Intrinsic shapes of massive elliptical galaxies

Stefano de Nicola

München 2022

Intrinsic shapes of massive elliptical galaxies

Stefano de Nicola

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Stefano de Nicola
aus Napoli

München, den 20.10.2022

Erstgutachter: P.D. Dr. Roberto Philip Saglia
Zweitgutachter: Prof. Dr. Andreas Burkert
Tag der mündlichen Prüfung: 23.11.2022

Contents

List of Figures

List of Tables

Zusammenfassung

"Brightest Cluster Galaxies" (BCGs) sind die massereichsten Galaxien des Universums. Sie bilden das massereiche Ende der Verteilung der Early-type Galaxies (ETGs) und beherbergen die massereichsten schwarzen Löcher. Verschiedene Beobachtungshinweise zeigen, dass diese Galaxien – im Gegensatz zu den kleineren ETGs – keine achsensymmetrischen sondern triaxiale Objekte sind. Deshalb benötigt man eine akkurate Software für triaxiale Systeme, um die intrinsische Struktur dieser Galaxien zu rekonstruieren. Diese wiederrum ist essentiell für die Vermeidung von systematischen Fehlern bei der Bestimmung der Massen der schwarzen Löcher, was seinerseits die Streuung in Skalenrelationen wie zwischen schwarzen Löchern und den sie beherbergenden Galaxien verringert, sowie die Streuung von Masse-zu-Leuchtkraft Verhältnissen Υ verringert, deren dynamische Bestimmung oft nicht mit unabhängigen Schätzungen aus der Analyse der stellaren Populationen übereinstimmt.
In dieser Dissertation wird eine komplett nicht-parametrische Software für triaxiale Galaxien präsentiert. Die Arbeit konzentriert sich vor allem auf meinen neuen Deprojektionscode `SHAPE3D` (de Nicola et al., 2020). Wesentlich für den Code ist, dass die Galaxie in Gestalt eines Ellipsoides beschrieben werden kann, was für massereiche ETGs eine gute Näherung ist. `SHAPE3D` bietet für die zwei Hauptprobleme, die bei Deprojektionen auftreten, gute Lösungen: Der Code erlaubt es, alternative Dichten für ein und dieselbe Sichtlinie zu finden, und er erlaubt es, die möglichen Orientierungen der Sichtlinie, welche mit einer bestimmten Photometrie kompatibel sind, signifikant einzuschränken. In de Nicola et al. (2022b) verwende ich den Code, um die intrinsischen Achsenverhältnisse der Lichtverteilung in einer Stichprobe aus 56 BCGs zu bestimmen. Ich zeige, dass diese Objekte extrem triaxial sind, sogar mehr als gewönliche ETGs. Ein Vergleich mit kosmologischen magneto-hydrodynamischen Simulationen zeigt, dass die Achsenverhältnisse in den BCGs mit denen von simulierten Halos gut übereinstimmen. Dies ermöglicht es, die Physik hinter der Entstehung von Dunkle-Materie (DM) Halos zu untersuchen und unterstreicht, wie viele wichtige Informationen die Deprojektionen enthalten.
`SHAPE3D` wird dann mit unserem neuen triaxialen Schwarzschild code `SMART` kombiniert. Ziel ist es, zu untersuchen, mit welcher Präzision die Sichtwinkel, die Achsenverhältnisse sowie die Verteilung der Sternbahnen in einer simulierten massereichen ETG bestimmt werden können (de Nicola et al., 2022a). Die Ergebnisse zeigen, dass die Sichtwinkel mit einer Präzision von ∼15-20° bestimmt werden können und dass die intrinsischen Achsenverhältnisse und die Anisotropie mit einer Genauigkeit von ≤ 0.1 rekonstruiert werden können. Bemerkenswert ist, dass schon allein mit der Deprojektion die oben angegebenen systematischen Fehler bei der Bestimmung der Achsenverhältnisse erreicht werden, während zusätzliche kinematische Informationen nötig

sind, um die Streuung der Messungen zu verringern. Außerdem ist die zentrale tangentielle Anisotropie, die für massenreiche ETGs typisch ist, sehr gut reproduziert.
Am Ende stelle ich die ersten Resultate der Modellierung der Galaxie NGC708 vor – der BCG des Galaxienhaufens Abell 262. Ich zeige, dass die Galaxie ein schwarzes Loch in ihrem Zentrum hat ($M_{\rm BH} \sim 10^{10}\,{\rm M}_{\odot}$) und präsentiere die rekonstruierten Profile der intrinsischen Achsenverhältnisse und der Anisotropie.

Abstract

Brightest Cluster Galaxies (BCGs) are the most massive galaxies in the Universe. They are found at the high-mass end of Early-type Galaxies (ETGs) and host the largest black holes. There are several indications pointing out that, differently from their smaller counterparts, these galaxies are not axisymmetric but triaxial objects. An accurate triaxial machinery is needed to recover the correct intrinsic shapes of these galaxies and to prevent biased estimations of black hole masses ($M_{\rm BH}$), needed to reduce the scatter in the black hole-hosts scaling relations, and mass-to-light ratios Υ, for which often a disagreement with estimations from stellar population analysis is found.

This work presents the first step towards a fully non-parametric triaxial machinery for fitting massive ETGs. In particular, it focuses on my novel semi-parametric triaxial deprojection code `SHAPE3D` (de Nicola et al., 2020). The code works under the approximation of a galaxy stratified onto concentric ellipsoids (which works well for massive ETGs) and can deal with the two main issues plaguing the deprojection: it can yield more than one solution for a given set of viewing angles and can reduce the number of possible orientations compatible with a given photometry. In de Nicola et al. (2022b) I apply the code to a sample of 56 BCGs to recover their intrinsic shapes, showing that these objects are extremely triaxial, even more than ordinary ETGs. A comparison with cosmological magneto-hydrodynamical simulations shows that BCG shapes match those of simulated halos well, allowing to investigate the physics behind Dark matter (DM) halo formation and stressing the importance of accurate deprojections.

`SHAPE3D` is then combined with the new triaxial Schwarzschild code `SMART` to study with which precision it is possible to recover the viewing angles, the intrinsic shape and the orbit distribution of a simulated massive ETGs (de Nicola et al., 2022a). The results show that the viewing angles can be recovered within ∼15-20°, while the intrinsic shape profiles and the anisotropy can be reconstructed with an accuracy $\leq 10\%$. In particular, the deprojection alone already estimates the intrinsic shape profiles accurately, while kinematical information is needed to reduce the scatter on these estimates. The central tangential anisotropy typical of massive ETGs is very well recovered.

Finally, I present preliminary results of the dynamical modeling of NGC708, the BCG of Abell 262, providing an estimate of its mass structure, in particular $M_{\rm BH}$ ($\sim 10^{10}$ $M_\odot$) as well as the intrinsic shape and anisotropy profiles.

Chapter 1

Introduction

Galaxies appear in a variety of shapes, some more spectacular, blue-coloured and with spiral arms, and some less, with redder colours and lacking any features. The first to introduce a classification scheme to take into account the different galaxy morphology was Hubble (1927), who was also the first one to realize that galaxies are not objects located in our Milky Way, but at much larger distance.

The Hubble classification scheme (Hubble, 1936) is shown in Fig. 1.1. It is based on how the galaxy appear on the plane of the sky. The two main galaxy groups are elliptical and spiral galaxies. Ellipticals appear as feature-less, smooth, nebulous, luminous objects consisting of a concentration of stars called bulge[1]. Nowadays, they are commonly called early-type galaxies (ETGs), because their old stellar populations imply that these galaxies must have formed during the earliest stages of our Universe. In the Hubble classification scheme, ellipticals are labeled by the letter E followed by a number calculated as $10 \times \varepsilon$, where $\varepsilon = 1 - b/a$ is the galaxy ellipticity and b, a are the projected axis lengths on the plane of the sky. Therefore, the higher the number, the flatter the galaxy: the flattest ETGs are classified as E7. Instead, spiral galaxies only have a small central bulge (if at all), and are dominated by a disk showing more or less prominent spiral arms, where star formation is still actively taking place. Given the young stellar population, spirals are commonly referred to as late-type galaxies. They are labeled with the letter S, followed by a letter going from a to c depending on the prominence of spiral arms (top-right part of Fig. 1.1). Moreover, some of them also host bar-shaped structures in the central bulges, and are labeled with an extra "B" (bottom-right part of Fig. 1.1). The main difference is that while classical bulges of non-barred spirals behave similarly as ellipticals, bulges found in barred spirals typically behave more like disks and are called pseudo-bulges (Kormendy & Kennicutt, 2004). A clear distinction between these two kinds of bulges remains a complicated challenge even today, as they can also coexist in the same galaxy (Erwin et al., 2015a).

The Hubble classification scheme also includes lenticular galaxies, which lie at the transition between ellipticals and spirals and are labeled as S0. In fact, they resemble spirals in that they consist of a bulge and a disk-like component. The difference is that lenticulars do not form stars and are thought to be remnants of spirals which lost their arms possibly because of the

[1]Elliptical were originally called *elliptical nebulae* by Hubble.

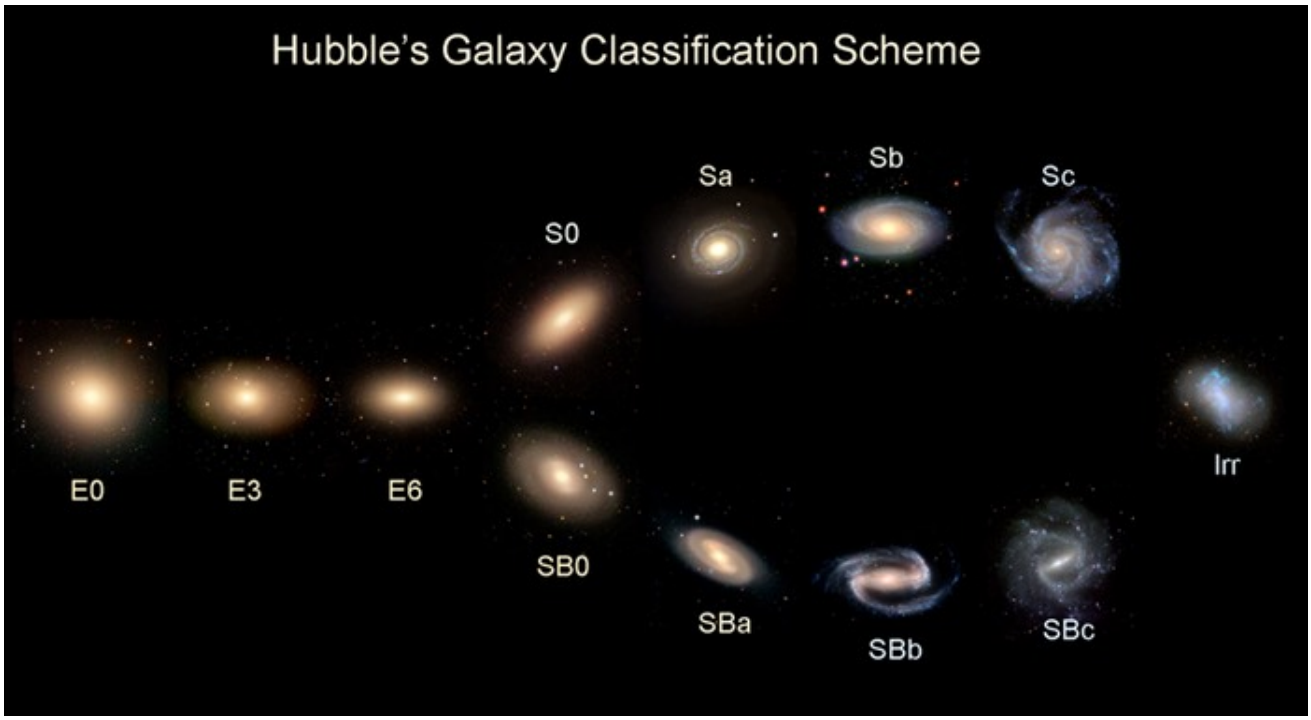


Figure 1.1: Hubble's classification scheme (also called Hubble tuning fork). "Elliptical nebulae" (nowadays called ellipticals) are on the left, while spiral galaxies are on the right (barred spirals below). Also shown is an example of irregular galaxy. Image credits: `http://astro.physics.uiowa.edu`

environment (Gunn & Gott, 1972; van Gorkom, 2004). Finally, we find irregular galaxies, which lack any ordered stellar structure or geometry.

1.1 Ellipticals

As previously written, elliptical galaxies appear as featureless, smooth galaxies on the plane of the sky. These galaxies do not typically form stars anymore: their stellar population is old. This is because ellipticals have little or no cold gas left. They are the largest galaxies in the Universe. In what follows their properties are described in detail.

1.1.1 Formation and evolution

Elliptical galaxies are the oldest galaxies in the Universe: in the most extreme cases, their stellar populations have ages of the same order of the age of the Universe. The most massive ETGs form in dense dark matter halos at redshift $z > 2$ (Genzel et al., 2006; Förster Schreiber et al., 2006). The common evolution picture for these objects sees them accreting mass at a high rate during the first evolutionary stages, favoured by the dense environment they form into. This generates

starbursts which, along with violent BH winds (Cattaneo et al., 2009; Fabian, 2012; King & Pounds, 2015), rapidly quench star formation (van Dokkum et al., 2008). Typically, the denser the halo, the more massive the galaxy is (Cimatti et al., 2006; Thomas et al., 2009). Then, during a second evolutionary phase, massive ETGs mostly increase their mass through several "dry" (gas-poor) merging events (Naab et al., 2009). Instead, smaller ETGs form at lower redshifts in less dense environments, thus accreting mass at a lower rate. This allows the star formation to last longer. Moreover, for these galaxies "wet" (gas-rich) mergers also play an important part in accreting their mass (Pulsoni et al., 2021).
Regardless of the environment, when galaxies form the conservation of angular momentum of the collapsing gas generates rotating systems (Fall & Efstathiou, 1980). However, while smaller galaxies are able to keep the rotation through gas replenishment in wet mergers, in massive ETGs the rotation is destroyed by dry mergers (Jesseit et al., 2009). This is only one of the differences between less and more massive ETGs: the topic is discussed further in Sec. 1.1.5.

1.1.2 Photometric properties

The Surface Brightness[2] (SB) of elliptical galaxies is generally well approximated by the Sersic law (Sersic, 1968):

$$I(R) = I_e \exp\left\{-\beta_n\left(\left(\frac{R}{R_e}\right)^{1/n} - 1\right)\right\} \tag{1.1}$$

where R_e is the effective radius, enclosing half of the total light, I_e is the SB value at R_e and $\beta_n \sim 2n - 0.324$. The index n is called Sersic index. A notable case is when $n = 4$ in eq. 1.1. This is called $R^{1/4}$ or De Vaucouleurs profile, and was introduced 20 years (de Vaucouleurs, 1948) before Sersic introduced eq. 1.1. The $R^{1/4}$ profile has for long been thought to describe all elliptical galaxies well. Nowadays high-resolution data allowed for the discovery of a correlation between the galaxy luminosity and n (Graham & Guzmán, 2003), namely that high-luminosity ETGs have higher Sersic indices. The values of n go from 0.5 up to 10 for the most massive ETGs.
When we look at ellipticals on the plane of the sky, the projected isophotes are very well approximated by ellipses. Therefore, in order to describe the isophotes one needs their ellipticity ε and also the position angle (PA[3]) of these on the plane of the sky. Recalling Hubble classification, the ellipticity lies in the range $[0, 0.7]$. Both variables can vary from one isophote to the other (i.e. as a function of radius). If this happens with the PA, the galaxy is said to have an isophotal twist.
By modeling the isophotes it is found that these are not perfect ellipses. In order to quantify the deviations of the actual shape from the best-fit perfect ellipse, we can expand these deviations as a function of the radius in Fourier series (Bender & Möllenhoff, 1987):

[2]The Surface Brightness is the amount of flux per unit solid angle.

[3]The PA is typically defined as the angle between the North and the isophotal major axis, measured counterclockwise.

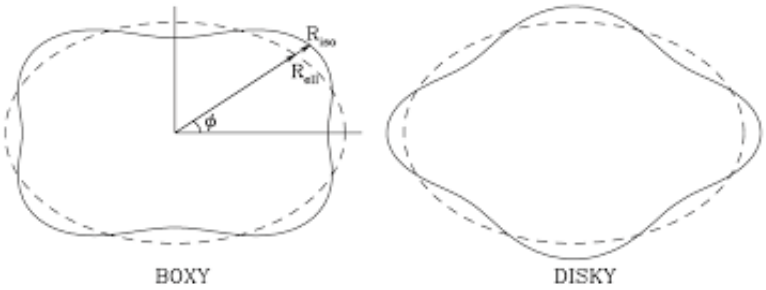


Figure 1.2: Example of boxy (left panel, negative a_4) and disky (right panel, positive a_4) isophotes. Image from Mo et al. (2008).

$$R_{iso}(\theta) - R_{ell}(\theta) = a_0 + \sum_{1}^{\infty}(a_n \cos n\theta + b_n \sin n\theta). \tag{1.2}$$

Here, the even and odd coefficients describe symmetric and asymmetric deviations, respectively. The most important coefficient is the 4th-order cosine term a_4. When a_4 is negative, we speak of boxy isophotes, otherwise of disky isophotes. As shown in Fig. 1.2, boxy isophotes have a light deficit along the projected principal axes, while the opposite happens for disky isophotes.

1.1.3 Intrinsic density profiles

The three-dimensional luminosity density $\rho(r)$ of ellipticals as a function of the distance from the galaxy center is well described by broken power-law models (Binney & Tremaine, 2008). A general form is (Zhao, 1996):

$$\rho(r) = \rho_0 \left(\frac{r}{r_0}\right)^{-\gamma} \left[1 + \left(\frac{r}{r_0}\right)^{\alpha}\right]^{(\gamma-\beta)/\alpha} \tag{1.3}$$

In particular, models with $(\alpha, \beta, \gamma) = (1, 3, 1)$ are called Navarro-Frenk-White (NFW) models (Navarro et al., 1997) and are commonly used to approximate Dark Matter (DM) profiles of elliptical galaxies. Profiles with $(\alpha, \beta) = (1, 4)$ are called Dehnen models (Dehnen, 1993). These profiles approximate the $R^{1/4}$ law if $\gamma = 3/2$. Dehnen models with $\gamma = 1$ are called Hernquist models (Hernquist, 1990), while models with $\gamma = 2$ are called Jaffe models (Jaffe, 1983). Typically, intrinsic density profiles of ETGs are well approximated by Dehnen models with γ in the range $[0.6, 2]$ (Binney & Tremaine, 2008). Dehnen models are useful in that the projected mass can be calculated analytically (Dehnen, 1993; Tremaine et al., 1994), with the gravitational potential also assuming a simple form. More massive ETGs typically have a lower α-value (see also Sec. 1.1.5).
The SB profiles corresponding to the models of eq. 1.3 is obtained by projecting these along the line-of-sight (LOS).

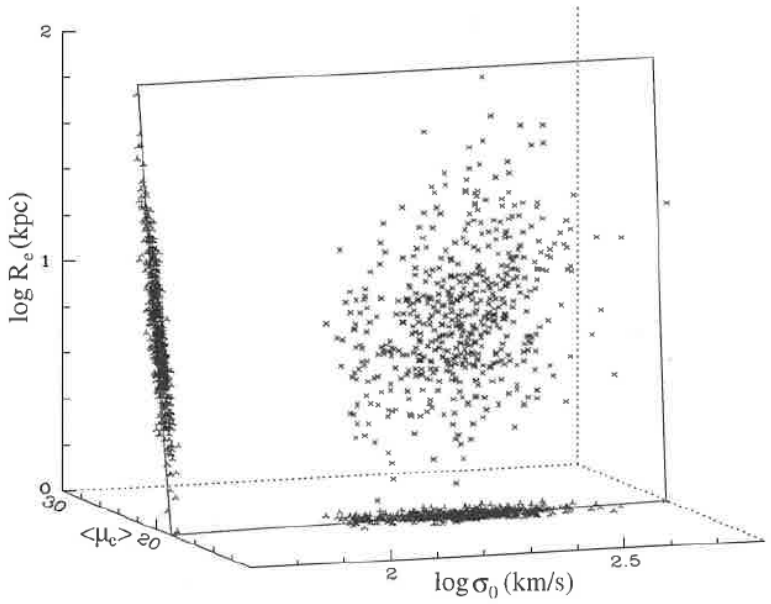


Figure 1.3: The FP of elliptical galaxies along with its 2D projections. While the FJ relation has a large scatter, the other two are much tighter, almost edge-on views. Image from Mo et al. (2008).

1.1.4 Scaling relations

Elliptical galaxies obey a tight correlation between size, luminosity and velocity dispersion. This relation is called Fundamental Plane (FP, Djorgovski & Davis 1987; Dressler et al. 1987):

$$\log R_{\rm e} = a \log \sigma_0 + b \log\langle I\rangle_e + ZP \tag{1.4}$$

where $R_{\rm e}$ is the effective radius, $\log\langle I\rangle_e$ is the mean SB *within* $R_{\rm e}$, σ_0 is the central velocity dispersion and ZP stands for zero-point. It is shown in Fig. 1.3. The slopes vary depending on the band one uses to derive the FP, with $a = 1.49 \pm 0.05$ and $b = -0.79 \pm 0.03$ in the r-band (Bernardi et al., 2003b). The variables are pairwise correlated, suggesting the possibility of a trivariate modeling of the FP, which has indeed been considered using a 3D Gaussian (Saglia et al., 2001; Bernardi et al., 2003b). Alternative parametrizations of the FP have been used. For example, Bender et al. (1992) proposed a parametrization based on rotated, orthogonal coordinates:

$$\begin{aligned}
\kappa_1/\sqrt{2} &\equiv \left(\log \sigma_0^2 + \log R_e\right) \\
\kappa_2/\sqrt{6} &\equiv \left(\log \sigma_0^2 + 2\log\langle I\rangle_e - \log R_e\right) \\
\kappa_3/\sqrt{3} &\equiv \left(\log \sigma_0^2 - \log\langle I\rangle_e - \log R_e\right)
\end{aligned} \tag{1.5}$$

which have the advantage of allowing for an easier interpretation of the parameters, since in this case κ_1 is directly related to the mass and κ_3 to the mass-to-light ratio.
The FP is a direct consequence of the virial theorem for self-consistent systems which would,

nevertheless, predict $a = 2$ and $b = -1$. The deviations of the measured slopes from the virial-predicted ones is called tilt. This is likely due to two different reasons. The first is non-homology: the Sersic index correlates with galaxy luminosity (cf. Sec. 1.1.2). The second is the variation of the mass-to-light ratio Υ within the galaxy population, which can arise both from age, metallicity and Initial Mass Function (IMF[4]) variations within stellar populations and from the DM influence on the total mass (Ciotti et al., 1996; Magorrian et al., 1998; Gerhard et al., 2001; Cappellari et al., 2006). It is still a matter of debate whether one of the two effects dominates over the other. The two notable projections of this fundamental plane are those on the I_e- $\sigma_e$and on the R_e-σ_e planes. The first is called the Faber-Jackson (FJ) relation (Faber & Jackson, 1976):

$$L \propto \sigma_0^4 \tag{1.6}$$

which implies that the most luminous galaxies have higher velocity dispersions. This equation predicts coefficients of $a = 2$ and $b = -1/2$ for the FP, thus implying that it is not an edge-on projection of the FP itself, as also shown in Fig. 1.3. The second relation reads:

$$log\langle I\rangle_e = R_e^{\nu} \tag{1.7}$$

and is called the Kormendy relation (Kormendy, 1977). Typically $\nu = 1.3 \pm 0.1$ (Bernardi et al., 2003a). Contrarily to the FJ relation, the Kormendy relation is an edge-on projection of the FP (see Fig. 1.3). Given that the exponent is smaller than 2, larger galaxies are more luminous and have lower SBs. The Kormendy relation is a prediction of the merger scenario (see Sec. 1.1.5). Both dry (gas-poor) and wet (gas-rich) mergers are required to explain the observed coefficients (Nipoti et al., 2003).

