

Bayesian methods for determining galactic star formation histories

R. J. Smethurst,¹

¹ *Oxford Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK*

24 September 2015

ABSTRACT

Does galaxy evolution proceed across the colour magnitude diagram via multiple pathways or as a single population? For my thesis I have attempted to answer this general question through the use of a new Bayesian determination of the most likely star formation history of a galaxy given its observed optical and NUV photometry. Using this new method and classifications from Galaxy Zoo I have been able to show that diverse morphologically dependant quenching mechanisms with a range of quenching rates are all instrumental in the formation of the present day red sequence. This is contrary to previous work highlighting radically different evolutionary pathways between early- and late-type galaxies. I have extended this work to explore the relationship between these quenched star formation histories and the presence of a type 2 AGN, finding evidence to show that a population of these AGN host galaxies have recently undergone a rapid drop in their star formation rate. This result provides strong observational support for AGN feedback in these systems, but this work also shows that feedback cannot be responsible for all of the quenching across the AGN host population. I shall discuss these results and how they will contribute to my thesis.

1 INTRODUCTION

Previous large scale surveys of galaxies have revealed a bimodality in the colour-magnitude diagram (CMD) with two distinct populations; one at relatively low mass, with blue optical colours and another at relatively high mass, with red optical colours (Baldry et al. 2004, 2006; Willmer et al. 2006; Ball, Loveday & Brunner 2008; Brammer et al. 2009). These populations were dubbed the ‘blue cloud’ and ‘red sequence’ respectively.

The sparsely populated colour space between these two populations, the so-called ‘green valley’, provides clues to the nature and duration of galaxies’ transitions from blue to red. These transitions must occur on rapid timescales, otherwise we would observe an accumulate of galaxies in the green valley. Green valley galaxies are therefore thought to be a transition population between the blue cloud and the ‘dead’ red sequence (Martin et al. 2007; Faber et al. 2007; Mendez et al. 2011; Gonçalves et al. 2012).

The intermediate colours of these green valley galaxies have been interpreted as evidence for recent quenching (suppression) of star formation (Salim et al. 2007). Star forming galaxies are observed to lie on a well defined mass-SFR relation, however quenching a galaxy causes it to depart from this relation (Noeske et al. 2007; Peng et al. 2010).

By studying the galaxies which have just left this mass-SFR relation, the quenching mechanisms by which this occurs can be probed. There have been many previous theories for the initial triggers of these quenching mechanisms, including negative feedback from AGN (Schawinski

et al. 2007), mergers (Darg et al. 2010), supernovae winds (Marasco, Fraternali & Binney 2012) and secular evolution (Masters et al. 2010a, 2011). In particular observations of AGN host galaxies reveal their location in the green valley, suggesting a particular role for AGN feedback in the process of quenching (Nandra et al. 2007; Hasinger 2008; Silverman et al. 2008; Schawinski et al. 2014). By investigating the *amount* of quenching that has occurred across the colour magnitude diagram and in a population of AGN host galaxies, constraints to these quenching mechanism theories can be applied.

To investigate the rates of quenching and therefore the possible mechanisms I have implemented a simple novel method utilising Bayesian statistics (for a comprehensive overview of Bayesian statistics see either MacKay 2003 or Sivia 1996) in order to find the typical model parameters of the star formation history of galaxies in a given population. Previous investigations have only implemented frequentist methods on small sample sizes; however here I have implemented Bayesian methods on large populations of galaxies to determine the evolution. I have assumed the simplest model possible in order to derive any useful information before implementing more sophisticated techniques in future work.

2 METHODS

STARPY¹ is a PYTHON code which allows the user to derive the quenched star formation history (SFH) of a galaxy through a Bayesian Markov Chain Monte Carlo method (Foreman-Mackey et al. 2013)² with the input of the observed $u-r$ and $NUV-u$ colours, a redshift, and the use of the stellar population models of Bruzual & Charlot (2003). The star formation history template is an exponential decline of the SFR and is described by two parameters $[t_q, \tau]$, where t_q is the time at which the onset of quenching begins [Gyr] and τ is the exponential rate at which quenching occurs [Gyr]. Under the simplifying assumption that all galaxies formed at $t = 0$ Gyr with an initial burst of star formation, the SFH can be described as:

