

THE BLACK HOLE MASS IN M87 FROM GEMINI/NIFS ADAPTIVE OPTICS OBSERVATIONS

KARL GEBHARDT¹, JOSHUA ADAMS¹, DOUGLAS RICHSTONE², TOD R. LAUER³, S. M. FABER⁴, KAYHAN GÜLTEKIN²,
JEREMY MURPHY¹, AND SCOTT TREMAINE⁵

¹ Department of Astronomy, University of Texas at Austin, 1 University Station C1400, Austin, TX 78712, USA; gebhardt@astro.as.utexas.edu, jjadams@astro.as.utexas.edu, murphy@astro.as.utexas.edu

² Department of Astronomy, Dennison Building, University of Michigan, Ann Arbor, MI 48109, USA; dor@umich.edu, kayhan@umich.edu

³ National Optical Astronomy Observatories, P. O. Box 26732, Tucson, AZ 85726, USA; lauer@noao.edu

⁴ UCO/Lick Observatory, UC Santa Cruz, Santa Cruz, CA 95064, USA; faber@ucolick.org

⁵ Institute for Advanced Study, Einstein Dr., Princeton, NJ 08540, USA; tremaine@ias.edu

Received 2010 September 10; accepted 2011 January 11; published 2011 February 16

ABSTRACT

We present the stellar kinematics in the central $2''$ of the luminous elliptical galaxy M87 (NGC 4486), using laser adaptive optics to feed the Gemini telescope integral-field spectrograph, Near-infrared Integral Field Spectrograph (NIFS). The velocity dispersion rises to 480 km s^{-1} at $0''.2$. We combine these data with extensive stellar kinematics out to large radii to derive a black hole mass equal to $(6.6 \pm 0.4) \times 10^9 M_{\odot}$, using orbit-based axisymmetric models and including only the NIFS data in the central region. Including previously reported ground-based data in the central region drops the uncertainty to $0.25 \times 10^9 M_{\odot}$ with no change in the best-fit mass; however, we rely on the values derived from the NIFS-only data in the central region in order to limit systematic differences. The best-fit model shows a significant increase in the tangential velocity anisotropy of stars orbiting in the central region with decreasing radius, similar to that seen at the centers of other core galaxies. The black hole mass is insensitive to the inclusion of a dark halo in the models—the high angular resolution provided by the adaptive optics breaks the degeneracy between black hole mass and stellar mass-to-light ratio. The present black hole mass is in excellent agreement with the Gebhardt & Thomas value, implying that the dark halo must be included when the kinematic influence of the black hole is poorly resolved. This degeneracy implies that the black hole masses of luminous core galaxies, where this effect is important, may need to be re-evaluated. The present value exceeds the prediction of the black hole–dispersion and black hole–luminosity relations, both of which predict about $1 \times 10^9 M_{\odot}$ for M87, by close to twice the intrinsic scatter in the relations. The high end of the black hole correlations may be poorly determined at present.

Key words: galaxies: elliptical and lenticular, cD – galaxies: individual (M87, NGC 4486) – galaxies: kinematics and dynamics

Online-only material: color figures

1. INTRODUCTION

The masses of central black holes (BHs) in galaxies appear to be closely related to the luminosity (Dressler 1989; Kormendy 1993; Kormendy & Richstone 1995; Magorrian et al. 1998) and stellar velocity dispersion (Ferrarese & Merritt 2000; Gebhardt et al. 2000b) of their host galaxies. These relationships, which are determined from local samples of galaxies, provide the means to assay the cosmological mass distribution function of massive BHs, and provide the empirical foundation for establishing the role of BHs in galaxy formation and evolution (e.g., Hopkins et al. 2007).

At present, the BH galaxy–property relationships are derived from several dozen BH mass determinations made over the last few decades (see Gültekin et al. 2009). The relationships remain poorly observed at both their high-mass and low-mass ends. Lauer et al. (2007) show, for example, that the $M_{\text{BH}}-\sigma$ and $M_{\text{BH}}-L$ relationships must be in conflict at high BH mass due to curvature in the Faber & Jackson (1976) relationship between galaxy velocity dispersion and luminosity. Small uncertainties in the high-mass end of the relations can lead to uncertainties of up to two orders of magnitude in the implied volume density of BHs with $M_{\text{BH}} > 10^9 M_{\odot}$, due to the high-end exponential cutoff of the galaxy luminosity and velocity-dispersion distribution functions. Such estimates also depend critically on knowledge

of the intrinsic scatter in the relationships (Gültekin et al. 2009). Thus, there remains a need to measure accurate BH masses in a sample of the most massive galaxies.

Apart from the need to enlarge the sample of galaxies used to define the BH galaxy–property relationships, it appears that we may also need to test and potentially revise some BH mass measurements already made, especially in the massive “core galaxies.” Recent work shows that BH masses are subject to several systematic errors that have not been generally incorporated in the models used for analyzing the data so far. Some of these are discussed in Gültekin et al. (2009) and include the radial variations in the mass-to-light ratio (M/L) due to changes in stellar populations or the presence of a dark halo, uncertainties in the deprojection of the surface brightness, and triaxiality, among others. The most important of these systematic effects are as follows.

Dark halo. Gebhardt & Thomas (2009) show that the measured BH mass for M87 increases by more than a factor of two when a dark halo is included in the models; the reason for the change is that the BH’s kinematic influence is poorly resolved in the data they use, so that there is substantial covariance between the BH mass and stellar M/L . In turn, the best-fit stellar M/L , assumed independent of the radius, is affected by whether or not a dark halo is included in the models. It is well understood that the mass-to-light profile for ellipticals changes with radius

and not including that trend biases the BH determination. An obvious, but challenging, solution to this degeneracy is to obtain data at radii where the kinematics are strongly dominated by the BH rather than by the stars.

Incomplete orbit library. Shen & Gebhardt (2010) find an increase of two in the BH mass for NGC 4649 when using a more complete orbital sampling compared to models using a less coverage (Gebhardt et al. 2003). They argue that the orbital structure near the BH is dominated by tangential orbits and that the older models do not have adequate coverage of these tangential orbits (as discussed in Thomas et al. 2004). Having too few tangential orbits (i.e., too many radial orbits) can be compensated by having a smaller BH mass.

Triaxiality. Van den Bosch & de Zeeuw (2010) find an increase of two in the measured BH mass for NGC 3379 by using triaxial models compared to triaxial models (although they find the same BH mass for M32).

All three of these systematic effects tend to increase the BH mass. The increases are generally larger than the statistical uncertainties and suggest that systematic effects still dominate. By observing stars close to the BH, many model assumptions are no longer needed. For example, if the gravitational potential is dominated by the BH, then the stellar contribution to the enclosed mass is not important; hence, uncertainties in the stellar M/L , which may arise from uncertainties in the dark halo properties, can be mitigated by probing well inside the influence region of the BH.

Since it is among the most luminous galaxies nearby, has the largest BH known (from spatially resolved kinematics), and has one of the nearest and best-studied active galactic nuclei (AGNs), M87 is a natural and important target. An accurate BH mass determination for M87 helps to pin down the sparsely sampled upper end of the BH mass distribution and provides insights into formation and evolution of the most luminous galaxies. The previous analysis of M87 from Gebhardt & Thomas (2009) is based on ground-based kinematic data taken in natural seeing under moderately good conditions (FWHM = $1''$). In this paper, we present kinematics based on the integral field spectrograph, Near-infrared Integral Field Spectrograph (NIFS), on the Gemini Telescope, taken with adaptive optics (AO) correction. The spatial FWHM of the kinematics is $0''.1$ on average, with the best seeing image at $0''.08$. At larger radii, we incorporate new kinematic data out to $245''$ or 2.5 effective radii that will appear in a companion paper. The extreme improvement in the data quality of M87 allows us to model BH mass with smaller systematic uncertainty. This paper focuses on the determination of the mass of the central BH; the analysis of the stellar M/L and the dark halo properties will be given in Murphy et al. (2011).

Obtaining the kinematics at spatial resolution down to $0''.1$, at the same signal to noise obtained here, would have required about 100 orbits (90 hr) of the *Hubble Space Telescope* (HST), due to the faint stellar surface brightness. This AO study using Gemini/NIFS took about 10 hr in total, highlighting one of the great advantages for ground-based AO.

We assume a distance to M87 of 17.9 Mpc. The value of the BH mass scales linearly with the assumed distance.

2. DATA

A large amount of data exists for M87, and we do not attempt to integrate all of these; we rely on those data that provide the highest spatial resolution, most complete spatial coverage, and

highest signal to noise. Gebhardt & Thomas (2009) combine the *HST* images of Lauer et al. (1992) with the ground-based data of Kormendy et al. (2009) at larger radii. These data determine the surface brightness and ellipticity from radii of $0''.02$ to $2000''$ (1.7 pc to 170 kpc). Gebhardt & Thomas deproject the surface brightness to obtain the stellar luminosity density, and we use their deprojected density profile in this paper. The deprojection assumes axisymmetry; we assume an edge-on projection and the deprojected ellipticity is generally close to zero, but rises in the central region to 0.2 and in the outer region to 0.5. The stellar light profile within $0''.05$ has large uncertainties both in the radial shape and the ellipticity. The best-fit profile is a power law with the exponent -0.26 in radius and an increase in the ellipticity inside of $0''.15$. We have run a variety of models including and excluding this ellipticity change, increasing and decreasing the stellar power law with a range of exponents from 0 to -0.5 . We find that the BH mass changes by less than the statistical uncertainties. Thus, the exact shape of the central stellar light profile does not appear to be important for the BH mass.

For the spectral data, we present new observations from the Gemini Telescope taken with laser AO correction with the integral field unit of NIFS. It is important to include kinematic data out to much larger radii in order to constrain the orbital structure and the dark halo. The main source of the stellar kinematics at larger radii is Murphy et al. (2011). They obtain spectra with the integral field unit on the McDonald Observatory 2.7 m telescope (VIRUS-P; Hill et al. 2008). These data extend to $245''$. VIRUS-P has $4''.1$ fibers, making it unable to provide high spatial resolution. Thus, the region between the edge of the NIFS field (radius of $1''.9$ and $10''$), where VIRUS-P provides adequate spatially resolved kinematics information, requires additional coverage. We therefore include kinematics from the SAURON integral field unit (see <http://www.strw.leidenuniv.nl/sauron>). Emsellem et al. (2004) present the SAURON data in terms of Gauss–Hermite polynomials. The dynamical models discussed below rely on fits to the line-of-sight velocity distribution (LOSVD). To convert the Gauss–Hermite polynomials into LOSVDs, we generate 1000 Monte Carlo realizations of the polynomials based on their reported uncertainties. From these realizations, we generate the average LOSVD and 68% uncertainty at each velocity bin used in the LOSVD.

