Mon. Not. R. Astron. Soc. **367,** 2–18 (2006) doi:10.1111/j.1365-2966.2005.09713.x

The central kinematics of NGC 1399 measured with 14 pc resolution

R. C. W. Houghton,[1⋆] J. Magorrian,[2] M. Sarzi,[1] N. Thatte,[1] R. L. Davies[1] and D. Krajnović[1]

[1] *University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH*
[2] *University of Oxford, Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford OX1 3RH*

Accepted 2005 October 7. Received 2005 September 30; in original form 2005 August 16

ABSTRACT

We present near-infrared (NIR) adaptive optics-assisted spectroscopic observations of the CO ($\Delta\mu = 2$) absorption bands towards the centre of the giant elliptical galaxy NGC 1399. The observations were made with NAOS-CONICA (on the European Southern Observatory's Very Large Telescope) and have a full width at half-maximum resolution of 0.15 arcsec (14 pc). Kinematic analysis of the observations reveals a decoupled core and strongly non-Gaussian line-of-sight velocity profiles in the central 0.2 arcsec (19 pc). NIR imaging also indicates an asymmetric elongation of the central isophotes in the same region.

We use spherical orbit-superposition models to interpret the kinematics, using a set of orthogonal 'eigen-velocity profiles' that allow us to fit models directly to spectra. The models require a central black hole of mass $1.2^{+0.5}_{-0.6} \times 10^9$ M$_\odot$, with a strongly tangentially biased orbit distribution in the inner 40 pc.

Key words: instrumentation: adaptive optics – galaxies: individual: NGC 1399 – galaxies: kinematics and dynamics.

1 INTRODUCTION

Supermassive black holes (SMBHs) are thought to be the only viable candidates for the massive dark object (MDO) observed at the centres of many nearby galaxies. Indeed, recent near-infrared (NIR) observations of the centre of the Milky Way have resolved individual stars orbiting in close proximity to the central MDO (which coincides with the radio source, Sgr A*) and these data rule out all other plausible explanations for an MDO, other than an SMBH (Schödel et al. 2003).

Significantly, a relationship between the mass of the SMBH and the bulge luminosity of the host galaxy was discovered (Kormendy & Richstone 1995), and subsequently a tighter correlation between the SMBH mass $M_\bullet$ and the velocity dispersion of the bulge σ (the $M_\bullet$–σ relation) was measured (Ferrarese & Merrit 2000; Gebhardt et al. 2000) of the form

$$\log(M_\bullet/\mathrm{M}_\odot) = \alpha + \beta \log(\sigma/\sigma_0), \quad (1)$$

with $\sigma_0 = 200$ km s^{-1}. Tremaine et al. (2002, hereafter T02) find $\alpha = 8.13$ and $\beta = 4.02$ and Ferrarese & Ford (2005, hereafter FF05) find $\alpha = 8.22$ and $\beta = 4.86$. Similar relations with low scatter have also been found between $M_\bullet$ and the infrared luminosity L_{IR} (Marconi & Hunt 2003) and between $M_\bullet$ and bulge mass (Häring & Rix 2004).

It is believed that the mass accretion history of an SMBH is linked to the formation and evolution of its host (Haehnelt & Kauffmann 2000; de Zeeuw 2004), and so such a precise relation, connecting quantities on vastly different scales, provides an important constraint on models of galaxy assembly. It would be particularly remarkable if it holds true for galaxies of different morphological types, which most likely underwent very different formation and evolution histories. For example, Faber et al. (1997) suggest that power-law ellipticals and spiral bulges formed dissipatively whereas core-like ellipticals formed from mergers, yet both appear to follow the same relation. In practice, the relation can also be used to measure the mass of a host galaxy black hole (BH) where a dynamical estimate is not possible (Aller & Richstone 2002; Yu & Tremaine 2002).

1.1 Contention

There has been considerable debate over the values of the parameters α and β (Ferrarese & Merrit 2000; Gebhardt et al. 2000; T02; FF05). The slope β is crucial for comparison with theoretical models that attempt to explain the $M_\bullet$–σ relation, but currently the sample of galaxies used by T02 (Fig. 1) and FF05 is somewhat limited and biased. Of the 31 galaxies in the sample of T02, 18 are elliptical, nine are lenticular and four are spiral (see Fig. 1). Although there are roughly equal numbers of power-law and core ellipticals, there remain few high-dispersion and low-dispersion galaxies where any deviations from the canonical slope will be most obvious. Such bias is not surprising considering that, until recently, only one telescope could perform dynamical mass estimates with the required spatial resolution: the *Hubble Space Telescope* (*HST*).

⋆E-mail: rcwh@astro.ox.ac.uk

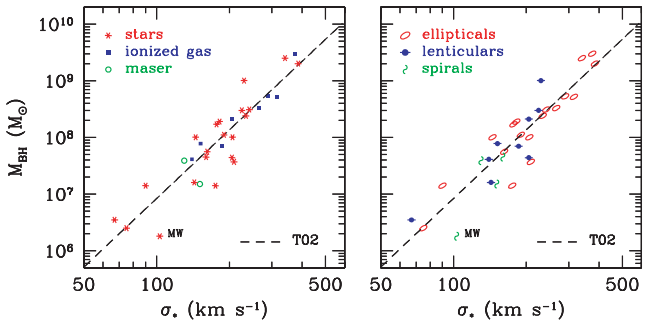

Figure 1. The sample of T02 and their best-fitting correlation of the form of (1) with $\alpha = 8.13$ and $\beta = 4.02$: the symbols indicate the technique used to derive the black hole mass (left) and the different morphological types (right).

In order to estimate the mass of the black hole reliably, it is necessary to resolve kinematics in the region of space where the black hole potential dominates over the potential of the stars, a point stressed by FF05. The radius of this sphere of influence (SoI) is of the order

$$r \sim \frac{GM_\bullet}{\sigma_\star^2}, \tag{2}$$

where G is the gravitational constant and $\sigma_\star$ is the average stellar velocity dispersion of the spheroidal component. Ground-based observations have, in the past, been limited by atmospheric seeing to a resolution of around 1 arcsec or more. There are very few galaxies close enough and with sufficiently large SMBHs to have an SoI larger than this, so nearly all of the sample of T02 and FF05 comes from *HST* data [diffraction limited full width at half-maximum (FWHM) of ∼0.047 arcsec at 5500 Å]. However, *HST* has its limitations.

Massive giant ellipticals which populate the high-dispersion region of the $M_\bullet$–σ plane are rare (so on average more distant) and have low surface brightness centres with flat core-like density profiles (Faber et al. 1997). Although the total luminosity of such galaxies is high, they have too low a surface brightness to achieve high signal-to-noise ratio (S/N) observations efficiently with *HST*'s 2.4-m primary. Conversely, high S/N nuclear observations of galaxies with steep power-law photometry (Faber et al. 1997) are feasible due to their large numbers (therefore proximity) and rapid increase in surface brightness towards the nucleus. Hence, less massive ellipticals with steep power-law surface brightness profiles are common in the sample of T02, but very massive ellipticals with core-like surface brightness profiles are rare.

The low-dispersion region of the $M_\bullet$–σ plane is populated by low-mass bulges in spirals that tend to be dusty and obscured at the centre. Extinction is reduced at longer wavelengths (the extinction at 2.3 μm is only 10 per cent of that at V-band), but *HST* is unable to probe the centres of such galaxies as it does not have a suitable IR spectrograph. Therefore, less massive dusty spirals are also underrepresented in the current sample.

The outcome of these selection effects is an uneven distribution of galaxies along the σ axis: 70 per cent of the T02 sample have $120\,\mathrm{km\,s^{-1}} < \sigma < 250\,\mathrm{km\,s^{-1}}$. It is imperative that galaxies with high and low velocity dispersions be investigated, not only to verify the uniformity of the $M_\bullet$–σ relation over different morphological types and a larger dispersion range, but also to better determine the nature of the relation. The slope of the $M_\bullet$–σ relation remains contentious (T02; FF05) because data points at the extremities, crucial for defining the slope, are sparse. Furthermore, some theories of galaxy evolution predict departures from the current power-law relation which would only be detectable with more data at high and low dispersions (Haehnelt & Kauffmann 2000; Zhao, Haehnelt & Rees 2002).

Modern adaptive optics (AO) facilities on large 8-m class ground-based telescopes (diffraction limited FWHM of 0.059 arcsec at 2.3 μm) are the key to solving this problem. The spatial resolution achieved with adaptive optics is limited primarily by the diffraction limit of the telescope. In the case of NIR observations with an 8-m primary, this can match the spatial resolution of *HST* in the optical. Furthermore, 8-m telescopes deliver high S/N observations of low surface brightness objects and the NIR spectrographs available on ground-based telescopes are better able to probe obscured dusty regions. We have undertaken a careful study to identify targets with bright reference stars close to the galaxy centre which are underrepresented in the current $M_\bullet$–σ plane.

1.2 NGC 1399

The first of these targets to be observed is the giant elliptical NGC 1399, which is the most luminous galaxy in Fornax and has a core-like surface brightness profile characteristic of a cD-type galaxy (Schombert 1986; Killeen & Bicknell 1988). Many dynamical studies in the visible have been undertaken in the past (Bicknell et al. 1989; Franx, Illingworth & Heckman 1989; Longo et al. 1994; D'Onofrio et al. 1995; Graham et al. 1998; Saglia et al. 2000) although none have been made in the NIR or with the spatial resolution available with an AO system.

The velocity dispersion of the spheroidal component $\sigma_\star$ is calculated in different ways by different authors (Ferrarese & Merrit 2000; Gebhardt et al. 2000). As highlighted by T02, Ferrarese & Merrit (2000) and subsequently FF05 calculate the rms dispersion within a circular aperture of radius $r_e/8$ while Gebhardt et al. (2000) and T02 use the (luminosity-weighted) rms dispersion within a slit aperture of length $2r_e$. Whether or not the different definitions affect the measured slope of the relation, it is important to measure the corresponding values for NGC 1399 so like can be compared with like.

The effective radius of NGC 1399 is reported to be 40 arcsec by de Vaucouleurs et al. (1991). Using the long-slit data of Graham et al. (1998) to give a slit aperture of length $2r_e$ centred on the galaxy (±40 arcsec) and the photometry of Lauer et al. (2005), we calculate the luminosity-weighted rms dispersion along the slit aperture to be $317 \pm 3\,\mathrm{km\,s^{-1}}$ (the quoted error is an estimate of random error only). NGC 1399 is therefore at the top of the T02 $M_\bullet$–σ plane, with three other galaxies that anchor the relation (M87, IC 1459 and NGC 4649). The predicted BH mass for NGC 1399 is $8.7 \times 10^8\,\mathrm{M_\odot}$, so the SoI would be 0.38 arcsec (37 pc), much smaller than even the best seeing at the best observatory sites.

Following Ferrarese & Merrit (2000), the rms dispersion within a circular aperture of radius $r_e/8$ was estimated from the central $2 \times 5\,\mathrm{arcsec^2}$ of Graham et al.'s (1998) data to be $329 \pm 4\,\mathrm{km\,s^{-1}}$ (as before, the quoted error is a measure of the random error). The relation of FF05 then predicts a BH mass of $1.89 \times 10^9\,\mathrm{M_\odot}$ and the SoI would be 0.78 arcsec (75 pc), still not well resolvable with the best seeing conditions.

Fortunately, NGC 1399 is ideally suited to AO-assisted observations. A bright ($m_v = 13.8$) reference star exists only 17.6 arcsec away from the galaxy centre. The galaxy is almost spherical, with little or no rotation, so the slit is free to be aligned to the galaxy centre and the AO reference star. This allows us to monitor the AO correction as a function of time.

One problem that plagues all stellar-dynamical estimates of black hole masses is the degeneracy between mass and anisotropy (Binney & Mamon 1982). Consider the case of a spherical galaxy. Using the Jeans equation, the mass enclosed within radius r can be written as (Kormendy & Richstone 1995)

$$M(r) = \frac{v^2 r}{G} + \frac{\sigma_r^2 r}{G}\left[- \frac{\mathrm{d}\ln j}{\mathrm{d}\ln r} - \frac{\mathrm{d}\ln \sigma_r^2}{\mathrm{d}\ln r} - \left(1 - \frac{\sigma_\theta^2}{\sigma_r^2}\right) - \left(1 - \frac{\sigma_\phi^2}{\sigma_r^2}\right)\right], \quad (3)$$

where v is the rotation velocity, σ_r, σ_θ and σ_ϕ are the radial and azimuthal components of the velocity dispersion, and j is the deprojected luminosity density. The first two terms in the square brackets can be estimated almost directly from observations. Both are positive for the vast majority of galaxies. The last two terms, however, depend on the the unknown anisotropy and can take either sign. Their effect on $M(r)$ is minimized for galaxies with steep $j(r)$ profiles, steep velocity dispersion profiles σ_r and rapid rotation $v \neq 0$, all of which tend to be satisfied in power-law galaxies. Core galaxies like NGC 1399, however, tend to be non-rotating with shallow density and dispersion profiles. For such galaxies it is particularly important to constrain the anisotropy by modelling at the detailed shape of the galaxy's line-of-sight velocity profiles (VPs), for which high S/N spectra are essential (Gerhard 1993).

Using the NAOS AO system coupled with the CONICA NIR imager/spectrograph at the European Southern Observatory's Very Large Telescope (ESO VLT), we have resolved the SoI of NGC 1399 and measured its stellar kinematics using the CO absorption bands at 2.3 μm and the Ca I absorption feature at 2.26 μm. The S/N of the spectra range from 90 to 20 over the region used to extract kinematics, with S/N ∼ 70 at the CO bandhead (2.3 μm). We use these kinematics to construct a spherical orbit superposition model for the galaxy to estimate the mass of the central MDO. Throughout this paper we assume a distance of 19.9 Mpc to NGC 1399 (Tonry et al. 2001); the reader is reminded that BH mass scales linearly with assumed distance.