$$SFR = \begin{cases} i_{sfr}(t_q) & \text{if } t < t_q \\ i_{sfr}(t_q) \times \exp\left(\frac{-(t-t_q)}{\tau}\right) & \text{if } t > t_q \end{cases} \quad (1)$$

where i_{sfr} is an initial constant star formation rate dependent on t_q (Schawinski et al. 2014; Smethurst et al. 2015). A smaller τ value corresponds to a rapid quench, whereas a larger τ value corresponds to a slower quench. The output of STARPY is probabilistic in nature and provides the posterior probability distribution across the entirety of the two parameter space for each individual galaxy. The probabilistic fitting methods to this SFH for an observed galaxy are described in full detail in Smethurst et al. (2015).

We apply this method to derive the quenched SFHs of a sample of 126,316 SDSS galaxies also observed with GALEX (to acquire NUV photometry) that have been classified by the Galaxy Zoo 2 project Willett et al. (2013). We refer to this as the GZ2-GALEX sample. The output is a 2-dimensional likelihood distribution across the $[t_q, \tau]$ parameter space for each individual galaxy. We sum individual galaxy likelihood distributions across the blue cloud, green valley and red sequence populations of the colour magnitude diagram, additionally weighting by the GZ2 morphologies to produce quenching parameter likelihood distributions for both disc- and smooth-dominated galaxy populations.

3 RESULTS TO DATE

3.1 Quenching across the Colour Magnitude Diagram

The SFH models were implemented with the STARPY package to produce Figures 1 & 2 for the red sequence, green valley and blue cloud populations of smooth and disc galaxies respectively. The percentages shown in Figures 1 & 2 are calculated as the fractions of the combined posterior probability distribution located in each region of parameter space for a given population.

The left panel of Figure 1 reveals that smooth galaxies with red optical colours show a preference (49.5%; see Figure 1) for rapid quenching timescales across all cosmic time. For these smooth red galaxies we see, at early times only, a preference for slow and intermediate timescales. The right panel of Figure 1 reveals that red disc galaxies show

are dominated by slow (44.1%) quenching timescales. The preference for *very* slow ($\tau > 3.0$ Gyr) quenching timescales (which are not seen for the green valley discs) suggests that these galaxies have only just reached the red sequence after a very slow evolution across the colour-magnitude diagram.

In Figure 2 intermediate quenching timescales become more significant (in agreement with the conclusions of Gonçalves et al. 2012), particularly for smooth-like galaxies (see the left panel of Figure 2).

The smooth galaxy parameters favour these intermediate quenching timescales (40.6%) with some preference for slow quenching at early times ($z > 1$). The preference for rapid quenching of smooth galaxies has dropped by over a half compared to the red galaxies. The disc galaxies of the green valley now overwhelmingly prefer slow quenching timescales (47.4%). If we compare Figure 2 to Figure 1 we can see quenching has occurred at later (more recent) cosmic times in the green valley for both morphological types. Therefore both morphologies are tracing the evolution of the red sequence.

In summary, rapid quenching is much more prevalent in smooth galaxies than disc galaxies, and the red galaxies are also much more likely to be characterised by a rapid quenching model than green valley galaxies. In the green valley there is also a distinct lack of preference for rapid quenching timescales with $\tau < 0.5$ Gyr, more so for the disc- than the smooth-like galaxies. This suggests that this rapid quenching mechanism causes a change in morphology from a disc- to a smooth-like galaxy as a galaxy moves across the colour magnitude diagram.

One simulation of interest by Springel, Di Matteo & Hernquist (2005) showed that feedback from black hole activity is a necessary component of destructive major mergers to produce such rapid quenching timescales. They find using hydrodynamical simulations that after ~ 1 Gyr the merger remnant has reddened to $u-r \sim 2.0$. This is in agreement with my simple quenching models which give that within ~ 1 Gyr the models with a SFH with $\tau < 0.4$ Gyr have reached the red sequence with $u-r \gtrsim 2.2$. In comparison the simulations of Springel, Di Matteo & Hernquist (2005) without any feedback from black holes, suggest that if even a small fraction of gas is not consumed in the starburst following a merger (either because the mass ratio is not large enough or from the lack of strong black hole activity) the remnant can sustain star formation for periods of several Gyrs. The remnants from these simulations take ~ 5.5 Gyr to reach red optical colours of $u-r \sim 2.1$. This is again in agreement with my simple quenching models with intermediate quenching timescales of $1.0 \lesssim \tau$ [Gyr] $\lesssim 2.0$ which take approximately 2.5–5.5 Gyr to reach these red colours.

