Triaxial Schwarzschild models of NGC 708: a 10-billion solar mass black hole in a low-dispersion galaxy with a Kroupa IMF

Stefano de Nicola [⬤],[1,2][⋆] Jens Thomas [⬤],[1,2] Roberto P. Saglia,[1,2] Jan Snigula,[1] Matthias Kluge[1,2] and Ralf Bender[1,2]

[1]*Max-Planck Institute for Extraterrestrial Physics, Giessenbachstrasse 1, D-85748 Garching, Germany*
[2]*Universitäts-Sternwarte Muenchen, Scheinerstrasse 1, D-81679 Munich, Germany*

ABSTRACT

We report the discovery of a $(1.0 \pm 0.28) \times 10^{10} \, M_\odot$ supermassive black hole (BH) at the centre of NGC 708, the Brightest Cluster Galaxy of Abell 262. Such high BH masses are very rare and allow to investigate BH–host galaxy scaling relations at the high mass end, which in turn provide hints about the (co)evolution of such systems. NGC 708 is found to be an outlier in all the canonical scaling relations except for those linking the BH mass to the core properties. The galaxy mass-to-light ratio points to a Kroupa IMF rather than Salpeter, with this finding confirmed using photometry in two different bands. We perform this analysis using our novel triaxial Schwarzschild code to integrate orbits in a five-dimensional space, using a semiparametric deprojected light density to build the potential and non-parametric line-of-sight velocity distributions (LOSVDs) derived from long-slit spectra recently acquired at Large Binocular Telescope (LBT) to exploit the full information in the kinematic. We find that the galaxy geometry changes as a function of the radius going from prolate, nearly spherical in the central regions to triaxial at large radii, highlighting the need to go beyond constant shape profiles. Our analysis is only the second of its kind and will systematically be used in the future to hunt supermassive BH in giant ellipticals.

Key words: galaxies: elliptical and lenticular, cD – galaxies: kinematics and dynamics – galaxies: structure.

1 INTRODUCTION

NGC 708 is the Brightest Cluster Galaxy (BCG) of the Abell 262 cluster. It is located at a distance of 68.48 Mpc and has an absolute magnitude $M_{g'} = -22.89$ (Kluge et al. 2020). The galaxy is commonly classified as cD (see e.g. Wegner et al. 2012). According to this classification, the galaxy is surrounded by a diffuse stellar envelope, whose origin is believed to lie in *ex situ* stellar accretion (Cooper et al. 2015; Pillepich et al. 2018), and whose evolution is tightly linked to the whole cluster rather than to the BCG itself (Gonzalez, Zabludoff & Zaritsky 2005). This stellar envelope probably traces the cluster potential and is typically referred to as intracluster light (ICL; Kluge et al. 2020). Kluge et al. (2021) attempt to dissect the BCG light from the ICL in several ways, finding that an accurate photometric decomposition cannot be obtained. A possible solution to the problem could be the comparison of SB profiles of BCGs with those of ordinary ETGs (Kluge & Bender 2023). Kluge et al. (2020) report as best-fit parameters for the SB profile of this galaxy a single Sersic profile with Sersic index $n = 2.96 \pm 0.11$ and effective radius $r_e = 54.26^{+1.32}_{-1.29}$ kpc.

HST images of the galaxy show the presence of a prominent, edge-on dust lane, extending ~ 3 arcsec in the southern direction and ~ 6 arcsec in the northern direction (Wegner et al. 2012). Its origin probably lies in a merger, a commonly observed phenomenon in

BCGs (Lauer et al. 2014). In fact, BCGs lie at the centre of potential wells in galaxy clusters, thus being in the ideal position to accrete material, and to experience several merging events. Another possible indication of a merger is given by the wiggles observed in the SB, ellipticity ε and PA profiles (see Section 2.1).

Perhaps the most interesting clue about merger(s) that this galaxy has experienced can be found in its light-deficient central core, which extends out to ~ 2.2 arcsec (~ 0.73 kpc, see Section 2.2). The most plausible formation mechanism is core scouring: stars on radial orbits, which come closest to the galaxy centre, are ejected via gravitational slingshot by a shrinking supermassive black hole (SMBH) binary (Ebisuzaki, Makino & Okumura 1991; Faber et al. 1997; Milosavljević & Merritt 2001; Merritt 2006; Thomas et al. 2014, 2016; Mehrgan et al. 2019) that formed after a major (dry) merger. The ejection of stars on radial orbits also generates a tangential anisotropy in the core, a commonly observed phenomenon in core-galaxies (Gebhardt et al. 2003; van den Bosch & de Zeeuw 2010; Thomas et al. 2014, 2016). Scaling relations linking the black hole mass $M_{\rm BH}$ to the core size (Rusli et al. 2013b; Saglia et al. 2016), the missing light with respect to an inward extrapolation of the SB profile (Kormendy & Bender 2009), and the SB of the central core itself (Mehrgan et al. 2019) have been discovered. For NGC 708, the core size predicts an $M_{\rm BH}$ value of $\geq 10^{10} \, M_\odot$. Hunting SMBHs with masses in this range will help in filling the high-mass end of the BH-host scaling relations: as it can be seen from fig. 11 in Mehrgan et al. (2019), $M_{\rm BH}$ estimates in this range are currently almost completely missing.

⋆ E-mail: denicola@mpe.mpg.de

Another interesting issue lies in the stellar initial mass function (IMF) of massive galaxies. Although there are massive ETGs following a lightweight IMF (Thomas et al. 2015, 2016), several studies (Thomas et al. 2011; Cappellari et al. 2012; Tortora et al. 2014) suggest that the most massive ETGs may follow a bottom-heavy (e.g. Salpeter 1955) IMF. To test this idea, mass-to-light ratios Υ calculated by dynamically modelling these galaxies are compared to estimations from stellar population analysis (Thomas, Maraston & Bender 2003; Maraston 2005; Maraston & Strömbäck 2011; Conroy, van Dokkum & Villaume 2017; Parikh et al. 2018) assuming a Kroupa (2001) IMF, finding systematically overestimated Υ with respect to Kroupa IMF, even if this could also signal Dark Matter (DM) tracing the stars. In Wegner et al. (2012), Υ estimates using both SSP (Υ_{SSP}) and dynamical models (Υ_*) for NGC 708 are published, finding evidence for a bottom-heavy IMF. However, Wegner et al. (2012) did not include a central BH in their axisymmetric dynamical modelling.

In order to obtain reliable estimates of the mass parameters, it is essential to robustly recover the intrinsic shape of a galaxy. BCGs have been shown to be extremely triaxial objects (de Nicola et al. 2022b), even more than ordinary early-type galaxies (Vincent & Ryden 2005). For NGC 708, de Nicola et al. (2022b) showed by deprojecting the surface brightness profile of the galaxy that, for the best-fit orientations, the galaxy is close to being spherical in the central regions, but becomes triaxial at large radii: the triaxiality parameter $T = (1 - p^2)/(1 - q^2)$, where $p = b/a$ and $q = c/a$ are the intrinsic axis ratios with $a > b > c$, reaches the maximum value of 0.5 at $r \sim 3$ kpc, before going down to 0.3 at larger radii. In this case, an axisymmetric approximation can possibly yield biased black hole mass M_{BH} estimates (van den Bosch & de Zeeuw 2010) as well as Υ estimates biased by 50 per cent (Thomas et al. 2007).

In two recent works, de Nicola et al. (2022a) and Neureiter et al. (2023a) tested a new sophisticated triaxial machinery, which combines the semiparametric deprojection algorithm SHAPE3D (de Nicola et al. 2020), the non-parametric code for kinematics extraction WINGFIT (Mehrgan et al. 2023), and the triaxial Schwarzschild code SMART (Neureiter et al. 2021), along with the novel model selection technique of Lipka & Thomas (2021) and Thomas & Lipka (2022), which avoids χ^2-biases linked to the different number of degrees-of-freedom (DOFs) in each model. The authors apply this to an *N*-body simulation (Rantala et al. 2018), which reproduces the formation of a typical core-galaxy, recovering the correct galaxy shape and orbit distribution with an accuracy $\Delta p, \Delta q, \Delta \beta < 0.1$, where β is the anisotropy. In a companion paper (Neureiter et al. 2023a), the same simulation is used to show that also M_{BH} and the mass-to-light ratio Υ can be recovered within 5–10 per cent.

The main goal of this paper is to apply our triaxial machinery to NGC 708 to investigate its intrinsic properties and in particular the mass of its SMBH. To this extent, we also make use of new LBT kinematical data and *K*-band photometry acquired at the Wendelstein observatory. The paper is structured as follows: Section 2 describes the photometric data and the deprojections, while Section 3 describes the spectroscopic data and kinematics extraction. Section 4 presents the results of the dynamical modelling, which are discussed in Section 5. Finally, we draw our conclusions in Section 6.

2 PHOTOMETRY

The photometry used in this work comes from three different image sources. The first is a g'-band image obtained with the Fraunhofer Telescope at the Wendelstein observatory (Hopp et al. 2010; Lang-Bardl et al. 2016) using the Wendelstein Wide Field

Imager (WWFI; Kosyra et al. 2014), with typical seeing of FWHM = (1.2 ± 0.2) arcsec and pixel size of 0.2 arcsec. The total field of view (FOV) of 27.6×28.9 arcmin allows to image the galaxy outskirts,[1] where the galaxy light mixes up with the ICL. The galaxy was imaged following a 52-step dither pattern; corrections for bias, flat-field, cosmic rays, background, and bright stars were applied. Technical details can be found in Section 3 of Kluge et al. (2020). A second image comes from high-resolution *HST* observations (see top panels in Fig. 1), with typical resolution of 0.1 arcsec. These were carried out using the Wide Field Planetary Camera 2 (WFPC2) using the filter F622W (GO program 10884; P.I. Wegner). The camera consists of a grid with dimensions 800×800 pixels, with pixelsize 0.0455 arcsec, yielding a FOV of 800×800 arcsec (Wegner et al. 2012). The galaxy belongs to a subset of 170 Brightest Cluster Galaxies (BCGs), observed down to a $SB_{g'} \sim 30$ mag asec^{-2} (Kluge et al. 2020). This allows to reach the galaxy regions, where the interaction with the ICL becomes visible in the SB profile itself (Kluge et al. 2021), as well as for a comparison between the galaxy intrinsic shape and the simulated DM halos (de Nicola et al. 2022b). Complementarily, the *HST* images give us the necessary resolution in the central regions to resolve the black hole sphere-of-influence (SOI), where the potential is dominated by the black hole itself. This is crucial to derive reliable M_{BH} estimates. Finally, we also use a *K*s-band Wendelstein image (see bottom panels in Fig. 1) to assess how much the presence of dust in the central region affects our analysis.

Because the two optical image sources are in different colour bands, we select the radii where we have data from both observation sets and combine them together. This is done by first interpolating the *HST* photometry at Wendelstein radii and then minimizing the differences between the two data sets. The resulting scaling factor is used to convert *HST* data to the g'-band. After the conversion has been made, we take *HST* data at $r < 15$ arcsec, and WWFI data at $r > 15$ arcsec. In this way, we have a very high resolution at the centre, which is what we need for a robust estimate of M_{BH}, and extending out to large radii (≥ 30 kpc). The resulting photometry is used to perform the triaxial deprojection described in Section 2.3.

2.1 Isophote features

We use the methods of Bender & Möllenhoff (1987) (for the g'-band image) and Kluge & Bender (2023), Kluge et al. (2023) (for the *K*s-band image) to extract the isophote parameters from the galaxy image. The two methods yield very similar results. The variables we are interested in are the SB, the ellipticity ε, the position angle (PA), and the Fourier coefficients a_n, b_n, which quantify the deviations of the isophotes from the best-fit ellipses. This is done *separately* for the two optical image sources before combining the isophotes as described above. The resulting g'- and *K*s-band isophotes are shown in Fig. 2. A few key-points to mention are:

(i) There are wiggles in both the SB and the ε. These are probably signs of a recent merger;

(ii) The ε and PA profiles show huge bumps at small radii due to the dust lane. This also generates wiggles in the SB profile at small radii;

(iii) With the exception of the innermost 10 arcsec, all variables follow similar trends as those observed in ordinary ETGs (Bender, Döbereiner & Möllenhoff 1988; Bender, Burstein & Faber 1992; Kormendy & Bender 1996): ε is low at small radii before increasing

---

[1]The camera itself consists of four CCDs, each one having dimensions 4096×4109 pixels.
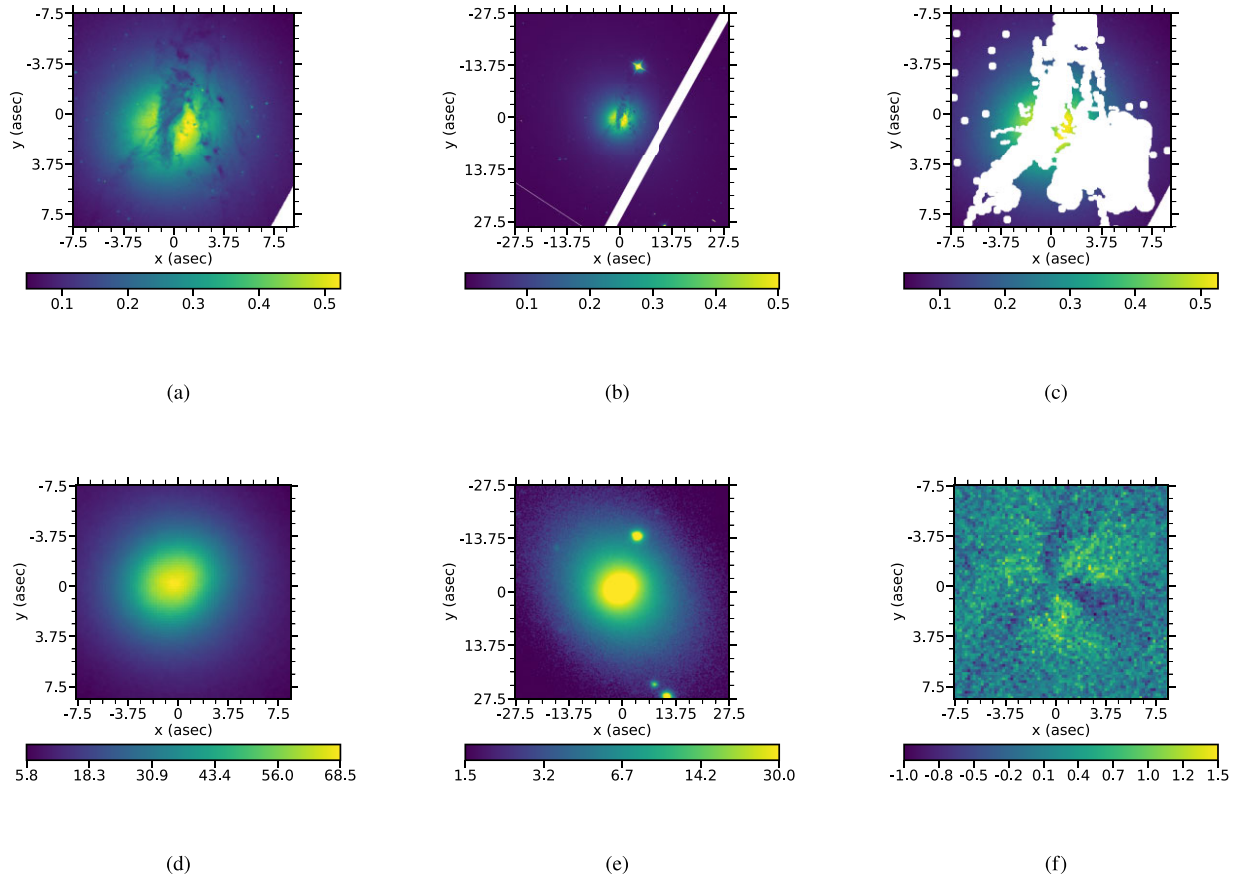
Figure 1. *HST*/Wendelstein images of NGC 708. Images are colour-coded according to the flux, which has been scaled for better visualization except for Figs 1(d) and (f) to show that the dust residuals are small. *Top-left: HST*-image, with the prominent dust lane clearly visible. *Top-middle HST* image with wider FOV. *Top-right*: masked *HST*-image used to generate the isophotes shown in left-hand panel of Fig. 2. *Bottom-left:* Wendelstein Ks-band image of NGC 708, no dust appearently visibile. *Bottom-middle:* Wendelstein Ks-image with wider FOV. Here, the isophote twist is evident. *Bottom-right: K*s-band residuals obtained by subtracting the fitted isophotes to the K-band image itself. A small residual due to dust is visible. In all images, North is up and East to the left.

at large radii, there is only a weak twist $<10°$ and the Fourier coefficients remain ≤ 1.5;

(iv) Given the dust lane, it is difficult to tell whether the galaxy is indeed round in the central regions as it is usually the case for BCGs (de Nicola et al. 2022b). Interestingly, ε goes towards zero already at 10 arcsec, i.e. farther out than the core size (see Section 2.2).

