Realizing the Full Potential: Detecting and Measuring Supermassive Black Holes in Triaxial Galaxies

by

Emily Rose Liepold

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Chung-Pei Ma, Chair
Professor Jessica Lu
Professor Daniel Kasen

Summer 2023

Realizing the Full Potential: Detecting and Measuring Supermassive Black Holes in Triaxial Galaxies

Abstract

Realizing the Full Potential: Detecting and Measuring Supermassive Black Holes in Triaxial Galaxies

by

Emily Rose Liepold

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Chung-Pei Ma, Chair

Supermassive black holes lie at the centers of all elliptical galaxies and are thought to play fundamental roles in guiding the evolution of their host galaxies. A number of well-established empirical scaling relations are known between the mass of the black hole to various galaxy properties, such as the stellar velocity dispersion, the break radius of the light profile, and the bulge mass of the galaxy. The behavior of these relations at their upper end is less well-established as only a handful of black holes above $10^9 M_\odot$ have been been detected and measured. Galaxies harboring these black holes are now known to frequently have kinematic features which indicate that the galaxies have triaxial intrinsic shapes, while the modelling that has produced these few black hole mass measurements has nearly always assumed axisymmetry. The assumed symmetry of a galaxy has potentially significant impact on the resulting inferred black hole mass. This dissertation follows the development of a code for triaxial Schwarzschild orbit modelling and application of that technique to measure the intrinsic three-dimensional triaxial shape of massive elliptical galaxies simultaneously with the masses of their central supermassive black holes and other parameters which describe their mass distributions.

We begin by exploring simpler axisymmetric Schwarzschild models. We found that our adopted Schwarzschild modelling code (TriOS) was frequently ill-behaved when running triaxial models with nearly axisymmetric shapes. To ensure that the code was robust once moving further from axisymmetry, we first focused on refining these axisymmetric models, including making changes to the code to allow for properly axisymmetric models, along the way developing further understanding of the conditions required for proper axisymmetry and then applying the updated code to the massive elliptical galaxy NGC 1453, a fast-rotating galaxy in the MASSIVE sample. With the updated code and procedures, we detected a black hole in NGC 1453 which was consistent with models using axisymmetric Jeans modelling and consistent with the $M_{\rm BH}$–σ relation.

We then turn to triaxiality. As with the axisymmetric case, we found that a number of procedures within the TriOS code and prior prescriptions for usage of the code lead to erroneous models. Quite significantly, one of the symmetries associated with triaxiality was mis-applied and the mass distributions were mis-constrained in the original code, leading to potential biases in the inferred parameters. We found a triplet of parameters which intuitively specify a particular deprojection of the light profile and allows for efficient searches over the space of allowed intrinsic shapes. We again apply our refinements to NGC 1453, finding consistent mass parameters with the axisymmetric case, but now with a constraint on the intrinsic shape which strongly rejects axisymmetry.

We finally focus on the massive elliptical galaxies M87, Holmberg 15A, and IC1101, performing observations on them over five years to build comprehensive maps of the stellar kinematics for each galaxy extending from the central region within the black hole's sphere of influence to deep within the dark matter halo. For Holmberg 15A and IC1101, the stellar kinematic maps extend to beyond 50 kpc from the center of the galaxy. Both M87 and Holmberg 15A's velocity fields exhibit prominent kinematic misalignments, a tell-tale indication of triaxiality. We use the TriOS code to perform triaxial Schwarzschild orbit modelling on all three galaxies, adding a new measurement of the black hole mass to the long history of measurements of M87's black hole and measuring the three-dimensional triaxial shape of its stellar halo for the first time, as well as significantly refining the mass measurement of Holmberg 15A* and making the most detailed measurements of the stellar kinematics of IC1101 to date.

For Tay

Contents

List of Figures

List of Tables

Acknowledgments

When I began my studies in Berkeley, I suffered from severe social anxiety, depression, and dysphoria. Today I can say that I'm extraordinarily lucky to have found an incredible collection of colleagues, collaborators, supporters, friends, and family over the past years. These communities have brought out the light and joy of life in ways I never knew possible and have instilled meaning where so recently there was none. There are far too many thanks to share here, but I'll mention a few.

First, thank you so much to my advisor Chung-Pei Ma, who's been my firmest advocate as I've grown as a scientist and has guided the way when the path forward has been foggy. Thank you to Jessica Lu, Dan Kasen, Adrian Lee, and Chung-Pei Ma for serving on my qualifying exam committee and Jessica, Dan, and Chung-Pei for serving on my thesis committee. Thank you to my collaborators, especially Jonelle Walsh, Matthew Quenneville, Jacob Pilawa, Ningyuan Xu, and Shaunak Modak.

Thank you to my friends, thankfully too many to fully enumerate here: to Alex Ekvik for sharing so many trails, concerts, albums, books, and puns; to Nikki and Reyna Bernstein for sharing with me the best pizza in the east bay; to Shayne Harwyn-Hawthorne for gracefully guiding me through my most difficult days (and for teaching me the art of espresso); to Annie Barry for reminding me of the lightness and levity of the dark; to Jay Klyman for a reminder of my past; and to so many others for glimpses of my future. Thank you for Jessica Rehak for help choosing a pithy title for this dissertation.

Thank you to the generations of transgender scientists who came before me–many of whom could not transition or live openly out of fear or prejudice–for building the world where I could. And to the generations of trans scientists who will follow for giving me hope that a more beautiful and just world is possible. And thank you to the queer and trans community of the east bay for teaching me the effervescence and power of queer joy.

Thank you to my family, who has stood by me, supported me, and unconditionally loved me through more than I could've ever imagined.

Thank you to Felix and Maude for teaching me to enjoy the stillness of a Sunday morning.

And thank you to Tay Apollo Nelson for seeing me, for sharing a life with me, for supporting me, for marrying me, and for loving me for so many years.

Chapter 1

Introduction

The broad goal of this dissertation is to identify and measure the masses of the supermassive black holes (SMBHs) that lie at the centers of a collection of massive galaxies. These objects are fascinating for a number of reasons which in no small part owe to their extreme nature—as some of the most massive objects in the universe they force our imaginations to their limits and inspire awe. The black holes which lie at the centers of galaxies are now understood to play essential roles in directing the evolution of the galaxies they live within. In turn, the growth of these supermassive black holes is impacted by the growth and evolution of their homes. This pattern of feedback and co-evolution between the two suggests that in order to deeply understand the long-term evolution of the galaxies in our universe we must also understand the relations between SMBHs and their host galaxies.

A number of relations between the mass of the central SMBH and various galaxy parameters have been empirically determined. For example, the $M_{\rm BH}$–σ relation connects that central mass to the velocity dispersion σ (the second moment of the velocity distribution; a measure of the spread in stellar velocities) [133]. This relation tells us that galaxies which have larger velocity dispersions and where stars tend to be moving more quickly tend to have more massive central SMBHs. Similarly, the $M_{\rm BH}$–$M_{\rm bulge}$ relation finds that galaxies with more massive bulges tend to have larger central black holes, with $M_{\rm BH} \sim 0.2\% M_{\rm bulge}$ [133, 108].

These relations hold over multiple orders of magnitude in the black hole mass, from M32* and Sgr A* at $\sim 10^{6.5} M_\odot$ [44, 85, 60, 89] to M87* at $\sim 10^{9.5} M_\odot$ [221, 77, 120], a thousand times larger. At the high end ($\gtrsim 10^9 M_\odot$) these relations become more tentative in part due to a paucity of black hole mass measurements in this region—at present only six SMBHs have been measured to have masses above $5 \times 10^9 M_\odot$ using stellar dynamical methods: NGC4486 [120], NGC2832 [136], NGC3842 [135], NGC1600, [192], NGC4889 [135], and Holmberg 15A [138].

All six of these measurements have overmassive central black holes which lie above the usual $M_{\rm BH}$–σ and those with reliable bulge mass or bulge luminosities lie above the $M_{\rm BH}$–$M_{\rm bulge}$ and $M_{\rm BH}$–$L_{\rm bulge}$. Deviation in the scaling relations in this regime is in part expected as the growth pathways for these most massive galaxies tend to shift somewhat

from those with lower mass. In particular, dry mergers are thought to have increased importance for these galaxies [8, 61, 145] and they're thought to be more likely to be among the brightest galaxies in their clusters and that the clusters they belong to are more likely to be very massive [112] and the mergers in these environments are thought to yield steeper scaling relations than the usual relations [16].

Given the sparseness of black hole mass measurements in this regime, two major strategies can be followed to improve the constraints on the relations between the central black holes and their host galaxies: additional points can be added and the existing points can be made more accurate. The work in this dissertation partially follows both pathways, working to establish techniques for properly modelling galaxies with triaxial stellar halos using the Schwarzschild orbit superposition method, then using that technique along with new stellar kinematic observations to make a new measurement of the mass of NGC1453* and to refine the previous mass measurements of M87* and Holmberg 15A*.

The galaxies harboring the most massive SMBHs tend to be elliptical in morphology, a result of their being the end product of series of mergers of less massive galaxies. These systems are known as 'early-type' galaxies, in part due to an arbitrary choice of nomenclature within Hubble's classification scheme [98, 100]. The central SMBHs cannot be observed directly, so their masses must be inferred from their impacts on the surrounding matter. In the work throughout this dissertation, the primary observable impact is in the motion and distributions of galaxy's stars, especially those near the central region. As this introduction will discuss, to infer the black hole mass from these properties, we must measure the light distribution of the galaxy, as well as the distributions of velocities of stars, then perform extensive modelling to determine the mass models which are consistent with those observations.

1.1 Observing Early-Type Galaxies

1.1.1 Photometric observations of early type galaxies

Photometric observations are among the most intuitive in astronomy – in short, it's simply photography. This class of observation has evolved rapidly over the past century from the emergence of extragalactic astronomy in the 1920s when distances to extragalactic objects were first measured [99, 98, 97] to today where large-scale galaxy surveys are commonplace. The distant nature of extragalactic astronomy pushes technologies to their limits as the angular size of objects shrinks with their distance from us, and observation on smaller angular scales requires overcoming atmospheric seeing which effectively blurs small-scale spatial features. In tandem with evolution of the photon collection evolving from photographic plates to simple photoelectric scans and photon counters (e.g. [39]) to modern CCDs, the telescopes available for such work have evolved tremendously, from the 100–inch class telescopes used at the turn of the century (including the 100–inch reflecting telescope used by Hubble through the 1920s) to the 10–meter class telescopes used today (including the Keck tele-

scopes used in this dissertation), and to higher and more remote locations, including at the peak of Maunakea, low Earth orbit (e.g., the *Hubble Space Telescope* whose archival data is used in this dissertation), and the Lagrange point L2 (e.g, the *James Webb Space Telescope*).

In these observations, a recording medium measures a quantity which is related to the number of photons received through a telescope from the astronomical object in the sky. With proper calibrations, that quantity is converted into a measure of the physical flux associated with the target. Typically a filter is used to only allow photons from a specific spectral range. In a similar manner to color photography, comparison of the flux in different bands gives a measure of the color of the object being viewed.

Measurement of the light profiles of early type galaxies and galaxies more generally has been a significant activity within astronomy over the past century. The profiles of elliptical galaxies are frequently described by the the ellipticity (their roundness or flatness), their PA (the orientation of the long axis of the ellipse on the sky), and their surface brightness (often expressed in magnitudes per square arcsecond). The projected morphology encodes a substantial amount of information about the internal behaviors of these galaxies. For example, if the intrinsic three-dimensional structure of the galaxy is well-described by axisymmetric ellipsoids, the isophotes of the projected light distribution will each have the same position angle. Accordingly, if the photometric PA is not constant one can immediately infer non-axisymmetry in the intrinsic shape of the galaxy [224, 184]. The degree of 'boxiness' in the isophotes, has been correlated with the relative amounts of coherent rotation and random motion in the stellar velocity field as well as the relative amount of rotation along the minor axis – galaxies which are more boxy tend to have more random motion, and much of the coherent rotation which is present is along the minor axis [106]. The overall brightness of the galaxies is also correlated with intrinsic properties. Efforts to infer the distributions of intrinsic three-dimensional shapes from populations of early-type galaxies have found that brighter ellipticals tend to have rounder isophotes than faint ellipticals and tend to be consistent with triaxial intrinsic shapes rather than axisymmetry [214, 195]. Additionally, the shapes of the isophotes have been correlated against the presence or absence of a central core in the brightness profile [113, 114], which in turn has been correlated against the presence and mass of a central supermassive black hole [192].

For ellipticals, the light profile can often be well-approximated by a profile in the Sérsic family of functions. The de Vaucouleurs profile was the first member of this family [40], followed by the general Sérsic profile [175]:

$$\begin{aligned} \ln I(R) &= \ln I_0 - kR^{1/n} \\ \ln I(R) &= \ln I_e - b_n \left[(R/R_e)^{1/n} - 1 \right] \end{aligned} \tag{1.1}$$

These expressions describe the intensity I as a function of two-dimensional radius R. The overall shape of the profile is set by the Sérsic index n. For $n = 1$, the profile is exponential, while for $n = 4$, the profile is the earlier de Vaucouleurs profile. The overall amplitude is set by the central intensity I_0 in the first expression or by the intensity I_e at the effective radius

R_e. The effective radius is the isophote which contains half the overall light. In the first form, the overall length scale of the profile is related to k. In the second form, the auxiliary quantity b_n appears. This is well-approximated by $b_n = 2n - 1/3$.

A somewhat more complicated but useful expansion of this form has been developed to accomodate galaxies which have power-law behavior in their central regions, but Sérsic-like behavior in their outer regions [84]:

$$I(R) = I' \left[1 + \left(\frac{R_b}{R}\right)^{\alpha}\right] \exp\left[-b_n \left(\frac{R^\alpha + R_b^\alpha}{R_e^\alpha}\right)^{1/(\alpha n)}\right] \tag{1.2}$$

This form introduces two additional parameters, the inner slope α and the break radius R_b which specifies the transition between the power-law and Sérsic behaviors.

For the most part in this thesis, we will focus on analytic but less easily interpretable surface brightness profiles in the form of Multi-Gaussian Expansions (MGEs; [24]), which are sums of several concentric Gaussians. These two-dimensional Gaussians each have potentially different widths, axis ratios, and orientations to flexibly fit a wide variety of real galaxies. The MGE parameterization of the surface brightness has the form

$$\Sigma(R', \theta') = \sum_{j=1}^{N} \frac{L_j}{2\pi\sigma_j'^2} \exp\left[-\frac{1}{2\sigma_j'^2}\left(x_j'^2 + \frac{y_j'^2}{q_j'^2}\right)\right] \tag{1.3}$$

where Σ is the surface brightness, R' and θ' are the polar coordinates on the sky plane, L_j, σ_j', and q_j' are the total luminosity, projected major axis length, and projected axis ratio for the j-th component of the fit. The orientations of each component are related to a series of position angles ψ_j' on the sky so that $x_j' = R'\sin(\theta' - \psi_j')$ and $y_j' = R'\cos(\theta' - \psi_j')$.

1.1.2 Deprojection

A primary problem in modelling and understanding the structure of early type galaxies is projection. Along any line of sight only the integrated light from stars along that line can be measured, so the intrinsically three-dimensional light distribution is collapsed into a two-dimensional image. While the two dimensional image places significant constraints on the the three-dimensional shape, assumptions about that shape must be made in order to *deproject* the light profile. The most significant assumption is typically the symmetry associated with the the three dimensional structure.

In the simplest case, one can imagine a spherically symmetric system where the three-dimensional light density $\nu(r)$ is a function of only the three-dimensional radial coordinate. The corresponding surface brightness profile $\Sigma(R)$ must also be azimuthally symmetric, resulting from the integration of the three-dimensional density along a line of sight with

$$\Sigma(R) = \int_{-\infty}^{\infty} \nu(\sqrt{R^2 + z^2}) dz = 2\int_R^{\infty} \frac{\nu(r) r\, dr}{\sqrt{r^2 - R^2}}, \tag{1.4}$$

where R is the two-dimensional radial coordinate and z traces along the line of sight. The second form on the right side comes from interpreting this integral as an Abel transform. This is valuable as it allows one to trivially perform a deprojection – that is, to infer ν from Σ – with the inverse Abel transform:

$$\nu(r) = \frac{-1}{\pi} \int_r^\infty \frac{d\Sigma}{dR} \frac{dR}{\sqrt{R^2 - r^2}} \tag{1.5}$$

Importantly, this deprojection is unique and a given radially symmetric surface brightness fully determines the form of the three-dimensional spherical density profile.

This notion can be extended trivially to the case of edge-on axisymmetric deprojections [80]. These deprojections depend only on the cylindrical coordinates with $\nu(r, y)$, where r is the cylindrical radial coordinate and y is the axial coordinate. Since the deprojection is edge-on, this axial coordinate y lies on the sky plane. We can index the line-of-sight with the coordinate z so that the x-axis also lies in the sky plane.

We can relate the two and three-dimensional profiles as before with

$$\Sigma(x, y) = \int_{-\infty}^{\infty} \nu(r = \sqrt{x^2 + z^2}, y) dz \tag{1.6}$$

This is identical in structure to the form shown above, though now with both densities indexed by the axial coordinate y. We can then repeat the deprojection as before.

$$\nu(r, y) = \frac{-1}{\pi} \int_r^\infty \frac{\partial \Sigma(x, y)}{\partial x} \frac{dx}{\sqrt{x^2 - r^2}} \tag{1.7}$$

Unfortunately, this case is a relatively uncommon occurrence. For randomly oriented galaxies, orientations within a degree of edge-on (that is, $\theta > 89°$) should appear in only about 1.7% of cases. Furthermore, except for very flattened systems the orientation can not be determined a priori – a thick system viewed edge-on may appear virtually identical to a flattened system viewed obliquely. In order to avoid these degeneracies, when the intrinsic three-dimensional shape of the galaxy is important it must be treated as an additional parameter to model.

A useful framework for early type galaxies is to assume that the intrinsic mass or light distribution stratified on similar ellipsoids, with an elliptical radius variable $r^2 = x^2 + y^2/p^2 + z^2/q^2$ and intrisic axis length ratios p and q, so that the intrinsic mass distribution can be expressed as a function of only this radius with $\nu(x, y, z) = \nu(r; p, q)$. In this case, a triplet of 'viewing angles' (θ, ϕ, ψ) are required to specify the coordinate transformation between the intrinsic (x, y, z) coordinate system of the galaxy and the projected (x', y', z') coordinate system associated with observation on the sky-plane [13, 41, 74]. Within this convention, the long, intermediate, and short axes of the three-dimensional ellipsoid are aligned with the intrinsic x, y, and z axes. Similarly, in the projected coordinate system, the long axis of the projected ellipse is aligned with the x' axis, the short axis with the y', and the z' axis points towards the us along the line of sight.

The viewing angles θ and ϕ define the vector pointing from the galaxy to the observer and are the usual spherical angles within the galaxy's reference frame ($\theta = 0°$ points along the z axis, $\theta = 90°$ points in the x–y plane; $\phi = 0°$ points in the x–z plane, $\phi = 90°$ points in the y–z plane). A third angle ψ is required to specify the rotation about the line of sight on the sky.

In projection, the corresponding two-dimensional surface brightness profile will also be ellipsoidal, with an elliptical radius variable $R'^2 = x'^2 + y'^2/q'^2$ and $\Sigma(x', y') = \Sigma(R'; q')$. With a choice of viewing angles, the projection can still be expressed as an Abel transform and the deprojection still expressed as the inverse Abel transform [184, 13]. Accordingly, for each triplet of viewing angles, there's a well-determined mapping from an observed q' axis ratio to the intrinsic axis ratios p and q and a dimensionless scale factor u. Alternatively, for a given q', there's a well-determined mapping from a choice of u, p, and q into a triplet of viewing angles, so long as the chosen intrinsic axis ratios obey the inequality $0 \leq q \leq uq' \leq p \leq u \leq 1$ [24, 202].

In realistic elliptical galaxies, the observed flattening of the surface brightness profile is typically variable. Often the central regions have relatively round isophotes, while the outer regions become increasingly flat. In triaxial systems, the position angle (PA) on the sky associated with the apparent long axis of the isophotes may also *twist* [9]. Both of these features suggest that a simple single-ellipsoid model may not be sufficient. A common approach to address this is to consider real galaxies to be built from a superposition of ellipsoids with different flattenings which possibly lie at different PAs. With a set of common viewing angles, these components with varied q' will deproject into a set of components with varied p and q, resulting in a three-dimensional profile which also has spatially-varying shape.

This opens yet another realm of non-uniqueness in the deprojection process – the particular decomposition of the two-dimensional image into a superposition of elliptical components may depend on both the parameterization of the light profile (that is, the functional form of $\Sigma(R'; q')$) and the fit parameters used to generate that functional form. A number of approaches and codes exist to address this problem (e.g. [131, 38]). Throughout this thesis, I'll use the commonly-used technique of Multi-Gaussian Expansion (MGE; [24]) to perform the decomposition. The MGE technique fits the two-dimensional surface brightness image with the form shown in Eqn. 1.3 which consists of a series of concentric Gaussian components, each with distinct projected axis ratio q'_j and length scale σ'_j. In turn, the three-dimensional density profile is also described by a series of Gaussians with varied intrinsic axis ratios p_j, q_j and length scales σ_j. As noted above, the mapping from the observed to intrinsic axis lengths $(\sigma'_j, q'_j\sigma'_j) \rightarrow (\sigma_j, \sigma_j p_j, \sigma_j q_j)$ is fully specified by a choice of viewing angles (θ, ϕ, ψ).

1.1.3 Spectroscopic observations of early type galaxies

While photometric observations typically collect all photons within some broad spectral window (the *band*), one can also perform spectroscopy to measure the flux as a function of the photon wavelength. This typically involves use of a diffraction grating to disperse multichromatic light along an axis. Historically, long-slit spectroscopy has been commonplace

in measurements of early-type galaxies. These observations capture light along a narrow region on the sky, essentially a one-dimensional image. This light is dispersed spectrally to form a two dimensional image, where one axis is spatial and the other is spectral. In recent decades, integral field spectroscopy has become standard, where the use of image slicers or fibers allows for construction of a three-dimensional datacube with two spatial axes and one spectral axis. The work in this thesis makes significant usage of the Gemini Multi Object Spectrograph (GMOS; [96]) and the Keck Cosmic Web Imager (KCWI; [142]).

Spectral fitting

As a rough approximation, the measured stellar spectra are convolutions of the distribution of line-of-sight velocities of the stars and the rest-frame spectra of the individual stars. In our work we are typically interested in measurement and use of the stellar line-of-sight velocity distribution (LOSVD), so additional processing of individual spectra must be performed to extract LOSVDs from observed spectra.

A conceptually simple early approach to this problem is the Fourier quotient method [171]. In short, given an observed spectrum $G(\lambda)$ and a stellar spectrum $S(\lambda)$, then the Fourier transform of the broadening function B which relates the two can be inferred by division of the Fourier transforms of the two spectra ($\tilde{G}(k)$ and $\tilde{S}(k)$). The parameters associated with the distribution (the line strength γ, mean velocity v and dispersion σ) can then be fit by finding the parameters which best reproduce the Fourier quotient $\tilde{Q}$ (see Sec 2.2 of [48], Eqn 1 reproduced here)

$$\tilde{Q}(k) = \frac{\tilde{G}(k)}{\tilde{S}(k)} \approx \tilde{B}(k) = \gamma \exp\left[-\frac{1}{2}\left(\frac{2\pi k\sigma}{N}\right)^2 + \left(\frac{2\pi v i k}{N}\right)\right] \tag{1.8}$$

A related technique is the cross-correlation scheme [194], where a stellar spectrum and galaxy spectrum are cross-correlated with $C(n) = \sum_m G(m)S(m-n)$. Tonry and Davis 1979 demonstrate that this cross-correlation can be related to the parameters of the broadening function with (see Sec 2.3 of [48], eqn 5 reproduced here)

$$C(v) \propto \frac{\gamma\tau^2}{\sigma^2 + 2\tau^2} \exp\left[-\frac{1}{2}\frac{(v-v_0)^2}{\sigma^2 + 2\tau^2}\right], \tag{1.9}$$

where τ is the broadening of the stellar spectrum, σ is the width of the velocity distribution, v_0 is the mean velocity of the distribution, and γ is the line strength.

Variations on these two techniques were developed over many years, including the Fourier correlation quotient (FCQ) method [7]. A major limitation of Fourier and cross-correlation techniques is their inability to efficiently incorporate masks on regions of the spectrum which are impacted by bad pixels, sky spectrum contamination, or gas emission. For the past decades performing fitting directly in the spectral pixel space has been standard practice. These techniques construct a model for the observed spectrum in each spectral pixel, then optimize the parameters of the model to best fit the observed spectrum. A standard code

used for this technique is the penalized pixel fitting code (pPXF; [22, 26]). The model used in pPXF has the following form (Eqn 11 of [26]), and other codes (see e.g., [78, 63, 139]) typically use some variation of this scheme:

$$G_{\rm mod}(x) = \sum_{n=1}^{N} w_n \left[[T_n(x) * \mathcal{L}(cx)] \sum_{k=1}^{K} a_k \mathcal{P}_k(x) \right] + \sum_{l=1}^{L} b_l \mathcal{P}_l(x) + \sum_{j=1}^{J} c_j S_j(x) \tag{1.10}$$

In this expression, a set of stellar templates T_n are superimposed with weights w_n and convolved against an LOSVD $\mathcal{L}$, whose spectral pixels are indexed by x. This convolution is multiplied against a 'multiplicative polynomial' $a_k \mathcal{P}_k$, which corrects for variations in the flux calibration and shape of the continuum between the stellar templates and observations or for reddening due to dust. An 'additive polynomial' $b_l \mathcal{P}_l$ corrects for additional additive components in the spectrum (sky background, AGN, etc), and explicit sky correction can be performed by including representative sky spectra $c_j S_j$. With this scheme, the best-fit parameters can be found through standard optimization methods.

1.2 Modelling Early-Type Galaxies

1.2.1 The collisionless Boltzmann equation and the Jeans equations

The basic problem of our work is inference of masses from the kinematic and photometric observations described in the previous section. In this section, we'll follow Binney and Tremaine Chapter 4 [14] to motivate various forms of the collisionless Boltzmann equation and Jeans equations.

It's useful to consider the state of an elliptical galaxy to be described by a distribution function $f(\mathbf{x}, \mathbf{v}, t)$. In the most literal sense, we can consider this to consist of delta functions at the true locations of each individual star in the phase space. That is,

$$f(\mathbf{x}, \mathbf{v}, t) \propto \sum_i m_i \delta(\mathbf{x_i}(t) - \mathbf{x}) \delta(\mathbf{v_i}(t) - \mathbf{v}). \tag{1.11}$$

However, it's more informative to interpret the distribution function as a probabilistic measure of how likely one is to locate a star in a particular location in the six-dimensional phase-space at a particular time. As a probability, the natural normalization for f is $\int f \, d^3\mathbf{x} d^3\mathbf{v} = 1$. With this interpretation, we can write down the conservation of probability for the phase-space

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (f \dot{\mathbf{x}}) + \frac{\partial}{\partial \mathbf{v}} \cdot (f \dot{\mathbf{v}}) = 0 \tag{1.12}$$

With use of Hamilton's equations, this can be substantially simplified:

$$\frac{\partial f}{\partial t} + \dot{\mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{x}} + \dot{\mathbf{v}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0 \tag{1.13}$$

We care about particles subject to the gravitational potential, generally expressed with $\Phi(\mathbf{x}, t)$, so $\dot{\mathbf{v}} = -\frac{\partial \Phi(\mathbf{x},t)}{\partial \mathbf{x}}$

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi(\mathbf{x}, t)}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0 \tag{1.14}$$

This is the collisionless Boltzmann equation, expressed in cartesian coordinates. The Jeans equations follow by taking the zero-th and first velocity moment of this equation. To take the zero-th moment, we integrate over all velocity, which results in the following:

$$\frac{\partial \nu}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\nu \overline{\mathbf{v}}) = 0 \tag{1.15}$$

The second Jeans equation comes from taking the first velocity moment of the CBE – that is, we multiply by velocity, then integrate over velocity. This results in the following:

$$\frac{\partial (\nu \overline{v_j})}{\partial t} + \frac{\partial (\nu \overline{v_i v_j})}{\partial x_i} + \nu \frac{\partial \Phi}{\partial x_j} = 0 \tag{1.16}$$

The commonly-used form of this equation subtracts the continuity equation and expresses the second velocity moments in terms of the velocity-dispersion tensor $\sigma_{ij}^2 = \overline{v_i v_j} - \overline{v}_i \overline{v}_j$:

$$\nu \frac{\partial \overline{v}_j}{\partial t} + \nu \overline{v}_i \frac{\partial \overline{v}_j}{\partial x_i} = -\nu \frac{\partial \Phi}{\partial x_j} - \frac{\partial (\nu \sigma_{ij}^2)}{\partial x_i} \tag{1.17}$$

This is the Jeans equation from the first velocity moment of the CBE, which relates the first and second moments of the velocity.

1.2.2 Estimating Masses with the Virial Theorem

The simplest mass estimation scheme follows from the virial theorem, which we can derive following the Jeans equations from above. The route here follows the description from Binney and Tremaine, Chapter 4 [14].

First, we multiply our second Jeans equation by $M x_k$ where M is the total mass of the system and integrate spatially. We can use the definition that the physical density is just $\rho = M\nu$

$$\begin{aligned} 0 &= \frac{\partial (\nu \overline{v_j})}{\partial t} + \frac{\partial (\nu \overline{v_i v_j})}{\partial x_i} + \nu \frac{\partial \Phi}{\partial x_j} \\ 0 &= \int d^3\mathbf{x}\, x_k \frac{\partial (\rho \overline{v_j})}{\partial t} + \int d^3\mathbf{x}\, x_k \frac{\partial (\rho \overline{v_i v_j})}{\partial x_i} + \int d^3\mathbf{x}\, x_k \rho \frac{\partial \Phi}{\partial x_j} \end{aligned} \tag{1.18}$$

The second term can be evaluated with the divergence theorem and is related to the kinetic energy tensor:

$$\int d^3\mathbf{x}\, x_k \frac{\partial (\rho \overline{v_i v_j})}{\partial x_i} = -\int d^3\mathbf{x}\, \delta_{ki} \rho \overline{v_i v_j} = -2K_{kj} \tag{1.19}$$

Similarly, the third term is the potential energy tensor W_{kj}. Note that $W_{jk} = W_{kj}$ and $K_{kj} = K_{jk}$ as both tensors are symmetric.

With a bit of algebra, the first component can be related to the moment of inertia tensor. First, since W and K are symmetric we can average the index-swapped versions of the previous equation and move the time derivative outside the integral.

$$0 = \frac{1}{2}\frac{d}{dt}\int d^3\mathbf{x}\,\rho(x_k\overline{v_j} + x_j\overline{v_k}) - 2K_{jk} - W_{jk} \tag{1.20}$$

Consider the time derivative of the moment of inertia tensor:

$$\frac{d}{dt}I_{jk} = \frac{d}{dt}\int d^3\mathbf{x}\rho x_j x_k = \int d^3\mathbf{x}\frac{\partial\rho}{\partial t}x_j x_k \tag{1.21}$$

The first Jeans equation from above has the form of a continuity equation and can be used to replace the derivative on the right side. The result can be evaluated with the divergence theorem.

$$\frac{d}{dt}I_{jk} = -\int d^3\mathbf{x}\frac{\partial(\rho\overline{v}_i)}{\partial x_i}x_j x_k = \int d^3\mathbf{x}\rho(x_k\overline{v}_j + x_j\overline{v}_k) \tag{1.22}$$

This is identical in form to the remaining term in our main equation, so we can plug it in to find the full tensor virial theorem.

$$0 = \frac{1}{2}\frac{d^2}{dt^2}I_{jk} - 2K_{jk} - W_{jk} \tag{1.23}$$

For a model galaxy we can assume time independence and drop the first term. We can also take the trace of this equation to form the usual scalar virial theorem.

$$0 = W + 2K \tag{1.24}$$

The kinetic energy is just $K = \frac{1}{2}M\langle v^2\rangle$, so this directly relates the second moment of the velocity to the potential energy.

$$\langle v^2\rangle = -W/M \tag{1.25}$$

We can understand this relation as describing how the motion of the stellar population relates to the potential surrounding them. This potential does not need to come from only the stellar mass distribution and can include contributions from non-stellar sources (e.g., dark matter and a central BH). For simple systems which consist of stars alone, the relation between W and M can be characterized by the gravitational radius $r_g = GM^2/|W|$. Different mass distributions will result in different values for r_g. Binney and Tremaine note that for many reasonable models (e.g., Jaffe, Hernquist, King), the half-light radius of the model and its gravitational radius are related with scale factors of $r_h \sim 0.4r_g$ to $0.5r_g$. Accordingly, this suggests a reasonable approximation with $\langle v^2\rangle \sim \kappa GM/r_h$ for $\kappa \sim 0.45$. In the case of systems with a spatially constant total mass-to-light ratio, we can interpret the mass M to be the total mass from all components.

This chain of reasoning provides a straightforward relation between the velocity dispersion, light profile, and mass. Of course, there are significant features lacking. Real galaxies have highly spatially variable velocity distributions and have velocity distributions which are not well described with only their second moment. Additionally it is not trivial (as we will see in the next section) to infer the three dimensional velocity dispersion from observations alone and assumptions must be made to relate the observed line of sight distribution to the three-dimensional velocity distribution. These limitations suggest that virial estimators are useful in their simplicity and intuitive behavior, but more comprehensive schemes must be used to model realistic galaxies.

1.2.3 Jeans (anisotropic) modelling

While the virial estimator discussed in the previous section provides relations between the total kinetic energy and total potential energy of the system, the Jeans equations describe how local quantities (ν, $\overline{v}_j$, $\overline{\sigma}^2_{ij}$, and Φ) vary throughout the galaxy. This is extremely valuable for constructing models which take into account the local variations in the potential (that is, differentiating between different mass profiles or assessing the presence of a central SMBH) and for incorporating the wealth of spatially resolved kinematic information we have available with modern instruments and techniques (see Sec. 1.1.3). Models which directly apply the Jeans equations are called Jeans models.

The Jeans Anisotropic Modelling (JAM, [18, 25]) code is well-established and widely used for this purpose. Other schemes and codes exist (e.g.,[12, 205, 185]), but as JAM gives a good picture of their general behavior, we'll restrict this discussion to that scheme. Two versions of the JAM code exist, which make different assumptions about the velocity anisotropy in the galaxy. For this discussion, we'll focus on the scheme described in Cappellari 2008 (hereafter C08; [18]), which assumes cylindrically aligned velocity ellpsoids. We'll follow the discussion in that paper to motivate the Jeans modelling strategy.

Eqn. 1.17 gave a general form of the Jeans equation, expressed in Cartesian coordinates. The cylindrically aligned JAM solution uses cylindrical coordinates, and imposes axial symmetry where the potential and distribution function are constant w.r.t. the azimuthal angle. Additionally, the system is assumed to be time-independent, so the time-derivative term is dropped. That equation can be rewritten in cylindrical coordinates as a pair of equations related to the R and z derivatives of the potential (C08 eqns 3 and 4):

$$\begin{aligned} \frac{\nu\overline{v_R^2} - \nu\overline{v_\phi^2}}{R} + \frac{\partial(\nu\overline{v_R^2})}{\partial R} + \frac{\partial(\nu\overline{v_R v_z})}{\partial z} &= -\nu\frac{\partial\Phi}{\partial R} \\ \frac{\nu\overline{v_R v_z}}{R} + \frac{\partial(\nu\overline{v_z^2})}{\partial z} + \frac{\partial(\nu\overline{v_R v_z})}{\partial R} &= -\nu\frac{\partial\Phi}{\partial z} \end{aligned} \tag{1.26}$$

In total this pair of equations has six variable quantities: ν, Φ, $\overline{v_R^2}$, $\overline{v_\phi^2}$, $\overline{v_z^2}$, $\overline{v_R v_z}$. Typically two of these quantities (Φ, ν) are fixed by assuming a specific mass model, so the model is still underspecified. JAM provides a closure to this system by assuming that the velocity

ellipsoids defined by $\overline{v_i v_j}$ terms are aligned with the cylindrical coordinate system – that is, that when cylindrical coordinates are used, that off-diagonal terms in that tensor vanish. This eliminates the $\overline{v_R v_z}$ term. The second major assumption is that in the R–z plane, the velocity ellipsoids have constant flattening with $\overline{v_R^2} = b\overline{v_z^2}$ which completes the closure. The equations above simplify significantly under these assumptions:

$$\begin{aligned} \frac{\nu(b\overline{v_z^2} - \overline{v_\phi^2})}{R} + \frac{\partial(\nu b\overline{v_z^2})}{\partial R} &= -\nu\frac{\partial\Phi}{\partial R} \\ \frac{\partial(\nu\overline{v_z^2})}{\partial z} &= -\nu\frac{\partial\Phi}{\partial z} \end{aligned} \tag{1.27}$$

These equations now have three model parameters (b,ν,Φ) and two unknown quantities ($\overline{v_z^2}$, $\overline{v_\phi^2}$)

These equations allow us to reconstruct $\overline{v_R^2}(R,z)$, $\overline{v_z^2}(R,z)$ and $\overline{v_\phi^2}(R,z)$, which characterize the velocity ellipsoids in the galaxy's coordinates. Following this, the intrinsic velocity distribution is projected along the line-of-sight. In Sec. 1.1.2 we discussed the general problem, noting that three angles are required to fully specify a projection and deprojection. When the intrinsic shape of the ellipsoid is axisymmetric (that is, $p = 1$), two of those angles are constrained and the projection-deprojection is fully specified by just the inclination θ, which is usually expressed as i in this context. This adds an additional JAM model parameter i, which specifies the projection. This projection performs the mapping from the internal model quantities $\overline{v_z^2}$ and $\overline{v_\phi^2}$ to the observable quantity $\overline{v_{\rm LOS}^2}$.

The basic strategy of practical usage of Jeans modelling is then to propose a trial model specified by (Φ,ν,b,i), infer the $\overline{v_{\rm LOS}^2}$ at numerous locations on the sky, then compare those dispersions to a set of observed LOSVDs at the same locations. This has the form of a standard curve-fitting optimization problem and can be treated as such. Jeans modelling is extremely valuable due to its relative flexibility and semi-analytic form. For a particular model, the goodness of fit to a set of observations can be evaluated in a fraction of a second in most cases, allowing for relatively short evaluation time even for long MCMC chains or nested sampling routines.

Unfortunately, the assumptions made for Jeans modelling are quite limiting. In the case of JAM, the global alignment of the velocity ellipsoids limits the ability of the model to reproduce the expected anisotropies in regions where the potential has different symmetries. Coarsely speaking, in regions where the potential is highly spherical, such as within the sphere of influence of a black hole or far into the dark matter halo, one may expect roughly spherical-coordinate-aligned ellipsoids while in regions where the potential is more axisymmetric (such as in the stellar-mass-dominated region of an axisymmetric galaxy), one may expect closer alignment with cylindrical coordinates [18, 25]. Additionally, the assumption of an axisymmetric stelllar component is less justified for the very massive early type galaxies explored through this thesis, as those systems tend to be at least somewhat triaxial [117, 102, 103, 54]. Finally, the line of sight velocity distributions which can now be measured with precision IFU spectroscopy often have distinctly non-Gaussian features and are not

well-characterized by only their dispersion. Since Jeans modelling typically only constrains the second moment of the velocity as it uses the first moment of the CBE, these higher order features cannot be captured in the current formalisms.

1.2.4 Schwarzschild modelling

Much of the work throughout this thesis uses the Schwarzschild orbit superposition method [173, 174], which is described at length elsewhere throughout this thesis (See Chap. 2.4, 3.2, 4.2, 5.4, and 6.5), so it will not be described at length here. In short, a mass distribution is proposed, a library of representative orbits which span the phase space associated with the potential is constructed, and the orbits within those libraries are superimposed in order to reproduce both the assumed mass distribution and a set of observables. By repeating this procedure for a large number of proposed mass distribution, one can find the mass distributions which best reproduce the observations.

1.3 Overview of this Dissertation

This thesis follows two broad trajectories, first expanding and demonstrating the robustness of a code for performing Schwarzschild orbit modelling in potentials with triaxial stellar components, then applying that code to real galaxies. Chapter 2 describes our usage of the TriOS triaxial orbit superposition to perform axisymmetric modelling of the massive elliptical NGC 1453. Chapter 3 discusses the behavior of the otherwise triaxial code in the axisymmetric limit and the changes to the code and its usage which are required to produce robust axisymmetric models. Chapter 4 extends our usage of TriOS to models with triaxial stellar halos, measruing the mass of the the SMBH in NGC 1453 simultaneously with its triaxial three-dimensinoal shape. Chapter 5 presents new measurements of the stellar kinematics of M87 using the Keck Cosmic Web Imager which are used to revise the stellar dynamical measurement of the central black hole. Finally, chapter 6 presents new measurements of the stellar kinematics of H15 and IC1101, two massive ellipticals with exceptionally large cores, as well as a re-measurement of the mass of H15's supermassive black hole.

Chapter 2

A Stellar Dynamical Mass Measurement of the Supermassive Black Hole in Massive Elliptical Galaxy NGC 1453

The broad goal of this thesis is to develop and expand techniques for simultaneously measuring the intrinsic three-dimensional triaxial shapes and mass parameters of massive early type galaxies, but we have begun our work with simpler axisymmetric models. The motivation for this comes from a few threads. First, axisymmetric models are fundamentally simpler models – since they have azimuthal symmetry, the effective dimensionality of the distribution function is reduced relative to triaxiality. Axisymmetric models have intuitively simpler orbital structures, consisting only of loop orbits which orbit about the symmetry axis, while triaxial models consist of multiple types of loops, box orbits, and other minor orbit families. Second, the parameter-space required to specify an axisymmetric deprojection is one-dimensional compared to the three dimensions of triaxial models, so exploration of the full parameterspace is substantially easier in axisymmetry. Third, we initially had quite a bit of difficulty getting physically reasonable results with triaxial models, even when those models had shapes relatively near axisymmetry. In our troubleshooting, bug-fixing, and head-scratching, we found that the most prudent path forward was to ensure that our TriOS Schwarzschild orbit modelling code could give reasonable results in axisymmetry before adding the complications of triaxiality.

These first two chapters describe our first explorations with the TriOS code with axisymmetric stellar components. This chapter discusses the detection and measurement of the black hole at the center of NGC 1453, a fast-rotator in the MASSIVE sample. This is our group's first black hole mass measurement using TriOS in the axisymmetric limit.

In addition to detection of the SMBH in NGC 1453 and discussion of the use of TriOS in the axisymmetric limit, this chapter also discusses considerations while using the Gauss-Hermite moments of the line-of-sight velocity distributions (LOSVDs) as constraints within

a Schwarzschild model, noting that prior practice typically underconstrained the velocity distributions, potentially introducing non-physical artifacts and weakening the constraint on the inferred model parameters (such as the black hole mass).

This chapter was originally published as

Christopher M. Liepold et al. "The MASSIVE Survey. XV. A Stellar Dynamical Mass Measurement of the Supermassive Black Hole in Massive Elliptical Galaxy NGC 1453". In: *Astrophys. J.* 891.1 (Feb. 2020), p. 4. DOI: 10.3847/1538-4357/ab6f71

Note that in the original publications of this chapter and the two following chapters, my first name and middle initial were woefully mis-spelled. For chapter 5 and beyond, this error has been corrected. Minor alterations to the text of this chapter have been made to ensure fluency throughout this dissertation.

2.1 Introduction

Making a direct dynamical measurement of the mass of a supermassive black hole (SMBH) using stellar or gas kinematics requires both exquisite observational datasets and extensive theoretical modeling. Over three decades of efforts by multiple research groups have accumulated about 100 dynamically determined masses for SMBHs at the centers of local galaxies out to a distance of about 120 Mpc, with varying degrees of accuracy (see compilations in, e.g., [133, 108, 165]).

The high-mass regime ($M_{\rm BH} \gtrsim 10^{9.5} M_\odot$) faces the additional challenge that the host galaxies are massive elliptical galaxies whose central stellar light profiles typically have flattened cores that differ significantly from the high-density and cuspy centers of less massive elliptical galaxies and bulges of disk galaxies. These stellar cores are a defining feature of the most massive ellipticals (e.g., [61, 112, 32, 66, 84]), indicating a significant deficit of stars, possibly due to three-body gravitational slingshots that scatter stars passing close to a SMBH binary to larger radii (e.g., [6]). These diffuse cores make it extremely difficult to obtain stellar absorption-line spectra of high-S/N quality that is needed for reliable $M_{\rm BH}$ measurements. Long-integration observations on large ground-based telescopes in excellent seeing conditions or with the assistance of adaptive optics are required.

In pursuit of a comprehensive study of the highest-mass regime of local SMBHs and galaxies, we have been conducting a volume-limited survey, MASSIVE, of the most massive galaxies in the local universe [126]. The MASSIVE survey targets ~ 100 early-type galaxies (ETGs) in the northern sky ($\delta > -6$ degrees) within a distance of 108 Mpc. Within this volume, it is designed to be complete to an absolute K-band magnitude of $M_K = -25.3$ mag, covering all ETGs with stellar mass $M^* \gtrsim 10^{11.5} M_\odot$ and with no selection cuts on galaxy size, velocity dispersion or environment. This parameter range is unexplored by ATLAS$^{\rm 3D}$, the previous volume-limited survey of 260 local ETGs out to a distance of 42 Mpc [19].

We have obtained comprehensive spectroscopic data using IFS on both sub-arcsecond and arcminute scales and performed uniform measurements of the spatially-resolved kinematics. Many results on the stellar kinematics and stellar populations of MASSIVE galaxies out to

a few effective radii from our wide-field IFS observations can be found in Veale et al. [209, 210, 211], Ene et al. [54], and Greene et al. [86, 87]. Results from finely-resolved stellar kinematics in the central ~ 2 kpc regions of 20 MASSIVE galaxies are presented in Ene et al. [53, 55]. In addition to the IFS data, we have also assembled an extensive array of multi-wavelength data of MASSIVE galaxies to study stellar light profiles [83], cold molecular gas [36, 37], warm ionized gas [149], and hot X-ray gas [82, 215].

In addition to studying the luminous baryonic components in massive ETGs, one major science goal of the MASSIVE survey is to perform simultaneous dynamical mass modeling of the SMBH, stars, and dark matter for a sample of cleanly selected massive ETGs using a uniform set of sub-arcsecond and wide-field IFS data and photometric data. To date, only 7 of the 100 galaxies in the MASSIVE survey have published SMBH masses that are determined from orbit mass modeling of stellar kinematic data. Three of the 7 galaxies are in the Virgo cluster: NGC 4486 (M87; Gebhardt et al. [77]; see also Walsh et al. [221] and EHT Collaboration [57]), NGC 4472 (M49; Rusli et al. [164]), and NGC 4649 (M60; Shen et al. [178]). Two others are the brightest cluster galaxies of rich clusters: NGC 4889 in the Coma cluster and NGC 3842 in the Leo cluster [134, 136]. The remaining two are the brightest galaxies in galaxy groups: NGC 1600 in a fossil-like group [192] and NGC 7619 in the Pegasus group [164]. Except for NGC 4649, the spectroscopic observations were all conducted with IFS on 8-10 meter telescopes. The measured $M_{\rm BH}$ spans an order of magnitude from $\sim 2 \times 10^9 M_\odot$ to $\sim 2 \times 10^{10} M_\odot$. More $M_{\rm BH}$ measurements in this mass range are clearly needed to quantify more robustly the upper end of the $M_{\rm BH}$-galaxy scaling relations for a better understanding of black hole feedback processes and massive galaxy evolution. We have acquired the spectroscopic and photometric data that are needed to perform dynamical modeling for the 20 galaxies reported in Ene et al. [53] and several other galaxies in the MASSIVE survey.

We turn to this goal in this paper and report the stellar dynamical measurement of the mass of a new SMBH at the center of the massive elliptical galaxy NGC 1453, a fast rotator in the MASSIVE survey. NGC 1453 is the brightest galaxy in its galaxy group, a typical environment for MASSIVE galaxies [210]. As listed in Table 3 of Ma et al. [126], the 2MASS "high-density contrast" group catalog [35] identified 12 galaxies as members in the NGC 1453 group, and estimated the virial mass of the group to be $10^{13.9} M_\odot$, presumably with large errors due to the small number of member galaxies. Our *HST* images of NGC 1453 show very regular elliptical isophotes (Figure 13 of Goullaud et al. [83]). The photometric and kinematic axes are also closely aligned [54, 53, 55], suggesting that the galaxy can be approximated as an axisymmetric system.

A distance measurement is needed to convert the observed angular scales to physical length and mass scales, and the inferred $M_{\rm BH}$ scales linearly with the assumed distance. For NGC 1453, we use our new determination of 51.0 Mpc from the MASSIVE-WFC3 project [83] using the surface-brightness fluctuation technique [101]. This new distance is about 10% smaller than 56.4 Mpc from group-corrected flow velocity in the 2MASS redshift survey. For a flat ΛCDM with a matter density of $\Omega_m = 0.315$ and a Hubble parameter of $H_0 = 70$ km s^{-1} Mpc^{-1}, 1 arcsec is 245 pc at 51.0 Mpc.

We perform Schwarzschild orbit modelling [173] using the triaxial implementation described by van den Bosch et al. [202]. We perform this modelling in the axisymmetric limit, and in Section 2.4.1 provide a prescription for how to achieve this limit properly in the triaxial code. The line-of-sight stellar velocity distributions (LOSVDs) are the main observational inputs in any stellar dynamical mass modeling of galaxies using orbit-based methods. It is a common practice, and the practice within this code, to expand the LOSVDs in a Gauss-Hermite series [132, 162]. The Gauss-Hermite expansion provides a natural way to express deviations from a Gaussian distribution since the terms in the series are orthogonal and linear. However, there has been little discussion in the literature about the appropriate order at which to truncate the series. To date, most published work on $M_{\rm BH}$ measurements that relied on the Gauss-Hermite expansion of the LOSVDs had measured only the lowest four moments from the stellar spectra (i.e., velocity V, dispersion σ, skewness h_3, and kurtosis h_4), using only these moments as observational constraints in subsequent orbit modeling and ignoring all higher moments. In this paper, we investigate the importance of including the higher moments for constraining $M_{\rm BH}$ in NGC 1453. When higher moments are left unconstrained, the LOSVDs predicted by the orbit models can contain large spurious contributions from these high moments.

In Sec. 2.2, we describe the spectroscopic observations and the resulting stellar kinematics from the Gemini Multi-Object Spectrograph (GMOS; [96]) IFS of the central $\sim$1.5 kpc by 2 kpc region of NGC 1453 and the wide-field coverage with the McDonald Mitchell IFS [93]. In Sec. 2.3, we describe our IR imaging observations of NGC 1453 from the *HST* Wide Field Camera 3 (WFC3) and the determination of the 2D light profile and the 3D deprojected stellar mass profile. The orbit modeling method is discussed in Sec. 2.4. The mass modeling results are given in Sec. 2.5, and the best-fit mass model is discussed further in Sec. 2.6. In Sec. 2.7, we discuss a number of relevant issues: the impact of Gauss-Hermite series truncation on the inferred $M_{\rm BH}$, the subtleties in achieving axisymmetry within the triaxial code, comparisons to results from Jeans modelling, implications for the black hole scaling relations, and connections to our previous observations of warm ionized gas in NGC 1453 [149].

2.2 Spectroscopic Data and Stellar Kinematics

As part of the MASSIVE survey, we obtained spatially-resolved stellar spectra for NGC 1453 with the Gemini Multi Object Spectrograph (GMOS; [96]) in the IFS mode on the 8.1 m Gemini North Telescope and the Mitchell/VIRUS-P IFS [93] on the 2.7 m Harlan J. Smith Telescope at McDonald Observatory. Here we summarize the observations, data reduction processes, and the procedures used to extract the stellar kinematics.

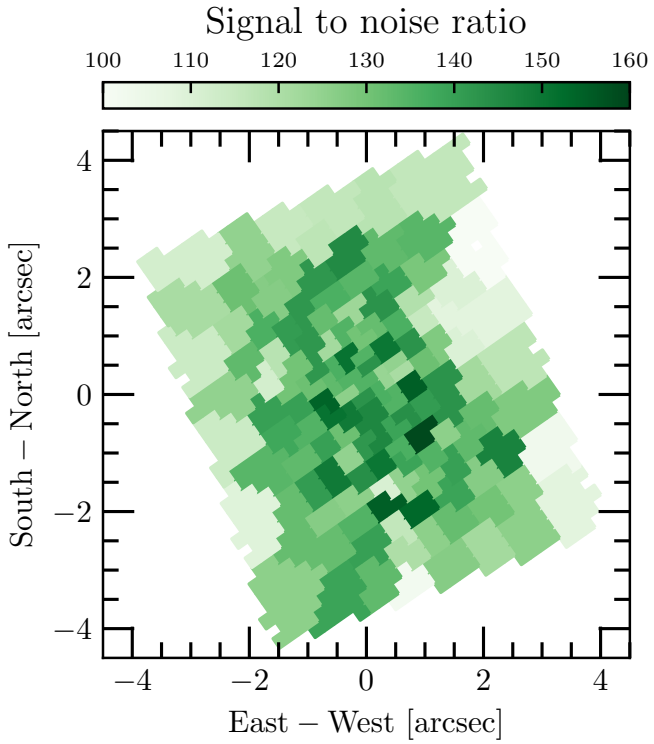

Figure 2.1: Signal-to-noise map of the Gemini GMOS IFS data for the 135 Voronoi bins in the central $5'' \times 7''$ of NGC 1453. A target S/N of 125 is used in the binning procedure. The S/N value for each bin scatters around the target with a typical RMS scatter of $\sim 10\%$, while the innermost bins achieve S/N up to ~ 150. Stellar kinematics from high-quality spectra are critical for measuring the gravitational effects of the central black hole. Our observations are able to achieve this high S/N over finely resolved spatially bins; both needed for kinematic extraction and black hole measurements.

2.2.1 Central kpc kinematics

We observed the central ∼1.5 kpc × 2 kpc region of NGC 1453 using GMOS in the 2015B semester. The two-slit mode of GMOS provided a field of view of $5'' \times 7''$ consisting of 1000 hexagonal lenslets, each with a projected diameter of $0''.2$. An additional 500 lenslets observed simultaneously a $5'' \times 3''.5$ region of the sky, which was offset by about $1'$ from the science field. The R400-G5305 grating and CaT filter combination was used to avoid spectral overlap on the detector and to provide a clean wavelength coverage of 7800-9330 Å. The spectral resolution of GMOS is determined from arc lamp lines for each lenslet with a mean value is 2.5 Å FWHM. Six science exposures, each of 850 seconds, were taken. The median seeing was $0''.7$ FWHM. Other details and our data reduction procedure are described in Ene et al. [53].

We use the CaII triplet absorption features over the rest wavelength range of 8420-8770 Å to measure the stellar kinematics. We apply the Voronoi binning algorithm [21] with a target S/N of 125 to determine how to spatially group the individual GMOS lenslets to achieve uniformly high-quality spectra. The procedure returns S/N values (per spectral pixel of 0.67 Å) that scatter about the target with an rms of $\sim 10\%$. Spectra from individual lenslets within a Voronoi bin are co-added as described in Ene et al. 2019 [53]. After fitting the spectra with pPXF, we re-estimate the S/N as the ratio of the median flux and the root-mean-square residual from the fit. The resulting S/N map for the 135 Voronoi bins is shown in Figure 2.1. The resulting CaII region of the spectra for three representative bins are shown (black curves) in Figure 2.2.

We measure the stellar line-of-sight velocity distribution (LOSVD) within each spatial bin using the penalized pixel-fitting (pPXF) method of Cappellari et al. [27]. The LOSVD is parameterized as a Gauss-Hermite series[1] up to order n

$$f(v) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi\sigma^2}}\left[1 + \sum_{m=3}^{n} h_m H_m(y)\right], \tag{2.1}$$

where $y = (v - V)/\sigma$, V is the mean velocity, σ is the velocity dispersion, and H_m is the m^{th} Hermite polynomial as defined in Appendix A of van der Marel and Franx [132].

For each spectrum, the stellar continuum is modeled with an additive polynomial of degree zero (i.e., an additive constant) and a multiplicative polynomial of degree three. A set of stellar template spectra are convolved with the instrumental line spread function and the LOSVD before adding and multiplying by these polynomials. The polynomial coefficients, template weights, and Gauss-Hermite moments are fitted simultaneously.

To test for potential issues with template mismatches, we compare two sets of stellar templates chosen from the Calcium Triplet (CaT) Library of 706 stars [30] and find negligible differences in the resulting kinematics. The first set contained 15 stellar templates of the same 15 stars used in the extensive tests in Barth et al. [5]. For the second set, we use all 360 G and K stars in the CaT Library for each bin. The resulting V and σ differ by an average of ~ 5 km s^{-1} and the higher moments by ~ 0.01, all well within the measurement errors. Our kinematic moments determined from the CaII triplet region are therefore robust to template choices, similar to the findings in Barth et al. [5]. The stellar spectra of the CaT library cover the wavelength range of 8348-9020 Å with a spectral resolution of 1.5 Å FWHM.

The resulting stellar template broadened by the best-fit LOSVD is shown for each of the three example bins in Figure 2.2 (red curves). We use a bootstrap approach to determine the error bars on the kinematic moments of each LOSVD. For comparison, we have also estimated the errors using the standard Monte Carlo method with 100 trial spectra per bin. The bootstrapped errors on the kinematic moments are typically 50% to 100% larger than the Monte Carlo errors. See Sec. 4 of Ene et al. [53] for a detailed discussion.

[1]Note that the pPXF method described in Cappellari et al. 2004 [22] only allows $n = 2$, 4, or 6. The version described in Cappellari et al. 2016 [27] allows arbitrary n.

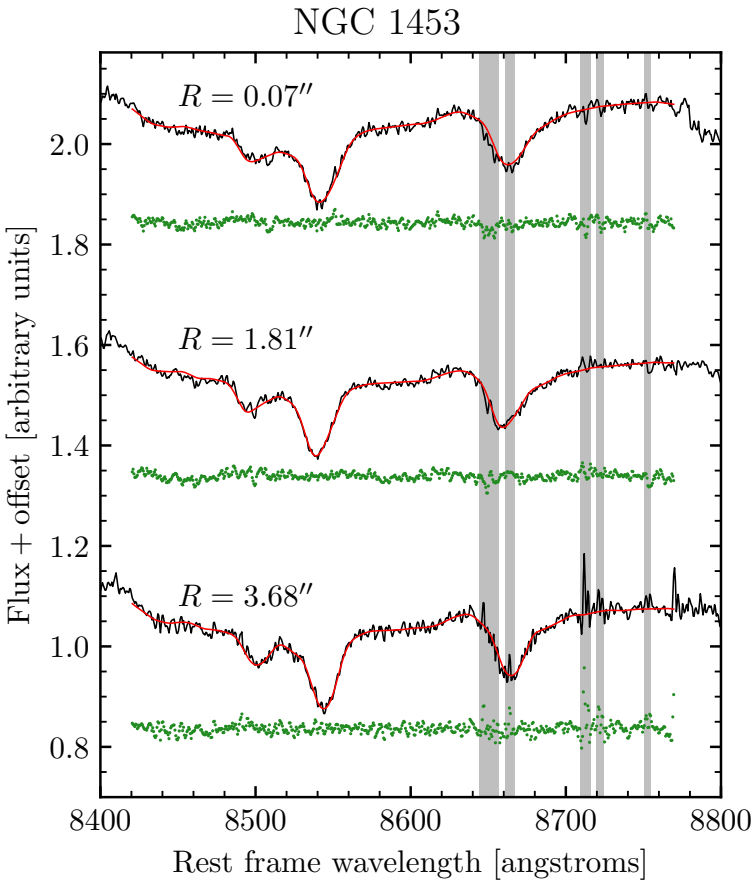

Figure 2.2: CaII triplet region of the Gemini GMOS IFS spectra (black) for three representative bins at three locations of NGC 1453: center with $S/N = 143$ (top), $1''.81$ from center with $S/N = 130$ (middle), and $3''.68$ from center with $S/N = 112$ (bottom). The stellar template broadened by the best-fit LOSVD (red) is overlaid on each observed spectrum. The fit is performed over the rest wavelength range of 8420-8770 Å centered around the CaII triplet absorption lines, excluding the grey shaded regions of improperly subtracted sky lines. The fit residuals (green dots) are shifted by an arbitrary amount for clarity.

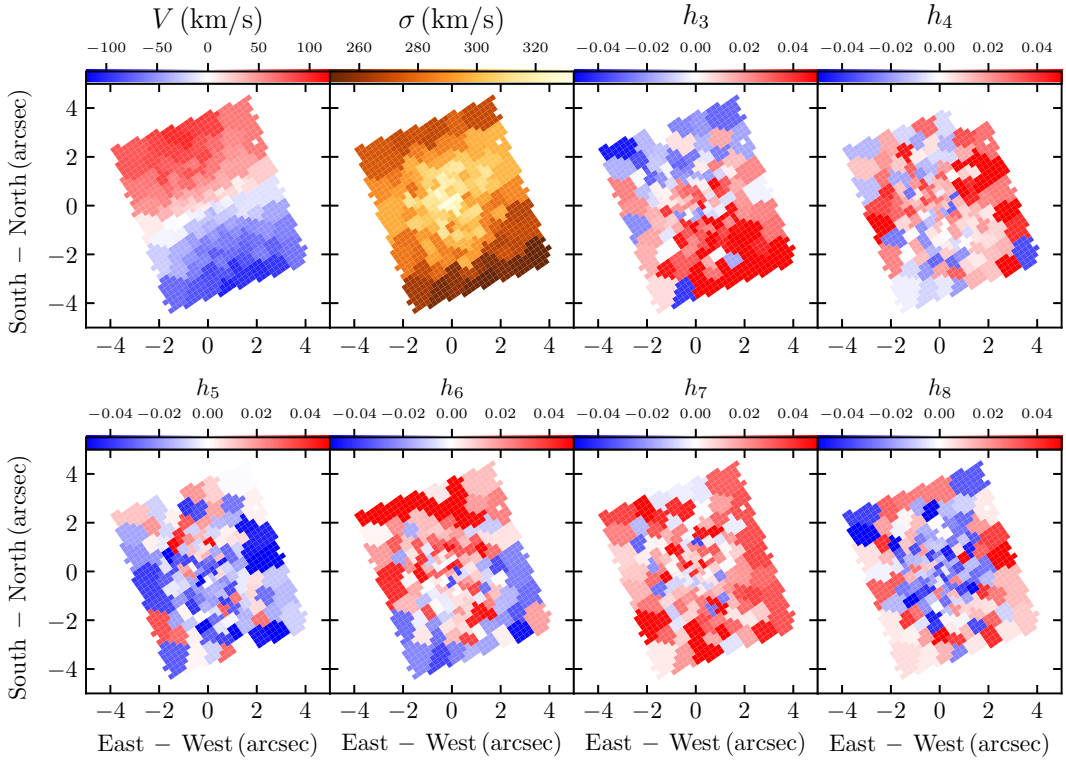

Figure 2.3: Maps of the stellar kinematics measured from the Gemini GMOS IFS over 135 spatial bins in the central $5'' \times 7''$ of NGC 1453. Each panel shows one of the eight velocity moments in the Gauss-Hermite expansion of the LOSVD: velocity V, velocity dispersion σ, and the h_3 to h_8 higher moments. The velocity map shows a regular rotation pattern with $|V|$ reaching ~ 100 km s^{-1} and the σ map shows a central peak. North is up and east is to the left.

