Spanning the halo: II. The mass profile of M87

L. J. Oldham*, M. W. Auger
Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

August 19, 2015

ABSTRACT

We model the mass profile of the giant elliptical galaxy M87, using multiple tracer populations to constrain the mass across a large radius range. We combine stellar kinematics in the central regions with the dynamics of 612 globular clusters out to 200 kpc and satellite Virgo galaxies on scales comparable with the virial radius of the cluster. A spherically symmetric Jeans analysis allows us to infer the stellar mass-to-light ratio as $\Upsilon_{\star,V} = 7.53 \pm 0.04$, the black hole mass as $\log(M_{\mathrm{BH}}/M_\odot) = 9.83 \pm 0.02$ and to set constraints on the structure of the dark matter halo. The high mass-to-light ratio allows us to rule out a Chabrier IMF. We explore three different halo models and conclude that the dark matter becomes cored in the centre, contrary to the inference based on GCs alone. We find a total mass within 150 kpc of $\log(M(R < 150\,\text{kpc})/M_\odot) = 13.21 \pm 0.23$ and a virial mass $\log(M_{\mathrm{vir}}/M_\odot) = 14.27^{+0.72}_{-0.38}$. We interpret this in terms of the gradual build-up of the galaxy and its associated cluster over long timescales and point to the possible role of M87’s AGN and supermassive black hole in the formation of the central core.

Key words: galaxies: elliptical and lenticular, cD – galaxies: kinematics and dynamics – galaxies: haloes – galaxies: individual: M87 – galaxies: structure

1 INTRODUCTION

The Λ-CDM paradigm of structure formation has been very successful in describing the Universe on large scales, but there remains some tension between cosmological simulations and observation regarding galaxy structure. For instance, one main prediction of cold, collisionless gravitational collapse is the formation of a central cusp in the density profile of the dark matter (DM) halos that envelope galaxies, with $\rho_{\mathrm{DM}} \sim r^{-\gamma}$ and $\gamma = 1$ at small radii (Navarro et al. 1997). However, this stands in contradiction to observational studies of a range of galaxy types, in which the halos are found to favour cored or only weakly cuspy central profiles. For instance, the recent local surveys THINGS and LITTLE THINGS (Hunter et al. 2007) found a large fraction of dwarf field galaxies to have DM density profiles that go as $\rho \propto r^{-0.4}$ within the central kiloparsec, and studies of low-surface-brightness galaxies also point to relatively flat central profiles with a large scatter (e.g. de Naray & Spekkens 2011). One main difference between DM-only cosmological simulations and real halos is that the latter also contain baryons, and it is possible that the imprint of baryonic physics on the halo is the cause of this discrepancy. For instance, feedback from supernovae and AGN as well as dynamical friction from infalling satellites could lead to some degree of heating and expansion (e.g. Mashchenko et al. 2006; Laporte et al. 2012; Governato et al. 2012; Velliscig et al. 2014). However, many questions remain about these processes and their effect on galaxy structure: understanding the systems we see, then, through sophisticated modelling, is a good place to start.

While it is comparatively easier to test these predictions in low-surface-brightness galaxies and dwarf spheroids, where DM dominates over the baryonic mass, the situation is much more complicated for their massive elliptical counterparts. Here, our ignorance about the stellar initial mass function (IMF) introduces a degeneracy between dark and luminous matter, which makes it hard to constrain the behaviour of the DM in the inner regions, while at large radii, the task of probing the gravitational potential is made hard by the fact that the outskirts of these galaxies are notoriously faint. However, one way of significantly alleviating these degeneracies is to use multiple dynamical tracer populations, spanning a range of galactocentric radii (e.g. Schuberth et al. 2010; Walker & Peñarrubia 2011; Napolitano et al. 2014). Massive elliptical galaxies are often home to large populations of planetary nebulae (PNe), globular clusters (GCs) and even, in the case of brightest cluster galaxies (BCGs), satellite galaxies. Each population, with its own signature spatial distribution and kinematic profile, can then be used as an independent probe of the gravitational potential. Though the pool of early-type galaxies (ETGs) for

* E-mail: ljo31@ast.cam.ac.uk

© 2015 RAS
which such an analysis has been carried out is still relatively small, the evidence so far adds to a picture in which the NFW profile does not provide a good universal fit: Schuberth et al. (2012), for instance, explored a range of halo models for the elliptical NGC 4636 and ultimately favoured a cored logarithmic profile over an NFW, while the multiple-population dynamical study of Pota et al. (2015) deduced a slope of $\gamma = -0.6$ for the elliptical NGC 1407, though they were not able to exclude an NFW slope given the uncertainties. In a similar way, Sand et al. (2004), Newman et al. (2011) and Newman et al. (2013) have combined strong and weak lensing and resolved stellar kinematics to find evidence for a flattening of the central density profile in clusters.

The massive elliptical M87, located at the centre of the Virgo cluster, is an ideal subject for continuing such studies: its estimated GC population, numbering $\sim 12,000$ (McLaughlin et al. 1994; Tamura et al. 2006), makes it one of the richest GC hosts in the local universe, and the sample of this population for which we have high-resolution kinematic data has been greatly expanded in recent years through the wide-field study of Strader et al. (2011). It is also understood to be a slow rotator and nearly spherical (e.g. Sarzi et al. 2013), indicating that it is suitable for mass modelling under the assumption of spherical symmetry. However, two recent studies of M87’s mass distribution, both primarily based on these data, are markedly inconsistent. The first of these, Agnello et al. (2014), divided the GC sample into three independent populations and used a virial analysis to infer a very cuspy density profile ($\gamma = 1.57$), while Zhu et al. (2014) combined SAURON central stellar kinematics with the Strader GC data and an additional GC sample from Hectospec, and modelled the density profile using a logarithmic potential, thus imposing a core. Their total inferred stellar masses differed by almost a factor of two. As M87 is one of the nearest and most well-observed BCGs, it seems unsatisfactory that its mass distribution should still be so poorly understood, and clearly there remains work to be done. The aim of this paper is therefore to use a synthesis of GC, satellite and stellar kinematic data in conjunction with flexible mass models in order to infer a density profile which is free to be cuspy or cored as the dynamics dictate.

The paper is organised as follows: in Section 2, we introduce the tracer populations used in our analysis and the associated datasets. In Section 3 we describe our mass model and Jeans analysis, the results of which are presented in Section 4. We discuss the implications of our findings in Section 5, and use Section 6 to summarise our main conclusions. Throughout this work, we assume a distance to M87 $D_L = 16.5$ Mpc.

