



The MASSIVE Survey. XV. A Stellar Dynamical Mass Measurement of the Supermassive Black Hole in Massive Elliptical Galaxy NGC 1453

Christopher M. Liepold^{1,2} , Matthew E. Quenneville^{1,2}, Chung-Pei Ma^{1,2} , Jonelle L. Walsh³ , Nicholas J. McConnell⁴,
Jenny E. Greene⁵, and John P. Blakeslee⁶

¹ Department of Astronomy, University of California, Berkeley, CA 94720, USA; cmlepold@berkeley.edu

² Department of Physics, University of California, Berkeley, CA 94720, USA

³ George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA

⁴ Institute for Scientist and Engineer Educators, University of California, Santa Cruz, CA, USA

⁵ Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

⁶ Gemini Observatory, Casilla 603, La Serena, Chile

Received 2019 December 14; revised 2020 January 21; accepted 2020 January 22; published 2020 February 27

Abstract

We present a new stellar dynamical mass measurement (M_{BH}) of the supermassive black hole (SMBH) in NGC 1453, a fast-rotating massive elliptical galaxy in the MASSIVE survey. We measure stellar kinematics in 135 spatial bins in the central $1.5 \text{ kpc} \times 2 \text{ kpc}$ region of the galaxy using high signal-to-noise ratio ($S/N \sim 130$) spectra from the Gemini-North GMOS integral field spectrograph (IFS). Combining with wide-field IFS kinematics out to ~ 3 effective radii and stellar light distributions from *Hubble Space Telescope* Wide Field Camera 3 images, we perform Schwarzschild orbit-based mass modeling in the axisymmetric limit to constrain the mass components in NGC 1453. The best-fit black hole mass is $M_{\text{BH}} = (2.9 \pm 0.4) \times 10^9 M_{\odot}$; the mass models without a central black hole are excluded at the 8.7σ level. The NGC 1453 black hole lies within the intrinsic scatter of the SMBH and galaxy scaling relations, unlike three other galaxies hosting $\gtrsim 10^{10} M_{\odot}$ SMBHs in the MASSIVE sample. The high- S/N GMOS spectra enable us to determine eight moments of the Gauss–Hermite expansion of the line-of-sight velocity distributions (LOSVDs), which are used as constraints in the orbit modeling. The stellar orbits in the mass models are further constrained to produce negligible h_9 through h_{12} to minimize spurious behavior in the LOSVDs. We show that truncating the series at h_4 , as was often done in prior work, leads to a much weaker constraint on the inferred M_{BH} for NGC 1453. Furthermore, we discuss precautions and modifications that are needed to achieve axisymmetry in triaxial orbit codes that use the Schwarzschild method to sample the start space of stellar orbits in triaxial gravitational potentials.

Unified Astronomy Thesaurus concepts: Elliptical galaxies (456); Galaxies (573); Galaxy dynamics (591); Galaxy evolution (594); Galaxy kinematics (602); Galaxy structure (622); Dark matter (353)

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 data sets 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., Kormendy & Ho 2013; McConnell & Ma 2013; Saglia et al. 2016).

The high-mass regime ($M_{\text{BH}} \gtrsim 10^{9.5} M_{\odot}$) 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., Faber et al. 1997; Graham et al. 2003; Ferrarese et al. 2006; Côté et al. 2007; Lauer et al. 2007), indicating a significant deficit of stars, possibly due to three-body gravitational slingshots that scatter stars passing close to an SMBH binary to larger radii (e.g., Begelman et al. 1980). These diffuse cores make it extremely difficult to obtain stellar absorption-line spectra of high signal-to-noise ratio (S/N) quality that is needed for reliable M_{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 (Ma et al. 2014). The MASSIVE survey targets ~ 100 early-type galaxies (ETGs) in the northern sky ($\delta > -6^\circ$) within a distance of 108 Mpc. Within this volume, it is designed to be complete to an absolute K -band magnitude of $M_K = -25.3 \text{ mag}$, 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^{3D}, the previous volume-limited survey of 260 local ETGs out to a distance of 42 Mpc (Cappellari et al. 2011).

We have obtained comprehensive spectroscopic data using IFS on both subarcsecond 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 integral field spectrograph (IFS) observations can be found in Veale et al. (2017a, 2017b, 2018), Ene et al. (2018), and Greene et al. (2015, 2019). Results from finely resolved stellar kinematics in the central $\sim 2 \text{ kpc}$ regions of 20 MASSIVE galaxies are presented in Ene et al. (2019, 2020).

In addition to the IFS data, we have also assembled an extensive array of multiwavelength data of MASSIVE galaxies to study stellar light profiles (Goullaud et al. 2018), cold molecular gas (Davis et al. 2016, 2019), warm ionized gas (Pandya et al. 2017), and hot X-ray gas (Goulding et al. 2016; Voit et al. 2018).

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 subarcsecond 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 seven galaxies are in the Virgo Cluster: NGC 4486 (M87; Gebhardt et al. 2011; see also Walsh et al. 2013; Event Horizon Telescope Collaboration et al. 2019), NGC 4472 (M49; Rusli et al. 2013), and NGC 4649 (M60; Shen & Gebhardt 2010). Two others are the brightest cluster galaxies of rich clusters: NGC 4889 in the Coma Cluster and NGC 3842 in the Leo Cluster (McConnell et al. 2011, 2012). The remaining two are the brightest galaxies in galaxy groups: NGC 1600 in a fossil-like group (Thomas et al. 2016) and NGC 7619 in the Pegasus group (Rusli et al. 2013). Except for NGC 4649, the spectroscopic observations were all conducted with IFS on 8–10 m telescopes. The measured M_{BH} spans an order of magnitude from $\sim 2 \times 10^9$ to $\sim 2 \times 10^{10} M_{\odot}$. More M_{BH} measurements in this mass range are clearly needed to quantify more robustly the upper end of the M_{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. (2019) 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 (Veale et al. 2017a). As listed in Table 3 of Ma et al. (2014), the Two Micron All Sky Survey (2MASS) “high-density contrast” group catalog (Crook et al. 2007) 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 *Hubble Space Telescope* (HST) images of NGC 1453 show very regular elliptical isophotes (Figure 13 of Goullaud et al. 2018). The photometric and kinematic axes are also closely aligned (Ene et al. 2018, 2019, 2020), 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_{BH} scales linearly with the assumed distance. For NGC 1453, we use our new determination of 51.0 Mpc from the MASSIVE-WFC3 project (Goullaud et al. 2018) using the surface brightness fluctuation technique (J. B. Jensen et al. 2020, in preparation). This new distance is about 10% smaller than 56.4 Mpc from group-corrected flow velocity in the 2MASS redshift survey. For a flat Λ CDM with a matter density of $\Omega_m = 0.315$ and a Hubble parameter of $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $1''$ is 245 pc at 51.0 Mpc.

We perform Schwarzschild orbit modeling (Schwarzschild 1979) using the triaxial implementation described by van den Bosch

et al. (2008). We perform this modeling in the axisymmetric limit, and in Section 4.1 we 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 (van der Marel & Franx 1993; Rix et al. 1997). 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_{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_{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 Section 2, we describe the spectroscopic observations and the resulting stellar kinematics from the Gemini Multi-Object Spectrograph (GMOS; Hook et al. 2004) IFS of the central $\sim 1.5 \text{ kpc} \times 2 \text{ kpc}$ region of NGC 1453 and the wide-field coverage with the McDonald Mitchell IFS (Hill et al. 2008). In Section 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 Section 4. The mass modeling results are given in Section 5, and the best-fit mass model is discussed further in Section 6. In Section 7, we discuss a number of relevant issues: the impact of Gauss–Hermite series truncation on the inferred M_{BH} , the subtleties in achieving axisymmetry within the triaxial code, comparisons to results from Jeans modeling, implications for the black hole scaling relations, and connections to our previous observations of warm ionized gas in NGC 1453 (Pandya et al. 2017).

2. Spectroscopic Data and Stellar Kinematics

As part of the MASSIVE survey, we obtained spatially resolved stellar spectra for NGC 1453 with the GMOS (Hook et al. 2004) in the IFS mode on the 8.1 m Gemini-North Telescope and the Mitchell/VIRUS-P IFS (Hill et al. 2008) on the 2.7 m Harlan J. Smith Telescope at McDonald Observatory. Here we summarize the observations, data reduction processes, and procedures used to extract the stellar kinematics.

2.1. Central Kiloparsec Kinematics

We observed the central $\sim 1.5 \text{ kpc} \times 2 \text{ kpc}$ 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 calcium triplet (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

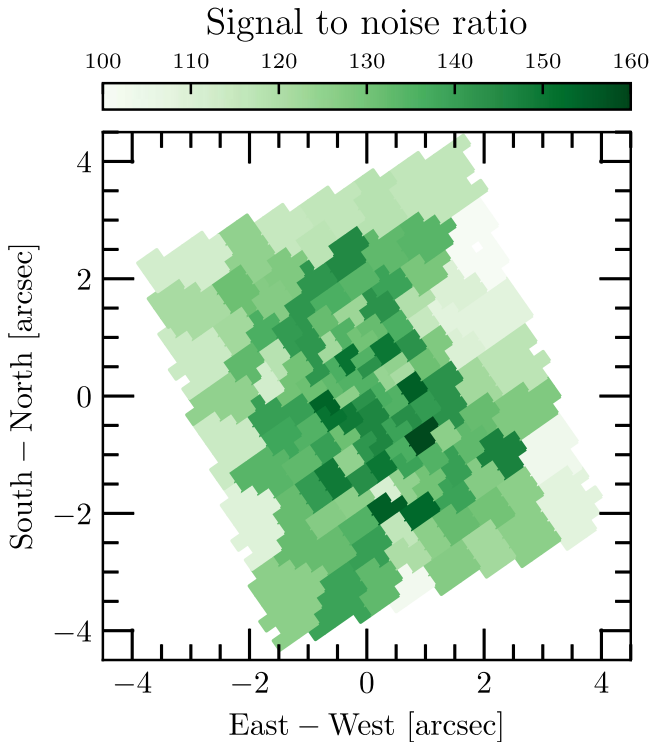


Figure 1. S/N map of the Gemini GMOS IFS data for the 135 Voronoi bins in the central $5'' \times 7''$ of NGC 1453. A target S/N of 125 is used in the binning procedure. The S/N value for each bin scatters around the target with a typical rms scatter of $\sim 10\%$, while the innermost bins achieve S/N 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 are needed for kinematic extraction and black hole measurements.

from arc lamp lines for each lenslet with a mean value of $2.5 \text{ \AA}$ FWHM. Six science exposures, each of 850 s, were taken. The median seeing was $0''.7$ FWHM. Other details and our data reduction procedure are described in Ene et al. (2019).