We only use the SAURON kinematics in this region ($2''.5$ – $11''$) even though they extend to $25''$. We do not use SAURON data within $2''.5$ because we want to provide an independent measure of the BH mass from the NIFS data alone. Including the SAURON data within this region does not change the best-fit BH mass, but does make the uncertainty smaller. We discuss these points further in Section 6.1. At radii beyond $11''$, we find a difference in the dispersion between the SAURON values and those from the VIRUS-P data. Murphy et al. (2011) argue that this difference is due to template issues in the kinematic extractions, and could be related to the limited wavelength coverage of SAURON, especially given the high alpha enhancement of M87; see Murphy et al. (2011) for a detailed discussion and analysis. The SAURON data used in this paper are available at the SAURON Web site.

One option for including the SAURON data at large radii would be to scale the velocity dispersions in order to make the overlap region consistent. We do not advocate this scaling. Primarily, the dynamical models use the full velocity profile and not just moments, and it is not clear whether a simple scaling of the dispersion is adequate. Furthermore, since the offset is likely due to template mismatch in the kinematic extraction (or

continuum placement), the scaling may not be constant with radius. Radial variations in the scaling could be due to template mix variations with radius, velocity dispersion changes with radius, or continuum differences with radius due to the AGN contribution. Within $11''$, Murphy et al. find consistency with the SAURON kinematics. Thus, we prefer to use the SAURON data where it is consistent and exclude it where there are differences.

We have performed several tests in which we exclude subsets of the data—removing the SAURON data, some of the large radii data, and some of the central NIFS kinematics—and find no effect on the best-fit BH mass.

2.1. Gemini NIFS Observations

We observed the central region of M87 in queue mode using the NIFS (McGregor et al. 2002) on the Gemini Telescope. We used AO corrections with the laser guide star system, ALTAIR (Boccas et al. 2006). The AGN in M87 provided the low-order corrections for tip-tilt and focus in a manner similar to Krajnović et al. (2009). An important feature in M87's center is the nearly point-like knot, located about $1''$ off the nucleus along the outbound jet, named HST-1 (Harris et al. 2003; Perlman et al. 2003; Cheung et al. 2007; Madrid 2009), which allows us to monitor the telescope's point spread function (PSF). The PSF serves as an important input to the kinematic modeling. The data were taken over five nights in 2008 April and May with 23 dithered positions of 10 minute exposure each on M87. The telluric standards HIP 59174 and HIP 61138 were observed to monitor and correct for atmospheric absorption. We used the K_G5605 grating that provides wavelength coverage from 2.00 to $2.43 \mu\text{m}$, with a spectral resolution of 5290 over a field of $3'' \times 3''$ sampled at $0''.04$ north–south and $0''.103$ east–west across the image slices.

We used the Gemini NIFS package of IRAF (Tody 1993) scripts (developed mainly by T. Beck) for the majority of the reductions. This package provides the flattening, registration (spatially and spectrally), hot pixel identification, and sky subtraction. This package does not yet handle the error frames appropriately, so we also passed our original uncertainty frames (dark current, read noise, and Poisson noise) through the same reductions as the science data as a first modification. Our second modification involves interpolating over the telluric Brackett γ line and dividing the telluric standard by a 10^4 K Planck function; this modification to the telluric correction retains a proper relative spectral shape in the science data. We also skipped the final NIFS script step where the data are resampled to equal x - and y -steps since we found that this interpolation enhanced the residual structure from our PSF fits.

We generally took sky exposures before and after the on-target frames, although there were some on-target frames that only had one sky exposure. We used the sky nearest in time for each on-target exposure. The sky subtraction usually worked well judging from the inspection of residuals under sky lines and tests with subtraction between different pairs of sky exposures. However, it is clear that some atmospheric emission-line variability occurred between our sky nods and caused uncertainty beyond our direct, statistical noise. There are techniques to bundle atmospheric emission lines into common transition families and fit a series of scalings between science and sky frames to minimize the residuals (Oliva & Origlia 1992; Rousselot et al. 2000; Davies 2007), but we have not employed them here. The public CO line lists extend only to $2.27 \mu\text{m}$ and therefore do not cover the CO bandheads through which we

measure all our kinematics. However, since we have so many sets of exposures, with each set at a different dither position, the problematic sky regions are mitigated. Furthermore, for the spectral extractions we mask out wavelengths near the CO bandheads that have large variations between sets of sky exposures above the thermal background. The final product from the NIFS reduction package is a wavelength-calibrated spectrum for each integral field unit (IFU) element at each of the 23 different dither positions. We next find the relative position for each exposure, the PSF, and remove the AGN and jet continua if present.

2.2. Determination of the PSF, Pointings, and Components

A crucial step for dynamical modeling with AO data is to determine the PSF and, in the case of M87, to remove non-stellar features from the spectra. For galaxies with shallow light profiles like M87, determination of the PSF is particularly important, since one has to know how much light from the outer regions is in the central spatial elements. Fortunately, for M87, we are in the excellent situation of having the point source (HST-1) within the field that can be used as a measure of the PSF. However, the central AGN is so bright that it significantly contaminates the stellar spectra within the central few spatial elements.

There are a few techniques to estimate the PSF with AO data outlined in Davies et al. (2004), which we considered. Davies et al. (2006) present observations where the PSF is measured from Brackett γ in an unresolved AGN. Since M87 has $H\alpha$ in the central regions, it should have some Brackett γ ; however, the redshifted line falls, to within one pixel, on one of the brightest sky lines at 21798 \AA . The residuals from the sky line are strong enough to not detect Brackett γ emission. The only observational handles we have on the AGN and jet flux contributions to any particular pixel are the spatial brightness variations and the change in the spectral slope. The AGN and jet are intrinsically much redder than the stellar populations. We wish to use this information without making assumptions about the stellar surface density. Thus, another goal is to study the stellar profile within the region dominated by the AGN; using the integrated light profile does not allow one to do this, whereas that information is contained in the spectra.

As opposed to measuring the light model (PSF and AGN fraction) from the reconstructed IFU data, we could use the stellar light profile as measured from *HST*. There are three reasons for not forcing the light profile from a previously measured *HST* image. First, the jet has knots that are moving on short timescales, and we cannot be sure that the AGN and jet fractional light and the spatial position will be the same. Second, we desire to use the spectral information to attempt to measure the underlying stellar component into the center. Third, the light profile and AGN fraction depend on the specific color. Data with *K*-band filters (Corbin et al. 2002) show color profiles that are flat with radius, although the analysis cannot easily go below $1''$ radii.

We make the assumptions of a constant stellar population in the center, with a spatially constant spectral slope and CO equivalent width (EW) only altered by the relative AGN/jet contamination. These assumptions form a closed constraint on the AGN/jet components and the PSF. Similar to Lauer et al. (1992) with visible light *HST* data and the approach of CLEAN algorithms (Högbom 1974), we model the inner AGN jet as a set of point sources. Additionally, we fit PSFs to the telluric standards as a verification of our in situ PSF models. We find in both that the sum of an inner, AO-corrected, non-circular

function and an outer, natural seeing function fit the data well without spatially coherent residuals. The two-component PSF is common with these types of data, but circular PSFs are commonly assumed (Neumayer et al. 2007).

It is important to get a robust spatial model for the AGN, jet, and stellar light since the kinematics in this spectral region are sensitive to the EW of the CO lines. Silge et al. (2005) show that a mismatch between the EWs of the velocity templates with the data can bias the velocity dispersion either high or low by up to 30%. Given that the additional continuum of the AGN will dilute the EWs, it is important to use as much information as possible to constrain the relative contribution.

Thus, (1) we assume that the stellar population (and therefore color and spectral slope) do not vary with radius near the center, and (2) we treat the AGN and jet as a set of point sources with unknown brightnesses with a flat continuum. We use the following operational definition of the CO EW:

$$\text{EW}_{\text{CO}} = \Delta\lambda \sum_{\lambda=2.29 \mu\text{m}}^{2.42 \mu\text{m}} \frac{D(\lambda) - a_{\text{tot}} \times \lambda^{\alpha_{\text{tot}}}}{a_s \times \lambda^{\alpha_s}}, \quad (1)$$

where $D(\lambda)$ are the counts in each pixel, a is the fitted zero point for the continuum (from a power-law fit), α is the fitted power-law exponent for the continuum, and a_s and α_s are the zero point and spectral index for the stellar light, respectively. Note that under assumption (1), this EW should not vary with position.

We begin the analysis of each science frame by considering all pixels outside of $0''.9$ from the center of the AGN and $0''.3$ from the center of HST-1; in this way the fit to the stellar model uses a relatively pure stellar signal. We make a least-squares power-law fit from 2.1 to $2.27 \mu\text{m}$ to each pixel and find a robust estimate for EW_{CO} and α_s from a bi-weight calculation (Beers et al. 1990). The estimate of any pixel's stellar continuum strength, a_s , can then be found with a least-squares power-law minimization for a_{tot} and α_{tot} followed by the application of Equation (1). The EW_{CO} and α_s over all frames are $220 \pm 26 \text{ \AA}$ and -3.11 ± 0.43 . We make continuum and EW maps for all pixels, subtract off the non-stellar continuum, and normalize by the stellar continuum extrapolated through the CO bandheads. This procedure produces the reduced spectra which are later used for extraction of the kinematics.

2.2.1. PSF Model

The PSF determination requires further analysis. We smooth the stellar continuum intensities using a $0''.2 \times 0''.2$ boxcar. We assume that the PSF has the form of an anisotropic Gaussian plus a power law (in the form of a Moffatt function), given by

$$\begin{aligned} \text{PSF}(x, y) &= N(x, y) + M(x, y), \\ N(x, y) &= \frac{a_1}{2\pi a_2^2 a_4} \exp\left(-\left(x_c^2 + \frac{y_c^2}{a_4^2}\right) / 2a_2^2\right) \\ M(x, y) &= (1 - a_1)(a_6 - 1) / \pi a_5 \left(1 + \frac{x^2 + y^2}{a_5^2}\right)^{a_6} \\ x_c &= x \cos(a_3) + y \sin(a_3) \\ y_c &= y \cos(a_3) - x \sin(a_3), \end{aligned} \quad (2)$$

where $\text{PSF}(x, y)$ is the value of the PSF at a given position x and y . $N(x, y)$ is the Gaussian model for the inner PSF, with normalization a_1 , width a_2 , position angle a_3 , and axis ratio a_4 . $M(x, y)$ is the Moffatt model for the outer PSF, with normalization $(1 - a_1)$, width a_5 , and exponent a_6 . In this model, the PSF

is assumed constant over the NIFS $2''$ field; this approximation is very accurate for a laser guide star system at a wavelength of $2 \mu\text{m}$.

