General Questions

1. Can you summarise your PhD in 5 minutes?
2. Did you enjoy your PhD?
3. What do you think the main result and impact of your thesis will be?
4. If you had to describe your thesis in a tweet, what would you say?
5. What are your plans for the future?
6. If you had the time, how would you follow up on this project?
7. What were the main research questions you were hoping to address?
8. What does Stephen Serjeant do? What work has he authored?
9. How does your work relate to the literature?
10. What is the Effectiveness of your Research?
11. What are the Limitations of this Thesis?
12. Are there existing similar works in the literature?
13. How has your research challenged or changed a topic?
14. Other approaches in your research?
15. What are the next steps for this area of research?

Abstract

1. Describe what is meant by hierarchical model.
2. What do you mean by the relation between physical processes and underlying parameters?
3. Explain a Bayesian Convolutional Neural Network. What are it’s component parts?
4. You talk about active galactic nuclei here, what is it? Why does it matter with interaction?
5. What is the dynamical timescale of an interaction? Does it have well defined beginnings and endings?
6. What is a confidence interval?
7. You mention Simulation Based Inference. What is that?

Introduction

1. You talk about large scale structure, what do you mean? (page 1)
2. Describe cosmic time to me. (page 1)
3. Can you fully elaborate on what you define as mutual interaction? (page 1)
4. Can you explain redshift? (page 1)
5. Define what you mean by a tidal feature. (page 1)
6. How else could we make this reliable identification of interacting systems? (page 2)
7. What do you mean by exploring the relationship between parameters and characteristics of systems? (page 2)
8. What is a super massive black hole? Why are they often at the centre of galaxies? (page 2)
9. Why may there not be supermassive black holes at the galactic centre? (page 2)
10. Could you describe the different components of the galactic disks. For instance, what is the thick and thin disk? (page 3)
11. What are the classifications for bulges? What is the difference between them? (page 3)
12. Can you name the different galaxies? What is the primary differences between them? What about their evolutionary pathways, what can you tell me about them? (page 3)
13. You say that the instabilities in the early Universe were caused by “physical processes”, what do you mean specifically? Why were they caused by the small scale of the Universe? (page 4)
14. The Universe was expanding. Why? (page 4)
15. Why did the rate of mergers and galaxy interaction go down? Why is it lowest now compared to the past? (page 5)
16. As above, but for cosmic star formation. Why is it lowest now compared to the early Universe? (page 5)
17. You say that mergers play a significant role in star formation of the past, but don’t really back this up. Why do you say this? (page 5)
18. You say that elliptical galaxies are purely an effect of mergers. Are there not secular processes which lead to the formation of ellipticals? (page 7)
19. Can you further expand on what a stellar population is? (page 7)
20. How does a stellar population relate to the star formation rate? (page 7)
21. Can you describe the process of gas cloud fragmentation into forming stars? (page 7)
22. Does the ultra-violet fall in our blue filters? (page 7)
23. Can discuss further the different states of gas. Why is it only molecular gas that can fragment and form stars? (page 9)
24. How was the Kennicutt-Schmidt relation found? Why was the number found to be n = 1.3? (page 9)
25. Could you speculate about the evolutionary pathway of red disks and blue spirals? (page 11)
26. You say we measure the SFRs of these galaxies. Could you explain how we do this? (page 11)
27. Can you confirm your definition of a filament galaxy? (page 11)
28. How would you define these different environmental classifications? (page 11)
29. Can you describe the parameters you’re referring to that change interaction? (page 12)
30. Could you describe how, observationally, we define a wet, dry and mixed merger? (page 13)
31. Could you explain why wet mergers have an increased AGN fraction? (page 13)
32. Why would micro mergers be a driver of cosmic star formation compared to major ones? (page 14)
33. Can you describe the work of Toomre and Toomre further? Why was their work so ground-breaking? (page 15)
34. Can you tell me what violent relaxation is? (page 15)
35. Also, can you expand on the definition of dynamical friction? (page 15)
36. What do you mean by changing gravitational fields? (page 16)
37. Do we see increases in the star formation rate in tidal features of interacting galaxies? (page 19)
38. What is the long, complicated process of accretion by the super massive black hole? (page 19)
39. Could you break down all the different classifications for AGN for us, and explain the difference between them? (page 20)
40. What is AGN flickering? (page 20)
41. Why would a delay in the AGN activation make sense? What would cause the delay? (page 20)
42. What other mechanisms could be in the AGN itself causing suppression? (page 20)
43. What are AGN winds and how do they blow out gas in the galaxy? (page 21)
44. Can you describe the structure of a neural network? How does it actually work? (page 23)
45. How are such classifications of a set actually made? (page 23)
46. What’s different about a Convolutional Neural Network and a regular neural network? (page 24)
47. Where does the inconsistency with training a neural network on simulations and applying to observations come from? (page 24)
48. Was it a good idea to name everything Chapter? (page 25)