(v) The ellipticity and PA profiles obtained from the Ks-band photometry also show large gradients. Indeed, after subtracting the best-fit elliptical model to the galaxy image, the residuals do show traces of dust, even if the flux loss due to dust is only ~ 1.4 per cent in the Ks-band image (Figs 1 d and f).

In practice, we are dealing with a (possibly still ongoing) merger with dust at the galaxy centre. To prevent dust from biasing our analysis, we set ε and the PA to constant values in the innermost regions (see left-hand panel of Fig. 2), while using the Ks-band image of the galaxy to discuss the impact of this assumption on the BH mass estimate.

2.1.1 ICL contamination

The light profile of NGC 708 is probably a superposition of the actual BCG light and an ICL component. According to Kluge et al. (2021), NGC 708 can be well described by a single Sersic component and, with the exception of the innermost regions, the ellipticity and

PA profiles are typical of a massive ETG. For instance, the galaxy becomes flatter with increasing radius and displays a mild twist. Hence, the regions that we include in the dynamical modelling show no sign of being affected by a structurally distinct ICL component. Nevertheless, even in BCGs like NGC 708 without obvious signs of structural changes in the light profile, there is a potential contribution of the ICL. Kluge et al. (2021) discusses several selection criteria that consistently yield typical ICL fractions around ~ 50 per cent. Among them, a SB'_g cut placed at 27 mag asec^{-2}: assuming that all the light fainter than this threshold belongs to the ICL on average results in 34 per cent of all diffuse light centred on the BCG belonging to the ICL. The outermost isophote used for our models has $SB_{g'} < 25$ mag asec^2, well brighter than this threshold. This confirms again that our models are unlikely to be affected by the dynamics of the ICL. Hence, even if the ICL is contaminating the SB profile, any effect on the dynamical models is expected to be small and not to produce biases in our analysis. We may overestimate the luminosity of the actual BCG by up to a factor of ~ 2 though since we do not subtract any potential ICL contribution to the galaxy's light profile.

2.2 Core radius

Like many other massive ellipticals, NGC 708 shows a light-deficient core with respect to an inward extrapolation of a single Sersic profile. Here, we focus on the core size using the (nearly) dust-free,
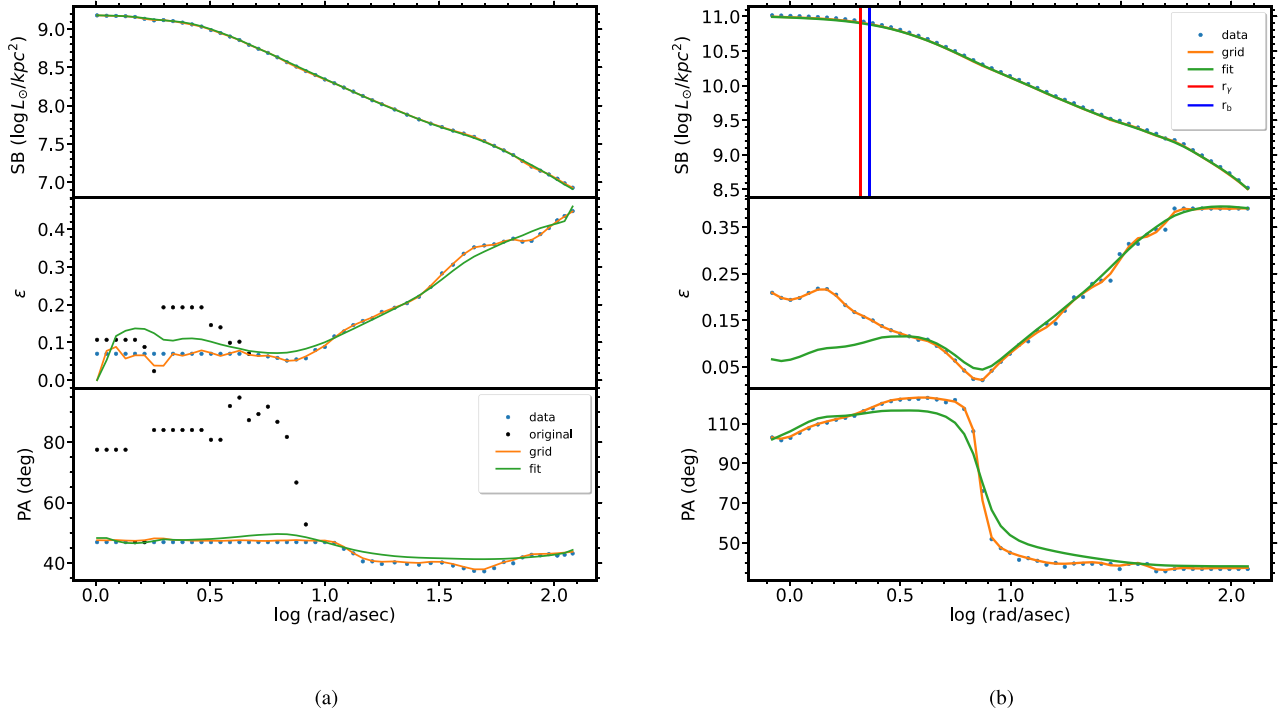
(a)   (b)

Figure 2. Isophotes (from top to bottom: SB, ellipticity, PA) of the galaxy NGC 708. The blue points represent the observed photometry, while the orange lines are computed by placing the SB on the grid and performing isophotal fits. The green lines come from the projection of the deprojected SB. *Right:* Same but using *K*s-band Wendelstein deconvolved isophotes. In the top-right-hand panel, we show the position of cusp-radius r_γ and break radius r_b, which we use to estimate the core size.

Table 1. Summary of the different core-radius estimates derived for NGC 708. *Col. 1:* Image used, where f110W/f814W means dust-corrected f110W image using f814W (and same for f110W/f555W). *Col. 2:* Method used – either cusp-radius or Core-Sersic (CS). *Col. 3–4:* Results in arcseconds and kpc. A value of 0.70 kpc implies $M_{BH} = 1.22 \times 10^{10}$ M$_\odot$ using the relation of Thomas et al. (2016).

Image	Method	Core radius (arcsec)	Core radius (kpc)
Wendelstein-*K*s	Cusp	2.10	0.70
	CS	2.29	0.76
f110W/f814W	Cusp	1.94	0.65
	CS	2.10	0.70
f110W/f555W	Cusp	2.08	0.69
	CS	2.43	0.81
	Avg Cusp	2.04 ± 0.09	0.68 ± 0.03
	Avg. CS	2.27 ± 0.16	0.75 ± 0.05

deconvolved *K*-band Wendelstein photometry of the galaxy, whose FWHM $\lesssim 1.2$ arcsec allows us to resolve the central core. The core size can be described both in terms of the cusp-radius r_γ, defined as the radius where d log Σ/d log $r = -1/2$, where $\Sigma = 10^{-0.4 \times SB}$, and of the break radius r_b, obtained by fitting a (double) core-Sersic law (Graham et al. 2003; Trujillo et al. 2004) to the SB profile. Adopting the first definition, we measure a core of 2.10 arcsec, corresponding to 0.70 kpc. The second method shows that the SB profile is better fitted by two components: a core-Sersic law, with Sersic index n_1 = 1.84 and effective radius $R_{e,1} = 0.50$ kpc, and a second simple Sersic profile with $n_2 = 1.59$ and $R_{e,2} = 2.59$ kpc. The size of the break radius is 2.29 arcsec, corresponding to 0.76 kpc, in line with the cusp-radius estimate. However, given that the galaxy might also host an homogeneous exponential dust disc, we exploit the *HST* WFPC3

f110W image (GO program 14219; P.I. Blakeslee) and correct for dust following Bender et al. (2015) (see also appendix A of Nowak et al. 2008) using both the *HST* f814W (GO program 5910; P.I. Lauer) and the *HST* f555W WFPC2 (GO program 7281; P.I. Fanti) images. Then, we measure the core radius from the two f110W dust-corrected images using both the cusp-radius r_γ and a core-Sersic fit to the SB profile to determine r_b, reporting the results in Table 1. The procedure is summarized in Appendix A. As we can see, regardless of the approach and/or image we choose, the result is stable around 0.70 kpc: according to the M_{BH}-core size relation (Thomas et al. 2016), this would imply a central BH with mass $> 1.0 \times 10^{10}$ M$_\odot$. Moreover, because the size of the sphere of influence of SMBHs in core galaxies is similar to the core size (Thomas et al. 2016), we expect the sphere of influence to be well resolved by our kinematic data.

2.3 Deprojection

One crucial ingredient needed when dynamically modelling a galaxy is the reconstructed triaxial light density ρ. To recover this, we deproject the SB profile using the software SHAPE3D extensively described in de Nicola et al. (2020). The code is a semiparametric method, which assumes the galaxy to be stratified on concentric ellipsoids:

$$m^{2-\xi(x)} = x^{2-\xi(x)} + \left[\frac{y}{p(x)}\right]^{2-\xi(x)} + \left[\frac{z}{q(x)}\right]^{2-\xi(x)}. \quad (1)$$

The exponent ξ can be used to generate discy ($\xi > 0$) or boxy ($\xi < 0$) bias. The four one-dimensional functions $p(x)$, $q(x)$, $\xi(x)$, along with the density on the *x*-axis $\rho_x(x)$, specify ρ at each point of the grid: the density values, which are not along the *x*-axis are updated through

a log-linear interpolation/extrapolation. Finally, the code fits $p(x)$, $q(x)$, $\xi(x)$, and $\rho_x(x)$, non-parametrically using a simulated annealing algorithm (Metropolis et al. 1953) and uses a one-dimensional radial smoothing to penalize against unsmooth solutions.

The code recovers the best-fit ρ by minimizing RMS= $\sqrt{\langle(\ln(I_{\mathrm{obs}}/I_{\mathrm{fit}}))^2\rangle}$, where I_{fit} is the projected surface luminosity corresponding to ρ for a given orientation. It allows to significantly restrict the range of viewing angles compatible with the observed photometry (de Nicola et al. 2020, 2022b) and to sample different solutions at the same orientation to explore the degeneracy intrinsically present in the deprojection (Rybicki 1987; Gerhard 1996; Gerhard & Binney 1996; van den Bosch 1997; Magorrian 1999; de Nicola et al. 2020).

For the deprojection, an intrinsic frame of reference $O(x, y, z)$ aligned with the galaxy principal axes is used, such that x, y, z are on the major, intermediate, and minor axis, respectively. The three viewing angles that are needed to determine the orientation of a triaxial body are the two angles (θ, ϕ), which specify the position of the line of sight (LOS), and a third misalignment angle ψ, defined as the angle between the projected z-axis on the plane of the sky and the x-axis, measured counterclockwise.

We place the SB on a circular 40×10 grid, with the innermost radius at ~ 1.02 arcsec and the outermost radius at ~ 120.4 arcsec, corresponding to 0.34 and 40 kpc, respectively. These limits are set at the points where, when doing isophotal fits, one or more parameters need to be set to a constant value to get a reliable estimate of the SB values (see Kluge et al. 2020 for details). In order to check that the SB placed on the grid represents the observed isophotes well, we determine the isophote parameters using the SB on the grid and compare these to the observations, obtaining $\mathrm{RMS}_{\log \Sigma} = 0.006$, $\mathrm{RMS}_\varepsilon = 0.008$, and $\mathrm{RMS}_{\mathrm{PA}} = 0.602^\circ$. For the intrinsic density ρ, we choose an ellipsoidal $60 \times 11 \times 11$ grid. The flattenings P and Q of this grid are determined for each set of viewing angles calculating the expected values for a perfect ellipsoid (see Appendix A of de Zeeuw & Franx 1989 or equations 2–3 of Cappellari 2002) and averaging these, unless the deprojection is along the principal axes.

We sample the viewing angles in 10° step. Because of the assumption of triaxiality, θ, ϕ only need to be sampled in $[0, 90]^\circ$, while ψ in $[0, 180]^\circ$. As shown in de Nicola et al. (2020), this does not guarantee the canonical $1 \geq p \geq q$ relation. Therefore, for every selected ρ (see below), we repeat the deprojection using the set of viewing angles such that the new deprojection is equivalent to the old one but the inequality $1 \geq p \geq q$ holds (see also Section 3.2 of de Nicola et al. 2022b). Finally, every time the solution is along (or within 10° of) the principal angles, we test more deprojections at the same orientation to account for the degeneracy arising in these unfavourable cases (see de Nicola et al. 2020). For example, along the y-axis, we let the code recover $q(r)$, while for $p(r)$, we use different initial guesses, given that in this case, the recovery of $p(r)$ is hampered by the position of the LOS.

From all deprojections, we select those for which RMS $\leq 1.2 \times \mathrm{RMS}_{\min}$, where $\mathrm{RMS}_{\min} = 0.027$ is the best-fit RMS. The viewing angles that provide the best deprojection within the RMS interval are $(\theta, \phi, \psi) = (70, 20, 130)^\circ$. The resulting light densities are those we use for the dynamical modelling of the galaxy. Finally, as a check of the goodness of the deprojections, we compute, for the best-fit case, the RMS between the parameters of the fitted isophotes and those of the recovered ones, finding $\mathrm{RMS}_{\log \mathrm{SB}} = 0.017$, $\mathrm{RMS}_\varepsilon = 0.027$, and $\mathrm{RMS}_{\mathrm{PA}} = 1.84^\circ$. The isophotal fits to the best-fit SB, superimposed to the observed photometry, are shown in Fig. 2(a).