## 2 DATA

To constrain M87’s density profile across a wide radius range, we use multiple dynamical tracers, combining stellar kinematics in the central regions with GC dynamics at large radii and satellite galaxies on cluster scales. We can then solve the Jeans equation for each population separately, provided that the underlying density distribution is known. We therefore take data from a number of sources, as summarised in Table 1.

### 2.1 Stars

The deprojected stellar surface brightness profile comes into our analysis at two points: firstly, the use of the stars as dynamical tracers in the Jeans equation requires us to know their 3D density distribution; second, our goal is to model M87’s mass as the sum of dark and luminous components, and the latter is simply the product of the integrated 3D luminosity density with some constant mass-to-light ratio, $\Upsilon_L$, which can be inferred from the data. We use the radial profile for M87 presented in Kormendy et al. (2009), in which 20 sets of observations across a range of radii and filter systems were synthesised into a single profile in the V band. As M87 is known to have a very extended cD envelope in addition to a stellar core (e.g. Chakrabarty 2007), its surface brightness profile cannot be accurately modelled by a Sérsic profile at both small and large radii. We therefore chose to model it using a more flexible Nuker profile, according to the following relation:

$$I(R) = I_0 \left( \frac{r}{r_b} \right)^{-\gamma} \left( 1 + \frac{r}{r_b} \right)^{-\alpha} \frac{\alpha-\gamma}{\alpha}$$

with amplitude $I_0$, break radius $r_b = 1.05$ kpc, inner slope $\gamma = 0.188$, outer slope $\beta = 1.88$ and break sharpness $\alpha = 1.27$. While this has the disadvantage of having no analytic deprojection of normalisation, it is much more flexible than a Sérsic profile as it allows both the inner and outer slopes greater freedom. Assuming spherical symmetry, we deproject this profile to give the 3D density shown in Figure 1.

The kinematics of the inner $33'' \times 41''$ of M87 have been observed with the IFU SAURON, and a catalogue of the first four moments of the Gauss-Hermite expansion of the line-of-sight velocity distribution is available online1. We use the velocity dispersions, which were obtained from the spectra using a direct pixel fitting routine (e.g. van der Marel 1994). The spectra were adaptively binned to ensure a signal-to-noise of at least 60 per spectral resolution element; uncertainties are generally less than $\sim 20$ kms$^{-1}$, with a mean uncertainty of $\sim 9$ kms$^{-1}$.

M87 is known to host a supermassive black hole (SMBH) of mass $\sim 6.6 \times 10^8 M_\odot$ (Gebhardt & Thomas 2009; Gebhardt et al. 2011), and this should make a significant contribution to the stellar velocity dispersions at the smallest radii. The SAURON dataset extends right down to the centre, though its resolution of $1''$ is too low to be able to set constraints on the SMBH mass. One option to deal with this would be to simply exclude the apertures within the central $\sim 3''$ from our analysis. However, an alternative is to include the black hole (BH) in our mass model and constrain it using high-resolution kinematics of the central 2'', as observed with the IFU NIFS on the Gemini Telescope (Gebhardt et al. 2011). As explained in that paper, these data were obtained from spectra which used laser adaptive optics corrections, and have a resolution of 0.08'' and a signal-to-noise generally greater than 50. The velocity moments are provided in radius and position angle bins, though our simplifying assumption of spherical symmetry allows us to combine bins azimuthally.

---

1 http://www.strw.leidenuniv.nl/sauron/
Spanning the halo: II. The mass profile of M87

<table>
<thead>
<tr>
<th>radial coverage</th>
<th>data type</th>
<th>instruments</th>
<th>sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>stars</td>
<td>0 - 30''</td>
<td>kinematics</td>
<td>SAURON/William Herschel Telescope Emsellem et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>0 - 2''</td>
<td>kinematics</td>
<td>NIFS/Gemini Telescope Gebhardt et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>0.02 - 2500''</td>
<td>photometry</td>
<td>multiple Kormendy et al. (2009)</td>
</tr>
<tr>
<td>GCs</td>
<td>1.3 - 240 kpc</td>
<td>photometry</td>
<td>MegaPrime/CFHT Oldham &amp; Auger (2015, in prep.)</td>
</tr>
<tr>
<td></td>
<td>2 - 200 kpc</td>
<td>kinematics</td>
<td>multiple (mostly SDSS) Strader et al. (2011)</td>
</tr>
<tr>
<td>satellite galaxies</td>
<td>35 - 1000 kpc</td>
<td>kinematics</td>
<td>multiple Kim et al. (2014); Blakeslee et al. (2009)</td>
</tr>
</tbody>
</table>

Table 1. Various sources of photometric and spectroscopic data for the different tracer populations used in the dynamical analysis

2.2 Globular clusters

We use the colour and radial profiles inferred in Oldham & Auger (2015, in prep.) for the GC populations, as shown in Figure 1. The details of the inference are explained fully in that paper, but in brief, archival CFHT/MegaPrime images in the $ugriz$ bands were used to compile a sample of 17620 GC candidates, selected according to their colours, magnitudes and apparent sizes, and the resulting catalogue was modelled to infer the radial and colour distributions and the relative fractions of the two GC populations. The surface density for each GC component was modelled as a Séracse profile

$$N(R) = N_0 \exp \left[ -k_n \left( \frac{R}{R_e} \right)^{1/n} \right]$$

with $k_n = 2n - 0.324$, and with the profiles normalised such that $\int_{R_{max}}^{R_{max}} N(R)dR = 1$, where $R_{max}$ is the radius of the outermost GC in the catalogue. The $g - r$, $g - i$ colours and the GC luminosity function for each GC component were modelled as Gaussians. The 3D deprojected profiles are also shown in Figure 1.

Strader et al. (2011) presents a spectroscopic catalogue of 737 GC candidates around M87 at radii from 2 kpc to 200 kpc (plus one object at 800 kpc, which we exclude because it lies outside the region over which our model from the photometry is strictly valid). The catalogue combines new measurements for 451 GCs – obtained using Keck/DEIMOS, Keck/LRIS and MMT/Hectospec – with literature data, and provides a ‘classification’ of objects as blue GCs, red GCs, UCDs and transient/unknown along with SDSS-band $g - r$ and $g - i$ colours. We cross-correlate the photometric and spectroscopic catalogues, selecting only the objects classified in Strader et al. (2011) as GCs, to obtain a sample of 612 GCs with complete luminosity, spatial, colour and kinematic information. In the analysis that follows, we choose to use the Oldham & Auger (2015, in prep.) photometry over that provided in Strader et al. (2011), for consistency with out GC colour distributions.