The structure of this paper is as follows. The data reduction techniques are discussed in Section 2; the kinematic analysis is discussed in Section 3; the imaging and kinematics are presented in Section 4; and the discussion of their implications is contained in Section 5. The dynamical modelling is described and discussed in Section 6. Finally, Section 7 concludes.

2 DATA AND REDUCTION

2.1 Observations

AO-assisted *K*-band images (*Ks* filter) and *K*-band long-slit spectra (*SK* filter) of the nuclear region of NGC 1399 were obtained with NAOS-CONICA (Rousset et al. 1998; Lenzen et al. 1998) on the nights of 2003 November 30 and December 01. The spectral range extended from 1.79 to 2.45 μm, with a scale of 0.972 nm pixel^{-1}, although atmospheric transmission limits high-S/N data to 1.95–2.45 μm. The spatial scale of the spectroscopy was 0.0543 arcsec pixel^{-1}; the scale of the imaging was 0.027 arcsec pixel^{-1}. The slit width was 0.172 arcsec (17 pc) corresponding to an instrumental resolution ($\lambda/\Delta\lambda$) of 880 at 2.3 μm (as measured from the width of the arc lines) and an instrumental broadening of $\sigma_{\mathrm{inst}} = 145$ km s^{-1}. The conditions while observing were excellent, with the seeing varying between 0.4 and 0.6 arcsec.

A total of 27 000 s (7.5 h) of useful on-source spectroscopic integration was achieved, with individual exposures lasting 300 s each. Airmass ranged from 1.0 to 1.4, although 85 per cent of exposures were made with airmass <1.2. The standard ABBA technique of nodding back and forth along the slit removed the need for separate sky exposures, and Fowler readout mode was used to minimize readout noise. AO-assisted images of NGC 1399 were also taken for a total of 80 s with the *Ks* filter together with an equivalent number of sky exposures.

In order to remove the complex atmospheric transmission curve (referred to as telluric absorption) from the object spectra, it was necessary to observe several *telluric standard stars* (Section 2.3), which have almost intrinsically featureless spectra in the NIR (e.g. very hot O or B stars). As the strength of telluric absorption depends on the airmass, on the first night telluric standards were observed at airmasses of 1.02 (HD 25631) and 1.40 (HD 25631), and on the second night telluric standards were observed at airmasses of 1.12 (HD 480) and 1.25 (HD 41814). The difference in airmass between telluric and galaxy exposures on each night was never more than 0.2 airmasses. Two kinematic template stars were also observed at low airmass: HD 11931 (K4III) and HD 25840 (M0III).

All stars (telluric standards and kinematic templates) were observed with an identical spectrograph configuration to the galaxy observations. However, for the stellar observations alone, it was necessary to reduce the higher-order gain of AO correction to ensure that the FWHM of the point spread function (PSF) was larger than the width of the slit to match the spectral resolution of the galaxy and stellar observations. With full AO correction, the FWHM of the PSF was ∼0.1 arcsec, which would have been significantly smaller than the 0.172 arcsec wide slit.

The position angle (PA) of the slit was 5.06° so as to include the galaxy centre and AO reference star in the slit. This allowed us to monitor and assess the AO correction. We assume that the SMBH lies at the most luminous region of the galaxy (also assumed to be at the centre of the galaxy), so it is important to position the slit to sample this region or to be able to quantify any offset from it. Prior to acquiring the galaxy, an image of the slit on the detector was made by removing the grism and illuminating the slit with the flat-field lamp. The slit image was then used as a bias for the subsequent acquisition images to align the slit precisely. Special care was taken with the first acquisition of the first night to ensure the PA of the slit intersected the AO reference star and the brightest part of the galaxy. The PA was subsequently held fixed for all observations and only shifts perpendicular to the slit length (i.e. along the slit's minor axis) were made to maintain the slit position on the star and the galaxy centre. On the second night, the same PA was verified to hold the star and the galaxy centre in the slit and then shifts were made along the slit's minor axis as before.

2.2 Reduction techniques

The data reduction was completed with the aid of the IRAF[1] and ECLIPSE [2] packages as well as custom IDL[3] scripts, incorporating use of the IDL Astronomy User's Library (Landsman 1993).[4]

[1] IRAF is distributed by National Optical Astronomy Observatories (NOAO). See http://iraf.noao.edu

[2] ECLIPSE is a reduction package developed by the European Southern Observatory (ESO). See http://www.eso.org/projects/aot/eclipse/

[3] Interactive Data Language, Research Systems, Inc.

[4] See http://idlastro.gsfc.nasa.gov/

The spectroscopic data were initially reduced following the standard ABBA technique for NIR data reduction,[5] which has many advantages: the time-scale on which the background subtraction is achieved is as short as possible, helping to correct for the variability of the NIR sky; any residual sky in a single A–B frame cancels with the residual in the B–A frame, assuming a uniform sky field; and the pixel-to-pixel subtraction is very well suited to removing systematic errors. However, such a technique does not optimize the random noise (sky, thermal, dark, readout) as a background exposure of the same duration as the object exposure is subtracted from each pixel.

In an effort to increase the S/N, the background level of each pixel was interpolated as a function of time from all the data frames. Pixels with significant source (galaxy) counts were excluded when fitting a third-order polynomial as a function of time to each pixel position. In order to gauge the change in the S/N from interpolation, the random noise of the sky-dominated region between galaxy and the AO reference star was measured as a function of wavelength and compared to the noise of the frames without background interpolation: the interpolated background showed a significant decrease in random noise ($\sim\sqrt{2}$ lower). This led to a significant increase in the final S/N of the data while maintaining the pixel-to-pixel subtraction to reduce systematics. However, the residuals from one object–sky pair do not necessarily cancel with the next pair, unlike the classical reduction technique for ABBA sequence observations (assuming a uniform illumination). The interpolation is also susceptible to bias from bad pixels, which can propagate into neighbouring frames. To counter this problem, the frames were cleaned of bad pixels prior to (object and sky frames), during (sky pixels only) and after the interpolation (background-subtracted pixels). Indeed, a consequence of interpolating the background was the identification of faulty pixels on the detector and the construction of an very accurate bad pixel map.

The timing and conditions of the spectroscopic data make it well suited to sky interpolation: 7.5 h of observations were performed on neighbouring nights with limited interruption between the exposures on each night and the atmospheric conditions during the two nights were excellent and stable.

After correction for the odd–even effect of the detector, sky subtraction, flat fielding, bad pixel correction, field transformation and wavelength calibration, the exposures were aligned along the spatial axis (to an accuracy of a few tenths of a pixel) by using the centroid of a Gaussian fitted to the light profile of the AO reference star. Reduction of telluric standard and kinematic template spectra followed the same sequence except with only standard ABBA sky subtraction.

The image data was reduced in the standard manner for NIR observations. After correcting for odd–even effect, the sky exposures were subtracted from the object frames, followed by flat-fielding. The centroid of a Gaussian, fitted to the AO reference star, was then used to align the individual exposures to a tenth of a pixel accuracy.

2.3 Telluric correction

Molecular gas in the Earth's atmosphere (CO_2, H_2O, etc.) produces an absorption spectrum known as telluric absorption, and this must be removed from astronomical observations. The galaxy and kinematic template observations were corrected for telluric absorption in the standard manner (Origlia, Moorwood & Oliva 1993), the details and caveats of which are described below.

The stellar continuum must first be removed from the telluric standard star spectrum. We therefore divide through by a blackbody spectrum of appropriate temperature to the star. This is a blind division; one cannot fit the temperature, as the atmospheric transmission curve is still imprinted on the spectrum (in addition to any instrumental response).

Telluric absorption varies with airmass, so the strength of the sharp prominent absorption features in the normalized telluric spectrum must be matched to those in the object spectrum (of similar airmass); division of this optimized standard then removes the prominent (high-frequency) features and, we assume, the low-frequency component of the atmospheric transmission curve. The telluric spectrum, now free from stellar continuum, is not yet normalized, so we divide by a linear fit to regions where we *believe* the atmospheric transmission to be $\simeq$100 per cent: (2.035–2.04), (2.09–2.15) and (2.21–2.22) μm. The increased strength and occurrence of telluric features beyond 2.3 μm causes atmospheric transmission to be consistently below unity, and falling with increasing wavelength, so one must extrapolate the continuum from shorter wavelengths. The absorption depth in the now normalized telluric spectrum $A(\lambda)$ is then amended using the expression

$$A'(\lambda) = 1.0 - F[1.0 - A(\lambda)], \quad (4)$$

where F is a free parameter, to account for small variations in airmass between the object and telluric standard star observations. Note that this linear correction for airmass is only an approximation to a more complicated response and is only effective for very small differences in airmass. The initial normalization was then removed by multiplying back by the original linear fit. Every telluric spectrum also required a small shift in wavelength (around a few tenths of a pixel) to compensate for differences in wavelength zero-point between the object and telluric spectra. The optimization of airmass and wavelength was initially automated by minimizing residuals around the strong telluric features at (2.0–2.1) μm in the telluric-divided object spectrum. Fine adjustments were then made to further optimize the removal of prominent telluric absorption features at the region of interest (the CO bandheads after 2.3 μm). Note that, in practice, while one can correct individual stellar exposures (due to the high S/N in each frame), galaxy exposures must be co-added in groups of similar airmass to increase the S/N before an accurate telluric correction can be determined.

Telluric correction is not without complication though. The telluric correction is optimized for the removal of sharp, prominent, high-frequency absorption features. Accordingly, there is likely to be minimal residuals from such in the corrected spectra. However, variation in the continuum normalization of a galaxy spectrum is known to introduce systematic effects into the derived kinematics (van der Marel et al. 1994). Owing to the blind division of a blackbody spectrum and the lack of continuum reference in the telluric spectra after 2.3 μm, error in telluric correction will most likely manifest itself as error in the continuum normalization of the telluric spectrum after the application of (4). This would propagate into an error in the continuum level of the object spectrum (galaxy or kinematic template). Such an error may not be uniform over the length of the spectrum: there is likely to be a higher chance of error where the normalization was extrapolated, after 2.3 μm. To help compensate for such systematic effects, we include a polynomial continuum correction when extracting kinematics (see Section 3), but the constraint for the correction is the minimization of χ^2_s (7), which does not guarantee to choose a solution free of systematics.

[5] For further information on NIR data reduction, the reader is referred to the ISAAC data reduction guide (the NACO data reduction guide is still being constructed).

Furthermore, this continuum correction is additive, whereas any real error would be divided into the object spectrum.

Different functions were used to normalize the galaxy and kinematic template spectra after telluric correction. The galaxy continuum appeared linear with wavelength and so, at every position along the slit, a linear fit to the continuum shortward of the CO bands was sufficient. The continuum of each kinematic template shortward of the CO bandhead was clearly non-linear but was well fitted (and removed) by a blackbody spectrum. As the continuum after 2.3 μm must be extrapolated, it was necessary to fit slowly varying functions with relatively little freedom. Note that, with all continuum fits, care was taken not to include obvious absorption or emission features, or areas of significant telluric absorption.

2.4 Kinematic templates

The choice of kinematic template(s) is important to extract kinematics accurately from the galaxy spectra. The systematic errors introduced into kinematics by use of a poorly matched template are well studied (van der Marel et al. 1994), although it is difficult to quantify such effects numerically. At best, we can say that different templates appear to introduce systematic offsets into the velocity profile (VP) parameters. Ideally, a large library of spectral types should be available so that an optimal mix can be found, but in this case, because of time constraints, only two templates were observed with the same instrumental setup as the NGC 1399 observations. However, it is possible to check if the templates are well matched to the luminosity-weighted population of the galaxy.

The CO (2–0) bandhead at 2.2935 μm is an indicator of stellar type: a linear relation has been found between the equivalent width (EW) of the first CO feature $W_{\mathrm{CO(2-0)}}$ and the stellar type (Kleinmann & Hall 1986; Origlia et al. 1993). Although this relation is based on observations of individual stars, it can be extended to galaxy populations by applying a correction based on the velocity dispersion of the system (Oliva et al. 1995; Thatte, Tecza & Genzel 2000). The relationship differs between giants and supergiants, but this should not be a problem for the old, giant-dominated, population of an elliptical galaxy such as NGC 1399. To estimate the galaxy EW, the (rest-frame) wavelength range over which the CO (2–0) EW is defined must be shifted to the velocity frame of the galaxy and the EW measurement must be corrected for the galaxy's velocity dispersion. The kinematic properties of the galaxy are not known a priori and depend on the template used to extract them. However, it is possible to account for this and calculate reasonable limits on the EW of the galaxy.

Such analysis was performed for NGC 1399. The variation of EW with velocity dispersion was simulated using the kinematic templates and the necessary quadratic correction (in σ) found; this was then applied to the galaxy EW measurements. Furthermore, all measurements were scaled by a constant factor (the correction of Oliva et al. 1995) to correct for the instrumental resolution of CONICA, allowing direct comparison with the relation of Origlia et al. (1993). The results are presented in Fig. 4 and discussed in Section 4.

2.5 AO correction and PSFs

The quality of the AO correction can be estimated from the reference star, which was observed in the slit simultaneously with the galaxy. The 1D profile of the star (calculated from summing the flux over the same wavelength range as that from which the kinematics are extracted) is well fitted by a double Gaussian, as shown in Fig. 2.