Finally, we repeat this exercise with the Ks-band photometry without masking anything and identify the best-fit viewing angles

$(\theta, \phi, \psi) = (77, 38, 141)^\circ$ for this case. These are slightly different from the g'-band deprojection given the twist in the central regions, which prevents the canonical $1 > p > q$ relation from holding at all radii. Fig. 2(b) shows the recovered photometry. Here, we see that the rise in ε cannot be entirely reproduced, as the recovered profile hits a ceiling at 0.1 close to the galaxy centre, similar to the deprojected ellipticity profile in g'-band. On the other hand, the twist is very well reproduced, meaning that it could be explained by triaxiality, and we do not necessarily need to assume that the galaxy has a substructure in the centre.

3 SPECTROSCOPY

In this section, we present the spectroscopic observations of NGC 708 and the extraction of the kinematics. Apart from providing the parameters to be fitted when performing the dynamical model, the kinematics is also important for a variety of reasons: it can provide evidence for triaxiality if the galaxy shows minor-axis rotation (Contopoulos 1956; Schechter & Gunn 1979; Davies & Birkinshaw 1986, 1988), the velocity dispersion σ can be used to get a prediction of M_{BH} using scaling relations (Ferrarese & Merritt 2000; Gebhardt et al. 2000; McConnell & Ma 2013; Saglia et al. 2016), while knowledge of the exact shape of the LOSVD is crucial to break the mass-anisotropy degeneracy, which can lead to biased M_{BH} (Binney & Mamon 1982). The few existing triaxial M_{BH} measurements in massive galaxies mostly used parametric LOSVDs (Gerhard 1993; van der Marel & Franx 1993), while we here use more flexible non-parametric LOSVDs.

3.1 MODS observations

Long-slit spectroscopic data for NGC 708 were already published by Wegner et al. (2012). These data come from observations carried out at the 2.4-m Hiltner telescope of the MDM Observatory at Kitt Peak. We re-observed the galaxy at the Large Binocular Array (LBT) observatory using the Multi-Object Double Spectrograph (MODS; Pogge et al. 2010). Its binocular configuration (MODS1 – MODS2) allows to place two slits at two different orientations in the plane of the sky. The observations were carried out in two runs, the first one in October 2019 (observers Jan Snigula and Stefano de Nicola) and the second one in October 2020 (remote observers Jan Snigula and Roberto Saglia), with PI Roberto Saglia. MODS has a field of view of 6×6 arcmin, works in the range [3200–10 000] Å and has spectral resolution $\lambda/\Delta\lambda = 5000/3.2 \sim 1500$ with slit width 0.8 arcsec in the blue. All science images were corrected for bias, dark, flat fields, and wavelength calibrated. Moreover, we acquired one sky image after each set up to allow for background subtraction. The pixel scale of each image is ~ 0.12 arcsec/pixel.

The galaxy was observed using two different configurations. First, the two slits, each one with width of 0.8 arcsec, were placed along the galaxy, projected principal axes on the plane of the sky (MAJOR and MINOR). The second configuration was obtained by rotating the slits by 45° (MAJOR+45, MINOR+45), as shown in Fig. 3. In both runs, we took data in the range 3200–8450 Å, splitting the spectra in a blue ($\lambda \in [3200$–$5700]$ Å) and a red ($\lambda \in [5700$–$8450]$ Å) part. The typical seeing is ~ 1.4 arcsec (FWHM). The relevant pieces of information of the four set ups for the blue part of the spectrum are reported in Table 2.

Because of the dust lane, the galaxy shows two apparent nuclei (see Fig. 4). While MAJOR and MAJOR+45 were centred on the brightest peak, MINOR and MINOR+45 were centred on the fainter

Table 2. Technical details about the four slit configurations that we used to observe NGC 708. *Col. 1:* Set up name. *Col. 2:* PA of the slit, measured from North to East. *Col. 3:* Centre of the slit coordinates (in arcseconds) in a NW frame of reference centred on the brightest peak. *Col. 4:* Number of spatial bins each slit is divided into. The details refer to the blue part of the spectrum since we did not use the red part in our fits (see text). *Col. 5:* Total exposure time in minutes adding up the spectra of the two runs together. All galaxies were observed for one hour in 2020, while MAJOR and MAJOR+45 were observed 10 arcmin longer in 2019.

Set up	PA (°)	Centre (arcsec)	Number of bins	Exp. time (arcmin)
MAJOR	35	(0, 0)	16	110
MINOR	125	(0.7369, 2.5913)	22	100
MAJOR+45	80	(0, 0)	15	110
MINOR+45	170	(0.7369, 2.5913)	23	100
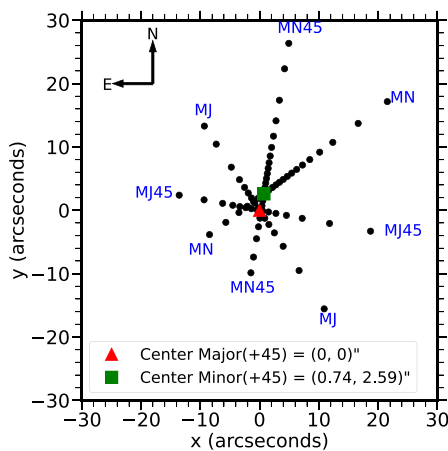
Figure 3. Illustration of the four slits configuration used throughout this work, whose centres are plotted as black dots. The 'true' centre of the galaxy is labelled with a red circle, while the green triangle shows where MINOR and MINOR+45 were centred.
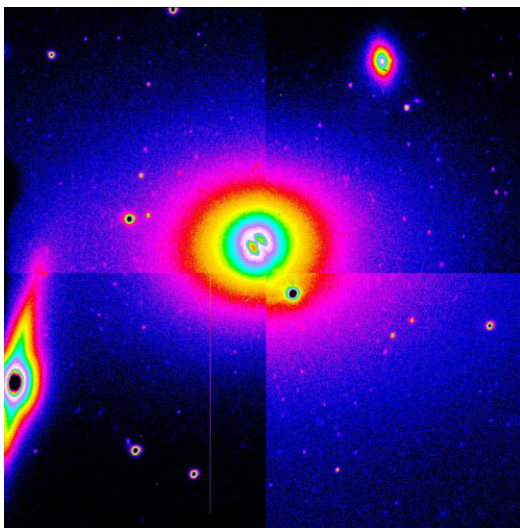
Figure 4. LBT acquisition image of NGC 708 before starting the spectroscopy. In the centre of the galaxy, there is a prominent 'double centre'; this effect is generated by the dust lane.

one. In a NW coordinate system centred on the brightest peak, the fainter has coordinates (0.7369, 2.5913) arcsec.

In order to achieve a nearly constant signal to noise of ~40 per pixel, for each one of the four set ups, we spatially binned the data (see Table 2) along the slit. In each spatial bin, we extracted the spectrum and measured the kinematics.

3.2 Extracting the kinematics

The extracted spectra show prominent emission lines: in the blue part, [O III] 5007 Å, [N I] 5199 Å, and Hβ are clearly visible. The red part of the spectrum is dominated by emission lines such as Hα, [N II] 6583 Å, and [S II] 6730 Å.

To extract the kinematic variables from the spectrum, we use the non-parametric LOSVD fitting routine WINGFIT (Thomas et al., in prep). The code exploits the full information contained in the LOSVD, allowing for accurate reconstruction of the wing profiles, a crucial ingredient to accurately recover M_{BH}. In fact, the well-known mass anisotropy degeneracy (Binney & Mamon 1982) can bias the measurement given that a radial anisotropy profile can also generate an increase in the velocity dispersion. The information contained in the wings can help lifting the degeneracy. The code uses a novel model selection technique to optimize the smoothing of the LOSVDs (Thomas & Lipka 2022). It is based on a generalization of the classical Akaike Information Criterion (AIC; Akaike 1974) and uses the concept of effective degrees of freedom (Lipka & Thomas 2021) to take into account the different number of DOFs each model has. Instead of minimizing the χ^2, the optimization is done in

$$\text{AIC}_p = \chi^2 + 2 * m_{\text{eff}}. \qquad (2)$$

In practice, equation (2) allows to find a model that fits the data well but is also smooth.

WINGFIT is inefficient when the number of templates is >50. To preselect a pool of template stars, we launch a preliminary fit using the parametric algorithm PPXF (Cappellari & Emsellem 2004; Cappellari 2017). The spectrum is fitted using all stars found in the MILES stellar library (Falcón-Barroso et al. 2011). Given the higher resolution of the library (FWHM ~ 2.5 Å), the stellar spectra are broadened to match the resolution of the observations (FWHM ~ 3.2 Å). We fit the blue spectrum in the range [4200–5400] Å. This allows us to fit the most prominent absorption features and exclude too noisy regions. We do not use the red part of the spectrum, which is dominated by emission lines and does not allow for a wavelength range large enough to obtain a reliable fit.

In PPXF, emission lines can be fitted along with the absorption features with separate templates, yielding two LOSVDs, one for the absorption and one for the emission lines. We tested two approaches obtaining comparable results: we first tried the fit as described above and then repeated the procedure fitting only the absorption features, masking the emission lines, obtaining similar results. One important caveat lies in the multiplicative polynomials. As shown in Mehrgan et al. (2023), their usage can generate artificially enhanced wings. Therefore, we limited ourselves to 4-th order multiplicative polynomials and avoided to use additive polynomials at all.

This procedure is repeated for every spatial bin. PPXF assigns weights to each template that it used to fit the galaxy spectrum. We take the best ~15–20 templates for each bin and pass these to WINGFIT to reconstruct the fully non-parametric LOSVD. An example of fitted spectrum along with its corresponding LOSVD is shown in Fig. 5.

The recession velocity of the galaxy barycentre needs to be subtracted from the actual estimate to set the zero-point of the
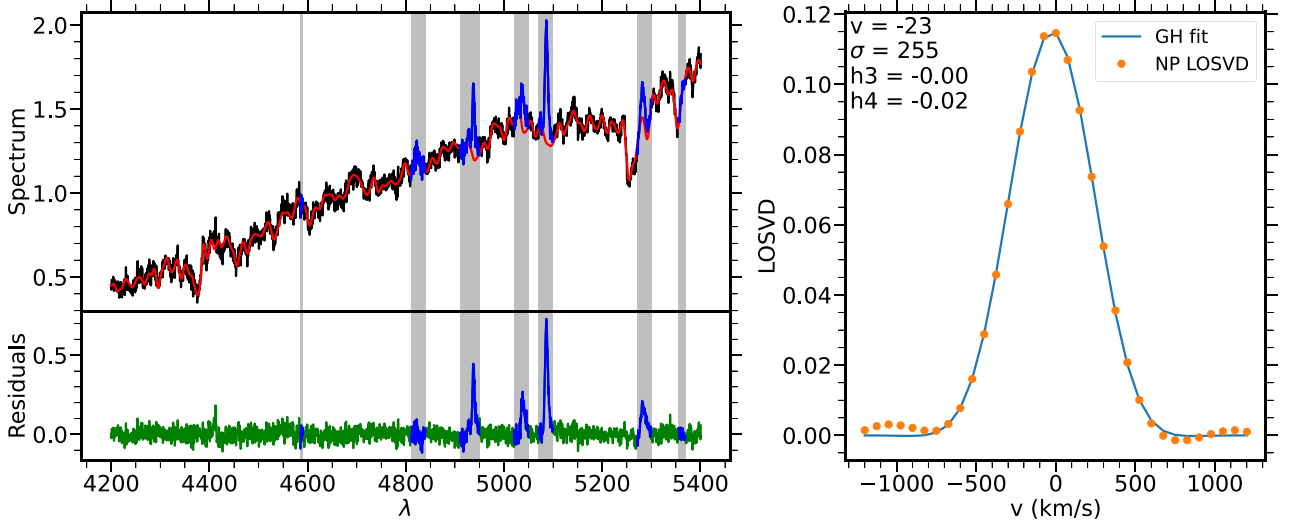
Figure 5. *Left:* Example of fitted spectrum by WINGFIT. The top panel shows the observed (black) and the fitted (red) spectrum, while the parts of the spectrum highlighted in blue and surrounded by grey rectangles are masked and omitted from the fit. Finally, the bottom panel shows the fit residuals. *Right:* Corresponding non-parametric LOSVD recovered by WINGFIT (orange points) along with the best-fit parametric LOSVD (solid blue line), whose coefficients are determined by fitting a 4-th order Gauss–Hermite polynomial to the recovered LOSVD. The bin belongs to the MAJOR slit (Table 2) and its coordinates are $(-0.8415, 1.2017)$ in a NW frame of reference.

systemic velocity at the galaxy centre. As first guess, we try $v = c \ln (1 + z)$, where z is the galaxy redshift. The fits are then repeated once more to fine-tune the velocity.

The resulting kinematics, parametrized in terms of standard Gauss–Hermite coefficients for illustration purposes, are shown in Fig. 6 as blue dots. This is in good agreement with the published values of Wegner et al. (2012). For the Schwarzschild fits, the LOSVDs are binned into $N_{\rm vel} = 25$ velocity bins. We discard bins with unrealistic Gauss–Hermite (GH) coefficients or spatial bins with too low S/N, retaining a total of 56 bins. We see that along the MINOR, there is a velocity variation $\Delta v \sim 100\,{\rm km\,s^{-1}}$, which is a clear indication of triaxiality. The fact that the rotation is only retrograde depends on the asymmetric position of the MINOR slit (see Section 3.1). The low h_4 values indicate that the wings are not very strong, while the slightly negative h_3 might indicate a small residual template mismatch (Bender, Saglia & Gerhard 1994; Mehrgan et al. 2023). Finally, the velocity dispersion σ hits a ceiling at $\sim 250\,{\rm km\,s^{-1}}$ in the central regions. This value is low compared to what is typically observed in large ETGs. Given that the core size of this galaxy predicts a black hole with mass $\sim 10^9\,{\rm M_\odot}$, while using the $M_{\rm BH}$–σ relation of Saglia et al. (2016), we infer a black hole mass smaller than $10^9\,{\rm M_\odot}$. Thus, the galaxy is a potential catastrophic outlier in the $M_{\rm BH}$–σ relation, similar to NGC1600 (Thomas et al. 2016).

4 DYNAMICAL MODELLING

We now put together the results of the previous two sections and turn to the dynamical modelling of NGC 708 under the assumption of triaxiality. The galaxy has already been modelled assuming it to be axisymmetric (see Wegner et al. 2012); the photometry does not show unambiguous signs of triaxiality. Nevertheless, several of our deprojections are strongly triaxial (see also de Nicola et al. 2022b), and even if the galaxy may be axisymmetric, triaxial solutions cannot

be excluded.[2] Moreover, the rotation along the minor axis at least excludes an oblate axisymmetric shape.

