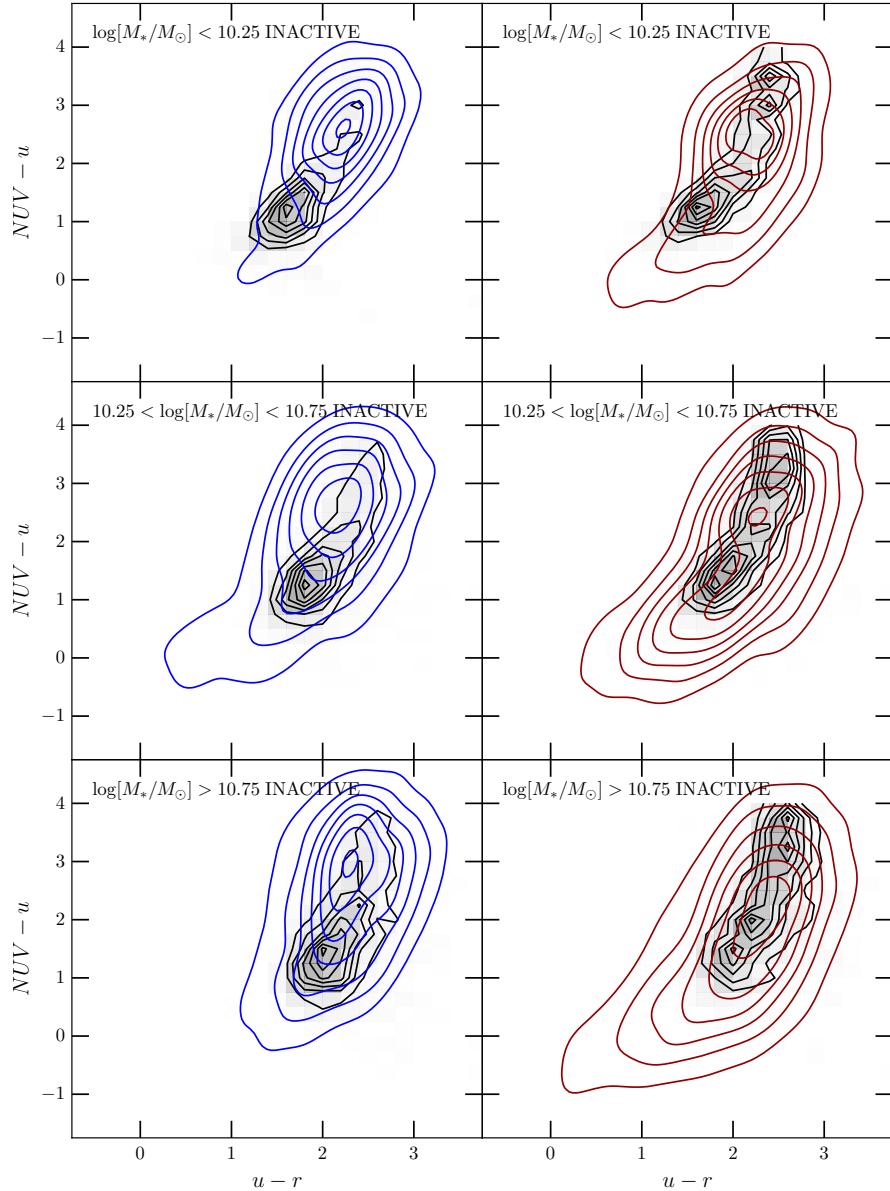


Figure 2.7: Optical-NUV colour-colour diagrams for the INACTIVE galaxies shown by the black contours, split into low mass (top), medium mass (middle) and high mass (bottom) galaxies weighted by  $p_d$  (left) and  $p_s$  (right). Kernel smoothing has been applied to the overlaid replica datasets, which are created by sampling from the **inferred 2 component Gaussian mixture model hierarchical parent distributions**. Gaussian random noise is also added to the inferred colours, with a mean and standard deviation of the errors on the observed colours of the respective sample. Contours are shown for samples taken from the disc (blue) and smooth weighted (red) inferred hierarchical distributions.

