Supermassive Black Hole Dynamics in Massive Early-Type Galaxies

Alexander Rawlings

ACADEMIC DISSERTATION

Faculty of Science
University of Helsinki
Helsinki, Finland

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Hall E204, Physicum (Gustaf Hällströmin katu 2, Helsinki), on Friday November 28th 2025, at one o'clock.

Helsinki 2025

Cover picture: Overlay of simulated orbital trajectories of five equal-mass supermassive black hole binaries starting from the same initial conditions, perturbed only by phase space discretisation. The data forms part of the Rawlings et al. (2023) study on stochasticity in supermassive black hole binary eccentricity.

Abstract

Almost all massive galaxies host supermassive black holes (SMBHs) at their centres. Observations of massive galaxies and their SMBHs provide compelling evidence for a strong evolutionary connection between both: an interplay that results in more massive SMBHs residing in more massive galaxies. Additionally, in the standard cosmological framework, galaxies grow through dynamical interactions, with the merging of near-equal mass galaxies generally producing an elliptical galaxy. Such systems often demonstrate a deficit of stellar light in the very centre, a phenomenon that is best explained by the interaction of the two central SMBHs brought in with the progenitor galaxies during the merger process.

As the two SMBHs lose energy and spiral towards their eventual coalescence, low-frequency gravitational waves are expected to be produced. The gravitational waves will hopefully be soon detectable by ongoing instruments such as pulsar timing arrays and the upcoming Laser Interferometer Space Antenna. Prior to their merging into a single SMBH, the asymmetry in spins and masses of the two SMBHs may impart a recoil kick to the merged SMBH by way of linear momentum conservation, potentially displacing the SMBH to large, kiloparsec-scale distances from the centre of the galaxy merger remnant. This presents an exciting target for multi-messenger astronomy, coupling gravitational wave observations with electromagnetic counterparts, to gauge the frequency and magnitude of these events.

This thesis uses simulations run with the KETJU code to explore supermassive black hole dynamics in the context of merging massive galaxies in two regimes: the pre-coalescence SMBH binary, and the post-coalescence recoiling SMBH.

In the first phase, small perturbations to the initial impact parameter of the two SMBHs due to numerical phase space discretisation can lead to the SMBH binary orbit, in particular the orbital eccentricity, diverging between repeat simulations. A consequence of this is an extreme variation in the predicted SMBH binary merger timescale. This highlights that single simulations of a given system are insufficient

in providing meaningful predictions for the expected timescale – and thus frequency – of SMBH binary merger events imperative for gravitational wave detectors.

In the second phase, the impact of a gravitational wave-induced recoiling SMBH on its host galaxy is systematically explored for discrete kick velocities. The growth in the stellar core by the kicked SMBH is quantified, as well as the change in stellar orbital structure, providing an 'archaeological' tool that may be applied to observations to ascertain the recoil magnitude a given SMBH likely experienced in its last major merger. From this, predictions about the likely magnitudes of SMBH recoil relative to the central stellar velocity dispersion are made for local galaxies.

Additionally, identifying a recoiling SMBH by way of enshrouding stars is also explored in the context of ongoing large-scale observational surveys such as Euclid. These 'black hole recoil clusters' are expected to appear as bright, offset stellar clusters with exceptionally high internal velocity dispersions, and would be most readily detectable for recoil velocities that displace the SMBH to distances approximately equal to the effective radius of the galaxy.

The thesis introduces novel analysis methods: stellar orbit modelling, mock integral field unit observations, and Bayesian methodology, to achieve a rich understanding of the complex dynamical processes at play between the host galaxy and its resident SMBH.

Acknowledgements

This thesis, and indeed the entire PhD journey, has encouraged me to become an independent researcher; this does not entail that the work has been done in isolation. For most of it, I have had persistent voices in my head: voices heard through my noise-cancelling, Dual Noise Sensor technology SONY WH-1000XM4 headphones, singing the sweet beats of techno, tech house, and electronica. Shout out to my most featured artists: Charlotte de Witte, Adriatique, Camelphat, Tinlicker, and Bonobo, for keeping me constant company.

Outside of music, I have also had the privilege of being in supportive company, company for which I care and respect deeply, and without which this journey would have been infinitely more difficult.

First, I would like to thank Peter Johansson for the opportunity to complete this PhD, his guidance and patience as my supervisor, and in particular for his effort in bringing together a fantastic research group in Helsinki for which I had the pleasure of being a part of. My thanks as well goes to Till Sawala (please answer my email), Matias Mannerkoski (who is, actually, a real person), Shihong Liao (one of the kindest people I've met), Dimitrios Irodotou (my fellow techno-loving companion), Francesco Paolo Rizzuto (who shared his gnocchi recipe with us), Jessica May Hislop (we love Tilly), Bastián Reinoso (my fellow gym-goer), Gábor Rácz (I only recently understood your work), and Toni Tuominen (who does the best impersonations) for their encouragement throughout my thesis. A special mention goes to Ruby Wright, my fellow Australian, for keeping me sane during the overwhelming times and her iconic sense of humour. I've also had the pleasure of working alongside three inspiring PhD students – Atte Keitaanranta, Max Mattero, and Jenni Häkkinen – and hope I have brought at least half the joy to them as they have to me. Thanks also to my office mates Emma Mannfors, Mikko Vuori, Vesa Björn, Job Vorster, and Atte (again), as well as the entire corridor – Devika Tharakkal, Bandi Harris, Anne Keski-Vakkuri, and everyone else – for making our workplace the fun space that it is.

In addition to the group in Helsinki, I've had the opportunity to work alongside collaborators outside of Finland. A special mention to Thorsten Naab and Antti Rantala, who both brought valuable insight and direction to my projects, and for their hospitality during each of my visits to the MPA in Garching.

Beyond research, I've been extremely fortunate to have met some amazing friends, many of whom I have formed lifelong friendships with. In particular, Marcel Niedermeier, Sanyi Maho, and Axelle Journet – who I met during my first week in Helsinki – have been a constant source of support and laughs for which I am eter-

You can never know everything, and part of
what you know is always wrong. Perhaps
even the most important part. A portion of
wisdom lies in knowing that. A portion of
courage lies in going on anyway.

Robert Jordan, *Winter's Heart*

List of publications

Publications included in this thesis

Paper I: Rawlings et al. (2023)

Paper II: Rawlings et al. (2025b)

Paper III: Rawlings et al. (2025a)

Paper IV: Mannerkoski et al. (2023)

Paper V: Rantala et al. (2024)

Publications not included in this thesis

Rawlings et al. (2020)
Mannerkoski et al. (2022)
Liao et al. (2023)
Liao et al. (2024a)
Liao et al. (2024b)
Sawala et al. (2025)
Fabricius et al. (2025, submitted)

List of abbreviations

CDF   cumulative distribution function
CoM   centre of mass
DCBH   direct collapse black hole
DM   dark matter
ELT   Extremely Large Telescope
ETG   early-type galaxy
FMM   Fast Multipole Method
FWHM   full width half maximum
GBS   Gragg-Bulirsch-Stoer
GW   gravitational wave
HMC   Hamiltonian Monte Carlo
IC   initial condition
ICDF   inverse cumulative distribution function
IFU   integral field unit
IMBH   intermediate mass black hole
IPTA   International Pulsar Timing Array
ISM   interstellar medium
KDK   kick-drift-kick
LIGO   Laser Interferometer Gravitational wave Observatory
LISA   Laser Interferometer Space Antenna
LOSVD   line-of-sight velocity distribution
LTG   late-type galaxy
MBH   massive black hole
MST   minimum spanning tree
NUTS   No U-Turn Sampler
PN   post-Newtonian
PTA   Pulsar Timing Array
SAMI   Sydney-AAO Multi-object Integral field spectrograph
SCF   self-consistent field
SDSS   Sloan Digital Sky Survey
SFR   star formation rate
sSFR   specific star formation rate
SKA   Square Kilometre Array
SMBH   supermassive black hole
SPH   smoothed particle hydrodynamics
SPS   stellar population synthesis

Contents

1 Introduction

1.1 A bit about galaxies

Let us begin with an insightful observation by renowned author Douglas Adams[1]:

> Space is big. Really big. You just won't believe how vastly, hugely, mind-boggingly big it is.

Owing to this, everyday units of measurement of distance, time, and mass do not suffice in our discussion of space. Henceforth, we will be discussing distance in terms of *parsecs*, which is equivalent to

$$1\,\mathrm{pc} = 3.26\,\text{light-year} = 3.09 \times 10^{16}\,\mathrm{m},$$

or about one billion round trips between Finland and Australia. Similarly, the timescales that we are interested in are of the order *gigayears*, where $1\,\mathrm{Gyr} = 10^9\,\mathrm{yr}$, or ten million (optimistic) human lifetimes. Additionally, the mass of objects that we will consider are also extreme, where we will be using *solar masses* ($\mathrm{M_\odot}$), with $1\,\mathrm{M_\odot} \simeq 2 \times 10^{30}\,\mathrm{kg}$, or more than three hundred thousand Earths. These numbers, just like the space they describe, are very big, and difficult to grasp an intuitive sense of.

The Earth orbits the Sun, which in turn is just one star amongst the 100 to 400 billion that collectively make up our Galaxy, the Milky Way (e.g. Licquia & Newman, 2015; Vallenari, 2022). In addition to stars, galaxies contain vast amounts of dark matter – material about which we know very little apart from its presence – and gas and dust. Holding all this material into what we call a galaxy is the force of gravity – a fundamental force of nature which is responsible for not only the existence of galaxies, but also moons orbiting a planet and apples falling to Earth. Gravity is an *attractive* force, which in the Newtonian approximation, brings two bodies together with a strength inversely-proportional to the square of their distance – closer bodies experience stronger gravitational attraction than distant ones.

---

[1]An excerpt from *The Hitchhiker's Guide to the Galaxy*, 1979.

Turning to the night sky with our current telescope capabilities, we are treated to an incredible sight: the Universe contains more than a trillion galaxies, all of different shapes, sizes, and ages. Galaxies are understood to form hierarchically through bottom-up formation: two or more galaxies interact with each other (a process called *merging*) to form a larger galaxy, gradually accreting increasingly-more material as they grow.

At this point, one might notice a conflict in our story. We have argued that galaxies grow by interactions, yet not all the material in the Universe is collapsed into a single point, as one might naïvely expect of gravity. There are two main reasons for this. On the galaxy-scale, the constituent components (stars, gas, etc.) possess angular momentum, which must be somehow lost for two objects to come arbitrarily close together from great distances. Conversely, on the extragalactic scale, different regions of the Universe are expanding away from each other at an accelerating rate, through a force provided by an unknown dark energy (Riess et al., 1998; Schmidt et al., 1998). Within the scientific community, we name the description of our Universe for which we know very little 'ΛCDM' – dark energy, represented by Λ, we do not understand, and cold dark matter (CDM), we can constrain but ultimately also have a very poor understanding of. With our most recent observational understanding, we can constrain the total energy budget of the Universe as consisting of approximately 70% dark energy and 30% matter, of which only 5% is in the form of baryonic matter – ordinary matter that makes up atoms (Planck Collaboration et al., 2020).

In special cases however, the force of gravity can form a singularity, a point of unbelievably extreme density, which we term a black hole.

1.2   A bit about black holes

Black holes, whilst popularised by science fiction as extreme behemoths, are among the simplest of physical laboratories in the Universe, being characterised by only three properties: their mass, their angular momentum, and their electrical charge. In pedestrian terms, a black hole can be considered as the collapse of matter into a compact object, which has a radius approaching zero, and consequently an extremely high density. The density is indeed so extraordinarily high that within a finite radius, not even photons of light can escape the gravitational pull of a black hole: this radius is known as the event horizon of the black hole. A related measure, a black hole's Schwarzschild radius, is defined as the radius of a sphere that has the same surface area in flat space as that of the event horizon of a non-spinning black hole of a given

mass $M_\bullet$.

The value of each of the properties of a black hole provides insight into how that particular black hole came to presently be. As a black hole accretes matter from its surroundings, its mass increases, the angular momentum of the accreted matter is inherited by the black hole, and any electrical charge of the accreted matter is inherited too. On average, a black hole will have a neutral charge, as overall the Universe is more or less electrically neutral, and so any accreted positive charge is likely to be rapidly cancelled out by a negative charge, and vice versa. Hence, we generally only have the mass and angular momentum of a black hole to guide our understanding of its history. As we cannot observe a black hole using electromagnetic radiation (light) – photons cannot reflect from the black hole, and the black hole itself does not emit photons – we are left with two options: infer the properties of the black hole from the effect it has on its surroundings, or observe the black hole using something other than light. This thesis, and the papers contained herein, uses both approaches.

We humans love to classify things and black holes are no exception. Generally we use the mass of a black hole to delineate three categories:

1. Stellar mass black holes, with masses $\sim 5\,\mathrm{M_\odot}$ to $10^2\,\mathrm{M_\odot}$,

2. Intermediate mass black holes (IMBHs), with masses $\sim 10^2\,\mathrm{M_\odot}$ to $10^5\,\mathrm{M_\odot}$, and

3. Supermassive black holes (SMBHs), with masses $\gtrsim 10^5\,\mathrm{M_\odot}$.

This thesis primarily focuses on SMBHs. How the first SMBHs came to be is an open and active area of scientific research, with two main channels proposed. The first channel is dynamical, where stellar mass black holes, formed through the collapse of the first stars, merge and form more massive black holes (Volonteri et al., 2003a,b; Volonteri & Rees, 2005). The second channel is direct collapse black holes (DCBH, e.g. Haehnelt & Rees, 1993; Umemura et al., 1993; Begelman et al., 2006), formed through the general relativistic collapse of a massive gas cloud in a metal free environment that has not been allowed to fragment due to cooling processes, notably by molecular hydrogen. In both scenarios, the 'seed' black hole for an SMBH has a mass that puts it in the IMBH regime. The IMBH then grows through gas accretion and mergers with other black holes to form the SMBHs presently observed.

The merger of SMBHs through the emission of gravitational waves has yet to be directly detected and is a key observational target for the upcoming Laser Space Interferometer Antenna (LISA, e.g. Amaro-Seoane et al., 2007), designed to be sensitive to SMBHs in the $10^4\,\mathrm{M_\odot} \lesssim M_\bullet \lesssim 10^7\,\mathrm{M_\odot}$ mass range, and pulsar timing arrays

(PTAs, e.g. Agazie et al., 2023; Xu et al., 2023; Zic et al., 2023; EPTA Collaboration et al., 2024), which are well-suited for SMBHs with $M_\bullet \gtrsim 10^8 \, \mathrm{M_\odot}$. Several gravitational wave detections of stellar mass black holes have already been made ($2.6 \, \mathrm{M_\odot}$ to $137 \, \mathrm{M_\odot}$, Abbott et al., 2016, 2023; The LIGO Scientific Collaboration et al., 2025), highlighting gravitational wave astronomy as an exciting new avenue by which we can observe the Universe.

1.3 A bit about simulations

Thus far we have discussed galaxies and black holes in terms of what we have observed. We know that galaxies dynamically interact with each other, and we can infer the presence of supermassive black holes at the centres of almost all massive galaxies. Whilst it stands to reason that SMBHs also interact with each other through gravity, the frequency of these events, the timescales they take, and the overall effect of such SMBH mergers on their host galaxy remains an open area of investigation. This is where numerical simulations of galaxy systems are of paramount importance.

We can consider astronomy as a conversation between two groups of people: observers, who collect data about objects like galaxies, and simulators, who aim to organise and provide an evolutionary picture that describes what the observers see. Sometimes this conversation goes smoothly, other times there is conflict between what observers see and what simulators predict – it is in the latter that new science and understanding is found.

Simulations are necessary to model and obtain an understanding of galaxy evolution for two primary reasons. The first is that the timescales that a galaxy evolves over far exceed a human lifetime. With simulations, we are able to fast-track the evolution of a galaxy system, turning ten billion years of evolution into mere weeks or months of computation. The second is that many physical processes that are important for a galaxy's overall evolution are highly non-linear or even chaotic: small perturbations to the system can be amplified, leading to large-scale differences between what would otherwise be identical starting conditions.

This thesis is concerned with simulations of galaxy and SMBH co-evolution, the aims of which we outline now.

1.4 Thesis aims

The aim of this thesis is to explore the dynamics of supermassive black holes in massive galaxies, including the interplay the SMBH has on the structural properties

of the galaxy, and vice versa. We consider two specific regimes of interest: the pre-SMBH coalescence phase (Phase I), and the post-SMBH coalescence phase (Phase II).

In Phase I, we explore the sensitivity of two SMBHs orbiting each other (termed an SMBH binary) to noise introduced in the simulation. Noise can arise from both round-off errors when numerically solving the equations of motion of a particle, and also from discretisation effects introduced when sampling the initial phase space of a system. Both sources can lead to changes in the orbit of an SMBH binary, and can have far-reaching consequences for the merger timescale of the system. Accurate merger timescale are a key prediction required for ongoing and upcoming gravitational wave detection experiments, notably PTAs and LISA.

In Phase II, we explore the effect of a post-merger SMBH undergoing gravitational wave-induced recoil on the host galaxy. These recoil events, where the SMBH is violently displaced from the centre of the galaxy at the time of coalescence by asymmetric gravitational wave emission, can have a long-lasting impact on the stellar structure of the galaxy. This in turn can increase the difficulty of future SMBH binaries to merge in the system. Additionally, most observations indicate the presence of a central SMBH in nearly all massive galaxies – reconciling this observation with predictions of SMBHs being removed from the centres of galaxies through recoil is an ongoing avenue of investigation.

All simulations used in this thesis have been run with the publicly available simulation code GADGET-4 (Springel et al., 2021) with KETJU, the latter being developed at the University of Helsinki (Rantala et al., 2017; Mannerkoski et al., 2023). The simulation studies presented have a focus on achieving *statistical* significance, by leveraging both large numbers of independent simulations and novel analysis techniques, to ensure that the conclusions drawn are robust and reliable in their predictions.

1.5 Thesis structure

This thesis is structured in six chapters, of which this introduction forms the first. An overview of extragalactic theory, encompassing galaxy properties, the observational effects of their supermassive black holes, and the various stages to supermassive black hole binary coalescence, is given in Chapter 2. Strategies to simulate galaxy and supermassive black hole evolution, including an overview of the GADGET-4 and KETJU codes, is found in Chapter 3. Novel analysis techniques used in this thesis are described in Chapter 4. A summary of the publications presented in this thesis

are then given in Chapter 5, followed by the final concluding remarks in Chapter 6. The publications themselves can be found after the bibliography.

2 Extragalactic theory

2.1 Observational properties of massive early-type galaxies

2.1.1 Defining galaxy morphology

There exist many schema for classifying observed galaxies, yet one of the first and most prevalent is the Hubble 'Tuning Fork' (Hubble, 1926), shown in Figure 2.1. Galaxies are broadly classified as belonging to one of two categories: early-type galaxies (ETGs), which lie along the stem of the fork, and late-type galaxies (LTGs), which form the tines of the fork. Each category is then subdivided further: ETGs consist of elliptical and lenticular galaxies, and LTGs consist of barred and unbarred spiral galaxies.

Elliptical galaxies are identified by a smooth surface brightness profile (e.g. Prugniel & Simien, 1996, 1997), where isophotes (contours of equal surface brightness) are ellipses. The degree of circularity in the central isophotes gives rise to the sub-classification of ellipticals, denoted by $n_{\text{elliptical}}$, and is calculated as:

$$n_{\text{elliptical}} = 10\left(1 - \frac{b}{a}\right) = 10(1 - \varepsilon), \tag{2.1}$$

where a and b are the semimajor and semiminor axes of the isophote, respectively, and ε is the ellipticity. Referring again to Figure 2.1, E0 ellipticals have almost perfectly round isophotes, whereas E6 ellipticals have elongated isophotes. Importantly, these isophotes describe the projected shape of the galaxy: inferring the intrinsic, three-dimensional shape of the galaxy is a highly non-trivial task.

Similar to elliptical galaxies, spiral galaxies of both barred and unbarred kind are further categorised, with the stratification depending on the number of spiral arms, and how tightly-wound those arms are. As this thesis is primarily focused on ETGs, we do not comment further on spiral galaxy morphological classification.

ETGs and LTGs also differ in properties other than just visual morphology, with
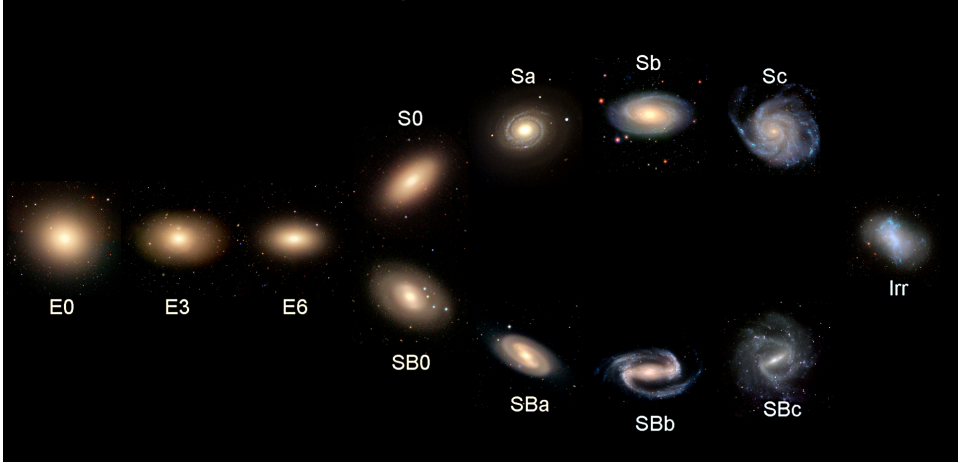
Figure 2.1: The Hubble galaxy morphology classification scheme, commonly referred to as the 'Hubble Tuning Fork'. Elliptical galaxies are to the left of the diagram, along the stem of the fork, whereas spiral galaxies are to the right, and divided along the two tines of the fork. The irregular ('Irr' in the figure) class is separate to the fork, owing to the lack of spiral arms and smooth isophotes. Image credit: Zooniverse.

a readily-obtained metric being the colour of the galaxy. Typically, colour is determined by observing the galaxy through a filter centred on one wavelength, calculating its magnitude, and then repeating for a second filter centred on a different wavelength. The commonly-used Sloan Digital Sky Survey (SDSS, York et al., 2000) passbands u, g, r, i, z are shown in the upper panel of Figure 2.2 (Fukugita et al., 1996; Rodrigo & Solano, 2020), where their normalised transmission $\tilde{T}$ follows a complex shape. We show the $g - i$ colours (i.e. difference between the g-band magnitude, centred at $\lambda_0 = 475\,\text{nm}$ with FWHM[1] $\Delta\lambda = 118\,\text{nm}$ and i-band magnitude, $\lambda_0 = 752\,\text{nm}$ and $\Delta\lambda = 68\,\text{nm}$) for a subset of galaxies with spectral information and morphological classification from SDSS in the lower left panel of Figure 2.2. A clear difference between classes is seen, with LTGs displaying a broad distribution of colour peaking at $g - i \sim 1.8$. Conversely, ETGs demonstrate a narrower distribution of colour, peaking at larger values of $g - i \sim 2.6$.

As colour is a difference in magnitudes, it also describes the ratio of fluxes through

---

[1]The full width at half maximum (FWHM) is a measure of the spread of a distribution or function at the points where the intensity is half the maximum intensity.
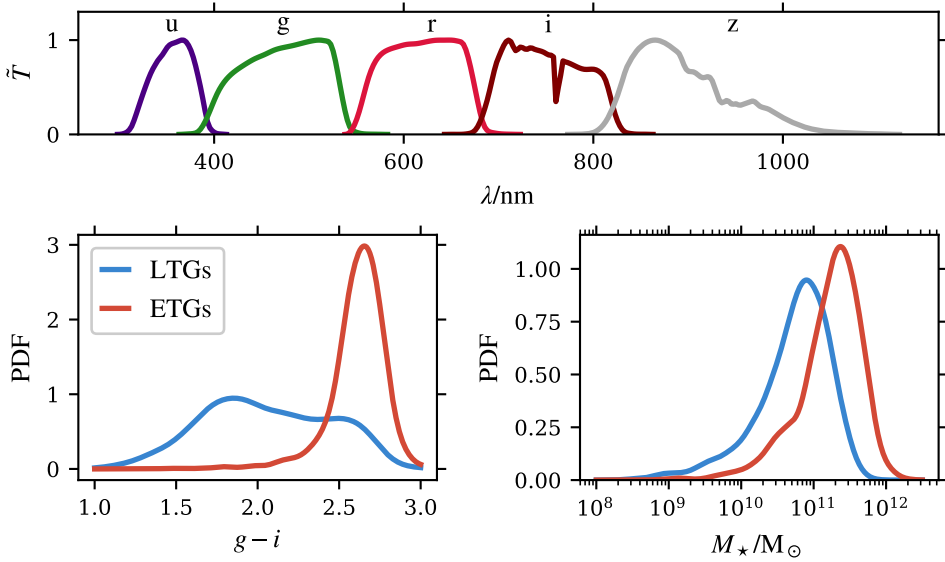
Figure 2.2: *Top panel*: Normalised transmission curves $\tilde{T}$ for the SDSS passbands as a function of wavelength. Note the complex shape of the curves, which are not top-hat filters. *Bottom left panel*: Distribution of $g - i$ colours for SDSS galaxies. LTGs peak at smaller values of $g - i$, and are typically bluer than ETGs. *Bottom right panel*: Stellar mass distribution of SDSS ETGs and LTGs. The ETG stellar mass distribution peaks at higher masses than LTGs by a factor of ~ 3.

the two used filters:

$$m_g - m_i = -2.5 \log_{10} \left(\frac{F_g}{F_i} \right), \tag{2.2}$$

where m_X and F_X represent the magnitude and flux through a filter X, respectively. In our case, a larger $g - i$ value indicates more light is transmitted through the i filter than the g filter, leading to those galaxies being termed 'red': the converse is true, where a lower value of $g - i$ corresponds to a 'blue' galaxy. In Figure 2.2, we see that ETGs are more red than LTGs.

A difference in colour is a clear indication of a difference in the underlying physical processes at play within a given galaxy (e.g. Mo et al., 2010), in particular the dominant stellar population – which is often the source for most of the emitted light from a galaxy – and dust content. Ongoing star formation typically produces a large population of massive stars, which are very hot, however live only for a short time owing to the efficiency of the nuclear fusion reaction fueling them. These hot, young

stars appear blue: a large population of such stars will result in the galaxy overall appearing blue. Conversely, an older stellar population which is not undergoing as rapid fusion as young stars and is not in the midst of ongoing star formation, are cooler and thus appear more red. Collectively, an old stellar population results in a predominantly red galaxy. It is for this reason that the moniker 'red and dead' is often used to described ETGs. The presence of dust can complicate this dichotomy however (e.g. Wise & Silva, 1996). If an LTG has a high dust content, then this galaxy too may appear red, due to the absorption and scattering of stellar light by the dust.

With an understanding of the stellar population within a galaxy and an assumption of the dust content, one may obtain a mass-to-light ratio Υ, which describes the amount of stellar mass per unit luminosity. Whilst for ETGs the population of old, red stars provides a relatively constant $\Upsilon = 3.5\,\mathrm{M_\odot}\,\mathrm{L_\odot}^{-1}$ to $4.0\,\mathrm{M_\odot}\,\mathrm{L_\odot}^{-1}$ (e.g. Bonfini & Graham, 2016), for LTGs Υ can be complex, with spatial variation apparent. Comparing the obtained total stellar masses $M_\star$ between LTGs and ETGs in SDSS in Figure 2.2, we see that ETGs are generally of higher mass than LTGs, peaking at $\sim 2 \times 10^{11}\,\mathrm{M_\odot}$, compared to $\sim 7 \times 10^{10}\,\mathrm{M_\odot}$ for LTGs.