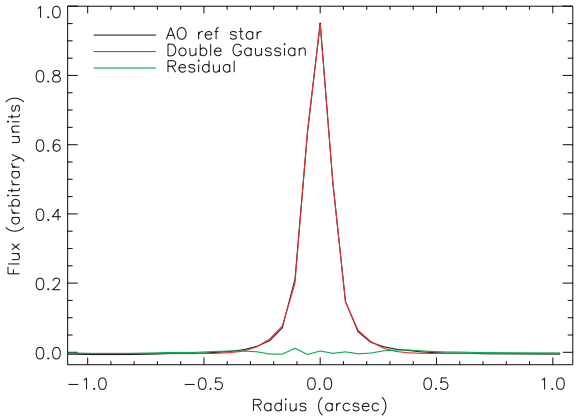

Figure 2. The profile of the AO reference star together with a double Gaussian fit and the residuals from this fit. The best-fitting σ for each of the Gaussians are 0.048 and 0.122 arcsec and the total fluxes (γ) are in the ratio 2:1, respectively. This 1D slit profile of the AO reference star is used to estimate the 2D on- and off-source PSFs.

However, the correction and subsequent PSF will vary further away from the reference star. Although the exact 'off-source' PSF at the centre of the galaxy is unknown, it can be estimated. Using the NAOS preparation software v1.74,[6] one is able to simulate how the PSF varies with seeing, airmass and distance from the reference star. However, the simulated PSFs do not account for slit effects, non-perfect data reduction (such as error in the field transformation or frame alignment) or drift of the the tip-tilt correction during long exposures. All these effects will further broaden the PSF. To account for such, we convolve simulated 2D on-source PSFs (with various airmass and seeing conditions) with a $\theta \times 3.2$ top hat (3.2 pixel is the slit width and θ accounts for imperfect data reduction and tip-tilt drift). We optimize θ and the atmospheric conditions by minimizing the difference between the 1D profile along the slit length ($y = 0$) and the observed 1D profile of the AO reference star (Fig. 2). Good agreement between the simulated and observed PSFs is found for 0.6 arcsec seeing, an airmass of 1.2 and $\theta = 2.2$ pixel. To estimate the off-source PSF at the galaxy centre, we simulate the 2D off-source PSF with identical atmospheric conditions and convolve it with the same 2D kernel (we assume that the broadening effects are uniform over the field of view). The resulting on- and off-source PSFs are both well fitted by a double 2D Gaussian

$$\mathrm{PSF}(x, y) = \frac{\gamma_1}{2\pi\sigma_{x1}\sigma_{y1}} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{x1}}\right)^2 + \left(\frac{y^2}{\sigma_{y1}}\right)^2\right] + \frac{\gamma_2}{2\pi\sigma_{x2}\sigma_{y2}} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{x2}}\right)^2 + \left(\frac{y^2}{\sigma_{y2}}\right)^2\right]. \quad (5)$$

The best-fitting parameters are given in Table 1 and are used to describe the 2D PSF in the dynamical modelling. The Strehl ratios also quoted in Table 1 are calculated from the unbroadened (pre-convolution) simulated PSFs: it makes no sense to compare the ideal Airy disc pattern with the slit-convolved PSF of the spectrograph. The FWHM of the off-source PSF along the x-axis is 0.15 arcsec (2.75 pixel), corresponding to 14 pc, which we adopt as our formal resolution.

The effect of contamination from the seeing-limited halo on the spectra and the knock-on effect on the VPs is not known precisely.

[6] See http://www.eso.org/observing/etc/naosps/doc/NAOS-PS-tool.html

Table 1. The characteristics of the on- and off-source PSFs: the best-fitting parameters (γ_1, σ_{x1}, σ_{y1}, γ_2, σ_{x2}, σ_{y2}) of a 2D double Gaussian (5) fit to the PSFs after broadening; the FWHM of both broadened simulated PSFs along the x- and y-axes (parallel and perpendicular to the slit, respectively); and the Strehl ratios of the simulated PSFs prior to broadening.

Parameter	On-source	Off-source
γ_1	0.141	0.135
σ_{x1}	0.956	1.071
σ_{y1}	1.214	1.280
γ_2	0.010	0.129
σ_{x2}	3.667	3.501
σ_{y2}	2.784	2.630
FWHM along x (arcsec)	0.135	0.150
FWHM along y (arcsec)	0.181	0.189
Strehl ratio (per cent)	38	30

According to (7), if the galaxy spectrum is a weighted sum of many different spectra with many different VPs, the problem is linear and χ^2_s will be minimized for a similarly weighted mean VP. The galaxy spectra and kinematics are therefore only expected to be 'diluted' by this effect. However, the dynamical modelling uses the estimate of the off-source PSF to account *fully* for the shape of the PSF; the peculiar shape of the AO-corrected PSF does not bias the models or the derived BH mass.

The quality of the PSF was assessed as a function of time from the AO reference star in the slit. The quality of the overall AO correction for the final co-added galaxy data can be marginally improved with frame selection. However, the cost in S/N was too high for any significant improvement in this case, so frame selection was not implemented.

3 KINEMATICS

We assume that each galaxy spectrum $G(\lambda)$ is a Doppler-broadened version of some underlying 'average' stellar template $T(\lambda)$. If we could completely remove all continuum features from G and T, then we would have that

$$G(\lambda) = T(\lambda) \otimes L(v), \tag{6}$$

where $L(v)$ is the unknown line-of-sight stellar velocity profile and $\otimes$ denotes convolution. In reality, it is impossible to remove all continuum contamination completely from observed spectra by eye, nor do we know the correct template T to use. Furthermore, both G and T are measured with finite S/N, making it impossible to use a simple deconvolution method to obtain $L(v)$ directly from (6).

Many methods have been proposed to address these problems. We follow Rix & White (1992), Saha & Williams (1994) and van der Marel (1994) and extract $L(v)$ using a direct pixel-fitting method. Taking the continuum-divided star and galaxy spectra, we find $L(v)$ and its associated uncertainties by minimizing

$$\chi^2_s \equiv \sum_{i=1}^{N_p} \left[\frac{G_i - k(T \otimes L)_i - \sum_{l=0}^{2} c_l (\ln \lambda_i)^l)}{\Delta G_i} \right]^2, \tag{7}$$

where ΔG_i is the measurement error in the ith galaxy spectrum pixel, N_p is the number of pixels being fitted and the third term corrects for any continuum that escapes our initial continuum division. The parameter $k \simeq 1$ accounts for differences in normalization between G and T: we want our VPs to be normalized with $\int L(v)\,dv = 1$. We further assume that $T(\lambda)$ is well approximated by a weighted average of known stellar templates. As we only have two stellar templates available for kinematic extraction, we define our optimal template to be

$$T(\lambda) = f T_1(\lambda) + (1 - f) T_2(\lambda), \tag{8}$$

where f defines the relative fraction of each of the two available templates (T_1, T_2).

Unlike Fourier methods (Richstone & Sargent 1972; Sargent et al. 1978; Franx et al. 1989; Bender 1990; van der Marel & Franx 1993), pixel fitting does not require any assumptions about window functions and allows us to propagate the pixel-to-pixel error estimates in our measured spectra directly into uncertainties in our line-of-sight velocity profiles (VPs). Note that the Na I doublet at 2.21 μm does *not* appear to be well fitted by our templates (the linestrength in the galaxy spectra is much higher than that in our templates). Thus the fitting range for extraction of kinematics is from 2.249 to 2.438 μm (rest-frame) to include the Ca I feature and the CO bands. We use two different parametrizations for $L(v)$, each of which is discussed separately below.

3.1 Gauss–Hermite parametrization of the VP

VPs are expected to be reasonably close to Gaussian. A convenient way of parametrizing a VP $L(v)$ is by using a truncated Gauss–Hermite expansion (Gerhard 1993; van der Marel & Franx 1993),

$$L_{GH}(v) = \frac{\gamma}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \left(\frac{v - V}{\sigma} \right)^2 \right] \sum_{i=0}^{N} h_i H_i \left(\frac{v - V}{\sigma} \right), \tag{9}$$

in which one starts from a Gaussian with scalefactor γ, mean V and dispersion σ and uses a weighted sum of Hermite polynomials H_i to quantify deviations of $L(v)$ from this underlying Gaussian. This parametrization is degenerate: there is a different set of coefficients $\{h_i\}$ for each choice of (γ, V, σ), and a number of different procedures have been used to fit spectra using (9). Before describing our own procedure, we review the motivation for these earlier methods.

3.1.1 Perfect data

Consider first an idealized situation in which we knew $L(v)$ perfectly. Then there is a unique set of expansion coefficients $\{h_i\}$ for any (sensible) choice of Gaussian (γ, V, σ). To see this, recall that the H_i satisfy the orthogonality relation (van der Marel & Franx 1993),

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-x^2) H_i(x) H_j(x)\,dx = \frac{1}{2\sqrt{\pi}} \delta_{ij}, \tag{10}$$

from which it is easy to see that choosing

$$h_i = \frac{1}{\sqrt{2}\gamma} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2} \left(\frac{v - V}{\sigma} \right)^2 \right] L(v) H_i \left(\frac{v - V}{\sigma} \right) dv \tag{11}$$

minimizes the mean-square deviation

$$\chi^2_0 \equiv \int_{-\infty}^{\infty} [L(v) - L_{GH}(v)]^2\,dv \tag{12}$$

of the expansion (9) from the VP $L(v)$. Since the H_i form a complete set, we can make χ^2_0 arbitrarily small for smooth $L(v)$ simply by

increasing the number of terms N included in the series, although the expansion (9) is of course useful in practice only if it can provide a good fit for small N.

We note that, when minimizing χ_0^2 at fixed (γ, V, σ):

(i) the Gauss–Hermite coefficients h_i are modified moments (11) of $L(v)$;

(ii) the Hessian $\partial^2\chi_0^2/\partial h_i\,\partial h_j$ is diagonal, meaning that the h_i are independent;

(iii) the h_i are also independent of the choice of N;

(iv) if we choose (γ, V, σ) to be the parameters of the Gaussian that minimize (12), then $(h_0, h_1, h_2) = (1, 0, 0)$ and (h_3, h_4) measure the lowest-order antisymmetric and symmetric deviations from this Gaussian.

Because of this last point, most Gauss–Hermite parametrizations of VPs fix $(h_0, h_1, h_2) = (1, 0, 0)$ and use $(\gamma, V, \sigma, h_3, h_4, \ldots, h_N)$ as the free parameters in the fit. For our purposes, this approach has the drawback of introducing non-trivial correlations between the Gaussian parameters (γ, V, σ) and the $\{h_i\}$, making these parameters cumbersome to use when comparing to dynamical models. Cappellari & Emsellem (2004) highlight these correlations.

3.1.2 Real data

In reality we do not have direct access to $L(v)$. Instead we constrain it by investigating how well a parametrized form, such as (9), affects the fit to the discretely sampled galaxy spectrum G_i using the χ_s^2 given by (7). As van der Marel & Franx (1993) point out, equation (7) reduces to (12) in the limit of high-resolution, finely sampled spectra and sharp template features. Here we consider the case where this limit does not apply. Let us assume that the continuum has been perfectly removed from G_i and that T is the correct stellar template. Then (7) becomes

$$\chi_s^2 = \sum_i \left\{ \frac{[T \otimes (L - L_{GH})]_i + n_i}{\Delta G_i} \right\}^2, \tag{13}$$

where $L(v)$ is the galaxy's real underlying VP and $n_i \equiv G_i - (T \otimes L)_i$ is the noise in the ith pixel. Now if we fix (γ, V, σ) and minimize (13) with respect to the $\{h_i\}$, then:

(i) in the absence of noise, the coefficients h_i are still the modified moments (11) of $L(v)$;

(ii) the h_i are not independent since the Hessian $\partial^2\chi_s^2/\partial h_i\,\partial h_j$ is no longer diagonal;

(iii) since the h_i are not independent, in the presence of noise there is a different set of h_i for each choice of N;

(iv) if we choose (γ, V, σ) to be the parameters of the best-fitting Gaussian to (7), then the minimum χ_s^2 will not occur at precisely $(h_0, h_1, h_2) = (1, 0, 0)$.

Our procedure for fitting Gauss–Hermite coefficients is motivated by these points and by our desire to have a set of parameters that depend linearly on $L(v)$. The procedure is as follows:

(i) Choose (γ, V, σ) to be the parameters of the best-fitting Gaussian to the VP. Find (γ, V, σ), template fraction f and continuum parameters c_l that minimize (7) with $k = h_0 = 1$ and $h_1 = h_2 = \ldots = 0$.

(ii) Holding (γ, V, σ) fixed at their best-fitting values, find the h_i, c_l and f that minimize (7). Having found these h_i, the normalization is $k = \gamma(h_0 + h_2/\sqrt{2} + h_4\sqrt{3/8} + \cdots)$.

(iii) Finally, divide γ and the h_i by k.

We use a standard Levenberg–Marquardt routine to carry out the minimizations in the first two steps. Our best-fitting parameters are $(h_0, \ldots, h_N)$, along with their covariances and the choice of (γ, V, σ). Note that there are no errors associated with (γ, V, σ) in our version of the Gauss–Hermite parametrization: they merely reflect the Gaussian around which we have chosen to expand $L(v)$. Choosing the best-fitting Gaussian here lets us make a straightforward comparison of our kinematics with earlier work. In practice we find that our method yields (h_0, h_1, h_2) that differ from (1, 0, 0) by around (0.1, 0.04, 0.04), but with very strongly coupled errors among all the even h_i. We describe how we deal with these covariances in Section 6.3 below.

We have tried using standard simultaneous $(N + 1)$-parameter fits to $(\gamma, V, \sigma, h_3, \ldots, h_N)$ (van der Marel & Franx 1993), but our VPs are so strongly non-Gaussian at the centre of the galaxy that the usual linear approximations among the errors in these parameters (van der Marel & Franx 1993) break down and we find multiple minima in χ_S^2 (7). For spectra at $(-0.08\,\text{arcsec}, 0.02\,\text{arcsec})$ this process can yield alarmingly low values of σ ($\sim$250 km s^{-1}) and correlated $h_4 > 0.6$ (Cappellari & Emsellem 2004) which describe a triple-peaked VP. In fact, for such values of h_4, it can be shown that even the idealized χ_0^2 (12) has multiple minima in σ and h_4.