Due to the large number of parameters needed to solve for the PSF (shape parameters for the PSF, multiple components for the AGN/jet, and stellar profile parameters), these fitting procedures involve minimizing a complicated function with local minima. We resort to simulated annealing as a global minimization tool (Press et al. 1992) with temperature schedules that reduce by 30% over each iteration and terminate with 10^{-4} fractional convergence. We first perform a simultaneous fit to the stellar-continuum-subtracted data with three central AGN/jet point source components, a fourth point source component for the HST-1 clump, and PSF parameters given by Equation (2).

We subsample each PSF evaluation over each pixel by five E-W and three N-S since the individual IFU elements do not properly Nyquist sample the PSF. This fit determines reasonable locations and strengths for all source components, but it improperly lets the central AGN/jet drive the PSF fit, and we know, in fact, that this feature is resolved given the multiple components. We therefore re-minimize the PSF terms and the HST-1 terms with data within $1''.2$ of the preliminary HST-1 position, but outside of $0''.5$ of the AGN center. The isolated clump, HST-1, then delivers a clean PSF determination. Finally, we re-minimize all source components, but hold the PSF terms fixed across the whole map. A representative decomposition example is shown in Figure 1 for one of the 23 data sets, where we show the raw IFU data, the stellar model, AGN/jet model, and residuals. The bottom right panel shows the residuals plotted on a linear scale. There remains some structure in the residuals, but the maximum residual is less than 1% of the measured value. We have tried a variety of PSF models and additional point source models, and find no improvement. Given the tightness in the HST-1 position determination, we register all frames off this cleanly isolated feature. From the median of all frames, we estimate AGN and HST-1 spectral indices of -0.67 ± 0.52 and -1.9 ± 2.6 , respectively.

This computation also delivers the range of PSFs for the 23 data sets. Most of the PSFs are close to the diffraction limit of $0''.06$ for the inner component, and the full range of the FWHM for the inner component is 0.055 – 0.19 . The fraction of light in the inner component ranges from 0.14 to 0.45 . The fraction of light in the central component is indicative of the Strehl ratio; however, in practice, the Strehl ratio is hard to measure given uncertainties in the PSF model (Gebhardt et al. 2000c). We make a two-dimensional image of each of the 23 analytic individual PSFs. From these 23 images, we then average to make the two-dimensional array which we use for the PSF in the dynamical modeling.

Figure 2 plots the flux, ellipticity, and position angle versus radius for the combined PSF. The values are reported in Table 1. This PSF has the inner Gaussian FWHM of around $0''.08$ with a fraction of the light in the central component of $a_1 = 0.38$. We use the analysis of the PSF as measured from the reconstructed IFU data, and we have also compared this PSF to that measured from the telluric standards. We find a similar FWHM for the inner and outer components, but the fraction of light in the inner component changes. For the PSF measured from the tellurics, the amount of light in the central component ranges from 60% to 70%, which is 1.5 to a factor of two larger than that determined from the science frames. We argue that using the telluric PSF is too optimistic for multiple reasons. First, the telluric star is

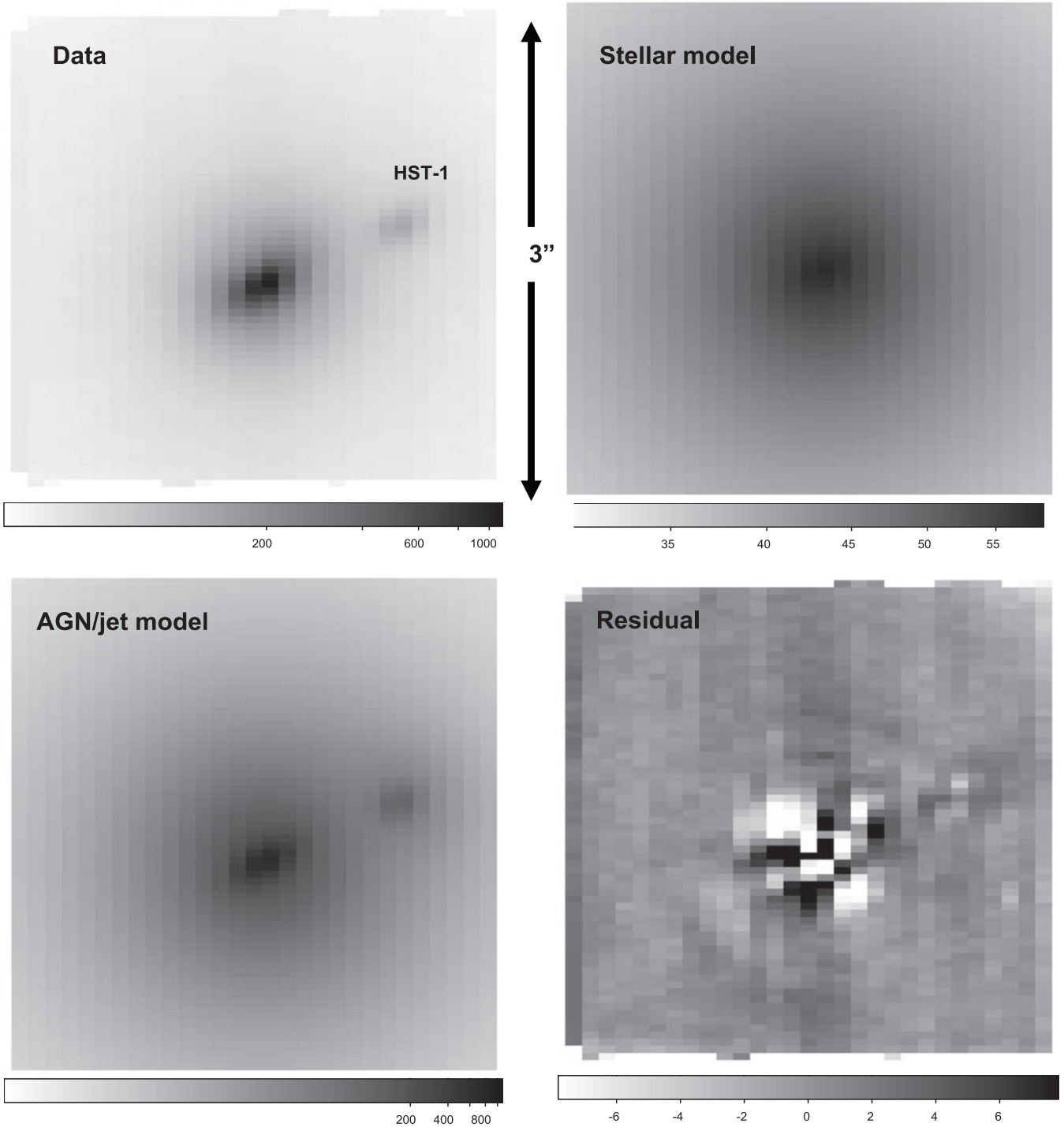


Figure 1. Stellar/AGN/clump decomposition and the PSF fit for one of the 23 data sets. In the upper left, we show the original data frame collapsed across wavelength. In the upper right, we show the fit to the stellar distribution by enforcing a constant stellar CO EW and spectral slope across the frame. In the lower left, we show the fit with three point sources to the central AGN/jet, one further point source for the HST-1 clump, and a six-parameter PSF. Finally, in the lower right we show the residuals. The final residual map is not entirely without structure, but further point source additions do not improve the χ^2 . Notice in the scales that different frames are displayed with log and linear stretched for clarity. The scales in the upper- and lower left are in the same (but arbitrary) units, and the scales of the residuals in the bottom right are in the same units as in the top-left figure. Thus, the residuals are less than 1%. In this image, the inner PSF is elongated with an axis ratio of 0.6, which can be seen in the left panels.

used as the reference star for the PSF and tip/tilt corrections (natural guide star mode), whereas the M87 data use the laser as the reference. Second, the M87 data use the nucleus for the tip/tilt corrections, which is more extended than the telluric star. Third, the M87 frames come from long exposures of 10 minutes compare to exposure times of only a few seconds for the telluric.

Thus, the M87 PSF is expected to be more extended. In any case, we also run a full set of dynamical models using the telluric PSF and find insignificant differences. It is encouraging to see that both PSFs give similar results; we attribute this robustness to the well-resolved kinematics in the region influenced by the BH.

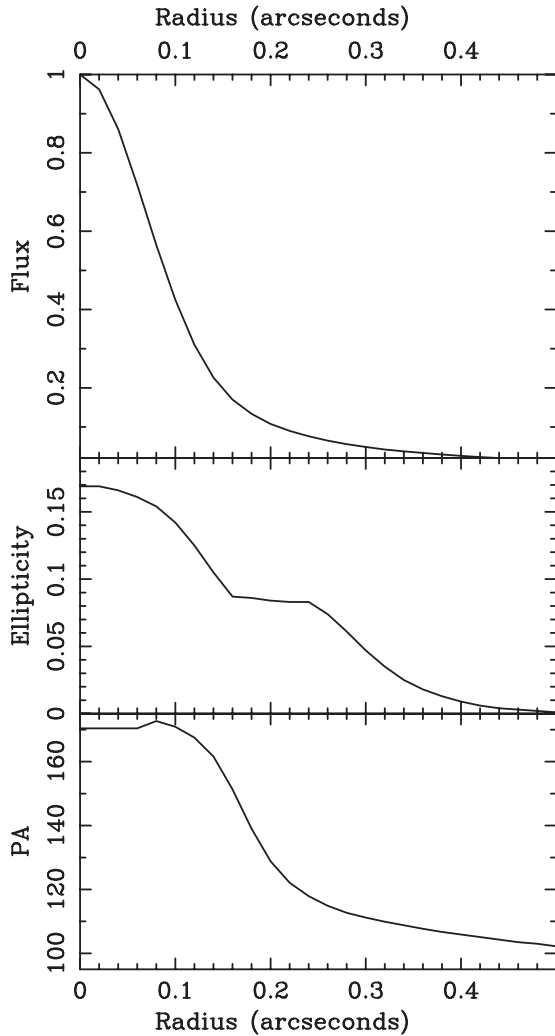


Figure 2. PSF used in the dynamical modeling. The top panel is the flux vs. radius, the middle panel is the ellipticity profile, and the bottom panel is the position angle (measured N to E). The combined PSF comes from the sum of the 23 individual two-dimensional PSFs. These values are reported in Table 1.