2.3 Satellite galaxies

The Extended Virgo Cluster Catalogue (EVCC) provides the redshifts and positions on the sky of 1589 galaxies in a footprint of 725 deg$^2$ centred on M87, extending to 3.5 times the virial cluster radius. The redshifts are compiled from the SDSS DR7 release and the NASA Extragalactic Database (NED), and each object is classified as either a certain cluster member, a possible member or a background.
source based on morphological and spectroscopic criteria. As it is important that our sample only contains galaxies moving in M87’s halo potential, we selected only those objects classed as certain members according to both criteria, and further cross-correlated these with the catalogue of Blakeslee et al. (2009), which used surface brightness fluctuations to calculate distance moduli. To avoid contamination from the W cloud, a slightly more distant component of the Virgo cluster at a characteristic distance of 23 Mpc, we imposed a distance cut of 20 Mpc; further, to separate the A cloud (centred on M87) from the B cloud (centred on M49) we also imposed a declination angle cut of 9.5 degrees and a radius cut of 1 Mpc. The spatial and velocity distributions of the resulting sample are shown in Figure 2: it comprises 60 galaxies, with radii relative to M87 ranging from 35 kpc to 1 Mpc.

We could use this tracer population in the same way as the stars and the GCs and require its velocity dispersion profile to satisfy the Jeans equation so as to obtain a further probe of the mass at scales comparable to the virial radius. However, as noted earlier, the Jeans equation requires us to know the tracer density distribution and anisotropy and, as our satellite sample is most likely highly incomplete, we do not have access to these quantities. Most importantly, the sample of satellites we use has been selected spectroscopically, and this imposes a non-trivial selection function which may alter the spatial distribution from the true underlying one, which would lead us to draw incorrect conclusions from a Jeans analysis. Instead, we choose to use these satellites to give an estimate of the total mass at large radii and so give a further constraint in our inference on the mass profile. We do this using the virial mass estimator developed in Watkins et al. (2010) and applied to a galaxy group in Deason et al. (2013), which is designed to be robust against simple approximations to the true distributions. We use the mass estimator given by

\[
M(R < R_{\text{out}}) = \frac{C}{G} \left( \frac{v_{\text{los}}^2}{r} \right)^\mu,
\]

where

\[
C = \frac{\mu + \nu - 2\beta}{I_{\mu,\nu}}
\]

and

\[
I_{\mu,\nu} = \frac{\pi^{\frac{5}{2}}}{4} \left( \frac{\mu + 1}{\nu + 1} \right) \left( \mu + 3 - \beta (\mu + 2) \right).
\]

Here \( \beta \) is the anisotropy parameter

\[
\beta = 1 - \frac{\sigma^2}{\sigma^2_r}
\]

for radial and tangential velocity dispersions \( \sigma^2 \) and \( \sigma^2_r \), and \( \mu \) and \( \nu \) are the slopes of the potential and tracer density respectively, both assumed to be scale-free. As mentioned previously, our satellite sample is likely to be incomplete, and this means we cannot infer \( \mu, \beta \) and \( \nu \) directly from the data; instead we calibrate their values using simulations. We use the \( z = 0 \) halo catalogue of the first MultiDark simulation, described in detail in Prada et al. (2012). This uses the WMAP5 cosmology and contains about 8.6 billions particles per Gpc/h^3; the halo finder used is the bound density maximum technique described in Klypin & Holtzman (1997). We identify all halos with more than 30 subhalos, and follow the principle of abundance matching in treating each subhalo as a distinct satellite galaxy. We then use the subhalo veloci-
Spanning the halo: II. The mass profile of M87

3 MASS MODEL

We model the galaxy density profile as the sum of luminous and DM components plus a BH:

$$\rho(r) = \rho_*(r) + \rho_{DM}(r) + \rho_{BH}(r).$$

The stellar density profile is obtained by deprojecting the stellar surface brightness and scaling by the stellar mass-to-light ratio $\Upsilon_*$, which we assume to be constant. To allow the DM density profile as much flexibility as possible and to explore the impact of changing the halo model on the mass inference, we carry out our analysis using three different halo profiles. The first is a generalised NFW (gNFW),

$$\rho_{DM}(r) = \frac{\rho_0}{4\pi} \left( \frac{r}{r_0} \right)^{-\gamma} \left( 1 + \frac{r}{r_0} \right)^{-\gamma-3},$$

where $\gamma$ is the inner slope, $\rho_0$ is a normalisation factor and $r_0$ is the break radius, governing the scale at which the slope transitions from going as $\rho \sim r^{-\gamma}$ to $\rho \sim r^{-3}$ as in the standard NFW scenario. This has the advantage of leaving the profile free to choose between cusps ($\gamma \geq 1$) and cores ($\gamma = 0$) in the centre, while still becoming NFW-like at large radii in a way that is consistent with both simulations and observations. This gives a cumulative mass profile related to the Gauss hypergeometric function $2F1$ as

$$M_{DM}(r) = \frac{\rho_0 r_0^2}{\omega} \left( \frac{r}{r_0} \right) \frac{\omega}{2F1[\omega, \omega, 1 + \omega; -\frac{r}{r_0}]}$$

where $\omega = 3 - \gamma$.

The second profile we use is a cored generalised NFW (cgNFW):

$$\rho_{DM}(r) = \rho_0 \left( 1 + \frac{r}{r_a} \right)^{-\gamma} \left( 1 + \frac{r}{r_0} \right)^{-\gamma-3},$$

in which $r_a$ is the break radius of the core, $\gamma$ is the intermediate profile slope, $r_0$ the outer break radius and $\rho_0$ a normalisation factor. As with the gNFW, this profile allows the data to choose between cusps ($r_a = 0, \gamma = 1$) and cores ($r_a > 0$, or $r_a = 0$ and $\gamma = 0$) but it additionally allows more flexibility at intermediate radii, though this comes at the cost of an extra parameter. We carry out the integration of the cumulative mass for this profile numerically.

Following Zhu et al. (2014), we also use a cored logarithmic potential (LOG) model:

$$\rho_{DM}(r) = \rho_0 \left( 1 + \frac{r}{r_0} \right)^{-\gamma} \left( 1 + \frac{r}{r_0} \right)^{-\gamma-3} \left( 1 + \frac{r}{r_a} \right)^{-\gamma},$$

with scale radius $r_0$ and density scale $\rho_0$. This model is inherently cored, but in contrast to the others has a $\rho \sim r^{-2}$ dependence at large radii: for a given core, then, this profile allows more dark matter to be placed at large radii. Its cumulative mass is given by

$$M(< r) = \frac{\rho_0 r_0^3}{1 + \frac{r}{r_0}}.$$  \hspace{1cm} (13)

The BH is simply a point mass at the origin,

$$\rho_{BH} = M_{BH} \delta(r).$$ \hspace{1cm} (14)

giving a constant term in the cumulative mass distribution

$$M_{BH}(r) = M_{BH}.$$ \hspace{1cm} (15)

As the SAURON dataset far outweighs the GC, NIFS and satellite datasets in terms of size, we regularise its contribution to the likelihood calculation by additionally fitting for a ‘noise’ parameter $\Delta_n$. This can be interpreted as accounting for scatter in the data not included in the uncertainties or, alternatively, as modifying the relative weight.
given to the SAURON data, and is added in quadrature to the measured uncertainties on the velocity dispersion.

