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.
   1. This is a model where smaller components combine and merge into larger ones. This happens over many times in a hierarchy of merging and accretion. In the context of the abstract, I mean LCDM which has within it a hierarchical model of galaxy evolution.
   2. I do not mean hierarchical observations. I mean that ‘hierarchical models’ and theories confirmed by observations.
2. What do you mean by the relation between physical processes and underlying parameters?
   1. An underlying process, in this context, is some process that leads from A to B. For example, star formation is an underlying process in a galaxy. These are then related to underlying parameters, for example the gas mass within the galaxy. Both can be parameterized, but one is describing movement of one thing to another while the other is simply a descriptive number.
3. Explain a Bayesian Convolutional Neural Network. What are its component parts?
   1. A Bayesian Convolutional Neural Network in this context is a deep neural network which is applied to images. A convolutional neural network takes an image, and breaks it down into smaller and more abstract parts. It passes each part, a representation, through many different layers which is altered by a weight which can be applied in a range of functions. These are passed through layer after layer until the final abstraction is then interpreted by a classification layer. However, in a Bayesian CNN, each weight is not simply a number which has been learned, it is a prior distribution of some kind. As we move through the BCNN, each distribution gives each abstraction a posterior distribution that can then be output in the final classification stage. Thus, we get our final answer but also a level of uncertainty within it.
4. You talk about active galactic nuclei here, what is it? Why does it matter with interaction?
   1. An active galactic nucleus is, essentially, the supermassive blackhole at the galactic centre which is growing. It has got an accretion disk around about it, which is incredibly hot, emitting high powered radiation and jets perpendicular to the plane of accretion. The high powered emission acts as a wind which goes through the galactic disk, driving outflows and causing ionization. This, in turn, leads to the cessation of star formation in the galaxy. Thus, the SMBH at the galactic centre is linked to the evolution of the galaxy as a whole. We call this co-evolution, and we see a distinct match between black hole mass and a host of galaxy properties. Most closely, the bulge and stellar mass of the galaxies.

One way in which AGN could be ignited is through galaxy interaction, and the rush of gas into the galactic core being the perfect way for material to accrete onto the black hole.

1. What is the dynamical timescale of an interaction? Does it have well defined beginnings and endings?
   1. The dynamical time, or the way I describe it, begins at the moment two galaxies are within each others dark matter haloes and dynamical friction is beginning to occur. It then ends when either the two galaxies have merged or they have escaped each others halos.
2. What is a confidence interval?
   1. A confidence interval is a measure of uncertainty about a defined parameter using Bayesian statistics. In a confidence interval, the true value of the parameter could be anywhere. It is often described by the percentiles about a value defined by assuming a statistical distribution. I’m worried that, in fact, I do mean a credible interval in this work and not a confidence interval.

A credible interval is measured using the mass of walker steps, as is done in my final Chapter. However, I have drawn the percentiles which dictate the different confidence interval, not the credible interval.

I have used credible intervals, not confidence intervals. Therefore, I need to re-write this part.

1. You mention Simulation Based Inference. What is that?
   1. Simulation based inference is a potential answer to the high computational cost of the MCMC approach described in this work. While I won’t really go into depth here, it involves front-loading the computational cost. We run our simulation across the parameter space, creating thousands of examples of images. These are then learned by an emulator or machine learning algorithm. This emulator then approximates the full probability space. It then can be applied to an observation, and recover the parameters of the observation.