2.1.2 Star formation in ETGs

Whilst we have explored star formation in the context of colour, more important is the rate of star formation in a galaxy. Stars form in pockets of dense, cold gas – for a star to form, the inward gravitational force of the gas cloud must at least exceed the outward force of the thermal pressure of the gas. Additionally, the presence of magnetic pressure force from the gas often increases the outward pressure force that gravity must overcome. The critical mass of the gas cloud for which collapse is inevitable is termed the Jeans mass (Jeans, 1902), and is defined:

$$M_\mathrm{J} = \frac{4\pi}{3}\rho\left(\frac{1}{2}\lambda_\mathrm{J}\right)^3,\tag{2.3}$$

where ρ is the density and λ_J is the Jeans length, the characteristic length scale above which perturbations lead to the collapse of the system, and is in turn defined:

$$\lambda_\mathrm{J} = c_\mathrm{s}\sqrt{\frac{\pi}{G\rho}} = \sqrt{\frac{5k_\mathrm{b}T}{3m_\mathrm{H}}}\sqrt{\frac{\pi}{G\rho}}.\tag{2.4}$$

In Equation 2.4, c_s is the sound speed of the gas, and m_H is the proton mass. Combining Equation 2.3 and Equation 2.4, we see that for collapse to occur the
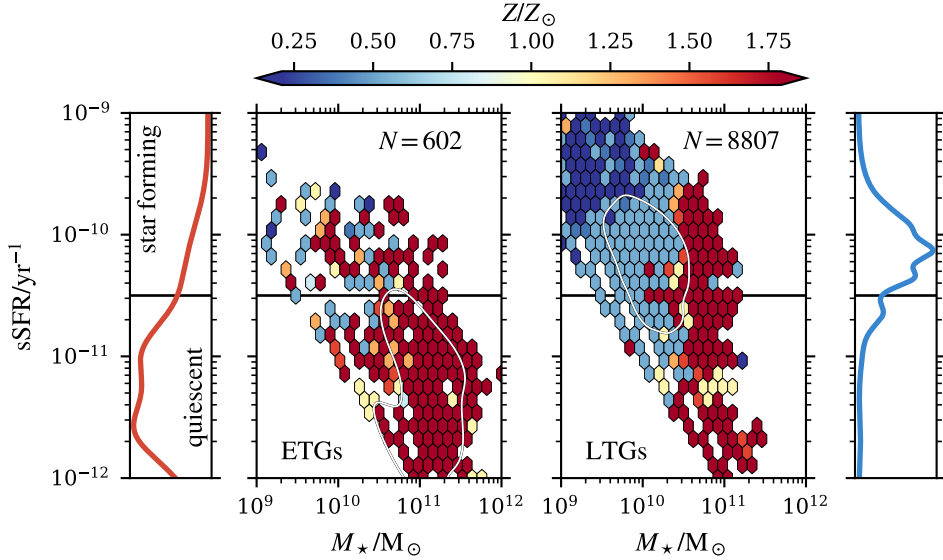
Figure 2.3: Specific star formation rate of a random sample of SDSS galaxies, using the stellar population synthesis code described in Maraston et al. (2009). Each hexbin depicts the median metallicity of all galaxies within that bin. The black horizontal line indicates the quiescent threshold, $\mathrm{sSFR}/\mathrm{yr}^{-1} < 10^{-10.5}$ (Bisigello et al., 2020), and the white loop indicates the 50% iso-density contour (i.e. 50% of the data lies within the loop). ETGs are typically more quiescent and metal-rich than LTGs. There are also typically more LTGs that ETGs: in this sample, there are 602 clearly-identified ETGs, and 8807 clearly-identified LTGs.

cloud mass M' must satisfy:

$$M' > M_{\mathrm{J}} \propto T^{3/2}\rho^{-1/2}. \tag{2.5}$$

Whilst these conditions for collapse are more readily found in LTGs than ETGs where there is a high density of low thermal energy gas (predominantly within the disc of the galaxy), environmental impacts, such as gas inflows (e.g. Cresci et al., 2010; Sánchez Almeida et al., 2013; Ceverino et al., 2016) and ram pressure (e.g. Kronberger et al., 2008; Kapferer et al., 2009), may still give rise to star formation by compressing the gas to high densities. If however the frequency and timescales that these conditions occur over are rare, then star formation, and the subsequent impact on the galaxy during the lifetime and demise of the star, is minimal.

To this end, we can estimate a star formation rate (SFR) using stellar population synthesis (SPS) models. A synthetic stellar population is created assuming an initial

stellar mass function (the probability a newly-formed star has a given mass) coupled with stellar evolution codes to predict, for a given initial state, what the stellar population properties are (e.g. Maraston et al., 2006; Conroy et al., 2009; Maraston et al., 2009). Typical properties we are interested in include the metallicity, age, and star formation rate. A large grid of template models are produced following this method, with each sampled initial state providing one point of the grid. By considering the integrated light from each grid point, and determining the best-fit template to the observed galaxy spectra (typically through Monte Carlo Markov Chain sampling), we gain insight into the likely evolution of the stellar population of a galaxy.

Such a modelling technique has been applied to SDSS galaxies (e.g. Cid Fernandes et al., 2005; Gallazzi et al., 2005; Maraston et al., 2013), a subset of which is shown in Figure 2.3. In the main panels is shown, as a function of total stellar mass $M_\star$, the *specific* star formation rate, sSFR:

$$sSFR = \frac{SFR}{M_\star}. \tag{2.6}$$

By normalising the SFR by the total stellar mass of the galaxy, we isolate the fraction of a galaxy's mass that constitutes new stellar material: in other words, how important is star formation to the overall galaxy and its evolution?

Galaxies for which their current star formation is not important for their overall evolution are termed quiescent: such galaxies satisfy (e.g. Bisigello et al., 2020):

$$sSFR/yr^{-1} < 10^{-10.5}, \tag{2.7}$$

and is shown as the black horizontal line in Figure 2.3.

For the sample of SDSS galaxies in Figure 2.3, immediately apparent is that the majority of ETGs have a sSFR which is less than the quiescent threshold: the peak in the marginal distribution is close to sSFR $\sim 3 \times 10^{-12}\,yr^{-1}$. As a result, the amount of newly formed stars is very low, and has a negligible impact on the overall evolution on the galaxy. This is contrasted to LTGs, for which the vast majority have sSFR $\sim 10^{-10}\,yr^{-1}$: star formation is an important physical process in these systems.

The low sSFR of ETGs indicates that there is generally a lack of star-forming gas in these systems (e.g. Roberts et al., 1991; Buson et al., 1993). Any gas that is present is hot and has a low density (e.g. Fabbiano et al., 1989; Mathews & Brighenti, 2003), such that the thermal pressure of the gas is able to withstand the contraction due to gravity over times much longer than a Hubble time. This inference is one of

the main reasons why ETGs can be well-modelled in numerical simulations of galaxy formation and evolution as being gas-free.

Another trend to be seen in Figure 2.3 is that massive star-forming systems typically have a metallicity less than solar, whereas massive quiescent systems typically have metal-rich stellar populations. This can be readily understood in the context of stellar evolution. For a massive galaxy to be quiescent, there cannot be large quantities of star-forming gas in the system: previous generations of star formation have already expended the available gas. A massive star is able to continue nuclear fusion until there is a significant build up of iron in the stellar core: at this point, the star explodes in a supernova, releasing its outer convective layers – which are metal-rich as a result of the different generations of nuclear fusion – into the interstellar medium (ISM) (e.g. Carroll & Ostlie, 2017). In a massive galaxy such as an ETG, the gravitational potential of the system is deep enough to retain the metals formed by supernovae, unlike less-massive (e.g. dwarf) systems (Dekel & Silk, 1986). These factors serve to increase the overall metallicity of the ISM, and give an indication on the number of generations of star formation that have occurred (assuming some initial stellar mass function). Consequently, those systems without star-forming gas, which have undergone multiple episodes of star formation which release metals into the ISM, have resulting high metallicities – explaining the metallicity trend in Figure 2.3.

2.1.3 Kinematics of ETGs

ETGs also show distinct differences to LTGs, as well as within-population variation, in terms of their kinematics. Observationally, one may obtain a line profile of a galaxy using either narrow slit spectroscopy, or integral field unit (IFU) spectroscopy. The former method captures a single measure of a galaxy's kinematics (e.g. Statler, 1991; Arnold et al., 1994; de Zeeuw, 1996); IFU spectroscopy supercedes narrow slit spectroscopy by obtaining a spectrum for a grid of pixels, allowing for local kinematic variation to be studied (e.g. Bacon et al., 1995, 2000, 2001; de Zeeuw et al., 2002; Croom et al., 2012; Sánchez et al., 2012; Bundy et al., 2015). The output from IFU spectroscopy is therefore a datacube containing two spatial dimensions and one spectral dimension. A given area element in the spatial dimensions with the corresponding information from the full spectral dimension is termed a 'spaxel' (spatial pixel). The construction of mock IFU images is discussed in Section 4.2.

After correcting for redshift, differences in spectral line positions relative to their rest frame positions provide valuable information about the internal motion of the galaxy, specifically the line-of-sight velocity distribution (LOSVD) of stars.
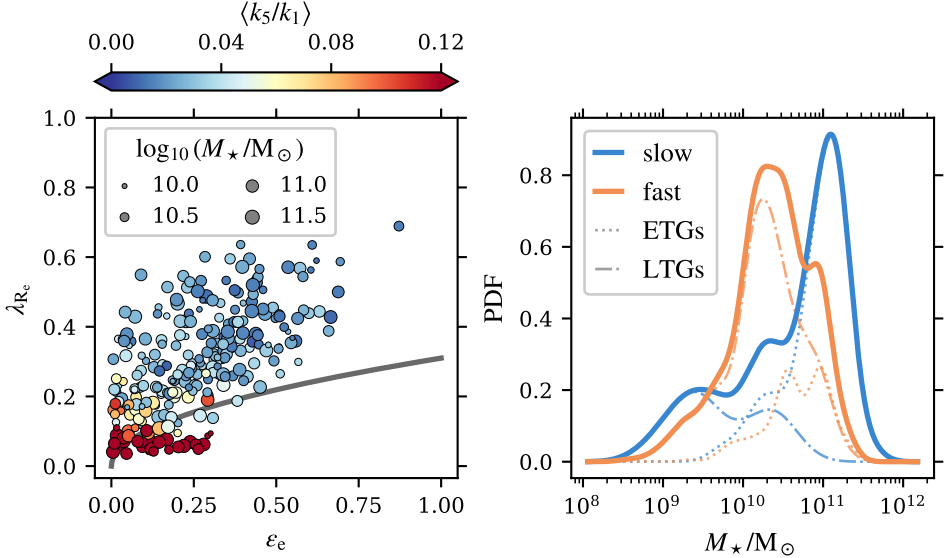
Figure 2.4: *Left panel*: Spin parameter as a function of galaxy ellipticity from van de Sande et al. (2017), with marker size indicating stellar mass. Points are coloured by the degree of kinematic asymmetry $\langle k_5/k_1 \rangle$, which measures the deviation of observed rotation from a cosine law (Equation 2.10). The grey line indicates the fast-slow rotator division from Emsellem et al. (2011). Regular rotators (blue points) are generally also fast-rotators, whereas slow-rotators often show non-regular rotation (red points). *Right panel*: Distribution of galaxy SAMI DR3 (Croom et al., 2021) stellar masses stratified by fast or slow rotation. The thick solid lines indicate all fast or slow rotators, with the dotted lines showing the contribution of galaxies identified as ETGs to the total, and dash-dotted lines the contribution of galaxies identified as LTGs to the total. Slow rotators are typically more massive than fast-rotators, and tend to be ETGs.

Typically, the LOSVD of a system is fit using a Gauss-Hermite series, where the first order moment is the mean of the best-fit Gaussian profile and the second order moment the standard deviation. Higher order moments j are given by the coefficients of the Hermite polynomials (van der Marel & Franx, 1993), defined:

$$H_j(w) = (-1)^j e^{w^2} \frac{\mathrm{d}^j e^{-w^2}}{\mathrm{d}w^j}.$$
(2.8)

14

Together, the LOSVD is parameterised as:

$$\mathscr{L}(w) = \frac{1}{\sqrt{2\pi}\sigma} e^{-w^2/2} \left\{ 1 + \sum_{j=3}^{N} h_j H_j(w) \right\}, \tag{2.9}$$

where $w \equiv (v_{\mathrm{LOS}} - V)/\sigma$. Whilst arbitrarily-high orders are possible for $\mathscr{L}$, typically a maximum of 4th order is used (e.g. van de Sande et al., 2017). A discussion on the interpretation of the different moments is given in Section 4.2.

Another powerful technique is the estimation of kinematic asymmetry of a system, such as that caused by counter-rotating subcomponents or kinematically-decoupled components (KDCs) at the centres of galaxies. Let us assume the projected velocity field is described by a cosine law along ellipses:

$$V(R, \theta) = V_{\mathrm{rot}}(R) \cos\theta, \tag{2.10}$$

on which we can perform a Fourier decomposition to investigate deviations from this profile (Krajnović et al., 2011), where each harmonic j in the Fourier series has an amplitude k_j. The rotational velocity V_{rot} is then described by the leading moment k_1, and the amplitudes of the higher odd-moments give the asymmetric deviations from this. The kinematic asymmetry is then a luminosity-weighted average of $\langle k_5/k_1 \rangle$ (Krajnović et al., 2011; van de Sande et al., 2017). Consequently, $k_{j|j>1} \approx 0$ indicates that there is little asymmetry in the velocity profile, and $|k_{j|j>1}| > 0$ indicates that there are higher-order harmonics present in the velocity profile: hints for complex inner kinematic substructure to the galaxy.

To quantify rotation, a commonly-used parameter by observers is the spin parameter within the effective radius[2] R_{e} (Emsellem et al., 2007, 2011):

$$\lambda_{R_{\mathrm{e}}} \equiv \frac{\langle R|V| \rangle}{\langle R\sqrt{V^2 + \sigma^2} \rangle} = \frac{\sum_i F_i R_i |V_i|}{\sum_i F_i R_i \sqrt{V_i^2 + \sigma_i^2}} \quad |R_i \leq R_{\mathrm{e}}. \tag{2.11}$$

The quantity is flux-averaged over IFU spaxels, hence the F_i term in the second equality.

To identify a galaxy as either fast rotating or slow rotating, Emsellem et al. (2011) describe fast-rotators as having a $\lambda_{R_{\mathrm{e}}}$ value exceeding that expected for its ellipticity ε_{e} from the empirical relation fit to the ATLAS$^{\mathrm{3D}}$ survey data:

$$\lambda_{R_{\mathrm{e}}} \geq 0.31\sqrt{\varepsilon_{\mathrm{e}}}, \tag{2.12}$$

---

[2]The effective radius is defined as the radius within which half of the total stellar light is contained.

where the factor 0.31 is the particular scaling when considering an aperture of $1\,R_{\mathrm{e}}$. This relation is shown as the grey line in the left panel of Figure 2.4. Also shown are a sample of galaxies with well-resolved kinematics but without a morphological classification from the Sydney-AAO Multi-object Integral field spectrograph (SAMI, Croom et al., 2012) taken from van de Sande et al. (2017). The galaxies are coloured by the kinematic asymmetry $\langle k_5/k_1 \rangle$, with 'regular rotation' corresponding to $\langle k_5/k_1 \rangle \leq 0.04$. Regular rotators make up the vast majority of this sample. There is however a population of non-regular rotators ($\langle k_5/k_1 \rangle > 0.08$) and quasi-regular rotators ($0.04 < \langle k_5/k_1 \rangle \leq 0.08$). Non-regular rotation may result from different physical processes, including kinematically-decoupled cores (i.e. an inner region that rotates differently to the outer region of the galaxy), counter-rotating discs, or irregular rotation (Krajnović et al., 2011). Non-regular rotators have a low $\lambda_{R_{\mathrm{e}}}$ and low ellipticity, making up the majority of the slow-rotator population in the van de Sande et al. (2017) sample.

When we consider the distribution of galaxy mass stratified by the galaxy belonging to either the slow-rotator or fast-rotator population (defined through Equation 2.12), we see that each population has a distinct peak in mass: fast-rotators at $M_\star \sim 10^{10}\,\mathrm{M}_\odot$, and slow-rotators $M_\star \sim 10^{11}\,\mathrm{M}_\odot$. This is closely correlated with the peak of the mass distributions for LTGs and ETGs, respectively. When we consider the make-up of the slow-rotator population (with ETGs represented as dotted lines, and LTGs as dash-dotted lines in Figure 2.4), we clearly see that the majority of slow-rotators in the SAMI (Croom et al., 2021) sample are ETGs, with LTGs dominating the slow-rotator population for $M_\star \lesssim 10^{10}\,\mathrm{M}_\odot$. Similarly for the fast-rotator population, LTGs are the dominant morphology over ETGs (e.g. Cappellari et al., 2011). These are of course general trends, and whilst slow-rotators *tend* to be ETGs, not all ETGs are necessarily slow-rotators (e.g. Emsellem et al., 2007).

2.1.4 Cored vs cuspy ETGs

Thus far, we have mainly considered the differences between ETGs and LTGs. There are however two important subclasses of ETGs identified through photometry: cored and cuspy.

The terms 'cored' and 'cuspy' refer to the projected luminosity (or equivalently mass) density profile at small radii from the galaxy centre, as shown in the schematic in the right panel of Figure 2.5. Historically, the surface brightness profile $I(R)$ of
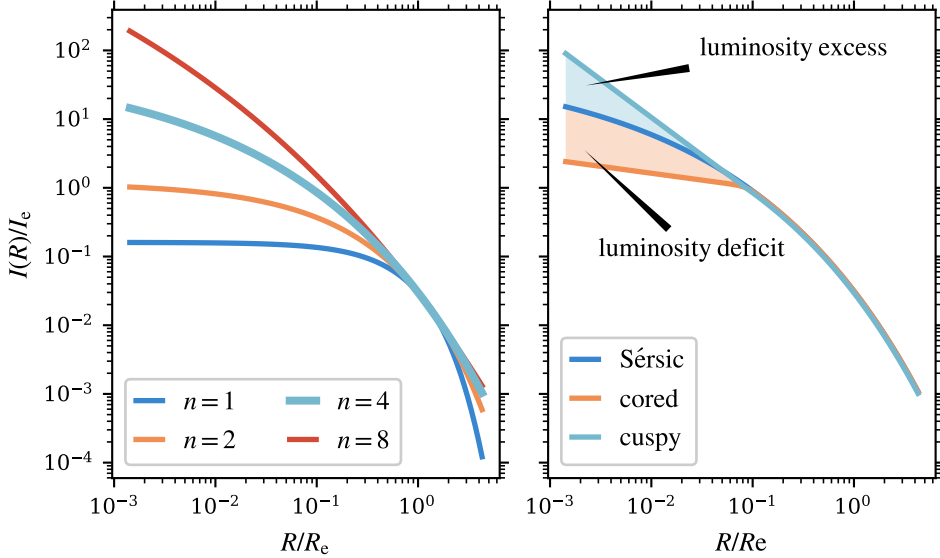
Figure 2.5: *Left panel*: Effect of changing the Sérsic index n on the luminosity profile given in Equation 2.14. The thick, light blue line corresponds to the de Vaucouleur's profile. *Right panel*: Schematic of a Sérsic (blue), cored (orange), and cuspy (light blue) luminosity profiles. In each profile, the Sérsic index $n = 4$, and the effective radius $R_\mathrm{e} = 7\,\mathrm{kpc}$. For a cored profile, the luminosity deficit is the integrated difference from the smallest-resolvable radius to the core radius compared to the extrapolated best-fit Sérsic profile. The luminosity excess is calculated similarly for the cuspy profile, however in this instance there is an excess of stellar light compared to the Sérsic profile.

ETGs was described by de Vaucouleurs (1948):

$$I(R) = I_\mathrm{e} \exp\left\{-7.669\left[\left(\frac{R}{R_\mathrm{e}}\right)^{1/4} - 1\right]\right\}, \tag{2.13}$$

where I_e is the surface brightness at the effective radius R_e. Equation 2.13 is a special case of the more general Sersic (1968) profile with $n = 4$:

$$I(R) = I_\mathrm{e} \exp\left\{-b_n\left[\left(\frac{R}{R_\mathrm{e}}\right)^{1/n} - 1\right]\right\}. \tag{2.14}$$

The parameter b_n is defined in terms of the gamma function and lower incomplete gamma function, but to a good approximation in the range $0.5 < n < 10$ is (Prugniel
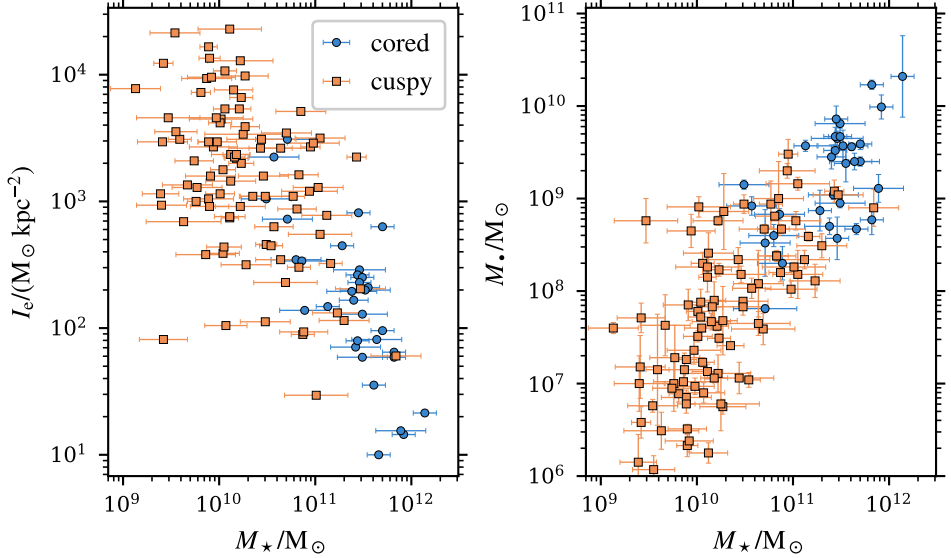
Figure 2.6: Properties of ETGs from Sahu et al. (2020) separated into cored and cuspy systems. *Left panel*: Stellar mass and bulge intensity (converted to mean surface mass density) at the effective radius. An increasing stellar mass is associated with a decreasing bulge intensity. *Right panel*: Stellar mass and SMBH mass. Converse to the I_e-$M_\star$ relation, there is a positive correlation between $M_\star$ and $M_\bullet$ – more massive SMBHs typically reside in more massive galaxies.

& Simien, 1997):

$$b_n = 2n - \frac{1}{3} + \frac{0.009876}{n}. \tag{2.15}$$

The Sérsic profile for different values of n is shown in the left panel of Figure 2.5.

It was soon noticed that the most luminous (and hence most massive) ETGs have a depleted central stellar luminosity, as shown in the left panel of Figure 2.6 for a sample of galaxies at $z < 0.04$ collated from various sources by Sahu et al. (2020). The cored stellar density profiles may be parameterised in a number of ways (e.g. Lauer et al., 1995; Graham et al., 2003), however most include a characteristic radius within which the luminosity profile begins to significantly deviate from the outer regions: this is often termed the core, or alternatively break, radius. A widely used parametrisation is the cored-Sérsic profile, which deviates from a Sérsic profile

at radii less than the core radius, and is defined (Graham et al., 2003):

$$I(R) = I' \left[1 + \left(\frac{R_{\rm b}}{R} \right)^\alpha \right]^{\gamma/\alpha} \exp \left[-b_n \left(\frac{R^\alpha + R_{\rm b}^\alpha}{R_{\rm e}^\alpha} \right)^{1/(\alpha n)} \right] \qquad (2.16)$$

where

$$I' = I_{\rm b} 2^{-\gamma/\alpha} \exp \left[b_n \left(2^{1/\alpha} \frac{R_{\rm b}}{R_{\rm e}} \right)^{1/n} \right]. \qquad (2.17)$$

Here $R_{\rm b}$ is the core radius, $I_{\rm b}$ is the projected surface density at the core radius, γ is the core slope, and α is the profile transition index, describing how abruptly the inner region deviates from a Sérsic profile at $R_{\rm b}$. Equation 2.16 is shown as the orange line in the right panel of Figure 2.5.

The most widely-accepted theory for the formation of a stellar core in massive galaxies is due to SMBH binary scouring, which we discuss in detail in Section 2.2. Supporting this theory, for a given ETG, a larger stellar mass is positively correlated with a larger SMBH mass (right panel of Figure 2.6). Consequently, those ETGs which are cored typically host a more massive SMBH than those galaxies identified as cuspy (e.g. Sahu et al., 2020).

The interaction of a galaxy with its central SMBH is complex, and is the primary theme of this thesis. We explore this interplay next.

2.2 Supermassive black holes and their scaling relations

We have seen that typically a larger stellar mass coincides with a larger SMBH mass (Figure 2.6). This positive correlation implies that galaxies and their SMBHs most likely grow together.

Aside from the merger of SMBHs (Section 2.3), the primary mechanism by which SMBHs grow is through the accretion of cold gas (Taniguchi et al., 1999; Haehnelt & Kauffmann, 2000; Volonteri et al., 2003a; Bromley et al., 2004; Silk, 2005). Assuming simple spherical accretion, an SMBH can accrete gas from a region where its gravitational potential overcomes the specific thermal energy of the gas (Bondi, 1952). The accretion rate onto the SMBH is then:

$$\frac{{\rm d}M_\bullet}{{\rm d}t} \sim 4\pi r_{\rm acc}^2 \rho_{\rm acc} c_{\rm s,acc}, \qquad (2.18)$$

where the accretion radius is given by:

$$r_{\rm acc} \sim \frac{GM_\bullet}{c_{\rm s}^2} \qquad (2.19)$$

Figure 2.7: Scaling relations with stellar velocity dispersion σ. *Left panel*: Stellar velocity dispersion as a function of the galaxy K-band (2.2 µm) luminosity L_K, which is a good proxy to the stellar mass $M_\star$. *Right panel*: SMBH mass as a function of stellar velocity dispersion. A larger velocity dispersion correlates with a higher SMBH mass, owing to the deeper gravitational potential of a more massive SMBH compared to a less massive SMBH. In both instances, the data is taken from van den Bosch (2016).

and $\rho_{\rm acc}$ is the gas density and $c_{\rm s,acc}$ the gas sound speed at the accretion radius. Substituting Equation 2.19 into Equation 2.18, and accounting for the relative velocity between the gas and the SMBH $v_{\rm rel}$, we arrive at the widely-used Bondi-Hoyle-Lyttleton accretion model[3]:

$$\frac{{\rm d}M_\bullet}{{\rm d}t} = \frac{4\pi G^2 M_\bullet^2 \rho}{\left(c_{\rm s}^2 + v_{\rm rel}^2\right)^{3/2}} \tag{2.20}$$

(Hoyle & Lyttleton, 1939; Bondi & Hoyle, 1944; Bondi, 1952). Most SMBHs are surrounded by an accretion disc of gaseous material, thus the accretion of low angular momentum gas is more probable than perfectly spherical accretion (e.g. Salpeter, 1964; Lynden-Bell, 1969; Novikov & Thorne, 1973; Pringle & Rees, 1972; Shakura & Sunyaev, 1973).

---

[3]In numerical modelling, Equation 2.20 is typically scaled by an artificial boost parameter α.