4.1 Our code

To compute the dynamical models of NGC 708, we use our recently developed triaxial Schwarzschild code SMART (Neureiter et al. 2021). SMART can deal with any deprojection or DM halo – parametric or non-parametric – and exploits a 5D orbital sampling space, allowing to characterize every orbit family (tubes, box orbits, Keplerian orbits), which may be found in a triaxial potential (Poon & Merritt 2001). Moreover, it uses the model selection technique described above (Lipka & Thomas 2021; Thomas & Lipka 2022) to optimize the smoothing individually in each trial mass model.

Specifically, SMART computes the potential by integrating Poisson's equation. The density is constructed as

$$\rho_{\rm TOT} = M_{\rm BH} \times \delta(r) + \Gamma \times \rho_* + \rho_{\rm DM}, \tag{3}$$

where $M_{\rm BH} \times \delta(r)$ is the point-like contribution from the central SMBH, $\Gamma \times \rho_*$ is the stellar contribution – obtained by multiplying the deprojected light density[3] with the mass-to-light ratio. While Γ is indeed the parameter that SMART fits, for reasons that will become clear later, we follow Mehrgan et al. (2024) and introduce the parameter Υ, computed at the minimum of $\Upsilon(r) = [\rho(r) + \rho_{\rm DM}(r)]/\rho_*$, where ρ is the stellar-mass density, obtained as $\rho_* \times \Gamma$. Finally, $\rho_{\rm DM}$ is the DM density, for which we assume a triaxial Zhao (1996) model

$$\rho_{\rm DM}(r) = \frac{\rho_0}{{\rm p_{DM}} \cdot {\rm q_{DM}} \cdot \left(\frac{r}{r_0}\right)^\gamma \left[1 + \left(\frac{r}{r_0}\right)^{1/\alpha}\right]^{(\beta-\gamma)/\alpha}}, \tag{4}$$

---

[2]An example is a triaxial galaxy with constant axis ratios as a function of the distance from the galaxy centre. In this case, a twist cannot be observed.
[3]Thus, in our code, the light density is not fitted, but treated as a constraint.
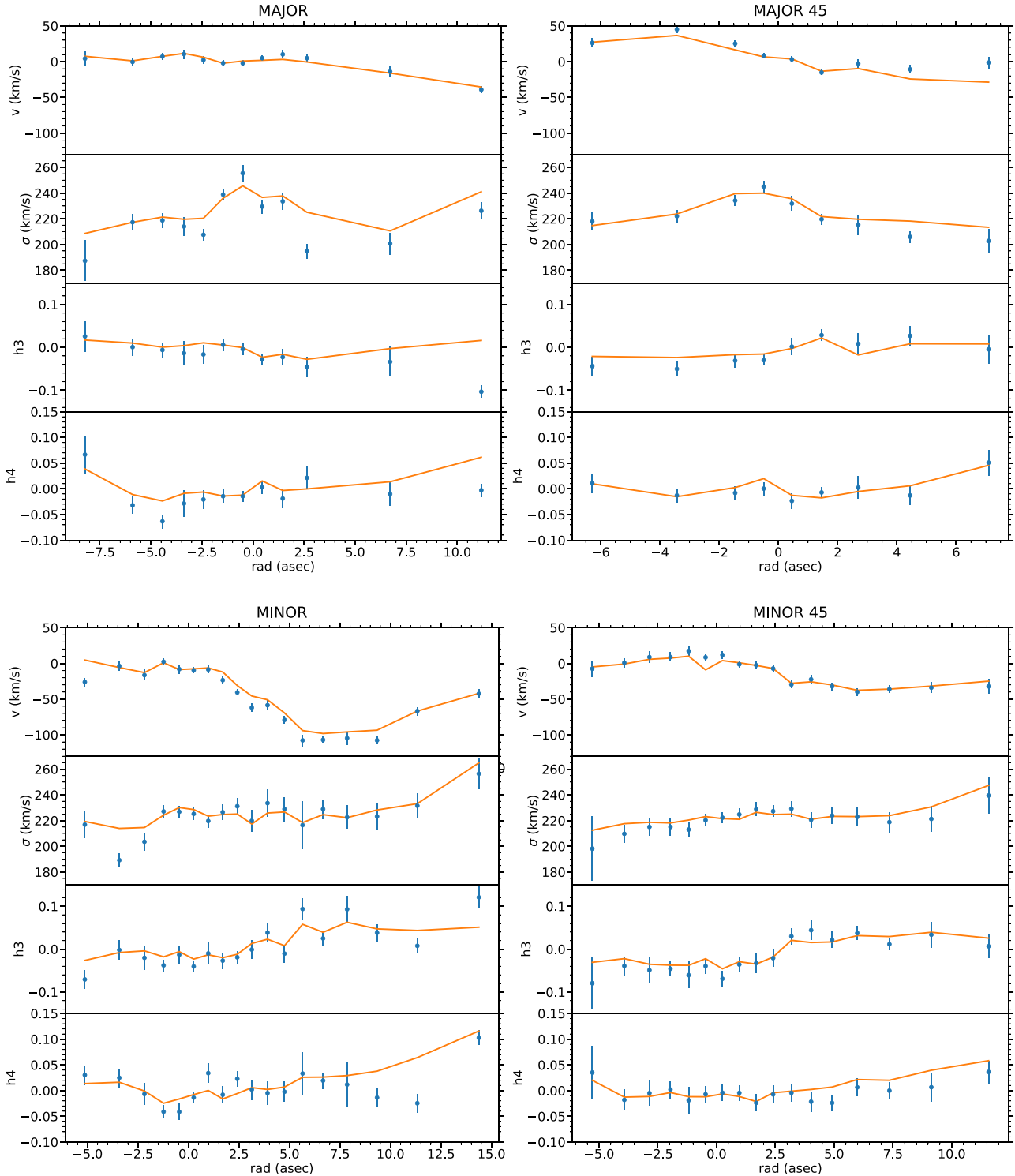
Figure 6. Kinematics of NGC 708 along the four different slits (see Table 2). Blue dots are the measured kinematics using WINGIFT, while orange lines show the result from the best-fit dynamical modelling using *g'*-band photometry. In particular, the minor-axis rotation is very well reconstructed. Radii values grow from East to West.

where p_{DM} and q_{DM} are the intrinsic flattenings of the DM halo, we fix $\alpha = 1$ and $r = \sqrt{x^2 + (y/p_{DM})^2 + (z/q_{DM})^2}$.

Once the potential has been calculated, a time-averaged orbit library is constructed, and time-averaged projections are computed

taking seeing into account. The stellar weights are then updated to maximize the entropy-like quantity

$$\hat{S} = S - \alpha \chi^2 \tag{5}$$

where S is related to the Shannon entropy, α is a smoothing term and χ^2 compares the differences between fitted and observed LOSVDs:

$$\chi^2 = \sum_{i=1}^{N_{\text{losvd}}} \sum_{j=1}^{N_{\text{vel}}} \left(\frac{\text{LOSVD}_{\text{model}}^{i,j} - \text{LOSVD}_{\text{data}}^{i,j}}{\Delta \text{LOSVD}_{\text{data}}^{i,j}} \right)^2 . \tag{6}$$

Here, the summations are carried over every i-th spatial bin and every j-th velocity bin (cfr. Section 3.1), while $\Delta \text{LOSVD}_{\text{data}}^{i,j}$ is the uncertainty on the i-th, j-th bin, yielded by WINGFIT. The smoothing optimization is defined in Thomas & Lipka (2022) and its implementation into SMART is described in Neureiter et al. (2023a).

SMART was tested by Neureiter et al. (2021) using an *N*-body simulation (Rantala et al. 2018), which aims at reconstructing the formation and evolution of massive core-galaxies. In particular, the simulation closely resembles NGC1600 (Thomas et al. 2016). When supplied with the 3D light and DM densities calculated from the particles, the code returned recoveries of M_{BH}, Υ and the normalization of the DM with unprecedented accuracy. Recently, Neureiter et al. (2023a) moved one step further and simulated the full modelling of a galaxy, deprojecting the SB profile using SHAPE3D and including the DM halo in the dynamical modelling, recovering all the mass parameters with an accuracy of 5–10 per cent. In a companion paper, de Nicola et al. (2022a) used the same approach to recover the intrinsic shape p, q, and the anisotropy β

$$\beta = 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2} \tag{7}$$

obtaining Δp, Δq, $\Delta \beta < 0.1$.

These results have been obtained using simulated high-resolution IFU kinematics. Given that the kinematics used in this work comes from long-slit data with seeing $\sim$1.4 arcsec, we repeated the same exercise, this time fitting a kinematic data set with the same geometry as that of NGC 708. The results are shown in Appendix B, where it can be seen that we still reach an accuracy of 10–20 per cent when recovering the mass parameters of the galaxy. Moreover, by fitting more than one mock kinematics, we can obtain a first-order estimate of the statistical uncertainties on the recovered mass parameters.

4.2 Modelling strategy

The parameters needed to build the trial gravitational potential (defined in equation (3)) are: M_{BH}, Γ, the halo normalization ρ_0, the inner and outer slopes γ and β and the break radius r_0. Moreover, since we use a triaxial halo, we also need to specify the two flattenings p_{DM} and q_{DM}. Finally, since we have several deprojections, the three viewing angles (θ, ϕ, ψ) also need to be varied for a total of 11 unknowns. As a preliminary step, we fix the break radius of the DM halo $r_0 = 50$ kpc and the outer slope $\beta = 4.5$. The value of β differs from the canonical value of 3 of the gNFW models: we assume that the halo progenitors can be modelled by a Hernquist (1990) sphere. Given that a large number of trial potentials must be tested, we use the software NOMAD (Audet & Dennis 2006; Le Digabel 2011; Audet & Warren 2017) to efficiently search for the global minimum by launching several models in parallel.

Our modelling strategy can be summarized as follows:

(i) We consider the *g'*-band deprojection launching a NOMAD run fixing the viewing angles at best-fit orientation $(\theta, \phi, \psi) = (70, 20, 130)^\circ$. The other six parameters are (log-)linearly sampled as reported in Table 3. With the exception of M_{BH}, which can be estimated using the BH-core size relation, we need to assume fiducial intervals. All the models are constructed by assuming a Gaussian PSF with FWHM

Table 3. Sampling we used for our NOMAD runs. *Col. 1:* Variable name. *Col. 2:* Sampled range (linear spacing). *Col. 3:* Number of sampled values. *Col. 4:* Step size.

Variable	Range	No. values	Step size
M_{BH}	$[4 \times 10^9$–$1.2 \times 10^{10}]\,\text{M}_\odot$	5	$4 \times 10^9\,\text{M}_\odot$
Γ	$[0.6$–$4.8]$	7	0.7
$\log \rho_0$	$[7.8$-$8.2]$	7	0.0667
γ	$[0.0$–$1.0]$	6	0.2
p_{DM}	$[0.8$–$1.0]$	3	0.1
q_{DM}	$[0.8$–$1.0]$	3	0.1

of 1.4 arcsec, measured from our spectroscopic observations. Given that the size of the central core is expected to approximate the sphere-of-influence radius r_{SOI} well and given that we model with a DM halo, we conclude that we can reliably estimate M_{BH} (see discussion in Rusli et al. 2013a). We verify this assumption below.

(ii) We fix the mass parameters to the values coming from step 1) and launch a second NOMAD run, this time *only* fitting for the viewing angles. For all the orientation along (or close to) the principal axes, we sample several deprojections to account for the degeneracy (see Section 4.3).

(iii) We launch a third and final NOMAD run fixing the viewing angles at the best-fit orientation found in step 2), sampling the mass parameters as in step 1) to determine our final estimate of the best-fit model.

(iv) Given that here we model all bins together, in order to get an estimate of the error bars on our best-fit estimates, we rely on the values we find in Appendix B for the *N*-body simulation. In that case, we have a total of nine configurations, so we take the standard deviation as (pessimistic) estimate of the error bars, scaled to the values we find for NGC 708.

(v) To evaluate the impact of the dust, we launch a further NOMAD run using the best-fit *K*s-band deprojection at $(\theta, \phi, \psi) = (77, 38, 141)^\circ$. Given that the photometry sets and the deprojections are similar in the outermost regions, we take only M_{BH} and Γ as free parameters while keeping the halo parameters fixed. As second test, we deproject the *g'*-photometry with masking at the best fit *K*s-band viewing angles and model this with fixed halo parameters. In this way, we can verify how much the BH mass remains stable and compare Γ estimates in different bands.

In de Nicola et al. (2022a), we have shown that by selecting deprojections using the approach described in Section 2.3 and computing the average shape profiles $p(r)$ and $q(r)$ among all these, the resulting profiles approximate the correct ones very well. The modelling strategy adopted here allows us to test this *a posteriori* by comparing the deprojection corresponding to the best-fit viewing angles found in step 2) with $<p(r)>$ and $<q(r)>$, where the average is performed over all selected deprojections.

4.3 Results

The shapes of the best-fit *g'*- and *K*s-band densities are shown in Fig. 7, where we leave out the dust-contaminated central regions of the galaxy for the *g'*-band. In the optical, the density is prolate in the central regions and then becomes triaxial at larger radii. The viewing angles for this density coming out from the second NOMAD run are $(\theta, \phi, \psi) = (80, 90, 135)^\circ$: for this orientation, close to the *y*-axis, we sampled three different deprojections using as starting
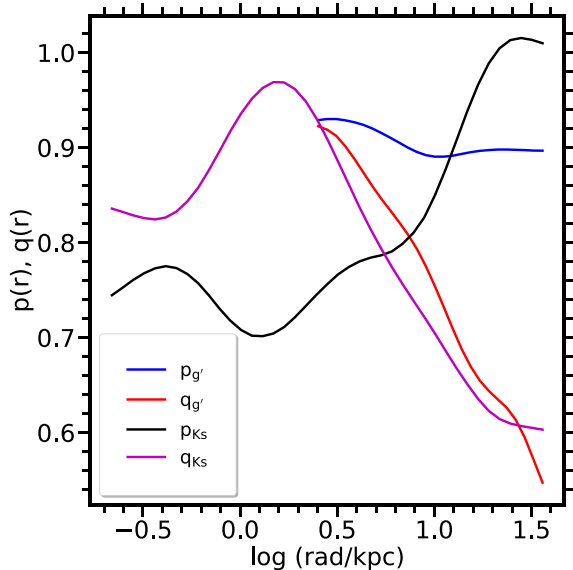
Figure 7. The recovered intrinsic shapes for the best-fit solutions found at $(\theta, \phi, \psi) = (80, 90, 135)°$ for the *g*'-band and at $(\theta, \phi, \psi) = (77, 38, 141)°$ for the *K*s-band. The innermost regions are not shown for the *g*'-band due to dust contamination; in the outermost region the galaxy becomes mildly triaxial.

values $p(r) = 1.0, 0.95, 0.9$ at all radii.[4] The last value provided the best-fit density. This agrees well with the findings of de Nicola et al. (2022b): the $p(r)$, $q(r)$, $T(r)$ intervals shown in fig. 9 in that paper overlap well with those found in this work. It should be noted that the best-fit solution does not appear in de Nicola et al. (2022b) because there only one deprojection[5] per orientation was tested. Instead, the recovered shape profile in *K*s-band intersect with each other in the central regions, which is expected given the large isophote twist. It is reassuring that the two $q(r)$ profiles agree nicely with each other, while for $p(r)$ deviations of ~ 0.1 are observed. Nevertheless, such scatter is expected (see de Nicola et al. 2022a).

