Institute of Cosmology and Gravitation

Human-Guided Four-Component Photometric Modelling of Spiral Galaxies

by

Timothy Kenneth Lingard

supervised by

Dr Coleman Krawczyk, Prof Karen Masters, Prof Robert Nichol

This thesis is submitted in partial fulfilment of
the requirements for the award of the degree of
Doctor of Philosophy of the University of Portsmouth.

October 25, 2020

Abstract

Multi-component modelling of galaxies is a valuable tool in the effort to quantitatively understand galaxy evolution, yet the use of the technique is plagued by issues of convergence, model selection and parameter degeneracies. These issues limit its application over large samples to the simplest models, with complex models being applied only to very small samples. I have attempted to resolve this dilemma of "quantity or quality" by developing a novel framework, *Galaxy Builder*, built inside the Zooniverse citizen science platform, to enable the crowdsourcing of complex photometric model creation (containing a disc, bulge, bar and spiral arms) for Sloan Digitial Sky Survey galaxies.

I have applied the method, including a final algorithmic optimisation step, on a sample of 198 galaxies, and examined its internal robustness using a small sample of synthetic galaxies and a repeated validation sample. I also compare its results to automated fitting pipelines, demonstrating that it is possible to consistently recover accurate models that either show good agreement with, or improve on, prior work.

I have made use of the crowdsourced spiral annotations from the *Galaxy Builder* project in a hierarchical Bayesian model to examine the relationship between spiral tightness (pitch angle) and central morphology (bulge strength, bar presence and strength), finding that central morphology does not significantly impact spiral pitch angle. I also made use of this Bayesian framework to test a simple model of spiral winding, finding support for the picture of spiral arms as transient and recurrent disc instabilities.

I conclude that citizen science is a promising technique for modelling images of complex galaxies, and release our catalogue of models.

Table of Contents

List of Tables

List of Figures

Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word count: 24,212 words.

Acknowledgements

The past few years have allowed me to work on a fantastic project, with a wonderfully supportive supervisory team. I will be forever grateful to Prof Karen Masters and Dr Coleman Krawczyk, who have been patient, understanding, honest, empathic and nothing short of brilliant. Without your expertise, understanding and mastery of mentorship, none of this work would have been possible.

I have had the good fortune to share office space with many enthusiastic, intelligent, kind people during my time at the ICG, and I am grateful to all of them for their good-humour, insightful observations, sympathetic ears and willingness to come for a half-hour tea & cake break at the drop of a hat. I wish you all the best with your future endeavours and am excited to see what you will achieve.

My family, dearest friends, and amazing partner have been instrumental in my continued sanity. I look forward to sharing a drink (or many) in the coming years, and reminisce over the mad endeavour that was this PhD.

The Galaxy Zoo and Zooniverse teams have been wonderfully insightful: helping with both the design and development of *Galaxy Builder* and by providing insight on improving the scientific exploitation of its results. I wish to send you all a heartfelt thanks, and look forward to the exciting times the field has coming!

As with most PhD studentships, there have been times where throwing in the towel seemed like the most sensible course of action, but the support of those mentioned above has been invaluable in helping me persevere and

create a piece of work I am immensely (self-indulgently) proud of.

I would like to thank the referee of Lingard et al. (2020a) for their input; suggesting numerous ways to improve the clarity and content of the paper, and providing avenues of discussion that had not previously been covered. I would also like to thank Dr Ross Hart for providing the target catalogue for the *stellar mass-complete sample*.

Lastly, a citizen science project is nothing without the volunteers kind enough to give up their spare time to help with it. I am very much in debt to the 2,340 volunteers who have submitted classifications to the *Galaxy Builder* project, especially user Elisabeth Baeten, whose presence on the *Galaxy Builder* forum on top of a large number of galaxies modelled was of immeasurable help.

This publication made use of SDSS-I/II data. Funding for the SDSS and SDSS-II was provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is `www.sdss.org/`. This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation.

This project was partially funded by a Google Faculty Research Award[1] to Karen Masters, and Timothy Lingard acknowledges studentship funding from the Science and Technology Facilities Council (ST/N504245/1).

---

[1] `ai.google/research/outreach/faculty-research-awards/`

Dissemination

Papers, talks and posters.

1. Poster presentation at the University of Portsmouth Research and Innovation conference, 2017.
2. Tuesday talk at the ICG, 2018.
3. Poster presentation at the European Week of Astronomy and Space Science, 2018.
4. Two Galaxy Zoo telecon presentations (2019, 2020).
5. Presentation at the MaNGA 2019 Conference.
6. Technical presentation to the Zooniverse team, 2019.
7. Galaxy Zoo Builder: Four Component Photometric decomposition of Spiral Galaxies Guided by Citizen Science (Lingard et al., 2020a), accepted for publication in the Astrphysical Journal, 2020.
8. Galaxy Zoo Builder: Morphological Dependence of Spiral Galaxy pitch angle (Lingard et al., 2020b), submitted to the Monthly Notices of the Royal Astronomical Society, 2020.

Nomenclature

χ^2_ν	Reduced Chi-squared
ΛCDM	Lambda Cold Dark Matter
°	Degree, measurement of a plane angle
API	Application Programming Interface
ApJ	The Astrophysical Journal
APO	Apache Point Observatory
APOGEE	Apache Point Observatory Galactic Evolution Experiment
BFGS alogirthm	Broyden-Fletcher-Goldfarb-Shanno alogirthm
BIC	Bayesian Information Criterion
CANDELS	Cosmic And Near-infrared Deep Extragalactic Legacy Survey
CNN	Convolutional Neural Network
DEC	Declination (J2000)
DES	Dark Energy Survey
EAGLE	Evolution and Assembly of GaLaxies and their Environments
ESO	European Southern Observatory
FFT	Fast Fourier Transform

GZ2	Galaxy Zoo 2
HI	Neutral Hydrogen
HTML	Hypertext Markup Language
ICG	The Institute of Cosmology and Gravitation, University of Portsmouth
IFU	Integral Field Unit
LSST	Legacy Survey of Space and Time (The Vera C. Rubin Observatory Project)
MaNGA	Mapping Nearby Galaxies at APO
MNRAS	Monthly Notices of the Royal Astronomical Society
MWA	Milky Way analogue
NN	Artificial Neural Network
NSA	The NASA-Sloan Atlas
P2DFFT	Parallel 2D Fast Fourier Transform
PCA	Principal Component Analysis
PSF	Point-spread function
QSDW theory	Quasi-Stationary Density Wave Theory
RA	Right ascension (J2000)
SAURON	Spectroscopic Areal Unit for Research on Optical Nebulae
SDSS	Sloan Digital Sky Survey
SGD	Stochastic Gradient Descent
SWAP	`Space Warps` analysis pipeline

WASER	Wave Amplification by Stimulated Emission of Radiation
WebGL	Web Graphics Library

"It's still magic even if you know how it's done"

Terry Pratchett, A Hat Full of Sky

Chapter 1

A Hitchhiker's guide to spiral galaxies

When most non-astronomers talk about a galaxy, they envision something akin to our own Milky Way: a beautiful swirling mass of stars and gas twinkling somewhere in the cosmos. Two-thirds of galaxies in the low-redshift universe do indeed show this spiral structure (Lintott et al., 2008), and it is in these galaxies that the vast majority of new stars are formed (Kennicutt, 1998).

Originally thought to be nebulae of gas and stars inside our own galaxy, we now know that our universe is host to at least 2 trillion of these galaxies (Conselice et al., 2016). A galaxy is a collection of stars, gas and dark matter, bound together by gravity and driven in its evolution by both complex internal dynamics and interactions with other massive objects (Buta, 2013). Galaxy structure is extremely variable, ranging from blob-like smooth galaxies (known as "early-types") to flattened, spinning discs with (or without) well-defined structures like bulges, bars and spiral arms ("late-types").

This chapter details the current understanding of how late-type galaxies are formed and what physical processes contribute to some of their distinct sub-structure, with a focus on the creation and maintenance of spiral arms. It presents a discussion of the different methods by which galaxies are classified into their morphological type, and efforts underway to leverage machine learning techniques to improve the scalability of morphological classification to next-generation surveys like the Vera C. Rubin Observatory Project (Ivezić et al., 2019) and *Euclid* (Laureijs et al. 2011; Amiaux et al. 2012). It introduces the tool of galaxy photometric modelling, whereby light from physically distinct galaxy sub-components are fit separately using analytical models, and details the issues faced with modelling spiral galaxies.

1.1 The formation of late-type galaxies

The fundamental physical processes governing how galaxies form, grow and evolve is still under much investigation (see Naab & Ostriker 2017 and Vogelsberger et al. 2020 for recent reviews). The timescales and distances involved are cosmological, and therefore the processes involved are greatly influenced by the cosmological framework used. These timescales also make it impossible to constrain theories using longitudinal studies; the entire lifespan of humanity is less than one-thousandth of the Milky Way's rotational period. Numerical simulations are therefore crucial in determining the importance of different physical processes, and what conditions are needed to produce the wonderfully diverse population of galaxies observed in the Universe.

Originally, galaxy formation was envisioned as occurring when colossal sheets of matter collapse and condense into protogalaxies (Eggen et al., 1962). This collapse would be counteracted by the rotation of a clump, creating a thin, spinning disc which would begin to form stars due to its increased density. This "monolithic collapse" model made predictions about differing orbits of generations of stars (original stars would be on more eccentric orbits), which at the time agreed with observations.

A decade and a half later, new observations about the stars in the outer reaches of our galaxy led researchers to a more chaotic method of galaxy formation. In this new model, small primordial overdensities create many tiny "dark" objects, which then interact and merge with each other through hierarchical clustering in a process which gradually builds the large-scale structure we observe today (Searle & Zinn 1978; White & Rees 1978). This mechanism is supported by the cold dark matter model augmented with a dark energy field (i.e. ΛCDM), and has become the standard model for galaxy formation over the past few decades (Naab & Ostriker, 2017) and is strongly supported by simulations (Springel et al., 2005).

The first step in the bottom-up formation mechanism is the creation of large dark matter halos, either through violent merger events or slower growth. If the formation process is slow, gas trapped inside the halos will collapse to form a spinning disc with a surface density mainly dependent on the rotation of the halo (Firmani & Avila-Reese, 2003). The increased gas density will allow star formation to begin, with most star formation in discs formed this way being driven by internal fluctuations and the infall of gas.

In the case of more violent major mergers, dramatic changes to morphology and kinematics of the galaxies will ensue (Barnes & Hernquist, 1996). Simulations indicate that major merger events can result in large, potentially warped discs, often with counter-rotating cores and similarly complex kinematics (Barnes, 2002). Merger events also have the effect of heating the stellar disc, transferring angular momentum and causing bursts of clumpy star formation due to the creation of pockets of very dense gas (Powell et al., 2013). Gas that has undergone strong shocks may form a central galactic bulge, while gas that has remained undisturbed can form an extended disc. Continued accretion from the remnants of the merger feed more angular-momentum rich gas onto the outskirts of the galaxy.

1.1.1 Secular evolution

The evolution of a galaxy is governed not only by the rate at which it acquires material from its environment, but also by its internal processes. Supernova(e) can cause galactic winds, allowing gas to escape from low-mass galaxies and suppressing their further growth (Vogelsberger et al., 2020). Similarly, active galactic nuclei may expel material from very high-mass galaxies limiting their formation (Silk & Mamon, 2012). This is illustrated by the difference between the observed galaxy mass function and the cosmic baryon fraction line in Figure 1.1. Modern simulations are excellent at recovering the true galaxy mass function, with the EAGLE project (Schaye et al., 2015) most accurately replicating the galaxy mass function of Li & White (2009) out of the simulations from Figure 1.1.

Kormendy & Kennicutt (2004) divide the evolution of galaxies into four categories (Figure 1.2):

1. Internal secular evolution driven by bar instabilities, dark matter halos, spiral structure etc...
2. Environmental secular evoution caused by gas infall, minor mergers and small-scale tidal interaction.
3. Rapid protogalactic collapse.
4. Galaxy mergers and ram-pressure stripping.

It has been argued that the Universe is currently in a state of transition from primarily interaction-based evolution to primarily isolated, secular evolution (Kormendy & Kennicutt, 2004). In the extreme future, galaxy evolution will occur purely via the interactions that occur inside a galaxy to redistribute matter. In an expanding Universe, this transition from chaotic, merger driven evolution to slower secular is inevitable, as the spacing between galaxies increases (though galaxy groups may persist).

1.2 The structure of late-type galaxies

Despite the staggering diversity of galaxies in the Universe, and as suggested by the presence of rigid classification schemes such as the Hubble tuning fork (Hubble, 1926), galaxies share common, distinctive sub-structures that can act as tracers of their evolutionary past. What follows is a curated list of

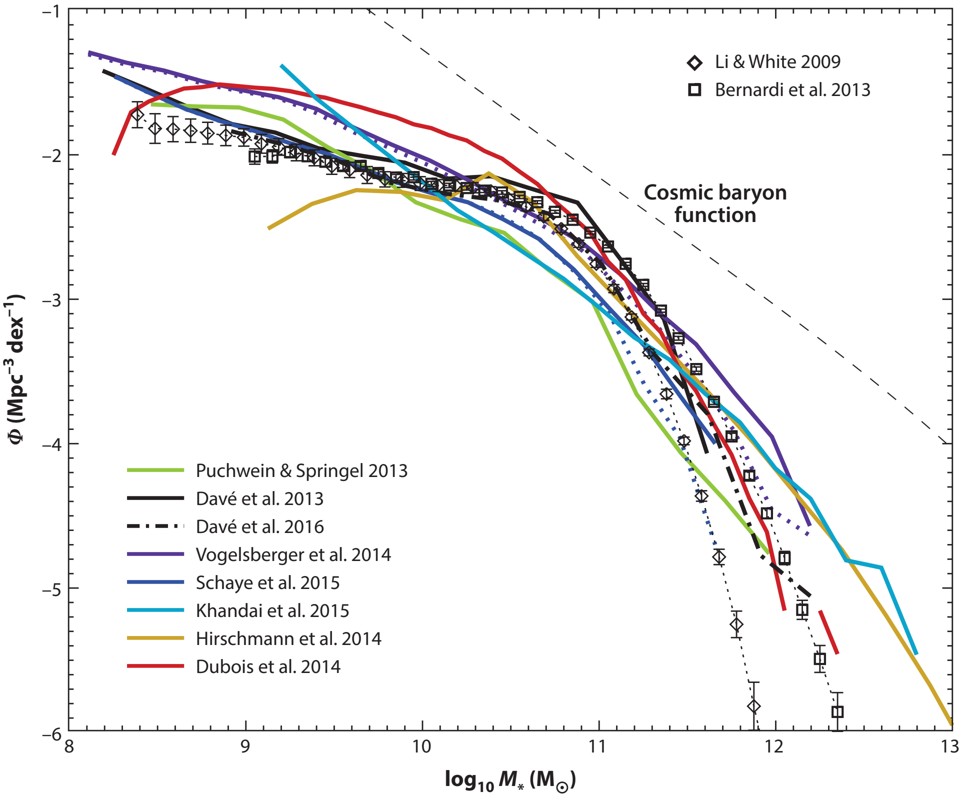


Figure 1.1: Figure 4 from Vogelsberger et al. (2020). Comparison of different galaxy stellar mass functions from recent cosmological simulations. All simulations except Davé et al. (2013) include tuned feedback mechanisms to more closely agree with observations of the galaxy mass function (Li & White 2009 and Bernardi et al. 2013). The straight dashed line illustrates the hypothetical galaxy mass function assuming the cosmic baryon fraction.

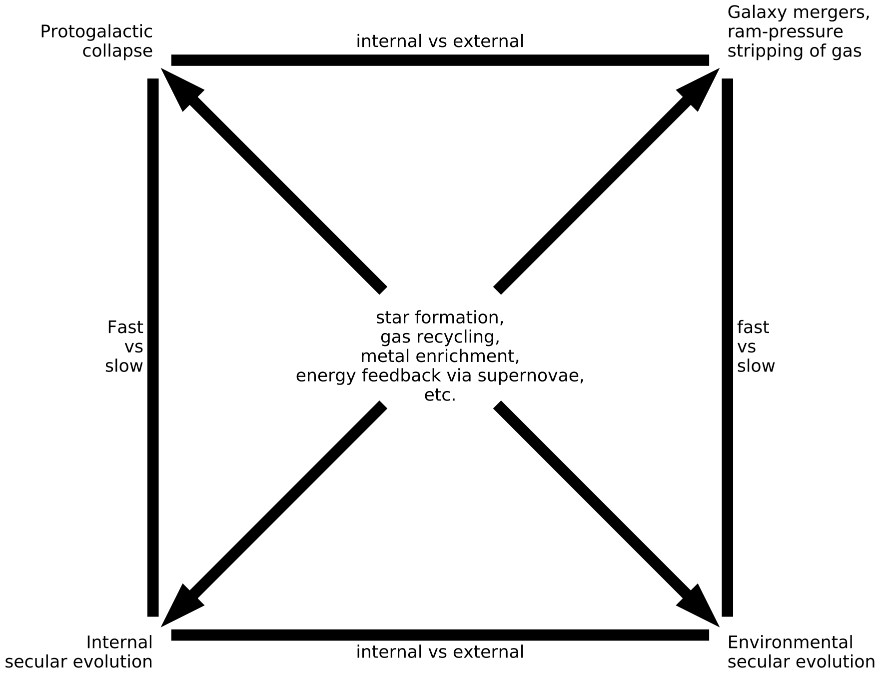


Figure 1.2: Morphological processes divided into fast or slow and internal or external. Central processes are aspects of all types of galaxy evolution. Adapted from Figure 1 from Kormendy & Kennicutt (2004).

common structures found in late-type galaxies, and a taster of the scientific work undertaken to understand them.

1.2.1 Galaxy bulges

Many disc galaxies have a central spheroidal structure generally comprised of older stars. These central "bulges" come in two distinct flavours, determined by their light profile and kinematics:

1. Classical bulges, which share a similar structure with early-type galaxies, with a light profile that obeys a de Vaucouleurs law (de Vaucouleurs, 1948) and chaotic stellar orbits (Minniti & Zoccali, 2008).
2. Pseudobulges, which behave more similarly to galaxy discs with regular rotation and an exponential light profile.

Detailed numerical simulations have demonstrated that merger events between two late-type galaxies of equivalent size causes the formation of a classical bulge (Toomre, 1977). Pseudobulges are instead expected to be the result of secular evolution, driven by non-axisymmetric disc instabilities and small-scale tidal interactions (Guedes et al., 2013).

Similar to most astrophysical processes, more complexity can be added to this picture. Boxy and peanut-shaped bulges are widely reported in both simulation and observations (Kormendy & Kennicutt 2004, including in the Milky Way, Minniti & Zoccali 2008).

1.2.2 Galaxy bars

Bars are highly ellipsoidal central structures present in roughly 30% of disc galaxies (Sellwood & Wilkinson, 1993; Masters et al., 2011). They are thought to be caused by density waves originating in the centre of the galaxy, building into a self-fuelling pattern (Sellwood & Wilkinson, 1993) capable of surviving many galactic rotational periods. N-body simulations have found that bars in more massive spiral galaxies most likely developed recently (Combes

& Sanders, 1981), which is supported by the observation that redder disc galaxies are more likely to have a bar (Masters et al., 2011). Bars are possibly triggered by tidal interactions (Elmegreen et al., 1990), which has observational support from morphological analysis of large samples of galaxies (Skibba et al. 2012; see Section 1.2.4).

Simulations suggest that the presence and strength of a bar increases with galaxy stellar mass and decreases with galaxy gas fraction, and that the development of a bar is linked to the cessation of star formation (Masters et al., 2010a; Zhou et al., 2020), a picture which is supported by HI observations (Masters et al., 2012; Newnham et al., 2020).

The asymmetric forces bars exert on the galaxy are thought to be responsible for changes in the angular momentum of both the baryonic and dark matter present, causing radial migration of stars and the formation of other morphological features such as a ring or lens (Sellwood & Wilkinson 1993; Kormendy & Kennicutt 2004), with the strength of the bar playing an important role in the galaxy's evolutionary pathway (Salo et al., 1999). Bars may boost central star formation (Hawarden et al. 1986; Lin et al. 2020), feed supermassive black holes (Jogee 2006; Galloway et al. 2015, though this has mixed observational support, Ho et al. 2009; Cheung et al. 2015); and govern the formation and evolution of spiral arms (Athanassoula, 2012).

1.2.3 Active Galactic Nuclei

Many galaxies show an incredibly bright central point, now thought to be supermassive black holes with hot accretion disks (Lynden-Bell, 1969) and possibly a relativistic jet (Blandford & Königl, 1979; Marscher, 1980). First observed by Seyfert (1943), these Active Galactic Nuclei (AGN) come in a wide variety of forms (Ulrich et al., 1997), and their presence can have a dramatic effect on a galaxy's evolution: massive enough AGN can expel gas and dust from the galaxy's nucleus, limiting the growth of the most massive galaxies (Silk & Mamon, 2012).

The method by which AGN grow is uncertain (e.g. Kormendy & Ho 2013;

Heckman & Best 2014 for reviews), but major merger events are considered to be a possible trigger for their activity (Sanders et al., 1988; Springel & Hernquist, 2005; Croton et al., 2006). Bars are more frequent in galaxies with AGN (Knapen et al., 2000; Galloway et al., 2015) and studies have shown a correlation between the mass of a galaxy's AGN and the stellar mass of its bulge.

1.2.4 Spiral arms

Spiral structure is present in a majority of massive galaxies (Buta 1989; Lintott et al. 2008) yet the formation mechanisms through which spiral structure originates are still hotly debated. Spirals are as diverse as the theories proposed to govern their evolution: from the quintessential pair of well-defined arcs of the grand design spiral; to the fragmented arm segments of the flocculent spiral; to the disjointed multi-armed spiral. These variations on structure account for 18%, 50% and 32% of the population respectively (Elmegreen et al. 2011; Buta et al. 2015, examples of each type are shown in Figure 1.3). The Hubble classification scheme (Hubble, 1926) and its revisions and expansions (Sandage 1961; de Vaucouleurs et al. 1991a) contain detailed variations of different types of spiral galaxy, divided by the presence of a bar and ordered by how obvious spiral arm patterns are, how tightly they are wound and the prominence of a central bulge.

A majority of the population of young stars in a galaxy are located in its spiral arms (Elmegreen, 2011), and it is possible the passage of a spiral arm triggers star formation (Cedrés et al., 2013; Kim et al., 2020). However, their main role may instead be to sweep up material, promoting the growth of Giant Molecular Clouds (Dobbs, 2014). Studies of spiral morphology have found interesting correlations to other galactic properties, such as a correlation between spiral tightness and central mass concentration (Yu & Ho 2019, Savchenko & Reshetnikov 2013 though Hart et al. 2017 found no such relation) and tightness and rotation curve shape (Seigar et al. 2005, with rising rotation curves creating more open spiral structure). These pre-

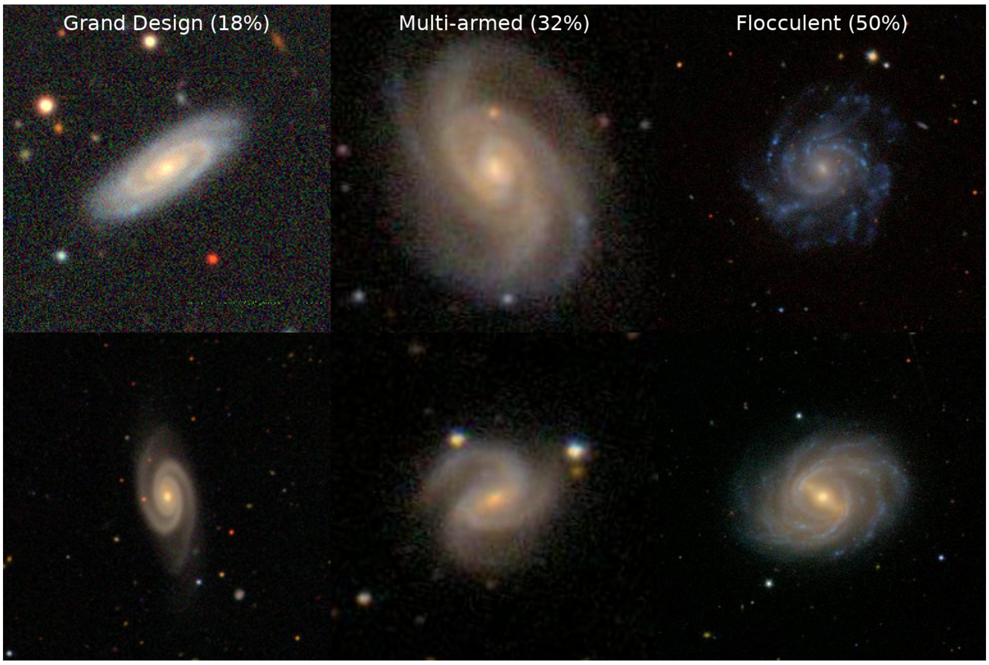


Figure 1.3: Examples of the different types of spiral galaxy present in the sky. The left column shows grand design spirals, the middle shows many-armed spirals and the right shows flocculent spirals. Images were taken with the Sloan Digital Sky Survey Telescope (Gunn et al., 2006).

dictions and observations provide compelling reasons for investigating their underlying rules and dynamics, as doing so is essential for understanding the secular evolution of disc galaxies.

Our current understanding of the mechanisms which drive spiral growth and evolution suggests that each of the different forms of spiral galaxy may be triggered primarily by different processes. Grand design spirals are thought to have undergone a tidal interaction, be driven by a bar (as seen in gas simulations, Sanders & Huntley 1976; Rodriguez-Fernandez & Combes 2008; and suggested for stars by Manifold theory, Romero-Gómez et al. 2006; Athanassoula et al. 2009a; Athanassoula et al. 2009b), or be obeying (quasi-stationary) density wave theory, in which spiral arms are slowly evolving, ever-present structures in the disc (Lin & Shu, 1964). Flocculent, patchy spirals are thought to be formed through swing amplification (shearing of small gravitational instabilities in the disc), and be transient and recurrent in nature (Julian & Toomre, 1966).

Static spirals

One of the fundamental assumptions of early work on spiral formation mechanisms was that the disc of a galaxy, if unstable to spiral perturbations, would create a stable, static wave which would exist unchanging for many rotational periods (Lin & Shu, 1964). The motivation for static waves with small numbers of arms (with a preference for $m = 2$) was primarily observational; most galaxies show spiral structure, suggesting that spirals exist for a long time or are continually rebuilt (Dobbs & Baba, 2014).

One of the first attempts to mathematically model the underlying dynamics of spiral arms, Lindblad (1927) considered the induction of highly eccentric orbits of stars at the outer edges of a flattened spheroid, eventually concluding that spiral arms are analogous to a harmonic wave in a flattened spheroid (Lindblad, 1940).

This idea of spiral arms as a wave of high-density material was later picked up by Lin & Shu (1964), who derived the properties of waves in a fluid disc (the Lin-Shu dispersion relation). In this picture, spiral arms were density waves propagating through the disk with a constant pattern speed, different from the pattern speed of the fluid disk itself. A popular analogy is the “phantom traffic jam”, in which a traffic jam (i.e. an overdense region of cars) propagates along a road[1].

This dispersion was derived under three approximations: linear perturbations, tightly winding spiral arms, and quasi-stationary structure in which the shape of a spiral pattern exists almost unchanging for many rotational periods. The result theory is referred to as quasi-stationary density wave theory (QSDW theory). Following the derivation of the dispersion relation of a fluid disc, Lin & Shu (1966) and Kalnajs (1965) derived the dispersion relation of a razor-thin stellar disc, which behaves similarly to a fluid disk for longer wavelength spirals. However, the collisionless nature of a stellar disc results in an upper limit on the frequencies of perturbations and therefore a minimum possible arm pitch angle (the angle between the spiral and the tangent to a circle centred on the galaxy, Binney & Tremaine 1987, illustrated in Figure 1.4).

The assumption of quasi-stationarity requires mechanisms by which QSDW spirals continually counteract the radial propagation of the energy and angular momentum of the density waves. Two such mechanisms are “wave amplification by stimulated emission of radiation” (WASER, Mark 1974; 1976) and swing-amplification (Goldreich & Lynden-Bell 1965; Julian & Toomre 1966; Toomre 1981). Each of these mechanisms cause waves to be amplified when crossing the co-rotation radius (the radius at which the stellar pattern speed matches that of the spiral pattern).

By examining spiral structure at different wavelengths, Pour-Imani et al. (2016) found evidence of a variation in spiral arm pitch angles at different wavelengths, as would have been suggested if the transit of a spiral density wave triggered star formation (though this is not evidence for long-lived density waves). However, Tenjes et al. (2017) found no such relationship in a

---

[1] `www.youtube.com/watch?v=Q78Kb4uLAdA`

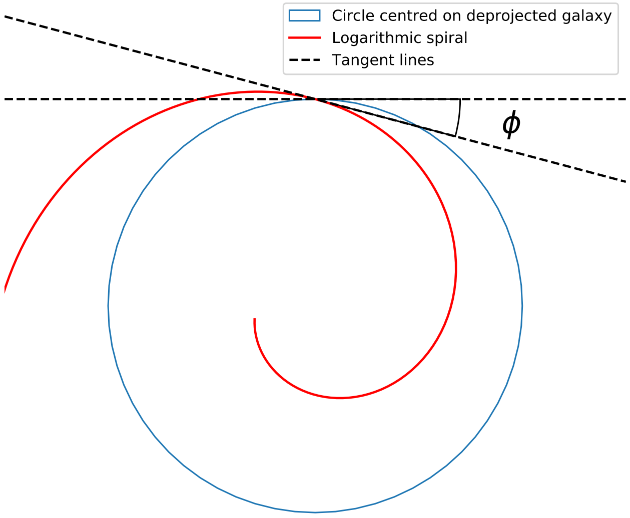


Figure 1.4: Illustration of the definition of pitch angle. It is given as $\phi = \tan^{-1}\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\, r^{-1}\right)$, or the angle between the spiral (red) and the tangent to a circle centred on the galaxy (blue).

detailed study of a well-resolved galaxy, M31. Sellwood (2011) posits that it is incredibly difficult to use observations of external galaxies to determine the lifetime of spiral patterns, and evidence from our Milky Way suggests that spirals have a much shorter lifetime than initially assumed, with in-plane velocity components of local stars being grouped into a number of separate streams, rather than the group motion a star-forming density wave would create (Dehnen & Binney 1998; Nordström et al. 2004; Bovy et al. 2009).