Our simple approximation notwithstanding, it stands to reason that a more massive galaxy with more cold gas is able to feed, and hence grow, its SMBH more rapidly than a lower-mass galaxy. The scenario of course changes if the gas is hot, thus making accretion harder, as is the case in massive ETGs – by this stage however, the SMBH is already generally very massive.

A consequence of an increased amount of mass interior to some radius r is an elevated circular velocity:

$$v_{\mathrm{circ}} = \sqrt{\frac{GM(< r)}{r}}. \tag{2.21}$$

Let us assume that we are interested in the inner regions of a galaxy, where the gravitational potential is dominated by the SMBH, and the stars in its vicinity are in a steady-state. In this instance, the mean velocity along some axis is zero, $\langle v_i \rangle = 0$. Recalling that variance of a quantity X is defined through the expectation operator E as:

$$\sigma_X^2 \equiv \mathrm{E}[X^2] - \mathrm{E}[X]^2, \tag{2.22}$$

then we have for our system $\sigma_i = \sqrt{\langle v_i^2 \rangle} \implies \sigma_t = \sqrt{\langle v_{\mathrm{circ}}^2 \rangle}$, assuming either cylindrical or spherical geometry. Hence, we expect for steady-state systems that not only should v_{circ} increase as M increases, but so too should the stellar velocity dispersion σ.

Our expectations are confirmed by observational data comparing the K-band (2.2 µm) luminosity (which is a good proxy of stellar mass, as it is less sensitive to dust than optical wavebands, and most of the stellar mass of a galaxy is in red dwarfs, e.g. Bell et al., 2003; Sureshkumar et al., 2021) of massive galaxies and the projected velocity dispersion, shown in the left panel of Figure 2.7. We see that a typical massive ETG of luminosity $L_K \sim 10^{11}\,\mathrm{L}_\odot$ has a stellar velocity dispersion in the range of $\sim 150\,\mathrm{km\,s^{-1}}$ to $200\,\mathrm{km\,s^{-1}}$ (van den Bosch, 2016).

When we consider just the SMBH mass, we see a tight correlation with stellar velocity dispersion in the right panel of Figure 2.7. As expected from the $M_\bullet$-$M_\star$ scaling relation shown earlier, coupled with our previous discussion on σ, a larger SMBH mass correlates with a higher velocity dispersion, with a $10^9\,\mathrm{M}_\odot$ SMBH corresponding to $\sigma \sim 200\,\mathrm{km\,s^{-1}}$ to $300\,\mathrm{km\,s^{-1}}$. This famous relation is simply called the $M_\bullet$-σ relation (e.g. Ferrarese & Merritt, 2000; Haehnelt & Kauffmann, 2000; Häring & Rix, 2004; McConnell & Ma, 2013), and is often used to estimate an SMBH mass if stellar kinematics data is available but no independent measure of the SMBH mass is possible.

The connection between SMBH mass and the stellar environment encodes the radial extent to which an SMBH may significantly affect the dynamics of stars. This
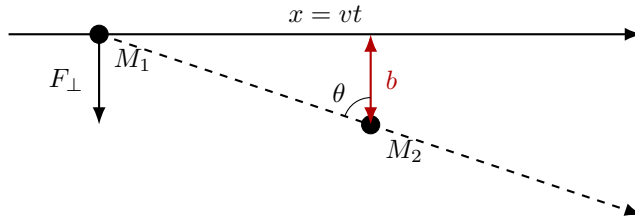
Figure 2.8: Schematic of the impact parameter b of a body M_1 passing by a second body M_2, approximated as a straight-line trajectory along the x-axis. The body M_1 experiences an impulse due to the perpendicular force exerted by M_2, $F_\perp$. In the case of dynamical friction acting on an SMBH, M_1 is a stellar particle and M_2 is the SMBH, thus $M_2 \gg M_1$.

radius, termed the influence radius r_{h}, is defined (e.g. Merritt, 2013) as:

$$r_{\mathrm{h}} = \frac{GM_\bullet}{\sigma^2}. \tag{2.23}$$

The orbits of stars at radii $r < r_{\mathrm{h}}$ are strongly influenced by – and show distinct signatures of – the presence of the SMBH. A more detailed exploration of stellar orbits is given in Section 4.1.

2.3 Supermassive black hole binaries

Thus far, we have considered galaxies in isolation. It is a well-documented observation however that galaxies merge: that is, mutual gravitational attraction brings two or more galaxies together, with the less massive galaxy being absorbed by the larger, or both galaxies being disrupted if they are of comparable mass. With our working hypothesis of a single SMBH at the centres of most if not all massive galaxies, it is to be expected that the SMBHs will also dynamically interact, such that they lose energy and angular momentum to 'sink' towards the centre of the merging galaxies. That the vast majority of observations indicate that there is a *single* SMBH at the centres of galaxies suggests this process of two SMBHs losing energy is an overall efficient process. Pioneering work by Begelman et al. (1980) describes the three broad steps to SMBH binary coalescence, which we discuss now in the context of two galaxies merging.

2.3.1 Dynamical friction

The first mechanism by which orbital energy is lost, and the most efficient over large distances, is dynamical friction. If we consider the two merging galaxies as point
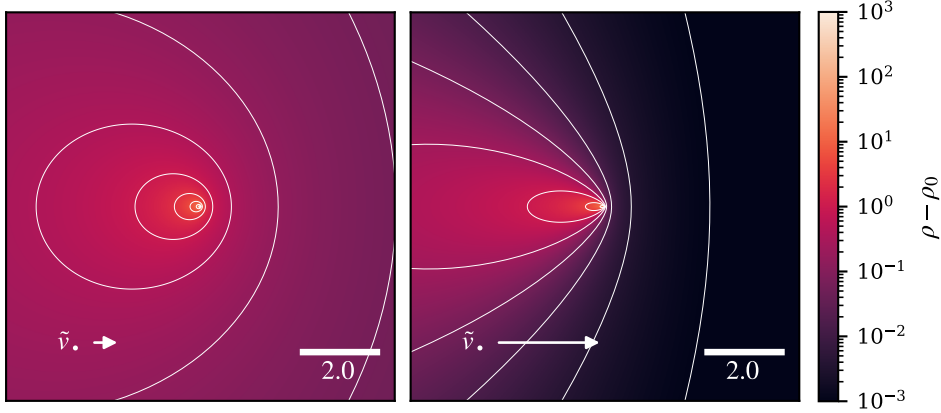
Figure 2.9: Dynamical friction on a massive body resulting from the induced overdensity through the surrounding medium (Mulder, 1983). Velocity units are scaled by the velocity dispersion of the medium, $\tilde{v}_\bullet = v_\bullet/\sigma$. The left panel shows an initial velocity of $\tilde{v}_\bullet = 1$ and the right panel $\tilde{v}_\bullet = 3$, indicated by the arrows in the figure. The induced overdensity, and hence the dynamical friction experienced by the body, is greater for a lower $\tilde{v}_\bullet$ than a higher $\tilde{v}_\bullet$.

masses[4] of M_1 and M_2, the distance of minimum approach between them is termed the impact parameter b. A schematic is shown in Figure 2.8. The more radial the original orbit of the two galaxies is, the smaller the impact parameter b is likely to be (Khochfar & Burkert, 2006). The close passage of the body M_1 causes an impulse from the perpendicular force exerted by M_2, altering the trajectory of M_1.

In doing so, the orbital kinetic energy of the galaxies as they pass by each other is transferred to random motion of the individual particles within the galaxy, rapidly leading the galaxies to merge. How the SMBHs behave, embedded within a stellar cusp owing to their deep potential wells, depends on the mass ratio between the merging galaxies. If we define the mass ratio q between the two merging galaxies as:

$$q \equiv \frac{M_{\star,2}}{M_{\star,1}} \quad \bigg| q \leq 1, \qquad (2.24)$$

---

[4]Realistically, galaxies are extended objects, and not point masses. A further discussion of this, and how M_1 and M_2 are determined in practice, is provided in Subsection 3.4.3.

then we have two classes of mergers: major mergers with $1/3 \lesssim q \leq 1$ and minor mergers with $q < 1/3$. In major mergers, the SMBHs remain well-bound within their stellar cusps until the nuclei of the progenitor galaxies merge (Gualandris & Merritt, 2012). Conversely, for minor mergers, the outer material of the less-massive galaxy is lost due to tidal stripping, and its SMBH slowly sinks to the centre of the merger remnant (Callegari et al., 2011). In both instances, the SMBHs sink to the centre of the merger through equipartition of kinetic energy with the surrounding medium:

$$M_\bullet \langle v_\bullet^2 \rangle = m_m \langle v_m^2 \rangle = \text{const.} \tag{2.25}$$

A full derivation of the rate of change of the kinetic energy K of the SMBH can be found in Binney & Tremaine (2008), which we state briefly here in terms of the diffusion coefficients $D[X]$ of a quantity X:

$$D[\Delta K] = M_\bullet \sum_{i=1}^{3} \left(v_i D[\Delta v_i] + \frac{1}{2} D[\Delta v_i \Delta v_i] \right) \tag{2.26}$$

$$= 16\pi^2 G^2 M_\bullet m_m \ln \Lambda \left[m_m \int_{v_\bullet}^{\infty} \mathrm{d}v_m v_m f_m(v_m) - M_\bullet \int_{0}^{v_\bullet} \mathrm{d}v_m \frac{v_m^2}{v_\bullet} f_m(v_m) \right]. \tag{2.27}$$

Here the term $\ln \Lambda$ is the Coulomb logarithm, with:

$$\Lambda \approx \frac{b_{\max}}{b_{90}}, \tag{2.28}$$

where $b_{\max}$ is the largest impact parameter (typically of the order of the system size R) and b_{90} is the impact parameter that corresponds to a $90°$ deflection of the SMBH. The term $f_m(v_m)$ describes the position-independent distribution function, which in general describes the probability of a particle being located with position coordinates $\boldsymbol{x} + \mathrm{d}\boldsymbol{x}$ and velocity coordinates $\boldsymbol{v} + \mathrm{d}\boldsymbol{v}$. In Equation 2.26, the first integral describes the growth of kinetic energy of the SMBH due to gravitational encounters, whereas the second integral describes reduction in kinetic energy due to diffusion. As indicated by the integration limits, stars and other particles with $v_m > v_\bullet$ contribute to the increase, and those with $v_m < v_\bullet$ to the decrease, of the SMBH kinetic energy. In this sense, the first integral is analogous to 'heating' the motion of the SMBH, whilst the second integral is analogous to 'cooling' the motion of the SMBH.

As the mass of the SMBH is much greater than that of individual stars, or gas or DM, $D[\Delta v_i] \gg D[\Delta v_i \Delta v_i]/v_\bullet$ as $D[\Delta v_i] \propto M_\bullet$, whereas $D[\Delta v_i \Delta v_i]/v_\bullet \propto m_m$. The

overall effect is for the SMBH to lose energy. From Equation 2.26 and using that $D[\Delta v_i] = v_i/v_\bullet D[\Delta v_{\parallel}]$, we then have:

$$D[\Delta K] \approx M_\bullet \sum_i^3 \frac{v_i^2}{v_\bullet} D[\Delta v_{\parallel}] \tag{2.29}$$

$$\implies D[\Delta \boldsymbol{v}] = \frac{\boldsymbol{v}_\bullet}{M_\bullet v_\bullet^2} D[\Delta K]. \tag{2.30}$$

Assuming the surrounding medium has an isotropic velocity distribution, the SMBH is then decelerated at a rate given by:

$$\frac{\mathrm{d}v_\bullet}{\mathrm{d}t} = D[\Delta \boldsymbol{v}] = -16\pi^2 G^2 M_\bullet m_m \ln \Lambda \left[\int_0^{v_\bullet} \mathrm{d}v_m v_m^2 f_m(v_m) \right] \frac{\boldsymbol{v}_\bullet}{v_\bullet^3}. \tag{2.31}$$

Equation 2.31 is commonly known as Chandarasekhar's formula (Chandrasekhar, 1943), where we see that only material moving with velocity $v_m < v_\bullet$ contributes to the frictional force.

If we further assume that the SMBH velocity is either very small or very large compared to the stellar velocity dispersion σ, two important approximations become relevant:

$$\frac{\mathrm{d}\boldsymbol{v}_\bullet}{\mathrm{d}t} \propto \begin{cases} \boldsymbol{v}_\bullet & \text{if } v_\bullet \ll \sigma \\ \boldsymbol{v}_\bullet/v_\bullet^3 & \text{if } v_\bullet \gg \sigma. \end{cases} \tag{2.32}$$

These two regimes are shown in the left and right panels of Figure 2.9, respectively, assuming the distribution function $f_m(v_m)$ is Maxwellian (Mulder, 1983). In both cases, the deflection of stars by the SMBH enhances the stellar density behind it, increasing the gravitational attraction and hence increasing the dynamical friction experienced. In the case $\tilde{v}_\bullet \equiv v_\bullet/\sigma = 1$ (Figure 2.9, left panel), the drag on the SMBH caused by the overdensity is large, as expected from the first case of Equation 2.32. Conversely, when $\tilde{v}_\bullet = 3$ (Figure 2.9, right panel) the overdensity is diminished, reducing the overall dynamical friction experienced by the SMBH.

In the instance of a major merger, the relative velocities between the galaxies (and also their respective SMBHs) is in general high when the impact parameter is small. As a result, the SMBHs experience dynamical friction in the high-velocity regime. As shown in the right panel of Figure 2.9, the overdensity formed behind the SMBH in this regime is concentrated into narrow contours extending behind the SMBH. As a result, the component of the dynamical friction force parallel to the direction of motion dominates, thus torquing the SMBH to a highly-radial orbit with small b, even if the initial approach had a larger impact parameter b. It is
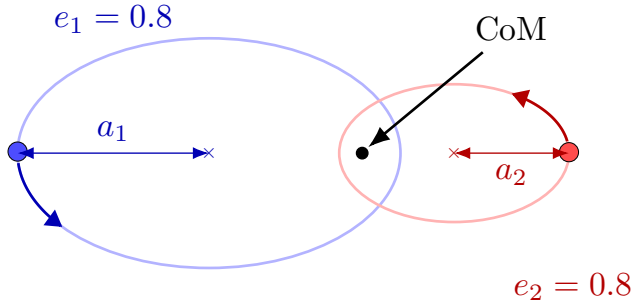
Figure 2.10: Schematic of a generic Keplerian binary system. Two bodies orbit a common centre of mass. The binary semimajor axis is given by $a = a_1 + a_2$, where a_1 and a_2 are the semimajor axes of each body's orbit. Conversely, the binary eccentricity is given by $e = e_1 = e_2$. Note the retrograde orbit of the bodies, and the 180° phase difference between them.

important to note though that in the case of a strongly axisymmetric system (such as spiral LTG), the relative orientation of the merging galaxies can greatly influence the trajectories of the component SMBHs, much more so than in the case of merging ETGs.

The loss of energy through dynamical friction eventually brings the two SMBHs to a proximity where they form a binary system (Milosavljević & Merritt, 2001), whereby the two body orbital energy becomes negative:

$$E = K - U < 0. \tag{2.33}$$

The radial separation at which this happens coincides with the influence radius r_{h} (Equation 2.23); from this point on, dynamical friction is no longer the most efficient means by which to lose orbital energy.

2.3.2 Three-body interactions

Following the dynamical friction phase, the SMBHs have sunk to the high density regions of the galaxy merger remnant, where energy is lost from the SMBH binary system through interactions with individual stars. At this moment, the binding energy of the binary is:

$$E_{\mathrm{bin}} = -\frac{GM_{12}\mu}{2a}, \tag{2.34}$$

where $M_{12} \equiv M_{\bullet,1} + M_{\bullet,2}$, $\mu = M_{\bullet,1}M_{\bullet,2}/M_{12}$ is the reduced mass, and $a = a_1 + a_2$ is the semimajor axis of the SMBH binary orbit, as shown in Figure 2.10. A 'hard'

binary is further defined if the binding energy per unit mass exceeds the local stellar velocity dispersion σ, and has a semimajor axis (e.g. Merritt, 2013):

$$a_{\rm h} = \frac{G\mu}{4\sigma^2}. \tag{2.35}$$

How circular the binary orbit is depends also on the angular momentum J, and is quantified by the eccentricity e, defined:

$$e = \sqrt{1 + \frac{2EJ^2}{G^2 M_{12}^2 \mu^3}}. \tag{2.36}$$

For a bound orbit, $E < 0 \implies 0 \leq e < 1$, with $e = 0$ corresponding to a circular orbit; a purely radial orbit has $e = 1$.

Only a particular subset of stars are able to interact with the SMBH binary to remove energy (Hills & Fullerton, 1980; Hills, 1983; Quinlan, 1996), and hence strengthen the binding energy $E_{\rm bin}$. These stars are said to belong to the 'loss cone' of the binary, and have an angular momentum J that satisfies:

$$J \leq J_{\rm lc} \simeq \sqrt{2GM_{12}\mathcal{K}a}, \tag{2.37}$$

where $\mathcal{K}$ is a dimensionless constant, often taken to be unity (Gualandris et al., 2017). Stars in the loss cone have centrophilic orbits: they are able to pass through the centre of mass (CoM) of the SMBH binary with arbitrarily small impact parameter. Each successive passage of a loss cone star removes some energy from the binary system, whereupon the star is eventually ejected at high velocity: a hypervelocity star. The typical velocity of such stars can be determined as (Sesana et al., 2006):

$$v_{\star,\rm hyp} \approx \frac{\sqrt{2C\eta}}{1+\eta}\left(\frac{GM_{12}}{a}\right)^{1/2} = \sqrt{\frac{8C\sigma^2}{1+\eta}\frac{a_{\rm h}}{a}}, \tag{2.38}$$

where C is a constant of order unity and η is the binary mass ratio[5], $\eta = M_{\bullet,2}/M_{\bullet,1}$ $|\eta \leq 1$. Using the virial theorem $2K + U = 0$ and an isotropic velocity dispersion tensor, we have that $\langle v^2 \rangle = 3\sigma^2$, and for a star to escape from the centre of a system $\langle v_{\star,\rm hyp}^2 \rangle \approx \langle (2v)^2 \rangle = 12\sigma^2 \implies v_{\star,\rm hyp} > 2\sqrt{3}\sigma$ (Spitzer, 1987; Quinlan, 1996). If we further assume an equal-mass binary ($\eta = 1$), a large fraction of hypervelocity stars are produced when $a \simeq a_{\rm h}/3$.

---

[5]If we assume that the SMBH mass is a constant fraction of a galaxy's stellar mass, then $q = \eta$.

The evolution of the SMBH binary orbital parameters a and e for semimajor axes $a < a_\mathrm{h}$ are well-described by two empirical relations (Quinlan, 1996). The SMBH binary hardens at an approximately constant rate:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{a}\right) = H\frac{G\rho}{\sigma},$$

(2.39)

where H is a dimensionless constant dependent on the stellar density distribution (typically in the range 15 to 20, Quinlan, 1996; Sesana et al., 2006; Sesana & Khan, 2015; Mannerkoski et al., 2019) and ρ is the stellar density, typically taken within the influence radius of the SMBH binary. Similarly, the eccentricity evolves approximately as:

$$\frac{\mathrm{d}e}{\mathrm{d}\ln(1/a)} = K,$$

(2.40)

where K is also a (small) dimensionless constant.

The shrinking of the SMBH binary orbit through stellar scattering relies upon there being a sufficiently-large reservoir of stars in the loss cone with which it may interact. Early simulation studies found that in a spherically-symmetric gravitational potential, there was insufficient replenishing of the loss cone reservoir to allow the SMBH binary to harden beyond approximately parsec-scale separations (Milosavljević & Merritt, 2001). This came to be known as the 'Final Parsec Problem'. Later studies however came to show that in triaxial potentials, such as those found in the aftermath of a galaxy merger, the loss cone population is refilled at a sufficient rate by torquing of centrophobic orbits to centrophilic orbits to prevent the SMBH binary separation from stalling (Berczik et al., 2006; Merritt & Vasiliev, 2011; Gualandris et al., 2017). Other processes that refill the loss cone, such as enhanced star formation, can also allow the SMBH binary to harden to separations less than a parsec (Liao et al., 2024a,b).

The removal of central stars through the three-body scattering process is a major contributor to the luminosity deficit in the central regions of ETGs (Begelman et al., 1980; Hills, 1983; Rantala et al., 2017; Rantala et al., 2018; Nasim et al., 2021; Khonji et al., 2024; Rawlings et al., 2025b, refer to Subsection 2.1.4), and is termed SMBH binary scouring. The amount of stellar mass removed due to binary scouring is found to approximately scale as:

$$M_\mathrm{def} \simeq 0.5 N_\mathrm{mergers} M_\bullet$$

(2.41)

(Merritt, 2006), hence repeated mergers tend to enlarge pre-existing stellar cores. Only a weak dependence on the SMBH binary mass ratio η and on the initial stellar density profile has been found.
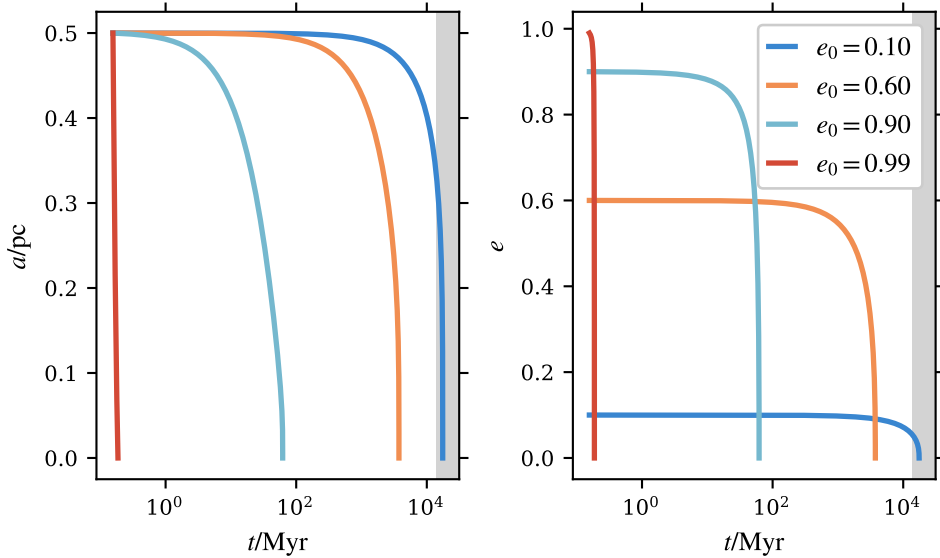
Figure 2.11: Evolution of the semimajor axis (left panel) and eccentricity (right panel) of an equal-mass SMBH binary due to gravitational wave emission. The initial configuration is of two SMBHs each of mass $10^9 \, \mathrm{M_\odot}$ separated by $0.5 \, \mathrm{pc}$, with different initial eccentricities. The time axis is logarithmic to highlight the difference the choice of initial eccentricity has on the merger timescale. For the $e_0 = 0.10$ case, the merger timescale is greater than the age of the Universe (indicated by the shaded grey region).

2.3.3 Gravitational wave emission

Once the SMBH binary has a semimajor axis of the order of milliparsecs, the loss cone has shrunk to such a small region in angular momentum space (Equation 2.37) that almost no stars belong to it. At this point, the emission of gravitational waves (GW) becomes the dominant, and final, means by which the SMBH binary can lose its remaining orbital energy and angular momentum (Peters & Mathews, 1963; Peters, 1964).

GW emission, in this instance, is a consequence of general relativistic effects stemming from the influence of the SMBHs on the surrounding space-time continuum. To gain insight into the role of GW emission, we can however consider the post-Newtonian (PN) formalism, which approximates general relativity in the weak-field limit. Whilst SMBHs are definitely *not* described by the weak-field limit, their motion through a weak gravitational field generated by distant particles is well captured by the PN approximation, provided we restrict our investigation to sufficiently

large distances from the SMBHs (Poisson & Will, 2014).

The PN formalism adds corrective terms in the form of an expansion parameter $\epsilon_{\mathrm{PN}} \sim (v/c)^2$, where the expansion is done about a small perturbation $h^{\alpha\beta}$ to the Minkowsi flat metric in general relativity. In this way, the acceleration of a body (here the SMBH) has its equation of motion modified such that:

$$\boldsymbol{a} = \boldsymbol{a}_{\mathrm{N}} + \sum_{i=1}^{3} \boldsymbol{a}_i + \sum_{i=1.5}^{3.5} \boldsymbol{a}_i + \mathcal{O}(4), \tag{2.42}$$

where each correction $\boldsymbol{a}_i$ is proportional to the expansion parameter, $\boldsymbol{a}_i \propto \epsilon_{\mathrm{PN}}^i \equiv i\mathrm{PN}$. In Equation 2.42, the first term is the Newtonian acceleration. Integer corrections (the first summation) are conserved quantities, whereas the half-integer corrections (the second summation) are non-conserved, or dissipative, corrections. GW emission is described, at lowest order, by the 2.5PN term.

The time-averaged evolution of the binary parameters a and e are relatively compact when only considering the 2.5PN contribution to the equations of motion, and are given by (Peters & Mathews, 1963; Peters, 1964):

$$\left\langle \frac{\mathrm{d}a}{\mathrm{d}t} \right\rangle = -\frac{64}{5} \frac{G^3 \mu M_{12}^2}{c^5 a^3 (1-e^2)^{7/2}} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4\right), \tag{2.43}$$

$$\left\langle \frac{\mathrm{d}e}{\mathrm{d}t} \right\rangle = -\frac{304}{15} \frac{G^3 \mu M_{12}^2}{c^5 a^4 (1-e^2)^{5/2}} e \left(1 + \frac{121}{304} e^2\right). \tag{2.44}$$

In Figure 2.11, we show the evolution of an equal-mass ($M_{\bullet,1} = M_{\bullet,2} = 10^9\,\mathrm{M}_\odot$) SMBH binary with an initial semimajor axis of $a_0 = 0.5\,\mathrm{pc}$, for different initial eccentricities e_0. Very clearly we see that the time to merger ($a \simeq 0\,\mathrm{pc}$), as a result of energy loss through GW emission, is highly dependent on the eccentricity of the binary. For highly-eccentric binaries ($e_0 = 0.99$), the time to merger is less than $1\,\mathrm{Myr}$. Conversely, for nearly-circular binaries, the merger timescale is well in excess of $10^4\,\mathrm{Myr}$ – longer than the age of the Universe. The separation at which GW emission dominates the binary orbital evolution can be estimated by equating Equation 2.39 and Equation 2.43, to obtain:

$$a_{\mathrm{GW}} = \left[\frac{64}{5} \frac{G^2 \mu M_{12}^2}{c^5} \frac{\sigma}{H\rho} (1-e^2)^{-7/2} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4\right)\right]^{1/5}. \tag{2.45}$$

A common feature irrespective of the initial eccentricity at $a = a_{\mathrm{GW}}$ is the circularisation of the SMBH binary, whereby the eccentricity tends to $e \to 0$ at merger, shown in the right panel of Figure 2.11. From this, it is expected that close to merger,
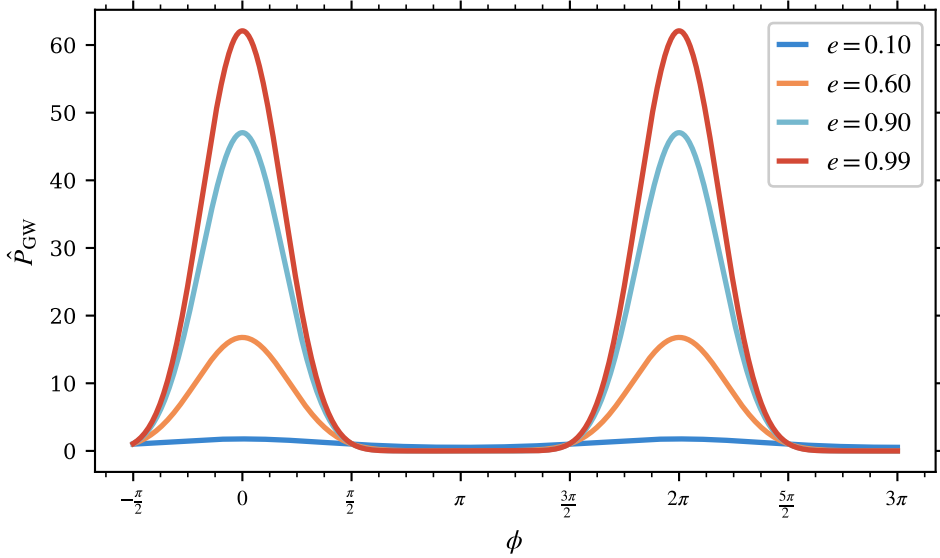
Figure 2.12: Normalised instantaneous power (Equation 2.46) radiated due to GW emission, for different binary eccentricities consistent with Figure 2.11. The radiated power is normalised such that the dependence on mass and semimajor axis is removed. A higher eccentricity results in larger accelerations of the SMBHs, increasing the power radiated as gravitational waves.

almost all SMBH binaries should be on nearly-circular orbits, with deviations from circularity possible due to environmental effects (e.g. Armitage & Natarajan, 2005; Matsubayashi et al., 2007; Merritt & Vasiliev, 2011).