The results of the third and final *g*'-band NOMAD run are shown in Fig. 8, for a total of 1617 models. For each variable, we marginalize the six-dimensional AIC_p distribution to recover the six one-dimensional functions $AIC_p(M_{BH})$, $AIC_p(\Gamma)$, $AIC_p(\rho_0)$, $AIC_p(\gamma)$, $AIC_p(p_{DM})$, and $AIC_p(q_{DM})$. The fit to the kinematics is shown in Fig. 6 as orange lines. Here, we see that the code is able to recover all four kinematic variables well;[6] in particular, the minor-axis rotation is well reproduced. Moreover, in Fig. 9, we show an example of a fit to an individual LOSVD, showing that SMART reproduces it very well: in this case, we have $\chi^2/N_{data} = 0.73$.

The most relevant result is the detection of a $(1.0 \pm 0.28) \times 10^{10}\,M_\odot$ SMBH in the centre of NGC 708. We calculate the size of its SOI in three ways: using the central velocity dispersion σ_0 as $r_{SOI} = GM_{BH}/\sigma_0^2$ and using the stellar mass derived from our models as $M_{tot}(r_{SOI}) = 2\,M_{BH}$ (Merritt 2004) and as $M_{tot}(r_{SOI}) = M_{BH}$ (Thomas et al. 2016). The Merritt (2004) definition is equivalent to the 'velocity dispersion' one if the mass density profile of the

galaxy can be described by a singular isothermal sphere (SIS). Using the three definitions, we find $r_{SOI} = 0.73, 1.83$, and 1.21 kpc, respectively.[7] The strong inconsistency points out that the SIS model does not work well for this galaxy.

The derived $\Gamma = 3.4 \pm 0.51$ is lower than the expected value for a Kroupa IMF: indeed, Wegner et al. (2012) found 4.17 ± 1.05 in the Kron–Cousins *R*-band for a Kroupa (2001) IMF using SSP models (Maraston 2005). We discuss this further in Section 5.3.

The DM halo has a normalization $\rho_0 = 10^{8.06}\,M_\odot\,kpc^{-3}$ and, in the central regions, is less steep ($\gamma = 0.2$) than predicted by a NFW model ($\gamma = 1$). Finally, we note that DM halo is oblate ($p_{DM} = 1.0$, $q_{DM} = 0.8$).

Finally, the upper panels of Fig. 10 show the results of the NOMAD run with fixed halo, using the best-fit *K*s-band deprojection. Here, the best-fit M_{BH} value is only 20 per cent smaller, which is well within our error bars. Furthermore, the value of $\Gamma_K = 0.4$ is also lower than the prediction assuming a Kroupa IMF, as we also found for the *g*'-band deprojection. In the bottom panels of Fig. 10, we show what we get when modelling, again with fixed halo, the *g*'-band photometry at the best-fit *K*s-band orientation: the results highlight that the black hole mass estimate also becomes 20 per cent smaller, while the Γ estimate remains the same. Given the uncertainty values, which we estimated using the *N*-body simulation (Table B1), we conclude that our M_{BH} estimate is robust.

5 DISCUSSION

We have presented the full data analysis of NGC 708, BGC of A262. This is the second BCG, which we modelled using our full non-parametric triaxial pipeline (SHAPE3D + WINGIFT + SMART): another work taking on dynamical modelling of a triaxial ETG is Neureiter et al. (2023b). In what follows, we discuss our results in more detail.

5.1 Intrinsic shape

As shown in de Nicola et al. (2022a), the deprojection alone suffices if one wants to recover the correct intrinsic shape of the galaxy: we need to compute the average shape profiles among all deprojections,[8] which we select for the dynamical modelling and look for the deprojection that is closest to it. Instead, the dynamical modelling is needed to reduce the scatter of this estimate.

Here, we tested this approach with a real galaxy: after obtaining a first estimate of the mass parameters, we used these to estimate the best-fit viewing angles and thus the best-fit deprojection, using more than one density for the degenerate cases. The best-fit solution, found at $(\theta, \phi, \psi) = (80, 90, 135)°$, is indeed remarkably close to the average $<p(r)>$, $<q(r)>$ profiles (see Fig. 11). We notice that $p(r)$ mostly oscillates around the grid flattening $P = 0.9$, which is expected given the orientation close to the *y*-axis. Instead, in this case, $q(r)$ is very close to the observed flattening $q' \equiv 1 - \varepsilon$: it starts decreasing at $r \sim 2.5$ kpc, as expected given the ε profile shown in Fig. 2, reaching $q(r) < 0.6$ at large radii. The corresponding triaxiality shows a monotonically decreasing profile, reaching the

---

[4]Note that also in this case the code *does* fit $p(r)$. However, given the orientation, the fitted profile is expected to be close to the initial value.

[5]For the line of sight along the *y*-axis, we tested an almost oblate projection with $p(r) = 0.95$ at all radii.

[6]Note the SMART fits the entire LOSVDs: Gauss–Hermite coefficients are only used for illustration.

[7]These values correspond to 2.20, 5.51, and 3.64 arcseconds, respectively.

[8]Intrinsic shape profiles, which cannot be fitted because of deprojections performed at orientations along or close to the principal axes, are not considered. For example, for the deprojection along the *y*-axis, we do not consider the $p(r)$ profile.
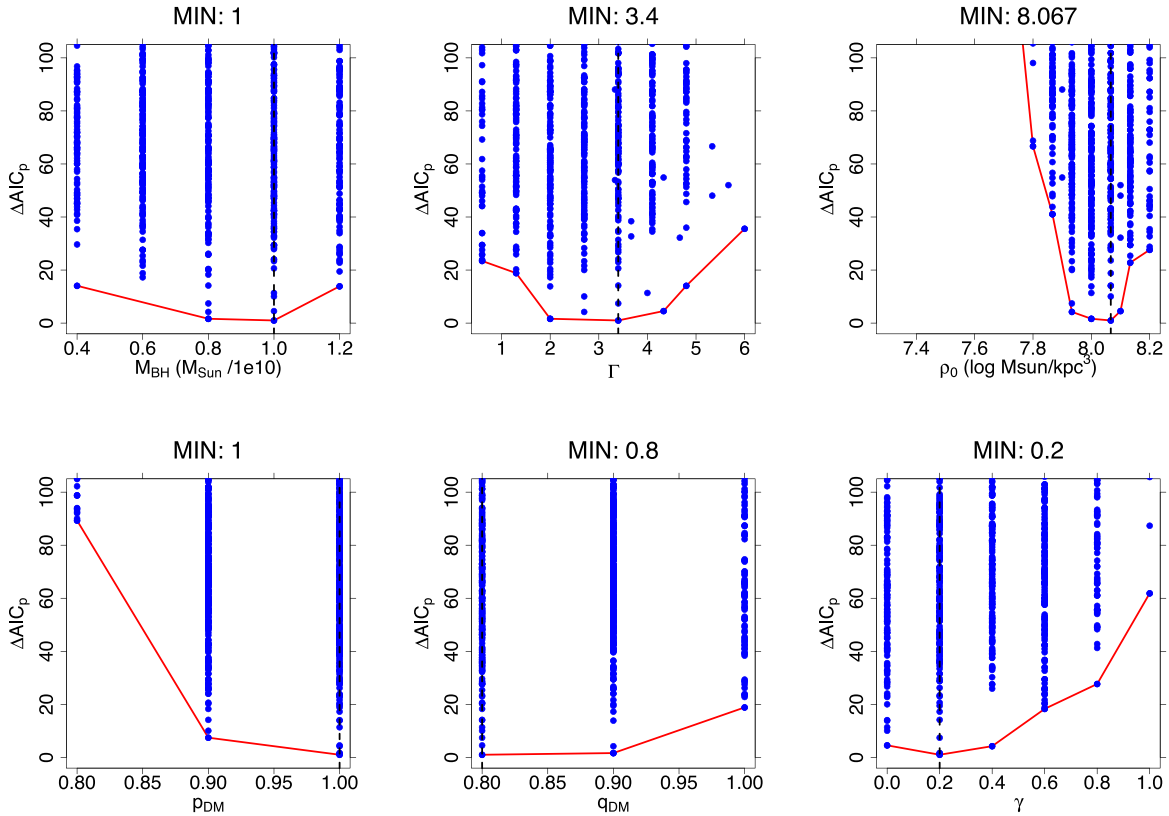
Figure 8. AIC$_p$ values, calculated following equation (2), plotted against the six variables fitted in our final NOMAD run. The blue points are the individual models. The majority of them are not shown as they fall outside the plotted range. The black dashed line labels the best-fitting value, while the red line follows the best-fit models for each tested value. *Left to right, top to bottom: M_{BH}, Γ, ρ_0, γ, p_{DM}, q_{DM}.*
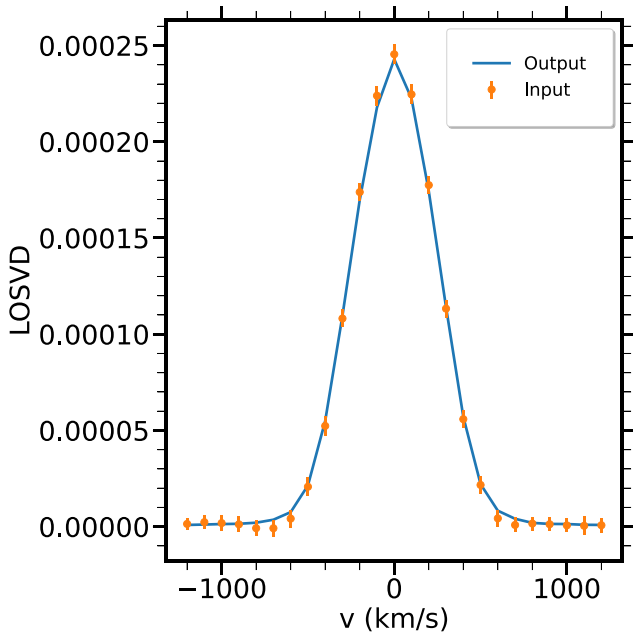

Figure 9. Example of SMART fit to a single LOSVD. The orange points are the input velocity bins, i.e. those provided by our code for kinematics extraction WINGFIT, while the light blue line is the recovered LOSVD. The bin is the same as that shown in Fig. 5, i.e. located at $(x, y) = (-0.8415, 1.2017)$ on the plane of the sky and belonging to the MAJOR slit (Table 2).

maximum triaxiality $T = 0.5$ at $\sim 10\,\mathrm{kpc}$. Averaging over all radii, we find $<T> =(1 - <p(r)>^2)/(1 - <q(r)>^2) = 0.455$.

5.2 The black hole: scaling relations, anisotropy, and orbital structure

Our dynamical modelling delivers a BH mass of $(1.0 \pm 0.28) \times 10^{10}\,\mathrm{M_\odot}$. Even using the smallest r_{SOI} value among the three we report in Section 4.3, we would still have 18 kinematical bins inside the BH SOI: this, along with the fact that we included a DM halo in our models, confirms that our estimate is secure. Objects with masses in this range are rare: there is an almost empty region between the most massive SMBH dynamically detected (Mehrgan et al. 2019) and SMBHs with masses $<10^{10}\,\mathrm{M_\odot}$. Measurements in this region include NGC4889 ($M_{BH} = 2.1 \times 10^{10}\,\mathrm{M_\odot}$, McConnell et al. 2012), NGC1600 ($M_{BH} = 1.7 \times 10^{10}\,\mathrm{M_\odot}$, Thomas et al. 2016), and the recently published combined mass ($M_{BH} = 1.0 \times 10^{10}\,\mathrm{M_\odot}$) of the two SMBHs of NGC5419 using triaxial models (Neureiter et al. 2023b).

At the very high-mass end, the scaling relation between SMBHs and the velocity dispersion of the host bulge σ saturates. This is linked to the evolution history of these galaxy: massive ETGs accrete mass through gas-poor ('dry') mergers which do not significantly alter σ (Naab, Johansson & Ostriker 2009) and, hence, generate M_{BH} values which are higher than the prediction of the canonical M_{BH}–σ relation, for which typically $M_{BH} \propto \sigma^{5.2 \div 5.4}$ (Saglia et al. 2016; van den Bosch 2016). This is expected: SMBHs correlate with the bulge parameters (Saglia et al. 2016; de Nicola, Marconi & Longo 2019), which are
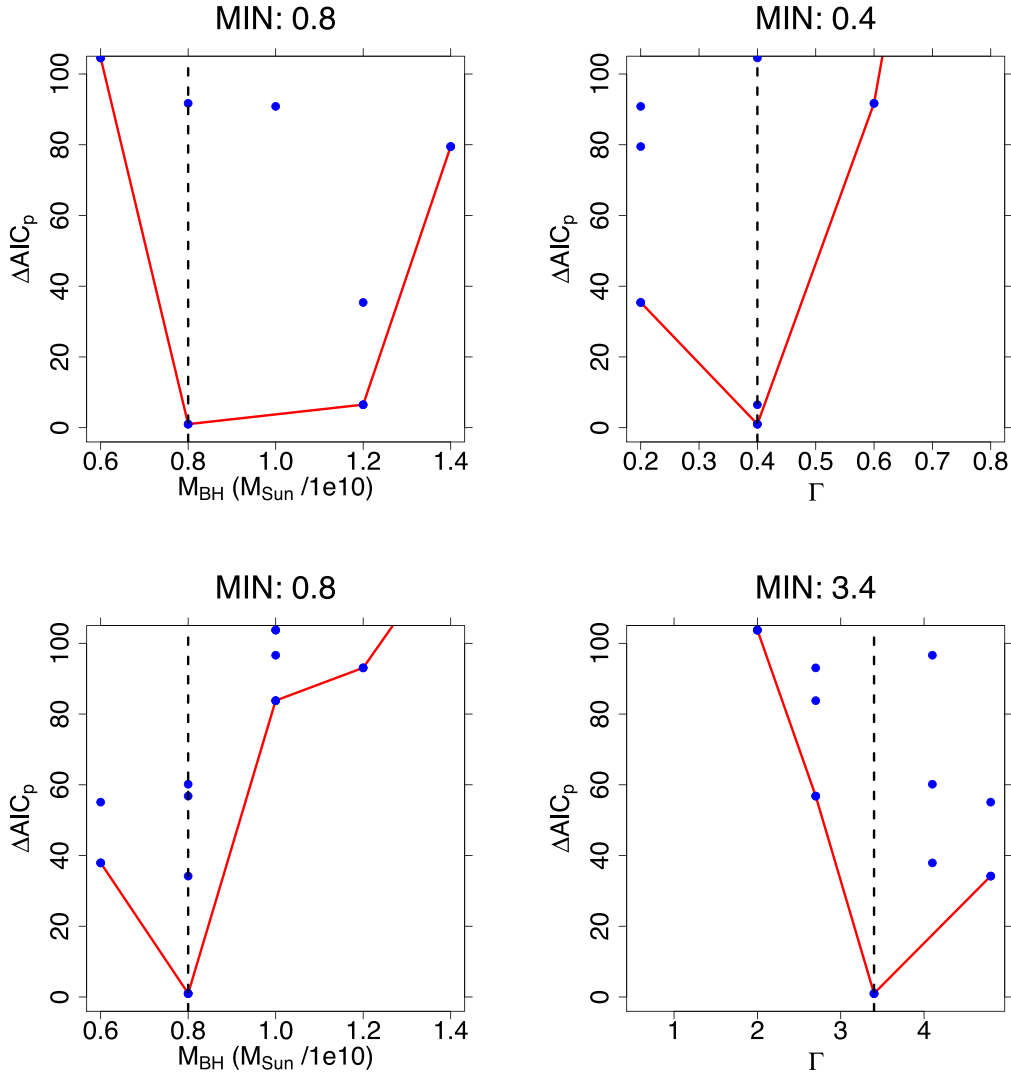
Figure 10. Dynamical estimates of BH mass and M/L in *K*s-band (top panels) and in *g'*-band (bottom panels) using the best-fit viewing angles found by deprojecting the Wendelstein *K*s-band photometry. In both cases, the BH mass gets lower by 20 per cent, well within our error bars. Both M/L values are too low with respect to SSP values assuming a Kroupa IMF, but this can be explained assuming that the DM halo traces the stars.

locked together through the Fundamental Plane (FP; Djorgovski & Davis 1987) and, for BCGs, different Faber & Jackson (1976) (FJ) and FP relations with respect to ordinary ETGs are found (Kluge et al. 2020).