Dynamic spirals

Many simulations observe that spirals arms do not maintain a constant pitch angle, and instead wind-up over time due to the differential rotation of the disc (Baba et al., 2013). Recent simulations suggest that spirals arms are transient in nature, and continually dissipate and re-form (Dobbs & Baba, 2014), in contrast to the assumptions of Lin & Shu (1964). These spirals can be maintained through the same mechanisms that drive QSDW spirals (such as WASER and swing amplification) but do not require the idealistic disc conditions needed for the formation and maintenance of a stationary wave.

Many overlapping mechanisms have been proposed to describe the generation of transient spiral structure. Sellwood (2011) highlights swing amplification, recurrent cyclic groove modes (Sellwood & Kahn, 1991), and global mode theory (Bertin & Lin, 1996) as three distinct, promising mechanisms to explain the complexities of dynamic spiral structure. Of these, swing-amplification has become a favourite among researchers thanks to the measurable predictions it makes regarding spiral arm number, formation and shape.

Julian & Toomre (1966) first showed that, through swing amplification, a small leading density fluctuation can cause a dramatic trailing spiral pattern. A clumpy mass distribution could lead to “a swirling hotch-potch of pieces of spiral arms” (Goldreich & Lynden-Bell, 1965), though the resulting spirals would be chaotic, with little symmetry or larger-scale periodicity. Evidence for swing amplification in galaxies comes from both simulations (D’Onghia et al. 2013; Baba et al. 2013; Michikoshi & Kokubo 2016; Michikoshi & Kokubo 2020) and observation of both the number of spiral arms present in galaxies and their tightness (Hart et al., 2018).

Some theories suggest that the pitch angles of these transient spiral arms will decrease due to the differential rotation of the disk, with the density of the arm peaking at a critical pitch angle, before dissipating to be reformed. Pringle & Dobbs (2019) proposes a simple test of spiral arm winding, assuming the cotangent of the pitch angle of a spiral arm evolves linearly with time. They found that the distribution of pitch angles in their sample of 86 galaxies was consistent with this prediction.

Morphological dependence of spiral evolution

The forces governing spirals are caused by a complex mix of small-scale instabilities, larger disc self-gravity, tidal interactions, and forces caused by other morphological components present in the galaxy, such as bulges and bars.

Bulge size has long been linked to spiral pitch angle, with the Hubble Tuning fork (Hubble, 1936) combining decreasing pitch angle and increasing bulge size in its classification of unbarred galaxies. This correlation has mixed support in observations, with Savchenko & Reshetnikov (2013) finding supporting evidence, Kennicutt (1981) finding only a weak correlation, and Masters et al. (2019) finding instead that galaxies with small bulges have arms with a large range of pitch angles.

Manifold theory is a comprehensive approach to modelling the effect of bars on the morphology of a galaxy (Romero-Gómez et al. 2006; Romero-Gómez et al. 2007; Romero-Gómez et al. 2009; Athanassoula et al. 2009b), specifically causing the formation of rings and spiral arms. It proposes that bars confine the chaotic orbits of stars, guiding them along invariant manifolds. These manifolds then provide the building blocks for the spirals and rings observed in barred galaxies. This theory has had great success in simulations (Athanassoula, 2012), but its predictions have not been demonstrated in observational studies (Díaz-García et al., 2019).

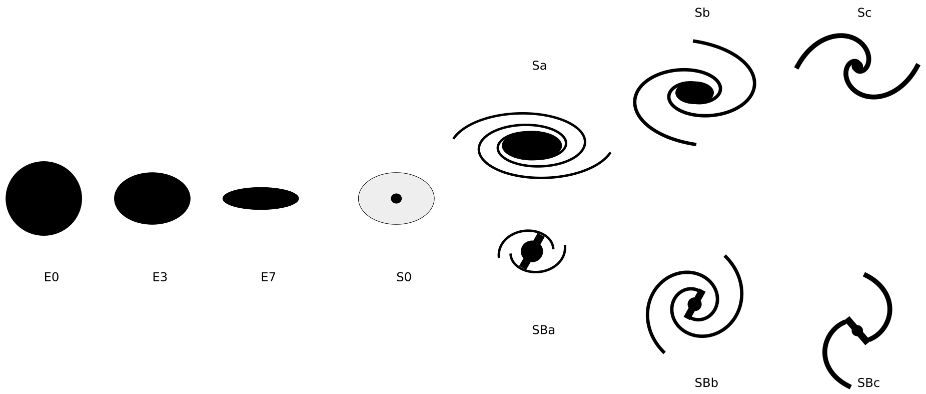


Figure 1.5: The Hubble tuning fork, showing early-type galaxies on the left and late-type galaxies on the right. Unbarred spirals are at the top right and barred spirals at the bottom right.

1.3 Methods of morphological classification

The understanding of a collection of objects often begins with an exploration into the presence of distinct subpopulations. This segregation provides avenues of investigation which help probe the underlying mechanisms causing the observed differences.

In an attempt to better categorise the diversity of observed galaxies, many classification schemes have been proposed by researchers: one of the original methods used is the Hubble tuning fork (Hubble 1936; Figure 1.5), which distinguished between "late-type" ("ellipticals" and "lenticulars") and "early-type" ("spiral" and "irregular") galaxies. These groups are then further sub-categorised based on ellipticity for early-types, and the presence of a galactic bar, the strength of a central spheroidal bulge and the tightness of any spiral arms present for late-types. This tuning fork has subsequently been expanded on and augmented, for instance by Sandage (1961) and Vaucouleurs (1959), maintaining the basic structure proposed by Hubble, but adding scope for galaxy rings, lenses and further sub-categorisation based on combinations and strengths of the different sub-components.

The pioneers of morphological classification spent hours poring over photographs taken from some of the most powerful telescopes at the time, painstakingly compiling lists of classifications such as the *Hubble Atlas of Galaxies* (Sandage, 1961), the *Carnegie Atlas of Galaxies* (Sandage & Bedke, 1994) and the *Third Reference Catalogue of Bright Galaxies* (de Vaucouleurs et al., 1991a).

Despite significant advancement in technology in the years since the inception of the Hubble classification scheme, visual inspection of galaxies is still a cornerstone of galaxy morphological classification. However, the era of big-data astronomy (including the Sloan Digital Sky Survey, SDSS, Blanton et al. 2017; the Dark Energy Survey, DES, Abbott et al. 2018; the Vera C. Rubin Observatory Project, Ivezić et al. 2019; and Euclid, Laureijs et al. 2011, Amiaux et al. 2012) has meant that small teams of researchers can no longer provide expert visual classifications for every observed galaxy in a survey. In addition to this issue, the subjectivity of galaxy classification results in less-than-consistent classification: Naim et al. (1995) found that even expert classification is subject to a high degree of scatter (only 1% of galaxies they tested received a consistent classification from all six of the experts in their study, and less than one quarter achieved a two-thirds consensus). One promising solution to these problems came in the form of citizen science, which is further discussed in Chapter 2.

1.3.1 Automating visual classification

The time-consuming nature of visual inspection has prompted many schemes for automated classification. These range from the use of simple proxies such as colour and surface brightness profiles, to more complex ones based on kinematics and unsupervised machine learning.

Morphological proxies, while being relatively easy to obtain, are inherently limited by unknown and often unquantifiable biases (for instance the presence of red spiral galaxies would confound a simple colour-based proxy, Masters et al. 2010a). Despite this, they have been widely employed as a stand-in for visual inspection of galaxies (Masters et al. 2019; for example, Bell et al. 2004; Weinmann et al. 2006; van den Bosch et al. 2008).

The use of supervised machine learning techniques to solve the problem of large-scale classification has been widely attempted. Storrie-Lombardi et al. (1992) first utilized a small feed-forward artificial neural network (NN) to attempt to distinguish basic Hubble types of a sample of 5217 galaxies (1700 used to train the NN, and 3517 to test its output) from the ESO-Uppsala catalogue (Lauberts & Valentijn, 1989). The input to the NN consisted of 13 derived parameters (such as axial ratio, parameters of a simple 2D photometric model fit and surface brightness at various radii), and the machine classifier achieved a modest 64% agreement to expert classification. This limited success is possibly due to the lack of a large-scale, sufficiently diverse, self-consistent training set, as well as the simple NN architecture employed.

Many studies have followed the pioneering example of Storrie-Lombardi et al. (1992), for instance, Ball et al. (2004) similarly used a small NN with 29 derived parameters to categorize 1875 galaxies into a modified numerical Hubble type (de Vaucouleurs, 1959), running from -5 to +11, with negative values representing early-type galaxies, and positive values late-type. Their classifier was correct to within ±1.5 types for most galaxies, despite the simple model and small samples used.

In more recent years there has been an explosion in the application of deep neural networks for image classification (Krizhevsky et al. 2012; Simonyan & Zisserman 2014; Szegedy et al. 2014; Szegedy et al. 2015; He et al. 2015). These "deep" networks involve millions of trainable parameters and thus require a large, diverse sample of training data in order to avoid overfitting. The Galaxy Zoo 2 catalogue of morphological classifications (Willett et al. 2013b; see Chapter 2 for further detail) provides one such training sample.

Dieleman et al. (2015) developed a rotation-invariant convolutional neural network (CNN) as part of the *Galaxy Challenge* competition[2], trained on a curated subset of the Galaxy Zoo 2 catalogue. They trained 17 variants

---

[2] `www.kaggle.com/c/galaxy-zoo-the-galaxy-challenge`

of a CNN model, applied 60 affine transformations to each galaxy image and averaged over the output classifications to form a final consensus. This ensemble classification algorithm replicated Galaxy Zoo 2 classifications with $> 99\%$ accuracy. Huertas-Company et al. (2015) extends the use of CNNs for morphological classification to high redshift, using a CNN to categorise 50,000 galaxies based of a training set of 8,000 visual morphologies published by the CANDELS collaboration (Grogin et al. 2011; Kartaltepe & CANDELS Collaboration 2014).

The use of transfer learning allows the adaptation of a machine classifier to new datasets and problems using significantly smaller amounts of training data than that which they were originally trained on. Domínguez Sánchez et al. (2019) successfully used transfer-learning to apply a CNN classifier, trained on SDSS data with Galaxy Zoo classifications, to DES data, using only a few hundred morphological classifications of the DES dataset.

The use of supervised machine learning for morphological classification is subject to a number of drawbacks: any bias or error in the training data which is not properly accounted for will be learnt by the machine classifier; unless special consideration is given to classification uncertainty, the uncertainty in machine classification will be a gross underestimate (except in the limit of very large training samples, Szegedy et al. 2015); advanced neural network architectures are difficult to interpret, and without careful regularization will overfit; machine classifiers do not respond well to novel data (for instance, an image which is unlike any it was trained on, Wang et al. 2017).

1.3.2 Beyond visual morphologies

As mentioned above, one major drawback of visual classification is its subjectivity, a subjectivity which will invariably impact any machine learning technique which uses human visual classification as a training or test set. An alternative approach is the use of unsupervised machine learning techniques for clustering of galaxies. These approaches have the advantage of being instantly applicable to new datasets without needing the labour-intensive creation of a training set; they are free from human error in labelling and can often generalise to previously unseen data.

Hocking et al. (2018) propose an elegant method of unsupervised morphological clustering and classification for galaxies from an entire survey, requiring no pre-annotation of data. They use a series of algorithms to identify the structure in an input data matrix (growing neural gas algorithm, Fritzke 1994), cluster structure into regions using hierarchical clustering (Hastie et al., 2009), and label sub-structures (galaxies) within the regions using connected-component labelling. They successfully use this algorithm to separate galaxies into what would conventionally be viewed as early- and late-types. Martin et al. (2020) make use of this method to identify 160 distinct "morphological clusters" of galaxy types; successfully differentiating between broad Hubble types, and even identifying high-redshift merger events.

Uzeirbegovic et al. (2020) made use of Principal Component Analysis (PCA) to identify 12 "eigengalaxies" representing basis images which, when combined, account for a majority of the variance in a population of imaged galaxies (a technique originally proposed by Sirovich & Kirby 1987 for human faces). They made use of 10,243 galaxies present in the "Galaxy Zoo: CANDELS" morphological catalogue (Simmons et al., 2017a), successfully accounting for 96% of the image variance. This method of morphological classification is a scalable alternative to visual inspection, which allows the clustering of similar galaxies, identification of novel images (outlier detection) and imputation of missing data.

A complimentary avenue to visual morphologies for large-scale automated galaxy classification is the separation of galaxies using measurements of their kinematics. Kinematic classification systems leverage the relatively new technology of integral field spectrography to calculate peculiar velocities of gas and stars in a galaxy by measuring the redshifting and line-widths of emission or absorption lines at different locations in the galaxy.

For example, Emsellem et al. (2007) made use of the 2D stellar kinematic measurements of the SAURON survey (Bacon et al. 2001; de Zeeuw et al.

2002) to group early-type galaxies into two distinct groups based on their stellar angular momentum per unit mass, naming these groups fast- and slow-rotators. This is of significant interest as it is complimentary to visual morphological classification, with many elliptical galaxies showing similar kinematic characteristics to S0s (Bender et al., 1988).

Krajnović et al. (2011) similarly used the ATLAS 3D (Cappellari et al., 2011a) sample of early-type galaxies. They separated galaxies into regular and non-regular rotators, with non-regular rotator early-type galaxies more likely to be found in dense neighbourhoods at higher mass than regular rotators. Emsellem et al. (2011) continues this work to characterise the distributions of fast- and slow-rotating galaxies using a larger sample than Emsellem et al. (2007). They conclude that fast rotators are mostly discy galaxies, with some being as flattened as late-type spiral galaxies, and Cappellari et al. (2011b) propose a revision to the Hubble tuning fork to better account for the large variation in bulge sizes of fast rotators. The kinematics of simulated galaxies have also been studied, in order to examine consistency with observation and predict evolutionary pathways for galaxies with a distinct rotation profile (Schulze et al., 2020).

The requirement of kinematic classification schemes for detailed IFU observations of galaxies and the potential pitfalls associated with obtaining rotation curves and velocity dispersions from IFU data (Noordermeer et al., 2008) is a major drawback of this novel approach to grouping galaxies based on their underlying dynamics.

Ultimately, the combination of complementary classification schemes and data sources (i.e. Nevin et al. 2019) will allow a more complete understanding of the plethora of differences between galaxies imaged in next-generation surveys.

1.4 Photometric modelling

Galaxy morphology can also be quantified and examined through the parametric fitting of light profiles to a galaxy image. Dating back to Reynolds (1913), it predates even the Hubble tuning fork. Models fit by photometry are often guided by visual morphological classification (e.g. Kruk et al. 2017), and can be combined with kinematic measurements to great effect (Binney, 1982).

Photometric fitting has become an important tool for extragalactic astronomers seeking to understand the formation and evolution of the galaxy population; from analysing bulge and bar structure (Elmegreen & Elmegreen 1985; de Jong 1996; Gadotti 2011; Mendez-Abreu et al. 2016; Gao & Ho 2017; Kruk et al. 2018) to the secular evolution of disc galaxies (Lilly et al. 1998; Barden et al. 2005; Allen et al. 2006) and general galaxy assembly and evolution (Simard et al. 2002; Bamford et al. 2011; Lackner & Gunn 2012; Rampazzo et al. 2019).

These fully quantitative methods allow researchers to obtain structural parameters of galaxy sub-components, which has been used in a variety of astrophysical and cosmological research, such as:

- The stellar mass found in discs and bulges places strong constraints on the galaxy merger tree from ΛCDM N-body simulations (Parry et al. 2009; Hopkins et al. 2010; Rodrigues et al. 2018).
- The strength of a galaxy's classical bulge is thought to be tied to the strength of merger events in its past (Springel & Hernquist, 2005; Kormendy et al., 2010).
- The mass of a central bulge relative to that of the disc may correlate with the mass of a central black hole (Simmons et al., 2017b; Davis et al., 2019; Sahu et al., 2019).

Over the years, the complexity of photometric models has gradually increased: from de Vaucouleurs (1948) demonstrating that many early-type elliptical galaxies display a $\boldsymbol{R}^{1/4}$ (commonly referred to as a de Vaucouleurs) profile, a specific case of the Sérsic profile ($\boldsymbol{R}^{1/n}$, Sérsic 1963; Graham & Driver 2005); to Freeman (1970) showing that disc profiles were exponential and the central bulges of disk galaxies could be described using a de Vaucouleurs profile, demonstrating that the observed light could be decomposed into two distinct components. Further work has expanded on these initial models, demonstrating the need for more nuance and freedom when performing photometric decomposition (Graham, 2014).

The usefulness of obtaining parametric models of a galaxy has motivated the creation of many image modelling and fitting suites, including GIM2D (Simard et al., 2002), GALFIT (Peng et al., 2002), MEGAMORPH (Bamford et al., 2011) and PROFIT (Robotham et al., 2016) for two-dimensional modelling, and PROFILER (Ciambur, 2016) for one-dimensional, to name a few. Using these tools, researchers have built large catalogues of model fits to galaxies. One of the largest photometric model catalogues is that of Simard et al. (2011), who performed automated 2D, two-component (bulge + disc) decomposition of 1,123,718 galaxies from the Legacy imaging of the Sloan Digital Sky Survey Data Release 7 (Abazajian et al., 2009). Simard et al. (2011) used two axisymmetric Sérsic components representing a bulge and disc, with the disc an exponential profile, and the bulge either a de Vaucouleurs profile or a general Sérsic profile.

Many other large catalogues of photometric fits exist (e.g. Lackner & Gunn 2012; Kelvin et al. 2012; van der Wel et al. 2012), but despite the usefulness of photometric fitting, and the presence of analytic profiles and methods for modelling more complex galaxy sub-components, relatively few studies have attempted to perform large-scale (1000s of galaxies) parametric decomposition of galaxies using more complicated models than that of Simard et al. (2011). Failing to take into account these “secondary” morphological features (such as a bar, ring and spiral arms) can impact detailed measurements of a galaxy’s bulge (Gao & Ho, 2017). Proper decomposition of secondary morphological features allows investigation into mechanisms behind the secular evolution of galaxies (Head et al. 2015; Kruk et al. 2018; Gao et al. 2018) and exploration of environmental effects on morphology, such as offset bars (Kruk et al., 2017).

A prominent issue when performing these detailed decompositions is the tendency for fitting functions to converge on unphysical results when not properly guided or constrained, for instance in a two-component model containing a Sérsic bulge and an exponential disc, the bulge may grow to encompass the galaxy's disc, as its extra parameter allows for a closer fit (as observed by Graham 2001 and Kruk et al. 2018). It is also the case that often, without well-chosen starting points, detailed model fits will fail to converge at all (Lange et al., 2016).

Compounding this, uncertainties reported by many software fitting packages (e.g. GALFIT and MEGAMORPH from the above list) are often lower estimates on the real uncertainty, due to secondary sources not being modelled, flat-fielding errors and incorrect statistical models (Peng et al. 2010; including the possibly incorrect assumption of Poisson noise). Other packages such as GIM2D and PROFIT attempt to fully model posterior distributions and so produce more representative uncertainties, however, this comes with a larger computational cost and configuration complexity. Formal uncertainties are measures of the likelihood space and therefore underestimates of the true error as an analytic model will rarely capture the nuanced light profile of a galaxy. Furthermore, the common use of reduced chi-squared (χ^2_ν) in photometric modelling is not well-motivated, as it is impossible to estimate the degrees of freedom of a nonlinear model (Andrae et al., 2010).

Another problem which needs to be addressed is whether a component should be present in the model at all. An automated fit will generally attempt to add as many components as possible to produce the closest-matching model. Each component can be selected from a wide variety of analytic profiles (e.g. the Sérsic profile; the core-Sérsic profile, Graham et al. 2003; the Nuker profile, Rest et al. 2001; and the modified Ferrer profile, Binney & Tremaine 1987, among others), providing further model selection complications. Many studies, therefore, need to select the most appropriate model by visual inspection of the resulting residuals or recovered parameters. For example, both Vika et al. (2014) and Kruk et al. (2018) inspected the resulting model and residual images for all of their parametric fits (163 and 5,282 respectively) to ensure physical results with the correct components present. The end result of most of these problems is that researchers will have to invest time to individually check many of their fits to ensure they have converged on a physical model. In the era of large sky surveys, the time required to do this becomes unsustainable and introduces concerns over human error if done by only a single, or small number of individuals.

Various methods have been devised and implemented with the aim of automatically choosing the correct model, ranging from the use of χ^2_ν values to perform an F-test (Simard et al., 2011); to taking the model with the lowest χ^2_ν (Lackner & Gunn, 2012); to attempting to implement a set of rules ("logical filter") by which models can be excluded from the possible pool (Allen et al., 2006). Méndez-Abreu et al. (2018) make use of both logical filtering and the Bayesian Information Criterion (BIC, Schwarz 1978) to study the bulges of S0 galaxies.

Another key consideration for any work fitting light profiles to galaxies is the choice of optimization algorithm and associated "goodness-of-fit" measure. It is common to minimize χ^2_ν using a gradient-descent based algorithm such as the Levenberg-Marquardt algorithm (Levenberg, 1944), or the BFGS algorithm (Broyden 1970; Fletcher 1970; Goldfarb 1970; Shanno 1970). While these algorithms are computationally efficient and often simple to implement, they will be trapped by local minima and often fail to converge to the globally optimum solution. Some work has incorporated more robust algorithms (Metropolis-Hastings, Simard et al. 2011; component-wise hit-and-run, Robotham et al. 2016), though this comes at a computational cost. One approach adopted by (Lange et al., 2016) is to perform a simple gradient-descent fit repeatedly, with many diverse initial conditions. This approach not only increases the chance that the global optimum is found, but also provides a measure of the likelihood space, allowing parameter uncertainties and degeneracies to be quantified.

An exciting recent development in light profile fitting is the use of CNNs to estimate photometric model parameters (Tuccillo et al., 2018). This technique provides a significant increase in speed over conventional model fitting, presenting a scalable solution to photometric modelling of existing and upcoming large-sky surveys.

Photometric modelling is a tricky problem space, but the potential rewards are high. When models are chosen carefully to well-represent the target galaxies, potential sources of error such as atmospheric effects and image stacking are properly accounted for, and care is taking in choice of initial conditions, optimization metric and algorithm choice, it is possible to gain incredibly detailed, quantitative insight into the internal structure of the galaxies being investigated.

1.5 Thesis plan

This thesis exhibits work undertaken by the candidate since October 2016; Chapter 2 provides a brief history and discussion of citizen science and crowdsourcing, and highlights existing work done to understand the motivations of citizen scientists and most effectively combine their efforts with machine learning to optimize project throughput.

Chapter 3 introduces the *Galaxy Builder* citizen science project, which forms the foundation of this work. It details the design, development and deployment of the project, and an examination of its results. It presents the selection of the main run of galaxies, how the *Galaxy Builder* web interface allowed volunteers to create photometric models of galaxies, and how machine learning techniques were used to aggregate many volunteer classifications into a single photometric model for each galaxy. This work is presented in Lingard et al. (2020a).

Chapter 4 leverages the results from the *Galaxy Builder* project to investigate the processes governing spiral arm evolution, using a hierarchical Bayesian modelling approach to robustly quantify the pitch angles of galaxies. This work is also presented in a paper (Lingard et al., 2020b).

Chapter 5 presents a summary of the progress achieved, proposes future scientific avenues where *Galaxy Builder* models could be utilized, and details improvements to the citizen science project that would improve a future iteration of the project.

Chapter 2

Many hands make light work

This Chapter contains an explanation of and a brief history of citizen science and crowdsourcing. It introduces Galaxy Zoo and the Zooniverse, examines work done to understand what drives volunteers to engage in citizen science, understand possible biases in volunteer classification, and presents work done to improve the efficiency of volunteer efforts.

2.1 Crowdsourcing

Crowdsourcing is the act of outsourcing work to a (decentralised) pool of individuals, generally dividing work among those involved to maximise productivity. Originally coined in 2006[1] in reference to the use of the internet by businesses to distribute tasks to a large pool of workers, the idea of using a group of independent opinions and ideas to come to an informed consensus has been around for centuries, including a UK governmental initiative from 1714 (the Longitude act) asking members of the public to propose new methods of calculating a ship's global position for a reward.

Crowdsourcing in the 21st Century comes in many flavours, from the aggregation of customer votes to drive business decision making (often through the use of social media) to the painstaking annotation of millions of images

---

[1] `https://www.wired.com/2006/06/crowds/`

to create machine learning training sets (for example, Yu et al. 2018; Tayyub et al. 2017). One could even consider the stock market to be a method which crowdsources an estimate of the value of businesses. These methodologies generally rely on the assumption that a large pool of independent classifiers will either converge to or peak at some desirable truth.

2.2 Citizen science

In the context of scientific research, crowdsourcing can often be found under the name Citizen Science, which is defined as the participation in scientific research of non-professional scientists. Historically the citizen scientist demographic was very different to what it is now; most of the well-known scientists prior to the 20th Century would be classed as citizen scientists. Greats such as Darwin, Descartes and Leibniz were all either self-funded or amateurs. These "gentlemen scientists" paved the way for the scientist profession we have today (Mims, 1999).

However, citizen science in modern times is more commonly linked to members of the general public undertaking some form of research, often prompted or guided by a professional scientist. One of the first recorded instances of this modern form of citizen science could be when Edmund Halley prompted people to observe the total eclipse of the sun on the 22nd of April 1715 (described as *"a request to the curious to observe what they could about it"*, Halley 1714). Halley used spatial variation of eclipse durations to create a map of the eclipse (Figure.2.1) as it passed over England, and as a test of the theory of Newtonian Gravity.

Using volunteers to collect data it would be impractical or impossible for researchers to collect themselves is a common form of citizen science, ranging from collecting acid rainfall across the US (Bolze & Beyea, 1989) to measuring Autumnal leaf colouring and fall in Austria to investigate climate change[2]. Other demonstrations of the power of the crowd include Fold-

---

[2] `eu-citizen.science/project/35`

Figure 2.1: A map of England produced by Edmund Halley showing the path of the 1715 total solar eclipse. Made available by the University of Cambridge, Institute of Astronomy Library

ing@Home (Shirts & Pande 2001; Larson et al. 2009), a distributed computing project for simulating protein dynamics in which citizen scientists volunteer their computers. In March 2020, as a result of the COVID-19 pandemic, the combined computing power of volunteer contributions exceeded 1.2 exaflops, making it the world's first exascale computing system.

It is commonly accepted that one of the major benefits of citizen science is the increased scientific literacy and engagement associated with those involved in the projects (Masters et al., 2016). A scientifically engaged population leads to more funding for scientific research (Miller, 2001) and more respect and understanding of the scientific method and results. With the recent influx of misinformation and falsities in the media, a scientifically literate audience is more capable of discerning what is factually correct and what is not (Scheufele & Krause, 2019).

Citizen science can be greatly mutually beneficial to all parties, and citizen science in the form of crowdsourcing has become invaluable with the influx of large scientific datasets which researchers would find prohibitively large to manage manually, and unfeasibly complex to categorize automatically.

2.3 Galaxy Zoo and the Zooniverse

The Sloan Digital Sky Survey (Blanton et al. 2017) was, at the time of its inception, "the most ambitious astronomical survey ever undertaken"[3]. One night of observing produced 200 gigabytes of data, and over the past 20 years, it has made photometric observations of over one million galaxies. The sheer volume of data output by the SDSS and other large-scale surveys presents galactic astrophysicists with a problem of person-power; it was no longer feasible to perform expert morphological classification of entire surveys, even for small teams (Schawinski et al., 2009). Many automated schemes have been proposed (see Section 1.3), but the nuances and lack of a clear "correct answer" to a galaxy's morphology make the creation of a self-consistent,

---

[3] `skyserver.sdss.org/dr1/en/sdss`

physically-motivated classification system nigh impossible.

Galaxy Zoo[4] was proposed as a solution to this problem of data deluge. Inspired by the success of Stardust@Home (Westphal et al., 2006), a crowdsourced project to locate particles of interstellar dust inside a collector returned from the NASA Stardust mission, Galaxy Zoo aimed to utilize a large, independent body of volunteers to classify galaxies into rough morphological bins. The project was immensely successful, and within one day was receiving over 70,000 classifications per hour. The sample of 900,000 galaxies (referred to here as Galaxy Zoo 1, or GZ1) was completed within half a year, with an average of 38 individual classifications per galaxy (Lintott et al., 2008).

The immense response from the community spurred a number of iterations and subsequent projects in the Galaxy Zoo ecosystem (including Galaxy Zoo 2, hereafter GZ2, Willett et al. 2013b; Galaxy Zoo: CANDELS, Simmons et al. 2014, Simmons et al. 2017a; Galaxy Zoo: Mergers, Holincheck et al. 2016b; and Galaxy Zoo: Hubble, Willett et al. 2017). The success of the initial project also prompted the creation of the Zooniverse Citizen Science platform[5] (Lintott, 2011). This platform is host to a wide variety of citizen science projects, including the most recent incarnation of Galaxy Zoo[6], other astrophysical projects such as exoplanet detection (Schwamb et al., 2012) and gravitational wave signal analysis (Zevin et al., 2017), as well as a variety of other projects such as classifying wildlife in camera traps, segmenting the nuclear envelope of cells and transcribing hand-written notes of anti-slavery activists in the 19th Century. The Zooniverse hosts over 100 “official” projects (as of June 2020), and is home to many more unofficial ones (which are user-created and not advertised through the official Zooniverse channels and project page). The scientific impact of the Zooniverse has been immense, with ~ 300 publications resulting directly from their projects[7], and many

---

[4] `zoo1.galaxyzoo.org`

[5] `zooniverse.org`

[6] `www.zooniverse.org/projects/zookeeper/galaxy-zoo/`

[7] `www.zooniverse.org/about/publications`

more making use of published data[8].

It is evident from the successes of platforms like the Zooniverse that opening up scientific methods to the crowd can help overcome many of the issues of large-scale datasets in science. Fewer restrictions on who can submit classifications, however, introduces the potential of bad actors (those who wish to disrupt the project), noisy classifications (including those who do not understand the task) and unpredictable biases necessitating detailed consideration of user experience and interface design.