The decrease in merger timescale for a higher eccentricity suggests that more energy is emitted for highly-eccentric orbits than low-eccentricity orbits. This dependence can be investigated by considering the instantaneous power P_{GW} radiated by the SMBH binary at the 2.5PN level as a function of its angular coordinate ϕ, where $\phi = 0$ coincides with the SMBH binary pericentre:

$$P_{GW} = \frac{32G^4}{5c^5} \frac{\mu^2 M_{12}^3}{a^5(1-e^2)^5}(1 + e\cos\phi)^4 \left[1 + 2e\cos\phi + \frac{e^2}{12}(1 + 11\cos^2\phi)\right], \quad (2.46)$$

as given in Poisson & Will (2014). From Equation 2.46 and shown in Figure 2.12, not only is the instantaneous power maximal when $\cos\phi = 1$, coinciding by definition with the binary pericentre, but there is also a strong dependence on the eccentricity e. Intuitively we can understand this (to first order) by considering the instantaneous
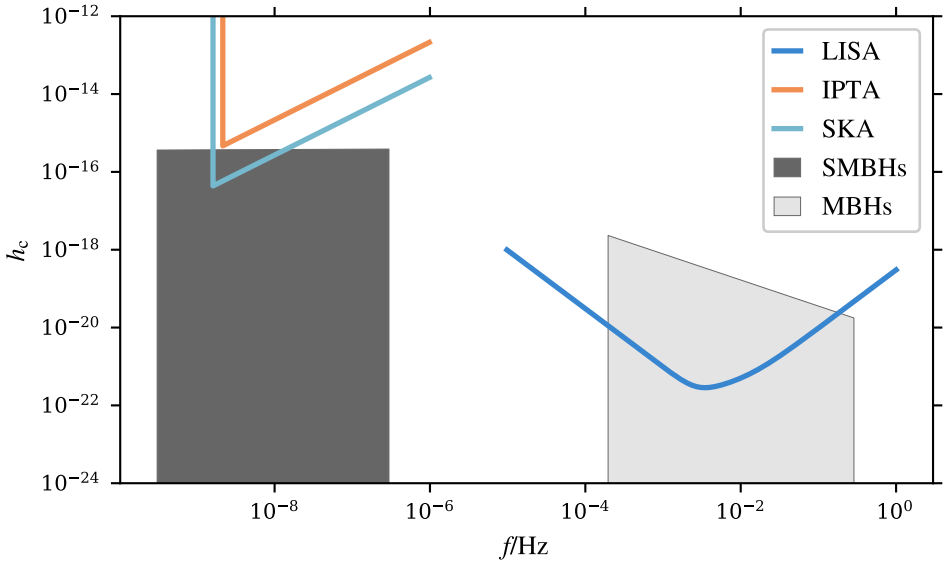
Figure 2.13: Characteristic strain h_c (Equation 2.53) for SMBHs ($M_\bullet \gtrsim 10^8 \, \mathrm{M_\odot}$) and massive black holes (MBHs, $10^5 \, \mathrm{M_\odot} \lesssim M_\bullet \lesssim 10^7 \, \mathrm{M_\odot}$). The sensitivity curves for three GW instruments: LISA, IPTA, and SKA, are shown. The latter two are pulsar timing array instruments. Critically, LISA will be sensitive to massive black holes, but PTAs will be better suited for detecting merging SMBHs.

acceleration in a Keplerian potential:

$$\boldsymbol{\nabla}\Phi \propto r(\phi)^{-2}\hat{r} \tag{2.47}$$

$$= \left(\frac{1 + e\cos\phi}{a(1 - e^2)}\right)^2 \hat{r} \tag{2.48}$$

where we note the $1-e^2$ term in the denominator. As a result, close to pericentre the SMBHs experience higher accelerations for higher eccentricities, leading to a greater loss of energy through GW emission.

Gravitational waves are detectable through their strain, or amplitude, as they pass through space-time – this quantity we denote as h. GW detectors typically aim to detect this strain as a frequency shift in an interferometer (e.g. LIGO, LISA), or a systematic delay in nearby pulsar signals (for pulsar timing arrays). For any GW source, including an inspiralling source such as an SMBH binary, we may measure an integrated signal over some time period, or equivalently, at some frequency f. Let

us define a characteristic strain h_c that accounts for this integration interval as:

$$[h_c(f)]^2 = 4f^2 \left|\tilde{h}(f)\right|^2 \tag{2.49}$$

following Finn & Thorne (2000), where $\tilde{h}(f)$ is the Fourier transform of $h(f)$. The number of cycles an inspiralling source completes at a frequency f is then:

$$N_{\mathrm{cycles}} = \frac{f}{2\pi}\frac{\mathrm{d}\phi}{\mathrm{d}f} = \frac{f^2}{\dot{f}}, \tag{2.50}$$

where ϕ is the orbital phase as before. If we consider a monochromatic (i.e. single frequency) GW source, the strain is then a modulation of constant amplitude h_0:

$$h(t) = \sqrt{2}h_0 \cos\left[\phi(t)\right], \tag{2.51}$$

and the Fourier transform of Equation 2.51 becomes (e.g. Moore et al., 2015):

$$\tilde{h}(f) \simeq \frac{h_0}{\sqrt{2\dot{f}}}. \tag{2.52}$$

Combining Equation 2.49 and Equation 2.52, we have the characteristic strain for an inspiralling SMBH binary as a function of the frequency:

$$h_c(f) = \sqrt{\frac{2f^2}{\dot{f}}}h_0. \tag{2.53}$$

The GW strain emitted by an inspiralling SMBHs is dependent on the masses of the binary components through Equation 2.46, and manifest in particular frequency bands. For SMBHs with masses $M_\bullet \gtrsim 10^8\,\mathrm{M_\odot}$, these frequencies are in the range $10^{-10}\,\mathrm{Hz}$ to $10^{-7}\,\mathrm{Hz}$, whereas for massive black holes (MBHs) with $10^5\,\mathrm{M_\odot} \lesssim M_\bullet \lesssim 10^7\,\mathrm{M_\odot}$, the frequencies are in the range $10^{-4}\,\mathrm{Hz}$ to $10^{-1}\,\mathrm{Hz}$ (Figure 2.13). Different instruments will be sensitive to black holes of different masses, with pulsar timing arrays (PTAs) being best-suited for SMBH binary detection.

Thus far we have considered symmetric SMBH binary systems: the SMBHs have equal-masses, and whilst not yet discussed, we have assumed the spins of both black holes to be zero (i.e. a Schwarzschild black hole). Typically however, the binary system is not perfectly symmetric: differences in masses or spin vectors is not only possible, but probable. As a result, linear momentum may be radiated from the system by GW emission (at the 3.5PN level), in addition to the afore-mentioned energy and angular momentum. This has important consequences following the coalescence of the SMBH binary, as discussed in Subsection 2.3.4.
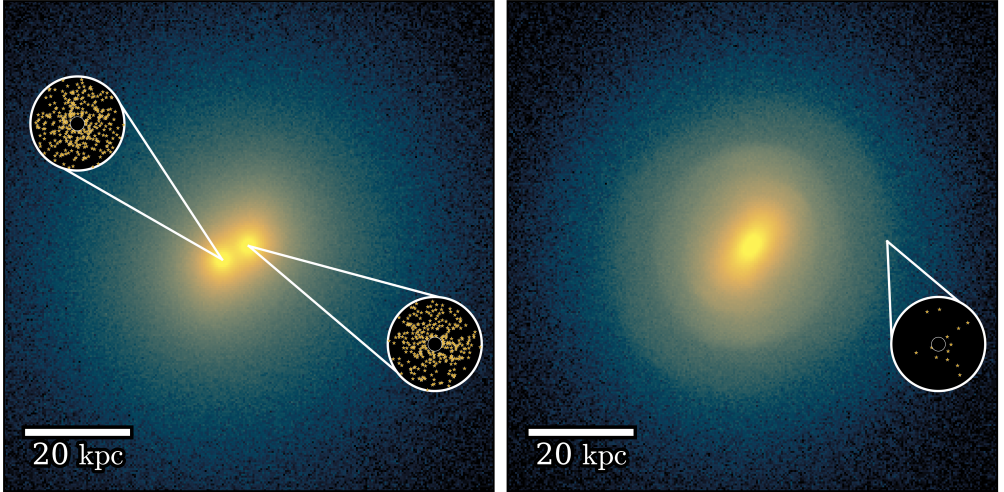
Figure 2.14: Simulated galaxy merger of two $M_\star \sim 10^{11}\,\mathrm{M_\odot}$ ETGs, each with an SMBH of $M_\bullet \sim 10^9\,\mathrm{M_\odot}$. *Left panel*: Prior to SMBH binary coalescence, the SMBHs are embedded within their respective stellar nuclei. *Right panel*: Following coalescence, the remnant SMBH experiences a recoil kick that takes it to the outskirts of the galaxy. In this particular instance, the recoil velocity was $v_{\mathrm{kick}} = 1020\,\mathrm{km\,s^{-1}}$.

2.3.4   Gravitational wave recoil

Just prior to coalescence, the radiation of linear momentum from an asymmetric SMBH binary system moves the binary centre of mass (e.g. González et al., 2007b). The linear momentum transferred from the binary increases the closer the SMBHs are to coalescence, with the most momentum being lost in the final few orbits. To conserve linear momentum, the newly-coalesced SMBH is imparted a velocity vector opposite to the direction of preferred GW emission (Bekenstein, 1973), resulting in the SMBH recoiling away from the place of coalescence. This sudden increase in velocity is termed a *recoil kick* (or recoil velocity or kick velocity), and is dependent on both the difference in masses between the SMBHs prior to merger, and also on their relative spin magnitudes and orientations (Campanelli et al., 2007; González et al., 2007b,a). Herein we will discuss SMBH spin in terms of the dimensionless parameter $\alpha_\bullet$, defined:

$$\alpha_\bullet = \frac{cJ_\bullet}{GM_\bullet^2} \quad |0 \leq \alpha_\bullet \lesssim 1, \tag{2.54}$$
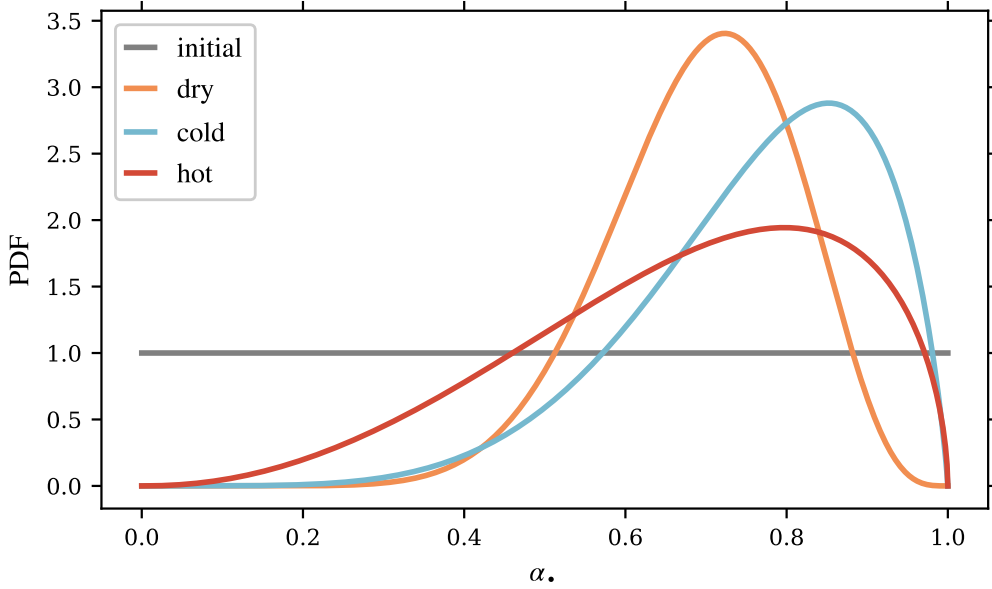
Figure 2.15: Distribution of SMBH spin magnitudes for different scenarios. The 'dry' model (orange line, Lousto et al., 2010) does not have spins affected by gas torquing, whereas 'cold' and 'hot' do (blue and red lines, respectively, Lousto et al., 2012). Monte Carlo sampling multiple merger generations transforms the initial condition to the various final states, under the respective model assumptions.

where $J_\bullet$ is the angular momentum magnitude of the SMBH.

The recoil kick can displace the SMBH a large distance from the nucleus of the galaxy merger remnant (González et al., 2007a; Tichy & Marronetti, 2007), as shown in the simulated merger in Figure 2.14. In this instance, prior to merger (left panel) the SMBHs are well-embedded within the stellar nuclei of their respective progenitor galaxies. Following coalescence (right panel), the remnant SMBH is displaced to the outskirts of the galaxy.

Intuitively, the distance a recoiling SMBH may travel from the galactic nucleus depends primarily on the recoil kick experienced. Modelling of the recoil velocity is however difficult, as the post-Newtonian approximations used previously break down in the final few orbits[6] (where the binary separation is $\lesssim 10R_\mathrm{S}$) where the greatest emission of linear momentum occurs, requiring the full field equations of general

---

[6]In this regime, the separation is given in terms of the Schwarzschild radius R_S, defined in Equation 3.13.

relativity to be solved.

Early work by Favata et al. (2004) and Merritt et al. (2004) suggested that the typical range of v_{kick} was $100\,\text{km}\,\text{s}^{-1}$ to $200\,\text{km}\,\text{s}^{-1}$, and not exceeding $500\,\text{km}\,\text{s}^{-1}$. These results were at first corroborated by numerical relativity studies of non-spinning black holes by González et al. (2007b), who showed that in the mass range of $\eta = M_{\bullet,2}/M_{\bullet,1} \in [0.25, 1]$ v_{kick} had a maximum value of $\sim 175\,\text{km}\,\text{s}^{-1}$, occurring when $\eta \simeq 0.36$.

However, when extending the numerical relativity studies to include spinning black holes, the range of recoil kicks increased, with maximal kicks in excess of $2500\,\text{km}\,\text{s}^{-1}$ (González et al., 2007a) and even up to $4000\,\text{km}\,\text{s}^{-1}$ (Campanelli et al., 2007; Tichy & Marronetti, 2007). To achieve such large recoil, special conditions for the binary SMBH prior to coalescence are required. In this instance, the SMBHs must be of equal-mass, maximally spinning ($\alpha_{\bullet} \simeq 1$) and anti-aligned, with the spin vectors lying in the orbital plane (Campanelli et al., 2007; Zlochower & Lousto, 2015). From these investigations, it is clear that the effect of SMBH spin is the dominant contributor to GW-induced recoil.

If we consider the environment within which the SMBH binary is coalescing – the centre of a massive ETG merger remnant – kicks of the order $v_{\text{kick}} \sim 4000\,\text{km}\,\text{s}^{-1}$ far exceed the typical escape velocities (see Section 3.2) of these systems, which are of the order $2000\,\text{km}\,\text{s}^{-1}$. As a result, whether or not the recoil velocity removes the SMBH from the galactic core has implications for the properties of the ETG itself (e.g. Nasim et al., 2021; Khonji et al., 2024; Rawlings et al., 2025b). A typical finding is that GW-induced recoil can enlarge the stellar core formed through SMBH binary scouring, even by up to a factor of ~ 3.

To understand the likelihood of SMBHs recoiling with high velocities, and any additional scouring of the stellar core, we require the distribution of SMBH spins and masses prior to coalescence. If we restrict ourselves to equal-mass SMBHs – which should produce the largest stellar cores from SMBH binary scouring – we may consider the distribution of spins in isolation.

The evolution of spin magnitude and direction prior to coalescence is largely dominated by the presence (or lack) of gas in the form of a circumnuclear disc that feeds the SMBH. Massive ETGs with little cold gas in the nucleus are often termed 'dry' mergers, and simulation studies evolve SMBH spins purely under the dynamical contribution from post-Newtonian effects. By assuming an initially-uniform distribution in spin magnitudes, Monte Carlo sampling of 10^7 systems, with multiple generations of mergers, resulted in a spin magnitude distribution (Lousto et al., 2010):

$$P(\alpha_{\bullet})_{\text{dry}} \simeq \text{Beta}(\alpha_{\bullet}; 10.5868, 4.66884), \tag{2.55}$$

where $\mathrm{Beta}(x; a, b)$ is the beta distribution, and is defined:

$$\mathrm{Beta}(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \qquad (2.56)$$

where $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t$ is the gamma function. The result of Equation 2.55 is shown in Figure 2.15.

A similar study to Lousto et al. (2010) was performed by Lousto et al. (2012), but incorporating the effects of gas accretion on the SMBH spins prior to coalescence – these are termed 'wet' mergers. The authors considered separately the accretion of two species of gas: 'cold', where the gas is assumed to have a polytropic equation of state with index $\gamma = 7/5$, and 'hot', where the equation of state index is $\gamma = 5/3$. Hydrodynamical simulations (Dotti et al., 2010) were used to follow the gas accretion of the SMBHs, and Monte Carlo sampling used to obtain a distribution of spin magnitudes for the gas species, shown also in Figure 2.15. The results also follow beta distributions:

$$P(\alpha_\bullet)_{\mathrm{cold}} \simeq \mathrm{Beta}(\alpha_\bullet; 5.935, 1.856), \qquad (2.57)$$
$$P(\alpha_\bullet)_{\mathrm{hot}} \simeq \mathrm{Beta}(\alpha_\bullet; 3.212, 1.563). \qquad (2.58)$$

The primary result of including gas accretion is to spin-up the SMBHs prior to merger compared to the dry-merger case: the peak in $\alpha_\bullet$ is > 0.8 for wet mergers, as opposed to $\alpha_\bullet \simeq 0.7$ for dry mergers.

In addition to the spin magnitude affecting the SMBH recoil velocity, so too does the spin direction relative to the orbital plane, and the relative separation between the SMBH binary spin vectors. Monte Carlo sampling by Zlochower & Lousto (2015) showed that for dry mergers with random spin alignments (uniform in $\cos(\theta)$, where θ is the angle between the SMBH binary spin vectors), 10.93% of SMBHs recoiled with $v_{\mathrm{kick}} > 1000\,\mathrm{km\,s^{-1}}$, and only 2.85% with $v_{\mathrm{kick}} > 2000\,\mathrm{km\,s^{-1}}$. This study indicates that the majority of SMBH recoil, whilst non-zero, does not lead to an SMBH escaping from the host galaxy.

3 Solving the N-body problem

3.1 Collisional or collisionless dynamics?

3.1.1 Timescales

The dynamical evolution of interacting galaxies is a highly non-linear process, and for almost all modern investigations, we must turn to simulations of galaxy evolution to answer our scientific questions.

Consider a typical massive elliptical galaxy with a stellar mass of $10^{11}\,\mathrm{M}_\odot$, and as a first-order approximation let us assume that this translates to 10^{11} (solar-mass) stars in the galaxy. Current computing architecture is not able to support such large numbers of particle data, as well as computing the time evolution of that data over timescales relevant for the investigation, typically of the order gigayears for ETGs. Consequently, approximations are required to be made. As with any approximation however, errors will accumulate, rendering the approximation increasingly-less representative of the original system we wished to investigate as time progresses. Hence, our first requirement is to understand the relevant timescales for our system.

Let us consider the stars in a galaxy as generating an infinitely smooth potential, not dissimilar to a fluid, which the stars trace in a Lagrangian sense. The potential experienced at a given volume element is the sum of the individual point-mass potentials from all the stars in the galaxy:

$$\Phi(\boldsymbol{r}_i) = -G \sum_{j,j\neq i}^{N} \frac{m_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}. \tag{3.1}$$

If we wish to follow the evolution of this fluid, we can ask what is the minimum number of tracer particles required that still captures the motion of the original particles under the overall influence of the self-generated gravitational potential for some time frame. Alternatively, we may ask *given* some number of particles that generate the gravitational potential, for how long is our representation faithful to the system being modelled. This standard exercise then asks what is the time for the perpendicular
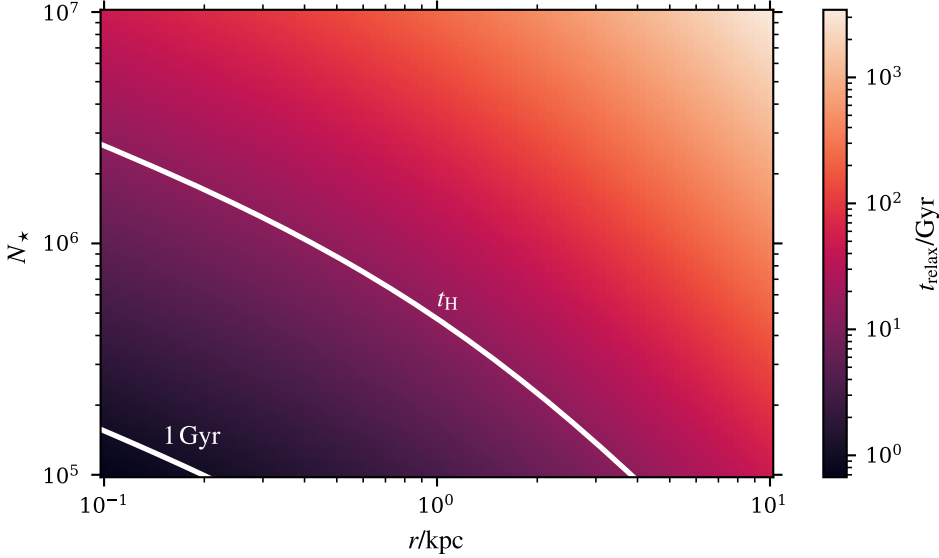
Figure 3.1: Relaxation time in Gyr for a $M_\star = 10^{11}\,\mathrm{M_\odot}$ galaxy at different radii for numbers of particles that trace the overall potential, from Equation 3.2. The mass interior to r, $M(< r)$, is assumed to follow a Hernquist profile (Section 3.2) with $r_\mathrm{s} = 1.5\,\mathrm{kpc}$. The contours for $t_\mathrm{relax} = 1\,\mathrm{Gyr}$ and $t_\mathrm{relax} = t_\mathrm{H}$, where t_H is the Hubble time, are shown. The relaxation time increases to larger radii, and for larger particle counts.

velocity component of a given particle orbiting in the gravitational potential of the galaxy to change by order of itself, if it undergoes a series of interactions with other particles that tend the system to a Markovian state. That is, given the present state of the system, the initial and final states are independent.

A particle orbiting in a collection of other particles will experience a series of velocity perturbations $\sum \delta v^2$ that acts to diffuse the motion of the particle: this effect is called two-body relaxation (e.g. Binney & Tremaine, 2008). By considering all impact parameters that the diffusion occurs over for an orbital period (or crossing time) $t_\mathrm{cross} \simeq R/v$ (where R is the radius of the galaxy), we find the number of crossings, and thus time, required to randomise the particle motion is:

$$n_\mathrm{relax} \simeq \frac{N}{10 \ln N} \implies t_\mathrm{relax} \simeq \frac{N}{10 \ln N} \frac{R}{\sqrt{GM(< R)/R}}, \tag{3.2}$$

where we have assumed circular velocity $v = \sqrt{GM(< R)/R}$ for the particle. The dependence of t_relax on radial distance and particle number is shown in Figure 3.1,

where we assume the mass profile follows a Hernquist sphere with scale radius $r_s = 1.5\,\mathrm{kpc}$ and total mass $M_\star = 10^{11}\,\mathrm{M_\odot}$ (see Section 3.2 for further discussion on gravitational potentials). Whilst the figure is illustrative, owing to the simplifying assumption in Equation 3.2 that particles are uniformly sampled in the space, smaller radii require more particles to reduce the effects of two-body relaxation as compared to larger radii. For simulations of galaxy mergers, Figure 3.1 suggests that having particle counts in excess of 2×10^6 are required to represent the galaxy for a Hubble time.

Another consideration is the collision between stars in the galaxy – fortunately, stellar collisions within a galaxy are rare. If we approximate all stars to be similar to the Sun, and make the simplifying assumption of no mutual gravitational attraction between stars, then the mean free path of a star between collisions is:

$$\lambda_{\mathrm{fp}} = \frac{1}{4n\pi R_\odot^2},\tag{3.3}$$

where n is the number density of stars uniformly distributed in a disc of radius R with thickness h:

$$n = \frac{N}{\pi R^2 h}.\tag{3.4}$$

For the Milky Way, $N \approx 10^{11}$, $R \approx 10\,\mathrm{kpc}$, and $h \approx 0.5\,\mathrm{kpc}$, resulting in $n \simeq 0.6\,\mathrm{pc^{-3}}$. Using $R_\odot \simeq 6.96 \times 10^8\,\mathrm{m}$, we have $\lambda_{\mathrm{fp}} = 2 \times 10^{14}\,\mathrm{pc}$. If we then consider the time interval between stellar collisions, $\tau \simeq \lambda/v$, and using a typical stellar velocity of $v = 50\,\mathrm{km\,s^{-1}}$ for stars in the solar neighbourhood, we have that $\tau \approx 5 \times 10^{18}\,\mathrm{yr} \approx 3 \times 10^8\,t_{\mathrm{H}}$.

Consequently, the chance of two stars colliding when they are both moving predominantly under the overall galaxy potential is exceedingly unlikely, and the timescale for this effect we can safely neglect. Hence, we require for our modelling that:

$$t_{\mathrm{simulation}} \leq t_{\mathrm{H}} \ll t_{\mathrm{relax}} \ll t_{\mathrm{collision}}.\tag{3.5}$$

Such systems that satisfy Equation 3.5 are said to be *collisionless*, whereas those with $t_{\mathrm{simulation}} > t_{\mathrm{collision}}$ are said to be *collisional*. An example of a collisional system is a dense stellar cluster.