Black holes

A commonly accepted paradigm is that every galaxy hosts a Supermassive Black Hole (SMBH) at its center. The black hole mass $M_{\rm BH}$ correlates with the properties of the host galaxy bulge. One of the most important relations is the $M_{\rm BH}$- σ relation (Fig. 1.4):

$$\log M_{\rm BH} = \alpha \log \sigma + \beta, \tag{1.8}$$

where the velocity dispersion is the luminosity-weighted average within a certain radius. The theoretical explanation behind this relation lies in the BH feedback: the energy coming from the in-falling gas onto the SMBH gets redistributed throughout the galaxy itself and, in this way, the BH communicates its presence to the galaxy. Observational evidence for this is provided by e.g. high-velocity winds observed in X-ray (Pounds et al., 2003a,b; Reeves et al., 2003). Depending on the exponent, the feedback can either be momentum-driven ($\alpha = 4$) or energy-driven ($\alpha = 5$). While in the first case the the BH thrust does not threat the bulge integrity, in the second case the energy can clear off the gas and quench star formation. This is analogous to stellar feedback, which likely also plays a major role in suppressing star formation and originates the FJ relation

[4]The Initial Mass Function yields the number of stars at the birth of a stellar population with masses in a certain range.

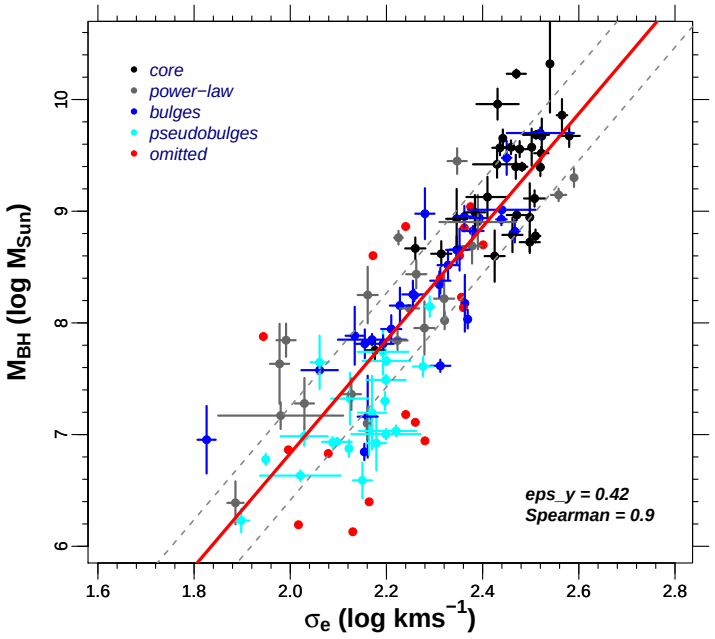


Figure 1.4: $M_{\rm BH}$- σ relation from de Nicola et al. (2019). Also shown are the different galaxy subgroups: core-ellipticals, which have a light deficit in the centre compared to a Sersic extrapolation, ordinary power-law ellipticals, classical (dynamically hot) bulges and dynamically cold pseudo-bulges. In this work a $M_{\rm BH} \propto \sigma^{5.23}$ dependence is found.

(Murray et al., 2005; Power et al., 2011). For the slope α a value larger than 5 is typically found (Gebhardt et al., 2000; McConnell & Ma, 2013; Saglia et al., 2016; van den Bosch, 2016; de Nicola et al., 2019). However, this might be affected by the heterogeneous samples used to derive this relation (King & Pounds, 2015), since the exponent is closer to 4 when more homogeneous galaxy subsets (e.g. only core or power-law ETGs) are considered (Saglia et al., 2016; de Nicola et al., 2019). The BH also correlates with the bulge mass, the bulge size and the bulge luminosity, suggesting a tight coevolution between galaxies and SMBHs. Given that velocity dispersion, luminosity and size are correlated through the FP, recent works have attempted to combine it with the BH-hosts scaling relations. For example, van den Bosch (2016) directly unified the $M_{\rm BH}$- σ with the FP, while de Nicola et al. (2019) used a multivariate approach to parametrize the FP. The results point to a scenario where the $M_{\rm BH}$- σ is the fundamental relation, while all other scaling relations are established through the FP. However, recent works (e.g. Sahu et al. 2022) claim that this is mostly due to inaccurate photometric decompositions.
Finally, tight scaling relations have been discovered within different galaxy subgroups. For instance, SMBHs in core-galaxies (see Sec. 1.1.5) follow a tight correlation with the core size (Rusli et al., 2013b), and anti-correlate with the central SB of the core itself (Mehrgan et al., 2019), in agreement with the paradigm which sees these light-deficient cores being formed as the results of gravitational slingshots following a dry mergers (Thomas et al., 2014).

1.1.5 Early-type galaxies bimodality

The Hubble classification only distinguishes ETGs based on the isophote flattening on the plane of the sky. Thus, contrarily to what happens with late-type galaxies, the Hubble classification scheme is not based on any physical property as far as ETGs are concerned. Kormendy & Bender (1996) proposed a revision of the Hubble classification based on isophote distortions. They noted that disky ellipticals rotate fast, while boxy ellipticals rotate slowly, thus suggesting a possible, further revised classification based on the velocity anisotropy (see Sec. 1.4.3). Indeed, this separation based on isophote distortions is only one of the differences between two families of ETGs. In what follows the fundamental aspects of this bimodality are discussed.

- Disky ellipticals tend to be less luminous, with magnitude $M_B > -20.6$ (or $M_V > -21.5$), while boxy ellipticals have $M_B < -20.6$ (or $M_V < -21.5$), even if there are works arguing that this is not a real dichotomy but rather a smooth transition (Ferrarese et al., 2006; Pasquali et al., 2007).

- Massive, boxy ETGs are on average rounder than less massive, disky ETGs. Moreover, by deprojecting the observed distributions of projected axis ratios and misalignment angles it has been found that massive ETGs are more triaxial (Tremblay & Merritt 1995, 1996; Weijmans et al. 2014; Foster et al. 2017; Ene et al. 2018, see Sec. 1.2.3).

- Less massive ETGs show an excess of light in the central regions compared to the prediction of a Sersic profile. Instead, massive ETGs show a shallower light profile, implying that there is a light deficit in the central regions (see left panel of Fig. 1.5). By defining a central slope $\gamma = dSB/dr$, we can distinguish between smaller, power-law ($\gamma > 0.3$) and core,

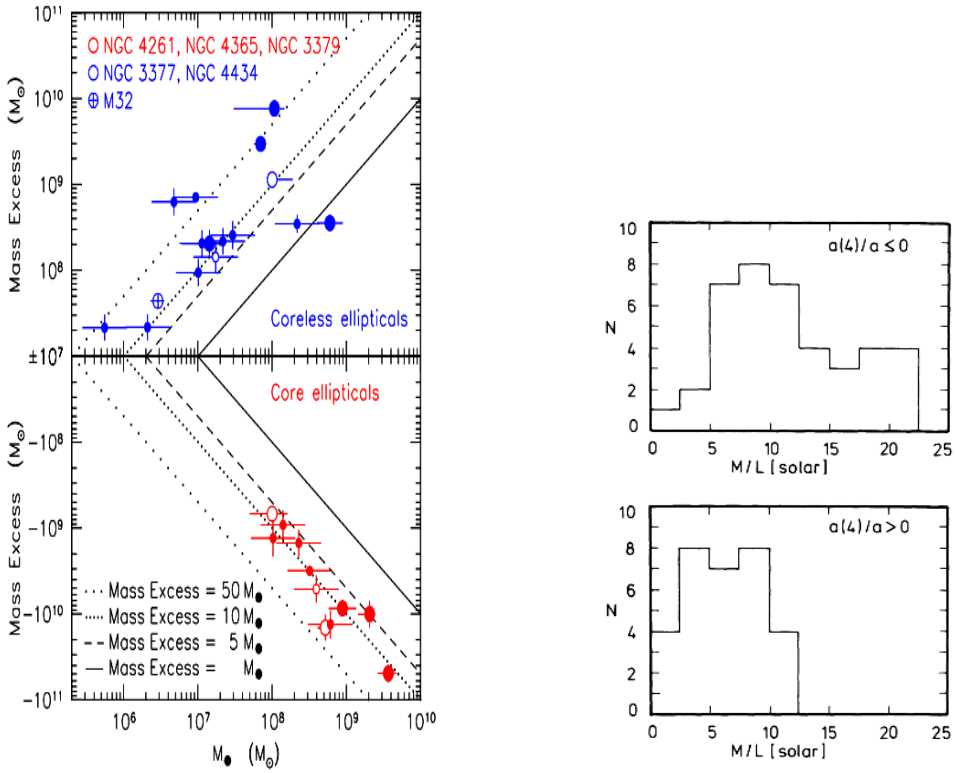


Figure 1.5: Examples of ETG bimodality. *Left:* The amount of extra-light in core-less ETGs (top) and of missing light in cored-ETGs (bottom) as a function of $M_{\rm BH}$ for all known galaxies in the Virgo Cluster (figure from Kormendy & Bender 2009). The amount of extra/missing light correlates well with $M_{\rm BH}$. *Right:* Number of boxy ($a_4 < 0$, top) and disky ($a_4 > 0$, bottom) galaxies as a function of the M/L ratio Υ from Bender et al. (1989). The higher Υ values are indeed found in boxy ETGs.

larger ($\gamma < 0.3$) galaxies (Lauer et al., 1995). This sharp transition has been questioned by more recent studies (Ferrarese et al., 2006; Côté et al., 2007). Cusps can be generated by starbursts occurring after wet mergers (Mihos & Hernquist, 1994). Particularly interesting are the cases of steep cusps ($\gamma > 1$), since these are theoretically predicted in stellar systems where the central black hole grows on timescales long compared to crossing times (Peebles, 1972; Quinlan et al., 1995). Instead, the missing light in the central regions of core-galaxies can be explained by the core-scouring mechanism, through which stars can be ejected by merging black holes via gravitational slingshot (Ebisuzaki et al., 1991; Trujillo et al., 2004; Rusli et al., 2013b; Thomas et al., 2014, 2016; Rantala et al., 2018; Mehrgan et al., 2019), and by the lack of gas replenishment since these galaxies form mostly via dry mergers.

- The Sersic index n correlates with the galaxy luminosity: the higher n, the more luminous the galaxy (Caon et al., 1993; Graham & Guzmán, 2003). In particular, less massive ETGs have n values as low as 0.5, while values as high as 10 are observed for the most massive galaxies.

- As already mentioned, massive ETGs show little or no rotation. This implies that they must be supported by velocity anisotropy. This is again a consequence of the formation scenario, since dry mergers do reduce rotation and possibly increase velocity dispersion (Jesseit et al., 2009; Naab et al., 2009). Instead, wet mergers allow less massive ETGs to keep rotating by exploiting the collapsing gas (Fall & Efstathiou, 1980).

- From a kinematical point of view, massive ETGs have less asymmetric LOSVDs (see Sec. 1.3.1) than flat and disky ones. This happens because the asymmetry of the LOSVDs correlates with the ratio between the rotational velocity v and the velocity dispersion σ (Bender et al., 1994): because massive ETGs are supported by anisotropy and rotate slowly, v/σ is lower than for less massive ETGs, which rotate fast.

- Kinematic misalignment, i.e. an angle between the photometric and the kinematic major axes, is observed in most slow-rotators ETGs, hinting at intrinsic triaxial shapes, whereas fast-rotators are generally well aligned. This is discussed in more detail in Sec. 1.2.3 & 1.1.6.

- The mass-to-light ratio Υ of boxy, massive ETGs is generally higher (Bender et al. 1989; Rix & White 1990, right panel of Fig. 1.5).

- The IMF depends on the velocity dispersion σ for power-law galaxies: the lower σ, the lower the fraction of low-mass stars. This is not observed in massive ETGs (Thomas et al., 2015): for these galaxies, it is a matter of debate whether the IMF is the same across the galaxy population or not. Accurate dynamical models (see e.g. Sec. 1.4.5) are essential to settle down this issue.

- Large galaxies typically have X-ray emission and are radio-loud. (Bender et al., 1989). The more natural mechanism to explain this is the presence of an AGN, which heats up the gas and prevents cooling (Cattaneo et al., 2009; Fabian, 2012; Werner et al., 2019).

- Perhaps the most significant difference lies in the formation and evolution history. As we saw in Sec 1.1.1, ETGs are thought to accrete mass via mergers (Toomre, 1977; Kormendy, 1984). Fainter ETGs are thought to form and grow in less dense DM halos and at lower redshifts w.r.t. massive ETGs. Furthermore, they accrete stars through wet mergers, where a significant amount of gas is present in the progenitors, whereas massive ETGs mostly increase their mass through dry mergers, with little or no gas involved (Bender, 1988; Barnes & Hernquist, 1992; Bender et al., 1992; Kormendy & Bender, 1996; Emsellem et al., 2007, 2011; Naab et al., 2014; Cappellari, 2016). As we saw above, this different formation scenario is able to explain most of the differences between these two galaxy groups.

Recent galaxy surveys (e.g. MASSIVE, Ma et al. 2014 or ATLAS, Cappellari et al. 2011b) which exploit Integral Field Spectroscopy (IFU) allowed to base this dichotomy on a kinematical properties rather than photometric ones. In this case, the main distinction is between fast and slow rotators, rather than between disky and boxy isophotes, depending on their position in the (λ, ε) diagram (see Sec. 1.3.2). Nevertheless, the kinematics-based classification mostly agrees with the photometry-based one, showing that rotation is not dynamically important in core-galaxies, while power-law galaxies are indeed consistent with the picture seeing them flattened by rotation.

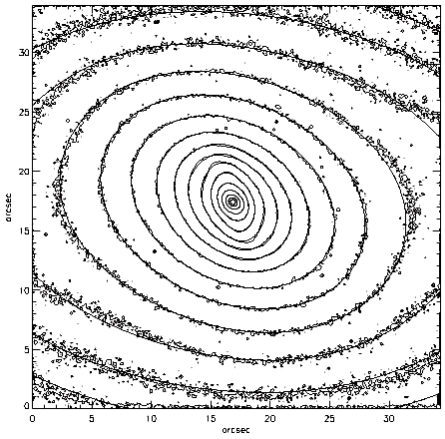
1.1.6 Evidence for triaxiality

Given that the isophotes of ETGs are well approximated by ellipses, the intrinsic three-dimensional shape is also of ellipsoidal shape (Binggeli, 1980; Binney, 1985; de Nicola et al., 2020). This approximation has been used in all pioneering works focusing on the geometry of these galaxies (e.g. Stark 1977; Binney 1978a, 1985; de Zeeuw & Franx 1989; Franx et al. 1991; Gerhard 1996). We define the axis ratios $p = b/a$ and $q = c/a$, where $a \geq b \geq c$ are the lengths of the three semi-axes of the ellipsoid. From this inequality[5], it follows that $1 \geq p \geq q > 0$. Eqs. C.1 and C.2 give the link between the observed flattening $q' = b'/a'$, where $a' \geq b'$ are the lengths of the projected principal axes of the isophotes, and the intrinsic axis ratios p, q. The two cases $1 = p \geq q > 0$ and $1 > p = q > 0$ correspond to an oblate and prolate geometry, respectively. Otherwise, the galaxy is triaxial. The degree of triaxiality can be quantified by the triaxiality parameter T (Franx et al., 1991), defined as

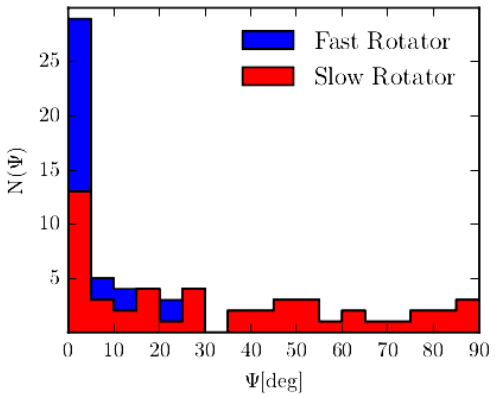
$$T = \frac{1 - p^2}{1 - q^2}. \tag{1.9}$$

Oblate and prolate objects have $T = 0$ and $T = 1$, respectively, while values in between signal a triaxial body. The maximum triaxiality is found for $T = 0.5$. For massive ellipticals, Vincent & Ryden (2005) found that $0.4 < T < 0.8$ by deprojecting the observed ellipticity distribution (see Sec. 1.2.3). Instead, faint ellipticals are consistent with oblate objects. Modeling a triaxial galaxy under the assumption of axisymmetry can lead to biased estimates of BH masses (van den Bosch & de Zeeuw, 2010) and Υ values (Thomas et al., 2007).

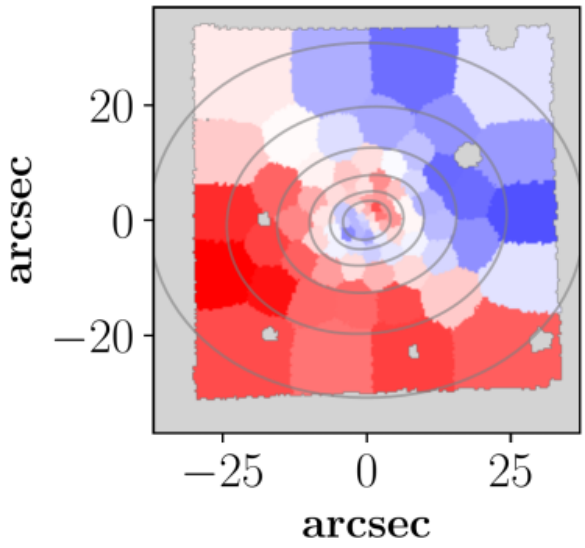
[5]From a physical point of view, $q \geq 0.2$.



(a)



(b)



(c)

Figure 1.6: Examples of commonly observed features in massive ETGs pointing to intrinsic triaxial geometry. *Top-left:* Isophotal twist for the galaxy NGC2950 (image from Cappellari 2002). *Top-right:* The kinematic misalignment (defined in Sec. 1.2.3) for a sample of 90 ETGs from the MASSIVE survey (image from Ene et al. 2018). This is mostly observed in massive, slowly rotating ETGs. *Bottom:* Example of kinematically decoupled core found in the massive ETG NGC5419. it is clearly seen from the image that the sense of rotation is reversed inside the core (image modified from the original provided by B. Neureiter).

The values of the triaxiality parameter T point out that massive ellipticals are not axisymmetric but rather triaxial objects. From a theoretical point of view, Schwarzschild (1979) demonstrated the existence of self-consistent triaxial systems in dynamical equilibrium. Nowadays, apart from the values of T, we have several more clues about the triaxiality of massive ellipticals. In fact, these objects often show observational features which an axisymmetric body cannot have:

- Isophotes of massive ellipticals often show isophotal twists (Bertola & Galletta, 1978), i.e. the PA is not the same for all isophotes[6]. Fig. 1.6a shows an example of this phenomenon for the galaxy NGC2950. For massive ellipticals the twists are not very large (Goullaud et al., 2018);

- Minor axis rotation, i.e. stellar rotation along the galaxy minor axis (Contopoulos, 1956; Kondratev & Ozernoi, 1979; Schechter & Gunn, 1979; Davies & Birkinshaw, 1986, 1988; Wagner et al., 1988; Jedrzejewski & Schechter, 1989);

- Kinematic misalignment, which means that the average photometric and kinematic PA (identified with the average direction of stellar motions) are shifted. Ene et al. (2018) study a sample of 90 ETGs, reporting 33% of them to have kinematic misalignment (see Fig. 1.6b).

- Kinematically decoupled cores, i.e. central cores rotating in a different direction with respect to the rest of the galaxy (Bender, 1988). They may be originated by BH mergers at galaxy centres. An example is provided by the ETG NGC5419 (Mazzalay et al. 2016, Neureiter et al. in prep., Fig. 1.6c).

- As we will see in Sec. 1.2.3, a triaxial galaxy population is required to match the observed distribution of projected flattenings of high-luminosity ETGs (e.g. Tremblay & Merritt 1995, 1996).

The evidence for triaxiality hints at the need of developing deprojection and dynamical modeling techniques which can deal with triaxial geometry. Easier approximations such as axisymmetry can lead to biased $M_{\rm BH}$ (van den Bosch & de Zeeuw, 2010) and Υ (Thomas et al., 2007) estimates. In particular, Υ can be underestimated up to 50%, causing problems when constraining the IMF. This is likely not to be a universal function as previously thought: Thomas et al. (2011) and Cappellari et al. (2012) report systematic variations within ETGs as a function of the effective velocity dispersion σ_e.
All these aspects point out that triaxial models are strongly needed. Indeed, one of the goals of this thesis is the introduction of a novel triaxial deprojection algorithm and the evaluation of its performance in recovering galaxy intrinsic shapes, as well as its usage in combination with our triaxial Schwarzschild code `SMART` (Neureiter et al., 2021).

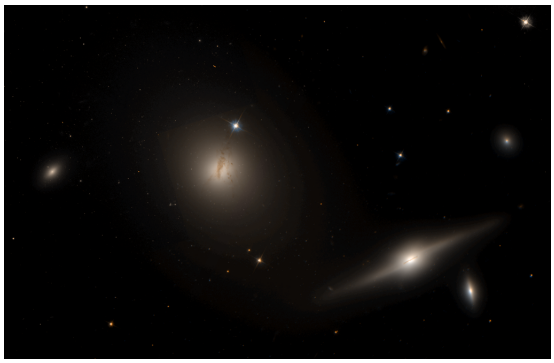


Figure 1.7: HST image of NGC708, the BCG of the cluster A262. A dust lane is clearly visible in the central regions. This is not observed in ordinary ETGs.

1.1.7 Brightest Cluster Galaxies

Galaxies are often not isolated objects, but are found in bound systems, called galaxy clusters. At the cluster centre, deep into the potential well of the cluster itself, lie the most massive galaxies in the Universe, the Brightest Cluster Galaxies (BCGs). These objects are typically, although not necessarily[7], far more luminous than any other galaxy of the cluster. Following Kluge et al. (2020), a BCG is defined as the closest galaxy to the geometrical and kinematical centre of the cluster. The position of the BCG allows it to increase its mass thanks to the in-falling material towards the cluster centre, for example through ram-pressure stripping (Gunn & Gott, 1972), harassment (Moore et al., 1996, 1998) or cannibalism (Gallagher & Ostriker, 1972; Ostriker & Tremaine, 1975; Richstone, 1976; Hausman & Ostriker, 1978). The last process is thought to be responsible for the creation of extended stellar envelopes surrounding the BCG. In this case, they are called cD galaxies (Matthews et al., 1964; Morgan & Lesh, 1965) although the nomenclature cD is sometimes used to identify *all* BCGs. In addition, a BCG can also be surrounded by Intracluster Light (ICL, Kluge et al. 2020). The difference between cD and ICL is that the first one is kinematically bound to the BCG, whereas the second one is not.
These galaxies possess several physical properties which make them interesting. First, differently from ordinary ellipticals, the SB profiles of BCGs often need two Sersic components rather than one to be fitted properly. Kluge et al. (2021) show that this is due to the fact that the outer component traces the ICL rather than the BCG itself. They often show twists larger than 20° as well as wiggled SB and ε profiles (Kluge et al., 2020; de Nicola et al., 2022b), which are typically generated by mergers. Moreover, BCGs follow both the FP and the Faber-Jackson relations, but with different slopes as those described in Sec. 1.1.4 (Kluge et al. 2022, in prep., see Fig. 1.8). Given that the Faber-Jackson relation can be used to predict the amount of light if one measures the velocity dispersion (cfr. eq. 1.6), the comparison between the two relations can be used to estimate the amount of ICL more reliably than using a simple BCG-ICL dissection method

[6]Binney (1978a) shows that a twist can be generated if $p(r), q(r)$ are not constant at all radii. Nevertheless, de Nicola et al. (2020) show that the variations must be large.