The maps of the 8 kinematic moments, V, σ, h_3, ..., h_8, are shown in Figure 2.3. The velocity map shows a regular rotation pattern with $|V|$ reaching ~ 100 km s^{-1}, and the σ map shows a central peak of ~ 325 km s^{-1}. The mean errors are 7.1 km s^{-1} for V and 8.4 km s^{-1} for σ. The mean errors for h_3 through h_8 are quite similar, varying from 0.018 to 0.023. The radial profiles of these moments are shown below in Figure 2.4.

2.2.2 Wide-field kinematics

We observed NGC 1453 as one of the 100 MASSIVE galaxies in 2013 trimester 3, using the Mitchell/VIRUS-P IFS. The Mitchell IFS consists of 246 evenly spaced fibers with a one-third filling factor. Each fiber has a $4''.1$ diameter, and the IFS covers a large $107'' \times 107''$ FOV. Three dither positions of equal exposure time were used to obtain contiguous coverage

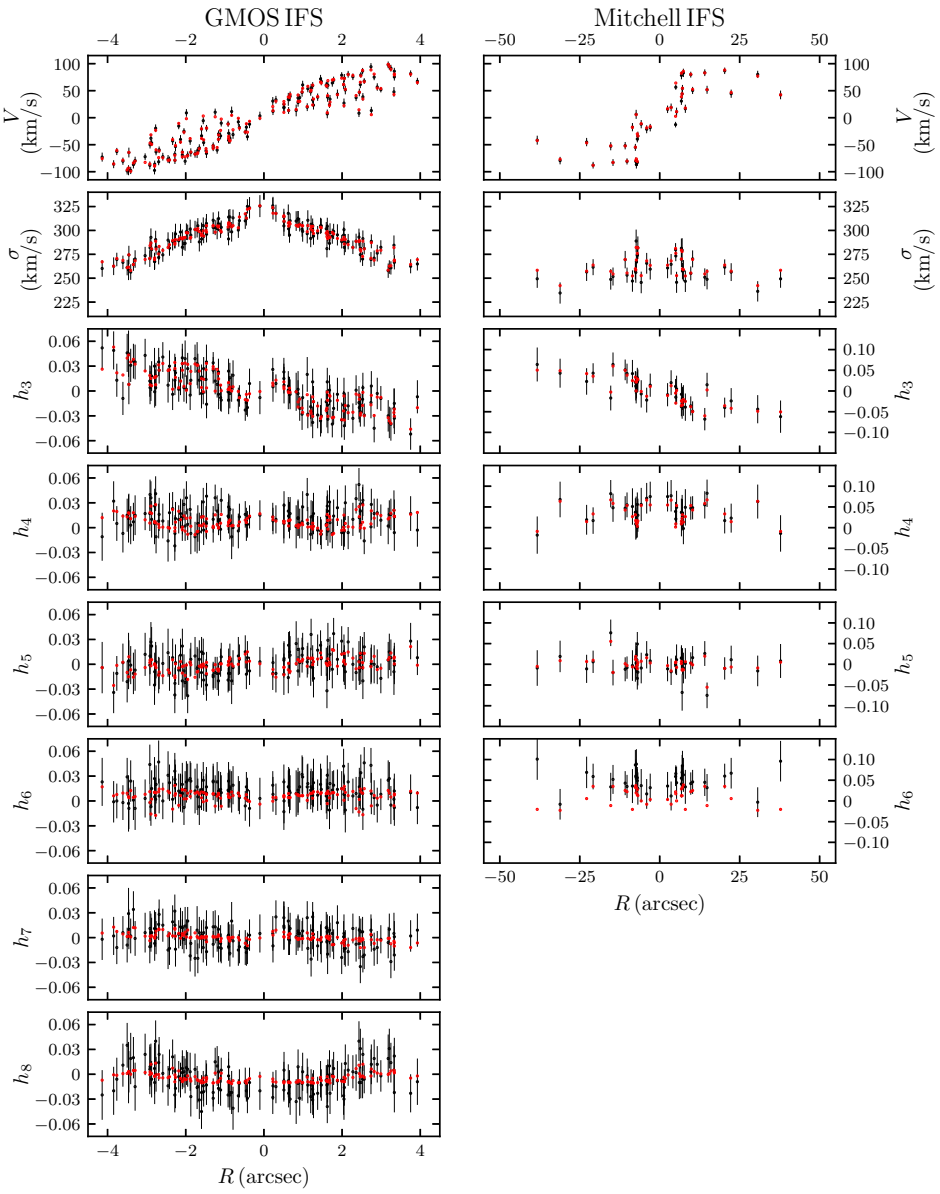

Figure 2.4: (Unfolded) radial profile of the stellar kinematics determined from GMOS (left) and Mitchell (right) observations (black), and kinematics predicted by the best-fit mass model (red) with black hole mass $M_{\rm BH} = 2.9 \times 10^9 M_\odot$, stellar mass-to-light ratio $M^*/L_{\rm F110W} = 2.09$ (in solar units), and enclosed dark matter halo mass (within 15 kpc) $M_{15} = 7 \times 10^{11} M_\odot$. The kinematic bins have been unfolded so that bins whose centers lie between $-90°$ and $+90°$ of the photometry PA are plotted with positive R and others are shown with negative R. The rotation in V and central values of σ are well-fit by this model, and the high moments $h_5 - h_8$ are close to 0 with some scatter.

of NGC 1453. We interleaved a 10-minute exposure on sky and two 20-minute exposures on target for a 2-hour total on-source exposure time. The spectral range spans 3650-5850 Å, covering the Ca HK region, the G-band region, Hβ, Mg I b, and several Fe absorption features.

Individual central fibers have S/N above 50, while the outer fibers are binned spatially to achieve a S/N threshold of 20 for the fainter outskirt of the galaxy. A similar procedure as in Sec. 2.2.1 is used to determine the stellar LOSVD for each of the 38 spatial bins. We used the MILES library of 985 stellar spectra ([166, 64]) as stellar templates and ran pPXF over the full library for each spectrum. Further details are described in Ma et al. [126] and Veale et al. [210].

As can be seen in Figure 2.4 here and Figure 21 of Ene et al. [53], the kinematic moments in the innermost Mitchell bins match well with the GMOS moments.

2.3 Photometric Data

To model the spatial distribution of the stellar component of NGC 1453, we use the IR imaging portion of the MASSIVE survey with the F110W filter of the *HST* Wide Field Camera 3 (WFC3/IR) [83]. The observations of NGC 1453 had a total exposure time of 2496 seconds, which was divided into five dithered exposures using a five-point subpixel dither pattern to improve the point spread function (PSF) sampling. The pixel scale at F110W is 0.128 arcsec pix^{-1} and is slightly undersampled for this wavelength. Details of the data reduction procedures, background sky measurement, mask construction, and isophotal fitting process were given in Goullaud et al. [83].

The isophotes of NGC 1453 are very regular (top panel of Fig. 2.5) with a mean ellipticity of 0.17 ± 0.001. The position angle (PA) changes with radius mildly from $(27.9 \pm 1.0)^\circ$ (east of north) at $1''$ to $36.1^\circ \pm 0.4^\circ$ at $79.5''$, with a luminosity-weighted average of $30.1^\circ \pm 0.2^\circ$. We fit the surface brightness using the Multi-Gaussian Expansion (MGE) method [51, 24] with a sum of 2D Gaussian components that share a common center and PA:

$$\Sigma(x', y') = \sum_{k=1}^{N} \frac{L_k}{2\pi\sigma_k'^2 q_k'} \exp\Big[-\frac{1}{2\sigma_k'^2}\Big(x'^2 + \frac{y'^2}{q_k'^2}\Big)\Big], \tag{2.2}$$

where x' and y' are projected coordinates measured from the galaxy center, with x' and y' being along the photometric major and minor axes, respectively. The subscript k labels the individual Gaussian components; L_k, σ_k', q_k' are the luminosity, projected width, and projected axis ratio of each Gaussian, respectively. To compare to WFC3 images, we convolve the model with a PSF composed of 5 nearly-circular gaussian components (with axis ratios > 0.98), obtained by fitting the PSF from Goullaud et al. [83]. The MGE fitting routine by default determines the PA using the central region of the galaxy. As a result, it chooses a PA of 28.5°, slightly different from the mean value 30.1° quoted in Goullaud et al. [83]. We repeated the MGE fit with the PA fixed to 30.1° and found a virtually identical fit. We choose to use 28.5°, the value from the MGE fitting routine.

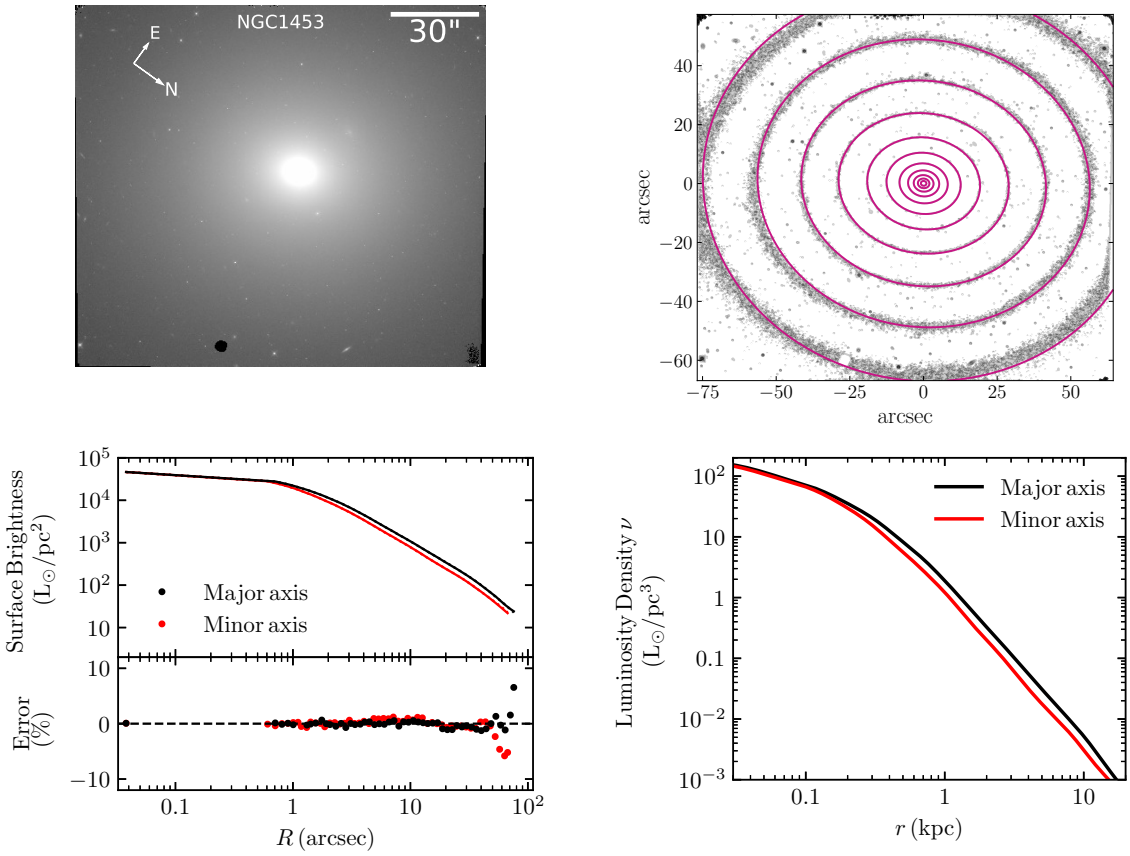

Figure 2.5: **(Upper left)** The F110W-band HST image of NGC 1453 used for our photometry [83]. The image is oriented so that the $+x$ axis lies at 30.1° east of north. **(Upper right)** Isophotes of the *HST* WFC3 IR image of NGC 1453 (black) and the best-fit MGE model (magenta). The isophotes have no measurable deviation from purely elliptical contours [83]. **(Lower left)** The surface brightness profiles along the major (black) and minor (red) axes are well-fit by the sum of 10 Gaussians with small fitting errors. The difference between the data (solid) and model (dotted) is not discernible in the plot, where the fractional error (lower half of the panel) is ∼1% except at large radii beyond 50″. **(Lower right)** Deprojected 3D luminosity density for an oblate axisymmetric model viewed edge-on for the best-fit MGE model.

Our best-fit MGE to the surface brightness of NGC 1453 consists of 10 Gaussian components, which are summarized in Appendix 2.A and plotted in Figure 2.5 (lower-left panel). The small fitting residuals (lower half of the panel) demonstrate that the MGE model agrees very well with the data. This MGE fit has an effective radius $R_e = 19\farcs6 \approx 4.8$ kpc, very similar to $R_e = 21\farcs9$ from Ene et al. [54] using our deep K-band photometry from CFHT.

The intrinsic and projected coordinate systems are related by a set of three viewing angles (θ, ϕ, ψ) [13]. The angles θ and ϕ specify how the line of sight is oriented relative to the principal axes of the galaxy, and ψ specifies the rotation of the galaxy around the line of sight, where an oblate axisymmetric potential is defined to have $\psi = 90°$. Given these viewing angles, an MGE fit to the light profile $\Sigma(x', y')$ can be deprojected into a 3D luminosity density $\nu(x, y, z)$ with x, y, z in the intrinsic coordinate system; see lower-right panel of Figure 2.5.

Dust was not observed in the central region of NGC 1453 in our WFC3 data. The mean optical and UV colors of NGC1453 are typical of those of evolved, red giant ellipticals of similar masses (e.g., [62, 124]). Annibali et al. [3] derived a mean age of 9.4 +/- 2.1 Gyr and metallicity [Z/H] = +0.22 dex within the central $3''$. Thus, the dominant stellar population is old and metal-rich. We find no significant gradient in the g-z color from PanSTARRS data [101].

2.4 Schwarzschild Orbit Models

We use the Schwarzschild orbit superposition method [173] through the implementation described by van den Bosch et al. [202]. In this method, a library of orbits with a wide range of initial conditions is constructed for a stationary potential due to a central black hole, a stellar component described by the MGE, and a dark matter halo. As each orbit passes through the region of the sky corresponding to a kinematic bin, its velocity is recorded to construct an LOSVD which is then decomposed in terms of Gauss-Hermite moments. A superposition of orbits is constructed with the QPB quadratic programming solver from the GALAHAD library [81], which minimizes the χ^2 associated with the kinematics under the constraint that both the projected mass within each aperture and the 3D mass distribution are fit within 1% of the MGE.

We have found several problems in the code during our tests and have fixed them as described in Chapter 3 [155]. We have also determined that additional modifications are required to achieve axisymmetry within the code. These changes are discussed briefly in the following subsections and more fully in Chapter 3 [155].

2.4.1 The axisymmetric limit

NGC 1453 is a fast rotator with regular elliptical isophotes (Fig. 2.5) and no significant misalignment between the projected rotation axis and photometric minor axis [54]. These properties suggest that NGC 1453 can be approximated as an oblate axisymmetric model.

However, we find the original version of the triaxial code by van den Bosch et al. [202] not to be able to achieve exact axisymmetry. Here we describe two precautions and one change that we implemented in order to achieve axisymmetry.

First, the box orbit library, which is generated by default in the original code, should be excluded when the code is to be used for axisymmetric gravitational potentials. Orbits in the box orbit library start from rest with $L_z = 0$. These orbits are important in triaxial potentials, but not in axisymmetric systems where L_z is an integral of motion. In an axisymmetric potential, box orbits cannot precess about the minor axis as they retain $L_z = 0$ for all time. As a result, they remain in their starting plane and do not exhibit axisymmetry. We therefore exclude these intrinsically non-axisymmetric orbits from our axisymmetric models[2].

The second precaution is to avoid generating long-axis tube orbits in the orbit library, a class of orbits not supported by axisymmetric stellar mass distributions. We find that these orbits can be eliminated only when the value for the viewing angle ψ is set to be sufficiently close to 90° in the input parameter file. As described in Section 2.3, the code uses three viewing angles (θ, ϕ, ψ) to relate the intrinsic and projected coordinate systems and to set the axis ratios of the stellar potential. An oblate axisymmetric potential is obtained when ψ is exactly 90° and the axis ratio p between the long and intermediate axes is exactly 1. Due to floating point imprecision, however, the code does not run when ψ is set to 90.0° with double precision. Earlier work typically chose $|\psi - 90°| = 10^{-3}$ or 10^{-2}, assuming these values were close enough to 90° to generate axisymmetry. For NGC 1453, however, we find even $|\psi - 90°| = 10^{-3}$ to be sufficiently far away from 90° to allow for long-axis tubes in the orbit start space, hence violating axisymmetry. We instead choose $\psi = (90 + 10^{-9})°$ in this work, which is far enough from 90° to avoid numerical issues but close enough that the potential is essentially axisymmetric for all available choices of the inclination.

Even after we excluded both the box and long-axis tube orbits, we still were unable to achieve axisymmetry with the original triaxial code. In the case of NGC 1453, we find that many orbits precess on timescales much longer than the default integration time, which is set to be 200 times the orbital periods in the code. These orbits should be symmeterized so that their contributions to the kinematics and mass grids are axisymmetric. To achieve this, we combine 40 copies of each orbit, each rotated slightly about the intrinsic minor axis of the galaxy; see Chapter 3 [155] for details of our implementation.

These changes allow the triaxial code to be properly run in the axisymmetric limit. We will compare results from the original code and our version in Section 2.7.2.

[2]For a triaxial potential, the "loop" orbit library generated by the code can contain some box orbits that have no overall sense of direction (see orbit start space in Fig. 2 of Schwarzschild 1993 [174]. However, as L_z is an integral of motion for an axisymmetric potential and all orbits in the library are initialized with $L_z \neq 0$, these orbits are also excluded.

2.4.2 Orbit library and phase space sampling

As described in Schwarzschild 1993 [174] and van den Bosch et al. [202], the orbits used for the models span a grid of energies (E) and starting positions (R, Θ) on the meridional plane of the galaxy. We choose 40 energies corresponding to the potential energies $\Phi(r, 0, 0)$ evaluated at a set of 40 radii that are logarithmically spaced between 0.01 and $10^{2.5}$ arcsec. These radii are chosen to span from roughly one order of magnitude below the pixel scale of our photometry to the radii where $\geq 99.999\%$ of the MGE mass is contained. We verify that orbits at the highest and lowest energies are given very low weight in the models. Our tests also verify that adding orbits starting at higher or lower radii does not impact our models. For each energy, we construct a grid of 9×9 starting positions spanning the radii between the inner and outer thin orbit radii for that energy and angles between 0° and 90°.

To improve the sampling of the phase space, the code allows orbit dithering where groups of orbits spanning a small volume in the (E, R, Θ) space are generated, combined, and given a single weight during orbit superposition. We use bundles of $5^3 = 125$ orbits for the final results below, and bundles of $3^3 = 27$ for numerous tests since they produce similar results and are less CPU-intensive. Our models also include a time-reversed copy of each orbit. In total, we use a library of 810,000 orbits (or 174,960 orbits for tests) for each mass model with $2 \times 40 \times 9 \times 9 = 6480$ independent weights.

As discussed in Section 2.4.1, we use $\psi = (90 + 10^{-9})°$ to run the triaxial code in the axisymmetric limit. In this limit, the ϕ viewing angle does not affect the MGE deprojection, but it sets the orientation of the plane of the orbit start space relative to the plane of the sky. As we axisymmetrize the orbits before projecting them onto the sky, the orientation between the start space and the sky becomes unimportant and we find that our model fits are independent of the viewing angle ϕ. We choose $\phi = 1°$. For reference, when the viewing angles $\psi = 90°$ and $\theta = 90°$ are used, the choices of $\phi = 0°$ and $\phi = 90°$ correspond to aligning the intrinsic x-axis and y-axis with the line of sight, respectively.

The potential due to the central black hole includes a softening length so that the potential at the origin is not singular. We set this length to 3×10^{-4}arcsec, which is roughly two and a half orders of magnitude smaller than the size of our central kinematic bin and one and a half orders of magnitude smaller than the peribothron of the most central orbits.

We convolve the integrated orbit trajectories in the models with PSFs while projecting the orbits onto the sky. This convolution is done separately for each kinematic dataset as they have different PSFs. For each, we assume a single circularly symmetric Gaussian with a FWHM of 0.″7 for the GMOS kinematics and 1.″2 for the Mitchell kinematics.

2.4.3 Input Gauss-Hermite moments

We use the first 12 moments in the Gauss-Hermite expansion of the LOSVDs as constraints in the orbit models. For the central region of NGC 1453, we use the first 8 moments V, σ, h_3, …, h_8 measured from the GMOS spectra as described in Section 2.2.1 and shown in Figure 2.3. The corresponding radial profile of each of the moments for all 135 GMOS

spatial bins is plotted (black points) in the left panel of Figure 2.4. The errors on h_3 through h_8 are quite similar from moment to moment and bin to bin. The mean errors on these moments range from 0.018 to 0.023, with a typical standard deviation of 0.003 over the spatial bins. To choose an appropriate number of moments to extract using pPXF, we performed the extraction with increasing numbers of moments (4, 6, 8). As the number of extracted moments is increased, we find that the typical value of the highest extracted moment becomes consistent with 0. For the GMOS spectra, this occured when 8 moments were extracted.

To prevent spurious behavior in the higher-order moments in the model, we further constrain the next four orders, h_9 to h_{12}, to be $0.0 \pm \delta$, where δ represents the typical errors in the higher moments. Since the size of errors is very similar from h_3 to h_8, we do not find the exact assigned values of δ to matter. Nonetheless, we try to mimic the mild bin-to-bin variations by assigning the measured errors for h_7 for a given bin to δ for the odd moments h_9 and h_{11} in that bin, and similarly for the even moments (i.e., using the h_8 errors for h_{10} and h_{12}).

For the wide-field data that have lower S/N, we use the first 6 Gauss-Hermite moments measured from the Mitchell spectra as constraints (Sec. 2.2.2). The radial profile of the moments for the 38 Mitchell spatial bins extending to a radius of $\sim 50''$ is shown in the right panel of Figure 2.4. We again constrain the 7th and 12th moments to be 0 with uncertainties equal to the measured errors for h_5 (for odd orders) or h_6 (for even orders). The errors on moments h_3 through h_6 from the Mitchell spectra are also quite uniform between moments. The mean errors on these moments range from 0.029 to 0.035 with a typical standard deviation of 0.006 over the spatial bins.

We discuss further the importance of constraining the higher Gauss-Hermite moments in Sec. 2.7.1 below.

2.5 Results: Mass Model Search

2.5.1 Mass Model

We investigate four mass model parameters – inclination θ, central black hole mass $M_{\rm BH}$, F110W-band stellar mass-to-light $M^*/L_{\rm F110W}$, and the enclosed mass of the dark matter halo at 15 kpc. We use a logarithmic halo with mass density

$$\rho_{DM}(r) = \frac{V_c^2}{4\pi G} \frac{3R_c^2 + r^2}{(R_c^2 + r^2)^2} \,. \tag{2.3}$$

We find the circular velocity V_c and the scale radius R_c to be highly degenerate for our data because the enclosed mass

$$M_{enc}(r) = \frac{V_c^2}{G} \frac{r^3}{r^2 + R_c^2} \tag{2.4}$$

scales with V_c^2/R_c^2 within the scale radius where most of our data points are located. We therefore choose to parameterize the halo with the enclosed mass within 15 kpc, M_{15}, where

15 kpc is the middle of the radial extent of the outermost Mitchell bins (spanning 9.4 kpc to 18.8 kpc).

2.5.2 Marginalization

Previous orbit modeling papers have often determined the 1σ (68%) and 3σ (95%) confidence intervals for each model parameter by finding the values at which the χ^2 rises by $\Delta\chi^2 = 1$ and 9 relative to the best-fit model. This method is only exactly correct when there is no covariance between the marginalized and free parameters and where the free parameter's χ^2 landscape is quadratic so that the likelihood is Gaussian. To avoid reliance on these assumptions, we compute best-fit values and confidence intervals through an interpolation and marginalization routine described in Appendix 2.B.

2.5.3 Inclination

In the oblate axisymmetric limit ($p = 1$), the MGE deprojection requires a single viewing angle, θ, which specifies the assumed inclination of the galaxy. An edge-on view of the system corresponds to $\theta = 90^\circ$ while face-on corresponds to $\theta = 0^\circ$. The inclination affects the axis ratios of the deprojected density distribution with $q_i = \sqrt{q_i'^2 - \cos^2\theta}/\sin\theta$, where q_i' is the observed axis ratio of the $i-$th component of the MGE fit described in Appendix 2.A, and q_i is the intrinsic axis ratio between short axis to the long axis in that component's deprojection.

Deprojection is only possible when $\cos^{-1}(\min q_i') < \theta < 90^\circ$, where $\min q_i'$ is the smallest axis ratio in the MGE fit. For the MGE used in this analysis, we have $\cos^{-1}(\min q_i') = 38^\circ$. When inclinations near this threshold are used, flattening of the the MGE component with the smallest q' changes significantly. For example, when $\theta = 40^\circ$, the component with $q' = 0.786$ has $q = 0.27$, and for $\theta = 50^\circ$, the component is flattened to $q = 0.59$.

To determine the inclination of NGC1453, we search coarsely over $M_{\rm BH}$, $M^*/L_{\rm F110W}$, and M_{15} but finely over the inclination. This grid includes 11 values of $M_{\rm BH}$ from 0.0 to $6 \times 10^9 M_\odot$ in steps of $6 \times 10^8 M_\odot$, 8 values of $M^*/L_{\rm F110W}$ from 1.6 to 2.3 (in solar units) in steps of 0.1, and 3 values of M_{15}: 2.8, 6.3, and 11.2 $\times 10^{11} M_\odot$, corresponding to $V_c = 400$, 600, and 800 km s^{-1} with $R_c = 15$ kpc. We use 12 values of θ from 40° to 89° in steps of 10° below 70° and 2.5° above. The code does not allow perfectly edge-on viewing angles, so the highest θ sampled was 89° rather than 90°. This grid contains $11 \times 8 \times 3 = 264$ models for each choice of θ and $264 \times 12 = 3168$ models in total.

Figure 2.6 shows that nearly edge-on viewing angles are strongly preferred. When $M_{15} = 6.3 \times 10^{11}\,{\rm M}_\odot$, $\theta = 89^\circ$ gave the lowest overall χ^2 with $\theta = 80^\circ$ and 70° lying $\Delta\chi^2 = 20.9$ and 22.3 higher. For each θ and M_{15}, we compute the best-fit $M_{\rm BH}$ and $M^*/L_{\rm F110W}$ with their 68% confidence intervals. When $\theta > 70^\circ$, the best-fit values depend only weakly on θ, and their confidence intervals coincide. This suggests that our recovered black hole mass and stellar mass-to-light ratio are relatively insensitive to the inclination within the edge-on

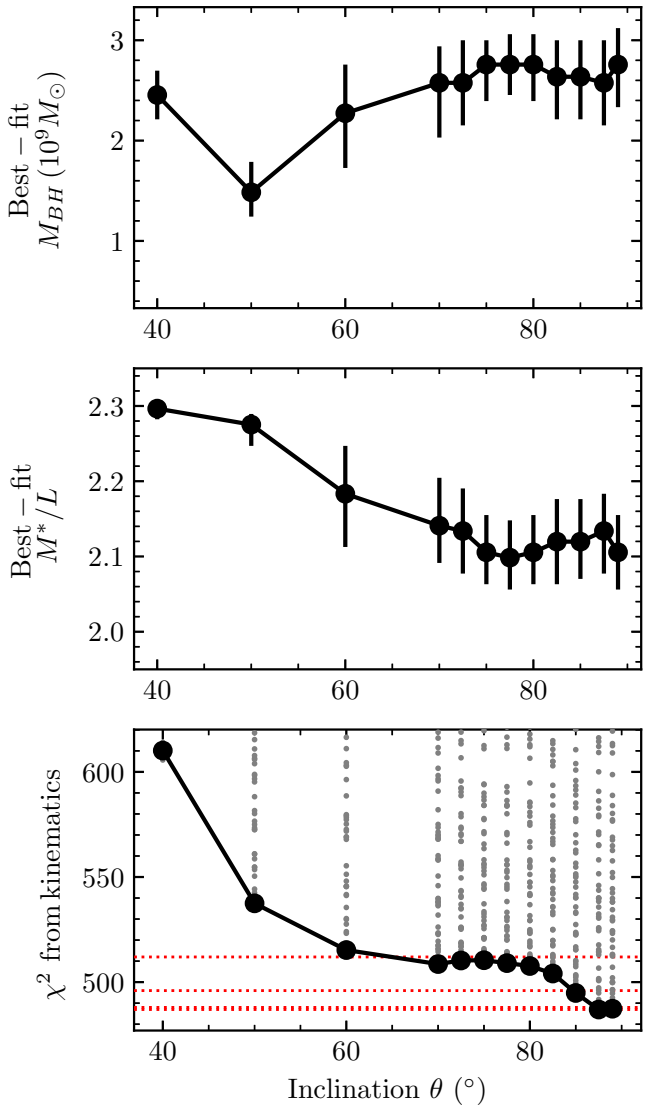

Figure 2.6: Best-fit $M_{\rm BH}$ (top) and $M^*/L_{\rm F110W}$ (middle) with 1σ confidence intervals and the corresponding marginalized 1-d χ^2 (bottom) as a function of the inclination angle θ. The grey points in the lower panel denote the χ^2 of individual models within the grid, and the red horizontal dashed lines denote the conventional $0, 1, 3, 5\sigma$ confidence levels corresponding to $\Delta\chi^2 = 0, 1, 9, 25$. The halo is fixed to $M_{15} = 6.3 \times 10^{11}\,{\rm M}_\odot$ in this plot for illustrative purposes; similar dependence is found for other halo masses we examined.

limit. We therefore fix the inclination to be nearly edge on with $\theta = 89°$ as we sample over halos below.

2.5.4 Black hole, stars, and dark matter halo

With the inclination fixed to be nearly edge-on with $\theta = 89°$, we search the three mass parameters, $M_{\rm BH}$, $M^*/L_{\rm F110W}$, and M_{15}, using two sets of grids. The primary grid covers the parameter ranges broadly and is then supplemented by a finer grid that zooms into the best-fit model of the primary grid with half the grid spacing in both $M_{\rm BH}$ and $M^*/L_{\rm F110W}$.

The primary grid has $16\times15\times13 = 3120$ models for $M_{\rm BH}$, $M^*/L_{\rm F110W}$ and M_{15}. This grid samples $M_{\rm BH}$ linearly from 0 to $6\times10^9 M_\odot$ in steps of $4\times10^8 M_\odot$, $M^*/L_{\rm F110W}$ from 1.60 to 2.30 (in solar units) in equal steps of 0.05, and the enclosed halo mass from $M_{15} = 2.79\times10^{11}\,{\rm M}_\odot$ to $11.16\times10^{11}\,{\rm M}_\odot$ by varying the circular velocity roughly linearly from $V_c = 400$ km s^{-1} to 800 km s^{-1} (for $R_c = 15$ kpc).

For the finer grid, we first determine the $M_{\rm BH}$ and $M^*/L_{\rm F110W}$ model that minimizes the χ^2 for each value of M_{15}. We then construct the fine grid around that model sampling another 16×15 values of $M_{\rm BH}$ and $M^*/L_{\rm F110W}$, where the spacing between models is half of that of the primary grid, and $M_{\rm BH}$ is sampled over a range of $3\times10^9 M_\odot$ in steps of $2\times10^8 M_\odot$, and $M^*/L_{\rm F110W}$ is sampled over a range of 0.35 in steps of 0.025. Many of these models overlap with those of the primary grid, so only $176\times13 = 2288$ additional models are run.

We perform the interpolation and marginalization described in Appendix 2.B to determine the best-fit values and uncertainties in $M_{\rm BH}$, $M^*/L_{\rm F110W}$ and M_{15} from these 5408 models. The resulting χ^2 landscapes are displayed in Figure 2.7. We find the best-fit mass parameters to be $M_{\rm BH} = (2.9\pm0.4)\times10^9 M_\odot$, $M^*/L_{\rm F110W} = (2.09\pm0.06) M_\odot/L_\odot$, and $M_{15} = (7.0\pm0.7)\times10^{11} M_\odot$. For comparison, if the best-fit parameters are chosen by finding the range of models where the χ^2 rises by $\Delta\chi^2 \le 1$ from the minimum value, as was frequently done in prior $M_{\rm BH}$ papers, we find comparable central values for the mass parameters for NGC 1453, but the error bars are underestimated by a factor of 1.5 to 2: $M_{\rm BH} = (3.0\pm0.2)\times10^9 M_\odot$, $M^*/L_{\rm F110W} = (2.06\pm0.03) M_\odot/L_\odot$, and $M_{15} = (7.4\pm0.4)\times10^{11} M_\odot$. These values are tabulated in Table 2.1.

Figure 2.8 presents a clear view of the χ^2 landscape over the wide range of $M_{\rm BH}$ covered by our grid. It shows that models with small black hole masses are highly disfavored. In particular, $M_{BH} = 0$ has $\Delta\chi^2 = 75.5$ above the minimum, corresponding to the 8.7σ confidence level. This result will be further discussed in Section 2.7.2.

In the best-fit mass model for NGC 1453, the enclosed stellar mass is equal to $M_{\rm BH}$, $2M_{\rm BH}$, $3M_{\rm BH}$ and $5M_{\rm BH}$ at radius 0.18 kpc (0.74″), 0.26 kpc (1.05″), 0.33 kpc (1.32″), and 0.45 kpc (1.83″), respectively. The enclosed stellar mass equals that of the dark matter at 8.4 kpc (34.1″). At the effective radius (5 kpc), the dark matter fraction is 0.27.

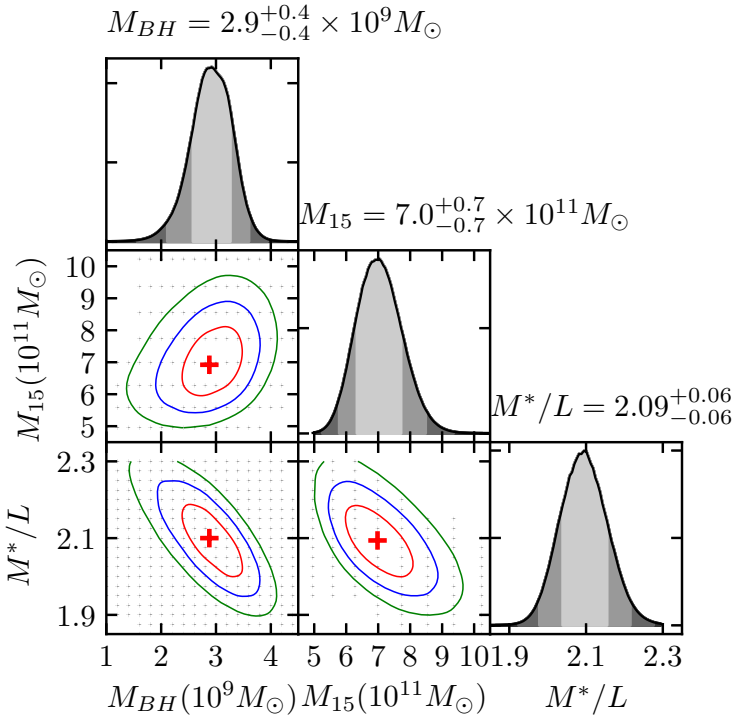

Figure 2.7: Marginalized 1-d and 2-d likelihood distributions from the grids of $M_{\rm BH}$, M_{15} and $M^*/L_{\rm F110W}$ described in Section. 2.5.4. The 1σ, 2σ, and 3σ confidence intervals corresponding to the 68, 97, and 99.5 percentile confidence levels are shown as red, blue, and green curves in the 2-d panels and as different shade of grey in the 1-d panels. The extracted best-fit values and 1σ confidence interval are shown above each 1-d panel.

Table 2.1: Best-fit values of the black hole mass, stellar mass-to-light in the F110W band, and dark matter mass enclosed within 15 kpc. The center column presents values determined though interpolation and marginalization as described in Appendix 2.B. The right column presents values determined through projection, where the confidence interval bounds all models within $\Delta\chi^2 \leq 1$ of the global minimum.

Mass parameters	Marginalized	Projected
$M_{\rm BH}\,(10^9 M_\odot)$	2.9 ± 0.4	3.0 ± 0.2
$M^*/L_{\rm F110W}\,(M_\odot/L_\odot)$	2.09 ± 0.06	2.06 ± 0.03
$M_{15}\,(10^{11} M_\odot)$	7.0 ± 0.7	7.4 ± 0.4

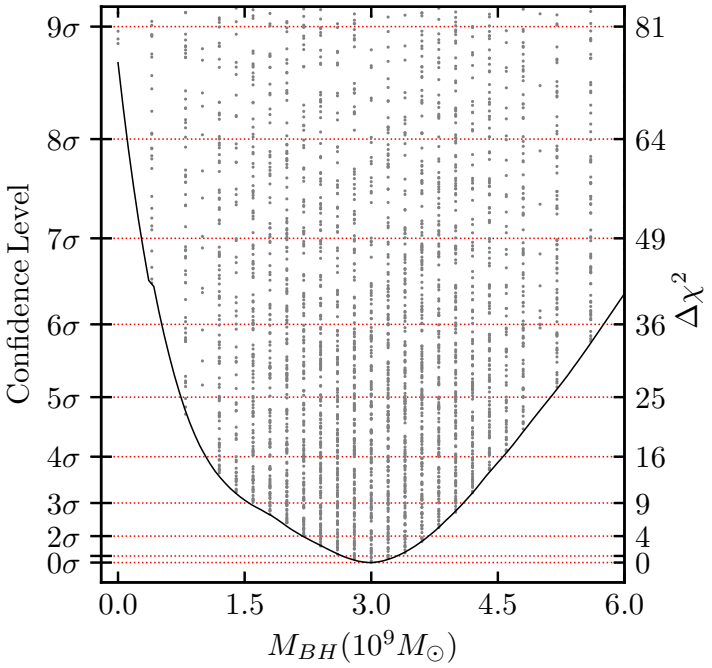

Figure 2.8: χ^2 as a function of $M_{\rm BH}$ for the full range of $M_{\rm BH}$ explored in this paper. The χ^2 is obtained by marginalizing over the other two mass parameters, M_{15} and $M^*/L_{\rm F110W}$, as described in Appendix 2.B. Models with $M_{\rm BH} = 0$ are highly disfavored with a $\Delta\chi^2 = 75.5$ relative to the best-fit $M_{\rm BH}$, corresponding to the 8.7σ confidence level.

2.6 Results: Best-Fit Mass Model

2.6.1 Stellar kinematics

Our best-fit mass model (red points in Fig. 2.4) provides an excellent fit to the observed stellar kinematics (black). Both the rotation V and the large central σ are well captured by this model. The total χ^2 for the best-fit model from all the kinematic moments is 493.0, where the bulk of this (471.5) comes from the moments extracted from data and only a small fraction (21.5) comes from the additional high moments that are constrained to be zero.

To estimate the reduced χ^2, we note that there are 8 measured moments for each of the 135 GMOS bins and 6 measured moments for each of the 38 Mitchell bins, for a total of 1308 data points. The kinematic maps of the odd moments have been point-anti-symmetrized and the even moments have been point-symmetrized according to the prescription in Appendix A of van den Bosch and de Zeeuw [199]. Our total reduced χ^2 from the moments extracted from data is therefore $471.5/1308 = 0.36$. The reduced χ^2 from the GMOS data alone is

$366.2/(135 * 8) = 0.34$ and the reduced χ^2 from Mitchell alone is $105.3/(38 * 6) = 0.46$. The high moments which were constrained to be zero have an associated reduced χ^2 of $21.5/768 = 0.03$.

2.6.2 Orbital structure

While computing the orbit libraries, the code constructs a 3D spherical grid containing the first and second velocity moments of the orbits. We use this velocity grid to compute the anisotropy parameter $\beta = 1 - \sigma_t^2/\sigma_r^2$ and the ratio of radial to tangential dispersions σ_r/σ_t. We note that various definitions of β have been adopted in prior papers, and at times it is unclear whether σ in such quantities is treated as a dispersion or a second moment of the velocity, i.e., whether $\sigma^2 = \langle v^2 \rangle - \langle v \rangle^2$ or $\sigma^2 = \langle v^2 \rangle$. We choose to define

$$\sigma_t^2 = \frac{\sigma_\theta^2 + \sigma_\phi^2}{2}\,, \qquad \beta = 1 - \frac{\sigma_t^2}{\sigma_r^2}\,,$$
$$\sigma_{rot}^2 = \frac{\sigma_\theta^2 + \sigma_\phi^2 + \langle v_\phi \rangle^2}{2}\,, \qquad \beta_{rot} = 1 - \frac{\sigma_{rot}^2}{\sigma_r^2}\,,$$

where the brackets denote a mass-weighted mean over θ and ϕ. These pairs of definitions are only expected to differ when there is significant contribution from the ordered flow velocity term $\langle v_\phi \rangle^2$. For reference, differing definitions and symbols were used in the literature, e.g., β from Thomas et al. [191], β_r from Peletier et al. [151], and σ_r/σ_t from Walsh et al. [220] all excluded the $\langle v_\phi \rangle^2$ term, while β_{rot} from Krajnović et al. [110] and Thomas et al. [191] and σ_r/σ_t from Gebhardt et al. [76] included this term.

The resulting velocity anisotropy as a function of spherical radius r for the best-fit model of NGC 1453 is shown in Figure 2.9. The orbits are tangential near the core, but become increasingly radially anisotropic beyond the effective radius (~ 5 kpc) Even though NGC 1453 exhibits rotation and is considered a fast rotator for an ETG, the maximal velocity observed in our kinematics is $\sim$100 km s^{-1}, which is much below the dispersion σ shown in Figure 2.4. The term $\langle v_\phi \rangle^2$ therefore has negligible impact on the value of σ_{rot} and β_{rot} at all radii, and $\sigma_{rot} \approx \sigma_t$ and $\beta \approx \beta_{rot}$ at all radii.

Thomas et al. [191] studied eleven massive elliptical galaxies with axisymmetric Schwarzschild models. Six of those galaxies had stellar cores and exhbited strongly tangential anisotropies ($\beta_{rot} < -0.5$) in the core regions and highly radial anisotropies ($\beta_{rot} \sim 0.5$) well outside the cores. Similar trends in the anisotropy were found in MASSIVE survey galaxy NGC 1600 [192]. This behavior is consistent with gravitational core scouring, where a central binary black hole preferentially ejects radial orbits from the core leaving an orbital structure which is tangentially biased [6]. We observe similar behavior in NGC 1453 suggesting that its core may have also been depleted through core-scouring.

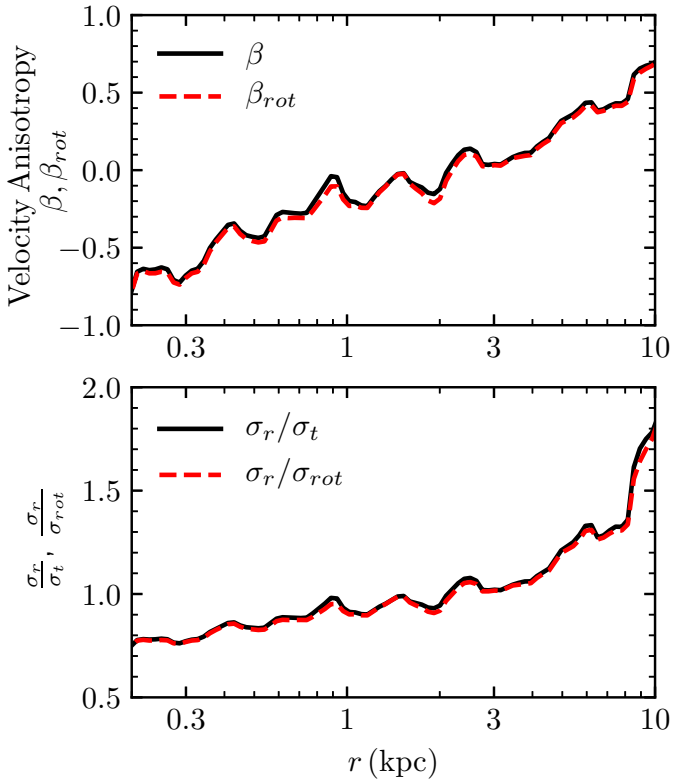

Figure 2.9: Velocity anisotropy β (top) and σ_r/σ_t (bottom) as a function of radius for the best-fit model of NGC 1453. The orbits within the central ~ 1 kpc are preferentially tangential with $\sigma_r/\sigma_t < 1$ and $\beta < 0$. The orbits become increasingly radial beyond the effective radius (≈ 5 kpc) Over all radii, β traces β_{rot} and σ_t traces σ_{rot} because $\langle v_\phi \rangle^2 \ll \sigma_\phi^2$.

2.7 Discussion

2.7.1 Gauss-Hermite series truncation and LOSVDs

As described above, the stellar LOSVD in each spatial bin is parameterized by a Gauss-Hermite series up to order n. Some care must be taken to ensure that the unconstrained higher moments beyond order n in the orbit models do not introduce spurious behavior in the predicted LOSVDs.

It is useful to begin the discussion by examining how the LOSVDs are obtained in the triaxial orbit code. During orbit integration, the code first computes the LOSVD of each orbit for a spatial bin as it passes through the aperture on the sky. The Gauss-Hermite moments are then determined from each LOSVD through the direct integration described by van der Marel and Franx [132], using the observed V and σ values for that bin. During the subsequent orbital weight finding process, the Gauss-Hermite moments of the superposition of orbits in each bin are easily computed due to their linearity.

To find the best-fit orbital weights for a mass model, the code places no constraints on moments beyond those provided to it. We are concerned that some of the unconstrained higher moments can add spurious non-Gaussian features to the LOSVDs. This is to be contrasted with how the LOSVDs are determined from the data. There, the pPXF algorithm determines the best-fit moments from the observed spectrum in the least-squares sense, choosing moments that minimize the residual contribution from higher moments.

To test the impact of unconstrained higher moments on $M_{\rm BH}$, we perform a series of controlled experiments in which we vary systematically the number of Gauss-Hermite moments determined from the GMOS spectra and used as constraints in the orbit model. We compare the results from grid searches for four cases here. For the first three cases, 4, 6, and 8 GH moments are measured from the GMOS data with pPXF and those 4, 6, and 8 GH moments are fit with orbit models to infer $M_{\rm BH}$. In the fourth case, 8 GH moments are determined from the GMOS data with pPXF and 12 GH moments are used as inputs into the dynamical models, with the 9th to 12th moments set to 0 and assigned uncertainties as described in Sec. 2.4.3. The fourth case, where 12 moments are used to constrain the dynamical models, corresponds to our production run reported in earlier sections. For each case, we then perform a grid search for the best-fit $M_{\rm BH}$ and $M^*/L_{\rm F110W}$. Our aim here is to test the effects on the measured $M_{\rm BH}$, so we keep the large-scale Mitchell kinematics unchanged and fix the halo to the best-fit value of $M_{15} = 7 \times 10^{11} M_\odot$ from our production run.

The resulting first 12 moments predicted by the best-fit model for each of the four cases of increasing truncation orders are shown in Figure 2.10. The corresponding LOSVDs for three representative GMOS bins in each case are shown in Figure 2.11. The marginalized χ^2 versus $M_{\rm BH}$ for the four cases are shown in Figure 2.12 (left panel), and the best-fit $M_{\rm BH}$ are listed in Table 2.2 under "Berkeley Version". We note that while the best-fit $M_{\rm BH}$ changes by only $\sim 20\%$ (in the range of 2.6 to $3.2 \times 10^9\,{\rm M}_\odot$) as the truncation order is varied, the confidence level is improved significantly when more input moments are used, and the errors

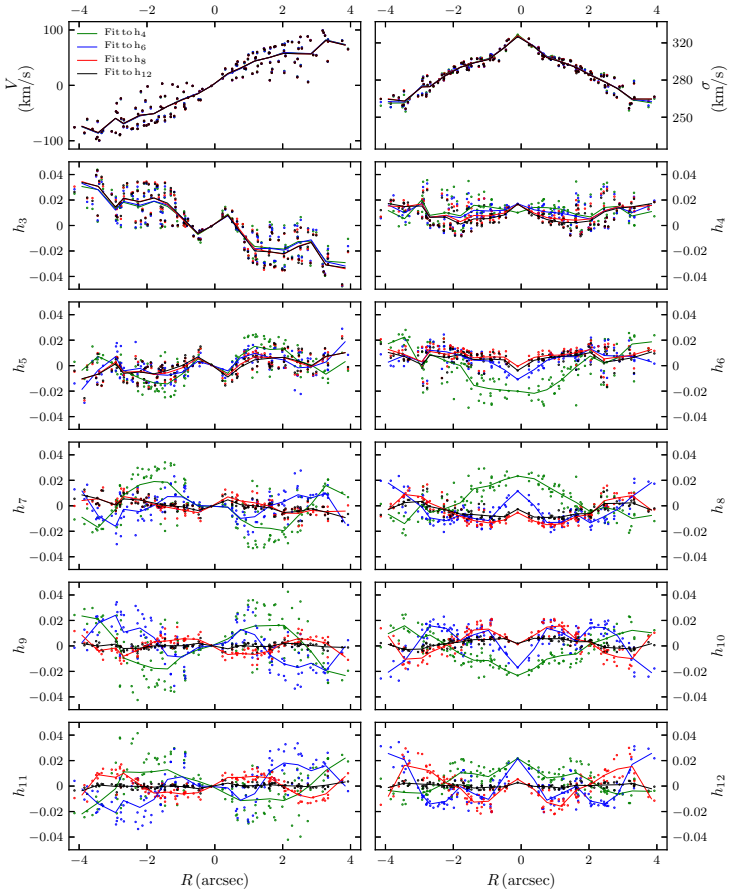

Figure 2.10: Dependence of the first 12 Gauss-Hermite moments predicted by the best-fit orbit model on the assumed truncation order applied to the GMOS data and used as input constraints. The four colors show the progression of increasing truncation order: up to h_4 (green), h_6 (blue), h_8 (red), and h_{12} (black; our production run). In each case, the moments beyond the truncation order are unconstrained in the orbit model and exhibit correlated deviations from 0. This is most clearly seen in the green curves in the unconstrained h_5 and above. An interpolating line has been added to each curve to guide the eye. The corresponding marginalized χ^2 versus $M_{\rm BH}$ for the four cases are shown in Figure 2.12 (left panel). See text in Sec. 2.7.1 for details.

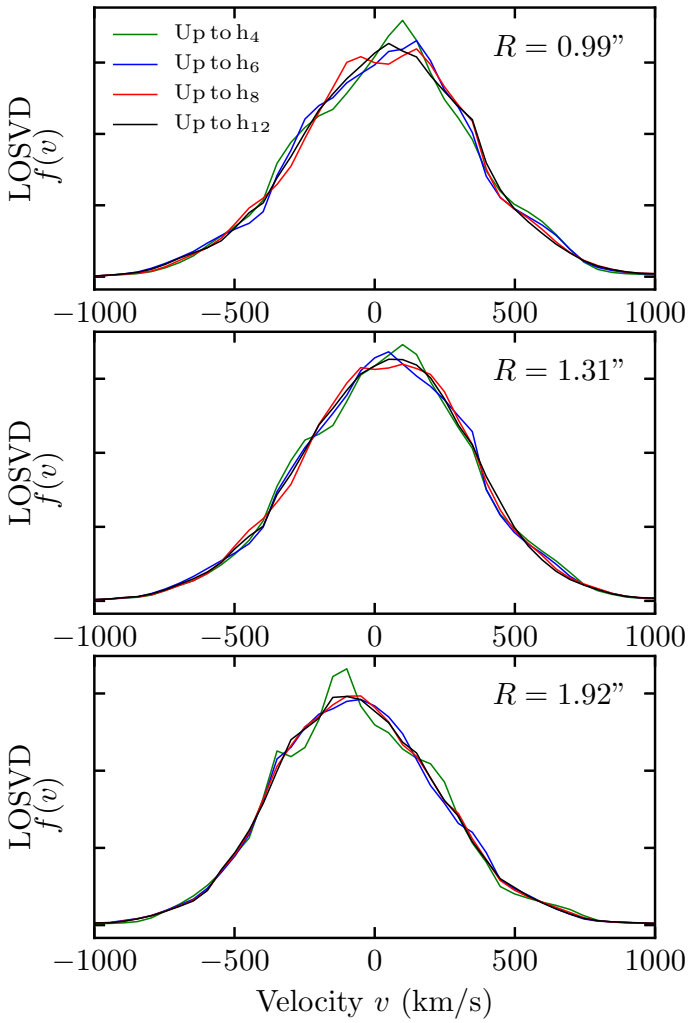

Figure 2.11: LOSVDs for three representative GMOS bins predicted by the best-fit orbit models. Each panel compares the LOSVDs from the four models described in Sec. 2.7.1 and shown in Figure 2.10, where the number of Gauss-Hermite moments fit by the dynamical models varies from 4 (green), 6 (blue), 8 (red), to 12 (black). When only 4 moments are constrained, the LOSVDs have the most pronounced irregular features due to the unconstrained h_5 and beyond seen in Figure 2.10. These unobserved features are gradually reduced when higher moments are used to constrain the model.

on $M_{\rm BH}$ are reduced by a factor of ~ 2.3 when we increase the truncation order from 4 to 12.

The detailed dependence of each of the 12 Gauss-Hermite moments on the truncation order can be clearly seen in Figure 2.10. These moments are determined from the LOSVD of each bin (e.g., Fig. 2.11) through the direct integration described by van der Marel and Franx [132]. Although only moments up to the truncation order are used for constraining the model LOSVDs, arbitrary higher moments can be computed. The lowest 4 moments V, σ, h_3 and h_4 predicted by the best-fit models are mostly independent of the truncation order we tested. This is not surprising since these 4 moments are fit during modelling in all cases. The predicted moments beyond h_4, however, start to show varying degrees of deviations. The case in which the series is truncated at h_{12} (black points) corresponds to our production run. It uses all 12 moments as constraints by design, so as expected, the best-fit model is well-behaved in all 12 panels. In comparison, when only 4 moments are fit by the orbit-based models (green points), the unconstrained 5th moment and beyond deviate strongly from the black points. Similarly, when 6 (blue) or 8 (red) moments are used as constraints during the modelling, the 7th or 9th moment and beyond also show deviations from the black points. Importantly, the deviation from 0 is not random; instead, the unconstrained moments are correlated spatially, being somewhat symmetric about $R = 0$" for even moments and antisymmetric about $R = 0$" for odd moments. The general trend that we observe in Figure 2.10 is that the lower the truncation order is, the more their higher moments show unobserved and correlated features.

We illustrate the spurious features in the shapes of the LOSVDs resulting from the unconstrained higher Gauss-Hermite moments in Figure 2.11. For all three representative GMOS bins shown, the model LOSVDs have the most pronounced irregular features when only 4 moments are used (green curve), and these features gradually go away as the truncation order is increased.

To date, a number of published dynamical $M_{\rm BH}$ measurements based on orbit modeling of stellar kinematics have used the method of Gauss-Hermite expansion to approximate the LOSVDs. Most have used the first four moments as constraints in the orbit models, e.g., [213, 199, 201, 219, 220, 217, 218, 1, 176, 110, 188, 189], while a few have used the first six moments, e.g., [23, 29, 111]. Our tests here are applied only to the triaxial Leiden code in the case of NGC 1453, so we cannot speak directly to the impact of higher Gauss-Hermite moments on $M_{\rm BH}$ in other work. However, we recommend that similar tests be performed in future work.

2.7.2 The $M_{BH} = 0$ model: comparison to Ahn et al. (2018)

Recently, Ahn et al. [1] used the same triaxial orbit code as in this work to perform axisymmetric mass modeling of the ultracompact dwarf galaxy M59-UCD3. They reported a puzzling global χ^2 minimum at $M_{\rm BH} = 0$, which was inconsistent with their best-fit $M_{\rm BH} = 4.2^{+2.1}_{-1.7} \times 10^6 M_\odot$ from Jeans modeling and an orbit code that is intrinsically axisymmetric. Various tests were performed but none explained the discrepancy. They speculated

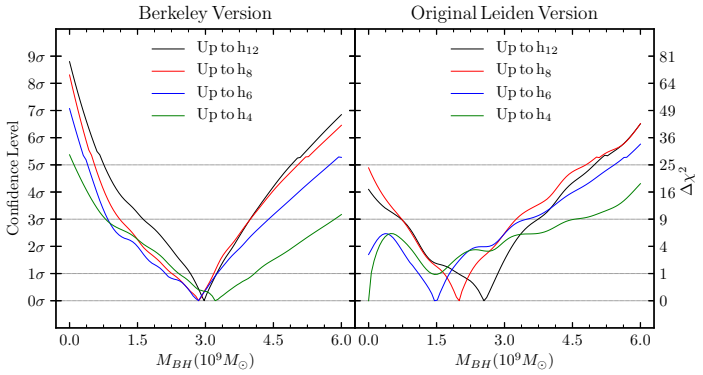

Figure 2.12: Illustration of the increasing constraints on $M_{\rm BH}$ provided by progressively higher Gauss-Hermite moments used to represent the LOSVDs. Two versions of the orbit code with different settings are shown: the Berkeley version described in Sec. 2.4.1 (left), and the original triaxial Leiden version (right) with typical choice of ψ, including box orbits, and without orbit axisymmetrization. In each panel, we vary systematically the number of Gauss-Hermite moments used as constraints in the orbit model from 4 (green), 6 (blue), 8 (red), to 12 (black). Our production run corresponds to the black curve in the left panel. The green curve in the right panel uses a similar setting as in Ahn et al. [1] for M59-UCD3 and prefers $M_{\rm BH} = 0$. The $M_{\rm BH} = 0$ minimum disappears as moments beyond h_4 are applied in the original code (right) but the best-fit $M_{\rm BH}$ is highly dependent on the truncation order. In comparison, the main effect of additional moments in the Berkeley version (left) is to tighten the error bars on $M_{\rm BH}$ while leading the central value largely unchanged. Note we use a linear scale in the confidence level for the y-axis here for a clearer illustration of the locations of the minima, while Fig. 2.8 uses a linear scale in χ^2.

about a "numerical artifact" in the triaxial Leiden code and favored the non-zero $M_{\rm BH}$ from Jeans and axisymmetric orbit modeling.

As discussed in Sec. 2.4 and 2.7.1, we have made a number of changes to the original triaxial code and typical settings to arrive at the "Berkeley version" results for NGC 1453 presented in Sec. 2.5 and 2.6. Even though our final outcome in Figure 2.8 shows $M_{\rm BH} = 0$ to be disfavored at the 8.7σ confidence level, we also encountered difficulties in constraining $M_{\rm BH}$ in the case of NGC 1453 when we ran the original triaxial code using similar settings as those of Ahn et al. [1]; that is, choosing $|\psi - 90°| \geq 10^{-3}$, including box orbits from the default library, not axisymmetrizing the loop orbits, and using 4 Gauss-Hermite moments as kinematic constraints. Our resulting χ^2 for this setting using the original code is represented by the green curve in the right panel of Figure 2.12. The overall constraint on the NGC 1453 $M_{\rm BH}$ is weak, with the lowest χ^2 occurring at $M_{\rm BH} = 0$, and another local χ^2 minimum at $M_{\rm BH} \sim 1.5 \times 10^9 M_\odot$. This is in stark contrast to the result from our version of the settings and code represented by the black curve in the left panel of Figure 2.12.

Table 2.2: Best-fit black hole mass and 1σ (68%) confidence intervals for the eight cases shown in Figure 2.12. The 4-moment Leiden run results in a χ^2 minimum at $M_{\rm BH} = 0$. For all other runs, the quoted confidence intervals here are properly marginalized through the routine described in Appendix 2.B.

# of constrained moments	Berkeley Version $M_{\rm BH}$ ($10^9 M_\odot$)	Original Leiden Version $M_{\rm BH}$ ($10^9 M_\odot$)
4	3.30 ± 0.81	0
6	2.64 ± 0.59	1.51 ± 0.49
8	2.63 ± 0.48	1.93 ± 0.41
12	2.91 ± 0.35	2.22 ± 0.55

In view of the importance of constraining higher Gauss-Hermite moments (Sec. 2.7.1), we have run further tests using the original code but increasing the number of input moments from 4 to 6, 8, and then 12. The results are plotted in the right panel of Figure 2.12. The χ^2 minimum at $M_{\rm BH} = 0$ in the case of h_4 disappears as the truncation order is increased, but the location of the χ^2 minimum depends sensitively on the number of moments, and the best-fit $M_{\rm BH}$ increases monotonically and shows no convergence even at order 12, as listed in Table 2.2. In comparison, models from the "Berkeley Version" in the left panel of Figure 2.12 and in Table 2.2 have better-behaved χ^2 contours.

2.7.3 Comparison to Jeans modeling

In Ene et al. [54], we applied the method of Jeans Anisotropic Modeling (JAM; [18]) to determine the mass parameters in NGC 1453. JAM is computationally cheap but is limited by the assumptions of axisymmetric potentials and cylindrically aligned velocity ellipsoids, and by the fact that its solutions could be unphysical. JAM has been shown to give consistent results as axisymmetric orbit models for regular fast rotators like NGC 1453 [28].

Previous studies with JAM have typically assumed a globally constant $\beta_z = 1 - \langle v_z^2 \rangle / \langle v_R^2 \rangle$, which quantifies the flattening of the velocity ellipsoid along the minor axis. In order to at least partially replicate orbit-type variation, we allowed two different values for β_z, one for the Gaussian components with $\sigma_k < 1''$, and the other for the Gaussian components with $\sigma_k > 1''$. The choice of $1''$ is motivated by the light profile of NGC 1453 which starts to fall off more rapidly at $R \gtrsim 1''$ (see bottom panel of Fig. 2.5).

The JAM Modelling in Ene et al. [54] used the distance 56.4 Mpc from the 2MASS redshift survey. After adjusting to 51.0 Mpc, the best-fit parameters from JAM are $M_{\rm BH} = (2.98 \pm 0.23) \times 10^9\ M_\odot$ and $M^*/L_{\rm F110W} = 2.28 \pm 0.04$ (with 1σ errors). The $M_{\rm BH}$ value is within the 1σ confidence interval of our best-fit value, while the $M^*/L_{\rm F110W}$ is higher than our best-fit value but is consistent with the 3σ interval.

The best-fit circular velocity for the dark matter halo from JAM is $V_c = 364 \pm 45$ km s^{-1} with $R_c = 6.5 \pm 2.5$ kpc. Assuming that the uncertainties in these two parameters are highly correlated, this corresponds to $M_{15} = (3.89 \pm 0.96) \times 10^{11} M_\odot$. This is roughly half

our preferred value of $M_{15} = (6.98 \pm 0.73) \times 10^{11} M_\odot$. At small radii, the enclosed dark matter in the JAM model is much larger than ours. The central density of JAM's halo is $(17.4 \pm 4.3) \times 10^7 M_\odot/\mathrm{kpc}^3$, while ours is half at $(9.9 \pm 1.0) \times 10^7 M_\odot/\mathrm{kpc}^3$. At 6.9 kpc, the enclosed masses of the two halos are identical.

The best-fit velocity anisotropy is $\beta_z^G(\sigma_k' < 1'') = -0.58 \pm 0.62$ for the inner part and $\beta_z^G(\sigma_k' > 1'') = 0.15 \pm 0.04$ for the outer part. The anisotropy in the central region is comparable to what we find in Section 2.6.2, but we find the the orbits to be much more radially biased in the outer region. The black hole and stellar mass distribution of the JAM best-fit model and our best-fit model are very similar, suggesting that similar velocity anisotropies are required to fit the kinematics. Conversely, the enclosed mass due to the halo in our best-fit model is much larger than that of JAM beyond 6.9 kpc, suggesting that our model must be more radially biased to similarly fit the kinematics, as we observe.