These intermediate quenching timescales are found to be equally prevalent across populations for both smooth and disc galaxies across cosmic time particularly in the green valley. Intermediate timescales are the prevalent mechanism for quenching smooth green valley galaxies, unlike the rapid quenching prevalent for red galaxies. We suggest that this intermediate quenching route must therefore be possible with routes that both preserve and transform morphology. It is this result of another route through the green valley that is in contradiction with the findings of Schawinski et al. (2014).

We speculate that the intermediate quenching timescales are caused by gas rich major mergers, major

¹ Publicly available: <http://github.com/zoouniverse/starpy>

² <http://dan.iel.fm/emcee/>

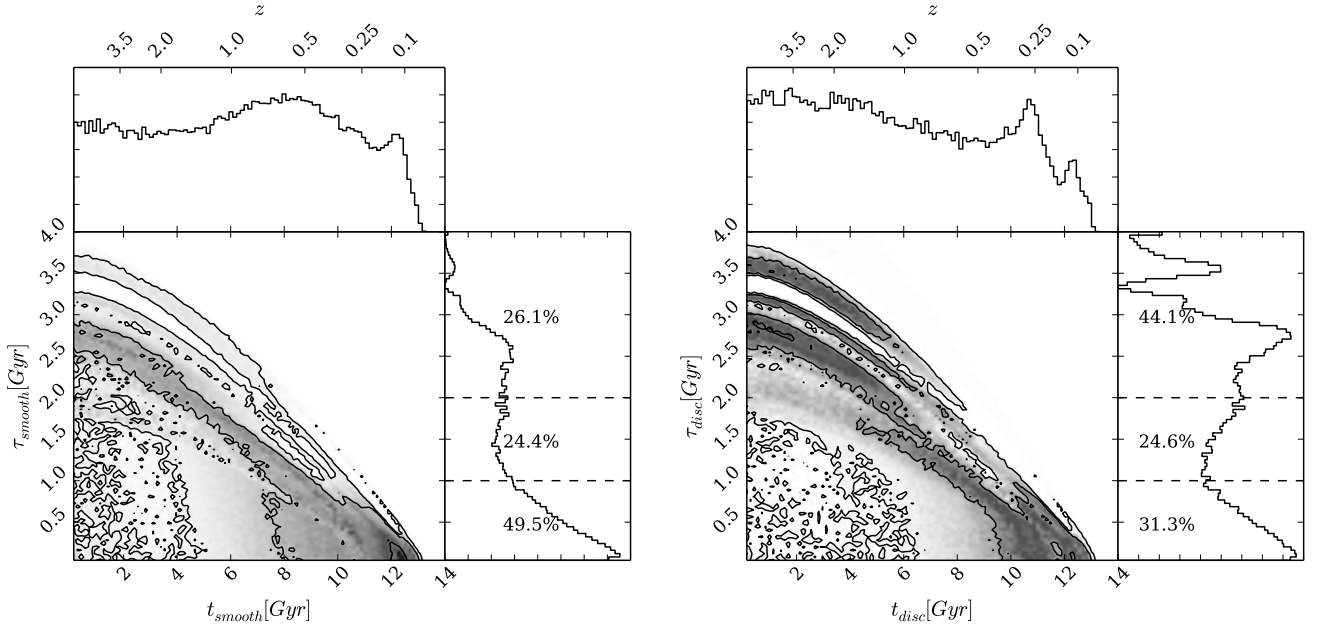


Figure 1. Contour plots showing the summed posterior probability distribution function weighted by the morphological vote fractions from GZ2 to give the areas of high likelihood in the model parameter space for both elliptical (left) and disc (right) dominated systems. The histograms show the projection into one dimension for each parameter. The dashed lines show the separation between rapid (τ [Gyr] < 1.0), intermediate ($1.0 < \tau$ [Gyr] < 2.0) and slow (τ [Gyr] > 2.0) quenching timescales with the fraction of the combined posterior probability distribution in each region shown