[7]An example is the Virgo Cluster, where the BCG, M87, is fainter than M49, which is located at the cluster outskirts.

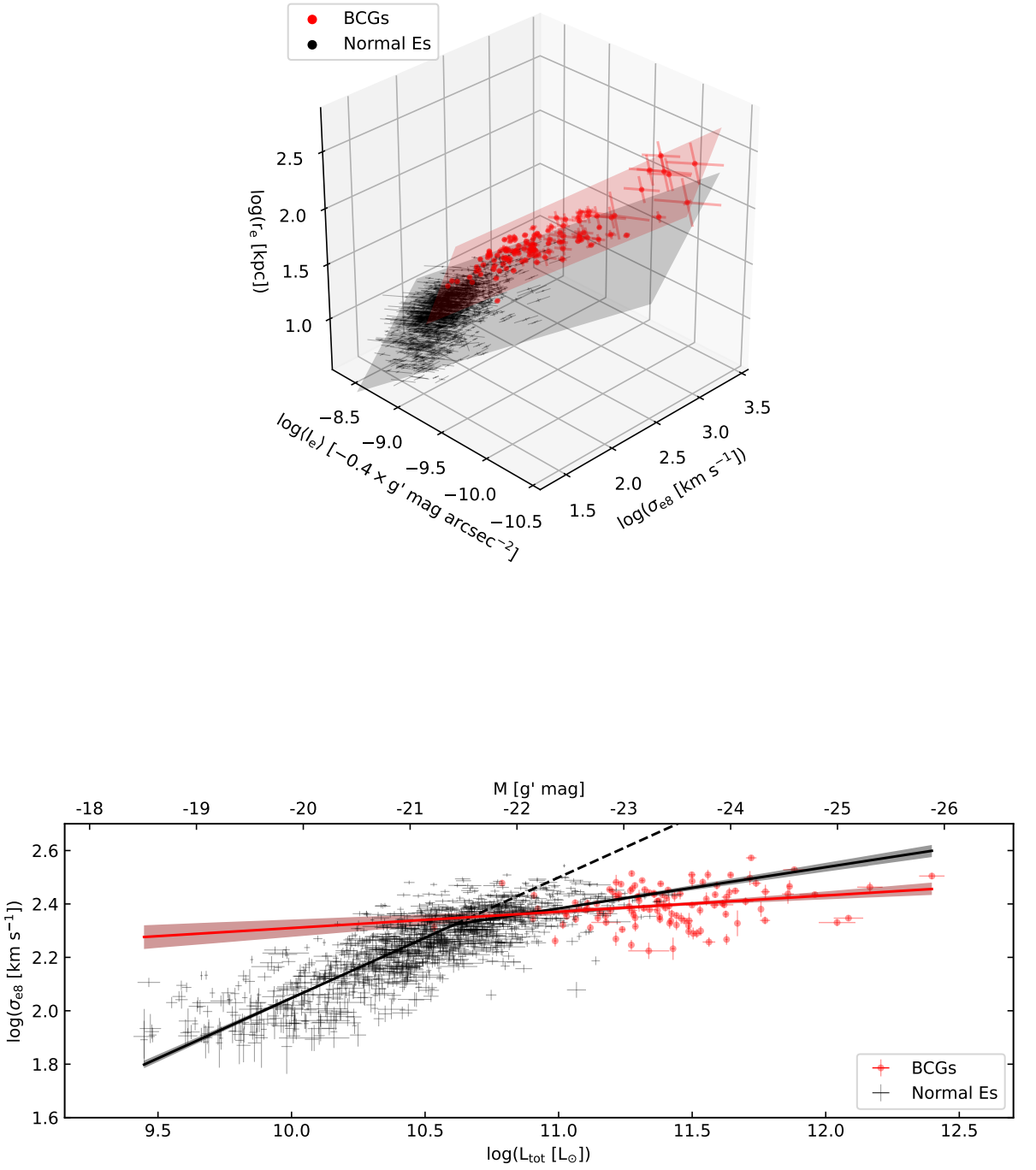


Figure 1.8: Fundamental Plane (eq. 1.4, top) and Faber-Jackson relation (eq. 1.6, bottom) for BCGs compared to those found for ordinary ETGs. The significant offset ($\sim 2.5\sigma$) is likely due to the ICL and offers a powerful way to dissect BCG from ICL light profiles. Images provided by M. Kluge.

(Kluge et al., in prep.). BCGs are also tracers of the cluster formation history: the brightness of the BCG together with its ICL correlates with cluster properties, showing that the growth of a BCG is tightly correlated with that of the cluster, as shown in Kluge et al. (2021).
Finally, BCGs appear to be very well suited for tracing the properties of DM halos. If one is able to separate a BCGs from the ICL, then this provides information about the Dark Matter halo surrounding the cluster, since the ICL is more aligned to the host cluster than the BCG itself. Another clue about DM comes from the intrinsic shape of these galaxies: de Nicola et al. (2022b) show that BCGs have intrinsic shapes which are remarkably close to DM halo shapes predicted by simulations (Hirschmann et al., 2014; Teklu et al., 2015; Marinacci et al., 2018; Pillepich et al., 2018; Springel et al., 2018). Simulations show that different assumptions about DM physics lead to different halo shapes (Robertson et al., 2019; Fischer et al., 2022), meaning that by deprojecting BCG photometric profiles (see chapter 3) one can make inferences about physical properties of DM halos.

1.2 Deprojection

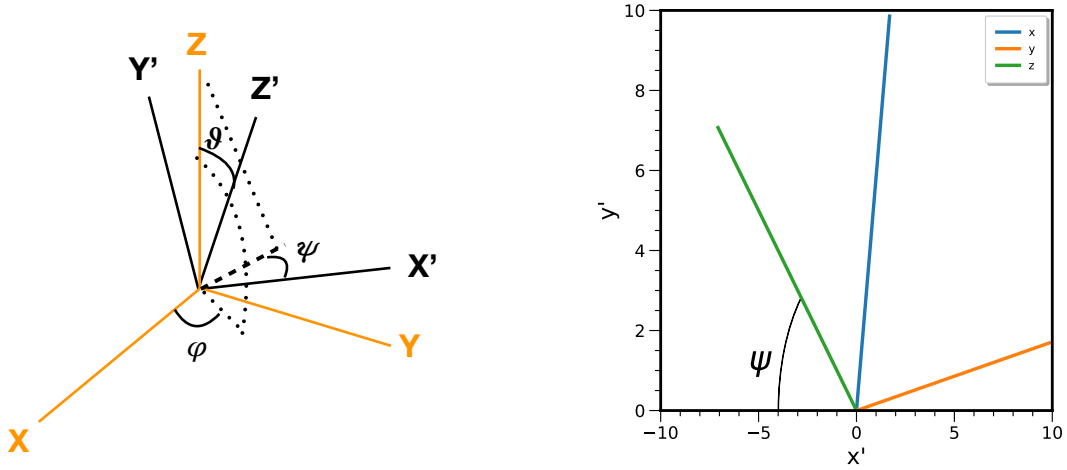
In the first section we saw that reconstructing intrinsic shapes not only gives us hints about the galaxy geometry, but also allows to speculate about DM physics by deprojecting BCGs surface brightness profiles. The importance of recovering the intrinsic light density from the observed photometry does not limit itself to this: it is, in fact, a crucial ingredient when dynamically modeling a galaxy. The aim of this section is to illustrate how the deprojection procedure works, present our knowledge of intrinsic shapes, the two main problems hidden behind deprojections and possible solutions to them.

1.2.1 Viewing angles

Conceptually, deprojections consist in recovering the luminosity density $\rho(x, y, z)$, defined in a coordinate system $O(x, y, z)$ aligned with the galaxy principal axes, which for a given galaxy orientation projects to the observed SB (x', y'):

$$\mathrm{SB}(x', y') = \int_{-\infty}^{+\infty} \rho(x, y, z) dz' \tag{1.10}$$

where $O'(x', y')$ is a coordinate system on the plane of the sky. The axis z', perpendicular to this plane, coincides with the LOS.
In the triaxial case, three angles (θ, ϕ, ψ) are needed to define the galaxy orientation. They are illustrated in Fig. 1.9. The first two angles (θ, ϕ) give the position of the LOS. The angle θ is the angle between the z'-axis and z, while ϕ is a rotation in the equatorial plane itself. Furthermore, a third angle ψ is required to specify a possible rotation about the LOS itself. It is defined as the angle between the projected z-axis on the plane of the sky and the x'-axis, measured counterclockwise (see Fig. 1.9b). In formulae, the deprojection can be written



(a) (b)

Figure 1.9: *Left:* Depiction of the three viewing angles (θ, ϕ, ψ). θ and ϕ specify the LOS position, while ψ gives a rotation around the LOS itself. *Right:* Position of the three principal axes on the plane of the sky when projected at $(\theta, \phi, \psi) = (45,45,45)°$. The angle ψ, defined as the angle between the projected z-axis and the x'-axis, is also shown for better clarification of its meaning.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{P}^t \cdot \mathbf{R}^t_\psi \cdot \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \quad \text{with} \quad \mathbf{R}^t_\psi = \begin{pmatrix} \sin\psi & \cos\psi & 0 \\ -\cos\psi & \sin\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.11}$$

and

$$\mathbf{P}^t = \begin{pmatrix} -\sin\phi & -\cos\phi\cos\theta & \cos\phi\sin\theta \\ \cos\phi & -\sin\phi\cos\theta & \sin\phi\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}. \tag{1.12}$$

In words, we first need a to apply a rotation specified by ψ using the matrix $\mathbf{R}^t_\psi$ to align the z-axis with the x'-axis. Then, the matrix $\mathbf{P}^t$ is used to specify the orientation of the LOS. This matrix is obtained by multiplication of the two matrices

$$\mathbf{R}^t_\phi = \begin{pmatrix} -\sin\phi & \cos\phi & 0 \\ \cos\phi & \sin\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{R}^t_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\cos\theta & \sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{1.13}$$

as $\mathbf{P}^t = \mathbf{R}^t_\phi \cdot \mathbf{R}^t_\theta$.
The projection equations are simply obtained by inverting eqs. 1.11 & 1.12 and are explicitly written in eqs. 2.12 & 2.13. In the axisymmetric case, ϕ does not matter, while ψ can be determined immediately because in this case the intrinsic axes project exactly onto the semi-major and semi-minor axes of the isophotes. The angle θ, commonly called i in this case, is

the only one required to specify the galaxy orientation, with the two extreme cases $i = 0, 90^\circ$ corresponding to face-on and edge-on projections, respectively.
For a perfect ellipsoid, the intrinsic axis ratio(s) can be computed using eqs. C.3 (or eq. C.5 for an axisymmetric body) using the observed photometry and viewing angles. Therefore, starting from the observed SB one could, in principle, invert eq. 1.10 and recover the intrinsic light distribution of a given galaxy. Unfortunately, viewing angles cannot be measured in general, and the issue gets particularly severe when dealing with massive, smooth elliptical galaxies, which are not perfect ellipsoids.

1.2.2 The Fourier Slice Theorem

In addition to the fact that the viewing angles cannot be measured, deprojections are plagued by the Fourier Slice Theorem. The theorem states that the following two operations on a two-dimensional function are interchangeable:

- Compute the 2D Fourier transform of this function then slice it parallel to the LOS;
- Project this function along the LOS then compute its 1D Fourier transform.

The theorem can easily be generalized to higher dimensions. Its mathematical proof is shown in App. B.1. This is particularly relevant when deprojecting a galaxy image. In fact, if the galaxy hosts a sub-structure (e.g. a disk, a bar, a ring, etc.) whose Fourier transform vanishes perpendicularly to the LOS, because of the Fourier Slice Theorem it will be invisible in projection. Therefore, the deprojection of a galaxy yields a non-unique solution.
This issue gets more or less severe depending on the galaxy geometry and on its inclination. In the axisymmetric case (Gerhard, 1996; Kochanek & Rybicki, 1996; van den Bosch, 1997; Magorrian, 1999), all LOSs which are obtained by rotating the galaxy around the symmetry axis project to identical images on the plane of the sky since the galaxy is axisymmetric, and this applies to the Fourier transform of the intrinsic density too. Thus, if a galaxy is seen exactly edge-on ($i = 90^\circ$), the deprojection can be uniquely recovered, because this rotation process covers the whole Fourier space. The situation gets increasingly worse going towards the face-on case ($i = 0^\circ$). In this last case, isophotes are circular, and the deprojection is completely unconstrained. At intermediate inclinations, Gerhard & Binney (1996) show that we can define a **cone of ignorance** of aperture $(90 - i)^\circ$, and any function whose Fourier transform is non-zero *only* in this cone of ignorance projects to nothing. Such functions are called **konus densities**, and there have been works in the past showing how to construct them analytically in the axisymmetric case (e.g. Gerhard & Binney 1996; Gerhard 1996).
In the triaxial case, the situation gets worse: Gerhard (1996) shows that the triaxial symmetry only enables us to uniquely derive Fourier transform of the the light density on four planes perpendicular to the LOS. The number of planes reduces to two if the LOS is along one of the principal planes[8] and to one if this is along one of the principal axes. To my knowledge, before this thesis was written no published work attempted to derive analytic models of triaxial density distributions which project to nothing (see Chapter 2).

[8]Principal planes are the equatorial plane (x, y), the plane (x, z) and the plane (y, z).

1.2.3 Statistics of intrinsic shapes

The statistical distribution of isophotes of observed ellipticals (Tremblay & Merritt, 1996; Vincent & Ryden, 2005; Ene et al., 2018) shows that ETGs have an average axis ratio $< q' > \equiv 1-\varepsilon \sim 0.8$. In particular, brighter ETGs show rounder isophotes ($< q' > \sim 0.85$) than fainter ones ($< q' > \sim 0.7$). However, it is perhaps more useful to figure out what this implies in terms of intrinsic shapes. An inversion of the observed of axis ratios probability distribution $P'(q')$ to recover the intrinsic distribution P yields an unique solution only in case of axisymmetry. Here, $P(q)$ is only a function of the minor-to-major axis ratio c/a and the observed axis ratio q' only depends on the inclination angle i and the intrinsic flattening q (see eq. C.5). Assuming that the galaxies are randomly oriented, $P(q)$ can be analytically calculated as (Fall & Frenk, 1983)

$$P(q) = \frac{2}{\pi}\sqrt{1-q^2} \begin{cases} \frac{1}{q}\int_0^q \frac{dP'}{dq'}\frac{q'dq'}{\sqrt{q^2-q'^2}} & \text{(oblate case)} \\ \\ \frac{1}{q^3}\int_0^q \frac{d(q'^3P')}{dq'}\frac{q'dq'}{\sqrt{q^2-q'^2}} & \text{(prolate case).} \end{cases} \tag{1.14}$$

Tremblay & Merritt (1995) developed a non-parametric algorithm to recover $P(q)$, finding that a distribution of pure axisymmetric galaxies cannot explain the observed $P'(q')$ because of a lack of galaxies with low ellipticity. A subsequent work (Tremblay & Merritt, 1996) shows that only massive ETGs require a triaxial geometry, whereas low-luminosity ETGs are indeed consistent with an axisymmetric population, as already mentioned in Sec. 1.1.5. However, the triaxial case is significantly more complicated, and requires further assumptions. In fact, in this case q' is a function of the viewing angles (θ, ϕ) as well as of the two intrinsic axis ratios p, q (eq. C.1). The only additional observable is the kinematic misalignment angle Ψ, between the photometric and the kinematic major axis, the latter being defined as the direction of the average orientation of stellar motion. For an axisymmetric galaxy, this coincides with the projected major axis. Instead, in triaxial galaxies this can be non-zero, and often close to 90° (minor-axis rotation). Thus, in general one has two observables and five variables (Foster et al., 2017):

$$\begin{aligned} \Psi &= \Psi(\Psi_{intr}, p, q, \theta, \phi) \\ q' &= q'(p, q, \theta, \phi) \end{aligned} \tag{1.15}$$

Commonly used assumptions include drawing p and q from (log-)normal distributions, assuming randomly oriented galaxies and Ψ_{intr} along the viewing direction which generates round isophotes (Weijmans et al., 2014; Foster et al., 2017), such that Ψ_{intr} and q' only depend on the galaxy shape. The distribution of intrinsic flattenings can then be recovered. This has been applied to a sample of galaxies from the MASSIVE survey (Ma et al., 2014; Ene et al., 2018), finding that massive ETGs are indeed consistent with triaxial shapes, with a broad distribution of observed kinematic misalignments.

All these approaches suffer from several drawbacks. For example, the assumption of constant axis ratios as a function of radius is unrealistic, because in this case according to eqs. C.1 and C.3 the ellipticity must be constant for all isophotes and there cannot be an isophotal twist, while ellipticity and PA gradients are commonly observed (e.g. Bender et al. 1988, 1989; Kormendy

& Bender 1996; Goullaud et al. 2018). In addition, it would be interesting to derive the statistics of **radially resolved** intrinsic shape profiles *for individual galaxies*, which is difficult given the degenerate deprojection. These aspects motivate the third chapter of this thesis, where the first analysis of this kind is presented.

1.2.4 Deprojection algorithms

Sec. 1.2.1 and 1.2.2 make clear that deprojections are a more difficult task than it may look at first glance. In the last ~25 years, both parametric and non-parametric deprojection algorithms have come out (Magorrian, 1999; Bissantz & Gerhard, 2002; Cappellari, 2002; de Nicola et al., 2020). This section highlights the strengths and weaknesses of each one of the two classes.

Parametric methods

These have the advantage of reducing the deprojection to a set of parameters, which can be obtained through a quick fit. In this case, one needs to approximate the SB distribution with an analytic formula for which eq. 1.11 can be inverted immediately. An example of analytic profile which well approximates SB profiles of cusp-ellipticals is the Jaffe (1983) profile. More general formulation such as eq. 1.3 can be used to model massive ETGs too. A quite general parametrization of the SB is the Multi-Gaussian Expansion (Bendinelli, 1991; Monnet et al., 1992; Emsellem et al., 1994; Cappellari, 2002), according to which the SB is written as a sum of gaussian functions[9]. The parameters needed to reproduce the SB are the luminosity, flattening, PA and dispersions of each component (see eq. 1 of Cappellari 2002). From the SB profile and *assuming* the viewing angles, one can immediately calculate the intrinsic shape through eqs. C.3 and consequently ρ through eq. 6 of Cappellari (2002). Although parametric methods are very fast, they do not satisfactorily address the two issues presented above. In fact, although they restrict the range of possible orientations since a (physically meaningful) solution is generally not found for all viewing angles (see e.g. eqs. C.3 for the ellipsoidal case), it is not possible to operate a further cut-off based on the goodness of fit since once the SB has been easily approximated, the deprojection is perfect by definition. Finally, no current implementation allows to sample more than one density for each set of viewing angles, thus not allowing for an investigation of the degeneracy.

Non-parametric methods

This class of algorithms aims at reproducing the full 3D density on a grid without making assumptions on the galaxy shape. They require penalty functions to rule out implausible (e.g. unsmooth) solutions. Compared with parametric methods here one trades computation speed (the number of points to fit is much larger) with a much more general approach (no assumptions on the shape are needed). Non-parametric methods have been tested in the axisymmetric approximation by Magorrian (1999) and in the triaxial case by Bissantz & Gerhard (2002). Although it is

[9] A single gaussian function deprojects to an ellipsoidal intrinsic density. By combining several gaussian functions, it is possible to approximate a wide variety of SB profiles, and to obtain non-ellipsoidal-shaped deprojections.

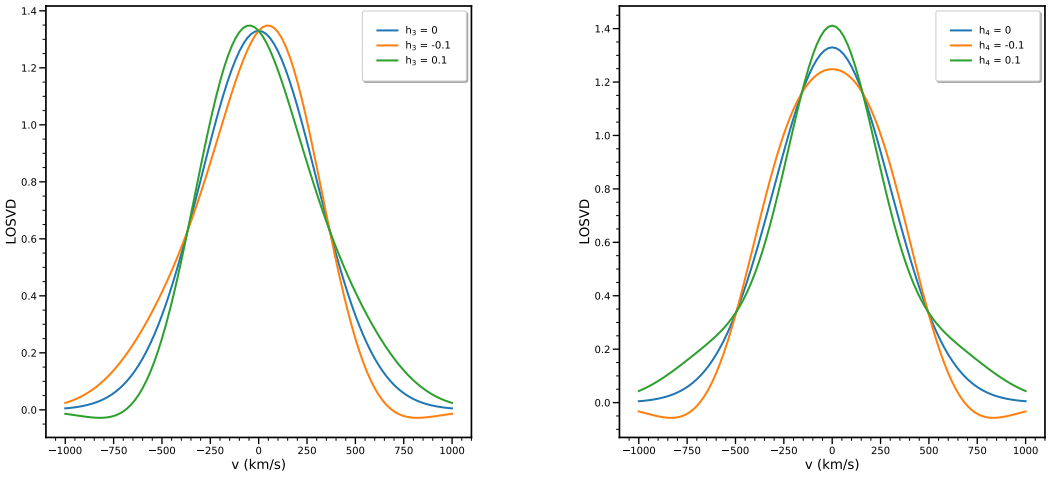


Figure 1.10: Example of LOSVDs illustrating the effects of Gauss-Hermite polynomials. In both panels, the blue line is a pure Gaussian. The left panel shows the meaning of h_3: negative values displace the peak towards positive velocity values, and make the retrograde wing steeper. Instead, the right panel focuses on h_4: positive values enhance the wings and the peak of the curve.

not possible to explore the full range of light densities compatible with a certain SB profile at a given orientation, these methods can yield a large number of solutions at the same (θ, ϕ, ψ). Nevertheless, even by applying some kind of regularisation, the degeneracy in the deprojection is hard to keep under control, since the number of invisible densities is very large already in the axisymmetric case (Gerhard & Binney, 1996), while in the triaxial case little is known (Gerhard, 1996). The viewing angles represent another problem, since a solution may be found at every orientation. At the time this thesis was written, the common belief was that the viewing angles cannot be determined photometrically. One may deproject at several different viewing angles, but sampling a large number of deprojected densities using dynamical models is too time consuming. From this discussion it appears that the optimal solution lies in between parametric and non-parametric methods. This motivated the development of such an algorithm, which I present in chapter 2.

1.3 Kinematics

This section focuses on galaxy spectroscopy and the extraction of kinematics from galaxy spectra. It explains the basic concepts of spectral fitting along with examples of the relation between kinematics and observed properties. Finally, a brief overview about fitting procedures is given.