2.7.4 Black Hole Scaling Relations

The SMBH at the center of NGC 1453 lies 0.32 dex above the mean M_{BH}-σ scaling relation from McConnell and Ma [133], which is within the 0.38 dex intrinsic scatter in that relation. For the other seven MASSIVE galaxies with stellar dynamical M_{BH}, four galaxies – NGC 4472, NGC 4486, NGC 4649 and NGC 7619 – are within 0.3 dex of the scaling relation, whereas the other three – NGC 1600, NGC 3842, and NGC 4889 – have M_{BH} that is overmassive by a factor of $\sim 3-6$ than predicted by their respective galaxy velocity dispersion. These 8 galaxies exhibit similarly large scatter in the scaling relation between M_{BH} and the bulge stellar mass, where M_* spans a factor of ~ 3 while M_{BH} spans a factor of ~ 10.

The stellar core radius of NGC 1453 from our photometry is $r_b = 0.97''$ (0.24 kpc). This value is obtained by fitting a 2D core-Sérsic profile, convolved with the PSF from Goullaud et al. [83]. This fit was performed using Imfit [56]. The scaling relation between M_{BH} and r_b for a sample of 21 massive cored ETGs is found to be $\log_{10} M_{\mathrm{BH}} = 10.27 + 1.17 \log_{10}(r_b/\mathrm{kpc})$ with an intrinsic scatter of 0.29 dex [192]. Our inferred M_{BH} for NGC 1453 is only 0.077 dex below this relation.

2.7.5 Gas Kinematics

In Pandya et al. [149], we observed the kinematics of warm ionized gas out to $\sim$ 8 kpc within NGC 1453 by tracing the 3727 Å [O II] emission line using the spectra obtained with the Mitchell IFS. This warm gas was found to rotate with a $\mathrm{PA_{gas}}\sim312.5°$, roughly perpendicular to the stellar rotation along $\mathrm{PA_{stars}}\sim35°$. This extreme misalignment suggests that the warm gas originated from external accretion rather than in-situ stellar mass loss. The gas was observed to have a rotation velocity of up to $\sim$200 km s^{-1} and a comparable velocity dispersion, giving an rms velocity of $\sim$300 km s^{-1}, similar to that of the stars studied in this paper.

2.8 Summary

We have presented a black hole mass determination of the MASSIVE survey galaxy NGC 1453 using high-spatial resolution stellar kinematic data from the GMOS IFS, wide-field kinematic data from the Mitchell IFS, and photometry from *HST* WFC3. Stellar kinematics are measured from the spectra to produce a truncated Gauss-Hermite parameterization of the LOSVDs. We determine the first eight moments of the LOSVDs from the high-S/N GMOS spectra and the first six moments from the Mitchell spectra (Figs. 3 and 4). The two sets of kinematic data together span about two orders of magnitude in radial extent, from $0\farcs3$ to $76''$ ($\sim$3 effective radii) with a total of 173 spatial bins of varied size.

In the production run described in this paper, we perform axisymmetric Schwarzschild orbit modelling for more than 8000 mass models to determine the mass parameters in NGC 1453. For each mass model, we use a library of up to 800,000 stellar orbits to sample the phase space, and then use a quadratic programming solver to find a superposition of orbits that minimizes the χ^2 associated with the observed kinematics and also fit the observed photometry to within 1%. This procedure is done for all mass models to produce likelihood distributions for the mass parameters (Figs. 2.7 and 2.8). The best-fit model for NGC 1453 has a black hole mass $M_{\rm BH} = (2.9 \pm 0.4) \times 10^9 M_\odot$, a stellar mass-to-light ratio (in F110W band) $M^*/L_{\rm F110W} = (2.09 \pm 0.06) M_\odot/L_\odot$, and an enclosed dark matter mass $M_{15} = (7.0 \pm 0.7) \times 10^{11} M_\odot$ at 15 kpc. The inclination is found to be nearly edge-on (Fig. 2.6).

We began the orbit modeling with the original triaxial Schwarzschild code of van den Bosch et al. [202] but determined that numerous changes must be made to properly model axisymmetric systems with that code. We found the gravitational potential not to be sufficiently axisymmetric when we adopted the typical setting of this code used in prior studies. As a result, the orbit start space includes box and long-axis orbits that are forbidden in truly axisymmetric potentials. Additionally, many of the integrated orbits near the black hole or far into the halo do not exhibit axisymmetry as their precession timescale is much longer than the code's default integration time. We introduced an additional axisymmetrizing step to enforce this symmetry. We also addressed several other issues and improved the computational efficiencies in the code. The changes leading to the Berkeley version of the code is discussed in Sec. 2.4; further details are described in Chapter 3 [155].

Another key finding of this paper is that care must be taken to properly handle the truncation of the Gauss-Hermite series used to describe the stellar LOSVDs. When the higher-order terms in the series are left unconstrained for NGC 1453, the resulting best-fit LOSVDs produced by the orbit models contain spurious features (Fig. 2.11), and the contributions from the unconstrained higher-order moments are not random but show spatial correlations (Fig. 2.10) When the Berkeley version of the orbit code is used, we find that the confidence level on the $M_{\rm BH}$ determination for NGC 1453 is significantly improved when at least 8 Gauss-Hermite moments are used as constraints: the 1-sigma confidence interval shrinks by a factor of ~ 2 relative to models with typical constraints on only V through h_4 (left panel of Fig. 12). By contrast, the χ^2 landscape is not as well behaved when the original

Table 2.3: Best-fit MGE parameters to the NGC 1453 *HST* WFC3 IR photometry. Each Gaussian component is parametrized by a central surface density $I_k = L_k/2\pi\sigma_k'^2 q_k'$ (calculated using an absolute solar AB magnitude $M_{\odot,F110W} = 4.54$), dispersion σ_k' (in arcseconds), and axis ratio q_k'.

I_k $[L_\odot/\mathrm{p}c^2]$	σ_k' $['']$	q_k'
6285.72	0.118	0.895
11089.5	0.323	0.928
15865.7	0.715	0.863
9393.34	1.392	0.794
5676.12	2.373	0.852
1824.78	3.846	0.791
1326.46	5.962	0.848
561.023	10.501	0.786
280.091	20.747	0.823
80.423	47.289	0.896

code is used with typical settings (right panel of Fig. 2.12). Tests on each individual galaxy would have to be performed to assess whether earlier M_{BH} determinations are similarly impacted.

A number of the findings and code changed discussed in this paper are also relevant when the orbit code is applied to a triaxial gravitational potential. In particular, the problem of insufficient integration time for the subset of orbits with long precession timescales occurs in both axisymmetric and triaxial models. We are currently investigating these issues with the aim to build equilibrium triaxial models for non-axisymmetric galaxies.

Appendix

2.A Multi-Gaussian Expansion parameters

We list in Table 2.3 the best-fit parameters of the 10 MGE components to our *HST* WFC3 IR photometry of NGC 1453 shown in Figure 2.5. The 10 Gaussians terms are assumed to have the same center and position angle of 28.5°.

We also performed an MGE fit to this photometry in Ene et al. [54]. The two MGEs differ in that the fit presented here was found by using the 'mge_fit_sectors()' function rather than the 'mge_fit_sectors_regularized()' function. We find that when the regularized fit is performed, the photometry is similarly well-fit. However, for the regularized fit we find a significant uptick in the model's surface brightness in the central $\sim 0''.1$, below the pixel scale of the photometry. To avoid this un-physical feature, we use the un-regularized fit here.

2.B Interpolation and Marginalization

We perform an interpolation with cubic radial basis functions (RBF) to promote our discrete sample of χ^2 evaluations at each model point to a continuous function over the parameter-space. We use a variation on the implementation described by [105]. The RBF interpolation is described by

$$\chi^2(\vec{x}) = \sum_{i=1}^{N} \lambda_i (||T(\vec{x} - \vec{x}_i)||)^3 + \vec{b} \cdot \vec{x} + a,$$

where $\vec{x}$ describes a point in the parameter-space and λ_i, $\vec{b}$, and a are uniquely defined from the criterion that the interpolation passes through all N sample points. T is initially the identity matrix.

A spatial rescaling is performed to improve the fit around the minimum of the landscape. This is done by evaluating the interpolating function at 10000 points drawn from a uniform distribution over the parameter-space. The covariance matrix of the 500 points with lowest predicted χ^2 is computed, then the eigenvalues α_i and eigenvectors $\vec{m}_i$ of that matrix are computed. Finally, T is constructed with $\vec{T}_i = \vec{m}_i / \sqrt{\alpha_i}$. Given this new T, λ_i, $\vec{b}$, and a are recomputed so that the interpolation once again passes through all N sample points.

To extract best-fit values and confidence intervals for each parameter we perform a straightforward marginalization. With marginalization we wish to reduce the interpolated $\chi^2(\vec{\theta}, \vec{\psi})$ to $\chi^2(\vec{\theta})$, where $\vec{\psi}$ are the parameters we wish to eliminate and $\vec{\theta}$ are those which remain. The likelihood is related to the χ^2 by $L = e^{-\chi^2/2}$ and likelihoods are marginalized in the same sense as probabilities. Therefore

$$L(\vec{\theta}) = \int d^N \psi L(\vec{\theta}, \vec{\psi})$$

and thus

$$\chi^2(\vec{\theta}) = -2 \ln \int d^N \psi e^{-\chi^2(\vec{\theta}, \vec{\psi})/2}$$

, where N is the number of parameters in ψ.

To obtain predictions for the best-fit and confidence interval for a parameter, we first construct the 1D likelihood function for that parameter:

$$L(\theta) = \int d^N \psi e^{-\chi^2(\theta, \vec{\psi})/2}.$$

For the best-fit, we determine the value where the cumulative-likelihood function is one-half:

$$\frac{\int_{-\infty}^{\theta} L(\theta') d\theta'}{\int_{-\infty}^{\infty} L(\theta') d\theta'} = \frac{1}{2}.$$

For the confidence intervals, we find the values where the cumulative-likelihood function reaches reaches the appropriate percentiles:

$$\frac{\int_{-\infty}^{\theta_\pm} L(\theta') d\theta'}{\int_{-\infty}^{\infty} L(\theta') d\theta'} = \frac{1 \pm \mathrm{erf}(k/\sqrt{2})}{2},$$

where θ_+ and θ_- yield the upper and lower bounds to the cumulative-likelihood and k sets the confidence level ($k = 1$ corresponds to the 68% level, $k = 2$ for 95%, and so on). We compute these integrals with the VEGAS Monte Carlo integrator implemented in the Python package 'vegas'.

Acknowledgements

The MASSIVE survey is supported in part by NSF AST-1411945, NSF AST-1411642, NSF AST-1815417, NSF AST-1817100, HST GO-14210, HST GO-15265 and HST AR-14573. M.E.Q. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), PGSD3-517040-2018. C.-P.M. acknowledges support from the Heising-Simons Foundation, the Miller Institute for Basic Research in Science, and the Aspen Center for Physics, which is supported by NSF grant PHY-1607611. J.L.W. is supported in part by NSF grant AST-1814799. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) at the San Diego Supercomputing Center through allocation AST180041, which is supported by NSF grant ACI-1548562. This work is based on observations obtained at the Gemini Observatory, processed using the Gemini IRAF package, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), Ministério da Ciência, Tecnologia e Inovação (Brazil), and Korea Astronomy and Space Science Institute (Republic of Korea).

Chapter 3

Axisymmetry in Triaxial Schwarzschild Orbit Superposition Models

The previous chapter describes our first results using a triaxial Schwarzschild orbit superposition code in the axisymmetric limit. This chapter describes in more detail the modifications of the code and prescriptions for settings which were required to generate consistent models within this limit. As in Chapter 2, we explore these modifications and prescriptions in the lens of NGC 1453, one of two fast-rotators within the MASSIVE sample.

This chapter was originally published as

Matthew E. Quenneville, Christopher M. Liepold, and Chung-Pei Ma. "Dynamical Modeling of Galaxies and Supermassive Black Holes: Axisymmetry in Triaxial Schwarzschild Orbit Superposition Models". In: *Astrophys. J. Suppl. Ser.* 254.2, 25 (June 2021), p. 25. DOI: 10.3847/1538-4365/abe6a0. arXiv: 2005.00542 [astro-ph.GA]

Minor alterations to the text have been made to ensure fluency throughout this dissertation.

3.1 Introduction

The Schwarzschild orbit superposition method [173] enables efficient construction of self-consistent and equilibrium mass models of galaxies. The basic procedure consists of two steps: integrating a representative set of orbits in a static triaxial gravitational potential, and finding weights for these orbits such that their superposition reproduces the assumed mass distribution.

The orbit superposition method has been extended to include kinematic information and used to determine mass distributions in real galaxies, starting with studies such as Pfenniger et al. [152], Richstone and Tremaine [160, 161], and Rix et al. [162]. From the quality of the fit to both kinematic and photometric data, this method can be used to assess the relative

likelihood of a range of mass models and to determine best-fit mass parameters such as $M_{\rm BH}$, stellar mass-to-light ratios, galaxy shapes, and dark matter halo parameters.

Due to the large number of orbits needed to sample the relevant phase space, the orbit superposition method is computationally expensive. To reduce the number of orbits and the dimensions of the model parameter space, a few orbit-based numerical codes have been developed for axisymmetric systems (e.g., [33, 79, 190, 197, 20]). Many dynamical measurements of $M_{\rm BH}$ from stellar kinematics have been obtained using these axisymmetric orbit codes.

Triaxiality allows for more general galaxy shapes and additional orbit types, but modeling orbits in triaxial potentials comes at the cost of increased complexity and computation time. van den Bosch et al. [202] presented a triaxial orbit-based code capable of comparing directly to observations, using an orbital sampling scheme based on that of Schwarzschild 1993 [174]. van de Ven et al. [212] performed recovery tests of this code for analytically tractable triaxial potentials (excluding central black holes). Only a handful dynamical determinations of $M_{\rm BH}$ have been obtained using triaxial models from this code [199, 219, 65]. Several additional $M_{\rm BH}$ were determined using this code in the (nearly) axisymmetric limit [176, 220, 217, 218, 1]. This code has also been used to construct axisymmetric and triaxial galaxy models to determine stellar dynamics and dark matter distributions for a wide range of galaxies (e.g., [230, 154, 103]). Vasiliev and Valluri [208] recently presented a new triaxial orbit-based code using a different method for phase space sampling and orbit initialization; the method was tested on mock data but had not been applied to real data.

An important test of the orbit superposition codes is the ability to produce consistent results between an axisymmetric code and a triaxial code in the axisymmetric limit. We note that the code by van den Bosch et al. [202] is written for triaxial potentials and "is not capable of making a perfectly axisymmetric model" [199]. Studies that attempt to run it near axisymmetry and then compare with results from axisymmetric codes have reached conflicting conclusions. For instance, van den Bosch and de Zeeuw [199] used their triaxial code to construct (nearly) axisymmetric models for M32 and NGC3379, and found the mass-to-light ratios and $M_{\rm BH}$ to be consistent with those from earlier studies using axisymmetric codes ([206, 104, 213] for M32; [79, 177] for NGC 3379). Ahn et al. [1], on the other hand, found a puzzling global χ^2 minimum at $M_{\rm BH} = 0$ while using this triaxial code to perform axisymmetric modeling of the ultracompact dwarf galaxy M59-UCD3. They found this minimum to be inconsistent with the best-fit non-zero $M_{\rm BH}$ from Jeans modeling and the axisymmetric orbit code of [20].

It is the purpose of Chapter 2 [119] and this chapter to investigate how to modify the van den Bosch et al. [202] code to enable it to handle properly both axisymmetric and triaxial systems. Since no galaxy in nature is likely to be exactly axisymmetric, it may appear that we are taking a step backwards in examining the axisymmetric limit of a triaxial code. While our next goal is indeed to adopt the more realistic triaxial potentials, we believe that one critical test of a triaxial code is its behavior in the simpler, axisymmetric limit. Such a study – the main goal of this paper – is a particularly important step in the quest for dynamical $M_{\rm BH}$ measurements in view of the facts that almost all existing $M_{\rm BH}$ measurements have

been obtained assuming exact axisymmetry, and that the aforementioned recent comparison of axisymmetric and triaxial codes have led to unresolved conflicting results.

In Chapter 2 [119], we described a set of recipes and code changes for achieving axisymmetry. We then performed proper axisymmetric orbit modeling using the revised code to obtain a new $M_{\rm BH}$ measurement for the massive elliptical galaxy NGC 1453, a fast rotator in the MASSIVE survey [126] well suited for axisymmetric orbit modeling. Similar to Ahn et al. [1], we had encountered difficulties in constraining $M_{\rm BH}$ in NGC 1453 when we used the original code with comparable settings. Through extensive testing, we came to two main conclusions: (1) higher Gauss-Hermite moments (beyond the typically used h_4) of the line-of-sight velocity distributions (LOSVDs) are needed to fully constrain the orbital weights, and (2) the orbit libraries need to be modified to satisfy axisymmetry. The use of higher moments is described in detail in Section 2.7.1 [119]. Here, we focus on the construction of axisymmetric orbit libraries in a triaxial orbit code.

In this paper, we provide a full discussion of the required steps to axisymmetrize the model and the various modifications that we have implemented to the triaxial code by van den Bosch et al. [202]. The code was never given a name; we will refer to it as the TriOS ("**Tri**axial **O**rbit **S**uperposition") code from this point on. In Section 3.2, we provide some background information about the implementation of the orbit superposition method in this code. We focus on four topics that are pertinent to subsequent discussions: the three major orbit types in a triaxial potential (Section 3.2.1), orbit sampling and initialization (Section 3.2.2), orbit integration (Section 3.2.3), and parameters used to quantify triaxial shapes (Section 3.2.4).

In Section 3.3, we give an in-depth discussion of the three main ingredients for axisymmetry listed in Section 2.4.1 [119]: axisymmetrization of short-axis tube orbits (Section 3.3.1), criteria for how to exclude long-axis tube orbits (Section 3.3.2), and exclusion of box orbits (Section 3.3.3).

We have made additional improvements and corrections to the code (Section 3.4). We identify a subset of slowly precessing quasi-planar orbits that are misclassified and are "mirrored" improperly in the orbit library (Section 3.4.1). We correct an issue with the zero point of the logarithmic potential for the dark matter halo that would otherwise render energy conservation checks ineffective in the code (Section 3.4.2). We are able to speed up the total runtime of a mass model by a factor of 2 to 3 by a simple modification to how the point spread function convolution is implemented in the code (Section 3.4.3). An improvement in setting the intrinsic mass grid used to constrain stellar density profiles is described in Section 3.4.4. Finally, we illustrate the effects of these changes in the case of NGC 1453 (Section 3.5).

Three appendices are included as well. Appendix 3.A derives a simple analytic criterion for the existence of long-axis tube orbits within a model. Appendix 3.B outlines a change in the thin orbit finding algorithm that must be made to the TriOS code in order to generate the correct orbit sampling. Finally, Appendix 3.C presents a mock recovery test demonstrating the ability of our revised TriOS code to recover the input mass parameters.

3.2 Orbit Modeling Background

A summary of the implementation of the Schwarzschild orbit superposition method in the TriOS code is given in Section 2.4 [119]. Here we focus on the topics relevant for subsequent discussions of axisymmetry (Section 3.3) and code modifications (Section 3.4).

In this paper, we use a Cartesian coordinate system in which the x, y, and z axes are directed along the intrinsic major, intermediate, and minor axes of the galaxy, respectively. The z-axis is therefore the symmetry axis of an oblate axisymmetric potential, and the x-axis is the symmetry axis of a prolate axisymmetric potential. We focus on oblate axisymmetric systems in this paper, although our discussions can be easily modified for the prolate axisymmetric case.

3.2.1 Orbit Types in a Triaxial Potential

In a static triaxial gravitational potential, time invariance is the only global continuous symmetry of the Hamiltonian, H. By Noether's theorem, this symmetry gives rise to conservation of energy as the only "classical" integral of motion. This conservation law restricts the allowed phase space for a given orbit from the full six phase space dimensions to a five dimensional subspace defined by the energy $H = E$. An integral that reduces the allowed phase space dimension in this way is referred to as an isolating integral.

Numerical studies have revealed that orbits in many potentials often conserve two additional "non-classical" isolating integrals of motion [173], which we refer to as I_2 and I_3. These additional integrals do not typically have simple analytical expressions nor correspond to global symmetries of H. Orbits that conserve three (or more) isolating integrals of motion are referred to as regular. These regular orbits often fall into one of three main orbit types: short axis tubes, long axis tubes, and boxes.

Both types of tubes have a fixed sense of rotation. For short-axis tubes, the component of angular momentum along the potential's minor axis, L_z, does not change sign. Similarly, for long-axis tubes, the component of angular momentum along the potential's major axis, L_x, does not change sign. For box orbits, all three components of angular momentum change sign, leaving no fixed sense of rotation. Box orbits also have the property of touching the equipotential surface, $\Phi(x, y, z) = E$, at some point during their trajectory. Intermediate axis tube orbits are typically unstable in triaxial models [91].

A triaxial system generally admits all three of these main orbit types. For oblate axisymmetric systems, the orbit structure is simpler because L_z is an integral of motion, and only short-axis tubes are present. Similarly, for prolate axisymmetric systems, L_x is an integral of motion and only long-axis tubes are present.

3.2.2 Orbit Sampling and Initialization

The set of initial conditions (referred to as a start space) should sample over all orbit types supported by the potential. Even though regular orbits in a triaxial potential conserve energy

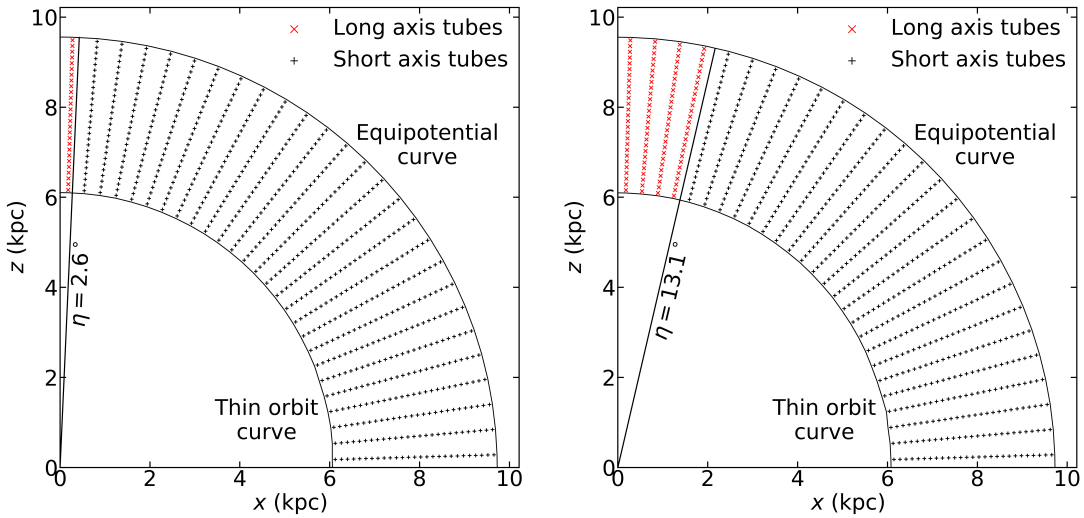

Figure 3.1: Two examples of the initial orbit locations in the x-z start space. Two nearly axisymmetric models for massive elliptical galaxy NGC 1453 are shown: (left) triaxiality parameter $T = 0.002$, (luminosity weighted) axis ratio $p = 0.9997$, and viewing angles $(\theta, \phi, \psi) = (89°, 45°, 90.001°)$; (right) $T = 0.05$, $p = 0.993$, and $(\theta, \phi, \psi) = (89°, 45°, 90.026°)$. Both models have the best-fit $M_{\rm BH}$, mass-to-light ratio, and dark matter halo from Chapter 2 [119] and assume the orbit sampling parameters $(N_\Theta, N_R, N_{\rm Dither}) = (9, 9, 3)$ (see Section 3.3.2). In each panel, one energy is shown, where the energy is chosen such that the potential is dominated by the stellar mass. Each symbol represents the initial location for a single trajectory, which are bundled with adjacent trajectories to form one dithered orbit. The long-axis tubes (red crosses) are all contained within the angle η of the z-axis for both values of T, where η and T are related by Equation (3.2). In general, more triaxial potentials contain a larger fraction of long-axis tubes in the x-z start space.

plus two additional integrals of motion, the non-classical integrals of motion, I_2 and I_3, may not be the same quantities for each orbit type [11, 14]. Thus, for a given energy, each orbit type can be sampled by a 2D start space, but the start spaces for the different orbit types cannot necessarily be combined into a single 2D start space.

Schwarzschild [174] argued that a 4D space can guarantee that all orbit types of a given energy are sampled, and further suggested that a pair of 2D start spaces is sufficient for sampling phase space in realistic galaxy potentials. The first of these start spaces, the x-z start space, is defined by sampling over a grid of points in the x-z plane, and setting $y = v_x = v_z = 0$ and v_y from $v_y^2 = 2[E - \Phi(x, 0, z)]$ for a given E. For simplicity, v_y is taken to be positive and a second copy is added to the orbit library with the velocity direction flipped. Two examples of this x-z start space are shown in Figure 3.1.

Typically, tube orbits will pass through the positive quadrant of the x-z plane perpendicularly at two points, separated by the thin orbit curve (see Figure 3.1). Orbits launched

along that curve will perpendicularly pass through the plane at a single point, so the curve can be found by iteratively launching orbits at different radii to identify those which pass through the x-z plane in a thin curve (see Appendix 3.B). Each orbit in the x-z start space passes once inside and once outside the thin-orbit radius, so the code avoids double counting by initializing orbits only between the thin-orbit curve and the equipotential where $E = \Phi(r)$, as shown by the crosses in the examples in Figure 3.1. All three main orbit types pass through this start space.

The second 2D start space proposed by Schwarzschild [174] is referred to as the stationary start space. In this start space, orbits are started from rest on the equipotential surface and are sampled over solid angle. Since tube orbits never come to rest, box orbits will be the only main orbit family in this start space. By combining the x-z start space that samples mainly tube orbits with the stationary start space that samples mainly box orbits, Schwarzschild [174] suggests that any remaining unsampled region of phase space is likely to be small.

The TriOS code is designed for static triaxial potentials that possess reflection symmetry along each of the three principal axes. Under this assumption, any orbital property only needs to be calculated in one octant; it can then be "mirrored" into the other seven octants by symmetry. Taking advantage of this symmetry, the code initializes orbits only in one octant ($x, y, z > 0$) and integrates only these orbits. Seven additional copies of each orbit are then created by simply mirroring along the three axes. The details are described in Section 4.5 of van den Bosch et al. [202] and the mirroring scheme is given in Table 2 there. A key feature to note in Table 2 is that the exact mirror procedure (i.e., how the signs of the velocity components are flipped in each octant) depends on whether the orbit is a short-axis tube, long-axis tube, or box. The orbits therefore must be classified first.

To classify an orbit, the code determines how the angular momentum components change sign over the course of its integrated trajectory and uses these rules: (1) short-axis tubes, if L_x and L_y flip signs while L_z does not, (2) long-axis tubes, if L_y and L_z flip signs while L_x does not, and (3) box orbits, if all three angular momentum components change signs. The velocities are mirrored in order to maintain the orbit's sense of rotation. If an orbit does not fall into any of these categories, its velocity is mirrored to have zero angular momentum.

3.2.3 Orbit Integration

The TriOS code uses the `DOP853` explicit Runga-Kutta integrator with order 8(5,3). The integrator performs adaptive time stepping to ensure that the relative error in the positions and velocities are below a set threshold, typically 10^{-5}. After each orbit is integrated, a relative energy tolerance is used to check energy conservation. If the change in energy exceeds this tolerance (typically set to 10%), it is re-integrated with a smaller integration error threshold.

The default integration time for each orbit is 200 dynamical times, where a dynamical time is set to the period of a closed elliptical orbit of the same energy. To enforce smoothness of the recovered distribution function, the orbital initial conditions can be "dithered" by combining $N^3_{\rm Dither}$ trajectories corresponding to nearby initial conditions. By merging

trajectories in this way, each orbit represents a small volume of the start space rather than a single point. This results in smoother orbital properties without a significant memory increase, since only the bundled orbital properties are stored.

After integration, the trajectory of each orbit is interpolated onto a set of points (typically 50,000) that are uniformly spaced in time. These interpolated points are then stored and used for computing orbital properties. Once the orbit libraries are constructed, weights are found for each orbit to reproduce the observed surface brightness (SB) distribution, the LOSVDs, and intrinsic 3D mass distribution.

3.2.4 Viewing Angles, Axis Ratios, and Triaxiality

Three viewing angles (θ, ϕ, ψ) can be used to relate the intrinsic and projected coordinate systems of a triaxial galaxy [13]. The two angles θ and ϕ describe the orientation of the observer's line of sight with respect to the intrinsic axes of the galaxy. The angle ψ specifies the remaining degree of freedom – rotation of the galaxy around the line of sight. The angle $\psi = 90^\circ$ corresponds to an oblate axisymmetric potential. In the oblate axisymmetric limit, θ is the inclination with $\theta = 90^\circ$ corresponding to edge-on, and ϕ describes rotations about the symmetry axis.

These three viewing angles are related to the intrinsic axis ratios p and q, where $p = b/a$ is the intrinsic intermediate-to-major axis ratio, $q = c/a$ is the intrinsic minor-to-major axis ratio, and a, b, c are the lengths of the three principal axes of a triaxial system (with $c \leq b \leq a$). A third parameter, $u = a'/a$, represents a compression factor due to projection, where a' is the major axis of the projected shape on the sky; $u = 1$ corresponds to the intrinsic major axis lying in the plane of the sky, while $u = p$ corresponds to the intrinsic intermediate axis lying in the plane of the sky. These quantities obey the inequality $0 \leq q \leq p \leq u \leq 1$. The relationship between the viewing angles, intrinsic axis ratios and observed axis ratio is given in Equations (7)-(10) of [202]. In addition, a triaxiality parameter is often used:

$$T = \frac{1 - p^2}{1 - q^2}\,. \tag{3.1}$$

This parameter ranges from 0 for oblate axisymmetry to 1 for prolate axisymmetry, with values in between indicating a triaxial shape.

The oblate axisymmetric limit can be achieved by setting either $p = 1$ or $\psi = 90^\circ$, but for numerical reasons, the code does not run when ψ is exactly 90°. As we discussed in Chapter 2 [119] and elaborate on below (Section 3.3.2), axisymmetry in the code can be achieved only with carefully chosen values of ψ or p.

3.3 Ingredients for Achieving Axisymmetry

In this section, we discuss a number of steps that need to be taken to generate orbit-superposition models in the oblate axisymmetric limit using the TriOS code. It is straightforward to modify these steps for the prolate axisymmetric limit. In Appendix 3.C, we test

the modified TriOS code on a mock dataset showing that we can accurately recover input parameters.

3.3.1 Axisymmetrize Short-Axis Tube Orbits

As we described in Section 3.2.2, a triaxial potential exhibits reflection symmetry along each principal axis, allowing the TriOS code to initialize orbits in only one octant of the x-z start space. These orbits are then mirrored via eight-fold reflections about the principal axes into each of the other seven octants. This setup is not meant for axisymmetric systems, in which the orbit library should respect azimuthal symmetry about the symmetry axis.

To enable modeling axisymmetric systems, we have implemented an axisymmetrized version of the orbit library by creating 80 copies of each short-axis tube orbit in the original loop library: 40 copies rotated evenly through an angle 2π about the short axis with velocities rotated to preserve L_z, and another 40 copies generated by flipping the sign of z and v_z in each of the 40 rotations. We choose 40 rotations, as this gives several copies per quadrant, with a comparable density to the start space grid sampling. Once we perform this operation, it is unnecessary to perform the eight-fold reflections in the original code. A similar rotation scheme was tested on mock data with no central SMBH in Hagen et al. [88].

The net result of our axisymmetrization process is to create a library of short-axis tube orbits in the TriOS code that samples the azimuthal angle uniformly with effectively equal orbital weights. In order for this procedure to be justified, the library should consist solely of short axis-tubes. In the next section, we show how to ensure that no long-axis tubes occur in this library.

3.3.2 Exclude Long-Axis Tube Orbits

In an oblate axisymmetric potential, the long-axis tube orbits become unstable since there is no longer a single preferred long axis. These orbits therefore should not be present in the orbit library. [1]

As we discussed in Section 3.2.4, the potential is oblate axisymmetric when ψ is set to 90° exactly, and long-axis tubes should be absent in this limit. For numerical reasons, however, the code does not run when ψ is set to 90° within machine accuracy. Prior work using this code for black hole mass measurements in the axisymmetric limit chose either $|\psi - 90°|$ between 0.001° and 0.01° [217, 1], or an axis ratio of $p = 0.99$ [176, 220, 218]. As we first pointed out in Chapter 2 [119], some of these values may not have been close enough to the desired axisymmetric values to exclude long-axis tubes. Here we provide a detailed explanation.

We use two examples of the x-z start space in Figure 3.1 to illustrate how long-axis tube orbits are initialized in the code. As shown in Appendix 3.A, long-axis tube orbits in many

[1] Similarly, in the case of a prolate axisymmetric potential, the short-axis tube orbits become unstable and should be absent.

realistic triaxial potentials are confined to pass through the x-z start space within an angle η from the z-axis. The angle η depends on the shape of the potential, and we find the relation between η and the triaxiality parameter T (Equation 3.1) to be well approximated by

$$\eta = \tan^{-1}\sqrt{\frac{T}{1-T}}\,. \tag{3.2}$$

This is demonstrated in Figure 3.1 where the black line at angle η separates the short-axis tube orbits (black crosses) from the long-axis tube orbits (red crosses). As the potential becomes more oblate axisymmetric ($T = 0.05$ in the right panel vs. $T = 0.002$ in the left panel), η becomes smaller and the area in the x-z start space occupied by long-axis tubes shrinks. To effectively achieve oblate axisymmetry, η needs to be small enough so that no orbits are sampled within an angle of η of the positive z-axis. Two additional mass models with higher triaxiality, ($T = 0.25$ and 0.75) are shown in Appendix 3.A and Figure 3.1. Equation (3.2) again provides an excellent approximation for the angle demarcating the long-axis and short-axis tube orbits in the x-z start space.

Whether orbits are sampled within the angle η on the x-z plane depends on the input parameters. For a given energy, the code starts the orbits on a grid of N_R radii between the inner and outer thin orbit radii and N_Θ angles between 0° and 90° in the positive quadrant on the x-z plane [202]. The code further allows for dithering, where $N^3_{\rm Dither}$ nearby initial conditions, adjacent in (E, R, Θ), are bundled together to improve the phase space sampling. Orbits are therefore sampled at a total of $N_\Theta \times N_{\rm Dither}$ angles, where the first angle from the z axis is chosen to start at half of the grid spacing (i.e., at an angle of $(\pi/2)/(2N_\Theta N_{\rm Dither})$ from the z-axis). The criterion to satisfy oblate axisymmetry is therefore

$$\frac{1}{2N_\Theta N_{\rm Dither}}\frac{\pi}{2} \gtrsim \eta. \tag{3.3}$$

The two examples of NGC 1453 shown in Figure 3.1 have $N_\Theta = N_R = 9$, $N_{\rm Dither} = 3$, and 27×27 orbits initialized in the x-z start space. The orbits closest to the z-axis are therefore at an angle of $\approx 1.67°$ away. These orbits lie within the demarcation angle η of Equation (3.2) for either model in Figure 3.1: $\eta = 2.56°$ for $T = 0.002$ (left) and $\eta = 12.9°$ for $T = 0.05$ (right). Both models therefore violate Equation (3.3) and contain long-axis tubes. This provides the physical explanation for our statement in Chapter 2 [119] that even $|\psi - 90°|$ as small as 0.001 (left panel) is not sufficiently close to 90° to achieve axisymmetry in our models.

To extend the discussion beyond the two specific mass models shown in Figure 3.1, we illustrate in Figure 3.2 the relation between T and ψ for nearly axisymmetric models of NGC 1453 (top panel), and the corresponding fraction of long-axis tubes that are initialized in the x-z plane (bottom panel). The inclination angle θ is assumed to be 89° here, and the shaded band indicates the additional dependence of T on ϕ. Figure 3.2 shows that $T \lesssim 5 \times 10^{-4}$ is needed to exclude long-axis tube orbits in this case. The corresponding requirement on ψ is $|\psi - 90°| \lesssim 8.7 \times 10^{-6}$ for $\phi \sim 1°, 89°$ and $|\psi - 90°| \lesssim 2.5 \times 10^{-4}$

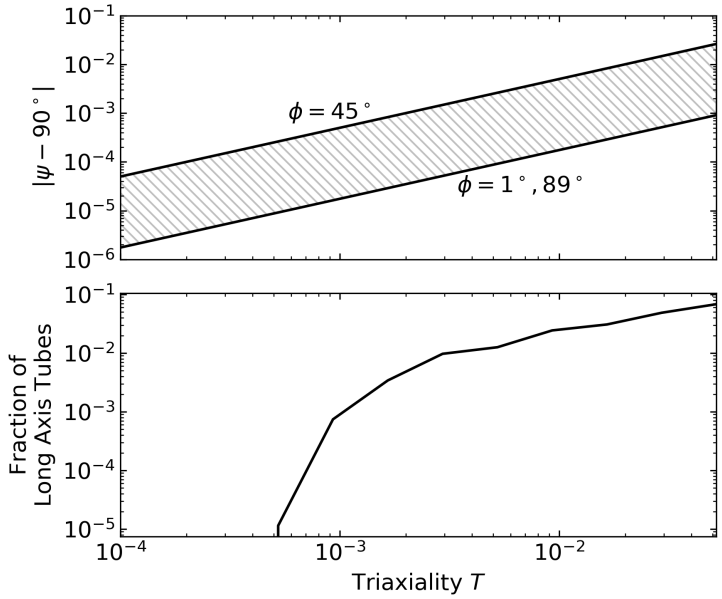

Figure 3.2: (Top panel) Relationship between the viewing angle ψ and the triaxiality of the deprojected stellar density. Exact oblate axisymmetry has $T = 0$ and $\psi = 90.0°$. The other viewing angle θ is taken to be 89°, and ϕ is varied from 1° to 89°. (Bottom panel) Fraction of long-axis tube orbits in the x-z start space as a function of the triaxiality of the stellar density near the oblate axisymmetric limit. The same mass model and orbit sampling parameters for NGC 1453 shown in Figure 3.1 is assumed here. In this example, long-axis tube orbits begin to appear when T is as small as $\sim 5 \times 10^{-4}$, or $|\psi - 90°|$ as small as $\sim 9 \times 10^{-6}$, and the fraction of these orbits increases monotonically as the potential becomes more triaxial, reaching $\sim 6\%$ at $T = 0.05$.

for $\phi \sim 45°$. We advocated $|\psi - 90°| = 10^{-9}$ in Chapter 2 [119], which safely excluded all long-axis tube orbits.

Earlier work using the code in the near axisymmetric limit does not typically satisfy the criterion in Equation (3.3). For M59-UCD3, Ahn et al. [1] used $(\theta, \phi, \psi) = (85°, -49.99°, 89.99°)$, which we find to correspond to $T = 0.004$ and $\eta = 3.64°$. The orbit sampling parameters were not explicitly given for the runs using the triaxial code. Assuming the same parameters used in their runs with the axisymmetric orbit code ($N_\Theta = 8, N_{\rm Dither} = 6$), we find that the innermost ray would be at an angle of 0.94° from the z-axis, which is well inside $\eta = 3.64°$, and therefore violates the criterion in Equation (3.3).

For M60-UCD1 [176], NGC 1271 [220], and Mrk 1216 [218], each paper quoted an axis ratio of $p = 0.99$. The minimum possible triaxiality with this value of p is $T = 1 - p^2 = 0.0199$ (in the unrealistic limit of a razor-thin disk with $q = 0$), leading to a minimum η of 8.1°. For NGC 1271 and Mrk 1216, $N_{\rm Dither} = 5$ was used, while N_Θ was set to 8 and 9 respectively.

Thus, orbits were sampled starting at 1.125° and 1° away from the z-axis, indicating that neither satisfies the criterion in Equation (3.3). For M60-UCD1, not enough information is given about the orbit sampling to determine whether the criterion is satisfied. However, for typical orbital sampling parameters quoted above, the criterion in Equation (3.3) would not be satisfied.

The modeling of the NGC 1277 black hole used $N_\Theta = 9$ and $N_{\rm Dither} = 5$ [217]; the innermost ray of initial orbits therefore lies at 1° from the positive z-axis. The complete shape information was not given in the paper, but private communication indicated that $(\theta, \phi, \psi) = (75.3°, 71.6°, 90.001°)$ was used. We find this set of viewing angles to correspond to $T = 0.0002$ and $\eta = 0.85°$, narrowly satisfying the criterion in Equation (3.3).

We note that the presence of the long-axis tube orbits in the orbit library does not necessarily imply that they receive significant weights after fitting to observational constraints for a given galaxy. Direct tests would need to be performed for each galaxy to assess the impact of these orbits on previous work.

3.3.3 Exclude Box Orbits

As we discussed in Section 3.2.1, all orbits in the (oblate) axisymmetric limit conserve L_z. Box orbits in this limit have $L_z = 0$ and therefore have similar properties as the tube orbits with small L_z. In this case, as long as angular momentum is sufficiently sampled by the tube orbits, there is no need to include box orbits explicitly.

The TriOS code devotes an entire library to box orbits and initializes them in the stationary start space (Section 3.2.2). One can modify the code to exclude this library when needed. We use a simpler approach without changing the code itself: we skip running the orbit integration routine `orblib_f.f90` for the stationary start space, and replace the box library with a copy of the x-z library in the input file for the weight-finding routine `triaxnnls.f90`. These modifications typically reduce the total computation time of the original code by more than half.

While box orbits are unnecessary in the axisymmetric limit, they also should be harmless and not affect the results if included. As a test, we have run our revised code including the box library for comparison. Since the box orbits launched at different azimuthal angles are allowed to have different weights in the triaxial code, we have to impose an additional constraint of equal weights to enforce axisymmetry in the box library. Once these weights are forced to be equal, we indeed find similar results as the case when the box library is excluded altogether. The case where the box library weights are free to differ between azimuthal angles is discussed in Section 3.5. To reduce computational cost, we recommend excluding the stationary start space for axisymmetric models.

For a triaxial potential, we note that box orbits can also occur in the x-z start space (e.g., Figure 1 of Schwarzschild 1993 [174]). However, the region in the x-z start space that would generate box orbits shrinks as the potential becomes increasingly axisymmetric. When exact axisymmetry is reached, only the orbits that begin exactly on the equipotential surface in the x-z start space have $L_z = 0$ (since they have zero initial velocities) and are box orbits.

The TriOS code does not sample orbits lying exactly on the equipotential curve in the x-z start space, so the number of box orbits will shrink to 0 as axisymmetry is approached. In other orbit-based codes that assume axisymmetry from the start, the $L_z = 0$ orbits also are not usually sampled, as they are presumed to be represented by the tube orbits with small but non-zero L_z (e.g., [33, 190]).

3.4 Additional Code Fixes and Improvements

We have made several modifications in the TriOS code in addition to those described in Section 3.3. These modifications include corrections, improvements and speedups that are general to the code regardless of the issue of axisymmetry. We describe these changes in this section.

3.4.1 Correct Orbit Misclassifications

As we described in Section 3.2.2, the TriOS code assumes the triaxial potential to possess reflection symmetry along each of the three principal axes and integrates only orbits that are initialized in one octant of space to save computation time. It then uses an eight-fold reflection scheme to generate seven more copies of each orbit. How the orbits are "mirrored" depends on whether the orbit is classified as a short-axis tube, long-axis tube, or box orbit.

We have discovered that the mirroring scheme in the original code misclassifies a subset of orbits for which the angular momenta vary on timescales slower than the integration time. We find this to happen in at least two situations. First, in nearly oblate axisymmetric models, many box orbits in the stationary start space tend to be misclassified as short axis tubes due to the near conservation of L_z. Because L_z varies slowly, it may not change sign throughout the integration time. However, these orbits have very low angular momentum, so it is unlikely that mirroring these orbits to preserve L_z would cause significant issues in the models themselves.

The second situation occurs in regions of space where the potential is nearly spherical, e.g., deep within the SOI of an SMBH, or in the outer part of a galaxy where a (spherical) dark matter dominates the potential. Some orbits in these regions follow quasi-planar rosettas or Keplerian-like ellipses with nearly constant angular momentum vectors, consistent with prior studies of orbits near a central point mass [182, 183, 198]. For the subset of orbits with precession time longer than the integration time, no component of their angular momentum changes sign over the entire integrated trajectories. These orbits therefore do not fall into any of the categories listed above and are mirrored incorrectly to have no net angular momentum.

These quasi-planar orbits will not be significant in most Schwarzschild models, as they are only present at extreme radii. We checked this in our models of NGC 1453, with the properly axisymmetrized code as described in Section 3.3 using the lowest four Gauss-Hermite kinematic moments as constraints. In this model, we find that $\sim 10\%$ of the total weight after orbital weight minimization is assigned to orbits that would have been quasi-planar

in the original version of the code ($\sim$ 10% of the mass within the Mitchell apertures and $\sim$ 2% of the mass within the GMOS apertures). These relatively low percentages suggest a minimal effect on the model for NGC 1453.

We expect the issues with orbit integration time and misclassification to be more severe for galaxies with data that resolve well within the black hole's sphere of influence (SOI), or well beyond the stellar half-light radius, e.g., M87 and the Milky Way black hole. The effect is also likely to be more significant if the galaxy has a net rotation at these radii.

We find a further issue with orbit classification in the orbital composition information outputted in the file `intrinsic_moments.out`. This file reports the mass fraction of box orbits for each bin in the intrinsic spatial grid described in Section 3.4.4. In this case, however, all orbits that are neither long-axis tubes or short axis-tubes are grouped together as box orbits. Since this includes the quasi-planar orbits discussed above, the reported fraction of true box orbits may be overestimated.

In our revised code for axisymmetric systems, these orbit misclassification issues are not present because we manually assign all orbits as short-axis tubes and exclude all other orbit types. We will discuss further these quasi-planar orbits in triaxial systems in Section 3.6.

3.4.2 Fix Zero-point Issues with the Logarithmic Halo

A logarithmic potential is often used to approximate the dark matter halo in prior orbit modeling work. The spherical version of a logarithmic halo is given by

$$\Phi(r) = \frac{1}{2} V_c^2 \ln\left(R_c^2 + r^2\right) + \Phi_0 \,, \tag{3.4}$$

where R_c is the core radius, and V_c is the circular velocity at large r:

$$V_c(r) = \frac{V_c\, r}{\sqrt{R_c^2 + r^2}} \,. \tag{3.5}$$

The zero point Φ_0 can in principle be chosen arbitrarily; the original code set $\Phi_0 = 0$. In practice, we find the choice of $\Phi_0 = 0$ and the use of physical units such as kilometers for all distances to create numerical problems. The cause is simple: unlike other commonly used dark matter potentials (e.g., Hernquist [92] and NFW [146]) that are negative at all locations and approach 0 at large r, the logarithmic potential with $\Phi_0 = 0$ is positive everywhere and grows unbounded at large r. Thus, for the other potentials, $|\Phi(r)|$ can be interpreted as the local maximum kinetic energy for a bound orbit, but the orbital binding energy is infinite in the logarithmic potential. Furthermore, with the choice of $\Phi_0 = 0$, $|\Phi(r)|$ is much larger than the kinetic energy for all orbits in a logarithmic halo. This is because the central potential energy value, $\Phi(0) = V_c^2 \ln(R_c)$, is much larger than the maximum possible kinetic energy sampled by the orbits, which is $\Phi(r_{\max}) - \Phi(0)$, where $r_{\max}$ is the largest equipotential radius of any orbit in a model.

To illustrate this point, we plot the ratio of $|\Phi(r)|$ and $|\Phi(r_{\max}) - \Phi(0)|$ for the best-fit logarithmic dark matter halo of NGC 1453 from Chapter 2 [119] in Figure 3.3 (dotted curve).

Additional contributions to the potential from the stars and black hole reduce the value of the potential energy and help lower this ratio (dot-dashed and dashed solid curves), but the ratio is well above unity for all relevant radii in all cases.

An unintended consequence of this large central offset is that even a $\sim 100\%$ change in the kinetic energy would contribute to only a tiny fraction of the total energy and would be difficult to detect. The energy conservation checks in the code are therefore effectively not performed for most orbits. While these numbers are worrying, we did not find the choice of $\Phi_0 = 0$ to affect significantly the best-fit mass parameters of NGC 1453 in Chapter 2 [119]. The reason for this particular case is that the orbit integrator happened to be accurate enough to satisfy the energy conservation tolerance (set to the default 10%) even when this conservation criterion was unchecked. There is, however, no guarantee that this would be true for other galaxies or for parameters outside the ranges that we had explored.

To ensure energy conservation is checked in the code for the logarithmic potential, we choose a different zero point

$$\Phi_0 = -\Phi(r = 2r_{\max}) \,, \tag{3.6}$$

so that $\Phi(r)$ is negative for the entire allowed radial range of the orbits and approaches 0 outside the largest equipotential radius $r_{\max}$. The resulting ratio of $|\Phi(r)|$ to $|\Phi(r_{\max}) - \Phi(0)|$ for the best-fit model of NGC 1453 is shown by the solid line in Figure 3.3.

Our choice of Φ_0 in Equation (3.6) also removes another issue that we have encountered with the original code: the orbit start space was sometimes not calculated correctly for mass models in which the black hole is either absent or has small mass compared to the stellar component and the logarithmic halo. As discussed in Section 3.2.2 and shown in Figure 3.1, the x-z start space of Schwarzschild 1993 [174] requires finding equipotential curves in the x-z plane. The code locates it by finding the equipotential radius for each of a series of angles in the plane. For each angle, the equipotential radius is found via bisection with a relative tolerance that is typically taken between 10^{-7} and 10^{-5}. For $\Phi_0 = 0$, this tolerance again is not enforcing the intended accuracy level due to the large central value of Φ. For NGC 1453, this issue exists only for a few central equipotential radii and thus did not have a significant impact on our results.

3.4.3 Speed Up Point Spread Function Implementation

The point spread function (PSF) of the relevant observations needs to be incorporated into a mass model before the model is fitted to data to determine the orbital weights. The TriOS code approximates the effect of the PSF by perturbing each trajectory at every stored time step with a pair of δx and δy randomly drawn from the PSF, which is assumed to be a single or multiple Gaussian functions. This scheme involves a large number of operations since an orbit is typically stored at 50,000 points along the trajectory (see Section 3.2.3), and up to $\sim 10^6$ orbits can be used to represent a single mass model.

The code generates each orbit perturbation by drawing two independent numbers, k_x and k_y, from a uniform distribution over the interval $(-1, 1)$ repeatedly until a pair with

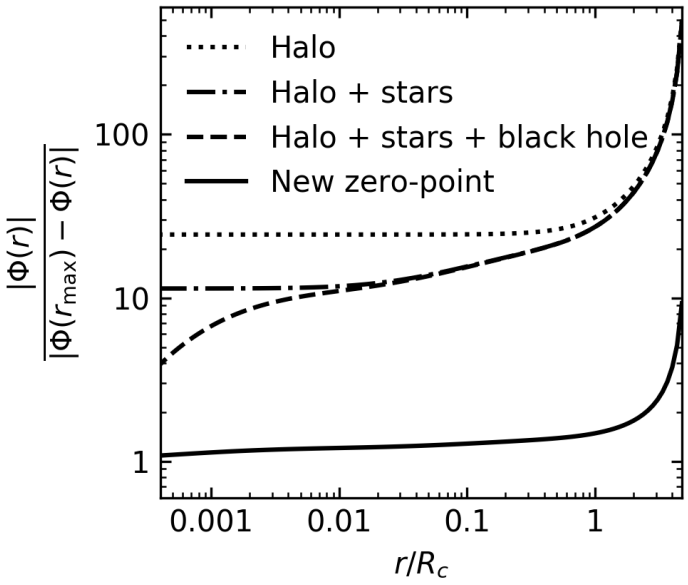

Figure 3.3: Illustration of the issue with setting the zero-point of the logarithmic potential to $\Phi_0 = 0$ in Equation (3.4), as is assumed in the original code. As an example, we use the best-fit mass model for NGC 1453 from Chapter 2 [119] with a logarithmic dark matter halo of $R_c = 15$ kpc and $V_c = 633$ km s^{-1}. The ratio of the potential energy to the maximum kinetic energy is plotted for this halo (dotted), halo plus stars (dot-dashed), and all three mass components (dashed). When this ratio is much larger than 1, as is shown for a large range of radius, even large errors in the kinetic energy would have little effect on the total energy. Energy conservation is therefore effectively not enforced in the original code for a logarithmic potential. The solid line shows the same ratio with all three mass components included, but with the halo zero point set according to Equation 3.6.

$k \equiv |\vec{k}| < 1$ is found. The perturbations $\delta x = \frac{k_x}{k}\sqrt{-2\ln(k^2)}$ and $\delta y = \frac{k_y}{k}\sqrt{-2\ln(k^2)}$ are then normally distributed. This large number of operations is not easily vectorized and is computed sequentially.

We are able to speed up this process significantly using instead the Box-Muller transform, which is easily vectorized. In this scheme, we draw a pair of independent numbers A and B from the uniform distribution over $(0, 1)$ and then construct the normal distribution with $\delta x = \sqrt{-2\ln A}\cos(2\pi B)$ and $\delta y = \sqrt{-2\ln A}\sin(2\pi B)$. We have tested that the resulting distributions of displacements are consistent with analytical PSFs to within the counting error from the finite number of timesteps, and the consistency increases as expected when the number of timesteps increases.

To benchmark the amount of speedup gained by our scheme, we note that PSF convo-

lution is one of several operations performed in the orbit library construction subroutine `orblib_f.f90` in the code. This subroutine first integrates the orbits and generates the necessary reflected or rotated copies of the orbits about the symmetry axes (see Section 3.3.1). It then computes each orbit's contribution to the 3D mass grid and projects each orbit onto the sky plane. The projected trajectories are then perturbed according to the PSF as described above. Finally, the subroutine determines each orbit's contribution to each observed kinematic aperture on the sky and stores the associated LOSVDs. The tasks performed in this subroutine consume the bulk ($> 90\%$) of the total runtime of the code (for one mass model); much of the remaining time is spent on performing minimizations to find optimal orbital weights.

To our surprise, our timing analysis of the various tasks executed in this subroutine (using $N_{\rm Dither} = 5$ and NGC 1453 as a test case) shows that the PSF portion of the code (before implementing orbit axisymmetrization in Section 3.3.1) takes up $\sim 55\%$ of the run time, while the orbit integration itself only contributes $\sim 20\%$, and sky projections contributes the remaining $\sim 25\%$. When we switch to the vectorized Box-Muller transform, the computation time for the PSF step becomes negligible. We are therefore able to reduce the total runtime of the code by a factor of ~ 2 in this test.

The speedup is even more dramatic in our axisymmetrized version when the orbits are copied azimuthally (Section 3.3.1). In this case, 80 (instead of 8) copies of each orbit are projected onto the sky and perturbed by the PSF. We find $\sim 70\%$ of runtime is spent on the PSF portion with the original scheme, while our new scheme reduces the runtime by a factor of ~ 3.

3.4.4 Improve Intrinsic 3D Mass Grid

The TriOS code uses an intrinsic 3D spatial grid to constrain the stellar component in a model to reproduce the 3D stellar density profile deprojected from the photometry of a galaxy. The code calculates the mass contributed by each orbit as it passes through a spatial bin and records this information during the stage of orbit library construction. At the subsequent stage of orbital weight optimization, the superposition of the orbits is required to match the input mass profile within a pre-specified precision (typically 1%) in each bin.

In each octant of this 3D spatial grid, the code uses azimuthal and polar bins for the two angles, each linearly spaced between 0 and 90°. The radial bins are logarithmically spaced between $r_{\rm min}$ and $r_{\rm max}/2$, where $r_{\rm min}$ and $r_{\rm max}$ are the innermost and outermost equipotential radii used to determine the orbital energies sampled in the model. The innermost bin is then extended down to $r = 0$ and the outermost is extended out to $100 r_{\rm max}$.

For the outer boundary of the innermost mass bin, we find it preferable not to base the value on $r_{\rm min}$, which is used for a different purpose of specifying the innermost equipotential radius for sampling orbital energy. Instead, we modify the code to make it an independent parameter, which we set to be of similar scale as the PSF of the photometric data since these are the data used to constrain the deprojected 3D mass density. To ensure that sufficient orbits are used to represent the innermost mass bins, we recommend that $r_{\rm min}$ be set to be

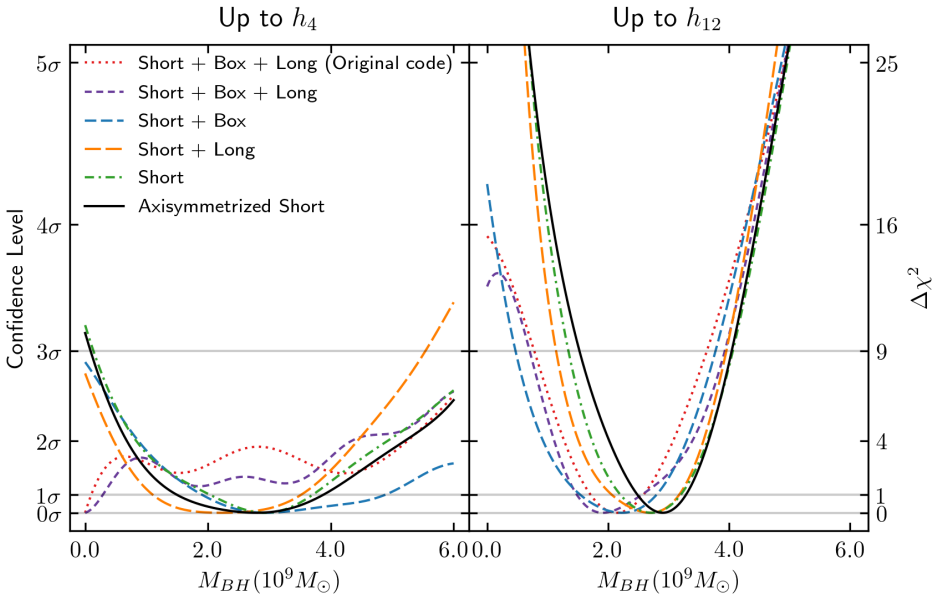

Figure 3.4: Illustration of the changing $M_{\rm BH}$ constraints in NGC 1453 as the orbit model goes through the step-by-step axisymmetrization procedure described in Sections 3.3 and 3.4. The starting case (red dotted) uses the original code with typical (near) axisymmetric parameters assumed in the literature ($\psi = 90.001°$; see Section 3.5.1 for details). The end case (black solid) uses our final axisymmetrized code including all changes from Sections 3.3 and 3.4. The four intermediate curves have all the code fixes described in Section 3.4, but have different combinations of orbit types according to Sections 3.3.2 and 3.3.3. The left panel is for models with orbital weights chosen by fitting to the first four Gauss-Hermite moments of the LOSVDs determined from kinematic data, as is typical in the literature. The right panel uses 12 moments as constraints and shows tighter constraints on $M_{\rm BH}$, as is reported in Chapter 2 [119]. The 1D χ^2 in $M_{\rm BH}$ is obtained by marginalizing over the stellar mass-to-light ratio using a smoothed 2D χ^2 landscape generated by Gaussian Process regression with a squared-exponential covariance function [150]. The dark matter halo is fixed to the best-fit logarithmic halo in Chapter 2 [119].

smaller than the outer boundary of the innermost mass bin. In the case of NGC 1453, we set the outer boundary of the innermost mass bin to be $0.03''$ and set $r_{\rm min}$ to $0.01''$.

For similar reasons, we allow the outermost mass bin's edges to also be set independently from the outermost equipotential radius, $r_{\rm max}$. The remaining bin boundaries are then logarithmically sampled between the outer boundary of the innermost bin and the inner boundary of the outermost bin.

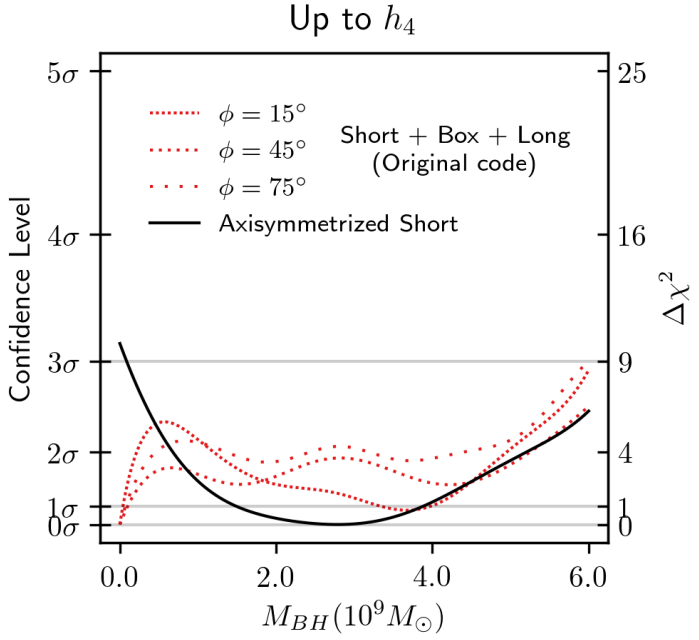

Figure 3.5: Same as the left panel of Figure 3.4 but showing the azimuthal dependence of the original code when ψ is chosen to be 90.001° and all three main orbit types are included (red curves). Our final axisymmetrized code does not depend on ϕ and obeys azimuthal symmetry.

3.5 A Case Study: NGC 1453

We use the massive elliptical galaxy NGC 1453 reported in Chapter 2 [119] to illustrate the effects of the modifications described thus far. In Chapter 2 [119], we demonstrated that using more than 4 Gauss-Hermite moments was essential for obtaining robust constraints on the model LOSVDs. Below we examine the effects in both the 4-moment and 12-moment cases, with the latter being our chosen configuration. We stress that the 4-moment case is included here only for comparison purposes since this is the typical configuration used in the literature. We have found the 4-moment case to lead to unconstrained higher moments and spurious features in the LOSVDs for NGC 1453 (Figs. 2.10 and 2.11; Figs. 10 and 11 of [119]); the resulting χ^2 in this case should therefore not be trusted.

3.5.1 Fitting up to h_4

We begin with the case labeled "up to h_4" and "original Leiden version" in Figure 2.12 (Figure 12 of [119]). This case is run with the original code, $N_{\rm Dither} = 3$, $N_\Theta = 9$, and the

viewing angles $(\theta, \phi, \psi) = (89°, 45°, 90.001°)$, corresponding to a nearly oblate axisymmetric potential with a triaxiality parameter of $T = 0.002$. As we discussed in Section 3.3.2, these parameters are chosen to resemble those used in earlier studies, and the models include both the x-z and stationary start spaces and contain all three major types of orbits: short-axis tubes, long-axis tubes and box orbits. The left panel of Figure 3.1 illustrates the starting locations of both short- and long-axis tube orbits in the x-z start space for one energy in this configuration.

The 1D χ^2 as a function of $M_{\rm BH}$ (marginalized over the mass-to-light ratio) is shown in the left panel of Figure 3.4 (red dotted curve). As first shown in Chapter 2 [119], the favored model in this case contains no black hole. The χ^2 minimum at $M_{\rm BH} = 0$ here resembles the finding for the dwarf galaxy M59-UCD3 by Ahn et al. [1], which also used four Gauss-Hermite moments as constraints and a set of viewing angles with a similar deviation from axisymmetry.