Regardless of which $M_{BH}-\sigma$ we assume, NGC 708 is an outlier (see Fig. 12a). The coefficients found by Saglia et al. (2016) omitting pseudobulges predict $M_{BH} = 5.75 \times 10^8 M_\odot$, making the galaxy an outlier *by almost a factor of 20*. We note that NGC 708 has a low velocity dispersion in comparison with most BCGs. This stays roughly constant at all radii, peaking at 250 km s^{-1} in the most central bins.

The same conclusions apply by considering the $M_{BH}-M_{bul}$ relation (Magorrian et al. 1998). As estimate of M_{bul}, we take the stellar mass profile, computed using the deprojected density multiplied by Υ, up to the largest radius used for the deprojection, yielding $M_{bul} = 2.8 \times 10^{11} M_\odot$. Using the coefficients of Saglia et al. (2016), again omitting pseudobulges, the galaxy is an outlier by a factor 10.5, while using the relation of Bogdán (2018), the galaxy is a factor 10.6 off. Thus, the galaxy is not only an extreme case among the galaxy

population, but also among core galaxies only.[9] Note that since we did not try to subtract a potential ICL component from the light of NGC 708 our M_{bul} are upper limits and the offsets from the relations might be even larger.

The commonly proposed formation mechanism for these central cores is the gravitational slingshot caused by SMBHs lying at the centre of the progenitors and forming a binary after the merging process. This phenomenon ejects stars, causing a light deficit in the central regions (i.e. the core, Ebisuzaki et al. 1991), and in dry mergers there is no gas that can replenish the centre. Therefore, scaling relations linking M_{BH} to the core properties are theoretically expected and have indeed been observed; these include a correlation with the missing mass (Kormendy & Bender 2009), with the core size r_γ (Rusli et al. 2013b; Thomas et al. 2016) and with the central surface

---

[9]Note that all scaling relations have intrinsic scatters. Using the (large) intrinsic scatter of the $M_{BH}-M_{bul}$ relation, $\varepsilon = 0.61$ from Bogdán (2018), NGC 708 remains an outlier by a factor of 2.6.
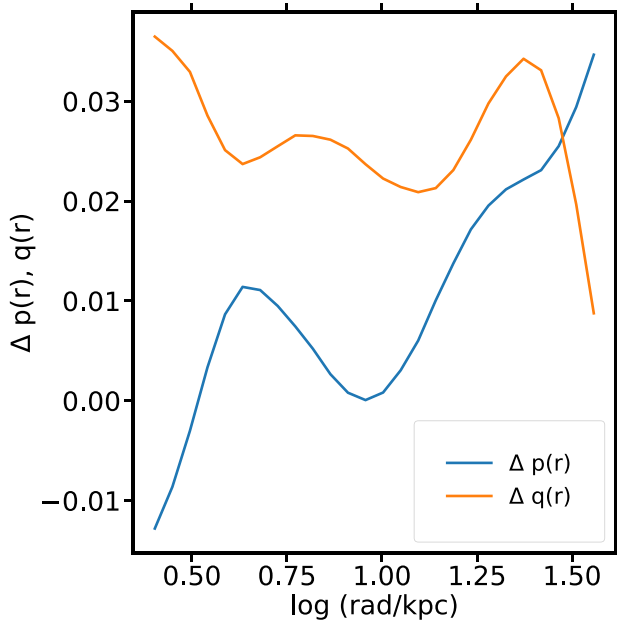
Figure 11. Differences between the intrinsic shape profiles $p(r)$ (blue) and $q(r)$ (orange) for the best-fit light density ρ found by NOMAD and the average shape profiles $<p(r)>$ and $<q(r)>$ among all deprojections selected for the dynamical modelling. The very small residuals confirm the findings of de Nicola et al. (2022a): only one deprojection is needed for the dynamical modelling.

brightness SB_0 of the core itself (Mehrgan et al. 2019). Indeed, the coefficients of Thomas et al. (2016) for the $M_{BH}-r_\gamma$ and those of Mehrgan et al. (2019) for the M_{BH}–SB relations[10] predict a BH with mass $\sim 1.5 \times 10^{10}$, a much better prediction for the M_{BH} value found in this work (see also Fig. 12b).

Another good predictor of M_{BH} for BCGs is the virial temperature kT (i.e. the total gravitating mass) of the host cluster (Bogdán et al. 2018). This happens because the mass drives the accretion onto the SMBH itself. Indeed, using the kT value of 2.25 ± 0.04 keV for A262 (Ilić, Blanchard & Douspis 2015) yields $M_{BH} = 6.5 \times 10^9$, also a reasonable estimate of the actual M_{BH} value.

The core scouring mechanism generates a tangential anisotropy ($\beta < 0$, equation (7)) in the central regions, because stars on radial orbits come closer to the SMBH and are more likely to be ejected. In the top panel of Fig. 13, we plot the anisotropy profile $\beta(r)$ for the best-fit model. Even if the profile shows the typical negative β, the less pronounced negative anisotropy with respect to other BCGs hints at a merger between two core-galaxies as a formation scenario for NGC 708. In the future, we will compare the derived $\beta(r)$-profile with simulations that reproduce the formation of cores in ETGs (Rantala et al. 2019; Frigo et al. 2021).

Instead, the bottom panel of Fig. 13 shows the corresponding orbit distribution, where we see that tubes dominate over box orbits. In particular, while in the innermost regions with tangential anisotropy, we find a similar amount of x- and z-tubes, orbits rotating around the minor axis take over at larger radii. This z-tubes dominance agrees with $p(r) \sim 0.9$, which indicates a galaxy shape close to an oblate spheroid, for which x-tubes cannot be observed at all.

---

[10]This relation was derived in the V-band, which is close to the g'-band photometry used in this work none the less.

5.3 Stars, DM, and gas

Mass-to-light ratio estimates for NGC 708 have been published by Wegner et al. (2012) by means of stellar population analysis. In that paper, the authors use the SSP method from Maraston (2005) to derive the mass-to-light ratio, assuming a Kroupa IMF. According to this value, we would need Γ to be *at least* 4.6 for a Kroupa IMF, assuming an old stellar population with twice solar metallicity. In a recent study (Mehrgan et al. 2024), the authors suggest to evaluate the mass-to-light ratio using the *total* mass-to-light profile, i.e. by using the parameter Υ introduced in Section 4.1. This is particularly relevant in this case because of the possible degeneracy between stars and DM: the dynamical mass-to-light ratio estimations assume that the whole mass comes from stars in the galaxy, while it may well be that the DM traces the stars and reduces M/L. Given that we do find a massive halo, we choose to use Υ rather than Γ to estimate M/L, showing the total mass-to-light profile in the upper panel of Fig. 14. We see that the profile shows a clear minimum, signaling the region, where stars dominate the potential at 2.7 kpc. Evaluating Υ then yields a g'-band mass-to-light ratio of 5.1, in good agreement with SSP models assuming a Kroupa IMF and with the findings of Wegner et al. (2012). A second consistency check is given by the Ks-band mass-to-light ratio Υ_K: as in g'-band, by evaluating Υ_K at the minimum of the M_{tot}/L curve (found at 2.4 kpc, see bottom panel of Fig. 14), we find a value to 1.1, in good agreement with the estimates in R and g'-bands, which would foresee $\Upsilon_K \in [1.0–1.5]$ for a Kroupa IMF and an old, metal-rich stellar population.

Following Mehrgan et al. (2024), we define an IMF mismatch parameter with respect to a Kroupa IMF as $\alpha = \Upsilon/\Upsilon_{Kroupa}^{SSP}$, which is known to correlate with σ (e.g. equation 6 of Posacki et al. 2015) and show in Fig. 15, the (α, σ) relation for the nine ETGs with MUSE kinematics from Mehrgan et al. (2024) and NGC 708 and the best-fit parabola of Posacki et al. (2015) obtained using SLACS + ATLAS^{3D} data. Both α and σ are evaluated at the radius where the M_{tot}/L profile for that particular galaxy has its minimum. This is called r_{main} in Mehrgan et al. (2024). While most ETGs fall below the best-fit curve, indicating that at high dispersion IMFs are more Kroupa-like,[11] the dispersion of NGC 708 is so low that the galaxy does fall within the scatter of the Posacki et al. (2015) relation.

Finally, from the best-fit potential yielded by SMART, we can reconstruct the rotation curve and compare it to the results of Wegner et al. (2012). While we did find a deprojected velocity profile agreeing with the circular velocity measured from our models, the inclination of the gas disc needed to match our results does not place it on any of the principal planes (x, y), (x, z), (y, z), stressing the need of IFU data to better clarify this issue.

6 SUMMARY AND CONCLUSIONS

We have obtained dynamical models of NGC 708, BCG of A262, combining high-resolution *HST* images with deep-photometry data taken with the WWFI at the Wendelstein observatory and using long-slit spectroscopy acquired with the MODS instrument at LBT. The analysis is done using a fully non-parametric triaxial pipeline, combining our deprojection code SHAPE3D, our code for kinematics

---

[11]Mehrgan et al. (2024) found evidence for mass-to-light gradients in eight of the nine galaxies they analysed, reporting central regions with Salpeter-like IMF. Nevertheless, these gradients are so concentrated that already beyond 2 kpc the IMF becomes Kroupa-like.
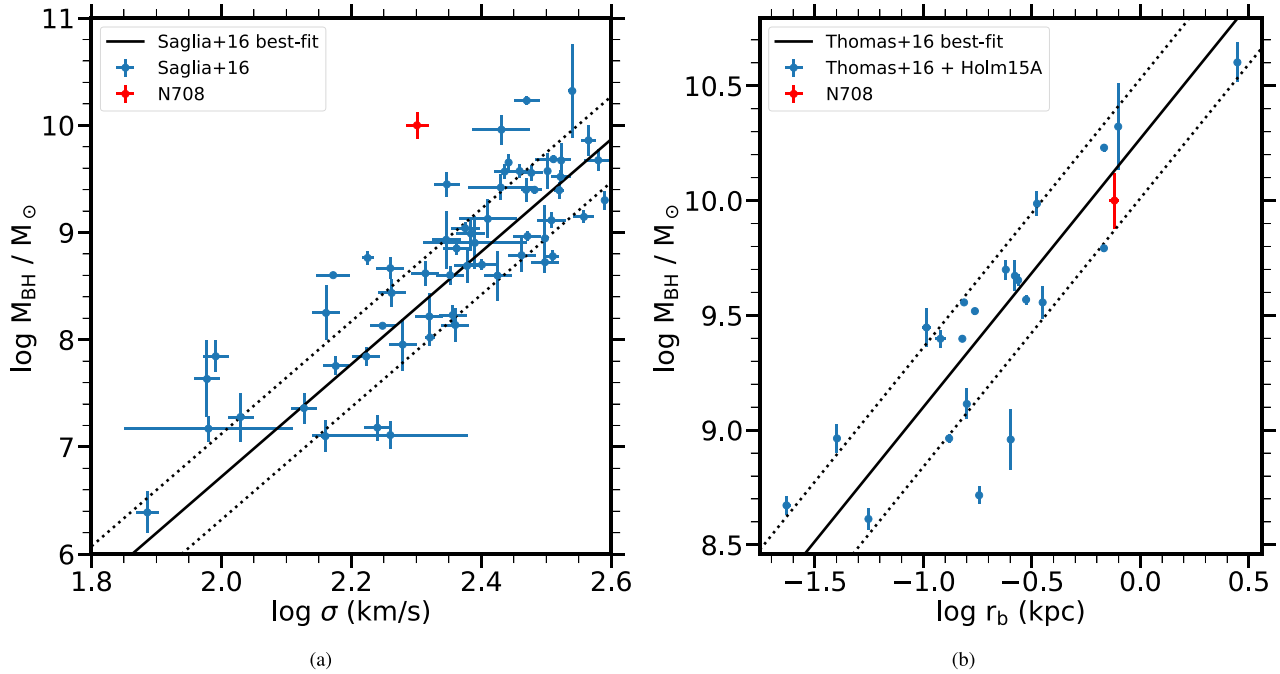
(a)

(b)