1.3.1 Line-of-Sight velocity profiles

Along with the light distribution, another crucial ingredient for dynamical modeling is the galaxy kinematics. For this, we need spectroscopic observations which provide us with the integrated

galaxy spectrum.[10] Since early-type galaxies are characterized by old stellar populations, they do not typically show emission lines, unless giant stars ionize photons through stellar winds (Sarzi et al., 2010) or there is a non-stellar component (e.g. an AGN). The spectrum consists of prominent absorption lines, Doppler-shifted according to the stellar velocity along the LOS, superimposed onto the continuum spectrum[11]. In practice, what we observe is the convolution of a large number of stellar spectra with the line-of-sight velocity distribution (LOSVD), which gives the fraction of stars within a certain velocity interval $v \pm dv$ (Binney & Merrifield, 1998). It used to be parameterized using a simple Gaussian function (Sargent et al., 1977), so that the only two parameters needed to describe it were the mean rotation velocity $\bar{v}$ and the velocity dispersion σ. For elliptical galaxies the velocity dispersions are in the range $[100-350]$ km s^{-1}, whereas the rotational velocities[12] are typically $\leq$150 km s^{-1} (Gerhard, 1994). These variables can be measured as a function of the distance from the galaxy centre to derive the velocity profile. New techniques such as wide-field spectrographs are able to sample the kinematics up to $\sim 3R_{\rm e}$ (Arnold et al., 2014). The presence of minor-axis rotation, i.e. rotation around the galaxy major axis, provides an evidence for triaxiality, unless the galaxy is prolate (cf. Sec. 1.1.6).
Nevertheless, the shape of the LOSVDs slightly deviates from a Gaussian shape (Dejonghe, 1987; Franx & Illingworth, 1988; Bender, 1990; Bender et al., 1994). This is theoretically expected in the presence of rotation, which makes LOSVDs asymmetric. It has become common practise to parameterize it using Gauss-Hermite moments (van der Marel & Franx, 1993; Gerhard, 1993), defined as

$$u(x) = \left(2^{n+1}\pi n!\right)^{1/2} H_n(x) e^{-\frac{x^2}{2}} \tag{1.16}$$

where $x = (v - \hat{v})/\hat{\sigma}$. $\hat{v}$ and $\hat{\sigma}$ are the fitted velocity rotation and dispersion, while H_n are the Hermite polynomials (Abramowitz & Stegun, 1972):

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \tag{1.17}$$

Gauss-Hermite moments have the advantage that they are not so sensitive to LOSVD wings in comparison with classical moments. Thus, the LOSVD can be parametrised as

$$LOSVD(v) = e^{\frac{-x^2}{2}} \left[1 + \sum_{n=3}^{+\infty} h_n H_n(x)\right]. \tag{1.18}$$

Of particular relevance are the polynomials of third and fourth order, expressing asymmetric and symmetric deviations from the Gaussian shape, respectively. Positive h_3 values indicate that the retrograde wing is less steep than the prograde one (Fig. 1.10, left panel), while positive h_4 values yield a more peaked LOSVD and more enhanced wings (Fig. 1.10, right panel). These coefficients are difficult to measure (see Sec. 1.3.3). Nevertheless, they provide a way to identify

[10] Individual stars cannot be resolved for individual galaxies, except for our Milky Way.

[11] Stellar absorption lines are broadened by effects such as collisional broadening, thermal broadening or Zeeman effect.

[12] To obtain a correct estimate, one needs to subtract the recession velocity of the galaxy barycentre from the actual estimate. See Cappellari (2017).

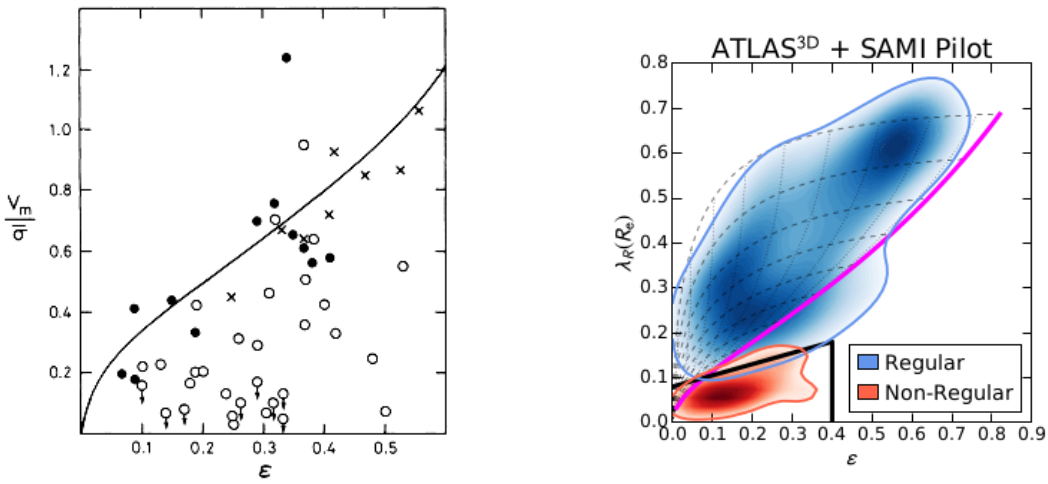


Figure 1.11: *Left:* the $(v/\bar{\sigma}, \varepsilon)$ diagram (modified from Davies et al. 1983). Massive and faint ETGs are shown as open and closed circles, respectively, while crosses are bulges of spirals. The solid line is the prediction for the isotropic rotator (Binney, 1978b). *Right:* the (λ, ε) (modified from Cappellari 2016). The blue and red regions enclose 80% of the total probability. The black line is the $\lambda = 0.08 + \varepsilon/4$ relation, used as threshold between the two galaxy classes, while the magenta line is the prediction for the isotropic rotator. The black dotted lines show how this relation depends on the assumed inclination through eq. 1.20.

features which may otherwise be difficult to spot. For example, a galaxy may host a disk which is invisible in projection (see previous section), but this generates a strongly negative $h_3 \leq -0.1$ (Scorza & Bender, 1995). Face-on disks also leave imprints in the LOSVD along the minor-axis, generating strongly positive $h_4 \geq 0.1$ (Magorrian, 1999).
An accurate reconstruction of the wings, in particular of the h_4 coefficient, is crucial in breaking the mass-anisotropy degeneracy (Gerhard, 1993), for which a radial orbit distribution with a constant Υ yields the same projected velocity dispersion as an isotropic orbit distribution with a Υ increasing towards the centre (Binney & Mamon, 1982). Moreover, a central BH also leaves imprints in the wings because of the high-velocity stars orbiting very close to it, meaning that not reconstructing the wings properly can yield biased estimates of $M_{\rm BH}$ when modeling a galaxy (cf. Sec. 1.4.5).

1.3.2 Kinematic properties

As we saw in Sec. 1.1.5, massive ETGs have very small rotational velocities. In that section, a separation between boxy and disky ellipticals based on kinematics properties was introduced. This is discussed here in more detail.
A very well known diagram is the $(v/\bar{\sigma}, \varepsilon)$ diagram. It compares the ratio of the maximum rotational velocity and the mean velocity dispersion with the ellipticity. For an isotropic, oblate rotator, $v/\bar{\sigma} \sim \sqrt{\varepsilon/(1-\varepsilon)}$ (Binney 1978b). The left panel of Fig. 1.11 reports this diagram, with

the expected prediction for the oblate rotator plotted as solid line. Since massive ETGs lie below the solid line, they cannot be flattened by rotation but rather by velocity anisotropy, while less luminous ETGs are consistent with the isotropic rotator.
This diagram has been revisited (Emsellem et al., 2011; Cappellari, 2016) using Integral Field Unit (IFU[13]) kinematics from large galaxy surveys such as ATLAS (Cappellari et al., 2011a) and SAMI (Bryant et al., 2015). This is plotted on the right side of Fig. 1.11. Here, an angular momentum parameter λ defined as

$$\lambda = \frac{\sum F_n R_n |V_n|}{\sum F_n R_n \sqrt{V_n^2 + \sigma_n^2}} \tag{1.19}$$

is introduced (Emsellem et al., 2007). F_n, R_n, V_n and σ_n are the mean flux, radial aperture, mean stellar velocity and velocity dispersion in the n-th spatial bin. For chaotic motions, $\lambda << 1$, while the opposite happens for ordered motions. Here, the threshold between anisotropy-flattened and rotationally-flattened galaxies is placed at $\lambda = 0.31\sqrt{\varepsilon}$ (Emsellem et al., 2011). More recently, Cappellari (2016) suggested $\lambda = 0.08 + \varepsilon/4$ for $\varepsilon < 0.4$ as threshold. This is also shown in Fig. 1.11. The (λ, ε) diagram allows us to add to the conclusions of above that rotationally flattened galaxies are mostly randomly oriented axisymmetric systems. In fact, for a general inclination

$$\frac{v}{\bar{\sigma}} = \sqrt{\frac{\varepsilon}{1-\varepsilon}} \frac{\sin i}{\sqrt{1 - \delta \cos^2 i}} \tag{1.20}$$

where i is the inclination angle introduced in Sec. 1.2.1 and δ is the global anisotropy (see Sec. 1.4.3). This explains the black lines in the right panel of Fig. 1.11.

1.3.3 Fitting methods

Fitting procedures which extract stellar kinematics need to deconvolve the LOSVD from the integrated spectrum (Bender 1990; Rix & White 1992; Cappellari & Emsellem 2004; Cappellari 2017; Falcón-Barroso & Martig 2021, Thomas et al. 2022, in prep.). Advanced methods for stellar fitting take into account sky subtraction as well as minimizing template mismatch by adding to eq. 1.18 additive polynomials and reduce the sensitivity of the fit from dust reddening by multiplying the convolved LOSVD of eq. 1.18 by orthogonal polynomials (for example Legendre polynomials). The degree of the polynomials must be chosen carefully. For example, h_4 is very sensitive to incorrect continuum subtraction, but using too high-order additive/multiplicative polynomials can artificially enhance the wings (Mehrgan et al. 2022, submitted).
To reproduce the observed spectra, large stellar libraries[14] are used. Two widely used libraries are the MILES[15] (Falcón-Barroso et al., 2011) library, and the more recent X-Shooter[16] library (Verro

[13] In contrast with long-slit spectroscopy, for which it is difficult to go beyond 1-2$R_e$because of the low signal-to-noise, IFU kinematics have the advantage of allowing the extraction of the kinematics up to 4R_e.

[14] A stellar library is a collection of stellar spectra. This should not be confused with an orbit library, which is a collection of stellar orbits and largely used in galaxy dynamics.

[15] `http://miles.iac.es/pages/stellar-libraries/miles-library.php`

[16] `http://xsl.u-strasbg.fr/`

et al., 2022). These libraries have resolutions which are typically much higher than the galaxy observations. Therefore, the template resolutions need to be artificially increased until they become consistent with the galaxy data; otherwise, the velocity dispersion will not be estimated correctly. In order to minimize template mismatch and obtain a reliable estimate of the h_3 and h_4 coefficients, a very large number (≥ 1000) of stars is needed, even if typically less than 40 stars are needed to acceptably fit a galaxy.
As for the deprojection, both parametric and non-parametric methods can be used to extract stellar kinematics. Parametric methods (e.g. Cappellari 2017) approximate the LOSVD with Gauss-Hermite polynomials, are therefore fast and can use a huge number of templates. Instead, non-parametric routines try to reconstruct the full LOSVD (Falcón-Barroso & Martig 2021; Thomas et al. 2022, in prep.). In this last case, the number of template must be small in order to reduce computational time and regularisation (see e.g. Lipka & Thomas 2021) must be used to rule out unphysical solutions, but the LOSVD (in particular the wings) can be reproduced with much more accuracy, which is crucial for galaxy modeling as discussed in Sec. 1.4.1. Notably, this has led to the discovery of the most massive SMBH so far (Mehrgan et al., 2019).

1.4 Dynamics

This section focuses on the main concepts of galaxy dynamics. It starts by introducing the collision-less Boltzmann and Jeans equation(s), which allow us to study stellar motions under the approximation that the stellar system is at equilibrium. After deriving the tensor-virial theorem, a short overview of stellar orbits in some notable gravitational potentials is presented. Finally, different techniques for modeling stellar orbits are presented, focusing on the Schwarzschild orbit superposition.

1.4.1 Collision-less Boltzmann equation

Given the huge number of stars in a galaxy, it is not possible to examine each orbit singularly. Instead, we can define a normalized distribution function $f(\boldsymbol{x}, \boldsymbol{v}, t)$[17] so that $f(\boldsymbol{x}, \boldsymbol{v}, t) d^3\boldsymbol{x} d^3\boldsymbol{v}$ is the probability that at a certain time t a randomly chosen star moves with a velocity in the range $d^3\boldsymbol{x} d^3\boldsymbol{v}$ around position $\mathbf{x}$ and velocity $\mathbf{v}$[18]. This distribution function obeys the collision-less Boltzmann equation (CBE)

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0 \tag{1.21}$$

where $\boldsymbol{\Phi}(\mathbf{x}, t)$ is the gravitational potential. This equation tells us that the distribution function around a given star does not change as a function of time. Thus, if in a certain spatial range the density is low but the stellar velocities are all similar to each other, then if the stars should come close to each other and thus increase the spatial density, the velocities will spread over a wider range.

[17] Quantities in bold are vectors.
[18] Recalling hamiltonian dynamics, $\boldsymbol{x}, \boldsymbol{v}$ are nothing else than the well known phase-space coordinates.

What makes the distribution function f important from a practical point of view is its link to observables. For example, the LOSVD is simply the fraction of stars within a certain velocity interval along the LOS. Moreover, eq. 1.21 together with Poisson's equation

$$\nabla^2\Phi = 4\pi G\rho \quad \text{with } \rho = \int f(\mathbf{x},\mathbf{v})d^3\mathbf{v} \tag{1.22}$$

where ρ is the probability density of finding a star at a certain position, links the gravitational potential with f itself. A system where eqs. 1.21 and 1.22 hold is said to be self-consistent.
Eq. 1.21 is very similar to the equation obeyed by an integral of motion, namely a function $F(\mathbf{x},\mathbf{v})$ for which $dF(\mathbf{x}(t),\mathbf{v}(t))/dt = 0$ along any orbit holds. This allows us to formulate Jeans theorem (Jeans, 1915):

Any function of the integrals of motion is a steady-state solution of the CBE, and any steady-state solution of the CBE depends on the phase-space coordinates only through integrals of motion.

The theorem proof is shown is App. B.2. According to it, if we manage to find any function only depending on one or more integrals, then this solves the CBE. However, not for every integral of motion an analytic expression can be written down ("non-classical" integral). A stronger statement (Lynden-Bell, 1962) is that if in a galaxy almost all orbits are regular, i.e. confined to invariant torii, then f depends on three independent isolating integrals, where an isolating integral reduces the motion in phase-space by one dimension.

1.4.2 Jeans equations

In practice, it is very difficult to solve eq. 1.21. Here we discuss how this can be done using moment equations. Other methods are described in Sec. 1.4.5. The 0-th moment equation of eq. 1.21, which can be obtained by integrating it over all velocities, reads

$$\frac{\partial\rho}{\partial t} + \frac{\partial\left(\rho\bar{v}_i\right)}{\partial x_i} = 0 \tag{1.23}$$

Eq. 1.23 is the analogous of the continuity equation and expresses the conservation of the probability density ρ defined in eq. 1.22. The first moment of eq. 1.21 can be written as

$$\rho\frac{\partial\bar{v}_j}{\partial t} + \rho\bar{v}_i\frac{\partial\bar{v}_j}{\partial x_i} = -\rho\frac{\partial\mathbf{\Phi}}{\partial x_j} - \frac{\partial(\rho\sigma_{ij}^2)}{\partial x_i} \tag{1.24}$$

which is the analogous of Euler's equations for fluids for ρ replacing the fluid velocity with the mean stellar velocity. The tensor $\sigma_{ij}^2 \equiv \overline{v_i v_j} - \bar{v}_i\bar{v}_j$ is called the velocity-dispersion tensor and describes the spread of stellar velocities around the mean velocity $\bar{\mathbf{v}}(\mathbf{x})$. The pressure applied on the stellar motions is described by the stress tensor $-\rho\sigma_{ij}^2$.
These equations are called Jeans equations, who was the first to apply them to stellar dynamics. However, these equations do not represent a closed set. In fact, we have a total of four equations: one is eq. 1.23 plus the three components of eq. 1.24. The problem here is that we have a

total of nine unknowns: the three components of the mean velocity $\bar{\mathbf{v}}$ and the six independent components[19] of σ_{ij}^2. Therefore, in order to solve the Jeans equations one needs to make assumptions about the geometry of the system. Two common assumptions are spherical and cylindrical symmetry. In the spherical case, a notable equation is obtained by multiplying the CBE in spherical coordinates by v_r:

$$\frac{\mathrm{d}(\rho\overline{v_r^2})}{\mathrm{d}r} + \rho\left(\frac{\mathrm{d}\mathbf{\Phi}}{\mathrm{d}r} + \frac{2\overline{v_r^2} - \overline{v_\theta^2} - \overline{v_\phi^2}}{r}\right) = 0. \tag{1.25}$$

We can now introduce the **anisotropy parameter** β

$$\beta = 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2}. \tag{1.26}$$

If $\beta > 0$, the model is said to have radial anisotropy, while if $\beta < 0$, we speak of tangential anisotropy. The intermediate case with $\beta = 0$ is the isotropic case. A relevant example of this is found in the central regions of the most massive galaxies: BH scouring ejects stars coming close to the galaxy centre via gravitational slingshot, generating a tangential anisotropy (Ebisuzaki et al., 1991; Thomas et al., 2014). Using eq. 1.26, eq. 1.25 can be recast as

$$\frac{\mathrm{d}(\rho\overline{v_r^2})}{\mathrm{d}r} + 2\frac{\beta}{r}\rho\overline{v_r^2} = -\rho\frac{\mathrm{d}\mathbf{\Phi}}{\mathrm{d}r} \tag{1.27}$$

and if one measures the radial velocity profile using spectroscopical information, then β can be calculated using 1.27. Accurate recovery of β is crucial to break the mass-anisotropy degeneracy. For example, tangential anisotropy can hide the presence of a SMBH, while radial anisotropy can reproduce the same effects that a central SMBH would have on the kinematics. In general, small variations of β can increase or decrease the central velocity dispersion, regardless of how large M_{BH} is (Binney & Mamon, 1982).

In the cylindrical case, it is convenient to define a (R, z, ϕ) coordinate system centered on the galaxy, with z being the vertical axis and (R, ϕ) on the symmetry plane. If we diagonalize the velocity-dispersion tensor and multiply the CBE by v_R we obtain

$$\frac{\partial\rho\overline{v_R^2}}{\partial R} + \rho\left(\frac{\overline{v_R^2} - \overline{v_\phi^2}}{R} + \frac{\partial\mathbf{\Phi}}{\partial R}\right) = 0 \tag{1.28}$$

while multiplication by v_z yields

$$\frac{\partial\rho\overline{v_z^2}}{\partial z} + \rho\frac{\partial\mathbf{\Phi}}{\partial z} = 0, \tag{1.29}$$

which constitutes a closed set of equations, given that v_R and v_z are the only two unknowns. This case is particularly useful to derive a robust estimate of the galaxy inclination, specified by only

[19]The velocity-dispersion tensor is manifestly symmetric.

one angle in the axisymmetric case. A widely used code implementation of the Jeans machinery for both spherical and cylindrical approximations can be found in Cappellari (2008).

1.4.3 Tensor-virial theorem

Starting from eq. 1.24 we can obtain, (for a full derivation see Binney & Tremaine 2008) multiplying it by x_i and integrating over the position space, the tensor-virial theorem:

$$\frac{1}{2}\frac{\mathrm{d}^2 I_{jk}}{\mathrm{d}t^2} = 2T_{jk} + \Pi_{jk} + W_{jk}, \tag{1.30}$$

valid under the assumption of self-gravitating systems. The term on the left-hand side is called moment of inertia tensor:

$$I_{jk} = \int \rho_m x_j x_k d^3\mathbf{x} \tag{1.31}$$

where ρ_m is the mass density. The first two terms on the right-hand side are

$$T_{jk} = \frac{1}{2}\int \rho_m \bar{v}_j \bar{v}_k d^3\mathbf{x}\ , \qquad \Pi_{jk} = \int \rho_m \sigma^2_{jk} d^3\mathbf{x} \tag{1.32}$$

and define the kinetic-energy tensor

$$K_{jk} = T_{jk} + \frac{1}{2}\Pi_{jk} \equiv \frac{1}{2}\int \rho_m \overline{v_j v_k} d^3\mathbf{x}. \tag{1.33}$$

Here, the term T_{jk} takes ordered motions into account, whereas Π_{jk} embeds the contributions from random motions. Finally, the term W_{jk} is called Chandrasekhar potential-energy tensor

$$W_{jk} = -\int \rho_m x_j \frac{\partial \mathbf{\Phi}}{\partial x_k} d^3\mathbf{x}, \tag{1.34}$$

whose trace yields the gravitational potential energy. Eq. 1.30 simply reduces to the well-known scalar virial theorem for systems in a steady state by taking the traces of the three tensors on the right-hand side.
An application of the tensor-virial theorem lies in its connection to observables. For example, in the case of an edge-on axisymmetric galaxy with z as short axis, thus with LOS perpendicular to the z-axis itself, then eq. 1.30 reduces to

$$\begin{aligned} 2K_{xx} + W_{xx} &= 0 \\ 2K_{zz} + W_{zz} &= 0 \end{aligned} \tag{1.35}$$

or equivalently, exploiting eqs. 1.32 and 1.33

$$\frac{\langle \bar{v}^2_{||} \rangle}{\langle \bar{\sigma}^2_{||} \rangle} = \frac{(1-\delta)W_{xx}/W_{zz} - 1}{\alpha(1-\delta)W_{xx}/W_{zz} - 1}, \tag{1.36}$$

where $\langle \bar{v}_{||}^2 \rangle, \langle \bar{\sigma}_{||}^2 \rangle$ are the v, σ components along the LOS, α is a scalar depending on how the stellar density ρ (eq. 1.22) and $\bar{\mathbf{v}}$ vary on the equatorial plane and δ is the **global anisotropy parameter**

$$\delta = 1 - \frac{\Pi_{zz}}{\Pi_{xx}} \tag{1.37}$$

analogous to the parameter β introduced in eq. 1.26. Given that $\langle \bar{v}_{||}^2 \rangle, \langle \bar{\sigma}_{||}^2 \rangle$ can be inferred from spectroscopical observations and that α varies weakly from one galaxy to the other (Binney, 2005), then by recovering the ratio W_{xx}/W_{zz} by deprojecting the SB profile of the galaxy one can derive the parameter δ from eq. 1.36.
Rewriting eq. 1.36 as

$$\frac{W_{xx}}{W_{zz}} = \frac{1}{1-\delta}\left(1 + \frac{1}{2}\frac{\langle \bar{v}_{||}^2 \rangle}{\langle \bar{\sigma}_{||}^2 \rangle}\right) \tag{1.38}$$

we immediately see that a galaxy can either be flattened by rotation ($\bar{v}_{||}^2 > 0$) or by anisotropy ($\delta > 0$), as we already saw in Sec. 1.1.5. In particular, in the case of a galaxy stratified on concentric, oblate ellipsoids, the ratio W_{xx}/W_{zz} only depends on the ellipticity ε (Roberts, 1962).