2.3.1 User weighting

The naïve approach to combining classifications from many volunteers would be to assign each volunteer an equal weighting and calculate the distribution of classifications for each galaxy. However, this does not account for those volunteers who for some reason (misguided or malign) do not provide high-quality classifications. Lintott et al. (2008) provide an iterative algorithm used to weight each volunteer based on how consistently they are in consensus with other volunteers. They note that for the vast majority ($> 99\%$) of galaxies, the resulting classification is unchanged. However, weighting users dramatically improves the confidence in galaxies that would otherwise not have received a consensus vote (if a galaxy was classified by an abnormally large number of inconsistent classifiers, weighting allows the smaller proportion of consistent classifiers to create a consensus).

Many subsequent algorithms for user weighting have been proposed (Bamford et al., 2009; Lintott et al., 2011; Simpson et al., 2012; Simmons et al., 2017a), each of which aims to best converge on a consensus result while accounting for the aptitudes and biases of individual classifiers. This is very much an ongoing field of research, especially given the popularity of ensemble methods in modern machine learning (Wang et al., 2020).

---

[8] Simply searching the ADS for “Galaxy Zoo” returns over five-hundred results

2.3.2 Correcting for biases in classifications

Lintott et al. (2008) also undertook a bias study, with the aim of quantifying the bias present in the morphologies reported by volunteers. As well as the regular SDSS colour image shown to volunteers, a mirror-image and a greyscale image were also presented. They observed that user behaviour was not fully consistent during this bias study (volunteers were more careful with their classifications as they knew they were being tested for bias), but some important effects were still able to be quantified (monochrome images were slightly more likely to be classified as early-type, and a bias towards anticlockwise spiral winding was noted). These results cement the need for careful consideration of how data is presented to volunteers, and how classifications can be validated.

Another important consideration in the morphological classification of galaxies is image quality (such as signal-to-noise, angular resolution and the presence of secondary sources and artefacts such as satellite trails). One example of such a problem in the GZ1 classifications is the increased prevalence of early-type galaxies at higher redshifts. This evolution with redshift is primarily caused by the poorer signal-to-noise, resolution and size relative to atmospheric effects of the galaxy, which makes it impossible to distinguish spiral features. Bamford et al. (2009) attempts to correct for this classification bias in the GZ1 catalogue using a statistical method to de-bias the classification likelihoods. This work was further expanded upon by Hart et al. (2016b) for the far more complex GZ2 question tree (Figure 2.2).

2.4 What motivates volunteers?

Citizen scientists come in two distinct flavours: either attempting a project once and then never returning (“transient”), or returning on one or more occasions to further engage and submit classifications (“regular”). Ponciano et al. (2014) found that 67% of volunteers who contributed to a project were transient volunteers, however, almost 80% of the classifications were sub-

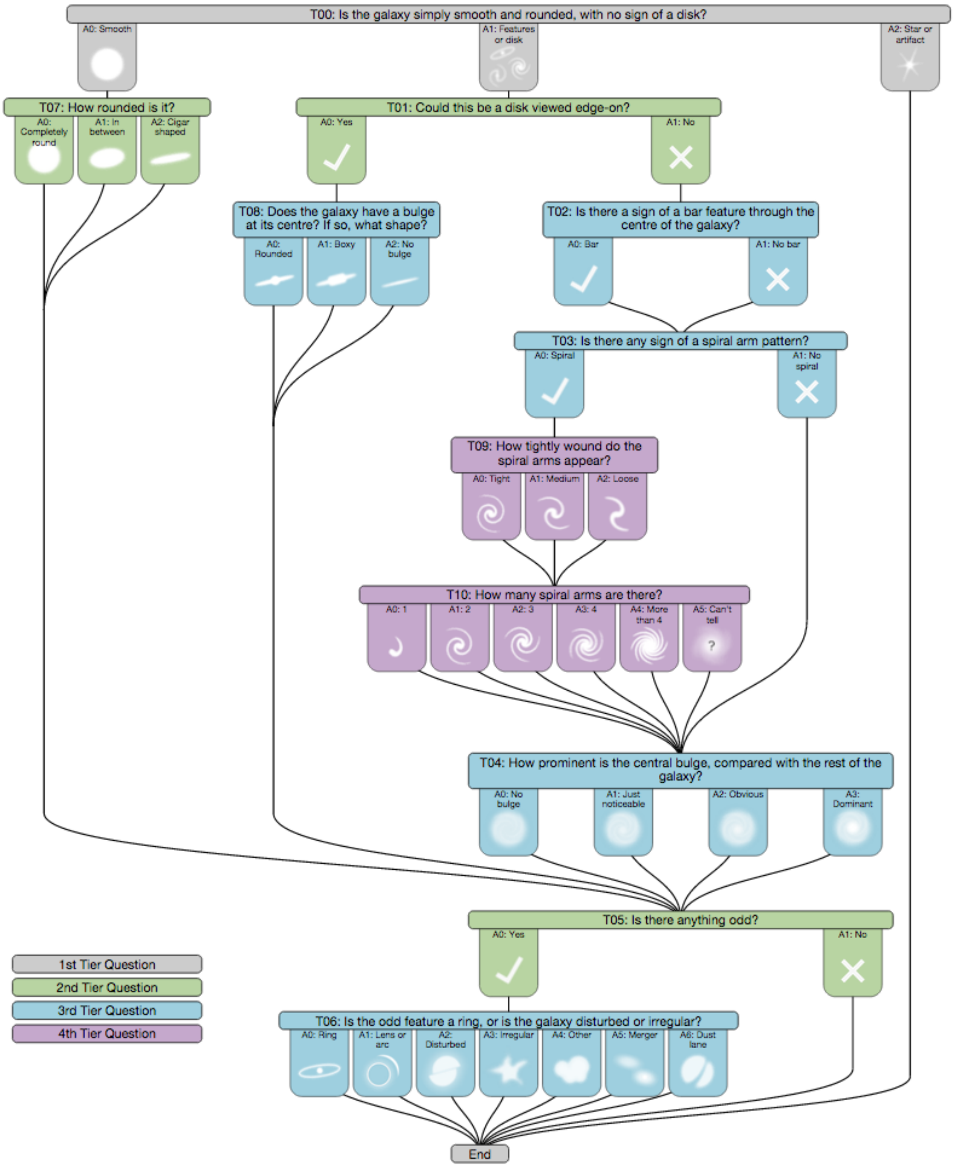


Figure 2.2: The descision tree for Galaxy Zoo 2. The visualization indicates what questions a volunteer would be asked given their previous responses for a given galaxy. This visualization was created by Coleman Krawczyk and is available at `data.galaxyzoo.org/gz_trees/gz_trees.html`

mitted by regulars, highlighting the importance of project design and active community management to encourage volunteers to return to the project.

Raddick et al. (2013) surveyed volunteers of the original Galaxy Zoo project, and found that the single most important shared motivator was the desire to contribute to active research. While this was often not the most important source of motivation, it was ranked very highly by most volunteers. Financial motivation has also been investigated in crowdsourcing; Mao et al. (2013a) studied the behaviour of Amazon Mechanical Turk workers found that financial reward can be used as a way of boosting volunteer speed, however, this comes at the expense of classification quality. They found that per-task pay results in rapid but less accurate classification, and quality suffers further when tasks are difficult. A "wage-like" pay, however, can produce much higher quality classifications, albeit at a slower rate.

One of the primary factors motivating a volunteer to stop engaging with a project is the feeling that their efforts are not helpful or of high-quality (Raddick et al. 2013; Mao et al. 2013a; Ponciano et al. 2014; Segal et al. 2016). Mao et al. (2013a) suggest that when financial incentives are being used, producing low-quality classifications can lead to being blacklisted and financial repercussions.

Volunteer engagement can be boosted in a number of ways; Eveleigh et al. (2014) suggest five design considerations to improve project accessibility to volunteers who are not willing to commit large amounts of time:

1. Allow volunteers to choose their own timeframe, and whether or not to interact with other volunteers.
2. Ensure tasks are short enough to be manageable within a busy lifestyle, by breaking down complex tasks and simplifying interfaces.
3. Publicise scientific outcomes to all volunteers, using social media or newsletters, to help rekindle interest in the project.
4. Use small tasks to pique volunteer interest and draw them into the

project ecosystem, rather than asking for large time investments from the start

5. Ensure volunteers know their contributions are valued, and that their responses are useful and accurate.

Mao et al. (2013b) designed a statistical model to predict volunteer attention and engagement in the Galaxy Zoo project, predicting the number of future tasks a volunteer will complete in a session. Segal et al. (2016) similarly modelled Galaxy Zoo volunteer engagement, and made use of the predictions to investigate the impact of interventions (messages) on volunteer behaviour. They found that a positive message reaffirming the usefulness of a volunteer's work, when timed at a point dictated by the engagement model, can significantly boost the amount of work a volunteer is willing to do in a session.

2.5 Making efficient use of volunteer classifications

Despite the significant increase in scalability of morphological classifications enabled by citizen science projects like Galaxy Zoo, the volume of data available to scientists is continuing to rise faster than our capacity to analyse it. Surveys such as DES (Flaugher, 2005), PanSTARRS (Kaiser et al., 2010) and Hyper Supreme-Cam (Aihara et al., 2018).

To confound this problem, the increased popularity of citizen science as a solution to scalability means that projects are competing more than ever for the attention and time of volunteers. This issue necessitates that volunteer time is spent in the most efficient way possible, through a combination of human and machine intelligence.

2.5.1 The SWAP algorithm

The Galaxy Zoo Express (GZX) framework (Beck et al., 2018) is one such attempt to increase classification efficiency. It makes use of the `Space Warps` analysis pipeline (SWAP, Marshall et al. 2016) algorithm to aggregate volunteer classifications of a galaxy and incorporates repeated training and validation of a random forest classifier, which is used to categorise any unclassified galaxies once it has reached an acceptable level of performance.

SWAP is a Bayesian algorithm first implemented for the gravitational-lens identification project `Space Warps`. It evaluates the accuracy of individual volunteers using subjects with known ground truths, learning a confusion matrix expressing that volunteers ability to discern the presence or absence of a given feature. When that volunteer provides a classification on a galaxy, the posterior probability of that galaxy is updated as per Bayes' theorem, and an image is “retired” (deemed to not require more classification to achieve a consensus) if its classification can be determined with a certainty above some pre-defined threshold.

This classification boost is enhanced using a random forest machine classifier (Breiman, 2001), an ensemble supervised learning technique consisting of a combination of decision tree predictors that is robust to noise and overfitting.

A combination of the SWAP algorithm and random forest classifier effectively reduced the classification time by a factor of eight over the pure GZ2 results, with negligible impact on accuracy and purity relative to the SWAP-only results. The machine does miss a significant number of false positives, though visual inspection suggests that the GZ2 classifications were more likely to be at fault than the machine (S0 galaxies would be considered “smooth” by volunteers but “featured with a disc” by the machine).

2.5.2 Galaxy Zoo enhanced

As detailed in Section 1.3.1, the use of CNNs to automate galaxy morphological classification has been hugely successful, and networks trained on one survey can be applied to others using transfer learning on a small number of labelled galaxies from the new survey (Khan et al., 2019). However, a major drawback is that these methods do not account for the uncertainty present in galaxy classifications, instead assuming no error in the training sample. The labelled galaxies to be used for training must also be compiled before training begins.

Walmsley et al. (2020) improves on this work by using Bayesian CNNs, accounting for varying uncertainty in volunteer responses, and returning galaxy classifications with a full posterior distribution. To account for uncertainty in CNN weights, they use Monte Carlo dropout (Gal et al., 2017) to marginalize over possible CNNs. The results are predictions of how a typical volunteer would have responded had they been presented with a galaxy image.

They combine these Bayesian CNNs with a subject prioritization algorithm: Bayesian Active Learning by Disagreement (Houlsby et al., 2011), an active learning strategy that selects subjects to label based off of the amount of information that subject provides over the model parameters (i.e. eliminate regions of parameter space as quickly as possible).

In this chapter, I have provided a brief history and overview of crowdsourcing and citizen science. I have discussed the evolution of Galaxy Zoo and the Zooniverse and elaborated on research conducted into the motivations of citizen scientists and efforts to best use volunteer classifications within a project. Alongside Chapter 1, which provided an overview of Galaxy formation and evolution, and methodologies by which galaxy morphology is categorised and quantified, this chapter provides the fundamental background to the work performed in this Thesis. The following chapter describes the citizen science project and resulting data analysis methodology that provides the backbone of my work, and allows the scientific exploitation present in Chapter 4.

Chapter 3

Galaxy Zoo Builder

This chapter presents the citizen science project, *Galaxy Builder*, which forms the foundation of this thesis. It is presented in Lingard et al. (2020a). *Galaxy Builder* is a citizen-science project built on the Zooniverse web platform which asks volunteers to perform detailed photometric modelling of spiral galaxies (potentially including bulge, disc, bar and spiral arm components). A project of this kind, allowing volunteers to interact with and model data, had never been attempted inside the current Zooniverse web platform before, so this project involved designing and implementing a model rendering[1] suite inside the existing Zooniverse front-end code-base. As with all citizen science solutions, we had to not only consider the accuracy of the resulting model, but also user experience and engagement in our design decisions. The entire process is summarised in flowchart form in Figure 3.1, with appropriate sections referenced therein.

The closest relative to this project within the Zooniverse ecosystem was the Galaxy Zoo: Mergers project (Holincheck et al., 2016a). This project asked volunteers to help match the morphological properties of an image of merging galaxies to a plethora of restricted three-body simulations. Galaxy Zoo: Mergers required volunteers to download a Java applet to take part in

---

[1] We use the term rendering in a similar manner to that used for computer graphics: to calculate an image from a model or set of rules.

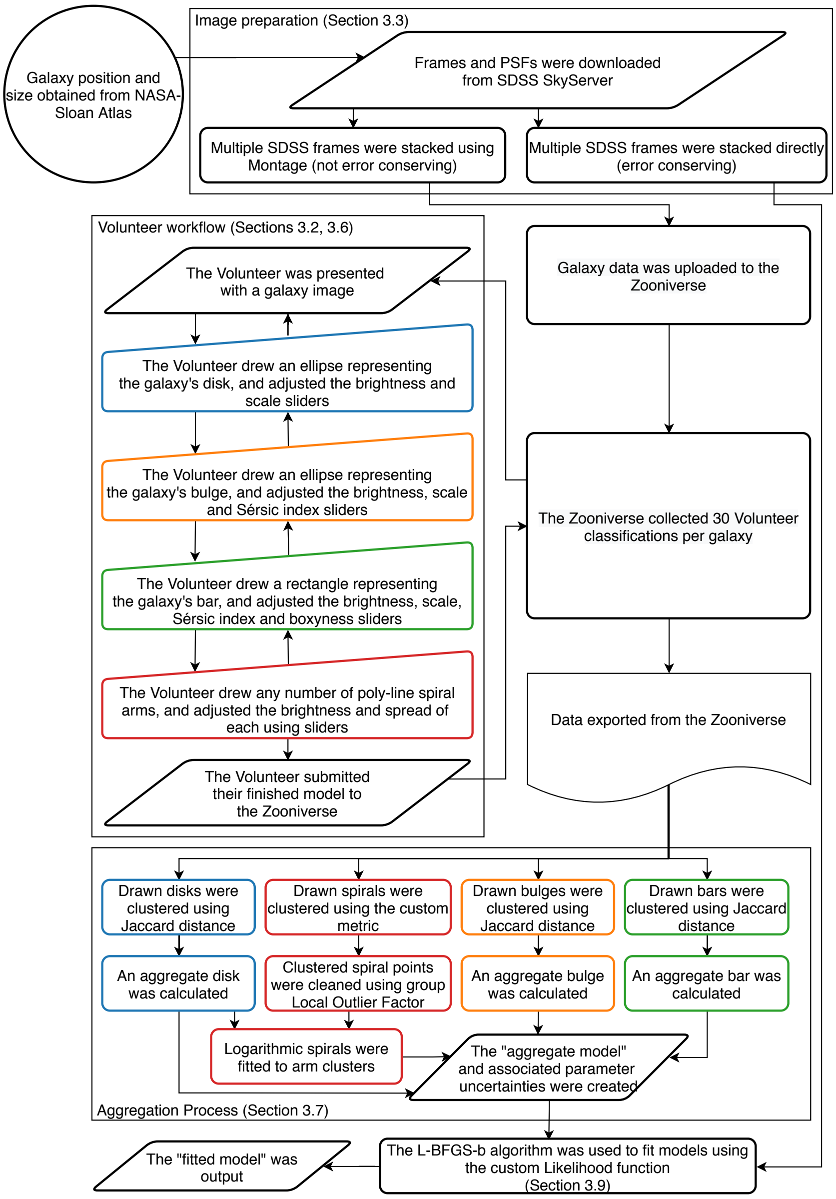


Figure 3.1: Flowchart detailing the entire *Galaxy Builder* process, from image creation, through classification collection using the Zooniverse, to model aggregation and fitting. Processes, manual input, data inputs and exports, and document exports are displayed distinctly. Colours distinguish between component-specific processes (disc in blue, bulge in orange, bar in green and spiral in red). Black nodes relate to the galaxy as a whole.

model selection, while *Galaxy Builder* operates purely inside a web page.

3.1 Project timeline and development

The *Galaxy Builder* project was built inside the Zooniverse's (Simpson et al., 2014) PANOPTES-FRONT-END[2] codebase, using the REACT.JS[3] framework, as well as custom WebGL[4] kernels to enable low-latency photometric galaxy model rendering.

Galaxy Builder entered a Zooniverse beta in late November 2017, this involved a user experience survey for which 260 Galaxy Zoo volunteers were recruited to give feedback on the work-in-progress interface. Screen recordings of volunteers using the interface were also taken, in order to examine the first reactions of novel volunteers to the interface and supplementary material present.

The feedback obtained through the beta led to many user experience improvements and code refactoring, including performing significant work towards upgrading the entire Zooniverse front-end codebase to a newer version of REACT.JS.

A major challenge during the development of the project was finding the right balance between keeping a simple and intuitive interface and workflow while also allowing the freedom and versatility to properly model galaxies. It was also a significant challenge to develop a compelling and simple tutorial for what is one of the most complex projects attempted on the Zooniverse platform. Feedback from expert users was essential to this process as part of the typical beta trial process for Zooniverse projects[5].

Following the beta test period, the project was launched as an official Zooniverse project on the 24th of April 2018. A blog post[6] and press bulletin[7]

---

[2] `github.com/zooniverse/Panoptes-Front-End`

[3] `reactjs.org`

[4] The Web Graphics Library, `www.khronos.org/webgl`

[5] `help.zooniverse.org/best-practices/`

[6] `blog.galaxyzoo.org/2018/06/18/galaxy-builder-results/`

[7] `phys.org/news/2018-06-galaxy.html`

was subsequently released to promote the project.

3.2 The project interface

The *Galaxy Builder* project prompts volunteers to work through the step-by-step creation of a photometric model of a galaxy (described in detail in Section 3.6). A screenshot of the interface can be seen in Figure 3.2, where a residual image is shown. The interface presents a volunteer with three views, which they can switch between at any time: a r-band cutout image of a spiral galaxy (see Chapter 3.3), the galaxy model they have created so far, and the residual between their model and image (shown in blue and yellow).

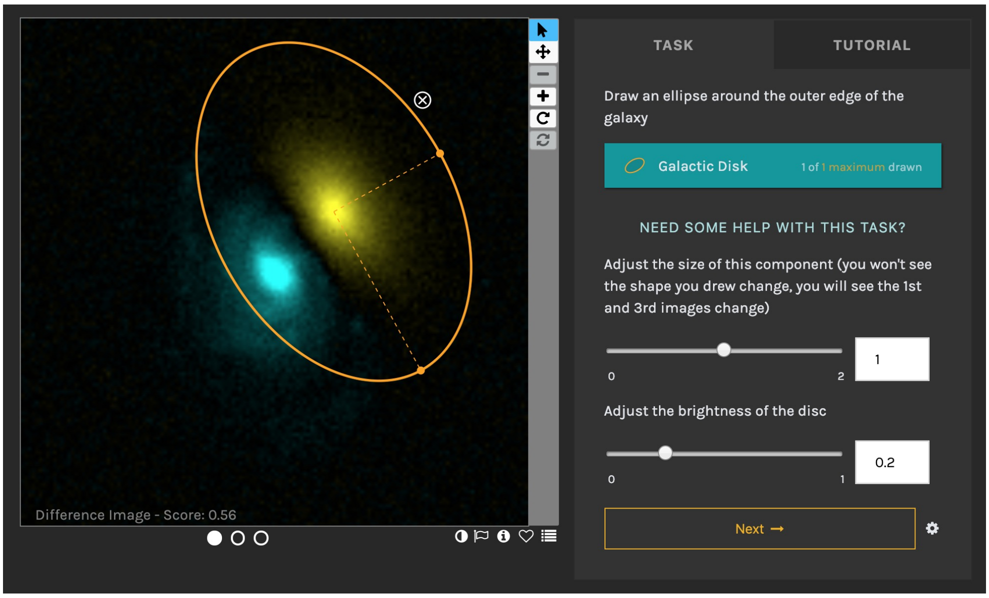
The workflow is designed so that volunteers slowly subtract increasing amounts of light from the galaxy, as is illustrated in Figure 3.3. A tutorial is available that contains a step-by-step guide to completing a classification. At each step, volunteers are asked to first draw a simple isophote, and then make use of a series of sliders to adjust the parameters of the model component (see Section 3.6 for more information).

Volunteers are also guided by a "score", which is tied to the residuals and chosen to increase from zero to some arbitrary value depending on the galaxy; a less noisy and more easily modelled galaxy will have a higher maximum score, with the highest possible score (for a perfectly blank residual) being 100.

To map a residual image to a final score shown to volunteers we use

$$S = 100\exp\left(\frac{-A}{N}\sum_{i=0}^{N}\frac{\operatorname{arcsinh}^2\left(\,|\mathrm{y}_i - M_i|\ /\ B\right)}{\operatorname{arcsinh} B}\right), \tag{3.1}$$

where N is the total number of pixels, y is the cutout of the galaxy, normalized to a maximum value of 1 (y = cutout/max(cutout)), M is the model calculated by volunteers and $A = 300$; $B = 0.6$ are scale factors chosen based on a handful of test galaxies.



Draw an ellipse around the outer edge of the galaxy

Adjust the size of this component (you won't see the shape you drew change, you will see the 1st and 3rd images change)

1

Adjust the brightness of the disc

0.2

Figure 3.2: The *Galaxy Builder* interface. The residual image is being shown, and the volunteer is on the "Disc" task. The drawn disc component (yellow) is offset from the galaxy image (blue) to demonstrate the positive and negative residuals. Where the image equals the model the residual is black. The dots below the residual image allow the user to switch images. The icons to the right allow panning and zooming of the image (rotation was not functional for this project). The icons to the bottom right of the image allow colour inversion of the galaxy cutout, flagging of the image as inappropriate, inspection of galaxy metadata (e.g. sky position, link to SDSS SkyServer), ability to save the image as a favourite and to add to a Zooniverse "collection". The Score shown in the bottom left of the image is calculated using Equation 3.1 and is a rough goodness-of-fit measure.

The primary motivation behind the use of Equation 3.1 is the speed at which it can be computed from the residual image shown to volunteers (which is Arcsinh-scaled in a manner described by Lupton et al. 2004). The worst bottleneck in the rendering pipeline is the transfer of data from the GPU (WebGL) to the CPU (Javascript); in order to maintain a low latency between the volunteer's model changing and the residual image updating (minimising perceived lag), the score calculation is delayed until after residual calculation is complete, meaning the Arcsinh-scaled residuals must be used.

This score comes with the significant drawback of being overly sensitive to small deviations of the model from the galaxy, and not incorporating pixel uncertainties.

3.2.1 Behind the scenes

The REACT.JS web framework allows the creation of a web page through the combination of multiple, hierarchical components, each of which has an internal "state" and can pass arguments ("props") to their children. The Zooniverse PANOPTES-FRONT-END[8] leverages this structure to separate the display of a "subject" (image, video or text shown to volunteers) to the classification interface with which volunteers interact. This design has the advantage that an update to the classification interface, such as a volunteer moving to the next stage of a workflow, does not cause the whole page to be recalculated and re-rendered.

The *Galaxy Builder* project necessitated a reworking of this architecture, as the input of a volunteer dynamically alters the contents of the subject area (residuals updating when the model changes). Such an alteration was non-trivial, as performance across the site would be significantly impacted by unnecessary updates. Once the required components had been introduced, it was possible to parse the work-in-progress volunteer model and render the result and residuals using an HTML canvas element, and the WebGL

---

[8] `github.com/zooniverse/Panoptes-Front-End/`

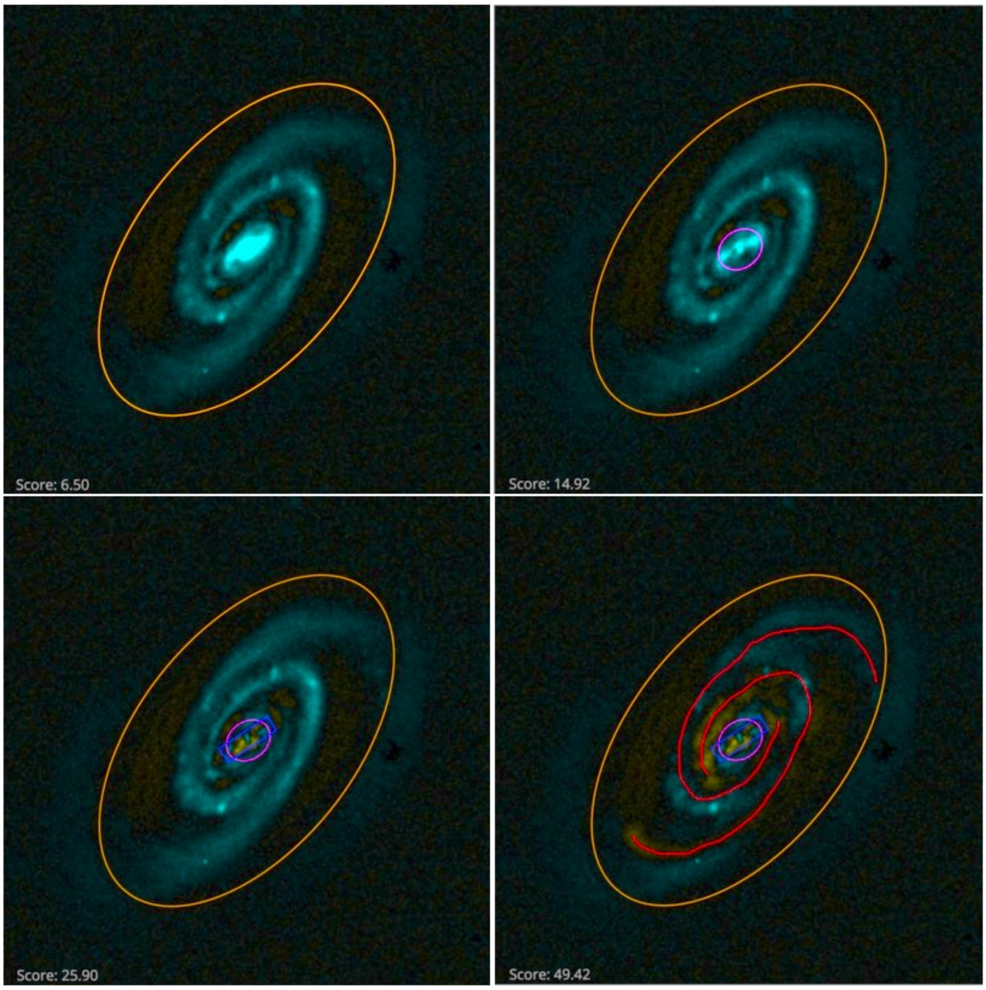


Figure 3.3: Figure demonstrating the desired result of each step of the modelling process, as seen from the residual image provided to volunteers. The top left panel shows the galaxy after only a disc component has been added: the top right contains a disc and a bulge; the bottom left has a disc, bulge and bar; the bottom right is the finished model with a disc, bulge, bar and spiral arms. The image shown is SDSS J104238.12+235706.8. The brightness and contrast of this image have been edited to improve visibility in print.

functional abstraction package REGL[9].

3.3 Sample selection: images and ancillary data

As a proposed solution to the problem of fitting multi-component and complex galaxies, *Galaxy Builder* finds a niche with a sample of disc galaxies with spiral features. One such sample is the *stellar mass-complete sample* in Hart et al. (2017), which is a sample of relatively face-on spiral galaxies ($b/a > 0.4$) with and without bars and selected to be complete across stellar masses $9.45 < \log(M_\star/M_\odot) < 11.05$. The test sample used for the Galaxy Builder project was therefore selected from the Hart et al. (2017) sample of relatively face-on spiral galaxies.

The morphological information required to select spiral galaxies came from the public data release of Galaxy Zoo 2 (Willett et al. 2013a, hereafter GZ2). Each response to a GZ2 morphology question is allocated a p value ranging from 0 to 1, where 0 indicates no volunteers responded positively to that question and 1 indicates all volunteers who classified that galaxy responded positively (for example, $p_{\rm bar} = 0.5$ would indicate 50% of volunteers said a bar was present in a galaxy). Photometric measurements used for selection came from the NASA-Sloan Atlas (Blanton et al. 2011, hereafter NSA). The *stellar mass complete sample* is constructed using the set of criteria detailed in Table 3.1.

The *stellar mass-complete sample* was split into smaller sub-samples, each containing 100 galaxies. In an iterative process, each sub-sample was chosen to contain the 60 lowest redshift unclassified galaxies, and 40 random unclassified galaxies. This was done to ensure an early sample would be available to work with given the *a priori* unknown rate at which volunteers would provide classifications. Due to time constraints, classifications were only collected for two unique sub-samples. The mass-redshift relation of galaxies in the *stel-*

---

[9] `regl.party`

Table 3.1. The selection criteria used in Hart et al. (2017) to create the *stellar mass-complete sample* of 6222 spiral galaxies.