3.2 VP histograms

One might expect that the putative BH in NGC 1399 would cause high-velocity wings in the central VPs, which might not be captured well by the low-order Gauss–Hermite parametrization above. Therefore, we also fit 'non-parametric' VPs, in which we choose n_v regularly spaced velocities $v_1 < v_2 < \cdots < v_{n_v}$ and parametrize $L(v)$ as the histogram

$$L_H(v) = \sum_{i=1}^{n_v} L_i S_i(v), \tag{14}$$

where the step function $S_i(v) = 1$ if $v_i < v < v_{i+1}$ and is zero otherwise.

Given parameters $L_1, \ldots, L_{n_v}$ it is straightforward to calculate the convolution (6) of this $L_H(v)$ with a stellar template. For any given galaxy spectrum G there will be many sets of parameters that produce good fits to the spectrum, but most of them will be unrealistically jagged. Therefore, instead of minimizing the χ_s^2 given by equation (7) directly, we minimize the penalized $\chi_p^2 = \chi_s^2 + P[L_i]$, where the penalty function

$$P[L_i] = \alpha \sum_i (L_{i+1} - 2L_i + L_{i-1})^2 \tag{15}$$

uses the mean-square second derivative of $L(v)$ as a measure of the jaggedness of the solution.

Our procedure for fitting L_i is simple:

(i) Find the best-fitting smooth L_i, continuum parameters c_l and template fraction f by minimizing the penalized $\chi_p^2 = \chi_s^2 + P[L_i]$ with $k = 1$.

(ii) Set the normalization factor to $k^{-1} = \sum_i (v_{i+1} - v_i)L_i$ and rescale the L_i.

This makes no attempt to impose the obvious non-negativity constraint on the L_i. While the resulting VP histograms are fine for 'by-eye' comparisons of one VP against another, they are not suitable for direct comparison against models. So, in Section 6.3 we describe a variation on this fitting procedure that takes account of the correlations among the L_i and removes the bias introduced by the penalty function.

Based on an average separation of 300 Å between the CO bands in the fitting range [^{12}CO (2–0), ^{12}CO (3–1), ^{12}CO (4–2), ^{12}CO (5–3), ^{12}CO (6–4)], the maximum relative velocity we can reasonably hope to measure is $\sim$1900 km s^{-1}. The systematic velocity of NGC 1399 is $\sim$ 1500 km s^{-1}, so we divide each LOSVD into $n = 50$ equispaced velocity points between $v_1 = -1000$ km s^{-1} and $v_n = 4000$ km s^{-1}. We choose $\alpha = 4 \times 10^7$, which is the minimum value required to give smooth non-parametric VPs consistent with our outer Gauss–Hermite VPs.

4 RESULTS: IMAGING AND KINEMATICS

An 80 s, AO-corrected, *Ks* exposure of NGC 1399 is shown in Fig. 3. It has the same PA as the long-slit observations, and a slit image (0.172 arcsec wide and centred on the AO reference star) has been overlaid. If the reference star is perfectly centred in the slit, there is a maximum position error of 0.3 pixel (1.6 pc) on the brightest region of the galaxy from error in the PA. However, the acquisition images (incorporating flat-fielding, etc., and using the slit images) indicate that the centring of the star along the minor axis of the slit was accurate to only 1 pixel (5.2 pc); the slit was approximately 3 pixel wide (0.172 arcsec or 17 pc). Hence, the dominant error in aligning the slit on the galaxy centre is from shifts along the minor axis, not from error in the PA. The former is expected to be random about the centre of the slit; the latter would give a systematic difference for all observations. The final position accuracy is therefore good: the slit was aligned on the brightest point of the galaxy to an accuracy of around $\pm$0.06 arcsec ($\pm$6 pc), with no significant systematic offset.

Two globular clusters are also seen in the image data close to the centre of the galaxy. The nearest is 1.15 arcsec to the east of the galaxy centre. A Gaussian fit to this globular cluster yields an estimate on the FWHM of the *image* PSF to be 0.078 arcsec (7.5 pc). This is equal to the *on-source* FWHM predicted by the NAOS preparation software in Section 2.5, suggesting that the single Gaussian was probably fitted to the diffraction-limited component of the globular cluster profile and ignored the seeing-limited halo, although the atmospheric conditions were excellent around the time of observation, with the seeing occasionally dropping below 0.4 arcsec.

The parametrized kinematics of the central few arcseconds of NGC 1399 are shown in Fig. 4. Where possible, previous data are overplotted (Graham et al. 1998; Longo et al. 1994). The flux is calculated by summing over all available wavelengths at each point along the slit. The parameters (γ, v, σ) are chosen to be the best-fitting Gaussian parameters and the Gauss–Hermite coefficients h_n

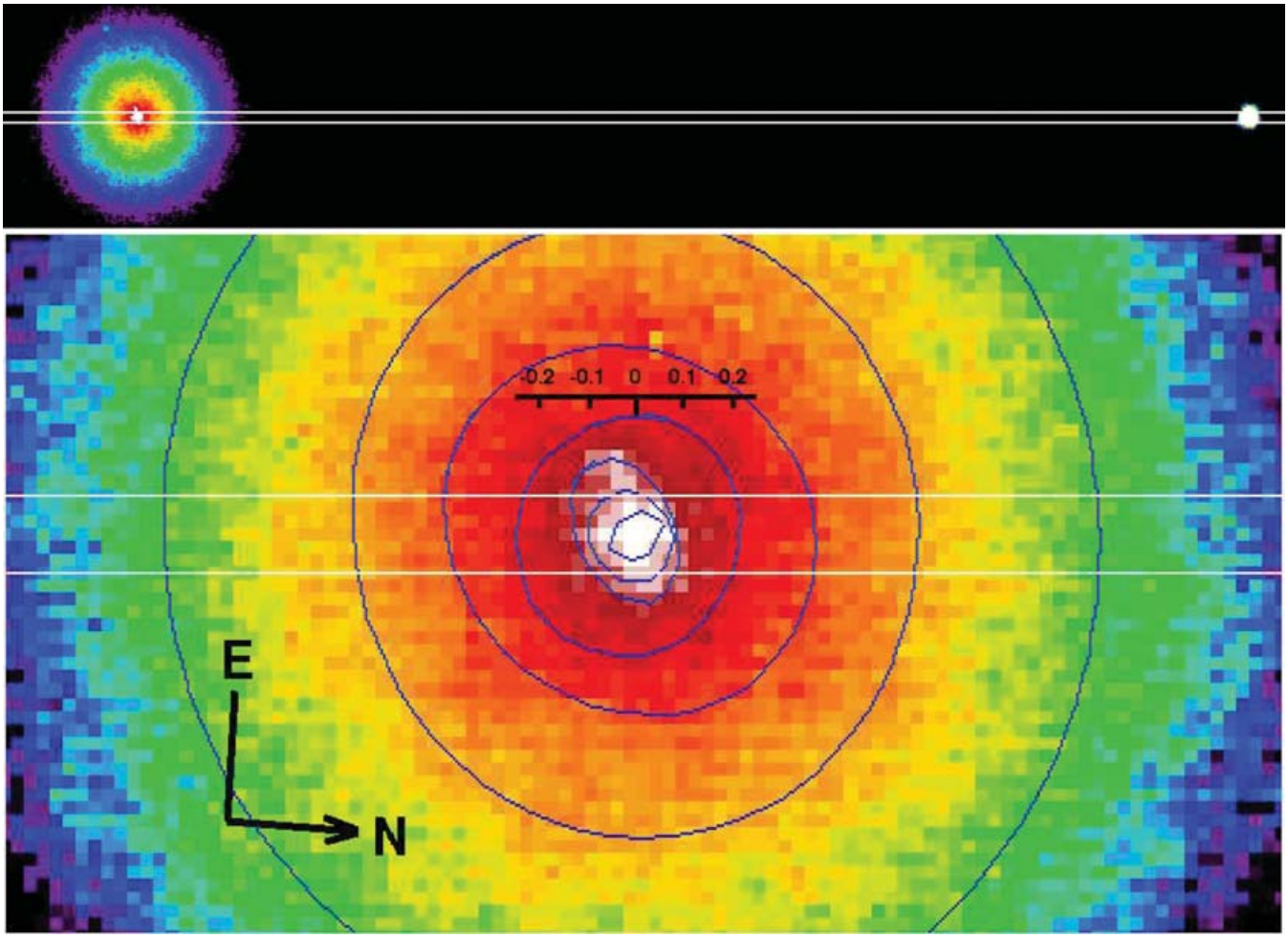

Figure 3. Top: A *Ks*-band image of NGC 1399 showing the slit alignment (white) on the galaxy and the AO reference star; the PA of the slit is 5.06°, so that the reference star is approximately due north of the galaxy nucleus; the nucleus of the galaxy and the star are separated by 17.6 arcsec. Bottom: The same as the above but magnified and centred on the nucleus of NGC 1399 with isophote ellipses overplotted in blue. Note the elongation of the nucleus to the south-east. Each pixel is 27 mas wide (the pixel size of the spectroscopic data was twice this at 54 mas). The angular scale is given in arcseconds.

are derived using these best-fitting Gaussian parameters. The velocity v is given relative to the mean heliocentric velocity of 1467 ± 4 km s^{-1}. Note that the error quoted for this velocity is an estimate of the random error only.

Fig. 5 compares the non-parametric and parametric VPs for the central arcsecond of the galaxy. Error bars are given for the non-parametric VPs. The galaxy spectra and best-fitting broadened template (constructed from the Gauss–Hermite VP) for the central arcsecond are also shown.

4.1 Highlights

Many interesting features are present in Figs 3, 4 and 5. The parametrized kinematics show a decoupled velocity structure across the galaxy centre, a double-peaked velocity dispersion across the centre separated by 0.2 arcsec or 19 pc (with the dip in σ half a pixel off the galaxy centre) and significant variations in the h_i with radius. The imaging shows offset asymmetric isophotes at the galaxy centre, and the non-parametric VPs (while generally in good agreement with the parametric VPs) show strong high-velocity wings at $r = -0.08$ arcsec and lopsided velocity structure at $r = 0.14$ arcsec, which are not detected with the parametrized fit.

4.2 Further details

The kinematic features described above persisted when the spectra were reduced using the standard sky subtraction technique for ABBA sequence observations (rather than using an interpolated sky), albeit with increased noise. Angular distances in Figs 4 and 5 have been shifted such that the brightest region of the galaxy lies at $r = 0$ arcsec. This position was determined by scaling to the (outer) centroid of a Nuker profile (Lauer et al. 1995) fitted to the flux of Fig. 4. The photometry is not symmetric at the centre, which could bias the calculation, although when the inner 0.5 arcsec of the light profile were omitted the change in the centroid was less than 0.1 spectroscopic pixels (5 mas or 0.5 pc). The centroids of the ellipses fitted to Fig. 3 are not constant and fluctuations of up to 0.5 image pixels (13.5 mas or 1.3 pc) from the mean are present. The scale on Fig. 3 is aligned to the mean x-axis centroid of the ellipses. Registration between the angular scale on the image and that of the kinematics may not be perfect as the zero-points were determined independently. However, it is likely to be no worse than 0.05 arcsec.

EW measurements for the two kinematic templates and the galaxy are shown in Fig. 4. The template EWs (orange for the M0III and blue for the K4III) allow for ±20 km s^{-1} errors in systemic velocity and instrumental dispersion. The corrected EW of the galaxy is shown in black, while the effect of ±20 km s^{-1} errors in velocity and dispersion are shown in red. The asymmetry in the error is associated with the error in velocity: large positive or negative changes in velocity move the CO (2–0) feature out of the predefined wavelength limits (Origlia et al. 1993), thus reducing the EW. The kinematics (v, σ) used to correct the galaxy EW are those shown in Fig. 4. Systematic differences in (v, σ) from using different templates in kinematic extraction are less than 20 km s^{-1} in each case. These EW measurements indicate that our kinematic templates are well matched to the luminosity-weighted stellar population of the galaxy. With reference to Origlia et al. (1993), one can see that the corrected EW measurements of the two templates (11.3 and 12.5 for the K4III and M0III, respectively) are typical of their spectral classifications. The galaxy EW (average value of 12.1) is closer to that of the M0III template; this is also reflected by the favouring of the M0III template in the kinematic extraction (Fig. 4). However,