Applying the code changes described in Section 3.4 results in minor changes in the χ^2 contour for NGC 1453 (purple short dashed curve in Figure 3.4), but the $M_{\rm BH} = 0$ minimum remains. In the next step, we exclude the box orbits and long-axis tube orbits as described in Section 3.3. The box orbits are eliminated by the simple procedure in Section 3.3.3. To remove the long-axis tube orbits, we choose a galaxy shape that is sufficiently axisymmetric, as discussed in Section 3.3.2. For NGC 1453, we simply change ψ from 90.001° to $(90+10^{-9})°$, as was done in Chapter 2 [119]. This new value is far enough from 90.0° to avoid numerical issues in the code but is close enough to 90.0° so that all of our orbits lie outside the long-axis tube region in the x-z start space shown in Figure 1.

The effect of excluding these orbits on the best-fit parameter values for NGC 1453 is significant. The preferred $M_{\rm BH}$ is changed from 0 to $2.8 \times 10^9 M_\odot$ (green dot-dashed curve in Figure 3.4a). Before their removal, box orbits generally accounted for less than 10-35% of total mass, while long axis tube orbits accounted for less than 2%. Removing box orbits (orange long dashed curve in Figure 3.4a) has a significant effect on $M_{\rm BH}$ because box orbits starting at different azimuthal angles are not forced to have equal weights in the original code (Section 3.3.3). Removing the long-axis tubes (blue dashed curve in Figure 3.4a) has a significant impact likely due to their ability to fit minor-axis rotation in triaxial potentials.

In addition to excluding the box and long-axis tubes, we describe in Section 3.3.1 the need to enforce axisymmetry in the code by generating many azimuthally rotated copies of each short-axis tube in the x-z start space. For NGC 1453, we find that the main effect on the χ^2 contour of this axisymmetrization procedure is to widen the minimum (black solid curve in Figure 3.4a), as a broader range of orbital weights are able to fit the mass constraint for each mass model.

The results presented thus far with the original version of the code all assumed a viewing angle of $\phi = 45°$. When the model galaxy is perfectly axisymmetric, this angle is irrelevant and the resulting χ^2 landscape should be independent of ϕ. As a test, we have repeated the run with the original code (using four Gauss-Hermite moments) with two other values of ϕ (15° and 75°) while keeping all other parameters fixed. The resulting χ^2 as a function of $M_{\rm BH}$ for the three values of ϕ are shown in Figure 3.5. The dependence on ϕ indicates that

the mass models are indeed not consistent with axisymmetry. All three values of ϕ exhibit the same preference for $M_{\rm BH} = 0$.

3.5.2 Fitting up to h_{12}

We now examine models in which the orbital weights are constrained to fit the first 12 Gauss-Hermite moments of the observed LOSVDs for NGC 1453. The first 8 moments are measured from spectroscopic observations, while the 9th-12th moments are set to 0 with an error bar based on the lower moments, as described in detail in Section 2.4.3 [119]. Even without any of the modifications described in this paper, Section 2.7.1 [119] showed that the original code performed better when 12, rather than 4, moments were used as constraints. The right panel of Figure 2.12 (Fig. 12 in [119]) illustrated how the best-fit black hole mass moved from $M_{\rm BH} = 0$ for 4 moments (green curve) to $M_{\rm BH} = 2.2 \times 10^9 M_\odot$ for 12 moments (black curve). The result from the original code, however, was highly dependent on the number of input moments and showed no convergence even at 12 moments. By contrast, after the orbit and code modifications were implemented, the main effect of increasing the constraining kinematic moments was to tighten the error bars while leaving the best-fit values largely unchanged (left panel of Fig. 2.12 or Fig, 12 in [119]).

Here we examine the progression of changes after each of the key modifications described in Sections 3.3 and 3.4 is implemented, all for the case of using 12 moments as constraints. The right panel of Figure 3.4 shows that implementing the code fixes described in Section 3.4 (purple dot-dot-dashed curve) and removing long-axis tubes (blue dot-dashed curve) move the best-fit $M_{\rm BH}$ by $\sim 10\%$ in comparison to $M_{\rm BH} \sim 2.2 \times 10^9 M_\odot$ from the original code (red dotted curve). Removing the box orbits increases $M_{\rm BH}$ to $\sim 2.9 \times 10^9 M_\odot$ (orange dot-dash-dashed and green dashed curve). The subsequent axisymmetrization of short-axis tubes (Section 3.3.1) has essentially no effect (black solid curve).

To ensure that the number of orbits included in the modelling is sufficient, we tested the effect of increasing the number of orbits. We increased the density of energy sampling by a factor of 4, from 40 energy values to 160 over the same range. With 4 times the number of orbits, the best-fit $M_{\rm BH}$ changed by less than 3%, and the 1σ error changed by less than 10%, demonstrating that our results do not depend on the exact number of orbits used.

3.6 Conclusion

We have presented a revised version of the triaxial orbit superposition code by van den Bosch et al. [202], which we refer to as the TriOS code, that is capable of properly modeling axisymmetric systems. The original code was designed for triaxial systems with (discrete) reflection symmetry along each of the three principal axes. The setup was not capable of modeling exactly axisymmetric systems in which the orbit library should respect (continuous) azimuthal symmetry about the symmetry axis.

We have implemented two main changes needed for modeling axisymmetric systems within the triaxial code: excluding all orbit types that are not allowed in an axisymmetric model, and enforcing axisymmetry among the allowed orbits. In the case of oblate axisymmetry, our recipe involves (1) axisymmetrizing the short-axis tube orbits by creating multiple copies of the orbits rotated about the symmetry axis (Section 3.3.1), (2) setting the viewing angle ψ to be sufficiently close to 90° to allow no long-axis tube orbits (Section 3.3.2), and (3) excluding the stationary start space used to generate box orbits (Section 3.3.3).

We have made further improvements and corrections to the code in general. We discussed an issue with slowly precessing quasi-planar orbits that are misclassified and are "mirrored" improperly in the orbit library (Section 3.4.1). We also corrected a problem with the logarithmic halo implementation that prevented checking energy conservation of the integrated orbits (Section 3.4.2). We achieved a factor of 2 to 3 speedup in the runtime of the code by adopting a different algorithm for modeling PSF convolution (Section 3.4.3). Finally, we allowed the orbit sampling and mass constraints to be set independently (Section 3.4.4).

For NGC 1453, we found the shape of the χ^2 contours for $M_{\rm BH}$ to vary significantly as we went through the step-by-step axisymmetrization procedure described in this paper (Figure 3.4). As we described in Chapter 2 [119], the orbit models favored no black holes when we used the original code with typical (near) axisymmetric parameters in the literature and four Gauss-Hermite moments to constrain the stellar LOSVDs. In contrast, we obtained a well constrained non-zero $M_{\rm BH}$ using our final axisymmetrized code including all the changes described in Sections 3.3 and 3.4.

One issue that warrants further investigation in triaxial models is the equilibrium behavior of quasi-planar orbits in regions where the potential is nearly spherical, e.g., well within a SMBH's SOI, or far outside the galaxy's effective radius in a spherical dark matter halo. As we discussed in Section 3.4.1, the subset of quasi-planar orbits with precession times longer than the integration time has a nearly constant $\vec{L}$ and is misclassified and mirrored incorrectly in the original code. Furthermore, the integration time for these orbits is not long enough to fill the allowed volume of phase space. For axisymmetric systems, we resolve these issues in our revised code described in this paper by including only short-axis tubes and enforcing axisymmetry in the orbits, while preserving L_z.

We also expect the severity of the orbit integration issue to vary from system to system: the better a SMBH's SOI is resolved by the available kinematic data, the more care is needed to test orbital integration time because quasi-planar orbits occupy a large fraction of the orbit library, and more orbits are deeper in the SMBH's potential and hence have longer precession times. For the NGC 1453 SMBH studied in Chapter 2, Liepold et al. [119], and here, since our kinematic data do not reach deep inside the SOI, orbits in our mass models with precession time exceeding 200 dynamical times account for less than 4% of luminosity within the central arcsecond. The integration issue (and the resulting misclassification) therefore does not significantly impact our results, as is evidenced by the similarity between the solid black and green dashed curves in Figure 3.4. We expect a different situation for better resolved systems such as the M87 and Milky Way SMBHs.

In future work, a straightforward solution to ensure that quasi-planar orbits are repre-

sentative of their equilibrium distributions is to extend the default integration time of 200 dynamical times in the code. Our preliminary tests suggest that integrating the orbits up to ~ 10 times longer is computationally feasible, but this may still be insufficient for the orbits closest to the SMBH and in the outermost part of the galaxy where the precession times are slowest. A more reliable treatment of these orbits would be needed.

Appendix

3.A Criterion for Existence of Long-axis Tubes

We use Stäckel potentials to gain insights into the existence of long-axis tubes. A potential is said to be in Stäckel form if it can be written as:

$$V(\lambda,\mu,\nu) = -\frac{F(\lambda)}{(\lambda-\mu)(\lambda-\nu)} - \frac{F(\mu)}{(\mu-\nu)(\mu-\lambda)} - \frac{F(\nu)}{(\nu-\lambda)(\nu-\mu)}, \tag{3.7}$$

for some function $F(\tau)$ where (λ,μ,ν) are ellipsoidal coordinates defined as the roots of τ in the equation

$$\frac{x^2}{\tau+\alpha} + \frac{y^2}{\tau+\beta} + \frac{z^2}{\tau+\gamma} = 1\,, \tag{3.8}$$

such that $-\gamma \leq \nu \leq -\beta \leq \mu \leq -\alpha \leq \lambda$. Here, (α,β,γ) are constants that define the coordinate system. Such a potential is said to be separable in these coordinates. When a density corresponding to a Stäckel potential is projected in any direction to give a 2D surface density, it will have no isophotal twists [74]. Thus, we can use the viewing angles (θ,ϕ,ψ) of Binney 1985 [13] to define the relationship between the primary axes of the projected and intrinsic densities. This set of viewing angles imposes a constraint on the allowed values of (α,β,γ) given by:

$$\frac{\sqrt{\beta-\alpha}}{\sqrt{\gamma-\beta}} = \sqrt{\frac{\sin^2\theta}{\cot 2\psi \sin 2\phi \cos\theta + \cos^2\phi(\cos^2\theta+1) - 1}}. \tag{3.9}$$

This expression follows from Equation (B9) of Franx 1988 [74]. Orbital structure in Stäckel potentials has been well studied [227]. This structure is what motivated the x-z start space described in Schwarzschild 1993 [174]. Long axis tube orbits pass through the x-z start-space above the focal curve, defined by

$$\frac{z^2}{\gamma-\beta} + \frac{x^2}{\alpha-\beta} = 1\,. \tag{3.10}$$

For large x, this curve is approximately a line given by $z \approx x\frac{\sqrt{\gamma-\beta}}{\sqrt{\beta-\alpha}}$. Therefore, the angle that this line forms with the z axis can be written simply in terms of the viewing angles as

$$\tan\eta = \sqrt{\frac{\sin^2\theta}{\cot 2\psi \sin 2\phi \cos\theta + \cos^2\phi(\cos^2\theta+1) - 1}}\,. \tag{3.11}$$

Any orbits launched initially between the focal curve and the positive z-axis in the x-z plane will be long axis tubes which violate axisymmetry. To effectively achieve axisymmetry, the angle η must be small enough for no orbits to be sampled above the focal curve. Since the line defined by the angle η is a lower bound to this curve, if all initial orbits in the positive x-z quadrant are launched outside of the approximate angular region between the z-axis and the angle η, there will be no long-axis tubes in the model.

This expression is derived for Stäckel potentials. However, in the absence of isophotal twists, we expect it to apply reasonably well to more realistic models as they can often be locally approximated by a Stäckel potential [168]. A central SMBH is inconsistent with a Stäckel potential and can thus destroy the ordered orbital structure. However, we suggest that Equation (3.11) could give a rough rule-of-thumb for where the boundary between long-axis and short-axis tubes will exist in models from the code, particularly at radii far from the SMBH.

The stellar mass distribution is represented by an MGE in our models. Each gaussian component is stratified on similar ellipsoids, and can thus be related to its deprojection via the equations given in Binney 1985 [13]. These equations can be rearranged to give

$$\frac{T}{1-T} = \frac{\sin^2\theta}{\cot 2\psi \sin 2\phi \cos\theta + \cos^2\phi(\cos^2\theta + 1) - 1}, \tag{3.12}$$

where $T = (1-p^2)/(1-q^2)$ of each MGE component. For an MGE with no isophotal twists, each MGE component has the same triaxiality parameter, T. Thus, in this case, the angle, η, can be written simply as:

$$\eta = \tan^{-1}\sqrt{\frac{T}{1-T}}, \tag{3.13}$$

where T is the triaxiality parameter for each MGE component. Two examples of triaxial start spaces for NGC 1453 models are shown in Figure 3.6. The boundary between long-axis tubes and short-axis tubes is well approximated by the angle η for a wide range of galaxy shapes.

3.B Thin Orbit Finding

The TriOS code uses the thin orbit curve to construct its start space. This curve has to be found numerically in the x-z plane. For a given angle in this plane, the thin orbit radius is found by integrating test orbits starting at different radii. For each orbit, the radius of the orbit is recorded each time it passes through the x-z plane. The thin orbit radius is found by minimizing the difference between the maximum and minimum of these radii.

This algorithm should work for triaxial models but needs some revision in the axisymmetric case, particularly when there is no central density cusp or mass concentration. In this case, when close enough to the center, the potential should be well approximated by a harmonic oscillator. When the potential is axisymmetric, the motion can be regarded as

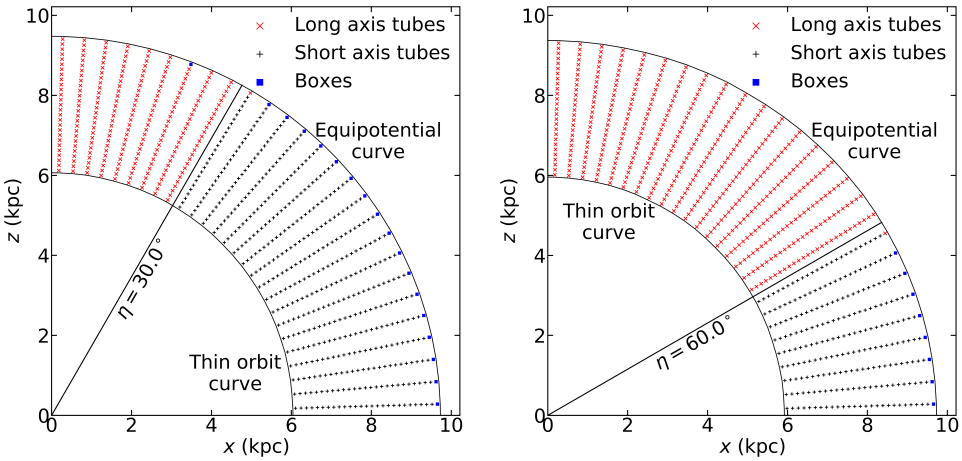

Figure 3.6: Same as Figure 3.1 but for two additional mass models with larger triaxiality: (left) triaxiality parameter $T = 0.25$, (luminosity averaged) shape parameters $(u, p, q) = (0.96, 0.95, 0.77)$, and viewing angles $(\theta, \phi, \psi) = (67.62°, -28.38°, 86.61°)$, and (right) $T = 0.75$, $(u, p, q) = (0.96, 0.85, 0.79)$, and $(\theta, \phi, \psi) = (48.74°, -51.33°, 67.15°)$. The diagonal black line in each panel represents the angle η given in Equation (3.2). As in Figure 3.1, this angle approximates well the boundary separating long-axis (red symbols) and short-axis (black symbols) tube orbits in the x-z start space.

two separate contributions: an oscillation in the z-axis and a closed elliptical orbit about the z-axis. Since the x-y motion constitutes a closed ellipse centered on the z axis, all orbits will pass through the x-z plane at a fixed x value, with some z value. The orbit width is then simply set by the maximum and minimum z values. Thus, for a given ray in the x-z plane, the orbital width in this plane can be minimized by simply taking the initial radius to be as small as possible. To solve this issue when running an axisymmetric model, we instead record radii when passing through the x-y plane. Closed ellipses will have a finite width in this plane while all thin orbits should pass through this plane in a circle of 0 width.

It is unclear how much this issue should affect the resulting orbit libraries. If orbits are sampled starting at the origin instead of the thin orbit, the result should be a less uniform sampling of angular momentum. There should also be some range of energies where the thin orbit radius is not estimated to be 0 or the correct value, but rather somewhere in between. This would result in a significantly non-uniform sampling of angular momentum since orbits passing through the x-z plane within this radius will be undersampled relative those that do not. This issue should be essentially resolved outside of the axisymmetric limit, or if a black hole or density cusp is included. However, axisymmetric studies that use this code with no central cusp may be affected [88].

3.C Mock Recovery Tests

In the above sections, we demonstrated that our changes to the TriOS code result in a consistent, non-zero SMBH mass estimate for our NGC 1453 dataset. Here, we show that the changes correctly recover the SMBH mass in a mock dataset with known parameter values. Mock tests have been performed within various other Schwarzschild codes [e.g. 197, 34, 130, 179, 208].

For our mock galaxy, we use a flattened version of the spherical potential introduced by Siopis et al. [179]. These models have an axisymmetric gravitational potential given by

$$\Phi(R,z) = \frac{1}{2}V_c^2 \ln\left(\frac{R^2 + z^2/q_\Phi^2}{1\ \mathrm{pc}^2}\right) - \frac{GM_{\mathrm{BH}}}{\sqrt{R^2+z^2}}, \tag{3.14}$$

where q_Φ is the flattening of the potential due to the extended mass distribution. The stellar DF is chosen to have a Michie-like form:

$$f = A \exp\left\{-\left[\frac{E + L_z^2/(2r_a^2)}{\sigma^2}\right]\right\} L_z^{2N} \sqcap (E_1, E, E_2), \tag{3.15}$$

where A is the normalization, and r_a, N, σ, E_1, and E_2 are parameters of the model: r_a is an anisotropy distance, N controls the L_z dependence, σ is a characteristic velocity dispersion, and E_1 and E_2 are energy cutoffs. The symbol $\sqcap$ denotes a step function defined by

$$\sqcap(E_1, E, E_2) = \begin{cases} 1, & \text{if } E_1 \le E \le E_2 \\ 0, & \text{otherwise.} \end{cases} \tag{3.16}$$

Because L_z only enters Eq. (3.15) in even powers, there is additional freedom in how f differs for positive and negative values of L_z. Here, we set a fixed fraction of stars to rotate in each positive direction. In this model, the stars are essentially regarded as massless tracers of the underlying potential in eq. 3.15. Even when the potential is chosen to be spherical, the stellar distribution function can be axisymmetric.

We use the same potential parameters as Siopis et al. [179], with $V_c = 220$ km s^{-1} and $M_{\mathrm{BH}} = 1.126 \times 10^8\ M_\odot$. We generated two models: one model with a spherical potential ($q_\Phi = 1$) to compare with Siopis et al. [179], and one model with a flattened potential ($q_\Phi = 0.95$). The models both have a sphere of influence of about 10 pc. We also use the same two component DF parameters as Siopis et al. [179]: the first component is a non-rotating nearly spherical bulge-like component which has $\sigma = 160$ km s^{-1}, $r_a = 600$ pc, $N = 0$, $E_1 = \Phi(10\ \mathrm{pc})$, $E_2 = \Phi(1000\ \mathrm{pc})$ with equal numbers of stars having positive and negative L_z; the second component is a rotating disk-like component w has $\sigma = 120$ km s^{-1}, $r_a = 200$ pc, $N = 2$, $E_1 = \Phi(10\ \mathrm{pc})$, $E_2 = \Phi(1000\ \mathrm{pc})$ with 3/4 of the stars having positive L_z and 1/4 having negative L_z. The two components have equal numbers of stars.

We draw points in phase space from this distribution function for each star to generate mock data. We use a nearly edge-on projection, with an inclination of $\theta = 89°$. For the model

with the flattened potential, we convolve the projected positions with a circular gaussian PSF with standard deviation 5 pc. We bin the stars into mock IFU data with a resolution of 10 pc, with a square FOV of 1000 pc. We fit an MGE to the projected surface brightness. We then run Voronoi binning on all bins with central radii > 20 pc, resulting in 12 inner unbinned kinematic points and 108 larger outer bins. In order to keep the bins between the two models fixed, we use the Voronoi bins derived from the spherical potential. Each LOSVD is fit with a Gauss-Hermite expansion up to h_{12}. Gaussian noise is added to each LOSVD bin, resulting in a scatter of about 0.03 in each moment and about $0.03\sqrt{2}\sigma$ in the average velocity and velocity dispersion for each bin. We draw 20 realizations of this noise, and run the updated TriOS code for each realization.

Figure 3.7 shows the resulting constraint on $M_{\rm BH}$ for each noise realization. The left panel is for the mock in the spherical potential, while the right panel is for the mock in the flattened potential. The kinematic contribution to the reduced χ^2 in these realizations ranges from 0.81 to 0.94 for the spherical potential and 0.71 to 0.84 for the flattened potential, indicating a good fit to the projected kinematics for all realizations.

The average SMBH masses and corresponding sample standard deviations from these combined 20 runs are $M_{\rm BH} = (1.17 \pm 0.09) \times 10^8\ M_\odot$ for the spherical potential, and $M_{\rm BH} = (1.18 \pm 0.13) \times 10^8\ M_\odot$ for the flattened potential. In both test cases, the estimated $M_{\rm BH}$ values are in excellent agreement with the true value of $M_{\rm BH} = 1.126 \times 10^8 M_\odot$.

Acknowledgements

We thank Jonelle Walsh for useful discussions and a critical reading of the manuscript, and Karl Gebhardt for discussions about mock tests. We also thank the anonymous referee for their helpful comments and suggestions. M.E.Q. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), PGSD3-517040-2018. C.-P.M. acknowledges support from NSF AST-1817100, HST GO-15265, HST AR-14573, the Heising-Simons Foundation, the Miller Institute for Basic Research in Science, and the Aspen Center for Physics, which is supported by NSF grant PHY-1607611. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) at the San Diego Supercomputing Center through allocation AST180041, which is supported by NSF grant ACI-1548562.

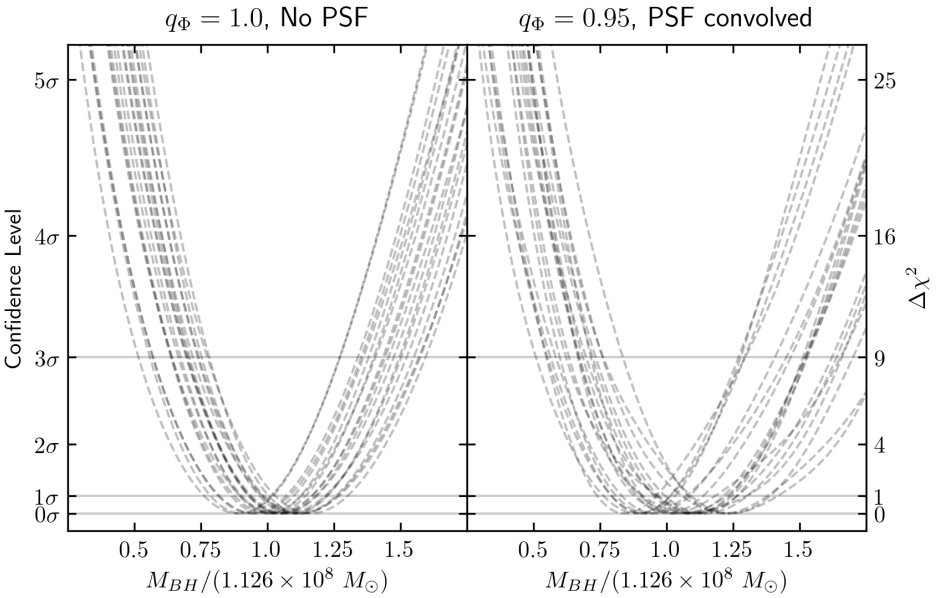

Figure 3.7: Illustration of the $M_{\rm BH}$ constraints for the mock datasets described in the text. Each dashed curve represents a separate realization of the noise. In the left panel, the potential is spherical, no PSF convolution is performed and each DF component has 5×10^8 stars. In the right panel, the model is flattened, projected stellar positions are convolved with a circular gaussian PSF with a standard deviation of 5 pc, and each DF component has 5×10^9 stars. The 1D χ^2 curves are obtained by marginalizing over V_c in the smoothed 2D χ^2 landscape generated by Gaussian Process regression with a squared-exponential covariance function [150].

Chapter 4

Triaxial Orbit-based Dynamical Modeling of Galaxies with Supermassive Black Holes and an Application to NGC 1453

The past two chapters have followed a perhaps paradoxical path – using a triaxial Schwarzschild orbit modelling code to produce axisymmetric models. That process allowed us to better understand in a practical sense a number of the forces at hand in practical usage of this code, in particular the orbit phase space sampling, the interplay between triaxiality and the various orbit families, and the impact of different constraints on the stellar line of sight velocity distributions.

Here we use these insights to apply the TriOS triaxial orbit superposition code to produce models with fully triaxial stellar mass distributions. Along the way we find a useful parameterization of the deprojection, characterized by a triplet of shape parameters (T, $T_{\rm maj}$, $T_{\rm min}$) which relatively intuitively map into either relations between the ratios of axis lengths or viewing angles. This parameterization allows for a more efficient and comprehensive search through the now 6-dimensional parameterspace of the problem. We also find additional prescriptions for properly constructing triaxial models with this code and fix a number of bugs.

We apply the revised code to the same galaxy that was explored with axisymmetry in Chapter 2, NGC 1453, a fast-rotating elliptical galaxy within the MASSIVE survey. These models prefer a relatively triaxial model and strongly rejects axisymmetry with $T = 0.33 \pm 0.06$, where T is the usual triaxiality parameter and $T = 0$ corresponds to oblate axisymmetry while $T = 1$ corresponds to prolate axisymmetry. The intermediate-to-long axis ratio is found to be $p = 0.933^{+0.014}_{-0.015}$ and the short-to-long axis ratio is found to be $q = 0.779 \pm 0.012$. Despite this very different shape from the axisymmetric case (which has $T = 0$ and $p = 1$), we find that the preferred central supermassive black hole mass, stellar mass-to-light ratio, and dark matter mass (enclosed within 15 kpc) is essentially unchanged

and is consistent with the axisymmetric case. Importantly, this work makes NGC 1453 one of only a small handful of galaxies where a black hole was measured along with a triaxial stellar potential.

This chapter was originally published as

Matthew E. Quenneville, Christopher M. Liepold, and Chung-Pei Ma. "Triaxial Orbit-based Dynamical Modeling of Galaxies with Supermassive Black Holes and an Application to Massive Elliptical Galaxy NGC 1453". In: *Astrophys. J.* 926.1, 30 (Feb. 2022), p. 30. DOI: 10.3847/1538-4357/ac3e68

Minor alterations to the text have been made to ensure fluency throughout this dissertation.

4.1 Introduction

Elliptical galaxies exhibit a wide range of isophotal shapes and surface brightness profiles. There is an intrinsic uncertainty in inferring the 3D stellar luminosity density from the observed 2D isophotes on the sky. When stellar kinematics from spectroscopic observations are combined with photometric information, stronger constraints can be placed on the intrinsic 3D shapes of elliptical galaxies (e.g., [13, 73]). An idealized galaxy obeying exact axisymmetry would, by construction, have a regular surface brightness distribution without any isophotal twists and have perfectly aligned photometric and kinematic axes. Triaxial systems, on the other hand, can have isophotal twists, misaligned photometric and kinematic axes, and other spatially varying kinematic features absent in an axisymmetric system. This consideration led Binney [13] to argue that triaxiality is common among elliptical galaxies.

Since then, a more detailed picture has emerged. Elliptical galaxies with lower stellar mass ($M_* \lesssim 10^{11.5} M_\odot$) tend to exhibit properties typical of axisymmetry [e.g., 49, 222, 27, 71]. Comparatively, elliptical galaxies with higher mass ($M_* \gtrsim 10^{11.5} M_\odot$) typically exhibit photometric twists, slow or no rotation, and misalignments between the photometric and kinematic axes, suggesting triaxial intrinsic shapes [e.g., 209, 210, 211, 110, 54, 83, 55]. Thus, it is vital to understand the role of triaxiality in dynamical galaxy modelling, particularly in studying massive elliptical galaxies and their central black holes in the local universe.

The most massive SMBHs observed in the nearby universe lie in centers of some of the most massive nearby elliptical galaxies [126]. However, few triaxial SMBH mass ($M_{\rm BH}$) measurements have been published thus far, perhaps because of the complexity in orbital structures, high-dimensional parameter space, and the associated computational cost required to model stellar orbits in triaxial potentials. To date, all published $M_{\rm BH}$ measurements based on triaxial orbit modeling have been performed using the code initially presented in van den Bosch et al. [202]. This code was first applied to determine the intrinsic shapes and $M_{\rm BH}$ of two fast-rotating elliptical galaxies M32 and NGC 3379 [199]. In this work, M32 was found to be near oblate axisymmetry with $M_{\rm BH} = (2.4 \pm 1.0) \times 10^6 M_\odot$, fully consistent with $M_{\rm BH}$ from earlier axisymmetric models [206, 104, 213]. NGC 3379, on the other hand, was found to be moderately triaxial, and the inferred $M_{\rm BH} = (4 \pm 1) \times 10^8 M_\odot$ was double the value

derived from axisymmetric models [79, 177]. In a subsequent application to the S0 galaxy NGC 3998 [219], the best-fit model was found to be moderately triaxial although oblate axisymmetry was not ruled out.

Feldmeier and Krause [65] applied the van den Bosch et al. [202] code to the nuclear star cluster and SMBH at the Galactic center. The cluster shape was strongly triaxial, and the inferred $M_{\rm BH}$ was consistent within $1\,\sigma$ of the values inferred from the orbit of the S2 star [85, 44].

More recently, den Brok et al. [42] used the van den Bosch et al. [202] code to model PGC 046832. This galaxy exhibits dramatic twists, and the resulting models preferred strong variations in triaxiality. However, while axisymmetric models suggested a central black hole mass of $6\times10^9 M_\odot$, the triaxial models prefer models with no central black hole. Instead they report an upper bound on the central black hole mass of $2\times10^9 M_\odot$. This differs significantly from the value determined from axisymmetric models.

In addition to these published triaxial $M_{\rm BH}$ values, the van den Bosch et al. [202] code has been used to determine several $M_{\rm BH}$ in the nearly axisymmetric limit [176, 220, 217, 218, 1]. It has also been used to estimate the intrinsic triaxiality of galaxies under the assumption of a fixed $M_{\rm BH}$ [e.g., 202, 115, 229, 230, 154, 226, 103].

We have been revamping the van den Bosch et al. [202] code for a systematic study of the SMBHs and other mass components in the ~ 100 most massive local early-type galaxies in the MASSIVE survey [126]. As a first step, we introduced a version of the code capable of achieving the exact axisymmetric limit in Chapters 2 and 3 [119, 156]. The original [202] code was (intentionally) not built to respect axisymmetry, but it had been used to perform (nearly) axisymmetric orbit modeling, leading to unexplained inconsistencies when the resulting $M_{\rm BH}$ values were compared to those from axisymmetric orbit codes (e.g., [1]). Our axisymmetrized version of the code has bridged this gap and now enables dynamical modeling of galaxies using stellar orbits that properly obey axisymmetry. We applied our axisymmetrized code to NGC 1453, a fast-rotating elliptical galaxy in the MASSIVE survey, and obtained a significant detection of its SMBH with $M_{\rm BH} = (2.9\pm0.4)\times10^9\ M_\odot$ (Chapter 2; [119]). Models without black holes were excluded at the 8.7σ level.

For clarity, we refer to the original code (which was unnamed) by the citation [202], and refer to our versions as the TriOS (Triaxial Orbit Superposition) code.

In this chapter, we move beyond the axisymmetric limit of Chapter 3 [156], and present a triaxial version of the TriOS code and a first application of this code. This triaxial TriOS code differs in a number of major ways from the original van den Bosch et al. [202] code. We have implemented these changes to correct a number of bugs and issues that we uncovered during extensive tests of the original code for triaxial potentials. As a start, we correct a major error in the orbit construction part of the code that incorrectly flips some velocity components for the tube orbits. Our tests indicate that for most viewing angles, correcting this mistake has a significant impact on the resulting orbital kinematics and galaxy model parameter recovery within the code. Other major changes include (i) modifying the acceleration table used for orbit integration to gain a significant speedup in runtime, (ii) resolving issues with insufficient orbit sampling that can result in spurious shape preferences, and (iii) using a more uniform

mass binning scheme to eliminate frequent problems in satisfying mass constraints. Details of these changes are described in Section 4.4.

In addition to these code changes, we introduce a new set of shape parameters in this chapter (Section 4.3) that are chosen to improve the efficiency of parameter searches in triaxial galaxy shapes and orientations. These parameters strike a balance between sampling in galaxy intrinsic shape and galaxy orientation, and result in fewer unrealistically flat galaxy shapes. To place these new parameters in context, we provide a summary (Section 4.2) of the parameters used in previous work to describe a triaxial galaxy's intrinsic and observed axis ratios, the relations of viewing angles and sky projections, and how an observed surface brightness is deprojected to obtain a 3D intrinsic shape within the TriOS code.

We apply our triaxial TriOS code to NGC 1453 in the final part of the chapter (Section 4.5). Since triaxial modeling typically involves at least five parameters (three for shapes and at least two for mass parameters), we introduce an efficient new search strategy for sampling this multi-dimensional parameter space. This new strategy does not rely on direct grid searches used in previous orbit modeling studies. Instead, we apply nested Latin hypercube sampling to a 6D parameter space and are able to converge to a best-fit model for NGC 1453 with an order-of-magnitude fewer sample points. The resulting best-fit triaxial model is compared to the best-fit axisymmetric model from Chapter 2 [119].

4.2 Modeling a Triaxial Galaxy

In this section we summarize the information relevant for modeling a triaxial galaxy, e.g., coordinate systems, intrinsic and apparent shape parameters, viewing angles, and sky projections.

4.2.1 Intrinsic Shapes and Axis Ratios

To describe the 3D structure of a galaxy, we use a Cartesian coordinate system centered at the galaxy's nucleus, in which the x, y, and z axes are directed along the intrinsic major, intermediate, and minor axes of the galaxy, respectively. The z-axis is therefore the symmetry axis of an oblate axisymmetric galaxy, and the x-axis is the symmetry axis of a prolate axisymmetric galaxy.

It is convenient to use a different coordinate system to describe properties projected on the sky. We follow the standard practice and take the x' and y' axes of this coordinate system to be along the major and minor axes of the projected surface brightness distribution of a galaxy. The z' axis is along the line-of-sight.

We use a, b, c to denote the lengths of the three principal axes of a triaxial ellipsoidal isodensity surface, assuming $c \leq b \leq a$. We use a' and b' to denote the lengths of the (observed) major and minor axes of the projected ellipse on the sky. Four useful axis ratios are

$$p = \frac{b}{a}\,, \quad q = \frac{c}{a}\,, \quad u = \frac{a'}{a}\,, \quad q' = \frac{b'}{a'}\,, \tag{4.1}$$

where p is the intrinsic intermediate-to-major axis ratio, q is the intrinsic minor-to-major axis ratio, u represents a compression factor between the intrinsic major axis and the apparent major axis on the sky due to projection, and q' is the flattening of the projected shape. These quantities obey the inequalities

$$\begin{aligned} & 0 \leq c \leq b' \leq b \leq a' \leq a\,, \\ \text{or} \quad & 0 \leq q \leq uq' \leq p \leq u \leq 1\,. \end{aligned} \tag{4.2}$$

The upper and lower limits of u correspond to the intrinsic major axis lying in the plane of the sky ($u = 1$ or $a' = a$) and the intrinsic intermediate axis lying in the plane of the sky ($u = p$ or $a' = b$), respectively.

The commonly used triaxiality parameter is

$$T = \frac{a^2 - b^2}{a^2 - c^2} = \frac{1 - p^2}{1 - q^2}\,, \tag{4.3}$$

which ranges between 0 for an oblate axisymmetric shape ($a = b$), and 1 for a prolate axisymmetric shape ($b = c$), with values between 0 and 1 indicating a triaxial shape.

4.2.2 Viewing Angles and Sky Projections

A line of sight between an observer and a galaxy is specified by two viewing angles (θ, ϕ), where θ and ϕ are the usual polar angles in the galaxy's intrinsic (x, y, z) coordinate system. Thus, $\theta = 0^\circ$ is for a line of sight along the intrinsic minor axis (i.e., a face-on view down the z-axis), and $\theta = 90^\circ$ is for lines of sight in the $x - y$ plane (i.e., an edge-on view with the intrinsic minor axis in the sky plane). Similarly, $\phi = 0^\circ$ is for lines of sight in the $x - z$ plane (i.e., the intrinsic intermediate axis is in the sky plane), and $\phi = 90^\circ$ is for lines of sight in the $y - z$ plane (i.e., the intrinsic major axis is in the sky plane).

Given a triaxial 3D density stratified on similar concentric ellipsoids, the viewing angle θ and ϕ are sufficient to project the 3D shape and determine the 2D projected coordinate system (x', y'). To de-project an observed 2D shape on the sky, however, a third angle, ψ, is needed to completely specify the intrinsic coordinate system. This third angle ψ specifies the remaining degree of freedom once θ and ϕ are fixed – a rotation of the galaxy around the line of sight. More precisely, ψ is defined as the angle between the y' axis, and the line defined by the intersection of the $x' - y'$ and $x - y$ planes. When $\psi = 0^\circ$, the $x - y$ plane and $x' - y'$ plane intersect along the y' axis; when $\psi = 90^\circ$, the $x - y$ plane and $x' - y'$ plane intersect along the x' axis.

Together, the three angles (θ, ϕ, ψ) uniquely specify the orientation of the intrinsic axes with respect to the projected axes. If the 3D density is stratified on similar concentric ellipsoidal surfaces, the axis ratios (p, q, u) of Equation (4.1) can be uniquely determined from the projected surface brightness and (θ, ϕ, ψ) using the equations from Appendix A of de Zeeuw and Franx [41].

4.2.3 Deprojecting Observed Surface Brightness

Within the TriOS code, the 3D stellar density distribution is described by a sum of multiple Gaussian components of varying widths and axis ratios using the Multi-Gaussian Expansion (MGE) scheme [24]. To determine these components, one first fits a 2D MGE to the observed surface brightness of the galaxy. Each MGE component is allowed to have its own projected flattening q' to account for radially varying ellipticity in the observed isophotes. In addition, each MGE component can have a different position angle (PA) to accommodate any observed isophotal twists.

In general, the deprojection of a 2D surface brightness distribution to give a 3D triaxial luminosity density is not unique. MGE is a parametric method of choosing one particular 3D density for a given 2D surface brightness and set of intrinsic axes. Non-parametric deprojection methods have also recently been developed for triaxial galaxies [38], but the TriOS code is not yet capable of using these deprojections.

For a set of (θ, ϕ, ψ) that specifies the alignment of the galaxy's intrinsic principle axes (x, y, z), one can determine the deprojection of each MGE component that shares these principle axes (if a valid deprojection exists). This deprojection is unique due to the assumption that each 2D gaussian corresponds with a 3D gaussian density with similar concentric ellipsoidal surfaces of constant density. The axis ratios p and q of each deprojected MGE component can have their own values. The triaxiality parameter T, on the other hand, has the convenient property that it is identical for all MGE components when the components share the same PA (i.e., no isophotal twists).[1]

4.3 New Parameters for Triaxial Space Sampling

4.3.1 Prior practice

As discussed in Section 4.2, either (p, q, u) or (θ, ϕ, ψ) can be used to specify the shape of a triaxial galaxy and its sky projections. One can in principle search in either space when running orbit models to determine a galaxy's intrinsic shape and mass parameters. In practice, however, prior triaxial orbit modeling studies favored (p, q, u) over the angles. In these studies, the orbit models were typically run for a grid of regularly spaced values of (p, q, u) [e.g., 203, 199, 219, 102]. In a few other triaxial studies, u was fixed to some value close to 1 while the parameter search was conducted over p and q in a regular 2D grid [e.g., 229, 230, 154]. Since $u \sim 1$ corresponds to the intrinsic major axis lying close to the sky plane, these studies did not search over all allowed viewing angles.

The argument used by van den Bosch and van de Ven [203] for favoring conducting parameter searches in (p, q, u) rather than (θ, ϕ, ψ) is that a change in the angles can result

[1] This is valid as long as the line-of-sight does not lie in a principal plane of the triaxial shape. If it does, then all aligned 3D ellipsoids will have parallel or perpendicular PAs when viewed in projection, and differences in T cannot be inferred from differences in projected PA. We do not consider any models with lines-of-sight lying directly in the principal planes.

in either a very small or very large change in axis ratios, depending on the angles being explored. We note, however, that the converse is also true: a change in the axis ratios can result in either a very small or very large change in the principal axes' alignment, depending on the values of these ratios. Two models with similar axis ratios, but viewed along very different lines of sight, can result in very different observables. An optimal sampling should consider both the intrinsic shape and the alignment of the line of sight.

4.3.2 Properties of new parameters

Here we propose a new set of variables to parameterize a galaxy's intrinsic triaxial shape and its sky projections. The advantages of conducting parameter searches in these variables over either (p, q, u) or (θ, ϕ, ψ) during triaxial orbit modeling will be discussed in Section 4.3.4.

For the first shape parameter, we choose the triaxiality parameter T (Equation 4.3). We define the next two parameters with forms analogous to T:

$$\begin{aligned} T &= \frac{a^2 - b^2}{a^2 - c^2} = \frac{1 - p^2}{1 - q^2}\,, \\ T_{\rm maj} &\equiv \frac{a^2 - a'^2}{a^2 - b^2} = \frac{1 - u^2}{1 - p^2}\,, \\ T_{\rm min} &\equiv \frac{b'^2 - c^2}{b^2 - c^2} = \frac{(uq')^2 - q^2}{p^2 - q^2}\,, \end{aligned} \tag{4.4}$$

where $T_{\rm maj}$ parameterizes the length of the projected major axis, a', relative to its allowed limits a and b, and $T_{\rm min}$ parameterizes the length of the projected minor axis, b', relative to its allowed limits b and c. It then follows from the inequalities in Equation (4.2) that $(T, T_{\rm maj}, T_{\rm min})$ form a unit cube, i.e.,

$$\begin{aligned} &0 \le T \le 1\,, \\ &0 \le T_{\rm maj} \le 1\,, \\ &0 \le T_{\rm min} \le 1\,. \end{aligned} \tag{4.5}$$

The limiting cases represented by each face of the unit cube has the following physical significance: (i) $T = 0$ and 1 correspond to oblate axisymmetric ($a = b$ or $p = 1$) and prolate axisymmetric ($b = c$ or $p = q$) shapes, respectively; (ii) $T_{\rm maj} = 0$ and 1 correspond to the intrinsic major axis lying in the sky plane ($a' = a$ or $u = 1$) and the intrinsic intermediate axis lying in the sky plane ($a' = b$ or $u = p$), respectively; (iii) $T_{\rm min} = 0$ and 1 correspond to the intrinsic minor axis lying in the sky plane ($b' = c$ or $uq' = q$) and the intrinsic intermediate axis lying in sky ($b' = b$ or $uq' = p$), respectively. While both the $T_{\rm maj} = 1$ and $T_{\rm min} = 1$ planes correspond to the intrinsic intermediate axis in the sky plane, they represent two complementary ranges of viewing angles such that b is equal to the projected *m*ajor axis a' for $T_{\rm maj} = 1$, whereas b is equal to the projected *m*inor axis b' for $T_{\rm min} = 1$.

Equation (4.4), along with the requirement that $q > 0$, yields the inequality

$$\frac{(1-T)T_{\rm min}}{1-TT_{\rm maj}} < q'^2\,, \tag{4.6}$$

implying that for an observed axis ratio q' on the sky, only the $(T, T_{\rm maj}, T_{\rm min})$ region satisfying the inequality has valid deprojections. When the projected shape is flattened ($q' < 1$), some models within the unit cube will result in negative (and thus invalid) values of the squared minor axis length, c^2. This volume surrounds the line $(T, T_{\rm maj}, T_{\rm min}) = (T, 1, 1)$, which does not have a valid deprojection for any flattened projected shape.

4.3.3 Relating $(T, T_{\rm maj}, T_{\rm min})$ to old parameters

While Equations (4.4) relate our new parameters to (p, q, u), it is often useful to do the inverse and convert a given set of $(T, T_{\rm maj}, T_{\rm min})$ to (p, q, u). To do so, we use these sequential expressions

$$\begin{aligned} 1-q^2 &= \frac{1-q'^2}{1-(1-T)T_{\rm min}-q'^2TT_{\rm maj}}\,,\\ 1-p^2 &= T(1-q^2)\,,\\ 1-u^2 &= T_{\rm maj}(1-p^2)\,. \end{aligned} \tag{4.7}$$

For a given set of $(T, T_{\rm maj}, T_{\rm min})$, these equations define the deprojection from an observed MGE component with flattening, q', to its 3D shape parameters, (p, q, u).

Similarly, it is useful to convert $(T, T_{\rm maj}, T_{\rm min})$ to the angles (θ, ϕ, ψ):

$$\begin{aligned} \cos^2\theta &= T_{\rm min}(1-T\,T_{\rm maj})\,,\\ \tan^2\phi &= \frac{1-T_{\rm maj}}{T_{\rm maj}}\frac{1-T_{\rm min}}{1-T_{\rm min}(1-T)}\,,\\ \tan^2\psi &= \frac{[1-T_{\rm min}(1-T)]\,(1-T\,T_{\rm maj})(1-T_{\rm min})}{T^2(1-T_{\rm maj})T_{\rm maj}\,T_{\rm min}}\,. \end{aligned} \tag{4.8}$$

We choose to use the branch where $0° \le \theta \le 90°$, $0° \le \phi \le 90°$, and $90° \le \psi \le 180°$, though other equivalent branches exist as well.[2] The inverse expressions relating $(T, T_{\rm maj}, T_{\rm min})$ and (θ, ϕ, ψ) are given in Appendix 4.A.

Equations (4.7) and (4.8), as well as Equations (4.11) and (4.12), follow directly from the definitions in Equation (4.4), and the general expressions for the deprojection of a triaxial density that is stratified on similar, concentric ellipsoids [e.g., 41, discussed further in appendix 4.A]. Furthermore, since Equation (4.8) and its inverse Equation (4.12) make no reference to the observed flattening, the same set of values $(T, T_{\rm maj}, T_{\rm min})$ can be used for a density that is composed of multiple such components with different flattening values. Thus,

[2]For instance, if one prefers $0° \le \psi \le 90°$ and $0° \le \theta \le 90°$, then ϕ obeys $-90° \le \phi \le 0°$.

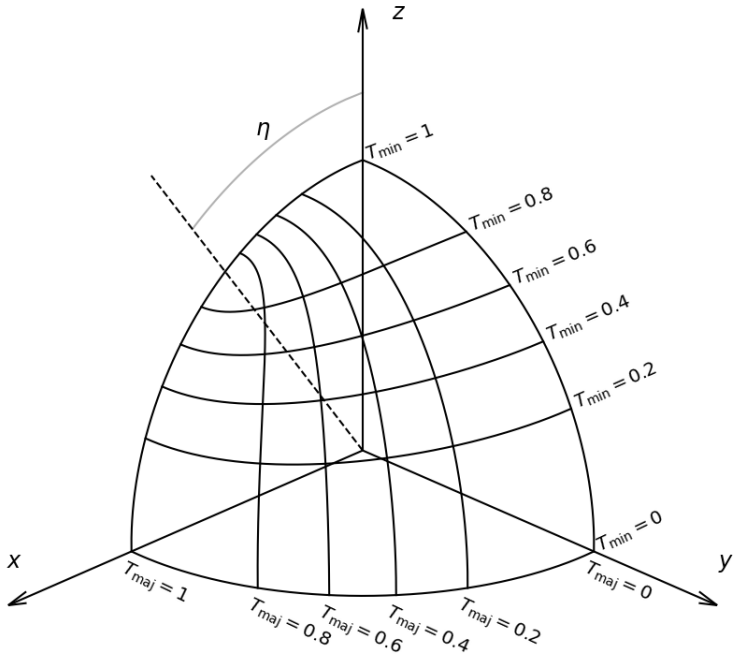

Figure 4.1: Isocontours of the new shape parameters, $T_{\rm maj}$ and $T_{\rm min}$, in a galaxy's coordinate system, where the x, y, and z axes are chosen to be the intrinsic major, intermediate, and minor axes, respectively. The triaxiality parameter, T, is assumed to be 0.35 here. The parameters $T_{\rm maj}$ and $T_{\rm min}$ are seen to change relatively uniformly with the line-of-sight direction, resulting in fewer unrealistically flattened models near non-deprojectable regions (see text).

$(T, T_{\rm maj}, T_{\rm min})$ and (θ, ϕ, ψ) are simply different parameterizations of the same space. Equations (4.7) and (4.8), along with the equations listed in Appendix 4.A, make no reference to the MGE formalism and are applicable to any triaxial system that meets these conditions. The existence and uniqueness of a valid deprojection are not affected by the choice of shape space parameterization outside the principal planes.

To illustrate the properties of $T_{\rm maj}$ and $T_{\rm min}$, we plot a set of lines of constant $T_{\rm maj}$ and $T_{\rm min}$ in a galaxy's intrinsic coordinate system x, y, and z in Figure 4.1. The corner points $(T_{\rm maj}, T_{\rm min}) = (1, 0)$, $(0, 0)$, and $(0, 1)$ correspond to viewing angles along the short, intermediate, and long axes, respectively. The point $(T_{\rm maj}, T_{\rm min}) = (1, 1)$ represents a line of sight lying along the line $\theta = \eta = \tan^{-1}(\sqrt{T/(1-T)})$ in the $x - z$ plane, which only results in a valid model for round projected shapes. For flattened shapes, there are no valid deprojections for lines of sight within a solid angle surrounding this direction. This

non-deprojectable region increases in size, as the projected shape becomes flatter.

4.3.4 Advantages of $T, T_{\rm maj}$ and $T_{\rm min}$

The parameters T, $T_{\rm maj}$, and $T_{\rm min}$ have a number of desirable properties. First, as Figure 4.1 illustrates, $T_{\rm maj}$ and $T_{\rm min}$ change relatively uniformly with the line-of-sight direction. This is in contrast to the axis ratio space, (p, q, u), in which tiny changes can result in large differences in the angles. For example, models with $p = 0.99$ and a fixed q would undergo a 90° rotation in ϕ when u is varied from 0.99 to 1.

Similarly, the galaxy shape varies much more uniformly with $(T, T_{\rm maj}, T_{\rm min})$ than with (θ, ϕ, ψ). Again, tiny changes in the latter can result in large differences in galaxy shape. For example, when an observed surface brightness (without isophotal twists) is deprojected into a 3D ellipsoidal shape with principle axes defined by $(\theta, \phi, \psi) = (89^\circ, 45^\circ, 90^\circ)$, Equation (A2) shows that the resulting 3D shape has $T = 0$, i.e., it is oblate axisymmetric. As ψ is increased from 90° by only $\sim 1^\circ$, however, the deprojected shape varies drastically, with oblate axisymmetry at $\psi = 90^\circ$ to prolate axisymmetry at $\psi \sim 91^\circ$, with the full range of triaxialities lying in between. From Equation (A2) with $\phi = 45^\circ$, we find prolate axisymmetry ($T = 1$) to occur when $\psi - 90^\circ = (90^\circ/\pi) \arctan\left(2\cos\theta/\sin^2\theta\right)$ on our chosen branch. As θ approaches 90°, the value of ψ that gives prolate axisymmetry approaches 90°. For $\theta = 89^\circ$ (and $\phi = 45^\circ$), prolate axisymmetry occurs at $\psi = 90.99985^\circ$.

The behavior in the example above arises from coordinate singularities in the (θ, ϕ, ψ) space. When the line-of-sight is chosen to lie in a principal plane (i.e., $\cos(\theta) = 0, 1$ or $\sin(2\phi) = 0$), it is impossible for continuous photometric twists to arise in projection as triaxiality is varied. One consequence of this is that the only valid values of ψ are 0° or 90°, meaning it is no longer an independent parameter. Thus, (θ, ϕ, ψ) are insufficient to fully specify the 3D projection. The parameters $(T, T_{\rm maj}, T_{\rm min})$, on the other hand, have no such singularity. In the above example, the proximity of the chosen value of θ to 90° causes the rapid shift in shape with ψ.

Another desirable property of $T_{\rm maj}$ and $T_{\rm min}$ is that, similar to T (see Section 4.2.3), they do not vary among MGE components with different axis ratios, so long as there are no isophotal twists.

This invariant property can be explained by identifying $T_{\rm maj}$ and $T_{\rm min}$ as the shifted and rescaled versions of the conical coordinates, $\mu_{\rm pro}$ and $\nu_{\rm pro}$ within the galaxy's intrinsic coordinate system [74], where $\mu_{\rm pro} = a'^2$ and $\nu_{\rm pro} = b'^2$. Since the coordinate surfaces of $\mu_{\rm pro}$ and $\nu_{\rm pro}$ are the same for all MGE components, the shifted and scaled quantities $T_{\rm maj}$ and $T_{\rm min}$ do not vary between components.

The advantages of T, $T_{\rm maj}$, and $T_{\rm min}$ are especially clear for systems not far from axisymmetry. Towards oblate axisymmetry ($T \approx 0$), we have $T_{\rm min} \approx \cos^2\theta$ and $T_{\rm maj} \approx \cos^2\phi$. Thus, a uniform sampling in $\sqrt{T_{\rm min}}$ and $\sqrt{T_{\rm maj}}$ will result in a nearly uniform sampling in the cosines of the inclination and the azimuthal angle. The same behavior holds towards prolate axisymmetry ($T \approx 1$) since the roles of $T_{\rm maj}$ and $T_{\rm min}$ are simply switched if the x and

Position	Long-axis tube	Short-axis tube	Intermediate-axis tube
(x, y, z)	(v_x, v_y, v_z)	(v_x, v_y, v_z)	(v_x, v_y, v_z)
$(-x, y, z)$	$(-v_x, v_y, v_z)$	$(v_x, -v_y, -\mathbf{v_z})$	$(v_x, -\mathbf{v_y}, -v_z)$
$(x, -y, z)$	$(-\mathbf{v_x}, v_y, -v_z)$	$(-v_x, v_y, -\mathbf{v_z})$	$(v_x, -v_y, v_z)$
$(x, y, -z)$	$(-\mathbf{v_x}, -v_y, v_z)$	$(v_x, v_y, -v_z)$	$(-v_x, -\mathbf{v_y}, v_z)$
$(-x, -y, z)$	$(\mathbf{v_x}, v_y, -v_z)$	$(-v_x, -v_y, v_z)$	$(v_x, \mathbf{v_y}, -v_z)$
$(-x, y, -z)$	$(\mathbf{v_x}, -v_y, v_z)$	$(v_x, -v_y, \mathbf{v_z})$	$(-v_x, v_y, -v_z)$
$(x, -y, -z)$	$(v_x, -v_y, -v_z)$	$(-v_x, v_y, \mathbf{v_z})$	$(-v_x, \mathbf{v_y}, v_z)$
$(-x, -y, -z)$	$(-v_x, -v_y, -v_z)$	$(-v_x, -v_y, -v_z)$	$(-v_x, -v_y, -v_z)$

Table 4.1: Corrected mirroring scheme of the three types of tube orbits in our TriOS code. Boldfaced velocity components have the opposite signs from the original scheme in Table 2 of van den Bosch et al. [202]. These components were flipped incorrectly in the original code.

z axis labels are interchanged. Thus, for nearly axisymmetric galaxies, a uniform sampling in $(T, \sqrt{T_{\rm maj}}, \sqrt{T_{\rm min}})$ results in fewer unrealistically flattened models.

4.4 Code Corrections and Improvements

In this section, we describe the key corrections, improvements, and speedups made to the van den Bosch. [202] code. See Section 3.4 (Section 4 of [156]) for other general changes that we had implemented (regardless of axisymmetry).

4.4.1 Correct orbital mirroring mistakes

The TriOS code is written for a static triaxial potential that is symmetric under reflection along each of the three principal axes of a triaxial system. Under this assumption, any orbital property only needs to be calculated in one octant of the orbit space; it can then be "mirrored" into the other seven octants by symmetry.

Taking advantage of this symmetry, the code initializes orbits in only one octant ($x, y, z > 0$) and integrates only these orbits. Seven additional copies of each orbit are then created by simply mirroring along the three axes. The recipe for how to flip the signs of the velocity components is given in Table 2 of van den Bosch et al. [202]. The exact procedure depends on whether the orbit is a short-axis tube, long-axis tube, or box. These orbits are classified as follow: throughout its trajectory, an orbit is labelled a box orbit if all three components of its angular momentum (L_x, L_y, L_z) change sign, and a tube orbit if exactly one component of angular momentum maintains its sign. The tube orbits are further classified according to the angular momentum component that maintains its sign, i.e., a long-axis (i.e. x-axis) tube maintains the sign of its L_x, an intermediate-axis (y-axis) tube maintains the sign of

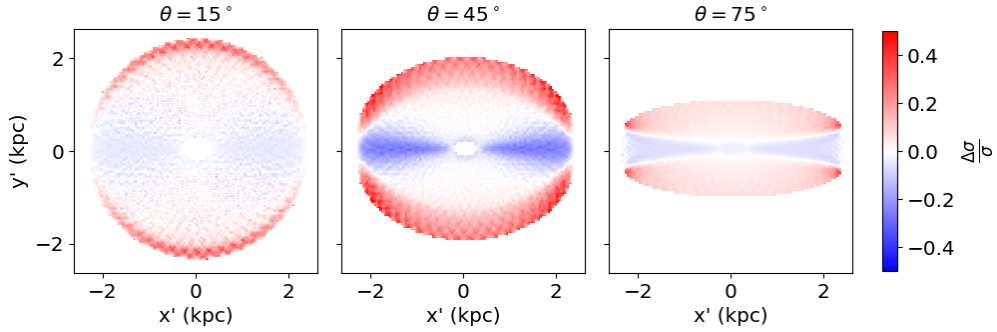

Figure 4.2: Illustration of the impact of the incorrect mirroring scheme in the van den Bosch et al. [202] code. We plot the fractional error between the incorrect and corrected schemes (see Table 1) in the kinematic map of the line-of-sight velocity dispersion, σ, for a single orbit. The orbit is chosen from the $x-z$ start space of a triaxial model with $T = 6 \times 10^{-6}$ for NGC 1453, but it is representative of typical short-axis tubes in a triaxial potential. Each panel represents a different viewing inclination angle θ. The fractional error is largest near $\theta = 45°$, reaching beyond 50% for some parts of the orbit.

L_y, and a short-axis (z-axis) tube maintains the sign of L_z. Orbits that don't fall into the tube or box orbit classifications are flipped in the same way as box orbits.

We discovered that the tube orbits are incorrectly flipped for four of the eight octants in Table 2 of van den Bosch et al. [202]. We indicate the incorrect components in boldface and give the corrected expressions in Table 4.1. The mistakes are such that the mirrored positions and velocities are inconsistent with one another, and the two do not combine to give a valid trajectory. A consequence of these mistakes is that the magnitude of each component of $\vec{L}$ is not always preserved by the mirroring, as it should be, and the resulting $|\vec{L}|$ is also not preserved. For instance, for the short-axis tube flip, the original recipe would change the amplitudes of L_x and L_y for 4 of the 8 copies, and the resulting total L would not be preserved. Similarly, L_y and L_z are incorrect for 4 copies of the long-axis tubes, and L_x and L_z are incorrect for 4 copies of the intermediate-axis tubes.

To illustrate the impact of the incorrect orbital flips, we plot the error in the line-of-sight velocity dispersion, σ, for a single short-axis tube orbit for three different viewing angles in Figure 4.2. We first integrate the trajectory of this orbit within the potential and then compute the 7 mirrored copies using the original and corrected flips in Table 4.1. The fractional difference in the projected σ between the two schemes is then plotted for three different values of viewing inclination angle θ. The errors vary across the plane of the sky and exceed 50% for $\theta = 45°$. For this orbit, the incorrect flip scheme tends to under-predict σ along the galaxy's projected major axis and over-predict σ near the edges. The orbit shown in Figure 4.2 is typical of short-axis tubes in triaxial potentials. Long-axis tube orbits exhibit

similar error patterns when the appropriate axis labels are switched. While the pattern of velocity dispersion error is different for each orbit, systematic errors with magnitudes of $10\% - 100\%$ are typical, with peak errors of over 1600% in some cases for orbital inclinations near $45°$.

To assess further the impact of the incorrect flips, we perform full orbit modeling for a grid of triaxial models for NGC 1453 using the original and then the corrected scheme. Overall, when the correct flips are used, we find that χ^2 is lowered by a wide range of values depending on the triaxiality and viewing angles. For instance, the value of χ^2 can decrease by more than 100 for strongly triaxial models, while it can change by less than 5 or even increase slightly for other models. The overall χ^2 landscape is therefore significantly altered by our corrections.

Due to the symmetry of the tube orbits, the errors in the orbital flips can cancel out when the galaxy is viewed along a principal axis. Nearly axisymmetric models that are viewed edge-on or face-on will be similarly unaffected. Outside of these special cases, the orbital kinematics have significant errors. The incorrect flips were not used in our axisymmetric modelling of NGC 1453 (Chapters 2 and 3; [119, 156]) since we used an axisymmetrization procedure in place of the flips in the TriOS code.

The discussion above is relevant only for tube orbits. For box orbits, we find the flips given in Table 2 of van den Bosch et al. [202] to be correct. However, in addition to this set of 8 mirrored orbits, we choose to include 8 more orbits for each point in the stationary start space (defined in section 4.4.3) that correspond to enforcing time reversal symmetry for the box orbits. This addition ensures that box orbits have the expected even parity in their line-of-sight velocity distributions (LOSVDs). In the cases that we have examined, these orbits already have small enough odd LOSVD components that this change makes very little difference.

4.4.2 Modify acceleration table for significant speedup

In order to speed up orbital integration, the orbit code pre-computes a lookup table of acceleration values over a spatial grid and performs a trilinear interpolation to closely approximate the true acceleration. If an orbit passes outside the radial range of this grid, the acceleration is then computed from scratch, which is multiple orders of magnitude slower than interpolating values from the lookup table. It is therefore prudent to choose the extent of the grid wisely because even a small number of orbits passing outside the table's coverage can dominate the total runtime and unnecessarily increase the computation time of the entire orbit library.