1.4.4 Motion in selected potentials

This section focuses on the stellar orbit in selected potentials describing systems which are commonly encountered in astrophysics: axisymmetric and triaxial galaxies. The results presented here refer to a closed system a equilibrium, where the energy is conserved. Moreover, we assume the distribution function f to be stationary.

Axisymmetric case

The axisymmetric approximation works well for fainter ETGs (cf. Sec 1.3.2). We use a cylindrical set of coordinates (R, z, ϕ) where z is the vertical and shorter axis and $R = \sqrt{x^2 + y^2}$, thus assuming the galaxy to be oblate. The potential $\mathbf{\Phi}(R, z)$ does not depend on the azimuthal angle ϕ, so Hamilton's equations read

$$\begin{aligned} \dot{p}_R &= \frac{p_\phi^2}{R^3} - \frac{\partial \mathbf{\Phi}}{\partial R} \\ \dot{p}_\phi &= \frac{\mathrm{d}}{\mathrm{d}t}\left(R^2\dot{\phi}\right) = 0 \\ \dot{p}_z &= -\frac{\partial \mathbf{\Phi}}{\partial z} \end{aligned} \tag{1.39}$$

where $p_R = \dot{R}$, $p_\phi = R^2\dot{\phi}$ and $p_z = \dot{z}$ are the momenta. We immediately see from the second equation that the z-component of the angular momentum, L_z, is a conserved quantity. Given that we already have the energy as integral of motion, the Hamiltonian is a function of two variables rather than four.
In particular cases, the trajectories on the surface of section[20] defined by $[H = E; z = 0]$ (thus on

[20] A surface of section (SOS) is a two-dimensional slice of the phase-space populated by stars.

the (R, p_R) plane) are closed curves (e.g. Thomas et al. 2004). This implies the presence of a third, non-classical integral of motion I_3 besides E and L_z. This integral constrains the orbit to short-axis tubes, i.e. with rotation in a fixed direction around the minor axis of the galaxy.
By defining the effective potential $\mathbf{\Phi}_{\rm eff} \equiv \mathbf{\Phi}(R, z) + \frac{L_z^2}{2R^2}$ we can write the following two equations

$$\begin{aligned} \ddot{R} &= -\frac{\partial \mathbf{\Phi}_{\rm eff}}{\partial R} \\ \ddot{z} &= -\frac{\partial \mathbf{\Phi}_{\rm eff}}{\partial z} \end{aligned} \tag{1.40}$$

which can be solved analytically if $\mathbf{\Phi}_{\rm eff}$ has particularly simple forms. Given that the Hamiltonian is the total energy of the system and that the kinetic energy is always positive, motions are only allowed in the region where $H \geq \mathbf{\Phi}_{\rm eff}$. The equality $H = \mathbf{\Phi}_{\rm eff}$ defines the zero-velocity curve, since the kinetic energy must be zero in this case.

Triaxial case

Massive ETGs often show photometric (e.g. isophotal twists) and kinematic (e.g. minor-axis rotation) features which exclude an axisymmetric geometry. In this case, the situation becomes more complicated. First, the only classical integral of motion is the total energy E of the system, possibly along with two non-classical integrals I_2, I_3, while the angular momentum is not conserved anymore along any direction. An analytic treatment like the one presented in the previous section is significantly more difficult, and we limit ourselves to a qualitative discussion of the possible orbits under the approximation of a galaxy stratified onto concentric ellipsoids. Analytic treatments can be found in Poon & Merritt (2001).
Let the galaxy shape be defined by a triaxial ellipsoid with axis lengths $a > b > c$. x, y, z are the major, intermediate and short axis, respectively. A first kind of orbits, which we cannot encounter in an axisymmetric system, is box orbits. These do not have a fixed rotation direction, but play a crucial role in supporting triaxial galaxies (Schwarzschild, 1979; Hunter & de Zeeuw, 1992). Then, we have tube orbits, which appear in a much broader variety in comparison with the axisymmetric case and have a fixed sense of rotation. Short-axis tubes rotate around the z-axis, while inner and outer long-axis tubes rotate around the x-axis. Instead, intermediate-axis tubes are orbits which rotate around the y-axis, but these orbits are unstable. Nevertheless, Neureiter et al. (2021) studied the time evolution of a simulated massive early-type galaxy, finding that intermediate-axis tubes can remain stable for up to ∼1 Gyr. Similar to box orbits are pyramid orbits, which have their major elongation parallel to the galaxy short axis.
The role of black holes is more important in triaxial gravitational potentials: in the case of axisymmetric potentials, the conserved angular momentum about the z-axis L_z places a constraint on the pericenter of stellar orbits, while in the triaxial case this condition does not hold (Gerhard & Binney, 1985). Therefore, many more stars can approach the central BH. In the region where the BH dominates the potential, i.e. when the pericenter r_p of stellar orbits is smaller than the BH gravitational radius $r_h = 2GM_{\rm BH}/\sigma^2$, the orbit space is populated by regular, pyramid orbits, Schwarzschild (1979); de Zeeuw & Merritt (1983), while at intermediate radii the orbits are mostly stochastic and fill the whole 5D space. The radius at which this transition occurs depends on the energy, the steepness of the density profile and (weakly) on the galaxy triaxiality

(Poon & Merritt, 2001). Finally, in the outermost region there is a mixture of stochastic and regular orbits, with box orbits being replaced by boxlet orbits such as resonant pyramid orbits ("banana-orbits") (Schwarzschild, 1993). A summary of possible orbits in a triaxial potential in illustrated in Fig. 1.12.

1.4.5 Dynamical modeling

Solving eq. 1.21 is of particular importance given how the solution f is linked to observables. However, f is, in the general case, a function of six variables: the three space and velocity coordinates x_i, v_i. Therefore, alternative methods are required to go around this issue. As for deprojection and the kinematics, both parametric and non-parametric methods can be used. An example of parametric method was presented in Sec. 1.4.2: by assuming a simple geometry (e.g. a spheroid or an axisymmetric galaxy) we can solve Jeans' equations and derive all variables of interest. Another parametric technique consists in assuming that all integrals of motion are classical (DF-based models), allowing to derive the galaxy properties of interest analytically. A notable case is again a galaxy stratified onto concentric ellipsoids. In this case, up to three classical integrals of motion can be found.
These methods are of limited applicability. Even if one solves Jeans' equations, there is no guarantee that the solution has physical meaning, while for the DF-based methods one needs to know the integral of motions. Instead, non-parametric methods can explore a much broader range of geometries and DFs. An example is the made-to-measure (M2M, Syer & Tremaine 1996a) method, which evolves in a self-gravitating, self-consistent N-body model. Here, the evolution of a large number of particles is studied by integrating the equations of motion of the particles, looking for the set of weights[21] which best reproduces the observables. A code implementation is provided by de Lorenzi et al. (2007).

Schwarzschild method

The technique which is most relevant for this thesis is the non-parametric Schwarzschild method of orbit superposition. Introduced in the pioneeering work of Schwarzschild (1979), this is a very general method which provides self-consistent solutions of the CBE and, in principle, does not require simplifying assumptions about the galaxy geometry. The keypoints are the following:

1. A trial gravitational potential is computed from the mass density by integrating Poisson's equation;

2. A stellar library[22] is integrated for the given potential for a time much larger than the crossing time;

3. A weight is assigned to each individual orbit, specifying the amount of light carried by the orbit itself. The orbits are combined and best solution for the weights is found.

[21] The weight of each particle quantifies the amount of mass of the particle itself.
[22] A stellar library is a collection of stellar orbits.

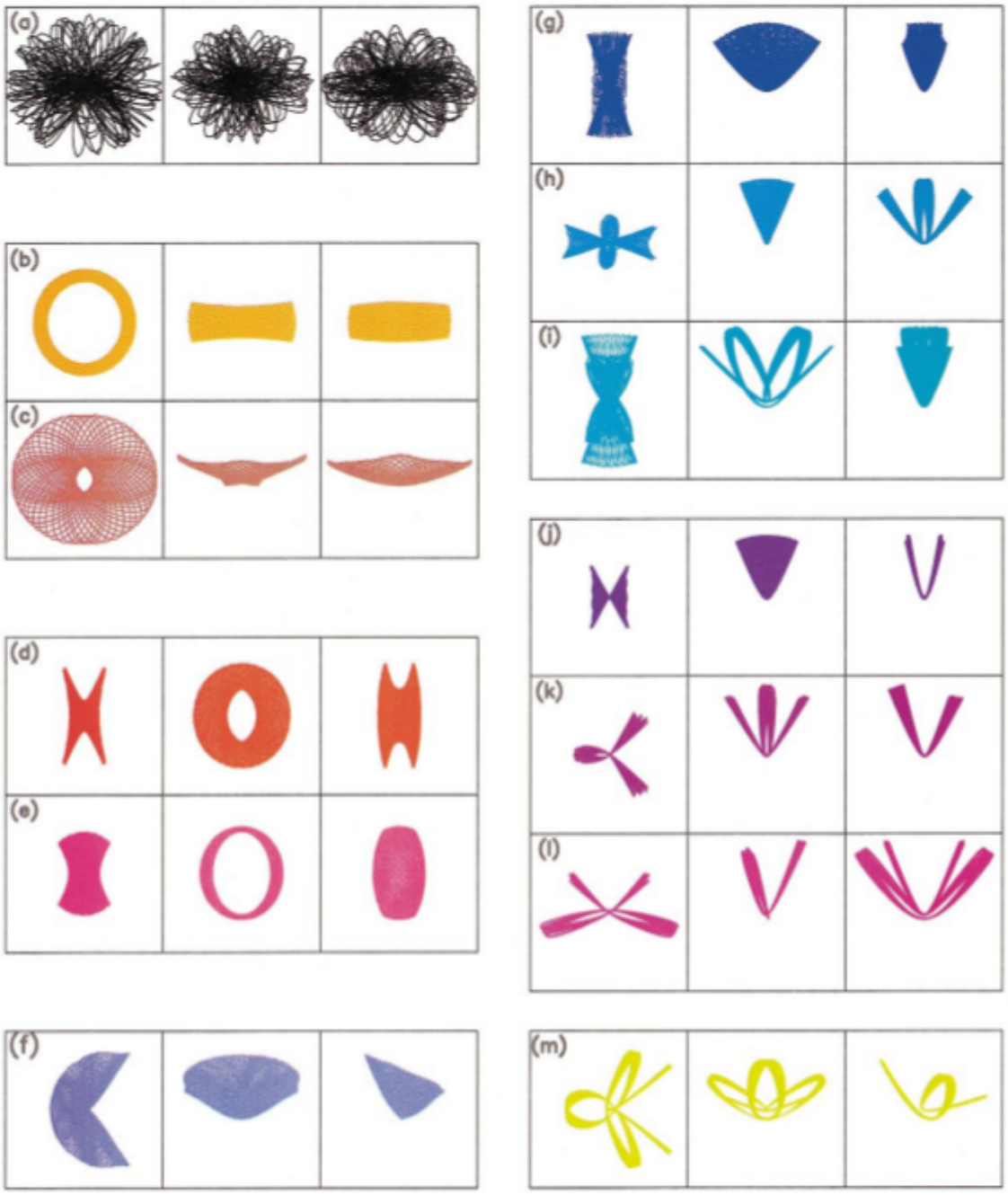


Figure 1.12: Collection of orbits in triaxial potentials with a central BH. (a): stochastic orbit. $(b-c)$: (resonant) short-axis tubes. $(d-e)$ inner/outer long-axis tubes. (g) pyramid orbit. (j) banana orbit. The other orbits $(f-h-i-k-l-m)$ are resonant pyramid orbits. For each orbit, the three panels show from left to right the projections onto the (x,y), (y,z) and (x,z) planes. Image from Poon & Merritt (2001).

The space occupied by the orbits is divided into cells, whose number is much smaller than the number of orbits. In this way, we have a large number of possible solutions, and a choice of an objective function to be minimized is necessary.
Starting from the pioneering work of Schwarzschild (1979) on orbit superposition technique, different codes have been developed to model galaxies. Implementations assuming spherical symmetry are Richstone & Tremaine (1984); Rix et al. (1997), while the widely used approximation of an axisymmetric galaxy has beed developed in van der Marel et al. (1998); Cretton et al. (1999); Gebhardt et al. (2003); Thomas et al. (2004); Valluri et al. (2004). Notably, this has led to publication of several BH mass estimates, with the new observations from the EHT confirming the robustness of this approach in the case of the galaxy M87 (Event Horizon Telescope Collaboration et al., 2019). Triaxial models are more difficult to handle, but codes (van den Bosch et al., 2008; Vasiliev & Valluri, 2020; Neureiter et al., 2021) have nonetheless been implemented and used in the last ∼15 years. The code `SMART` used in this thesis to derive the results presented in Chapter 4 was recently published (Neureiter et al., 2021). It works as follows:

- The trial gravitational potential is built as

$$\mathbf{\Phi} = M_{\rm BH} \times \delta(r) + \Upsilon \times \rho + \rho_{\rm DM} \tag{1.41}$$

  where $M_{\rm BH} \times \delta(r)$ is the contribution from the central BH, $\Upsilon \times \rho$ is the contribution from the stars and $\rho_{\rm DM}$ comes from dark matter. The stellar density is treated as a constraint, and can be recovered by deprojecting the observed SB profile assuming a certain galaxy orientation.

- `SMART` uses a 5D sampling space, which allows for an exhaustive sampling of all types of orbits encountered in a triaxial potential with a central BH (cf. Fig. 1.12) and, consequently, to explore a wide range of models compatible with the observations.

- The stellar orbits are combined and best solution for the weights is found by maximizing

$$\hat{S} = S - \alpha\chi^2. \tag{1.42}$$

  This is called maximum-entropy technique. In eq. 1.42, S is an entropy term related to the Shannon entropy[23], α is a regularization parameter controlling the strength of the entropy penalty and χ^2 compares the model and data LOSVDs:

$$\chi^2 = \sum_{i=1}^{N_{\rm losvd}} \sum_{j=1}^{N_{\rm vel}} \left(\frac{\mathrm{LOSVD}^{i,j}_{\rm model} - \mathrm{LOSVD}^{i,j}_{\rm data}}{\Delta\mathrm{LOSVD}^{i,j}_{\rm data}} \right). \tag{1.43}$$

- Schwarzschild models are noisy (see e.g. Binney & Tremaine 2008, Sec. 4.7.2.). It is important to estimate the smoothing accurately to avoid both overfitting and oversmoothing. `SMART` has been extended to include a novel technique to achieve this. To find the right value of α it minimizes

$$\mathrm{AIC}_p = \chi^2 + 2\mathrm{m}_{\rm eff}. \tag{1.44}$$

[23] Another choice (see e.g. Thomas et al. 2004 is the Boltzmann entropy.

This generalises the classical Akaike Information Criterion (AIC, Akaike 1974) by using the concept of effective free parameters ($m_{\rm eff}$, Lipka & Thomas 2021; Thomas & Lipka 2022). The higher the smoothing, the less flexible the model will be, thus resulting in a smaller number of effective free parameters. The number of effective parameters $m_{\rm eff}$ does not only vary for different smoothings but also from one mass model to another (Lipka & Thomas, 2021). A pure χ^2 minimization does not take into account the different number of Degrees-of-Freedom (DOF) each model has, and a reduced χ^2 is unreliable given the strong non-linearity of the models (Andrae et al., 2010).

- Steps 1-4 are repeated for different choices of $\mathbf{\Phi}$.

Neureiter et al. (2022, submitted) have shown that by combining `SHAPE3D`,`WINGFIT` and`SMART` it is possible to recover estimates of the mass parameters with unprecedented accuracy (within 10%). The question which remains open is: can this accuracy be reached also when attempting to constrain the galaxy intrinsic shape and its orbit distribution? As we saw in this chapter, this is crucial to break the mass-anisotropy degeneracy: I show in Chapter 4 that this is indeed the case.

1.5 This thesis

Regardless of the technique one chooses to model galaxies, the deprojected 3D light density is needed to build up the potential. Parametric methods such as the MGE (Cappellari, 2002) or fully non-parametric algorithms (e.g. de Nicola et al. 2020) are either not able to explore the degeneracy or cannot keep it under control. Moreover, the deprojection itself should rule out most of the implausible galaxy orientations: recovering the correct viewing angles using the dynamical modeling is too time consuming, but assuming the wrong galaxy orientation biases the results. A satisfactory algorithm able to deal with both issues did not exist at the time this work was started: the common assumption was that the galaxy orientation, and thus its intrinsic shape, could be recovered only with kinematical data, if at all. Only the average intrinsic shapes for galaxy samples have been investigated so far by inverting the observed flattening and misalignment distributions and making assumptions about the distribution of viewing angles of the galaxy population (Tremblay & Merritt, 1996; Weijmans et al., 2014; Foster et al., 2017; Ene et al., 2018). Finally, accurate deprojection, kinematics (Thomas et al 2022, in prep.) and dynamical modeling (Neureiter et al., 2021) codes can help to bring the errors on the shape and mass parameters found in recent works (e.g. Jin et al. 2019) down to 5-10%. This is what is needed to reduce the scatter in the BH-host scaling relations (Magorrian et al., 1998; Gebhardt et al., 2000; Saglia et al., 2016; Mehrgan et al., 2019) and to settle down the issue about the (non-)universality of the IMF (Thomas et al., 2011; Cappellari et al., 2012).
In this thesis I introduce a novel deprojection code (`SHAPE3D`) which lies in between the deprojection algorithms already found in the literature: it assumes the density to be stratified on concentric ellipsoids but recovers the shape non-parametrically, allows to explore the range of ellipsoidal shapes compatible with a certain photometry for a given set of viewing angles using a smoothing function to derive physically plausible light distributions and, at the same time, is able to find the *best-fit, unique* ellipsoidal deprojection. It can be used to quantify the goodness

of fit as a function of the viewing angles, which is not possible with the MGE. Also, the code shows that it is *possible to constrain the viewing angles with photometric data alone*.
I demonstrate the power of this algorithm by applying it to a sample of BCGs, which are particularly interesting since their properties differ from those of ordinary ETGs (Kluge et al., 2020, 2021), to derive the first *radially resolved* shape measurements using deep-photometric data and show how deprojecting BCGs Surface Brightness profiles could be used to investigate the shape of DM halos surrounding these galaxies and, possibly, the physics behind their formation. Finally, I combine my code with the novel triaxial Schwarzschild code `SMART` (Neureiter et al., 2021) and the non-parametric code for kinematics extraction `WINGFIT` (Thomas et al 2022, in prep.) to obtain the first complete triaxial, non-parametric pipeline for dynamically modeling massive ellipticals. I show that this allows to recover the right orientation, shape and orbit distribution of a simulated massive core-galaxy with unprecedented accuracy and present preliminary results of the modeling of the BCG NGC708.
The thesis is structured as follows:

- Chapter 2 presents the new deprojection code `SHAPE3D` along with several examples developed to test it, notably how the `SHAPE3D` recovers the right intrinsic density if the viewing angles are known and how the range of these can be significantly shrunk. A fully non-parametric deprojection algorithm is also presented, showing why this cannot be reliably used. In this chapter I explain how commonly observed phenomena such as twists can arise and show a scheme to construct triaxial density distributions which are invisible in projection. This is published as de Nicola et al. (2020).

- Chapter 3 presents an application of `SHAPE3D` on a sample of 56 BCGs. Here deep-photometry, extending out to 100 kpc data are used to derive radially resolved shape profiles, showing that BCGs are close to being maximally triaxial. A comparison with magneto-hydrodynamical simulation is carried out, showing that BCGs trace the properties of the DM halos surrounding them. This is published as de Nicola et al. (2022b).

- Chapter 4 presents the first application of our non-parametric pipeline (`SHAPE3D`, `WINGFIT` and `SMART`) in recovering the orientation, shape and orbit distribution of a simulated N-body massive elliptical, all of which are achieved with unprecedented accuracy, highlighting the requirements needed to obtain these results. Finally, I show that the deprojection alone suffices in recovering the galaxy shape, but the dynamical modeling allows to further increase the precision. This is published as de Nicola et al. (2022a).

- Chapter 5 presents preliminary results of the triaxial dynamical modeling of the NGC708, BCG of Abell 262. The approach presented in chapter 4 is here used to derive $M_{\rm BH}$ and Υ along with the anisotropy profile, the shape and the mass distribution. This will be the subject of a forthcoming paper.

- Finally, in chapter 6 I summarize the results of this thesis and draw my conclusions.

Chapter 2

Non-parametric Triaxial Deprojection of Elliptical Galaxies

2.1 Introduction

[1]Photometric observations of galaxies provide us with the surface brightness (SB) profile and the shapes of the isophotes, possibly as a function of wavelength. From here, we need to constrain the intrinsic, three-dimensional luminosity and, possibly, stellar mass density ρ, as a starting point to study the dynamics of galaxies. This step can be performed fitting the galaxy kinematics through the powerful Schwarzschild (1979) method, where the stellar mass density, along with a dark matter (DM) profile and a black hole (BH) mass, is used to build a potential by integrating the Poisson's equation through which the orbits are computed (Gebhardt et al., 2003; Thomas et al., 2004, 2005). Another potent dynamical modelling approach is the made-to-measure N-body technique (Syer & Tremaine, 1996b; de Lorenzi et al., 2007; Dehnen, 2009), where an N-body model is adapted to fit the data subject to any constraints. A less general but popular alternative is to solve the Jeans equation, typically assuming cylindrical symmetry (Cappellari, 2008).

Computer tomography solves the problem of reconstructing the three dimensional structure of a body by combining a number of two-dimensional projections taken at different angles covering a semi-circle. Astronomers have access to only one line of sight (LOS). For an axisymmetric system this means that the true density can be reconstructed only when the object can be assumed to be seen edge-on (Rybicki, 1987; Gerhard & Binney, 1996). In general, at any assumed inclination angle i (defined as the angle between the LOS and the equatorial plane) a 'cone of ignorance' of opening $90^\circ - i$ is generated in Fourier space, such that any density inside this cone will project to nothing along the assumed LOS. Such *conus densities* are unphysical on their own, since they are necessarily negative somewhere. However, to some extent, they make the deprojection at any assumed $i < 90^\circ$ non-unique. Gerhard & Binney (1996), van den Bosch (1997) and Kochanek & Rybicki (1996) discuss extensively conus densities. Gerhard (1996) considers the extension of the Fourier slice theorem (Rybicki, 1987) to the triaxial case, where the degeneracy of the

[1]Published paper: de Nicola S., Saglia R.P., Thomas J., Dehnen W., Bender R. - Non-parametric Triaxial Deprojection of Elliptical Galaxies, MNRAS, 2020, 496, 3076

problem is increased further, since only 4 planes in Fourier space are constrained by the measured surface brightness.

Although deprojecting SBs is a mathematically ill-posed question, a number of parametric and non-parametric approaches have been implemented to sample the space of possible three dimensional density distributions of galaxies. Parametric algorithms have the natural advantage of yielding smooth solutions and being fast, while non-parametric methods trade off naturally smooth solutions and very short computational time for an approach that can find a much broader family of solutions. In both cases, exploiting additional *statistical* information, like the ellipticity distribution determined from the observations of millions of galaxies on the sky, can help reducing the ambiguity of the deprojection of individual objects.