2.2.2. Central Stellar Distribution and Offset With AGN

Bagnuolo & Chambers (1987) and Lauer et al. (1992) both find power-law profiles with exponent -0.26 that stays constant into the smallest radius measured. Subject to the constant EW and spectral slope assumptions of our decomposition model, we investigate our stellar profile and that as determined by Lauer et al. (1992). We find a similar slope of -0.2 , and find no evidence for a change in the power-law profile near the BH.

The multiple components for the AGN and jet included in the model provide a measure as to whether the stellar centroid is consistent with the AGN. Batcheldor et al. (2010) report an offset of 6.8 ± 0.8 pc using Advanced Camera for Surveys images on *HST*. The displacement they report is along the jet axis, which they attribute to either a recoil event from the BH or a BH binary. The AO data presented here have similar spatial resolution ($0''.075$ for the *HST* data they used versus $0''.08$ for the AO data). The very large advantage of the AO data, however, is that the spectral information provides a further constraint on the relative amount of the AGN and stellar contribution. The average difference between the stellar centroid and the brightest AGN component is 3.2 ± 6.0 pc, consistent with no offset. Our statistical uncertainty is $8\times$ larger than that of Batcheldor

Table 1
PSF Parameters

R "	Flux	Ellipticity	P.A. °
0.00	1.00	0.169	170.4
0.02	0.96	0.169	170.4
0.04	0.86	0.166	170.4
0.06	0.72	0.161	170.4
0.08	0.56	0.154	172.7
0.10	0.42	0.142	170.9
0.12	0.31	0.125	167.5
0.14	0.22	0.105	161.6
0.16	0.17	0.087	151.3
0.18	0.13	0.086	139.0
0.20	0.11	0.084	128.8
0.22	0.090	0.083	122.1
0.24	0.076	0.083	117.9
0.26	0.065	0.074	114.9
0.28	0.056	0.061	112.7
0.32	0.043	0.035	109.9
0.36	0.034	0.018	107.7
0.40	0.027	0.009	105.9
0.44	0.021	0.004	104.3
0.48	0.017	0.002	103.0
0.52	0.014	0.001	101.7
0.56	0.011	0.000	100.6
0.60	0.010	0.000	99.6
0.64	0.008	0.000	98.7
0.68	0.006	0.000	97.7
0.72	0.005	0.000	96.6
0.76	0.005	0.000	96.1
0.80	0.004	0.000	94.3
0.86	0.003	0.000	92.2
0.94	0.002	0.000	90
1.00	0.001	0.000	90

et al.: $0''.08$ accuracy versus $0''.01$, respectively. Since the AO data should provide a better measure of the AGN and stellar contribution, it is not understood why the Batcheldor et al. uncertainty is $8\times$ lower. We suspect that important details such as the jet having multiple components (that may move with time) and having less observational constraints (imaging only versus imaging plus spectra) led to the result and small uncertainty reported in Batcheldor et al. We find no evidence that the AGN is offset from the galaxy center.

2.3. Kinematic Fits

To align the data with the kinematic axis at large radius from the fit of Kormendy et al. (2009), we take a position angle of -25° (E of N) for the major axis. Figure 3 shows the radial and angular bins used in the modeling, following the same binning of Gebhardt & Thomas (2009). Figure 3 only plots the spatial region for the NIFS data (the model goes out to $2000''$), with a gray-scale image from one of the 23 reconstructed IFU images. Spectra on opposite sides of the major axis are combined. We generally have 10 spectra at a given radius, since we have five angular bins on each side of the minor axis. The two central radial bins are not used due to AGN contamination (discussed below), and the next two outer radial bins require a sum over all angles in order to obtain adequate signal-to-noise ratio (S/N). The total number of spectra from the NIFS data that are used for dynamical modeling is then 40. Table 2 provides the locations for these 40 bins. We also include the signal to noise per pixel for each bin.

Table 2
NIFS Kinematics as Measured from the LOSVDs^a

R (")	Angle ^{b,c} (°)	V (km s ⁻¹)	ΔV (km s ⁻¹)	σ (km s ⁻¹)	$\Delta\sigma$ (km s ⁻¹)	h_3	Δh_3	h_4	Δh_4	S/N (pixel ⁻¹)
0.185–0.315	0–360	11.693	12.623	479.721	17.795	0.089	0.035	-0.013	0.025	63
0.315–0.515	0–360	5.576	6.672	466.656	13.211	0.042	0.023	-0.020	0.012	91
0.515–0.785	0–12	3.992	11.123	461.232	18.444	0.045	0.033	-0.043	0.015	64
0.515–0.785	12–24	-12.650	10.711	438.757	14.323	0.090	0.034	-0.002	0.019	68
0.515–0.785	24–37	4.525	9.517	421.979	14.776	0.116	0.032	-0.005	0.024	67
0.515–0.785	37–53	2.607	9.225	445.912	12.694	0.103	0.027	0.020	0.020	77
0.515–0.785	53–90	8.578	8.407	445.619	12.103	0.084	0.031	0.009	0.019	83
0.515–0.785	90–127	6.372	8.854	445.399	10.590	0.050	0.027	-0.014	0.013	78
0.515–0.785	127–143	0.174	9.401	438.214	9.015	0.060	0.024	-0.021	0.018	76
0.515–0.785	143–156	-6.244	9.969	445.479	9.902	0.060	0.029	-0.019	0.017	70
0.515–0.785	156–168	6.560	9.540	444.045	12.367	0.070	0.031	-0.017	0.015	71
0.515–0.785	168–190	10.741	11.112	431.575	15.242	0.099	0.034	0.029	0.024	67
0.785–1.165	0–12	-14.076	12.217	429.930	14.445	0.032	0.028	-0.007	0.016	63
0.785–1.165	12–24	12.211	10.777	454.078	13.451	0.068	0.036	-0.010	0.016	77
0.785–1.165	24–37	8.222	9.548	458.669	12.431	0.048	0.031	-0.012	0.017	78
0.785–1.165	37–53	19.721	7.987	463.289	16.540	0.106	0.027	0.015	0.021	91
0.785–1.165	53–90	6.845	6.650	449.497	9.592	0.054	0.022	-0.024	0.014	99
0.785–1.165	90–127	4.880	7.805	414.727	12.954	0.102	0.028	-0.031	0.015	83
0.785–1.165	127–143	-8.899	7.953	433.607	11.828	0.061	0.031	-0.021	0.014	83
0.785–1.165	143–156	-15.409	7.443	429.344	12.668	0.057	0.026	-0.018	0.018	76
0.785–1.165	156–168	-15.741	7.633	419.106	9.512	0.013	0.026	-0.036	0.012	82
0.785–1.165	168–190	-14.161	6.063	403.280	8.514	0.065	0.028	-0.010	0.018	90
1.165–1.695	0–12	-8.112	7.671	422.901	12.390	0.035	0.023	-0.013	0.015	78
1.165–1.695	12–24	-16.421	11.238	385.838	16.543	0.063	0.032	-0.037	0.027	83
1.165–1.695	24–37	-14.880	10.918	443.343	17.179	0.068	0.027	0.033	0.025	85
1.165–1.695	37–53	-35.708	7.213	371.643	11.054	-0.022	0.027	-0.002	0.023	89
1.165–1.695	53–90	-5.784	7.162	432.182	13.779	0.081	0.028	0.017	0.019	90
1.165–1.695	90–127	-8.583	7.950	442.494	12.169	0.083	0.029	0.015	0.017	91
1.165–1.695	127–143	-13.390	8.538	411.492	10.236	0.021	0.029	-0.033	0.012	93
1.165–1.695	143–156	-28.217	7.040	383.476	8.524	-0.007	0.017	-0.057	0.011	75
1.165–1.695	156–168	6.600	6.800	416.661	9.681	0.059	0.027	-0.018	0.013	89
1.165–1.695	168–190	-11.790	7.154	395.357	7.200	0.043	0.024	0.002	0.015	85
1.695–2.435	0–12	-44.146	14.064	337.386	13.033	-0.019	0.027	-0.037	0.011	75
1.695–2.435	12–24	-19.657	14.405	400.640	17.146	0.039	0.034	0.010	0.020	57
1.695–2.435	24–37	-16.149	14.549	342.453	13.296	0.013	0.029	-0.054	0.018	54
1.695–2.435	37–53	-32.272	41.911	316.492	42.518	0.011	0.058	-0.044	0.048	40
1.695–2.435	53–90	15.907	6.637	410.784	8.774	0.108	0.025	0.014	0.020	93
1.695–2.435	90–127	-8.944	13.338	371.176	21.056	0.030	0.036	-0.009	0.031	50
1.695–2.435	143–156	13.834	17.756	413.824	21.267	0.081	0.042	0.026	0.030	32
1.695–2.435	156–168	-17.332	9.840	392.510	12.277	0.101	0.030	-0.018	0.016	56

Notes.

^a The dynamical models use the LOSVDs, and the Gauss–Hermite coefficients are measured from the LOSVDs.

^b The spatial extractions are symmetrized about the major axis (i.e., bins ranging from 0° to 12° and from 0° to -12° are added, bins ranging from 12° to 24° and from -12° to -24° are added, etc.)

^c For angles over 90°, the LOSVDs have been flipped about zero velocity; thus, V and h_3 have opposite signs from the original value.

Figure 4 shows the spectra for three different spatial regions. The top spectrum comes from radii $0''.08 < R < 0''.18$, where we have summed over all angular bins; the middle is from $0''.18 < R < 0''.3$, and the bottom is from a radius of $0''.6$. The wavelength range shown here is the region used for the kinematic extractions. There are no significant absorption lines on either side, although the blueward region is used to help estimate the relative contribution of AGNs and stars (as discussed previously). The spectrum in the middle panel of Figure 4 represents the innermost point used in the dynamical models. The black line shows the data, and the dashed black lines are regions not used in the kinematic extractions due to large variations in the night sky. In this wavelength region, the residuals due to sky subtraction tend to be positive, even though the subtracted sky frames use the same exposure time as the

on-target frames. We have tried different sky subtraction levels and find little difference in the kinematic extractions because we excise the regions with sky lines. In the other wavelength regions, the sky residuals average to zero. The red line in the plots is the fitted LOSVD convolved with the template. The template comes from a library of 10 stars observed with GNIRS, as reported in the GNIRS template library (Winge et al. 2009). We select stellar types from G dwarfs to M supergiants. We rely on the GNIRS template library as opposed to the NIFS template library, since the GNIRS library contains a larger range of spectral types. The template library is an important consideration for this wavelength region (Silge et al. 2005).