Chapter 1: Zoobot

1. What are the debates about the fueling of AGN? (page 28)
2. What more could be done to remove contamination by close pairs? (page 30)
3. Could you describe how ESA Datalabs works? (page 31)
4. What other ways will ESA Datalabs impact the field? (page 31)
5. What future work would you do with ESA Datalabs? (page 31)
6. Why did you choose to use ACS, *F814W*, etc for your dataset? (page 31)
7. What time is ESA Datalabs actually saving you? Why can’t you just use TAP services to download the cutouts? (page 31)
8. Could you further discuss the applicability of the Shapely Python package? (page 32)
9. What do you mean by affects of interpolation? (page 32)
10. Describe more the functionality of Zoobot. How does it specifically work? (page 33)
11. Explain representation learning, as I’m unsure what you specifically mean here? (page 34)
12. In the original W+22 work, they used flattened 3-colour images. How does using only a single-band image affect your classifications? (page 34)
13. If you are not using the prediction score as a probability score here, then what does it mean when you have a cutoff of 0.95? (page 34)
14. Why do we need a smaller training set size when we are conducting finetuning? (page 36)
15. Why are the image contrasts changing with source size? What affect might this have had on training? (page 38)
16. In general, what is active learning? (page 40)
17. Is conducting this image augmentation and adding to your training set valid? What impacts does this have on your accuracy? (page 41)
18. Why is your validation set not completely balanced as well? (page 41)
19. What effect does this bi-modality tell you about your results? (page 41)
20. What effects on your results would be reducing you 0.95 cutoff? 0.95 does seem very stringent? (page 43)
21. If your value does not correspond to probability, what does it actually mean when you have a cutoff of 0.95? (page 43)
22. What do you mean by ‘despite removing 50% of the catalogue’? Do you think you would be finding 63 million interacting galaxies? (page 43)
23. As you say, using a balanced dataset leads you to be biased towards classifying a galaxy as interacting. Why didn’t you use an un-balanced dataset? (page 46)
24. Can you extend what you say about hierarchical clustering? How does it actually work? (page 47)
25. Describe Euclidean Linkage. Why didn’t you use a different linkage system? (page 47)
26. You talk about representation learning. What is this? (page 48)
27. Explain Principal Component Analysis. (page 48)
28. How do the representations relate to the morphology of the galaxy? (page 48)
29. Explain an AutoEncoder. What is it doing and how does it take the 40-dimensional representation and reduce it to a 2-dimensional projection? (page 49)
30. What would happen to your results if you changed your X and Y mappings? Why did you select these specific values? (page 49)
31. Why would each of these sources be in the HSC, but not in any other archives? (page 53)
32. Does a source not existing in Simbad, ViZier or NED really mean that it’s unknown? What other databases exist where these could be within them? (page 53)
33. What is bootstrapping? How did you actually conduct this here? (page 55)
34. Why was there final contamination at the end? Why not remove it using further visual classification? (page 55)
35. You’ve mentioned them, so I’m going to ask. What are: submillimeter galaxies, quasars, jellyfish galaxies, galactic jets, gravitational lenses, Lyman-alpha emitters, transitional stellar objects and supernova remnants? (page 57)
36. 5” for a match seems rather large… Why did you use such a wide criteria to match to Simbad or ViZier? (page 57)
37. Have you done the further work in order to confirm the classifications of these extra gem systems requiring multi-wavelength data? (page 57)
38. You are looking for hard and soft X-Ray emission for AGN. What’s the difference? What do they tell us about the system being observed? (page 57)
39. What do you mean by heterogeneous selection and analysis procedures? (page 58)
40. Why would you expect a second locus in this parameter space? How can you be so sure you’re not just picking up high starforming objects due to the *F814W* filters dependence on the UV emission? (page 62)
41. Explain Figure 2.13. What are you talking about? What does this actually show that we would expect about this sample? (page 64)
42. What value would you use to split the red and blue sample here? For each panel? (page 64)
43. In your conclusions, you don’t really discuss that these are morphologically identified interacting galaxies. Why? (page 65)