The most widely used parametric method is the Multi-Gaussian Expansion (MGE, Bendinelli 1991; Emsellem et al. 1994; Cappellari 2002). This routine can be directly applied to images and fits a SB distribution with a sum of Gaussians. This Multi-Gaussian *model* can be deprojected analytically for a given set of viewing angles (see the next section for their geometric definition), under the assumption that each 2D Gaussian component of the SB deprojects into a 3D Gaussian component of the density (see Section 2.4.1). This approach has several benefits, however, it just yields one deprojection per set of viewing angles, and there is no guarantee that this intrinsic deprojected density is correct. By construction, all these deprojections project *exactly* to the same Multi-Gaussian model, for the respective viewing angles. Thus, this approach does not allow to rank different deprojections based on their different relative likelihoods. Although the range of possible viewing angles is limited by the requirement that the axis ratios p and q should always follow the constraint $0 \leq q \leq p \leq 1$ and can be further constrained by assumptions about the minimal physically plausible intrinsic flattening etc.

Non-parametric algorithms are the best choice with respect to the intrinsic degeneracy issue. However, the ability to find any mathematically possible solution comes to the expense that these algorithms need some kind of penalized approach in order to keep the deprojection under control and filter out non-physical solutions. One well-working non-parametric axisymmetric algorithm is the one presented in Magorrian (1999), hereafter M99. This algorithm implements a penalized Metropolis Monte-Carlo algorithm that starting from an isophotal table deprojects the SB under the assumption of axisymmetry, also allowing for penalty functions that make the solution smoother or bias it towards a more boxy/discy shape. Being axisymmetric, the code has the limitation of not allowing for isophotal twist, a typical indication of triaxiality.

In this work, we implement a triaxial version of this fully non-parametric approach. Unlike for the axisymmetric case, finding suitable smoothing or penalty functions turns out difficult in the extension to triaxiality and we here follow a different approach to "penalize" the deprojection: we take advantage of the empirical fact that iso-density contours of massive ellipticals do not deviate strongly from ellipsoidal shapes. This suggests a smoothing towards ellipsoidal intrinsic shapes and we develop a constrained non-parametric tool, where the density is stratified onto deformed (discy-boxy) ellipsoids. We show that exactly ellipsoidal deprojections are unique when the line-of-sight (LOS) is known (and different from one of the principal axes). We use our ellipsoidal code to explore, for the first time, how tightly the viewing angles of triaxial objects can be constrained from surface photometry only.

The paper is organized as follows. Section 2.2 provides the mathematical background of

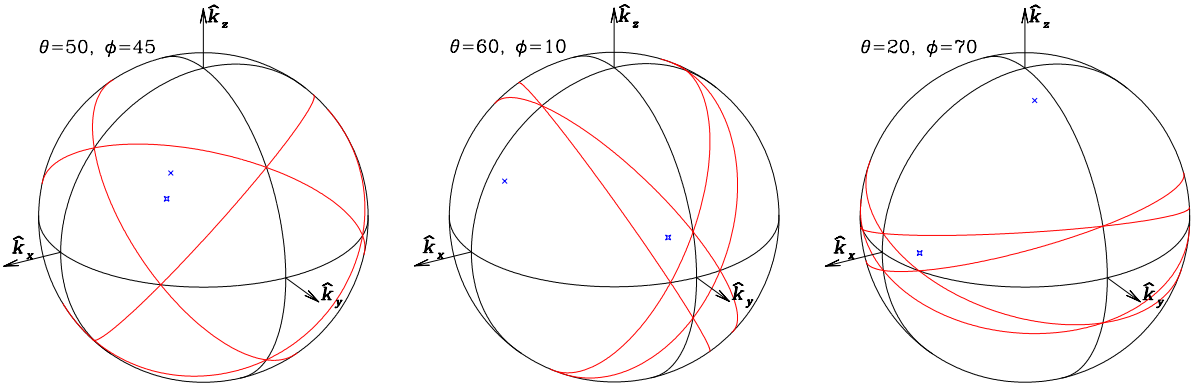


Figure 2.1: The projection along various LOSs (as indicated, also by the cross) constrains the 3D density only on four planes in Fourier space, shown as great circles (red) on a unit sphere. If the LOS is near a fundamental plane or axis, large parts of Fourier space are completely unconstrained. The planes dissect Fourier space into funnel-shaped regions: a four-sided funnel around each fundamental axis and a three-sided funnel between them in each octant around the reciprocal LOS (star).

the triaxial deprojection problem and introduces the concept of **cloaked densities**, the triaxial analogous of the konus densities encountered in the axisymmetric case. In Section 2.3 we show the details of the transformation of the M99's code from the axisymmetric to the triaxial case and discuss the degeneracies connected to non-parametric deprojections. Section 2.4 demonstrates that ellipsoidal deprojections of density distributions stratified on ellipsoidals are unique if the viewing angles are known and different from the principal axes. Section 2.4.2 presents the modifications implemented to constrain the solution on deformed ellipsoids. Section 2.5 explores the range of observables (ellipticities, position angle twists, a_4 coefficients) generated by projection effects along with illustrating the reliability of the algorithm, while Section 2.6 explores how tightly the viewing angles can be constrained. Section 2.7 compares the performances of our approach with the MGE strategy. Section 2.8 summarizes our findings and conclusions. The Appendices present a number of analytic *cloacked densities* (see Section 2.2) and discuss how to deal with the presence of discs.

2.2 The Fourier-slice theorem & cloaked densities

It is well known that the deprojection of an axisymmetric density is not unique, unless the object is viewed precisely edge-on (Gerhard & Binney, 1996). The deprojection of a triaxial galaxy is even less constrained: let $\rho(\boldsymbol{r})$ be a 3D density distribution of a transparent galaxy and

$$\Sigma(x, y) = \int \rho(\boldsymbol{r})\, \mathrm{d}z \tag{2.1}$$

its surface density when projected along the z axis. Now consider the Fourier transforms of both ρ and Σ:

$$\hat{\rho}(\boldsymbol{k}) = \int \rho(\boldsymbol{r}) \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, \mathrm{d}\boldsymbol{r}, \tag{2.2}$$

$$\begin{aligned}\hat{\Sigma}(k_x, k_y) &= \int \Sigma\, \mathrm{e}^{-i(k_x x + k_y y)}\, \mathrm{d}x\, \mathrm{d}y \\ &= \int \rho(\boldsymbol{r}) \mathrm{e}^{-i(k_x x + k_y y)}\, \mathrm{d}\boldsymbol{r},\end{aligned} \tag{2.3}$$

where equation (2.1) has been used. Thus (Fourier-slice theorem, Rybicki 1987),

$$\hat{\Sigma}(k_x, k_y) = \hat{\rho}(\boldsymbol{k})\big|_{k_z=0}. \tag{2.4}$$

Of course, nothing is special about the choice of the z-axis and a more general form of the theorem states that the Fourier transform $\hat{\Sigma}$ of the projection of ρ along the LOS direction

$$\boldsymbol{\ell} = (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta)^t \tag{2.5}$$

equals $\hat{\rho}$ in the plane $\boldsymbol{\ell}\cdot\boldsymbol{k} = 0$. If ρ is assumed/known to be triaxial, then so is $\hat{\rho}$ and knowledge of Σ constrains $\hat{\rho}$ on the four planes $\boldsymbol{\ell}_i \cdot \boldsymbol{k} = 0$ with

$$\boldsymbol{\ell}_1 \equiv \boldsymbol{\ell}, \quad \boldsymbol{\ell}_2 \equiv \begin{pmatrix} -\ell_x \\ \ell_y \\ \ell_z \end{pmatrix}, \quad \boldsymbol{\ell}_3 \equiv \begin{pmatrix} \ell_x \\ -\ell_y \\ \ell_z \end{pmatrix}, \quad \boldsymbol{\ell}_4 \equiv \begin{pmatrix} -\ell_x \\ -\ell_y \\ \ell_z \end{pmatrix}, \tag{2.6}$$

which are the reflections of $\boldsymbol{\ell}$ off the symmetry axes. These four planes dissect the space into distinct regions as depicted in Fig. 2.1. Of course, if $\rho(\boldsymbol{r})$ is triaxial, the conditions that $\hat{\rho}(\boldsymbol{k}) = 0$ on any one of these planes are mutually identical.

Hereafter, we denote a density distribution that projects to $\Sigma = 0$ at all sky positions a **cloaked density**. We also denote as **cloak** the set of LOS $\boldsymbol{\ell}$ such that the projection of $\rho(\boldsymbol{r})$ is invisible. For every non-trivial density, the cloak can only cover a small but possibly continuous set of directions.

A projection along of one of the principal axes provides the least amount of information and does not constrain the density distribution along those axes. On the other hand, a projection along a line of sight far from any principal axis does not constraint the Fourier transform $\hat{\rho}$ near these axes, such that distributions ρ with $\hat{\rho} \neq 0$ only around one principal axis are cloaked. Adding or subtracting such cloaked densities typically adds or subtracts a disc perpendicular to the respective principal axis. In Appendix 2.A we study cloaked densities in more detail and present several straightforward and elegant schemes for constructing them as well as near-invisible densities.
Some mathematical properties of these cloaked densities are as follows. Let $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$ be two cloaked densities and $f(\boldsymbol{r})$ an arbitrary function, then the following are also invisible when projected along $\boldsymbol{\ell}$.

1. Linear combinations of $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$, whereby the cloak shrinks to the intersection of the cloaks of ρ_1 and ρ_2;

2. any linear differential of $\rho_1(\boldsymbol{r})$ with respect to either $\boldsymbol{r}$ or any parameters (or both);

3. a convolution of $\rho_1(\boldsymbol{r})$ with $f(\boldsymbol{r})$;

4. a convolution of $\rho_1(\boldsymbol{r})$ with $\rho_2(\boldsymbol{r})$, whereby the cloak extends to the union of the cloaks of ρ_1 and ρ_2.

The high degree of degeneracy in the triaxial deprojection problem suggests to approach the problem in a non-parametric fashion. For the axisymmetric case, such code already exists (Magorrian, 1999) and has been successfully applied to many galaxies. In the next section we present our triaxial extension of the axisymmetric code of M99.

2.3 Non-parametric triaxial deprojection

2.3.1 Extension to the triaxial case

We start with a short overview of J.Magorrian's algorithm. The most significant points are:

- Both the SB and the intrinsic density are placed onto elliptical polar grids. In the axisymmetric case, a natural choice for the flattenings of the two grids is given by the inclination angle i and by the relation[2]

$$q' = \sqrt{q^2 \sin^2 i + \cos^2 i}, \;\; q = b/a \tag{2.7}$$

  where q' is the mean value of $1 - \varepsilon$, ε being the measured ellipticity.

- The program minimizes a likelihood function

$$\mathcal{L} = -\frac{1}{2}\chi^2 + P \tag{2.8}$$

  where χ^2 is given by

$$\chi^2 = \sum_{i=0}^{n_{m'}} \sum_{j=0}^{n_{\theta'}} \left(\frac{S_{ij} - \hat{S}_{ij}}{\Delta S_{ij}} \right)^2 , \tag{2.9}$$

  and P is a penalty term used to penalize against unsmooth solutions or to drift the solution towards a certain shape. In equation (2.9), S_{ij} and $\hat{S}_{ij}$ refer to the observed and the model SB, respectively, while ΔS_{ij} is the error coming from the observations. The grid has dimensions $n_{m'} \times n_{\theta'}$.

- Using a Metropolis algorithm (Metropolis et al., 1953), the program starts from an initial guess given by a double-power-law profile (2.16) to seek an intrinsic density projecting to a good fit to the observed SB profile.

[2]We use primes to denote coordinates and quantities defined in projection, i.e. on the sky.

In the triaxial case, some modifications are needed. First, we choose to represent the SB onto a grid of the form

$$x'_{ij} = m'_i \cos\theta'_j, \quad y'_{ij} = \eta m'_i \sin\theta'_j \tag{2.10}$$

where η is used to flatten the grid along y ($\eta < 1$), along x ($\eta > 1$) or to keep it circular ($\eta = 1$). Typically we sample m'_i with 50 points and θ'_j with 11 points from 0 to π.

The triaxial intrinsic density $\rho_{ijk} \equiv \rho(x_{ijk}, y_{ijk}, z_{ijk})$ is sampled onto an ellipsoidal grid of the form:

$$\begin{aligned} x_{ijk} &= m_i \sin\theta_j \cos\phi_k, \\ y_{ijk} &= P m_i \sin\theta_j \sin\phi_k, \\ z_{ijk} &= Q m_i \cos\theta_j. \end{aligned} \tag{2.11}$$

Hereafter we define $\mathcal{R} \equiv \log\rho$. The radial variable m_i ranges the semi-minor axis of the innermost isophote to a few (∼4) times the semi-major axis of the outermost isophote with typically 50 logarithmic steps, θ and ϕ go from 0 to $\pi/2$ with 11 linearly spaced steps, and P, Q are the two flattenings of the grid. P and Q can be chosen freely, their values do not influence the solutions discussed in Section 2.4.2, but have an impact on the computing time needed to achieve the final solution.

The two coordinate systems (x', y', z') and (x, y, z) are related by a rotation. Instead of using Euler angles, we follow the convention of Binney (1985) and de Zeeuw & Franx (1989) and use the polar coordinates $(\theta,\ \phi)$ of the LOS (2.5) plus a rotation in the plane of the sky to parameterise this coordinate transform. Then

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{R}_\psi \cdot \mathbf{P} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \text{with} \quad \mathbf{R}_\psi = \begin{pmatrix} \sin\psi & -\cos\psi & 0 \\ \cos\psi & \sin\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.12}$$

and the projection matrix (de Zeeuw & Franx 1989, equation 3.2)

$$\mathbf{P} = \begin{pmatrix} -\sin\phi & \cos\phi & 0 \\ -\cos\phi\cos\theta & -\sin\phi\cos\theta & \sin\theta \\ \cos\phi\sin\theta & \sin\phi\sin\theta & \cos\theta \end{pmatrix}. \tag{2.13}$$

The inverse transform is simply $\boldsymbol{r} = \mathbf{P}^t \cdot \mathbf{R}^t_\psi \cdot \boldsymbol{r}'$. ψ is the angle between the projection of the z-axis onto the sky and the x'-axis, measured counterclockwise, see also Fig. 2.2.

In the axisymmetric case, the orientation of the SB major axis determines the appropriate value for ψ, while owing to axial symmetry the angle ϕ has no effect, so that only the inclination $i = \theta$ is of importance. Conversely, for the triaxial case all three viewing angles must be considered.

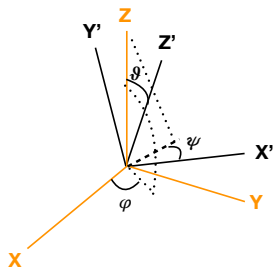


Figure 2.2: Geometric meaning of the viewing angles θ and ϕ, which determine the LOS direction $\hat{z}'$, and ψ, which is a rotation around the LOS itself.

The penalty function

The penalty function consists of two terms. The first one

$$
\begin{aligned}
P_{\rm sm} = & \frac{C}{\lambda_m} \sum_{i,j,k} \left[\frac{\mathcal{R}_{i+1,j,k} - 2\mathcal{R}_{i,j,k} + \mathcal{R}_{i-1,j,k}}{\Delta \log m} \right]^2 \\
& + \frac{C}{\lambda_\theta} \sum_{i,j,k} \left[\frac{\mathcal{R}_{i,j+1,k} - 2\mathcal{R}_{i,j,k} + \mathcal{R}_{i,j-1,k}}{\Delta\theta} \right]^2 \\
& + \frac{C}{\lambda_\phi} \sum_{i,j,k} \left[\frac{\mathcal{R}_{i,j,k+1} - 2\mathcal{R}_{i,j,k} + \mathcal{R}_{i,j,k-1}}{\Delta\phi} \right]^2 ,
\end{aligned} \tag{2.14}
$$

(with $C = -\sqrt{2 n_{m'} n_{\theta'}}$) penalizes un-smooth solutions and extends equation (9) of M99 to the triaxial case. We use typically values for λ_m between 0.5 and 1.2, which is up to an order of magnitude smaller than the default value $\lambda_m = 6$ for the axisymmetric code, while for λ_θ and λ_ϕ we usually adopt a value of 0.5 (Magorrian, 1999).

The second term of the penalty function

$$
\begin{aligned}
P_{\rm nn} = & \frac{C_{nn}}{n_m} \sum_{i,k} \max\left\{0, \left(\mathcal{R}_{i,2,k} - \mathcal{R}_{i,1,k}\right)\right\}^2 \\
& + \frac{C_{nn}}{n_m} \sum_{i,k} \max\left\{0, \left(\mathcal{R}_{i,n_\theta,k} - \mathcal{R}_{i,n_\theta-1,k}\right)\right\}^2 \\
& + \frac{C_{nn}}{n_m} \sum_{i,j} \max\left\{0, \left(\mathcal{R}_{i,j,2} - \mathcal{R}_{i,j,1}\right)\right\}^2 \\
& + \frac{C_{nn}}{n_m} \sum_{i,k} \max\left\{0, \left(\mathcal{R}_{i,j,n_\phi} - \mathcal{R}_{i,j,n_\phi-1}\right)\right\}^2 ,
\end{aligned} \tag{2.15}
$$

(with $C_{nn} = -4C$) generalizes equation (6) of M99 and penalizes models whose isocontours have negative $\partial\mathcal{R}/\partial\theta$ at $\theta = 0$ and $\pi/2$ (and the same for ϕ).

The Metropolis algorithm works in the same way as in Magorrian's code. The problem here is that since we go up one dimension, there will be a significant larger amount of points that shall be modified by the code. For instance, in the axisymmetric case we sample $\log \rho$ on a 50×11 grid, while in the triaxial case we take a $50 \times 11 \times 11$ grid. Since in the Magorrian's code *all* points of the SB grid are recomputed after each iteration, even a modest increase in the grid dimension leads to a significant increase in computational time. To speed up things, after the initial guess for $\mathcal{R}$ has been computed, we vary each $\mathcal{R}_{ijk}$ by a large factor, say 100, project the intrinsic density along the LOS and verify which points of the SB grid are actually varied by a factor larger than 0.1%. We tested that by using this mapping on the axisymmetric code we can get a triaxial `Python` code nearly as fast as the axisymmetric `C` code of M99.

Seeing Convolution

When the distance from the galaxy centre is significantly larger than the resolution of the observations, one can neglect point-spread-function (PSF) effects; however, when studying the central regions of a galaxy, correct BH masses can be derived only when the BH sphere of influence is well resolved and the PSF effects are taken into account (Rusli et al., 2013a). In our code, we added the option to perform the PSF convolution at every step of the Metropolis before comparing the projection to the observations. Typical dimensions of the (non-parametric) PSF matrix we use are about 100×100, but can be adapted to the specific photometric data; the PSF is supposed to be sampled from $-3\sigma_{\rm obs}$ to $3\sigma_{\rm obs}$, where $\sigma_{\rm obs}$ is the seeing of the observations. The PSF convolution is by far is the most time-consuming step and is the only step that has been parallelized. We postpone a detailed discussion of this part to the code to upcoming first applications to real galaxies.

2.3.2 Exploring non-parametric triaxial deprojections

A benchmark model

As a first step towards the testing of our deprojection algorithm, we consider as a benchmark a Jaffe (1983) model, which corresponds to the case $\alpha = 2$ and $\beta = 4$ of the double-power-law models (Binney & Tremaine, 2008, equation 2.64),

$$\rho(r) = \frac{\rho_0}{(r/s)^\alpha (1 + r/s)^{\beta-\alpha}}, \tag{2.16}$$

stratified on coaxial ellipsoids with specified radial profiles of the axis ratios. The values we chose for the total luminosity and the scale radius are $10^{10}\, L_\odot$ and $s = 1\,$kpc, whereas the grids (2.10) and (2.11) have dimensions 30×7 and $50 \times 11 \times 11$, respectively. The SB grid extends from 0.1 to 10 kpc with bin size $\sim 0.16\,$dex, while the ρ grid reaches out to 30 kpc and has bin size $\sim 0.115\,$dex. We modelled the p, q profiles to be cubic polynomials with coefficients such that p increases from ~ 0.55 to ~ 1 from the innermost to the outermost density contour, while q increases from ~ 0.45 to ~ 0.8, see Fig. 2.3. We also show the triaxiality parameter

$$T = \frac{1 - p(r)^2}{1 - q(r)^2}. \tag{2.17}$$

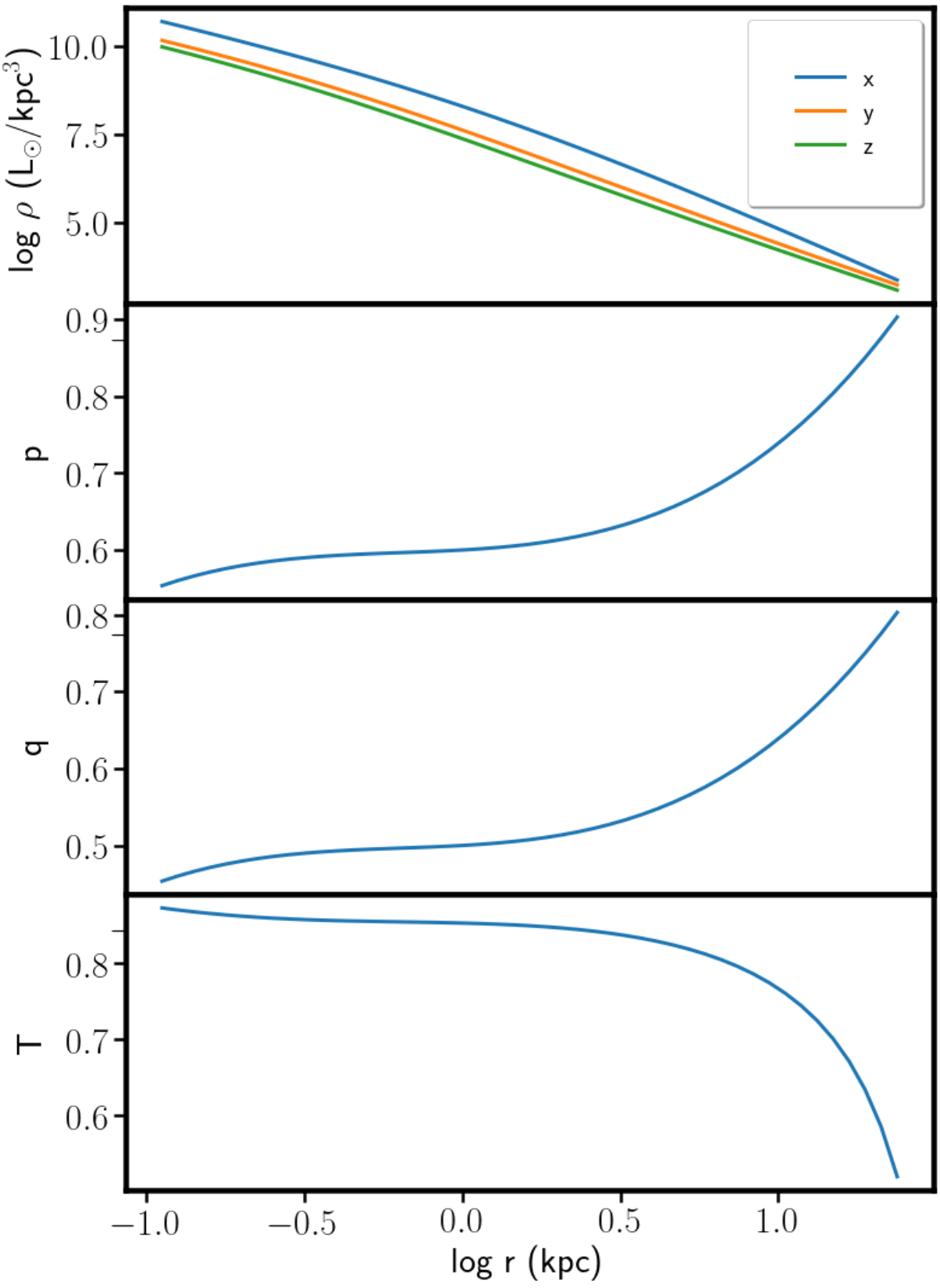


Figure 2.3: The *ELLIP* model of Section 2.3.2, see also Table 2.1. *From top to bottom:* radial profiles of the density along the principal axes, p, q, and the triaxiality parameter T (equation 2.17).