Our overall model therefore has a number of free parameters dependent on the halo model in question, but which varies between five and seven. In common in all cases are four free parameters: the mass-to-light ratio $T_\star$, the normalisation of the DM halo log($\rho_0$), the BH mass log($M_{BH}$) and the noise in the SAURON data, $\Delta m$. The gNFW, cgNFW and LOG halo models then add the free parameters ($r_0$, $\gamma$), ($r_0$, $\gamma$, $r_a$) and ($r_0$) respectively. In our notation in the following sections, we use the gNFW parameter set whenever we write out the model parameters explicitly, but this should be understood as standing in for any of the halo models.

### 3.1 Jeans analysis

Given the stellar velocity dispersion and the GC and satellite velocities, we want to infer the posterior probability distribution on M87’s density profile. For the satellite galaxies, this involves a direct comparison of the mass calculated from our virial estimator and that obtained by integrating the Jeans equation. Assuming spherical symmetry and dynamical equilibrium, the Jeans equation has the simple form

$$\frac{d}{dr} (l(r)^2 + \frac{2}{r} \beta(r) l(r)^2) = l(r) \frac{GM(r)}{r^2}$$

(16)

where $l(r)$ is the luminosity density of the tracer, $\sigma_r(r)$ the radial velocity dispersion and $\beta(r)$ the anisotropy parameter defined in Equation 6. Assuming isotropy, such that $\beta = 0$, the solution simplifies to

$$l(r) \sigma_r^2 = \int_{r}^{\infty} M(s) \frac{ds}{s^2}$$

(17)

which can be projected along the line of sight to give $\sigma_{los}^2(R)$ as

$$\sigma_{los}(R)^2 = \frac{2G}{l(R)} \int_{R}^{\infty} \frac{\sqrt{r^2 - R^2}}{r^2} l(r)M(r)dr$$

(18)

as given in Mamon & Lokas (2005). A comprehensive derivation of many relations for a range of anisotropy parameterisations is provided in the Appendix of that paper, and we refer the reader there for further details. As explained in Section 2, we use a deprojected Nuker profile for the stellar luminosity density, and deprojected Sérsic profiles for the GCs. For the latter, we follow the prescription of Prugniel & Simien (1997) to perform the deprojection analytically: the Nuker profile, on the other hand, must be deprojected numerically.

We can then use the methods of Bayesian analysis to infer the posterior probability distribution of our density profile parameters, given the data. Bayes theorem states that the posterior distribution is proportional to the product of the likelihood function of the data given the model and the priors on the model, and so the task here is to construct sensible likelihood functions for each dataset. The total likelihood is then, in turn, the product of these, as each constitutes an independent set of measurements.

First, as we have velocity dispersion measurements for the stars, the likelihood of observing a particular velocity dispersion at radius $R$ is assumed to be Gaussian, with a standard deviation equal to the uncertainty. Thus the $k^{th}$ stellar velocity dispersion measurement gives a contribution to the likelihood:

$$\ln L_{*k}(\sigma_k, R_k|\bar{M}) = -0.5 \left( \frac{\sigma_k^2 - \sigma_{los}^2}{\delta \sigma_k^2} \right)^2 - 0.5 \ln(2\pi \delta \sigma_k^2)$$

(19)

for uncertainty $\delta \sigma_k$ and model prediction $\sigma_{los}^2$, and model parameters $\bar{M} = (T_\star, M_{BH}, \Delta m, \rho_0, \gamma, r_a)$. As the observations in each aperture are independent, the total log likelihood of observing the ensemble is just the sum:

$$\ln L_* = \sum_k \ln L_{*k} = -0.5 \sum_k \left( \left( \frac{\sigma_k^2 - \sigma_{los}^2}{\delta \sigma_k^2} \right)^2 + \ln \left( 2\pi \delta \sigma_k^2 \right) \right)$$

(20)

Note that, for measurement uncertainty $\Delta \sigma_k$, the regularisation of the SAURON data gives a total uncertainty $\delta \sigma_k^2 = \Delta \sigma_k^2 + \sigma_{los}^2$. For the NIFS data, on the other hand, $\delta \sigma_{los}^2 = \sigma_{los}^2$.

The virial mass estimate from the satellites can be treated in a similar way, though here we only have one measurement:

$$\ln L_{sat} = -0.5 \left( \frac{M_{sat} - M_{mod}}{\delta M_{sat}} \right)^2 - 0.5 \ln \left( 2\pi \delta M_{sat}^2 \right)$$

(21)

for viral mass estimate $M_{sat}$, model mass $M_{mod}$ and the uncertainty from the mass estimator, $\delta M_{sat}$. This is a direct comparison without the need for recourse to the Jeans equation.

For the GCs, we have line-of-sight velocities rather than velocity dispersions, and as our sample size is small with respect to its radial coverage, an attempt to bin the data radially and construct a velocity dispersion profile would be very noisy, and weak from a statistical point of view. Further, while we can assign probabilities, based on the colour and position information provided in the photometric catalogue, of each GC belonging to either the red or the blue population, we are not able to classify them with certainty. As the two populations are assumed to be dynamically decoupled, we do not expect their velocity dispersion profiles to be the same. In contrast to the other tracer populations, then, we calculate the likelihood of observing a GC with a particular velocity under the assumption that the GCs’ velocities of each GC population can be described by a Gaussian with a standard deviation given by the velocity dispersion, such that

$$\ln L_{GC,k} = -0.5 \left( \frac{(v_k - v_{gal})^2}{\delta v^2 + \sigma_{m}^2} \right)^2 - 0.5 \ln \left( 2\pi (\delta v^2 + \sigma_{m}^2) \right)$$

(22)

where the velocity dispersion $\sigma_m$ is modelled separately for the red and the blue populations, based on their distinct luminosity densities, and is a function of projected radius.
$R$, and $v_{\text{helio}} = 1284 \text{ kms}^{-1}$ is the heliocentric velocity of M87 (Cappellari et. al. 2011).