The kinematic extraction simultaneously fits a non-parametric LOSVD and template weights for the individual stars. The template composition is allowed to vary spatially.

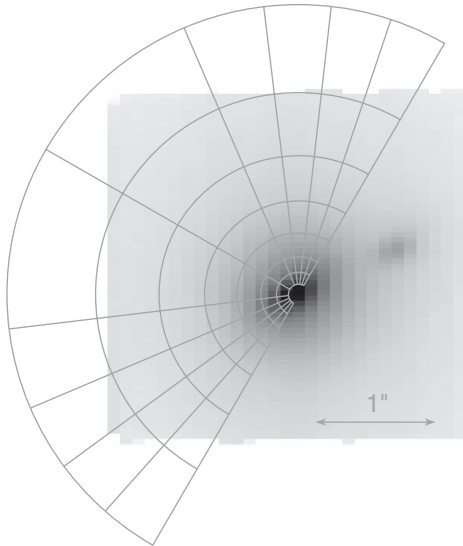


Figure 3. Binning scheme in M87 for the NIFS data only. Although this particular frame does not have data for each bin, the dithered set fills all bins. Data in the mirror bins around the major axis are added to the bins shown.

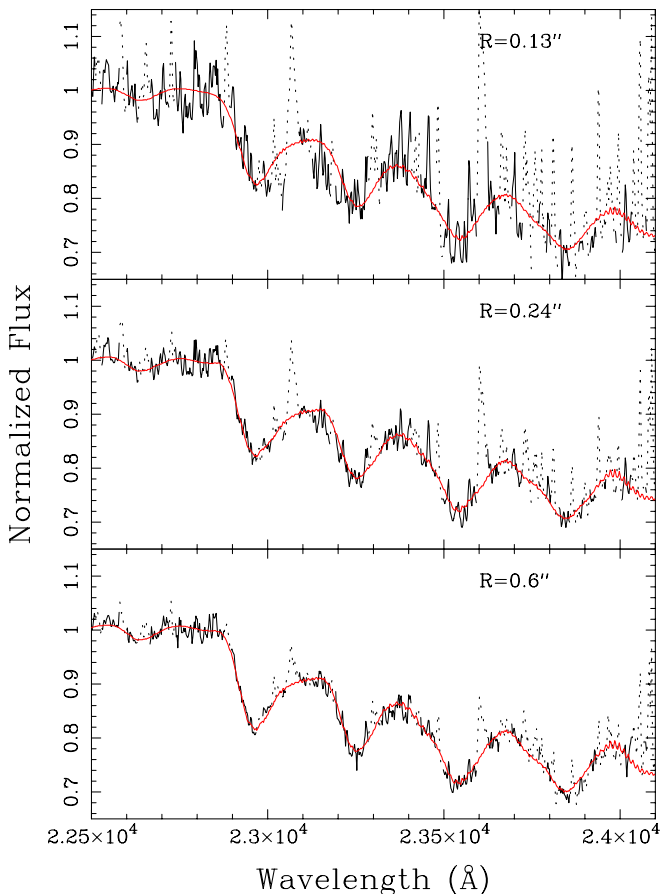


Figure 4. Spectra at three different radii. The top is from $0''.08 < R < 0''.18$, the middle is from $0''.18 < R < 0''.3$, and the bottom is from $R = 0''.6$. The black line is the spectrum and the smooth red line is the best-fit template convolved with the best-fit LOSVD. The dashed lines are those regions excluded from the fit due to high sky contamination. The spectrum at the top, which comes from the central region, is not used in the fit due to AGN contamination. The velocity dispersion obtained from the fits shown in red, from top to bottom, is 480 km s^{-1} , 480 km s^{-1} , and 445 km s^{-1} , and the S/N per pixel in each is 30, 63, and 91.

(A color version of this figure is available in the online journal.)

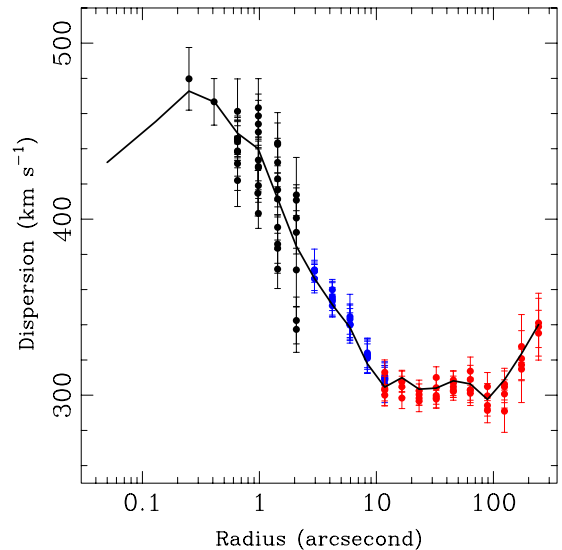


Figure 5. Velocity dispersion vs. radius for M87. The black points are the NIFS data. The red points are the VIRUS-P data from Murphy et al. (2011), and the blue points are from the SAURON data. The multiple points at each radius represent the various position angles. The solid line is the best-fit model, convolved to the appropriate PSF. For the dynamical model, we include the predicted dispersion within $0''.18$.

(A color version of this figure is available in the online journal.)

The technique is described in Gebhardt et al. (2000a) and Pinkney et al. (2003). The LOSVD is defined in 15 velocity bins of 260 km s^{-1} . There is a smoothing parameter applied to the LOSVD, but given the high S/N for most of the spectra, the smoothing has little effect on the extractions; thus, there is only a modest correlation between adjacent velocity bins. We use Monte Carlo simulations to determine the uncertainties in the LOSVD. The S/N of each spectrum determines the noise to be used in the Monte Carlo simulations; from 1000 realizations of each spectrum, we generate an average LOSVD and the 68% uncertainty.

The dynamical modeling uses the non-parametric LOSVD directly. However, it is sometimes convenient to express the LOSVD in a parameterized form as Gauss–Hermite moments, to show the radial run of the kinematics, and to compare the data with the models. Table 2 shows the first four Gauss–Hermite moments for the NIFS data. Figure 5 plots the velocity dispersion versus radius, where the dispersion is measured from a Gauss–Hermite fit to the LOSVDs. Figure 5 plots all of the data at each radius, and there are between 1 and 10 angular bins at each radii; thus, there are multiple points at nearly all radii in the figure. There is no rotation seen at a significant level in the NIFS data.

We input 107 LOSVDs in the dynamical models. These LOSVDs come from 40 spatial bins from the NIFS data, 25 from the SAURON data, and 42 from the large radial data of Murphy et al. (2011). The data in Murphy et al. come from the IFU VIRUS-P, where we have nearly complete angular coverage. The S/N of those data is very high (50–100 per resolution element). The solid line in Figure 5 plots the velocity dispersion from the best-fit dynamical model. The model generates LOSVDs, and their dispersions come from Gauss–Hermite fits to those LOSVDs. For the dynamical model dispersions, we average along angles at a given radius for clarity. In Figure 5, we plot both the NIFS and VIRUS-P dispersions, which have different PSFs. The model is convolved to each of the PSFs, and the plotted dispersions include the convolution.

2.4. Spectra at $R < 0''.18$

Data within the central $0''.18$ are excluded. Within $R < 0''.08$, the number of individual NIFS spatial pixels is only 50, whereas the number of spatial pixels for bins used in the models ranges from 250 to 3000. Given the shallow surface brightness profile for M87 and the low number of NIFS pixels, the signal from the central stars is low, and contamination from the AGN is high. We have tried a variety of models for the AGN, PSF, and stellar light profile; in all cases, we find that little information is contained in the central spectrum. We do not further discuss this spectrum.

The spectrum coming from the radial region $0''.08 < R < 0''.18$ has higher S/N, but still low enough that the kinematic extraction is highly uncertain. Figure 4 plots this spectrum in the top panel. It has lower stellar S/N compared to any spectra we use, and is further compromised by the uncertainty in the AGN subtraction. However, we still attempt a kinematic extraction. The red line in Figure 4 is the best-fit convolved template from the region one radial bin further in radius, from $0''.18 < R < 0''.3$. This region has a velocity dispersion of 480 km s^{-1} . There are wavelength regions, for example at $2.31 \mu\text{m} < \lambda < 2.35 \mu\text{m}$, where the model and data are offset. We do not attribute this offset to poor LOSVD modeling, but instead to poor AGN and stellar light discrimination. We use this region as an example and extracted LOSVDs including and excluding various regions. The range in velocity dispersion over all tests is from 350 to 620 km s^{-1} . The uncertainties on the dispersion for each individual extraction, coming from Monte Carlo simulations, are much smaller than this range, indicating that we are dominated by systematics as opposed to noise. For these reasons, we exclude this spectrum. We note that our best-fit dynamical model predicts a velocity dispersion of 451 km s^{-1} at this location (the solid line in Figure 5 shows the predicted dispersion from the model). This value is in the middle of our range of dispersions using the various extractions. For the central radial bin ($R < 0''.08$), the model prediction is 430 km s^{-1} .

3. DYNAMICAL MODELS

We use the orbit-based modeling algorithm described in Gebhardt et al. (2000a, 2003), Thomas et al. (2004, 2005), and Siopis et al. (2009). These models are based on Schwarzschild's (1979) method, and similar models are presented in Richstone & Tremaine (1984), Rix et al. (1997), Cretton et al. (1999), and Valluri et al. (2004).