We use the Ca II triplet absorption features over the rest wavelength range of $8420\text{--}8770 \text{ \AA}$ to measure the stellar kinematics. We apply the Voronoi binning algorithm (Cappellari & Copin 2003) 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 $0.67 \text{ \AA}$) 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). After fitting the spectra with penalized pixel fitting (pPXF), we reestimate the S/N as the ratio of the median flux and the rms residual from the fit. The resulting S/N map for the 135 Voronoi bins is shown in Figure 1. The resulting Ca II regions of the spectra for three representative bins are shown (black curves) in Figure 2.

We measure the stellar LOSVD within each spatial bin using the pPXF method of Cappellari (2016). The LOSVD is parameterized as a Gauss–Hermite series⁷ up to order n ,

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

⁷ Note that the pPXF method described in Cappellari & Emsellem (2004) only allows $n = 2, 4$, or 6 . The version described in Cappellari (2016) allows arbitrary n .

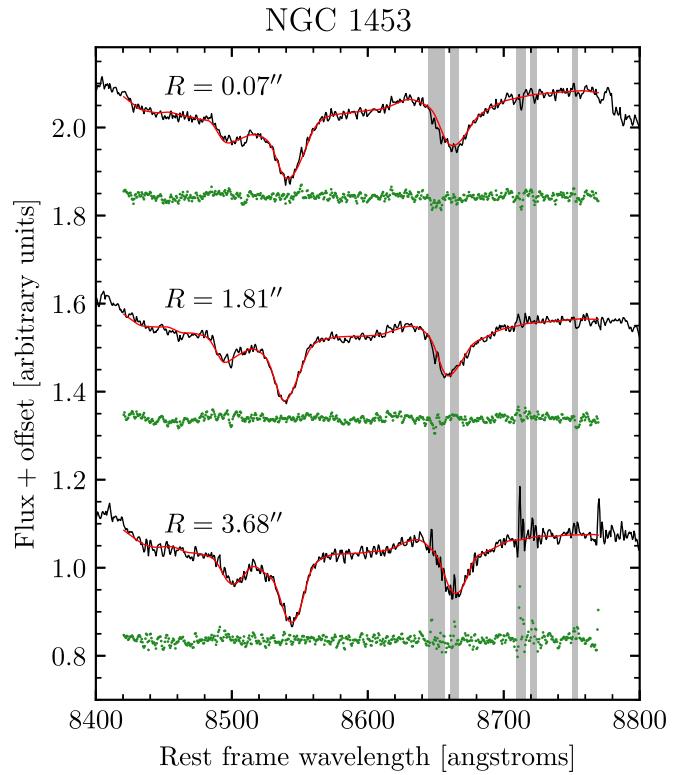


Figure 2. Ca II triplet region of the Gemini GMOS IFS spectra (black) for three representative bins at three locations of NGC 1453: center with S/N = 143 (top), $1''.81$ from center with S/N = 130 (middle), and $3''.68$ from center with S/N = 112 (bottom). The stellar template broadened by the best-fit LOSVD (red) is overlaid on each observed spectrum. The fit is performed over the rest wavelength range of $8420\text{--}8770 \text{ \AA}$ centered around the Ca II triplet absorption lines, excluding the gray shaded regions of improperly subtracted sky lines. The fit residuals (green circles) are shifted by an arbitrary amount for clarity.

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 & Franx (1993).

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 CaT library of 706 stars (Cenarro et al. 2001) 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. (2002). 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 $\sim 5 \text{ km s}^{-1}$ and the higher moments by ~ 0.01 , all well within the measurement errors. Our kinematic moments determined from the Ca II triplet region are therefore robust to template choices, similar to the findings in Barth et al. (2002). The stellar spectra of the CaT library cover the wavelength range of $8348\text{--}9020 \text{ \AA}$ with a spectral resolution of $1.5 \text{ \AA}$ FWHM.

The resulting stellar template broadened by the best-fit LOSVD is shown for each of the three example bins in Figure 2 (red curves). We use a bootstrap approach to determine the error bars on the kinematic moments of each LOSVD. For

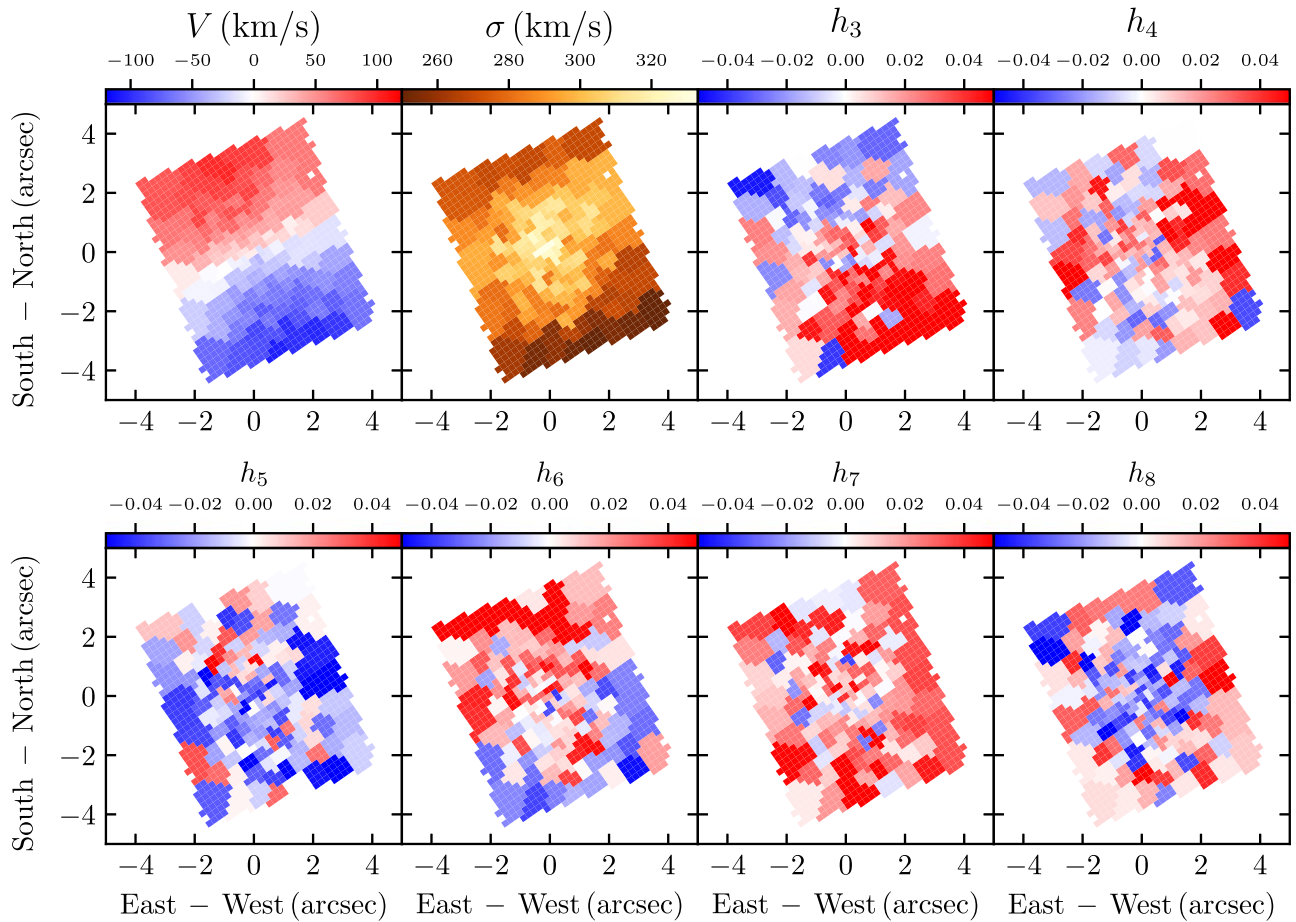


Figure 3. Maps of the stellar kinematics measured from the Gemini GMOS IFS over 135 spatial bins in the central $5'' \times 7''$ of NGC 1453. Each panel shows one of the eight velocity moments in the Gauss–Hermite expansion of the LOSVD: velocity V , velocity dispersion σ , and the h_3 to h_8 higher moments. The velocity map shows a regular rotation pattern with $|V|$ reaching $\sim 100 \text{ km s}^{-1}$, and the σ map shows a central peak of $\sim 325 \text{ km s}^{-1}$. North is up and east is to the left.

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%–100% larger than the Monte Carlo errors. See Section 4 of Ene et al. (2019) for a detailed discussion.

The maps of the eight kinematic moments, V , σ , h_3 , ..., h_8 , are shown in Figure 3. The velocity map shows a regular rotation pattern with $|V|$ reaching $\sim 100 \text{ km s}^{-1}$, and the σ map shows a central peak of $\sim 325 \text{ km s}^{-1}$. 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 4.