Description	Value
Face-on spiral morphological selection.	GZ2 $p_{\mathrm{features}} \times p_{\mathrm{not\ edge\ on}} \times p_{\mathrm{spiral}} \geq 0.5$
Redshift limits.	$0.02 < z < 0.055$
Face-on galaxy selection using g-band axial ratio.	$(b/a)_g > 0.4$
Mass limits for rough volume limited sample.	$9.45 < \log(M_*/M_\odot) \leq 11.05$
Mass limits for complete sample[a]	$2.07\log(z) + 12.64 < \log(M_*/M_\odot) < 2.45\log(z) + 14.05$

[a] Stellar masses from Mendel et al. (2014)

[a] Using the method of Pozzetti et al. (2009) and limits calculated by Hart et al. (2017)

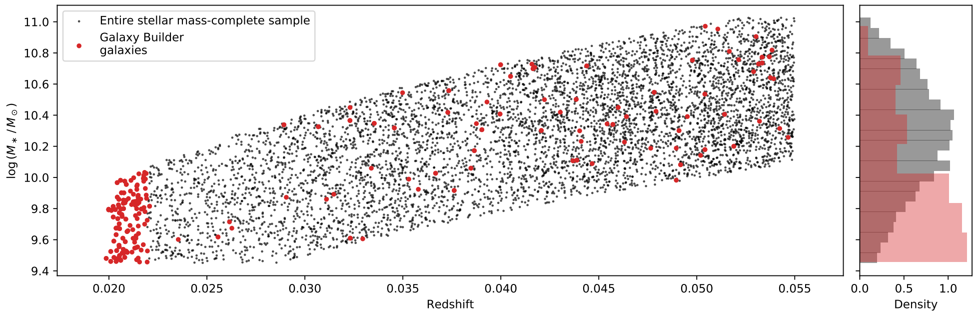


Figure 3.4: Redshift against total galaxy stellar mass for all galaxies in the *stellar mass-complete sample*, with the 198 galaxies considered here highlighted in red. The distribution of stellar masses is shown in the right panel for the total sample and for the galaxies considered here. It is evident that the galaxies for which classifications were collected are not complete in stellar mass, but it is possible to select a further subset that would be.

lar mass-complete sample from Hart et al. (2017) can be seen in Figure 3.4, with galaxies present in this work highlighted in red. Stellar Masses were calculated by Mendel et al. (2014).

In the first two sets of 100 galaxies, 1% of galaxies (i.e. 2 images) failed to run through the image preparation process, due to an error when attempting to montage multiple frames. The root cause of this error is unknown, but it leaves a sample of 198 galaxies with images (the *test sample*, 98 of which are repeated in a validation subset) that are considered in this thesis, in order to explain the method used and test the reliability of the models obtained.

3.3.1 Image and modelling metadata extraction

The galaxy data shown to volunteers in the *Galaxy Builder* project came in two forms: A grey-scale image cutout of the galaxy and a JSON file containing rendering information for the web-interface.

Both forms of data were obtained using a similar process:

1. A montage of multiple r-band corrected frames from the SDSS DR13

(Albareti et al., 2017) data release was created. To combine multiple FITS images, Astropy (Astropy Collaboration et al., 2018) and the MONTAGE (Jacob et al., 2010) software packages were used.

2. This montage was cropped to four times the Petrosian radius of the galaxy.
3. The SEXTRACTOR software (Bertin & Arnouts, 1996) was used to identify regions containing secondary sources (foreground stars, other galaxies) and generate a mask.
4. A PSF was obtained from the relevant Sloan r-band `psField` file, extracted at the central position of the galaxy (Stoughton et al., 2002).
5. The JSON file was written containing the cut-out data and the 2D boolean mask obtained from the source extraction process. This file also contained other metadata needed for the rendering process (PSF, the size of the PSF array, and the width and height of the image).
6. Another JSON file containing simply the information used to render the volunteer's model (image size and PSF) was created.
7. An arcsinh-stretch was applied to the masked cutout (as described by Lupton et al. 2004). It was then saved as a grey-scale image.

The decision to use r-band images in our subject set was due to its higher signal-to-noise than other bands.

Once a sub-sample had been created, the Zooniverse's PANOPTES-PYTHON-CLIENT[10] was used to upload them as a subject-set to the Zooniverse.

The reprojection performed by MONTAGE has a smoothing effect on the data, and thus does not conserve errors. We, therefore, create a separate stacked image, sigma image and corresponding pixel mask, using the same r-band corrected frames present in the montage. These images were not shown to volunteers but were used for model fitting and comparison.

---

[10] `github.com/zooniverse/panoptes-python-client`

3.4 Stacking of multiple SDSS frames

The reprojection performed by MONTAGE has a smoothing effect on the data, and thus does not conserve errors. A separate stacked image, sigma image and corresponding pixel mask was, therefore, created using the same r-band corrected frames present in the montage. These images were not shown to volunteers but were used for model fitting and comparison.

All data required for sigma image creation for stacked frames came from the corrected frames, as detailed in the frame datamodel[11]. For each pixel in an SDSS frame, we have

$$\frac{I}{C} = \frac{n}{g} - S + V, \tag{3.2}$$

where I represents the sky-subtracted, corrected image (in units of nanomaggies[12]), C reprents the calibration image, n is the number of electrons captured, g is the gain, S is the Sky value (data units) and V is the dark current, $V = 0 \pm \sqrt{v}$ (v being the dark variance).

Given Poisson error,

$$\sigma_n = \sqrt{n}. \tag{3.3}$$

If multiple frames are stacked, given N observations of a pixel

$$\begin{aligned} n_{\mathrm{total}} &= \sum_i n_i = \sum_i g_i \left(\frac{I_i}{C_i} + S_i - V_i\right), \\ &= \sum_i \frac{g_i}{C_i} I_i + \sum_i g_i \left(S_i - V_i\right) = \sigma^2_{n_{\mathrm{total}}}. \end{aligned} \tag{3.4}$$

---

[11] `data.sdss.org/datamodel/files/BOSS_PHOTOOBJ/frames/RERUN/RUN/CAMCOL/frame.html#example`

[12] A nanomaggy is a linear unit of flux defined as the flux of a source relative to a standard source, where the standard source defines the zero point. The SDSS standard source is close to that of the AB magnitude ($\sim 3631\,\mathrm{Jy}$), meaning that one nanomaggy is approximately $3.631 \times 10^{-6}\,\mathrm{Jy}$.

This is ideal, and is the level that many fitting software packages work at. As we wish to return to working in units of nanomaggies on a stacked image, further calculation is needed:

$$I = \frac{1}{N}\sum_i I_i, \tag{3.5}$$

$$I = \frac{1}{N}\sum_i C_i\left(\frac{n_i}{g_i} - S_i + V_i\right), \tag{3.6}$$

And so, assuming no covariances between the uncertainties,

$$\sigma_I^2 = \frac{1}{N^2}\sum_i \frac{C_i^2}{g_i^2}\sigma_{n_i}^2 + \frac{1}{N^2}\sum_i C_i^2\sigma_{S_i}^2 + \frac{1}{N^2}\sum_i C_i^2\sigma_{V_i}^2. \tag{3.7}$$

We treat the sky value as a constant, such that $\sigma_{S_i}^2 = 0$. Substituting $\sigma_{n_i}^2 = n_i$ gives

$$\sigma_I^2 = \frac{1}{N^2}\sum_i \frac{C_i^2}{g_i^2}n_i + \frac{1}{N^2}\sum_i C_i^2 v_i, \tag{3.8}$$

$$\sigma_I = \frac{1}{N}\sqrt{\sum_i C_i^2\left(\frac{n_i}{g_i^2} + v_i\right)}. \tag{3.9}$$

Note that this is identical to saying

$$\sigma_I^2 = \frac{1}{N^2}\sum_i \sigma_{I_i}^2. \tag{3.10}$$

Pixel covariances introduced by the PSF are not accounted for, which can vary significantly between frames. An image with reduced noise is highly desirable for photometric model fitting, and since this method does not involve the reprojections performed by MONTAGE, we concluded the resulting image and associated uncertainties generated are the best available option moving forwards.

3.5 Choice of retirement limit

The number of independent answers needed to create reliable and reproducible aggregate classifications was not known at the start of this project. An initial experiment with collecting 10 classifications per galaxy demonstrated that this was insufficient; further experimentation with a diverse range of galaxy types (most with prominent spiral features including grand-design and flocculent arms) revealed 30 classifications per galaxy was sufficient.

The entire *test sample* of 198 galaxies was then presented to users, with 30 classifications collected per galaxy. In addition, one of the subsets was presented a second time, thus providing a validation subset to measure consistency between sets of 30 classifications on the same galaxies.

Nine synthetic images of galaxies were also created, containing various combinations of components available to volunteers and a spread of possible parameters. These synthetic galaxies were based off of a set of target galaxies from *Galaxy Builder* and designed to be as realistic as possible, including the addition of realistic noise and pixel masks. This set of synthetic images is shown in Figure 3.5 and was used to calibrate our aggregation and fitting methodology and thus is referred to as the *calibration subset*.

3.6 The galaxy model

Our chosen galaxy model was largely based on components described in Peng et al. (2002). The modelling code ignores masked regions identified as secondary sources by SExtractor. It over-samples the bulge, disc and bar components by a factor of five and performs PSF convolution using a PSF obtained from the relevant Sloan r-band `psField` file, extracted at the central position of the galaxy (Stoughton et al., 2002). The model created by a volunteer could be chosen from

1. One exponential, ellipsoidal disc.

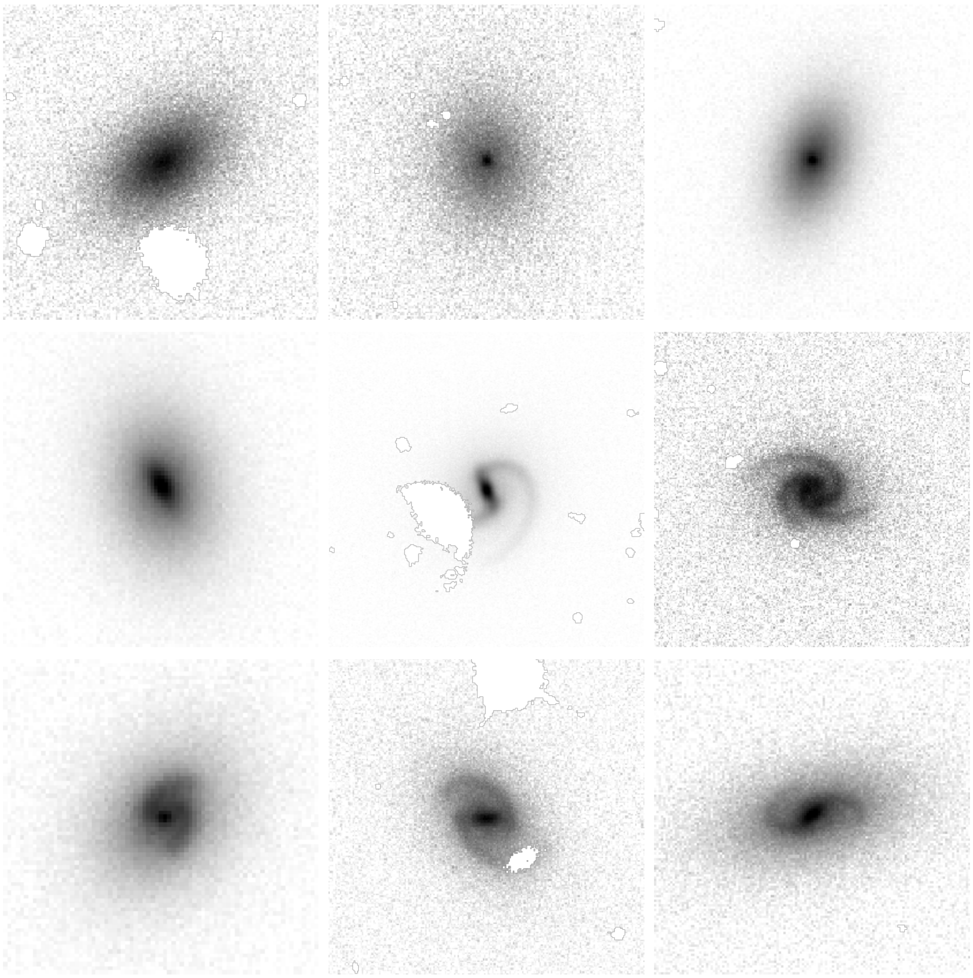


Figure 3.5: Arcsinh-stretched images of the synthetic galaxies present in the *calibration subset*. These galaxies were designed to look as realistic as possible, while being described perfectly by the model available to volunteers.

2. One ellipsoidal Sérsic bulge, with n chosen by volunteers.
3. One Sérsic bar with a "boxiness" modifier (as described in Peng et al. 2002), with n and c chosen by volunteers.
4. Any number of freehand poly-line[13] spiral arms, as described below.

3.6.1 Spiral arm model

Each spiral arm is modelled using a poly-line drawn by the volunteer. The brightness of a spiral arm at any point is given by the value of a Gaussian centred at the nearest point on any drawn poly-line, with volunteers able to choose the Gaussian width and peak brightness using sliders. Radial falloff was added by multiplying by the value of the previously added exponential disc, though volunteers could change the half-light radius of this falloff disc.

3.7 Classification aggregation methodology

In this section, we will use the galaxy UGC 4721, a two-armed barred spiral galaxy at $z = 0.02086$ classified by de Vaucouleurs et al. (1991b) as SBcd, to illustrate the data reduction and aggregation methodology. For UGC 4721 32 classifications were received, containing 28 discs, 24 bulges, 17 bars and 47 drawn spiral arm poly-lines (four classifications did not contain spirals, seven contained one spiral arm, fourteen contained two arms, six contained three arms and one contained four arms). These annotations can be seen in Figure 3.6, overlaid on the greyscale r-band image of the galaxy.

3.7.1 Aggregation of volunteer models

Aggregate model calculation was done on a component-by-component basis, rather than per classification, i.e. clustering of discs was performed independently to that of bulges, bars and spirals. None of the slider values were

---

[13] a poly-line, or polygonal chain, is a series of connected line segments.

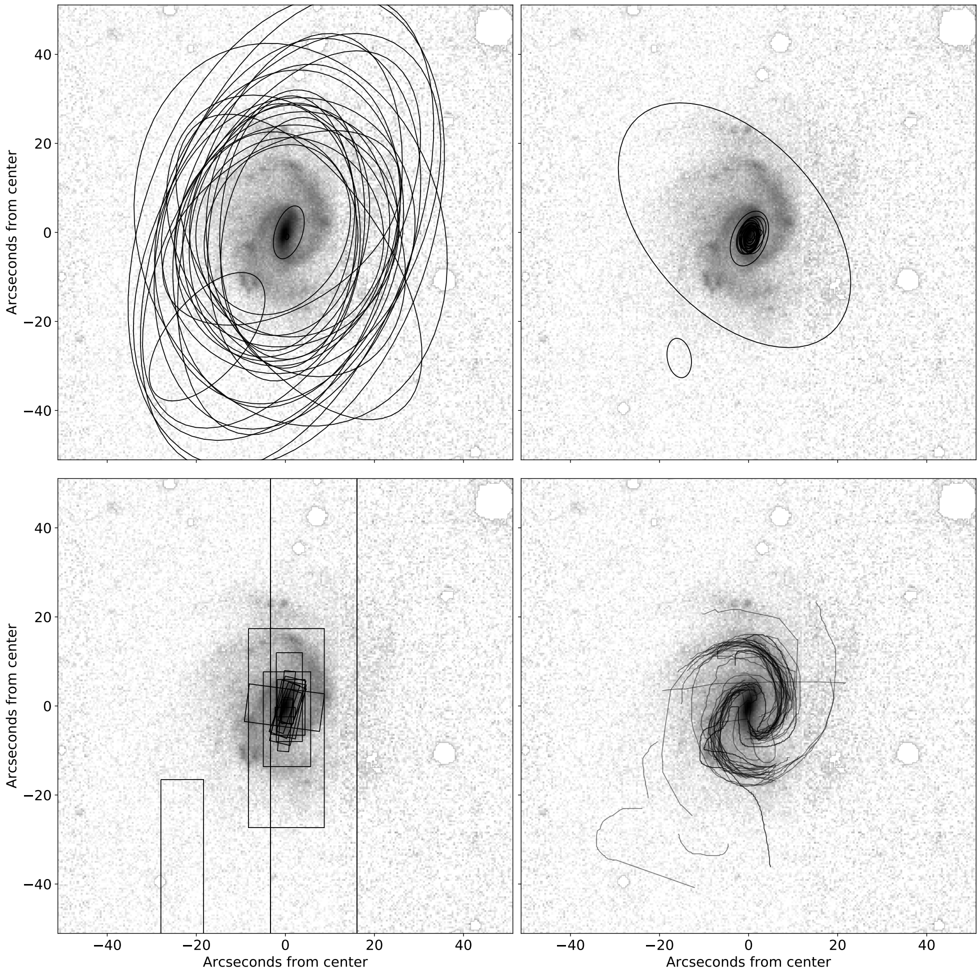


Figure 3.6: Components drawn by volunteers for UGC 4721. The top left panel shows drawn discs, top right shows drawn bulges, bottom left shows drawn bars and bottom right shows drawn spiral arms. Discs, bulges and bars are displayed at twice their effective radii. These raw marks are subsequently aggregated to produce a consensus value for each galaxy component.

taken into account, only the shape drawn by the volunteers. Disk classifications were doubled in effective radius to correct for a systematic error in disk size observed in the classifications received for the *calibration subset*. Model parameters were restricted to be within the limits shown in Table 3.2 (deemed to be the physically acceptable bounds). All components were transformed from the coordinate space of the MONTAGE-created images to the more accurate stacked images created for fitting. Clustering was performed using the Jaccard distance measure (also known as the intersect-over-union distance, or IOU distance), which is a simple metric determining the relative shared area of two sets:

$$d_J(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|}. \tag{3.11}$$

The algorithm chosen to perform clustering was the Density-Based Spatial Clustering of Applications with Noise (DBSCAN, Boonchoo et al. 2018) algorithm, due to its robustness and speed. Scikit-learn (Pedregosa et al., 2011) was used to implement the algorithm. In DBSCAN the core of a cluster is defined as a group of at least N_{min} items that are all within a distance ϵ of each other. Additionally, any points within a distance ϵ of a cluster's core are also associated with the cluster.

3.7.2 Disc, bulge and bar clustering

The disc clustering hyperparameters were selected such that a disc is clustered for all galaxies, and the bulge hyperparameters to most successfully recover the morphology of galaxies in the *calibration subset*. The value of ϵ used to cluster bars was tuned such that the aggregate model best agreed with GZ2 p_{bar} ($p_{\text{bar}} < 0.2$ implying no bar and $p_{\text{bar}} > 0.5$ implying a definite bar; as discussed in Masters et al. 2012; Skibba et al. 2012 and used by Willett et al. 2013b; Kruk et al. 2018). The values chosen for ϵ were 0.3, 0.4, 0.478 for the disc, bulge and bar; N_{min} was set to 4 for all three of these components.

Table 3.2. The maximum, minimum and default values for model parameters. Model parameters are defined in Section 3.9. Note that some parameters were allowed to overflow when fitting, for instance an axis ratio greater than 1 (signifying a swap of major and minor axis) was allowed, and corrected for once fitting reached completion. This helped avoid the optimizer encountering parameter bounds and failing to converge. Component position angle (ψ) and spiral pitch angle (ϕ) were similarly unconstrained.

Component	Parameter	Tuning Minimum Bound	Tuning Maximum Bound
disc, bulge, bar	μ_x	-inf	inf
	μ_y	-inf	inf
	ψ	-inf	inf
disc	q	0.25	1.2
	R_e	0	inf
	I_e	0	inf
bulge	q	0.6	1.2
	$R_e\ /\ R_{e,\text{disc}}$	0.01	1
	$(B/T)_r)$	0	0.99
	n	0.5	5
bar	q	0.05	0.5
	$R_e\ /\ R_{e,\text{disc}}$	0.05	1
	$(B/T)_r$	0	0.99
	n	0.3	5
	c	1	6
spiral	I_s	0	inf
	A	0	inf
	spread	0	inf
	ϕ	-85	85
	$\theta_\text{min},\ \theta_\text{max}$	-inf	inf

The aggregate component is defined to be the shape that minimises the sum of Jaccard distances to each of the members of the cluster. For our example galaxy, UGC 4721, clustered and aggregate components can be seen in Figure 3.7.

3.7.3 Spiral arm clustering

To cluster drawn spiral arms, we define a custom separation measure to represent how far away one poly-line is from another. This measure was chosen to be the mean of the squared distances from each vertex in a poly-line to the nearest point (vertex or edge) of another poly-line, added to the mean of the squared distances from the second poly-line to the first. This separation measure was used inside the DBSCAN algorithm to cluster these drawn lines, after removing any self-intersecting drawn arms (as this was deemed an easy method to filter out "bad" classifications). Values of 0.001 and 4 were used for the ϵ and `min_samples` hyper-parameters respectively.

Once spiral classifications on a galaxy have been clustered into the physical arms they represent, the points are deprojected using the axis ratio and position angle of the aggregated disc. The deprojection method assumes a thin disc and stretches the ellipsoidal minor axis to match the major axis.

Deprojected points within each drawn poly-line are converted to polar coordinates and unwound to allow model fitting. These unwound points are then cleaned using the Local-outlier-factor algorithm (LOF, Breunig et al. 2000). For each drawn poly-line in the cluster, the LOF algorithm was trained on all points not in that arm, and then used to predict whether each point in the arm should be considered an outlier. In this way, our data is cleaned while respecting its grouped nature. The points removed as outliers for the example galaxy are shown in the bottom right panel of Figure 3.7.

For each arm cluster in each galaxy, a logarithmic spiral model was fitted using Bayesian Ridge Regression, performed using the Scikit-learn Python package. A logarithmic spiral was chosen due to its simple form with a constant pitch angle. Hyperpriors on the noise parameter were chosen by

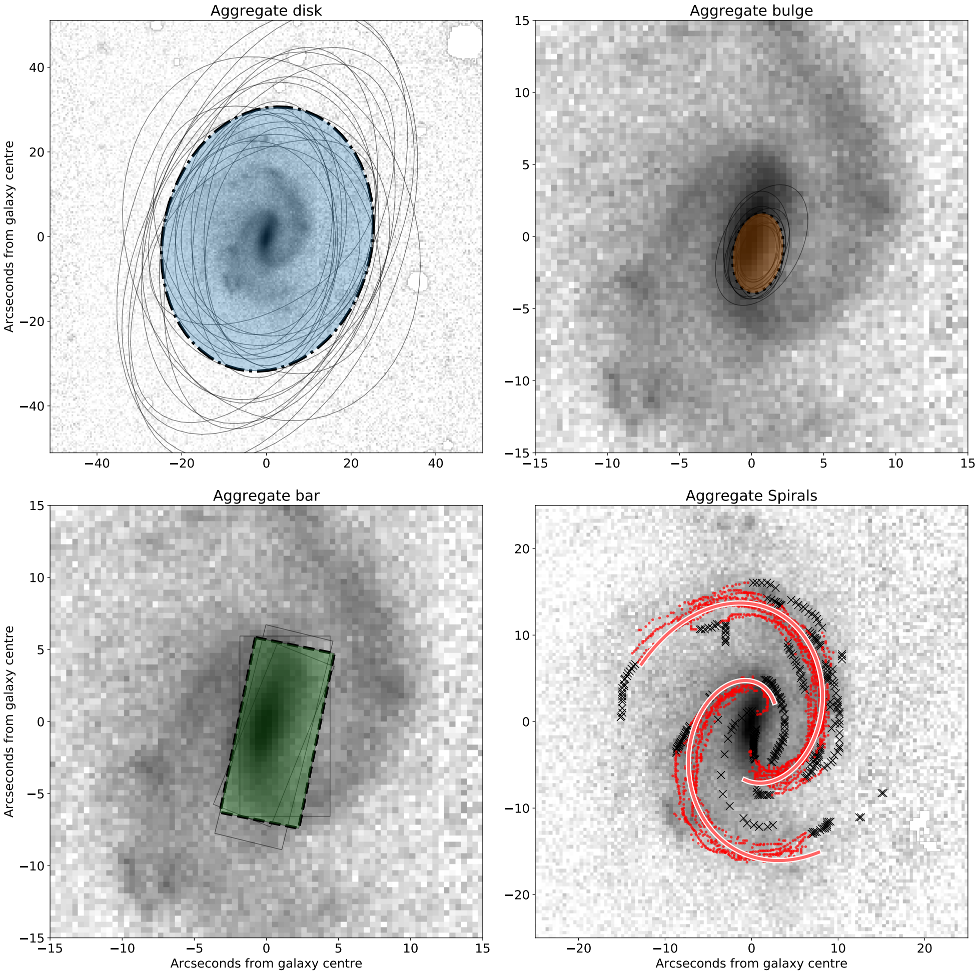


Figure 3.7: Calculated aggregate components for UGC 4721. The aggregate disc is shown using a dot-dashed line and blue fill in the upper left panel, the aggregate bulge with a dotted line and orange fill in the upper right panel, the aggregate bar using a dashed line and green fill in the lower-left panel and the aggregate spiral arms are plotted as red lines in the lower right panel. Sérsic components are displayed at twice their effective radii. Black crosses in the lower right panel indicate spiral arm points that were identified as outliers and removed during cleaning (described in Section 3.7.3). The aggregated components agree well with the underlying morphology, despite the noisiness of the classifications received.

fitting a truncated gamma distribution (Zaninetti, 2014) to the spiral width slider values returned by volunteers (ignoring sliders left at the default or moved to the extremes of allowed values). Any logarithmic spirals within a distance of 0.0005 (given by the clustering metric) were deemed to be from the same arm and thus their classifications were merged and a log-spiral recalculated.

We do not assume that every arm in a galaxy has the same pitch angle. To obtain a single value for the pitch angle of a galaxy, we take the length-weighted average pitch angle of all arms detected in the galaxy (as used by Davis & Hayes 2014a).

The galaxy model for UGC 4721 obtained through aggregation can be seen in the bottom left panel of Figure 3.8.

3.8 Error estimation of aggregate models

As all components in a cluster can be viewed as volunteers' attempts at modelling the true underlying component, the sample variance of the parameters of these shapes can be used as a measure of confidence in the parameters present in the aggregate result. These are highly sensitive to clustering hyperparameters, and are only valid for a component's position, size and shape. Figure 3.7 illustrates the variance in clustered shapes for our example galaxy (UCG 4721); we see a large variation in the clustered discs, and much closer agreement on the bulge and bar size and shape.

3.9 Model fitting

The final step in creating *Galaxy Builder* models is a numerical fit to fine-tune parameters. This fitting was performed using the L-BFGS-b algorithm (Byrd et al. 1995; the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm with bounds), implemented in SCIPY (Jones et al., 2001). The L-BFGS-b is a limited-memory quasi-Newton optimizer which uses an estimate of the inverse Hessian matrix to guide its exploration of parameter space, it scales well with increasing numbers of parameters, and efficiently incorporates simple box constraints on variables.

We minimize a custom likelihood function that assumes Gaussian error on pixel values and incorporates the priors on parameters obtained from clustering. We use the same model as used by volunteers in the online interface (with altered limits), with spiral arms restricted to being logarithmic spirals relative to the disc, and without the ability to change the relative falloff of spiral arms.

Fitting *Galaxy Builder* models using only χ^2_ν, or some other measure of residuals (mean squared error, median absolute error, the R^2 coefficient of determination), does not make use of the information available in the spread of volunteer classifications. Initial tests showed that models would still converge to unphysical results, even when provided physical and visually well-motivated starting points. Instead, we assume Normal priors on component parameters determined from clustering (μ_x, μ_y, q, Re), with the spread given by the spread in the clustered values. We therefore have that our final log-likelihood (to be maximised) is the sum of the gaussian log-likelihood of the residuals given the pixel uncertainty and the gaussian log-likelihood of the variation in parameters, given their uncertainty. This loss function incorporates our prior beliefs on parameters, and helps constrain models to physical regions of parameter space.

The model being rendered is the PSF-convolved sum of the separate components and outputs an (N_x, N_y) image. The disc, bulge and bar are variations on the boxy Sérsic profile:

$$I_{\text{sersic}}(\vec{P}) = I_e \exp\left\{-b_n\left[\left(\frac{r\,(\vec{P})}{R_e}\right)^{1/n} - 1\right]\right\} \tag{3.12}$$

where

$$r\,(\vec{P}) = \left|\begin{pmatrix}\frac{1}{q} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\cos\psi & -\sin\psi\\ \sin\psi & \cos\psi\end{pmatrix}\left(\vec{\mu}-\vec{P}\right)\right|_c. \tag{3.13}$$

The disc is resticted to $n = 1; c = 2$, bulge to $n \in (0.5, 6); c = 2$ and bar to $n \in (0.5, 6); c \in (0.5, 6)$.

The Sérsic components are actually rendered at 5x the image resolution, and downsampled using the mean pixel brightness. This is a widely used method of approximating the true pixel value, which is an integration over the area of sky inside the pixel: for a pixel of size (δ_x, δ_y),

$$I_{\mathrm{pix}}(\vec{P}) = \frac{1}{\delta_x\delta_y}\int_{-\delta_y/2}^{\delta_y/2}\int_{-\delta_x/2}^{\delta_x/2}\mathrm{d}x\mathrm{d}y\; I_{\mathrm{sersic}}\left(\vec{P}+\begin{pmatrix}\delta_x\\ \delta_y\end{pmatrix}\right). \tag{3.14}$$

Spiral arms were restricted to be logarithmic with respect to the inclined, rotated disc. They were rendered in a similar manner to the online interface; using the nearest distance from a pixel to a calculated logarithmic spiral.

An inclined, rotated log spiral requires parameters brightness I_s, spread s, minimum and maximum θ (θ_{min} and θ_{max}), an amplitude A, pitch angle ϕ, position $\vec{\mu}$, position angle ψ and axis ratio q, where $\vec{\mu}$, ψ and q are inherited from the disc component.