The M87 models use the spherical geometric layout described in Murphy et al. (2011). We use 28 radial bins and five angular bins for the spatial sampling, and 15 velocity bins. The smallest spatial bin goes from 0 to $0''.05$. The velocity bins are 260 km s^{-1} wide. The average number of orbits per model is 40,000. The orbital sampling follows the design in Thomas et al. (2004, 2005), and is the same as that used in Shen & Gebhardt (2010). The latter paper illustrates the importance of a densely sampled orbital library: the mass obtained for NGC 4649, a galaxy similar to M87, is a factor of two larger than was found in earlier papers (e.g., Gebhardt et al. 2003). These papers did not include enough low-eccentricity polar orbits; for those galaxies that require significant tangential anisotropy in the central regions, this lack of circular orbits will bias the orbital structure and hence the BH mass. If a galaxy is dominated by tangential orbits in the central regions, the projected velocity dispersion will drop (for purely tangential orbits, the dispersion goes to zero at the galaxy center). Thus, if the central dispersion is

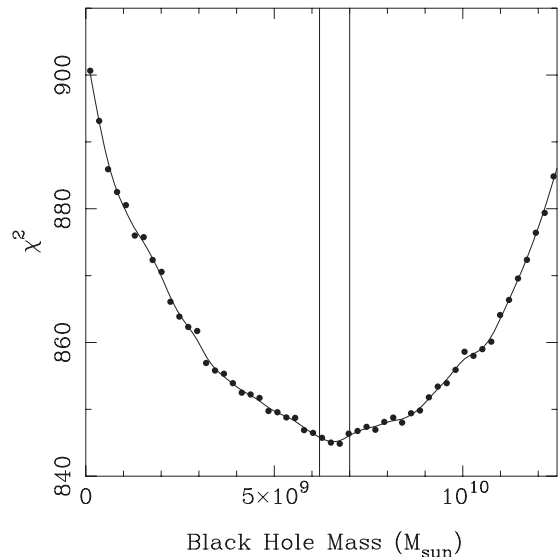


Figure 6. χ^2 vs. BH mass. Each point represents the χ^2 at that particular BH mass, and the line is a smoothed curve fitted to the points. The best-fit BH mass is $(6.6 \pm 0.4) \times 10^9 M_\odot$. The vertical lines represent the 68% range for the BH mass. The stellar M/L and the dark halo parameters have been fixed to the values reported in Murphy et al. (2011).

smaller than, for example, an isotropic distribution, this drop can be accommodated by either a lower BH mass or a tangential bias with a higher BH mass. In fact, the dynamical model predicts a drop in the dispersion in the central region (solid line in Figure 5); as discussed later, this drop is likely due to a tangential bias in the orbital distribution.

Gebhardt & Thomas (2009) find a strong correlation between the BH mass and the circular speed of the dark halo, both of which are anticorrelated with the stellar M/L . These correlations arise because of the limited spatial resolution of their data. This paper is based on data of higher quality, in particular with higher spatial resolution at the center ($0''.1$ compared to $1''.0$) and extending to larger radii ($245''$ compared to $30''$ for the stellar kinematics). Using the present data, there are no significant correlations between the BH mass, stellar M/L , and dark halo parameters. Therefore, in this paper we focus on the BH mass, deferring a discussion of the stellar M/L and dark halo parameters to Murphy et al. (2011). Following Murphy et al., we adopt $M/L = 9.1$ in V (solar units) and a spherical dark halo with potential $\Phi(r) = \frac{1}{2} V_c^2 \log(r^2 + R_c^2)$, where $V_c = 800 \text{ km s}^{-1}$ and $R_c = 36 \text{ kpc}$.

Figure 6 presents the χ^2 versus BH mass. Each of the 107 LOSVDs that we use in the dynamical models has 15 velocity bins, sampling velocities from -1820 to 1820 km s^{-1} . Given the dispersion profile, the outer two velocity bins at each end (i.e., four bins) are zero in the models and in the data, and since the uncertainties in the data are still significant for the large velocities, effectively they add nothing to χ^2 . Thus, we have only 11 LOSVD bins that contain signal (i.e., we could have limited the velocity range to $\pm 1400 \text{ km s}^{-1}$ with 11 bins and would have the same χ^2). The total number of data points in the kinematic fits is therefore around 1100. There is a small correlation between the adjacent velocity bins due to the smoothing used in the LOSVD extraction; this smoothing is set small enough and the velocity bins are large enough (260 km s^{-1}) that the correlation only mildly reduces the number of degrees of freedom. The best-fit model has $\chi^2 = 848$, so the reduced χ^2 is slightly less than unity, as expected given the

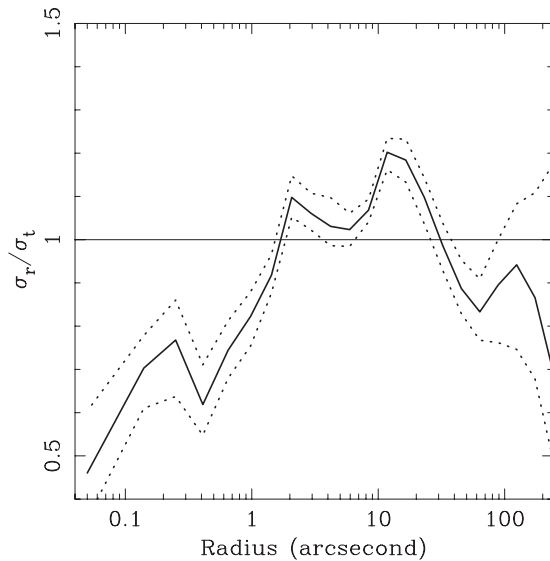


Figure 7. Radial to tangential velocity dispersion vs. radius. We average over polar angles in this plot since the variations in σ_r/σ_t between the angular bins are small. The average ratio (solid line) and 68% confidence band (dotted lines) come from all models that are within the 68% uncertainties for the four fitting parameters (BH mass, stellar M/L , dark halo circular velocity, and dark halo core radius). An isotropic distribution would have the ratio equal to unity.

correlation in the LOSVD bins. The points in Figure 6 are the χ^2 values from the individual models, and the line is a smoothing spline. We find a BH mass of $(6.6 \pm 0.4) \times 10^9 M_\odot$.

Figure 7 plots the ratio of the radial velocity dispersion to the tangential dispersion for the best-fit model. The tangential dispersion is defined as $\sigma_t^2 = 0.5 \times (\sigma_\phi^2 + \sigma_\theta^2 + V_\phi^2)$, where ϕ and θ are the spherical coordinates, and V_ϕ is the streaming motion in the ϕ direction. This ratio does not depend systematically on the polar angle, and so Figure 7 plots the angular average at a given radius. The confidence band comes from the range of models that are within the 68% uncertainties of the mass model, based on the uncertainties of the four parameters: BH mass, stellar M/L , dark halo circular velocity, and dark halo core radius. There is a sharp drop in this ratio at the center, implying a significant amount of tangential anisotropy (or similarly a lack of radial motion). As seen in Figure 5, the predicted projected velocity dispersion falls strongly inside of $0''.2$ (for orbits with no radial dispersion, the projected dispersion in the central region would fall to zero) due to the stronger tangential anisotropy. At radii beyond about $30''$, the orbital structure is close to isotropic. The tendency toward stronger tangential anisotropy in the central region has been seen in previous analyses for other galaxies (Gebhardt et al. 2003, 2007; Cappellari & McDermid 2005; Shapiro et al. 2006; Cappellari et al. 2007; Krajnović et al. 2009). Theoretical models (Quinlan et al. 1995; Quinlan & Hernquist 1997; Milosavljević & Merritt 2001) predict increased tangential anisotropy in the central regions due to a destruction of stars on radial orbits from ejection by or accretion onto the central BH, leaving only those stars on tangential orbits. Additionally, binary BHs will result in a significantly increase tangential anisotropy (Milosavljević & Merritt 2001), similar to the amount seen here in M87. While it appears that the large amount of tangential anisotropy seen here is due to a binary BH, a proper analysis requires a simulation tuned to the surface brightness profile and kinematics of M87. In particular, it is important to include a large range of initial conditions in the simulations for the stellar orbital structure in order to use the measured anisotropy profile to

determine the evolutionary history. Given that there are now many galaxies with well-measured central orbital structures, this analysis would be worthwhile.

4. MODELS WITHOUT A DARK HALO

We also ran models without a dark halo to investigate the sensitivity of our results to assumptions about the halo. In these fits, we include kinematic data out to a radius of $100''$, compared to $245''$ for the full data set; we do not use kinematic data between 100 and $245''$, because in this region the kinematics are likely to be dominated by the dark halo. We find that the best-fit mass decreases to $6.4 \times 10^9 M_\odot$, only 0.5σ or 2% smaller than the mass we obtain from models with a dark halo that use all the kinematic data. We conclude that the details of how we model the dark halo have negligible effects on the BH mass. In contrast, when using data with much lower spatial resolution ($1''$ versus $0''.08$ in this paper), Gebhardt & Thomas (2009) find a large change in the BH mass, around a factor of 2.5, between models with and without a dark halo. As suggested by them, the reason for this difference is that we now have high S/N kinematic data from well within the region influenced by the BH, so the covariance between the BH mass and stellar M/L is negligible.

5. M87 AND THE BH- σ AND BH- L RELATIONS

The M87 BH mass derived here and in Gebhardt & Thomas (2009) is significantly larger than most of the previous determinations (with the notable exception of Sargent et al. 1978, which we discuss in the following section). It is thus interesting to re-evaluate M87's position in the correlations of BH mass versus velocity dispersion (BH- σ) and BH mass versus luminosity (BH- L).

5.1. The Effective Velocity Dispersion

In Gültekin et al. (2009), we assign M87 a velocity dispersion of $\sigma_e = 375 \text{ km s}^{-1}$, which in turn comes from the analysis of Gebhardt et al. (2000b). In Figure 5, however, we see that this value is reached only at $r < 2''$, a location that is clearly within the inwardly rising portion of the dispersion profile associated with the BH's kinematic influence; this value unlikely represents the M87 galaxy overall.

The velocity-dispersion parameter used in the BH- σ relation is the effective velocity dispersion, σ_e , which in Gebhardt et al. (2000b) is defined as $\sigma_e^2 = \int_0^{R_e} I(r)V^2(r)dr / \int_0^{R_e} I(r)dr$, along the major axis, where I is the surface brightness, V is the projected second moment of the velocity distribution, and R_e is the half-light radius, which for M87 is $100''$ as reported in Lauer et al. (2007) and Kormendy et al. (2009). This operational definition of σ_e appears to provide a good correlation with BH mass, but there are many different ways in which one can integrate the kinematics in order to provide one number for the galaxy. With this definition $\sigma_e = 352 \text{ km s}^{-1}$, using the kinematics and surface brightness profile presented in this paper. The previous value of 375 km s^{-1} results from using the older kinematic and light profiles. It is clear that σ_e contains a substantial contribution from the light inside where the BH influences the kinematics. If instead we exclude radii within this region (defined as $r_s = GM_{\text{BH}}/\sigma^2$ and equal to $2''.1$ for our models) from the integral that determines σ_e , we find $\sigma_e = 324 \text{ km s}^{-1}$, about 8% smaller. We choose 324 km s^{-1} as our best estimate of σ_e , with a range from 312 km s^{-1} to 352 km s^{-1} .