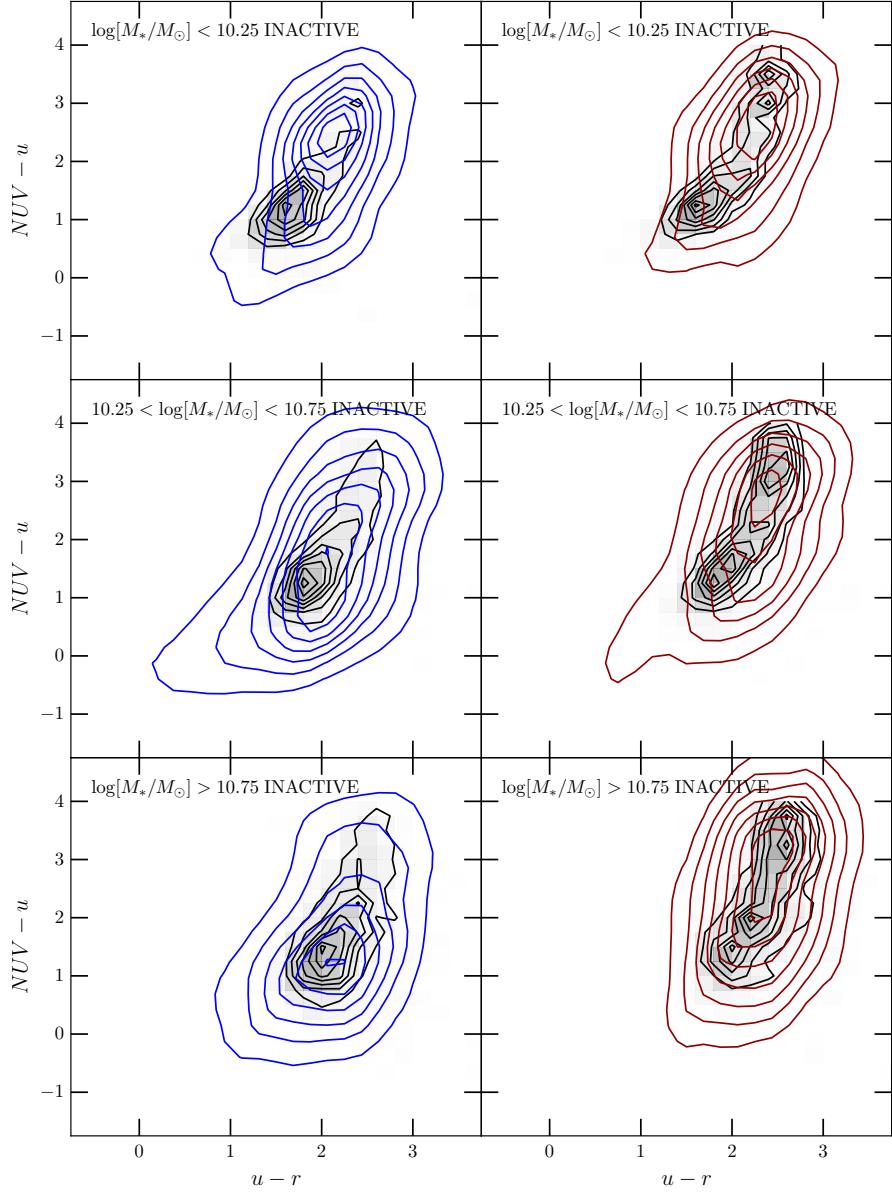


Figure 2.8: Optical-NUV colour-colour diagrams for the INACTIVE galaxies shown by the black contours, split into low mass (top), medium mass (middle) and high mass (bottom) galaxies weighted by  $p_d$  (left) and  $p_s$  (right). Kernel smoothing has been applied to the overlaid replica datasets, which are created by sampling from the **popstarpy population density distributions described in Section 2.5**. Gaussian random noise is also added to the inferred colours, with a mean and standard deviation of the errors on the observed colours of the respective sample. Contours are shown for samples taken from the disc (blue) and smooth weighted (red) inferred hierarchical distributions.

Lahav et al., 2000). For example a two component Gaussian mixture model in  $[t, \tau]$  space is described by eight hyper-parameters for a single morphology,  $\vec{\theta}' = [\mu_{t,1}, \sigma_{t,1}, \mu_{\tau,1}, \sigma_{\tau,1}, \mu_{t,2}, \sigma_{t,2}, \mu_{\tau,2}, \sigma_{\tau,2}]$ . This approach assumes no covariance between hyper-parameters for simplicity. The equations outlined above, combined with MCMC methods can be used to infer these  $8\vec{\theta}'$  parameters from which the hierarchical population distribution can be determined.

In order to test whether this assumption of a multi-component Gaussian mixture model is appropriate, I sampled the inferred hierarchical distributions to produce replica datasets in optical-NUV colour space. These are shown here in Figure 2.7 in comparison to the observed colour-colour distributions of the INACTIVE sample (a subset of  $\sim 6,000$  galaxies from the GZ2-GALEX sample, see Section 4.1). For all masses and morphologies the replicated  $u - r$  and  $NUV - u$  colours do not accurately match the observed data.

I also varied the value of  $N_s$  and found that increasing the number of samples drawn did not improve this fit for the INACTIVE population. Similarly increasing the number of components in the Gaussian mixture model did not immediately improve the accuracy of the fit. I therefore concluded that this functional form of the population distribution was unsatisfactory.

The POPSTARPY approach described in section 2.5 was motivated by the investigation increasing the number of samples,  $N_s$  drawn from the posterior of each galaxy,  $k$ , until the point where all the samples were drawn. Instead of attempting to infer parameters to describe this distribution, as above, I presented the distribution itself (as described in Section 2.5). The distributions produced by this visualisation method reveal the complexity that the parent distribution must describe which, as concluded earlier, cannot be effectively modelled.

I also tested whether the POPSTARPY method is reasonable by producing replica datasets in optical-NUV colour space, as before, by drawing 1000  $[t, \tau]$  values from the population density distributions derived for the INACTIVE sample (see Section 4.1). These replica datasets are shown here in Figure 2.8 in comparison to the observed colour-colour distributions of the INACTIVE sample. Comparing these replica colours in Figure 2.8, with those produced by drawing from the inferred hierarchical distributions, shown in Figure 2.7, they can be seen to produce a more accurate match to the observed data for the majority of masses and morphologies.

Considering these issues with assuming a functional form for the hierarchical parent distribution, an expansion on this approach would be to perform ‘heat map optimization’, similar to image reconstruction, to determine the parent distribution for a given population. Each pixel would need a prior (e.g. a basic entropic prior) and the heat map would sum to unity. This is a significant expansion upon the work presented here and is something the author wishes to investigate in future work.

For the results presented in the following chapters, I therefore use the POPSTARPY method to visualise the population distribution, rather than quoting inferred values to describe it.