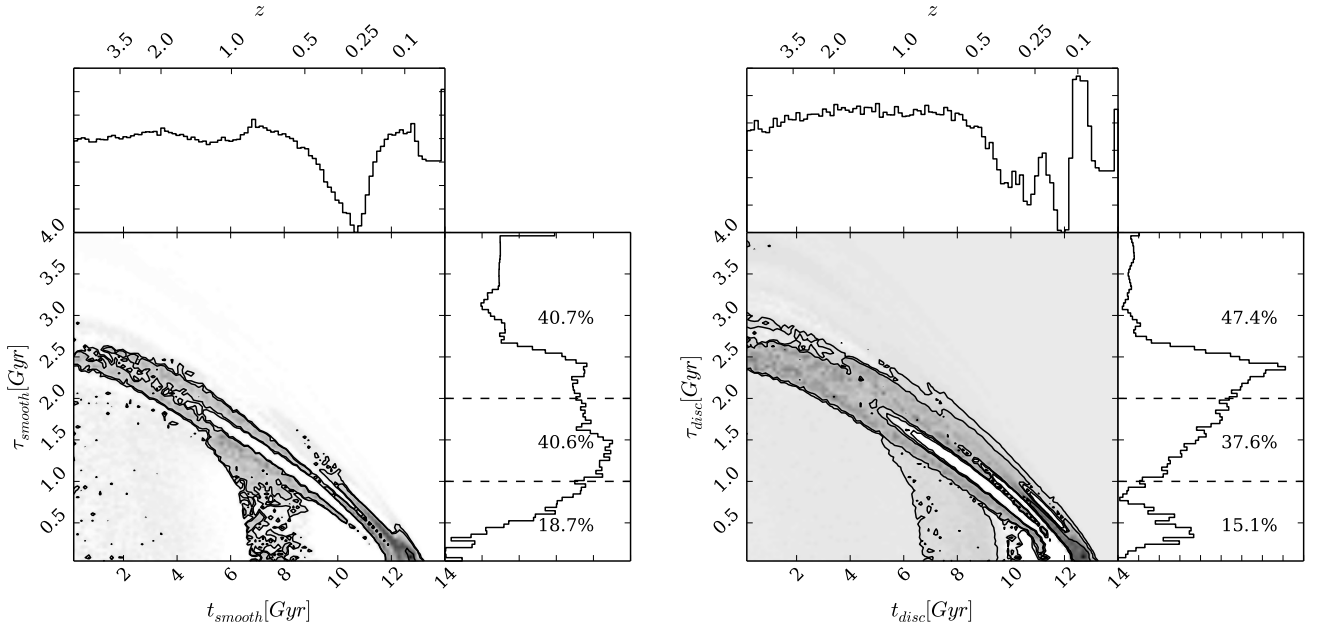


Figure 2. Contour plots showing the summed posterior probability distribution function weighted by the morphological vote fractions from GZ2 to give the areas of high likelihood in the model parameter space for both elliptical (left) and disc (right) dominated systems. The histograms show the projection into one dimension for each parameter. The dashed lines show the separation between rapid (τ [Gyr] < 1.0), intermediate ($1.0 < \tau$ [Gyr] < 2.0) and slow (τ [Gyr] > 2.0) quenching timescales with the fraction of the combined posterior probability distribution in each region shown