Introduction

1. You talk about large scale structure, what do you mean? (page 1)
   1. The large scale structure is the overall filamentary structure we observe in the Universe. We get these nice redshift maps which show the filamentary structure of the universe on a grand scale. At a smaller scale, we then get over- and under-dense environments. Galaxy clusters, nodes, voids and the field. Each of these are made up, fundamentally, of galaxies.
2. Describe cosmic time to me. (page 1)
   1. Cosmic time describes the evolution of galaxies and the Universe’ age. It is from the Big Bang until now, and the Universe has expanded, changed and evolved with it. Cosmic time goes with the Hubble Flow, and assumes that the Universe is the same density everywhere at all times.
3. Can you fully elaborate on what you define as mutual interaction? (page 1)
   1. Mutual interaction is when two galaxies have a gravitational effect upon one another. This can be from dynamical friction, to the complete disturbance of their disks via a flyby.
4. Can you explain redshift? (page 1)
   1. Redshift is a result of light moving through an expanding Universe. As the universe expands, the distance between coordinates increases. This leads to a stretching in the wavelength of the light – a reddening effect of it. When applied to a galaxy, we see the internal spectrum being shifted by some amount in wavelength. As we know the true spectrum, and the speed of light, we can convert the redshift measurements into distances to the galaxies.
5. Define what you mean by a tidal feature. (page 1)
   1. A tidal feature is a structure in or about a galaxy which has been created by the effect of another galaxies gravity. For instance, as two galaxies flyby each other, they can shear off the outer regions of a galaxy and form a tidal tail and arm.
6. How else could we make this reliable identification of interacting systems? (page 2)
   1. Really, what we need is some measure of the distance towards them to show that they are truly at the same distance from us. We could potentially try to model them with simulations, and recreate their morphology. However, this may not be definitive. We could look for signs of increased star formation and tidal disruption. However, if we’re getting that we might as well measure the redshift of the system. Same with the velocity. What we really do need is the redshift more than anything.
7. What do you mean by exploring the relationship between parameters and characteristics of systems? (page 2)
   1. It’s sort of as I’ve said before. Interacting systems are characterized by increased star formation rates, potentially increased AGN fractions and tidal disruption in their disks. These are physical processes occurring within the galaxies that lead to these differences. I want to then map out the parameters which make this happen. Whether that be how long the two galaxies have been interacting for, the stellar masses of the systems, and the orientation angles of the interaction.
8. What is a super massive black hole? Why are they often at the centre of galaxies? (page 2)
   1. A supermassive blackhole is a black hole with a mass greater than 10^5 solar masses. Significantly larger than a stellar black holes (5 – 10 times solar mass). They are often, but not always, found at the centre of the galaxy. This is due to dynamical friction, where large mass objects sink in the gravitational potential due to friction with lots of smaller mass objects. Thus, we do no know if galaxies formed about SMBHs, or if galaxies formed with SMBHs within them, that then began to move into the galactic core due to dynamical friction.
9. Why may there not be supermassive black holes at the galactic centre? (page 2)
   1. This is a bit of a tricky one. The paper I’ve cited here is not saying there is no black hole at the centre of M33, but that using the known co-evolution arrangements between central velocity dispersion and black hole mass yields a mass between 0 and 1,500 solar masses. Therefore, there is something there, but there is clearly a space at the lower end of the co-evolution relation which we do not yet understand. They also do not really comment on why one might not be present. One could be there used to be many there, and they merged / interacted and ejected each other. However, these authors do not comment on this.
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 explain the KS and AD tests? What’s the difference? What does each test actually measure? (page 94)
28. Could you describe clumpy galaxies to us? As you have written about it. (page 97)
29. 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)
30. 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)
31. Does the weighting scheme you have used in stellar mass and SFR keep your results valid? (page 101)
32. How does balancing based on the stellar mass lead you to have equal number counts in each bin? (page 101)
33. Could you summarise the error analysis you used from Cameron (2011)? What are you talking about with the beta function? (page 105)
34. Fully break down and explain equation 3.1. You say it’s from Aird et al (2019), but you do not explain the fundamentals of this. (page 106)
35. Are these results what are expected for the different stages of interaction? If so, why? (page 110)
36. Why have you not used the confirmed merging galaxies in your pair sample? Could these not be assumed to be all at 0kpc separation? (page 110)
37. You change your bin widths through the projected separation space. How does this affect your results? Is this something you can legitimately do? (page 110)
38. How does taking the average measure the excess of the SFR from interaction? (page 110)
39. What about only the pericenter and apocenter measurements on Figure 3.21? (page 111)
40. How did you apply the Cameron (2011) methodology with respect to the projected separation? (page 112)
41. Is the projected separation distribution with SFE for the full sample as expected? Why? (page 112)
42. Should you not say something here about the timelines of the AGN and interaction? (page 113)
43. Your weighting scheme in the AGN fractions is not clear… Did I actually do this two step fraction? (page 113)
44. Do we expect that the AGN fraction in the approaching pair phase is actually supposed to be this high? (page 113)
45. How do the low counts in this sample affect your results? Would you expect these results to remain consistent with a higher number counts? (page 113)
46. What other tests could be conducted which lead to a more optimal examination of these low count AGN results? (page 113)
47. Explain the physical mechanisms behind a delayed AGN ignition. What’s going on there? (page 116)
48. Could you define what you mean by a delay in AGN ignition? Surely this is just because you have no mergers in this sample? (page 118)
49. Could you expand on the change in the SFR other works have observed? Do these often use observations or simulations? (page 119)
50. You argue from your results that you have found evidence of no starburst, but actually of a slow increase in star formation rate with time. How are you able to do this? Why do your results reflect this conclusion? (page 119)
51. I’m going to need you to really expand on the small increase in the star formation rates over a longer time on page 120. How do you get this from your results? (page 120)
52. Is this idea of long term decline of SFE in escaped galaxies and quick decline of SFE in merging galaxies supported in the literature? (page 120)
53. How do you know that post-merger galaxies are quenched? You haven’t got many in your sample? (page 120)
54. You said you might have contamination from post-merger galaxies in your final sample. Doesn’t this mean that you’re finding high star formation rates in these as well? (page 120)
55. Please explain the difference between delayed AGN ignition and AGN flickering? (page 121)
56. What is a delayed AGN? (page 121)
57. What is AGN flickering? (page 121)
58. Using the projected separation of the interaction, you make a temporal argument for AGN flickering or delay. Could you explain how you made this leap? (page 122)
59. Could you explore, for a moment, what would occur if the two galaxies were overlapping and merging? Would your measure of star formation and mass change? (page 123)
60. How significant is finding that the projected separation is not a perfect proxy for stage? Should we use something else besides interaction stage? (page 124)