Figure 12. *Left:* M_{BH}–σ relation for ETGs; the grey line is the best-fit line relation for ETGs from Saglia et al. (2016). *Right:* M_{BH}–r_b relation with best-fit line from Thomas et al. (2016). We also add the estimate of Holm15A (Mehrgan et al. 2019). Here, the size of error bars on r_b estimates is too small to be clearly visible in the plot. In both plots, NGC 708 is shown as a red dot, while the dotted lines enclose the intrinsic scatters of the linear relations.
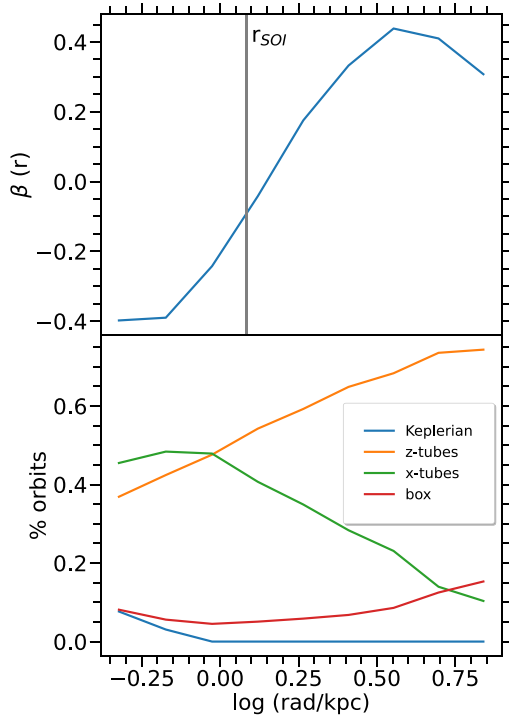
Figure 13. *Top:* The recovered anisotropy of NGC 708 for the best-fit model. We find tangential anisotropy in the central regions. This is a fingerprint of BH core scouring and therefore provides evidence for past merger(s). The fact that the tangential bias is not so pronounced as in most core-galaxy might indicate that the progenitors of NGC 708 were core-galaxies themselves. At larger radii, the galaxy shows a small radial anisotropy. The core radius computed using the definition of Thomas et al. (2016) is also plotted. *Bottom:* Orbit distribution for the best-fit model. *x*- and *z*-tubes dominate in the central region with tangential anisotropy, while *z*-tubes alone take over at large radii.
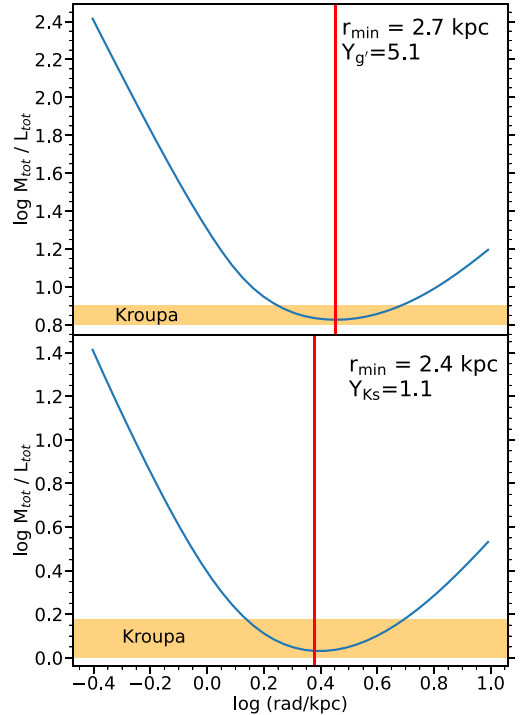
Figure 14. The total mass-to-light profile for the best-fit *g'*-band model (top) and *K*s-band model (bottom) of NGC 708. The value at the minimum is what we use to derive an estimate of Υ_* under the assumption that the halo partially traces stars. The shaded regions show the expected values range under the assumption of Kroupa IMF in order to match the *R*-band value of Wegner et al. (2012).
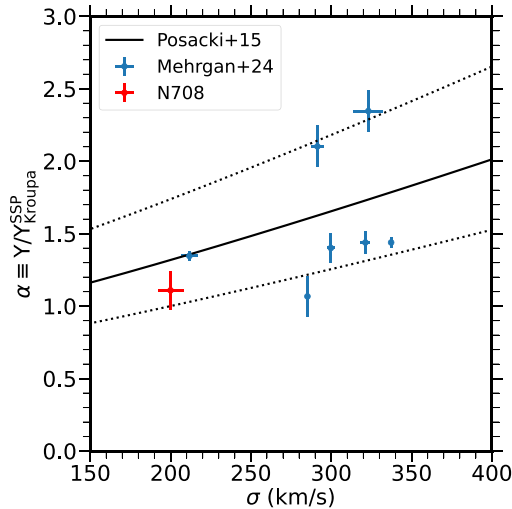
Figure 15. Similar as in Fig. 12 for the (α, σ) relation evaluated at the radius $r_{\rm main}$ for the nine ETGs of Mehrgan et al. (2024) with the best-fit quadratic relation and its scatter of Posacki et al. (2015). NGC 708 is shown as a red dot; despite having a Kroupa-like IMF, the dispersion is so low that the galaxy lies within the scatter of Posacki et al. (2015)'s relation.

extraction WINGFIT and, finally, our triaxial Schwarzschild code SMART.

The galaxy shows several interesting features. We detect a SMBH with mass $(1.0 \pm 0.28) \times 10^{10}\,{\rm M}_\odot$, one of the few measurements in this mass range, which makes the galaxy a strong outlier in the classical $M_{\rm BH}$–σ and $M_{\rm BH}$–$M_{\rm bul}$ relation, but not in all scaling relations linking $M_{\rm BH}$ with core properties (central SB, core size) when intrinsic scatters of these relations are taken into account. In these respects, the galaxy resembles NGC1600 (Thomas et al. 2016).

The typical anisotropy profile of cored galaxies, tangential inside the core and then radial at larger radii, is found. Nevertheless, the value of β inside the core is typical of galaxy mergers where the two progenitors are, themselves, core-galaxies. This is similar to what it has been observed for Holm15A, which is also an outlier in the scaling relations (Mehrgan et al. 2019). For NGC 708, we find a g'-band Υ_* value of 3.4 ± 0.51, which hints at an uncharacteristic lightweight IMF. However, this seems to be due to a degeneracy in the total mass profile; computing this value at the minimum of the curve yields an estimate in agreement with a Kroupa IMF.

We find that the galaxy is observed close to the intermediate axis. Given that the galaxy shows a small isophotal twist as well as minor-axis rotation, it must be triaxial, but the large amount of z-tubes found at large radii suggests that the triaxiality is not so pronounced. This is what we get when looking at the best-fit shape from the dynamical model: the shape of the galaxy is indeed remarkably close to an oblate spheroid, especially in the outermost regions. Interestingly, also our best-fit DM halo shows a similar geometry. Finally, we also attempt to derive the inclination and the PA of the gas disc by comparing its kinematics with our dynamical modelling-derived rotation curve, finding that the gas disc is not on any of the principal planes, even if the lack of IFU data makes it difficult to say more.

The successful modelling of a difficult galaxy such as NGC 708 motivates us to systematically extend this analysis to every BCG for which we have photometric and kinematical data. Given that these objects follow different scaling relations with respect to ordinary ETGs, recovering their intrinsic properties can help us to shred light on their evolution history.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Akaike H., 1974, IEEE T. Autom. Contr., 19, 716
Audet C., Dennis J. E., 2006, SIAM J. Optimiz., 17, 188
Audet C., Warren H., 2017, Derivative-Free and Blackbox Optimization. Springer, Cham
Bender R., Möllenhoff C., 1987, A&A, 177, 71
Bender R., Döbereiner S., Möllenhoff C., 1988, A&AS, 74, 385
Bender R., Burstein D., Faber S. M., 1992, ApJ, 399, 462
Bender R., Saglia R. P., Gerhard O. E., 1994, MNRAS, 269, 785
Bender R., Kormendy J., Cornell M. E., Fisher D. B., 2015, ApJ, 807, 56
Binney J., Mamon G. A., 1982, MNRAS, 200, 361
Bogdán Á., Lovisari L., Volonteri M., Dubois Y., 2018, ApJ, 852, 131
Cappellari M., 2002, MNRAS, 333, 400
Cappellari M., 2017, MNRAS, 466, 798
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cappellari M. et al., 2012, Nature, 484, 485
Conroy C., van Dokkum P. G., Villaume A., 2017, ApJ, 837, 166
Contopoulos G., 1956, ZA, 39, 126
Cooper A. P., Parry O. H., Lowing B., Cole S., Frenk C., 2015, MNRAS, 454, 3185
Davies R. L., Birkinshaw M., 1986, ApJ, 303, L45
Davies R. L., Birkinshaw M., 1988, ApJS, 68, 409
de Nicola S., Marconi A., Longo G., 2019, MNRAS, 490, 600
de Nicola S., Saglia R. P., Thomas J., Dehnen W., Bender R., 2020, MNRAS, 496, 3076

de Nicola S., Neureiter B., Thomas J., Saglia R. P., Bender R., 2022a, MNRAS, 517, 3445

de Nicola S., Saglia R. P., Thomas J., Pulsoni C., Kluge M., Bender R., Valenzuela L. M., Remus R.-S., 2022b, ApJ, 933, 215

de Zeeuw T., Franx M., 1989, ApJ, 343, 617

Djorgovski S., Davis M., 1987, ApJ, 313, 59

Ebisuzaki T., Makino J., Okumura S. K., 1991, Nature, 354, 212

Faber S. M., Jackson R. E., 1976, ApJ, 204, 668

Faber S. M. et al., 1997, AJ, 114, 1771

Falcón-Barroso J., Sánchez-Blázquez P., Vazdekis A., Ricciardelli E., Cardiel N., Cenarro A. J., Gorgas J., Peletier R. F., 2011, A&A, 532, A95

Ferrarese L., Merritt D., 2000, ApJ, 539, L9

Frigo M., Naab T., Rantala A., Johansson P. H., Neureiter B., Thomas J., Rizzuto F., 2021, MNRAS, 508, 4610

Gebhardt K. et al., 2000, ApJ, 539, L13

Gebhardt K. et al., 2003, ApJ, 583, 92

Gerhard O. E., 1993, MNRAS, 265, 213

Gerhard O., 1996, in Minniti D., Rix H.-W., eds, Spiral Galaxies in the Near-IR. Springer-Verlag, Berlin, Heidelberg, p. 138

Gerhard O. E., Binney J. J., 1996, MNRAS, 279, 993

Gonzalez A. H., Zabludoff A. I., Zaritsky D., 2005, ApJ, 618, 195

Graham A. W., Erwin P., Trujillo I., Asensio Ramos A., 2003, AJ, 125, 2951

Hernquist L., 1990, ApJ, 356, 359

Hopp U. et al., 2010, in Stepp L. M., Gilmozzi R., Hall H. J., eds, SPIE Conf. Ser. Vol. 7733, Ground-based and Airborne Telescopes III. p. 773307

Ilić S., Blanchard A., Douspis M., 2015, A&A, 582, A79

Kluge M., Bender R., 2023, ApJS, 267, 41

Kluge M. et al., 2020, ApJS, 247, 43

Kluge M., Bender R., Riffeser A., Goessl C., Hopp U., Schmidt M., Ries C., 2021, ApJS, 252, 27

Kluge M., Remus R.-S., Babyk I. V., Forbes D. A., Dolfi A., 2023, MNRAS, 521, 4852

Kormendy J., Bender R., 1996, ApJ, 464, L119

Kormendy J., Bender R., 2009, ApJ, 691, L142

Kosyra R., Gössl C., Hopp U., Lang-Bardl F., Riffeser A., Bender R., Seitz S., 2014, Exp. Astron., 38, 213

Kroupa P., 2001, MNRAS, 322, 231

Lang-Bardl F. et al., 2016, in Evans C. J., Simard L., Takami H., eds, SPIE Conf. Ser. Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI. p. 990844

Lauer T. R., Postman M., Strauss M. A., Graves G. J., Chisari N. E., 2014, ApJ, 797, 82

Le Digabel S., 2011, ACM T. Math. Software, accessed on 23-04-2023

Lipka M., Thomas J., 2021, MNRAS, 504, 4599

Magorrian J., 1999, MNRAS, 302, 530

Magorrian J. et al., 1998, AJ, 115, 2285

Maraston C., 2005, MNRAS, 362, 799

Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785

McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184

McConnell N. J., Ma C.-P., Murphy J. D., Gebhardt K., Lauer T. R., Graham J. R., Wright S. A., Richstone D. O., 2012, ApJ, 756, 179

Mehrgan K., Thomas J., Saglia R., Mazzalay X., Erwin P., Bender R., Kluge M., Fabricius M., 2019, ApJ, 887, 195

Mehrgan K., Thomas J., Saglia R., Parikh T., Bender R., 2023, ApJ, 948, 79

Mehrgan K., Thomas J., Saglia R., Parikh T., Neureiter B., Erwin P., Bender R., 2024, ApJ, 961, 127

Merritt D., 2004, in Ho L. C., ed, Coevolution of Black Holes and Galaxies. Cambridge University Press, Cambridge, p. 263

Merritt D., 2006, ApJ, 648, 976

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, J. Chem. Phys., 21, 1087

Milosavljević M., Merritt D., 2001, ApJ, 563, 34

Naab T., Johansson P. H., Ostriker J. P., 2009, ApJ, 699, L178

Neureiter B. et al., 2021, MNRAS, 500, 1437

Neureiter B., de Nicola S., Thomas J., Saglia R., Bender R., Rantala A., 2023a, MNRAS, 519, 2004

Neureiter B., Thomas J., Rantala A., Naab T., Mehrgan K., Saglia R., de Nicola S., Bender R., 2023b, ApJ, 950, 15

Nowak N., Saglia R. P., Thomas J., Bender R., Davies R. I., Gebhardt K., 2008, MNRAS, 391, 1629

Parikh T. et al., 2018, MNRAS, 477, 3954

Pillepich A. et al., 2018, MNRAS, 473, 4077

Pogge R. W. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy VI. p. 77350A

Poon M. Y., Merritt D., 2001, ApJ, 549, 192

Posacki S., Cappellari M., Treu T., Pellegrini S., Ciotti L., 2015, MNRAS, 446, 493

Rantala A., Johansson P. H., Naab T., Thomas J., Frigo M., 2018, ApJ, 864, 113

Rantala A., Johansson P. H., Naab T., Thomas J., Frigo M., 2019, ApJ, 872, L17

Rusli S. P. et al., 2013a, AJ, 146, 45

Rusli S. P., Erwin P., Saglia R. P., Thomas J., Fabricius M., Bender R., Nowak N., 2013b, AJ, 146, 160

Rybicki G. B., 1987, in de Zeeuw P. T., ed, IAU Symp. Vol. 127, Structure and Dynamics of Elliptical Galaxies. Dordrecht, p. 397

Saglia R. P., Bertschinger E., Baggley G., Burstein D., Colless M., Davies R. L., McMahan R. K., Wegner G., 1993, MNRAS, 264, 961

Saglia R. P. et al., 2016, ApJ, 818, 47

Salpeter E. E., 1955, ApJ, 121, 161

Schechter P. L., Gunn J. E., 1979, ApJ, 229, 472

Thomas J., Lipka M., 2022, MNRAS

Thomas D., Maraston C., Bender R., 2003, MNRAS, 339, 897

Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Corsini E. M., Wegner G., 2007, MNRAS, 382, 657

Thomas J. et al., 2011, MNRAS, 415, 545

Thomas J., Saglia R. P., Bender R., Erwin P., Fabricius M., 2014, ApJ, 782, 39

Thomas J., Saglia R., Bender R., Erwin P., Fabricius M., 2015, in Cappellari M., Courteau S., eds, Galaxy Masses as Constraints of Formation Models. Vol. 311. Cambridge University Press, Cambridge, p. 36

Thomas J., Ma C.-P., McConnell N. J., Greene J. E., Blakeslee J. P., Janish R., 2016, Nature, 532, 340

Tortora C., Romanowsky A. J., Cardone V. F., Napolitano N. R., Jetzer P., 2014, MNRAS, 438, L46

Trujillo I., Erwin P., Asensio Ramos A., Graham A. W., 2004, AJ, 127, 1917

van den Bosch F. C., 1997, MNRAS, 287, 543

van den Bosch R. C. E., 2016, ApJ, 831, 134

van den Bosch R. C. E., de Zeeuw P. T., 2010, MNRAS, 401, 1770

van der Marel R. P., Franx M., 1993, ApJ, 407, 525

Vincent R. A., Ryden B. S., 2005, ApJ, 623, 137

Wegner G. A., Corsini E. M., Thomas J., Saglia R. P., Bender R., Pu S. B., 2012, AJ, 144, 78

Zhao H., 1996, MNRAS, 278, 488

APPENDIX A: DUST CORRECTION

The availability of *HST* images acquired using different filters (in our case, f110W, f555W, and f814W) allows us to derive dust-corrected images and, hence, multiple estimates of the core radius of NGC 708 for a consistency/stability check with the values derived using the *K*s-image acquired at Wendelstein Observatory. The purpose of this appendix is to briefly summarize, following appendix A Nowak et al. (2008) and section 3.3 of Bender et al. (2015), the adopted procedure. We use as an example the correction of the *J*-band f110W image by means of the *I*-band f814W one, but the procedure is the same also for the case where we use f555W to dust-correct f110W.