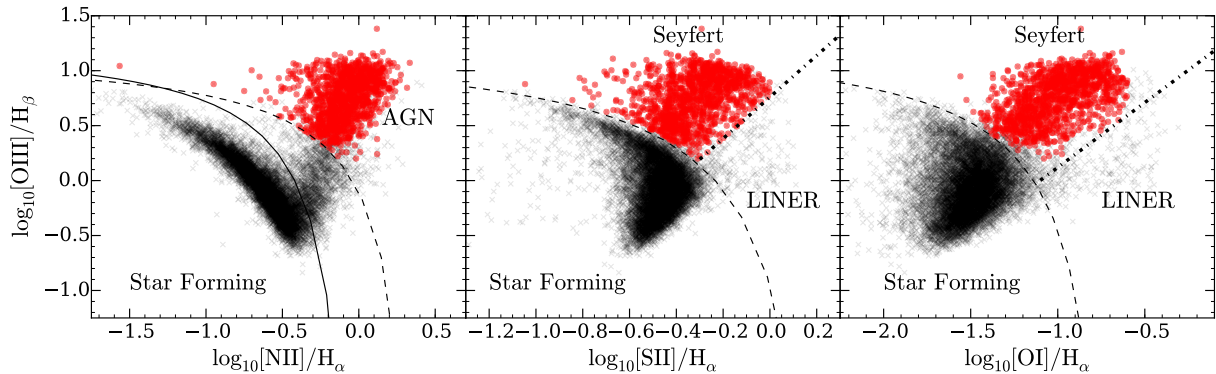


Figure 3. BPT diagrams for galaxies in the GZ2-GALEX sample (black crosses) with $S/N > 3$ for each emission line. Inequalities defined in: Kewley et al. (2001) to separate SF galaxies from AGN (dashed lines), Kauffman et al. (2003b) to separate SF from composite SF-AGN galaxies (solid line) and Kewley et al. (2006) to separate LINERS and Seyferts (dotted lines). Galaxies are included in the AGN-HOST sample (red circles) if they satisfy all the inequalities to be classified as Seyferts.

mergers without black hole feedback and from minor mergers, the latter of which is the dominant mechanism.

Although intermediate and rapid quenching timescales are the dominant mechanisms across the colour-magnitude diagram, together they cannot completely account for the quenching of disc galaxies. There is also a significantly lower preference for smooth galaxies to undergo such slow quenching timescales; suggesting that the evolution (or indeed creation) of typical smooth galaxies is dominated by processes external to the galaxy.

Bamford et al. (2009) using GZ1 vote fractions of galaxies in the SDSS, found a significant fraction of high stellar mass red spiral galaxies in the field. As these galaxies are isolated from the effects of interactions from other galaxies, the slow quenching mechanisms present in their preferred star formation histories are most likely due to secular processes (i.e. mechanisms internal to the galaxy, in the absence of sudden accretion or merger events; Kormendy & Kennicutt 2004; Sheth et al. 2012). Bar formation in a disc galaxy is such a mechanism, whereby gas is funnelled to the centre of the galaxy by the bar over long timescales where it is used for star formation (Masters et al. 2012; Saintonge et al. 2012; Cheung et al. 2013), consequently forming a ‘pseudo-bulge’ (Kormendy et al. 2010; Simmons et al. 2013).

If we believe that these slow quenching timescales are due to secular evolution processes, this is to be expected since these processes do not change the disc dominated nature of a galaxy.

3.2 Type 2 AGN Host Quenching Histories

The prevalence of rapid quenching rates found for smooth-like galaxies in the red sequence and their connection to black hole activity discussed above led to an investigation of the star formation histories across a population of Type 2 AGN host galaxies. Type 2 AGN were used to investigate this connection due to their photometric obscuration and were selected using a BPT diagram (Baldwin, Phillips & Terlevich 1981) using line and continuum strengths for [OIII], [NII], [SII] and [OI] obtained from the MPA-JHU catalogue (Kauffman et al. 2003a; Brinchmann et al. 2004)

for the galaxies studied above. We then required the $S/N > 3$ for each emission line as in Schawinski et al. (2010). Those galaxies which satisfied all of the inequalities defined in Kewley et al. (2001) and Kauffman et al. (2003b) to be classified as Seyfert galaxies were selected giving 1,244 AGN host galaxies and are shown in red in Figure 3. LINERs (shown not to be primarily powered by AGN) and composite galaxies were removed for purity. We refer to this sample as the AGN-HOST sample.

We constructed a sample of inactive galaxies by removing from the GZ2-GALEX sample all galaxies in the AGN-HOST sample, as well as sources identified as Type 1 AGN by the presence of broad emission lines (Oh et al. 2015). We refer to this sample as the INACTIVE sample.

We also split both the AGN-HOST and INACTIVE samples into low, medium and high mass ranges (see Figures 4 & 5 to investigate any trends in the SFH with mass. Masses were calculated using the $(u-r)$ colour and absolute r -band magnitude with the method outlined in Baldry et al. (2006).

We apply the method outlined in Section ?? to obtain a 2-dimensional likelihood distribution across the $[t_q, \tau]$ parameter space. We combine individual galaxy likelihood distributions within the AGN-HOST and INACTIVE galaxy samples, additionally weighting by GZ2 morphologies to produce quenching parameter likelihood distributions for both disc- and smooth-dominated galaxy populations as in section 3.1. In Figures 4 and 5 we show only the one dimensional histograms across each parameter.