We have noticed that some orbits can indeed pass outside the radial range used in the original code and result in a significant slow down. To eliminate this situation, we have made a simple modification to the radial range used for the acceleration table. In van den Bosch

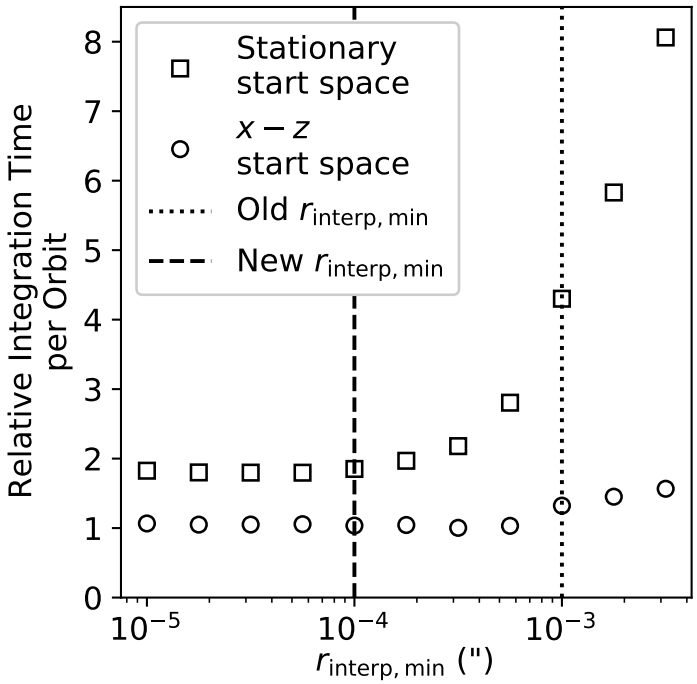

Figure 4.3: Average orbital integration time (per orbit) as a function of the inner interpolation radius, $r_{\rm interp,min}$, used to tabulate the accelerations. The stationary start space contains mostly box orbits that pass near the galaxy center. The box orbit integration time increases drastically with $r_{\rm interp,min}$, and the value used in the van den Bosch et al. [202] code is typically not small enough to minimize the integration time.

et al. [202], the acceleration is pre-computed over a grid spanning the radial range

$$\begin{aligned} r_{\rm interp,min} &= \min\left[0.1 \times \min\left(\sigma_i'\right), 0.01\, r_{\rm min}\right], \\ r_{\rm interp,max} &= \max\left[6 \times \max\left(\sigma_i\right), 1.05\, r_{\rm max}\right], \end{aligned} \tag{4.9}$$

where σ_i' is length of the semi-major axis of the ith projected MGE component, σ_i is the length of the semi-major axis of the corresponding intrinsic MGE component, and $r_{\rm min}$ and $r_{\rm max}$ are the innermost and outermost orbital equipotential radii in the model. Thus, the lowest and highest energy orbits included in the model have energies $\Phi(x = r_{\rm min}, y = 0, z = 0)$ and $\Phi(x = r_{\rm max}, y = 0, z = 0)$, where Φ is the gravitational potential of the model.

In practice, we find that the second conditions in Equation (4.9) typically determine the range of the acceleration table, i.e., $r_{\rm interp,min} = 0.01\, r_{\rm min}$ and $r_{\rm interp,max} = 1.05\, r_{\rm max}$. The outer boundary is never exceeded because energy conservation prevents orbits from passing outside $r_{\rm max}$ and therefore $r_{\rm interp,max}$. The inner boundary of $r_{\rm interp,min} = 0.01\, r_{\rm min}$,

however, can be problematic because centrophilic box orbits can pass well within $0.01\,r_{\rm min}$. The `DOP853` Runga-Kutta integrator in the TriOS code uses adaptive timesteps, tuning them to minimize errors in the position and velocity between timesteps. In this scheme many acceleration evaluations are required in regions of the trajectory where the timestep is smaller, namely, when the trajectory passes closest to the central black hole where the orbits are most likely to reach below $0.01\,r_{\rm min}$. The fraction of acceleration evaluations within this boundary is somewhat model-dependent and may be higher when box orbits are launched from well within the SMBH's sphere of influence because the potential felt by those orbits is largely spherical and supportive of highly centrophilic box orbits. For a typical case of $r_{\rm min} = 0.1''$ and $r_{\rm interp,min} = 0.01\,r_{\rm min} = 0.001''$, we find as many as a sixth of the acceleration evaluations during the box orbit integrations to lie outside the lookup table. This minority of acceleration evaluations take up more than 50% of the total time when constructing the orbit library.

To enable a more efficient use of the acceleration table, we choose to decouple $r_{\rm interp,min}$ and $r_{\rm interp,max}$ from $r_{\rm min}$ and $r_{\rm max}$ which are used to determine the range of orbital energy sampling. When $r_{\rm interp,min}$ is allowed to be smaller than $0.01\,r_{min}$, we find the total time to integrate orbits can be reduced by a factor of a few, with a negligible change in accuracy. This speed-up is illustrated in Figure 4.3. As the acceleration table is extended to smaller radii, fewer orbits fall outside the radial coverage of the table, and the average integration time for box orbits drops significantly with decreasing $r_{\rm interp,min}$. For the example shown in Figure 4.3, choosing $r_{\rm interp,min} \sim 0.0001''$ would reduce the orbit integration time by a factor of $\gtrsim 2$ compared with the original setting of $r_{\rm interp,min} = 0.01\,r_{min} = 0.001''$. Since energy conservation prevents orbits from passing outside $r_{\rm max}$, setting $r_{\rm interp,max}$ to be slightly larger than $r_{\rm max}$ minimizes the integration time while maximizing the interpolation accuracy.

Since we don't typically vary the interpolation boundaries by more than 1 dex, the density of points in the interpolation grid does not change dramatically, and we find that the accuracy of the interpolated potential is sufficient. However, if the boundaries are changed more drastically, the number of radial interpolation points should be adjusted to maintain the desired accuracy.

4.4.3 Resolve issues with insufficient orbit sampling

The TriOS code samples orbit initial conditions from two separate spaces, referred to as start spaces [174, 202]. In the first start space ("stationary start space"), all orbits start from rest on the equipotential surface for a given energy. This start space contains only box, box-like, and chaotic orbits.

The second start space ("$x-z$ start space") contains mainly tube orbits and samples orbits in the $x-z$ plane, with velocity vectors pointing along the y-axis. As illustrated in Figure 4.4, orbits of a given energy in this space are sampled over the region bounded by the equipotential and thin-orbit curves. Typically, $N_{\rm I_2} = 9$ rays of orbits are sampled uniformly in polar angles from 0 to $\pi/2$ in the positive x and z quadrant; along each ray, $N_{\rm I_3} = 9$ orbits are uniformly spaced between thin-orbit curves and equipotential curve. Additionally,

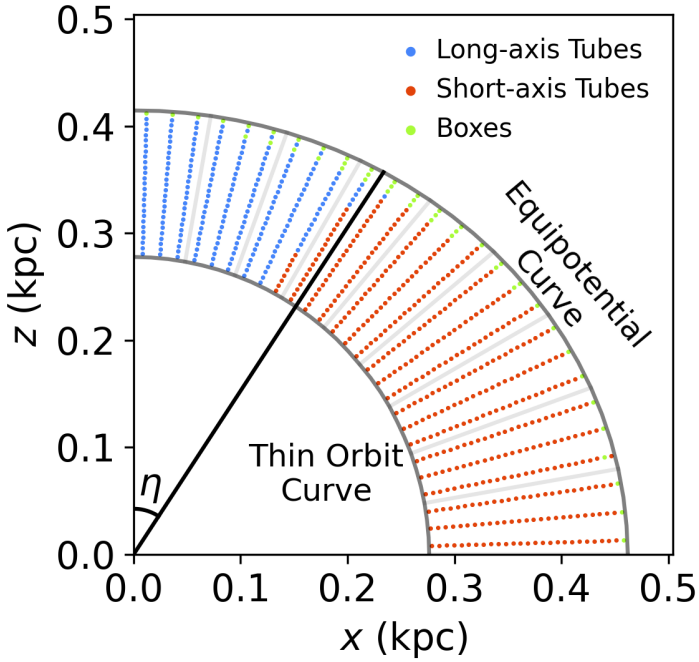

Figure 4.4: An example of the initial orbit locations in the $x-z$ start space for a single energy value in the triaxial TriOS code. Orbits are launched from within the thin-orbit curve (inner grey arc) and equipotential curve (outer grey arc). The orbit initial conditions are sampled with $N_{\mathrm{I}_2} = 9$ radial rays uniformly spaced in the polar angle from the z-axis to the x-axis, $N_{\mathrm{I}_3} = 9$ points along each ray, and $N_{dither} = 3$ to further improve the sampling, resulting in a total of 27×27 orbits. Each of the 27×27 color dots indicates the initial locations of an orbit (color coded by the type of orbits). The black line at angle η (see text) approximates the boundary between long-axis and short-axis tube orbits within this start space. Model χ^2 values are sensitive to the alignment between the angle η and orbit cell boundaries.

the code allows for dithering, where orbits with N_{dither} adjacent initial conditions in each dimension are integrated and then bundled together to form each of the 9×9 orbits in order to improve phase space sampling. Figure 4.4 illustrates the case of $(N_{\mathrm{I}_2}, N_{\mathrm{I}_3}, N_{dither}) = (9, 9, 3)$, where 27×27 tube orbits are launched in the positive quadrant of the $x - z$ start space for a given energy.

For a triaxial model, the short-axis tubes (red points in Figure 4.4) and long-axis tubes (blue points) occupy two regions of the $x - z$ start space separated by the focal curve. As derived in Appendix 3.A (Appendix A of [156]), the focal curve is roughly approximated by a line at angle

$$\eta = \tan^{-1}\sqrt{\frac{T}{1-T}}\,. \tag{4.10}$$

Thus, as T increases from 0 to 1, the focal curve moves smoothly from the z-axis to the x-axis, and the composition of the tube orbits changes from being all short-axis tubes (for an oblate axisymmetric potential) to all long-axis tubes (for a prolate axisymmetric potential).

When orbits are well sampled, model properties such as the goodness-of-fit (χ^2) should vary smoothly as η (and hence T) is varied. In our test runs for NGC 1453, however, we find that on top of a smooth variation, χ^2 varies periodically with T with a frequency matching the spacing between dithered orbits, $(\pi/2)/N_{\mathrm{I}_2}$, resulting in multiple spurious local minima at different values of T. Further testing reveals that these local minima arise from insufficient orbit sampling: as T increases, the focal curve approximated with η crosses rays of orbits in a periodic manner, resulting in the artificial oscillations in χ^2 with that same period. Since the periodic behavior is coherent as other model parameters are changed, it can have a significant impact on the recovered value of T and its uncertainty. Other parameter values are mainly impacted through their correlations with T.

We are able to eliminate the spurious oscillations in χ^2 vs. T by increasing N_{I_2}, which increases the number of radial rays in the $x - z$ start space and therefore improves the sampling in the polar angle. For the models presented in Section 4.5, we find that increasing N_{I_2} from the default value of 9 to 15 and beyond removes the oscillations and also yields convergent results. We choose $N_{I_2} = 18$ for the $x - z$ start space.

We do not find similar issues for the other start space. Nonetheless, we increase N_{I_2} to 18 for the stationary start space as well so as to maintain equal sizes for the tube and box orbit libraries. In summary, we use $(N_E, N_{I_2}, N_{I_3}, N_{\mathrm{dither}}) = (40, 18, 9, 3)$ for both start spaces. This results in a total $40 \times 18 \times 9 \times 3^3 \times 3 = 524,880$ integrated orbits in each galaxy model, where the last factor of 3 accounts for the 3 orbit libraries (the $x - z$ start space, its time-reversed copy, and the stationary start space).

4.4.4 Improve intrinsic mass binning scheme

In addition to kinematic constraints, the TriOS code enforces self-consistency of the mass model by requiring that the orbital weights be chosen to reproduce an input mass distribution (e.g., deprojected surface brightness profile of a galaxy). This is done by binning the mass

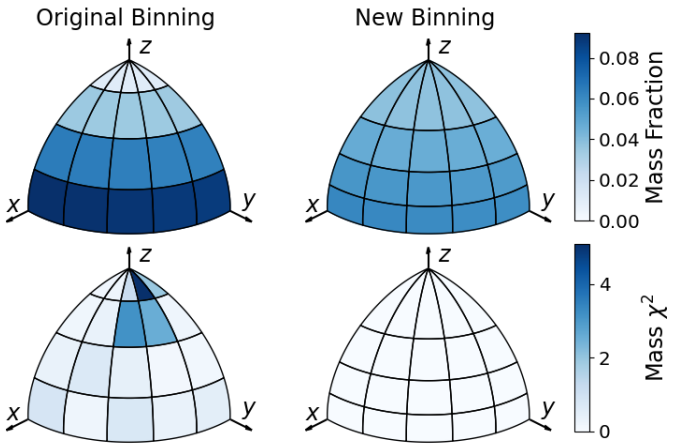

Figure 4.5: Comparison of the original (left) and new (right) mass binning scheme in the TriOS code. The top row shows that the bins near the $x-y$ plane contain far more mass than the bins near the z axis due to the significant difference in bin volume in the original scheme (top left). Our new binning scheme evens out the mass considerably (top right). The color scale here indicates the fraction of mass that falls within a given angular bin, summed over radius. The bottom row shows an example of the resulting χ^2 in the mass fits for a triaxial galaxy for the two binning schemes. The color scale here indicates χ^2 from attempting to fit a particular mass model, summed over radius. Only the 3D mass distribution is fit, with an error of 1% assumed on each bin. The most significant contributions to the mass χ^2 are from bins near the z-axis that contain very little mass. The triaxial mass model shown here has $M_{\rm BH} = 2.9 \times 10^9 M_\odot$, $M^*/L_{\rm F110W} = 2.0$, $T = 0.10$, $q = 0.96q'$, $T_{\rm maj} = 0.95$, and $T_{\rm min} = 0.12$.

in spherical coordinates (r, θ, ϕ), and requiring that the mass in each bin be reproduced to within a pre-specified precision (typically 1%). van den Bosch et al. [202] uses linearly spaced bins between 0° and 90° for θ and ϕ, and logarithmically spaced bins between $r_{\rm min}$ and $r_{\rm max}/2$ for r, where $r_{\rm min}$ and $r_{\rm max}$ are the innermost and outermost equipotential radii discussed in Section 4.4.2.

In the axisymmetrized TriOS code (Chapter 3; [156]), we changed the radial binning scheme above to ensure sufficient orbits are used to represent the innermost and outermost mass bins. During our subsequent tests for triaxial systems, however, we noticed occasional problems with mass misfits in which a handful mass bins would have difficulty satisfying the 1% precision and/or contribute disproportionately high values to the total χ^2 of the galaxy model under examination. We are able to trace the problem to uneven bin sizes in θ used in the original code: the bins near the poles contained much less mass, as shown in the left panel of Figure 4.5. Because of this, the mass near the z axis was subject to much more stringent constraints than elsewhere, leading to frequent difficulties in satisfying the 1% fitting criterion. Even in the absence of kinematic constraints, spurious variations would

arise in the χ^2 landscape, as illustrated in the right panel of Figure 4.5. The more oblate ($T \lesssim 0.1$) and round ($q \gtrsim 0.9q'$) systems are more prone to this issue.

We find that this mass misfitting problem can be easily resolved by using mass bins linearly spaced in $\cos(\theta)$ and ϕ, rather than in θ and ϕ. The resulting bins at a given radius then occupy the same volume, and the mass in each bin is much more uniform, with the bin-to-bin variations representing the galaxy's intrinsic deviation from spherical symmetry. Correspondingly, the pre-specified mass constraint criterion is enforced more uniformly.

For clarity, we have chosen to illustrate the mass misfitting issue in Figure 4.5 without imposing any kinematic constraints. When kinematic constraints are added in full orbit modeling (see Section 4.5), the total χ^2 returned by the code includes contributions from fits to the masses as well as kinematics. In this case, models with significant mass misfits due to uneven binning schemes would have disproportionately larger χ^2 values, leading to potential biases in the recovered galaxy parameters.

4.5 Triaxial Orbit Models of NGC 1453

4.5.1 NGC 1453

We apply the updated TriOS code described in the previous section to NGC 1453, a massive elliptical galaxy targeted by the MASSIVE survey [126]. In Chapter 2 [119], we performed orbit modeling of NGC 1453 using the axisymmetrized TriOS code. A detailed description of the input kinematic and photometric data is presented in that chapter. In brief, the stellar kinematics are measured over 135 spatial bins from our high-spatial resolution Gemini GMOS IFS data [53, 55] and wide-field McDonald Mitchell IFS data [209, 210, 211]. The first eight Gauss-Hermite moments are measured from the IFS spectra and used to constrain the stellar LOSVD in each kinematic bin; see Figure 2.4 (Figure 4 of [119]).

The MGE components representing the galaxy's mass distribution (see Section 4.2.3) are obtained from deprojections of our *HST* WFC3 photometry [83]. Here we use the same input data but relax the assumption of axisymmetry in the orbit models. In order to ensure that all trajectories within the model are representative of their equilibrium distributions, we integrate each orbit in the $x - z$ start space for 2000 times the orbital period for a thin tube orbit of the same energy. For orbits in the stationary start space, we integrate for 200 times the orbital period, as is typical of previous studies using the van den Bosch et al. [202] code.

Due to the regular isophotes of NGC 1453 (Figure 2.5; Figure 5 of [119]), we use the same PA for all MGE components and do not model isophotal twists. This is a common simplifying assumption [e.g., 199, 219, 65] and it enables us to explore the galaxy's shape using the new scheme outlined in Section 4.3.

For the distance to NGC 1453, we adopt our new determination of 51.0 Mpc from the MASSIVE-WFC3 project [83] using the surface-brightness fluctuation technique [101]. At

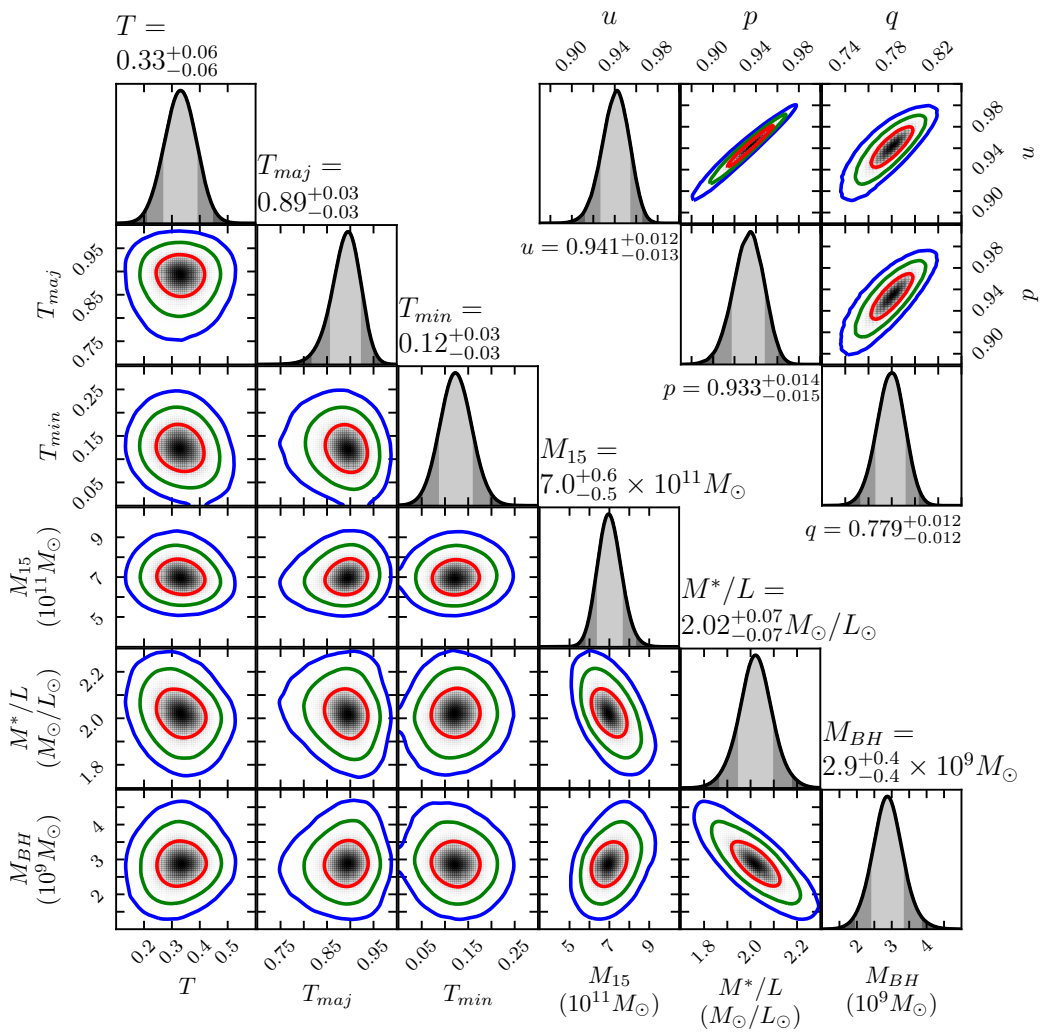

Figure 4.6: (Left) 5D likelihood landscape for orbit models of NGC 1453. As described in the text, the models are sampled in T, $\sqrt{T_{\rm maj}}$, $\sqrt{T_{\rm min}}$, $M_{\rm BH}$, $M^*/L_{\rm F110W}$, and M_{15}, and the 1D and 2D likelihood landscapes are obtained by marginalizing over a smoothed 5D landscape generated by Gaussian process regression. The red, green, and blue curves represent the 1σ, 2σ, and 3σ contours, respectively. (Right) 3D likelihood in axis ratio space, (p, q, u), marginalized over $M_{\rm BH}$, $M^*/L_{\rm F110W}$, and M_{15}. All three axis ratios are significantly correlated with one another, in particular between p and u. This degeneracy is significantly reduced when our new shape parameters $T, T_{\rm maj}$, and $T_{\rm min}$ are used.

this distance, 1″ is 245 pc for a flat ΛCDM model with a matter density of $\Omega_m = 0.315$ and a Hubble parameter of $H_0 = 70$ km s^{-1} Mpc^{-1}.

4.5.2 Parameter Search Using Latin Hypercube Sampling

We conduct the search for the best-fit galaxy shape in the new triaxial parameters $(T, T_{\rm maj}, T_{\rm min})$ introduced in Section 4.3. The dark matter halo is modelled as a logarithmic potential. We parameterize it through its mass within 15 kpc, M_{15}, which is roughly the central radius of the outermost kinematic bins, following Chapter 2 [119]. As in Chapter 2, we fix the scale radius of the dark matter halo to 15 kpc. Combining the three shape parameters with the three mass parameters $M_{\rm BH}$, $M^*/L_{\rm F110W}$, and M_{15}, we sample the 6D parameter space of galaxy models.

We determine the best-fit parameters by minimizing a χ^2 that includes terms for each LOSVD moment within each aperture, the projected light within each aperture, as well as the binned 3D mass density in order to enforce self-consistency for the stellar density. For each model, the best-fit set of weights are used to calculate the χ^2 differences between models. Lipka and Thomas [121] recently suggested that recovery of the inclination of axisymmetric models can be biased unless the intrinsic flexibility of the models is accounted for. However, a triaxial exploration of model flexibility is beyond the scope of the present chapter.

Instead of conducting model searches on a regular grid as was done in previous studies, we use the more efficient method of Latin hypercube sampling [137]. There are many techniques for ensuring spatial uniformity in multidimensional spaces. We adopt the scheme described in Deutsch and Deutsch [43], as implemented in the `LHSMDU` python package [143]. This procedure results in models that span a more continuous range of values than a regular grid, and are more uniformly spaced than random sampling. This approach allows a more representative sampling of the 6 dimensions with many fewer points than a regular grid.

We initially use a hypercube consisting of 1000 models spanning the range of $M^*/L_{\rm F110W} \in [1.7, 2.3]$, $M_{15} \in [3.5, 10.5] \times 10^{11} M_\odot$, and $M_{\rm BH} \in [1, 5] \times 10^9 M_\odot$, and the full range between 0 and 1 for $(T, \sqrt{T_{\rm maj}}, \sqrt{T_{\rm min}})$. Of these models, 927 resulted in valid deprojections. We then use a rejection-based scheme to choose subsequent sets of model points. A Gaussian process interpolation of the 6-dimensional χ^2 surface is computed from the previously-run models. We use this interpolation to estimate the χ^2 for $O(10^4)$ points chosen using the LHS scheme described above in the original volume and select points where the estimated χ^2 is within $\Delta\chi^2 = 20.06$ (3σ for 6 parameters) of the estimated global minimum. To avoid premature optimization we perform this routine 10 times where random subsets of half of all previously-run models are used to build the interpolation function. With this scheme we select roughly 1000 model points which are expected to lie near the global χ^2 minimum to evaluate with the TriOS code. We perform two iterations of this rejection scheme, yielding roughly 3000 total model evaluations.

The resulting 6D likelihood landscape is shown in Figure 4.6. To determine the best-fit value and uncertainties, we fit the χ^2 landscape using Gaussian process regression with a squared-exponential covariance function [150]. To make the 2D contours shown in Fig-

ure 4.6, we transform this smoothed surface from $(T, \sqrt{T_{\rm maj}}, \sqrt{T_{\rm min}})$ to $(T, T_{\rm maj}, T_{\rm min})$, or (p, q, u). The marginalized 1D likelihood is also shown for each parameter. The shapes of the 2D contours in Figure 4.6 clearly demonstrate that $(T, T_{\rm maj}, T_{\rm min})$ do not have the strong degeneracy apparent in (p, q, u).

The standard values of $\Delta\chi^2 = 1, 4, 9$ are used to define the 1σ, 2σ, and 3σ confidence intervals for 1 degree of freedom when considering the marginalized landscape for each variable individually. For the 2D contours, we use the values for 2 degrees of freedom, giving $\Delta\chi^2 \approx 2.3, 6.2, 11.8$. This is different from most previous work using the van den Bosch et al. [202] code, where typically $\Delta\chi^2 = \sqrt{2N_{\rm obs}}$ is used to define the 1σ confidence interval, where $N_{\rm obs}$ is the number of apertures on the sky, multiplied by the number of moments fitted within each aperture. This value is chosen to represent the intrinsic noise in the χ^2 values for each model, and is much larger than our values. However, while this is true when the input data are varied according to its noise level as discussed in Vasiliev and Valluri [208], the noise level in the χ^2 values between models are significantly smaller when the input data are fixed.

4.5.3 Best-fit Triaxial Model

The best-fit values and the uncertainties for each NGC 1453 parameter are listed in Table 4.2. For each parameter, all other dimensions have been marginalized over. The best-fit $M_{\rm BH}$ is consistent with the value determined from axisymmetric modelling in Chapter 2 [119]. The value of $M^*/L_{\rm F110W}$ has shifted down slightly, but is still consistent within 2σ of the axisymmetric value. To provide the closest comparison between the axisymmetric and triaxial modelling we use the same configuration of kinematic observations as in Chapter 2 [119]. We use 8 moments to describe the observed LOSVDs while fitting the GMOS spectra and six moments to describe the Mitchell spectra. For all bins the remaining moments up to h_{12} are constrained to be nearly zero in our models to reduce spurious behavior in the LOSVDs. This procedure is described in Sec. 2.4.3.

The best-fit shape, on the other hand, is inconsistent with axisymmetry. It is useful to compare our best-fit values of $p = 0.93$ and $q = 0.78$ with those inferred statistically from the observed distributions of ellipticity and misalignment angle between the kinematic and photometric axes for 49 slowly-rotating massive elliptical galaxies with measurable kinematic axes in the MASSIVE survey [54]. In that sample, 56% of the galaxies have $p > 0.9$ with a mean value of 0.88, and the mean value of q is 0.65. Our best-fit shape for NGC 1453 indicates this fast-rotating galaxy is relatively oblate like the MASSIVE slow rotators and is slightly less flattened than the mean of that population.

The orbital composition of the best-fit triaxial model is shown in Figure 4.7 (top panel). Long-axis tubes and box orbits – two orbit types that are present only in triaxial potentials – together account for $\sim 30\%$ of the orbital weights in the inner part and $\sim 45\%$ in the outer part of NGC 1453. Quasi-planar orbits account for a small fraction of the total mass at small and large radii and are excluded from the plot. While long-axis tubes contribute

Parameter	Value
$M_{\rm BH}$ $(10^9 M_\odot)$	2.9 ± 0.4
$M^*/L_{\rm F110W}$ $(M_\odot/L_\odot)$	2.02 ± 0.07
M_{15} $(10^{11} M_\odot)$	$7.0^{+0.6}_{-0.5}$
T	0.33 ± 0.06
$T_{\rm maj}$	0.89 ± 0.03
$T_{\rm min}$	0.12 ± 0.03
u	$0.941^{+0.012}_{-0.013}$
p	$0.933^{+0.014}_{-0.015}$
q	0.779 ± 0.012
θ (°)	73 ± 3
ϕ (°)	19 ± 3
ψ (°)	$92.7^{+0.7}_{-0.8}$

Table 4.2: Best-fit triaxial model parameters for NGC 1453 from the 6D likelihood landscape in Figure 4.6. For each parameter, all other dimensions have been marginalized over.

a significant fraction of the mass, the projected model has fairly little minor axis rotation, due in part to the LOS being close to the intrinsic major axis.

The orbital velocity anisotropy of the best-fit model (bottom panel of Figure 4.7) is mildly tangential ($\beta < 0$) in the inner part and becomes increasingly radial outward. The radial profile has a similar shape to the axisymmetric model presented in Chapter 2 [119].

4.5.4 Triaxial vs. Axisymmetric Best-fit Models

The best-fit triaxial model presented above matches the observed kinematics significantly better than the best-fit axisymmetric model in Chapter 2 [119].

Even though the best-fit χ^2 values in the two cases – 493.0 for axisymmetric versus 382.7 for triaxial – differ by ~ 110, they should not be compared directly because triaxial potentials require a new library of box orbits, and different numbers of orbits are used (6480 independent weights for axisymmetric versus 19440 for triaxial). Nonetheless, within triaxial modeling, our best-fit triaxiality of $T = 0.33$ is preferred over nearly oblate axisymmetric models with $T \approx 0$ at a confidence level of about 5σ. To understand why non-axisymmetric models are favored, we examine the 2D maps of V and the lowest 3 even Gauss-Hermite moments in the GMOS data in Figure 4.8 (first row). We recall that axisymmetric models by construction produce only bisymmetric kinematics about the photometric major axis on the sky, meaning that the LOSVDs would be symmetric for points mirrored across the projected major axis and anti-symmetric for points mirrored across the projected minor axis. Any observed systemic deviation from bisymmetry would then indicate triaxiality.

For this reason, we decompose each GMOS moment map into a bisymmetrized component

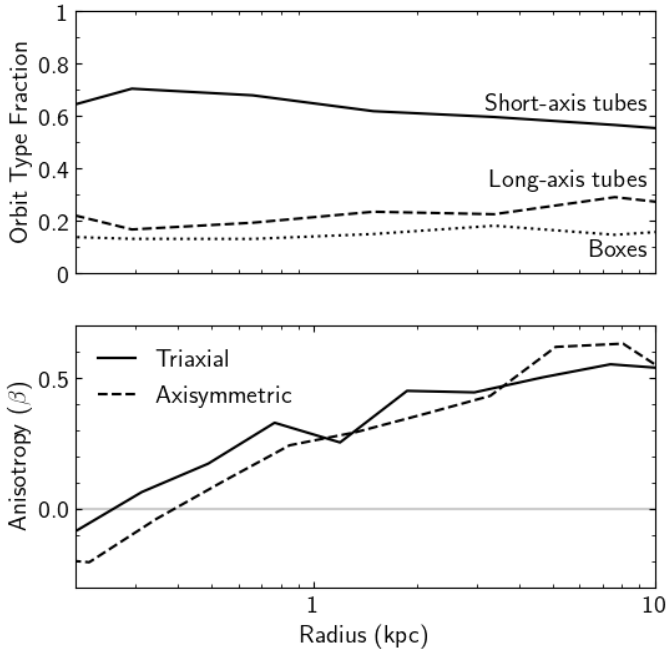

Figure 4.7: Orbital composition (top) and velocity anisotropy (bottom) of the best-fit triaxial model of NGC 1453 as a function of radius. Short-axis tubes (solid) are dominant throughout the model, with significant contributions from long-axis tubes (dashed) and box orbits (dotted) that are present only in triaxial potentials. The velocity anisotropy parameter, β, has a similar radial profile for the best-fit triaxial (solid) and axisymmetric (dashed) models, being mildly tangentially anisotropic in the inner part and becoming more radially anisotropic in the outer part.

(second column) and a non-bisymmetrized component (third column). The latter exhibits clear systemic deviations from bisymmetry. The most obvious feature is the residual minor axis rotation indicative of kinematic misalignment. These maps assume a bisymmetrization along the projected photometric major axis used by our dynamical models, with a PA of 28.5°. The residual pattern persists and can not be "rotated away" even if the PA is within uncertainties in the PA determination determined from the isophotal profile from Goullaud et al. [83]. An axisymmetric model (consistent with the photometry) would be incapable of fitting these non-bisymmetric features in the data. To confirm this point, we plot the residual maps (fourth column) between the GMOS data and the best-fit axisymmetric model from Chapter 2 [119]. Indeed, the axisymmetric model exhibits similar residual patterns as in the data (third column). In comparison, the best-fit triaxial model is able to fit these non-bisymmetric features to a large extent, producing essentially random residuals (fifth column).

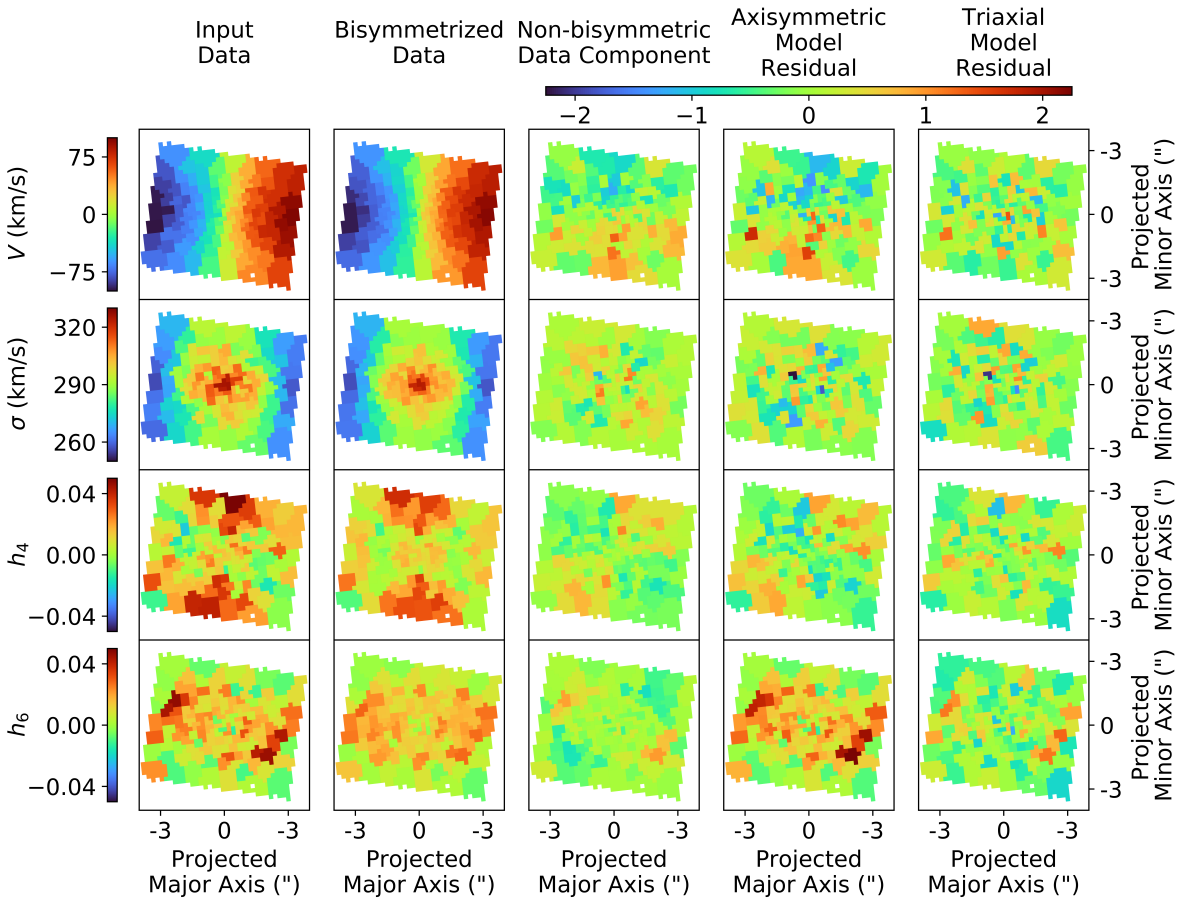

Figure 4.8: Maps of the stellar kinematics from the Gemini GMOS IFS in 135 spatial bins of the central $5'' \times 7''$ of NGC 1453. Four velocity moments are shown (from top down): V, σ, h_4 and h_6. The maps are oriented such that the horizontal and vertical axes are aligned with the galaxy's projected major and minor photometric axes, respectively. The data (first column) are decomposed into a bisymmetric component (second column) and a non-bisymmetric component (third column). To accentuate systematic patterns, we plot the non-bisymmetric component normalized by the moment uncertainty. Since an axisymmetric model can only produce bisymmetric kinematic maps, the residuals from the best-fit axisymmetric model (fourth column) show similar patterns to the bisymmetrized residuals. h_6 shows additional residuals that are consistent with bisymmetry, but unable to be fit by an axisymmetric model. A triaxial model (right column) is able to capture most of the systematic behaviour in the input map, resulting in largely random residuals. The residuals have been normalized by the moment uncertainty.

Figure 4.8 indicates that the preference for triaxiality is driven by the non-axisymmetric features in the NGC 1453 kinematics. Even though the non-bisymmetric features are somewhat subtle, they lead to detectable triaxiality, which we find to be best fit with $p = 0.933$, $q = 0.779$, and $T = 0.33$. Thus, despite being a fast rotator with regular isophotal and kinematic features, NGC 1453 is best fit by a triaxial model. This is further evidence for widespread triaxiality in massive elliptical galaxies.

Importantly, however, the best-fit black hole mass $M_{\rm BH} = 2.9 \times 10^9 M_\odot$ is unchanged from that in the axisymmetric model. The stellar mass-to-light ratio and dark matter mass within 15 kpc agree to within a 1σ confidence level.

4.6 Conclusions

In this paper we have presented a revised code and a revamped approach for performing dynamical modeling of triaxial galaxies and their central SMBHs using the orbit superposition technique. We discussed a new triaxial version of the TriOS code that is capable of modeling triaxial systems while avoiding several shortcomings of the original van den Bosch et al. [202] code. As a first application of this code, we performed triaxial orbit modeling of the massive elliptical galaxy NGC 1453 and presented the best-fit galaxy shape and mass parameters. This work complements Chapters 2 and 3 [119, 156], in which we introduced a properly axisymmetrized version of the TriOS code.

We discovered and corrected a major error in the orbit kinematics in the van den Bosch et al. [202] code: the tube orbits had wrong signs in certain mirrored velocity components in the orbit library (Table 1), resulting in incorrect projected kinematics. The magnitude of the kinematic errors varies spatially and depends on the viewing angles (Figure 4.2). This issue impacts all triaxial models that are not viewed along a principal axis, and all nearly axisymmetric models that are not viewed edge-on. How this error affects the best-fit galaxy shapes and mass parameters would have to be assessed on a galaxy-by-galaxy basis by rerunning the models with the corrected orbital flips in Table 4.1. In the case of NGC 1453, we find the χ^2 landscapes to be altered drastically, with χ^2 values changing non-uniformly by more than 100 for some models.

Following Chapter 3 [156], we continued to find ways to speed up the code. In this updated version of the TriOS code, we achieved another significant speedup (of up to ~ 50 %; Figure 4.3) in orbit integration time by a simple extension of the interpolation table used to evaluate orbit accelerations (Section 4.3). The reduction in integration time is particularly pronounced for centrophilic orbits.

We have made two other adjustments in the code that significantly improve the sampling of long-axis tube orbits (Section 4.4.3) and enforce more uniformly the 3D mass constraints (Section 4.4.4). After these changes, the behavior of χ^2 vs. T (triaxiality parameter) no longer exhibits spurious oscillations, and the orbit code is able to find reasonable solutions for some mass models that were previously strongly disfavored.

The rest of this paper is devoted to new and improved strategies for searching the multi-dimensional parameter space required to specify triaxial galaxy models. We introduced a new set of shape parameters (Section 3) as well as a novel sampling technique (Section 4.4.3), which together lead to a remarkable gain in parameter searching efficiency. Searching in the new parameters T, $T_{\rm maj}$, and $T_{\rm min}$ (Equation 4.3) avoids significant non-uniformities associated with other parameters used in earlier work. Our Latin hypercube sampling scheme results in an order-of-magnitude reduction in needed sampling points compared with conventional grid searches.

We applied the TriOS code and triaxial sampling scheme to the fast-rotating massive elliptical galaxy NGC 1453 in the MASSIVE survey (Section 4.5). NGC 1453 has a relatively small twist in the isophotes, and the kinematic and photometric axes are nearly aligned. Despite these properties that are typically invoked to justify the use of axisymmetric orbit codes, we find the best-fit model to have a triaxiality value of $T = 0.33$, with intrinsic axis ratios $p = 0.933$ and $q = 0.779$. This best-fit triaxial model is able to match the observed kinematic maps significantly better than the best-fit axisymmetric model in Chapter 2 [119]. The improvement is mainly due to the ability of triaxial models to account for non-bisymmetric features in the data (Figure 4.8). Most other galaxies in the MASSIVE survey exhibit less (or no) rotation and more twists in their photometric and kinematic maps compared to NGC 1453. This is further evidence that massive elliptical galaxies have triaxial intrinsic shapes.

$M_{\rm BH}$ in the best-fit triaxial model for NGC 1453 is unchanged from the value measured with the axisymmetrized TriOS code from Chapter 2 [119]. Among the many dozens of stellar dynamical $M_{\rm BH}$ measurements in local galaxies (e.g., [133]), NGC 1453 is only one of a handful galaxies whose central SMBH is studied with the full triaxial orbit modeling technique not limited to axisymmetry. In four other galaxies (Section 4.1), M32 had consistent $M_{\rm BH}$ from axisymmetric and triaxial modeling, the NGC 3379 $M_{\rm BH}$ increased by a factor of ~ 2 when axisymmetry was relaxed, the PGC 046832 $M_{\rm BH}$ decreased enough to be consistent with 0, while NGC 3998 was only modeled with the triaxial code so no comparison can be made. All four systems were modeled with the original van den Bosch et al. [202] code, which used the incorrect mirroring scheme. Triaxial orbit modeling of more galaxies is needed for a full assessment of the systematic effects on stellar dynamical $M_{\rm BH}$ measurements when the commonly-made assumption of axisymmetry is relaxed.

Appendix

4.A Relating new and old parameters

The expressions given in Equation (4.8) can be written in a simpler form when expressed sequentially:

$$\begin{aligned} \cos^2\theta &= T_{\rm min}(1 - TT_{\rm maj})\,, \\ \sin^2\phi &= \frac{(1-T_{\rm maj})(1-T_{\rm min})}{\sin^2\theta}\,, \\ \tan\psi &= \frac{-(1-T_{\rm min})\cos\theta}{(T_{\rm min}-\cos^2\theta)\tan\phi}\,. \end{aligned} \tag{4.11}$$

The inverse expressions are then

$$\begin{aligned} T &= \frac{\sin^2\theta}{\cos\theta\sin 2\phi\cot 2\psi + \cos^2\phi - \cos^2\theta\sin^2\phi}\,, \\ T_{\rm maj} &= 1 - \sin^2\phi(1 - \cos\theta\cot\phi\cot\psi)\,, \\ T_{\rm min} &= 1 - \sin^2\theta(1 - \cos\theta\cot\phi\cot\psi)^{-1}\,. \end{aligned} \tag{4.12}$$

The deprojection equations, giving the intrinsic shape in terms of the projected flattening and angles (θ, ϕ, ψ) are given by [24]:

$$\begin{aligned} 1 - q^2 &= \frac{\delta'[2\cos 2\psi + \sin 2\psi(\sec\theta\cot\phi - \cos\theta\tan\phi)]}{2\sin^2\theta[\delta'\cos\psi(\cos\psi + \cot\phi\sec\theta\sin\psi) - 1]} \\ p^2 - q^2 &= \frac{\delta'[2\cos 2\psi + \sin 2\psi(\cos\theta\cot\phi - \sec\theta\tan\phi)]}{2\sin^2\theta[\delta'\cos\psi(\cos\psi + \cot\phi\sec\theta\sin\psi) - 1]} \\ u^2 &= \frac{1}{q'}\sqrt{p^2\cos^2\theta + q^2\sin^2\theta(p^2\cos^2\phi + \sin^2\phi)}, \end{aligned} \tag{4.13}$$

where $\delta' = 1 - q'^2$. While Cappellari [24] presents these expressions in the context of the MGE formalism, they are more broadly applicable to all densities that are stratified on similar concentric ellipsoids. This is demonstrated in de Zeeuw and Franx [41]. The first two expressions in equation 4.13 are listed as their equation A8. The third expression giving the projection axis ratio, u, follows from expressions in this paper as well. Following appendix A of this paper, combining their equations 3.37, 3.38, and 3.49 gives:

$$\begin{aligned} a'^2 + b'^2 &= 2c^2 + (a^2 - c^2)(\sin^2\phi + \cos^2\phi\cos^2\theta) + (b^2 - c^2)(\cos^2\phi + \sin^2\phi\cos^2\theta) \\ (a'^2 - b'^2)^2 &= [(a - c^2)(\sin^2\phi - \cos^2\phi\cos^2\theta) + (b^2 - c^2)(\cos^2\phi - \sin^2\phi\cos^2\theta)]^2 \\ &\quad + 4(a^2 - b^2)^2\sin^2\phi\cos^2\phi\cos^2\theta. \end{aligned} \tag{4.14}$$

Here, (α, β, γ) in the original expressions have been set to $(-a^2, -b^2, -c^2)$ in order to consider a perfect ellipsoid. The first of these expressions is explicitly given in equation A6

of the original paper. Squaring the first expression and subtracting the second gives (after significant simplification):

$$4a'^2b'^2 = 4a^2b^2\cos^2\theta + 4a^2c^2\sin^2\phi\sin^2\theta + 4b^2c^2\cos^2\phi\sin^2\theta. \tag{4.15}$$

Substituting the definitions of the axis ratios reduces this expression to the third line of equation 4.13 above.

Equation 4.8 follows from equations 3.39 and 3.42 of [41], together with the definitions given in equation 4.4. Equation 4.7 then follows from equation 4.13, together with equation 4.8. As in appendix A1 of [41], while these expressions are derived in the context of a perfect ellipsoid, the results are independent of the assumed profile and are thus valid for all densities stratified on similar concentric ellipsoids.

Acknowledgements

We thank Shaunak Modak and Jonelle Walsh for useful discussions. M.E.Q. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), PGSD3-517040-2018. C.-P.M. acknowledges support from NSF AST-1817100, HST GO-15265, HST AR-14573, the Heising-Simons Foundation, the Miller Institute for Basic Research in Science, and the Aspen Center for Physics, which is supported by NSF grant PHY-1607611. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) at the San Diego Supercomputing Center through allocation AST180041, which is supported by NSF grant ACI-1548562.

Chapter 5

Keck Integral-Field Spectroscopy of M87 Reveals an Intrinsically Triaxial Galaxy and a Revised Black Hole Mass

The history of supermassive black hole measurements and much of the development within the field can be seen very directly through the history of measurements of M87's central black hole M87*. Among the first murmurs of the existence of supermassive black holes from stellar dynamical arguments came from Sargent in 1978 [170] who used long slit spectroscopic observations, the Fourier quotient method to infer velocity distributions, and fitting with a simple Jeans model to infer a central mass of $M_{\rm BH} \sim 5.5 \times 10^9 M_\odot$ (after converting to the modern adopted distance). In 1990, Dressler and Richstone [46] used long slit spectroscopy with Fourier methods to infer the stellar kinematics, but now used spherically symmetric Schwarzschild models. A few years later, van der Marel [204] advanced to fitting the long-slit spectra in the pixel space, allowing for extraction of the first six Gauss-Hermite moments, then modelled that data with spherical Jeans models. In 1998, Magorrian et al used the same data but now applied axisymmetric Jeans models [129]. In 2009, Gebhardt and Thomas used significantly more sophisticated data from the SAURON integral field spectrograph and globular cluster kinematics and applied axisymmetric Schwarzschild models [75]. Two years later they followed up small-scale spectra enhanced by adaptive optics on NIFS and large scale kinematics from the VIRUS-P spectrograph and again modelled the data using axisymmetric Schwarzschild models [77].

In parallel, M87 has a history of gas dynamical measurements, including work by van der Marel [204], HST spectroscopy by Harms et al [90], Macchetto et al [127], and Walsh et al [221]. This literature has intriguingly consistently measured the central supermassive black hole mass to be much smaller than stellar dynamical measurements – below $4 \times 10^9 M_\odot$, vs $4–7 \times 10^9 M_\odot$. A further parallel literature has emerged through the work surrounding the Event Horizon Telescope, which imaged the radio emission surrounding the central SMBH

in 2019 [57].

This chapter extends the efforts of the previous chapters, which worked to build a practical and theoretical understanding of proper usage of the TriOS triaxial orbit superposition code and uses that code to model the three-dimensional kinematical structure and mass distribution of the galaxy. This development in modelling is paired with our first published observations using the Keck Cosmic Web Imager, a relatively new instrument on Keck II. This allowed us to construct a large, $250'' \times 300''$ mosaic field of view with both high spatial resolution and signal-to-noise. This map is among the most comprehensive views of the stellar kinematics of M87 to date.

This chapter was originally published as

Emily R. Liepold, Chung-Pei Ma, and Jonelle L. Walsh. "Keck Integral-field Spectroscopy of M87 Reveals an Intrinsically Triaxial Galaxy and a Revised Black Hole Mass". In: *Astrophys. J. Lett.* 945.2, L35 (Mar. 2023), p. L35. DOI: 10.3847/2041-8213/acbbcf. arXiv: 2302.07884 [astro-ph.GA]

Minor alterations to the text have been made to ensure fluency throughout this dissertation. Note also that a significant fraction of the text associated with this chapter is presented as Appendix 5.5, in line with how it was originally published.

5.1 Introduction

Some of the earliest dynamical evidence for the presence of a supermassive black hole (SMBH) came from M87 [170]. A bright asymmetric ring of radio emission around the M87 SMBH was imaged in 2019 [57]. The black hole mass ($M_{\rm BH}$) inferred from the ring features is consistent with the value determined from stellar dynamics based on axisymmetric orbit modeling [77], but it is nearly twice the mass inferred from dynamics of a gas disk around the hole [221].

M87 is classified as an elliptical galaxy based on the two-dimensional shape of the stellar light projected on the sky. However, its three-dimensional intrinsic shape has never been determined. A galaxy's intrinsic shape is a fundamental property that encodes the galaxy's past merger history and provides information about the mass ratios of the progenitor galaxies, the merger orbital parameters, gas fractions, and fraction of stars formed ex-situ. Whether a galaxy is intrinsically spherical, axisymmetric, or triaxial also impacts dynamical determinations of its SMBH mass and stellar mass, as well as any mass reconstructions based on the method of gravitational lensing.

Thus far, almost all information about galaxy intrinsic shapes has been inferred statistically by inverting distributions of observed galaxy properties [73, 222, 72, 54, 118]. Here we use the Keck Cosmic Web Imager (KCWI; [142]) on the 10 m Keck II telescope to obtain a spatially-resolved two-dimensional map of the stellar kinematics of M87 over a $250'' \times 300''$ field of view. The resulting kinematics span a radial range of $\sim 0\farcs6$–$150''$, corresponding to a physical range of 50 pc–12 kpc at a distance of 16.8 ± 0.7 Mpc to M87, the value adopted in [59] and in this work (an angular size of $1''$ corresponds to a physical length of 81.1 ± 3.3

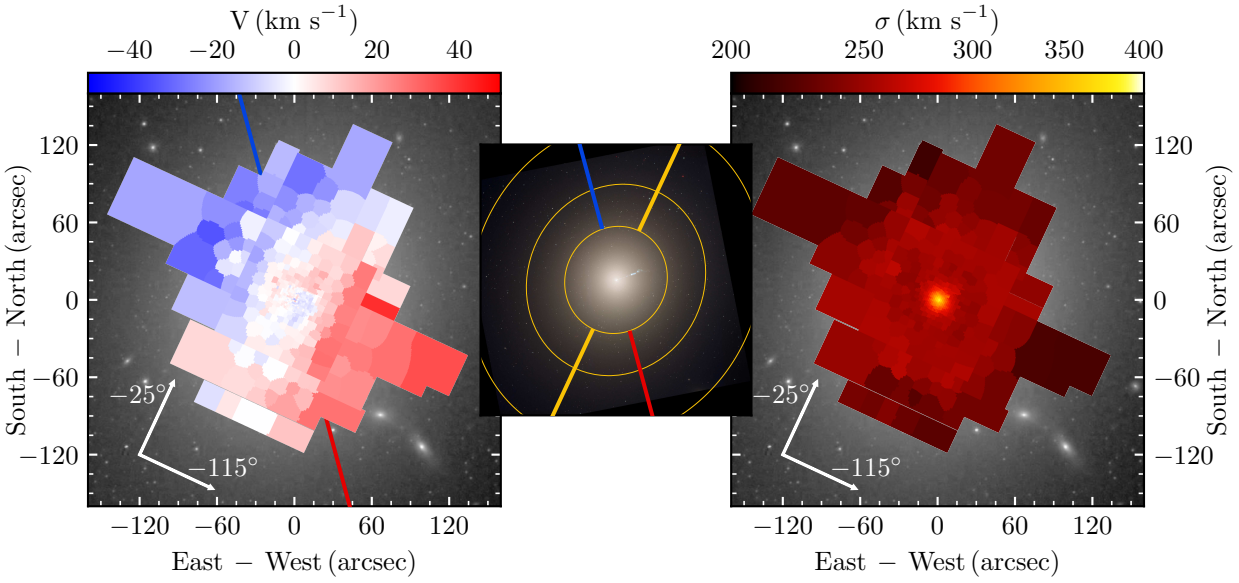

Figure 5.1: (Left and Right) Stellar kinematic portraits of M87 from Keck KCWI spectra in 461 spatial bins. The line-of-sight velocities (left) and velocity dispersions (right) of stars in M87 are shown over a 250″ by 300″ field of view centered at the galaxy's nucleus. The systemic velocity of M87 has been removed in the left panel. North is up and east is to the left. The two orthogonal white arrows indicate the orientations of the photometric major axis (PA of −25°) and minor axis (PA of −115°), as determined from the mean position angle of the galaxy's major axis between a radius of 50″ and 250″ in photometric data [109]. The red and blue lines in the left panel mark the measured kinematic axis (PA of −165°) outward of 70″ (see Figure 5.2). (Middle) HST composite image of the central 200″ by 200″ FOV of M87, illustrating the misalignment of the photometric major axis (yellow) and kinematic axis (red-blue) beyond 50″ along with sample isophotes of the stellar light distribution (yellow contours).

pc). We perform triaxial Schwarzschild orbit modeling using the detailed stellar kinematic measurements as constraints to determine M87's shape and mass parameters. Our models include a radially declining profile for the stellar mass-to-light ratio ($M^*/L_{\mathrm{F110}W}$) inferred from stellar population measurements [172].

5.2 Keck observations of M87

We observed M87 with Keck KCWI in May 2020, May 2021, March 2022, and April 2022. With the large slicer and BL grating of the integral-field unit (IFU), we obtained spectra between 3500 and 5600 Å at 62 pointings, which provide contiguous two-dimensional spatial coverage of the nucleus and the outer parts of M87 (Figure 5.1). The data span about 20 kpc (250″) across the photometric major axis (−25° east of north) and about 24 kpc (300″) across the photometric minor axis (−115° east of north).

We co-add spectra from individual KCWI spaxels to reach high signal-to-noise ratios (S/Ns), forming 461 spatial bins. Within each spatial aperture, we measure the line-of-sight stellar velocity distributions (LOSVDs) from the shapes of the absorption lines. Further details about the observations, data reduction procedures, spectral fitting processes, and stellar kinematic determination are provided in Appendices 5.A and 5.B.

5.3 Stellar Kinematic Maps

5.3.1 Misalignment between kinematic and photometric axes

The KCWI map for the line-of-sight velocity V (left panel of Figure 5.1) shows a prominent rotational pattern at large radii, in which the northeast side of the galaxy is blueshifted and the southwest side is redshifted. The kinematic axis that connects the maximal receding and approaching velocities, however, is not aligned with the photometric major axis, as it would be for an axisymmetric rotating galaxy.

To quantify the amplitude and axis of rotation, we model the velocity field as a cosine function, with $V(R, \Theta) = V_1(R) \cos[\Theta - \Theta_0(R)]$, where R is the projected radius from the galaxy's center and Θ is the azimuthal angle on the sky. The model parameters $V_1(R)$ and $\Theta_0(R)$ are the the amplitude of rotation and the position angle (PA) of the kinematic axis at radius R, respectively. With increasing radius, the velocity curve shows a systematic shift in phase and an increase in rotational amplitude (Figure 5.2). Within a radius of 3 kpc, the PA of the kinematic axis changes rapidly clockwise with radius (lower right panel of Figure 5.2), representing the kinematically distinct core mapped out by the Multi Unit Spectroscopic Explorer (MUSE) on the Very Large Telescope [50]. Beyond 3 kpc, where the MUSE data end (at about 35″), we find that the PA of the kinematic axis continues to change clockwise and crosses the PA of the photometric minor axis, plateauing at −165° between 6 and 12 kpc. Hence, there is a 40° misalignment between the stellar kinematic axis and photometric major axis in M87.

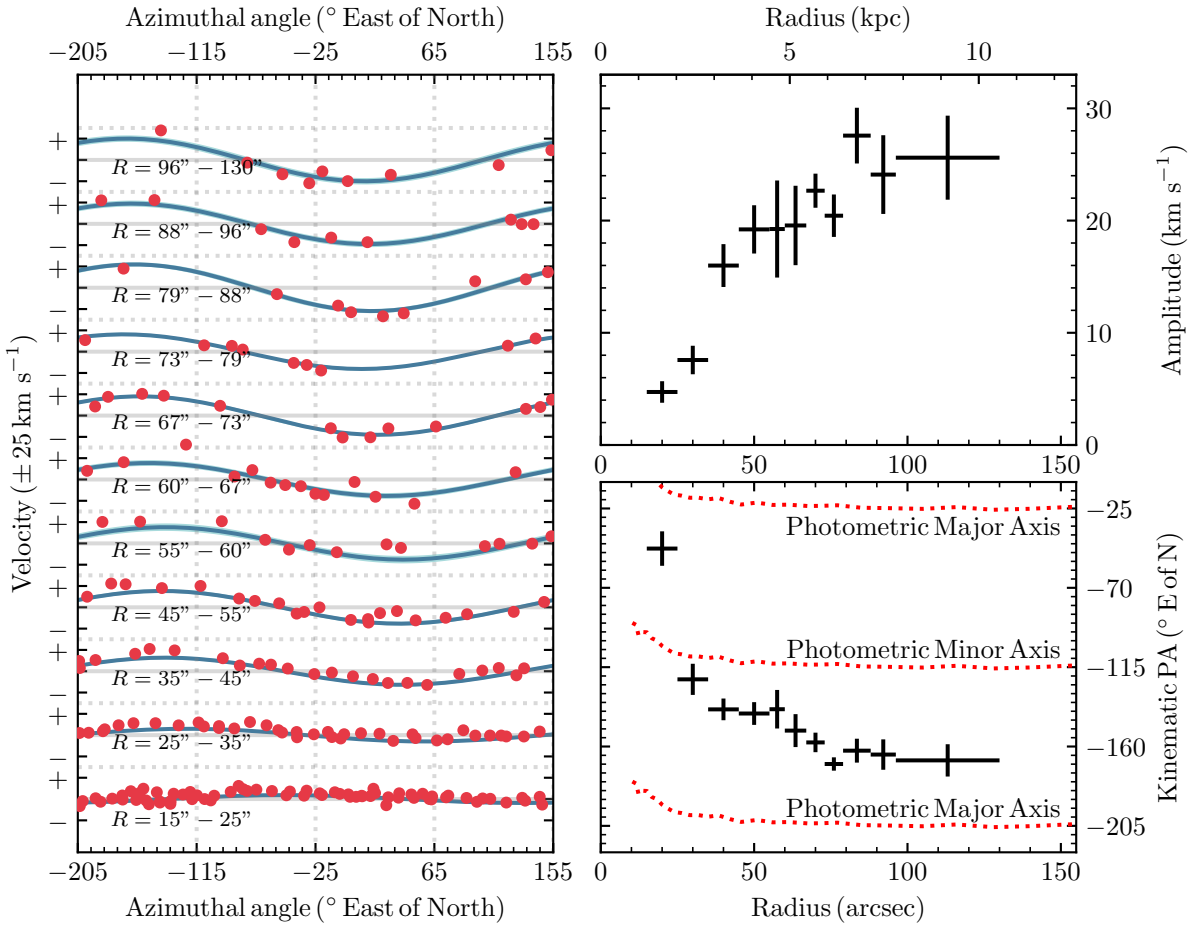

Figure 5.2: Line-of-sight stellar velocities as a function of projected radius and azimuthal angle on the sky. (Left) Line-of-sight velocity as a function of azimuthal angle on the sky for 11 radial shells spanning $R = 15''$–$130''$. The velocities in each shell (red data points) are well fit (blue) by a sinusoidal function of the form $V(R, \Theta) = V_1(R) \cos[\Theta - \Theta_0(R)]$. (Upper right) The amplitude of rotation, $V_1(R)$, increases with radius and reaches 25 km s^{-1} around 6 kpc. (Lower right) The phase of the velocity function, Θ_0, measures the orientation of the kinematic axis and varies significantly with radius. It plateaus to $-165°$ beyond 6 kpc, indicating a 40° misalignment between the kinematic axis and the photometric major axis (red dashed curves; [109]) in M87.

5.3.2 Stellar velocity dispersion

The KCWI map (right panel of Figure 5.1) and radial profile (Figure 5.3) of the stellar velocity dispersion σ exhibit several features. Towards the center of M87, σ increases rapidly from 250 km s^{-1} at a radius of 2 kpc to 370 km s^{-1} at 100 pc from the nucleus. This is a clear signature of the gravitational influence of the central black hole on the motions of the stars in its vicinity. The velocity dispersion stays at about 250 km s^{-1} between 2 and 5 kpc and then shows a gentle 10% decline between 5 kpc and the outermost reach of our data at 12 kpc. The stellar σ at the edge of our field connects smoothly to the latest determinations of the velocity dispersions of discrete dynamical tracers (lower panel of Figure 5.3) such as red globular clusters and planetary nebulae in the outer parts of M87 [228, 122]. Beyond about 10 kpc, sub-populations of planetary nebulae have been reported to have distinct kinematics [122]: σ of "intra-cluster" planetary nebulae rises to 800 km s^{-1} at 100 kpc, whereas those in the galaxy halo component have a relatively flat σ profile out to 100 kpc, similar to that of the red population of globular clusters [31, 186, 228].

5.4 Determination of mass and shape parameters from triaxial Schwarzschild modeling

We use the full LOSVDs from Keck KCWI, along with photometric observations from the Hubble Space Telescope (HST) and ground-based telescopes [109], to measure M87's mass distribution and intrinsic shape. We perform triaxial Schwarzschild orbit modeling with the TriOS code [156, 157] based on an earlier code [202], and use more than 4000 observational constraints to simultaneously determine six parameters: $M_{\rm BH}$, $M^*/L_{\rm F110W}$, dark matter content, and the three-dimensional intrinsic shape. As described below, we implement a new capability in the code to model spatial variations in $M^*/L_{\rm F110W}$ and use a radially declining $M^*/L_{\rm F110W}$ profile that closely approximates the variation inferred from stellar population and dynamics studies of M87 [148, 172].

5.4.1 Galaxy model and orbit sampling

Each galaxy model has three mass components: a central SMBH, stars, and a dark matter halo. The three-dimensional stellar density in the TriOS code is represented as a sum of multiple Gaussian functions of differing widths and axial ratios. To determine these functions, we first fit a two-dimensional Multi-Gaussian Expansion (MGE; [24]) to the surface brightness distribution of M87 (see Appendix 5.C). Each MGE component is allowed an independent flattening parameter (q' in Table 5.2) to model any radially changing ellipticity observed on the sky.