The relation between the *J*-band absorption $A_J \equiv J - J_0$ and the extinction $E(I - J)$ in the $(I - J)$ colour is

$$A_J = \alpha E(I - J) \equiv \alpha[(I - J) - (I - J)_0], \tag{A1}$$

where the subscript 0 denotes dust-free quantities and $\alpha = 1/(A_I/A_J - 1)$ is the absorption coefficient. The strongest correction is
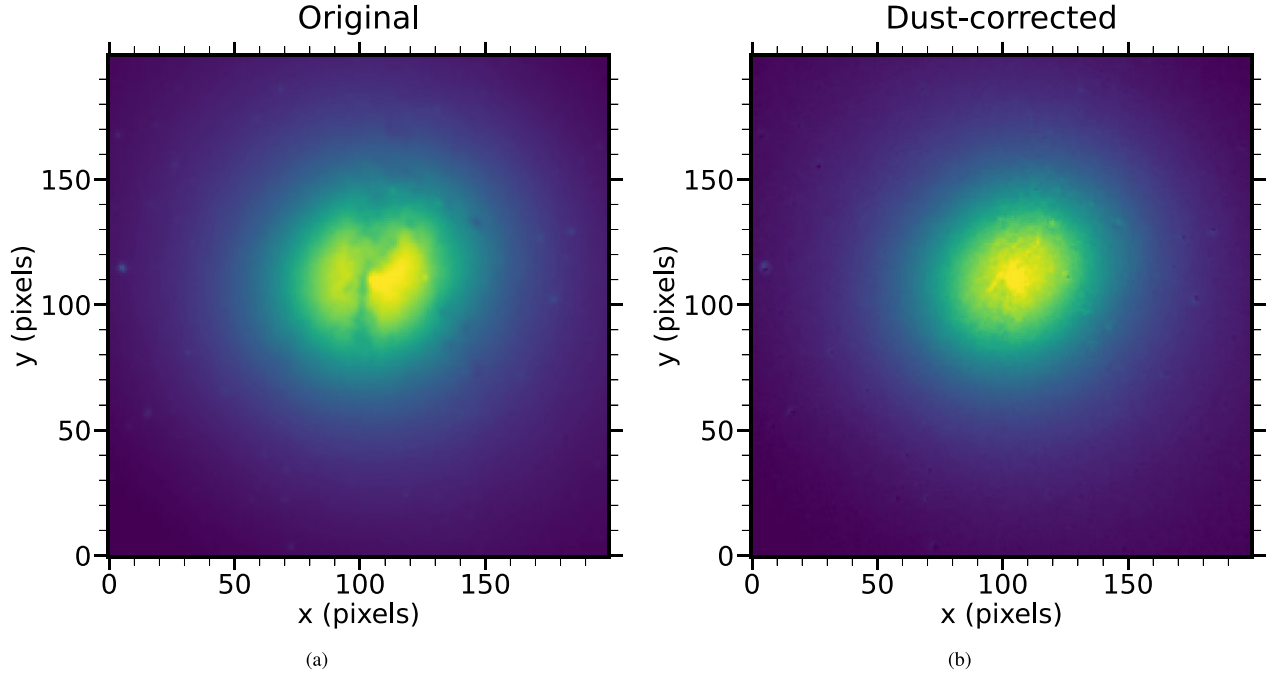
Figure A1. The central 18 arcsec of the *HST* f110W image of NGC 708 before (left) and after (right) dust-correcting using f814W. A value of 0.9 for the absorption coefficient α was adopted.
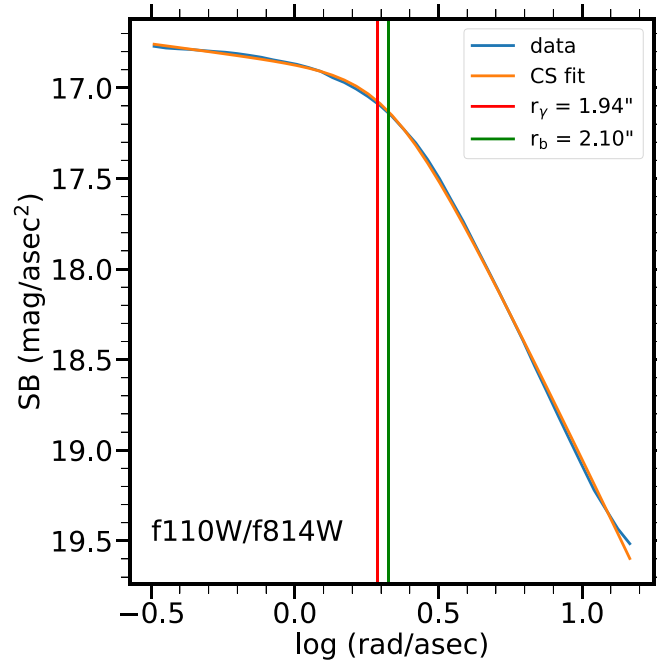

Figure A2. Central SB profile of NGC 708 derived from the dust-corrected f110W image (Fig. A1b). The blue line is the derived profile using the procedure described in Kluge & Bender (2023), while the orange line is the best-fit Core-Sersic (CS) profile. Finally, the vertical lines show the cusp-radius r_γ and the break radius r_b.

obtained for $\alpha = 1$, while $\alpha = 0$ implies no correction. Given that the images provide us with fluxes, we can write using equation (A1)

$$f_{J,0} \propto \frac{f_J^{1+\alpha}}{f_I^{\alpha}}, \qquad (A2)$$

where we have assumed that the stellar population gradient is negligible, thus implying $f_{I,0}/f_{J,0} \sim$ constant.

After registering the images and interpolating them onto the same pixel grid (f110W has a different pixel scale and a broader field of view), we test different values of α, finding that a quite high value
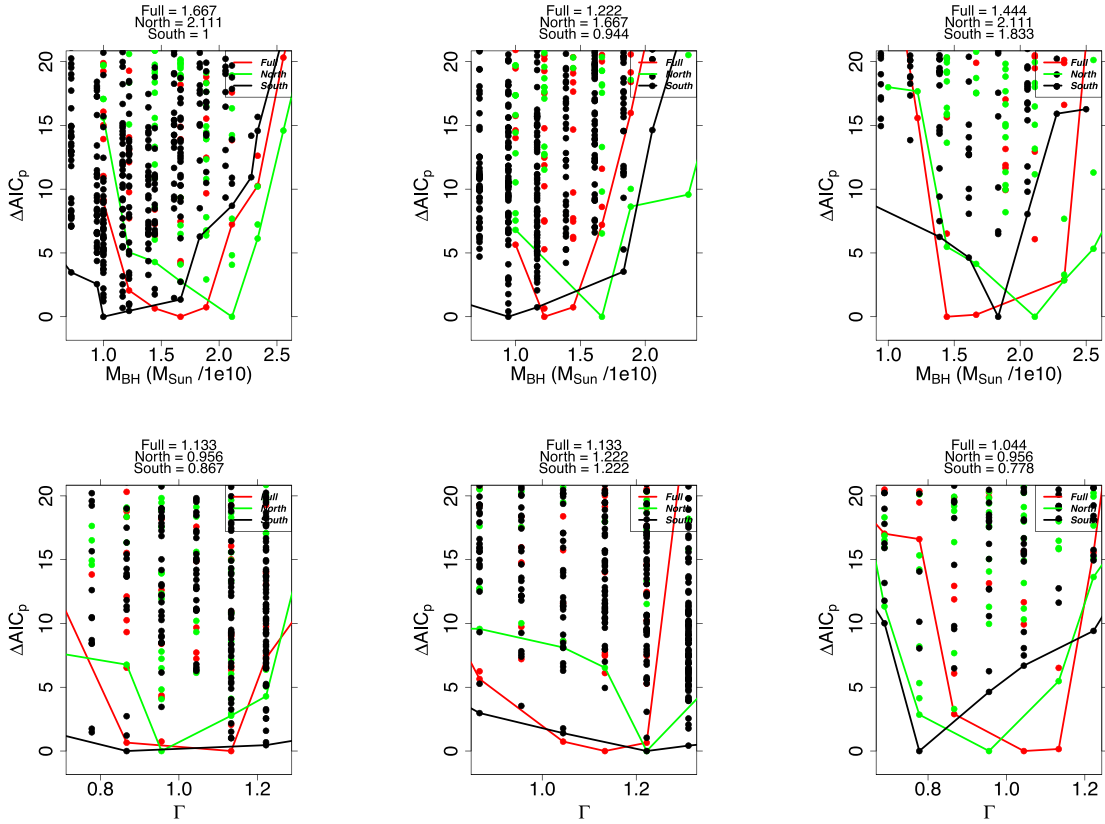
Figure B1. Similarly, as in Fig. 8, here we show the results of the NOMAD runs for the three mock kinematics, one for each column. The first row shows the M_{BH} recovery, whereas the second row shows Γ. In each plot, the red, green, and black lines follow the best-fit model for each tested value for the Full, North, and South configurations (see Table B1), respectively.
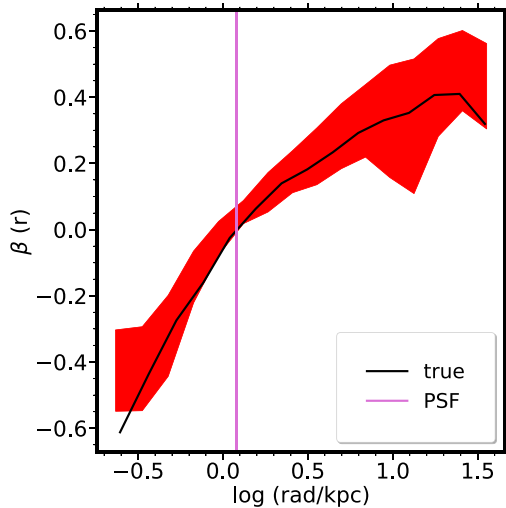

Figure B2. Anisotropy intervals that we get considering all nine best-fit models. The solid line shows the true *N*-body profile, whereas the vertical line labels the PSF.

of $\alpha = 0.9$ is required to remove the dust (see Fig. A1) and yield a smooth central SB profile (see Fig. A2), which is what we use to estimate the core radius by means of both the cusp-radius r_γ and the break radius r_{b}.

APPENDIX B: TRIAXIAL MODELS USING LONG-SLIT KINEMATICS

This appendix presents the result of the application of SMART to the same *N*-body simulation discussed in de Nicola et al. (2022a) and Neureiter et al. (2023a), this time using a simulated long-slit kinematics with the same geometry as the one used in this work for NGC 708. This exercise is not only useful to evaluate how accurate our triaxial machinery is with long-slit data, but also to get an estimate of the uncertainties on our mass parameters. Given the low number of bins, we need to model the galaxy fitting all bins together rather than split them in two halves and model these separately. Therefore, here we generate three mock kinematics and run independent sets of dynamical models for each one of these.

The mock kinematics are generated as follows: The spatial bins for the kinematics are taken at the exact same locations on the plane of the sky as those of NGC 708. After identifying the *N*-body particles found in the *i*-th spatial bin, we sample the *i*-th LOSVD using 15 velocity bins equally spaced in [−1500, 1500] km s^{-1}. The particle count inside the *i*-th spatial bin with velocities inside the *j*-th velocity bin gives us LOSVD$_{\mathrm{data}}^{i,j}$. For a given LOSVD, we assume a constant uncertainty for all velocity bins, equal to 3 per cent of the maximum number of counts among all velocity bins themselves. This gives us error bars comparable to those of NGC 708. We also add random Gaussian numbers N$_{\mathrm{rand}}^{i,j}$ to each *ij*-th bin to add noise to the data. Finally, we convolve the kinematics with a Moffat profile (see equation (3) of (Saglia et al. 1993)) with FWMH = 1.4 arcsec and β

Table B1. Best-fitting M_{BH} and Υ values for the *N*-body simulation when modelled using simulated long-slit kinematics. The true values are $M_{BH} = 1.7 \times 10^{10}\,M_\odot$ and $\Upsilon = 1.0$. *Col. 1:* Set up name. *Col. 2:* Whether we model all bins together (Full) or we split the bins in two subsets, one for each galaxy half (North and South). *Col. 3–4:* Best-fitting M_{BH}, Υ values.

	Configuration	$M_{BH}/10^{10}M_\odot$	Γ
true values	–	1.7	1.0
mock 1	Full	1.67	1.13
	North	2.11	0.95
	South	1.00	0.95
mock 2	Full	1.22	1.13
	North	1.67	1.22
	South	0.94	1.22
mock 3	Full	1.44	1.04
	North	2.11	0.96
	South	1.89	0.78
Result	–	1.56 ± 0.44	1.03 ± 0.15

$= 2.5$, similar to our MODS observations. Repeating this procedure thrice gives us the three kinematics sets.

Besides modelling all bins together as we did for NGC 708, we also modelled the two halves separately, given that for the simulated kinematics, the S/N is much higher, for a total of nine model runs.

Modelling the two halves separately would be the usual strategy, provided that the number of bins and the signal-to-noise are large enough. For each one of the nine model runs, we compute a total of ~1000 models. We do not fit the orientation, but assume it to be $(\theta, \phi, \psi) = (80, 90, 135)°$, i.e. the best-fit orientation yielded by SMART for NGC 708.

The results of the NOMAD runs are shown in Fig. B1 and summarized in Table B1. As it can be seen, averaging the results of the three mocks gives estimates that are well within 10 per cent of the true value. Moreover, having nine estimates per variable coming from the three mocks, we can also get an estimate on the statistical uncertainties on each variable by taking the standard deviation of the nine values we got. These uncertainties are what we assume for our NGC 708 mass parameter estimates. Finally, in Fig. B2, we consider all nine best-fit models and plot the resulting anisotropy intervals, comparing this to the true anisotropy profile of the *N*-body simulation. The fact that the true profile lies well in between the interval is a further proof of the goodness of our triaxial machinery even using long-slit data only.

This paper has been typeset from a T<small>E</small>X/L<small>A</small>T<small>E</small>X file prepared by the author.