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 $107'' \times 107''$ field of view. Three dither positions of equal exposure time were used to obtain contiguous coverage of NGC 1453. We interleaved a 10-minute exposure on sky and two 20-minute exposures on target for a 2 hr total on-source exposure time. The spectral range spans 3650–5850 Å, covering the Ca HK region, the G -band region, H β , Mgb, and several Fe absorption features.

Individual central fibers have S/N above 50, while the outer fibers are binned spatially to achieve an S/N threshold of 20 for the fainter outskirts of the galaxy. A similar procedure to that in Section 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 (Sánchez-Blázquez et al. 2006; Falcón-Barroso et al. 2011) as stellar templates and ran pPXF over the full library for each spectrum. Further details are described in Ma et al. (2014) and Veale et al. (2017a).

As can be seen in Figure 4 here and Figure 21 of Ene et al. (2019), the kinematic moments in the innermost Mitchell bins match well with the GMOS moments.

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* WFC3/IR (Goullaud et al. 2018). The observations of NGC 1453 had a total exposure time of 2496 s, 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 $0''.128 \text{ pixel}^{-1}$ 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. (2018).

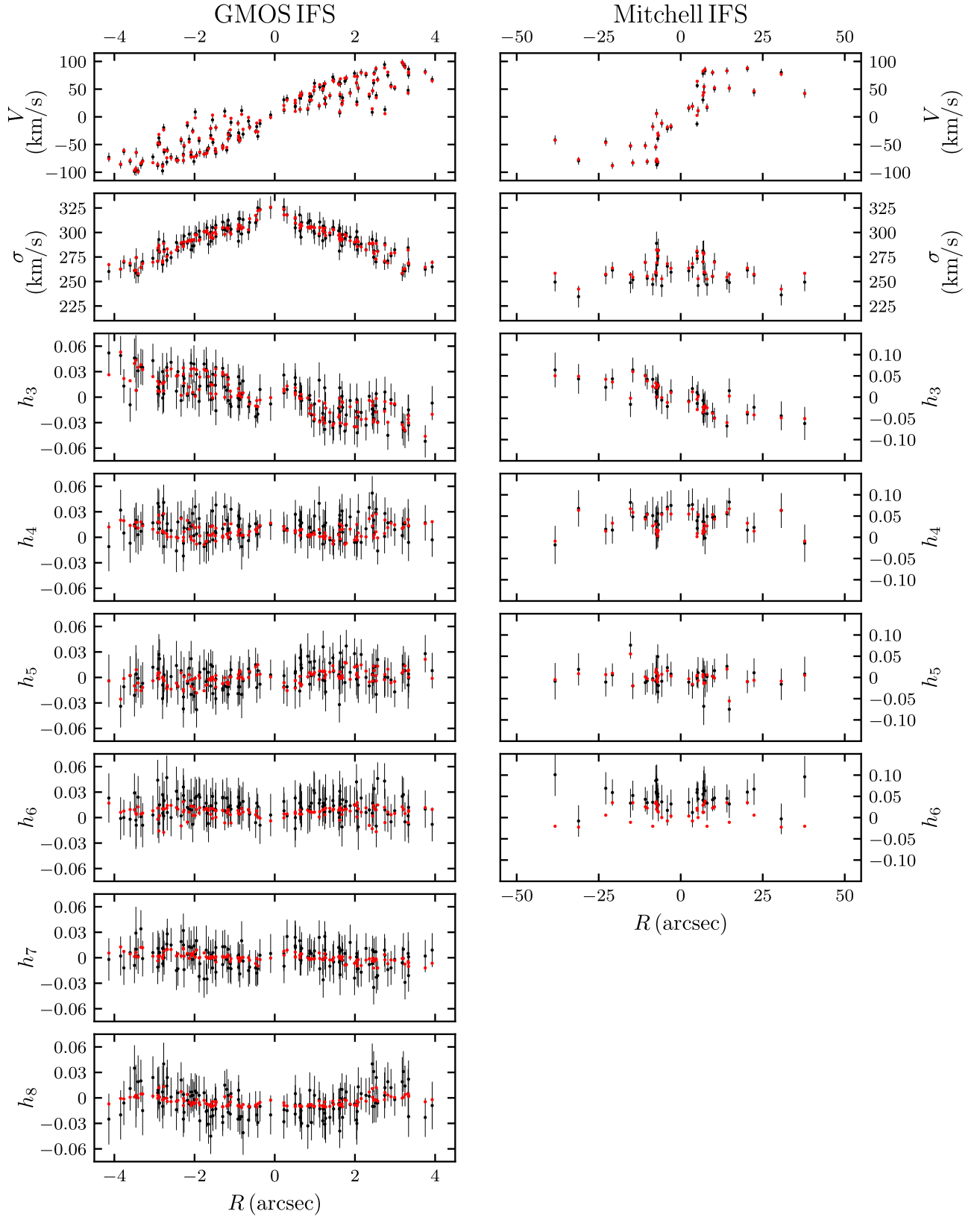


Figure 4. (Unfolded) radial profile of the stellar kinematics determined from GMOS (left) and Mitchell (right) observations (black), and kinematics predicted by the best-fit mass model (red) with black hole mass $M_{\text{BH}} = 2.9 \times 10^9 M_{\odot}$, stellar mass-to-light ratio $M^*/L_{\text{F110W}} = 2.09$ (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 $+90^\circ$ of the photometry position angle (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.

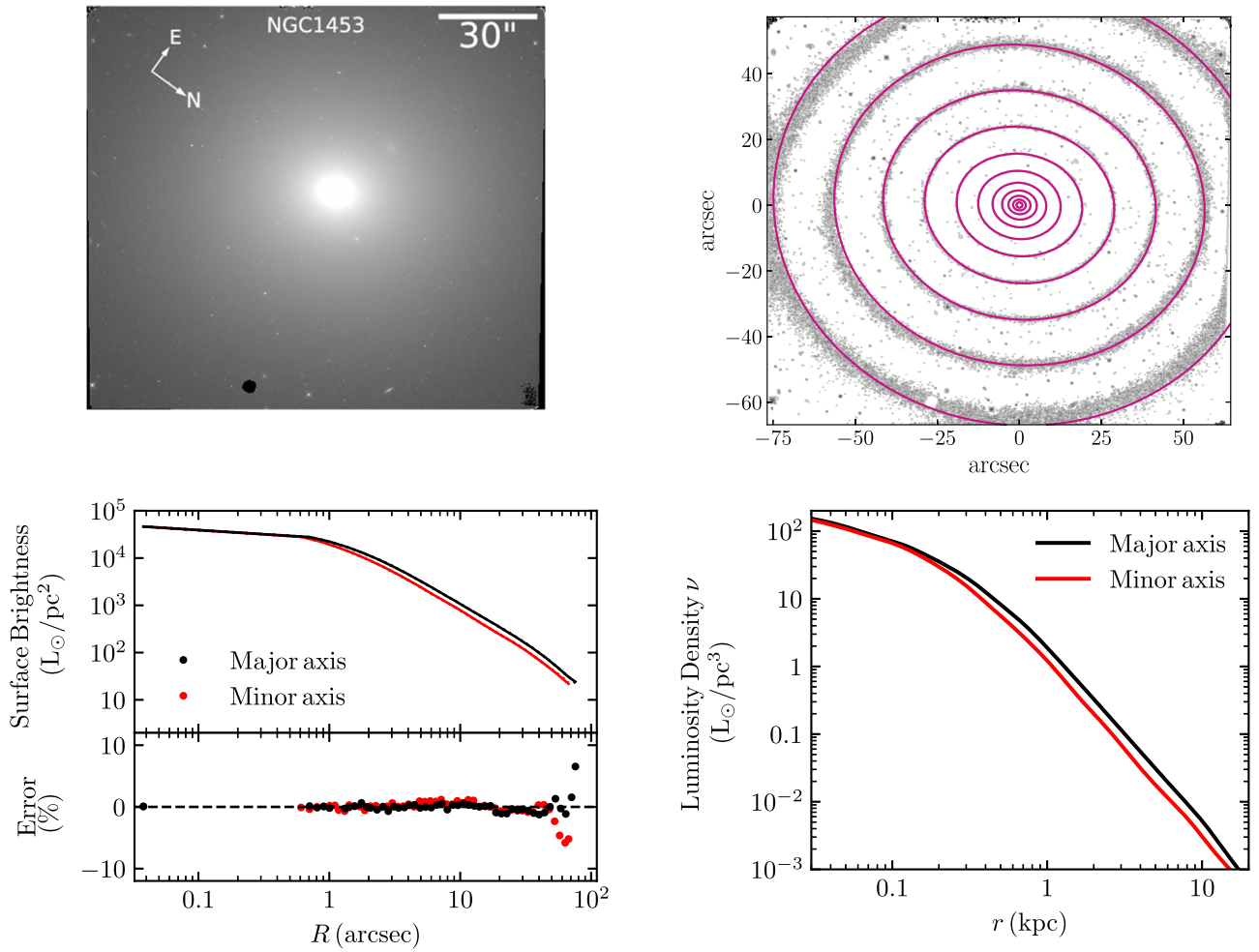


Figure 5. Top left: F110W-band *HST* image of NGC 1453 used for our photometry (Goullaud et al. 2018). The image is oriented so that the $+x$ -axis lies at $30^\circ 1$ east of north. Top right: isophotes of the *HST* WFC3 IR image of NGC 1453 (black) and the best-fit MGE model (magenta). The isophotes have no measurable deviation from purely elliptical contours (Goullaud et al. 2018). Bottom left: 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 $\sim 1\%$ except at large radii beyond $50''$. Bottom right: deprojected 3D luminosity density for an oblate axisymmetric model viewed edge-on for the best-fit MGE model.