Chapter 3: Inferring Galactic Parameters

1. Can you describe in a bit more detail cosmological simulations? How do they work? (page 127)
2. Can you expand on and describe the Galaxy Zoo: Mergers process? (page 127)
3. You randomly bring up the Vera C. Rubin observatory here. Why? Why does the age of the Vera C. Rubin observatory matter? (page 127)
4. You say samples of thousands of interacting galaxies are going to be found a week with LSST. Is this really true? (page 127)
5. What is a probability distribution? (page 128)
6. You discard 8 of these systems because they are not in SDSS. Can you not just use the best fit underlying parameters that GZM found and incorporate these into your sample? (page 129)
7. Also, could you not simply assume the redshift of these systems and gain another 3 systems in your sample? (page 129)
8. What effects on the morphology of the system would combining the white images be? (page 129)
9. Why did you conduct the block reduce to 100 x 100 pixels? (page 129)
10. Could you provide a summary of the JSPAM code for us? Why is it very efficient and fast? (page 131)
11. Define a reasonable time resolution. (page 132)
12. Could you explain a fourth order Runge-Kutta methodology? (page 132)
13. What is a softened point mass approximation? (page 132)
14. In this context, what is an N-body approximation? (page 132)
15. How does the force change based on these approximations? (page 132)
16. You say you elect a softened-point mass over the N-body approximation for computational efficiency. How different are the different approximations? Are then any other physical reasons that you might choose one over the other? (page 132)
17. You mention genetic algorithms. What are these? (page 132)
18. How would a genetic algorithm be used in the context of finding the parameters of interacting galaxies? (page 132)
19. What are semi-analytic models? (page 133)
20. You state you use a Bruzual & Charlot simple stellar population. What is this? What does it look like? (page 133)
21. You then use a Chabrier IMF. How do you use it specifically? What is the initial mass function in general? (page 133)
22. What is the e-folding timescale? (page 133)
23. What is a spectral energy distribution? (page 133)
24. You say this matches a bunch of works. Why does it have this value of about 1.7 Gyrs? (page 133)
25. You define a gas fraction in these galaxies as 0.15. Is this reasonable? Where did this number come from? (page 134)
26. Is the initial galaxy age of 10Gyrs reasonable? (page 134)
27. You say you model new stellar populations as the interaction continues. What do you mean? How do you do that? (page 134)
28. How is your semi-analytic SFR enhancement similar to that done in CIGALE? (page 134)
29. What is CIGALE? (page 134)
30. Does your enhancement parameter have any basis in the literature? It seems rather arbitrary. (page 135)
31. You define an SFR\_{galaxy} in equation 4.3, but do not define this. (page 136)
32. Why have you not defined a gas distribution model? (page 136)
33. Can you show that your assumptions do mimic the simulations direct calculations? (page 136)
34. What are the limitations of removing particles with no neighbours to zero? (page 138)
35. Your mock observations appear to still have holes in their tidal features. So, did you flux distribution algorithm not work? (page 142)
36. Why did you opt to use MCMC here? And not something like, say, simulation based inference? (page 143)
37. Please explain how a typical MCMC would work. (page 145)
38. What is a posterior distribution? (page 145)
39. What is a confidence interval? (page 145)
40. Please explain how emcee works in the context of your work. (page 145)
41. Please explain what a Differential and Differential Evolution move is? How do these differ from a typical walker move? (page 145)
42. What motivated the choice of moves for the walkers? (page 145)
43. Why are these moves good for high dimensional parameter space? (page 145)
44. Define a prior. (page 145)
45. What motivated the decision to have uniform priors? (page 145)
46. By adding to the prior as you describe, it sounds like you do not have uniform priors? (page 146)
47. Explain further this constant C, in equation 4.6. (page 147)
48. You state that you find the underlying probability distribution is not Gaussian. Therefore, how does this affect your results? Could making another decision here lead to a better outcome? (page 147)
49. You bring up GALFIT here, what is that? What’s it for? (page 147)
50. Why does taking the log help with computation time? (page 148)
51. How should I publish all the probability distributions? Currently, on a GoogleDrive. (page 148)
52. Why does the best fit and observation images look so different in Figure 4.2? (page 149)
53. What is the Corner Python package? How does it work? (page 150)
54. Why do we find degeneracy in the orientations? Why is it a 4-fold degeneracy? (page 152)
55. Why did you use the most massive system in your sample as your example? (page 152)
56. Why is the mass the easiest to constrain? (page 152)
57. Are there other approaches you could have taken besides a chi-squared approach? Why weren’t these considered? (page 152)
58. Would this movement of the secondary in the image actually change the errors you have? (page 153)