3.1.2 Gravitational Brownian motion

Related to two-body relaxation discussed earlier, a massive particle, say an SMBH, will undergo a random walk motion as a result of encounters with field stars with masses $m \ll M_\bullet$ – a form of gravitational Brownian motion. The random motion
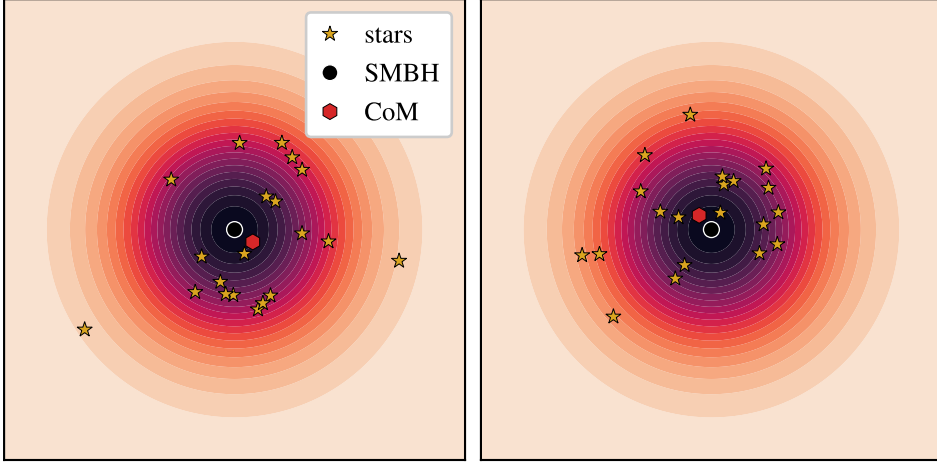
Figure 3.2: Demonstration of gravitational Brownian motion introduced by discrete particle sampling. Both panels show stars drawn from a standard multivariate normal distribution ($\boldsymbol{\mu} = 0$, cov $= 1$) centred about an SMBH at the origin, with contours indicating areas of equal probability density. Different sampling realisations of the underlying distribution shift the stellar centre of mass (indicated by the red hexagon), leading to the SMBH experiencing different accelerations between realisations.

is largely dominated by distant encounters (Merritt, 2001; Bortolas et al., 2016) and leads to energy equipartition (Equation 2.25). This random walk motion leads to the SMBH wandering some small distance from the centre of the galaxy. If the precise positions of the field stars are sampled from some underlying distribution (see Section 3.2), differences in positions between sampling realisations can potentially lead to a different sequence of interactions of the SMBH between the two otherwise-identical setups. This is demonstrated in Figure 3.2.

Brownian motion has important consequences for SMBH binaries. Following Binney & Tremaine (2008), if we consider two SMBHs of equal-mass $M_\bullet$ that encounter one another in the context of a galaxy merger, the resulting velocity impulse is related to the impact parameter b by:

$$\Delta v \approx \frac{GM_\bullet}{bv}. \tag{3.6}$$

If we were to consider an almost identical setup but with the field stars of the galaxies – although sampled from the same underlying distribution – perturbed slightly, then

we have a perturbed impact parameter δb between the two SMBHs in the galaxy merger, leading to:

$$\delta(\Delta v) \approx \frac{GM_\bullet \delta b}{b^2 v}.$$ (3.7)

If the next encounter between the SMBHs (say on the second pericentre passage) occurs after a time τ with impact parameter b', we have:

$$\frac{\delta b'}{\delta b} \approx 1 + \frac{GM_\bullet \tau}{b^2 v} \simeq e^{\tau/t_{\rm L}},$$ (3.8)

where $t_{\rm L}$ is the Liapunov timescale:

$$t_{\rm L} \approx \frac{\tau}{\ln\left[1 + \left(\frac{\tau}{t_{\rm cross}}\right)^2\right]}.$$ (3.9)

The Liapunov timescale describes the growth of perturbations δb to an orbit, provided that δb is much smaller than the typical interparticle separation (Valluri & Merritt, 2000; Hut & Heggie, 2002). As a result, when considering discrete particles drawn from some underlying distribution, the locations of individual particles hold no significance. The underlying statistical properties of the distribution that the particles collectively trace *does* however hold significance, as we explore next.

3.2 Potentials of spherical systems

The density distribution $\rho(\boldsymbol{x})$ of particles in a galaxy is a consequence of the gravitational potential $\Phi(\boldsymbol{x})$ acting on those particles. This is formally captured by the Poisson equation:

$$\nabla^2 \Phi(\boldsymbol{x}) = 4\pi G \rho(\boldsymbol{x}).$$ (3.10)

An important quantity that can be determined from the potential of a system is the escape velocity, defined:

$$v_{\rm esc} = \sqrt{2|\Phi(\boldsymbol{x})|},$$ (3.11)

and describes the required kinetic energy a particle must have to exceed the gravitational potential energy of the system.

If we consider ETGs, which to a good approximation may be considered as being spherically symmetric (at least prior to an interaction with another galaxy), the potential of the system depends only on the radial coordinate r, hence $\Phi(\boldsymbol{x}) \to \Phi(r)$. This leads to the family of spherical potentials, of which the ones relevant for this thesis are explored.
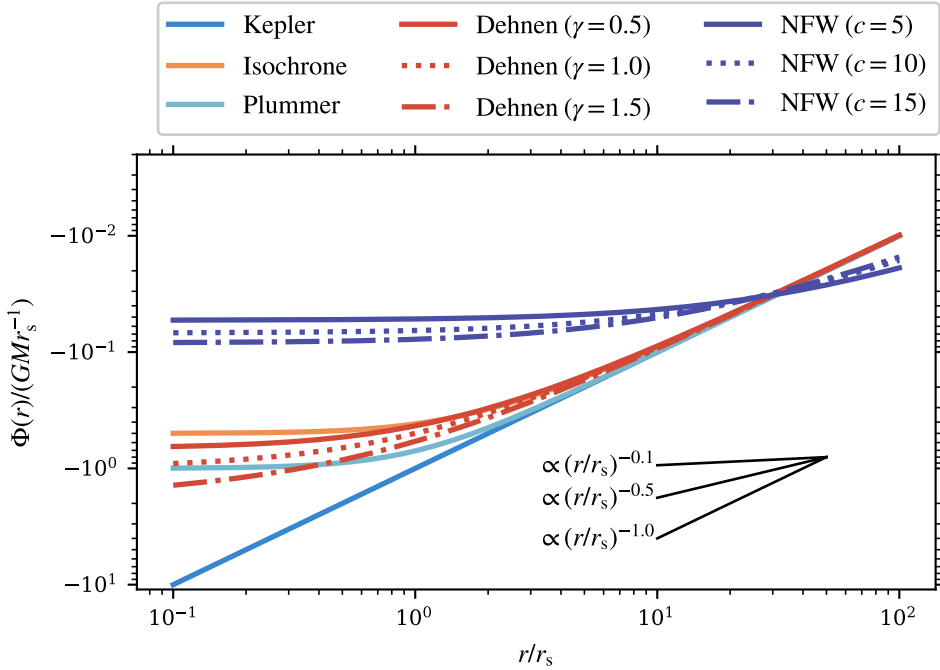
Figure 3.3: Gravitational potentials $\Phi(r)$ generated by different spherical systems. The radial coordinate is scaled by $r_{\mathrm{s}} = 1\,\mathrm{kpc}$, and the potential by the Keplerian potential at r_{s}. Note that apart from the NFW profile, the other potentials converge to the Kepler potential for $r \gg r_{\mathrm{s}}$. The Dehnen profile with $\gamma = 1$ is commonly referred to as a Hernquist profile.

Arguably the simplest of potentials is that generated by a point mass: an example of which is an SMBH. This potential, termed the Kepler potential, is calculated as:

$$\Phi(r) = -\frac{GM}{r}, \tag{3.12}$$

where M is the mass. The potential is shown as the blue line in Figure 3.3, and tends to negative infinity as $r \to 0$. If we consider a non-rotating SMBH, we recover

the Schwarzschild radius[1] of the SMBH when $v_{\text{esc}} = c$:

$$R_{\text{S}} = \frac{2GM_{\bullet}}{c^2}. \tag{3.13}$$

Related to the Kepler potential is the Plummer potential, which modifies the Kepler potential to ensure that the potential is finite at all radii. This is achieved by adding a term to the denominator:

$$\Phi(r) = -\frac{GM}{\sqrt{r^2 + r_{\text{s}}^2}}, \tag{3.14}$$

where r_{s} is some characteristic scale radius. As shown in Figure 3.3, the deviation between the Plummer potential (orange) and the Kepler potential is omnipresent, but diminishes as the radius increases.

Similarly related to the Plummer potential is the isochrone potential:

$$\Phi(r) = -\frac{GM}{r_{\text{s}} + \sqrt{r^2 + r_{\text{s}}^2}}. \tag{3.15}$$

For a particle orbiting in the isochrone potential, its radial period is independent of its angular momentum, and thus depends only on the energy of the particle. The constant radial period – a unique property of this potential – gives the isochrone potential its name.

Compared to the afore-mentioned potentials, the stellar component of ETGs are often well-modelled by a two-part power law. A commonly used class of such models is the Dehnen potential family (Dehnen, 1993), which is parametrised by a shape parameter γ in addition to some characteristic scale radius r_{s}:

$$\Phi(r) = -\frac{GM}{r_{\text{s}}} \times \begin{cases} \frac{1}{2-\gamma}\left[1 - \left(\frac{r}{r+r_{\text{s}}}\right)^{2-\gamma}\right], & \gamma \neq 2 \\ \ln\left(\frac{r+r_{\text{s}}}{r}\right), & \gamma = 2. \end{cases} \tag{3.16}$$

Dehnen profiles with different γ are shown in Figure 3.3, with an increasing γ resulting in a decreasing central potential. A special case of the Dehnen profile is the Hernquist profile, which corresponds to $\gamma = 1$ (Hernquist, 1990).

---

[1] The definition of the Schwarzschild radius arises from general relativity as a constant in the metric that uniquely describes a vacuum spacetime outside a spherically-symmetric, non-rotating mass when solving Einstein's field equations. The constant (and hence the relation in Equation 3.13) is found in terms of physical parameters by considering the weak-field limit of general relativity, where as $r \to \infty$ the potential outside of the sphere approaches the Kepler potential.

Compared to the stellar component of ETGs, the potential of the DM component is often modelled using a Navarro-Frenk-White (NFW, Navarro et al., 1997) profile, which also has the form of a two-part power law:

$$\Phi(r) = -\frac{GM_{\mathrm{vir}}}{r} \frac{\ln\left(1 + \frac{r}{r_{\mathrm{s}}}\right)}{\ln(1 + c) - \frac{c}{1+c}}, \tag{3.17}$$

where M_{vir} is the DM halo virial mass and c is a concentration parameter defined as $c = R_{\mathrm{vir}}/r_{\mathrm{s}}$. Typically, $c \in [4, 16]$ (e.g. Ludlow et al., 2014), with a higher value of c resulting in a deeper potential compared to a lower value (shown in Figure 3.3).

3.3 The distribution function

Previously, we introduced the concept that so long as the simulation time is less than the relaxation time, we may approximate the collection of stars in a galaxy as generating a smooth gravitational potential. More formally, we can say that the stars are distributed in six-dimensional phase space (i.e. in position and velocity space) at a given moment t according to some distribution function $f(\boldsymbol{x}, \boldsymbol{v})$. Hence, the probability of finding a given star in a region Ω of phase space becomes:

$$P = \int_{\Omega} f(\boldsymbol{x}, \boldsymbol{v}) \mathrm{d}\boldsymbol{x}^3 \mathrm{d}\boldsymbol{v}^3, \tag{3.18}$$

where $P = 1$ if Ω corresponds to the full space, and we impose the requirement that $f \geq 0$ everywhere. The distribution function however evolves with time, such that $f = f(\boldsymbol{x}, \boldsymbol{v}, t)$. As with any mathematical distribution function, its integral must be unity, and hence we have a conservation of probability: this is given by the collisionless Boltzmann equation, namely:

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} f - \boldsymbol{\nabla}\Phi \cdot \boldsymbol{\nabla}_v f = 0. \tag{3.19}$$

An important assumption of the collisionless Boltzmann equation is the persistence of the particles tracing the distribution function: real stars, that undergo star formation and supernovae events, do not in general satisfy this condition. If however we are concerned primarily with an old, evolved stellar population (as is the case with an ETG), then Equation 3.19 is a good approximation to the system.

If we turn our consideration to spherical systems, then we may represent the total potential that is generated by the particles as a smooth potential that depends only

on the radial coordinate as $\Phi(r)$ as in Section 3.2. The system is said to be ergodic if the distribution function is a function of the relative energy $\mathcal{E}$ only:

$$\mathcal{E} = \Psi - \frac{1}{2}v^2, \tag{3.20}$$

where Ψ is the relative potential defined as:

$$\Psi = \Phi - \Phi_0, \tag{3.21}$$

and we impose Φ_0 to be such that $f > 0$ when $\mathcal{E} > 0$, and $f = 0$ elsewhere. As the distribution function depends only on the magnitude of v (through $\mathcal{E}$) and not on the direction of v, we have an isotropic system:

$$\sigma_r^2 = \sigma_\theta^2 = \sigma_\phi^2, \tag{3.22}$$

namely that the velocity variance along different coordinate axes are equal.

We can define a measure of velocity anisotropy β:

$$\beta = 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2} \tag{3.23}$$

which tends to $\beta \to -\infty$ when particles only have circular motion, and $\beta = 1$ when particles have pure radial motion; for an ergodic system, $\beta = 0$. We can then introduce a level of velocity anisotropy to the distribution function by defining a new variable Q (Osipkov, 1979; Merritt, 1985):

$$Q = \mathcal{E} - \frac{J^2}{2r_a^2}, \tag{3.24}$$

where J is the angular momentum vector and r_a is some characteristic anisotropy radius. The distribution function then transforms such that $f(\mathcal{E})$ (ergodic) $\to f(Q)$ (anisotropic). In this sense, we may consider an ergodic system as one where the anisotropy radius $r_a \to \infty$.

Knowledge of the distribution function is critical for generating initial conditions for galaxy simulations, which we explore next.

3.4 Creating initial conditions

In preparation to simulate the merger of two galaxies, we first have to create each model galaxy in isolation. First, we consider the different types of particles required

by the simulation. For ETGs, these are stars, DM, and an SMBH. As discussed in Subsection 2.1.2, the minimal amount of star-forming gas observed in ETGs allows us to neglect gas from the simulation entirely, and still capture the overall dynamical evolution of the system to a high degree of accuracy.

The initial conditions (ICs) are generated from Eddington's formula (Eddington, 1916), which given the density and potential profiles of a particle species i, allows the distribution function $f_i(\mathcal{E})$ to be constructed as[2]:

$$f_i(\mathcal{E}) = \frac{1}{2\sqrt{2}\pi^2} \int_{\Phi_T=0}^{\Phi_T=\mathcal{E}} \frac{d^2\rho_i}{d\Phi_T^2} \frac{d\Phi_T}{\sqrt{\mathcal{E} - \Phi_T}}. \tag{3.25}$$

Here, Φ_T is the total gravitational potential, $\Phi_T = \sum_i \Phi_i$.

In practice, evaluating Equation 3.25 is broken down into a series of steps, which we discuss now.

3.4.1 Component profiles

For each component (e.g. stellar, DM, or SMBH) in the ETG, we require both the density, the potential, and their derivatives. Let us define the normalised density ν such that:

$$\nu(\Psi) = \frac{\rho}{M}, \tag{3.26}$$

where ρ is the mass density of the component, and M is the total mass enclosed within some sufficiently large radius, $r_{\max} \to \infty$, and Ψ is the normalised potential from Section 3.3. As we assume the system extends to $r \simeq \infty$, we have $\Phi_0 = 0$ and hence $\Psi = \Phi$. We then compute (analytically if possible) $\frac{d\nu}{dr}$ and $\frac{d^2\nu}{dr^2}$ from the normalised density, and both $\frac{d\Psi}{dr}$ and $\frac{d^2\Psi}{dr^2}$ from the relative potential. Equation 3.25 is thus:

$$f_i(\mathcal{E}) = \frac{1}{2\sqrt{2}\pi^2} \int_0^{\mathcal{E}} \frac{d\Psi}{\sqrt{\mathcal{E} - \Psi}} \frac{d^2\nu}{d\Psi^2}. \tag{3.27}$$

3.4.2 Sampling the distribution function

With the necessary derivatives in hand, we now set about sampling the distribution function. We begin by determining the integration limits, set by the minimum and

---

[2]In this work, we concern ourselves purely with the case of ergodic ICs, although the ability to generate anisotropic ICs following Section 3.2 was explored in early work of the thesis.

maximum radii (which should be close to 0 and ∞, respectively) and finding the corresponding values of Ψ. Using the chain rule, we construct:

$$\frac{\mathrm{d}\nu}{\mathrm{d}\Psi} = \frac{\mathrm{d}\nu}{\mathrm{d}r}\left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-1} \tag{3.28}$$

$$\frac{\mathrm{d}^2\nu}{\mathrm{d}\Psi^2} = \frac{\mathrm{d}^2\nu}{\mathrm{d}r^2}\left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-2} - \frac{\mathrm{d}\nu}{\mathrm{d}r}\left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-3}\frac{\mathrm{d}^2\Psi}{\mathrm{d}r^2} \tag{3.29}$$

Next, we perform a variable transformation to avoid the $(\mathcal{E} - \Psi)^{-1/2}$ term in Equation 3.27:

$$\psi \equiv 2\sqrt{\mathcal{E} - \Psi} \implies \frac{\mathrm{d}\psi}{\mathrm{d}\Psi} = -\frac{1}{\sqrt{\mathcal{E} - \Psi}}, \tag{3.30}$$

and numerically integrate:

$$f_i(\mathcal{E}) = \frac{1}{2\sqrt{2}\pi^2}\int_0^{2\sqrt{\mathcal{E}}}\frac{\mathrm{d}^2\nu}{\mathrm{d}\Psi^2}\,\mathrm{d}\psi. \tag{3.31}$$

As $f_i(\mathcal{E})$ is a *distribution*, we can then calculate the *cumulative* distribution function:

$$F_i(\mathcal{E}) = \int_0^{\mathcal{E}} f_i(\mathcal{E}')\,\mathrm{d}\mathcal{E}', \tag{3.32}$$

which maps to the space $[0, 1]$. We then employ transformation sampling (Section 4.3) to Monte Carlo sample[3] the distribution function with discrete particles.

This process is repeated for each particle species in the simulation, allowing us to create an isolated, spherically-symmetric, ergodic representation of an ETG.

3.4.3 Merger setup

In all simulations run for this thesis, we are interested in the merger of two ETGs. Having generated the ICs for each ETG in isolation, we then set about constructing a merger of these two systems by setting them on a bound orbit.

In Section 2.3 we introduced the concept of two point masses having a bound orbit, as defined through the binding energy E_{bin}, and shown in Figure 2.10. We use a similar philosophy for constructing the merger of two galaxies, however an important difference surfaces: galaxies are not point masses, but rather extended

---

[3]Monte Carlo sampling is a numerical technique to draw random samples from a distribution, with each sample being statistically independent from other samples. It relies on the Law of Large Numbers for convergence.

objects. Hence, we approximate the galaxies as point masses, each with a mass enclosed within some radius $\tilde{r}$, such that $M \approx M(< \tilde{r})$.

The two galaxies are separated by some initial separation r_0, and the amount of mass $M(< \tilde{r})$ assigned to each galaxy is determined by setting $\tilde{r} = \xi r_0$, where ξ is a fractional number less than unity. There are two remaining free parameters required for the system: the initial eccentricity e_0, and the pericentre distance at first passage r_{peri}. Care must be taken when selecting these parameters: too small a value of r_{peri} results in the galaxies passing through each with such high relative velocities that dynamical friction is unable to bring the galaxies to merger on an appreciable timescale. In practice, one normally selects either r_{peri} or e_0, and then motivates the other through relations from statistical studies of merging haloes from cosmological simulations, such as that found in Khochfar & Burkert (2006), which has the form:

$$e_0 \simeq \left(1 + \left(\frac{1}{0.320} \frac{r_{\text{peri}}}{R_{\text{vir}}} \right)^{1.629} \right)^{-0.176} . \tag{3.33}$$

With this complete, we now have a simulation ready to be run for investigation.

3.5   GADGET-4: capturing the big picture

All the simulations in this thesis have been run with the KETJU extension of the GADGET-4 code (Springel et al., 2021). GADGET-4 is a widely-used, multipurpose galaxy simulation code with numerous configuration options, each suited and tunable for the specific scientific question being investigated. The main features, and the relevant configuration options used in this thesis work, we detail below in two broad categories: gravity calculation, and time integraton[4]. The additions that KETJU brings to GADGET-4 are discussed in Section 3.6.

3.5.1   Gravity calculation

Naïvely, one may expect that to determine the gravitational acceleration acting on a particle i from all other particles $j \neq i$, we may simply determine the acceleration as:

$$\boldsymbol{a}_i = G \sum_{j \neq i}^{N-1} \frac{m_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^2}, \tag{3.34}$$

---

[4]GADGET-4 also has the capability for hydrodynamical calculations using smoothed particle hydrodynamics (SPH). As we limit ourselves to the gas-free case, we do not include a discussion of SPH here.
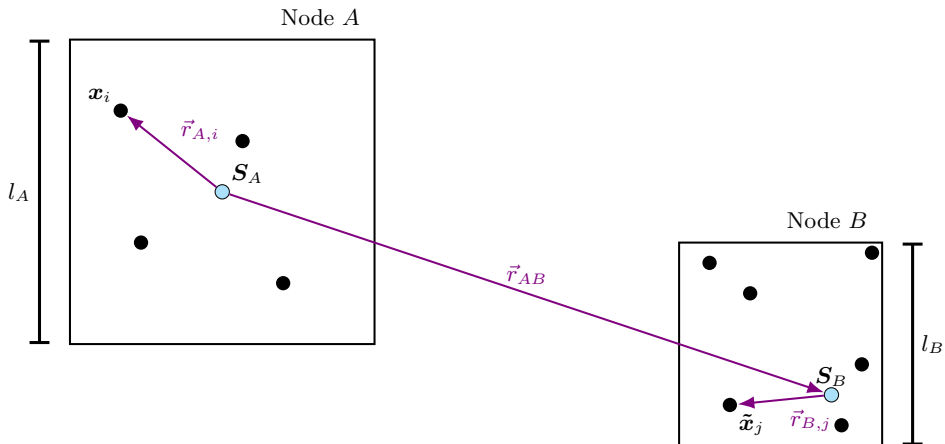
Figure 3.4: Schematic of a fast multipole method interaction between two nodes A and B, which have lengths l_A and l_B, respectively. Whilst the sketch is shown in 2D, the method generalises to 3D. Each node contains a number of particles, with the centre of mass of node A marked as S_A, and similarly S_B for node B, with the distance between the centres given as $\vec{r}_{AB}$. Distance vectors from the centre of mass of each node to the constituent particles are denoted $\vec{r}_{A,i}$ for node A, and $\vec{r}_{B,i}$ for node B.

a technique termed direct summation. However, upon inspection, we see that Equation 3.34 scales as $\mathcal{O}(N^2)$: prohibitively expensive if we consider the number of particles required to faithfully represent an ETG (see Section 3.1). A widely-used scheme that reduces the computational cost to $\mathcal{O}(N \log N)$ is the Barnes & Hut (1986) oct-tree, which approximates collections of particles sufficintly-distant from a particle i as a single source of mass. In its simplest form, the oct-tree recursively halves the simulation volume along each axis – where a sub-volume is termed an 'octant' of the parent 'node' – until a given octant contains only a single particle. In GADGET-4, this requirement has been updated so that instead of the recursive-division algorithm terminating when there is a single particle in a given octant, splitting only occurs when the number of particles in an octant exceeds some user-defined threshold; we use $N_{\mathrm{split}} = 4$.

Let us term the particle we wish to determine the gravitational force for as the *sink* particle, and those particles providing the gravitational force as the *source* particles. Traditionally, the force acting on the sink particle from the source particles is approximated as coming collectively from the tree node (of side length l) containing

the source particles if the node is far enough away from the sink particle:

$$\frac{l}{r} < \theta_c, \tag{3.35}$$

where r is the radial separation between the sink particle and the node centre of mass, and θ_c is a user-defined tolerance angle. A disadvantage of this 'one-sided' tree force calculation approach is that momentum is not necessarily conserved, due to the asymmetry in the use of nodes to represent the force from source particles. If we denote the node centre of mass as $\boldsymbol{S}$, we have for the potential at the sink particle position $\boldsymbol{x}$:

$$\Phi(\boldsymbol{x}) = -G \sum_{j \in \text{node}} m_j g(\boldsymbol{x}_j - \boldsymbol{x}) \rightarrow -G \sum_{n=0}^{p} \frac{1}{n!} \boldsymbol{Q}_n \cdot \boldsymbol{D}_n(\boldsymbol{S} - \boldsymbol{x}) + \dots . \tag{3.36}$$

In the first equality, the function $g(\boldsymbol{x}_j - \boldsymbol{x})$ is the Green's function of the interaction, which is simply $1/|\boldsymbol{x}|$ in the Newtonian case[5]. The second equality shows the conversion to a multipole expansion of order p about the node centre of mass, where:

$$\boldsymbol{Q}_n \equiv \sum_{j \in \text{node}} m_j (\boldsymbol{x} - \boldsymbol{s})^{(n)}, \tag{3.37}$$

$$\boldsymbol{D}_n = \nabla^{(n)} g(\boldsymbol{x}) \tag{3.38}$$

are the Cartesian multipole expansion and derivative tensors, respectively. The notation $\boldsymbol{y}^{(n)}$ denotes the nth tensor outer product with itself, i.e. $\boldsymbol{y} \otimes \boldsymbol{y} \otimes \cdots \otimes \boldsymbol{y}$. GADGET-4 supports multipole expansion up to order 5.

Computational speed-up, as well as the attractive benefit of momentum conservation, can be achieved by using the Fast Multipole Method (FMM), and is used in all simulations for this thesis. A schematic of the FMM scheme is given in Figure 3.4. We begin by labelling particles in node A (the sink node) as having coordinates $\boldsymbol{x}_i$, and the particles in node B (the source node) as $\tilde{\boldsymbol{x}}_j$. Equation 3.36 then becomes:

$$\Phi(\boldsymbol{x}_i) = -G \sum_{j \in \text{node}} m_j g(\tilde{\boldsymbol{x}}_j - \boldsymbol{x}_i) \tag{3.39}$$

where $g(\tilde{\boldsymbol{x}}_j - \boldsymbol{x}_i)$ is again the Newtonian interaction term. Following the notation of Figure 3.4, let the vector between the node centres of mass ($\boldsymbol{S}_A$ and $\boldsymbol{S}_B$) be

---

[5]The Green's function $g(\boldsymbol{x}_j - \boldsymbol{x})$ becomes more complex when we include gravitational softening. For now, considering the Newtonian case serves our purposes.

$\vec{r}_{AB} = \boldsymbol{S}_B - \boldsymbol{S}_A$. We then have:

$$\boldsymbol{x}_i = \boldsymbol{S}_A + \vec{r}_{A,i} \tag{3.40}$$

$$\tilde{\boldsymbol{x}}_j = \boldsymbol{S}_B + \vec{r}_{B,j} \tag{3.41}$$

which leads to $\tilde{\boldsymbol{x}}_j - \boldsymbol{x}_i = \vec{r}_{AB} + \vec{r}_{B,j} - \vec{r}_{A,i}$. This then allows for the potential at $\boldsymbol{x}_i$ to be expressed, analogous to Equation 3.36 as:

$$\Phi(\boldsymbol{x}_i) \simeq -G \sum_{k=0}^{p} \frac{(-1)^k}{k!} \vec{r}_{A,i}^{(k)} \cdot \sum_{n=0}^{p-k} \frac{1}{n!} \boldsymbol{D}_{n+k} \cdot \boldsymbol{Q}_n^B, \tag{3.42}$$

where $\boldsymbol{D}_n$ has the same meaning as before, and and $\boldsymbol{Q}_n^B \equiv \sum_j m_j \vec{r}_{B,j}^{(n)}$. In this work, we use order 2 multipoles, and hence the expression for $\boldsymbol{D}_n = \boldsymbol{D}_2$ becomes:

$$\boldsymbol{D}_2 = \nabla^{(2)} g(\vec{r}_{AB}) = \left(\frac{1}{r}\frac{d}{dr}\right) g(r)\delta_{ij} + \left(\frac{1}{r}\frac{d}{dr}\right)^2 g(r) r_i r_j, \tag{3.43}$$

where $r \equiv |\vec{r}_{AB}|$, $g(r)$ is again the Green's function, δ_{ij} is the Kronecker delta, and r_i is the ith component of $\vec{r}_{AB}$.