The isophotes of NGC 1453 are very regular (top panel of Figure 5) with a mean ellipticity of 0.17 ± 0.001 . The PA changes with radius mildly from $27^\circ 9 \pm 1^\circ 0$ (east of north) at $1''$ to $36^\circ 1 \pm 0^\circ 4$ at $79'' 5$, with a luminosity-weighted average of $30^\circ 1 \pm 0^\circ 2$. We fit the surface brightness using the Multi-Gaussian Expansion (MGE) method (Emsellem et al. 1994; Cappellari 2002) with a sum of 2D Gaussian components that share a common center and PA:

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

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' , and 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 five nearly circular Gaussian components (with axis ratios > 0.98), obtained by fitting the PSF from Goullaud et al. (2018). The MGE fitting routine by default determines the PA

using the central region of the galaxy. As a result, it chooses a PA of $28^\circ 5$, slightly different from the mean value $30^\circ 1$ quoted in Goullaud et al. (2018). We repeated the MGE fit with the PA fixed to $30^\circ 1$ and found a virtually identical fit. We choose to use $28^\circ 5$, the value from the MGE fitting routine.

Our best-fit MGE to the surface brightness of NGC 1453 consists of 10 Gaussian components, which are summarized in Appendix A and plotted in Figure 5 (bottom 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 $R_e = 19'' 6 \approx 4.8$ kpc, very similar to $R_e = 21'' 9$ from Ene et al. (2018) 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) (Binney 1985). 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 $\psi = 90^\circ$. 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 the bottom right panel of Figure 5.

Dust was not observed in the central region of NGC 1453 in our WFC3 data. The mean optical and UV colors of NGC 1453 are typical of those of evolved, red giant ellipticals of similar masses (e.g., Faber et al. 1989; Loubser & Sánchez-Blázquez 2011). Annibali et al. (2007) 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 (J. B. Jensen et al. 2020, in preparation).

4. Schwarzschild Orbit Models

We use the orbit superposition method of Schwarzschild (1979) through the implementation described by van den Bosch et al. (2008). 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 (Gould et al. 2003), 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 M. E. Quenneville et al. (2020, in preparation). 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 M. E. Quenneville et al. (2020, in preparation).

4.1. The Axisymmetric Limit

NGC 1453 is a fast rotator with regular elliptical isophotes (Figure 5) and no significant misalignment between the projected rotation axis and photometric minor axis (Ene et al. 2018). 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. (2008) 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 nonaxisymmetric orbits from our axisymmetric models.⁸

⁸ 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 Figure 2 of Schwarzschild 1993). 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.

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 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 $|\psi - 90^\circ| = 10^{-3}$ or 10^{-2} , assuming that these values were close enough to 90° to generate axisymmetry. For NGC 1453, however, we find even $|\psi - 90^\circ| = 10^{-3}$ to be sufficiently far away from 90° to allow for long-axis tubes in the orbit start space, hence violating axisymmetry. We instead choose $\psi = (90 + 10^{-9})^\circ$ 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 symmetrized 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 M. E. Quenneville et al. (2020, in preparation) 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 7.2.

4.2. Orbit Library and Phase-space Sampling

As described in Schwarzschild (1993) and van den Bosch et al. (2008), 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 4.1, we use $\psi = (90 + 10^{-9})^\circ$ 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 $\phi = 1^\circ$. For reference, when the viewing angles $\psi = 90^\circ$ and $\theta = 90^\circ$ are used, the choices of $\phi = 0^\circ$ and $\phi = 90^\circ$ 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 data set, as they have different PSFs. For each, we assume a single circularly symmetric Gaussian with an FWHM of $0''.7$ for the GMOS kinematics and $1''.2$ for the Mitchell kinematics.

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 eight moments V , σ , h_3 , ..., h_8 measured from the GMOS spectra as described in Section 2.1 and shown in Figure 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 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 occurred when eight 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 six Gauss–Hermite moments measured from the Mitchell spectra as constraints (Section 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 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 Section 7.1 below.

5. Results: Mass Model Search

5.1. Mass Model

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

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

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

$$M_{\text{enc}}(r) = \frac{V_c^2}{G} \frac{r^3}{r^2 + R_c^2} \quad (4)$$

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 to 18.8 kpc).

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 B.

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 A, and q_i is the intrinsic axis ratio between short axis and 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 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 NGC 1453, we search coarsely over M_{BH} , M^*/L_{F110W} , and M_{15} but finely over the

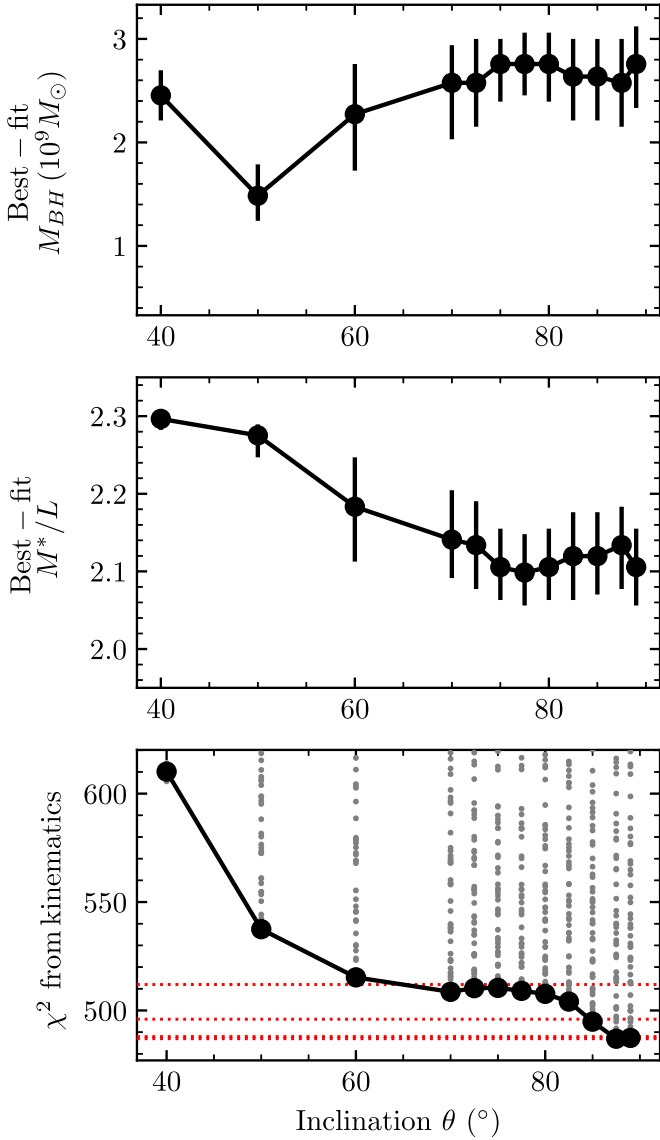


Figure 6. Best-fit M_{BH} (top) and M^*/L_{F110W} (middle) with 1σ confidence intervals and the corresponding marginalized 1D χ^2 (bottom) as a function of the inclination angle θ . The gray points in the bottom panel denote the χ^2 of individual models within the grid, and the red horizontal dashed lines denote the conventional 0, 1σ , 3σ , and 5σ confidence levels corresponding to $\Delta\chi^2 = 0, 1, 9$, and 25 . The halo is fixed to $M_{15} = 6.3 \times 10^{11} M_\odot$ in this plot for illustrative purposes; similar dependence is found for other halo masses we examined.

inclination. This grid includes 11 values of M_{BH} from 0.0 to $6 \times 10^9 M_\odot$ in steps of $6 \times 10^8 M_\odot$, eight values of M^*/L_{F110W} from 1.6 to 2.3 (in solar units) in steps of 0.1 , and three values of M_{15} : $(2.8, 6.3, 11.2) \times 10^{11} M_\odot$, corresponding to $V_c = 400, 600$, and 800 km s^{-1} with $R_c = 15 \text{ kpc}$. 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 6 shows that nearly edge-on viewing angles are strongly preferred. When $M_{15} = 6.3 \times 10^{11} M_\odot$, $\theta = 89^\circ$ 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_{BH} and M^*/L_{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 limit. We therefore fix the inclination to be nearly edge-on with $\theta = 89^\circ$ as we sample over halos below.

5.4. Black Hole, Stars, and Dark Matter Halo

With the inclination fixed to be nearly edge-on with $\theta = 89^\circ$, we search the three mass parameters, M_{BH} , M^*/L_{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_{BH} and M^*/L_{F110W} .

The primary grid has $16 \times 15 \times 13 = 3120$ models for M_{BH} , M^*/L_{F110W} , and M_{15} . This grid samples M_{BH} linearly from 0 to $6 \times 10^9 M_\odot$ in steps of $4 \times 10^8 M_\odot$, M^*/L_{F110W} from 1.60 to 2.30 (in solar units) in equal steps of 0.05 , and the enclosed halo mass from $M_{15} = 2.79 \times 10^{11} M_\odot$ to $11.16 \times 10^{11} M_\odot$ by varying the circular velocity roughly linearly from $V_c = 400$ to 800 km s^{-1} (for $R_c = 15 \text{ kpc}$).

For the finer grid, we first determine the M_{BH} and M^*/L_{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_{BH} and M^*/L_{F110W} , where the spacing between models is half of that of the primary grid, and M_{BH} is sampled over a range of $3 \times 10^9 M_\odot$ in steps of $2 \times 10^8 M_\odot$, and M^*/L_{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 $176 \times 13 = 2288$ additional models are run.