(Binney & Tremaine, 2008). This model, hereafter referred to as *ELLIP*, appears in several figures, listed in Table 2.1 (on page 53), which summarizes all models considered in this study. Although the p, q, T profiles of this model are probably not representative of bright ellipticals, their mean values $< q > \approx 0.6$ and $< T > \approx 0.7$ are in line with the observed ranges $0.6 \lesssim q \lesssim 0.8$ and $0.4 \lesssim T \lesssim 0.8$ (Tremblay & Merritt, 1996; Vincent & Ryden, 2005; Weijmans et al., 2014; Foster et al., 2017; Ene et al., 2018).

The SB of *ELLIP* is placed onto a grid with $\eta = 0.8$ (equation 2.10) while the flattenings of the ρ grid are $P = 0.7$ and $Q = 0.6$ (equation 2.11). These values for η, P, and Q will be used throughout the paper for all tests using a Jaffe density profile. As M99, we first compare the analytic expression of the Jaffe profile with our numerical projection, getting an RMS of ~0.03%, good enough for our purposes since it is smaller than typical uncertainties. The RMS can be simply obtained by multiplying equation (2.9) by ΔS_{ij}^2 and dividing the result by the number of grid points before taking the square root.

Due to the existence of cloaked densities, it is not a surprise that our non-parametric deprojection algorithm reconstructs a variety of densities, depending on many factors, such as the random seeds values and the shape of grid imposed by the choice of P and Q.

An example of a cloaked density

We illustrate this effect by deprojecting the projection of model *ELLIP* for $\theta = \phi = \psi = 45°$. In Section 2.4 we show that a density that is stratified on ellipsoids, such as *ELLIP*, admits a unique deprojection onto ellipsoids (unless it is viewed along one of the principal axes). But our non-parametric algorithm can find many more solutions (which are necessarily not stratified on ellipsoids), for whatever choice of viewing angles. Fig. 2.4, bottom, shows the percentage differences between one of these solutions obtained using the true viewing angles ($\theta = \phi = \psi = 45°$) and the true density along the three principal axes. In projection this model agrees with the true SB to a striking 0.0007% (Fig. 2.4, top), but differs from the true space density by up to 60% (Fig. 2.4, bottom). Moreover, it is physically plausible, when compared to the properties of low and high luminosity ellipticals. Re-projecting it along a variety of viewing angles generates SBs with ellipticity lower than 0.5, twists[3] τ exceeding 10° only for low ellipticities (Fig. 2.5, left) and higher-order shape coefficients a_4 (Fig. 2.5, right) and a_6[4] spanning the range observed in discy or boxy ellipticals Bender et al. (1988, 1989). However, for a few viewing directions (close to the z- or x axes: black and blue points in Fig. 2.5, right) the isophotes are discy, which would rule out such a model for massive ellipticals, which are only observed with elliptical and boxy isophotes (see discussion in Section 2.5.1).

Figs. 2.6 & 2.7 show the actual cloaked density, i.e. the difference between the true density and the deprojection solution. It is a flattened structure almost orthogonal to the z-axis with negative density at low z, reminiscent of a (reversed) disc, causing discy isophotes in projection when seen near the $z-$axis. A qualitative equivalent of the bottom plot of Fig. 2.6 (left) and the

[3]We define the *twist angle* τ as the maximal variation across the position angles of the isophote major axis.

[4]Throughout this study, we adopt the definition of Bender & Möllenhoff (1987) for the isophote shape coefficients, normalizing them to the major axis value a as $a_n \equiv a_n/a \times 100$.

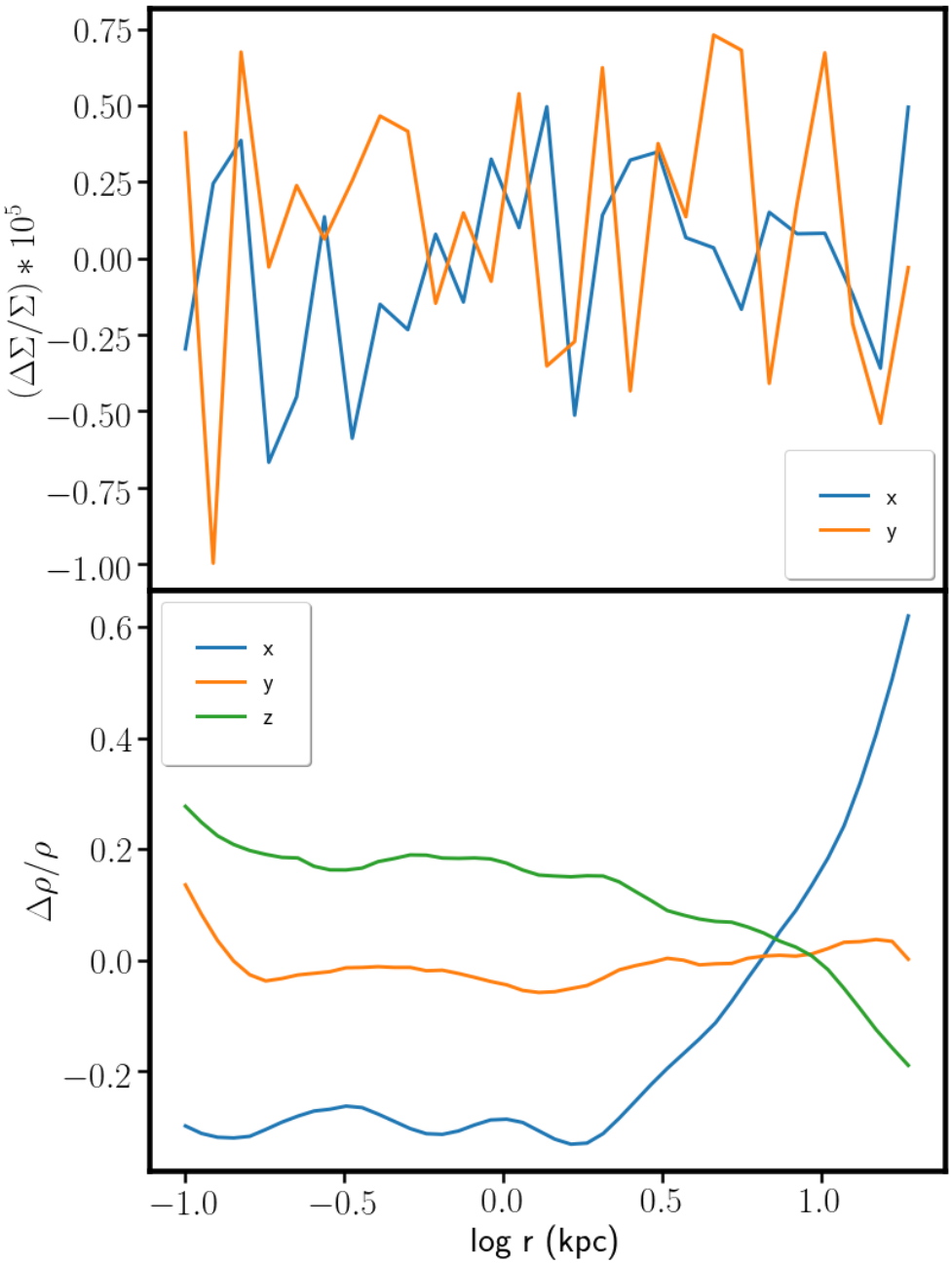


Figure 2.4: Relative differences between the true density along the principal axes (bottom) as well as SB (top) along the apparent major (blue) and minor (orange) axes of model *ELLIP* projected for $\theta = \phi = \psi = 45°$ and those obtained by our non-parametric deprojection. Although the fit to the observed SB is very good, the intrinsic density is far off the true value, a consequence of the non-uniqueness of triaxial deprojection.

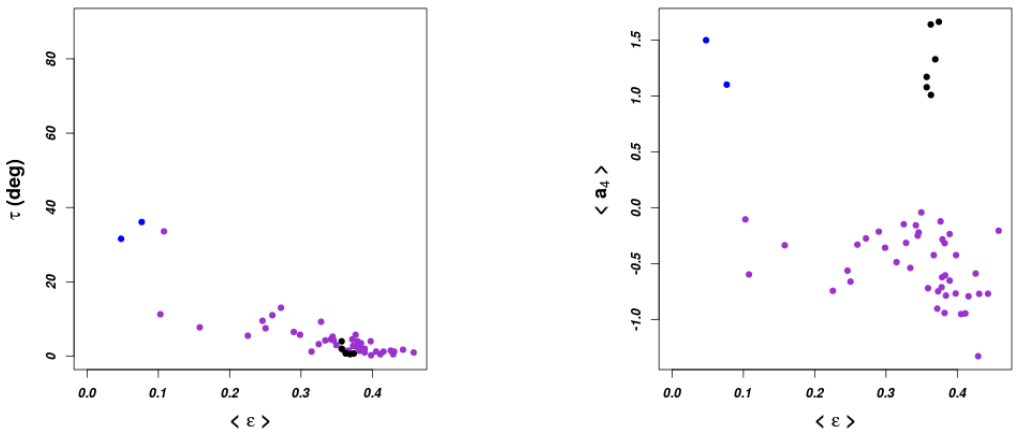


Figure 2.5: Correlations between mean ellipticity $< \varepsilon >$ and twist angle τ (*left*) and $< \varepsilon >$ and mean a_4 (*right*), when re-projecting an intrinsic density recovered by non-parametric deprojection for model *ELLIP*. The discy isophotes ($< a_4 >> 0$) obtained for some viewing directions are very unusual for massive ellipticals. In the right panel, black points have $(\theta,\ \phi) \leq (24, 22)^\circ$ (near the z-axis), while blue points have $\theta \approx 70^\circ$ and $\phi \approx 6^\circ$ (near the x-axis).

middle plot of Fig. 2.7 is given in Fig. 2.A.2 for one of the possible analytical descriptions of cloacked densities discussed in Appendix 2.A

This example shows that, although the code does its job very well in producing a good fit to the observations, an efficient mechanisms is needed to suppress solutions that are unrealistic for massive ellipticals.

2.4 Deprojection assuming approximately ellipsoidal isodensity contours

Although a non-parametric approach is desirable for exploring the broadest possible range of densities, it suffers from the large ambiguity in triaxial deprojections. Observationally, however, we know that the isophotes of massive elliptical galaxies do not deviate strongly from ellipses. This suggests, that the intrinsic density distributions of these galaxies are approximately ellipsoidal. As we will show in this Section, the assumption of ellipsoidal density distributions makes the deprojection problem a lot more tractable.

2.4.1 Ellipsoidal projection and deprojection

First note that the projection

$$\Sigma_s(R) = \int_{-\infty}^{\infty} \rho_s\left(\sqrt{R^2+z^2}\right) \mathrm{d}z = 2\int_R^{\infty} \frac{\rho_s(r)\, r\ \mathrm{d}r}{\sqrt{r^2-R^2}} \tag{2.18}$$

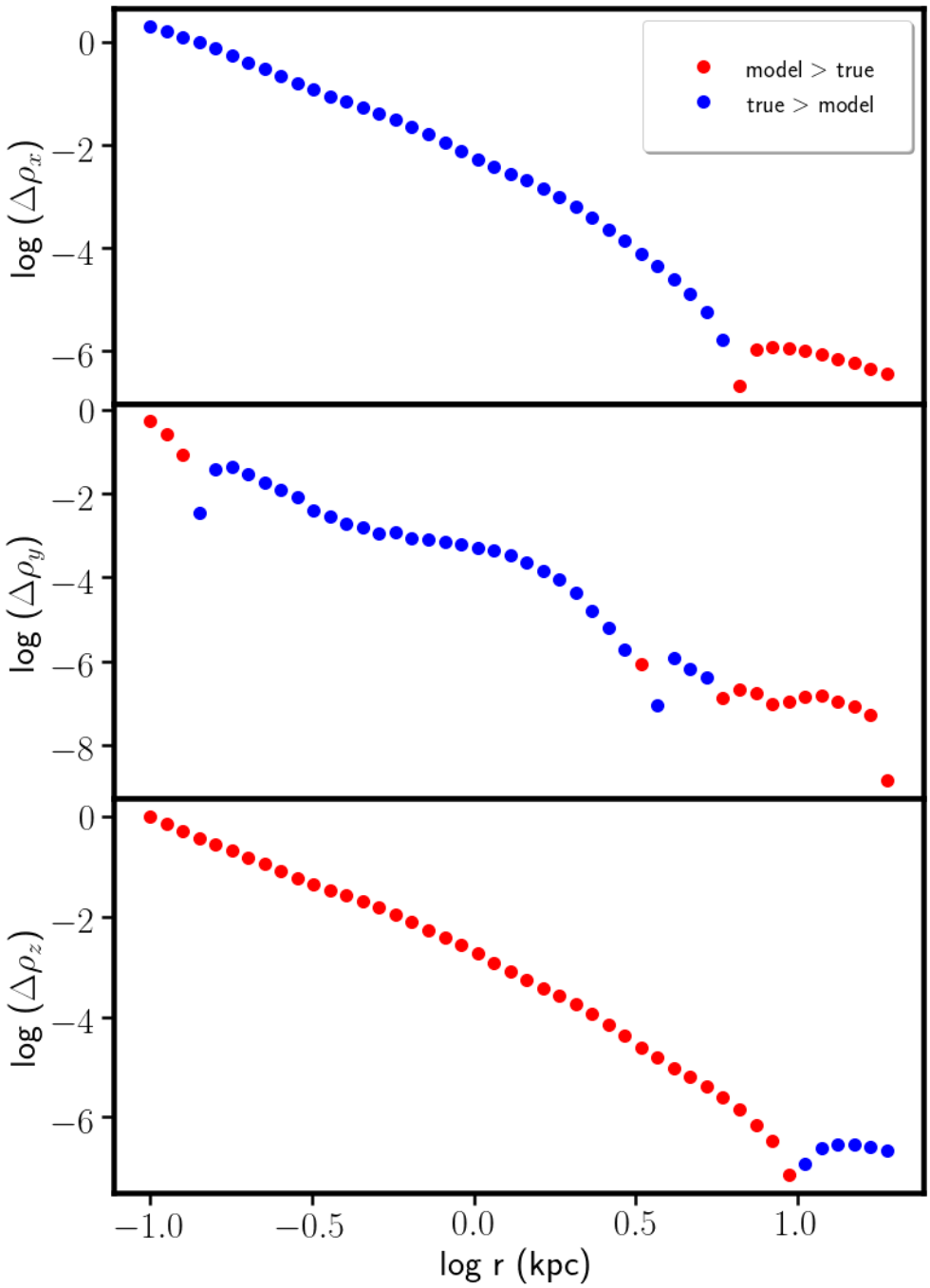
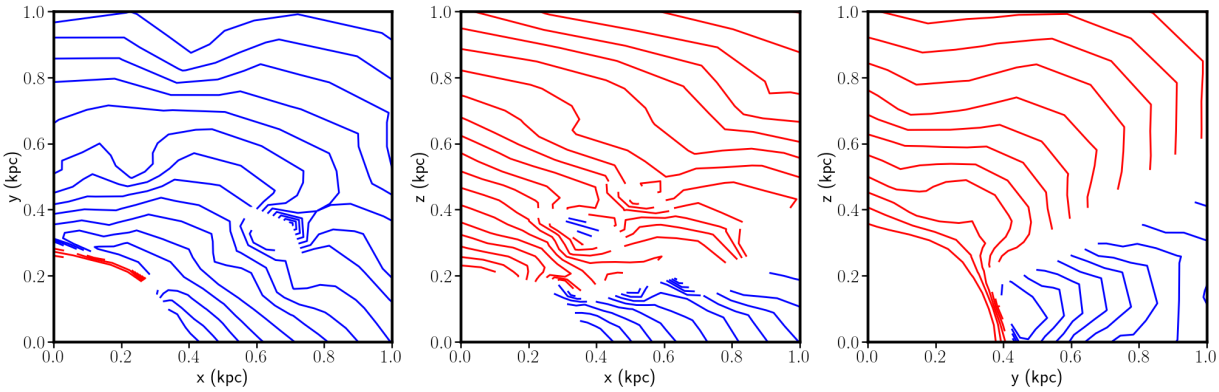
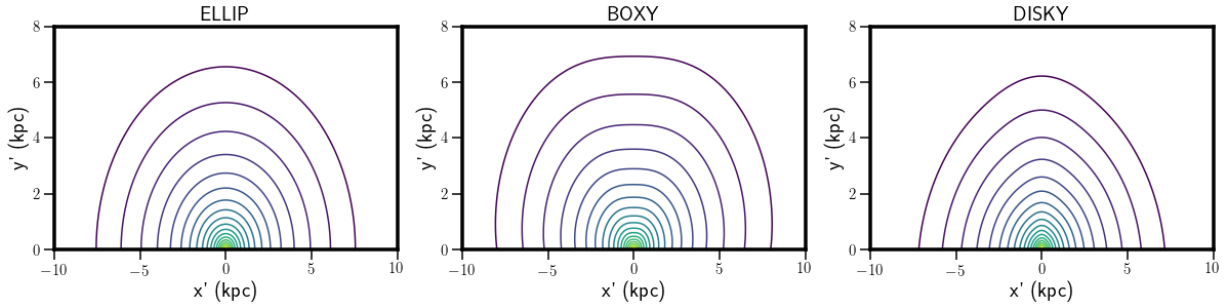


Figure 2.6: Radial profiles of the cloaked density hidden in the non-parametric deprojection of the *ELLIP* model (see Fig. 2.4) along the three principal axes (*top: major; middle: intermediate; bottom: minor*). Red/blue colours indicate that the recovered density is larger/smaller than the true density.



(a) (x, y) plane. (b) (x, z) plane. (c) (y, z) plane.

Figure 2.7: Contours of the cloaked density shown in Fig. 2.6 on the (x,y) plane (*left*), (x,z) plane (*middle*), and (y,z) plane (*right*). The contours line are red (blue) where the recovered density is larger (smaller) than the true.



(a) $\xi = 0$, exact ellipsoids. (b) $\xi = -0.5$, boxy bias. (c) $\xi = 0.3$, discy bias.

Figure 2.8: Jaffe models stratified on deformed ellipsoids (2.29) projected along the minor axis: model *ELLIP* with $\xi = 0$ (Fig. 2.8a), model *BOXY* with $\xi = -0.5$ (Fig. 2.8b) and model *DISCY* with $\xi = 0.3$ (Fig. 2.8c).

of a spherical galaxy with density $\rho_s(r)$ can (in principle) always be de-projected to obtain $\rho_s(r)$ via (e.g. Binney & Tremaine, 2008, problem 1.2)

$$\rho_s(r) = -\frac{1}{\pi}\int_r^\infty \frac{\mathrm{d}R}{\sqrt{R^2-r^2}}\frac{\mathrm{d}\Sigma_s}{\mathrm{d}R}. \tag{2.19}$$

Now, given parameters $a,\ b,\ c > 0$, an ellipsoidal version of the galaxy has density

$$\rho(\boldsymbol{r}) = (abc)^{-1}\rho_s(m) \tag{2.20}$$

with

$$m^2 = \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} = \boldsymbol{r}^t\cdot\mathbf{C}^{-1}\cdot\boldsymbol{r}, \tag{2.21}$$

where $\mathbf{C} \equiv \mathrm{diag}(a^2, b^2, c^2)$. Complement the LOS direction $\boldsymbol{\ell} = \hat{\boldsymbol{z}}'$ with the two perpendicular unit vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ spanning the sky, which we take to be the first and second rows of matrix $\mathbf{P}$ (2.13). The rotated coordinates are then

$$\tilde{\boldsymbol{r}} \equiv (\xi,\eta,\ell)^t = \mathbf{P}\cdot\boldsymbol{r} \tag{2.22}$$

and the projection integral becomes

$$\Sigma(\xi,\eta) = \int \mathrm{d}\ell\,\rho\left(\sqrt{\tilde{\boldsymbol{r}}^t\cdot\tilde{\mathbf{C}}^{-1}\cdot\tilde{\boldsymbol{r}}}\right), \tag{2.23}$$

where $\tilde{\mathbf{C}} \equiv \mathbf{P}\cdot\mathbf{C}\cdot\mathbf{P}^t$. Using the Fourier slice theorem[5] Σ can be expressed in terms of the spherical projection as

$$\Sigma(\xi,\eta) = \left|\bar{\mathbf{C}}\right|^{-1/2}\Sigma_s\Big((\xi,\eta)^t\cdot\bar{\mathbf{C}}^{-1}\cdot(\xi,\eta)\Big), \tag{2.26}$$

where the 2×2 matrix $\bar{\mathbf{C}}$ is the ξ-η part of $\tilde{\mathbf{C}}$ with components

$$\bar{\mathsf{C}}_{\xi\xi} = a^2\sin^2\phi + b^2\cos^2\phi, \tag{2.27a}$$

$$\bar{\mathsf{C}}_{\xi\eta} = (a^2-b^2)\sin\phi\cos\phi\cos\theta, \tag{2.27b}$$

$$\bar{\mathsf{C}}_{\eta\eta} = (a^2\cos^2\phi + b^2\sin^2\phi)\cos^2\theta + c^2\sin^2\theta \tag{2.27c}$$

[5]The Fourier transform of ρ is

$$\hat{\rho}(\boldsymbol{k}) = \int \rho(\boldsymbol{r})\,\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r} = \int \rho_s(m)\,\mathrm{e}^{-i\boldsymbol{\kappa}\cdot\boldsymbol{m}}\,\mathrm{d}\boldsymbol{m} = \hat{\rho}_s(\kappa)$$
$$\text{with}\qquad \boldsymbol{m}\equiv\mathbf{C}^{-1/2}\cdot\boldsymbol{r},\quad \boldsymbol{\kappa}\equiv\mathbf{C}^{1/2}\cdot\boldsymbol{k}. \tag{2.24}$$

Thus, $\hat{\rho}(\boldsymbol{k})$ is an ellipsoidal function, but with axis ratios inverted from those in real-space. According to the Fourier slice theorem

$$\hat{\Sigma}(k_\xi,k_\eta) = \hat{\rho}(\boldsymbol{k})\big|_{k_\ell=0} = \hat{\rho}_s\left(\sqrt{(k_\xi,k_\eta)^t\cdot\bar{\mathbf{C}}\cdot(k_\xi,k_\eta)}\right), \tag{2.25}$$

where the 2×2 matrix $\bar{\mathbf{C}}$ is the ξ-η part of $\tilde{\mathbf{C}}$, and equation (2.26) follows.

and

$$\begin{aligned} |\bar{\mathbf{C}}| &= |\mathbf{C}|\,(\boldsymbol{\ell}^t \cdot \mathbf{C}^{-1} \cdot \boldsymbol{\ell}) \\ &= a^2 b^2 \cos^2\theta + (a^2 \sin^2\phi + b^2 \cos^2\phi) c^2 \sin^2\theta. \end{aligned} \tag{2.28}$$

Thus, **the projection of an ellipsoidal density is elliptic**. Moreover if the LOS is not within any of the fundamental planes, then the parameters of the matrix $\mathbf{C}$ can be recovered from $\bar{\mathbf{C}}$: an elliptic surface density has a unique ellipsoidal density[6]. In contrast, for $\theta = 0$ (the projection along the minor axis) $\bar{\mathbf{C}}$ does not depend on c; for $\theta = \pi/2$ and $\phi = 0$ (the projection along the major axis) $\bar{\mathbf{C}}$ does not depend on a; for $\theta = \pi/2$ and $\phi = \pi/2$ (the projection along the intermediate axis) $\bar{\mathbf{C}}$ does not depend on b. Therefore, in these cases the deprojection of an elliptical SB onto an ellipsoidal density is not unique.