The criterion of when to use a node for the gravitational interaction is an adapted version of Equation 3.35:

$$\frac{l_A + l_B}{r} < \theta_c \tag{3.44}$$

if we use a purely-geometric criterion, or if we use the recommended force-error controlled version:

$$\frac{\max(M_A, M_B)}{r^2} \left(\frac{\max(l_A, l_B)}{r}\right)^{p-1} < \alpha |\min(|\boldsymbol{a}|_A, \boldsymbol{a}|_B)|. \tag{3.45}$$

Here M_i is the mass of node i, $|\boldsymbol{a}|_i$ is the minimum acceleration occuring between any of the particles in either node, p is the the multipole expansion order, and α is a parameter controlling the force accuracy.

Up until now, we have considered the Green's function $g(\boldsymbol{x})$ to be the Newtonian potential, $1/|\boldsymbol{x}|$. In practice, we need to consider the simulation particles not as point masses, but rather a rigid spherical collection of particles representing a given family (e.g. stars or DM). This leads to the concept of gravitational softening: the main result being a finite force at all particle separations.

The simplest mechanism to introduce a softened potential is the Plummer potential, already introduced in Equation 3.14. Whilst this ensures a finite force, the
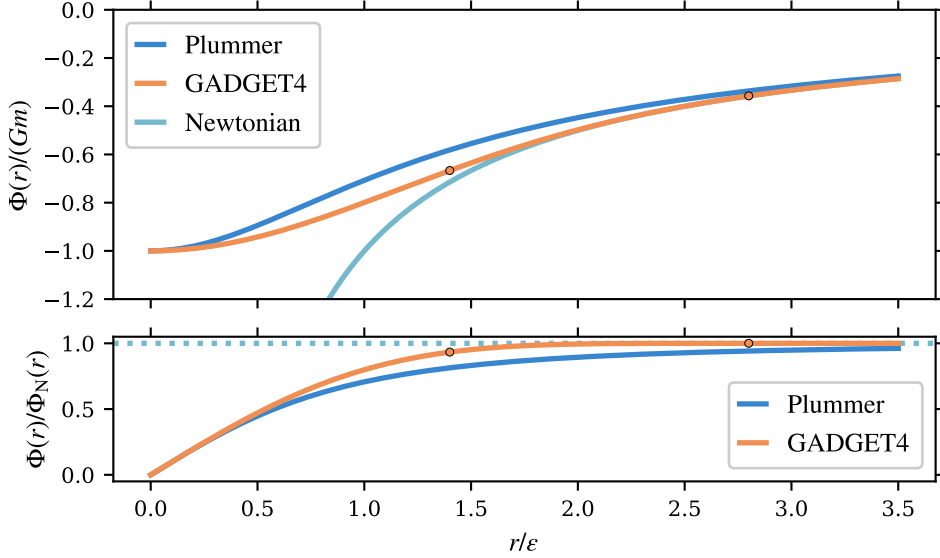
Figure 3.5: *Top panel*: Normalised potential for the Plummer model, the spline kernel used by GADGET-4, and the Newtonian potential. The radial coordinate r is scaled by the particle softening length ε. Markers on the orange line indicate the different interfaces of the spline kernel. Note that the Newtonian potential diverges as $r \to 0$. *Bottom panel*: Relative difference between the Plummer and spline model to the Newtonian potential. Importantly, the spline kernel used by GADGET-4 is exactly Newtonian beyond $r/\varepsilon = 2.8$, unlike the Plummer potential, which is only asymptotically Newtonian.

force is by construct never Newtonian, hence introducing a small but constant error in each force calculation, even when particles are well separated. In GADGET-4, the softening kernel used instead is a three-part spline kernel, which has support on $r < h = 2.8\varepsilon$, where ε is the Plummer-equivalent softening length, such that at $r = 0$ the potential is equivalent to that of a Plummer sphere. Defining the variable $u = r/h$, the spline kernel is:

$$\Phi(r) = -Gm \begin{cases} h^{-1}\left(-\frac{16}{3}u^2 + \frac{48}{5}u^4 - \frac{32}{5}u^5 + \frac{14}{5}\right), & r < \frac{h}{2} \\ h^{-1}\left(-\frac{1}{15u} - \frac{32}{3}u^2 + 16u^3 - \frac{48}{5}u^4 + \frac{32}{15}u^5 + \frac{16}{5}\right), & \frac{h}{2} \leq r < h \\ \frac{1}{r}, & r \geq h. \end{cases}$$

$$(3.46)$$

A comparison of the Plummer, spline, and Newtonian potentials is shown in the top panel of Figure 3.5. The ratios of the Plummer and spline potentials to the Newtonian potential are shown in the bottom panel, where we see that beyond $r/\varepsilon = 2.8$,

$\Phi(r)/\Phi_{\mathrm{N}} = 1$ for the spline potential, whereas the ratio is only asymptotically 1 as $r \to \infty$ for the Plummer potential.

3.5.2 Time integration

Simulations of merging galaxies take the initial conditions at a time t_0 and integrate the system forward until some final time t_{f}. The integration is performed by advancing the system by discrete timesteps Δt, with the gravity calculations described in Subsection 3.5.1 being performed at each timestep of the simulation.

GADGET-4 uses a kick-drift-kick (KDK) leapfrog integration scheme to advance particles in time, which for fixed timesteps is a symplectic (phase space conserving) scheme (e.g. Springel, 2005). Kicks correspond to updates in velocities, whereas drifts correspond to updates in position. The timestep evolution operator $E(\Delta t)$ consists of two kick operators K and one drift operator D, and for a given timestep Δt is[6]:

$$E(\Delta t) = K\left(\frac{\Delta t}{2}\right) \circ D\left(\Delta t\right) \circ K\left(\frac{\Delta t}{2}\right) \tag{3.47}$$

such that:

$$E(\Delta t)\left[(\boldsymbol{x}(t), \boldsymbol{v}(t - \tfrac{1}{2}\Delta t))\right] \to (\boldsymbol{x}(t + \Delta t), \boldsymbol{v}(t + \tfrac{1}{2}\Delta t)). \tag{3.48}$$

In GADGET-4 there is the option for hierarchical time integration, which we make use of for the simulations in this thesis. Consider a particle system represented by the Hamiltonian H such that it can be divided into a kinetic and a potential contribution:

$$H = H_{\mathrm{kin}} + H_{\mathrm{pot}}. \tag{3.49}$$

Further, let us assume that at a given timestep Δt, there exists two groups of particles: those for which an integration step of Δt would be sufficiently accurate, and another group which would require a smaller timestep, say $\Delta t/2$. The first group we term 'slow' (S), and the second group 'fast' (F), and the union of these sets we denote P. Note that particles for which a timestep of (for example) $2\Delta t$ would be sufficiently accurate, do not belong to the set P. The Hamiltonian can thus be represented as:

$$H = H_{\mathrm{kin}}^{\mathrm{S}} + H_{\mathrm{pot}}^{\mathrm{S}} + H_{\mathrm{kin}}^{\mathrm{F}} + H_{\mathrm{pot}}^{\mathrm{F}} + H_{\mathrm{pot}}^{\mathrm{FS}} \tag{3.50}$$

$$= H^{\mathrm{S}} + H^{\mathrm{F}} + H_{\mathrm{pot}}^{\mathrm{FS}} \tag{3.51}$$

where the last term is an interaction term between the fast and slow particle sets.

---

[6]The notation $f \circ g$ denotes a composite operator, such that $(f \circ g)(x) = f(g(x))$.
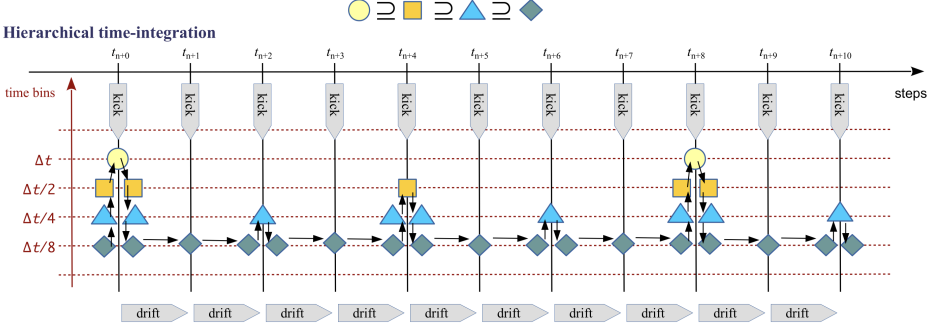
Figure 3.6: Schematic demonstrating the method of hierarchical time-stepping in GADGET-4. In the sketch, there are four active timebins: the circle is active on the fiducial Δt timebin, the squares are active on the next lowest timebin $\Delta t/2$, triangles on the $\Delta t/4$ timebin, and finally diamonds occupy the lowest timebin $\Delta t/8$. To provide the correct kick operation, the force contribution to a particle from particles on its timebin and all lower timebins is required. For example, the square point requires the triangle and and diamond point for its kick operation, but not the circle point. Similarly, to update the diamond point, only particles on the lowest timebin are required. All forces are calculated pairwise, thus by definition the scheme is momentum-conserving. Figure adapted from Springel et al. (2021).

Let us now introduce a new notation for the kick operator $K \rightarrow K_A^B$, which indicates that the forces *due to* particles in B are applied to the particles in A. The time evolution operator in Equation 3.47 thus becomes:

$$E(H, \Delta t) \simeq K_P^P \left(\frac{\Delta t}{2} \right) K_F^F \left(-\frac{\Delta t}{2} \right)$$
$$\times \ E \left(H_F, \frac{\Delta t}{2} \right) D_S(\Delta t) E \left(H_F, \frac{\Delta t}{2} \right) \ \times \ K_F^F \left(-\frac{\Delta t}{2} \right) K_P^P \left(\frac{\Delta t}{2} \right). \quad (3.52)$$

The benefit of Equation 3.52 is that for fast particles (set F), new forces are calculated only from those particles in that set – we have no K_F^S term. The scheme can now be applied recursively, such that those particles in F form a new system P', for which if a subset of particles F' require a smaller timestep than $\Delta t/2$ for sufficiently accurate integration, we can apply Equation 3.52 to P'. This leads to the hierarchical time integration scheme, which is demonstrated in Figure 3.6. Critically, all operators in Equation 3.52 are of the form K_A^A – we have momentum conservation for the local time integration. An additional advantage of the hierarchical time integration scheme is that the gravity tree computation in Subsection 3.5.1 need only be done for those particles in the topmost active timebin at the given timestep: at most Δt, but
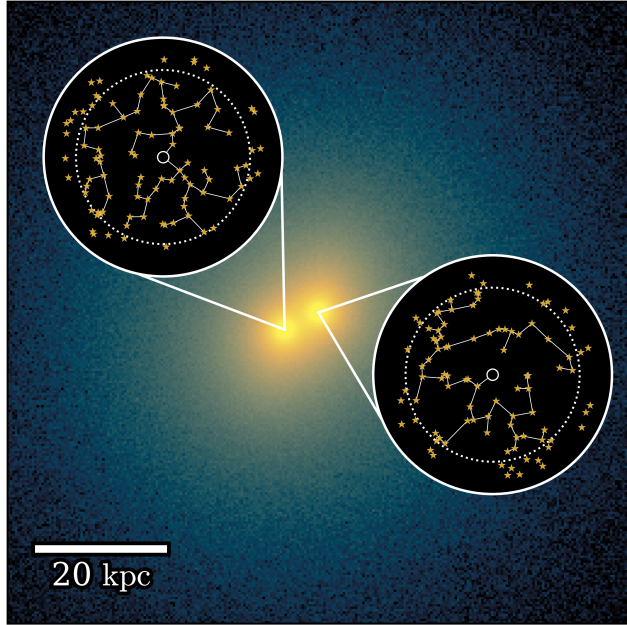
Figure 3.7: Illustrative figure of KETJU. Spherical regions of radii $r_{\text{ketju}} = 2.8\varepsilon$ are introduced around each SMBH (shown as the dotted circle), within which the MSTAR integrator is used to integrate the orbits of the SMBHs and stars to high accuracy. MSTAR uses a minimum spanning tree coordinate system, shown as the white connector lines between stars and SMBHs in the inset panels.

if recursive cycling has been done, this can be lower, as demonstrated in Figure 3.6.

3.6 KETJU: Regularised dynamics in large-scale simulations

In Equation 3.46, we motivated the use of a softened potential when integrating massive tracer particles that represent a *collection* of physical particles, such as stars. This methodology is not appropriate for SMBH interactions, which are physically point-mass particles that undergo hard scattering interactions. If we were to integrate SMBHs with softened interactions (where the particle softening length is typically of the order of tens of parsecs), the SMBH dynamics are not able to be accurately captured during the binary phase, and three-body interactions with stars are unable to extract the correct energy and angular momentum from the SMBH binary. This

can lead to SMBH binary stalling (see Subsection 2.3.2) at the softening length (e.g. Milosavljević & Merritt, 2001).

A naïve solution might be to reduce the softening length used to arbitrarily-small values, to better resolve the force between particles at small physical separations. Aside from introducing an increased numerical error, smaller softening lengths necessitate smaller timesteps to capture the dynamics to a given precision – practically, the simulation becomes unrunnable within a reasonable wallclock time. For this reason, alternative methods are needed for accurately modelling SMBH dynamics.

One solution is the use of specialised N-body codes that are designed for simulations with very small, or even zero, softening (Berentzen et al., 2009; Dehnen, 2014; Khan et al., 2016; Nasim et al., 2021). These codes are however generally limited to particle numbers of $N \lesssim 10^6$ due to the $\mathcal{O}(N^2)$ scaling they entail during the direct summation calculation of particle accelerations. Some mass refinement schemes (e.g. Zhang & Magorrian, 2008; Attard et al., 2024) aim to emulate higher-resolution[7] simulations, but are still fundamentally limited in the number of particles able to be feasibly run.

Another solution is to introduce a new integration scheme only for a *subset* of particles for which accurate dynamics are a necessity. This is the fundamental philosophy of the KETJU code (Rantala et al., 2017; Mannerkoski et al., 2023), which combines the MSTAR integrator (Rantala et al., 2020) with GADGET-4 to achieve accurate SMBH dynamics integration within a large-scale galaxy simulation. Such a hybrid scheme is not limited to $N \lesssim 10^6$ particles, and being an extension to GADGET-4, allows for the wider environment (such as dark matter haloes that typically enshroud the stellar component of a galaxy) to be captured and influence the SMBH dynamics. Other hybrid codes with a similar philosophy (e.g. Harfst et al., 2008; Iwasawa et al., 2015; Karl et al., 2015) also exist in the literature.

The hybrid scheme introduced by KETJU defines a spherical region of radius r_{ketju} about each SMBH in a simulation, which must satisfy:

$$r_{\text{ketju}} \overset{!}{\geq} 2.8\varepsilon_\star. \tag{3.53}$$

This requirement is a consequence of the smoothing spline used in GADGET-4: from Equation 3.46, the potential generated by particles at distances $r \geq 2.8\varepsilon_\star$ from a given particle is exactly Newtonian. The integration scheme of KETJU, MSTAR, then ensures that particle accelerations within r_{ketju} of the SMBH are not softened, allowing for accurate modelling of SMBH dynamics. As the computational cost of

---

[7]Resolution here refers to the mass per particle: a 'higher' resolution therefore corresponds to a smaller particle mass than a 'lower' resolution.

KETJU scales similar to N-body codes, in practice r_{ketju} is chosen to be $2.8\varepsilon_\star$ and not larger. We discuss the main features of KETJU and MSTAR below.

MSTAR introduces a new coordinate system to reduce numerical errors in subtracting large floating point numbers. The coordinate system is based on a minimum spanning tree (MST, see Figure 3.7), where particle separations and velocities are stored so as to minimise the sum of the separations. The MST is constructed using Prim's algorithm (Prim, 1957) with the regular Cartesian coordinates at the start of the integration, and then updated at regular intervals throughout. The Cartesian coordinates are updated at the end of each integration from the MST coordinates.

Additionally, MSTAR introduces algorithmic regularisation to remove the coordinate singularity $1/|\boldsymbol{x}|$ in the Newtonian potential. This is achieved by parameterising the equations of motion with a new variable s, which is related to the time variable t as:

$$\frac{\mathrm{d}t}{\mathrm{d}s} = \frac{1}{T + B}, \tag{3.54}$$

where T is the kinetic energy and $B = U - T$ the binding energy (where U is the potential energy) of the system. The equations of motion then become:

$$\frac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}s} = \frac{1}{T + B}\boldsymbol{v}_i \tag{3.55}$$

$$\frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}s} = \frac{1}{U}(\boldsymbol{a}_i + \boldsymbol{f}_i + \boldsymbol{g}_i(\boldsymbol{v})) \tag{3.56}$$

$$\frac{\mathrm{d}B}{\mathrm{d}s} = -\frac{1}{U}\sum_i m_i \boldsymbol{v}_i \cdot (\boldsymbol{f}_i + \boldsymbol{g}_i(\boldsymbol{v})), \tag{3.57}$$

where $\boldsymbol{a}_i$ is the Newtonian acceleration of the particles, $\boldsymbol{f}_i$ encodes extra contributions to the acceleration depending only the particle position, and $\boldsymbol{g}_i(\boldsymbol{v})$ denotes extra contributions to the acceleration depending on the particle velocity, such as the post-Newtonian contribution. The regularised integration scheme is used only for SMBH-SMBH and SMBH-stellar interactions. As DM particles typically have masses comparable to the SMBH, and we wish to model DM as a continuous fluid that provides some background potential rather than assuming a particular particle cross section, all DM-SMBH interactions are softened. Additionally, to reduce energy errors between stellar particles exiting and entering the KETJU region and collisional dynamics between stellar particles, all stellar-stellar interactions are softened as well.

The integration scheme of MSTAR makes use of Gragg-Bulirsch-Stoer (GBS) extrapolation (Gragg, 1965; Bulirsch & Stoer, 1966). Given a timestep Δt, the algorithm integrates the timestep with successively smaller steps $\Delta t/n$, and then uses polynomial extrapolation to determine the integration for $n \to \infty$.
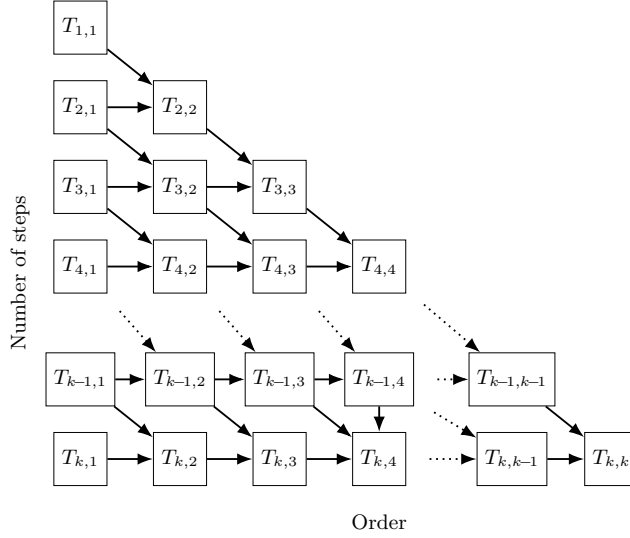
Figure 3.8: Schematic of the Gragg-Bulirsch-Stoer extrapolation algorithm, given in Equation 3.59. At each order of the extrapolation, the integration interval is divided into a number of steps $\Delta t/(2k)$, until the relative error between successive steps is below some user threshold η_{ketju}.

For a given number of substeps n, the evolution operator (following Subsection 3.5.2) becomes:

$$E(\Delta t)^{(n)} = D\left(\frac{\Delta t}{2n}\right) K\left(\frac{\Delta t}{n}\right) \left[D\left(\frac{\Delta t}{n}\right) K\left(\frac{\Delta t}{n}\right)\right]^{n-1} D\left(\frac{\Delta t}{2n}\right). \qquad (3.58)$$

Having completed the integration for $n_i = 2i$ substeps with $i = 1, \ldots k$, the results for $T_{i,1}$ are used to form the first column of a lower-triangular matrix, shown in Figure 3.8. Successive columns of the matrix are then populated using polynomial extrapolation for $i > 1$:

$$T_{i,j+1} = T_{i,j} + \frac{T_{i,j} - T_{i-1,j}}{(n_i/n_{i-j})^2 - 1}, \qquad (3.59)$$

where the index $j \in [1, \ldots, i]$. Convergence is deemed to have occured when the extrapolated result falls below a relative error:

$$\left|\frac{T_{k,k-1} - T_{k,k}}{T_{k,k}}\right| < \eta_{\text{ketju}}, \qquad (3.60)$$

59

where η_{ketju} is a tolerance parameter defined by the user, and is typically set to $\eta_{\text{ketju}} = 10^{-8}$. Irrespective of whether or not Equation 3.60 is satisfied, the timestep Δt_{next} is chosen such that:

$$\Delta t_{\text{next}} = 0.94 \left(\frac{0.65 \eta_{\text{ketju}}}{\epsilon} \right)^{\frac{1}{2k-1}} \Delta t, \tag{3.61}$$

where ϵ is the largest error in any of the coordinates of the system (Hairer et al., 1993).

The dynamics of SMBHs as their separation decreases will at some point deviate from Newtonian mechanics. For this reason, Ketju includes post-Newtonian corrections to the equations of motion (see Subsection 2.3.3) up to and including order PN3.5 for SMBH-SMBH interactions. At some point however, even the PN corrections are inadequate at describing the motion of the SMBH binary: this defines the minimum spatial scale that Ketju can accurately resolve SMBH dynamics to. At the point the SMBH binary has a separation of:

$$r_{\text{merge}} = \frac{12G \left(M_{\bullet,1} + M_{\bullet,2} \right)}{c^2}, \tag{3.62}$$

equivalently four times the innermost stable circular orbit, $4R_{\text{ISCO}}$, the SMBH binary is merged to form a single SMBH. In practice, this generally translates to the final few orbits before the SMBH would merge under full general relativistic calculations, keeping in mind that at this moment the orbit of the SMBH binary is almost always circular.

Additionally, Ketju incorporates the fitting formulae from Zlochower & Lousto (2015) to determine the loss of mass from the coalesced SMBH following merger, as well as the remant SMBH spin and any potential recoil velocity. The addition of these formulae allow, in principle, for the following of the SMBH after a merger has occurred. A key assumption in the Zlochower & Lousto (2015) modelling however is that the *distribution* of SMBH spin is the same at large radii as at small radii, though the particular-sampled quantity of that distribution might be different.

4 Novel analysis techniques

Thus far we have discussed the motivation and practical implementation of simulations of supermassive black holes in galaxies. This chapter focuses on some of the novel techniques employed and developed throughout the thesis to achieve an understanding of the relationship between the SMBH and its host galaxy.

4.1 Orbit modelling

The first analysis tool we discuss is orbit modelling. Orbits are the building blocks of the stellar component of a galaxy; an understanding of orbits allows for exploring the inner structure and dynamics of galaxies from both an observational (e.g. Schwarzschild, 1979) and theoretical perspective. In particular, to gain a deeper understanding of the influence of an SMBH on the surrounding stellar dynamics, a knowledge of orbits is necessary.

In broad terms, there are two main categories of orbits: box orbits and tube orbits (Binney & Tremaine, 2008). Box orbits (shown in the top left of Figure 4.1) do not conserve net angular momentum, and may approach arbitrarily small distances to the centre of the gravitational potential – these orbits are centrophilic. Conversely, tube orbits (shown in the bottom left of Figure 4.1) conserve net angular momentum, and as a result are not able to explore the full coordinate space, being bounded by an inner and outer annulus – these orbits are centrophobic. Both box and tube orbits are comprised of different subclasses, which are identified through the resonances (or lack thereof) of the particle along the galaxy axes. The main steps in the analysis are outlined below.

First, the galaxy is rotated so that the z-axis coincides with the minor axis of the reduced inertia tensor (Gerhard, 1983; Bailin & Steinmetz, 2005) as measured using the top 50 per cent most bound stellar particles to the stellar centre of mass:

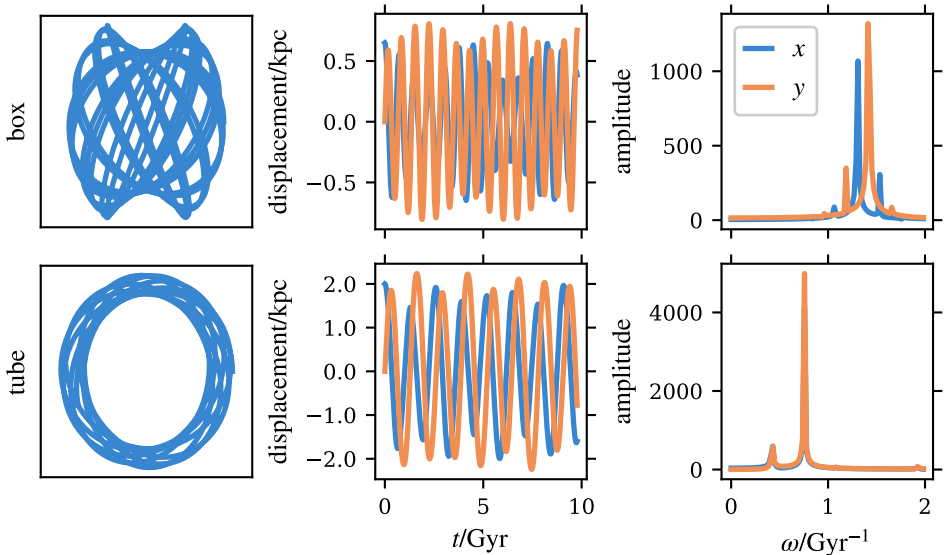$$\tilde{I}_{i,j} = \sum_k \frac{r_{k,i} r_{k,j}}{r_k^2}. \tag{4.1}$$

Figure 4.1: Representative overview of orbit modelling in 2D in an analytical potential. Top row shows a centrophilic box orbit, whereas the bottom row shows a centrophobic tube orbit. *Left column*: The *x*-*y* position of the particle. *Centre column*: Time series of the *x* (blue) and *y* (orange) coordinates. *Right column*: Fourier transform of the positional coordinates, with colours consistent with the centre column. The frequency coinciding with the maximum peaks for the *x* and *y* coordinates are found, and the ratio used to determine the orbital family.