We perform the interpolation and marginalization described in Appendix B to determine the best-fit values and uncertainties in M_{BH} , M^*/L_{F110W} , and M_{15} from these 5408 models. The resulting χ^2 landscapes are displayed in Figure 7. We find the best-fit mass parameters to be $M_{\text{BH}} = (2.9 \pm 0.4) \times 10^9 M_\odot$, $M^*/L_{\text{F110W}} = (2.09 \pm 0.06) M_\odot/L_\odot$, and $M_{15} = (7.0 \pm 0.7) \times 10^{11} M_\odot$. For comparison, if the best-fit parameters are chosen by finding the range of models where the χ^2 rises by $\Delta\chi^2 \leq 1$ from the minimum value, as was frequently done in prior M_{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 – 2 : $M_{\text{BH}} = (3.0 \pm 0.2) \times 10^9 M_\odot$, $M^*/L_{\text{F110W}} = (2.06 \pm 0.03) M_\odot/L_\odot$, and $M_{15} = (7.4 \pm 0.4) \times 10^{11} M_\odot$. These values are tabulated in Table 1.

Figure 8 presents a clear view of the χ^2 landscape over the wide range of M_{BH} covered by our grid. It shows that models with small black hole masses are highly disfavored. In particular, $M_{\text{BH}} = 0$ has $\Delta\chi^2 = 75.5$ above the minimum, corresponding to the 8.7σ confidence level. This result will be further discussed in Section 7.2.

In the best-fit mass model for NGC 1453, the enclosed stellar mass is equal to M_{BH} , $2M_{\text{BH}}$, $3M_{\text{BH}}$, and $5M_{\text{BH}}$ 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 .

6. Results: Best-fit Mass Model

6.1. Stellar Kinematics

Our best-fit mass model (red points in Figure 4) provides an excellent fit to the observed stellar kinematics (black). Both the

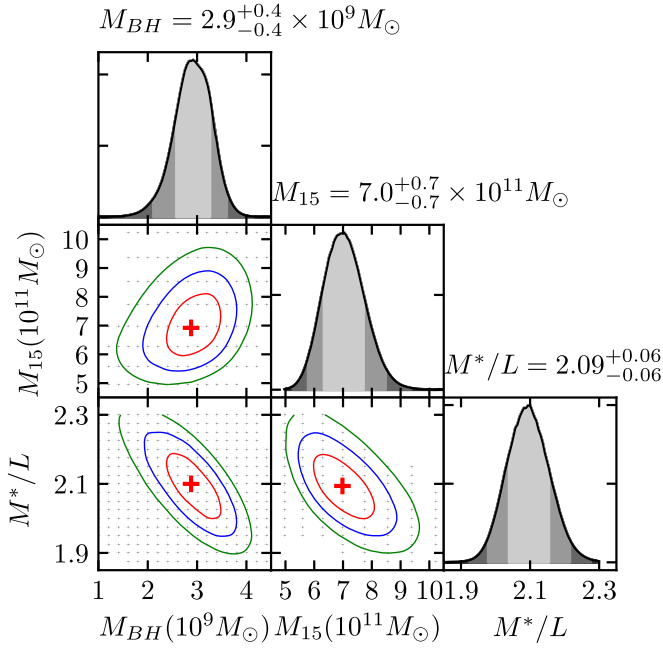


Figure 7. Marginalized 1D and 2D likelihood distributions from the grids of M_{BH} , M_{15} , and M^*/L_{F110W} described in Section 5.4. The 1σ , 2σ , and 3σ confidence intervals corresponding to the 68, 97, and 99.5 percentile confidence levels are shown as red, blue, and green curves in the 2D panels and as a different shade of gray in the 1D panels. The extracted best-fit values and 1σ confidence interval are shown above each 1D panel.

Table 1

Best-fit Values of the Black Hole Mass, Stellar Mass-to-light Ratio in the F110W Band, and Dark Matter Mass Enclosed within 15 kpc

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

Note. The center column presents values determined through interpolation and marginalization as described in Appendix 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.

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 eight measured moments for each of the 135 GMOS bins and six 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-antisymmetrized and the even moments have been point-symmetrized according to the prescription in Appendix A of van den Bosch & de Zeeuw (2010). 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 $366.2/(135 \times 8) = 0.34$, and the reduced χ^2 from Mitchell alone is $105.3/(38 \times 6) = 0.46$. The high moments that were constrained to be zero have an associated reduced χ^2 of $21.5/768 = 0.03$.

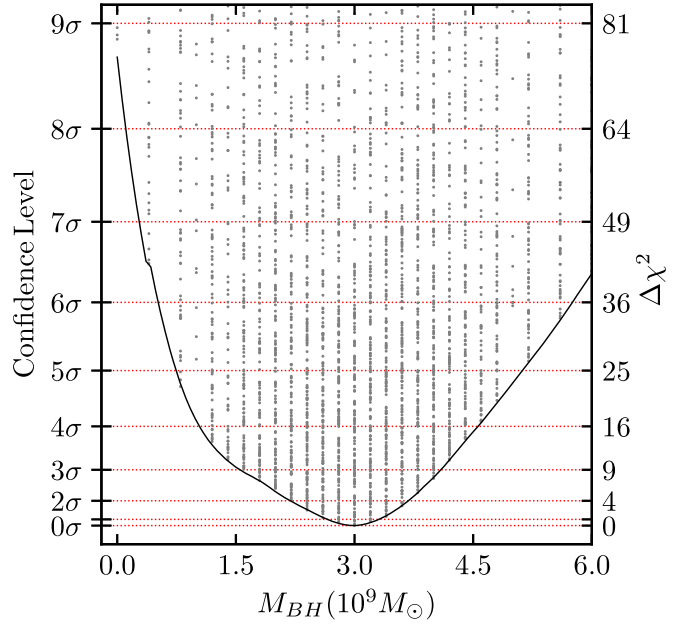


Figure 8. χ^2 as a function of M_{BH} for the full range of M_{BH} explored in this paper. The χ^2 is obtained by marginalizing over the other two mass parameters, M_{15} and M^*/L_{F110W} , as described in Appendix B. Models with $M_{\text{BH}} = 0$ are highly disfavored, with a $\Delta\chi^2 = 75.5$ relative to the best-fit M_{BH} , corresponding to the 8.7σ confidence level.

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

$$\sigma_t^2 = \frac{\sigma_\theta^2 + \sigma_\phi^2}{2}, \quad \beta = 1 - \frac{\sigma_t^2}{\sigma_r^2},$$

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

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. (2014), β_r from Peletier et al. (2007), and σ_r/σ_t from Walsh et al. (2015) all excluded the $\langle v_\phi \rangle^2$ term, while β_{rot} from Krajnović et al. (2018) and Thomas et al. (2014) and σ_r/σ_t from Gebhardt et al. (2003) 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 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 $\sim 100 \text{ km s}^{-1}$, which is much below the dispersion σ shown in Figure 4. The term $\langle v_\phi \rangle^2$

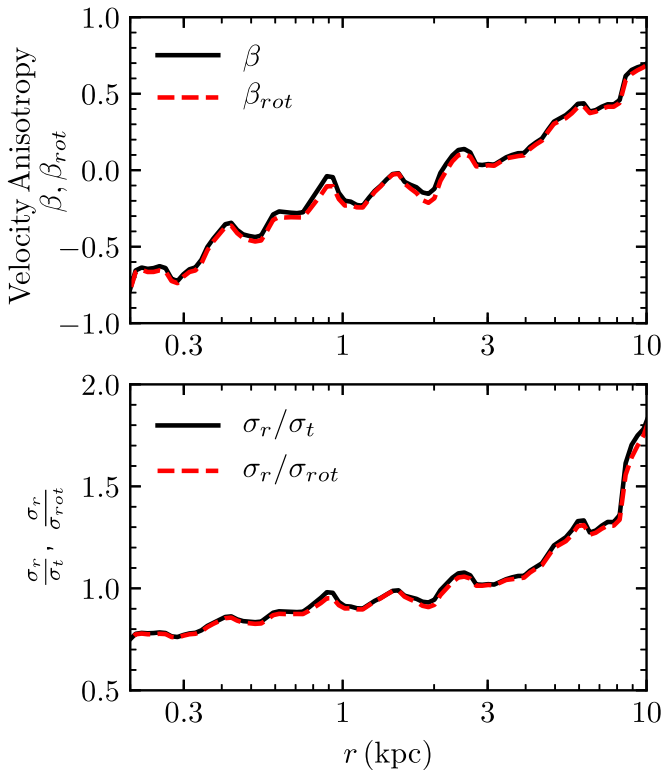


Figure 9. Velocity anisotropy β (top) and σ_r/σ_t (bottom) as a function of radius for the best-fit model of NGC 1453. The orbits within the central ~ 1 kpc are preferentially tangential with $\sigma_r/\sigma_t < 1$ and $\beta < 0$. The orbits become increasingly radial beyond the effective radius (≈ 5 kpc). Over all radii, β traces β_{rot} and σ_t traces σ_{rot} because $\langle v_\phi \rangle^2 \ll \sigma_\phi^2$.

therefore has negligible impact on the value of σ_{rot} and β_{rot} at all radii, and $\sigma_{\text{rot}} \approx \sigma_t$ and $\beta \approx \beta_{\text{rot}}$ at all radii.

Thomas et al. (2014) studied 11 massive elliptical galaxies with axisymmetric Schwarzschild models. Six of those galaxies had stellar cores and exhibited strongly tangential anisotropies ($\beta_{\text{rot}} < -0.5$) in the core regions and highly radial anisotropies ($\beta_{\text{rot}} \sim 0.5$) well outside the cores. Similar trends in the anisotropy were found in MASSIVE survey galaxy NGC 1600 (Thomas et al. 2016). 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 that is tangentially biased (Begelman et al. 1980). We observe similar behavior in NGC 1453, suggesting that its core may have also been depleted through core scouring.