Churazov et al. (2010) show that there exists a radial “sweet spot”, where the velocity dispersion at that radius is robustly related to the circular velocity. By providing a dispersion value that is indicative of the galaxy as a whole, this estimate may correlate well with the BH. Based on the kinematics from Murphy et al. (2011), the dispersion value of the “sweet spot” for M87 is $312 \pm 10 \text{ km s}^{-1}$. Furthermore, Cappellari et al. (2007) measure a value of 306 km s^{-1} by integrating the two-dimensional data within a radius of $30''$. There are a variety of ways to represent a velocity dispersion for a galaxy, and until there is a physically motivated model it is not obvious which measure is optimal. Thus, in order to be consistent with uses of σ_e for other galaxies and the current incarnation of the BH– σ correlation, one should use σ_e as reported above (324 km s^{-1}), but other correlations should be studied.

We note that contamination of σ_e by the light from stars within BH’s kinematic influence is likely to be less important for most other galaxies, since M87 is both close and has an unusually massive BH. At the same time, it may be prudent that this issue be examined for all galaxies in the context of refining the BH– σ relation overall.

5.2. Estimated Black Hole Mass in M87

Gültekin et al. (2009) present two BH– σ relations, one for all galaxies, and one for elliptical galaxies alone. Using $\sigma_e = 324_{-12}^{+28} \text{ km s}^{-1}$ gives $\log(M_{\text{BH}}) = 9.0_{-0.2}^{+0.4}$ in the first case and $\log(M_{\text{BH}}) = 9.1_{-0.2}^{+0.4}$ in the second. Likewise, evaluating the Gültekin et al. BH– L relation with $M_V = -22.71$ (Lauer et al. 2007) gives $\log(M_{\text{BH}}) = 9.0 \pm 0.2$. Both relations thus give $M_{\text{BH}} = 1 \times 10^9 M_{\odot}$, in contrast to our determination of $M_{\text{BH}} = 6.6 \times 10^9 M_{\odot}$. Thus, our measurement differs from the predictions of this BH– σ and BH– L relations by 0.82 dex. However, the intrinsic scatter in these relations is estimated by Gültekin et al. to be 0.44 (BH– σ , all galaxies), 0.31 (BH– σ , ellipticals only), and 0.38 (BH– L , ellipticals only). Adding this scatter in quadrature gives estimates of $\log(M_{\text{BH}}) = 9.0_{-0.5}^{+0.6}$, $9.1_{-0.4}^{+0.5}$, 9.0 ± 0.4 , respectively. Thus, our measured value of $\log(M_{\text{BH}}) = 9.82 \pm 0.03$ differs from the predictions by $1.4-2\sigma$. Given the present uncertain state of knowledge of the high-mass ends of both the BH– σ and BH– L relations, we do remark on the larger significance of M87 deviation from the relations, except to say that it does highlight the need to improve the sample of BH mass determinations from the most massive galaxies.

6. DISCUSSION

6.1. M87 Specific Results

Our best-fit BH mass is $(6.6 \pm 0.4) \times 10^9 M_{\odot}$. Sargent et al. (1978) report a BH mass of $6 \times 10^9 M_{\odot}$ (after scaling to our assumed distance), which is within 1σ of our reported value. Their model is based on lower spatial resolution data (about $1''.5$), assumes that the velocity distribution is isotropic, and does not include a dark halo. It is impressive that after three decades of improvement in data quality, modeling, and understanding, there is essentially no change in the measured BH mass. Part of the reason for the robustness of the Sargent et al. result is that the radial influence on the projected kinematics from the BH extends to nearly $10''$ (see Figure 5), so the influence of the BH was clearly visible in their kinematic data. They also use isotropic models, whereas we run axisymmetric models with no restrictions on the anisotropy. To study the effect

of the assumption of isotropy, we fit isotropic models to the kinematic data presented in this paper. The comparison between the projected dispersions of the isotropic models and the data is poor, with an increase in χ^2 by over a factor of two. The poor fit makes it difficult to assign a best-fit mass and the range of equally poor-fitting models has BH masses that range from 6×10^9 to $8 \times 10^9 M_{\odot}$, consistent with the models of Section 3, which show significant tangential anisotropy (Figure 7). Thus, in M87, the assumption of isotropy does not have a significant effect on the measurement of the BH mass, although isotropic models provide a poor fit to the data. Sargent et al. also do not include a dark halo, which has been shown to cause the BH to be underestimated. Their velocity dispersions at large radii are lower than ours (245 compared to 300 km s^{-1}), which is most likely because their template library was incomplete and their spectra had lower S/N. The lower dispersion causes the assumed M/L of the stars to be lower, an error of the opposite sign to the error caused by neglect of the dark halo. Thus, the impressive agreement between our value and that of Sargent et al. (1978) appears to be due in part to the competing effects of observational errors (dispersions too small, which makes the stellar M/L too low and the BH mass too large) and oversimplified models (no dark halo or velocity anisotropy, both of which make the BH mass too small). Another often-quoted BH mass determination from stellar kinematics comes from Magorrian et al. (1998) who report a value of $4.2 \times 10^9 M_{\odot}$ (for our distance). The likely reason for the difference is that they do not include a dark halo and thus overestimate the stellar M/L .

The BH mass reported here is nearly the same as that reported in Gebhardt & Thomas (2009), within 4%. There is very little kinematic data in common between the two studies. The kinematic data in Gebhardt & Thomas come from older long-slit data at a spatial resolution of $1''.0$ (van der Marel 1994), while in this paper we use two-dimensional coverage at a spatial resolution of $0''.08$. We further use ground-based data from Murphy et al. (2011) that have excellent S/N and radial extent. There is some data from SAURON (Emsellem et al. 2004) in common between the two studies, but this provides only 10% of the LOSVDs used in the models. Thus, the dynamical models from the two studies use nearly independent kinematic data sets, and give approximately the same answer.

The uncertainties on the BH mass from these two studies are similar even though the data presented here are superior in many ways: the previous uncertainty is 0.5×10^9 , whereas the uncertainty with the NIFS data is 0.4×10^9 . In order to keep the BH mass measures independent, the models presented in this paper do not include the SAURON data inside of $2''.5$. The similarity in the BH mass uncertainty is then due primarily to the fact that the two sets of data have similar accuracy on the kinematics in the central $2''.5$. Combining all NIFS data, the accuracy on the velocity dispersion is 0.2% (1 km s^{-1}). Combining all SAURON data within $2''.5$ provides the same accuracy. Thus, as long as one has a reliable PSF and no systematic differences in the kinematic extractions, then it is expected that the uncertainty on the BH mass is similar using either data set. We have run a subset of models including both the NIFS data and all SAURON data; in this case, the uncertainty on the BH mass decreases to 0.25×10^9 (with no change in the best-fit mass). We report and utilize the result using only the NIFS data within $2''.5$ in order to (1) provide as independent result as realistically possible and (2) control potential systematic differences in the kinematic extractions. Murphy et al. (2011) find a difference in the velocity dispersion of the SAURON

data at large radii compared to their measurements, which they attribute to template issues. While we do not find an offset in the dispersion values in the central region, we desire to maintain the independence. The major difference, however, is that there is no degeneracy with the stellar M/L using the NIFS data, whereas the degeneracy is very strong otherwise. Thus, the systematic uncertainty from the M/L profile is effectively removed with the AO data, making the result on the BH mass and orbital structure much more robust.

For M87, the AO data have removed the systematics due to the M/L profile, but the systematics due to the extraction of the kinematics remain important. These systematics include continuum placement, template mismatch, and removal of the AGN contribution. The first two are general and the latter is specific to M87. Getting any of these controlled to better than 1% of the velocity dispersion will be very difficult.

Corrected to our distance, the BH masses reported from gas kinematics are $(2.9 \pm 0.8) \times 10^9 M_{\odot}$ in Harms et al. (1994) and $(3.8 \pm 1.1) \times 10^9 M_{\odot}$ in Macchetto et al. (1997). As discussed in Gebhardt & Thomas (2009) the mass reported here is in conflict with these by about 2σ . Possible reasons for the differences are discussed in Gebhardt & Thomas, with the most likely reason being uncertainty in the inclination of the gas disk. Macchetto et al. assume a value of 51° based on the gas kinematics. Harms et al. assume a value of 42° based on the imaging of the gas emission. The reported difference provides a measure of the systematic uncertainty in the inclination (i.e., whether the gas kinematics or the gas distribution are more affected by non-gravitational forces). Applying this 9° difference in the inclination changes the Macchetto et al. BH mass from $(3.8 \pm 1.1) \times 10^9$ to $(5.4 \pm 1.3) \times 10^9 M_{\odot}$, which would lead to an insignificant difference of 0.6σ between our result and theirs. Of course, the analysis is more complicated than this simple application since one would need to re-model the gas kinematics with a different inclination. A proper treatment would be to include the gas kinematics with the stellar dynamical models. Our focus in this paper is on the stellar kinematics, and we do not attempt to merge the gas kinematic analysis.

6.2. General Implications for Black Hole Mass Measurements

While the kinematics obtained from the AO study produce effectively the same BH mass and its uncertainty from kinematics taken in native image quality, the robustness of the measures is greatly strengthened. For example, the BH mass is not dependent on the assumption of constant M/L . Trying to generalize this result to other galaxies with BH mass determinations is difficult since the measure of the BH mass depends on many aspects. There are two observational extremes that we highlight as examples. The first is having a measure of BH mass that comes from observations that resolve well the kinematic influence of the BH. In the most extreme case, high S/N spectra could potentially see the high-velocity wings in the LOSVD due to the BH (as discussed in van der Marel 1994). The second example would be to allow the poorer resolution of the BH, but provide a very accurate measure of the mass-to-light profile. In this paper, we rely on the first strategy; Gebhardt & Thomas rely on the second. That the two strategies give consistent results, at least for M87, suggests that both may be reliable.