The distance from a pixel to a logarithmic spiral is given by

$$D_{\mathrm{s}}(\vec{P}) = \min_{\theta\in[\theta_{\mathrm{min}},\theta_{\mathrm{max}}]}\left\|\vec{P}-\vec{\mu}-Ae^{\theta\tan\phi}\begin{pmatrix}\cos\psi & \sin\psi\\ -\sin\psi & \cos\psi\end{pmatrix}\begin{pmatrix}q\cos\theta\\ \sin\theta\end{pmatrix}\right\|^2. \tag{3.15}$$

In practice the spiral distance was approximated using the distance to a poly-line with 200 vertices, as solving the above minimization for each pixel at each fitting step is computationally intractable. We also adjust A, θ_{min} and θ_{max} to account for the rotation of the disc component from its starting value, in order to prevent spirals inadvertently moving far from starting locations for face-on discs (which have poorly constrained position angles). These adjustments are

$$\begin{aligned}
A' &= A e^{\Delta\psi \tan\phi}, \\
\theta'_{\text{min}} &= \theta_{\text{min}} - \Delta\psi, \\
\theta'_{\text{max}} &= \theta_{\text{max}} - \Delta\psi.
\end{aligned} \tag{3.16}$$

The pixel brightness is then calculated as

$$I_{\text{spiral}}(\vec{P}) = I_{e,\,\text{disc}}(\vec{P}) \times I_s \exp\left(\frac{-D_{\text{s}}(\vec{P})}{2s^2}\right). \tag{3.17}$$

For the fit, disc I_e is parametrized as the Sérsic total luminosity, given by

$$L_{\text{tot}} = I_e R_e^2 \; 2\pi n \frac{e^{b_n}}{(b_n)^{2n}} \Gamma(2n). \tag{3.18}$$

Bulge (bar) I_e is reparametrized as “bulge (bar) fraction”, which is defined as

$$F_{\text{bulge}} = \frac{L_{\text{bulge}}}{L_{\text{disc}} + L_{\text{bulge}}}, \tag{3.19}$$

and is limited to be between 0 and 1. Disc luminosity is allowed to take any value greater than or equal to zero.

Similarly, bulge and bar effective radius are reparametrized as their scale relative to the disc ($R_e = R_e / R_{e,\,\text{disc}}$). Bulge and bar are also restricted to have the same position.

The model rendering and fitting code was written up using Google’s JAX package (Bradbury et al., 2018), which allows GPU-optimization and automatic gradient calculation, enabling quick and accurate calculation of the jacobian matrix needed for the L-BFGS-b minimization algorithm.

We initially fit only for the brightnesses of components, and then simultaneously for all free parameters of all components. The result of the fit,

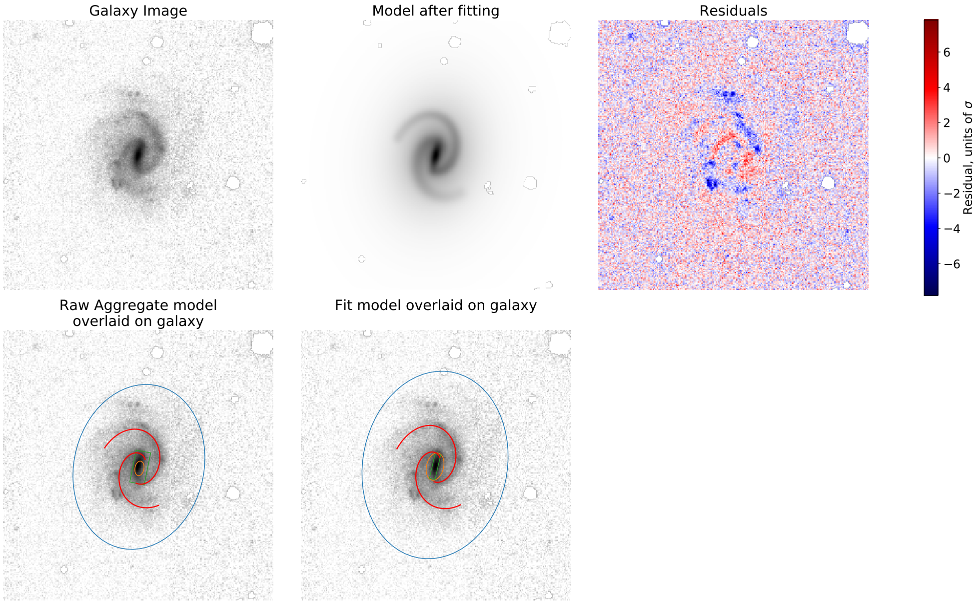


Figure 3.8: Effect of fitting on the aggregated models. The top left panel shows an Arcsinh-scaled image of the galaxy being fit (UGC 4721), the top middle shows the final model obtained (with the same limits and scaling as the galaxy image) and the top right shows the difference between the two images, in units of pixel uncertainty. The bottom panels show a simple representation of the model before and after tuning, overlaid on the galaxy image from the top-left panel. With minimal change to the aggregated components, we recover a detailed model that matches the galaxy exceptionally well, as evident in the residuals.

including the final photometric model for UGC 4721, can be seen in Figure 3.8. The secondary components have been accounted for well, and the model has a sensible reduced chi-squared value of 1.176, where we have approximated degrees of freedom as the number of unmasked pixels present in the galaxy image (similar to Galfit).

We use the errors described in Section 3.8 as parameter uncertainties, as we feel an approach based on the local curvature of the likelihood-space (as used by `Galfit`) would likely fall foul of the issues described in the introduction and thus be an under-estimate. This decision means we do not have uncertainties for some parameters.

We remove two models for which a fit did not converge.

3.10 Results

In this section we present *Galaxy Builder* models for 198 galaxies, from the aggregation of user classifications (aggregate models), and with parameters fine-tuned by a numerical fit (fitted models). We explore the consistency with which volunteers modelled galaxies, the accuracy of the aggregate models, and compare the aggregate and fitted models to comparable results in the literature.

3.10.1 The calibration set

The calibration subset was a set of nine synthetic galaxy images created from *Galaxy Builder* models, which were then re-run through the *Galaxy Builder* process. These galaxies were used to fine-tune clustering and fitting hyperparameters (See Section 3.7.1), as the ground truth was known. Our ability to recover morphology accurately is essential validation for our ability to recover good photometric models of galaxies.

The scatter between true and measured parameters is shown in Figure 3.9; these results highlight the importance of good priors to obtain accurate fits of complex photometric models. In more detail, the models recovered for the nine synthetic galaxy images demonstrate that:

1. Model parameters were generally recovered to a high degree of accuracy

2. We successfully recover all spiral arms present, and do not receive any false positives. The spiral pitch angles obtained through aggregation vary by $< 9°$ from the true values, with fitting improving this error slightly.

3. Volunteers systematically use elongated bulges to model bar components. This resulted in two false positives for bulge presence in the

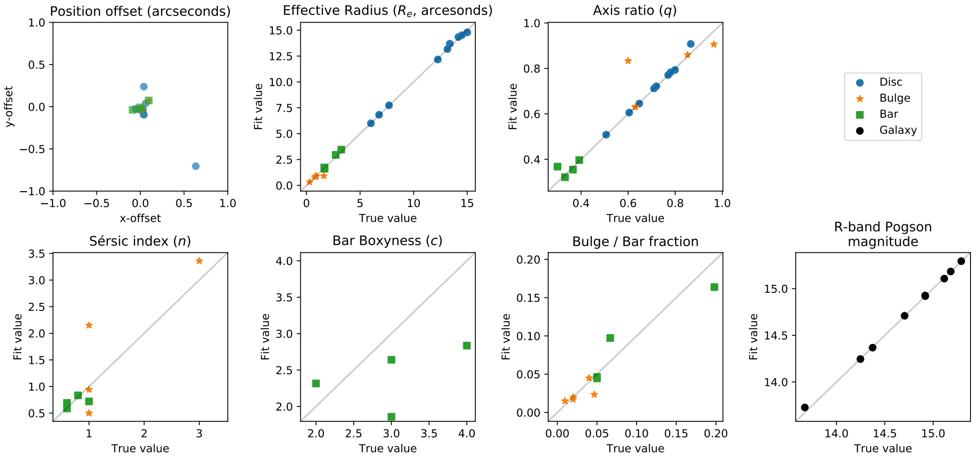


Figure 3.9: Plots examining the accuracy of fit parameters for the calibration subset of galaxies. Most parameters are recovered to a high degree of accuracy, however Sérsic index and boxinxess are difficult to determine only using gradient descent, as they do not significantly impact the goodness of fit (Lackner & Gunn, 2012). The error in the fit values reflects this problem.

aggregate models. This feature (switching light between model components) is a common issue in all photometric fitting methods (Kruk et al., 2018).

4. The Jaccard metric is unstable to small changes in rotation for highly ellipsoidal components (i.e. bars). This resulted in one false negative of bar presence in the aggregate model.

The fitting step for this subset of images highlighted the benefit of obtaining a rough starting point through clustering of user classifications; the method struggled to recover structural parameters for which we did not obtain such a starting point (Sérsic index and bar boxiness). These parameters are difficult to identify using gradient descent (Lackner & Gunn, 2012), suggesting future work should attempt to obtain priors on these parameters from volunteers and make use of a more robust fitting algorithm.

3.10.2 Examination of volunteer consistency

We aggregate two independent models for a set of 98 galaxies based on "original" or repeat ("validation") classifications, obtained with the same retirement limit (see Chapter 3.3 for more on this selection).

One of the simplest choices the volunteers have is whether to include a model component or not. Figure 3.10 illustrates the consistency with which volunteers made use of a component in their model for a galaxy. We see that volunteer classification is very consistent, with scatter as predicted by the Binomial uncertainty on the mean. Volunteers almost always make use of a disc and bulge (as seen in the *calibration subset*), and bulge, bar and spiral arm usage is consistent within Binomial error. One common challenge is that some volunteers used a very ellipsoidal bulge and the ends of spiral arms to model light that other users modelled with a bar. This caused some scatter in aggregate models.

In the end, the aggregated validation model is identical to the original aggregated model in around 40% of galaxies. The most common changes are a missing bar component or a missing single spiral arm. This may suggest that more than 30 classifications should be collected per galaxy, or could be an artefact of the lack of consensus among volunteers for galaxies with difficult-to-determine components.

After selecting a component, the volunteer sets its shape and size. The variation in axial ratios and effective radii for the aggregate discs, bulges and bars are shown in Figure 3.11. The aggregate discs and bulges are consistent within errors, however, bars show more scatter. Bars are one of the most challenging components to aggregate consistently. This is partly because even a strongly barred galaxy with 30 classifications overall might receive only 15 or so drawn bars, and lower numbers of classifications result in more scatter. In addition, the aggregation method is more sensitive to rotation of highly elongated shapes. Both factors probably contribute to lower consistency in bar components.

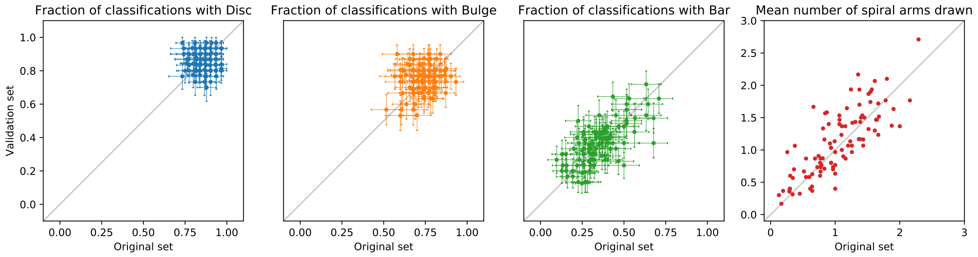


Figure 3.10: Comparison of frequency of use of component in volunteer models between the original and validation sets of classifications. Errors shown on the disc, bulge and bar arise from Binomial error estimation. We see that classifications are generally consistent within errors, validating our assumption of volunteer independence.

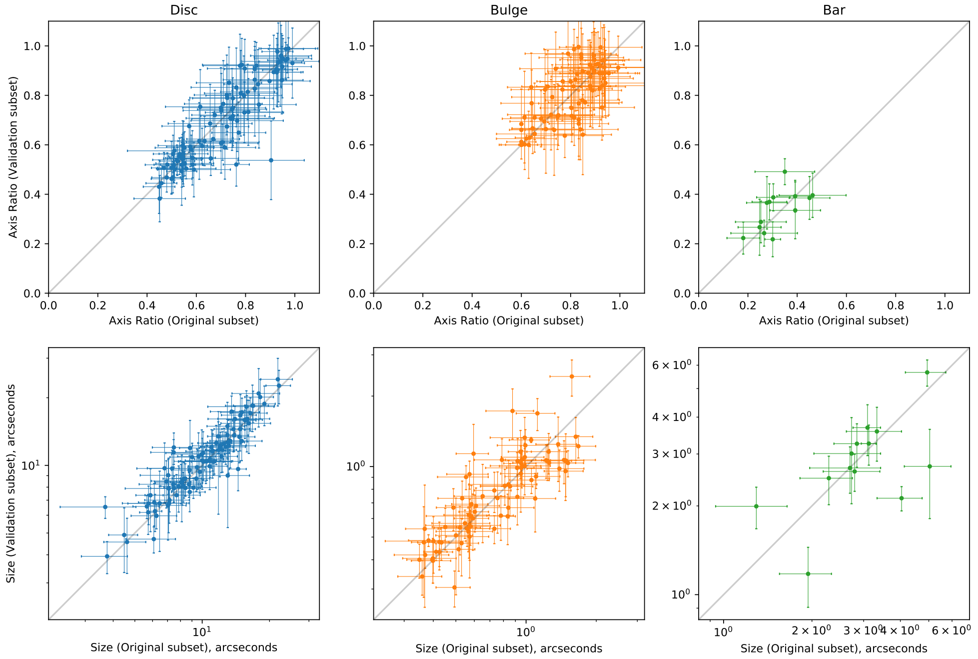


Figure 3.11: Comparison of component shape in aggregate models between the original and validation sets. Errors are obtained through the sample variance of clustered components, as detailed in Section 3.8. We see close agreement between aggregate components from the two sets, suggesting that the clustering method is robust to the scatter in classifications.

3.10.3 Comparison to results in the literature

After having aggregated and fitted models for our galaxies, we examine how our models compare to other results in the literature. Part of the motivation for exploring the *Galaxy Builder* method was that there exists no published large sample of galaxies with four-component photometric fits. This means we can only make comparisons for individual or subsets of model components (e.g. just disc and bulge) and by design *Galaxy Builder* models will differ as we have attempted to fit bulge-disc-bar-spiral models to all our galaxies. The reader is therefore cautioned against treating literature models as any kind of "ground truth" since deviation from these simple models is part of the goal of this project. We provide these comparisons not to check how well our models work, but to provide data on how they compare with other well known, but much simpler photometric models.

Comparison to Galaxy Zoo morphology

The simplest comparison we can make to external results is to examine whether our models respect the existing morphological classifications present in the literature. We make use of GZ2 results, including the redshift debiasing described in Hart et al. (2016a) and spiral properties calculated in Hart et al. (2016a).

When comparing the probability of a volunteer's classification containing a bar component against a galaxy being classed as strongly-barred or as having no bar (as defined in Masters et al. 2010b), we see reasonable agreement. Classifications of GZ2 strongly-barred galaxies ($p_{\text{bar}} > 0.5$) are more likely to contain a bar than GZ2 unbarred galaxies (0.47 ± 0.15 vs. 0.29 ± 0.11). While there is some overlap in these probabilities, the Pearson correlation between GZ2's p_{bar} and the bar likelihood in *Galaxy Builder* is 0.56, implying a significant correlation. We also note that GZ2 bar classifications exclude most weak bars (Kruk et al., 2017).

We also compare the number of spiral arms aggregated by *Galaxy Builder* with the responses to the GZ2 "number of arms" question (of which the possible responses were one, two, three, four, more than four or “Can’t tell”). We attempt to account for the spread in volunteer answers to this question by binning responses, rather than using the mean or modal response. The results of this comparison can be seen in Figure 3.12. The area of each circle can be seen as the level of agreement between *Galaxy Builder* aggregate models and GZ2 classifiers, it is defined as

$$A_{i,j} \propto \sum_{k}^{N_g} \frac{1}{M_k} \sum_{m}^{M_k} \begin{cases} 1, & \text{if } n_k = i \text{ and } C_{k,m} = j \\ 0, & \text{otherwise} \end{cases}, \tag{3.20}$$

where n_k is the number of aggregate arms for galaxy k (out of N_g galaxies), $C_{k,m}$ is the m-th answer for galaxy k (out of M_k answers).

The circle with the largest area for each possible GZ2 response is highlighted, and agrees with the number of spiral arms aggregated here for $m = 1, 2, 3, 4$. No aggregate model contained more than four spiral arms, and when galaxies have an uncertain number of spiral arms (the “Can’t tell” GZ2 response) we mostly do not aggregate any spiral arms.

It is not uncommon in *Galaxy Builder* for one spiral arm to have been broken into two smaller segments. We also occasionally identify two distinct clusters that represent the same physical arm. These two reasons account for a majority of cases where GZ2 classifications suggest a galaxy has two spiral arms and we have clustered a larger number. Improved project user experience would be crucial in correcting these errors.

Comparison to one-component fit - axis ratio

We compare the axis ratios of the discs of *Galaxy Builder* aggregate models (without fitting) to the axis ratio of a 2D Sérsic fit to the r-band SDSS image of each galaxy (as provided in the NSA catalog, Blanton et al. 2011). The resulting scatter is shown in Figure 3.13; for these untuned models there is an error of ~ 0.1, consistent with our expected errors (derived in Section 3.8).

We observe a clustering of outlying values around $b/a = 0.5$. This is

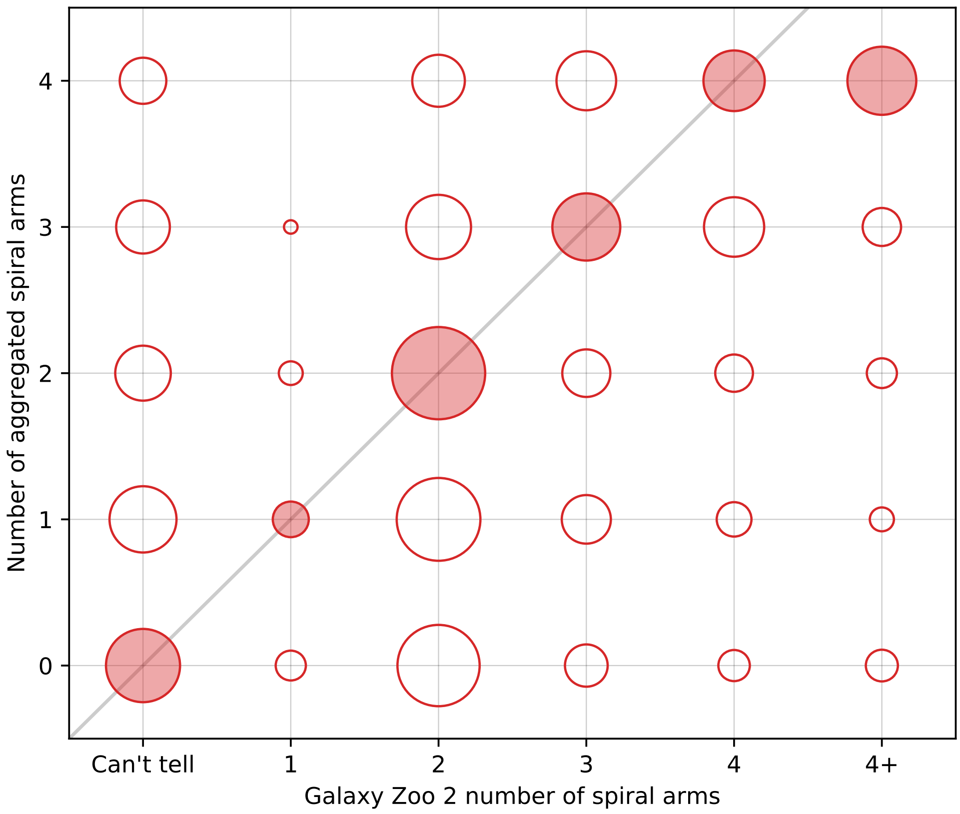


Figure 3.12: Density plot of GZ2 vote counts for spiral arm number vs the number of spiral arms obtained through aggregation. The area of each circle can be seen as the level of agreement between *Galaxy Builder* aggregate models and GZ2 classifiers, and is defined by Equation 3.20. The circle with the largest area for each possible GZ2 response is highlighted by shading. The 1:1 relationship suggests the clustering method is correctly recovering the behaviour of volunteers.

almost certainly due to the drawing tool ellipse having a default axis ratio of 0.5. Where this default is a “good enough” fit we hypothesise that volunteers are less likely to modify it, while if it needs to move a long way they find a more refined value. Overall we see that 36% of all disc components drawn by volunteers were left at the default axis ratio. We recommend that future projects should carefully consider their interface design to minimize this bias (e.g. forcing volunteers to draw both the major and minor axis), however, the fitting process we implement on the aggregate models successfully removes the bias, and the overall scatter does not change significantly.

As we account for light in spiral arms and bars, we expect that disc axis ratios fit by *Galaxy Builder* should be more physical than those from models that do not account for how these non-axisymmetries can bias measurements of ellipticity.

Comparison to disc-bulge models

A strong motivation for performing multi-component modelling is the desire to measure the fraction of a galaxy’s light being emitted by its central components (such as bulge fraction, defined as the ratio of bulge luminosity to total luminosity). Gao & Ho (2017) demonstrate that modelling secondary central components is essential for recovering an accurate measure of bulge fraction. The difficulty of measuring bulge fraction is further compounded by the complex degeneracies present in even two-component fits, meaning that many gradient-descent based solvers often fail to find the globally optimum solution (Robotham et al., 2016), especially when bulge Sérsic index is a free parameter.

One of the largest catalogues of 2D multi-component fits is Simard et al. (2011), which performed simultaneous, two-bandpass decompositions of 1,123,718 galaxies in the Legacy area of the SDSS DR7 using GIM2D. Three variations of models were fitted: a pure Sérsic model, an Exponential disc and de-Vaucouleurs bulge model (hereafter exp+deV), and an Exponential disc and a Sérsic bulge model (exp+S). Fitting was performed using the Metropolis

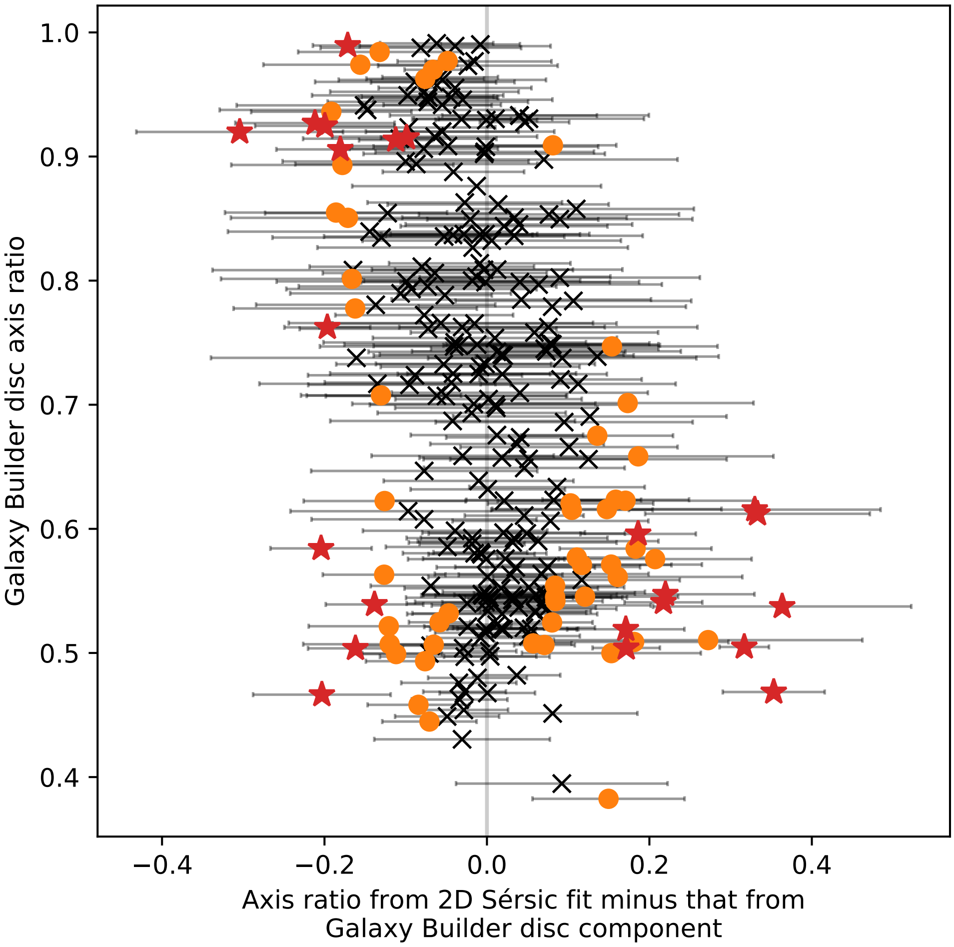


Figure 3.13: Difference between the axis ratios of the aggregated disc component (before fitting) to the results of an r-band Sérsic profile fit. Points between one- and two-sigma are highlighted as orange squares, points outside 2σ are shown as red stars. While the overall relationship is good, the increase prevalence of points outside 2σ is a clear indication of bias caused by the *Galaxy Builder* online user interface.

algorithm, which is resilient to local minima and therefore suitable for the complex likelihood space of galaxy photometric modelling. Lackner & Gunn (2012) similarly fitted two models to SDSS main-sample galaxies: an exponential disc and exponential bulge (exp+exp), and an exponential disc and de Vaucouleurs bulge. They used a Levenberg-Marquardt gradient descent algorithm, with initial parameters taken from previous SDSS analysis.

We compare our central component fraction (the flux of the bulge and bar relative to the total model flux) to bulge fraction from Simard et al. (2011) where their analysis indicated genuine bulge+disc systems ($P_{pS} \leq 0.32$). We compare to Lackner & Gunn (2012) bulge fractions only when their model selection criteria determined that model was the best-fit model. We see a strong correlation with significant scatter (Figure 3.14). The relationship to exp+deV models appears to be less than 1:1, while the relationship to exp+exp models is greater than 1:1, highlighting the dependence of bulge fraction on Sérsic index. Taking *Galaxy Builder* results as ground truth implies that exp+deV puts too much light into the bulge, while exp+exp puts too little.

The amount of scatter (and lack of consistent 1:1 relationships) between bulge fractions between any two of the published two-component models is comparable to the scatter we see between any one of them and our more complex model. Bulge fractions for complex multi-component galaxies fit with any method should be used with caution.

Another comprehensive catalogue of 2D two-component fits is that of Meert et al. (2015), who fit identical models to Simard et al. (2011) on $\sim 7 \times 10^5$ galaxies imaged by SDSS, using GALFIT and PYMORPH (Vikram et al., 2010). They made use of a set of logical filters to distinguish between model fits, allowing them to identify cases where the model did not converge to a physically meaningful result. There is an overlap of 86 galaxy builder galaxy models with their "intermediate catalogue", and we see some scatter between measured parameters (see Figure 3.15). The modelling of spiral arms does not appear to impact measured disk parameters, with disk size

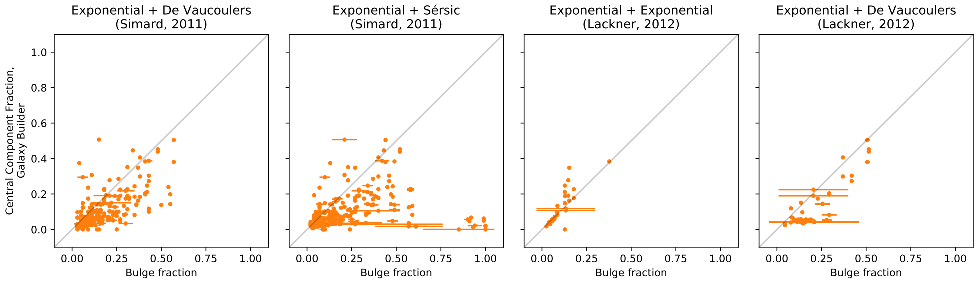


Figure 3.14: Scatter plots comparing the ratio of flux from central components (bulge and bar) to the total flux between fitted models from *Galaxy Builder* and two-component models in the literature. Our models are broadly consistent with their results, but should be more accurate for complex galaxies, as we account for galaxy bars.

and ellipticity showing strong agreement between the catalogues. We see significant scatter in bulge Sérsic index, especially when a bar is present. Total luminosity is not strongly affected by the addition of detail to the model.

Comparison to disc-bulge-bar models

Kruk et al. (2018) performed multi-component (up to three), multi-band decompositions of a selection of SDSS galaxies, 23 of which were also classified in *Galaxy Builder* (with 16 in the repeated validation subset). Figure 3.16 compares the axis ratios and effective radii of bulges, discs and bars in Kruk et al. (2018) to those present in the fitted models. We see good consistency in effective radii of all components in the majority of galaxies. There is more scatter in the fit axis ratios of components. In particular, we observe many of the *Galaxy Builder* bulges reaching the imposed lower boundary. Comparing the central component fraction between *Galaxy Builder* models and those in Kruk et al. (2018), we see next to no scatter.

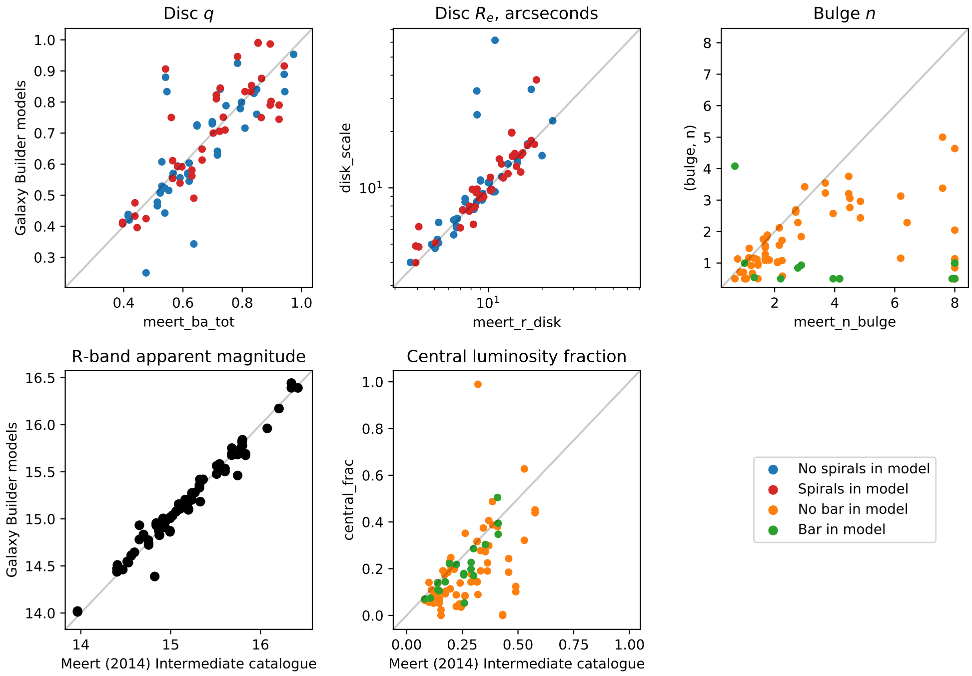


Figure 3.15: Scatter plots comparing measured model parameters between Meert et al. (2015, x-axis) and *Galaxy Builder* (y-axis). We note that adding spirals to a model does not strongly impact disc parameters, but the presence of a bar has a significant impact on bulge Sérsic index measurement.