We then need to fit the gravitational potential of the system using a self-consistent field (SCF) potential (Hernquist & Ostriker, 1992; Jesseit et al., 2005; Röttgers et al., 2014), representing the potential $\Phi(\boldsymbol{x})$ generated by the particles in the simulation as a combination of harmonics with expansion coefficients A_{nlm}:

$$\Phi(\boldsymbol{x}) = \sum_{nlm} A_{nlm}\Phi_{nlm}(\boldsymbol{x}), \tag{4.2}$$

where each individual harmonic independently satisfies the Poisson equation (Equation 3.10). The coefficients A_{nlm} describe the order n the radial component of the potential is expanded to, and to which order l, m the angular components are expanded; the angular dependence is a (somewhat complicated) function of the spherical harmonics $Y_{lm}(\theta, \phi)$, where $|m| < l$ as usual. The zeroth order of the potential expansion $\Phi_{000}(\boldsymbol{x})$ is the model upon which we wish to construct our expansion, and a common choice for this is the Hernquist (1990) profile, which is an acceptable

Table 4.1: Classification of orbital families

Family	$n_x : n_y : n_z$	Description
rosette	$1 : 1 : 1$	Typical orbit in a point-mass dominated spherically-symmetric potential
inner x-tube	$n_x : 1 : 1$	Rotate about the major axis of the galaxy, but move radially when far from centre (concave shaped)
outer x-tube	$n_x : 1 : 1$	Rotate about the major axis of the galaxy (convex shaped)
z-tube	$1 : 1 : n_z$	Rotate about the minor axis of the galaxy
π-box	None	Non-resonant motion with no net angular momentum (radial motion)
boxlet	$m : n \,\|m \neq n$	Resonant motion along any two axes with no net angular momentum
irregular	–	No integrals of motion
unclassified	–	Orbits unable to be classified

representation of many massive elliptical galaxies. Following Frigo et al. (2021), the expansion is limited to $n_{\max} = 18$ and $l_{\max} = 7$, as this captures the potential of the galactic merger remnant with sufficient accuracy.

Whilst the chosen basis functions describe the stellar and dark matter components of the galaxy, they are unable to describe a point-mass potential. This complicates simulations containing an SMBH. To overcome this, we centre the snapshot on the SMBH position, but exclude the SMBH from the SCF routine. This provides us with a potential field without the SMBH contribution; the potential from the SMBH is then added as a point-mass potential following the SCF potential construction. We now have a representation of the potential in an analytical form in which the orbits of individual stellar particles can be integrated. The potential of the merger remnant is checked for stability by comparing the potential from the SCF method to the potential computed from the particle data by GADGET-4 during the simulation, and ensuring the ratio of the two is ~ 1 at all radii.

Each stellar particle within some aperture of the centre is integrated for a fixed number of orbits (centre column in Figure 4.1). The Fourier transform of the particle motion along each axis is then calculated to determine the dominant (i.e. highest amplitude) frequencies of the orbit (right column in Figure 4.1). The ratio of the dominant frequencies along each axis are then compared to determine the orbital

family of the particle, with a resonance being defined when:

$$n_x\omega_x + n_y\omega_y + n_z\omega_z = 0 \tag{4.3}$$

following Carpintero & Aguilar (1998) and Jesseit et al. (2005), where $n_i \in \mathbb{Z}$. The orbital resonances define the different families of orbits, as given in Frigo et al. (2021), and are listed in Table 4.1. It should be noted that $1 : n_y : 1$ orbits (intermediate axis orbits, or y-tube orbits), whilst quasi-stable for some period of time (e.g. Neureiter et al., 2021), are not stable in a triaxial potential.

An important caveat for the orbit analysis is that as each stellar particle is integrated in a fixed potential, the effect of two-body interactions are not captured by this method.

4.2 IFU modelling

Direct observations of the internal dynamics of galaxies are achieved by taking spectra of the galaxy. From a spectrum, it is possible to ascertain the line-of-sight velocity distribution (LOSVD) of stars or gas, but not the transverse motion. The LOSVD of a galaxy is in general complex, and is commonly parameterised as a Gauss-Hermite series of the form given in Equation 2.8 and Equation 2.9 (van der Marel & Franx, 1993). The effect of varying the mean velocity V, the standard deviation σ, and the first two higher-order moments h_3 and h_4 are shown in Figure 4.2.

Whilst V and σ have the same meaning as in regular Gaussian statistics, h_3 and h_4 are non-trivial. The coefficient h_3 describes asymmetric deviations from a Gaussian distribution, and has the effect of introducing skew to the distribution. For more extreme values of h_3, an off-centre secondary mode is introduced to the distribution. Physically, $h_3 \neq 0$ can arise from the superposition of a slowly-rotating bulge with a more rapidly-rotating disc (e.g. Halliday et al., 2001) or other kinematically decoupled components (KDCs), particularly in fast-rotator galaxies (Krajnović et al., 2008).

Similarly, the coefficient h_4 describes symmetric deviations from a Gaussian distribution. For $h_4 < 0$, the effect is to broaden the central peak of the distribution, whereas for $h_4 > 0$, the effect is to narrow the central peak and broaden the tails of the distribution. Features in h_4 are readily understood in terms of the different stellar orbits. For box orbits, the zero velocity at their apocentre (typically coinciding with large radii) results in a positively-biased h_4 as there is an overrepresentation of $v_{\mathrm{LOS}} \sim 0\,\mathrm{km\,s^{-1}}$ stars. As these same stars then pass through their pericentre (typically at small separations from the centre of the galaxy) with their maximal
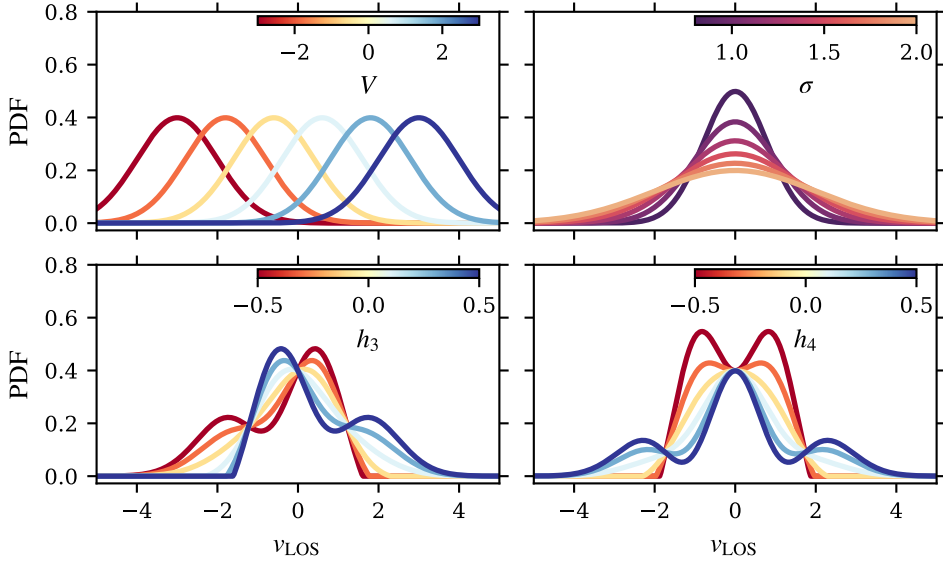
Figure 4.2: Systematic variation of the first four moments of the Gauss-Hermite series. *Top left panel*: Changing the mean of the distribution has the effect of shifting the peak of the distribution. *Top right panel*: Increasing the standard deviation broadens the distribution. *Bottom left panel*: A non-zero first order asymmetric deviation h_3 creates a multimodal distribution, with two peaks of different heights. The location of the smaller peak relative to $v_{\mathrm{LOS}} = 0$ coincides with $\mathrm{sign}(h_3)$. *Bottom right panel*: A non-zero first order symmetric deviation h_4 acts to either flatten the centre of the distribution ($h_4 < 0$) or broaden the wings of the distribution ($h_4 > 0$). For extreme values of $|h_4|$, multiple modes may be introduced to the distribution.

velocity, the various projections of the radial orbits serves to broaden the peak of the velocity distribution, seen as a negatively-biased h_4. A negative h_4 can also indicate a counter-rotating disc structure for tube orbits – two spatially-coherent groups of stars with overall v_{LOS} towards and away from the observer.

Thus far, we have discussed a system's LOSVD assuming data obtained from some spectra. Historically, such spectra were taken by using a long slit extending over a large fraction of the full galaxy. Modern advances however have introduced integral field unit (IFU) spectroscopy, where the spectra at many locations of a galaxy can be taken simultaneously. This allows for LOSVD maps to be constructed, as opposed to a single LOSVD for the full galaxy.

Mock IFU observations feature in Rawlings et al. (2025a) and Rawlings et al. (2025b) of this thesis. The method of constructing these maps is shown in Figure 4.3
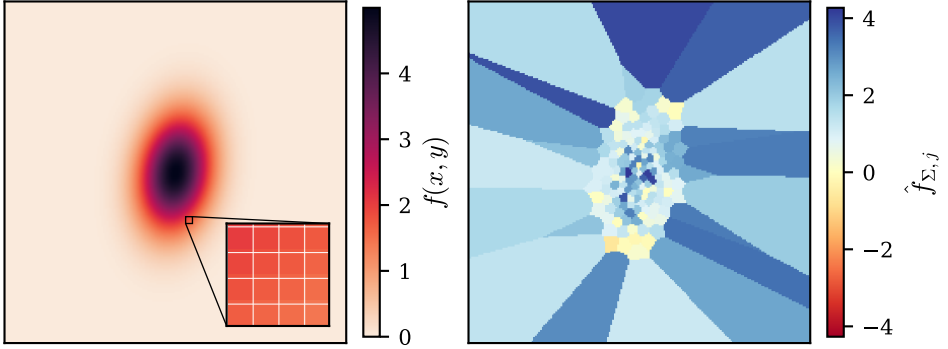
Figure 4.3: Overview of Voronoi tesselation for an arbitrary multinormal function. Particles are gridded onto a regular rectangular mesh (left panel), and then grid cells added to achieve a desired particle count per Voronoi bin (right panel). Note that the Voronoi bins are smaller towards the centre of the image, coinciding with higher particle counts, and that the value of $\hat{f}_{\Sigma,j}$ within each Voronoi bin is $\gtrsim 0$.

for an arbitrary distribution, and follows Voronoi tesselation. First, the data is binned onto a regular rectangular grid – for mock IFU observations, the width of these rectangular bins is set by the spatial sampling and resolution of the particular instrument one assumes. If we consider the value in each rectangular bin however as representing some signal – be it photons incident on a detector, or stellar particles in a simulation – to fairly compare statistical properties (such as the LOSVD) across bins, we require the bins to contain a minimum amount of signal to reduce the effect of poor sampling. To this end, we can perform Voronoi tesselation, whereby starting from the centre of the image, adjacent rectangular bins are joined to form irregular bins that have a signal in excess of some threshold S:

$$f_{\Sigma,j} = \sum_{i \in \text{Voronoi}} f(x_i, y_i) \gtrsim S \tag{4.4}$$

$$\hat{f}_{\Sigma,j} = \frac{f_{\Sigma,j} - \mathrm{E}\left(f_{\Sigma,j}\right)}{\sqrt{\mathrm{var}\left(f_{\Sigma,j}\right)}} \tag{4.5}$$

where $\hat{f}_{\Sigma,j}$ is the normalised excess count, as shown in the right panel of Figure 4.3, where all cells have $\hat{f}_{\Sigma,j} \gtrsim 0$. A result, the Voronoi bins are generally smaller towards the centre of the graph than at its outer edges.
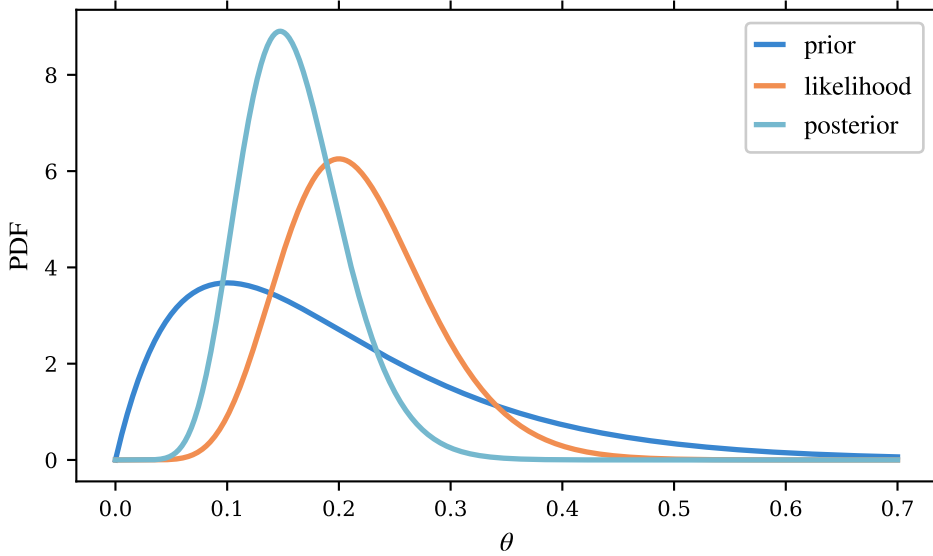
Figure 4.4: The principle tenet of Bayesian analysis: prior belief about the distribution of a variable is updated via the likelihood function (here normalised to integrate to unity) to form the posterior distribution of the parameter. In this example, both the prior and posterior distribution follow a gamma distribution, with the data generating function following an exponential distribution. In this case, the form of the posterior is analytically tractable. Instances where the prior and posterior distributions have the same form are termed 'conjugate distributions'.

4.3 Bayesian methods

A large focus of this thesis has been on the incorporation of rigorous statistical modelling through Bayesian methods (e.g. Gelman et al., 2015). Bayesian statistics assigns a prior belief to a parameter of interest θ, and is then updated given observations $y = (y_1, y_2, \ldots, y_n)$ from a data generating process (which is some function of θ) to form the posterior distribution of θ. Mathematically, this is encoded through Bayes' formula:

$$p(\theta|y) \propto p(\theta)p(y|\theta). \tag{4.6}$$

Here, $p(\cdot)$ is a probability distribution function, and $p(\cdot|\cdot)$ is a conditional distribution. Hence, in Equation 4.6, we have the (unnormalised) posterior distribution $p(\theta|y)$ ("probability of θ given data y"), the prior distribution $p(\theta)$, and the sampling distribution $p(y|\theta)$ ("probability of observing y assuming some θ"). If we regard

$p(y|\theta)$ as a function of θ for some fixed set of y, we then have the likelihood function $\mathcal{L}(\theta|y)$, defined:

$$\mathcal{L}(\theta|y) = \prod_i^n p(y_i|\theta) \tag{4.7}$$

for independent and identically-distributed data.

Thus far, we have assumed a simple data generating process that is characterised by a single parameter θ. In general, constraining the posterior density distribution is a tedious task, except in some special instances where the posterior takes on the same form as the prior (a conjugate distribution, used in Figure 4.4). Additionally, more often than not, we wish to constrain a *set* of parameters, such that θ represents a vector of parameters to estimate. For this reason, most Bayesian analysis in practice is performed using Monte Carlo methods.

The Bayesian methods employed in this thesis have made use of the software STAN (Stan Development Team, 2018), which uses the No U-Turn Sampler (NUTS) variation of Hamiltonian Monte Carlo (HMC). In numerically sampling an arbitrary target distribution[1] $\pi(\theta)$, we are required to sample from the *typical set*, or those regions of the D-dimensional target distribution which provide the dominant contribution to an expectation value. Simplistically, this may be achieved using any Markov generating process, such as the widely-used random walk Metropolis-Hastings algorithm, albeit at the expense of inefficient and biased sampling. The Metropolis-Hastings algorithm generates from a starting position q a proposal for the next position (or state) q', with proposals accepted with some probability $Q(q'|q)$, where Q is often chosen to be the normal distribution, and rejected otherwise. The Markov generating process forms a Markov chain, where the next proposed state depends only on the current state, and once in equilibrium:

$$\lim_{n \to \infty} P(\theta_n = \theta) = \pi(\theta). \tag{4.8}$$

With parallel processing on modern computer architecture, we run multiple chains in parallel to sample the target distribution. Sampling with increased efficiency can be achieved by using information about the geometry of the typical set of the target distribution. This is the premise of HMC – an excellent overview of which is given in Betancourt (2018) – which uses Hamiltonian mechanics, autodifferentiation, and symplectic integration to efficiently sample the target distribution. The advantage of efficient sampling becomes particularly apparent as we consider models with high numbers of dimensions, such as hierarchical models.

---

[1]Generally, the target distribution is the posterior distribution, in the context of Bayesian analysis.
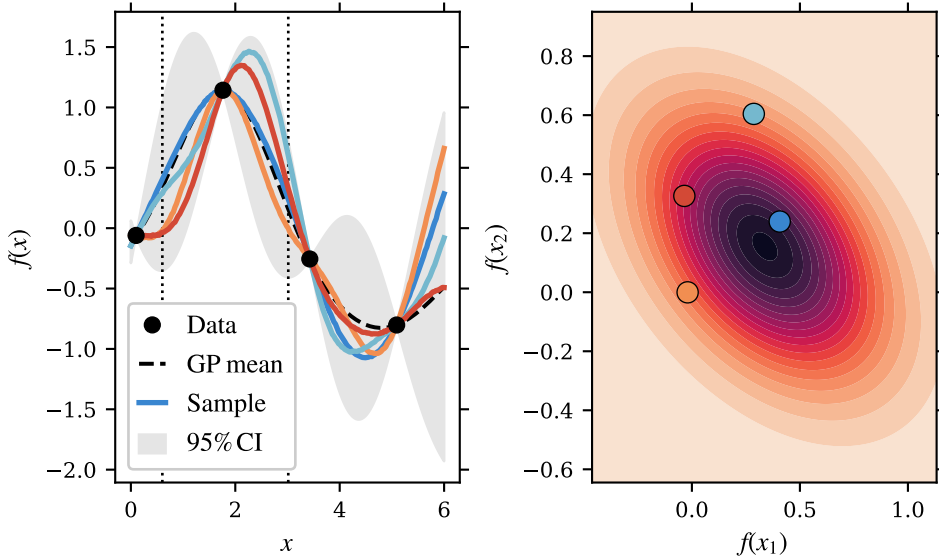
Figure 4.5: Example Gaussian process regressor. *Left panel*: Regression is performed on data, with four example draws shown in coloured lines, to obtain a confidence interval estimate for the best-fit regression model of the data. The mean regression line is shown as the dashed black line. The two vertical dotted lines indicate two arbitrarily-chosen x values from the regression. *Right panel*: The joint distribution of the two vertical dotted lines in the left panel is shown, with contours indicating the probability density function. The coloured points indicate the four example draws of the distribution corresponding to the same coloured lines in the left panel. A 2D space is shown for visual clarity, however the full distribution is of a higher order dimension D corresponding to the number of sampled points in the left panel.

Models of a data generating process often contain information on different levels of the observation units: for example, the projected density of a galaxy from different viewing angles. In the case that there are groups of data that are subsets of some global population, and these subsets are exchangeable in the sense that there is no distinguishing order to them, we can model each subset of data y_j as having its own value of the model parameter θ_j. The distribution of all values of θ_j can then be modelled as samples from a global distribution, termed a hyperdistribution, with hyperparameter(s) ϕ: this is where the hierarchical nature of the model is introduced. A prior distribution can then be assigned to the hyperparameter ϕ, $p(\phi)$. As such,
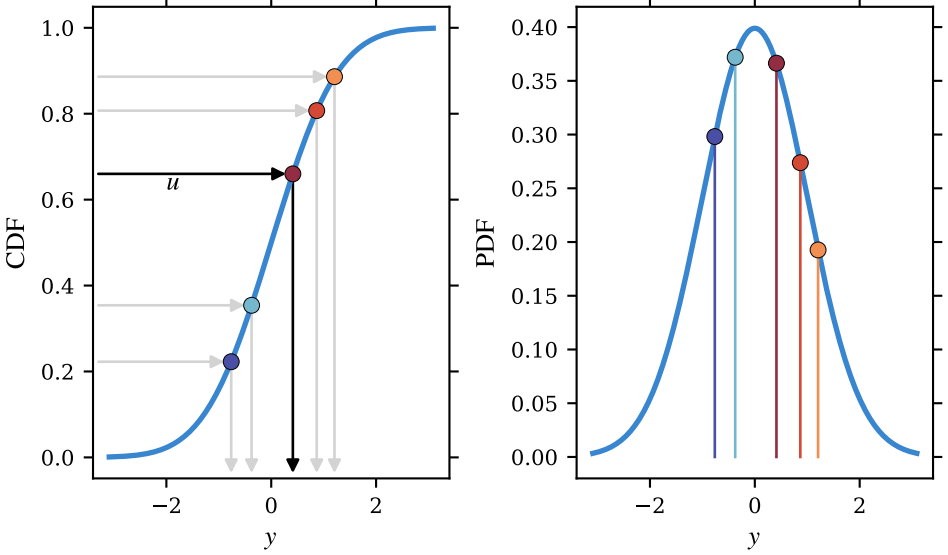
Figure 4.6: Representation of inverse transformation sampling. Given some data, we may construct an empirical cumulative distribution function (CDF), which has support on $[0, 1]$. Drawing a random sample u from a standard uniform distribution, we may transform to the distribution we wish to sample from through the mapping $\text{ICDF}(u) \to y$, where ICDF is the inverse CDF. Repeating this process many times allows us to build up a distribution of the quantity of interest, y.

Equation 4.6 then takes the form:

$$p(\theta, \phi|y) \propto p(\phi)p(\theta|\phi)\mathcal{L}(\theta, \phi|y). \tag{4.9}$$

Herein, we will revert to our original notation of θ representing the parameters to be constrained for simplicity, however the discussions naturally extend to hierarchical models.

Care must be taken when selecting the prior distribution for θ, and throughout this work we employ weakly-informative priors, as well as perform post-sampling prior sensitivity checks (Kallioinen et al., 2024). Prior distributions are chosen to have a mode centred about expected beliefs on the parameter, but are simultaneously broad, and encode knowledge of the known physics governing the parameter. As an example, when fitting a projected stellar mass density profile (Equation 2.16), we may assign as a weakly-informative prior for the effective radius R_e a Rayleigh distribution (which describes the magnitude of a two-dimensional vector) that has a peak around

7 kpc, but is broad enough that the interquartile range[2] of the distribution spans from 3 kpc to 12 kpc. In this way, we encode prior information from previous studies about the parameters, but do not constrain it so tightly that we recover the prior distribution in the posterior sampling.

After constraining the posterior distribution, we often wish to sample from it, particularly to simulate new data $\tilde{y}$ that is not part of the observed data y that the posterior was conditioned on. Mathematically,

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)\mathrm{d}\theta. \tag{4.10}$$

Such samples $\tilde{y}$ are said to come from the *posterior predictive* distribution.

An instance where sampling from the posterior predictive distribution is highly desirable is in the use of Gaussian process regression for interpolation, employed in Rawlings et al. (2025a), and is the infinitely-dimensional generalisation of a Gaussian distribution (e.g. Gelman et al., 2015). Given data x, the regressor aims to fit the data with a regression model $\mu(x) \sim \mathrm{GP}(m, k)$, where m is the mean function, and k is the covariance function. An example Gaussian process regressor is shown in the left panel of Figure 4.5, trained on four observed points shown in black. Different forms of the covariance function exist; the one used in this thesis is the exponentiated quadratic covariance which takes two unknown parameters τ (the function magnitude) and l (the smoothness) to be inferred through the Bayesian framework:

$$k(x, x') = \tau^2 \exp\left(-\frac{|x - x'|^2}{2l^2}\right). \tag{4.11}$$

A defining feature of a Gaussian process, namely that any finite-dimensional marginal distribution is also Gaussian, is shown in the right panel of Figure 4.5 for two arbitrarily chosen x values. Gaussian process regressors have the advantage of being a non-parametric model that can capture complex relationships between a set of independent variables on response variables. Being a probabilistic fit to the data, Gaussian processes also encode uncertainty that can be assessed using standard Bayesian techniques.

The final statistical technique to be discussed is transformation sampling, which lends itself nicely to the Monte Carlo nature of Bayesian posterior sampling. Often, we wish to obtain a distribution of some unknown quantity y, for which we do not have direct samples of. Let us assume we have access to random samples u, and that

---

[2]The interquartile range describes the central 50% of the data, i.e. from the 25th to the 75th percentile.

we have a mapping $f : u \mapsto y$. We can then 'push through' our sample of u variates to obtain a distribution of y: this process is termed transformation sampling. A common application of this procedure is to draw samples from standard distributions, such as the normal distribution, shown in Figure 4.6. We take standard uniformly-sampled variates $u \in [0, 1]$, and with knowledge of the inverse cumulative distribution function (ICDF) of the normal distribution, we can obtain normally-distributed variates y by applying $\mathrm{ICDF}(u)$. In this instance, as we have used the *inverse* CDF, we have applied inverse transformation sampling. Transformation sampling is a powerful technique that allows for generating of random samples given some existing variates, that leverages the Monte Carlo sampled draws obtained through Bayesian analysis. In particular, non-linear error propagation becomes trivial with transformation sampling, compared to the much more cumbersome method of analytically deriving the error propagation of a variable common to frequentist methods.

5 Summary of the publications

The thesis consists of five journal publications:

- **Paper I:** Rawlings, Mannerkoski, Johansson, & Naab (2023). 'Reviving stochasticity: uncertainty in SMBH binary eccentricity is unavoidable'.

- **Paper II:** Rawlings, Keitaanranta, Mattero, Soininen, Wright, Kallioinen, Liao, Rantala, Johansson, Naab, & Irodotou (2025b). 'Identifying supermassive black hole recoil in elliptical galaxies'.

- **Paper III:** Rawlings, Johansson, Naab, Rantala, Thomas, & Neureiter (2025a). 'Caught in the act: detections of recoiling supermassive black holes from simulations'.

- **Paper IV:** Mannerkoski, Rawlings, Johansson, Naab, Rantala, Springel, Irodotou, & Liao (2023). 'KETJU – resolving small-scale supermassive black hole dynamics in GADGET-4'.

- **Paper V:** Rantala, Rawlings, Naab, Thomas, & Johansson (2024). 'The supermassive black hole merger-driven evolution of high-redshift red nuggets into present-day cored early-type galaxies'.

The papers are summarised below. The author's contribution to each paper is described at the end of each section.

5.1 Paper I

In Paper I we study supermassive black hole binary eccentricity of equal-mass galaxy mergers in N-body simulations with KETJU. The simulations follow the same galaxy initial conditions as Nasim et al. (2020) – the modelling of the central bulge of an ETG – allowing for a direct comparison between the studies. A merger suite of varying $M_\bullet/m_\star$ with initial eccentricity $e_0 = 0.99$ is included in addition to the

$e_0 = 0.90$ case presented in Nasim et al. (2020). For these high eccentricity galactic merger orbits (which are well motivated through cosmological simulations), the hard binary eccentricity e_h is found to be a non-linear function of the impact parameter in the SMBH orbit during the final, nearly-radial close encounter between the SMBHs before they form a bound binary. This mapping between the deflection angle and the binary eccentricity has no apparent resolution dependence in our simulations spanning the resolution range of $N_\star = 1 \times 10^5$–8×10^6 particles per galaxy. The mapping is also captured using a simple model with an analytic two-dimensional potential, indicating that it is driven by the interplay between a smooth asymmetric stellar background potential and dynamical friction acting on the SMBHs. Due to the non-linearity of this mapping, in eccentric major merger configurations small, parsec-scale variations in the merger orbit can result in SMBH binary eccentricities varying in nearly the full possible range between $e = 0$ and $e = 1$. In idealised simulations, such variations are caused by finite resolution effects, and convergence of the binary eccentricity may be achieved with increasing resolution for a finite domain of the impact parameter space. Our results indicate that the distribution of these effectively random eccentricities can be studied using even moderate resolution simulations.