For a given set of three angles, θ, ϕ, and ψ, that relate the intrinsic and projected coordinate systems of a galaxy [13], we deproject each MGE component, multiply by a radially varying $M^*/L_{\rm F110W}$ (see below), and add the deprojected Gaussians to obtain the three-

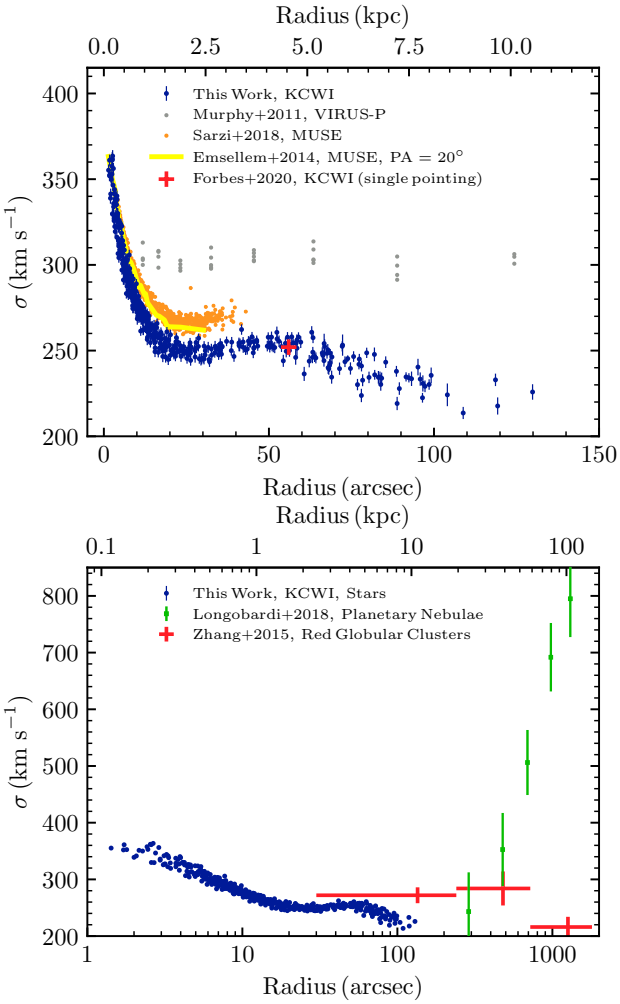

Figure 5.3: Radial profile of stellar velocity dispersion of M87 in the inner 10 kpc (on a linear scale; top) and out to 100 kpc (on a logarithmic scale; bottom). All 461 KCWI bins are shown (blue) but many overlap. (Top) The KCWI values within 1 kpc agree well with those from MUSE on the Very Large Telescope ([50, 172]; yellow and orange respectively), while the MUSE values are 10–20 km s^{-1} larger than KCWI between 1 and 3 kpc. At 4.5 kpc, our KCWI measurements match the single data point (red) from an independent KCWI observation [70]. The VIRUS-P values ([144]; grey), which were used in the axisymmetric stellar-dynamical measurement of the M87 black hole [77], are 30–50 km s^{-1} higher than all other measurements. [144] had noted a similar offset between their values and earlier IFU measurements [52] in the inner 2 kpc. (Bottom) Red globular clusters have similar σ (red) as stars and appear to belong to M87's stellar halo [228], whereas the intra-cluster component of planetary nebulae have sharply rising σ ([122]; green).

dimensional stellar density. Each deprojected MGE component can have its own axis ratios p, q, and u, where $p = b/a$ is the intrinsic middle-to-long axis ratio, $q = c/a$ is the intrinsic short-to-long axis ratio, and u is the apparent-to-intrinsic long axis ratio. When the best-fit p, q, and u are quoted below, each value is luminosity averaged over the MGE components. Further details of the relations between the apparent and intrinsic shape parameters and the deprojections can be found in Section 2 of [157].

The $M^*/L_{\mathrm{F110}W}$ we use to obtain the stellar density varies radially, following a logistic curve given by

$$\frac{M^*}{L}(R) = \left(\frac{M^*}{L}\right)_{\mathrm{outer}} \left[\frac{\delta + (R/R_0)^k}{1 + (R/R_0)^k}\right], \tag{5.1}$$

where δ is the ratio of the inner and outer $M^*/L_{\mathrm{F110}W}$, and R_0 and k parameterize the location and sharpness of the transition. We choose $\delta = 2.5$, $R_0 = 10''$, and $k = 2$, which together well approximate (Figure 5.7) the spatial $M^*/L_{\mathrm{F110}W}$ profile of M87 determined from [172]. We leave the overall normalization—the outer $M^*/L_{\mathrm{F110}W}$—as a free parameter. A similar form as Equation (5.1) was used in an axisymmetric Jeans dynamical study of M87 globular cluster and stellar kinematics data [148]. We implement this spatial variation in our models by choosing distinct $M^*/L_{\mathrm{F110}W}$ ratios for each component of the MGE such that the profile is reproduced.

The dark matter halo is described by a generalized Navarro-Frenk-White density profile [146]

$$\rho(r) = \frac{\rho_0}{(r/r_s)^\gamma (1 + r/r_s)^{3-\gamma}}, \tag{5.2}$$

where ρ_0 is the density scale factor and r_s is the scale radius. This form of the dark matter halo is used by [116] when fitting axisymmetric Jeans models to M87 globular cluster and stellar kinematics data. They determine that $r_s = 15.7^{+2.3}_{-2.0}$ kpc for the cored $\gamma = 0$ model but find no significant preference for $\gamma = 0$ over $\gamma = 1$. [147], on the other hand, find a strong preference for flat cores with $\gamma \lesssim 0.13$. We have tested models with $\gamma = 0$, 0.5, and 1, and find that the models with a $\gamma = 0$ halo are a better description of the data, with the goodness of fit (χ^2) lower by at least 100. We therefore adopt the flat core, $\gamma = 0$ dark matter halo. Since the KCWI stellar kinematics extend to a projected radius of 12 kpc, we expect r_s and ρ_0 to be quite degenerate; we choose to fix $r_s = 15$ kpc and keep ρ_0 as a free parameter in the models.

For each galaxy model, we compute the trajectories of a library of around 500,000 stellar orbits that sample 120 values of energy, 54 and 27 values of the second integral of motion for the loop and box orbit libraries, and 27 values of the third integral of motion over logarithmically spaced radii from $0\farcs01$ to $316''$. The loop and box orbits are integrated for 2000 and 200 dynamical times, respectively. We project the stellar orbits onto the sky and compute the LOSVDs, accounting for the KCWI point-spread function (PSF) and spatial binning. Using a non-negative least-squares optimization, we determine the orbital weights such that the linear superposition of orbits reproduces the luminous mass (to an accuracy of 1%) and the observed kinematics in each spatial bin. As described below, the procedure is

repeated for a large suite of galaxy models to determine the best combination of the galaxy model parameters.

5.4.2 Parameter search

The best-fit model parameters and uncertainties are determined as follows. We use an iterative grid-free Latin hypercube scheme to select sampling points in the six-dimensional model parameter space [119, 157, 153]. In each iteration, the TriOS code is run to assess the χ^2 of each of the sampled galaxy models. The χ^2 of a model is determined by comparing the data and uncertainties for the lowest eight kinematic moments in each of the 461 spatial bins to the model predictions. An additional set of constraints is imposed on kinematic moments h_9 to h_{12}, in which the value of each moment is required to be zero with error bars comparable to the errors in h_3 to h_8. As shown in Chapter 2 [119], these additional constraints help eliminate spurious behavior in the LOSVDs predicted by the models.

The goodness-of-fit landscape is then approximated using Gaussian process regression (GPR; [159, 150]) with a Matérn covariance kernel. To map the high-likelihood region in finer detail, we run the TriOS code again for a next set of models selected by uniformly sampling a zoom-in volume that lies within the 3-σ confidence level for six parameters in the previous regression surface. A more accurate GPR surface is then obtained from all the available models. After multiple iterations we again use GPR to construct a smooth likelihood surface from all available models (nearly 20,000 in total). Finally, we use the dynamic nested sampler `dynesty` [181] to sample from this surface to produce Bayesian posteriors assuming a uniform prior for all parameters.

Following [157], we search over a different set of shape parameters, T, $T_{\rm maj}$, and $T_{\rm min}$, instead of angles θ, ϕ, and ψ. Such a parameterization maps the deprojectable volume in the viewing-angle space into a unit cube in the shape-parameter space, allowing for simpler and more efficient searches. The definitions of $(T, T_{\rm maj}, T_{\rm min})$ and the relationships with (θ, ϕ, ψ) are given in Section 4.3 [157].

The final posterior distributions yield clear constraints on all six model parameters: $M_{\rm BH}$, outer $M^*/L_{\rm F110W}$, dark matter density ρ_0, T, $T_{\rm maj}$, and $T_{\rm min}$ (Figure 5.4). Instead of the halo density parameter ρ_0, we describe the dark matter halo in terms of the ratio of dark matter to total matter enclosed within 10 kpc, f_{10}. The posterior distributions for the more intuitive (luminosity-averaged) axis ratios p, q, and u are also shown. The best-fit parameters are summarized in Table 5.1.

5.4.3 Black hole mass and stellar mass-to-light ratio

The mass of the M87 black hole has been determined with two other independent methods [221, 57] in addition to the stellar-dynamical method used here. Compared to our value $M_{\rm BH} = (5.37^{+0.37}_{-0.25} \pm 0.22) \times 10^9\ M_\odot$, the value $M_{\rm BH} = (6.5 \pm 0.2 \pm 0.7) \times 10^9\ M_\odot$ inferred from the crescent diameter by the Event Horizon Telescope (EHT) team [59] is 21% higher, but the difference is within 1.5-σ of their uncertainties. A recent re-analysis of EHT observations

M87 Property (units)	Inferred value
Black hole mass $M_{\rm BH}$ (10^9 $M_\odot$)	$5.37^{+0.37}_{-0.25} \pm 0.22$
Outer $M^*/L_{\rm F110W}$ (V-band; $M_\odot/L_\odot$)	$3.46^{+0.04}_{-0.06} \pm 0.15$
Inner $M^*/L_{\rm F110W}$ (V-band; $M_\odot/L_\odot$)	$8.65^{+0.10}_{-0.15} \pm 0.38$
Dark matter fraction at 10 kpc f_{10}	0.67 ± 0.02
Total mass within 10 kpc ($10^{11} M_\odot$)	5.77 ± 0.12
Shape parameter T	0.65 ± 0.02
Shape parameter $T_{\rm maj}$	$0.46^{+0.03}_{-0.02}$
Shape parameter $T_{\rm min}$	0.61 ± 0.02
Average middle-to-long axis ratio p	0.845 ± 0.004
Average short-to-long axis ratio q	0.722 ± 0.007
Average apparent-to-intrinsic long axis ratio u	0.935 ± 0.004
Line-of-sight direction θ, ϕ (°)	$48.9^{+1.1}_{-1.0}$, $37.5^{+1.4}_{-1.3}$
Rotation about line-of-sight ψ (°)	$-61.3^{+1.4}_{-1.7}$

Table 5.1: Mass and shape properties of M87. The search over galaxy parameters in the triaxial orbit modeling in this paper is performed over $M_{\rm BH}$, outer $M^*/L_{\rm F110W}$, halo scale density ρ_0, and the shape parameters T, $T_{\rm maj}$, and $T_{\rm min}$. All other parameters in the table are computed from the posteriors of those six parameters. For the two primary mass parameters $M_{\rm BH}$ and $M^*/L_{\rm F110W}$, the second set of errors denotes systematic uncertainties (68% confidence levels) due to the uncertainty in the distance to M87: 16.8 ± 0.7 Mpc [59].

[17] revised the black hole mass to $M_{\rm BH} = (7.13 \pm 0.39) \times 10^9\ M_\odot$, which is 33% above our value, but [193] cautioned the false-positive tendency of the method used in the re-analysis and found that significant systematic uncertainties were not taken into account. The ionized gas-dynamical determination of $M_{\rm BH} = (3.45^{+0.85}_{-0.26}) \times 10^9\ M_\odot$ (after scaling to our adopted distance of 16.8 Mpc) is 36% below our value [221, 59].

Before this work, the most recent mass measurement of the M87 black hole that also used orbit-based stellar dynamics obtained [77, 59] $M_{\rm BH} = (6.14^{+1.07}_{-0.62}) \times 10^9\ M_\odot$ (after scaling to our adopted distance of 16.8 Mpc), which is 14% above our value. Despite the apparent consistency, there are many differences between the two measurements. In this work, the stellar spectra are obtained in a homogeneous manner from the latest IFU at the Keck Telescope over a contiguous $250 \times 300''$ field and have S/N of around 100 per Å for the outermost bins and above 200 per Å for central bins. The observed stellar velocity dispersions used to constrain the orbit models in this work are about 20% lower than [77, 144] beyond 1 kpc (Figure 5.3; top panel), but this work is in broad agreement with other recent measurements [50, 172, 70]. The orbit modeling in this work allows for triaxiality, and the $M_{\rm BH}$ is obtained from a full six-dimensional model parameter search with posteriors measured using a Bayesian framework. Furthermore, [77] adopts a spatially constant V-band $M^*/L_{\rm F110W}$ of 9.7 $M_\odot/L_\odot$ (scaled to our distance of 16.8 Mpc). However, a recent detailed stellar population analysis of M87 reported a negative radial gradient due to a changing stellar initial mass function [172]. When incorporating the shape of this $M^*/L_{\rm F110W}$ gradient into our stellar-dynamical models, we find the V-band $M^*/L_{\rm F110W}$ declines from 8.65 $M_\odot/L_\odot$ at the center to an outer value of 3.46 $M_\odot/L_\odot$.

Using either $M^*(< r_{\rm SOI}) = M_{\rm BH}$ or $M^*(< r_{\rm SOI}) = 2M_{\rm BH}$ as the definition of a black hole's gravitational sphere of influence (SOI), we find the SOI radius of the M87 SMBH to be $r_{\rm SOI} = 4\farcs4$ (0.36 kpc) or $6\farcs1$ (0.50 kpc).

5.4.4 Dark matter mass

At the outer reach of our data, at a radius of 10 kpc, we find the enclosed dark matter mass to be $M_{\rm DM}(< 10\,{\rm kpc}) = (3.88 \pm 0.12) \times 10^{11}\ M_\odot$, constituting about 67% of the total mass of the galaxy (f_{10} in Table 5.1). A similar dark matter fraction (73% at 14.2 kpc) is obtained from Jeans modeling of the kinematics of globular clusters [116]. A lower dark matter fraction (about 30% at 11 kpc) is estimated from axisymmetric orbit-based modeling of the kinematics from stars and globular clusters [144]. This lower fraction arises mainly from their high estimate of $M^*/L_{\rm F110W}$ discussed in the previous paragraph.

Our inferred total mass of M87 within 10 kpc is $M_{\rm tot}(< 10\,{\rm kpc}) = (5.77 \pm 0.12) \times 10^{11}\ M_\odot$. Dynamical modeling of globular clusters under the assumption of spherical symmetry yields very similar value at the same radius [225, 163] but with large modeling uncertainties [225]. Estimates from axisymmetric orbit models find a 15% lower value [144]. Jeans modeling studies [147, 116] incorporating a radially declining $M^*/L_{\rm F110W}$ find a total mass enclosed within 10 kpc to be in the range of (3–7.5) $\times 10^{11}\ M_\odot$.

5.4.5 M87's intrinsic shape

Our orbit modeling results show that M87 is strongly triaxial, where the lengths of the short and middle principal axes are 72% and 85% of the length of the long axis, corresponding to q and p, respectively. A triaxiality parameter often used to quantity the ratios of a galaxy's principal axes is $T = (1-p^2)/(1-q^2) = (a^2-b^2)/(a^2-c^2)$. This parameter ranges between $T = 0$ for an oblate axisymmetric shape ($p = 1$ or $a = b$) and $T = 1$ for a prolate axisymmetric shape ($p = q$ or $b = c$), with values between 0 and 1 indicating a triaxial shape. Our inferred value for M87 is $T = 0.65 \pm 0.02$, strongly excluding the possibility that M87 is an axisymmetric galaxy.

The shape parameters p, q, and u in Table 5.1 are related to a set of angles θ, ϕ, and ψ that uniquely specify the orientation of M87's intrinsic axes with respect to its projected axes on the sky [202, 157]. The angles θ and ϕ specify the direction of the line-of-sight from M87 to the observer; they are the usual polar angles in M87's intrinsic coordinate system. The inclination angle $\theta = 0°$ corresponds to a face-on view of M87 along its intrinsic short axis, and $\theta = 90°$ corresponds to an edge-on view with the short axis in the sky plane. The azimuthal angle $\phi = 0°$ places the intrinsic middle axis in the sky plane and $\phi = 90°$ places the intrinsic long axis in the sky plane. Once the line of sight is described by θ and ϕ, the third angle ψ specifies the remaining degree of freedom for the rotation about the line of sight. Our best-fit angles for M87 are $(\theta, \phi, \psi) = (48.^\circ9, 37.^\circ5, -61.^\circ3)$. Thus, we are viewing M87 from a direction that is roughly equidistant from all three principal axes.

5.4.6 Angular momentum vector and origin of kinematic misalignment

To gain physical insight into the origin of the observed misalignment between the kinematic axis and photometric major axis of M87 on the sky (Figure 5.2; lower right), we examine the direction of the total angular momentum vector, $\boldsymbol{L}$, of the stars predicted by our best-fit orbit model and how it would be projected on the sky. To do this, we sum the individual contributions to the angular momentum from the superposition of stellar orbits and compute the total $\boldsymbol{L}$. Among the three major orbital types computed in the TriOS code, the box orbits supported by a triaxial gravitational potential, by construction, have zero angular momentum, whereas the short-axis and long-axis tube orbits have net $\boldsymbol{L}$ along the intrinsic short axis and long axis, respectively [173, 202, 157]. The direction of the total $\boldsymbol{L}$ is therefore determined by the relative contributions from the two types of tube orbits [73].

The rotational velocity of M87 reaches sufficiently high amplitudes beyond about 5 kpc (Figures 5.1 and 5.2) for us to determine the direction of $\boldsymbol{L}$ robustly. We find it to point approximately 60° off of the intrinsic short axis. Using the best-fit viewing angles to project $\boldsymbol{L}$ on the sky, we find it to lie at a PA of approximately $-60°$. Because the projected $\boldsymbol{L}$ is orthogonal to the kinematic axis of the projected velocity field, this simple calculation indicates that the PA of the kinematic axis predicted by the model is around $-150°$, very similar to the observed kinematic axis. The observed kinematic misalignment of M87 on the

sky is therefore a result of both projection effects of a triaxial galaxy and a physical offset between the total angular momentum vector and the intrinsic short axis of the galaxy.

M87's jet lies at a PA of approximately $-72°$ oriented about 17° off of the line of sight [141, 216]. Our $\boldsymbol{L}$ is oriented roughly 30° off of the line of sight, or only about 15° off of the jet axis. This is a remarkably close alignment, as only about 1.7% of randomly oriented vectors would be expected to lie within 15° of each other. On a much larger scale, the principal axis of the Virgo cluster has a projected PA of roughly $-70°$ and lies 20–40° off of the line of sight [223, 140]. This axis is only about 11° off of our model's net $\boldsymbol{L}$. These alignments suggest that the persistent intrinsic sense of rotation within M87's stellar halo is related in part to both the overall structure of the Virgo cluster and the orientation of the jet, which in turn is thought to be related M87*'s spin [58]. The origins and implications of this multiscale collimation are as of yet unclear but are under active investigation.

5.5 Conclusions

With 4000 constraints from Keck KCWI and our latest triaxial orbit modeling code and procedure for sampling high-dimensional parameter spaces even with computationally intensive models, we are able to relax the common assumption of axisymmetry and present the most comprehensive stellar-dynamical study of the M87 galaxy and its central black hole. This work is one of only a small number of studies that have produced constraints on all three intrinsic shape parameters for individual galaxies [103, 169]. Even fewer galaxies have been observed with sufficient angular resolution, field of view, spectral coverage, and S/N for a simultaneous determination of the intrinsic shape, supermassive black hole mass, and galaxy mass [200, 219, 42, 157, 153]. As demonstrated in this work, further advancements have only been made possible by the installations of wide-field and highly sensitive IFUs on large ground-based telescopes.

Moving forward, it is crucial to apply triaxial stellar-dynamical orbit models to larger samples of galaxies, thereby advancing this method from a rarity to a standard technique. This is especially pertinent for massive elliptical galaxies such as M87 because the majority of them—when a rotational pattern can be detected in the stellar velocity field—show some degree of misalignment between the kinematic and photometric major axes, extending to the half-light radius and beyond [54, 110, 55]. Such an offset indicates triaxiality [13, 73]; an axisymmetric galaxy would, by symmetry, produce only aligned kinematic and photometric major axes.

When direct comparisons between axisymmetric and triaxial modeling were made on the same galaxy, the black hole mass from axisymmetric models has ranged from about 50% [200] to 170% [153] of the mass when triaxiality was allowed; and in two galaxies, the black hole mass did not change appreciably [200, 119, 157]. Overall, triaxial models were able to match the observed stellar kinematics significantly better than axisymmetric models [157, 153].

More secure black hole masses could result in significant changes to the local black hole census and the shapes of the scaling relations between black holes and host galaxies, thereby impacting our understanding of black hole fueling and feedback physics, as well as binary black hole merger physics used to forecast and eventually interpret gravitational wave signals for Pulsar Timing Arrays [187] and space-based detectors [2]. In terms of black hole imaging studies, since the photon ring diameter ranges from about 9.6 to 10.4 gravitational radii depending on the black hole spin [59], future analyses combining direct imaging with stellar kinematic measurements such as that presented this paper have the potential to significantly improve the prospects for measuring black hole spins.

Appendix

5.A Keck KCWI data reduction and analysis

We observed M87 using the integral-field spectrograph KCWI on Keck. We used the BL grating centered on 4600 Å and the Kblue filter to obtain the widest wavelength coverage, reducing possible template mismatch during the subsequent extraction of the stellar kinematics. The integration time per exposure varied from 300 s for the central pointings to 1500 s for the outermost pointings with low surface brightness. We periodically acquired offset sky exposures in between the on-source galaxy exposures, each roughly half the integration time of the adjacent galaxy exposures. Only data taken in good observing conditions are used in this analysis; the on-source and sky exposure times total 13 hr and 2.8 hr, respectively.

5.A.1 Data reduction

The KCWI Data Extraction and Reduction Pipeline [142] is actively maintained on a publicly accessible GitHub repository. We use the `IDL` version of the pipeline with its default settings to reduce each frame. The main steps include overscan and bias removal, cosmic ray rejection, dark and scattered light subtraction, solving for the geometric distortion and wavelength solution, flat-fielding, correction for vignetting and the illumination pattern, sky subtraction, and the generation of datacubes using the spatial and spectral mappings determined previously. The pipeline then corrects for differential atmospheric refraction and applies a flux calibration using a standard star.

In addition to the default pipeline, we perform custom steps to improve the quality of the processed data. Some cosmic rays are improperly removed by the KCWI pipeline, leaving sharp features at certain wavelengths in a small number of spaxels in our datacubes. We therefore scan through each wavelength slice of the cubes, mask the impacted pixels, and perform an interpolation to replace their values with those of neighboring pixels. Furthermore, beyond about $100''$, the KCWI spectra are sky-dominated and subtle mis-subtraction of the sky can result in significant reduction of the S/N of the galaxy spectra. The sky subtraction stage of the KCWI pipeline uses b-spline interpolation to build a "noise-free" model

of the sky in each pixel that is subtracted from the corresponding object exposure. We find that this routine does not capture highly space- or time-variant sky features, so we further remove residual sky features using the combination of a principal component analysis (PCA) and the penalized pixel-fitting (pPXF; [26]) method, as described in Appendices 5.A.4 and 5.B.

In the final step, we merge the on-source M87 datacubes. Roughly half of the pointings were taken with the long axis of KCWI aligned with a PA of $-25°$ and half were oriented perpendicular to this with a PA of $-115°$. We construct a pair of datacubes, one for each of the two orientations using the `nifcube` and `gemcube` IRAF tasks that are part of Gemini's data reduction software. We input the fully calibrated KCWI datacubes (the "_icubes.fits" files) and map the cubes onto a shared grid with a spacing of $0\farcs3 \times 1\farcs4 \times 1$ Å. This choice of spaxel size matches the native scale of the individual KCWI datacubes for our observational setup.

5.A.2 Line-spread function

We find that our selected spectrograph configuration produces a line-spread function (LSF) that is distinctly non-Gaussian (Figure 5.5). The LSF is instead well described by the convolution of a Gaussian function and a top-hat function of the form

$$\begin{aligned}\mathcal{L}(\lambda) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\lambda-\tau)^2}{2\sigma^2}} \Pi\left(\frac{\tau}{\Delta}\right) d\tau \\ &= \frac{1}{2}\left[\mathrm{erf}\left(\frac{\lambda+\Delta/2}{\sigma}\right) - \mathrm{erf}\left(\frac{\lambda-\Delta/2}{\sigma}\right)\right],\end{aligned} \tag{5.3}$$

where $\Pi(x) = 1$ if $|x| \leq 1/2$ and 0 otherwise, Δ is the full width of the top-hat component, and σ is the standard deviation of the Gaussian. To measure the widths of the Gaussian and top-hat components of the LSF, we simultaneously fit 31 lines of an FeAr arc lamp spectrum between 4500 and 5000 Å and determine $\Delta = 5.105$ Å and $\sigma = 0.627$ Å. Repeating this procedure on different spectral or spatial regions yields comparable best-fit parameters.

5.A.3 Point-spread function

During the first night of observations, we took KCWI data of the inner region of M87 and the atmospheric seeing was estimated to be $0\farcs63$ by the differential image motion monitor at the nearby Canada France Hawaii Telescope weather station. This estimate is consistent with the broadening of point sources measured from exposures taken with the guider camera. During the other four nights, we observed the outer regions of M87 and measured similar seeing. While running stellar-dynamical models, described in Section 5.4, we use a PSF that is a Gaussian with a full width at half maximum (FWHM) of $0\farcs63$ ($\sigma = 0\farcs28$).

5.A.4 PCA decomposition of sky features

As part of the process to remove residual sky features seen in the reduced M87 datacubes, we perform a PCA decomposition of the sky spectra. For each sky cube, we apply a conservative spatial masking of possible sources in the field and coadd the unmasked pixels to obtain a high S/N sky spectrum. A weighted expectation-maximization PCA [4] is then applied to each of the sky spectra between 3800 and 5650 Å. Since the amplitudes of the 4861 Å Hβ, 5200 Å [N I], and the 5577 Å [O I] lines are highly variable and are not well captured with a PCA decomposition [207], we mask these features. The first PCA component is effectively the mean sky spectrum. The second and fourth components capture slight variations in the shape of the continuum and the Ca H and K features. The third and fifth components capture variations in the numerous OH lines. While we obtain measurements of the first ten components, the fifth component and beyond are consistent with noise. A similar routine was previously applied to KCWI observations [207] and this method is similar in spirit to the Zurich Atmospheric Purge (ZAP; [180]) used for MUSE observations.

5.A.5 Spectral and spatial masking

We mask nine spectral features, which together span a total of 274 Å (Figure 5.6). The masked features include emission lines that are prominent at the nucleus, as well as the 4861 Å Hβ, 5200 Å [N I], and 5577 Å [O I] lines that are masked in the PCA decomposition. The Mg I b region (5184–5234 Å) is also masked because it is contaminated by the 5200 Å [N I] skyline and is coincident with Fe emission features at M87's redshift.

We also apply a spatial mask to exclude potentially contaminant spaxels. This is done by collapsing the datacubes spectrally, flagging regions of spaxels with substantially higher surface brightness than their surroundings, and then masking the brightest spaxels in those regions. This process removes the spaxels that are contaminated by the prominent jet, the central $\sim 0\farcs85$ that is affected by the active galactic nucleus (AGN), and numerous bright globular clusters.

5.A.6 Spatial binning

We use the `vorbin` package [21] to construct spatial bins and obtain coadded KCWI stellar spectra with uniformly high S/N. By default, `vorbin` calculates the S/N of each coadded spectrum based on values of the signal and the noise of the individual spaxel spectrum given by the user, adding the signals linearly and the noise in quadrature. Instead of this default setting, we modify the `sn_func()` routine in `vorbin`'s `voronoi_2d_binning` to nonanalytically recompute the S/N from the M87 datacube while binning. This approach improves the uniformity of the resultant S/N across the bins as it naturally incorporates spatial correlations in the signal and noise between spaxels. We estimate the S/N by first smoothing the spectrum with a Gaussian kernel with FWHM = 4 Å, comparable to the LSF. The noise is taken to be the root-mean-square (rms) difference between the raw and smoothed spectra,

while the signal is taken to be the median flux of the raw spectrum. We apply the spectral masks described above before smoothing to avoid contamination from sharp features in the spectra.

This procedure results in 461 spatial bins and a coadded spectrum for each of the bins. The S/N per Å ranges from about 200 in the central regions to about 100 in the outer regions. Figure 5.6 shows a series of representative KCWI spectra (black curves) for ten of the 461 spatial bins located at projected radii of 1″–130″.

5.B Stellar kinematic determination

We measure the stellar LOSVD for each of the 461 binned spectra using pPXF [26]. With pPXF, we convolve a linear combination of template stars with an LOSVD, parameterized by V, σ, and high-order Gauss-Hermite moments h_3–h_8 that account for asymmetric and symmetric deviations from a Gaussian velocity distribution [132]. The S/N of our data enable the measurement of high-order Gauss-Hermite moments. We find that truncating the series at h_4 (or h_6) results in elevated values for h_4 (or h_6), but when fitting to h_8 or h_{12}, the values of h_4 and h_6 converge and the highest extracted moments become consistent with 0, as seen in past work [153, 119]. In addition, we find it important to constrain the kinematic moments beyond h_4 in dynamical modeling. When those moments are not constrained in orbit models, the models are prone to producing LOSVDs with unphysical features due to large values in the high-order moments, potentially biasing the preferred model parameters [156, 119].

For stellar templates, we use the MILES library [64, 167] but select 485 spectra out of the full 985 templates that have well-identified spectral types and luminosity classifications. These stellar templates have a higher spectral resolution than our observations and are degraded to match the KCWI (non-Gaussian) LSF before fitting with pPXF.

During the kinematic fit, we use an additive polynomial of degree one and a multiplicative polynomial of degree 15 to model the stellar continuum. We also supply the PCA components that describe the sky background to pPXF. This procedure results in a weighted combination of the PCA components, which is included as an additional additive term to match the residual sky features in the M87 spectra that remained after the KCWI pipeline's default sky subtraction. Ultimately, we use the first ten PCA components, but find that the extracted Gauss-Hermite moments are unchanged as long as at least the first five PCA components are included in the fit.

Because we excluded three highly variable sky lines during the PCA decomposition process, we also mask those spectral regions when running pPXF, as well as emission lines associated with M87 and the Mg I b region, as described previously. In contrast to the other masked regions, we find that the extracted Gauss-Hermite moments depend strongly on the endpoints of Mg I b mask and only stabilize once the entire 5184–5234 Å region is excluded from the fit.

$I_k\,[L_\odot/\mathrm{pc}^2]$	σ'_k [arcsec]	q'_k
2382.4	0.039	0.860
2460.8	0.206	0.906
1598.8	0.508	0.959
973.48	1.468	1.000
1830.9	4.558	1.000
1515.4	9.851	0.980
592.67	22.228	0.942
180.27	54.299	0.910
35.555	126.213	0.735
9.8327	293.034	0.650
1.6948	567.511	1.000

Table 5.2: Best-fit MGE parameters for the surface brightness of M87. For each of the 11 two-dimensional Gaussian components, the first column lists the central surface brightness density, the middle column lists the dispersion of the Gaussian, and the last column lists the axis ratio, where primed variables denote projected quantities. We obtain the MGE by fitting to the V-band light profile in [109]. To impose a $M^*/L_{\mathrm{F110}W}$ gradient in the dynamical models, the I_k values are adjusted to reproduce the profile in Figure 5.7.

Altogether, we fit the Gauss–Hermite moments, polynomial coefficients, template weights, and sky weights simultaneously. The stellar templates broadened by the best-fit LOSVD provide excellent fits to each of the observed spectra, as illustrated by the red curves for the ten representative spectra shown in Figure 5.6.

The measurement uncertainties on the LOSVDs are determined as follows. After an initial fit to each binned spectrum, we perturb the spectrum at a given wavelength by drawing a random number from a Gaussian distribution centered on the spectrum and with a dispersion equal to the rms of the pPXF residuals from the preliminary fit at that wavelength. We perform 1000 such perturbed fits with the pPXF bias parameter set to 0 and determine the mean and standard deviation of each moment over those 1000 realizations, which we adopt as the kinematic value and its 1-σ uncertainty. For bins in the central $100'' \times 100''$ region, the mean error on V is 2.6 km s^{-1} and on σ is 3.0 km s^{-1}. The mean errors on h_3 through h_8 are similar, spanning from 0.009 to 0.016. The typical errors in the outer bins are slightly larger with mean errors on V, σ, and h_3 through h_8 of 2.6 km s^{-1}, 3.4 km s^{-1}, and 0.012–0.022, respectively.

5.C Surface brightness of M87

Besides the stellar kinematics, another constraint used in the dynamical models is the galaxy's luminosity density. We use a previously published V-band light profile, along with

measurements of the ellipticity and PA of the isophotes [109]. The profile extends from $0''.017$ to $2400''$ and comes from a combination of ground-based data and high-resolution HST images, which have been deconvolved to remove the effects of the PSF as well as the AGN.

We fit the sum of multiple two-dimensional Gaussians to the composite surface photometry. These MGE [24] approximations are commonly used because they are able to match the surface brightnesses of galaxies while also enabling analytical deprojections to obtain intrinsic luminosity densities. Our best-fit MGE reproduces the surface brightness between $0''.1$ and $500''$ within 10%. This MGE has 11 Gaussian components that share the same center and PA of $-25°$. While the value of the PA in [109] varies within $50''$, the isophotes between $1''$ and $50''$ are very round with ellipticity $\epsilon \lesssim 0.08$; using a constant PA in our MGE therefore does not affect the quality of the fit. The MGE parameters are given in Table 5.2.

5.D Orbit modeling

Radial profiles of the $M^*/L_{\mathrm{F110}W}$ ratio and stellar kinematics used in the orbit modeling in this work are shown in Figure 5.7 and Figure 5.8, respectively.

Acknowledgements

We dedicate this work to the late Wal Sargent, who reported the first observational evidence for the M87 black hole and was a mentor to generations of scientists including C.-P.M. We thank Scott Tremaine, Michael Johnson, Charles Gammie, and the referee for insightful comments. E.R.L. and C.-P.M. are supported by NSF AST-1817100 and AST-2206307. J.L.W. is supported by NSF AST-1814799 and AST-2206219. C.-P.M. acknowledges the support of the Heising-Simons Foundation and the Miller Institute for Basic Research in Science. The spectroscopic data presented in this paper were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. This work used observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) at the San Diego Supercomputing Center through allocation AST180041, which is supported by NSF grant ACI-1548562.

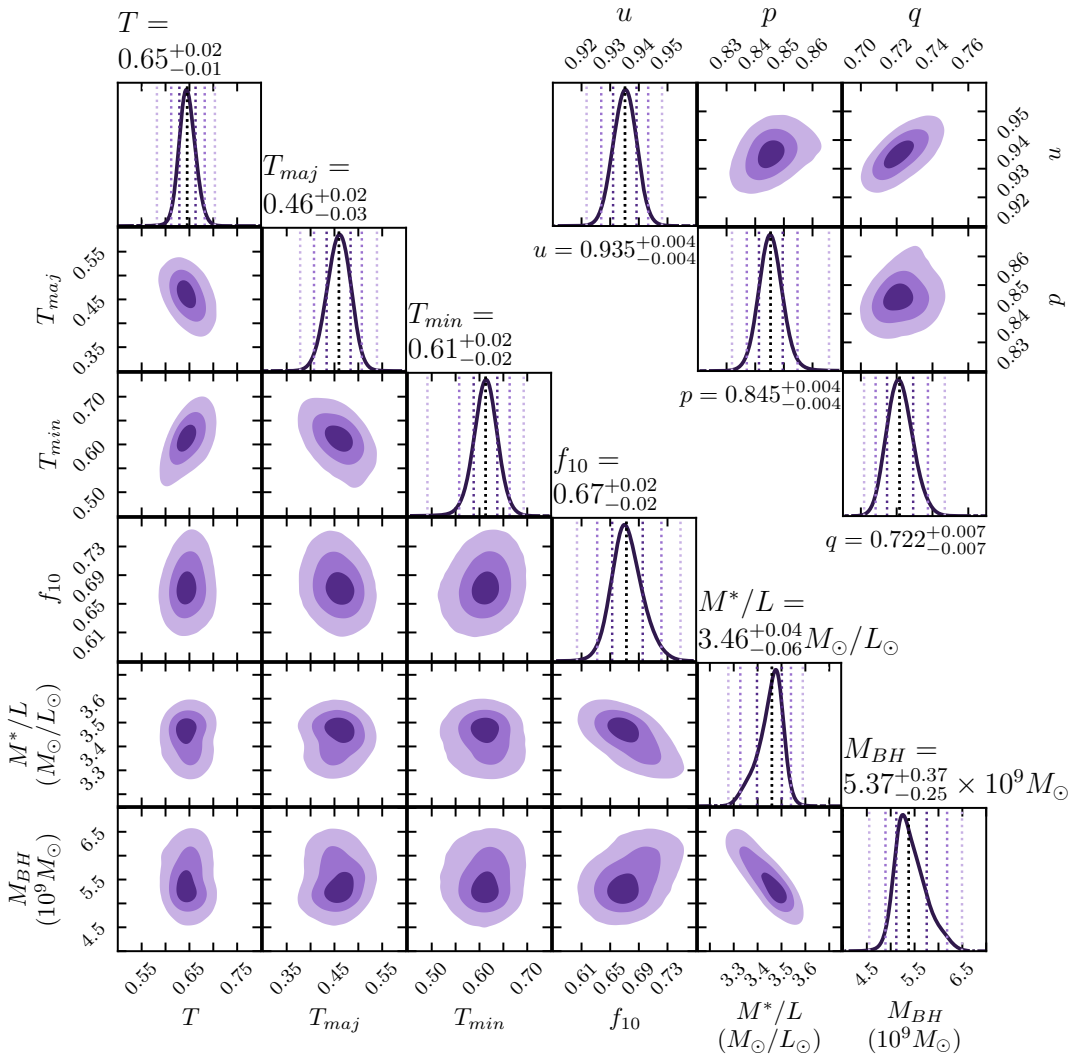

Figure 5.4: Posterior distributions of six parameters from triaxial Schwarzschild orbit modeling of M87: black hole mass $M_{\rm BH}$, outer stellar mass-to-light ratio $M^*/L_{\rm F110W}$, dark matter fraction enclosed within 10 kpc f_{10}, and shape parameters T, $T_{\rm maj}$, and $T_{\rm min}$. The posterior distributions of the luminosity-averaged axis ratios u, p, and q are shown in the upper right. The three levels of purple shading bound the 1σ, 2σ, and 3σ regions (68%, 95%, and 99.7% confidence levels, respectively) of the parameters. The vertical lines in each 1-dimensional distribution indicate the median and the corresponding three confidence levels.

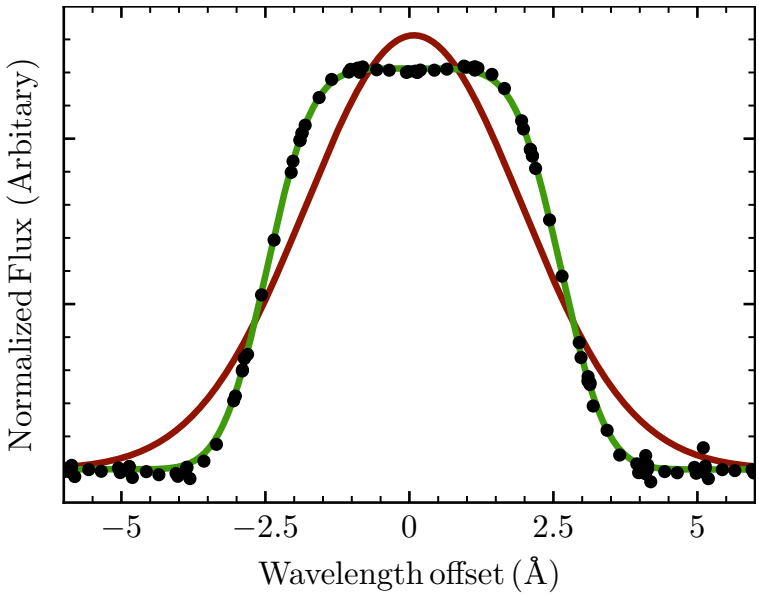

Figure 5.5: Line Spread Function of Keck KCWI with BL grating. We find the LSF of KCWI BL grating to be well approximated by a Gaussian function convolved with a top-hat function, as shown in Equation 5.3. To measure the shape of the LSF, we simultaneously fit 31 lines of an FeAr lamp spectrum as described in Appendix 5.A.2. Here we plot a superposition of the nine most prominent of those lines. Black points mark the flux in the lamp spectrum around each line after normalizing for each line's amplitude. Our best-fit LSF model (green) has a top-hat function of width $\Delta = 5.105$ Å convolved with a Gaussian function of $\sigma = 0.627$ Å. A single Gaussian function, as is typically assumed, would provide a very poor fit to the KCWI LSF (red).

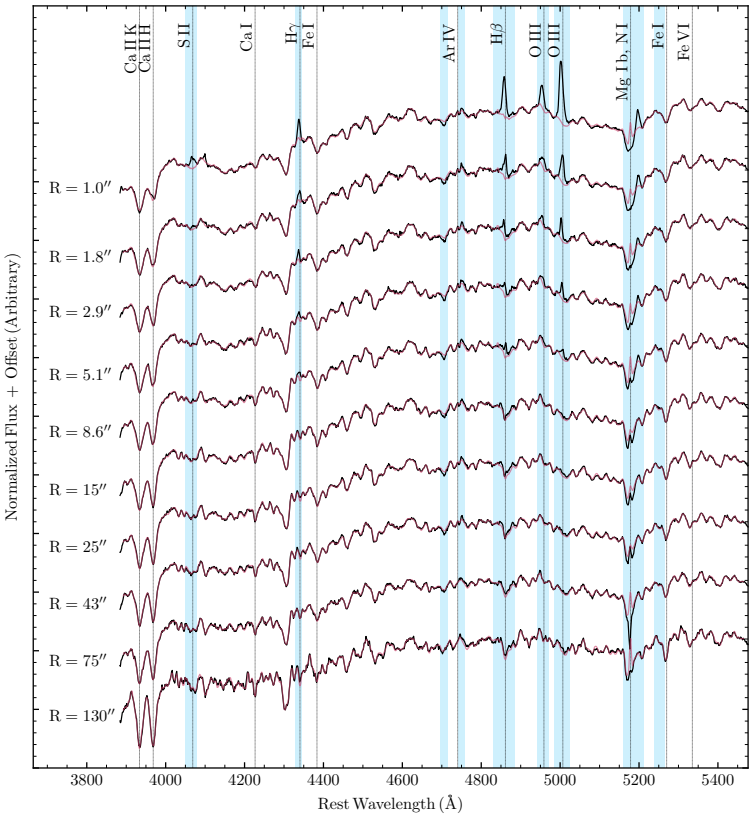

Figure 5.6: Representative KCWI spectra of M87. Sky-subtracted galaxy spectra (black curves) for ten representative spatial bins located at projected radii from 1″ to 130″ are shown. A total of 461 binned spectra are used in this work. The S/N of these co-added spectra range from about 100 to 200 per Å. The stellar template broadened by the best-fit LOSVD is overlaid (red curves) on each spectrum. Typical fitting residuals are comparable to the line widths. Shaded blue regions indicate masked spectral regions excluded from the analysis. At M87's redshift, the Mg I *b* region is contaminated by the 5200 Å [N I] sky line (marked). The central bins exhibit strong AGN emission lines, especially from [OIII] (4959 Å and 5007 Å), [SII] (4069 Å and 4076 Å), Hβ (4861 Å), and Hγ (4330 Å).

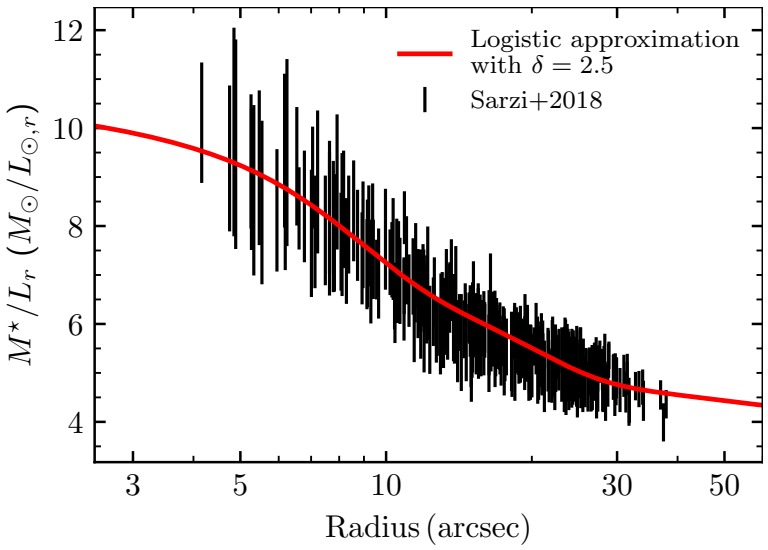

Figure 5.7: Radial profile of $M^*/L_{\mathrm{F110}W}$ ratio used in this work. The logistic approximation (red) used in our modeling, given by Equation (5.1), is chosen to match the shape of the r-band $M^*/L_{\mathrm{F110}W}$ (black) in Figure 11 of [172]. The inner $M^*/L_{\mathrm{F110}W}$ is $\delta = 2.5$ times the outer $M^*/L_{\mathrm{F110}W}$ ratio, and the transition is centered around 10 arcsec. Our dynamical model prefers an outer V-band $M^*/L_{\mathrm{F110}W}$ of $3.46^{+0.04}_{-0.06} M_\odot/L_\odot$ and inner $M^*/L_{\mathrm{F110}W}$ of $8.65^{+0.10}_{-0.15}\ M_\odot/L_\odot$.

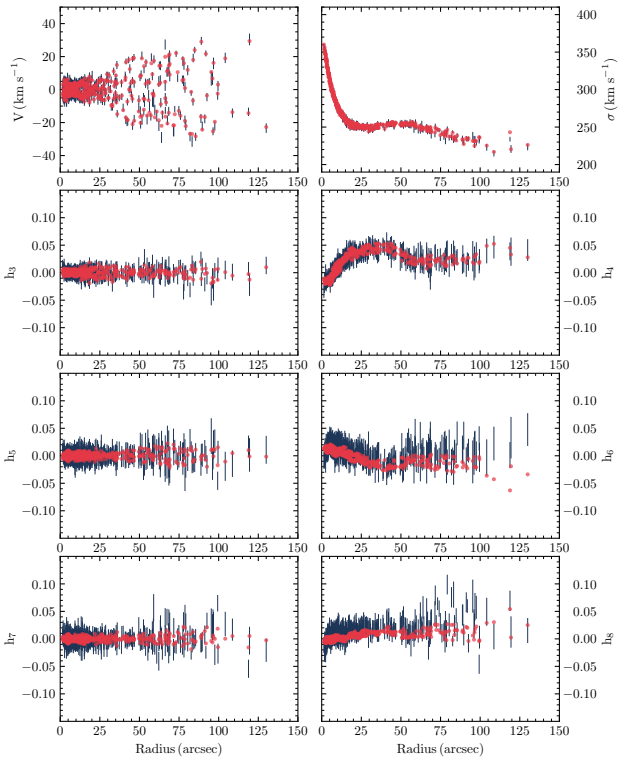

Figure 5.8: Radial profiles of the first eight moments of the stellar LOSVDs. The observed Keck KCWI moments (gray) are well matched by the moments predicted by the best-fit model (red) given by Table 5.1. The triaxial orbit models produce point-symmetric LOSVDs, so we have point-symmetrized the kinematic moments before fitting.

Chapter 6

Unveiling the 22-Billion Solar Mass Black Hole in Holmberg 15A with KCWI Spectroscopy and Triaxial Modelling

This dissertation closes by turning to my recent and ongoing work on the brightest cluster galaxies H15 and IC1101. These two galaxies have exceptionally large and diffuse central regions which suggest the presence of exceptionally large central supermassive black holes. This chapter presents our work to map the stellar kinematics of these two galaxies using the Keck Cosmic Web Imager (KCWI), expanding on techniques which were learned and developed in Chapter 5. For H15, I present the results of triaxial Schwarzschild modelling using TriOS, following the lessons and usage procedures developed through the earlier chapters of this dissertation. We find that the supermassive black hole at the center of H15 is roughly half as massive as previously measured. This galaxy, along with others with very massive SMBHs, lies far above the $M_{\rm BH}$–σ relation, but the size of the black hole's sphere of influence is relatively well-predicted by the size of the apparent core in the photometry. Modelling of IC1101 is ongoing and will be presented in future work, but I use the deviations from the r_b–$r_{\rm SOI}$ which were found for H15 to estimate the mass of IC1101's central SMBH.

This chapter is adapted from a manuscript in preparation by Emily R. Liepold, Jonelle L. Walsh, and Chung-Pei Ma.

6.1 Introduction

Supermassive black holes (SMBHs) are understood to lie at the center of all massive galaxies, and are thought to play an instrumental role in directing the evolution of their host galaxies through various modes of SMBH feeding and feedback. Models of this co-evolution takes are part informed by so-called scaling relations which relate various galaxy properties, including

the central SMBH mass.. Famously, the $M_{\rm BH}$–σ relation which connects the SMBH mass to the stellar velocity dispersion ties the two quantities across four decades in the black hole mass – from $10^6 M_\odot$ to $10^{10} M_\odot$ – with relatively little scatter (see e.g. [133]). The relative robustness of this relation, along with others (e.g., $M_{\rm BH}$ vs Bulge mass or luminosity, stellar light deficit, cusp or break radius) suggest that those relations can be used to constrain models for the co-evolution of galaxies and their central black holes. With this picture in hand, precise SMBH mass measurements are vital for drawing accurate relations and those accurate relations are vital for inferring accurate models of co-evolution. This paper focuses on the first link in that chain.

At present the upper end of the scaling relations is less well-constrained. Only six SMBHs with masses above 5×10^9 have been measured dynamically: NGC4486 ($5.4 \times 10^9 M_\odot$, [120]), NGC2832 ($< 9 \times 10^9 M_\odot$, [136]), NGC3842 ($9.7 \times 10^9 M_\odot$, [135]), NGC1600 ($17 \times 10^9 M_\odot$,[192]), NGC4889 ($21 \times 10^9 M_\odot$, [135]), and Holmberg 15A ($40 \times 10^9 M_\odot$, [138]). The five of these with two-sided constraints lie above the $M_{\rm BH}$–σ relation [133]. This is a small sample, but may suggest that either these masses are overestimated or that the same scaling relation does not hold in the high-mass region.

Thomas et al. [192] demonstrated that for a collection of high-mass galaxies with well-determined core radii from core-Sérsic fits, the size of the sphere of influence of the central black hole is closely correlated with the core radius with a relation which is statistically indistinguishable from $r_{\rm b} = r_{\rm SOI}$. Accordingly, for this sample the core radius is also closely correlated with the mass of the central SMBH. This paper examines two brightest cluster galaxies which have been claimed to have exceptionally large cores, Holmberg 15A (hereafter H15) and IC1101, the BCGs of Abell 85 and 2029, respectively. H15 has a redshift of $z \sim 0.0555$ [128]. For a flat ΛCDM with $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_m = 0.31$, this corresponds to an angular diameter distance of 222.3 Mpc and an angular size of 1" corresponds to a physical size of 1.078 kpc. IC1101 has a redshift of $z \sim 0.0777$. For the same cosmology, this corresponds to an angular diameter distance of 303.1 Mpc and 1" corresponds to 1.470 kpc.

H15 was first cataloged by Holmberg in 1937 [95] and its photometry was first examined by Hoessel in 1980 [94]. This work fit modified Hubble profiles to a collection of 108 BCGs of Abell clusters and identified H15 to have the largest core radius of the sample, 6.2 kpc (corrected with the H_0 from above). In recent years, there has been active discussion about the size of this core and whether or not a core is actually present. In 2014 López-Cruz et al. performed Nuker profile fits to R-band photometry from KPNO and r′-band photometry from CFHT [123]. For both sets of fits, they find a cusp radius $r_\gamma = 4.57$ kpc. To estimate the light deficit in the central region, they compare the integrated light from a sum of two Sérsic profiles fit to their photometry against the integrated light from a de Vaucouleurs profile fit to their data. Their inferred light deficit was $L_{\rm V,def} \simeq 5 \times 10^{11} L_{\rm V,\odot}$. This cusp radius suggests a central SMBH mass of $M_{\rm BH} \sim 3.1 \times 10^{11} M_\odot$ [112, 192] while the light deficit suggests $M_{\rm BH} \sim 2.6 \times 10^{11} M_\odot$ [107]. The $M_{\rm BH}$–σ relation [108] for their adopted $\sigma = 310$ km s^{-1} suggests a black hole which is two orders of magnitude smaller at just $2.1 \times 10^9 M_\odot$.

In 2015, Bonfini et al [15] refuted this picture by analyzing the same CFHT data with

variations on Sérsic, core-Sérsic, and exponential profiles, finding that a good fit to the data can be found without including an obvious core. The following year, Madrid and Donzelli [128] followed up with high resolution photometry using the Gemini Multi Object Spectrograph (GMOS) in its imaging mode. They found that in the central arcsecond of their image there is a *surplus* of light compared to what would be inferred by extrapolating the image inward, in full contradiction of López-Cruz et al. They also found consistent fitting behavior with Bonfini et al., where an explicit core is not required to fit the data well. Notably, they state "Our results thus nullify the existence of a supermassive black hole based solely on the presence of a depleted core, which this galaxy, in fact, does not have."

In contrast, the history of claims surrounding IC1101's central SMBH is relatively short. IC1101 has a sparse history of spectroscopic observations including Dressler's famous observation of a rising dispersion profile [45] and confirmation of that behavior by Fisher et al [67]. Dullo et al. [47] observed the galaxy with *HST* photometry, finding that in contrast to H15 a core is required to fit the light profile. They infer a core radius of 4.2 kpc and a central light deficit of $L_{\mathrm{F702W,def}} \simeq 10^{11} L_{\mathrm{F702W},\odot}$. As in analysis of H15 in [123], these suggest extremely large central SMBHs, while the M_{BH}–σ relation suggests a value which is an order and a half of magnitude lower.

We have obtained extensive spectroscopic observations of both H15 and IC1101 using the Keck Cosmic Web Imager to explore the stellar kinematics of the central $\sim 50''$ of both galaxies. We observed H15 in November 2018, 2019, 2020, 2021 and September 2021 and we observed IC1101 in June 2019, May 2020, and March 2022. In Section 6.2 we complement these observations with archival photometric measurements first presented in Madrid and Donzelli [128] and Dullo et al. [47]. Our spectroscopic observations are described and discussed in Sections 6.3 and 6.4. In Section 6.5, we describe our procedures for performing triaxial Schwarzschild orbit modelling on H15. And finally we present, discuss, compare, and contextualize the results our modelling of H15 in Section 6.6.

6.2 Photometric Observations

6.2.1 GMOS Photometry of Holmberg 15A

To characterize the stellar mass distribution of H15, we use archival observations from the Gemini Multi Object Spectrograph (GMOS) located at the Gemini South telescope (PI Madrid [128]). The instrument was operating in its imaging mode with the filter *r_G0326*. In those observations, two exposures of 200 s were obtained with an effective pixel scale of 0.16 arcsec per pixel and a seeing of 0.56 arcsec. We use data from only the central GMOS CCD chip to ensure that the background sky is uniform. This results in a $282'' \times 159''$ field, centered on H15 and with the long axis of the field aligned with north. We perform spatial masking to exclude contaminants and companion objects. The companion galaxy 75" to the northwest of the center of H15 has been modelled and subtracted from the image before performing the primary fitting of H15.

We use the multi Gaussian expansion (MGE) fitting routine `mgefit` [26] to parameterize the surface brightness profile of H15. We use 9 Gaussian components with a fixed PA of $-34°$ to fit the image. The fits tend to reproduce the observations well, with residuals of less than 10% for sectors within 100" of the center.

As previously discussed in [119], we found that some configurations of the `mgefit` fitting routine produced centrally peaky deprojected light distributions. This occurs because for the innermost component of the fit, the width σ is poorly constrained below the scale of the PSF. On the other hand, after deprojection the central density scales with $\rho(r=0) \propto \Sigma(R=0)/\sigma$ (where Σ is the surface brightness). We find that often $\sigma \ll \sigma_{\rm PSF}$ is selected by the fit, resulting in exceptionally large and unconstrained central densities. To ameliorate this issue, we place a constraint on the smallest σ allowed in the fit. For H15 we find that requiring $\sigma > 0.96''$ well balances the quality of fit to the centralmost pixels and the shape of the deprojected light profile.

The photometry, fit, and deprojected light profile for our best-fit model (See Section 6.6) is shown in Figure 6.1. The parameters of the MGE are tabulated in Table 6.1. The MGE becomes increasingly flattened with radius, with $\epsilon = 0.08$ at 1", rising to $\epsilon = 0.32$ at 50". The outermost MGE components are round with $q' = 0.80$ and 0.99, so beyond 50", the MGE surface brightness profile becomes increasingly round. The most flat component of this MGE has $q'_{\rm min} = 0.639$ and the luminosity-averaged flattening is $\langle q' \rangle = 0.82$.

6.2.2 *HST* Photometry of IC1101

For IC1101, we use archival WFPC2 images from the HST archive (Proposal 6228; PI: Trauger). Two exposures were captured in the F702W band with a total exposure time of 2100 s. Two additional F450W exposures were taken with a total exposure time of 2300 s. We use the two F702W images for our analysis as they have somewhat higher S/N than the F450W images. This data was also used for the analysis presented in [47].

The two F702W images are combined with scale of 0.1 arcsec per pixel. We construct a composite PSF image from a collection of point sources in the image. The PSF is fit with an MGE and is well characterized by three components with $\sigma = 0.068''$, $0.120''$, and $0.318''$ with 44.3%, 42.7%, and 13.0% of the PSF mass, respectively. The spatially masked F702W image is well-fit with an MGE, though we find the same issue as with H15 for the central component – if no constraints are placed on the fit, then centralmost Gaussian σ is much smaller than the PSF scale and the deprojected light density is inflated. We impose a constraint that $\sigma > 0.4''$, which provides both a good fit to the data and removes the central peakiness. The image, fit, and deprojected light profiles are presented in Figure 6.2 and the parameters of the MGE are tabulated in Table 6.2.

IC1101's light profile is very flattened, with an ellipticity which grows with radius. At 1", the best-fitting MGE has $\epsilon = 0.09$, which rises dramatically to 0.35 by 10", 0.45 by 20" and 0.52 at 50". The flattest MGE component has $q'_{\rm min} = 0.416$, and the luminosity-averaged flattening is $\langle q' \rangle = 0.656$.

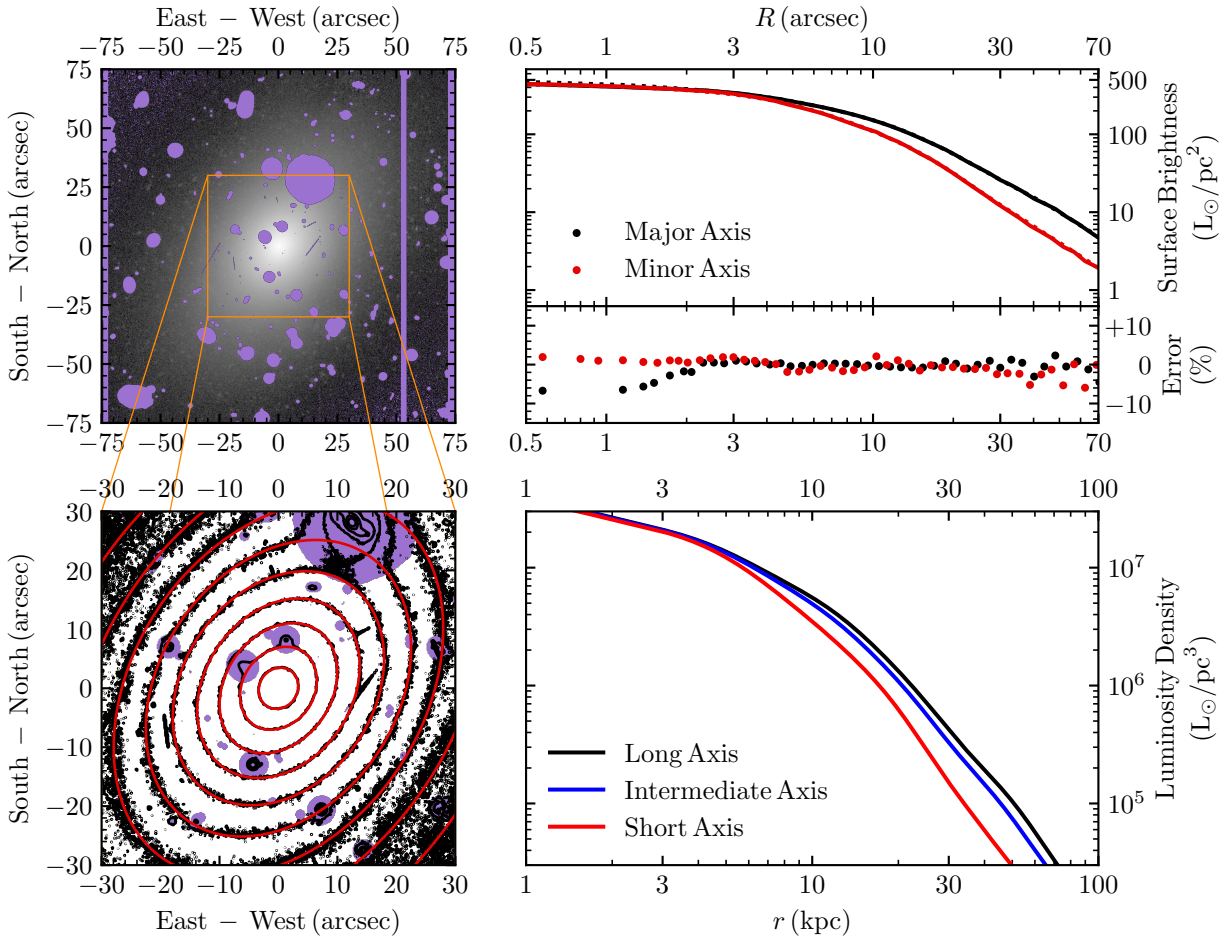

Figure 6.1: **(Upper left)** The *r_G0326*-band GMOS image of H15 used for our photometry [47]. The image is oriented with north up and east left. Masked regions are marked in purple. **(Lower right)** Isophotes of the central region (orange inset) of the GMOS image of H15 are marked in black at 0.5 mag/arcsec^2 intervals. Contours of the MGE fit to the image at the same intervals are marked in red. Masked regions are marked in purple. **(Upper right)** The surface brightness along the major (black) and minor (red) axes are well-reproduced by the MGE model. The fractional difference between the observed surface brightness (solid lines) and the MGE model (dotted lines) is marked in the lower subpanel. **(Lower right)** Deprojected 3D luminosity density for the best-fit model as described in Sec. 6.6. The viewing angles associated with this deprojection are $\theta = 89.31°$ and $\phi = 89.08°$ with $T = 0.352$.