At all masses, the distribution of likelihood for the AGN-HOST population across the quenching time t_q parameter (left panels of Figure 4) is different from that of the inactive galaxies (right panels of Figure 4). Recent quenching ($t > 11$ Gyr) of AGN-HOST galaxies is the dominant history for low and medium mass galaxies, particularly for the smooth galaxy population. However, this effect is less dominant in higher mass galaxies where quenching at earlier times has significant likelihood.

The distributions of likelihood for the quenching rate, τ , in Figure 5 show the dominance of rapid quenching ($\tau < 1$ Gyr) across the AGN-HOST population, particularly for smooth galaxies. With increasing mass the dominant

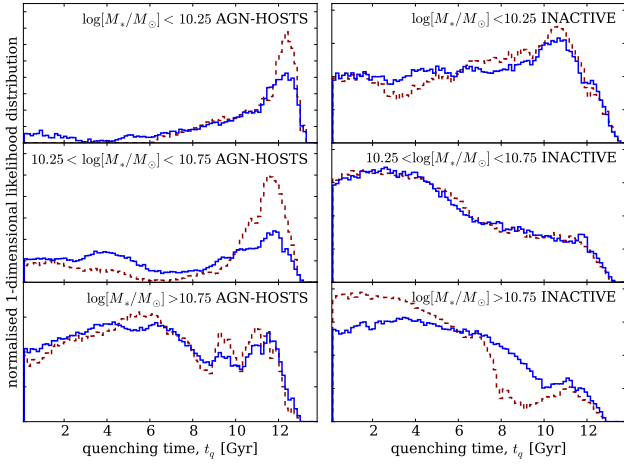


Figure 4. Likelihood distribution for the quenching time, t_q parameter normalised so that the areas under the curves are equal. AGN-HOST (left) and INACTIVE (right) galaxies are split into low (top), medium (middle) and high (bottom) mass for smooth (red dashed) and disc (blue solid) galaxies. A low value of t_q corresponds to the early Universe and a high value to the recent Universe.

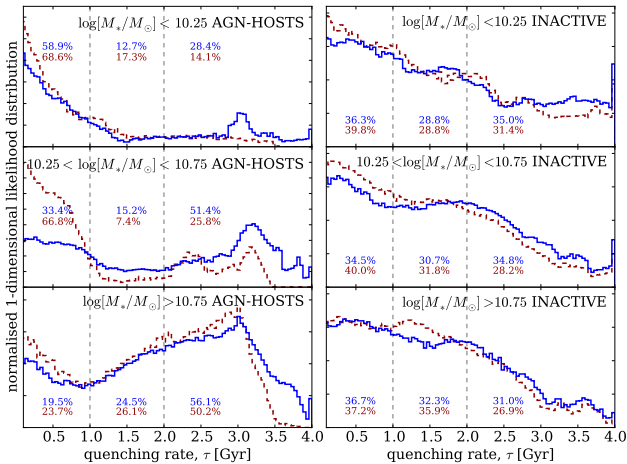


Figure 5. Likelihood distribution for the quenching rate, τ normalised so that the areas under the curves are equal. AGN-HOST (left) host and INACTIVE (right) galaxies are split into low (top), medium (middle) and high (bottom) mass for smooth (red dashed) and disc (blue solid) galaxies. The black dashed lines show the separation between rapid ($\tau < 1.0$ Gyr), intermediate ($1.0 < \tau$ [Gyr] < 2.0) and slow ($\tau > 2.0$ Gyr) quenching timescales with the percentage likelihood distribution in each region for disc (blue) and smooth (red) populations. A small (large) value of τ corresponds to a rapid (slow) quench.

quenching rate becomes slow ($\tau > 2$ Gyr) especially for disc galaxies hosting an AGN. Similar trends in likelihood are observed for the INACTIVE population but the overall distribution of likelihood is very different.

The likelihood distribution for the AGN-HOST galaxies therefore shows evidence for the dominance of rapid, recent quenching having occurred across the entire population.

This result implies the importance of AGN feedback for the evolution of these galaxies.