7. Discussion

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 & Franx (1993), 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 owing 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_{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, four, six, and eight GH moments are measured from the GMOS data with pPXF, and those four, six, and eight GH moments are fit with orbit models to infer M_{BH} . In the fourth case, eight 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 Section 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_{BH} and M^*/L_{F110W} . Our aim here is to test the effects on the measured M_{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 10. The corresponding LOSVDs for three representative GMOS bins in each case are shown in Figure 11. The marginalized χ^2 versus M_{BH} for the four cases are shown in Figure 12 (left panel), and the best-fit M_{BH} are listed in Table 2 under “Berkeley Version.” We note that while the best-fit M_{BH} changes by only $\sim 20\%$ (in the range of $(2.6\text{--}3.2) \times 10^9 M_\odot$) as the truncation order is varied, the confidence level is improved significantly when more input moments are used, and the errors on M_{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 10. These moments are determined from the LOSVD of each bin (e.g., Figure 11) through the direct integration described by van der Marel & Franx (1993). Although only moments up to the truncation order are used for constraining the model LOSVDs, arbitrary higher moments can be computed. The lowest four 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 four moments are fit during modeling 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 four moments are fit by the orbit-based models (green

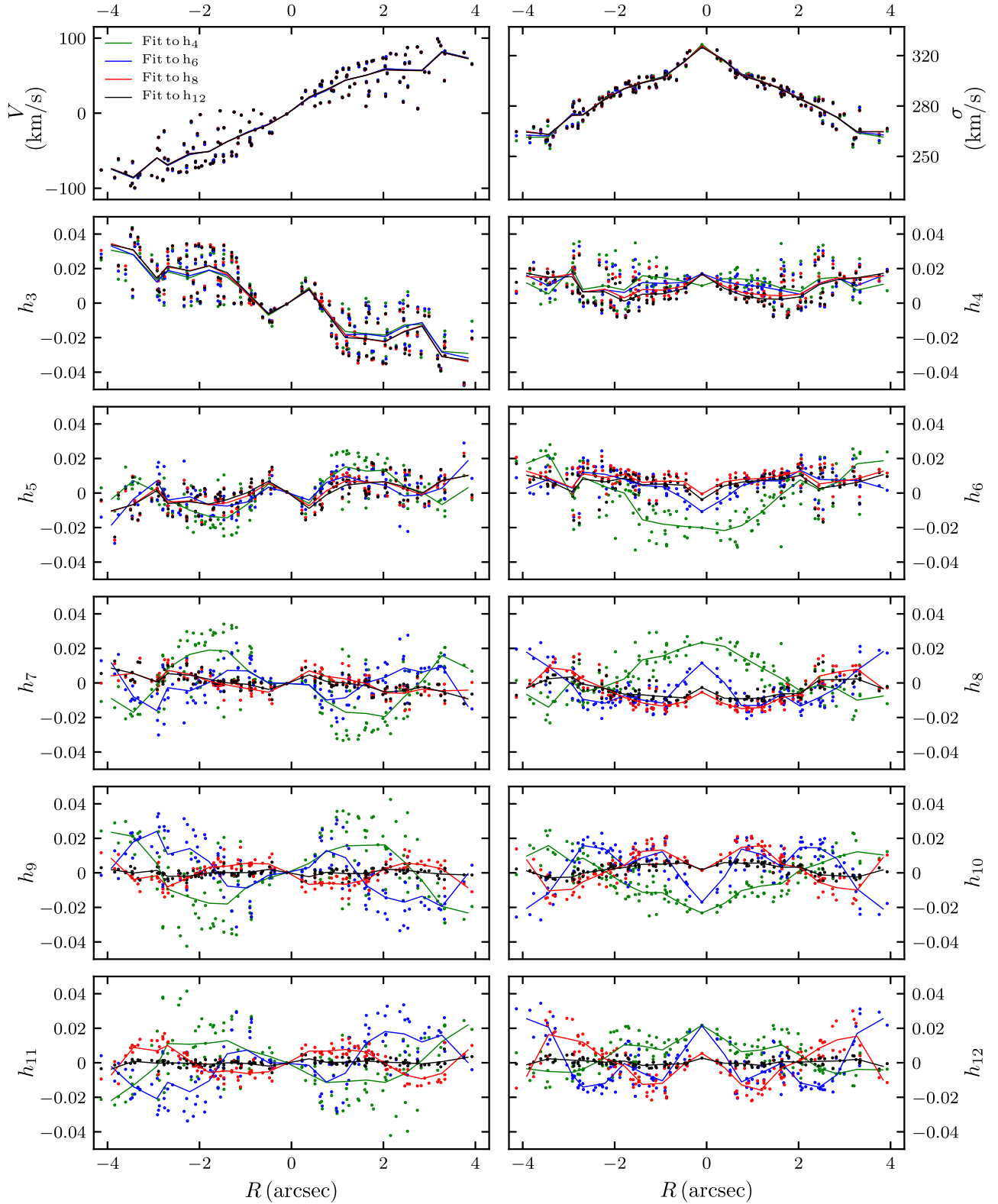


Figure 10. Dependence of the first 12 Gauss–Hermite moments predicted by the best-fit orbit model on the assumed truncation order applied to the GMOS data and used as input constraints. The four colors show the progression of increasing truncation order: up to h_4 (green), h_6 (blue), h_8 (red), and 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 vs. M_{BH} for the four cases are shown in Figure 12 (left panel). See text in Section 7.1 for details.

points), the unconstrained fifth moment and beyond deviate strongly from the black points. Similarly, when six (blue) or eight (red) moments are used as constraints during the

modeling, the seventh or ninth moment and beyond also show deviations from the black points. Importantly, the deviation from 0 is not random; instead, the unconstrained moments are

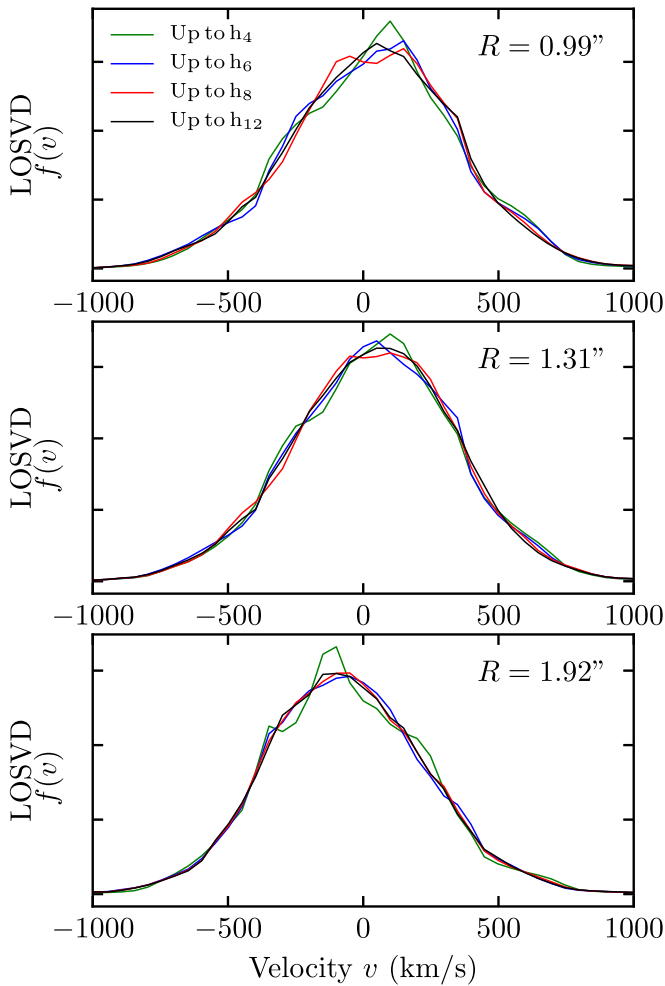


Figure 11. LOSVDs for three representative GMOS bins predicted by the best-fit orbit models. Each panel compares the LOSVDs from the four models described in Section 7.1 and shown in Figure 10, where the number of Gauss-Hermite moments fit by the dynamical models varies from 4 (green), 6 (blue), 8 (red), to 12 (black). When only four moments are constrained, the LOSVDs have the most pronounced irregular features owing to the unconstrained h_5 and beyond seen in Figure 10. These unobserved features are gradually reduced when higher moments are used to constrain the model.

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 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 11. For all three representative GMOS bins shown, the model LOSVDs have the most pronounced irregular features when only four 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_{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., Verolme et al. 2002; van den Bosch & de Zeeuw 2010; van den Bosch et al. 2012; Walsh et al. 2012, 2015, 2016, 2017; Seth et al. 2014; Thater et al. 2017, 2019; Ahn et al. 2018; Krajnović et al. 2018), while a

few have used the first six moments (e.g., Cappellari et al. 2002, 2009; Krajnović et al. 2009). 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_{BH} in other work. However, we recommend that similar tests be performed in future work.

7.2. The $M_{\text{BH}} = 0$ Model: Comparison to Ahn et al. (2018)

Recently, Ahn et al. (2018) 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 $M_{\text{BH}} = 0$, which was inconsistent with their best-fit $M_{\text{BH}} = 4.2^{+2.1}_{-1.7} \times 10^6 M_{\odot}$ from Jeans modeling and an orbit code that is intrinsically axisymmetric. Various tests were performed, but none explained the discrepancy. They speculated about a “numerical artifact” in the triaxial Leiden code and favored the nonzero M_{BH} from Jeans and axisymmetric orbit modeling.