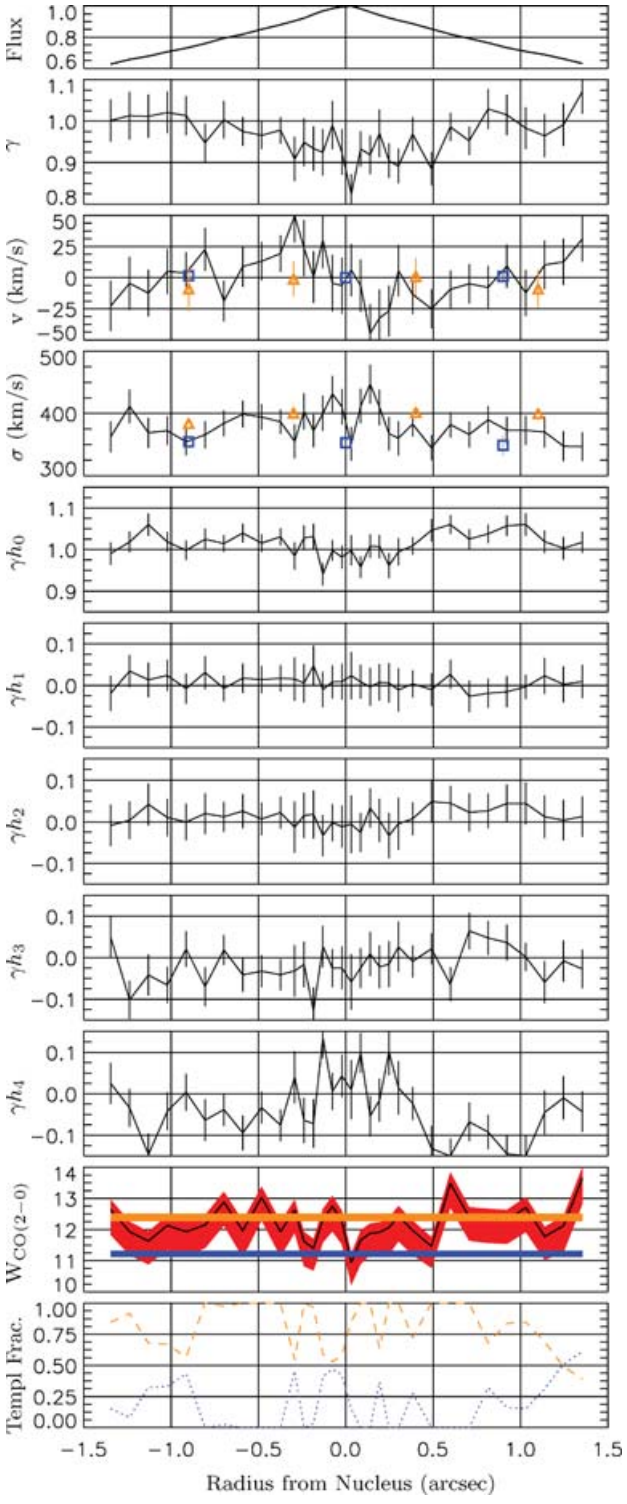

Figure 4. The nuclear kinematics of NGC 1399. The top plot indicates the flux received at each position on the slit; the next eight plots show the variations in the parameters of (9); and the lowermost two show the template and galaxy EWs and the optimal kinematic template composition as a function of radius from the nucleus. Where applicable, previous data are shown as orange triangles (Longo et al. 1994) and blue squares (Graham et al. 1998). In the EW plot, the kinematic template EWs are shown in orange (M0III) and blue (K4III); the width represents a 20 km s^{-1} error in v and $\sigma_{\rm inst}$. The EW of NGC 1399 is in shown in black (red shows the combined effect of 20 km s^{-1} errors in v and σ). The PA of the slit was 5.06°, so that the positive radius is northward.

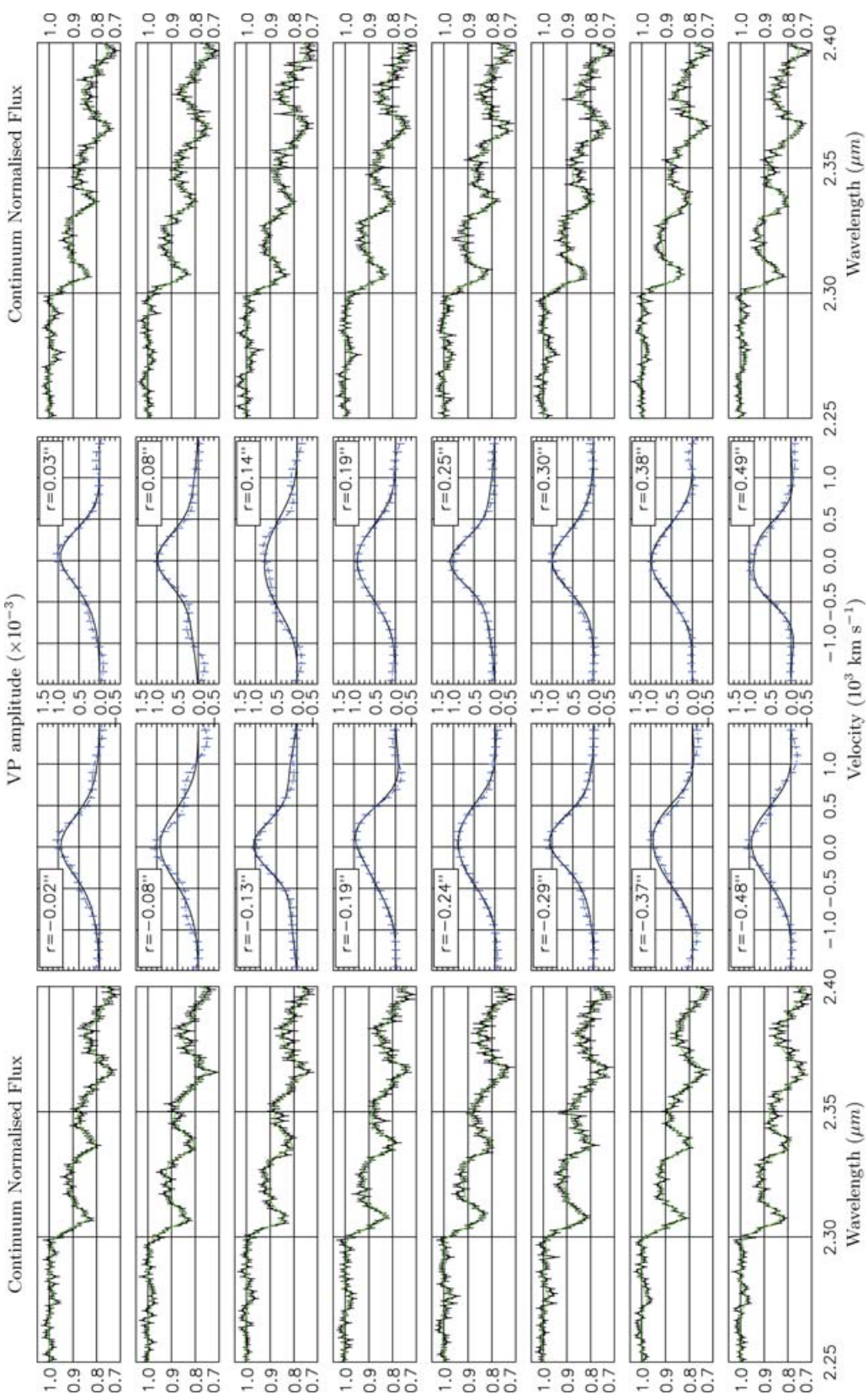

Figure 5. The spectra and VPs of the central arcsecond of NGC 1399. The centroid of the light profile along the slit (taken to be the centre of the galaxy) is roughly in between two pixels allowing for the VPs to be constructed at approximately equidistant intervals to the left and right of the nucleus, without interpolation. For every point along the slit, two VPs are plotted: black is the Gauss–Hermite VP (9) and blue is the non-parametric VP. The velocity axis of the VPs has been corrected to the mean systematic velocity of the system. Accompanying every spectrum is the best-fitting broadened template (green) reconstructed from the Gauss–Hermite VP.

one should be apprehensive about the variations of the galaxy's EW with radius: the dispersion correction assumes a Gaussian VP, which we have seen is not always the case. Overall, the analysis of EWs indicates that the templates are well matched to the galaxy population, so we should not expect problems associated with template mismatch to be present in the kinematics. Additionally, previous (seeing-limited) data on the central velocity dispersion of NGC 1399 agree with our data (Fig. 4), given the different resolutions. Saglia et al. (2000) report slightly negative h_4 towards the galaxy centre, which we confirm, although our rise in h_4 in the very central 0.5 arcsec is beyond the (seeing-limited) spatial resolution of their data.

To maintain a high S/N at larger radii ($|r| > 0.3$ arcsec), the spectra were binned into pairs prior to kinematic extraction. However, after binning, χ^2_s (7) rose, on average, from ∼160 to ∼230 (an approximate change of $\sqrt{2}$) for a fit to 174 data points. This indicates that systematic errors are starting to become comparable to the random errors. There are many possible causes of systematic errors. The optimal template convolved with the best VP will not reproduce all the features in the galaxy spectrum; although the galaxy and template spectra have comparable CO (2–0) EWs, there may be absorption or even emission features in the galaxy spectra that are not obvious and are not accounted for in equation (7), but still contribute to the differences between the galaxy and templates and to the systematic 'noise'. Telluric correction is only estimated to be accurate to 1 per cent, increasing to 2–3 per cent at regions where the telluric absorption shows sharp prominent features. In addition, the stars used to obtain a telluric spectrum are unlikely to have featureless spectra at high S/Ns. The S/N of the data binned at 0.3 arcsec over the wavelength range used to extract kinematics falls from 110 (2.25 μm) to 30 (2.45 μm) per pixel due to the thermal background (the S/N at the unbinned galaxy centre falls from 90 to 20). This S/N (and corresponding error estimates) quantify random error alone and do not account for systematics introduced by telluric correction. Binning will reduce the random noise but not the systematic noise. Thus as χ^2_s (7) is weighted by the random noise estimates, when the random error in each spectrum is reduced by binning, χ^2_s will increase so long as the systematic noise persists.

5 DISCUSSION

The interpretation of the results is discussed below in appropriate sections.

5.1 Decoupled kinematics

There is a strong rotation gradient in v within a radius of 0.5 arcsec (48 pc), which is clearly decoupled from the kinematics at larger radii. The magnitude of the central rotation (taking the difference of the maximum velocity in each direction) is $\sim 70\,\mathrm{km\,s^{-1}}$. The absence of high-resolution data perpendicular to our long-slit position prevents us from concluding if the system is truly counterrotating, or just decoupled. Previous publications considered the possibility of a decoupled system (D'Onofrio et al. 1995; Saglia et al. 2000), but the poorer spatial resolution of the data prevented a reliable detection. Kinematically decoupled cores are often (but not exclusively) found in spherical, high-dispersion systems with core-like photometry (Emsellem et al. 2004).

5.2 The central dispersion profile

The drop in σ half a pixel off the galaxy centre is accompanied by less convincing dips in γ and the CO (2–0) EW. These features are close to the limit of the spatial resolution and thus may be unresolved. It is well known that random errors in γ and σ are statistically correlated (Efstathiou, Ellis & Carter 1980) and these features are only clearly seen when continuum correction is incorporated into the minimization of χ^2_s (7); if continuum correction is left out of the minimization [i.e. calculated once from the continuum shortward of the CO (2–0) feature], γ shows no obvious change (and is noisier in general) and the dispersion becomes flat at the centre.

The decrease in γ and σ could be correlated effects from fitting constrained parametrized VPs to more complicated profiles; γ is expected to fall at the centre (van der Marel 1994). However, the non-parametric VP at $r = 0.03$ arcsec matches the parametric form well. Alternatively, the dip in σ could reflect a genuine fall in the dispersion of the stars at the centre of the galaxy, but this is the exact opposite to what one would expect when approaching the BH.

An alternative explanation is the presence of young stellar population at the centre, although γ and σ may not necessarily decrease in this case: a young stellar population may include younger stellar types with reduced CO (2–0) EWs, but it would also have an increased fraction of supergiants, which have a higher CO (2–0) EW than giants for the same MK spectral type (Origlia et al. 1993). In addition, one would expect a correlated change in the optimal template used to extract the kinematics, which is not the case here. One would expect a different stellar population to produce a colour gradient but unfortunately there is insufficient homogeneous data at the required resolution to confirm this.

Central dips in σ have been seen in other core galaxies such as M87 and NGC 4649 (van der Marel 1994; Pinkney et al. 2003), which both have nuclear activity. If NGC 1399 contained a central *nucleus* of light from a non-thermal, non-stellar object (van der Marel 1994), one would expect the CO bands to show a lower γ relative to the continuum level, which may introduce correlated dips in γ and σ. New photometry by Lauer et al. (2005) shows a slight excess of light at the galaxy centre after subtraction of the best-fitting Nuker profile and therefore suggests the presence of a very faint nucleus. However, the authors emphasize the possible dangers and unknown systematics of extrapolating the Nuker fit into the central regions.

5.3 Offset photometry and peculiar VPs

The *Ks* band image of NGC 1399 shows a departure from spherical symmetry in the central 0.5 arcsec with an elongation of the surface brightness towards the east-south-east, approximately 0.2 arcsec (19 pc) in length. The image is a relatively short exposure and the S/N is low. However, the elongation of the central isophotes covers more than 16 pixel; the probability of finding 16 neighbouring pixels above the mean, given the noise statistics of the image, is negligible, so we can rule out the possibility of random error causing such an artefact. Obvious systematic errors that may cause elongation of the central isophotes would be either an error in frame alignment, or an unusual PSF from the AO correction. However, both these effects can be dismissed: the globular cluster found 1.15 arcsec from the galaxy centre is circular and well fitted by a 2D Gaussian of equal width in each dimension; no significant residual is seen after subtraction. We note that *HST H*-band images exist for NGC 1399, but owing to a poorer spatial resolution (0.13 arcsec) and PSF sampling, these structures are not observed. The non-spherical isophotes at the centre of NGC 1399 are a genuine feature of the galaxy. There appear to be peculiar kinematic features associated with the photometric anomaly: at $r = -0.08$ arcsec, the non-parametric VP contains

substantial high-velocity wings; and at $r = 0.14$ arcsec, the non-parametric VP is asymmetric with an excess of receding velocity structure. We discuss possible explanations for these features below.

Isophote twists are common in core regions of core galaxies (Lauer et al. 2005). However, such twists are generally smooth and seen on large scales, which is not the case here. The elongation could be a projection effect from the obscuration by a non-uniform distribution of dust. However, no obvious dust signature can be seen in either this *Ks* band image or archival *HST V*-band images. The density distribution of the dust would need to vary rapidly to cause such a sudden effect in the NIR on such small spatial scales (0.3 arcsec or 29 pc). However, neither of these explanations would give rise to the associated kinematic features.