Tortora et al. (2009) model the effects of negative AGN feedback on a typical early type (i.e. smooth) galaxy and find the time between the current galaxy age in their simulation, t_{gal} and the time that the feedback began, t_{AGN} , peaks at $t_{gal} - t_{AGN} \sim 0.85$ Gyr. This agrees with the location of the peak in Figure 4 for low mass galaxies, where the difference between the peak of the likelihood and the average age of the population (galaxy age is calculated from the redshift by assuming all galaxies form at $t = 0$) is ~ 0.83 Gyr. This implies that this dominant recent quenching history is caused directly by AGN feedback, as opposed to the AGN being a consequence of an alternative quenching mechanism.

Rapid quenching is particularly dominant for low-to-medium mass smooth galaxies. Smethurst et al. (2015) suggest that incredibly rapid quenching rates could be attributed to mergers of galaxies in conjunction with AGN feedback, which are thought to be responsible for creating the most massive smooth galaxies (Conselice et al. 2003; Springel, Di Matteo & Hernquist 2005; Hopkins et al. 2008). This dominance of rapid quenching across the smooth AGN-HOST population supports the idea that a merger, having caused a morphological transformation to a smooth galaxy, can also trigger an AGN, causing feedback and cessation of star formation (see Figure 14 of Schawinski et al. 2014).

4 THESIS DEVELOPMENT

The two main results discussed above in Sections 3.1 & 3.2 will form the basis of my thesis with an in depth discussion of the limitations of STARPY and the uncertainties associated with it. I am also currently investigating the connection between the morphology and the cluster environment of galaxies in the GZ2-GALEX sample, the results of which I hope will populate another chapter.

A plan for my thesis is highlighted below with two line abstracts for each chapter.

- (i) **Introduction:** The morphological evolution of galaxies across the colour magnitude diagram is intrinsically tied to the quenching of star formation. There has been much speculation on the possible processes driving this evolution.
- (ii) **Galaxy Zoo:** Galaxy Zoo is a citizen science project which asks users to classify the morphology of SDSS images of galaxies using a simple web interface. The output from this project is statistically robust, reliable morphological classifications for 250,000+ galaxies.
- (iii) **STARPY :** STARPY is a piece of software which allows the user to determine the likelihood distribution for the 2D star formation history of a galaxy given two photometric colour measurements and a redshift. This is determined using Bayesian statistics and Markov Chain Monte Carlo methods to produce robust statistical results.
- (iv) **Quenching across the Colour Magnitude Diagram:** Green valley galaxies are thought to be the evolutionary missing link between the red sequence of red, dead, massive galaxies and the blue cloud of blue, star forming, less massive galaxies. Recent work showed highlighted radically different pathways of quenching for different morphological types, here I show, using STARPY, that diverse quenching

mechanisms with a large range of quenching rates are instrumental in forming the present day red sequence.

- (v) **Type 2 AGN Host Quenching Histories:** The co-evolution of galaxies and their central super massive black holes has long been theorised with ideas of negative feedback from AGN causing quenching in galaxies. Using STARPY I demonstrate that a population of Type 2 AGN galaxies shows a strong likelihood for rapid quenching occurring in the last 2 Gyr which is not found for a population of currently inactive galaxies.
- (vi) **Environmental Quenching Dependancies:** The density in which a galaxy evolves has a significant effect on the morphology of a galaxy by either restricting, providing or rationing the gas available for star formation. Using STARPY I (hope to) demonstrate how the rate of quenching is dependent on a galaxies position in a cluster as either a central or satellite.
- (vii) **Observational Evidence:** Observing runs were completed at the Isaac Newton Telescope, La Palma, Canary Islands and at the Caltech Submillimeter Observatory, Hawaii. Observations of both bulgeless galaxies with growing black holes and blue elliptical galaxies were taken as observational support of the theories discussed in support of the research findings above.
- (viii) **Discussion & Conclusion:** There is a clear difference between the quenching timescales of smooth and disc dominated galaxies suggesting that the processes driving quenching are morphologically dependant. However, whether the properties of galaxies observed (e.g. the presence of an AGN, tidal tails, location in cluster etc.) are the cause or consequence of these quenching mechanisms is still under discussion here.