As discussed in Sections 4 and 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 Sections 5 and 6. Even though our final outcome in Figure 8 shows $M_{\text{BH}} = 0$ to be disfavored at the 8.7σ confidence level, we also encountered difficulties in constraining M_{BH} in the case of NGC 1453 when we ran the original triaxial code using similar settings to those of Ahn et al. (2018), that is, choosing $|\psi - 90^\circ| \geq 10^{-3}$, including box orbits from the default library, not axisymmetrizing the loop orbits, and using four 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 12. The overall constraint on the NGC 1453 M_{BH} is weak, with the lowest χ^2 occurring at $M_{\text{BH}} = 0$ and another local χ^2 minimum at $M_{\text{BH}} \sim 1.5 \times 10^9 M_{\odot}$. 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 12.

In view of the importance of constraining higher Gauss-Hermite moments (Section 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 12. The χ^2 minimum at $M_{\text{BH}} = 0$ 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_{BH} increases monotonically and shows no convergence even at order 12, as listed in Table 2. In comparison, models from the “Berkeley Version” in the left panel of Figure 12 and in Table 2 have better-behaved χ^2 contours.

7.3. Comparison to Jeans Modeling

In Ene et al. (2018), we applied the method of Jeans Anisotropic Modeling (JAM; Cappellari 2008) 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 with those of axisymmetric orbit models for regular fast rotators like NGC 1453 (Cappellari et al. 2010).

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

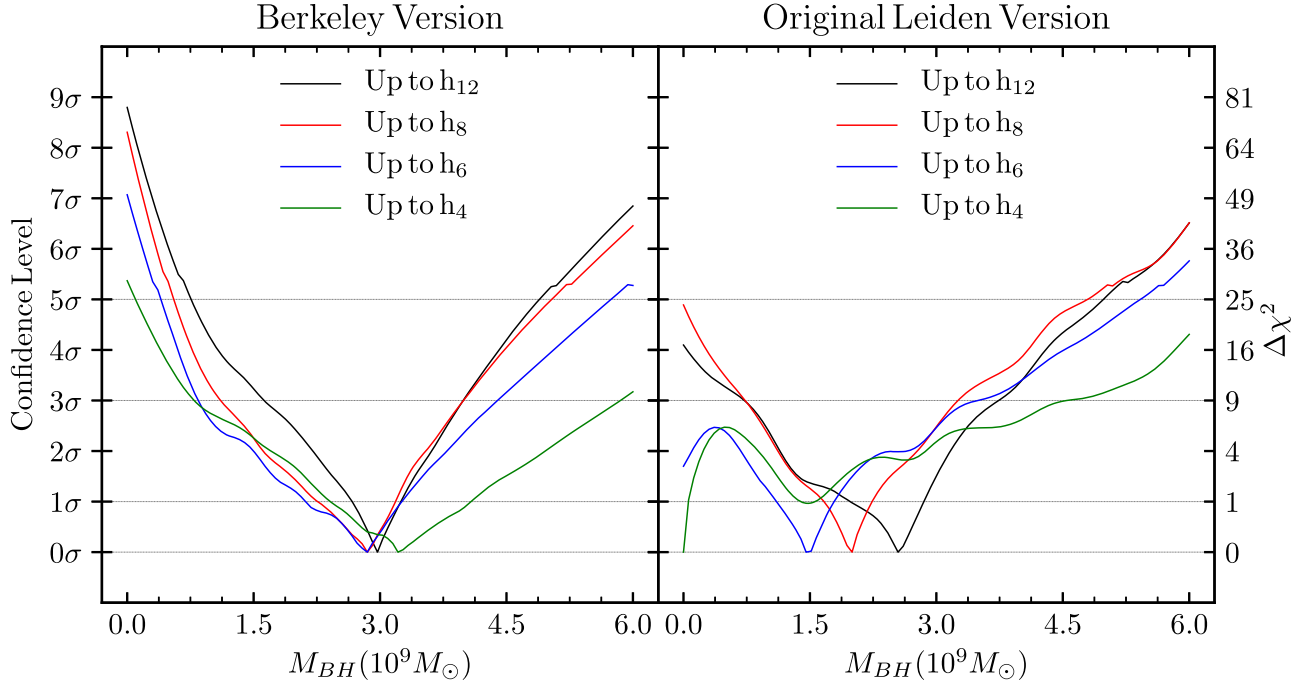


Figure 12. Illustration of the increasing constraints on M_{BH} provided by progressively higher Gauss–Hermite moments used to represent the LOSVDs. Two versions of the orbit code with different settings are shown: the Berkeley version described in Section 4.1 (left), and the original triaxial Leiden 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 to that in Ahn et al. (2018) for M59-UCD3 and prefers $M_{\text{BH}} = 0$. The $M_{\text{BH}} = 0$ minimum disappears as moments beyond h_4 are applied in the original code (right), but the best-fit M_{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_{BH} while leaving the central value largely unchanged. Note that 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 Figure 8 uses a linear scale in χ^2 .

Table 2
Best-fit Black Hole Mass and 1σ (68%) Confidence Intervals for the Eight Cases Shown in Figure 12

No. of Constrained Moments	Berkeley Version $M_{\text{BH}} (10^9 M_{\odot})$	Original Leiden Version $M_{\text{BH}} (10^9 M_{\odot})$
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

Note. The four-moment Leiden run results in a χ^2 minimum at $M_{\text{BH}} = 0$. For all other runs, the quoted confidence intervals here are properly marginalized through the routine described in Appendix B.

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 Figure 5).

The JAM modeling in Ene et al. (2018) used the distance 56.4 Mpc from the 2MASS redshift survey. After adjusting to 51.0 Mpc, the best-fit parameters from JAM are $M_{\text{BH}} = (2.98 \pm 0.23) \times 10^9 M_{\odot}$ and $M^*/L_{\text{F110W}} = 2.28 \pm 0.04$ (with 1σ errors). The M_{BH} value is within the 1σ confidence interval of our best-fit value, while the M^*/L_{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 $V_c = 364 \pm 45 \text{ km s}^{-1}$ with $R_c = 6.5 \pm 2.5 \text{ kpc}$. 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 $(17.4 \pm 4.3) \times 10^7 M_{\odot} \text{ kpc}^{-3}$, while ours is half at $(9.9 \pm 1.0) \times 10^7 M_{\odot} \text{ kpc}^{-3}$. 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 6.2, but we find 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.

7.4. Black Hole Scaling Relations

The SMBH at the center of NGC 1453 lies 0.32 dex above the mean $M_{\text{BH}}-\sigma$ scaling relation from McConnell & Ma (2013), 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 ~ 3 –6

than predicted by their respective galaxy velocity dispersion. These eight 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. (2018). This fit was performed using Imfit (Erwin 2015). The scaling relation between M_{BH} and r_b for a sample of 21 massive cored ETGs is found to be $\log_{10} M_{\text{BH}} = 10.27 + 1.17 \log_{10}(r_b/\text{kpc})$ with an intrinsic scatter of 0.29 dex (Thomas et al. 2016). Our inferred M_{BH} for NGC 1453 is only 0.077 dex below this relation.

7.5. Gas Kinematics

In Pandya et al. (2017), we observed the kinematics of warm ionized gas out to ~ 8 kpc within NGC 1453 by tracing the [O II] $\lambda 3727$ emission line using the spectra obtained with the Mitchell IFS. This warm gas was found to rotate with a $\text{PA}_{\text{gas}} \sim 312.5^\circ$, roughly perpendicular to the stellar rotation along $\text{PA}_{\text{stars}} \sim 35^\circ$. 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 $\sim 200 \text{ km s}^{-1}$ and a comparable velocity dispersion, giving an rms velocity of $\sim 300 \text{ km s}^{-1}$, similar to that of the stars studied in this paper.

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 (Figures 3 and 4). The two sets of kinematic data together span about two orders of magnitude in radial extent, from $0''.3$ to $76''$ (~ 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 modeling 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 (Figures 7 and 8). The best-fit model for NGC 1453 has a black hole mass $M_{\text{BH}} = (2.9 \pm 0.4) \times 10^9 M_\odot$, a stellar mass-to-light ratio (in F110W band) $M^*/L_{\text{F110W}} = (2.09 \pm 0.06) M_\odot/L_\odot$, 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 (Figure 6).

We began the orbit modeling with the original triaxial Schwarzschild code of van den Bosch et al. (2008) 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 are discussed in Section 4; further details are described in M. E. Quenneville et al. (2020, in preparation).

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 (Figure 11), and the contributions from the unconstrained higher-order moments are not random but show spatial correlations (Figure 10). When the Berkeley version of the orbit code is used, we find that the confidence level on the M_{BH} determination for NGC 1453 is significantly improved when at least eight Gauss–Hermite moments are used as constraints: the 1σ confidence interval shrinks by a factor of ~ 2 relative to models with typical constraints on only V through h_4 (left panel of Figure 12). By contrast, the χ^2 landscape is not as well behaved when the original code is used with typical settings (right panel of Figure 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 changes 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 nonaxisymmetric galaxies.

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).

Table 3
Best-fit MGE Parameters to the NGC 1453 *HST* WFC3 IR Photometry

I_k ($L_\odot/\text{pc}^2$)	σ'_k (arcsec)	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

Note. Each Gaussian component is parameterized by a central surface density $I_k = L_k/2\pi\sigma_k'^2 q_k'$ (calculated using an absolute solar AB magnitude $M_{\odot, \text{F110W}} = 4.54$), dispersion σ'_k (in arcseconds), and axis ratio q'_k .

Appendix A MGE Parameters

We list in Table 3 the best-fit parameters of the 10 MGE components to our *HST* WFC3 IR photometry of NGC 1453 shown in Figure 5. The 10 Gaussians terms are assumed to have the same center and position angle of $28^\circ 5$.

We also performed an MGE fit to this photometry in Ene et al. (2018). 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 unphysical feature, we use the unregularized fit here.