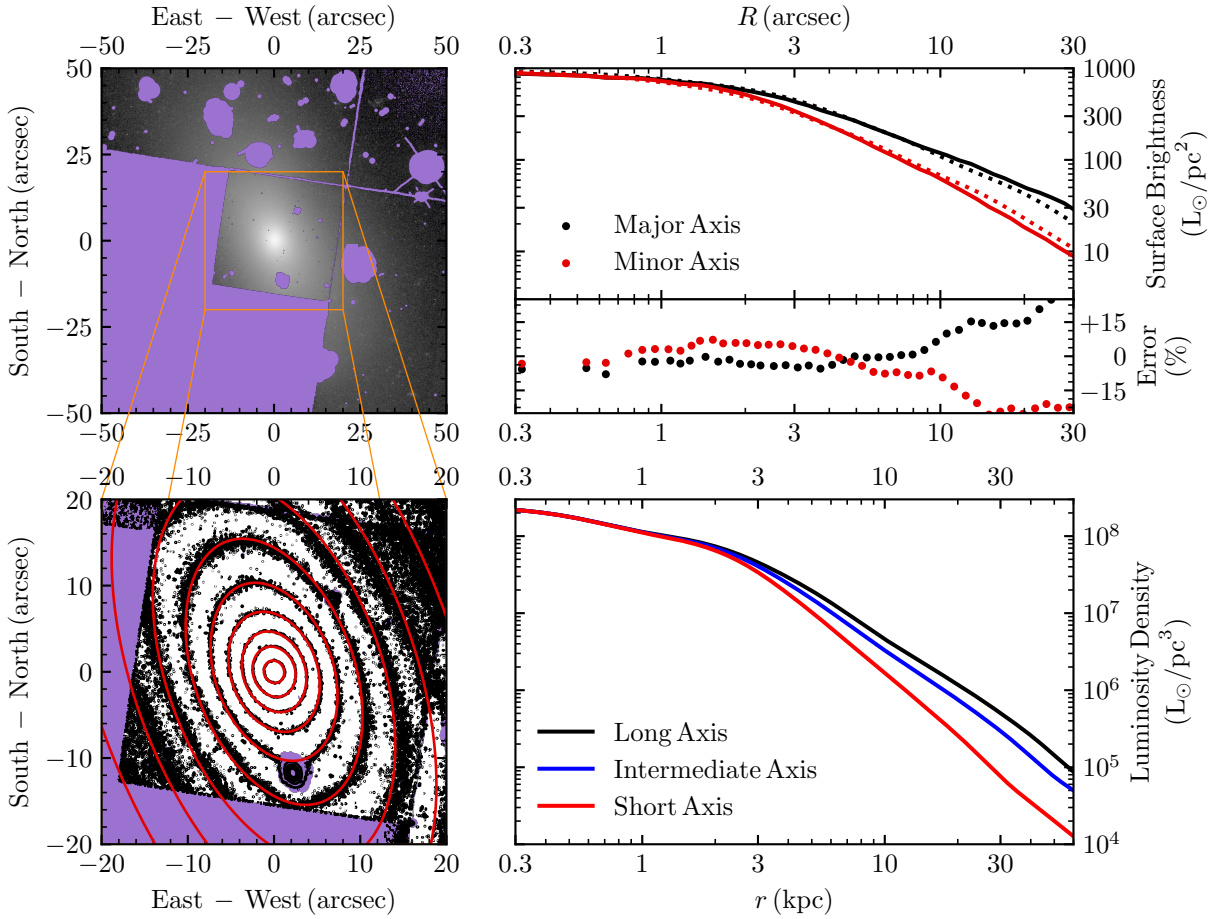

Figure 6.2: Similar to Fig. 6.1, but for the F702W *HST* photometry of IC1101. The lower right panel shows an example deprojection with $\theta = \phi = 90°$ with $T = 0.5$.

6.3 Spectroscopic Observations

6.3.1 KCWI observations of H15

We observed H15 using Keck KCWI during November 2018, 2019, 2020, September 2021, and November 2021. During the November 2018 observations we used the small slicer with the BL grating to observe the central region and during all other observing runs we used the large slicer with the same grating to achieve broad spatial coverage.

The small slicer configuration has an $8\farcs25 \times 20''$ field of view, with $0\farcs34 \times 0\farcs30$ spatial pixels. Seven exposures totalling 150 min were taken at slightly dithered positions. Three sky exposures totalling 30 min were taken, following a standard OSO pattern. The field of view was oriented with the $20''$ axis lying in the -34° east of north direction. The seven exposures were combined as described in [120] using a resampled pixel scale of $0\farcs17 \times 0\farcs15$.

The atmospheric seeing during the November 2018 small slicer observations was $0\farcs60$, as measured by the differential image motion monitor at the nearby Canada–France–Hawaii Telescope, so the spatial resolution is seeing-limited. This configuration has high spectral resolution, with a line spread function which is well approximated as Gaussian with a FWHM of 1.34Å in our spectral window. As discussed in [120], the KCWI line spread function can be better described as the convolution of a top-hat function and a Gaussian function. For the small slicer configuration the top-hat component has a width of 1.22Å and the Gaussian component has $\sigma = 0.421$Å. The spectral axis is sampled at 0.5Å intervals.

The large slicer configuration has a field of view which is four times larger, $33'' \times 20''$ with $1\farcs38 \times 0.30$ spatial pixels. We use a mosaic of 30 exposures at 10 locations totalling 9.9 hours to cover the central $\sim 50''$ ($\sim$56kpc). The seeing measurements for the four nights where the large slicer was used were similar, between $0\farcs57$ and $0\farcs70$. The line-spread function for this configuration is distinctly non-Gaussian with a top-hat component with a width of 2.55Å and a Gaussian component with $\sigma = 0.62$Å. The spectral axis is sampled at 1.0Å intervals.

For both configurations, we use the BL grating with a central wavelength of 4600Å to obtain spectra between 3600Å and 5600Å.

6.3.2 KCWI observations of IC1101

We observed IC1101 using Keck KCWI during June 2019 and May 2022. The June 2019 observations used the KCWI small slicer while the May 2022 observations used the large slicer. We used an identical configuration to our H15 observations with the BL grating and a central wavelength of 4600Å.

Eight object exposures totalling 160 min and four sky exposures totalling 40 min were obtained using the small slicer. The object exposures were centered at the center of IC1101 at slightly dithered locations with the 20" axis of the $8\farcs25 \times 20''$ small slicer field of view was aligned with the +20° photometric major axis. The DIMM seeing estimate from the nearby CFHT was $0\farcs39$ during these observations.

Exposures using the large slicer were obtained at six positions spanning a mosaic field of view of $120'' \times 33''$ with the long axis of the mosaic field aligned with the $+20°$ photometric major axis. These object exposures totalled 240 min. Five sky exposures were interspersed, totalling 44 min. These observations were made on two consecutive nights, with positions to the northeast of the center observed on the first night and positions to the southwest observed on the second. The mean seeing was $0''.65$ during the first night of observations and $0''.60$ during the second.

6.3.3 Data reduction and kinematic measurement

Data Reduction Pipeline

The numerous KCWI exposures were reduced using the `IDL` version of the KCWI Data Extraction and Reduction Pipeline [142]. This pipeline performs overscane and bias removal, cosmic ray rejection, dark and scattered light subtraction, solving for the geometric distortion and wavelength solution, flat-fielding, correction for vignetting and the illumination pattern, sky subtraction, and the generation of three-dimensional datacubes. Finally, the cubes are corrected for differential atmospheric correction and flux calibrated by use of a standard star.

We perform the same additional reduction stages as described in [120]. First, a small number of cosmic rays are improperly corrected by the KCWI pipeline. We correct for these by scanning through each wavelength slice of the datacubes, identifying the sharp peaks or dips resulting from this mis-correction, masking those regions, and replacing the erroneous flux in those spaxes with an interpolation from the unimpacted neighboring spaxels. Second, we the standard sky subtraction routine in the KCWI pipeline does not fully eliminate highly time- or space-variant features in the sky spectrum, which may be subtly different between the object exposures and the adjacent offset sky exposures. This is corrected for by constructing a principal component analysis of all available sky spectra and including the ten most significant components as additional additive components in fits to the spectra (See Sec 6.4.1). Finally, we perform a mosaicing of the numerous science frames which cover different regions of each galaxy using the `nifcube` and `gemcube` IRAF tasks which are part of Gemini's data reduction software. The resulting spaxels lie on an $0''.3 \times 1''.4 \times 1$Å grid for the large slicer exposures and an $0''.15 \times 0''.17 \times 0.5$Å grid for the small slicer exposures.

Spectral and Spatial Masking

We follow similar spectral and spatial masking procedures as in Chapter 5 [120].

In Chapter 5, we discussed the use of a PCA decomposition of the sky spectra to correct for residual sky contamination while performing spectral fitting (see also the following Sec. 6.4.1). We construct the decomposition between 3800 and 5650Å using the weighted expectation-maximization PCA [4] routine, excluding three strong and highly variable sky lines (4861 Å Hβ, 5200 Å [N I], and 5577 Å [O I]) as they tend to dominate the decompo-

sition. Accordingly, we also mask those regions in our spectral fits to the galaxy spectra for both H15 and IC1101.

For H15, we observe emission in the central $\sim 3''$. These emission features are not captured by the stellar templates in our spectral fitting, so we mask four regions in the small slicer data for H15: Hγ (4340Å rest wavelength), Hβ (4861Å), [O III] (5007Å), and [N I] (5200Å). For spectra without apparent emission lines, we find that the inferred kinematic moments are negligibly impacted by masking these regions, so we apply this masking to all small slicer H15 spectra for self-consistency. For the large-slicer H15 spectra and all IC1101 spectra, we do not apply any spectral masks beyond the regions excluded from the PCA decomposition.

For both H15 and IC1101, there are several regions within our mosaics which are contaminated by spatially compact foreground objects. To mask these regions, we collapse the data cubes spectrally, locate regions with substantially higher flux than their surroundings, then iteratively masking the brightest pixels in the region until the contaminant is fully removed.

Spatial Binning

We co-add neighboring pixels using voronoi binning to achieve uniform signal to noise in each spatial bin. We estimate the S/N while binning by the routine described in [120]. The spectrum is smoothed with a Gaussian kernel with width comparable to the line spread function, then the noise is taken to be the root-mean-square difference between the original and smoothed spectra and the signal is taken to be the median flux from the original spectrum. This binning is performed separately for the large and small slicer regions. For both galaxies, the region observed with the large slicer configuration overlaps with the central region observed with the small slicer. We use the bins in the overlapping region to test for consistency between the two datasets but exclude the central large slicer bins which overlap the small slicer region in the discussion, analysis, and modelling described in the following sections.

For H15, we bin to achieve a target signal to noise of 130 per spectral pixel for the small slicer region and 60 per spectral pixel in the large slicer region, resulting in 97 and 216 bins in the small slicer and non-overlapping large slicer regions. For IC1101, we bin to 100 and 60 per spectral pixel for the small and large slicer regions, resulting in 157 and 101 bins.

6.4 Stellar Kinematics of H15 and IC1101

6.4.1 Stellar Kinematic Fitting

We use pPXF [26] to determine the line-of-sight velocity distributions (LOSVDs) associated with the stars in each spatial bin. We use similar procedures for both H15 and IC1101 and for the large and small slicer data. The pPXF model convolves a linear combination of stellar spectra with an LOSVD. To account for small mismatches in the shape of the continuum, we use a multiplicative polynomial of order 7. Finally, to accommodate slight residual sky features which remain after the coarse sky subtraction described in Section 6.3.3, we include

the first ten most significant PCA sky components as additional additive terms whose weights are chosen during fitting. The LOSVD is characterized as a Gauss-Hermite series of order 8 and the coefficients of that series (V, σ,h_3,...,h_8) are parameters in the fit. The full number of parameters in each pPXF fit is 8 (LOSVD coefficients) + 7 (multiplicative polynomial) + 10 (PCA sky components) + $N_{\text{templates}}$ (Number of stellar templates).

We use the Indo-US stellar template library [196]. Similar to as discussed in [120], we exclude templates without well-identified spectral types in Table 3 of Valdes et al. [196]. This exclusion reduces the library from 1273 templates to 671. For each galaxy we choose a subset of these 671 templates to use for fitting the numerous binned spectra. The subset is found by first coadding all unmasked spectra associated with each distinct observing run, then fitting those spectra allowing all 671 templates. Each of these fits places positive weight on only a small number, typically $\sim 15-20$, of the full set of templates for each galaxy. We choose the subset of the full library to use from the union of those selected by these individual fits. For H15 we use a set of 22 templates and for IC1101 we use 25 templates. To match the instrumental spectral resolution of the stellar templates and our observations, we convolve each template spectrum with a top-hat function and Gaussian of proper width. This in effect broadens the Indo-US line spread function to match the large and small-slicer LSFs described in Section 6.3.1.

For each binned spectrum, we perform an initial pPXF fit to find the optimal LOSVD coefficients, template weights, polynomial coefficients, and sky component weights. Overall the fits are quite good with fit residuals comparable to the noise in the spectrum. We determine the uncertainties on the LOSVD coefficients by perturbing the spectrum by a series of Gaussian-distributed random numbers with dispersion set by the rms residual from the initial fit. We then re-fit that perturbed spectrum, keeping all fit parameters aside from the the LOSVD coefficients fixed to their values from the initial fit. We perform 1000 perturbed fits and take the mean and standard deviation of each LOSVD to be the accepted central value and uncertainty.

6.4.2 H15 Stellar Kinematic Features

Only a few prior measurements of H15's stellar kinematics have been made. Most significantly, [138] used the Multi Unit Spectroscopic Explorer (MUSE) on the Very Large Telescope to observe the central 60"×60" region. After spatial masking, their coverage is a roughly octagonal region with outermost kinematic bins lying 28″ from the center of H15. They fit the spectra in this region with both pPXF and a non-parametric scheme. [67] performed slit spectroscopy along a PA of 157° (23° off of the photometric major axis) and performed Fourier fitting to infer the stellar velocity and velocity dispersion in five bins out to 4.6". [69] and [68] using the Sydney-AAO Multi-object IFS (SAMI) to examine the velocity and dispersion of the central 15"×15" of a number of ETGs in Abell 85, including H15 (there denoted as J004150.46-091811.2).

The first eight Gauss-Hermite moments of our H15 LOSVDs are shown in Figure 6.3. For the central small-slicer bins, we find average errors on V, σ, and h_3–h_8 of 3.5 km s^{-1},

3.5 km s^{-1}, and 0.0095, respectively. For the large-slicer bins, the average errors on V, σ, and h_3–h_8 are 5.2 km s^{-1}, 6.9 km s^{-1}, and 0.023, respectively.

The stellar velocity dispersion profile is shown in Fig 6.4. H15 exhibits a centrally peaked dispersion, with $\sigma \sim 330$ km s^{-1} in our central kinematic bins. The dispersion quickly drops from this peak to be roughly constant at 290 km s^{-1} beyond 4 kpc. Beyond 15 kpc, the profile begins to rise, reaching 350–375 km s^{-1} in our outer data points at ~ 40 kpc.

In the central region, we find the Gauss-Hermite moments to be near or consistent with 0, suggesting that the LOSVDs in that region are nearly Gaussian. Mehrgan et al. [138] observed H15 with MUSE on the VLT and performed parametric and non-parametric fits to the spectra. They found both higher central σ than our measurements as well as higher h_4 (see Fig 6.4). For an LOSVD characterized by only σ and h_4, when $h_4 \lesssim 0.1$ the full-width at half-maximum of the LOSVD is well-approximated by FWHM $\approx 2.355\sigma(1 - 1.28h_4)$. That is, an LOSVD with $h_4 \sim 0.1$ is about 12.8% narrower than an LOSVD with the same σ but $h_4 = 0$. Because of this, we find that the LOSVDs reported both here and in Mehrgan et al. [138] have comparable widths, but very different shapes – LOSVDs with higher h_4 have both sharper central peaks and wider wings extending to high velocities.

The velocity field of H15 exhibits rotation beyond 20". To quantify the rotation, we follow the scheme described in Section 3.1 of [120] where we model the velocity field with $V(R, \Theta) = V_1(R)\cos[\Theta - \Theta_0(R)]$. We perform fits to find $V_1(R)$ and $\Theta_0(R)$ for groups of bins lying at different distances R from the center of the field. 10-20 spatial bins are included in each fit and we perform 20 such fits.

From 20" to our outermost kinematic bin at 44", the kinematic PA (defined as the PA where the velocity is maximized) of this rotation is stable at $(28 \pm 7)^\circ$. This is misaligned with both the photometric major (-34°) and minor ($+56^\circ$) axes, suggesting an intrinsically non-axisymmetric shape as axisymmetric models cannot produce kinematic misalignments. The amplitude of this rotation rises to to the outermost bins to amplitude of approximately 25 km s^{-1}. Within 20", there is no obvious sense of coherent rotation. Fits to the azimuthal velocity profile in this region yield amplitudes of $\lesssim 5$ km s^{-1} with poor constraints on the phase of the rotation. Fogarty et al. [68] describes the rotation at $\sim 8''$ to have a PA of 31.5° but negligible amplitude. The data presented in Mehrgan et al. [138] also has a faint indication of rotation with a similar orientation as reported here, though the apparent amplitude of that rotation was not much larger than the uncertainty in that data.

6.4.3 IC1101 Kinematic Features

The stellar kinematics of IC1101 have previously only been measured through long-slit observations and those few measurements have only explored the velocity and dispersion along those slits. Dressler performed used a slit with a PA of -135° and found a stellar dispersion profile which rises with distance from the galaxy's center [45]. They find a central dispersion of ~ 375 km s^{-1} which rises to ~ 500 km s^{-1} $\sim 50''$ from the center in a relatively continuous manner. Fisher et al. seemingly confirmed this result, finding a dispersion profile which rose slightly from 362 km s^{-1} at the center to 486 km s^{-1} 11" along a slit with a PA of 31°

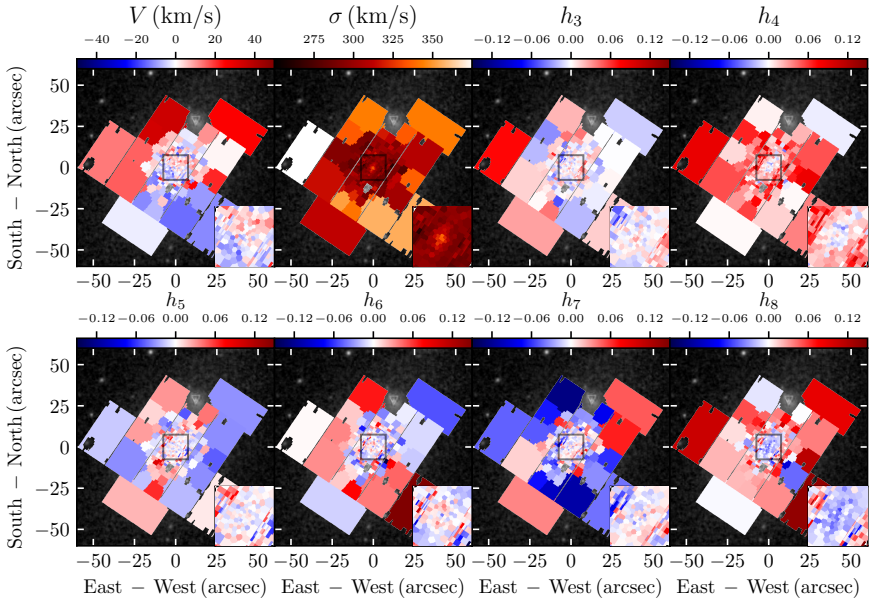

Figure 6.3: Maps of the stellar kinematics of H15 as measured from Keck KCWI spectra in 313 spatial bins. The inner $20'' \times 8''$ region is sampled using the small slicer mode of KCWI to achieve high spatial resolution. Each map shows one of the first eight Gauss-Hermite moments of the line-of-sight stellar velocity distribution. The fields are oriented so that north is up and east is left. The background image is an SDSS image of the region surrounding H15. The mosiac is aligned with an axis with a position angle of $-34°$, the photometric major axis of H15.

[67]. Loubser et al. did not find a rising profile as their data only extended to $\sim 5''$, below the scale where the rise was previously observed [125].

Our dispersion measurements exhibit qualitatively the same features as these previous observations. We find a central dispersion of $\sim$ 380 km s^{-1}, which drops gradually to a minimum of $\sim$ 350 km s^{-1} at 3" from the center. Beyond this, the dispersion rapidly and continuously rises to $\sim$ 500–600 km s^{-1} with large noise in our outermost data beyond 15". Similar to with H15, we find $h_4 \approx 0$ in our central region. For IC1101, h_4 rises rapidly outside the central $\sim 3''$ to lie at roughly 0.08 over most of our data.

Ouside the central region, the velocity field does not show a coherent sense of rotation. Within the central 5-10", the velocity field has a $\sim$ 10 km s^{-1} rotation which which is approximately aligned with the photometric long axis (see inset of 'V' panel of Fig. 6.5).

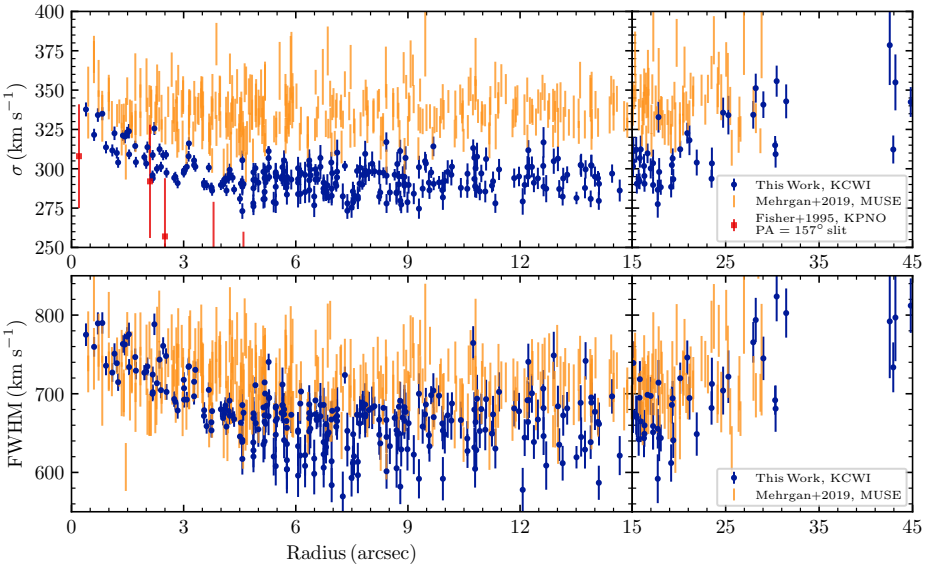

Figure 6.4: **(Top)** Velocity dispersion (σ) measurements from this work (Blue), Mehrgan et al. [138] (orange), and Fisher et al. [67] (red). The points from Mehrgan et al. are the Non-parametric data points shown in their Figure 18. The σ measurements from Mehrgan et al. lie systematically above the bulk of our data points between 3" and 15". **(Bottom)** The full-width at half-maximum of our new KCWI LOSVDs and from those in Mehrgan et al. Figure 18 as calculated from σ and h_4. The h_4 values measured in Mehrgan et al. tend to be larger than those measured in this work. The inferred FWHM from both datasets still differ, but with less tension than σ or h_4 alone, suggesting that the widths of LOSVDs coarsely agree but their shapes do not.

6.5 Triaxial Schwarzschild Orbit Modelling

We use the TriOS code [156, 158] to construct triaxial Schwarzschild orbit superposition models of H15 and IC1101. Models in TriOS consist of a central black hole, a stationary stellar component parameterized by the deprojection of a two-dimensional MGE, and a dark matter halo. A pair of libraries of orbits spanning the phase space are integrated and their contributions to the stellar LOSVDs and light distribution are calculated. Finally, the orbits in the libraries are superposed with non-negative weights to reproduce the observed stellar kinematics and light distribution.

The two orbit libraries in TriOS are the x–z and stationary start spaces. For the x–z start space which consists primarily of loop orbits, we integrate the orbits for 2000 dynamical times and construct orbits with 120 energies, 54 rays in the x–z plane, and 27 positions along those rays. A retrograde orbit library is constructed by time-reversing the orbits in the prograde x–z library. For the stationary start space which consists primarily of box orbits, we integrate

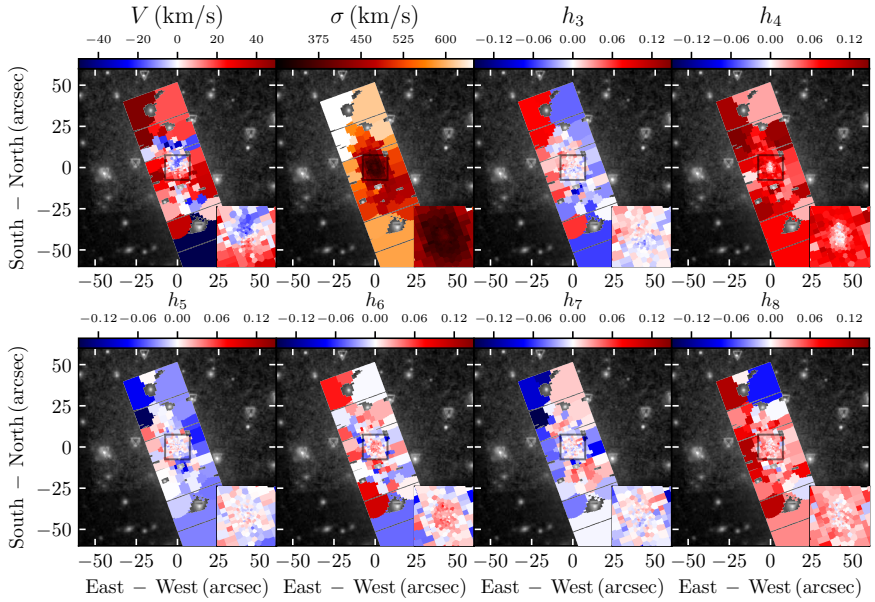

Figure 6.5: Similar to Fig. 6.3, but for IC1101. At the scale of these maps, the light profile of IC1101 is very flattened, with $\epsilon \sim 0.5$. The mosiac is aligned with an axis with a position angle of 20°, the photometric major axis of IC1101. The inset in the velocity map shows a slight rotation aligned with this photometric axis, but that rotation does not continue coherently to larger radii.

the orbits for 200 dynamical times and initialize them with 120 energies along a 27×27 array of positions on the equipotential. We bundle groups of $3^3 = 27$ neighboring orbits in the phase space together and average their observable properties to enforce smoothness in the resulting distribution function and to reduce the size of the weight optimization problem. Additionally, we perform a series of mirroring operations to produce eight versions of each orbit which lie at the same location in phase space but which have different initial conditions [158]. In total, our models contain ∼440,000 orbits whose superposition is described by ∼15,000 independent weights.

Deprojection of the two-dimensional MGE is described by a single parameter (the inclination) for axisymmetric deprojections or three parameters for triaxial deprojections. For triaxial deprojections, a triplet of angles (θ, ϕ, ψ) associated with the viewing direction from the galaxy to us and with the orientation of the galaxy on the sky are used to specify the deprojection. Alternatively, a triplet of shape parameters $(T, T_{\rm maj}, T_{\rm min})$ can be used [158]. This triplet of shape parameters is convenient to use while searching the parameter-space as it maps all valid deprojections into a unit cube and minimizes degeneracies between the three parameters. In this paper, we'll use a mixture of these parameters (T,θ, ϕ) for intuitive simplicity.

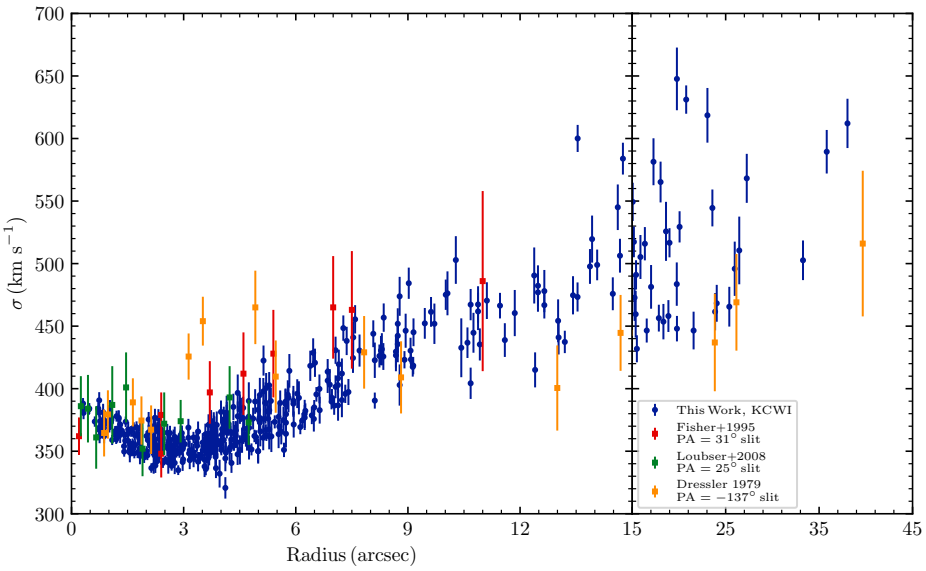

Figure 6.6: Velocity dispersion profile of IC1101 from our KCWI observations with comparisons to prior measurements. We find qualitative agreement with each prior measurement of the dispersion of IC1101. The values from Dressler's 1979 observations (orange; [45]) bracket ours coarsely tracing the same shape. Fisher' 1995 observations (red; [67]) lie slightly above the bulk of our points between 3" and 12" but the large errorbars in those observations provides consistence with our observations.

6.5.1 Parameter search and model selection

The parameter-space associated with our model selection searches has six dimensions: ($M_{\rm BH}$, M^*/L, ρ_0, T, θ, ϕ). The M^*/L is in the filter of the photometric observations–*r_G0326* for H15. The dark matter halo is parameterized by a generalized Navarro-Frenk-White profile [146] with

$$\rho(r) = \frac{\rho_0}{(r/r_s)^\gamma (1 + r/r_s)^{3-\gamma}}, \tag{6.1}$$

where ρ_0 is the density scale factor, r_s is the scale radius, and γ is the central slope. We choose a $\gamma = 0$ model which has finite central density. For this form, the halo's density changes only very gradually within r_s, dropping to half the central density at $r \approx 0.25 r_s$ and an eighth the central density by $r = r_s$. We find that our models are not able to constrain both halo parameters independently, so we fix r_s to values roughly preferred in coarse searches over a seven-dimensional parameter-space (the six parameters enumerated above, plus r_s). For H15 we fix $r_s = 150$ kpc.

To find the preferred models for each galaxy, we follow the procedure described in [120]. First, the parameter-space is broadly searched by points drawn from a Latin hypercube sampling. Models at these points are evaluated, along with versions of those models where

the velocities have been scaled up and down by small factors, corresponding to small multiplicative shifts in the three mass parameters ($M_{\rm BH}$, M^*/L, and ρ_0). After these models have been evaluated, we construct 12 random subsets of the models, then use Gaussian process regression to construct a smooth interpolation of the χ^2 values associated with those subsets. We then draw new points from the parameter-space, but only accept those which have $\Delta\chi^2 < \Delta\chi^2_{\rm cutoff}$ from at least one of the 12 regression surfaces. Initially we use a large cutoff $\Delta\chi^2_{\rm cutoff}$ to search broadly, then reduce the cutoff with subsequent iterations.

To perform parameter inference, we perform similar Gaussian process regression on the χ^2 values which lie within $\Delta\chi^2 < 100$ of the global minimum. We use a Matérn $\nu = 1/2$ kernel for the regression. This χ^2 surface is passed to the `dynesty` dynamic nested sampling code as a log-likelihood, and we use that code to infer parameter posteriors given uniform priors on each parameter. The final posterior distribution for the model search for H15 is shown in Fig. 6.7.

6.6 Results and Discussion

6.6.1 Best-fit Triaxial Models of H15

Our sampling of the six-dimensional parameter-space ($M_{\rm BH}$, M^*/L, ρ_0, T, θ, ϕ) for H15 consists of approximately 23,000 models. As described in Sec. 6.5.1, to infer parameter values from these discrete model points, we model the log-likelihood surface associated with the roughly 2,000 of the best-fit models in this space using Gaussian process regression and sample that continuous log-likelihood surface using `dynesty`, resulting in the posterior distribution shown in Fig. 6.7. Here we discuss these results.

First, we find a central SMBH with $M_{\rm BH} = 21.7^{+1.8}_{-1.9} \times 10^9 M_\odot$. This value is substantially smaller than the prior measurement ($40 \pm 8 \times 10^9 M_\odot$) using MUSE data and axisymmetric Schwarzschild modelling from Mehrgan et al. [138]. Their value was inferred by splitting their kinematic observations into four quadrants delineated by the photometric major and minor axes and fitting their models to the four quadrants independently. Their adopted value and uncertainty comes from the mean and standard deviation across the locations of the four best-fit models at (3, 4, 4, and 5)$\times 10^{10} M_\odot$ (see Fig. 9 of [138]).

The enclosed mass, local density, and enclosed mass fraction for the best-fit models and the models of Mehrgan et al. [138] are shown in Fig. 6.8. Notably, the enclosed stellar mass profiles for our preferred models very closely align with those of the prior work, lying within the error band from their Figure 15 everywhere within 15 kpc and differing by only $\sim$10% within 100 kpc. On the other hand, the preferred halo in our modelling contributes considerably less mass within the central 10 kpc than that of Mehrgan et al. Within the central 10 kpc, the enclosed dark matter mass in our modelling is $5 \times 10^{10} M_\odot$, while the models from Mehrgan et al. contain $\sim 10^{11} M_\odot$ within the same radius. Accordingly, to produce a comparable dark matter halo mass, the scale density would have to double. As shown in our modelling posterior (Fig. 6.7), the upper 99.7% credible region extends only

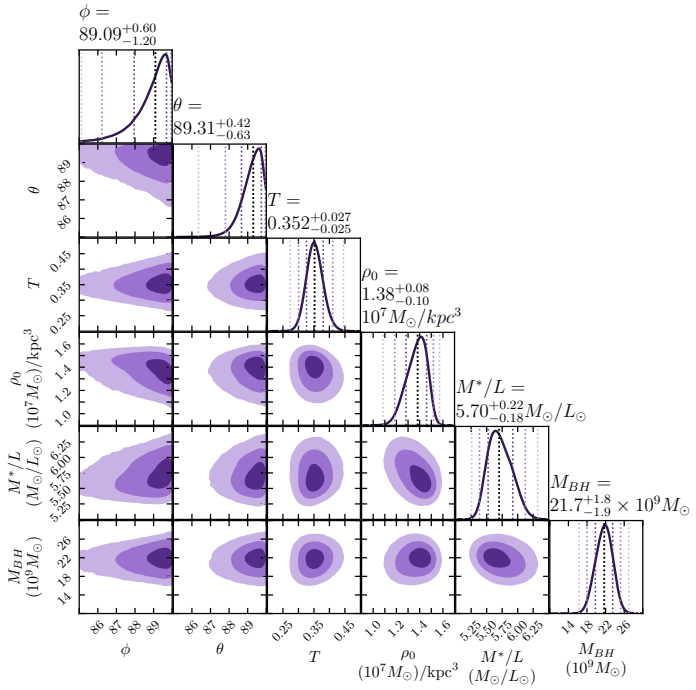

Figure 6.7: Posterior distributions of the six parameters in our triaxial Schwarzschild models of H15. The six parameters explored are the central black hole mass $M_{\rm BH}$, the *r_G0326*-band stellar mass-to-light ratio M^*/L, the dark matter halo scale density ρ_0, the triaxiality parameter T, and the viewing angles θ and ϕ. The three dashed and shaded regions mark the 68%, 95%, and 99.7% credible regions in 1- and 2-D. The central value reported for each parameter represents the 50th percentile, while the stated upper and lower uncertainty bound the central 68% of the posterior mass.

up to about $\rho_0 = 1.6 \times 10^7 M_\odot/\mathrm{kpc}^3$, about about 15% above the posterior median value. Similarly, the covariance between the scale density and stellar mass-to-light ratio in our posterior traces $(M^*/L)/(M_\odot/L_\odot) \sim 9.1 - 2.5(\rho_0/(10^7 M_\odot/\mathrm{kpc}^3))$. For a halo scale density which produces a comparable halo mass to Mehrgan et al., this suggests $M^*/L \sim 2.1 M_\odot/L_\odot$, which is unrealistically small for this band.

The preferred viewing angles (θ,ϕ) are both near their upper limits, corresponding to a viewing direction near H15's intermediate axis, and our modelling does not clearly determine an upper limit for those parameters aside from $\theta \leq 90°$ and $\phi \leq 90°$ as required by geometrical constraints. This arrangement of viewing angles maximizes the length of the short axis c and minimizes the length of the long axis a. The relative length of intermediate axis between these two values is set by the remaining shape parameter T (often referred to as the triaxiality parameter), which prefers an intermediate triaxiality of $0.352^{+0.027}_{-0.025}$. A

triaxiality parameter value of $T = 0$ corresponds to oblate axisymmetry where the long and intermediate axes have the same length while $T = 1$ corresponds to prolate axisymmetry where the short and intermediate axes have the same length. Intermediate values correspond to triaxial shapes where the intermediate axis has a length between the other two axis lengths. The preferred triaxiality here strongly excludes axisymmetry. This preference is consistent with the kinematic misalignment described in Sec. 6.4.2. Due to the symmetry of their constituent orbits, axisymmetric models cannot exhibit a kinematic misalignment where the apparent axis of rotation in the projected velocity field is misaligned with the principal axes of the projected light distribution.

To test whether the lack of an upper bound on the viewing angles is likely to impact the inferred mass parameters, we've repeated our parameter inference using models whose viewing angles lie at various distances from the intermediate axis. We find that the inferred mass parameters are broadly consistent even when the viewing orientation is fixed to be up to 15 degrees from the preferred location. For these models, the preferred black hole mass shifts downward from the global best-fit value by only about 1σ to $19.5 \times 10^9 M_\odot$.

6.6.2 Scaling relations

The luminosity averaged dispersion across our H15 observations is 306 ± 1 km s^{-1}. Using the $M_{\rm BH}$–σ relation for early type galaxies from McConnell and Ma [133], this suggests $\log_{10}(M_{\rm BH}/M_\odot) \sim 9.35$ or $M_{\rm BH} \sim 2.25 \times 10^9 M_\odot$. This is an order of magnitude smaller than the value found in our dynamical models described in the previous section ($\log_{10} M_{\rm BH}/M_\odot \sim 10.34$). The intrinsic scatter in that relation is estimated to be 0.34 dex, so this measurement indeed lies far above the relation.

Our modelling of H15 suggests that its SMBH is among most massive BH dynamically measured to date. The prior measurement of H15* in Mehrgan et al. [138] placed its mass at nearly twice that of the next largest SMBH (NGC4889 with $M_{\rm BH} = (21 \pm 15) \times 10^9 M_\odot$; [135]). Our revised measurement lies near NGC4889's central value though with substantially smaller errorbars which allow it to anchor the upper end of the $M_{\rm BH}$–σ relation. The $M_\odot$–σ relation is thought to 'saturate' in the high-mass regime as dry major mergers simply combine the black hole masses but gradually raise velocity dispersion [112]. The overmassive SMBH in H15 supports this picture.

In light of the systematically overmassive black holes in these systems, potential sources of bias in modelling must be explored. Li et al. examines a sample of of massive galaxies within the Illustris simulation, finding that for galaxies in their sample some degree of triaxiality is common [117]. Especially for the most massive systems in their sample ($M_* > 10^{12} M_\odot$), prolate system where $p - q < 0.2$ are common. The stellar halo in our preferred model has $M_* = 10^{12.3} M_\odot$. Beyond the scope of the MGE fit and deprojection, the stellar density will tend to be underestimated in our modelling, so this should be understood as an underestimate of H15's true stellar mass. Similarly, Ene et al. found that for slowly rotating galxies within the MASSIVE survey sample, some degree of kinematic misalignment is common, suggesting non-axisymmetric intrinsic shapes [54]. As described in Sec. 6.2.1, we find such

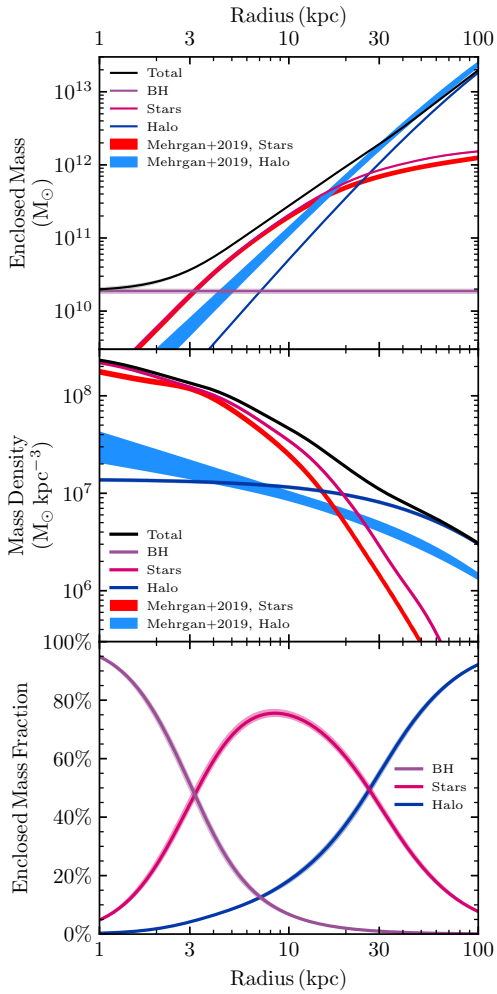

Figure 6.8: Enclosed mass, local density, and enclosed mass fraction for H15. For the curves from this work, we draw mass models from the posterior shown in Fig. 6.7, then mark the median and 68% credible regions with a solid curve and shaded region, respectively. For the stellar mass and dark matter mass components, the shaded region is comparable to the width of the solid line. The stellar mass and halo components from Figure 15 of Mehrgan et al. [138] are reproduced here (red and blue shaded regions).

a misalignment for H15. Of the six SMBH's with $M_{\rm BH} > 5 \times 10^9 M_\odot$ whose masses have been measured through stellar dynamical observations, only one (NGC4486, [120]) has been determined in models which include triaxial symmetry. Only a small handful of galaxies have now been explored with both axisymmetric and triaxial Schwarzschild modelling, sometimes resulting in very different black hole masses: M32 and NGC1453's masses were consistent with the two methods [157, 199]; NGC4486, PGC 046832, and NGC2693 [120, 42, 153] have smaller masses when triaxiality is allowed; and NGC3379 has a larger mass with triaxiality [199]. This work finds that the inferred mass of H15's central SMBH while using triaxial models is nearly halved compared to the previous measurement which assumed axisymmetry. Of course, there are a multitude of other differences between our data and modelling and that of Mehrgan et al., so further analysis will be necessary to determine whether these differences are due to differences in modelling, observations, or both.

As shown in Fig 6.8, the black hole's sphere of influence, defined to be the radius where the enclosed stellar mass matches the central SMBH mass, is $r_{\rm SOI} = 3.19 \pm 0.13$ kpc or $2.96 \pm 0.12''$. The $r_{\rm SOI}$–r_γ and $r_{\rm SOI}$–r_b relations from Thomas et al. [192] relate this sphere of influence to the Nuker cusp radius and break radius from a sample of galaxies with 15 core galaxies with dynamically measured black hole masses. López-Cruz et al. performed Nuker fits to KPNO and CFHT images of H15, finding $r_\gamma = 4.57$ kpc for both fits with $r_b = 18.48$ and 20.50, respectively. Madrid and Donzelli [128] performed a Nuker fits to the Gemini image used in this Chapter (see Section 6.2.1), finding $r_\gamma = 5.02$ kpc and $r_b = 11.56$ kpc.

These relations slightly over-estimate the size of the sphere of influence of H15's SMBH. The two cusp radii noted above predict $r_{\rm SOI}$ to be 4.79 kpc and 5.23 kpc (for $r_\gamma = 4.57$ kpc and 5.02 kpc, respectively). These over-estimates are 1.1 and 1.3 times the intrinsic scatter in the relation. Similarly, the three break radii predict $r_{\rm SOI} = 12.2$, 13.5, and 7.6 kpc (for $r_b = 18.48$, 20.50, and 11.56 kpc), which lie above our measured value by 3.6, 3.0, and 2.4 times the intrinsic scatter. Notably, the relations in Thomas et al. [192] were fit to a sample with $0.05 \lesssim r_{\rm SOI}/{\rm kpc} \lesssim 1$, so the exceptionally large sphere of influence of H15 lies nearly half a dex beyond the sample. Given this, it's perhaps surprising that the relations (especially $r_{\rm SOI}$–r_γ) continue to coarsely hold under such an extrapolation.

Interestingly, the $M_{\rm BH}$–r_b and $M_{\rm BH}$–r_γ relations for Nuker fits substantially over-estimate the SMBH mass for H15, estimating $\log_{10} M_{\rm BH} = 11.16$–$11.70$ for the Nuker fits from [123] and [128], while our dynamical modelling suggests a value which is an order of magnitude smaller at $\log_{10} M_{\rm BH} = 10.34$.

6.6.3 Implications for IC1101

Dynamical modelling of IC1101 is ongoing and the results of that modelling will be reported in future work, but we can apply the scaling relation behaviors from H15 to estimate reasonable parameters for IC1101.

Dullo et al. [47] found the core-Sérsic core radius of IC1101 to be 4.2 kpc. The core-Sérsic $r_{\rm SOI} - r_b$ relation from Thomas et al. [192] suggests $r_{\rm SOI} \sim 3.8$ kpc. If the overestimation in this relation is similar to that of the cusp radius relation when applied to H15 we would ex-

pect $r_{\rm SOI} \sim 2.3$–2.5 kpc for IC1101. The isophotes in this central region are relatively circular, so the enclosed light within these radii in deprojections of the MGE parameterization of the light profile depend only slightly on the assumed viewing angles. The enclosed light within 2.3-2.5 kpc in an edge-on axisymmetric deprojection of our MGE is 3.8–$4.6 \times 10^9 L_{\rm F702W,\odot}$. For $M_\odot/L_{\rm F702W,\odot} = 5$–$7$, this in turn suggests $M_{\rm BH} \sim 19$–$32 \times 10^9 M_\odot$. This is substantially smaller than the mass one would infer from the $M_{\rm BH}$–r_b relation, $\sim 10^{11} M_\odot$.

6.7 Conclusions

In this chapter we've used the Keck Cosmic Web Imager to construct detailed maps of the stellar kinematics of the brightest cluster galaxies H15 and IC1101. These two galaxies have exceptionally large and diffuse central regions, which have been speculated to harbor exceptionally massive central supermassive black holes. The stellar kinematic maps for both galaxies extend to roughly 50 kpc from the galaxy center. In the central regions, the small slicer configuration of KCWI is used to achieve a high spatial resolution at the scale of the black holes' spheres of influence.

For both galaxies, we find intriguing kinematic features. H15's velocity field has a substantial kinematic misalignement, where beyond 20" the kinematic axis which describes the sense of net rotation is roughly 28° misaligned with the photometric minor axis, suggesting that H15's intrinsic shape is neither oblately or prolately axisymmetric. The measured velocity dispersion σ in our work is somewhat smaller than that of Mehrgan et al. [138], but when the higher moment h_4 is taken into account the overall widths of the velocity distributions are comparable. However, the shapes of the distributions are very different, potentially resulting in different inferred model parameters.

We find IC1101 to have the same rising velocity dispersion profile which was first described by Dressler [45]. Our data is the first integral field unit exploration of IC1101, linking the small and large scale kinematic behaviors with high spatial resolution and high signal-to-noise. Our observations agree with the prior measurements by Dressler [45], Fisher [67], and Loubser [125]. The rising σ profile in our data extends to our outermost data points, from ~ 350 km s^{-1} at 3" to 500–600 km s^{-1} beyond 15".

We've used the TriOS triaxial orbit superposition code to model the kinematics of H15. We perform over 23,000 models, simultaneously exploring the central supermassive black hole mass, the stellar mass-to-light ratio, the halo scale density, and three parameters which specify the three-dimensional intrinsic shape. We find that the mass of the central supermassive black hole is $21.7^{+1.8}_{-1.9} \times 10^9 M_\odot$. This is roughly half the previous value measured by Mehrgan et al., though we find that models with substantially larger black holes are incompatible with our data. H15 is found to be highly triaxial with the triaxiality parameter $T = 0.352^{+0.027}_{-0.025}$.

The various scaling relations between H15's black hole and other parameters yield a wide range of inferred $M_{\rm BH}$, in part due to uncertainties in prior parameterizations of the light profile. We find that the size of the black hole's sphere of influence is well predicted from

H15's Nuker cusp radius by the relation described in Thomas et al. [192], but the $M_{\rm BH}$–r_γ relation over-estimates the black hole mass by an order of magnitude. Similarly, the $M_{\rm BH}$–σ relation for early-type galaxies from McConnell and Ma [133] underestimates the black hole mass by about an order of magnitude.

Finally, we consider the implications of these over and under-estimates to IC1101, whose data will be modelled in future work. The $M_{\rm BH}$–r_γ and $M_{\rm BH}$–σ relations continue to predict black hole masses which differ by roughly an order of magnitude. However, if IC1101 deviates from the $r_{\rm SOI} - r_b$ relation in a similar manner to H15, we should expect it to have a central supermassive black hole with $M_{\rm BH} = 19 - 32 \times 10^9 M_\odot$, which would be amongst the most massive black holes to be measured through stellar dynamical observation and modelling.

Appendix

6.A Multi-Gaussian Expansion Parameters

The Multi Gaussian Expansion (MGE) fits to the surface brightness profiles described in Sec 6.2 are presented in Tables 6.1 and 6.2. The MGE parameterization is has the form

$$\Sigma(R', \theta') = \sum_{j=1}^{N} \frac{L_j}{2\pi\sigma_j^2 q'_j} \exp\left[-\frac{1}{2\sigma_j^2}\left(x_j'^2 + \frac{y_j'^2}{q_j'^2}\right)\right], \tag{6.2}$$

which is presented as Eqn 1 in Cappellari 2002 [24]. R' and θ' are polar coordinates on the sky plane and $x'_j = R' \sin(\theta' - \psi_j)$ and $y'_j = R' \cos(\theta' - \psi_j)$. The MGE consists of N components, each with a corresponding σ_j, q'_j, and ψ_j. We choose a parameterizations where ψ_j is fixed for all components, which keeps a constant photometric position in the fit. We use $\psi = -34°$ for H15 and $\psi = 20°$ for IC1101. For H15, the data is very well fit within 70" with residuals of less than 10% and there is no clear photometric twist in the data. For IC1101, Dullo et al. [47] finds evidence of a slight photometric twist. In their two-dimension fits which included core-Sérsic, Sérsic, and exponential components, they found that the central core-Sérsic spheroid had a position angle of 18.4° and the Sérsic component had 21.1°, while the outer exponential halo was roughly twenty degrees misaligned with a PA of 38°. Indications of this slight twist appear in our fits, where surface brightness isophotes near the edge of the PC1 chip of the WFPC2 image are apparently oriented slightly clockwise of the MGE fit isophotes (see lower left panel of Fig 6.2). Despite this slight twist, fits with the position angle fixed to 20° are still relatively good, with $< 10\%$ residuals within 10", a $\sim 15\%$ light deficit along the major axis and a $\sim 15\%$ light surplus along the minor.

Acknowledgements

E.R.L. and C.-P.M. are supported by NSF AST-1817100 and AST-2206307. J.L.W. is supported by NSF AST-1814799 and AST-2206219. C.-P.M. acknowledges the support of the

$I_k\,[L_\odot/\mathrm{p}c^2]$	σ'_k [arcsec]	q'_k
11.219	0.96	0.955
35.767	0.96	0.911
129.90	2.969	1.000
125.53	6.531	0.735
110.40	11.46	0.792
28.854	24.43	0.639
7.0301	36.89	0.802
1.5427	95.15	1.000

Table 6.1: Best-fit MGE parameters for the surface brightness of H15. For each of the 8 two-dimensional Gaussian components, the first column lists the central surface brightness density, the middle column lists the dispersion of the Gaussian, and the last column lists the axis ratio, where primed variables denote projected quantities.

$I_k\,[L_\odot/\mathrm{p}c^2]$	σ'_k [arcsec]	q'_k
138.96	0.400	1.000
207.73	1.403	1.000
170.61	2.130	0.778
230.89	3.449	0.746
89.652	6.234	0.666
56.690	11.07	0.558
41.694	21.82	0.498
11.985	56.46	0.415
2.6436	87.50	0.936

Table 6.2: Best-fit MGE parameters for the surface brightness of IC1101. For each of the 9 two-dimensional Gaussian components, the first column lists the central surface brightness density, the middle column lists the dispersion of the Gaussian, and the last column lists the axis ratio, where primed variables denote projected quantities.

Heising-Simons Foundation and the Miller Institute for Basic Research in Science. The spectroscopic data presented in this paper were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. This work used observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) at the San Diego Supercomputing Center through allocation AST180041, which is supported by NSF grant ACI-1548562.

Chapter 7

Closing Discussion and Future Directions

This dissertation has broadly explored the development of a number of techniques and strategies for performing Schwarzschild orbit modelling of galaxies with central supermassive black holes and triaxial stellar halos with the goal of simultaneously measuring the mass of the central black hole and the intrinsic three-dimensional triaxial shape.

A number of background topics were discussed in Chapter 1.

Chapters 2 and 3 described our first usage of the TriOS triaxial orbit superposition code in the axisymmetric limit to make the first Schwarzschild modelling measurement of the mass of the black hole in NGC1453 ($2.9 \times 10^9 M_\odot$). To do this, we needed to establish the conditions under which the TriOS code would build robust models near the axisymmetric limit. Along the way we also found a number of issues with previous prescriptions for usage of the code, including, for example, that the line-of-sight velocity distributions of the stars tended to be under-constrained.

Chapter 4 discussed the usage of the TriOS for models with full-fledged triaxiality. We identified a set of shape parameters that conveniently describe the transformation between the intrinsic coordinate system of the galaxy and the projected coordinate system while also describing the relative values of the three intrinsic axis lengths. We also found a number of major issues relevant to robust modelling of triaxial galaxies: we fixed an error in the application of a mirroring symmetry that exists in static triaxial potentials; we improved the orbit sampling density to reduce artifacts from the discrete orbit initial conditions; and we fixed a number of smaller bugs. We applied the improved code again to NGC1453 and found that our measurement of the black hole mass was essentially unchanged as the symmetry was relaxed from axisymmetry to triaxiality.

Chapter 5 presented new measurements of the stellar kinematics of M87 and performed triaxial Schwarzschild modelling of that data with TriOS. We found that the intrinsic shape of M87 is indeed highly triaxial and that the data is fully incompatible with axisymmetry. This is in large part related to our detection of a prominent kinematic misalignment in the velocity field of M87, where the apparent axis of rotation is about halfway between

both principal photometric axes. The mass of M87* as inferred from our data with triaxial TriOS models is somewhat smaller than measured in previous work that used axisymmetric Schwarzschild modelling at $5.37 \times 10^9 M_\odot$.

Chapter 6 expands on the previous chapter by presenting new measurements of the stellar kinematics of Holmberg 15A and IC1101 using observations from the Keck Cosmic Web Imager. These two galaxies have extremely diffuse central cores, suggesting that they should have extremely massive central supermassive black holes. We perform triaxial Schwarzschild modelling of Holmberg 15A using TriOS, finding an intrinsically triaxial shape and a black hole whose mass is only half of that from a prior measurement that used axisymmetric Schwarzschild modelling ($21.7 \times 10^9 M_\odot$).

There are a number of astrophysical motivations which underlaid this work. First, the total number of dynamically measured black hole masses in the regime explored in this work is still quite small – only 6 black hole masses above $5 \times 10^9 M_\odot$ have been measured dynamically. This work adds a new mass measurement just below this threshold (NGC1453, $2.9 \times 10^9 M_\odot$) and revises two prior measurements above this threshold (M87, $5.4 \times 10^9 M_\odot$; Holmberg 15A, $21.7 \times 10^9 M_\odot$).

Beyond simply adding additional counts to the high-mass black hole census, the measurements presented in this work are among the first assuming full-fledged triaxiality in the stellar potential and fully exploring the space of allowed intrinsic shapes while modelling. This is significant for a number of reasons. First, the very massive galaxies that are targeted by both this work and the MASSIVE survey more broadly are likely to be intrinsically triaxial. Though this triaxiality only slightly changes the shape of the galaxy, compressing the intermediate axis of all three galaxies we've explored here by $\lesssim 15\%$, it fundamentally changes the sense of symmetry of the underlying potential. Under triaxial potentials additional orbit families are allowed that don't exist in axisymmetry, including a second family of loop orbits and box orbits that have no net sense of rotation and tend to be highly radially anisotropic near the origin of the potential. The additional nuances associated with triaxial potentials make robust triaxial Schwarzschild modelling technically difficult, as described in Chapter 4.

A primary advantage of Schwarzschild modelling over other simpler techniques is its ability to reproduce the internal orbital structure of the galaxy. A major difficulty in relating the observed motion of the stars and the potential is the fact that only motion along the line of sight is observable, so the three-dimensional motion must be inferred assuming some velocity anisotropy. Schwarzschild modelling avoids explicit assumptions on the velocity anisotropy by explicitly constructing self-consistent three-dimensional velocity distributions from integrated orbits in the potential. However, this advantage may be partially lost when the Schwarzschild models cannot reproduce the true symmetry associated with the galaxy under consideration as the velocity anisotropies may differ significantly and consequentially from the true anisotropy and orbital structure. In particular, since triaxial models include box and box-like orbits which can be centrophilic and highly radial near the central black hole those models may have additional flexibility to be more radially anisotropic near the origin than axisymmetric models. The central anisotropy has long been understood to be strongly

related to the inferred enclosed mass. Famously, the first black hole mass measurement of M87* from stellar kinematics by Sargent et al [170] was disputed by Binney and Mamon who re-fit the data without a black hole by allowing for a highly anisotopic velocity dispersion [10].

A major open question that this dissertation has begun to address is whether the apparently overmassive black hole mass measurements in the highest mass regime ($\gtrsim 5 \times 10^9 M_\odot$) in the various scaling relations (e.g., $M_{\rm BH}$–σ) is in part due to systematic biases associated with assumptions of axisymmetry instead of triaxiality. To present there are 7 cases where both axisymmetry and triaxiality have been used to measure the black hole mass of the same galaxy. Of these, three are presented in this dissertation (NGC1453 in Chapters 2 and 4; M87 in Chapter 5; Holmberg 15A in Chapter 6) and a fourth is presented in adjacent work [153]. For these seven cases, two have consistent masses in the two measurements (M32, [199]; NGC1453, Chapters 2 and 4 [119, 157]), one has a larger mass when triaxiality is assumed (NGC3379, [199]), and four have smaller masses when triaxiality is allowed (PGC 046832, [42]; NGC2693, [153]; M87, Chapter 5 [77, 120]; Holmberg 15A, Chapter 6 [138]). Both of these cases which lie above $5 \times 10^9 M_\odot$ (M87 and Holmberg 15A) have a smaller black hole in work using triaxial models rather than axisymmetry, though the multitude of differences between the prior works which assumed axisymmetry and the new work which assumes triaxiality make these comparisons inexact.

Several of the scaling relations between the central black hole mass and other galaxy properties are expected to have different slopes in the high-massive regime for a number of reasons, including the increased importance of dry mergers in forming the host galaxies [8, 61, 145], shifts in the types of galaxies in the mergers [112] and in the environments surrounding the mergers [16]. In light of this, it's vitally important to ensure that the most massive measured black hole masses are accurate in order to produce the most meaningful constraints on models for evolution and merging of the most massive galaxies in our universe. The work presented in this dissertation lays the groundwork for substantial movement along this axis. In addition to the data presented in this dissertation, we have collected a KCWI observations on a number of other distant galaxies with ultra diffuse cores, which are expected to hold similarly massive black holes to Holmberg 15A. These, along with the observations taken as part of the MASSIVE survey, present us now with both plentiful data and robust tools to greatly expand our current supermassive black hole census.

Bibliography

[1] C. P. Ahn et al. "The Black Hole in the Most Massive Ultracompact Dwarf Galaxy M59-UCD3". In: *Astrophys. J.* 858, 102 (May 2018), p. 102. DOI: 10.3847/1538-4357/aabc57. arXiv: 1804.02399.

[2] Pau Amaro-Seoane et al. "Astrophysics with the Laser Interferometer Space Antenna." In: 2203.06016 (Mar. 2022), Preprint at https://arxiv.org/abs/2203.06016.

[3] F. Annibali et al. "Nearby early-type galaxies with ionized gas. III. Analysis of line-strength indices with new stellar population models". In: *Astron. & Astrophys.* 463 (Feb. 2007), pp. 455–479. DOI: 10.1051/0004-6361:20054726. eprint: arXiv:astro-ph/0609175.

[4] Stephen Bailey. "Principal Component Analysis with Noisy and/or Missing Data". In: *Pub. Astron. Soc. Pacific* 124.919 (Sept. 2012), p. 1015. DOI: 10.1086/668105.

[5] A. J. Barth, L. C. Ho, and W. L. W. Sargent. "A Study of the Direct Fitting Method for Measurement of Galaxy Velocity Dispersions". In: *Astron. J.* 124 (Nov. 2002), pp. 2607–2614. DOI: 10.1086/343840. eprint: astro-ph/0209564.

[6] M. C. Begelman, R. D. Blandford, and M. J. Rees. "Massive black hole binaries in active galactic nuclei". In: *Nature* 287 (Sept. 1980), pp. 307–309. DOI: 10.1038/287307a0.

[7] R. Bender. "Unraveling the kinematics of early-type galaxies. Presentation of a new method and its application to NGC 4621." In: *Astron. & Astrophys.* 229 (Mar. 1990), pp. 441–451.

[8] Ralf Bender, David Burstein, and S. M. Faber. "Dynamically Hot Galaxies. I. Structural Properties". In: *Astrophys. J.* 399 (Nov. 1992), p. 462. DOI: 10.1086/171940.

[9] J. Binney. "On the rotation of elliptical galaxies". In: *Mon. Not. R. Astron. Soc.* 183 (May 1978), pp. 501–514.

[10] J. Binney and G. A. Mamon. "M/L and velocity anisotropy from observations of spherical galaxies, of must M 87 have a massive black hole ?" In: *Mon. Not. R. Astron. Soc.* 200 (July 1982), pp. 361–375. DOI: 10.1093/mnras/200.2.361.

[11] J. Binney and D. Spergel. "Spectral stellar dynamics. II - The action integrals". In: *Mon. Not. R. Astron. Soc.* 206 (Jan. 1984), pp. 159–177. DOI: 10.1093/mnras/206.1.159.

[12] J. J. Binney, Roger L. Davies, and Garth D. Illingworth. "Velocity Mapping and Models of the Elliptical Galaxies NGC 720, NGC 1052, and NGC 4697". In: *Astrophys. J.* 361 (Sept. 1990), p. 78. DOI: 10.1086/169169.

[13] James Binney. "Testing for triaxiality with kinematic data". In: *Mon. Not. R. Astron. Soc.* 212.4 (Feb. 1985), pp. 767–781. ISSN: 0035-8711. DOI: 10.1093/mnras/212.4.767. eprint: http://oup.prod.sis.lan/mnras/article-pdf/212/4/767/3792766/mnras212-0767.pdf. URL: https://doi.org/10.1093/mnras/212.4.767.