5 CONCLUSION

I have used morphological classifications from the Galaxy Zoo 2 project to determine the morphology-dependent star formation histories across the colour magnitude diagram and in a population of AGN host galaxies. We determined the most likely parameters for the quenching onset time, t_q , and exponential quenching rate, τ , and find clear differences in the combined population likelihoods between red sequence and green valley galaxies, providing evidence for the possible driving mechanisms behind this quenching including secular evolution, galaxy interactions and mergers with and without AGN. the inactive and AGN host galaxy populations. In particular this relationship between the quenched star formation histories and the presence of a type 2 AGN has been shown with evidence that a population of AGN host galaxies having recently undergone a rapid drop in their star formation rate. This result provides strong observational support for AGN feedback in these systems, but this work also shows that feedback cannot be responsible for all of the quenching across the AGN host population

REFERENCES

- Baldry, I. K. et al., 2004, ApJ, 600, 681
- Baldry, I. K. et al., 2006, MNRAS, 373, 469

- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
- Ball, N. M., Loveday, J. & Brunner, R. J., 2008, MNRAS, 383, 907
- Bamford, S. P. et al., 2009, MNRAS, 393, 1324
- Brammer, G. B. et al., 2009, ApJ, 706, 173
- Brinchmann, J. et al., 2004, MNRAS, 351, 1151
- Bruzual, G. & Charlot, S., 2003, MNRAS, 344, 1000
- Cheung, E. et al., 2013, ApJ, 779, 162
- Conselice, C. J. et al., 2003, AJ, 126, 1183
- Darg, D. et al., 2010a, MNRAS, 401, 1552
- Faber, S. M. et al., 2007, ApJ, 665, 265
- Foreman-Mackey, D., Hogg, D. W., Lang, D., Goodman, J., 2013, PASP, 125, 306
- Gonçalves, T. S. et al., 2012, ApJ, 759, 67
- Hasinger, G., 2008, A&A, 490, 905
- Hopkins, F. et al., 2008, ApJSS, 175, 390
- Kauffman, G. et al., 2003, MNRAS, 341, 33
- Kauffman, G. et al., 2003, MNRAS, 346, 1055
- Kewley, L. J. et al., 2001, ApJ, 556, 121
- Kewley, L. J. et al., 2006, MNRAS, 372, 961
- Kormendy, J. & Kennicutt, R. J., 2004, ARA&A, 42, 603
- Kormendy, J. et al., 2010, ApJ, 723, 54
- MacKay, D. J. C., 2003, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, ISBN 978-0-521-64298-9
- Marasco, A., Fraternali, F. & Binney, J. J., 2012, MNRAS, 419, 1107
- Martin, D. C. et al., 2007, ApJS, 173, 342
- Masters, K. L. et al., 2010, MNRAS, 405, 783
- Masters, K. L. et al., 2011, MNRAS, 411, 2026
- Masters, K. L. et al., 2012, MNRAS, 424, 2180
- Mendez, A. J. et al., 2011, ApJ, 736, 110
- Nandra, K. et al., 2007, ApJ, 660, L11
- Noeske, K. G. et al., 2007, ApJ, 660, L43
- Oh, K. et al. 2015, arXiv: 1504.07247
- Peng, Y. et al., 2010, ApJ, 721, 193
- Saintonge, A. et al., 2012, ApJ, 758, 73
- Salim, S. et al., 2007, ApJSS, 173, 267
- Schawinski, et al., 2007, MNRAS, 382, 1415
- Schawinski, K. et al., 2010, MNRAS, 711, 284
- Schawinski, K. et al., 2014, MNRAS, 440, 889
- Sheth, K. et al., 2012, ApJ, 758, 136
- Silverman, J. D., et al. 2008, ApJ, 675, 1025
- Simmons, B. D. et al., 2013, MNRAS, 429, 2199
- Sivia, D. S., 1996, *Data Analysis: A Bayesian Tutorial*, Oxford University Press, ISBN 0-19-851889-7
- Springel, V., Di Matteo, T. & Hernquist, L., 2005, ApJ, 620, L79
- Smethurst, R. J. et al., 2015, MNRAS, 450, 435
- Tortora, C. et al., 2009, MNRAS, 369, 61
- Willett, K. et al., 2014, MNRAS, 435, 2835
- Willmer, C. N. A. et al., 2006, ApJ, 647, 853