Appendix B Interpolation and Marginalization

We perform an interpolation with cubic radial basis functions (RBFs) 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 Knysh & Korkolis (2016). The RBF interpolation is described by

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

where $\mathbf{x}$ describes a point in the parameter space and λ_i , $\mathbf{b}$, 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 10,000 points drawn from a uniform distribution over the parameter space. The covariance matrix of the 500 points with lowest predicted χ^2 is computed, and then the eigenvalues α_i and eigenvectors $\mathbf{m}_i$ of that matrix are computed. Finally, T is constructed with $T_i = \mathbf{m}_i/\sqrt{\alpha_i}$. Given this new T , λ_i , $\mathbf{b}$, 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 $\chi^2(\boldsymbol{\theta}, \boldsymbol{\psi})$ to $\chi^2(\boldsymbol{\theta})$, where $\boldsymbol{\psi}$ are the parameters we wish to eliminate and $\boldsymbol{\theta}$ are those that 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,

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

and thus

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

where N is the number of parameters in $\boldsymbol{\psi}$.

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

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

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 the appropriate percentiles:

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

where θ_+ and θ_- yield the upper and lower bounds to the cumulative likelihood, respectively, 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.”

ORCID iDs

Christopher M. Liepold  <https://orcid.org/0000-0002-7703-7077>
 Chung-Pei Ma  <https://orcid.org/0000-0002-4430-102X>
 Jonelle L. Walsh  <https://orcid.org/0000-0002-1881-5908>
 John P. Blakeslee  <https://orcid.org/0000-0002-5213-3548>

References

- Ahn, C. P., Seth, A. C., Cappellari, M., et al. 2018, *ApJ*, **858**, 102
 Annibali, F., Bressan, A., Rampazzo, R., Zeilinger, W. W., & Danese, L. 2007, *A&A*, **463**, 455
 Barth, A. J., Ho, L. C., & Sargent, W. L. W. 2002, *AJ*, **124**, 2607
 Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Natur*, **287**, 307
 Binney, J. 1985, *MNRAS*, **212**, 767
 Cappellari, M. 2002, *MNRAS*, **333**, 400
 Cappellari, M. 2008, *MNRAS*, **390**, 71
 Cappellari, M. 2016, *ARA&A*, **54**, 597
 Cappellari, M., & Copin, Y. 2003, *MNRAS*, **342**, 345
 Cappellari, M., di Serego Alighieri, S., Cimatti, A., et al. 2009, *ApJL*, **704**, L34
 Cappellari, M., & Emsellem, E. 2004, *PASP*, **116**, 138
 Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, *MNRAS*, **413**, 813
 Cappellari, M., McDermid, R. M., Bacon, R., et al. 2010, in AIP Conf. Proc. 1240, *Hunting for the Dark: The Hidden Side of Galaxy Formation*, ed. V. P. Debattista & C. C. Popescu (Melville, NY: AIP), 211
 Cappellari, M., Verolme, E. K., van der Marel, R. P., et al. 2002, *ApJ*, **578**, 787
 Cenarro, A. J., Cardiel, N., Gorgas, J., et al. 2001, *MNRAS*, **326**, 959

- Côté, P., Ferrarese, L., Jordán, A., et al. 2007, *ApJ*, **671**, 1456
- Crook, A. C., Huchra, J. P., Martimbeau, N., et al. 2007, *ApJ*, **655**, 790
- Davis, T. A., Greene, J., Ma, C.-P., et al. 2016, *MNRAS*, **455**, 214
- Davis, T. A., Greene, J. E., Ma, C.-P., et al. 2019, *MNRAS*, **486**, 1404
- Emsellem, E., Monnet, G., & Bacon, R. 1994, *A&A*, **285**, 723
- Ene, I., Ma, C.-P., McConnell, N. J., et al. 2019, *ApJ*, **878**, 57
- Ene, I., Ma, C.-P., Veale, M., et al. 2018, *MNRAS*, **479**, 2810
- Ene, I., Ma, C.-P., Walsh, J. L., et al. 2020, arXiv:2001.11046
- Erwin, P. 2015, *ApJ*, **799**, 226
- Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, *ApJL*, **875**, L1
- Faber, S. M., Tremaine, S., Ajhar, E. A., et al. 1997, *AJ*, **114**, 1771
- Faber, S. M., Wegner, G., Burstein, D., et al. 1989, *ApJS*, **69**, 763
- Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, *A&A*, **532**, A95
- Ferrarese, L., Côté, P., Jordán, A., et al. 2006, *ApJS*, **164**, 334
- Gebhardt, K., Adams, J., Richstone, D., et al. 2011, *ApJ*, **729**, 119
- Gebhardt, K., Richstone, D., Tremaine, S., et al. 2003, *ApJ*, **583**, 92
- Gould, N. I. M., Orban, D., & Toint, P. L. 2003, *ACM Trans. Math. Softw.*, **29**, 353
- Goulding, A. D., Greene, J. E., Ma, C.-P., et al. 2016, *ApJ*, **826**, 167
- Goullaoud, C. F., Jensen, J. B., Blakeslee, J. P., et al. 2018, *ApJ*, **856**, 11
- Graham, A. W., Erwin, P., Trujillo, I., & Asensio Ramos, A. 2003, *AJ*, **125**, 2951
- Greene, J. E., Janish, R., Ma, C. P., et al. 2015, *ApJ*, **807**, 11
- Greene, J. E., Veale, M., Ma, C.-P., et al. 2019, *ApJ*, **874**, 66
- Hill, G. J., MacQueen, P. J., Smith, M. P., et al. 2008, *Proc. SPIE*, **7014**, 701470
- Hook, I. M., Jørgensen, I., Allington-Smith, J. R., et al. 2004, *PASP*, **116**, 425
- Knysh, P., & Korkolis, Y. 2016, arXiv:1605.00998
- Kormendy, J., & Ho, L. C. 2013, *ARA&A*, **51**, 511
- Krajinović, D., Cappellari, M., McDermid, R. M., et al. 2018, *MNRAS*, **477**, 3030
- Krajinović, D., McDermid, R. M., Cappellari, M., & Davies, R. L. 2009, *MNRAS*, **399**, 1839
- Lauer, T. R., Faber, S. M., Richstone, D., et al. 2007, *ApJ*, **662**, 808
- Loubser, S. I., & Sánchez-Blázquez, P. 2011, *MNRAS*, **410**, 2679
- Ma, C.-P., Greene, J. E., McConnell, N., et al. 2014, *ApJ*, **795**, 158
- McConnell, N. J., & Ma, C.-P. 2013, *ApJ*, **764**, 184
- McConnell, N. J., Ma, C.-P., Graham, J. R., et al. 2011, *ApJ*, **728**, 100
- McConnell, N. J., Ma, C.-P., Murphy, J. D., et al. 2012, *ApJ*, **756**, 179
- Pandya, V., Greene, J. E., Ma, C.-P., et al. 2017, *ApJ*, **837**, 40
- Peletier, R. F., Falcón-Barroso, J., Bacon, R., et al. 2007, *MNRAS*, **379**, 445
- Rix, H.-W., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. 1997, *ApJ*, **488**, 702
- Rusli, S. P., Thomas, J., Saglia, R. P., et al. 2013, *AJ*, **146**, 45
- Saglia, R. P., Opitsch, M., Erwin, P., et al. 2016, *ApJ*, **818**, 47
- Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, *MNRAS*, **371**, 703
- Schwarzschild, M. 1979, *ApJ*, **232**, 236
- Schwarzschild, M. 1993, *ApJ*, **409**, 563
- Seth, A. C., van den Bosch, R., Mieske, S., et al. 2014, *Natur*, **513**, 398
- Shen, J., & Gebhardt, K. 2010, *ApJ*, **711**, 484
- Thater, S., Krajinović, D., Bourne, M. A., et al. 2017, *A&A*, **597**, A18
- Thater, S., Krajinović, D., Cappellari, M., et al. 2019, *A&A*, **625**, A62
- Thomas, J., Ma, C.-P., McConnell, N. J., et al. 2016, *Natur*, **532**, 340
- Thomas, J., Saglia, R. P., Bender, R., Erwin, P., & Fabricius, M. 2014, *ApJ*, **782**, 39
- van den Bosch, R. C. E., & de Zeeuw, P. T. 2010, *MNRAS*, **401**, 1770
- van den Bosch, R. C. E., Gebhardt, K., Gültekin, K., et al. 2012, *Natur*, **491**, 729
- van den Bosch, R. C. E., van de Ven, G., Verolme, E. K., Cappellari, M., & de Zeeuw, P. T. 2008, *MNRAS*, **385**, 647
- van der Marel, R. P., & Franx, M. 1993, *ApJ*, **407**, 525
- Veale, M., Ma, C.-P., Greene, J. E., et al. 2017a, *MNRAS*, **471**, 1428
- Veale, M., Ma, C.-P., Greene, J. E., et al. 2018, *MNRAS*, **473**, 5446
- Veale, M., Ma, C.-P., Thomas, J., et al. 2017b, *MNRAS*, **464**, 356
- Verolme, E. K., Cappellari, M., Copin, Y., et al. 2002, *MNRAS*, **335**, 517
- Voit, G. M., Ma, C. P., Greene, J., et al. 2018, *ApJ*, **853**, 78
- Walsh, J. L., Barth, A. J., Ho, L. C., & Sarzi, M. 2013, *ApJ*, **770**, 86
- Walsh, J. L., van den Bosch, R. C. E., Barth, A. J., & Sarzi, M. 2012, *ApJ*, **753**, 79
- Walsh, J. L., van den Bosch, R. C. E., Gebhardt, K., et al. 2015, *ApJ*, **808**, 183
- Walsh, J. L., van den Bosch, R. C. E., Gebhardt, K., et al. 2016, *ApJ*, **817**, 2
- Walsh, J. L., van den Bosch, R. C. E., Gebhardt, K., et al. 2017, *ApJ*, **835**, 208