We have already seen that globular clusters (GCs) have been resolved near the centre of NGC 1399. Although the elongation of the nucleus does not appear to be separate from the central maximum, it is plausible that a GC happens to lie in the line of sight between ourselves and the galaxy centre and is unresolved from the central photometric maximum of the galaxy. There would be a kinematic effect from such a chance alignment: one would expect a 'spike' in the VP at the systematic velocity of the GC caused by the low dispersion velocity structure of the globular cluster. However, we see high-velocity wings in the VP at $r = -0.08$ arcsec and a lopsided structure at $r = 0.14$ arcsec, which is not the same.

The elongation of the central isophotes is more likely to be an offset centre as seen in a handful of other *core* galaxies (Lauer et al. 2005). In fact, Lauer et al. (2005) argue that the offset centres detected in new WFPC2 photometry are all eccentric discs, analogous to that of M31 (Tremaine 1995). Certainly, this would produce a strong kinematic signature in the form of high-velocity wings if our slit bisected the eccentric disc. Our slit does not completely bisect the peculiar photometry and nor does it coincide with the PA of the eccentric isophotes. However, the alignment of the AO reference star in the slit varied by around 1 pixel and this translates to a potential *wandering* of the slit, perpendicular to its length, by two image pixels. Thus, there is likely to be considerable contribution from the elongated photometry in the spectroscopic data.

Keeping in mind the problems with registration between the image and kinematics, an eccentric disc at the centre of NGC 1399 would create high velocity wings at $r = -0.08$ arcsec, where the slit is closest to bisecting the anomaly. If the non-parametric VP at this point is to be trusted, the velocity dispersion of just the wings (ignoring the central bulk) is in excess of 1000 km s^{-1}. Furthermore, if the stars in the eccentric disc rotate in the same direction, one would expect a lopsided VP at the pericentre of the ellipse if the disc was viewed near edge-on. We do see a strong excess or 'hump' of structure receding away from the observer at $r = 0.14$ arcsec, which is possibly the closest pixel to the pericentre of the elliptical photometry. Although Lauer et al. (2005) report no offset photometry in NGC 1399 with new WFPC2 photometry, the isophote ellipticity is seen to jump from approximately 0 to 0.2 between $0.09 < r < 0.1$ arcsec.

We note that this eccentric disc hypothesis cannot explain the decoupled kinematics discussed in Section 5.1. A coherent disc can only survive well within the SoI of the BH, which we estimate to be $\sim$0.3 arcsec.

6 DYNAMICAL MODELLING

To obtain preliminary constraints on the mass of any BH in NGC 1399, we ignore the mild rotation gradient and elongated central isophotes and fit spherical dynamical models to our central kinematics combined with Graham (1998) velocity dispersion profile. The latter extends to 70 arcsec and provides important constraints on the galaxy's mass-to-light ratio.

We assume that mass follows light, except at the galaxy centre where there can be a BH. The mass density distribution is then

$$\rho(r) = M_\bullet \delta(r) + \Upsilon\, j(r). \tag{16}$$

Our goal is to find the range of BH masses $M_\bullet$ and stellar mass-to-light ratios Υ that are consistent with our kinematics. We take $j(r)$ from the models of Magorrian et al. (1998), which was obtained by deprojecting a composite surface brightness profile constructed by combining *HST* and ground-based photometry. Having this $j(r)$ it is straightforward to calculate the gravitational potential $\psi(r)$ corresponding to (16) for any choice of $M_\bullet$ and Υ. Throughout this paper all mass-to-light ratios are in the *V* band.

Our modelling procedure is a straightforward adaptation of the extended Schwarzschild method described by Rix et al. (1997) and Cretton et al. (1999):

(i) choose trial values for $M_\bullet$ and Υ and calculate the corresponding potential $\psi(r)$;
(ii) follow a representative sample of orbits in this $\psi(r)$;
(iii) find the weighted combination of orbits that minimizes the χ^2 of the fit between the model and the observations, subject to the constraint that each orbit carries non-negative weight;
(iv) assign the likelihood $\exp(-\frac{1}{2}\chi^2)$ to the potential ψ.

We do not impose any regularization in the third step. By considering a range of plausible trial potentials and comparing their relative likelihoods, we obtain constraints on both $M_\bullet$ and Υ.

6.1 Orbit distribution

Apart from the parameters ($M_\bullet$, Υ) describing the potential, the other unknown in our models is the distribution function (DF) $f(\boldsymbol{x}, \boldsymbol{v})$, defined such that $f(\boldsymbol{x}, \boldsymbol{v})\, \mathrm{d}^3\boldsymbol{x}\, \mathrm{d}^3\boldsymbol{v}$ is the luminosity of stars in some small volume $\mathrm{d}^3\boldsymbol{x}\, \mathrm{d}^3\boldsymbol{v}$ of phase space. Jeans' theorem tells us that if our models are to be in equilibrium their DFs can depend on $(\boldsymbol{x}, \boldsymbol{v})$ only through the integrals of motion $\mathcal{E}$ and $\boldsymbol{J}$, the binding energy and angular momentum per unit mass. We make the stronger assumption that the local stellar velocity distribution is symmetric about the v_r-axis, so that the DF depends on $\boldsymbol{J}$ only through its magnitude. We then discretize the DF as a sum of delta functions,

$$f(\mathcal{E}, J^2) = \sum_{i=1}^{n_\mathcal{E}} \sum_{j=1}^{n_J} f_{ij}\, \delta(\mathcal{E} - \mathcal{E}_i)\, \delta\left(J^2 - J_{ij}^2\right), \tag{17}$$

on a regular $n_\mathcal{E} \times n_J$ grid in phase space. The points $\mathcal{E}_i$ are chosen through $\mathcal{E}_i = \psi(r_i)$ with the r_i spaced logarithmically between 1 pc and 100 kpc. This ensures that orbit apocentres are approximately uniformly distributed among each decade in radius. For each $\mathcal{E}_i$ there are n_J values of angular momentum, with J_{ij}^2 running linearly between 0 and $J_\mathrm{c}^2(\mathcal{E}_i)$, the angular momentum of a circular orbit of energy $\mathcal{E}_i$. To avoid a rash of indices we also write the double sum (17) as a single sum over $n \equiv n_\mathcal{E} \times n_J$ points:

$$f(\mathcal{E}, J^2) = \sum_{k=1}^{n} f_k\, \delta(\mathcal{E} - \mathcal{E}_k)\, \delta\left(J^2 - J_k^2\right). \tag{18}$$

6.2 Observables

Having a trial potential $\psi(r)$ and a set of DF components $(\mathcal{E}_k, J_k^2)$, we calculate the unnormalized, PSF-convolved VP histogram of

each component at the projected radius R_i of each of our $n_{\rm AO} = 31$ kinematical data points:

$$L_{ij}^{(k)} = \int_{v_j}^{v_{j+1}} \mathrm{d}v_z \int \mathrm{d}x\,\mathrm{d}y\,\mathrm{PSF}(R_i - x, -y) \times \int \mathrm{d}v_x\,\mathrm{d}v_y\,\delta(\mathcal{E} - \mathcal{E}_k)\,\delta\left(J^2 - J_k^2\right). \quad (19)$$

Here we use a rectangular coordinate system (x, y, z) with origin O at the galaxy centre and Oz-axis parallel to lines of sight. For our standard models each histogram has $n_v = 24$ bins of width $v_j - v_{j-1} = 50\ \mathrm{km\,s^{-1}}$, with innermost bin edge at $v_1 = 0\ \mathrm{km\,s^{-1}}$, the galaxy's systemic velocity, and outermost edge at $v_{25} = 1200\ \mathrm{km\,s^{-1}}$. The function PSF $(\Delta x, \Delta y)$ is the two-dimensional off-source point spread function of Section 2.5 and Table 1. Armed with the L_{ij}^k, the unnormalized VP histogram of a model with DF $(f_1, \ldots, f_n)$ is simply

$$L(R_i; v_j, v_{j+1}) = L_{ij} = \sum_{k=1}^{n} f_k L_{ij}^{(k)}. \quad (20)$$

The normalization constant is the PSF-convolved surface brightness,

$$I(R_i) = I_i = \sum_{k=1}^{n} f_k I_i^{(k)}, \quad (21)$$

where, by analogy with (19),

$$I_i^{(k)} = \int \mathrm{d}x\,\mathrm{d}y\,\mathrm{PSF}(R_i - x, -y) \int \mathrm{d}^3\boldsymbol{v}\,\delta(\mathcal{E} - \mathcal{E}_k)\,\delta\left(J^2 - J_k^2\right). \quad (22)$$

We explain how we evaluate the multiple integrals (19) and (22) in Magorrian et al. (in preparation). The spherical symmetry of our models means that we can afford to calculate these projection coefficients for each of our 31 observed radii directly. More sophisticated axisymmetric models (e.g. Cretton et al. 1999; Gebhardt et al. 2003) usually resort to introducing subgrids to store intermediate quantities in this calculation.

Our treatment of Graham et al.'s (1998) velocity dispersion profile is similar. We assume that each of their data points measures the second moment of the VP convolved with a Gaussian PSF with FWHM 2 arcsec. The (unnormalized) second moment of each DF component is given by (22) with an extra factor of v_z^2 inside the innermost integral.

6.3 Fitting models to observations

6.3.1 Gauss–Hermite coefficients

Calculating the Gauss–Hermite coefficients of our models is simple: the (unnormalized) contribution of the kth DF component to the ith VP is given by $L_{ij}^{(k)}$, with $j = 1, \ldots, n_v$, from which equation (11) allows us to calculate this component's contribution $h_{ij}^{(k)}$ to $\{h_j\}$. However, as we have explained in Section 3.1, the observed Gauss–Hermite coefficients are not independent. It takes only a little effort to include the effects of the covariances among the $\{h_j\}$ into the modelling.

Let us define the column vector $\boldsymbol{h} \equiv (h_0, \ldots, h_N)^{\rm T}$ and let $\hat{\boldsymbol{h}}$ be the vector of coefficients that minimize the $\chi_{\rm s}^2$ of equation (7). For fixed normalization k and continuum parameters c_i, this $\chi_{\rm s}^2$ is a quadratic form in the h_i:

$$\chi^2(\boldsymbol{h}) \simeq \chi^2_{\rm min} + \tfrac{1}{2}(\boldsymbol{h} - \hat{\boldsymbol{h}})^{\rm T} \cdot M \cdot (\boldsymbol{h} - \hat{\boldsymbol{h}}) = \chi^2_{\rm min} + \frac{1}{2}\sum_{i=0}^{N} \lambda_i [\boldsymbol{e}_i \cdot (\boldsymbol{h} - \hat{\boldsymbol{h}})]^2, \quad (23)$$

where the Hessian $M_{ij} \equiv \partial^2 \chi_{\rm s}^2 / \partial h_i\,\partial h_j$ has eigenvalues $\lambda_0, \ldots, \lambda_N$ with corresponding eigenvectors $\boldsymbol{e}_0, \ldots, \boldsymbol{e}_N$. Therefore, the new parameters

$$h'_j \equiv \boldsymbol{e}_j \cdot \boldsymbol{h} \quad (24)$$

have independent errors $\Delta_j \equiv \sqrt{2/\lambda_j}$.

Summing the results from (23) for each of our measured VPs,

$$\chi^2_{\rm GH} = \sum_{i=1}^{n_{\rm AO}} \sum_{j=0}^{N} \left[\frac{\gamma_i \hat{h}'_{ij} - (1/\hat{I}_i)\gamma_i \sum_{k=1}^{n} f_k h'^{(k)}_{ij}}{\Delta_{ij}} \right]^2, \quad (25)$$

where we have dropped the $\chi^2_{\rm min}$ terms and have introduced the $h'^{(k)}_{ij}$, which are related to the $h_{ij}^{(k)}$ by (24). The model's predicted VPs in (25) are normalized by the 'observed' local surface brightnesses $\hat{I}_i$. The latter are difficult to extract from spectroscopic observations. Instead, we obtain them by convolving the photometric profile used to obtain our $j(r)$ profile with the effective spectroscopic PSF. Of course, a reasonable model must also fit this $I(R)$, so we add

$$\chi_I^2 = \sum_{i=1}^{n_{\rm AO}} \left(\frac{\hat{I}_i - \sum_{k=1}^{n} f_k I_i^{(k)}}{\Delta I_i} \right)^2 \quad (26)$$

to (25), somewhat arbitrarily assigning errors $\Delta I_i = 10^{-3}\hat{I}_i$. Finally, we add one more term, $\chi_{\rm G}^2$, to measure how well our models fit both the unnormalized second moments from Graham et al.'s (1998) data and the corresponding PSF-convolved surface brightnesses, again assuming fractional errors of 10^{-3} in the latter. Adding all these measurements together, the final χ^2 of our model is

$$\chi^2_{\rm m}[\psi, f_k] = \chi^2_{\rm AO} + \chi_I^2 + \chi^2_{\rm G}, \quad (27)$$

which, for fixed potential ψ, depends quadratically on the orbit weights f_k. We use a standard non-negative least-squares algorithm (Lawson & Hanson 1974) to find the non-negative set of f_k that minimize it. Unlike most other orbit-superposition methods, we do not include the luminosity density $j(r)$ in this fit explicitly, but use the surface brightness profile $I(R)$ instead. Our reason for this is that, unlike $j(r)$, $I(R)$ is (in principle) directly measurable and can therefore be assigned meaningful error bars. We do not expect real galaxies to have perfectly constant mass-to-light ratios and, in the absence of anything better, use $j(r)$ merely to estimate the stellar contribution to the overall potential.