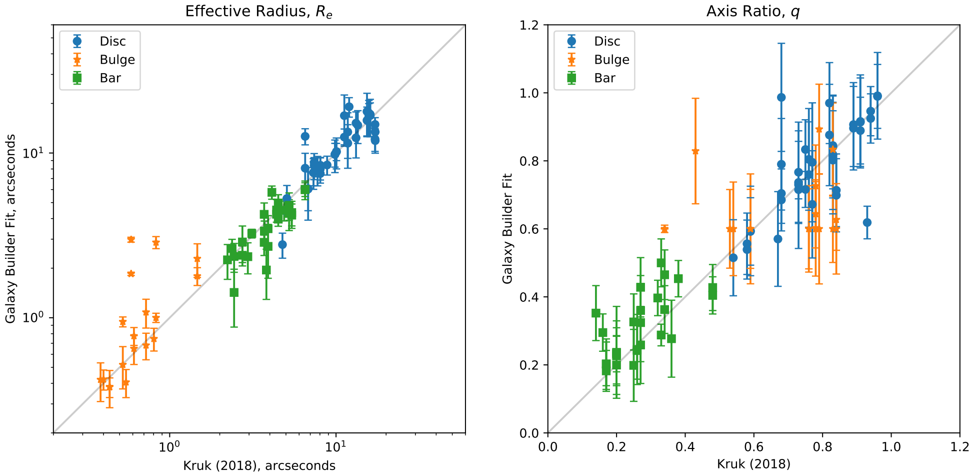


Figure 3.16: Comparison between *Galaxy Builder* fitted models and the result of 3-component, multiwavelength fits performed by Kruk et al. (2018). Discs, Bulges and Bars are shown as blue circles, orange stars and green squares respectively. The left panel compares components' effective radii, the right panel compares the component axis ratio. The components match well, with bulges showing the most scatter. Bulges in *Galaxy Builder* fit models often get stuck at the lower allowed value, despite the physically motivated initial conditions.

Comparison to disc-bulge-bar-spiral models

To the best of our knowledge, no photometric models exist for the *Galaxy Builder* sample that contain spiral arm structure. The closest comparable result is that produced by Gao & Ho (2017), however, the galaxies they used are not in the Sloan footprint.

In order to provide a comparison for our novel method of spiral parameter (pitch angle and amplitude) extraction, we compare the result of our galaxy length-weighted pitch angles to the relationship obtained by Hart et al. (2016a) between GZ2 classification and galaxy pitch angle. Their fit was obtained by using the Zooniverse project *Spiral Spotter* to filter good vs bad spiral arm segments identified using an automated spiral arm detection and fitting tool, SpArcFiRe (Davis & Hayes, 2014a), whereas *Galaxy Builder* asks volunteers to provide their own opinion on spiral arm number, location and tightness. *Galaxy Builder* pitch angles are within the (large) uncertainties on the Hart et al. (2016a) fit.

Many researchers (Davis & Hayes 2014a; Díaz-García et al. 2019 to name a few) have noted that many galaxies show large inter-arm variations in pitch angle, suggesting that obtaining a single value of a galaxy's pitch angle is highly dependent on which arms have been identified. We plan to further explore this issue in future work.

3.11 Summary

This chapter presented a novel method for modelling of galaxy images, *Galaxy Builder*, which was conceived with the goal of solving the "quality or quantity" dilemma facing galaxy image modelling, which, despite advances in computation, still typically requires significant human interaction to achieve quality fits. In future work, we use this sample to investigate spiral arm formation mechanisms.

Galaxy Builder leverages the power of crowdsourcing for the hardest to automate parts of image fitting, namely determining the appropriate number of model components to include, and finding regions of parameter space close to the global optima.

The use of a small sample of synthetic images to calibrate and test our model clustering and fitting code has demonstrated our ability to recover galaxy morphology in the majority of cases. For example, our spiral arm fitting recovered spiral pitch angles to within 9 deg. This set of 9 synthetic images revealed a systematic tendency for volunteers to incorporate more bulges and fewer bars than necessary for photometric models of strongly barred spirals. Future work might implement an improved clustering algorithm and an improved user interface to address the failures of bar model clustering we observed in a small fraction of galaxies.

Some parameters are not recovered well (bulge and bar Sérsic n, bar boxiness), we hypothesise that this is because a wide range of values fit the light profile well. As a result, we are unable to obtain reliable physical results with our optimization algorithm (gradient descent-based methods are subject to being trapped in local minima, or not converging for parameters with flat likelihoods). A solution to this would be performing a full Bayesian optimization with priors obtained from volunteer input, or using a more robust algorithm (such as Basin-Hopping; Wales & Doye 1998). This work is beyond the scope of the current study.

We have demonstrated our ability to obtain physically motivated models with comparable reduced chi-squared values (between 1 and 5) to results in the literature. We obtain errors on parameters where possible through the sample standard deviation of component clusters, which is less likely to be an under-estimate than approximations using the local curvature of the Likelihood-space.

We compare these new models to existing results in the literature. We find good agreement where the models or parameters are comparable, and suggest that where differences are found, *Galaxy Builder* should generally provide superior models because of the more realistic modelling of the galaxy morphologies.

Upcoming survey missions such as the Legacy Survey of Space and Time (Ivezić et al., 2019) and Euclid (Laureijs et al. 2011; Amiaux et al. 2012) present a rich source of astrophysical data. However, *Galaxy Builder* will not be sufficient to deal with the volume of galaxies these surveys will image (twenty billion and two billion respectively, though a large proportion of these will not benefit from detailed photometric modelling). Tools such as *Galaxy Builder* may serve an important role in the generation of training catalogues for scalable machine learning techniques, in an analogous manner to that currently employed for visual morphological classification in Galaxy Zoo: Enhanced (Walmsley et al., 2020).

We were able to obtain aggregate models for 296 images with an average rate of one galaxy per day, and fit photometric models for 294 images. At the time of writing and to the best of our knowledge, the number of photometric models obtained here is still significantly larger than the largest sample obtained through purely computational photometric fitting of a disc, bulge, bar and spiral arms in galaxies (10 galaxies, Gao & Ho 2017, who also included rings, disc-breaks and further components).

The software used to generate image cutouts; perform clustering and aggregation of volunteer models, and fit photometric models is available under a GNU general public licence on GitHub[14]. All models created as part of the *Galaxy Builder* project will be available on the Galaxy Zoo website[15].

---

[14] `github.com/tingard/gzbuilder_analysis`

[15] `data.galaxyzoo.org`

Chapter 4

Morphological dependence of spiral galaxy pitch angle

This chapter explores the use of models obtained through *Galaxy Builder* to measure spiral tightness in a Bayesian hierarchical framework. It is presented in Lingard et al. (2020b).

4.1 Measuring spiral arm structure

Traditionally, measuring spiral structure is a difficult process requiring human tracing of arms. Many methodologies have been proposed and implemented to measure spiral arm properties, including visual inspection (Herrera-Endoqui et al. 2015), Fourier analysis (e.g. 2DFFT, Davis et al. 2012), texture analysis (such as SpArcFiRe, Davis & Hayes 2014b, which attempts to identify "spiral arm segments" using aligned regions of bright pixels; and Ganalyzer, Shamir 2011, which identifies maxima in radial intensity at sequential annuli), combinations of automated methods and human classifiers (Hart et al. 2017, Hewitt & Treuthardt 2020) and photometric model fitting using tools such as GALFIT (Peng et al., 2010).

A common assumption when measuring galaxy pitch angle is that observed spiral arms have a constant pitch angle with radius (e.g. Davis et al.

2012; Savchenko & Reshetnikov 2013; Davis & Hayes 2014b). Spirals of this kind are known as logarithmic spirals and are described by

$$r = A\,e^{\theta \tan\phi}, \tag{4.1}$$

where ϕ is the arm's pitch angle, A is an amplitude coefficient and θ is the polar coordinate. One method used to obtain a pitch angle of a galaxy is to fit logarithmic spirals to individually identified arm segments and take the weighted mean of their pitch angles (which often vary by upwards of 10°, Davis & Hayes 2014b). Weighting is determined by the length of the arc segment, with longer being assigned higher weights, i.e. for a galaxy where we have identified N arm segments, each with length L_i and pitch angle ϕ_i

$$\phi_{\rm gal} = \left(\sum_{i=1}^{N} L_i\right)^{-1} \sum_{i=1}^{N} L_i \phi_i. \tag{4.2}$$

The most commonly used measurement of uncertainty of length-weighted pitch angles is the unweighted sample variance between the arm segments which were identified.

A notable drawback of length-weighted pitch angle is sensitivity to the number and quality of the spiral arm segments; Hart et al. (2017) found that only 15% of the arm segments which were identified using SpArcFiRe (Davis & Hayes, 2014b) were identified as "good" matches to real spiral arms by citizen science classifiers.

Fourier analysis in one- and two-dimensions (as performed by Díaz-García et al. 2019, Davis et al. 2012, Mutlu-Pakdil et al. 2018) is another widely used method of computationally obtaining galaxy pitch angles. Two-dimensional Fourier methods generally decompose a deprojected image of a galaxy into a superposition of logarithmic spirals between inner and outer annuli (Davis et al., 2012) and reports the pitch angle with the highest amplitude as the galaxy's pitch angle, with a number of different methodologies proposed for uncertainty measurement (such as measuring the "stability" of the mean pitch angle for a variety of inner annuli, Davis et al. 2012). Hewitt & Treuthardt (2020) combined Fourier analysis of spiral galaxies with the visual tracing of spiral arms, successfully eliminating observed bias in a sample of toy images of galaxies. It is unclear how the variation between pitch angles of individual arms impacts this measurement.

4.2 The Galaxy Sample

The galaxies analysed in this chapter are those for which photometric models were obtained in Chapter 3. These are a subset of the *stellar mass-complete sample* in Hart et al. (2017), a sample of low-redshift ($0.02 < z < 0.055$) face-on spiral galaxies selected using data from the NASA-Sloan Atlas (Blanton et al., 2011) and Galaxy Zoo 2 (Willett et al., 2013a). The *stellar mass-complete sample* ranged in stellar mass from $9.45 < \log(M_*/M_\odot) < 11.05$.

We combine the 30 classifications of galaxies in the validation subset (see Chapter 3.5) with the 30 original classifications. Clustering of drawn spiral arms and cleaning of points was then performed as detailed in Chapter 3. We remove any galaxies for which no spiral arms were identified, resulting in a hierarchical data structure of 139 galaxies, 261 spiral arms and 239,947 points.

Spiral arm points are deprojected to a face-on orientation using the disk inclination and position angle obtained through photometric model fitting. Arms are individually corrected to all have the same chirality (a pitch angle greater than or equal to zero) using the logarithmic spiral fit in Chapter 3. This was achieved by multiplying the polar coordinate θ by -1 for arms identified as winding counter-clockwise.

4.3 Bayesian modelling of spiral arms in *Galaxy Builder*

In this section, we lay out our Bayesian hierarchical model for galaxy pitch angle. We fit directly to clustered, cleaned points from polylines drawn in *Galaxy Builder*, deprojected and unwrapped to polar coordinates. We fit a logarithmic spiral to each clustered spiral arm (examples are shown in Figure 4.1), with the pitch angles of multiple arms in a single galaxy being drawn from a single parent distribution.

We wish to utilize the logarithmic spiral's desirable properties of a constant pitch angle and a small number of free parameters, therefore, we make use of it here without an explicit comparison to other models. A simple visual inspection of the fitted logarithmic spirals suggests that it is an appropriate model, however, a comparison of a logarithmic spiral profile to other spiral forms (i.e. Archimedian or polynomial) is another important piece of work, outside of the scope of this research, as it has been reported that galaxy arms do not have constant pitch angles (Kennicutt 1981; Ringermacher & Mead 2009).

We assume that a galaxy has some value for pitch angle, ϕ_{gal}, and that the pitch angles of spiral arms in that galaxy, ϕ_{arm}, are constant with radius (giving logarithmic spirals) and drawn from a normal distribution centred on ϕ_{gal}, with some spread σ_{gal} common to all galaxies. We truncate this normal distribution between the physical limits of 0° (a ring) and 90° (a "spoke"), giving

$$\phi_{\mathrm{arm}} \sim \mathrm{TruncatedNormal}(\phi_{\mathrm{gal}}, \sigma_{\mathrm{gal}}, \mathrm{min} = 0, \mathrm{max} = 90). \tag{4.3}$$

The choice to assume all galaxies show the same inter-arm variation in pitch angle (represented by a common value of σ_{gal} across all galaxies) was motivated by our small sample size and the low number of arms measured per galaxy.

We assume that the observed points in a *Galaxy Builder* spiral arm, once deprojected, follow a logarithmic spiral with gaussian radial error σ_r,

$$\widetilde{r_\mathrm{arm}} = \exp\left(\overrightarrow{\theta_\mathrm{arm}} \tan\phi_\mathrm{arm} + c_\mathrm{arm}\right). \tag{4.4}$$

Where $\widetilde{r_\mathrm{arm}}$ is the model's predictions for the radii of the deprojected points in a *Galaxy Builder* arm ($\overrightarrow{r_\mathrm{arm}}$), c_arm is the amplitude parameter (equivalent to A in Equation 4.1), and $\overrightarrow{\theta_\mathrm{arm}}$ is the polar angles of the points.

We choose hyperpriors over ϕ_gal, σ_gal, c_arm and σ_r of

$$\begin{aligned} \phi_\mathrm{gal} &\sim \mathrm{Uniform}(\mathrm{min} = 0, \mathrm{max} = 90), && (4.5)\\ \sigma_\mathrm{gal} &\sim \mathrm{InverseGamma}(\alpha = 2,\ \beta = 20), && (4.6)\\ c_\mathrm{arm} &\sim \mathrm{Cauchy}(\alpha = 0,\ \beta = 10), && (4.7)\\ \sigma_r &\sim \mathrm{InverseGamma}(\alpha = 2, \beta = 0.5). && (4.8) \end{aligned}$$

The inverse gamma distribution is used to aid the convergence of the Hamiltonian Monte Carlo (HMC) algorithm used (discussed later). The Cauchy distribution is equivalent to the Student's t-distribution with one degree of freedom, and was chosen due to its fatter tails than the normal distribution. Our likelihood function for N arms, each with n_arm points, is

$$\mathcal{L} = \prod_{\mathrm{arm}=1}^{N} \left(2\pi\sigma_r^2\right)^{-n_\mathrm{arm}/2} \exp\left(-\frac{||\overrightarrow{r_\mathrm{arm}} - \widetilde{r_\mathrm{arm}}||^2}{2\sigma_r^2}\right). \tag{4.9}$$

We assume that the radial error is Gaussian for simplicity of analysis, however, Shapiro-Wilk tests on the residuals of the logarithmic spirals fit in Chapter 3 suggest that this is not a good assumption, and a more robust likelihood (such as the Student's t-distribution) would possibly more appropriate.

To perform inference, we make use of the No-U-Turn-Sampler (NUTS, Hoffman & Gelman 2011), implemented in PYMC3[1], an open-source probabilistic programming framework written in Python (Salvatier et al., 2016).

---

[1] `docs.pymc.io`

To aid the convergence of MC chains, we scale the radii of deprojected points to have unit variance.

4.4 Results

4.4.1 Constraints on Galaxy Pitch angle

Our hierarchical model identifies the pitch angle of individual arms (ϕ_{arm}) with less than 1.6° of uncertainty for 95% of arms, assuming no error on disc inclination and position angle. This is illustrated well by the small uncertainties on fit spiral arms in Figure 4.1. The pitch angle of a galaxy as a whole (ϕ_{gal}), however, is not well constrained. This is primarily a result of only having pitch angles measurements for a small number of arms per galaxy, and reflects the difficulty in providing a single value for the pitch angle of a galaxy containing individual arms with very different pitch angles. For galaxies with two arms identified in *Galaxy Builder*, we have a mean uncertainty of ($\sigma_{\phi_{\mathrm{gal}}}$) of 7.9°, which decreases to 6.8° and 6.0° for galaxies with three and four arms respectively. This is roughly consistent with the standard error on the mean for a galaxy with N arms,

$$\sigma_{\phi_{\mathrm{gal}}} = \frac{\sigma_{\mathrm{gal}}}{\sqrt{N}}, \tag{4.10}$$

where σ_{gal} is our measure of inter-arm variability of pitch angle and has a posterior distribution of $11.0° \pm 0.9°$. This inter-arm variability is similar to that found by Kennicutt (1981) and Davis & Hayes (2014b) and emphasises the need for fitting algorithms to not assume all arms have the same pitch angle. Examples of galaxies containing arms with a large spread of pitch angles are shown in Figure 4.2. Often two arms will be at a similar pitch angle, with a third arm at a significantly higher pitch angle. The spread of arm pitch angle from the mean galaxy pitch angle can be seen in Figure 4.3, with points colour-coded by the number of arms measured for a galaxy. We see a slight drop in the expectation values of galaxy pitch angle ($E[\phi_{\mathrm{gal}}]$)

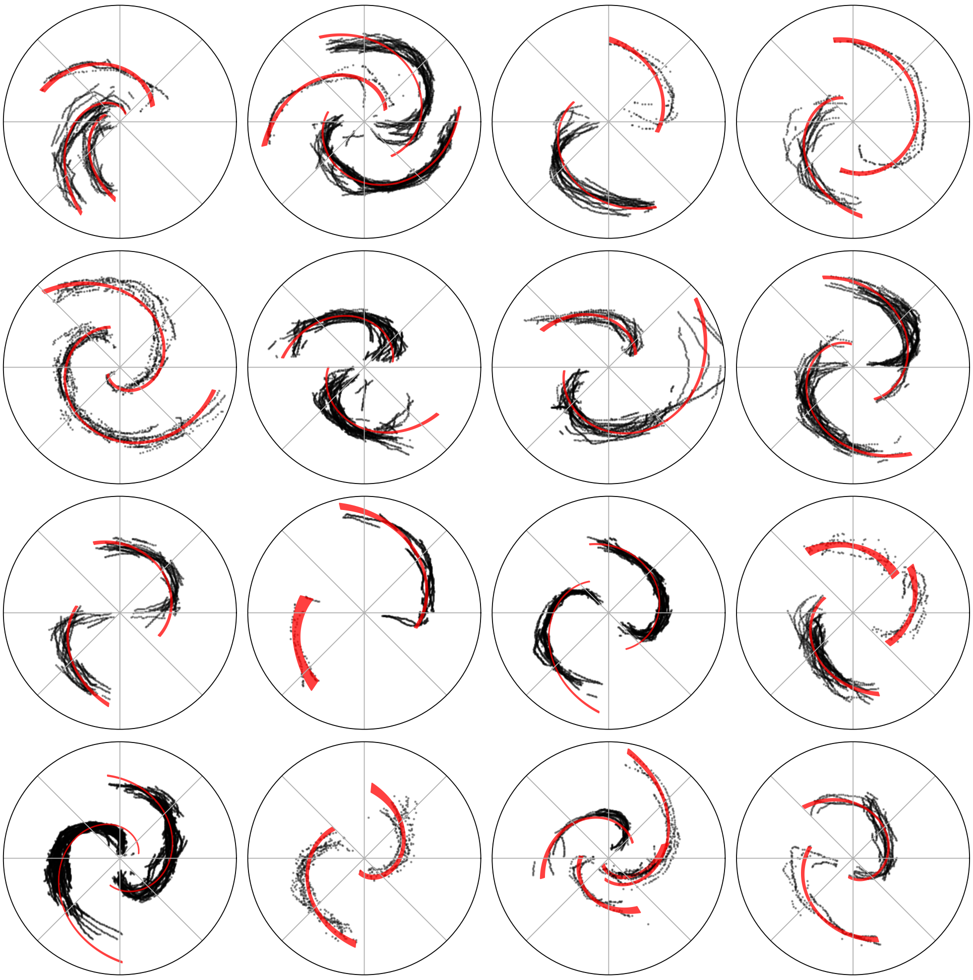


Figure 4.1: Examples of spiral profiles fit using the hierarchical model described in Section 4.3. Deprojected points from *Galaxy Builder* clustered, cleaned spiral arms are shown in black; fit logarithmic spiral arms are shown in red, with the width of the line corresponding to the 2σ interval on predicted values of $\widetilde{r_{\rm arm}}$.

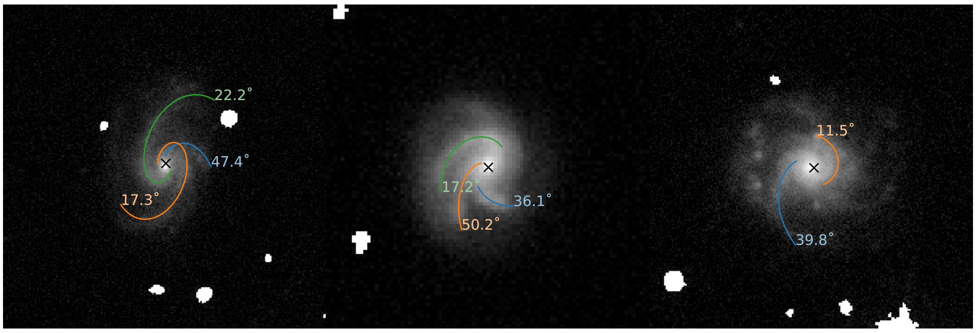


Figure 4.2: Example logarithmic spiral arm fits overlaid on $\sinh^{-1}$-stretched r-band SDSS images. This plot illustrates the wide spread in arm pitch angles inside a single galaxy, and underlines the importance of properly accounting for this variability in any analysis.

compared to the expectation of arm pitch angles ($E[\phi_{\rm arm}]$) at small galaxy pitch angles, due to the truncation of $\phi_{\rm gal}$ at 0°.

4.4.2 Dependence of pitch angle on Galaxy Morphology

In order to test the possible progenitor distribution of our estimated arm pitch angles, we repeatedly perform an Anderson-Darling test (Stephens 1974, implemented in Scipy, Jones et al. 2001) over each draw present in the MC trace, resulting in a distribution of Anderson-Darling statistics. We will refer to this test as the *marginalized Anderson-Darling test*. We also make use of the two-sample Anderson-Darling (Scholz & Stephens, 1987) test in a similar manner.

Pitch angle vs. Bulge size

Morphological classification commonly links bulge size to spiral tightness, and such a link is implied by the Hubble Sequence (Sandage 2005, Gadotti 2009, Buta 2013). Some studies have indeed reported a link between measured spiral galaxy pitch angle and bulge size (e.g. Hart et al. 2017, Davis

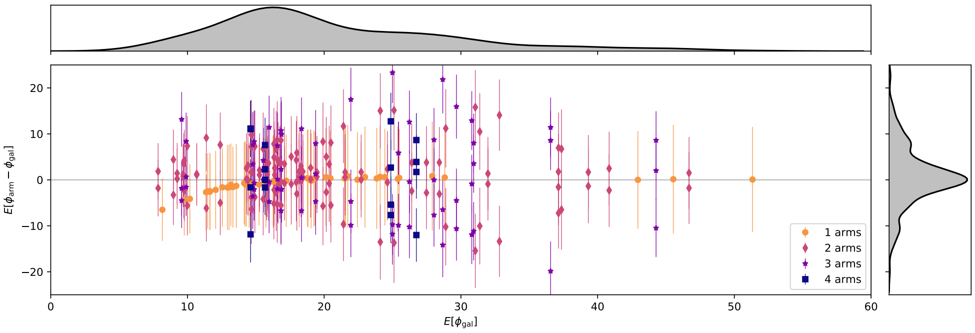


Figure 4.3: Scatter plot showing how arm pitch angle compares to galaxy pitch angle for galaxies with different pitch angles and number of arms. The top panel shows a Gaussian KDE for $E[\phi_{\rm gal}]$, and the right panel shows a Gaussian KDE for $E[\phi_{\rm arm} - \phi_{\rm gal}]$. The galaxy pitch angle is consistent with the mean of its arms, with large scatter and a slight bias against values near the lower bound of 0 due to the lower limit applied.

et al. 2019), while others have not found any significant correlation (Masters et al., 2019). We investigate this relationship here using a measure of bulge prominence from Galaxy Zoo 2, as Equation 3 in Masters et al. (2019):

$$B_{\rm avg} = 0.2 \times p_{\rm just\ noticeable} + 0.8 \times p_{\rm obvious} + 1.0 \times p_{\rm dominant}, \tag{4.11}$$

where $p_{\rm just\ noticeable}$, $p_{\rm obvious}$ and $p_{\rm dominant}$ are the fractions of classifications indicating the galaxy's bulge was "just noticeable", "obvious" or "dominant" respectively.

We see no correlation between galaxy pitch angle derived from the hierarchical model and $B_{\rm avg}$ (evidenced in Figure 4.4). The Pearson correlation coefficient between the expectation value of galaxy pitch angle ($E[\phi_{\rm gal}]$) and $B_{\rm avg}$ is 0.00 (with a p-value of 0.95).

We separate our sample into galaxies with weaker bulges ($B_{\rm avg} < 0.28$, 83 galaxies) and those with stronger bulges ($B_{\rm avg} \geq 0.28$, 54 galaxies), in order to test whether their pitch angles could be drawn from significantly different distributions. A marginalized two-sample Anderson-Darling test

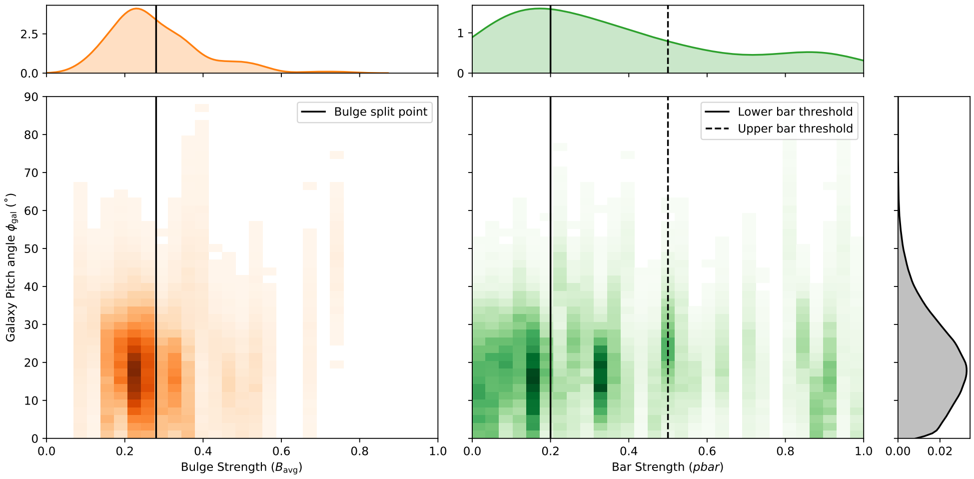


Figure 4.4: Density plot showing bulge strength (B_{avg}; left, orange) and bar strength (p_{bar}; right, green) against galaxy pitch angle (ϕ_{gal}). Split points for the marginalized Anderson-Darling tests are labelled. There is no statistically significant relationship for either bulge or bar strength.

comparing the distributions of ϕ_{gal} for the samples does not find evidence that galaxy pitch angles were drawn from different distributions: we reject the null hypothesis at the 1% level for only 1% of the samples. Similarly comparing arm pitch angles for galaxies in the different samples results in not rejecting the null hypothesis at the 1% level for any of the samples. The distributions of the Anderson-Darling test statistic for ϕ_{gal} and ϕ_{arm} are shown in the upper panel of Figure 4.5 in blue and orange respectively.

One limitation of this result is that our sample does not contain many galaxies with dominant bulges: B_{avg} only varied from 0.09 to 0.75 (the allowed maximum being 1.0), with only four galaxies having $B_{\text{avg}} > 0.5$. The split point of 0.28 was also chosen to produce evenly sized comparison samples rather than from some physical motivation. However, the lack of any form of correlation implies that there is no evidence in our data for the link between bulge size and pitch angle predicted by the Hubble sequence and observed in other studies (e.g. Hart et al. 2017, Davis et al. 2019).

Pitch angle vs. Bar Strength

One of the predictions of Manifold theory is that pitch angle increases with bar strength (Athanassoula et al., 2009b). In order to investigate this relationship in our data, we make use of Galaxy Zoo 2's bar fraction (p_{bar}), which has been demonstrated to be a good measure of bar length (Willett et al., 2013a) and bar strength (Skibba et al., 2012; Masters et al., 2012; Kruk et al., 2018) and therefore a good measure of the torque applied on the disc gas and stars.

We do not observe a correlation between p_{bar} and $E[\phi_{\text{gal}}]$ (Pearson correlation coefficient of -0.05, with a p-value of 0.54; a lack of correlation is clearly visible in Figure 4.4). Following Masters et al. (2012) and Skibba et al. (2012), we separate the sample into galaxies without a bar ($p_{\text{bar}} < 0.2$), with a weak bar ($0.2 \leq p_{\text{bar}} \leq 0.5$) and with a strong bar ($p_{\text{bar}} > 0.5$). Performing marginalized three-sample Anderson-Darling tests does not find that pitch angles (ϕ_{gal} or ϕ_{arm}) of galaxies with different bar strengths were drawn from different distributions; we do not reject the null hypothesis at the 1% level for any samples for the test of ϕ_{gal}, and at the 10% level for the test of ϕ_{arm}. The distributions of the Anderson-Darling test statistic is shown in the lower panel of Figure 4.5.

The fact that we do not find any link between bar strength and pitch angle suggests that the primary mechanism driving the evolution of the spirals in our sample is not Manifold theory.