5.2 Paper II

In Paper II we study stellar core growth in simulations of merging massive ($M_\star > 10^{11}\,\mathrm{M_\odot}$) elliptical galaxies by a supermassive black hole displaced by gravitational wave-induced recoil velocity. We run one simulation to obtain an SMBH merger with KETJU, and then spawn from this 31 clone simulations just prior to the SMBH merger, where in each simulation a unique SMBH recoil velocity is prescribed, in the range $0\,\mathrm{km\,s^{-1}}$ to $1800\,\mathrm{km\,s^{-1}}$ in $60\,\mathrm{km\,s^{-1}}$ increments, with one additional recoil at $2000\,\mathrm{km\,s^{-1}}$ – above the escape velocity of the galaxy. The simulations are run until either the kicked SMBH settles at the centre of the galaxy, or three dynamical times have elapsed. We fit a core-Sérsic profile with Bayesian hierarchical modelling to the projected stellar mass density of those galaxies for which the SMBH settled, and find that the core radius originally formed by SMBH binary scouring can grow by a factor of 2-3 when the recoil velocity exceeds ~ 50 per cent of the central escape velocity, and the mass deficit grows by up to a factor of ~ 4. Using Bayesian inference we predict the distribution of stellar core sizes formed through this process to peak at $\sim 1\,\mathrm{kpc}$. Additionally, we perform an orbital decomposition of stellar particles within the core, revealing that radial orbits dominate over tube orbits when

the recoil velocity exceeds the velocity dispersion of the core, whereas tube orbits dominate for the lowest recoil kicks. A change in orbital structure is reflected in the anisotropy parameter, with a central tangential bias present only for recoil velocities less than the local stellar velocity dispersion. Emulating current integral field unit observations of the stellar line-of-sight velocity distribution, we uncover a distinct signature in the Gauss-Hermite symmetric deviation coefficient h_4 that uniquely constrains the core size due to binary scouring. This signature is insensitive to the later evolution of the stellar mass distribution due to SMBH recoil, providing a novel method to estimate the increase in an ETG's core size due to GW-induced recoil velocity alone. The implications of this study is that the majority of massive ETGs most likely experienced recoil velocities less than the stellar velocity dispersion of the core.

5.3 Paper III

In Paper III, we study the detectability of the recoiling supermassive black holes in Paper II as they are settling, as opposed to inferring their history as was done in Paper II. We find that the ejected SMBHs carry clusters of bound stellar material (black hole recoil clusters, BRCs) with masses in the range of $10^6 \lesssim M_{\mathrm{BRC}} \lesssim 10^7 \, \mathrm{M_\odot}$ and sizes of several $10 \, \mathrm{pc}$ during their excursion from the centre following the recoil kick. For recoil velocities up to 60% of the galaxy escape velocity, the BRCs are detectable in mock photometric images at a Euclid-like resolution up to redshift $z \sim 1.0$. We use Gaussian process regression to model the relationship between recoil velocity and apocentre distance, and Monte Carlo sampling to determine the apocentre distribution. We then use transformation sampling to determine the minimum angle from the line-of-sight axis a recoiling SMBH can have and still be detectable. We predict that in $\sim 20\%$ of instances the BRCs are photometrically detectable, most likely for kicks with SMBH apocenters less than the galaxy effective radius. We also find that BRCs occupy distinct regions in the stellar mass/velocity dispersion vs. size relations of known star clusters and galaxies. An enhanced velocity dispersion in excess of $\sigma \sim 600 \, \mathrm{km \, s^{-1}}$ coinciding with the SMBH position provides the best evidence for an SMBH-hosting stellar system, effectively distinguishing BRCs from other faint stellar systems. Such detections will be most readily-achievable with future Extremely Large Telescope (ELT) instruments such as HARMONI and MICADO. BRCs are thus promising candidates to observe the aftermath of the yet-undetected mergers of the most massive SMBHs and we estimate that up to 8000 BRCs might be observable below $z \lesssim 0.6$ with large-scale photometric surveys such

as Euclid, and upcoming high-resolution imaging and spectroscopy with the ELT.

5.4 Paper IV

In Paper IV we present the public version of the KETJU supermassive black hole dynamics module, as implemented into GADGET-4. We present tests of the code showing that it correctly captures, at sufficient mass resolution, the sinking driven by dynamical friction and binary hardening driven by stellar scattering. We also present an example application demonstrating how the code can be applied to study the dynamics of SMBHs in mergers of multiple galaxies and the effect they have on the properties of the surrounding galaxy. We envision that by making the presented KETJU SMBH dynamics module public, it can also be straightforwardly incorporated into other codes similar to GADGET-4 (such as AREPO), which would allow coupling small-scale SMBH dynamics to the rich variety of galactic physics models that exist in the literature.

5.5 Paper V

In Paper V, we use sequences of major and minor merger N-body simulations starting with compact spherical and disc-like progenitor models to investigate the impact of supermassive black holes on the evolution of progenitor galaxies to the observed 'red nugget' galaxies – very compact ($R_e \lesssim 1\,\mathrm{kpc}$) massive quiescent galaxies. With KETJU we accurately follow the collisional interaction of the SMBHs with the nearby stellar population and the collisionless evolution of the galaxies and their dark matter halos. We show that only models including SMBHs can simultaneously explain the formation of low-density cores up to sizes of $R_b \sim 1.3\,\mathrm{kpc}$ with mass deficits in the observed range and the rapid half mass evolution. In addition, the orbital structure in the core region (tangentially biased orbits) is consistent with observation-based results for local cored ETGs. The displacement of stars by the SMBHs boost the half mass evolution by up to a factor of two and even fast rotating progenitors (compact quiescent disks) lose their rotational support after 6–8 mergers. We conclude that the presence of SMBHs is required for merger driven evolution models of high redshift red nuggets into local ETGs.

5.6 Author's contributions

In Paper I, I created, ran, and analysed all of the simulations for the study. Matias Mannerkoski developed the simple analytical potential model that we compared our results against in the 'infinte resolution' case. I completed the majority of the writing, with contribution from the other authors.

In Paper II, I created and ran all of the simulations. I developed the Bayesian modelling required to fit the density profiles and provide estimates for core radii distributions. I also performed the orbital decomposition, and wrote the analysis code to fit arbitrary-orders of an LOSVD, parallelising the routine over Voronoi cells in mock IFU maps. Atte Keitaanranta identified when the SMBHs had settled, and when during their excursion the majority of core growth occurred. Max Mattero determined the velocity anisotropy parameter for each simulation, and Sonja Soininen determined the mass deficit for each simulation. The majority of the writing was done by myself, with contributions to the original text by Atte Keitaanranta, Max Mattero, and Sonja Soininen, and later text contributions from the other authors. This project was largely student-led, where I guided the work of the other students (who formed the first three coauthors).

In Paper III, I used many of the analysis tools developed in Paper II, with the additional development of Gaussian process regression within a Bayesian framework, and generating mock observations motivated by present and future instruments (e.g. Euclid, MICADO, etc.). The majority of the writing was done by myself, with contributions from the other authors.

In Paper IV, my main contribution was the design, running, and analysis of N-body simulations to study the convergence of SMBH binary hardening rate H as a function of mass resolution. A simple permutation test was used to assess convergence of H, and the contributing number of stars in the SMBH binary loss-cone was assessed. I contributed original text to accompany this analysis. The majority of the analysis and writing was carried out by Matias Mannerkoski.

In Paper V, my main contribution was the orbital analysis to determine the effect of multiple mergers on the interior orbital structure of the remnant galaxy. I also contributed original text to accompany this analysis. The majority of the analysis and writing was carried out by Antti Rantala.

6 Concluding remarks

The work presented in this thesis covers two main areas of galaxy simulation study: a deeper foray into the inherent uncertainty of supermassive black hole orbital parameters, and the first systematic studies of GW-induced SMBH recoil with KETJU. An overarching theme of this thesis has been on *statistical* significance, by using large numbers of simulations and rigorous analysis techniques to ensure that the conclusions drawn are robust and quantify the inherent level of uncertainty in our modelling process. As a result, the main aim of the thesis has been on understanding the outcome of our numerical experiments, rather than developing the simulation code itself – an often-overlooked aspect when there is perhaps more general excitement in developing new models.

As with most scientific explorations, the work completed here does not mark the end of our journey to understand galaxy evolution, and the role of SMBHs in that, but rather is a necessary foundation for answering a plethora of further questions that we have just started to scratch the surface of.

From our study on the stochasticity of SMBH binary eccentricity, a key finding has been the sensitivity of our numerical experiments to the particular initial conditions used. As a consequence, a shift in group mentality has occurred to designing numerical experiments with multiple realisations of the simulation, as opposed to assuming that one simulation captures 'the' answer – almost any physical quantity we are interested in investigating has uncertainty in it, and each simulation probes just one point in that distribution. The case of SMBH binary eccentricity is of particular importance, as we have seen that eccentricity is arguably the strongest driver of GW-induced coalescence. An uncertainty in eccentricity has the knock-on effect of an uncertainty for SMBH binary merger timescales and merger rates, confounding predictions necessary for gravitational wave detectors, namely ongoing pulsar timing arrays and the upcoming Laser Interferometer Space Antenna (LISA). This presents challenges that will need to be overcome with new and innovative solutions. Whilst we can efficiently explore merger stochasticity with targeted, isolated simulations such as those presented in this thesis, we are effectively considering a

conditional distribution, where we investigate the distribution of merger timescales *conditional* on a given set of initial conditions. To explore the full marginal distribution of merger timescales, we require knowledge of the distribution of SMBH masses, SMBH binary mass ratios, galaxy triaxialities, amongst other variables, available through full-volume cosmological simulations. Studying stochasticity for a cosmological volume however is a challenge owing to the computational expense of such simulations, and is an avenue for future investigation.

A similar argument can be made for the SMBH recoil studies in this thesis, which whilst explored *typical* massive galaxy merger remnants with *typical* SMBHs, did not explore the full distribution of either. These explorations were however outside the scope of the investigation, as we were interested in the effect of a recoiling SMBH on a galaxy's stellar mass distribution. From the presented thesis work, we found that recoils $\lesssim 50\%$ of the galaxy's escape velocity were sufficient in producing noticeable effects in a galaxy's stellar mass distribution, provided that the SMBH exited the binary-scoured stellar core. Already with mild recoil velocities of the order of $300\,\mathrm{km\,s^{-1}} \simeq 0.15 v_{\mathrm{esc}}$ a significantly-altered stellar phase space was produced. As a result, an interesting future study would be to turn away from *extreme* recoil events (such as those for equal-mass galaxy mergers) to *most-probable* recoil events. Many studies agree on the recoil velocity distribution following a right-skewed distribution, whereby the peak recoil velocity is at lower rather than higher values. Future work considering the instance of unequal-mass mergers, and in particular minor mergers – which are by definition more probable to occur compared to equal-mass or major mergers – would be of relevance here. Assuming that the SMBH mass is proportional to the galaxy stellar mass, recoil events resulting in kicks of $\sim 300\,\mathrm{km\,s^{-1}}$ are highly probable for minor mergers, and likely to be detectable by the same methods explored in this thesis. How the distribution of core radii, for example, would change as a result of these most-probable mergers would be of particular interest when attempting to constrain the dynamical history of SMBHs in nearby local massive galaxies.

In conclusion, the work presented in this thesis has furthered our understanding of SMBH dynamics by a small yet meaningful step. The avenues of further research, as briefly discussed here, are plentiful – highlighting the unfortunate but amusing paradox of the more one learns, the more one realises there is left to learn.

Bibliography

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102

Abbott, R., Abbott, T. D., Acernese, F., et al. 2023, Physical Review X, 13, 011048

Adams, D. 1979, The Hitchhiker's Guide to the Galaxy, first edition edn. (London: Pan Books)

Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023, ApJ, 951, L8

Amaro-Seoane, P., Gair, J. R., Freitag, M., et al. 2007, Classical and Quantum Gravity, 24, R113

Armitage, P. J., & Natarajan, P. 2005, ApJ, 634, 921

Arnold, R., de Zeeuw, P. T., & Hunter, C. 1994, MNRAS, 271, 924

Attard, K., Gualandris, A., Read, J. I., & Dehnen, W. 2024, MNRAS, 529, 2150

Bacon, R., Emsellem, E., Copin, Y., & Monnet, G. 2000, in Astronomical Society of the Pacific Conference Series, Vol. 195, Imaging the Universe in Three Dimensions, ed. W. van Breugel & J. Bland-Hawthorn, 173

Bacon, R., Adam, G., Baranne, A., et al. 1995, A&AS, 113, 347

Bacon, R., Copin, Y., Monnet, G., et al. 2001, MNRAS, 326, 23

Bailin, J., & Steinmetz, M. 2005, ApJ, 627, 647

Barnes, J., & Hut, P. 1986, Nature, 324, 446

Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307

Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, MNRAS, 370, 289

Bekenstein, J. D. 1973, ApJ, 183, 657

Bibliography

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Berczik, P., Merritt, D., Spurzem, R., & Bischof, H.-P. 2006, ApJ, 642, L21

Berentzen, I., Preto, M., Berczik, P., Merritt, D., & Spurzem, R. 2009, ApJ, 695, 455

Betancourt, M. 2018, A Conceptual Introduction to Hamiltonian Monte Carlo, arXiv

Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press)

Bisigello, L., Kuchner, U., Conselice, C. J., et al. 2020, MNRAS, 494, 2337

Bondi, H. 1952, MNRAS, 112, 195

Bondi, H., & Hoyle, F. 1944, MNRAS, 104, 273

Bonfini, P., & Graham, A. W. 2016, ApJ, 829, 81

Bortolas, E., Gualandris, A., Dotti, M., Spera, M., & Mapelli, M. 2016, MNRAS, 461, 1023

Bromley, J. M., Somerville, R. S., & Fabian, A. C. 2004, MNRAS, 350, 456

Bulirsch, R., & Stoer, J. 1966, Numerische Mathematik, 8, 1

Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7

Buson, L. M., Sadler, E. M., Zeilinger, W. W., et al. 1993, A&A, 280, 409

Callegari, S., Kazantzidis, S., Mayer, L., et al. 2011, ApJ, 729, 85

Campanelli, M., Lousto, C., Zlochower, Y., & Merritt, D. 2007, ApJ, 659, L5

Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 416, 1680

Carpintero, D. D., & Aguilar, L. A. 1998, MNRAS, 298, 1

Carroll, B., & Ostlie, D. 2017, An Introduction to Modern Astrophysics, 2nd edn. (Shaftesbury Road, Cambridge CB2 8BS, UK: Cambridge University Press)

Ceverino, D., Sánchez Almeida, J., Muñoz Tuñón, C., et al. 2016, MNRAS, 457, 2605

Chandrasekhar, S. 1943, ApJ, 97, 255

Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. 2005, MNRAS, 358, 363

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486

Cresci, G., Mannucci, F., Maiolino, R., et al. 2010, Nature, 467, 811

Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, MNRAS, 421, 872

Croom, S. M., Owers, M. S., Scott, N., et al. 2021, MNRAS, 505, 991

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

de Zeeuw, P. T. 1996, in Gravitational dynamics, ed. O. Lahav, E. Terlevich, & R. J. Terlevich, 1

de Zeeuw, P. T., Bureau, M., Emsellem, E., et al. 2002, MNRAS, 329, 513

Dehnen, W. 1993, MNRAS, 265, 250

Dehnen, W. 2014, Computational Astrophysics and Cosmology, 1, 1

Dekel, A., & Silk, J. 1986, ApJ, 303, 39

Dotti, M., Volonteri, M., Perego, A., et al. 2010, MNRAS, 402, 682

Eddington, A. S. 1916, MNRAS, 76, 572

Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, MNRAS, 379, 401

Emsellem, E., Cappellari, M., Krajnović, D., et al. 2011, MNRAS, 414, 888

EPTA Collaboration, InPTA Collaboration, Antoniadis, J., et al. 2024, A&A, 685, A94

Fabbiano, G., Gioia, I. M., & Trinchieri, G. 1989, ApJ, 347, 127

Favata, M., Hughes, S. A., & Holz, D. E. 2004, ApJ, 607, L5

Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9

Finn, L. S., & Thorne, K. S. 2000, Phys. Rev. D, 62, 124021

Frigo, M., Naab, T., Rantala, A., et al. 2021, MNRAS, 508, 4610

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41

Gelman, A., Carlin, J. B., Stern, H. S., et al. 2015, Bayesian Data Analysis, 3rd edn. (Chapman and Hall/CRC)

Gerhard, O. E. 1983, MNRAS, 203, 19P

González, J. A., Hannam, M., Sperhake, U., Brügmann, B., & Husa, S. 2007a, Phys. Rev. Lett., 98, 231101

González, J. A., Sperhake, U., Brügmann, B., Hannam, M., & Husa, S. 2007b, Phys. Rev. Lett., 98, 091101

Gragg, W. B. 1965, SIAM Journal on Numerical Analysis, 2, 384

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

Gualandris, A., & Merritt, D. 2012, ApJ, 744, 74

Gualandris, A., Read, J. I., Dehnen, W., & Bortolas, E. 2017, MNRAS, 464, 2301

Haehnelt, M. G., & Kauffmann, G. 2000, MNRAS, 318, L35

Haehnelt, M. G., & Rees, M. J. 1993, MNRAS, 263, 168

Hairer, E., Nørsett, S. P., & Wanner, G. 1993, Springer Series in Computational Mathematics, Vol. 8, Solving Ordinary Differential Equations I: Nonstiff Problems, 2nd edn. (Springer)

Halliday, C., Davies, R. L., Kuntschner, H., et al. 2001, MNRAS, 326, 473

Harfst, S., Gualandris, A., Merritt, D., & Mikkola, S. 2008, MNRAS, 389, 2

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

Hernquist, L. 1990, ApJ, 356, 359

Hernquist, L., & Ostriker, J. P. 1992, ApJ, 386, 375

Hills, J. G. 1983, AJ, 88, 1269

Hills, J. G., & Fullerton, L. W. 1980, AJ, 85, 1281

Hoyle, F., & Lyttleton, R. A. 1939, Proceedings of the Cambridge Philosophical Society, 35, 405

Hubble, E. P. 1926, The Astrophysical Journal, 64, 321

Hut, P., & Heggie, D. C. 2002, Journal of Statistical Physics, 109, 1017

Iwasawa, M., Portegies Zwart, S., & Makino, J. 2015, Computational Astrophysics and Cosmology, 2, 6

Jeans, J. H. 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Jesseit, R., Naab, T., & Burkert, A. 2005, MNRAS, 360, 1185

Kallioinen, N., Paananen, T., Bürkner, P.-C., & Vehtari, A. 2024, Statistics and Computing, 34, 57

Kapferer, W., Sluka, C., Schindler, S., Ferrari, C., & Ziegler, B. 2009, A&A, 499, 87

Karl, S. J., Aarseth, S. J., Naab, T., Haehnelt, M. G., & Spurzem, R. 2015, MNRAS, 452, 2337

Khan, F. M., Fiacconi, D., Mayer, L., Berczik, P., & Just, A. 2016, ApJ, 828, 73

Khochfar, S., & Burkert, A. 2006, A&A, 445, 403

Khonji, N., Gualandris, A., Read, J. I., & Dehnen, W. 2024, ApJ, 974, 204

Krajnović, D., Bacon, R., Cappellari, M., et al. 2008, MNRAS, 390, 93

Krajnović, D., Emsellem, E., Cappellari, M., et al. 2011, MNRAS, 414, 2923

Kronberger, T., Kapferer, W., Ferrari, C., Unterguggenberger, S., & Schindler, S. 2008, A&A, 481, 337

Lauer, T. R., Ajhar, E. A., Byun, Y. I., et al. 1995, AJ, 110, 2622

Liao, S., Irodotou, D., Johansson, P. H., et al. 2024a, MNRAS, 528, 5080

Liao, S., Irodotou, D., Johansson, P. H., et al. 2024b, MNRAS, 530, 4058

Liao, S., Johansson, P. H., Mannerkoski, M., et al. 2023, MNRAS, 520, 4463

BIBLIOGRAPHY

Licquia, T. C., & Newman, J. A. 2015, ApJ, 806, 96

Lousto, C. O., Nakano, H., Zlochower, Y., & Campanelli, M. 2010, Phys. Rev. D, 81, 084023

Lousto, C. O., Zlochower, Y., Dotti, M., & Volonteri, M. 2012, Phys. Rev. D, 85, 084015

Ludlow, A. D., Navarro, J. F., Angulo, R. E., et al. 2014, MNRAS, 441, 378

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

Mannerkoski, M., Johansson, P. H., Pihajoki, P., Rantala, A., & Naab, T. 2019, ApJ, 887, 35

Mannerkoski, M., Johansson, P. H., Rantala, A., et al. 2022, ApJ, 929, 167

Mannerkoski, M., Rawlings, A., Johansson, P. H., et al. 2023, MNRAS, 524, 4062

Maraston, C., Daddi, E., Renzini, A., et al. 2006, ApJ, 652, 85

Maraston, C., Strömbäck, G., Thomas, D., Wake, D. A., & Nichol, R. C. 2009, MNRAS, 394, L107

Maraston, C., Pforr, J., Henriques, B. M., et al. 2013, MNRAS, 435, 2764

Mathews, W. G., & Brighenti, F. 2003, ApJ, 590, L5

Matsubayashi, T., Makino, J., & Ebisuzaki, T. 2007, ApJ, 656, 879

McConnell, N. J., & Ma, C.-P. 2013, ApJ, 764, 184

Merritt, D. 1985, AJ, 90, 1027

Merritt, D. 2001, ApJ, 556, 245

Merritt, D. 2006, ApJ, 648, 976

Merritt, D. 2013, Dynamics and Evolution of Galactic Nuclei (Princeton University Press)

Merritt, D., Milosavljević, M., Favata, M., Hughes, S. A., & Holz, D. E. 2004, ApJ, 607, L9

Merritt, D., & Vasiliev, E. 2011, ApJ, 726, 61

Milosavljević, M., & Merritt, D. 2001, ApJ, 563, 34

Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution (The Edinburgh Building Building, Cambridge CB2 8RU, UK: Cambridge University Press)

Moore, C. J., Cole, R. H., & Berry, C. P. L. 2015, Classical and Quantum Gravity, 32, 015014

Mulder, W. A. 1983, A&A, 117, 9

Nasim, I., Gualandris, A., Read, J., et al. 2020, MNRAS, 497, 739

Nasim, I. T., Gualandris, A., Read, J. I., et al. 2021, MNRAS, 502, 4794

Nasim, I. T., Petrovich, C., Nasim, A., Dosopoulou, F., & Antonini, F. 2021, MNRAS, 503, 498

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Neureiter, B., Thomas, J., Saglia, R., et al. 2021, MNRAS, 500, 1437

Novikov, I. D., & Thorne, K. S. 1973, in Black Holes (Les Astres Occlus), ed. C. Dewitt & B. S. Dewitt, 343–450

Osipkov, L. P. 1979, Pisma v Astronomicheskii Zhurnal, 5, 77

Peters, P. C. 1964, Physical Review, 136, 1224

Peters, P. C., & Mathews, J. 1963, Physical Review, 131, 435

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Poisson, E., & Will, C. M. 2014, Gravity (The Edinburgh Building Building, Cambridge CB2 8RU, UK: Cambridge University Press)

Prim, R. C. 1957, Bell System Technical Journal, 36, 1389

Pringle, J. E., & Rees, M. J. 1972, A&A, 21, 1

Prugniel, P., & Simien, F. 1996, A&A, 309, 749

Prugniel, P., & Simien, F. 1997, A&A, 321, 111

Quinlan, G. D. 1996, New A, 1, 35

Rantala, A., Johansson, P. H., Naab, T., Thomas, J., & Frigo, M. 2018, ApJ, 864, 113

Rantala, A., Pihajoki, P., Johansson, P. H., et al. 2017, ApJ, 840, 53

Rantala, A., Pihajoki, P., Mannerkoski, M., Johansson, P. H., & Naab, T. 2020, MNRAS, 492, 4131

Rantala, A., Rawlings, A., Naab, T., Thomas, J., & Johansson, P. H. 2024, MNRAS, 535, 1202

Rawlings, A., Johansson, P. H., Naab, T., et al. 2025a, ApJ, 991, 83

Rawlings, A., Mannerkoski, M., Johansson, P. H., & Naab, T. 2023, MNRAS, 526, 2688

Rawlings, A., Foster, C., van de Sande, J., et al. 2020, MNRAS, 491, 324

Rawlings, A., Keitaanranta, A., Mattero, M., et al. 2025b, MNRAS, 537, 3421

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Roberts, M. S., Hogg, D. E., Bregman, J. N., Forman, W. R., & Jones, C. 1991, ApJS, 75, 751

Rodrigo, C., & Solano, E. 2020, in XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society, 182

Röttgers, B., Naab, T., & Oser, L. 2014, MNRAS, 445, 1065

Sahu, N., Graham, A. W., & Davis, B. L. 2020, ApJ, 903, 97

Salpeter, E. E. 1964, ApJ, 140, 796

Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012, A&A, 538, A8

Sánchez Almeida, J., Muñoz-Tuñón, C., Elmegreen, D. M., Elmegreen, B. G., & Méndez-Abreu, J. 2013, ApJ, 767, 74

Sawala, T., Delhomelle, J., Deason, A. J., et al. 2025, Nature Astronomy

Schmidt, B. P., Suntzeff, N. B., Phillips, M. M., et al. 1998, ApJ, 507, 46

Schwarzschild, M. 1979, ApJ, 232, 236

Sersic, J. L. 1968, Atlas de Galaxias Australes

Sesana, A., Haardt, F., & Madau, P. 2006, ApJ, 651, 392

Sesana, A., & Khan, F. M. 2015, MNRAS, 454, L66

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Silk, J. 2005, MNRAS, 364, 1337

Spitzer, L. 1987, Dynamical evolution of globular clusters

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Pakmor, R., Zier, O., & Reinecke, M. 2021, MNRAS, 506, 2871

Stan Development Team. 2018, The Stan Core Library

Statler, T. S. 1991, ApJ, 382, L11

Sureshkumar, U., Durkalec, A., Pollo, A., et al. 2021, A&A, 653, A35

Taniguchi, Y., Ikeuchi, S., & Shioya, Y. 1999, ApJ, 514, L9

The LIGO Scientific Collaboration, the Virgo Collaboration, & the KAGRA Collaboration. 2025, arXiv e-prints, arXiv:2507.08219

Tichy, W., & Marronetti, P. 2007, Phys. Rev. D, 76, 061502

Umemura, M., Loeb, A., & Turner, E. L. 1993, ApJ, 419, 459

Vallenari, A. 2022, in EAS2022, European Astronomical Society Annual Meeting, 2358

Valluri, M., & Merritt, D. 2000, in The Chaotic Universe, ed. V. G. Gurzadyan & R. Ruffini, 229–246

van de Sande, J., Bland-Hawthorn, J., Fogarty, L. M. R., et al. 2017, ApJ, 835, 104

van den Bosch, R. C. E. 2016, ApJ, 831, 134

van der Marel, R. P., & Franx, M. 1993, ApJ, 407, 525

Volonteri, M., Haardt, F., & Madau, P. 2003a, ApJ, 582, 559

Volonteri, M., Madau, P., & Haardt, F. 2003b, ApJ, 593, 661

Volonteri, M., & Rees, M. J. 2005, ApJ, 633, 624

Wise, M. W., & Silva, D. R. 1996, ApJ, 461, 155

Xu, H., Chen, S., Guo, Y., et al. 2023, Research in Astronomy and Astrophysics, 23, 075024

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Zhang, M., & Magorrian, J. 2008, MNRAS, 387, 1719

Zic, A., Reardon, D. J., Kapur, A., et al. 2023, PASA, 40, e049

Zlochower, Y., & Lousto, C. O. 2015, Phys. Rev. D, 92, 024022