To account for the uncertainty in assigning each GC to either the red or the blue population, for each set of model parameters $M_k$ we draw 1000 Monte Carlo samples which stochastically explore the population distribution based on the colour, magnitude and position information for the individual GCs. We then marginalise over the samples to give a final contribution to the likelihood. While some GCs have either very high or very low probabilities of belonging to one of the GC populations, with small uncertainty, there also exists a significant fraction with comparable probabilities of belonging to either: for these objects, it would not be meaningful to simply assign them to one population or the other. (This stochastic sampling also allows us to explore different combinations of red and blue GCs within our sample.)

As each tracer population is independent, the final log-likelihood of any set of model parameters is the sum over all contributions:

$$\ln L = \sum \ln L_* + \sum \ln L_{\text{GC}} + \ln L_{\text{M sat}}.$$  

We explore the parameter space using the ensemble-sampling code emcee (Foreman-Mackey et al. 2013). We report the maximum-likelihood values of our inferred posterior distributions in Table 2 and present the 1D and 2D posterior distributions for each halo model in Figures 4, 5 and 6.

4 RESULTS

Our inference on the model parameters for the three halo models are shown in Figures 4, 5 and 6, and the maximum-likelihood parameters are presented in Table 2. All three models have reduced $\chi^2 < 1.16$ and are sufficiently flexible to reproduce the data convincingly – we therefore cannot discount any of them and do not choose a ‘best’ parameterisation of the profile. We compare the inferred circular velocity curves for each model in Figure 7. In what follows, all quantitative results are given with respect to the best gNFW model.

Our best-fitting (gNFW) mass profile is shown in Figure 8, with the measurements from a number of previous studies overplotted. We find that our model is generally consistent with earlier work, though it is perhaps slightly larger at intermediate radii than found previously. The residuals on the SAURON and NIFS data are shown in Figure 9, and the DM fraction is shown as a function of radius in Figure 7. We infer a black hole mass of $\log(M_{\text{BH}}/M_\odot) = 9.83 \pm 0.02$, which is very consistent with the findings of Gebhardt & Thomas (2009); Gebhardt et al. (2011). Perhaps our most interesting result is that M87’s dark matter halo is cored.

5 DISCUSSION

5.1 Comparison with other studies

Part of the impetus for this study was to resolve the discrepancies between recent models of M87’s mass structure. The two recent analyses of Zhu et al. (2014) and Agnello et al. (2014), both of which relied heavily on the Strader et al. (2011) GC kinematics, disagreed on the total stellar mass of the system by almost a factor of two, with Agnello et al. (2014) finding a luminous mass $M(R < 135 \text{ kpc}) = 5.5^{+1.5}_{-2.0} \times 10^{11} M_\odot$ and Zhu et al. (2014) inferring $M(R < 135 \text{kpc}) \sim 1 \times 10^{12} M_\odot$. There are a number of possible reasons for this: first, while both studies over-lapped in the majority of their GC data, each used different mass models, and it is possible that these may have unnecessarily constrained the inference on the mass. Specifically, Zhu et al. (2014) chose to use a logarithmic potential model for the dark matter, so enforcing a core, whereas Agnello et al. (2014) used a power law for the dark halo and inferred a cusp with $\gamma \sim 1.6$. Further, while Agnello et al. (2014) relied solely on the GCs, separating them into three independent populations based on their velocities, positions and colours, Zhu et al. (2014) treated all the GCs as a single tracer population, but used the same SAURON data as in this study to constrain the stellar mass-to-light ratio. Thus it is possible that Agnello et al. (2014), though using a more flexible mass model in the centre, lacked the data coverage.

© 2015 RAS, MNRAS 000, 000–000
Table 2. Final inference on the parameters in the different $\beta = 0$ models. We report the maximum-likelihood values of our inferred posterior distributions, along with the 16th and 84th percentiles as a measure of our uncertainty. All quantities are measured in units of solar mass, solar luminosity and kiloparsecs.

<table>
<thead>
<tr>
<th>halo model</th>
<th>$M_*/L$</th>
<th>$\log(\rho_{DM})$</th>
<th>$r_0$</th>
<th>$\gamma$</th>
<th>$\log(M_{BH})$</th>
<th>$\Delta\sigma^2$</th>
<th>$r_\alpha$</th>
<th>$\log(M_{vir})$</th>
<th>$R_{vir}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gNFW</td>
<td>7.53 ± 0.04</td>
<td>7.93 ± 0.08</td>
<td>121 ± 14</td>
<td>0.01$^{+0.04}_{-0.01}$</td>
<td>9.83 ± 0.02</td>
<td>14.3 ± 0.3</td>
<td>--</td>
<td>14.27$^{+0.74}_{-0.38}$</td>
<td>1480$^{+220}_{-210}$</td>
</tr>
<tr>
<td>cgNFW</td>
<td>7.54 ± 0.03</td>
<td>6.87 ± 0.07</td>
<td>123 ± 12</td>
<td>0.03$^{+0.03}_{-0.03}$</td>
<td>9.81 ± 0.06</td>
<td>14.2 ± 0.2</td>
<td>15$^{+0.19}_{-0.15}$</td>
<td>14.27$^{+0.12}_{-0.19}$</td>
<td>1475$^{+100}_{-215}$</td>
</tr>
<tr>
<td>LOG</td>
<td>7.57 ± 0.04</td>
<td>7.33 ± 0.06</td>
<td>78 ± 9</td>
<td>--</td>
<td>9.81 ± 0.02</td>
<td>14.1 ± 0.3</td>
<td>--</td>
<td>14.28$^{+0.77}_{-0.35}$</td>
<td>1490$^{+310}_{-200}$</td>
</tr>
</tbody>
</table>

Figure 4. Inference on the gNFW model parameters. There is some covariance between the structural parameters, and we also note that the posterior on the inner slope, $\gamma$, hits the lower limit imposed by the prior.
in these central regions that would have permitted a reliable distinction between cusps and cores.