4.4.3 Spiral Winding

For transient and recurrent spiral arms driven by self-gravity, Pringle & Dobbs (2019) suggest that spiral patterns form at some maximum pitch angle (ϕ_{max}), continually wind up over time and finally dissipate at some minimum pitch angle (ϕ_{min}). They propose that, under a set of very simple

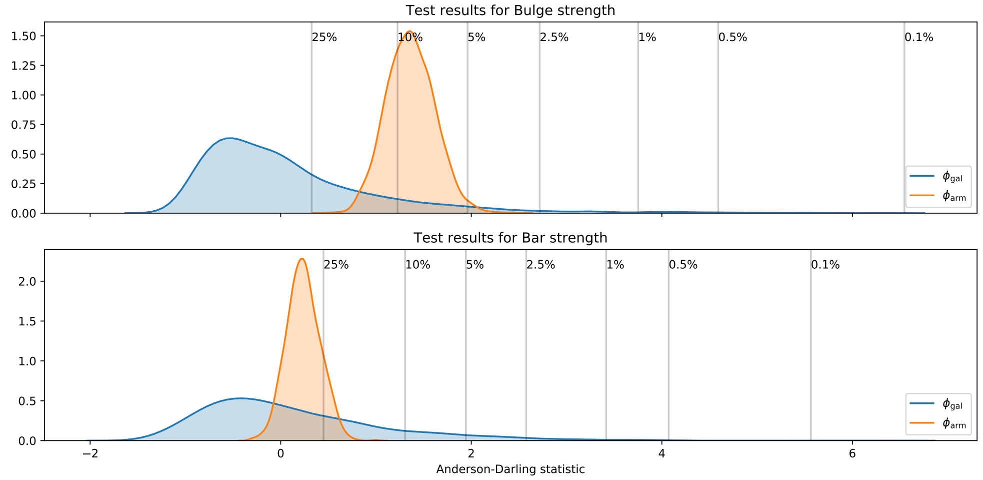


Figure 4.5: The results of marginalized two-sample Anderson-Darling tests examining whether pitch angles (ϕ_{gal} in blue and ϕ_{gal} in orange) for galaxies with $B_{\text{avg}} < 0.28$ and $B_{\text{avg}} \geq 0.28$ are drawn from the same distribution (top panel), and the results of marginalized three-sample Anderson-Darling tests for galaxies with no bar ($p_{\text{bar}} < 0.2$), a weak bar ($0.2 \leq p_{\text{bar}} \leq 0.5$) and a strong bar ($p_{\text{bar}} > 0.5$) (bottom panel). Confidence intervals are shown, with moving rightwards indicating more confidence in rejecting the null hypothesis that the compared values were drawn from the same parent distribution. We cannot reject the null hypotheis at the 1% level for any of the tests conducted, meaning there is no evidence in this sample that bulge size or bar strength impacts pitch angle.

assumptions, the evolution of pitch angle would be governed by

$$\cot\phi = \left[R\frac{\mathrm{d}\Omega_p}{\mathrm{d}R}\right](t - t_0) + \cot\phi_{\mathrm{max}}, \tag{4.12}$$

where Ω_p is the radially dependant pattern speed of the spiral arm and t_0 is the initial time at which it formed.

In QSDW theory, the pattern speed Ω_p is a constant in R, as spiral arms obey rigid-body rotation. If Ω_p instead varies with radius we would expect $\cot\phi$ to be uniformly distributed between $\cot\phi_{\mathrm{max}}$ and $\cot\phi_{\mathrm{min}}$.

In order to test this theory, Pringle & Dobbs (2019) used a Kolmogorov-Smirnov test to examine whether a sample of observed galaxy pitch angles was likely to have been drawn from a distribution uniform in its cotangent. Pitch angles were measured using discrete Fourier transformations in one- and two-dimensions, and as such do not account for inter-arm variations. They chose limits of $\cot\phi \in [1.00, 4.75]$ (roughly $11.9° < \phi < 45.0°$), motivated by examination of their data.

We perform a similar test in this work, using our sample and methods. We will make use of the marginalized Anderson-Darling test described above, and examine winding on a per-arm basis, as well as a per-galaxy basis. Observation of the distribution of arm pitch angles in our sample (Figure 4.6) suggests limits of $15° < \phi < 50.0°$.

Galaxy Pitch angle

Testing the uniformity of $\cot\phi_{\mathrm{gal}}$ between 15° and 50° using a marginalized Anderson-Darling test results in rejecting the null hypothesis at the 1% level for just 5% of samples, with a large spread in observed test values. The full distribution of Anderson-Darling statistics can be seen in the upper panel of Figure 4.7. The large spread in results is caused by the large uncertainties in ϕ_{gal}.

This result suggests that we cannot rule out a cot-uniform source distribution for galaxy pitch angle, but the large uncertainty in ϕ_{gal} makes it difficult

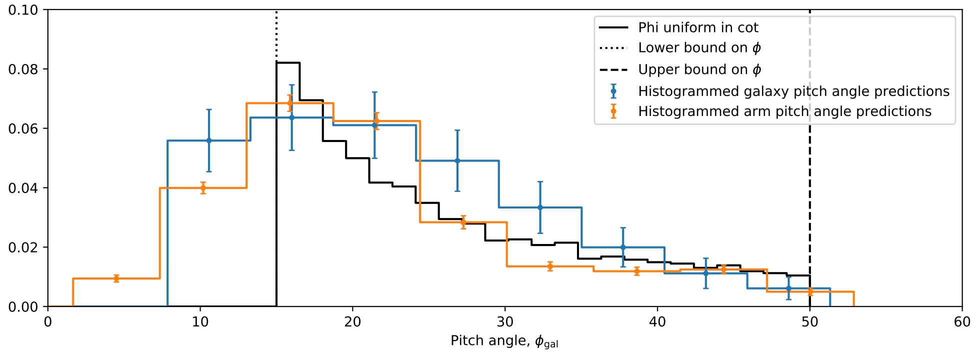


Figure 4.6: The distributions of pitch angles (blue and orange) relative to one uniform in $\cot\phi$ (black). Histograms have been normalised by the area between the limits such that they are comparable. The histogram was re-calculated with identical bins for each posterior sample of ϕ_{gal} and ϕ_{arm}, we plot the mean value of each bin, with the sample standard deviation shown as error bars. It is evident that the distributions are very similar between the chosen limits.

to make any conclusive statements. This result is also highly sensitive to the lower limit of ϕ: decreasing it to 10° results in us rejecting the cot-uniform model at greater than the 0.1% level for 96% of the posterior samples. As we have no information available on the selection biases present for classification of extremely loose or tight spiral arms in *Galaxy Builder*, we choose to keep the less strict limit of 15°.

Arm Pitch angle

The inconclusive result for ϕ_{gal} is perhaps unsurprising: were we to assume that spiral arms are transient and recurrent instabilities, there is little reason for all of the arms to be at precisely the same evolutionary stage at the same time. This is supported by the large observed spread in inter-arm pitch angles (Section 4.4.1).

If we assume instead that spirals form and wind independently inside a galaxy, and that their evolution over time can be described by Equation 4.12,

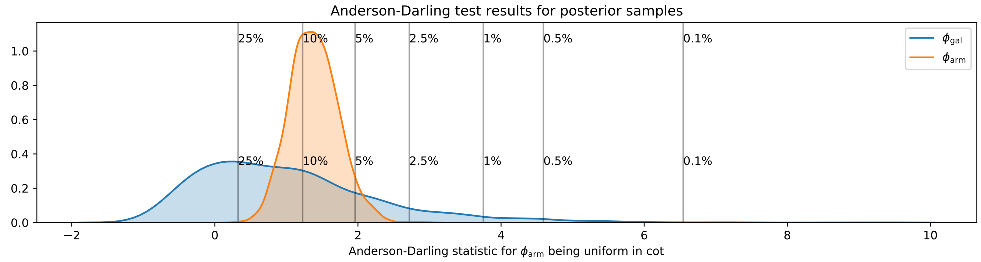


Figure 4.7: The results of a marginalized Anderson-Darling test for uniformity in $\cot$ for $\phi_{\rm gal}$ (blue) and $\phi_{\rm arm}$ (orange), with values corresponding to various confidence intervals shown. Moving rightwards on the x-axis implies greater confidence in rejecting the null hypothesis that the sample was drawn from a distribution uniform in $\cot$ between $15° < \phi < 50.0°$. In this instance, we would not be able to reject the null hypothesis at the 1% level for either $\phi_{\rm gal}$ or $\phi_{\rm arm}$.

the distribution of the cotangent of pitch angles of individual arms should be uniform between our limits, rather than that of the galaxy's pitch angle as a whole.

Using the marginalized Anderson-Darling test we cannot reject the null hypothesis at even the 5% level for any of the possible realizations of arm pitch angle. The resulting distribution of Anderson-Darling statistics is shown in in the lower panel of Figure 4.7. This result is highly consistent with the model for spiral winding proposed by Pringle & Dobbs (2019) and can be interpreted as evidence that spirals are formed through local disc perturbation(s), and are primarily governed by local forces.

4.5 Summary

This chapter presents a new Bayesian approach to estimate galaxy pitch angle, making use of citizen science results to measure spiral arms through photometric modelling. We introduce an adaptation of the Anderson-Darling test, which we name the *marginalized Anderson-Darling test*, to incorporate

]

full Bayesian posterior probabilities and utilize this test to investigate theories governing spiral formation and evolution.

The hierarchical Bayesian approach implemented in this chapter allows a more thorough examination of pitch angle than length-weighted pitch angle calculation; obtaining posterior distributions of measured parameters. It better accounts for the large variations observed in inter-arm pitch angle than Fourier analysis, which assumes all arms in a given mode have the same pitch angle. In this work, we find that the mean inter-arm difference in pitch angle is $11.0° \pm 0.9°$.

There is no evidence in our data for the link between bulge size and pitch angle predicted by the Hubble sequence and observed in other studies (see Section 4.4.2).

We do not find any link between bar strength and pitch angle suggests that the primary mechanism driving the evolution of the spirals in our sample is not Manifold theory (see Section 4.4.2).

Our results are consistent with spiral winding of the form described by Pringle & Dobbs (2019), in which spiral arms are transient and recurrent, evolve through mechanisms such as swing-amplification (Goldreich & Lynden-Bell, 1965) and which wind up over time. However, the assumptions of this model of spiral winding are highly simplistic, and it leaves many unanswered questions: what determines the limits on ϕ? Is the spiral arm equally apparent at all pitch angles, or is a selection effect present? This result is also not evidenced against QSDW, as it is possible that our distribution of pitch angles is dictated by other factors such as disk shear.

In this work, we assume that spiral arms are equally likely to be identified and recovered at all pitch angles. This is not an unfair assumption given the amount of human effort that went into obtaining spiral arm measurements (more so than any other pitch angle measurement method, with each galaxy receiving 30 human classifications). The galaxy sample used is not guaranteed to be representative of the general spiral population, but is comparable in size to those used in other similar studies (Savchenko & Reshetnikov 2013, Yu & Ho 2019, Pringle & Dobbs 2019).

The methodology proposed here is a robust solution to the problems facing investigation of spiral morphology, namely that of reliably identifying spiral arms, and properly accounting for the spread in pitch angles of arms within a galaxy. As with most analyses, the most impactful improvement it would be possible to make here would be to increase the quality and volume of data analysed. A larger sample would make possible further comparisons, such as splitting galaxies into spiral type (grand design / many-armed / flocculent) and examining the differences between populations.

The processes governing the formation and evolution of spiral arms are immensely complicated, but the prevalence of spiral galaxies in the Universe, and the spiral nature of our own Milky Way, makes investigating their dynamics of fundamental importance to the scientific aims of understanding, predicting and explaining the nature of the cosmos.

Chapter 5

Future work and conclusions

The *Galaxy Builder* citizen science project and subsequent data reduction and fine-tuning I have developed, detailed in Chapter 3, explores a novel method by which researchers can work with citizen scientists to decompose a galaxy into a set of distinct, complex, physically meaningful subcomponents. I have demonstrated that this method produces results which are consistent with those in the literature, and comment on cases where our model should produce more accurate results (Chapter 3; Lingard et al. 2020a). I have made use of the spiral arm classifications from these models to explore the formation and evolution of spiral arms in a statistically rigorous manner using Bayesian hierarchical modelling, finding results which support a picture of transient, recurrent spiral arms caused by disc instabilities rather than central galaxy morphology (Chapter 4; Lingard et al. 2020b). No link was found between spiral tightness and bar strength, contrary to the prediction of the Manifold theory of spiral arm evolution; and no relationship was found between spiral tightness and bulge strength, suggesting that popular scaling-relations between pitch angle and black hole mass may need to be reevaluated.

The citizen science project developed in this Thesis is a unique blend of the social science and design requirements of a citizen science project; the exploratory data analysis and pipeline design of a machine learning project; and the complex data management, code optimization and statistical rigour of an astrophysical photometric modelling suite. The tools developed have room for improvement (see Section 5.2), but have already delivered one of the largest catalogues of detailed photometric models (including spiral arms) available, with wide-ranging potential avenues for scientific exploitation, some of which are presented in the following section.

5.1 Potential science cases for *Galaxy Builder* models

The rich dataset provided by *Galaxy Builder* can be utilized to investigate a plethora of possible scientific questions, for which volunteer classifications have already been collected and fitted photometric models obtained.

5.1.1 An investigation into spiral arm profiles

Throughout this work, I have made the assumption that logarithmic spirals are a valid model for spiral arms. This assumption is widely shared in the literature, but not necessarily well-founded (Ringermacher & Mead, 2009). One could make use of the Bayesian hierarchical modelling approach described in Chapter 4 to rigorously compare the predictive power of logarithmic spirals against other spiral profiles, such as the Archimedian spiral,

$$R = a\theta^{\frac{1}{n}}, \tag{5.1}$$

or the “scaffold” description proposed by Ringermacher & Mead (2009),

$$r = A\left[\log\left(B\tan\frac{\theta}{2N}\right)\right]^{-1}. \tag{5.2}$$

I propose the use of Bayesian leave-one-out cross-validation using Pareto-smoothed importance sampling (Vehtari et al., 2015) to examine the relative accuracy of each candidate spiral profile. Given a large enough sample of galaxies, segmentation by morphology would allow a detailed investigation into the effect of secondary components on spiral structure (for instance, does the presence of a bar cause arms to deviate from a purely logarithmic spiral?).

5.1.2 Testing the swing-amplification mechanism of spiral arm formation

Swing amplification is a popular mechanism to describe spiral arm formation. Discussed in Section 1.2.4, swing amplification involves the growth of small density perturbation (wave) due to the differential rotation of the disc. If spirals are generated by swing amplification, then it is expected that the measured pitch angle corresponds to the wave which is most strongly amplified (Julian & Toomre, 1966).

The shear rate of a galaxy is given by

$$\Gamma = \frac{2A}{\Omega} = -\frac{\mathrm{d}\log\Omega}{\mathrm{d}\log R} = 2 - \frac{\kappa^2}{2\Omega^2}, \tag{5.3}$$

where R is the distance from the galaxy's centre, Ω is the circular frequency, A is the Oort constant (Oort, 1927) and κ is the epicyclic frequency.

If we assume a target galaxy has a flat rotation curve,

$$\Omega = \frac{a\tanh(bR)}{2\pi R}, \tag{5.4}$$

where a and b are free parameters. Substituting this into Equation 5.3 gives

$$\Gamma = -\frac{\mathrm{d}\log\Omega}{\mathrm{d}\log R} = 1 - \frac{4bRe^{2bR}}{e^{4br} - 1}. \tag{5.5}$$

Michikoshi & Kokubo (2014) make use of N-body simulations of stellar discs to derive a relationship between the shear rate and spiral pitch angle,

$$\tan\phi \approx \frac{2}{7}\frac{\sqrt{4 - 2\Gamma}}{\Gamma}. \tag{5.6}$$

Combining the above with Equation 5.5, we obtain a differential equation for R:

$$\begin{aligned}\frac{R'}{R} &= \tan\phi \approx \frac{2}{7}\frac{\sqrt{4-2\Gamma}}{\Gamma} \\ &= \frac{8\sqrt{2}}{7}\frac{\sqrt{1+\frac{2bR}{\sinh(2bR)}}}{1-\frac{2bR}{\sinh(2bR)}}.\end{aligned} \tag{5.7}$$

Using data from the MaNGA survey (Smee et al., 2013; Drory et al., 2015; Bundy et al., 2015), it is possible to calculate galaxy rotation curves to a high degree of accuracy (Pilyugin et al., 2019). I propose to fit flat rotation curves to a sample of MaNGA galaxies with reliably measured rotation curves. The resulting estimated b parameter could then be used to fit the above spiral profile to spiral arm points in *Galaxy Builder* models, comparing the fit to a simple logarithmic spiral.

A sample of 196 galaxies (different to that described in Chapter 3) with reliable rotation curve measurements have already been analysed by *Galaxy Builder* volunteers, allowing us to test this prediction of swing amplification and investigate the potential driving mechanisms behind spiral structure.

5.1.3 Detailed photometric decomposition of Milky Way-like galaxies

Detailed photometric decomposition provides insight into the physical processes occurring in a galaxy. The MaNGA survey has collected IFU data on a sample of 40 Milky Way analogue (MWA) galaxies, selected via stellar mass and bulge-to-total ratios (Boardman et al. 2020, displayed in Figure 5.1). This sample of galaxies was chosen such that the distribution of their stellar masses and star formation rates matched the posterior distributions of our own Milky Way. This project aimed to investigate the Milky Way using extragalactic astrophysics, combining the insights learned from the APOGEE dataset (Majewski et al., 2017; Wilson et al., 2019) with the “external view” provided by MaNGA observations.

Photometric models of 32 out of the 40 MWA galaxies have been collected using *Galaxy Builder*, in a new subject set separate to that described in Chapter 3. This subset of MWA galaxies were chosen to have distinct spiral features and so be most appropriate for *Galaxy Builder* classification. By combining *Galaxy Builder* models with the kinematic and stellar population models available through IFU data, a rich understanding of the internal structure of Milky-way like galaxies is possible.

5.1.4 Re-examination of the black hole mass-pitch angle relation

A number of studies have reported a link between galaxy supermassive black hole mass and spiral tightness (e.g. Seigar et al. 2008; Davis et al. 2017; Al-Baidhany et al. 2019). This is partially predicted by the link between bulge size and black hole mass (Magorrian & Tremaine, 1999; Marconi & Hunt, 2003; Häring & Rix, 2004; Wandel, 2004), and the ties between bulge size and spiral arm tightness present in the Hubble sequence. Given the lack of correlation between pitch angle and bulge size reported in Chapter 4, I propose a re-examination of the relationship between galaxy supermassive black hole mass and spiral tightness, using the rigorous statistical methodology detailed in Chapter 4.

Hewitt & Treuthardt (2020) have previously demonstrated success in combining human annotation and automated spiral arm measurement. Their work motivated the creation of the *Spiral Graph*[1] citizen science project, which asks volunteers to trace spiral arm profiles and combines these annotations into an image to be fed into the P2DFFT automated spiral pitch angle measurement tool. They plan to compare the measured pitch angles with a catalogue of measured supermassive black hole masses, in order to re-examine the correlations reported in the literature.

---

[1] `zooniverse.org/projects/astro-lab-ncmns/spiral-graph`

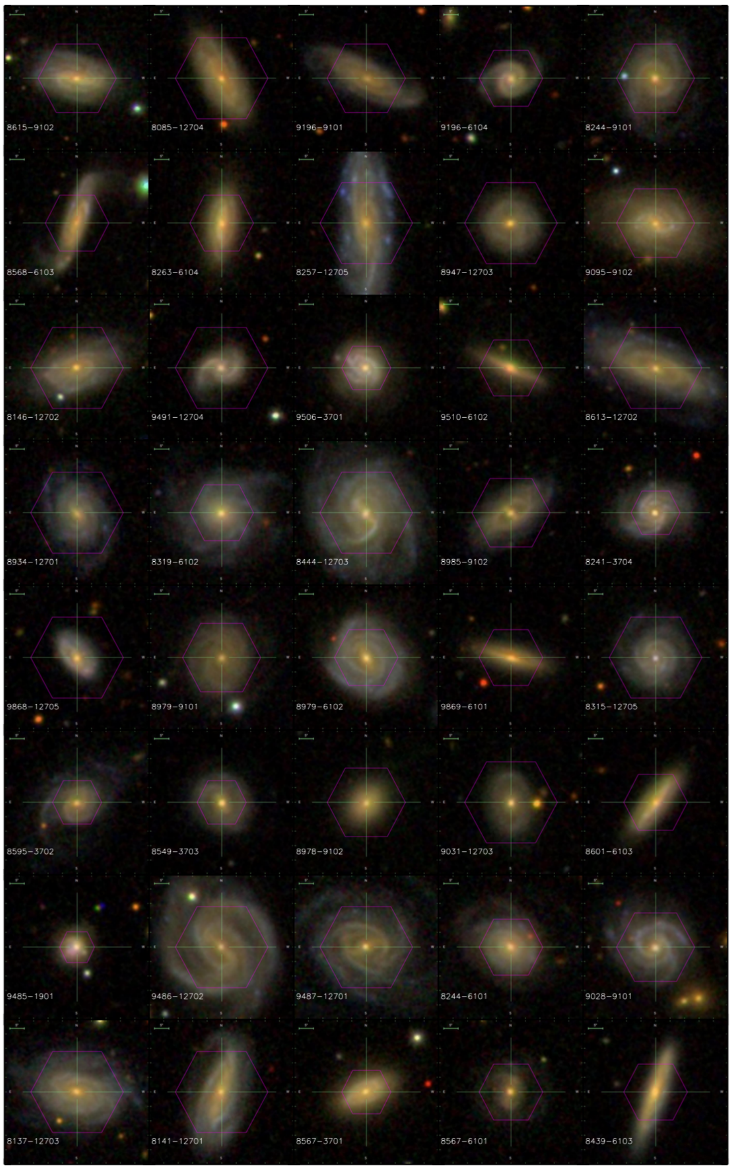


Figure 5.1: SDSS thumbnails of the Milky-Way analogue galaxies detailed in Boardman et al. (2020), with MaNGA field of view overlaid. Figure created by Nick Boardman. Of these 40 galaxies, the 32 with prominent spiral features were uploaded to *Galaxy Zoo* for volunteer classification.

The classification data collected in *Spiral Graph* is compatible with the analysis methodology described in Chapter 4, and our methodology makes fewer assumptions and better accounts for inter-arm variability than the one proposed in Hewitt & Treuthardt (2020). In order to minimize the duplication of volunteer effort, the implementation of the methodology in Chapter 4 to data from the *Spiral Graph* project would likely lead to an improvement in predictive accuracy.

I suggest a collaborative effort to implement the Bayesian model described above with *Spiral Graph* data.

5.2 Galaxy Builder 2

The *Galaxy Builder* project enables the creation of detailed photometric models of galaxies on a scale not previously possible. The user interface provides a set of simple engaging tools with (relatively) intuitive controls, a far abstraction from the complex configuration-heavy interfaces of similar photometric modelling suites. During the development of the project and exploration of the results, a number of issues presented themselves which cannot easily be fixed in the current project architecture. This subsection allows me to put forward a list of changes that would push the project forwards in terms of its scalability, reliability and scientific value.

The current form of *Galaxy Builder* has scope for improvement:

- Removing default values from parameters would reduce the potential for biased classifications (see Chapter 3.10.3).
- The Zooniverse was not designed to have interactive subjects and as such the *Galaxy Builder* code will soon be incompatible with the next-generation of the website.
- Improving the precision of the rendering code used would allow more precise model calculation.

- User experience improvements, including gradually building model complexity, would make the web-interface more friendly for newcomers.
- Introducing interactive human-machine cooperation would, hopefully, speed up the classification process and volunteer satisfaction by providing better real-time feedback.

The net result of these consideration motivates a novel piece of software, here referred to as *Galaxy Builder 2*, closely tied to but distinct from the Zooniverse website, that can both be more scientifically rigorous and less daunting to newcomers. My proposed architecture is detailed below:

5.2.1 Website infrastructure

The Zooniverse is an incredibly valuable resource, and the current work incorporating machine learning for project optimization is a desirable addition to any citizen science project. *Galaxy Builder 2* should incorporate the Zooniverse servers and API for subject uploading and management, and classification storage and aggregation (using their versatile Caesar codebase[2]).

The front-end should be a statically-served site[3], which authenticates users with the Zooniverse using OAuth[4], makes use of simple Javascript for communication with the Zooniverse backend to request subjects and send volunteer classifications. Suggested frameworks include NEXT.JS, GATSBY.JS (both of which are REACT.JS frameworks), NUXT.JS, HUGO or simply a custom hand-coded HTML page. Serving a static site maximises the number of volunteers who can access the project, increases the responsiveness of the interface (essential to keeping users engaged) and better respects data download limits by only sending required code to the browser.

---

[2] `github.com/zooniverse/caesar`

[3] A "static" website is one in which the site is delivered to the user's browser as-is, rather than being generated from a bundled Javascript package

[4] `oauth.net`

As *Galaxy Builder 2* would be purely a front-end interface, and uses the Zooniverse for storage and authentication, neither a database or persistent server infrastructure is needed.

5.2.2 User interface

Similar to the current interface, *Galaxy Builder 2* components should be created and modified using a combination of visual annotation (such as drawing ellipses, rectangles and poly-lines), and numeric inputs. Care should be taken not to provide default values for parameters where possible, or to ensure those default values will never provide a “good enough” result, encouraging volunteers not to perform further adjustments. For advanced users, restrictions on the components' profile should be removed, with them being able to choose from a catalogue of profiles similar to GALFIT. These “expert models” could be used to benchmark the reliability of models obtained through aggregation of models created by citizen scientists.

Galaxy Builder 2 should re-examine the methods by which spiral arms are modelled, examining whether constraining spirals to a known analytic profile (e.g. a logarithmic spiral or Archimedes' spiral) reduces the potential for users to misinterpret instructions and create unphysical models. This decision should be motivated by research such as that mentioned in Section 5.1.1, as well as focus-group testing.

5.2.3 Human-computer interaction

The original *Galaxy Builder* made use of stacked WebGL textures in order to render models. This results in a tradeoff between speed, precision and code readability and is a significant limitation of the current approach. Since the inception of the *Galaxy Builder* project, significant work has been done to enable the training and evaluation of machine learning models (specifically Neural Networks) inside the browser environment using JavaScript. The

TENSORFLOW.JS[5] project is one such endeavour, allowing highly optimized, GPU-accelerated linear algebra to be executed in a browser environment.

Galaxy Builder 2 should leverage the work done by TENSORFLOW.JS developers, specifying models through its easily understandable API. An important benefit of this migration is the inclusion within TENSORFLOW.JS of a number of optimization algorithms (SGD; SGD with momentum, Sutskever et al. 2013). These optimizers would allow an iterative workflow where users could specify parts of their model, fine-tune parameters and then continue adding components, or correct the fitted result to a more physical region of parameter space. It is our hope that the inclusion of human-guided optimization will result in significantly increased volunteer satisfaction, engagement and classification rate. This "snap-to-fit" approach was discussed extensively during the development of *Galaxy Builder*, but its inclusion in the project was not deemed to be a suitable use of the limited time available.

5.2.4 Aggregation and model fitting

Much of the aggregation methodology can be kept from the original *Galaxy Builder* project, though I recommend altering the clustering metric for bars to be more resilient, perhaps by adding a penalty term for differences in component centre-of-mass.

Unwrapping poly-lines into polar coordinates proved to be a difficult task, as ensuring each line unwrapped to the correct region of θ space is not trivial. For example, one poly-line may be unwrapped to $\theta \in [-0.1, \pi]$, while an identical line would be unwrapped to $\theta \in [2\pi - 0.1, 3\pi]$. Fitting a log spiral to points in cartesian space would solve this issue, and is an avenue worth exploring. Fitting logarithmic spirals using the Bayesian method detailed in Chapter 4 would be another desirable addition, as being able to provide a prior on spiral parameters for photometric fitting would further improve the reliability of fitting.

---

[5] `tensorflow.org/js`

Photometric model fitting should attempt to use a gradient-descent algorithm which is more resilient to local minima, such as basin-hopping (Wales & Doye, 1998), or perform a fully Bayesian analysis.

5.2.5 Impact on Scientific Value

The cumulative result of the changes proposed above should be a project which provides the scientific community with high-quality decompositions of large catalogues (1000s) of galaxies. This would be made possible by the increased reliability of an individual classification (thanks to decreased bias, interactive computer optimization and increased rendering precision), and the increased rate of classification from an improved user experience.

By making these proposed changes, *Galaxy Builder 2* would provide significant scientific value to future large-scale surveys, both directly and by providing a training set by which scalable machine learning approaches can be trained to perform photometric decomposition of galaxies.

5.3 Conclusions

This Thesis has presented a novel methodology, by which researchers can obtain detailed photometric models of the light from a galaxy. It discusses the design and implementation of the project and accompanying data reduction and analysis, and makes use of the models created to investigate the evolution of spiral arms with respect to galaxy morphology. This work has produced one published paper (Lingard et al., 2020a), and one submitted paper (Lingard et al., 2020b).

The *Galaxy Builder* project created as part of this thesis was a novel blend of human and computer optimization, and has demonstrated the possibility of recovering detailed photometric models, including spiral arms, for large samples of well-resolved galaxies. The initial dataset of 198 spiral galaxies has been exploited to further our understanding of spiral evolution, suggesting that spiral arms are not strongly influenced by the presence of central structures such as a bulge or bar (see Section 4.4.2), and are transient features that wind with the differential rotation of the disc (see Section 4.4.3).

A number of potential future works have been proposed which delve into the physics governing the evolution of late-type galaxies such as our own Milky Way, much of the data for which is already collected. These include further investigation into spiral arm profiles and formation mechanisms; the use of IFU data for dynamical modelling and examination of star formation histories; and the re-examination of the black-hole mass - pitch-angle relationship observed in other studies.

Final thoughts

Citizen science is a fantastic solution for problems too labour-intensive for small research teams, yet too nuanced or complex for computational approaches. Photometric modelling of detailed galaxy structures fits this problem space perfectly: even with excellent initial conditions, automated fits regularly fail; and manually guiding a complex fit is an exercise in extreme patience and self-flagellation. By leveraging the scientifically-engaged volunteer base of the Zooniverse we have the potential to rapidly generate large numbers of robust, detailed photometric fits, an essential improvement given the increased angular resolution and scale of next-generation surveys such as LSST.