Chapter 2: Mergers in COSMOS

1. In your Introduction, you define the dynamical timescale. Is this actually correct? (page 67)
2. Can you define the project separation in this context? (page 67)
3. Why do you not use the 3D separation between your systems, and map this out? (page 67)
4. Which COSMOS2020 catalogue did you use? Why this one? (page 69)
5. What is broadband photometry? (page 70)
6. How is it used to measure these ancillary parameters? (page 70)
7. If at redshift 1.2, you are saying that tidal features are difficult to identify, how could Zoobot in the previous chapter? (page 70)
8. You choose limits of stellar mass down to 6.5 and star formation rate down to -5. Is this not incredibly low? Can a galaxy actually have a stellar mass down to 10^6.5? (page 70)
9. Why do degeneracies appear in the plots when you use the same software? What does this tell you about your results? (page 71)
10. What are the limitations of the visual approach taken in the COSMOS project? (page 73)
11. What could you do differently to find more secondaries? (page 73)
12. The large % of your catalogue not having an identified secondary is incredible surprising… Expand on why you think this is acceptable. (page 73)
13. You use a very large separation to identify your galaxies. Why did you do this? (page 73)
14. Are the fractions of secondaries with stage expected? Break down each one. (page 74)
15. Could you explain what you mean by de-duplication? (page 75)
16. What did you do to check that your control galaxy wasn’t interacting? (page 75)
17. When you are finding additional interacting systems, why were they not picked up by your catalogue? (page 75)
18. What do you mean the mass-limited sample gives you uniform sensitivity across the volume? (page 77)
19. How does the environment affect the SFR within a galaxy? Why could it match an interacting galaxy? (page 79)
20. How would you go about breaking this degeneracy in the dynamical timescale? (page 80)
21. Are you biased in the inclination of your systems? (page 81)
22. Your COSMOS cutouts are 30” by 30”. What distance would the secondary have moved to not be within the cutout anymore? (page 82)
23. Why would post-mergers be classified in your merger stage if a key criteria is that a secondary must be present? (page 82)
24. Why does a galaxy in a cluster environment have, on average, a higher SFR than a field galaxy? (page 86)
25. Could you describe the different methods of measuring the galactic environment? Why did you stick with Darvish et al’s method? (page 86)
26. What are you talking about with bootstrapping and approximating the selection effects in your paired sample? (page 90 - 91)
27. Could you describe clumpy galaxies to us? As you have written about it. (page 97)
28. Could you expand on the relations you expect to uncover from the different mass ratios? How would this affect your results if you were dominated by major interactions rather than micro or minor ones? (page 99)
29. In your description of your samples, you have jumped between the mass and volume- limited samples. You need to summarise this better in Chapter 3. (page 99)

Chapter 3: Inferring Galactic Parameters

1. Place Holder

Conclusions

1. Place Holder