By exploring multiple mass models and combinations of data, we are able to reproduce the results of both studies. First, excluding the stellar kinematics and carrying out the inference using just the satellites and GCs in a way more similar to that of Agnello et al. (2014), we infer a gNFW profile with $\gamma \sim 1.6$, thus reproducing their finding of a cusp. This data combination also provides only a weak constraint on the stellar mass-to-light ratio, as might be expected given the dearth of GCs in the centre. When we compare the predictions of the best-fit model in this case, however, we find it significantly overpredicts the stellar velocity dispersions in the centre. Adding in the stellar data, then, requires an excavation of the central regions, giving rise to the core as modelled in Zhu et al. (2014) and a comparable stellar mass of $M_*(R < 135 \text{ kpc}) = (1.03 \pm 0.03) \times 10^{12} M_\odot$. We also find a significantly lower DM fraction than Agnello et al. (2014) in the central regions: while they measure $f_{DM} \sim 0.2$ at a radius of 6 kpc, we find $f_{DM} < 0.1$. However, at large radii the agreement is much better, with both finding $f_{DM} \sim 0.94$ at 135 kpc. Interestingly, even the cgNFW, with its greater

Figure 5. Inference on the LOG model parameters. Note the unsurprising covariance between the normalisation $\rho_0$ and the break radius $r_0$ of the halo.
Figure 6. Inference on the cgNFW model parameters. The picture that emerges here is that the cgNFW model is actually recovering the best gNFW model, with a break radius of around 100 kpc, inner slope $\gamma \sim 0$ and a core radius $r_0$ that is relatively unconstrained.

flexibility in intermediate regions, seems to converge onto a similar solution to the gNFW, with $\gamma \sim 0$ and $r_0 \sim 100$ kpc. The solution of the discrepancy between the two studies is presumably then that it is simply not possible to constrain the whole mass profile using GC data alone. This really shows the importance of combining multiple tracer populations with different characteristic radii.

Unlike those studies, our dataset also includes the high spatial-resolution kinematics in the very centre, allowing us to infer the black hole mass. As we are using the same NIFS data as Gebhardt et al. (2011) here, we should not expect our result to differ significantly from theirs: however, in that study, they used a logarithmic potential for the dark halo and fixed the parameters of both the halo and the stellar mass-to-light ratio to fit solely for the black hole, whereas we fit for all parameters at once, so it is an interesting consistency check to compare our results. We infer $\log(M_{BH}/M_\odot) = 9.83 \pm 0.02$, which is consistent with their results and has a similar uncertainty. However, we find a smaller mass-to-light ratio than used in that study - compare our $\Upsilon_{*,V} = 7.5$ with their $\Upsilon_{*,V} = 9.1$. Our best LOG halo also has a slightly lower circular velocity of 750 kms$^{-1}$
compared to their 800 km s$^{-1}$, and a significantly larger scale radius of 78 kpc compared to their $r_0 = 36$ kpc. This is interesting, as it seems to support the claim in that paper that the resolution of the NIFS data is sufficient to eliminate any significant correlations between the BH mass and the halo and stellar parameters. As we should also expect the BH mass to be degenerate with anisotropy, this further suggests that our assumption of isotropic stellar orbits is reasonable.

5.2 Comparison with other galaxies

It is also interesting to compare our findings for the halo structure with other ETGs; as explained in the Introduction, this is important for developing our understanding of the role of baryons in shaping the halo and the diversity of structures that can arise.

A number of other studies have found elliptical galaxies to have DM profiles which are flatter than NFW; Newman et al. (2011), for instance, combined stellar dynamics with weak and strong gravitational lensing and X-ray data to infer a density slope $\gamma < 1$ for the BCG Abell 383 with 95% confidence, while Newman et al. (2013) fitted gNFW halo models to a sample of seven massive BCGs and found a mean slope of $\gamma = 0.5 \pm 0.1$, inconsistent with the NFW prediction. However, our inferred gNFW slope of $\gamma = 0$ is significantly flatter than this; Schuberth et al. (2012) found a comparable result for the Virgo elliptical NGC 4636, for which a cored logarithmic halo was preferred over NFW and Burkert profiles (though the latter two could be no means be ruled out); also, Pota et al. (2015) used a Jeans analysis for the elliptical NGC 1407 and found the slope of the halo to be relatively weakly constrained, but consistent with cored profiles. This is also in keeping with the studies of the halos in spiral galaxies and dwarf spheroids (e.g. Herrmann & Ciardullo 2009; Walker & Peñarrubia 2011; Amorisco & Evans 2012), although studies of these galaxy types are currently more numerous and tend to be able to rule out NFW slopes with higher confidence due to the fewer degeneracies involved. The picture that emerges, then, is that of a diversity of halo structures as opposed to some universal, one-size-fits-all law, thus highlighting the complexity of the baryonic processes at work.

5.3 Implications of a cored profile

While the cored DM halo that we infer is inconsistent with DM-only simulations, there have been a number of recent simulations in which baryonic effects have been included, and we note some similarities with these. Laporte & White (2015) ran zoom-in simulations of two BCGs from redshift $z = 2$ to $z = 0$ and demonstrated that the effect of repeated dissipationless mergers is to soften the otherwise-NFW-like cusp by $\Delta \gamma \sim 0.3$–0.4 on scales of the stellar half-light radius. In this paradigm, infalling satellite galaxies experience dynamical friction from DM as they move through the halo, and this transfers energy to the DM, causing it to expand and so become less dense in the central regions. While we do not measure a core radius explicitly, in Figure 10 we plot the density slope as a function of radius, $d \log \rho / d \log r$, and see that the slope becomes sub-NFW at radii $r < 60$ kpc and that $\gamma \sim 0.6$ (equivalent to $\Delta \gamma \sim 0.4$) at $r \sim 25$ kpc, which is comparable to the scale radius of the starlight measured by Kormendy et al. (2009). However, we find a cored centre while these simulations find only $0.6 < \gamma < 1$ down to the innermost resolvable radius: while they suggest that the merging of SMBHs could then give rise to cores in the central $3$–$4$ kpc – with the binary BH system spiralling inwards and transferring energy to the surrounding matter – their simulations do not include the effects of SMBHs and so they are unable to test this. It therefore remains unclear.

Figure 7. Left: the mass profiles inferred using different halo models are fairly similar across the radius range in which we have most data. Here we plot a measure of the squared circular velocity, $M(r)/r$, in order to highlight the differences. At small radii, all are dominated by the Nuker profile, while at large radii they differ significantly, with the LOG density profile decaying as $r^{-2}$ while the others go as $r^{-3}$. Right: the DM fraction as a function of radius. In the central regions, the mass is totally baryon-dominated, whereas by a projected radius of $\sim 200$ kpc, the tables have turned.
as to whether merging events would be capable of totally erasing the cusp.