I am hugely optimistic regarding the prospect of projects like *Galaxy Builder* to assist in large-scale photometric modelling projects. The crowd is a resource which we are only beginning to utilize fully, with many organizations recognising the value of human annotation both as a precursor to machine learning tools as well as for validation and edge-case examination. Browser environments become more complex and optimized daily, with Javascript code (running in the V8 engine) already easily outpacing pure Python in terms of performance. Recent advances in tooling, improvements in internet connectivity and speed, and the growing body of research on gamification, human-computer interaction and human-machine optimization make this a fantastically exciting problem space, with a bright, spiral-like future.

Bibliography

Abazajian K. N., et al., 2009, ApJS, 182, 543

Abbott T. M. C., et al., 2018, ApJS, 239, 18

Aihara H., et al., 2018, PASJ, 70, S4

Al-Baidhany I. A., Chiad S. S., Jabbar W. A., Hussein R. A., Hussain F. F. K., Habubi N. F., 2019, in Materials Science and Engineering Conference Series. p. 012118, doi:10.1088/1757-899X/571/1/012118

Albareti F. D., et al., 2017, ApJS, 233, 25

Allen P. D., Driver S. P., Graham A. W., Cameron E., Liske J., de Propris R., 2006, MNRAS, 371, 2

Amiaux J., et al., 2012, in Proc. SPIE. p. 84420Z (`arXiv:1209.2228`), doi:10.1117/12.926513

Andrae R., Schulze-Hartung T., Melchior P., 2010, arXiv e-prints, p. arXiv:1012.3754

Astropy Collaboration et al., 2018, AJ, 156, 123

Athanassoula E., 2012, MNRAS, 426, L46

Athanassoula E., Romero-Gómez M., Masdemont J. J., 2009a, MNRAS, 394, 67

Athanassoula E., Romero-Gómez M., Bosma A., Masdemont J. J., 2009b, MNRAS, 400, 1706

Baba J., Saitoh T. R., Wada K., 2013, ApJ, 763, 46

Bacon R., et al., 2001, MNRAS, 326, 23

Ball N. M., Loveday J., Fukugita M., Nakamura O., Okamura S., Brinkmann J., Brunner R. J., 2004, MNRAS, 348, 1038

Bamford S. P., et al., 2009, MNRAS, 393, 1324

Bamford S. P., Häußler B., Rojas A., Borch A., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, Astronomical Society of the Pacific Conference Series Vol. 442, Astronomical Data Analysis Software and Systems XX. p. 479

Barden M., et al., 2005, ApJ, 635, 959

Barnes J. E., 2002, MNRAS, 333, 481

Barnes J. E., Hernquist L., 1996, ApJ, 471, 115

Beck M. R., et al., 2018, MNRAS, 476, 5516

Bell E. F., et al., 2004, ApJ, 608, 752

Bender R., Doebereiner S., Moellenhoff C., 1988, A&AS, 74, 385

Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697

Bertin E., Arnouts S., 1996, A&AS, 117, 393

Bertin G., Lin C. C., 1996, Spiral structure in galaxies a density wave theory

Binney J., 1982, Annual Review of Astronomy and Astrophysics, 20, 399

Binney J., Tremaine S., 1987, Galactic dynamics

Blandford R. D., Königl A., 1979, ApJ, 232, 34

Blanton M. R., Kazin E., Muna D., Weaver B. A., Price-Whelan A., 2011, AJ, 142, 31

Blanton M. R., et al., 2017, AJ, 154, 28

Boardman N., et al., 2020, MNRAS, 491, 3672

Bolze D., Beyea J., 1989, Environmental Science & Technology, 23, 645

Boonchoo T., Ao X., He Q., 2018, CoRR, abs/1801.06965

Bovy J., Hogg D. W., Roweis S. T., 2009, ApJ, 700, 1794

Bradbury J., Frostig R., Hawkins P., Johnson M. J., Leary C., Maclaurin D., Wanderman-Milne S., 2018, JAX: composable transformations of Python+NumPy programs, `http://github.com/google/jax`

Breiman L., 2001, Machine Learning, 45, 5

Breunig M. M., Kriegel H.-P., Ng R. T., Sander J., 2000, in Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. SIGMOD '00. ACM, New York, NY, USA, pp 93–104, doi:10.1145/342009.335388, `http://doi.acm.org/10.1145/342009.335388`

Broyden C. G., 1970, IMA Journal of Applied Mathematics, 6, 76

Bundy K., et al., 2015, ApJ, 798, 7

Buta R., 1989, Galaxy Morphology. p. 151

Buta R. J., 2013, Galaxy Morphology. p. 155

Buta R. J., et al., 2015, VizieR Online Data Catalog, p. J/ApJS/217/32

Byrd R., Lu P., Nocedal J., Zhu C., 1995, SIAM Journal on Scientific Computing, 16, 1190

Cappellari M., et al., 2011a, MNRAS, 413, 813

Cappellari M., et al., 2011b, MNRAS, 416, 1680

Cedrés B., Cepa J., Bongiovanni Á., Castañeda H., Sánchez-Portal M., Tomita A., 2013, A&A, 560, A59

Cheung E., et al., 2015, MNRAS, 447, 506

Ciambur B. C., 2016, Publ. Astron. Soc. Australia, 33, e062

Combes F., Sanders R. H., 1981, A&A, 96, 164

Conselice C. J., Wilkinson A., Duncan K., Mortlock A., 2016, apj, 830, 83

Croton D. J., et al., 2006, MNRAS, 365, 11

D'Onghia E., Vogelsberger M., Hernquist L., 2013, ApJ, 766, 34

Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645

Davé R., Thompson R., Hopkins P. F., 2016, MNRAS, 462, 3265

Davis D. R., Hayes W. B., 2014a, ApJ, 790, 87

Davis D. R., Hayes W. B., 2014b, ApJ, 790, 87

Davis B. L., Berrier J. C., Shields D. W., Kennefick J., Kennefick D., Seigar M. S., Lacy C. H. S., Puerari I., 2012, ApJS, 199, 33

Davis B. L., Graham A. W., Seigar M. S., 2017, MNRAS, 471, 2187

Davis B. L., Graham A. W., Cameron E., 2019, ApJ, 873, 85

Dehnen W., Binney J. J., 1998, MNRAS, 298, 387

Díaz-García S., Salo H., Knapen J. H., Herrera-Endoqui M., 2019, arXiv e-prints, p. arXiv:1908.04246

Dieleman S., Willett K. W., Dambre J., 2015, MNRAS, 450, 1441

Dobbs C. L., 2014, in Feltzing S., Zhao G., Walton N. A., Whitelock P., eds, IAU Symposium Vol. 298, Setting the scene for Gaia and LAMOST. pp 221–227 (`arXiv:1307.7133`), doi:10.1017/S1743921313006406

Dobbs C., Baba J., 2014, Publ. Astron. Soc. Australia, 31, e035

Domínguez Sánchez H., et al., 2019, MNRAS, 484, 93

Drory N., et al., 2015, AJ, 149, 77

Dubois Y., et al., 2014, MNRAS, 444, 1453

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748

Elmegreen B. G., 2011, in Charbonnel C., Montmerle T., eds, EAS Publications Series Vol. 51, EAS Publications Series. pp 19–30 (`arXiv:1101.3109`), doi:10.1051/eas/1151002

Elmegreen B. G., Elmegreen D. M., 1985, ApJ, 288, 438

Elmegreen D. M., Elmegreen B. G., Bellin A. D., 1990, ApJ, 364, 415

Elmegreen D. M., et al., 2011, ApJ, 737, 32

Emsellem E., et al., 2007, MNRAS, 379, 401

Emsellem E., et al., 2011, MNRAS, 414, 888

Eveleigh A., Jennett C., Blandford A., Brohan P., Cox A., 2014, Conference on Human Factors in Computing Systems - Proceedings

Firmani C., Avila-Reese V., 2003, in Avila-Reese V., Firmani C., Frenk C. S., Allen C., eds, Revista Mexicana de Astronomia y Astrofisica Conference Series Vol. 17, Revista Mexicana de Astronomia y Astrofisica Conference Series. pp 107–120 (`arXiv:astro-ph/0303543`)

Flaugher B., 2005, International Journal of Modern Physics A, 20, 3121

Fletcher R., 1970, The Computer Journal, 13, 317

Freeman K. C., 1970, ApJ, 160, 811

Fritzke B., 1994, in Proceedings of the 7th International Conference on Neural Information Processing Systems. NIPSâĂŹ94. MIT Press, Cambridge, MA, USA, p. 625âĂŞ632

Gadotti D. A., 2009, MNRAS, 393, 1531

Gadotti D. A., 2011, MNRAS, 415, 3308

Gal Y., Islam R., Ghahramani Z., 2017, arXiv e-prints, p. arXiv:1703.02910

Galloway M. A., et al., 2015, MNRAS, 448, 3442

Gao H., Ho L. C., 2017, ApJ, 845, 114

Gao H., Ho L. C., Barth A. J., Li Z.-Y., 2018, ApJ, 862, 100

Goldfarb D., 1970, Mathematics of Computation, 24, 23

Goldreich P., Lynden-Bell D., 1965, MNRAS, 130, 125

Graham A. W., 2001, AJ, 121, 820

Graham A. W., 2014, in Seigar M. S., Treuthardt P., eds, Astronomical Society of the Pacific Conference Series Vol. 480, Structure and Dynamics of Disk Galaxies. p. 185 (`arXiv:1311.7207`)

Graham A. W., Driver S. P., 2005, Publ. Astron. Soc. Australia, 22, 118

Graham A. W., Erwin P., Trujillo I., Asensio Ramos A., 2003, AJ, 125, 2951

Grogin N. A., et al., 2011, ApJS, 197, 35

Guedes J., Mayer L., Carollo M., Madau P., 2013, ApJ, 772, 36

Gunn J. E., et al., 2006, AJ, 131, 2332

Halley E., 1714, Philosophical Transactions of the Royal Society of London Series I, 29, 245

Häring N., Rix H.-W., 2004, ApJ, 604, L89

Hart R. E., et al., 2016a

Hart R. E., et al., 2016b, MNRAS, 461, 3663

Hart R. E., et al., 2017, MNRAS, 472, 2263

Hart R. E., Bamford S. P., Keel W. C., Kruk S. J., Masters K. L., Simmons B. D., Smethurst R. J., 2018, MNRAS, 478, 932

Hastie T., Tibshirani R., Friedman J., 2009, The Elements of Statistical Learning, pp 520–528

Hawarden T. G., Mountain C. M., Leggett S. K., Puxley P. J., 1986, MNRAS, 221, 41P

He K., Zhang X., Ren S., Sun J., 2015, arXiv e-prints, p. arXiv:1512.03385

Head J. T. C. G., Lucey J. R., Hudson M. J., 2015, MNRAS, 453, 3729

Heckman T. M., Best P. N., 2014, ARA&A, 52, 589

Herrera-Endoqui M., Díaz-García S., Laurikainen E., Salo H., 2015, A&A, 582, A86

Hewitt I. B., Treuthardt P., 2020, MNRAS, 493, 3854

Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304

Ho L. C., Kim M., Peng C. Y., Barth A. J., Im M., Martini P., 2009, in American Astronomical Society Meeting Abstracts #213. p. 421.07

Hocking A., Geach J. E., Sun Y., Davey N., 2018, MNRAS, 473, 1108

Hoffman M. D., Gelman A., 2011, arXiv e-prints, p. arXiv:1111.4246

Holincheck A. J., et al., 2016a

Holincheck A. J., et al., 2016b, MNRAS, 459, 720

Hopkins P. F., et al., 2010, The Astrophysical Journal, 724, 915

Houlsby N., Huszár F., Ghahramani Z., Lengyel M., 2011, arXiv e-prints, p. arXiv:1112.5745

Hubble E. P., 1926, ApJ, 64, 321

Hubble E. P., 1936, Realm of the Nebulae

Huertas-Company M., et al., 2015, ApJS, 221, 8

Ivezić Ž., et al., 2019, ApJ, 873, 111

Jacob J. C., et al., 2010, arXiv e-prints, p. arXiv:1005.4454

Jogee S., 2006, The Fueling and Evolution of AGN: Internal and External Triggers. p. 143, doi:10.1007/3-540-34621-X_6

Jones E., Oliphant T., Peterson P., et al., 2001, SciPy: Open source scientific tools for Python, `http://www.scipy.org/`

Julian W. H., Toomre A., 1966, ApJ, 146, 810

Kaiser N., et al., 2010, The Pan-STARRS wide-field optical/NIR imaging survey. p. 77330E, doi:10.1117/12.859188

Kalnajs A. J., 1965, PhD thesis, HARVARD UNIVERSITY.

Kartaltepe J. S., CANDELS Collaboration 2014, in American Astronomical Society Meeting Abstracts #223. p. 139.08

Kelvin L. S., et al., 2012, MNRAS, 421, 1007

Kennicutt R. C. J., 1981, AJ, 86, 1847

Kennicutt Robert C. J., 1998, ARA&A, 36, 189

Khan A., Huerta E. A., Wang S., Gruendl R., Jennings E., Zheng H., 2019, Physics Letters B, 795, 248

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

Kim W.-T., Kim C.-G., Ostriker E. C., 2020, arXiv e-prints, p. arXiv:2006.05614

Knapen J. H., Shlosman I., Peletier R. F., 2000, ApJ, 529, 93

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Kormendy J., Kennicutt Robert C. J., 2004, ARA&A, 42, 603

Kormendy J., Drory N., Bender R., Cornell M. E., 2010, The Astrophysical Journal, 723, 54

Krajnović D., et al., 2011, MNRAS, 414, 2923

Krizhevsky A., Sutskever I., Hinton G. E., 2012, in Pereira F., Burges C. J. C., Bottou L., Weinberger K. Q., eds, , Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp 1097–1105, `http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf`

Kruk S. J., et al., 2017, MNRAS, 469, 3363

Kruk S. J., et al., 2018, MNRAS, 473, 4731

Lackner C. N., Gunn J. E., 2012, MNRAS, 421, 2277

Lange R., et al., 2016, MNRAS, 462, 1470

Larson S. M., Snow C. D., Shirts M., Pande V. S., 2009, arXiv e-prints, p. arXiv:0901.0866

Lauberts A., Valentijn E. A., 1989, The surface photometry catalogue of the ESO-Uppsala galaxies

Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193

Levenberg K., 1944, Quarterly of Applied Mathematics, 2, 164

Li C., White S. D. M., 2009, MNRAS, 398, 2177

Lilly S., et al., 1998, ApJ, 500, 75

Lin C. C., Shu F. H., 1964, ApJ, 140, 646

Lin C. C., Shu F. H., 1966, Proceedings of the National Academy of Science, 55, 229

Lin L., et al., 2020, arXiv e-prints, p. arXiv:2005.09853

Lindblad B., 1927, MNRAS, 87, 553

Lindblad B., 1940, ApJ, 92, 1

Lingard T. K., et al., 2020a, Galaxy Zoo Builder: Four Component Photometric decomposition of Spiral Galaxies Guided by Citizen Science, (in press)

Lingard T. K., et al., 2020b, Galaxy Zoo Builder: Morphological Dependence of Spiral Galaxy pitch angle, Manuscript submitted for publication

Lintott C., 2011, Nature, 478, 320

Lintott C. J., et al., 2008, MNRAS, 389, 1179

Lintott C., et al., 2011, MNRAS, 410, 166

Lupton R., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A., Wherry N., 2004, PUBL ASTRON SOC PAC, 116, 133

Lynden-Bell D., 1969, Nature, 223, 690

Magorrian J., Tremaine S., 1999, MNRAS, 309, 447

Majewski S. R., et al., 2017, AJ, 154, 94

Mao A., Kamar E., Chen Y., Horvitz E., Schwamb M. E., Lintott C. J., Smith A. M., 2013a, in Proceedings of the 1st AAAI Conference on Human Computation and Crowdsourcing (HCOMP'13). `http://econcs.seas.harvard.edu/files/econcs/files/mao_hcomp13.pdf`

Mao A., Kamar E., Horvitz E., 2013b, in First AAAI Conference on Human Computation and Crowdsourcing. `https://www.microsoft.com/en-us/research/publication/why-stop-now-predicting-worker-engagement-in-online-crowdsourcing/`

Marconi A., Hunt L. K., 2003, ApJ, 589, L21

Mark J. W. K., 1974, ApJ, 193, 539

Mark J. W. K., 1976, ApJ, 205, 363

Marscher A. P., 1980, ApJ, 239, 296

Marshall P. J., et al., 2016, MNRAS, 455, 1171

Martin G., Kaviraj S., Hocking A., Read S. C., Geach J. E., 2020, MNRAS, 491, 1408

Masters K. L., et al., 2010a, MNRAS, 405, 783

Masters K. L., et al., 2010b, Monthly Notices of the Royal Astronomical Society, 411, 2026

Masters K. L., et al., 2011, MNRAS, 411, 2026

Masters K. L., et al., 2012, MNRAS, 424, 2180

Masters K., Oh E. Y., Cox J., Simmons B., Lintott C., Graham G., Greenhill A., Holmes K., 2016, arXiv e-prints, p. arXiv:1601.05973

Masters K. L., et al., 2019, MNRAS, 487, 1808

Meert A., Vikram V., Bernardi M., 2015, MNRAS, 446, 3943

Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, ApJS, 210, 3

Mendez-Abreu J., et al., 2016

Méndez-Abreu J., et al., 2018, MNRAS, 474, 1307

Michikoshi S., Kokubo E., 2014, ApJ, 787, 174

Michikoshi S., Kokubo E., 2016, ApJ, 821, 35

Michikoshi S., Kokubo E., 2020, arXiv e-prints, p. arXiv:2005.07367

Miller S., 2001, Public Understanding of Science, 10, 115

Mims F. M., 1999, Science, 284, 55

Minniti D., Zoccali M., 2008, in Bureau M., Athanassoula E., Barbuy B., eds, IAU Symposium Vol. 245, Formation and Evolution of Galaxy Bulges. pp 323–332 (`arXiv:0710.3104`), doi:10.1017/S1743921308018048

Mutlu-Pakdil B., Seigar M. S., Hewitt I. B., Treuthardt P., Berrier J. C., Koval L. E., 2018, MNRAS, 474, 2594

Naab T., Ostriker J. P., 2017, ARA&A, 55, 59

Naim A., et al., 1995, Monthly Notices of the Royal Astronomical Society

Nevin R., Blecha L., Comerford J., Greene J., 2019, ApJ, 872, 76

Newnham L., Hess K. M., Masters K. L., Kruk S., Penny S. J., Lingard T., Smethurst R. J., 2020, MNRAS, 492, 4697

Noordermeer E., Merrifield M. R., Aragón-Salamanca A., 2008, MNRAS, 388, 1381

Nordström B., et al., 2004, The Messenger, 118, 61

Oort J. H., 1927, Bull. Astron. Inst. Netherlands, 3, 275

Parry O. H., Eke V. R., Frenk C. S., 2009, MNRAS, 396, 1972

Pedregosa F., et al., 2011, Journal of Machine Learning Research, 12, 2825

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097

Pilyugin L. S., Grebel E. K., Zinchenko I. A., Nefedyev Y. A., Vílchez J. M., 2019, A&A, 623, A122

Ponciano L., Brasileiro F., Simpson R., Smith A., 2014, Computing in Science and Engineering, 16, 52

Pour-Imani H., Kennefick D., Kennefick J., Davis B. L., Shields D. W., Shameer Abdeen M., 2016, ApJ, 827, L2

Powell L. C., Bournaud F., Chapon D., Teyssier R., 2013, MNRAS, 434, 1028

Pozzetti L., et al., 2009, Astronomy & Astrophysics, 523, A13

Pringle J. E., Dobbs C. L., 2019, arXiv e-prints, p. arXiv:1909.10291

Puchwein E., Springel V., 2013, MNRAS, 428, 2966

Raddick M. J., et al., 2013, arXiv e-prints, p. arXiv:1303.6886

Rampazzo R., et al., 2019, arXiv e-prints, p. arXiv:1910.02664

Rest A., van den Bosch F. C., Jaffe W., Tran H., Tsvetanov Z., Ford H. C., Davies J., Schafer J., 2001, AJ, 121, 2431

Reynolds J. H., 1913, Monthly Notices of the Royal Astronomical Society, 74, 132

Ringermacher H. I., Mead L. R., 2009, MNRAS, 397, 164

Robotham A. S. G., Taranu D. S., Tobar R., Moffett A., Driver S. P., 2016

Rodrigues M., Puech M., Flores H., Hammer F., Pirzkal N., 2018, MNRAS, 475, 5133

Rodriguez-Fernandez N. J., Combes F., 2008, A&A, 489, 115

Romero-Gómez M., Masdemont J. J., Athanassoula E., García-Gómez C., 2006, A&A, 453, 39

Romero-Gómez M., Athanassoula E., Masdemont J. J., García-Gómez C., 2007, A&A, 472, 63

Romero-Gómez M., Athanassoula E., Masdemont J. J., García-Gómez C., 2009, Astrophysics and Space Science Proceedings, 8, 85

Sahu N., Graham A. W., Davis B. L., 2019, ApJ, 876, 155

Salo H., Rautiainen P., Buta R., Purcell G. B., Cobb M. L., Crocker D. A., Laurikainen E., 1999, AJ, 117, 792

Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, PeerJ Computer Science, 55

Sandage A., 1961, The Hubble Atlas of Galaxies

Sandage A., 2005, ARA&A, 43, 581

Sandage A., Bedke J., 1994, The Carnegie atlas of galaxies. Vol. 638

Sanders R. H., Huntley J. M., 1976, ApJ, 209, 53

Sanders D. B., Scoville N. Z., Soifer B. T., 1988, ApJ, 335, L1

Savchenko S. S., Reshetnikov V. P., 2013, MNRAS, 436, 1074

Schawinski K., et al., 2009, MNRAS, 396, 818

Schaye J., et al., 2015, MNRAS, 446, 521

Scheufele D. A., Krause N. M., 2019, Proceedings of the National Academy of Sciences, 116, 7662

Scholz F. W., Stephens M. A., 1987, Journal of the American Statistical Association, 82, 918

Schulze F., Remus R.-S., Dolag K., Bellstedt S., Burkert A., Forbes D. A., 2020, MNRAS, 493, 3778

Schwamb M. E., et al., 2012, ApJ, 754, 129

Schwarz G., 1978, Annals of Statistics, 6, 461

Searle L., Zinn R., 1978, ApJ, 225, 357

Segal A., Gal Y., Kamar E., Horvitz E., Bowyer A., Miller G., 2016, in Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. IJCAIâĂŹ16. AAAI Press, p. 3861âĂŞ3867

Seigar M. S., Block D. L., Puerari I., Chorney N. E., James P. A., 2005, MNRAS, 359, 1065

Seigar M. S., Kennefick D., Kennefick J., Lacy C. H. S., 2008, ApJ, 678, L93

Sellwood J. A., 2011, MNRAS, 410, 1637

Sellwood J. A., Kahn F. D., 1991, MNRAS, 250, 278

Sellwood J. A., Wilkinson A., 1993, Reports on Progress in Physics, 56, 173

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Seyfert C. K., 1943, ApJ, 97, 28

Shamir L., 2011, ApJ, 736, 141

Shanno D. F., 1970, Mathematics of Computation, 24, 647

Shirts M., Pande V., 2001, Science (New York, N.Y.), 290, 1903

Silk J., Mamon G. A., 2012, Research in Astronomy and Astrophysics, 12, 917

Simard L., et al., 2002, ASTROPHYS J SUPPL S, 142, 1

Simard L., Mendel J. T., Patton D. R., Ellison S. L., McConnachie A. W., 2011, ApJS, 196, 11

Simmons B. D., et al., 2014, MNRAS, 445, 3466

Simmons B. D., et al., 2017a, MNRAS, 464, 4420

Simmons B. D., Smethurst R. J., Lintott C., 2017b, MNRAS, 470, 1559

Simonyan K., Zisserman A., 2014, arXiv e-prints, p. arXiv:1409.1556

Simpson E., Roberts S., Psorakis I., Smith A., 2012, arXiv e-prints, p. arXiv:1206.1831

Simpson R., Page K. R., De Roure D., 2014, in Proceedings of the 23rd International Conference on World Wide Web. WWW '14 Companion. ACM, New York, NY, USA, pp 1049–1054, doi:10.1145/2567948.2579215, `http://doi.acm.org/10.1145/2567948.2579215`

Sirovich L., Kirby M., 1987, Journal of the Optical Society of America A, 4, 519

Skibba R. A., et al., 2012, MNRAS, 423, 1485

Smee S. A., et al., 2013, AJ, 146, 32

Springel V., Hernquist L., 2005, ApJ, 622, L9

Springel V., et al., 2005, Nature, 435, 629

Stephens M. A., 1974, Journal of the American Statistical Association, 69, 730

Storrie-Lombardi M. C., Lahav O., Sodre L. J., Storrie-Lombardi L. J., 1992, MNRAS, 259, 8P

Stoughton C., et al., 2002, AJ, 123, 485

Sutskever I., Martens J., Dahl G., Hinton G., 2013, in Proceedings of The 30th International Conference on Machine Learning. pp 1139–1147

Szegedy C., et al., 2014, arXiv e-prints, p. arXiv:1409.4842

Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z., 2015, arXiv e-prints, p. arXiv:1512.00567

Tayyub J., Hawasly M., Hogg D. C., Cohn A. G., 2017, arXiv e-prints, p. arXiv:1709.03456

Tenjes P., Tuvikene T., Tamm A., Kipper R., Tempel E., 2017, A&A, 600, A34

Toomre A., 1977, in Tinsley B. M., Larson Richard B. Gehret D. C., eds, Evolution of Galaxies and Stellar Populations. p. 401

Toomre A., 1981, in Fall S. M., Lynden-Bell D., eds, Structure and Evolution of Normal Galaxies. pp 111–136

Tuccillo D., Huertas-Company M., Decencière E., Velasco-Forero S., Domínguez Sánchez H., Dimauro P., 2018, MNRAS, 475, 894

Ulrich M.-H., Maraschi L., Urry C. M., 1997, ARA&A, 35, 445

Uzeirbegovic E., Geach J. E., Kaviraj S., 2020, arXiv e-prints, p. arXiv:2004.06734

Vaucouleurs G., 1959, Handbuch der Physik, 53, 275

Vehtari A., Gelman A., Gabry J., 2015, arXiv e-prints, p. arXiv:1507.04544

Vika M., Bamford S. P., HÃd’uÃ§ler B., Rojas A. L., 2014, Monthly Notices of the Royal Astronomical Society, 444, 3603

Vikram V., Wadadekar Y., Kembhavi A. K., Vijayagovindan G. V., 2010, MNRAS, 409, 1379

Vogelsberger M., et al., 2014, MNRAS, 444, 1518

Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Reviews Physics, 2, 42

Wales D., Doye J., 1998, arXiv e-prints, pp cond–mat/9803344

Walmsley M., et al., 2020, MNRAS, 491, 1554

Wandel A., 2004, in Storchi-Bergmann T., Ho L. C., Schmitt H. R., eds, IAU Symposium Vol. 222, The Interplay Among Black Holes, Stars and ISM in Galactic Nuclei. pp 65–69 (`arXiv:astro-ph/0407399`), doi:10.1017/S1743921304001474

Wang W., Wang A., Tamar A., Chen X., Abbeel P., 2017, arXiv e-prints, p. arXiv:1711.08534

Wang Z., Yang L., Yin F., Lin K., Shi Q., Luo Z.-Q., 2020, arXiv e-prints, p. arXiv:2006.04097

Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2

Westphal A. J., et al., 2006, in AGU Fall Meeting Abstracts. pp P52B–08

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Willett K. W., et al., 2013a

Willett K. W., et al., 2013b, MNRAS, 435, 2835

Willett K. W., et al., 2017, MNRAS, 464, 4176

Wilson J. C., et al., 2019, PASP, 131, 055001

Yu S.-Y., Ho L. C., 2019, ApJ, 871, 194

Yu F., Xian W., Chen Y., Liu F., Liao M., Madhavan V., Darrell T., 2018, arXiv e-prints, p. arXiv:1805.04687

Zaninetti L., 2014, arXiv e-prints, p. arXiv:1401.0287

Zevin M., et al., 2017, Classical and Quantum Gravity, 34, 064003

Zhou Z.-B., Zhu W., Wang Y., Feng L.-L., 2020, ApJ, 895, 92

de Jong R. S., 1996, A&AS, 118, 557

de Vaucouleurs G., 1948, Annales d'Astrophysique, 11, 247

de Vaucouleurs G., 1959, Handbuch der Physik, 53, 275

de Vaucouleurs G., de Vaucouleurs A., Corwin Herold G. J., Buta R. J., Paturel G., Fouque P., 1991a, Third Reference Catalogue of Bright Galaxies

de Vaucouleurs G., de Vaucouleurs A., Corwin Jr. H. G., Buta R. J., Paturel G., Fouqué P., 1991b, Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between 0^h and 12^h. Volume III: Data for galaxies between 12^h and 24^h.

de Zeeuw P. T., et al., 2002, MNRAS, 329, 513

van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79

van der Wel A., et al., 2012, ApJS, 203, 24

FORM UPR16

Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Research Degrees Operational Handbook for more information

Postgraduate Research Student (PGRS) Information		**Student ID:**	834331
PGRS Name:	Timothy Lingard		
Department:	ICG, Tech	**First Supervisor:**	Dr. Coleman Krawczyk
Start Date: (or progression date for Prof Doc students)		1/10/2016	
Study Mode and Route:	Part-time ☐ Full-time ☐	MPhil ☐ PhD ☐	MD ☐ Professional Doctorate ☐

Title of Thesis:	Human-Guided Four-Component Photometric Modelling of Spiral Galaxies
Thesis Word Count: (excluding ancillary data)	24,212

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study. Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-do/code-of-practice-for-research/)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES ☐ NO ☐
b) Have all contributions to knowledge been acknowledged?	YES ☐ NO ☐
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES ☐ NO ☐
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES ☐ NO ☐
e) Does your research comply with all legal, ethical, and contractual requirements?	YES ☐ NO ☐

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):	A046-8002-4949-5AC8-98AB-0968-CBFD-1C7E

If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

Signed *(PGRS)*:		**Date:** 30/06/2020