6.3.2 VP histograms

Unlike the Gauss–Hermite coefficients above, there is no simple way to compare the VP histograms found in Section 3.2 against models: in addition to the unavoidable correlations among the L_i, there are complicated biases introduced by the penalty function (15). So, we simply refit the VP histograms using the same n_v velocity bins for which we calculate the models' VPs in Section 6.2, reflecting the VPs about $v = 0$ in order to make them symmetric. The procedure is as follows:

(i) First find the best-fitting template fraction, continuum level and normalization. Find the f, c_i and smoothed L_i that minimize

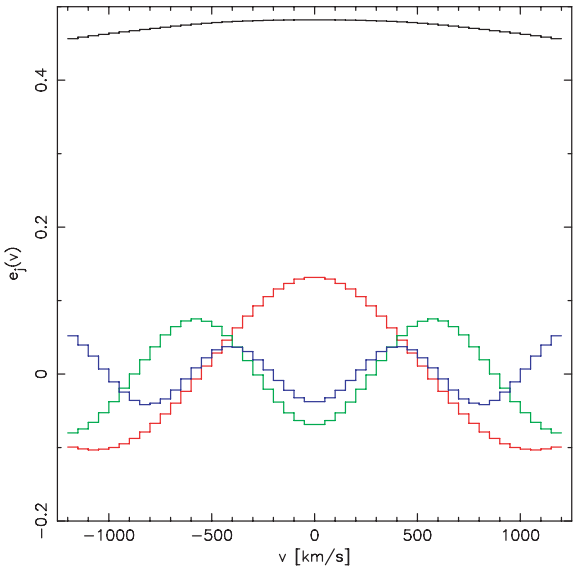

Figure 6. Plot of the first four eigen-VPs of one of our spectra obtained by diagonalizing χ_s of equation (7). Any (symmetric) VP can be expressed as a weighted sum of eigen-VPs (equation 29), in which case the errors in the weights are independent. We have divided each of the VPs plotted here by its corresponding eigen-error so that the scale of each gives a direct indication of the uncertainty in its weight.

the penalized $\chi^2_p = \chi^2_s + P[L_i]$ for $k = 1$, where χ^2_s and the penalty function P are given by equations (7) and (15) respectively. The best-fitting normalization factor is then $k^{-1} = \sum_i (v_{i+1} - v_i) L_i$.

(ii) Holding k, the c_i and f fixed, find the *formal* best-fitting VP histogram $(\hat{L}_1, \ldots, \hat{L}_{n_v})$ to the unpenalized χ^2_s (7).

The best-fitting VP $\hat{\boldsymbol{L}}$ is in general unphysical, with many $\hat{L}_i < 0$, but we use it only because it locates the minimum of the χ^2_s quadratic form. Let us write our formal best-fitting VP histogram as the column vector $\hat{\boldsymbol{L}} \equiv (\hat{L}_1, \ldots, \hat{L}_{n_v})^{\mathrm{T}}$ and consider another $\boldsymbol{L} \equiv (L_1, \ldots, L_{n_v})^{\mathrm{T}}$. Once k, f and the c_i have been fixed, equation (7) becomes

$$\chi^2_s(\boldsymbol{L}) = \chi^2_{\min} + \sum_{i=1}^{n_v} \left(\frac{\boldsymbol{e}_i \cdot \boldsymbol{L} - \boldsymbol{e}_i \cdot \hat{\boldsymbol{L}}}{\Delta_i} \right)^2, \tag{28}$$

where $\boldsymbol{e}_i$ are the eigenvectors of the Hessian $\partial^2 \chi^2_s / \partial L_i \, \partial L_j$ and the eigen-errors Δ_i are related to the eigenvalues λ_i through $\Delta_i \equiv \sqrt{2/\lambda_i}$. So by taking the projections

$$L'_i \equiv \boldsymbol{e}_j \cdot \boldsymbol{L}, \tag{29}$$

of $\boldsymbol{L}$ along the full set of 'eigen-VPs' $\boldsymbol{e}_j$, one tests directly how well it reproduces the observed galaxy spectrum. We note that our method is essentially a restatement of the work of De Rijcke & Dejonghe (1998) using the language of eigen-VPs. Fig. 6 plots the first few eigen-VPs of one of our spectra. Unlike the terms in a Gauss–Hermite expansion (9), they do not taper off rapidly at high velocities.

Summing the results of (28) for each VP and neglecting the $\chi^2_{\min}$ terms, a model with DF $(f_1, \ldots, f_n)$ has

$$\chi^2_H = \sum_{i=1}^{n_{AO}} \sum_{j=1}^{n_v} \left[\frac{\hat{L}'_{ij} - (1/\hat{I}_i) \sum_{k=1}^{n} f_k L'^{(k)}_{ij}}{\Delta_{ij}} \right]^2, \tag{30}$$

where $L'^{(k)}_{ij}$ is obtained from $L^{(k)}_{ij}$ through (29). Apart from replacing χ^2_{GH} in equation (27) by (30), our procedure for fitting models to VP histograms is identical to that for Gauss–Hermite coefficients.

6.4 Results

We have calculated the projection coefficients $L^{(k)}_{ij}(\psi)$ and $I^{(k)}_i(\psi)$ for $n_{\mathcal{E}} \times n_J = 200 \times 40$ DF components (equation 17) in a range of potentials ψ. Our main results below are obtained using coefficients calculated for BH masses $M_\bullet / 10^9\ \mathrm{M}_\odot = 0, 0.25, \ldots, 3.75$ with a single mass-to-light ratio, $\Upsilon_0 = 8.5 \Upsilon_\odot$. Since the projection coefficients scale straightforwardly with mass, we can use the method of Section 6.3 to fit models with other values of Υ provided we remember to scale $M_\bullet$ by Υ / Υ_0 and the bins v_j of the velocity histograms (19) by $\sqrt{\Upsilon / \Upsilon_0}$.

On Fig. 7 we plot the result of using (27) to fit the Gauss–Hermite coefficients $(h_0, h_2, \ldots, h_N)$ for the cases $N = 4$ and 6. In each case we have extracted VPs from spectra assuming that $h_i = 0$ for $i > N$. Marginalizing Υ, the former yields a moderately firm BH mass $M_\bullet \simeq 2.3^{+1.1}_{-0.9} \times 10^9\ \mathrm{M}_\odot$. Including h_6 in the fit, however, results only in the upper bound, $M_\bullet < 1.6 \times 10^9\ \mathrm{M}_\odot$, statistically consistent with the $N = 4$ result. To test whether this could be a result of the best-fitting template fraction f, continuum c_i or normalization k (Section 3) changing as we change N, we have repeated our $N = 6$ model fit omitting the coefficient h'_j (equation 24) with the largest eigen-error Δ_j. The resulting $\chi^2_m(M_\bullet, \Upsilon)$ contours are broadly similar to the $N = 4$ case, suggesting that the shift in the location of the minimum in χ^2_m is a genuine effect caused by the extra information contained in the h_6 parameter, and not by systematic changes in the fitted continuum or normalization.

These results are perhaps unsurprising when one recalls that our results in Section 4 indicate that NGC 1399 has very strongly non-Gaussian central VPs. Although it would be a mildly interesting exercise to test the effects of fitting to h_8 and even higher-order terms, we now simply drop the Gauss–Hermite parametrization and turn to fitting eigen-VPs. Fitting models to the full set of eigen-VPs yields the result plotted in the bottom right of Fig. 8. The best-fitting BH mass, $1.2^{+0.5}_{-0.6} \times 10^9\ \mathrm{M}_\odot$, is consistent with the results from the Gauss–Hermite fits, but has smaller error bars. The other panels on the figure show the effect of rearranging the eigen-VPs in order of increasing eigen-error and fitting only the first N for each spectrum. The case $N = 3$ yields results that look qualitatively similar to our sixth-order Gauss–Hermite fit, while adding one more eigen-VP introduces a useful lower bound on the BH mass. The results for $N = 5$ are very similar to fitting the full $N = 24$: in fact, it is impossible to distinguish between $N = 6$ and 24 by eye.

Since the the $N = 24$ model is essentially a direct fit to the galaxy spectrum, we adopt its best-fitting $M_\bullet = 1.2^{+0.5}_{-0.6} \times 10^9\ \mathrm{M}_\odot$ as our best estimate of the BH mass in NGC 1399. Fig. 9 plots the anisotropy parameter of one of the best-fitting models in this range. Our models have moderate radial anisotropy ($\beta \approx 0.3$) between 2 and 30 arcsec, similar to the results found by Saglia et al. (2000) in their models of NGC 1399. At larger radii the orbit distribution in our models becomes tangentially biased, which is what one expects from fitting a constant mass-to-light ratio model to a galaxy with a massive dark halo. Much more surprisingly, however, our models become extremely tangentially biased in the innermost arcsecond, which is where we find the interesting kinematic and photometric features.

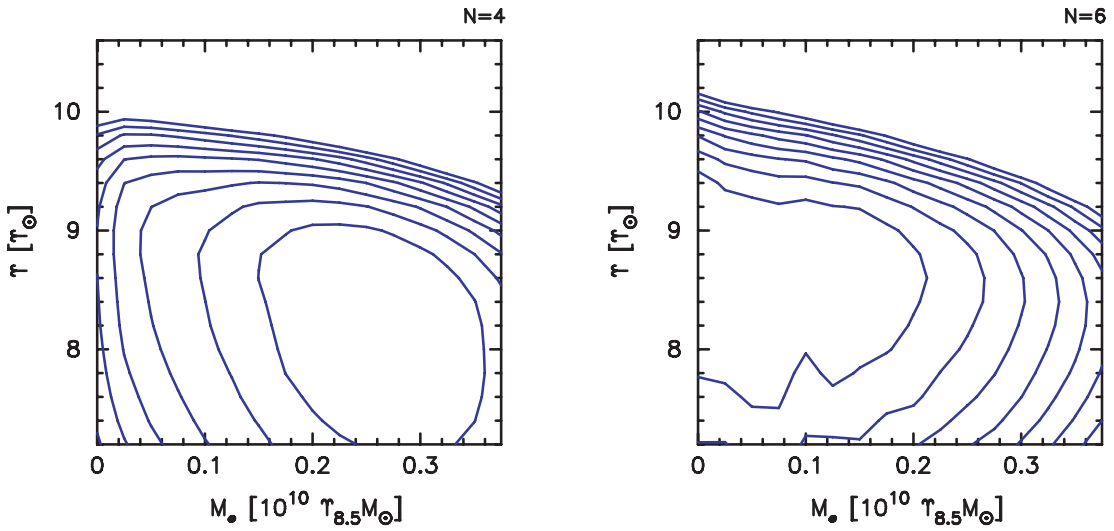

Figure 7. The $\chi^2_{\rm m}$ ($M_\bullet$, Υ) contours (27) obtained by fitting dynamical models to Gauss–Hermite parametrizations (Section 3.1) of our spectra and to the velocity dispersion profile measured by Graham et al. (1998). Successive contour levels have $\Delta\chi^2_{\rm m} = 1$. The models have 200 × 40 DF components (17) and fit $(h_0, h_2, \ldots, h_N)$ extracted from spectra under the assumption that $h_i = 0$ for $i > N$. The left panel shows results for $N = 4$, the right for $N = 6$. The x-axis is both cases is the BH mass scaled by $\Upsilon_{8.5} \equiv \Upsilon/8.5\Upsilon_\odot$. Mass-to-light ratios Υ are V band.

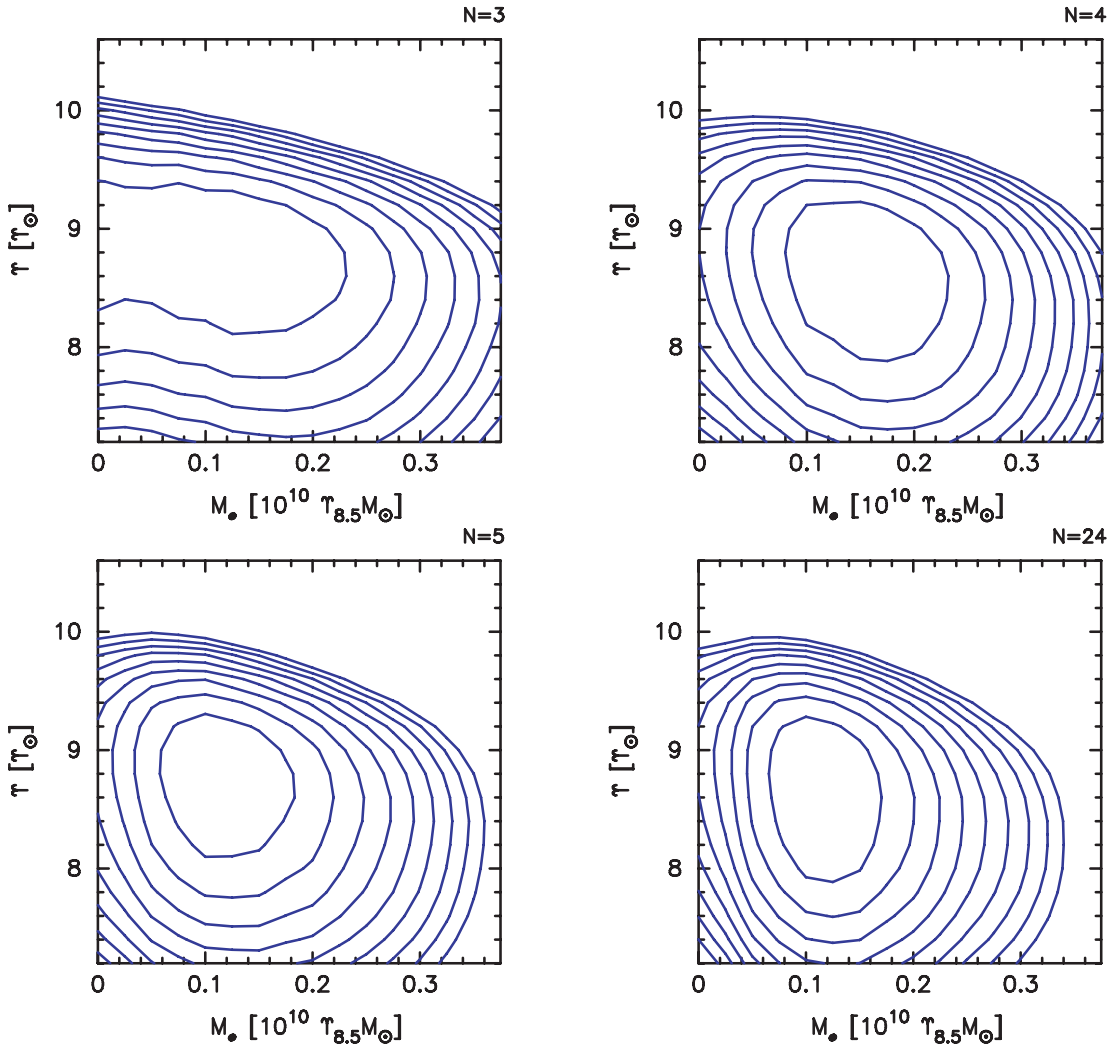

Figure 8. As for Fig. 7, but fitting to eigen-VPs (Section 6.3.2) instead of Gauss–Hermite coefficients. The first three panels show the results of fitting only to the first 3, 4 and 5 eigen-VPs of each spectrum. The last shows the results of fitting to the full spectrum by using all 24 eigen-VPs.