On the other hand, M87 has a large AGN at its centre and this may also contribute to core formation. The study of Martizzi et al. (2012), for instance, specifically simulated Virgo-like ETGs using a recipe for AGN feedback and found cores to develop within the inner 10 kpc. In this scheme, outflows of gas due to the AGN are able to irreversibly modify the gravitational potential of the halo and so cause expansion of both the dark and luminous matter. In that study, they suggest that a combination of AGN feedback and SMBH effects are the main contributing processes, though they also note that the large amount of gas expelled by the AGN increases the efficiency of the energy transfer due to dynamical friction. This role for AGN is in line with previous findings: our stellar mass is very similar to that inferred by Zhu et al. (2014), while the virial mass and radius implied by their logarithmic halo model are slightly smaller than ours, with log(M_{vir}/M_\odot) = 13.98, R_{vir} = 1.18 Mpc, making their stellar-to-halo mass ratio even more anomalous. Murphy et al. (2011) infers both a larger stellar mass (with Υ_⋆/V_\odot = 9.1 ± 0.2) and a larger halo mass (log(M_{vir}/M_\odot) = 14.37, R_{vir} = 1.6 Mpc). The inference is, then, that M87 appears to have an unusually high stellar mass relative to its halo, and it is interesting to speculate as to why this might be the case. As M87 lies at the centre of a massive cluster, we might look to the role of merger and accretion events.

5.5 Comparison with stellar population models

A basic principle of stellar evolution is that stellar populations fade as they age, and this means we can use our inferred ratio Υ_⋆ to make some statements about the evolutionary stage of M87’s stars. The stellar evolution tracks of Padova (1994) (provided as part of the stellar population synthesis package of Bruzual & Charlot 2003) provide ages, metallicities and V-band T values in tabulated form for both Salpeter and Chabrier IMFs: when we compare these with our mass-to-light ratio, we find that we are consistent with a solar metallicity (Z = 0.02) and log(T/yr) = 10.14, or supersolar metallicity (Z = 0.05) and log(T/yr) = 9.97, but only under the assumption of a Salpeter IMF. The Chabrier IMF cannot generate such high mass-to-light ratios at any metallicity. While this is a very basic way to constrain the age – assuming as it does a single stellar population and interpreting its global properties – it is nevertheless consistent with other, more targeted studies. In Montes et al. (2014), for instance, the inner ~ 8 kpc are studied with high-resolution spectra in 16 bands, and are found to imply log(T/yr) = 9.8 – 10.2 and Z = 0.02 – 0.05 (though with a radial gradient in the metallicity), while Kuntschner et al. (2010) analysed the SAURON data to find a metallicity of 0.13 [Z/H] (i.e. Z = 0.03) within an adopted effective radius of 8.8 kpc. Equally, the favouring of a Salpeter IMF over the bottom-light Chabrier IMF corroborates the findings of Auger et al. (2010) for other massive ETGs and those of Grillo & Gobat (2010) for massive ETGs in the Coma cluster – a comparable environment to Virgo – and reasserts the idea that massive ETGs tend to be better described by bottom-heavy IMFs.

To get another angle on these physical properties, we consider the DM fraction f_{DM} as shown in Figure 7. The central DM fraction in ETGs is understood to correlate with stellar mass, anti-correlate with stellar age and further depend on the form of the IMF, increasing as the IMF is made interesting to see how our results compare with these expectations.

We find M87 to have a high stellar mass relative to its halo mass. While our Nuker profile for the projected light is not strictly normalisable, we take the half-light radius of R_e = 16 kpc from Kormendy et al. (2009) and define the total stellar mass as twice the mass within this radius. This gives a stellar mass log(M_{star}/M_\odot) = 11.95 ± 0.30. Compared with the virial mass log(M_{vir}/M_\odot) = 14.27, this places M87 above the abundance-matching curves of Behroozi et al. (2010), even allowing for uncertainties. This appears to be in line with previous findings: our stellar mass is very similar to that inferred by Zhu et al. (2014), while the virial mass and radius implied by their logarithmic halo model are slightly smaller than ours, with log(M_{vir}/M_\odot) = 13.98, R_{vir} = 1.18 Mpc, making their stellar-to-halo mass ratio even more anomalous. Murphy et al. (2011) infers both a larger stellar mass (with Υ_⋆/V_\odot = 9.1 ± 0.2) and a larger halo mass (log(M_{vir}/M_\odot) = 14.37, R_{vir} = 1.6 Mpc). The inference is, then, that M87 appears to have an unusually high stellar mass relative to its halo, and it is interesting to speculate as to why this might be the case. As M87 lies at the centre of a massive cluster, we might look to the role of merger and accretion events.

5.4 Relation to expectations from abundance matching

Abundance matching is a well-established technique for predicting the virial halo mass of a galaxy given its stellar mass. As we infer both dark and luminous mass components, it is
increasingly bottom-light (Napolitano et al. 2010; Auger et al. 2010). Comparing our results with the ETG sample of Napolitano et al. (2010), we find that our DM fraction within the adopted effective radius of 16 kpc is low compared with their sample assuming a Kroupa IMF but is consistent with the predictions of their toy models, given M87’s stellar mass, if a Salpeter IMF is assumed. In both cases, the implication is that the stellar populations are old, though we are unable to put any meaningful quantitative constraints on the age from this comparison.

5.6 Importance of correctly modelling the underlying tracer populations

The Jeans equation requires us to know the underlying tracer density for each population, and this, along with the mass profile and anisotropy, determines the velocity dispersion that is measured. Therefore any uncertainty or bias in the parameterisation of the tracer density might manifest itself as uncertainty or bias in the resulting velocity dispersions and, in a study such as this in which the velocity dispersions are the primary tool for the inference, also as uncertainty or bias in the mass profile. A particular worry here are the distributions of the GCs, as the spectroscopic catalogue of Strader et al. (2011) is only a subsample of the wider population and, by virtue of its being selected for spectroscopy, is subject to some non-trivial selection function that may change the apparent distribution from that of the true underlying one. It is therefore important to model the distribution using an independent photometric sample. Indeed, this was the motivation for the initial photometric study of Oldham & Auger (2015, in prep.).

To demonstrate the impact on the inference of using the density of the spectroscopic subsample in the Jeans analysis rather than that of the more representative photometric sample, we repeat our analysis using the three-population decomposition given in Agnello et al. (2014). This decomposition was carried out based solely on the Strader et al. (2011) spectroscopic catalogue, and separates the GCs into a compact red component along with two much more extended blue and intermediate-colour components. Our analysis differs from Agnello et al. (2014)’s in that they focused mainly on virial decompositions and used the GC data exclusively, whereas we use a spherical Jeans analysis and combine the four datasets, so this comparison is not intended as a comment on their work, but merely aims to show that, in this particular analysis, the use of this approximation to the true distribution is important. When we carry out the inference using the three populations, we find that we underpredict the halo mass in the region in which the GCs dominate the fit by ~ 0.25 dex. $M_{\text{vir}}$ also decreases by ~ 10%, though the effect is presumably smaller here because the GC data only extend out to a fraction of the virial radius. Unsurprisingly, our inference on the stellar mass-to-light ratio is robust against the GC Sérsic distributions, as this is mainly determined by the properties of the stellar population, which is fixed by our fit to the surface brightness profile. This shows, then, that using the incorrect tracer density for any of the tracer populations has a significant systematic effect on the mass inference.