Other studies have reported robust measures of the BH mass from ground-based studies that only poorly resolve the BH's kinematic influence. Shapiro et al. (2006) measure a BH for NGC 3379 from SAURON data that is consistent with that measured from *HST* data using both stars (Gebhardt et al. 2000a)

and gas kinematics. Kormendy (2004) summarizes the history of BH mass measures for many galaxies and finds that, in general, the differences are within the reported uncertainties. If one has sufficient signal to noise and two-dimensional coverage (e.g., SAURON or VIRUS-P), then it should be possible to measure a BH mass robustly. Thus, it is not necessarily required to resolve the region influenced by the BH.

Being able to use data that do not resolve well the BH's influence on the kinematics allows us to study BHs that are either distant or of low mass. Both of these regimes are important for understanding the physical nature of the BH correlations with the host galaxy. For example, McConnell et al. (2011) measure a BH mass in NGC 6086, which is 133 Mpc distant. The kinematic influence of the BH is barely resolved, and the degeneracy between the BH mass and M/L profile is strong. However, as demonstrated for M87, as long as one properly characterizes the mass profile at large radii, high signal-to-noise data can measure the BH mass accurately.

It is possible that systematic uncertainties bias the current crop of BH correlations. One obvious consequence could be that without accounting for the effect of systematic uncertainties, the measured intrinsic scatter would increase. Gültekin et al. (2009) measure the scatter of 0.44 dex for the full sample of galaxies with measured BH masses and 0.31 dex for ellipticals. Once systematic effects are understood and included, the intrinsic scatter may decrease. Other consequences include increasing the mass density of BHs, if BH masses are all underestimated, and changing the slope or curvature of any correlation. Schulze & Gebhardt (2011) re-analyze the set of 12 galaxies from Gebhardt et al. (2003) including a dark halo. They find an increase of 50% in the BH mass, due primarily to improved dynamical modeling (more complete orbit sampling) and partly to including a dark halo. The increase correlates with BH mass. It is important to re-evaluate all BH mass estimates.

The key to understanding all of these effects comes from high spatial resolution data. Data from *HST* (mainly from Space Telescope Imaging Spectrograph, STIS) are generally regarded as providing the most significant results for BH mass studies. The small and stable PSF is a central aspect for the robustness of the data from *HST*. Future uses of STIS will play an important role for quantifying BH masses. The main obstacle for *HST* though is that it is a relatively small mirror and requires substantial observing time. For example, in order to measure the BH mass in M87 at the same accuracy presented here would require nearly 100 orbits. While this amount of time could be justified for a small number of objects, going to a much larger sample using *HST* is difficult. Fortunately, AO observations are in a mature stage where they can provide much larger samples.

We are grateful to the excellent staff at the Gemini Telescope, both for providing the observations and for significant support in the data analysis. The NIFS with the laser AO system, ALTAIR, is a powerful resource. We thank the referee for excellent and thoughtful comments which improved the manuscript. K.G. acknowledges Ralf Bender, the Max-Planck-Institut für Extraterrestrische Physik, and the Carnegie Institute of Washington for their excellent support and hospitality. The project would not have been possible without the facilities at the Texas Advanced Computing Center at The University of Texas at Austin, which has allowed access to over 5000 node computers where we ran all of the models. K.G. acknowledges support from NSF-0908639. The observations were obtained at the Gemini Observatory, 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), the Science and Technology Facilities Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciência e Tecnologia (Brazil), and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), under program GN-2008A-Q-12. The paper uses data obtained at the William Herschel Telescope, operated by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias

REFERENCES

- Bagnuolo, W. G., Jr., & Chambers, R. W. 1987, *Nature*, **326**, 681
- Batcheldor, D., Robinson, A., Axon, D. J., Perlman, E. S., & Merritt, D. 2010, *ApJ*, **717**, L6
- Beers, T. C., Flynn, K., & Gebhardt, K. 1990, *AJ*, **100**, 32
- Boccas, M., et al. 2006, *Proc. SPIE*, 6272, 62723
- Cappellari, M., & McDermid, R. 2005, *Class. Quantum Grav.*, **22**, 347
- Cappellari, M., et al. 2007, *MNRAS*, **366**, 1126
- Cheung, C. C., Harris, D. E., & Stawarz, L. 2007, *ApJ*, **663**, L65
- Churazov, E., et al. 2010, *MNRAS*, **404**, 1165
- Corbin, M. R., O’Neil, E., & Rieke, M. J. 2002, *AJ*, **124**, 183
- Cretton, N., de Zeeuw, P. T., van der Marel, R., & Rix, H.-W. 1999, *ApJS*, **124**, 383
- Davies, R. I. 2007, *MNRAS*, **375**, 1099
- Davies, R. I., Tacconi, L., Genzel, R., Ott, T., & Rabien, S. 2004, *Proc. SPIE*, **5490**, 473
- Davies, R. I., et al. 2006, *ApJ*, **646**, 754
- Dressler, A. 1989, in *Proc. IAU Symp. 134, Active Galactic Nuclei*, ed. D. E. Osterbrock & J. S. Miller (Dordrecht: Kluwer), 217
- Emsellem, E., et al. 2004, *MNRAS*, **352**, 721
- Faber, S. M., & Jackson, R. E. 1976, *ApJ*, **204**, 668
- Ferrarese, L., & Merritt, D. 2000, *ApJ*, **539**, L9
- Gebhardt, K., Pryor, C., O’Connell, R. D., Williams, T. B., & Hesser, J. E. 2000a, *AJ*, **119**, 1157
- Gebhardt, K., & Thomas, J. 2009, *ApJ*, **700**, 1690
- Gebhardt, K., et al. 2000b, *ApJ*, **539**, L13
- Gebhardt, K., et al. 2000c, *AJ*, **119**, 1268
- Gebhardt, K., et al. 2003, *ApJ*, **583**, 92
- Gebhardt, K., et al. 2007, *ApJ*, **671**, 1321
- Gültekin, K., et al. 2009, *ApJ*, **698**, 198
- Harms, R., et al. 1994, *ApJ*, **435**, L35
- Harris, D. E., Biretta, J. A., Junor, W., Perlman, E. S., Sparks, W. B., & Wilson, A. S. 2003, *ApJ*, **586**, L41
- Hill, G. J., et al. 2008, *Proc. SPIE*, 7014, 701470
- Högbom, J. A. 1974, *A&AS*, **15**, 417
- Hopkins, P. F., Hernquist, L., Cox, T. J., Robertson, B., & Krause, E. 2007, *ApJ*, **669**, 45
- Kormendy, J. 1993, in *The Nearest Active Galaxies*, ed. J. Beckman, L. Colina, & H. Netzer (Madrid: CSIC), 197
- Kormendy, J. 2004, in *Coevolution of Black Holes and Galaxies*, ed. L. C. Ho (Cambridge: Cambridge Univ. Press), 1
- Kormendy, J., Fisher, D., Cornell, M., & Bender, R. 2009, *ApJS*, **182**, 216
- Kormendy, J., & Richstone, D. 1995, *ARA&A*, **33**, 581
- Krajnović, D., McDermid, R. M., Cappellari, M., & Davies, R. L. 2009, *MNRAS*, **399**, 1839
- Lauer, T., et al. 1992, *AJ*, **103**, 703
- Lauer, T., et al. 2007, *ApJ*, **662**, 808
- Macchetto, F., Marconi, A., Axon, D., Capetti, A., Sparks, W., & Crane, P. 1997, *ApJ*, **489**, 579
- Madrid, J. P. 2009, *AJ*, **137**, 3864
- Magorrian, J., et al. 1998, *AJ*, **115**, 2285
- McConnell, N., Ma, C. P., Graham, J., Gebhardt, K., Lauer, T., Wright, S., & Richstone, D. 2011, *ApJ*, **728**, 100
- McGregor, P., et al. 2002, *Proc. SPIE*, 4841, 178
- Milosavljevic, M., & Merritt, D. 2001, *ApJ*, **563**, 34
- Murphy, J., Gebhardt, K., & Adams, J. 2011, *ApJ*, in press
- Neumayer, N., Cappellari, M., Reunanen, J., Rix, H.-W., van der Werf, P., deZeeuw, P. T., & Davies, R. 2007, *ApJ*, **671**, 1329
- Oliva, E., & Origlia, L. 1992, *A&A*, **254**, 466
- Perlman, E. S., Harris, D. E., Biretta, J. A., Sparks, W. B., & Macchetto, F. D. 2003, *ApJ*, **599**, L65
- Pinkney, J., et al. 2003, *ApJ*, **596**, 903
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in FORTRAN: The Art of Scientific Computing* (2nd ed.; Cambridge: Cambridge Univ. Press)
- Quinlan, G., & Hernquist, L. 1997, *New Astron.*, **2**, 533
- Quinlan, G., Hernquist, L., & Sigurdsson, S. 1995, *ApJ*, **440**, 554
- Richstone, D., & Tremaine, 1984, *ApJ*, **286**, 27
- Rix, H.-W., et al. 1997, *ApJ*, **488**, 702
- Rousselot, P., Lidman, C., Cuby, J.-G., Moreels, G., & Monnet, G. 2000, *A&A*, **354**, 1134
- Sargent, W. L. W., Young, P., Lynds, C. R., Boksenberg, A., Shorridge, K., & Hartwick, F. D. A. 1978, *ApJ*, **221**, 731
- Schulze, A., & Gebhardt, K. 2011, *ApJ*, in press
- Schwarzschild, M. 1979, *ApJ*, **232**, 236
- Shapiro, K., et al. 2006, *MNRAS*, **370**, 559
- Shen, J., & Gebhardt, K. 2010, *ApJ*, **711**, 484
- Silge, J. D., Gebhardt, K., Bergmann, M., & Richstone, D. 2005, *AJ*, **130**, 406
- Siopis, C., et al. 2009, *ApJ*, **693**, 946
- Thomas, J., Saglia, R., Bender, R., Thomas, D., Gebhardt, K., Magorrian, J., Corsini, E., & Wegner, G. 2005, *MNRAS*, **360**, 1355
- Thomas, J., Saglia, R., Bender, R., Thomas, D., Gebhardt, K., Magorrian, J., & Richstone, D. 2004, *MNRAS*, **353**, 391
- Tody, D. 1993, in *ASP Conf. Ser. 52, Astronomical Data Analysis Software and Systems II*, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (San Francisco, CA: ASP), 173
- Valluri, M., Merritt, D., & Emsellem, E. 2004, *ApJ*, **602**, 66
- van den Bosch, R., & de Zeeuw, P. T. 2010, *MNRAS*, **401**, 1770
- van der Marel, R. 1994, *MNRAS*, **270**, 271
- Winge, C., Riffel, R., & Storchi-Bergmann, T. 2009, *ApJS*, **185**, 186