[14] Jams Binney and Scott Tremaine. *Galactic Dynamics, 2nd edition), p. 75*. 2008.

[15] Paolo Bonfini, Bililign T. Dullo, and Alister W. Graham. "Too Big to Be Real? No Depleted Core in Holm 15A". In: *Astrophys. J.* 807.2, 136 (July 2015), p. 136. DOI: 10.1088/0004-637X/807/2/136. arXiv: 1506.08560 [astro-ph.GA].

[16] M. Boylan-Kolchin, C.-P. Ma, and E. Quataert. "Dissipationless mergers of elliptical galaxies and the evolution of the fundamental plane". In: *Mon. Not. R. Astron. Soc.* 362 (Sept. 2005), pp. 184–196. DOI: 10.1111/j.1365-2966.2005.09278.x. eprint: astro-ph/0502495.

[17] Avery E. Broderick et al. "The Photon Ring in M87*". In: *Astrophys. J.* 935.1, 61 (Aug. 2022), p. 61. DOI: 10.3847/1538-4357/ac7c1d.

[18] M. Cappellari. "Measuring the inclination and mass-to-light ratio of axisymmetric galaxies via anisotropic Jeans models of stellar kinematics". In: *Mon. Not. R. Astron. Soc.* 390 (Oct. 2008), pp. 71–86. DOI: 10.1111/j.1365-2966.2008.13754.x. arXiv: 0806.0042.

[19] M. Cappellari et al. "The ATLAS^{3D} project - I. A volume-limited sample of 260 nearby early-type galaxies: science goals and selection criteria". In: *Mon. Not. R. Astron. Soc.* 413 (May 2011), pp. 813–836. DOI: 10.1111/j.1365-2966.2010.18174.x. arXiv: 1012.1551 [astro-ph.CO].

[20] M. Cappellari et al. "The SAURON project - IV. The mass-to-light ratio, the virial mass estimator and the Fundamental Plane of elliptical and lenticular galaxies". In: *Mon. Not. R. Astron. Soc.* 366 (Mar. 2006), pp. 1126–1150. DOI: 10.1111/j.1365-2966.2005.09981.x. eprint: arXiv:astro-ph/0505042.

[21] M. Cappellari and Y. Copin. "Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations". In: *Mon. Not. R. Astron. Soc.* 342 (June 2003), pp. 345–354. DOI: 10.1046/j.1365-8711.2003.06541.x. eprint: astro-ph/0302262.

[22] M. Cappellari and E. Emsellem. "Parametric Recovery of Line-of-Sight Velocity Distributions from Absorption-Line Spectra of Galaxies via Penalized Likelihood". In: *Pub. Astron. Soc. Pacific* 116 (Feb. 2004), pp. 138–147. DOI: 10.1086/381875. eprint: arXiv:astro-ph/0312201.

[23] M. Cappellari et al. "The Counterrotating Core and the Black Hole Mass of IC 1459". In: *Astrophys. J.* 578.2 (Oct. 2002), pp. 787–805. DOI: 10.1086/342653. URL: https://doi.org/10.1086%2F342653.

[24] Michele Cappellari. "Efficient multi-Gaussian expansion of galaxies". In: *Mon. Not. R. Astron. Soc.* 333.2 (2002), pp. 400–410. DOI: 10.1046/j.1365-8711.2002.05412.x. eprint: /oup/backfile/content_public/journal/mnras/333/2/10.1046_j.1365-8711.2002.05412.x/2/333-2-400.pdf. URL: http://dx.doi.org/10.1046/j.1365-8711.2002.05412.x.

[25] Michele Cappellari. "Efficient solution of the anisotropic spherically aligned axisymmetric Jeans equations of stellar hydrodynamics for galactic dynamics". In: *Mon. Not. R. Astron. Soc.* 494.4 (June 2020), pp. 4819–4837. DOI: 10.1093/mnras/staa959. arXiv: 1907.09894 [astro-ph.GA].

[26] Michele Cappellari. "Improving the full spectrum fitting method: accurate convolution with Gauss-Hermite functions". In: *Mon. Not. R. Astron. Soc.* 466.1 (Apr. 2017), pp. 798–811. DOI: 10.1093/mnras/stw3020. arXiv: 1607.08538 [astro-ph.GA].

[27] Michele Cappellari. "Improving the full spectrum fitting method: accurate convolution with Gauss–Hermite functions". In: *Mon. Not. R. Astron. Soc.* 466.1 (Nov. 2016), pp. 798–811. ISSN: 0035-8711. DOI: 10.1093/mnras/stw3020. eprint: http://oup.prod.sis.lan/mnras/article-pdf/466/1/798/10865080/stw3020.pdf. URL: https://doi.org/10.1093/mnras/stw3020.

[28] Michele Cappellari et al. "Testing Mass Determinations of Supermassive Black Holes via Stellar Kinematics". In: *AIP Conference Proceedings* 1240.1 (2010), pp. 211–214. DOI: 10.1063/1.3458489. eprint: https://aip.scitation.org/doi/pdf/10.1063/1.3458489. URL: https://aip.scitation.org/doi/abs/10.1063/1.3458489.

[29] Michele Cappellari et al. "The mass of the black hole in Centaurus A from SINFONI AO-assisted integral-field observations of stellar kinematics". In: *Mon. Not. R. Astron. Soc.* 394.2 (2009), pp. 660–674. DOI: 10.1111/j.1365-2966.2008.14377.x. eprint: /oup/backfile/content_public/journal/mnras/394/2/10.1111/j.1365-2966.2008.14377.x/2/mnras0394-0660.pdf. URL: http://dx.doi.org/10.1111/j.1365-2966.2008.14377.x.

[30] A. J. Cenarro et al. "Empirical calibration of the near-infrared Ca ii triplet - I. The stellar library and index definition". In: *Mon. Not. R. Astron. Soc.* 326 (Sept. 2001), pp. 959–980. DOI: 10.1046/j.1365-8711.2001.04688.x. eprint: astro-ph/0109157.

[31] Patrick Côté et al. "Dynamics of the Globular Cluster System Associated with M87 (NGC 4486). II. Analysis". In: *Astrophys. J.* 559.2 (Oct. 2001), pp. 828–850. DOI: 10.1086/322347.

[32] Patrick Côté et al. "The ACS Fornax Cluster Survey. II. The Central Brightness Profiles of Early-Type Galaxies: A Characteristic Radius on Nuclear Scales and the Transition from Central Luminosity Deficit to Excess". In: *Astrophys. J.* 671.2 (Dec. 2007), pp. 1456–1465. DOI: 10.1086/522822. arXiv: 0711.1358 [astro-ph].

[33] N. Cretton et al. "Axisymmetric Three-Integral Models for Galaxies". In: *Astrophys. J. Suppl. Ser.* 124.2 (Oct. 1999), pp. 383–401. DOI: 10.1086/313264. arXiv: astro-ph/9902034 [astro-ph].

[34] Nicolas Cretton and Eric Emsellem. "On the reliability of the black hole mass and mass-to-light ratio determinations with Schwarzschild models". In: *Mon. Not. R. Astron. Soc.* 347.2 (Jan. 2004), pp. L31–L35. DOI: 10.1111/j.1365-2966.2004.07374.x. arXiv: astro-ph/0312307 [astro-ph].

[35] A. C. Crook et al. "Groups of Galaxies in the Two Micron All Sky Redshift Survey". In: *Astrophys. J.* 655 (Feb. 2007), pp. 790–813. DOI: 10.1086/510201. eprint: arXiv:astro-ph/0610732.

[36] T. A. Davis et al. "The MASSIVE survey - III. Molecular gas and a broken Tully-Fisher relation in the most massive early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 455 (Jan. 2016), pp. 214–226. DOI: 10.1093/mnras/stv2313. arXiv: 1510.00729.

[37] T. A. Davis et al. "The MASSIVE survey - XI. What drives the molecular gas properties of early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 486 (June 2019), pp. 1404–1423. DOI: 10.1093/mnras/stz871. arXiv: 1903.08884.

[38] Stefano de Nicola et al. "Non-parametric triaxial deprojection of elliptical galaxies". In: *Mon. Not. R. Astron. Soc.* 496.3 (Aug. 2020), pp. 3076–3100. DOI: 10.1093/mnras/staa1703. arXiv: 2006.05971 [astro-ph.GA].

[39] G. de Vaucouleurs. "Photoelectric photometry of the Andromeda Nebula in the UBV system." In: *Astrophys. J.* 128 (Nov. 1958), p. 465. DOI: 10.1086/146564.

[40] Gerard de Vaucouleurs. "Recherches sur les Nebuleuses Extragalactiques". In: *Annales d'Astrophysique* 11 (Jan. 1948), p. 247.

[41] Tim de Zeeuw and Marijn Franx. "Kinematics of Gas in a Triaxial Galaxy". In: *Astrophys. J.* 343 (Aug. 1989), p. 617. DOI: 10.1086/167735.

[42] Mark den Brok et al. "Dynamical modelling of the twisted galaxy PGC 046832". In: *Mon. Not. R. Astron. Soc.* 508.4 (Dec. 2021), pp. 4786–4805. DOI: 10.1093/mnras/stab2852. arXiv: 2109.14640 [astro-ph.GA].

[43] Jared L. Deutsch and Clayton V. Deutsch. "Latin hypercube sampling with multidimensional uniformity". In: *Journal of Statistical Planning and Inference* 142.3 (2012), pp. 763–772. ISSN: 0378-3758. DOI: https://doi.org/10.1016/j.jspi.2011.09.016. URL: https://www.sciencedirect.com/science/article/pii/S0378375811003776.

[44] Tuan Do et al. "Relativistic redshift of the star S0-2 orbiting the Galactic Center supermassive black hole". In: *Science* 365.6454 (Aug. 2019), pp. 664–668. DOI: 10.1126/science.aav8137. arXiv: 1907.10731 [astro-ph.GA].

[45] A. Dressler. "The dynamics and structure of the cD galaxy in Abell 2029." In: *Astrophys. J.* 231 (Aug. 1979), pp. 659–670. DOI: 10.1086/157229.

[46] Alan Dressler and Douglas O. Richstone. "New Measurements of Stellar Kinematics in the Core of M87". In: *Astrophys. J.* 348 (Jan. 1990), p. 120. DOI: 10.1086/168218.

[47] Bililign T. Dullo, Alister W. Graham, and Johan H. Knapen. "A remarkably large depleted core in the Abell 2029 BCG IC 1101". In: *Mon. Not. R. Astron. Soc.* 471.2 (Oct. 2017), pp. 2321–2333. DOI: 10.1093/mnras/stx1635. arXiv: 1707.02277 [astro-ph.GA].

[48] G. Efstathiou, R. S. Ellis, and D. Carter. "Spectroscopic observations of three elliptical galaxies." In: *Mon. Not. R. Astron. Soc.* 193 (Dec. 1980), pp. 931–946. DOI: 10.1093/mnras/193.4.931.

[49] E. Emsellem et al. "The SAURON project - IX. A kinematic classification for early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 379 (Aug. 2007), pp. 401–417. DOI: 10.1111/j.1365-2966.2007.11752.x. eprint: arXiv:astro-ph/0703531.

[50] E. Emsellem, D. Krajnovic, and M. Sarzi. "A kinematically distinct core and minor-axis rotation: the MUSE perspective on M87." In: *Mon. Not. R. Astron. Soc.* 445 (Nov. 2014), pp. L79–L83. DOI: 10.1093/mnrasl/slu140.

[51] E. Emsellem, G. Monnet, and R. Bacon. "The multi-gaussian expansion method: a tool for building realistic photometric and kinematical models of stellar systems I. The formalism". In: *Astron. & Astrophys.* 285 (May 1994), pp. 723–738.

[52] E. Emsellem et al. "The SAURON project - III. Integral-field absorption-line kinematics of 48 elliptical and lenticular galaxies". In: *Mon. Not. R. Astron. Soc.* 352 (Aug. 2004), pp. 721–743. DOI: 10.1111/j.1365-2966.2004.07948.x. eprint: astro-ph/0404034.

[53] I. Ene et al. "The MASSIVE Survey XIII. Spatially Resolved Stellar Kinematics in the Central 1 kpc of 20 Massive Elliptical Galaxies with the GMOS-North Integral Field Spectrograph". In: *Astrophys. J.* 878, 57 (June 2019), p. 57. DOI: 10.3847/1538-4357/ab1f04. arXiv: 1904.08929.

[54] Irina Ene et al. "The MASSIVE Survey - X. Misalignment between kinematic and photometric axes and intrinsic shapes of massive early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 479.2 (Sept. 2018), pp. 2810–2826. DOI: 10.1093/mnras/sty1649. arXiv: 1802.00014 [astro-ph.GA].

[55] Irina Ene et al. "The MASSIVE Survey XIV – Stellar Velocity Profiles and Kinematic Misalignments from 200 pc to 20 kpc in Massive Early-type Galaxies". In: (2020). arXiv: 2001.11046 [astro-ph.GA].

[56] Peter Erwin. "IMFIT: A Fast, Flexible New Program for Astronomical Image Fitting". In: *Astrophys. J.* 799.2, 226 (Feb. 2015), p. 226. DOI: 10.1088/0004-637X/799/2/226. arXiv: 1408.1097 [astro-ph.IM].

[57] Event Horizon Telescope Collaboration et al. "First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole". In: *Astrophys. J. Lett.* 875.1, L1 (Apr. 2019), p. L1. DOI: 10.3847/2041-8213/ab0ec7. arXiv: 1906.11238 [astro-ph.GA].

[58] Event Horizon Telescope Collaboration et al. "First M87 Event Horizon Telescope Results. V. Physical Origin of the Asymmetric Ring". In: *Astrophys. J. Lett.* 875.1, L5 (Apr. 2019), p. L5. DOI: 10.3847/2041-8213/ab0f43. arXiv: 1906.11242 [astro-ph.GA].

[59] Event Horizon Telescope Collaboration et al. "First M87 Event Horizon Telescope Results. VI. The Shadow and Mass of the Central Black Hole". In: *Astrophys. J. Lett.* 875.1, L6 (Apr. 2019), p. L6. DOI: 10.3847/2041-8213/ab1141.

[60] Event Horizon Telescope Collaboration et al. "First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way". In: *Astrophys. J. Lett.* 930.2, L12 (May 2022), p. L12. DOI: 10.3847/2041-8213/ac6674.

[61] S. M. Faber et al. "The Centers of Early-Type Galaxies with HST. IV. Central Parameter Relations." In: *Astron. J.* 114 (Nov. 1997), p. 1771. DOI: 10.1086/118606. eprint: arXiv:astro-ph/9610055.

[62] S. M. Faber et al. "Spectroscopy and Photometry of Elliptical Galaxies. VI. Sample Selection and Data Summary". In: *Astrophys. J. Suppl. Ser.* 69 (Apr. 1989), p. 763. DOI: 10.1086/191327.

[63] J. Falcón-Barroso and M. Martig. "BAYES-LOSVD: A Bayesian framework for non-parametric extraction of the line-of-sight velocity distribution of galaxies". In: *Astron. & Astrophys.* 646, A31 (Feb. 2021), A31. DOI: 10.1051/0004-6361/202039624. arXiv: 2011.12023 [astro-ph.GA].

[64] J. Falcón-Barroso et al. "An updated MILES stellar library and stellar population models". In: *Astron. & Astrophys.* 532, A95 (Aug. 2011), A95. DOI: 10.1051/0004-6361/201116842. arXiv: 1107.2303.

[65] A. Feldmeier-Krause et al. "Triaxial orbit-based modelling of the Milky Way nuclear star cluster". In: *Mon. Not. R. Astron. Soc.* 466.4 (2017), pp. 4040–4052. DOI: 10.1093/mnras/stw3377. eprint: /oup/backfile/content_public/journal/mnras/466/4/10.1093_mnras_stw3377/4/stw3377.pdf. URL: http://dx.doi.org/10.1093/mnras/stw3377.

[66] L. Ferrarese et al. "The ACS Virgo Cluster Survey. VI. Isophotal Analysis and the Structure of Early-Type Galaxies". In: *Astrophys. J. Suppl. Ser.* 164 (June 2006), pp. 334–434. DOI: 10.1086/501350.

[67] David Fisher, Garth Illingworth, and Marijn Franx. "Kinematics of 13 brightest cluster galaxies". In: *Astrophys. J.* 438 (Jan. 1995), pp. 539–562. ISSN: 0004-637X. DOI: 10.1086/175100.

[68] L. M. R. Fogarty et al. "The SAMI Pilot Survey: stellar kinematics of galaxies in Abell 85, 168 and 2399". In: *Mon. Not. R. Astron. Soc.* 454 (Dec. 2015), pp. 2050–2066. DOI: 10.1093/mnras/stv2060. arXiv: 1509.02641.

[69] L. M. R. Fogarty et al. "The SAMI Pilot Survey: the kinematic morphology-density relation in Abell 85, Abell 168 and Abell 2399". In: *Mon. Not. R. Astron. Soc.* 443.1 (Sept. 2014), pp. 485–503. DOI: 10.1093/mnras/stu1165. arXiv: 1406.3899 [astro-ph.GA].

[70] Duncan A. Forbes et al. "Keck Cosmic Web Imager (KCWI) spectra of globular clusters and ultracompact dwarfs in the halo of M87". In: *Mon. Not. R. Astron. Soc.* 497.1 (Sept. 2020), pp. 765–775. DOI: 10.1093/mnras/staa1924.

[71] C. Foster et al. "The SAMI Galaxy Survey: the intrinsic shape of kinematically selected galaxies". In: *Mon. Not. R. Astron. Soc.* 472.1 (Nov. 2017), pp. 966–978. DOI: 10.1093/mnras/stx1869. arXiv: 1709.03585 [astro-ph.GA].

[72] C. Foster et al. "The SAMI Galaxy Survey: the intrinsic shape of kinematically selected galaxies". In: *Mon. Not. R. Astron. Soc.* 472.1 (Nov. 2017), pp. 966–978. DOI: 10.1093/mnras/stx1869.

[73] M. Franx, G. Illingworth, and T. de Zeeuw. "The ordered nature of elliptical galaxies - Implications for their intrinsic angular momenta and shapes". In: *Astrophys. J.* 383 (Dec. 1991), pp. 112–134. DOI: 10.1086/170769.

[74] Marijn Franx. "The projection of galaxy models with a Staeckel potential". In: *Mon. Not. R. Astron. Soc.* 231 (Mar. 1988), pp. 285–308. DOI: 10.1093/mnras/231.2.285.

[75] K. Gebhardt and J. Thomas. "The Black Hole Mass, Stellar Mass-to-Light Ratio, and Dark Halo in M87". In: *Astrophys. J.* 700 (Aug. 2009), pp. 1690–1701. DOI: 10.1088/0004-637X/700/2/1690. arXiv: 0906.1492 [astro-ph.CO].

[76] K. Gebhardt et al. "Axisymmetric Dynamical Models of the Central Regions of Galaxies". In: *Astrophys. J.* 583 (Jan. 2003), pp. 92–115. DOI: 10.1086/345081. eprint: astro-ph/0209483. URL: http://adsabs.harvard.edu/abs/2003ApJ...583...92G.

[77] K. Gebhardt et al. "The Black Hole Mass in M87 from Gemini/NIFS Adaptive Optics Observations". In: *Astrophys. J.* 729, 119 (Mar. 2011), p. 119. DOI: 10.1088/0004-637X/729/2/119. arXiv: 1101.1954 [astro-ph.CO].

[78] Karl Gebhardt, R. M. Rich, and Luis C. Ho. "A 20,000 M_{solar} Black Hole in the Stellar Cluster G1". In: *Astrophys. J. Lett.* 578.1 (Oct. 2002), pp. L41–L45. DOI: 10.1086/342980. arXiv: astro-ph/0209313 [astro-ph].

[79] Karl Gebhardt et al. "Axisymmetric, Three-Integral Models of Galaxies: A Massive Black Hole in NGC 3379". In: *Astron. J.* 119.3 (Mar. 2000), pp. 1157–1171. DOI: 10.1086/301240. arXiv: astro-ph/9912026 [astro-ph].

[80] Ortwin E. Gerhard and James J. Binney. "On the deprojection of axisymmetric bodies". In: *Mon. Not. R. Astron. Soc.* 279 (Apr. 1996), p. 993. DOI: 10.1093/mnras/279.3.993. arXiv: astro-ph/9508116 [astro-ph].

[81] Nicholas I. M. Gould, Dominique Orban, and Philippe L. Toint. "GALAHAD, a Library of Thread-safe Fortran 90 Packages for Large-scale Nonlinear Optimization". In: *ACM Trans. Math. Softw.* 29.4 (Dec. 2003), pp. 353–372. ISSN: 0098-3500. DOI: 10.1145/962437.962438. URL: http://doi.acm.org/10.1145/962437.962438.

[82] A. D. Goulding et al. "The MASSIVE Survey. IV. The X-ray Halos of the Most Massive Early-type Galaxies in the Nearby Universe". In: *Astrophys. J.* 826, 167 (Aug. 2016), p. 167. DOI: 10.3847/0004-637X/826/2/167. arXiv: 1604.01764.

[83] Charles F. Goullaud et al. "The MASSIVE Survey. IX. Photometric Analysis of 35 High-mass Early-type Galaxies with HST WFC3/IR". In: *Astrophys. J.* 856, 11 (Mar. 2018), p. 11. DOI: 10.3847/1538-4357/aab1f3.

[84] Alister W. Graham et al. "A New Empirical Model for the Structural Analysis of Early-Type Galaxies, and A Critical Review of the Nuker Model". In: *Astron. J.* 125.6 (June 2003), pp. 2951–2963. DOI: 10.1086/375320. arXiv: astro-ph/0306023 [astro-ph].

[85] Gravity Collaboration et al. "A geometric distance measurement to the Galactic center black hole with 0.3% uncertainty". In: *Astron. & Astrophys.* 625, L10 (May 2019), p. L10. DOI: 10.1051/0004-6361/201935656. arXiv: 1904.05721 [astro-ph.GA].

[86] Jenny E Greene et al. "THE MASSIVE SURVEY. II. STELLAR POPULATION TRENDS OUT TO LARGE RADIUS IN MASSIVE EARLY-TYPE GALAXIES". In: *Astrophys. J.* 807.1 (2015), p. 11. ISSN: 15384357. DOI: 10.1088/0004-637X/807/1/11.

[87] Jenny E. Greene et al. "The MASSIVE Survey. XII. Connecting Stellar Populations of Early-type Galaxies to Kinematics and Environment". In: *Astrophys. J.* 874.1 (Mar. 2019), p. 66. DOI: 10.3847/1538-4357/ab01e3. URL: https://doi.org/10.3847%2F1538-4357%2Fab01e3.

[88] Jorrit H. J. Hagen, Amina Helmi, and Maarten A. Breddels. "Axisymmetric Schwarzschild models of an isothermal axisymmetric mock dwarf spheroidal galaxy". In: *Astron. & Astrophys.* 632, A99 (Dec. 2019), A99. DOI: 10.1051/0004-6361/201936196. arXiv: 1907.00156 [astro-ph.GA].

[89] N. Häring and H.-W. Rix. "On the Black Hole Mass-Bulge Mass Relation". In: *Astrophys. J. Lett.* 604 (Apr. 2004), pp. L89–L92. DOI: 10.1086/383567. eprint: astro-ph/0402376.

[90] Richard J. Harms et al. "HST FOS Spectroscopy of M87: Evidence for a Disk of Ionized Gas around a Massive Black Hole". In: *Astrophys. J. Lett.* 435 (Nov. 1994), p. L35. DOI: 10.1086/187588.

[91] G. Heiligman and M. Schwarzschild. "On the nonexistence of three-dimensional tube orbits around the intermediate axis in a triaxial galaxy model." In: *Astrophys. J.* 233.3 (Nov. 1979), pp. 872–876. DOI: 10.1086/157449.

[92] Lars Hernquist. "An Analytical Model for Spherical Galaxies and Bulges". In: *Astrophys. J.* 356 (June 1990), p. 359. DOI: 10.1086/168845.

[93] G. J. Hill et al. "Design, construction, and performance of VIRUS-P: the prototype of a highly replicated integral-field spectrograph for HET". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 7014. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Aug. 2008. DOI: 10.1117/12.790235.

[94] J. G. Hoessel. "The photometric properties of brightest cluster galaxies. II. Sit and CCD surface photometry." In: *Astrophys. J.* 241 (Oct. 1980), pp. 493–506. DOI: 10.1086/158364.

[95] Erik Holmberg. "A Study of Double and Multiple Galaxies Together with Inquiries into some General Metagalactic Problems". In: *Annals of the Observatory of Lund* 6 (Jan. 1937), pp. 1–173.

[96] I. M. Hook et al. "The Gemini-North Multi-Object Spectrograph: Performance in Imaging, Long-Slit, and Multi-Object Spectroscopic Modes". In: *Pub. Astron. Soc. Pacific* 116 (May 2004), pp. 425–440. DOI: 10.1086/383624.

[97] E. P. Hubble. "A spiral nebula as a stellar system, Messier 31." In: *Astrophys. J.* 69 (Mar. 1929), pp. 103–158. DOI: 10.1086/143167.

[98] E. P. Hubble. "Extragalactic nebulae." In: *Astrophys. J.* 64 (Dec. 1926), pp. 321–369. DOI: 10.1086/143018.

[99] E. P. Hubble. "NGC 6822, a remote stellar system." In: *Astrophys. J.* 62 (Dec. 1925), pp. 409–433. DOI: 10.1086/142943.

[100] E. P. Hubble. "The classification of spiral nebulae". In: *The Observatory* 50 (Sept. 1927), pp. 276–281.

[101] Joseph B. Jensen et al. "Infrared Surface Brightness Fluctuation Distances for MASSIVE and Type Ia Supernova Host Galaxies". In: *arXiv e-prints*, arXiv:2105.08299 (May 2021), arXiv:2105.08299. arXiv: 2105.08299 [astro-ph.CO].

[102] Yunpeng Jin et al. "Evaluating the ability of triaxial Schwarzschild modelling to estimate properties of galaxies from the Illustris simulation". In: *Mon. Not. R. Astron. Soc.* 486.4 (July 2019), pp. 4753–4772. DOI: 10.1093/mnras/stz1170. arXiv: 1904.12942 [astro-ph.GA].

[103] Yunpeng Jin et al. "SDSS-IV MaNGA: Internal mass distributions and orbital structures of early-type galaxies and their dependence on environment". In: *Mon. Not. R. Astron. Soc.* 491.2 (Jan. 2020), pp. 1690–1708. DOI: 10.1093/mnras/stz3072. arXiv: 1911.00777 [astro-ph.GA].

[104] C. L. Joseph et al. "The Nuclear Dynamics of M32. I. Data and Stellar Kinematics". In: *Astrophys. J.* 550.2 (Apr. 2001), pp. 668–690. DOI: 10.1086/319781. arXiv: astro-ph/0005530 [astro-ph].

[105] Paul Knysh and Yannis Korkolis. "Blackbox: A procedure for parallel optimization of expensive black-box functions". In: *arXiv e-prints*, arXiv:1605.00998 (May 2016), arXiv:1605.00998. arXiv: 1605.00998 [cs.MS].

[106] J. Kormendy and R. Bender. "A Proposed Revision of the Hubble Sequence for Elliptical Galaxies". In: *Astrophys. J.* 464 (June 1996), p. L119. DOI: 10.1086/310095.

[107] J. Kormendy and R. Bender. "Correlations between Supermassive Black Holes, Velocity Dispersions, and Mass Deficits in Elliptical Galaxies with Cores". In: *Astrophys. J. Lett.* 691 (Feb. 2009), pp. L142–L146. DOI: 10.1088/0004-637X/691/2/L142. arXiv: 0901.3778 [astro-ph.GA].

[108] J. Kormendy and L. C. Ho. "Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies". In: *Ann. Rev. Astron. Astrophys.* 51 (Aug. 2013), pp. 511–653. DOI: 10.1146/annurev-astro-082708-101811. arXiv: 1304.7762 [astro-ph.CO].

[109] J. Kormendy et al. "Structure and Formation of Elliptical and Spheroidal Galaxies". In: *Astrophys. J. Suppl. Ser.* 182 (May 2009), pp. 216–309. DOI: 10.1088/0067-0049/182/1/216. arXiv: 0810.1681.

[110] D. Krajnović et al. "A quartet of black holes and a missing duo: probing the low end of the M_{BH}-σ relation with the adaptive optics assisted integral-field spectroscopy". In: *Mon. Not. R. Astron. Soc.* 477 (July 2018), pp. 3030–3064. DOI: 10.1093/mnras/sty778. arXiv: 1803.08055.

[111] Davor Krajnović et al. "Determination of masses of the central black holes in NGC 524 and 2549 using laser guide star adaptive optics". In: *Mon. Not. R. Astron. Soc.* 399.4 (2009), pp. 1839–1857. DOI: 10.1111/j.1365-2966.2009.15415.x. eprint: /oup/backfile/content_public/journal/mnras/399/4/10.1111/j.1365-2966.2009.15415.x/2/mnras0399-1839.pdf. URL: http://dx.doi.org/10.1111/j.1365-2966.2009.15415.x.

[112] T. R. Lauer et al. "The Masses of Nuclear Black Holes in Luminous Elliptical Galaxies and Implications for the Space Density of the Most Massive Black Holes". In: *Astrophys. J.* 662 (June 2007), pp. 808–834. DOI: 10.1086/518223. eprint: arXiv:astro-ph/0606739.

[113] T. R. Lauer et al. "The Centers of Early-Type Galaxies with HST.I.An Observational Survey". In: *Astron. J.* 110 (Dec. 1995), p. 2622. DOI: 10.1086/117719.

[114] Tod R. Lauer et al. "The Centers of Early-Type Galaxies with Hubble Space Telescope. V. New WFPC2 Photometry". In: *Astron. J.* 129.5 (May 2005), pp. 2138–2185. DOI: 10.1086/429565. arXiv: astro-ph/0412040 [astro-ph].

[115] Gigi Y. C. Leung et al. "The EDGE-CALIFA survey: validating stellar dynamical mass models with CO kinematics". In: *Mon. Not. R. Astron. Soc.* 477.1 (June 2018), pp. 254–292. DOI: 10.1093/mnras/sty288. arXiv: 1803.02259 [astro-ph.GA].

[116] Chao Li et al. "A discrete chemo-dynamical model of M87's globular clusters: Kinematics extending to ∼400 kpc". In: *Mon. Not. R. Astron. Soc.* 492.2 (Feb. 2020), pp. 2775–2795. DOI: 10.1093/mnras/staa027.

[117] H. Li et al. "The origin and properties of massive prolate galaxies in the Illustris simulation". In: *Mon. Not. R. Astron. Soc.* 473 (Jan. 2018), pp. 1489–1511. DOI: 10.1093/mnras/stx2374. arXiv: 1709.03345.

[118] Hongyu Li et al. "SDSS-IV MaNGA: The Intrinsic Shape of Slow Rotator Early-type Galaxies". In: *Astrophys. J. Lett.* 863.2, L19 (Aug. 2018), p. L19. DOI: 10.3847/2041-8213/aad54b.

[119] Christopher M. Liepold et al. "The MASSIVE Survey. XV. A Stellar Dynamical Mass Measurement of the Supermassive Black Hole in Massive Elliptical Galaxy NGC 1453". In: *Astrophys. J.* 891.1 (Feb. 2020), p. 4. DOI: 10.3847/1538-4357/ab6f71.

[120] Emily R. Liepold, Chung-Pei Ma, and Jonelle L. Walsh. "Keck Integral-field Spectroscopy of M87 Reveals an Intrinsically Triaxial Galaxy and a Revised Black Hole Mass". In: *Astrophys. J. Lett.* 945.2, L35 (Mar. 2023), p. L35. DOI: 10.3847/2041-8213/acbbcf. arXiv: 2302.07884 [astro-ph.GA].

[121] Mathias Lipka and Jens Thomas. "A novel approach to optimize the regularization and evaluation of dynamical models using a model selection framework". In: *Mon. Not. R. Astron. Soc.* (Apr. 2021). DOI: 10.1093/mnras/stab1092. arXiv: 2104.10168 [astro-ph.GA].

[122] A. Longobardi et al. "Kinematics of the outer halo of M 87 as mapped by planetary nebulae". In: *Astron. & Astrophys.* 620, A111 (Dec. 2018), A111. DOI: 10.1051/0004-6361/201832729.

[123] O. López-Cruz et al. "The Brightest Cluster Galaxy in A85: The Largest Core Known So Far". In: *Astrophys. J. Lett.* 795.2, L31 (Nov. 2014), p. L31. DOI: 10.1088/2041-8205/795/2/L31. arXiv: 1405.7758 [astro-ph.GA].

[124] S. I. Loubser and P. Sánchez-Blázquez. "The ultraviolet upturn in brightest cluster galaxies". In: *Mon. Not. R. Astron. Soc.* 410.4 (Jan. 2011), pp. 2679–2689. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.17666.x. eprint: https://academic.oup.com/mnras/article-pdf/410/4/2679/6297232/mnras0410-2679.pdf. URL: https://doi.org/10.1111/j.1365-2966.2010.17666.x.

[125] S. I. Loubser et al. "Radial kinematics of brightest cluster galaxies". In: *Mon. Not. R. Astron. Soc.* 391.3 (Dec. 2008), pp. 1009–1028. DOI: 10.1111/j.1365-2966.2008.13813.x. arXiv: 0808.1521.

[126] C.-P. Ma et al. "The MASSIVE Survey. I. A Volume-limited Integral-field Spectroscopic Study of the Most Massive Early-type Galaxies within 108 Mpc". In: *Astrophys. J.* 795, 158 (Nov. 2014), p. 158. DOI: 10.1088/0004-637X/795/2/158. arXiv: 1407.1054.

[127] F. Macchetto et al. "The Supermassive Black Hole of M87 and the Kinematics of Its Associated Gaseous Disk". In: *Astrophys. J.* 489.2 (Nov. 1997), pp. 579–600. DOI: 10.1086/304823. arXiv: astro-ph/9706252 [astro-ph].

[128] Juan P. Madrid and Carlos J. Donzelli. "The Abell 85 BCG: A Nucleated, Coreless Galaxy". In: *Astrophys. J.* 819.1, 50 (Mar. 2016), p. 50. DOI: 10.3847/0004-637X/819/1/50. arXiv: 1603.00078 [astro-ph.GA].

[129] J. Magorrian et al. "The Demography of Massive Dark Objects in Galaxy Centers". In: *Astron. J.* 115 (June 1998), pp. 2285–2305. DOI: 10.1086/300353. eprint: astro-ph/9708072.

[130] John Magorrian. "Constraining black hole masses from stellar kinematics by summing over all possible distribution functions". In: *Mon. Not. R. Astron. Soc.* 373.1 (Nov. 2006), pp. 425–434. DOI: 10.1111/j.1365-2966.2006.11054.x. arXiv: astro-ph/0609238 [astro-ph].

[131] John Magorrian. "Kinematical signatures of hidden stellar discs". In: *Mon. Not. R. Astron. Soc.* 302.3 (Jan. 1999), pp. 530–536. DOI: 10.1046/j.1365-8711.1999.02135.x. arXiv: astro-ph/9902033 [astro-ph].

[132] Roeland P. van der Marel and Marijn Franx. "A new method for the identification of non-Gaussian line profiles in elliptical galaxies". In: *Astrophys. J.* 407 (Apr. 1993), p. 525. ISSN: 0004-637X. DOI: 10.1086/172534. URL: http://adsabs.harvard.edu/doi/10.1086/172534.

[133] N. J. McConnell and C.-P. Ma. "Revisiting the Scaling Relations of Black Hole Masses and Host Galaxy Properties". In: *Astrophys. J.* 764, 184 (Feb. 2013), p. 184. DOI: 10.1088/0004-637X/764/2/184. arXiv: 1211.2816 [astro-ph.CO].

[134] N. J. McConnell et al. "The Black Hole Mass in the Brightest Cluster Galaxy NGC 6086". In: *Astrophys. J.* 728, 100 (Feb. 2011), p. 100. DOI: 10.1088/0004-637X/728/2/100. arXiv: 1009.0750 [astro-ph.CO].

[135] N. J. McConnell et al. "Two ten-billion-solar-mass black holes at the centres of giant elliptical galaxies". In: *Nature* 480 (Dec. 2011), pp. 215–218. DOI: 10.1038/nature10636. arXiv: 1112.1078 [astro-ph.CO].

[136] Nicholas J. McConnell et al. "Dynamical measurements of black hole masses in four brightest cluster galaxies at 100 Mpc". In: *Astrophys. J.* 756.2 (2012). ISSN: 15384357. DOI: 10.1088/0004-637X/756/2/179. arXiv: 1203.1620. URL: http://iopscience.iop.org/article/10.1088/0004-637X/756/2/179/pdf.

[137] M. D. McKay, R. J. Beckman, and W. J. Conover. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code". In: *Technometrics* 21.2 (1979), pp. 239–245. ISSN: 00401706. URL: http://www.jstor.org/stable/1268522.

[138] Kianusch Mehrgan et al. "A 40 Billion Solar-mass Black Hole in the Extreme Core of Holm 15A, the Central Galaxy of Abell 85". In: *Astrophys. J.* 887.2, 195 (Dec. 2019), p. 195. DOI: 10.3847/1538-4357/ab5856. arXiv: 1907.10608 [astro-ph.GA].

[139] Kianusch Mehrgan et al. "Detailed Shapes of the Line-of-sight Velocity Distributions in Massive Early-type Galaxies from Nonparametric Spectral Models". In: *Astrophys. J.* 948.2, 79 (May 2023), p. 79. DOI: 10.3847/1538-4357/acbf2e. arXiv: 2303.01495 [astro-ph.GA].

[140] Simona Mei et al. "The ACS Virgo Cluster Survey. XIII. SBF Distance Catalog and the Three-dimensional Structure of the Virgo Cluster". In: *Astrophys. J.* 655.1 (Jan. 2007), pp. 144–162. DOI: 10.1086/509598. arXiv: astro-ph/0702510 [astro-ph].

[141] F. Mertens et al. "Kinematics of the jet in M 87 on scales of 100-1000 Schwarzschild radii". In: *Astron. & Astrophys.* 595, A54 (Oct. 2016), A54. DOI: 10.1051/0004-6361/201628829. arXiv: 1608.05063 [astro-ph.HE].

[142] Patrick Morrissey et al. "The Keck Cosmic Web Imager Integral Field Spectrograph". In: *Astrophys. J.* 864.1, 93 (Sept. 2018), p. 93. DOI: 10.3847/1538-4357/aad597.

[143] Sahil Moza. *sahilm89/lhsmdu: Latin Hypercube Sampling with Multi-Dimensional Uniformity (LHSMDU): Speed Boost minor compatibility fixes*. Version 1.1.1. July 2020. DOI: 10.5281/zenodo.3929531. URL: https://doi.org/10.5281/zenodo.3929531.

[144] J. D. Murphy, K. Gebhardt, and J. J. Adams. "Galaxy Kinematics with VIRUS-P: The Dark Matter Halo of M87". In: *Astrophys. J.* 729 (Mar. 2011), pp. 129–+. DOI: 10.1088/0004-637X/729/2/129. arXiv: 1101.1957 [astro-ph.CO].

[145] Thorsten Naab, Sadegh Khochfar, and Andreas Burkert. "Properties of Early-Type, Dry Galaxy Mergers and the Origin of Massive Elliptical Galaxies". In: *Astrophys. J. Lett.* 636.2 (Jan. 2006), pp. L81–L84. DOI: 10.1086/500205. arXiv: astro-ph/0509667 [astro-ph].

[146] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. "The Structure of Cold Dark Matter Halos". In: *Astrophys. J.* 462 (May 1996), p. 563. DOI: 10.1086/177173. arXiv: astro-ph/9508025 [astro-ph].

[147] L. J. Oldham and M. W. Auger. "Galaxy structure from multiple tracers - II. M87 from parsec to megaparsec scales". In: *Mon. Not. R. Astron. Soc.* 457.1 (Mar. 2016), pp. 421–439. DOI: 10.1093/mnras/stv2982.

[148] Lindsay Oldham and Matthew Auger. "Galaxy structure from multiple tracers - III. Radial variations in M87's IMF". In: *Mon. Not. R. Astron. Soc.* 474.3 (Mar. 2018), pp. 4169–4185. DOI: 10.1093/mnras/stx2969.

[149] Viraj Pandya et al. "The MASSIVE Survey VI: The Spatial Distribution and Kinematics of Warm Ionized Gas in the Most Massive Local Early-type Galaxies". In: *Astrophys. J.* 837 (2017), pp. 1–16. DOI: 10.3847/1538-4357/aa5ebc. arXiv: 1701.08772.

[150] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[151] Reynier F. Peletier et al. "The SAURON project – XI. Stellar populations from absorption-line strength maps of 24 early-type spirals". In: *Mon. Not. R. Astron. Soc.* 379.2 (July 2007), pp. 445–468. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.11860.x. eprint: http://oup.prod.sis.lan/mnras/article-pdf/379/2/445/3372070/mnras0379-0445.pdf. URL: https://doi.org/10.1111/j.1365-2966.2007.11860.x.

[152] D. Pfenniger. "The velocity fields of barred galaxies." In: *Astron. & Astrophys.* 141 (Dec. 1984), pp. 171–188.

[153] Jacob D. Pilawa et al. "The MASSIVE Survey. XVII. A Triaxial Orbit-based Determination of the Black Hole Mass and Intrinsic Shape of Elliptical Galaxy NGC 2693". In: *Astrophys. J.* 928.2, 178 (Apr. 2022), p. 178. DOI: 10.3847/1538-4357/ac58fd.

[154] Adriano Poci et al. "Combining stellar populations with orbit-superposition dynamical modelling: the formation history of the lenticular galaxy NGC 3115". In: *Mon. Not. R. Astron. Soc.* 487.3 (Aug. 2019), pp. 3776–3796. DOI: 10.1093/mnras/stz1154. arXiv: 1904.11605 [astro-ph.GA].

[155] Matthew E. Quenneville, Christopher M. Liepold, and Chung-Pei Ma. "Dynamical Modeling of Galaxies and Supermassive Black Holes: Axisymmetry in Triaxial Schwarzschild Orbit Superposition Models". In: *Astrophys. J. Suppl. Ser.* 254.2, 25 (June 2021), p. 25. DOI: 10.3847/1538-4365/abe6a0. arXiv: 2005.00542 [astro-ph.GA].

[156] Matthew E. Quenneville, Christopher M. Liepold, and Chung-Pei Ma. "Dynamical Modeling of Galaxies and Supermassive Black Holes: Axisymmetry in Triaxial Schwarzschild Orbit Superposition Models". In: *Astrophys. J. Suppl. Ser.* 254.2, 25 (June 2021), p. 25. DOI: 10.3847/1538-4365/abe6a0. arXiv: 2005.00542 [astro-ph.GA].

[157] Matthew E. Quenneville, Christopher M. Liepold, and Chung-Pei Ma. "Triaxial Orbit-based Dynamical Modeling of Galaxies with Supermassive Black Holes and an Application to Massive Elliptical Galaxy NGC 1453". In: *Astrophys. J.* 926.1, 30 (Feb. 2022), p. 30. DOI: 10.3847/1538-4357/ac3e68.

[158] Matthew E. Quenneville, Christopher M. Liepold, and Chung-Pei Ma. "Triaxial Orbit-based Dynamical Modeling of Galaxies with Supermassive Black Holes and an Application to Massive Elliptical Galaxy NGC 1453". In: *The Astrophysical Journal* 926.1 (Feb. 2022), p. 30. DOI: 10.3847/1538-4357/ac3e68. URL: https://doi.org/10.3847/1538-4357/ac3e68.

[159] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. 2006.

[160] D. O. Richstone and S. Tremaine. "A general method for constructing spherical galaxy models". In: *Astrophys. J.* 286 (Nov. 1984), pp. 27–37. DOI: 10.1086/162572.

[161] D. O. Richstone and S. Tremaine. "Dynamical models of M 87 without a central black hole." In: *Astrophys. J.* 296 (Sept. 1985), pp. 370–378. DOI: 10.1086/163455.

[162] Hans-Walter Rix et al. "Dynamical Modeling of Velocity Profiles: The Dark Halo around the Elliptical Galaxy NGC 2434". In: *Astrophys. J.* 488.2 (Oct. 1997), pp. 702–719. DOI: 10.1086/304733. arXiv: astro-ph/9702126 [astro-ph].

[163] Aaron J. Romanowsky and Christopher S. Kochanek. "Dynamics of Stars and Globular Clusters in M87". In: *Astrophys. J.* 553.2 (June 2001), pp. 722–732. DOI: 10.1086/320947.

[164] S. P. Rusli et al. "The Influence of Dark Matter Halos on Dynamical Estimates of Black Hole Mass: 10 New Measurements for High-σ Early-type Galaxies". In: *Astron. J.* 146, 45 (Sept. 2013), p. 45. DOI: 10.1088/0004-6256/146/3/45. arXiv: 1306.1124 [astro-ph.CO].

[165] R. P. Saglia et al. "The SINFONI Black Hole Survey: The Black Hole Fundamental Plane Revisited and the Paths of (Co)evolution of Supermassive Black Holes and Bulges". In: *Astrophys. J.* 818, 47 (Feb. 2016), p. 47. DOI: 10.3847/0004-637X/818/1/47.

[166] P. Sánchez-Blázquez et al. "Medium-resolution Isaac Newton Telescope library of empirical spectra". In: *Mon. Not. R. Astron. Soc.* 371 (Sept. 2006), pp. 703–718. DOI: 10.1111/j.1365-2966.2006.10699.x. eprint: astro-ph/0607009.

[167] P. Sánchez-Blázquez et al. "Medium-resolution Isaac Newton Telescope library of empirical spectra". In: *Mon. Not. R. Astron. Soc.* 371 (Sept. 2006), pp. 703–718. DOI: 10.1111/j.1365-2966.2006.10699.x. eprint: astro-ph/0607009.

[168] Jason L. Sanders and James Binney. "A fast algorithm for estimating actions in triaxial potentials". In: *Mon. Not. R. Astron. Soc.* 447.3 (Mar. 2015), pp. 2479–2496. DOI: 10.1093/mnras/stu2598. arXiv: 1412.2093 [astro-ph.GA].

[169] Giulia Santucci et al. "The SAMI Galaxy Survey: The Internal Orbital Structure and Mass Distribution of Passive Galaxies from Triaxial Orbit-superposition Schwarzschild Models". In: *Astrophys. J.* 930.2, 153 (May 2022), p. 153. DOI: 10.3847/1538-4357/ac5bd5.

[170] W. L. W. Sargent et al. "Dynamical evidence for a central mass concentration in the galaxy M87." In: *Astrophys. J.* 221 (May 1978), pp. 731–744. DOI: 10.1086/156077.

[171] W. L. W. Sargent et al. "Velocity dispersions for 13 galaxies." In: *Astrophys. J.* 212 (Mar. 1977), pp. 326–334. DOI: 10.1086/155052.

[172] Marc Sarzi et al. "MUSE observations of M87: radial gradients for the stellar initial-mass function and the abundance of sodium". In: *Mon. Not. R. Astron. Soc.* 478.3 (Aug. 2018), pp. 4084–4100. DOI: 10.1093/mnras/sty1092.

[173] M. Schwarzschild. "A numerical model for a triaxial stellar system in dynamical equilibrium". In: *Astrophys. J.* 232 (Aug. 1979), pp. 236–247. DOI: 10.1086/157282.

[174] Martin Schwarzschild. "Self-consistent Models for Galactic Halos". In: *Astrophys. J.* 409 (June 1993), p. 563. DOI: 10.1086/172687.

[175] J. L. Sérsic. "Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy". In: *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina* 6 (Feb. 1963), pp. 41–43.

[176] Anil C. Seth et al. "A supermassive black hole in an ultra-compact dwarf galaxy". In: *Nature* 513.1 (2014), p. 398. DOI: 10.1038/nature13762. URL: https://doi.org/10.1038/nature13762.

[177] Kristen L. Shapiro et al. "The black hole in NGC 3379: a comparison of gas and stellar dynamical mass measurements with HST and integral-field data". In: *Mon. Not. R. Astron. Soc.* 370.2 (Aug. 2006), pp. 559–579. DOI: 10.1111/j.1365-2966.2006.10537.x. arXiv: astro-ph/0605479 [astro-ph].

[178] J. Shen and K. Gebhardt. "The Supermassive Black Hole and Dark Matter Halo of NGC 4649 (M60)". In: *Astrophys. J.* 711 (Mar. 2010), pp. 484–494. DOI: 10.1088/0004-637X/711/1/484. arXiv: 0910.4168 [astro-ph.CO].

[179] Christos Siopis et al. "A STELLAR DYNAMICAL MEASUREMENT OF THE BLACK HOLE MASS IN THE MASER GALAXY NGC 4258". In: *Astrophys. J.* 693.1 (Mar. 2009), pp. 946–969. DOI: 10.1088/0004-637x/693/1/946.

[180] Kurt T. Soto et al. "ZAP – enhanced PCA sky subtraction for integral field spectroscopy". In: *Monthly Notices of the Royal Astronomical Society* 458.3 (Mar. 2016), pp. 3210–3220. ISSN: 0035-8711. DOI: 10.1093/mnras/stw474.

[181] Joshua S. Speagle. "DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences". In: *Mon. Not. R. Astron. Soc.* 493.3 (Apr. 2020), pp. 3132–3158. DOI: 10.1093/mnras/staa278.

[182] S. Sridhar and J. Touma. "Stellar dynamics around black holes in galactic nuclei". In: *Mon. Not. R. Astron. Soc.* 303.3 (Mar. 1999), pp. 483–494. DOI: 10.1046/j.1365-8711.1999.02218.x. arXiv: astro-ph/9811304 [astro-ph].

[183] S. Sridhar and Jihad Touma. "Stellar dynamics around a massive black hole – I. Secular collisionless theory". In: *Mon. Not. R. Astron. Soc.* 458.4 (Mar. 2016), pp. 4129–4142. ISSN: 0035-8711. DOI: 10.1093/mnras/stw542. eprint: http://oup.prod.sis.lan/mnras/article-pdf/458/4/4129/18509203/stw542.pdf.

[184] Antony A. Stark. "Triaxial Models of the Bulge of M31". In: *Astrophys. J.* 213 (Apr. 1977), pp. 368–373. DOI: 10.1086/155164.

[185] Thomas S. Statler, Herwig Dejonghe, and Tammy Smecker-Hane. "The Three-Dimensional Mass Distribution in NGC 1700". In: *Astron. J.* 117.1 (Jan. 1999), pp. 126–139. DOI: 10.1086/300688. arXiv: astro-ph/9810046 [astro-ph].

[186] Jay Strader et al. "Wide-field Precision Kinematics of the M87 Globular Cluster System". In: *Astrophys. J. Suppl. Ser.* 197.2, 33 (Dec. 2011), p. 33. DOI: 10.1088/0067-0049/197/2/33.

[187] Stephen R. Taylor. "The Nanohertz Gravitational Wave Astronomer". In: *arXiv e-prints*, arXiv:2105.13270 (May 2021), arXiv:2105.13270.

[188] S. Thater et al. "A low upper mass limit for the central black hole in the late-type galaxy NGC 4414". In: *Astron. & Astrophys.* 597 (2017), A18. DOI: 10.1051/0004-6361/201629480. URL: https://doi.org/10.1051/0004-6361/201629480.

[189] Sabine Thater et al. "Six new supermassive black hole mass determinations from adaptive-optics assisted SINFONI observations". In: *A&A* 625 (2019), A62. DOI: 10.1051/0004-6361/201834808. URL: https://doi.org/10.1051/0004-6361/201834808.

[190] J. Thomas et al. "Mapping stationary axisymmetric phase-space distribution functions by orbit libraries". In: *Mon. Not. R. Astron. Soc.* 353.2 (Sept. 2004), pp. 391–404. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2004.08072.x. eprint: http://oup.prod.sis.lan/mnras/article-pdf/353/2/391/3858810/353-2-391.pdf.

[191] J. Thomas et al. "The Dynamical Fingerprint of Core Scouring in Massive Elliptical Galaxies". In: *Astrophys. J.* 782, 39 (Feb. 2014), p. 39. DOI: 10.1088/0004-637X/782/1/39. arXiv: 1311.3783 [astro-ph.GA].

[192] Jens Thomas et al. "A 17-billion-solar-mass black hole in a group galaxy with a diffuse core". In: *Nature* 532 (Apr. 2016), pp. 340–342. DOI: 10.1038/nature17197.

[193] Paul Tiede et al. "Measuring Photon Rings with the ngEHT". In: *Galaxies* 10.6 (Dec. 2022), p. 111. DOI: 10.3390/galaxies10060111. arXiv: 2210.13498 [astro-ph.HE].

[194] J. Tonry and M. Davis. "A survey of galaxy redshifts. I. Data reduction techniques." In: *Astron. J.* 84 (Oct. 1979), pp. 1511–1525. DOI: 10.1086/112569.

[195] B. Tremblay and D. Merritt. "Evidence From Intrinsic Shapes for Two Families of Elliptical Galaxies". In: *Astron. J.* 111 (June 1996), p. 2243. DOI: 10.1086/117959. eprint: astro-ph/9601038.

[196] Francisco Valdes et al. "The Indo-US Library of Coudé Feed Stellar Spectra". In: *Astrophys. J. Suppl. Ser.* 152.2 (June 2004), pp. 251–259. DOI: 10.1086/386343. arXiv: astro-ph/0402435 [astro-ph].

[197] Monica Valluri, David Merritt, and Eric Emsellem. "Difficulties with Recovering the Masses of Supermassive Black Holes from Stellar Kinematical Data". In: *Astrophys. J.* 602.1 (Feb. 2004), pp. 66–92. DOI: 10.1086/380896. arXiv: astro-ph/0210379 [astro-ph].

[198] Monica Valluri et al. "A Unified Framework for the Orbital Structure of Bars and Triaxial Ellipsoids". In: *Astrophys. J.* 818.2, 141 (Feb. 2016), p. 141. DOI: 10.3847/0004-637X/818/2/141. arXiv: 1512.03467 [astro-ph.GA].

[199] R. C. E. van den Bosch and P. T. de Zeeuw. "Estimating black hole masses in triaxial galaxies". In: *Mon. Not. R. Astron. Soc.* 401 (Jan. 2010), pp. 1770–1780. DOI: 10.1111/j.1365-2966.2009.15832.x. arXiv: 0910.0844 [astro-ph.CO].

[200] R. C. E. van den Bosch and P. T. de Zeeuw. "Estimating black hole masses in triaxial galaxies". In: *Mon. Not. R. Astron. Soc.* 401 (Jan. 2010), pp. 1770–1780. DOI: 10.1111/j.1365-2966.2009.15832.x.

[201] R. C. E. van den Bosch et al. "An over-massive black hole in the compact lenticular galaxy NGC1277". In: *Nature* 491 (Nov. 2012), pp. 729–731. DOI: 10.1038/nature11592. arXiv: 1211.6429 [astro-ph.CO].

[202] R. C. E. van den Bosch et al. "Triaxial orbit based galaxy models with an application to the (apparent) decoupled core galaxy NGC 4365". In: *Mon. Not. R. Astron. Soc.* 385 (Apr. 2008), pp. 647–666. DOI: 10.1111/j.1365-2966.2008.12874.x. arXiv: 0712.0113.

[203] Remco C. E. van den Bosch and Glenn van de Ven. "Recovering the intrinsic shape of early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 398.3 (Sept. 2009), pp. 1117–1128. DOI: 10.1111/j.1365-2966.2009.15177.x. arXiv: 0811.3474 [astro-ph].

[204] R. P. van der Marel. "Velocity profiles of galaxies with claimed black holes - III. Observations and models for M 87." In: *Mon. Not. R. Astron. Soc.* 270 (Sept. 1994), pp. 271–297. DOI: 10.1093/mnras/270.2.271.

[205] Roeland P. van der Marel, James Binney, and Roger L. Davies. "Models of Elliptical Galaxies - NGC3379 NGC4261 NGC4278 and NGC4472". In: *Mon. Not. R. Astron. Soc.* 245 (Aug. 1990), p. 582. DOI: 10.1093/mnras/245.4.582.

[206] Roeland P. van der Marel et al. "Improved Evidence for a Black Hole in M32 from HST/FOS Spectra. II. Axisymmetric Dynamical Models". In: *Astrophys. J.* 493.2 (Jan. 1998), pp. 613–631. DOI: 10.1086/305147. arXiv: astro-ph/9705081 [astro-ph].

[207] Pieter van Dokkum et al. "Spatially Resolved Stellar Kinematics of the Ultra-diffuse Galaxy Dragonfly 44. I. Observations, Kinematics, and Cold Dark Matter Halo Fits". In: *Astrophys. J.* 880.2, 91 (Aug. 2019), p. 91. DOI: 10.3847/1538-4357/ab2914.

[208] Eugene Vasiliev and Monica Valluri. "A New Implementation of the Schwarzchild Method for Constructing Observationally Driven Dynamical Models of Galaxies of All Morphological Types". In: *Astrophys. J.* 889.1, 39 (Jan. 2020), p. 39. DOI: 10.3847/1538-4357/ab5fe0. arXiv: 1912.04288 [astro-ph.GA].

[209] Melanie Veale et al. "The MASSIVE Survey - V. Spatially resolved stellar angular momentum, velocity dispersion, and higher moments of the 41 most massive local early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 464 (Jan. 2017), pp. 356–384. DOI: 10.1093/mnras/stw2330. arXiv: 1609.00391 [astro-ph.GA].

[210] Melanie Veale et al. "The MASSIVE Survey - VII. The relationship of angular momentum, stellar mass and environment of early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 471 (Oct. 2017), pp. 1428–1445. DOI: 10.1093/mnras/stx1639. arXiv: 1703.08573 [astro-ph.GA].

[211] Melanie Veale et al. "The MASSIVE survey - VIII. Stellar velocity dispersion profiles and environmental dependence of early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 473 (Feb. 2018), pp. 5446–5467. DOI: 10.1093/mnras/stx2717. arXiv: 1708.00870 [astro-ph.GA].

[212] G. van de Ven, P. T. De Zeeuw, and R. C. E. Van Den Bosch. "Recovery of the internal orbital structure of galaxies". In: *Mon. Not. R. Astron. Soc.* 385.2 (Feb. 2008), pp. 614–646. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2008.12873.x. eprint: http://oup.prod.sis.lan/mnras/article-pdf/385/2/614/18226009/mnras0385-0614.pdf.

[213] E. K. Verolme et al. "A SAURON study of M32: measuring the intrinsic flattening and the central black hole mass". In: *Mon. Not. R. Astron. Soc.* 335 (Sept. 2002), pp. 517–525. DOI: 10.1046/j.1365-8711.2002.05664.x. arXiv: astro-ph/0201086 [astro-ph].

[214] R. Anthony Vincent and Barbara S. Ryden. "The Dependence of Galaxy Shape on Luminosity and Surface Brightness Profile". In: *Astrophys. J.* 623.1 (Apr. 2005), pp. 137–147. DOI: 10.1086/428765. arXiv: astro-ph/0501221 [astro-ph].

[215] G. M. Voit et al. "A General Precipitation-limited L $_X$-T-R Relation among Early-type Galaxies". In: *Astrophys. J.* 853, 78 (Jan. 2018), p. 78. DOI: 10.3847/1538-4357/aaa084. arXiv: 1708.02189.

[216] R. Craig Walker et al. "The Structure and Dynamics of the Subparsec Jet in M87 Based on 50 VLBA Observations over 17 Years at 43 GHz". In: *Astrophys. J.* 855.2, 128 (Mar. 2018), p. 128. DOI: 10.3847/1538-4357/aaafcc. arXiv: 1802.06166 [astro-ph.HE].

[217] J. L. Walsh et al. "A 5 x 10^9 Msun Black Hole in NGC 1277 from Adaptive Optics Spectroscopy". In: *Astrophys. J.* 817, 2 (Jan. 2016), p. 2. DOI: 10.3847/0004-637X/817/1/2. arXiv: 1511.04455.

[218] J. L. Walsh et al. "A Black Hole Mass Determination for the Compact Galaxy Mrk 1216". In: *Astrophys. J.* 835, 208 (Feb. 2017), p. 208. DOI: 10.3847/1538-4357/835/2/208. arXiv: 1612.02015.

[219] J. L. Walsh et al. "A Stellar Dynamical Mass Measurement of the Black Hole in NGC 3998 from Keck Adaptive Optics Observations". In: *Astrophys. J.* 753, 79 (July 2012), p. 79. DOI: 10.1088/0004-637X/753/1/79. arXiv: 1205.0816.

[220] J. L. Walsh et al. "The Black Hole in the Compact, High-dispersion Galaxy NGC 1271". In: *Astrophys. J.* 808, 183 (Aug. 2015), p. 183. DOI: 10.1088/0004-637X/808/2/183. arXiv: 1506.05129.

[221] J. L. Walsh et al. "The M87 Black Hole Mass from Gas-dynamical Models of Space Telescope Imaging Spectrograph Observations". In: *Astrophys. J.* 770, 86 (June 2013), p. 86. DOI: 10.1088/0004-637X/770/2/86. arXiv: 1304.7273 [astro-ph.CO].

[222] A.-M. Weijmans et al. "The ATLAS ^{3D} project - XXIV. The intrinsic shape distribution of early-type galaxies". In: *Mon. Not. R. Astron. Soc.* 444 (Nov. 2014), pp. 3340–3356. DOI: 10.1093/mnras/stu1603. arXiv: 1408.1099.

[223] Michael J. West and John P. Blakeslee. "The Principal Axis of the Virgo Cluster". In: *Astrophys. J. Lett.* 543.1 (Nov. 2000), pp. L27–L30. DOI: 10.1086/318177. arXiv: astro-ph/0008470 [astro-ph].

[224] T. B. Williams and M. Schwarzschild. "A photometric determination of twists in three early-type galaxies." In: *Astrophys. J.* 227 (Jan. 1979), pp. 56–63. DOI: 10.1086/156703.

[225] Xiaoan Wu and Scott Tremaine. "Deriving the Mass Distribution of M87 from Globular Clusters". In: *Astrophys. J.* 643.1 (May 2006), pp. 210–221. DOI: 10.1086/501515.

[226] Meng Yang et al. "Mapping the dark matter halo of early-type galaxy NGC 2974 through orbit-based models with combined stellar and cold gas kinematics". In: *Mon. Not. R. Astron. Soc.* 491.3 (Jan. 2020), pp. 4221–4231. DOI: 10.1093/mnras/stz3293. arXiv: 1911.11058 [astro-ph.GA].

[227] Tim de Zeeuw. "Elliptical galaxies with separable potentials". In: *Mon. Not. R. Astron. Soc.* 216.2 (Sept. 1985), pp. 273–334. ISSN: 0035-8711. DOI: 10.1093/mnras/216.2.273. eprint: http://oup.prod.sis.lan/mnras/article-pdf/216/2/273/3500591/mnras216-0273.pdf.

[228] Hong-Xin Zhang et al. "The Next Generation Virgo Cluster Survey. VI. The Kinematics of Ultra-compact Dwarfs and Globular Clusters in M87". In: *Astrophys. J.* 802.1, 30 (Mar. 2015), p. 30. DOI: 10.1088/0004-637X/802/1/30.

[229] Ling Zhu et al. "Orbital decomposition of CALIFA spiral galaxies". In: *Mon. Not. R. Astron. Soc.* 473.3 (Jan. 2018), pp. 3000–3018. DOI: 10.1093/mnras/stx2409. arXiv: 1709.06649 [astro-ph.GA].

[230] Ling Zhu et al. "The stellar orbit distribution in present-day galaxies inferred from the CALIFA survey". In: *Nature Astronomy* 2 (Jan. 2018), pp. 233–238. DOI: 10.1038/s41550-017-0348-1. arXiv: 1711.06728 [astro-ph.GA].