59. Why does the secondary disk contribute more to the tidal debris? (page 153)
60. Why does the simulation being in the frame of the primary mean that the 3D velocity of the secondary is more? (page 153)
61. What are you talking about with the backward integration? It’s not very clear. (page 155)
62. Why does the simulation give non-physical results when we start at or very close to pericenter? (page 155)
63. Is your time parameter actually within this confidence interval? This does not seem constrained at all. (page 155)
64. Why does the skewing of the time parameter to small values seem excellent compared to the skewing of the position and velocity parameters? (page 155)
65. Is your gamma parameter really necessary? If it is preventing you finding the best fit parameters of these galaxies, then surely it is a hindrance? (page 156)
66. What do you mean by changing each parameter by a small amount based on the posterior? (page 156)
67. You say that your results appear to have converged. Is this what you really mean? (page 157)
68. You’ve described the Geweke diagnostic. How does it actually work in practice in ChainConsumer? (page 157)
69. Why does the Geweke diagnostic work? What does comparing the first 10% and the last50% of each chain tell us? (page 157)
70. Why does the degeneracies in the orientation space lead to a disruption of the Geweke diagnostic? (page 157)
71. What does it mean to flatten and thin the walker chain? What does this actually do to it? (page 159)
72. What effect would throwing away more steps have on your results? What about less? (page 159)
73. What does it mean to ‘burn-in’ the walker chain? (page 159)
74. Is assuming the circular velocity at different radii not enough to ascertain the rotational velocity? What do you think the rotational velocity is? (page 161)
75. Do you expect to do anything with WEAVE in the future on this project? (page 162)
76. What are IFUs? How do they work? (page 162)
77. Why is your ability to constrain dependent on the stage of the interaction? (page 162)
78. Is it a bit of a cop out to discuss Arp 240 again in this section? (page 162)
79. Why do all of you mass probability distributions have a banana shape? (page 162)
80. What do you mean by stage 2, 3 and 4 here? (page 162)
81. You’re saying that two overlapping disks is the same as an ongoing merger or merger remnant? That doesn’t sound right… (page 162)
82. You seem to consistently find under-estimates for the time and mass parameters. Why do you think this might be? (page 165)
83. What further investigation have you made into exploring a more representative parameter space in the velocity? (page 167)
84. What do you mean by folding the parameter space? (page 167)
85. This is a very insightful discussion on the outputs of your simulation. So, how could we go about fixing this? Is it reasonable in a small timescale? How could you account for the over-estimates due to the ring-distribution of your particles in the galaxy? (page 167)
86. Why would the systems require a high beta parameter to increase the time of the closest approach? (page 168)
87. Where does this discrepancy between the scales of the H16 images and your observations come from? (page 169)
88. How are the positions of the secondary galaxies different between H16 and the observations? (page 169)
89. Why does having noise surrounding the galaxy lead to higher uncertainty in the radial probability distribution? (page 171)
90. How do you make a conversion of total to stellar mass in the Arp 240 best fit simulation? (page 171)
91. What are the errors within the He et al. (2020) paper for their measurements on the Arp 240 stellar mass? (page 171)
92. What about the other parameters of Arp 240? Can you make any further comparisons? (page 171)
93. Why do you say that some systems may simply not be able to be recreated with restricted N-body simulations? (page 174)
94. In your observational corner plots, there almost seems like a lot of binomial distributions. What’s going on there? (page 172 - 173)
95. What would happen if your MCMC tried to constrain an image at the wrong resolution? (page 174)
96. You say your MCMC needs a reasonable number of degrees of freedom to reach convergence. What do you mean? (page 175)
97. Explain different methods of improving computational efficiency. You talk about GPUs, but why would they actually achieve increased computational efficiency? (page 176)
98. In your conclusions, you say you had to scale up the observed images. What do you mean by this? (page 178)
99. You say in your conclusions that this took thirty days, but in the main text forty days. Which is it? (page 178)
100. You talk about gaussian processes, machine learning and simulation based inference in the conclusions of this Chapter. What are these? Could you expand on how these could help you? (page 178)

Conclusions

1. Do conflicting results emerge from using projected separation and morphological stage with projected separation? You found the same overall behavior with project separation anyway? (page 180)
2. Is it confirmation of AGN flickering? Please explain this further. (page 180)
3. What is amortization? (page 184)
4. Classic question, you explain the next steps of the development of the MCMC algorithm. Are you actually going to carry them out? (page 184)