5.7 On the assumption of isotropy

One limitation of our model is the assumption of isotropy. While this is the simplest case, and motivated on the grounds of M87’s apparent symmetry and slow rotation, deviations from isotropy could cause a bias in our results. For instance, by enforcing isotropy, we may be disfavouring orbits with extreme api- and periheia, and this could have an impact on our inferred inner halo slope. To assess the importance of this degeneracy on our inference, we note that previous work suggests a mild radial anisotropy for the stars (Romanowsky et al. 2011; Murphy et al. 2011) and largely isotropic orbits for the GCs (Côté et al. 2001; Zhu et al. 2014): we therefore run a series of models in which we fix the stellar anisotropy at some non-zero constant value while keeping the GCs isotropic, and compare the results.

To impose sensible limits on the anisotropy, we reason that it would be unlikely for the radial component of the characteristic stellar velocity to be either greater than twice the tangential component, or less than half: we therefore only consider anisotropies in the range $-1 < \beta < 0.5$. We find that, for the $gNFW$ and $cgNFW$ models, our inference on $\gamma$ is robust against changes in anisotropy – $\gamma = 0$ is still favoured – but note the appearance of two sweetspots or pinch radii in the mass profile at which the inferred enclosed mass is virtually independent of anisotropy, as shown in Figure 11. The existence of a pinch radius has been demonstrated by Wolf et al. (2010) for Jeans modelling based on a single tracer population, and shown to occur at approximately the 3D effective radius of the tracer population, where the dependence of the enclosed mass profile on anisotropy is minimized. Indeed, we do find the three $\beta$-curves to intersect at a radius of ~ 14 kpc, comparable to the effective radius of the stars. However, in our models we see a second pinch point at $R \sim 218$ kpc. This is much too large to be associated with the effective radius of the starlight, and may be an indication of where the dependence of the enclosed mass on the model parameters – as opposed to the anisotropy – is minimized. In this picture, the degeneracy at small radii (where we would expect the changing stellar anisotropy to have most effect on the inferred mass) puts some constraints on the parameters of the model, which influence the inferred mass at larger radii: there may then exist a radius further out at which the dependence of the mass profile on changing the model parameters is minimized, giving rise to the second pinch radius that we observe. This is an interesting idea which will warrant further investigation. For now, though, we simply report the total enclosed mass at each of the pinch radii for the $gNFW$ model in Table 5.7, and note that the extra scatter introduced by the mass-anisotropy degeneracy should be borne in mind when interpreting the main results of this study, which assume isotropic motions. We also note that, as can be seen in Figure 11, the mass inferred from the isotropic model has larger uncertainties than the other two, suggesting that this is the model with the most flexibility. This generally supports our choice to model the tracers isotropically.
L. J. Oldham and M. W. Auger

Figure 11. Varying the stellar anisotropy has an effect on the inferred mass profile at all radii except for the two pinch radii, clearly visible at $\sim 14$ kpc and $\sim 218$ kpc, where the dependence of the mass on the anisotropy is minimised.

$$\frac{R_p}{\text{kpc}} \quad \log(\frac{M(<R_p)}{M_\odot})$$

<table>
<thead>
<tr>
<th>$R_p$/kpc</th>
<th>$\log(\frac{M(&lt;R_p)}{M_\odot})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>11.72 ± 0.17</td>
</tr>
<tr>
<td>218</td>
<td>13.48 ± 0.4</td>
</tr>
</tbody>
</table>

Table 3. When the anisotropy is changed from zero, two pinch radii emerge at radii $R_p$ at which the dependence of the enclosed mass on the anisotropy is minimised. These masses are reported in the table, and should be robust against our assumption of isotropy.

6 CONCLUSIONS

We have modelled the mass profile of M87 using a combination of stellar, GC and satellite galaxy kinematics in a Jeans analysis, and our main conclusions are as follows:

(i) M87 is a massive BCG with a cored DM halo. It has a virial mass of $\log(M_{\text{vir}}/M_\odot) = 14.27^{+0.72}_{-0.38}$ and a virial radius $R_{\text{vir}} = 1.48^{+0.29}_{-0.21}$ Mpc and a stellar mass $\log(M_{\text{st}}/M_\odot) = 11.95 \pm 0.30$, placing it at the top end of the galaxy mass distribution and above the stellar-to-halo mass relations expected from abundance matching.

(ii) M87 has a relatively high stellar mass-to-light ratio of $\Upsilon_\star = 7.53 \pm 0.04$ in the V-band, consistent with a picture in which its stellar populations are old and with a metallicity at least equal to the solar value. Making a simple comparison of our inferred $\Upsilon_\star$ with the stellar evolution tracks of Padova (1994), we find that such a large mass-to-light ratio is inconsistent with a Chabrier IMF, but can be accommodated by a Salpeter IMF with metallicity $Z > 0.02$ and age $\log(T/\text{yr}) > 9.96$. This is consistent with a picture of star formation followed by relatively passive evolution, and supports other findings that the Salpeter IMF provides a good description for massive ETGs.

(iii) The inclusion of tracers at a variety of spatial scales has a significant impact on the inference. Modelling the mass using the GCs and satellite galaxies alone - that is, without the stars - we infer a cuspy halo, with inner slope $\gamma \sim 1.6$, but when we include the stars in the inference, the halo becomes decisively excavated in the centre and we infer $\gamma =$ 0. This shows the importance of consistently modelling the profile across a range of spatial scales.

(iv) It is important to properly characterise the distributions of the underlying tracer populations as opposed to those of their kinematic subsamples. We have shown that the use of GC colour and spatial distributions based on the kinematic dataset alone leads to a systematic underprediction of the total mass of the system.

(v) Assuming isotropic orbits for all the tracers, as done here, does not appear to impose artificial constraints on the halo parameters. When we relax this assumption and rerun the inference with different (constant) values for the stellar anisotropy between -1 and 0.5, we find that, while the inferred mass profile changes, a cored halo is always preferred. Two pinch radii also emerge at roughly 14 kpc and 218 kpc, and the inferred mass is robust against changes in the anisotropy at these points.
References


Grillo C., Gorini C., 2010, MNRASLett., 402, L67


Murphy J. D., Gebhardt K., Adams J. J., 2011, p. 29