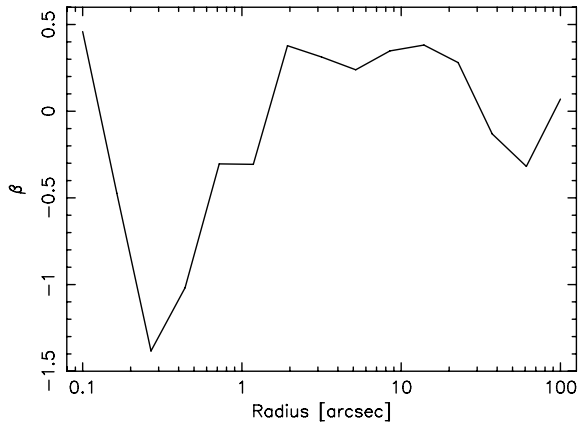

Figure 9. Anisotropy parameter $\beta \equiv 1 - \sigma_\theta^2/\sigma_r^2$ of a typical model ($M_\bullet = 10^9\,\mathrm{M}_\odot$, $\Upsilon = 8.5\Upsilon_\odot$), averaged over five shells per decade in radius.

6.5 Tests

Our best-fitting BH mass remains unchanged if we change the number of DF components we use in the models: models using 100 × 20 components instead of 200 × 40 yield the same result. We find no evidence for the flat-bottomed χ^2_m profiles claimed by Valluri, Merritt & Emsellem (2004) (albeit for the axisymmetric case), even though our models sample ($\mathcal{E}$, J^2) phase space a factor ∼20 more densely than theirs. We have tried varying the extent and widths of the velocity bins we use when we fit VPs, but the BH mass is unchanged whether we take 24 fine $25\,\mathrm{km\,s^{-1}}$ wide bins extending to $1200\,\mathrm{km\,s^{-1}}$ or 20 coarse $100\,\mathrm{km\,s^{-1}}$ wide bins extending to $2000\,\mathrm{km\,s^{-1}}$.

The assumed width of the velocity bins does affect the models in a slightly more subtle way, however. When we rescale our base $\Upsilon_0 = 8.5\Upsilon_\odot$ models to a new mass-to-light ratio Υ, we scale their velocity bins by an amount $\sqrt{\Upsilon/\Upsilon_0}$ and re-extract VPs from spectra using the the new bin widths. Therefore, values of $\chi^2_\mathrm{s}(\boldsymbol{L})$ from (28) for different Υ are not strictly directly comparable. To test the effects of this, we have calculated coarser 100 × 20 component models using 24 bins of fixed width $50\,\mathrm{km\,s^{-1}}$ on a grid of potentials with $\Upsilon/\Upsilon_\odot$ running from 7.2 to 10.6 in 0.2 steps and $M_\bullet/10^9\,\mathrm{M}_\odot$ running from 0 to 3.75 in 0.25 steps. The only notable difference between the resulting χ^2_m ($M_\bullet$, Υ) and the plots in Fig. 8 is that the range of acceptable Υ increases to $(9 \pm 1)\Upsilon_\odot$. The BH mass is unaffected.

6.6 Caveats on BH mass

Despite these reassurances, there nevertheless are some shortcomings of the models and data upon which this BH mass is based:

(i) We follow the usual 'extended Schwarzschild' procedure from which the majority of existing stellar-dynamical BH masses have been obtained and consider only the very best orbit distribution f_k for each potential. This best-fitting distribution is typically very spiky, with only ∼140 non-zero f_k out of 200 × 40! One way around this would be to apply some kind of regularization to the f_k (Thomas et al. 2005), but the biases introduced by this procedure are not well understood (Valluri et al. 2004).

(ii) Because of the huge freedom in fitting the f_k, the best-fitting model should have a very low χ^2_m. For example, in Monte Carlo experiments with synthetic data sets of toy galaxy models we typically find $\chi^2_\mathrm{m} \sim N_\mathrm{data}/3$, where N_data is the number of data points fitted in the models. For our real NGC 1399 data, however, our best-fitting χ^2_GH and χ^2_H are usually at least as big as the number of parameters we use to describe our kinematics. This is probably due to our neglect of the systematic errors (Sections 2.3 and 4.2). In contrast, the fit to the full surface brightness profile and to the outer dispersion profile is astonishingly good, with $\chi^2_I + \chi^2_\mathrm{G} \sim 3$, making our total χ^2_m relatively low.

(iii) Finally, our models assume that the galaxy is spherical and non-rotating, despite the clear evidence to the contrary. Nevertheless, the models do seem to require a strong bias towards circular orbits in the central 0.5 arcsec.

7 CONCLUSIONS

Using NAOS-CONICA at the VLT, we have successfully measured the central kinematics within the SoI of NGC 1399 ($r \sim 0.34$ arcsec) with a resolution (FWHM) of ∼0.15 arcsec (14 pc, Section 2.5) using adaptive optics correction on a bright reference star 17.6 arcsec away.

Alone, the kinematics extracted from the Ca I feature at 2.26 μm and the CO bands after 2.3 μm establish the presence of velocity gradient within a radius of ∼0.5 arcsec (48 pc), suggestive of a kinematically decoupled core.

Ks band imaging reveals offset and elongated isophotes within a radius of 0.2 arcsec (19 pc) that are not visible in *H*-band *HST* images. Such inner structure is reminiscent of that seen in other core ellipticals. The non-parametric VPs corresponding to this region also show an unusual velocity structure that may be consistent with the presence of an eccentric disc around the BH, akin to that of M31.

We have demonstrated that errors in the Gauss–Hermite coefficients h_j extracted from real galaxy spectra are *not* independent, and have shown a simple way of taking the covariances among the h_j into account when fitting dynamical models. The VPs near the nucleus of NGC 1399 are strongly non-Gaussian, however, and are not well described by a low-order Gauss–Hermite expansion. We show that the 'eigen-VPs' obtained by diagonalizing (7) are a more useful way of describing VPs, at least for numerical purposes.

Subject to the caveats of Section 6.6, our best estimate for the mass of the BH in NGC 1399 is $1.2^{+0.5}_{-0.6} \times 10^9\,\mathrm{M}_\odot$, obtained by fitting spherical dynamical models directly to our observed spectra. The models are based on the usual extension of Schwarzschild's method to the problem of potential estimation. Taken at face value, they place this galaxy on the $M_\bullet$–σ plane mid-way between the predictions of T02 and FF05, being consistent with both. The best-fitting model also becomes extremely tangentially anisotropic in the innermost 0.5 arcsec.

We have demonstrated that AO observations are a viable alternative to *HST* when measuring black hole masses and can break the mass–anisotropy degeneracy even in the most massive, non-rotating elliptical galaxies. The difficulty in interpreting these long-slit data emphasizes the need for high-S/N AO-assisted integral-field observations to further understand the kinematic and photometric features discovered here.

ACKNOWLEDGMENTS

We thank the referee, Laura Ferrarese, for her constructive comments. This research is based on observations collected at the European Southern Observatory, Chile (ESO Programme 072.B-0763). We acknowledge use of the Simbad astronomical data base and the HyperLeda data base. The authors are funded by the Particle Physics and Astronomy Research Council (PPARC) and the Royal Society.

REFERENCES

Aller M. C., Richstone D., 2002, AJ, 124, 3035
Baird S. R., 1981, ApJ, 245, 208
Bender R., 1990, A&A, 229, 441
Bicknell G. V., Bruce T. E. G., Carter D., Killeen N. E. B., 1989, ApJ, 336, 639
Binney J., Mamon G. A., 1982, MNRAS, 200, 361
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cretton N., de Zeeuw P. T., van der Marel R. P., Rix H.-W., 1999, ApJS, 124, 383
De Rijcke S., Dejonghe H., 1998, MNRAS, 298, 677
de Vaucouleurs G., de Vaucouleurs A., Corwin H. G., Buta R. J., Paturel G., Fouque P., 1991, Third Reference Catalogue of Bright Galaxies, Vol. 1–3, XII. Springer, Berlin
de Zeeuw P. T., 2004, in Ho L. C., ed., Carnegie Observatories Astrophysics Ser. Vol. 1, Coevolution of Black Holes and Galaxies. Cambridge Univ. Press, Cambridge, p. 460
D'Onofrio M., Zaggia S. R., Longo G., Caon N., Capaccioli M., 1995, A&A, 296, 319
Efstathiou G., Ellis R. S., Carter D., 1980, MNRAS, 193, 931
Emsellem E. et al., 2004, MNRAS, 352, 721
Faber S. M. et al., 1997, AJ, 114, 1771
Ferrarese L., Ford H., 2005, Space Sci. Rev., 116, 523 (FF05)
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Franx M., Illingworth G., Heckman T., 1989, ApJ, 344, 613
Gebhardt K. et al., 2000, ApJ, 539, L13
Gebhardt K. et al., 2003, ApJ, 583, 92
Gerhard O. E., 1993, MNRAS, 265, 213
Graham A. W., Colless M. M., Busarello G., Zaggia S., Longo G., 1998, A&AS, 133, 325
Haehnelt M. G., Kauffmann G., 2000, MNRAS, 318, L35
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Killeen N. E. B., Bicknell G. V., 1988, ApJ, 325, 165
Kleinmann S. G., Hall D. N. B., 1986, ApJS, 62, 501
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Kuntschner H., 2000, MNRAS, 315, 184
Landsman W. B., 1993, in Hanisch R. J., Brissenden R. J. V., Barnes J., eds, ASP Conf. Ser. Vol. 52, Astronomical Data Analysis Software and Systems II. Astron. Soc. Pac., San Francisco, p. 246
Lauer T. R. et al., 1995, AJ, 110, 2622
Lauer T. R. et al., 2005, AJ, 129, 2138
Lawson C. L., Hanson R. J., 1974, Solving Least Squares Problems. Prentice-Hall, Englewood Cliffs, NJ
Lenzen R., Hofmann R., Bizenberger P., Tusche A., 1998, Proc. SPIE, 3354, 606
Longo G., Zaggia S., Busarello G., Richter G., 1994, VizieR Online Data Catalog, 410, 50433
Madore B. F. et al., 1999, ApJ, 515, 29
Magorrian J. et al., 1998, AJ, 115, 2285
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Oliva E., Origlia L., Kotilainen J. K., Moorwood A. F. M., 1995, A&A, 301, 55
Origlia L., Moorwood A. F. M., Oliva E., 1993, A&A, 280, 536
Pinkney J. et al., 2003, ApJ, 596, 903
Richstone D., Sargent W. L. W., 1972, ApJ, 176, 91
Rix H., White S. D. M., 1992, MNRAS, 254, 389
Rix H.-W., de Zeeuw P. T., Cretton N., van der Marel R. P., Carollo C. M., 1997, ApJ, 488, 702
Rousset G. et al., 1998, Proc. SPIE, 3353, 508
Saglia R. P., Kronawitter A., Gerhard O., Bender R., 2000, ApJ, 119, 153
Saha P., Williams T. B., 1994, AJ, 107, 1295
Sargent W. L. W., Young P. J., Lynds C. R., Boksenberg A., Shortridge K., Hartwick F. D. A., 1978, ApJ, 221, 731
Schödel R., Ott T., Genzel R., Eckart A., Mouawad N., Alexander T., 2003, ApJ, 596, 1015
Schombert J. M., 1986, ApJS, 60, 603
Simkin S. M., 1974, A&A, 31, 129
Smith R. J., Lucey J. R., Hudson M. J., Schlegel D. J., Davies R. L., 2000, MNRAS, 313, 469
Thatte N., Tecza M., Genzel R., 2000, A&A, 364, L47
Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Corsini E. M., Wegner G., 2005, MNRAS, 360, 1355
Tonry J. L., Dressler A., Blakeslee J. P., Ajhar E. A., Fletcher A. B., Luppino G. A., Metzger M. R., Moore C. B., 2001, ApJ, 546, 681
Tremaine S., 1995, AJ, 110, 628
Tremaine S. et al., 2002, ApJ, 574, 740 (T02)
Valluri M., Merritt D., Emsellem E., 2004, ApJ, 602, 66
van der Marel R. P., 1994, MNRAS, 270, 271
van der Marel R. P., Franx M., 1993, ApJ, 407, 525
van der Marel R. P., Rix H. W., Carter D., Franx M., White S. D. M., de Zeeuw T., 1994, MNRAS, 268, 521
Yu Q., Tremaine S., 2002, MNRAS, 335, 965
Zhao H., Haehnelt M. G., Rees M. J., 2002, New Astron., 7, 385

This paper has been typeset from a TEX/LATEX file prepared by the author.