2.4.2 Developing the algorithm

Exploiting this result, we modified the fully non-parametric code in order to find the best solution on shells of a given shape (ellipsoids with possible boxy or discy deformations). Instead of searching for the density values on the three dimensional grid, we assume that four one-dimensional functions $\rho(x)$, $p(x)$, $q(x)$, and $\xi(x)$ of the distance x along the major axis describe at every point the density $\rho(x, y, z)$ to be stratified on shells of the form:

$$m^{2-\xi(x)} = x^{2-\xi(x)} + \left[\frac{y}{p(x)}\right]^{2-\xi(x)} + \left[\frac{z}{q(x)}\right]^{2-\xi(x)}. \tag{2.29}$$

$\xi = 0$ obtains perfect ellipsoids, while $\xi > 0$ and $\xi < 0$ give discy and boxy deformations, respectively. Fig. 2.8 shows examples of Jaffe models with different values of ξ: the $\xi = 0$ *ELLIP* case (left) described in the last section, the model *BOXY* with $\xi = -0.5$ (middle), and the model *discy* (right) with $\xi = 0.3$.

The algorithm makes random changes to the function values $\rho(m_i)$, $p(m_i)$, $q(m_i)$, and $\xi(m_i)$ on the radial grid m_i introduced in Section 2.3.1, and then uses linear interpolation (and extrapolation) in $\log r$ along each of the grid directions $(j,\ k)$ to update the density values on the three-dimensional grid. Changes that result in an intersection of density shells (2.29) for any of the grid directions are rejected.

The initialization follows closely Section 2.2 of M99: the SB is placed onto an elliptical polar grid and the initial guess for the intrinsic density is found by fitting an ellipsoidal double-power-law model (2.16) via the Levenberg-Marquardt algorithm with the axis ratios fixed to those of the grid ($p = P,\ q = Q$).

Then, for each Metropolis iteration the code randomly chooses one of $\mathcal{R} = \log\rho$, p, q or ξ as the variable X to change and applies a random change to one of its elements picked at random. As in M99, the change is made by setting $X_l \leftarrow X_l + r\Delta X_l$, where $r \in [-1, 1]$ is a uniformly

[6]This is the basis of the **multi-Gaussian deprojection** method, when $\Sigma(\xi, \eta)$ is decomposed into a superposition of elliptical Gaussian surface densities, each of which is then deprojected into the corresponding unique ellipsoidal Gaussian density.

Table 2.1: The models considered in this study and the figures where they feature.

Name	Property	Figures
ELLIP	Jaffe with $\xi = 0$	2.3, 2.8 (left), 2.12, 2.13, 2.18
DISCY	Jaffe with $\xi = 0.3$	2.8 (middle)
BOXY	Jaffe with $\xi = -0.5$	2.8 (right), 2.9,2.10
DISCYBOXY	Jaffe with ξ from 0.3 to -0.5	2.12, 2.13, 2.14, 2.15, 2.19, 2.22
PQCROSS	Jaffe with $\xi = 0$ and p & q profiles crossing	2.11
NBODY	N-body model	2.17, 2.20
LARGEDISC	50% Jaffe with $\xi = 0$ plus 50% disc (equation 2.57)	2.A.3, 2.B.3
SMALLDISC	85% Jaffe with $\xi = 0$ plus 15% disc (equation 2.57)	2.A.3, 2.B.2, 2.B.1

Table 2.2: The correspondence between p and q values and position of the principal axes of a galaxy according to the convention used in our model.

p and q values	Major	Intermediate	Minor
$q < p < 1$	x	y	z
$p < q < 1$	x	z	y
$q < 1 < p$	y	x	z
$1 < q < p$	y	z	x
$p < 1 < q$	z	x	y
$1 < p < q$	z	y	x

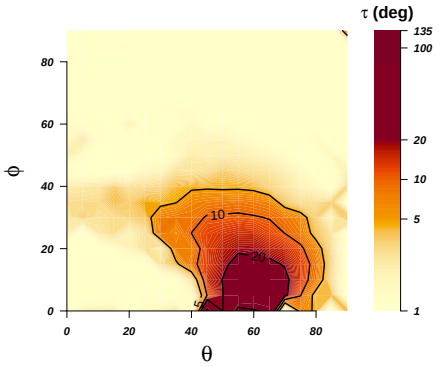
distributed random number, while the Δ arrays determine the maximum possible change. Initially, we set $\Delta\mathcal{R} = 0.5$, $\Delta\xi = 0.1$ and $\Delta p = \Delta q = 0.05$ for all l, but multiply (divide) an element by 1.5 after a change that was rejected (accepted).

Since the density contours are constrained to be deformed ellipsoids, the smoothness penalty function (2.14) is better expressed directly in terms of the four functions actually fitted. Therefore, we replace (2.14) with

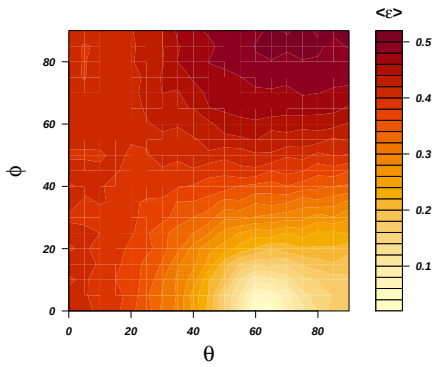
$$P_{\rm sm} = \frac{C}{\lambda_\rho}\sum_l\left[\frac{\mathcal{R}_{l+1}-2\mathcal{R}_l+\mathcal{R}_{l-1}}{\Delta\log m}\right]^2+\frac{C}{\lambda_p}\sum_l\left[\frac{p_{l+1}-2p_l+p_{l-1}}{\Delta\log m}\right]^2 + \frac{C}{\lambda_q}\sum_l\left[\frac{q_{l+1}-2q_l+q_{l-1}}{\Delta\log m}\right]^2+\frac{C}{\lambda_\xi}\sum_l\left[\frac{\xi_{l+1}-2\xi_l+\xi_{l-1}}{\Delta\log m}\right]^2, \tag{2.30}$$

where typically $\lambda_\rho \sim 10\lambda_\xi \sim 100\lambda_{p,q}$, when the four terms in (2.30) are of comparable magnitude.

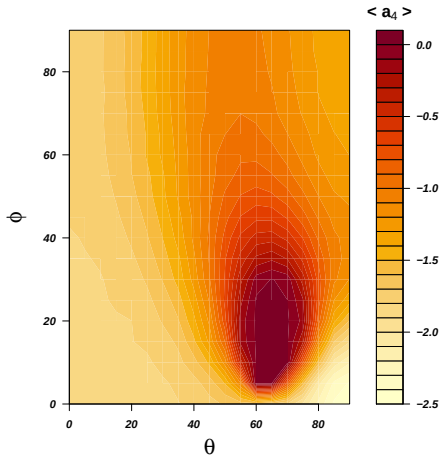
Since we cannot know which of the principal axes is major, minor or intermediate at any distance from the centre, we do not constrain the values of p and q a priori. Crossing p and q profiles result in changes of the relative axis ranking, summarized in Table 2.2. As discussed in Section 2.3.2, some p and q profiles generate strong twists when re-projected and can be discarded a posteriori as unlikely.



(a) Maximum twist τ.



(b) Mean ellipticity $< \varepsilon >$.



(c) Mean a_4 coefficient.

Figure 2.9: Maps of the twist angle τ, of the mean ellipticity $< \varepsilon >$ and of the Fourier coefficient a_4 as a function of the projection angles θ and ϕ for model *BOXY* described in Section 2.3.2, which has $\xi = -0.5$ (equation 2.29). In (a) contours are drawn for twists of 5°, 10°, and 20°. For octants not shown ($90° < \theta < 180°$ and $90° < \phi < 360°$) the results are identical by triaxial symmetry.

2.5 Tests of the ellipsoidal isodensity assumption for massive ellipticals

In this section we argue that the deformed-ellipsoidal model described in the previous section when projected on the sky is able to match reasonably well the general properties of massive ellipticals (see e.g. Foster et al. 2017; Goullaud et al. 2018; Kluge et al. 2020). Such galaxies have ellipticity distributed in the range $[0, 0.5]$ (unless sub-structures are present), tend to have slight boxy biases ($-1.5 \lesssim a_4 \lesssim 0$) and not very large twists ($\tau \lesssim 10°$). All these observables depend on the viewing angles θ and ϕ *and* on the p, q profiles (see e.g. equation 7-8 of Cappellari 2002), but not on ψ, since a rotation around the LOS rotates all isophotes by the same amount. Then, building on the result discussed in Section 2.4.1 about the uniqueness of density deprojections stratified on ellipsoids, we will verify the performances of our numerical algorithm.

2.5.1 The range of isophote shapes of triaxial elliptical galaxies

Reproducing real massive ellipticals

We consider the model *BOXY*, similar to *ELLIP* described in Fig. 2.3, but with $\xi = -0.5$, in order to reproduce the boxy bias observed in most massive ellipticals. We map the twist τ, the mean ellipticity $< \varepsilon >$ and the mean a_4 as functions of θ and ϕ (and assuming $\psi = 45°$). The results are shown in Fig. 2.9. For most of the angles θ and ϕ, the twist τ (Fig. 2.9a) is smaller than 5°; larger values are obtained when observing the model between the intrinsic long and the intrinsic short axes. At these viewing directions, the compression of the short axis relative to the intermediate axes near the centre tends to elongate the isophotes along the projected direction of the ("longer") intermediate axis. The compression of the short axis relative to the long axis in the outer parts likewise tends to elongate the isophotes, again along the direction of the "longer" axis. However, this time, the "longer" axis is the intrinsic long axis and therefore points to different projected direction. This gives rise to isophote twists, which depend on the exact profiles of $p(r)$ and $q(r)$. As pointed out in Section 2.3.2, the ones of the *ELLIP* model are not necessarily representative of bright ellipticals, but illustrative of how twists can be generated. Moreover, their mean values match the observed ones reasonably well: the mean ellipticity (Fig. 2.9b) is indeed in the range 0-0.5 and is roughly *anti-correlated* with the twist, being lower when the twist is higher. As expected, it reaches the highest values for high θ and ϕ, thus close to projections along the intermediate axis. Finally, the mean a_4 (Fig. 2.9c) spans the range -2.5-0.1, which is what we expected given the ξ profile we have chosen.

As a final check, in Fig. 2.10 we show the analogous of Fig. 2.5 for the model considered here: similarly to what observed by e.g. Bender et al. (1988, 1989); Foster et al. (2017) for real bright ellipticals, Fig. 2.10 shows that the strongest twist happens where the isophotes are rounder and that the mean a_4 becomes more negative (i.e., the isophothes are boxier) as the isophotes become more flattened. Differently from what discussed in Fig. 2.5, here we do not find any viewing directions yielding strongly discy isophotes.

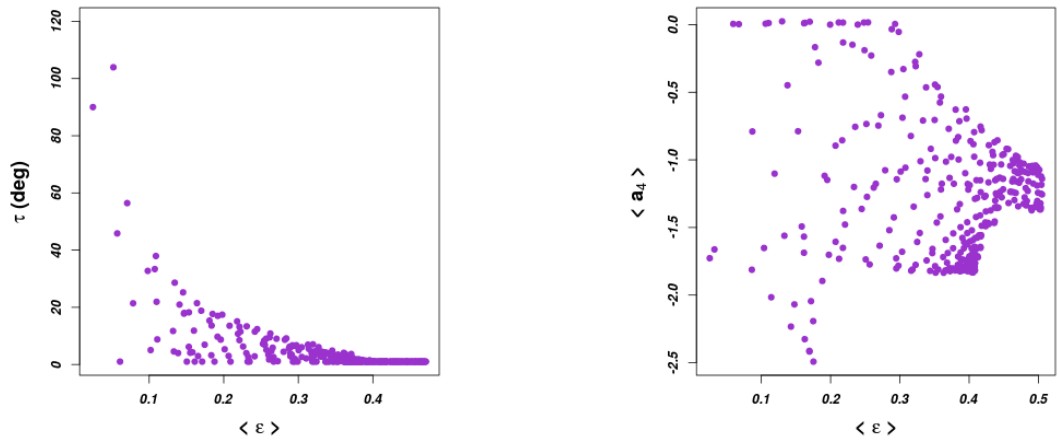


Figure 2.10: Same as Fig. 2.5 for model *BOXY*. No particularly discy projections are present and the distributions are qualitative consistent with what observed for massive ellipticals Bender et al. (1988, 1989).

A case with strong twist

Strong twists can be obtained when the orientation of the intrinsic long axis changes with radius. Fig. 2.11 shows the mapping of τ for model *PQCROSS*: this is similar to model *ELLIP*, but with p decreasing linearly with $\log(r)$ from 1.3 to 0.6 and q increasing linearly in $\log(r)$ from 0.6 to 1.3. In this case the intrinsic long axis of the model is along the y-direction in the inner regions, but along the z-direction at large radii; the model is near-spherical in the transition region. Since the orientation of the long and short axes changes with radius, there is now more than one region in the viewing-angle plane, where twists can occur.

The comparison of Figs. 2.9a and 2.11 illustrates how the expected occurance rate of isophote twists is closely related to the direction stability of the long and short axes in triaxial galaxies. The above examples suggest that if the orientation of the intrinsic long axes would change with radius in many real massive galaxies, then we should observe strong isophote twists very often. While such strong twists indeed exist in individual galaxies (e.g. Mazzalay et al. 2016), they are not characteristic for massive elliptical galaxies as a class (Kluge et al., 2020; Ma et al., 2014; Goullaud et al., 2018). Thus, in the following we will often restrict the analysis to the case of $p(r) > q(r)$. However, our code can also deproject without this condition, to cover individual galaxies where strong twists may be real.

In summary, an ellipsoidal density distribution with generically $q(r) < p(r)$ as just described qualitatively reproduces the observed properties of massive elliptical galaxies for any random projection angles.

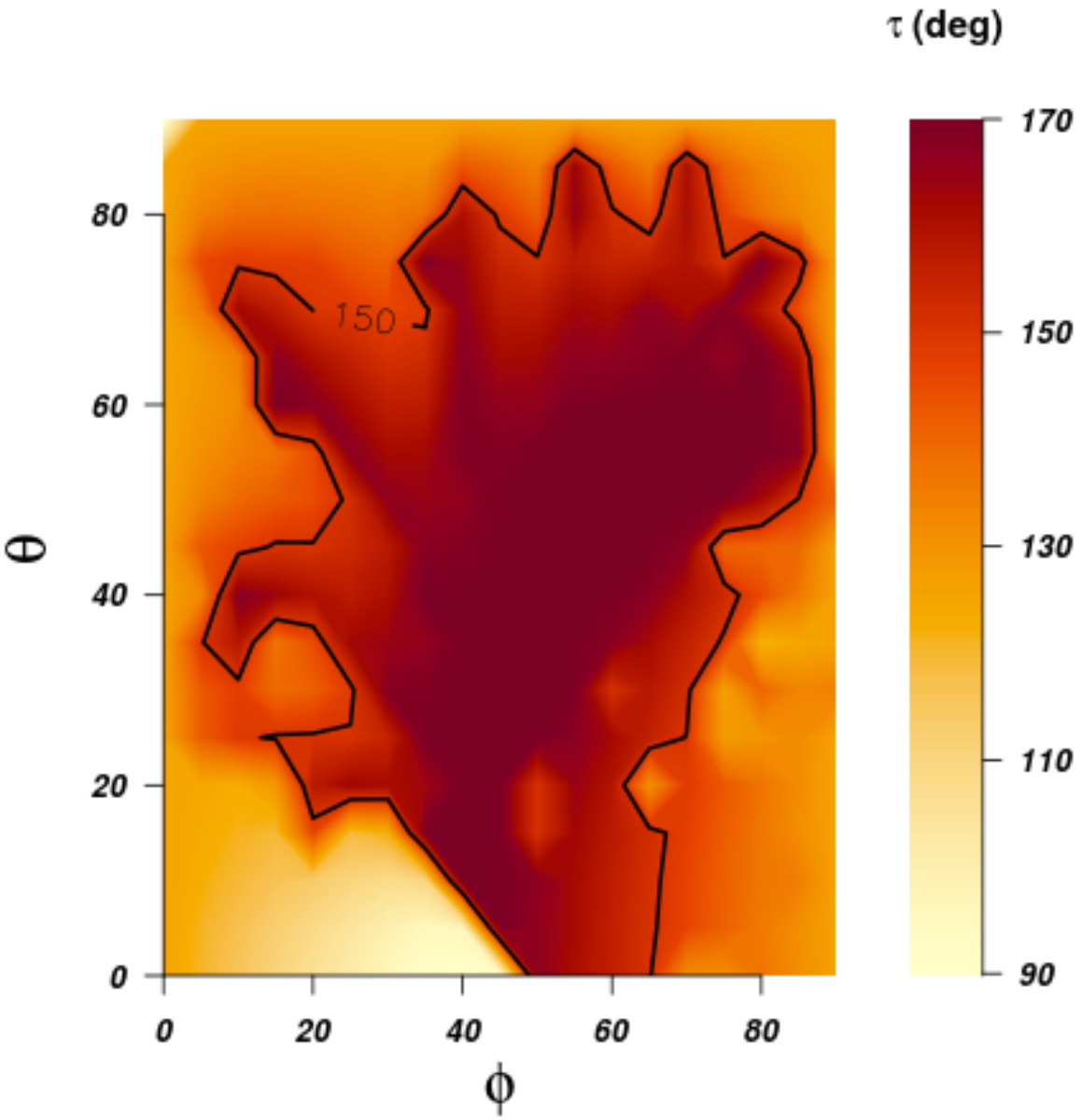
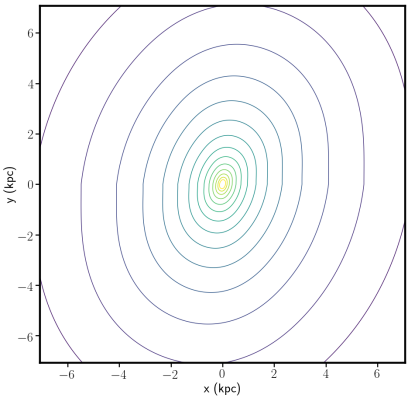
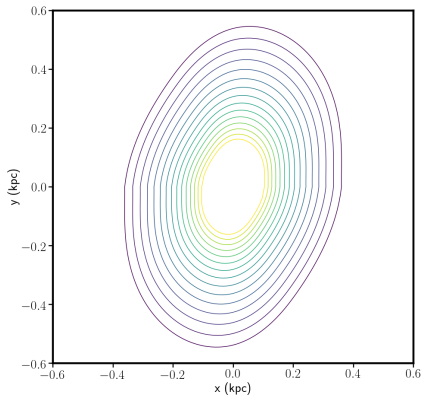


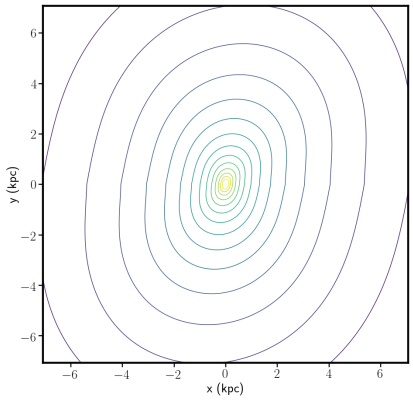
Figure 2.11: Map of the twist angle τ for model *PQCROSS* (see Section 2.5.1).



(a) *ELLIP*.



(b) *DISCYBOXY*, inner regions.



(c) *DISCYBOXY*, outer regions.

Figure 2.12: Contours of the projected SB (using $\theta = \phi = \psi = 45°$) of *ELLIP* (left panel) and *DISCYBOXY* (middle and right panels). *DISCYBOXY* has discy isophotes near the centre and boxy ones in the outer regions. The contour are colour reflects the SB value.

Hidden discs

In the axisymmetric case, much of the deprojection degeneracy can be traced back to disc-like conus densities, which become quickly unidentifiable when the inclination is far enough from edge-on. Lower-luminosity elliptical galaxies often show discy isophotal distortions and are intrinsically flattened strong rotators; for this class of objects consideration of embedded axisymmetric discs is indeed important. But for massive ellipticals, less flattened and mainly boxy objects, discs are either not present or contribute very little to the total density and therefore in the following we will ignore them. We discuss the effects of the superposition of an axisymmetric disc and a triaxial spheroidal body in Appendix 2.B. There we also show that the presence of important discs also tends to produce strong isophote twists that are not observed in most massive ellipticals, as discussed above.

2.5.2 Testing deprojections with constrained shapes

We consider the *ELLIP* Jaffe model described in Section 2.3.2 with $\xi = 0$ and the *DISCYBOXY* model, where ξ decreases linearly $\log(r)$ from 0.3 to -0.5, i.e. discy in the innermost regions and boxy outside, and project them along the direction $\theta = \phi = \psi = 45°$. Fig. 2.12 shows their surface brightness contours, and Fig. 2.13 their ellipticity, PA and a_4 profiles. First, we provide the code with the correct viewing angles, p, q and ξ profiles and let it search for the density profile. Secondly, we let the code also search for p, q non-parametrically, starting from an initial guess of $p = q = 1$ across the whole grid. Finally, we let the code recover also the ξ profile for *DISCYBOXY*

We show in Fig. 2.14 the results of the deprojection for this last case, i.e. when all parameters need to be recovered. In the top panel, the three lines are the percentage differences between the true model and what the code reconstructs, computed along the principal axes. Also shown are the true p, q and ξ profiles superimposed with the reconstructed ones. Finally, in the bottom panel we show that the fit to observed SB is excellent. In all cases, the density is recovered well, within an accuracy of 1%, out to the maximum radius sampled by the SB and down to a radius of the order of the resolution of the grid. For the very innermost and outermost points, all profiles start to deviate significantly from the true shapes. This is mostly due to the extrapolation to large radii: we have repeated the test highlighted in Fig. 2.14 stopping the radial grid first at 20 kpc and then extending it out to 80 kpc. In Fig. 2.15 we superimpose these results to those obtained for the grid extended out to 30 kpc, showing that the point at which the radial profiles start becoming unreliable also decreases. The last inner reliable point is set by a combination of radial extent and resolution of the grid. We will discuss it in detail in a future paper in combination with the PSF convolution.

Finally, provided that many ellipticals have cores, we also tested our algorithm with a Hernquist (1990) model (equation 2.16 with $\alpha = 1$ and $\beta = 4$) using the same parameters as above but with $\alpha = 1$ (see Section 2.3.2). We show in Fig. 2.16 that we do not find significant differences with the results presented above.

These findings go even beyond what stated in Section 2.4, namely that the density can be uniquely recovered if it is stratified on perfect ellipsoids and we know the viewing angles. Here

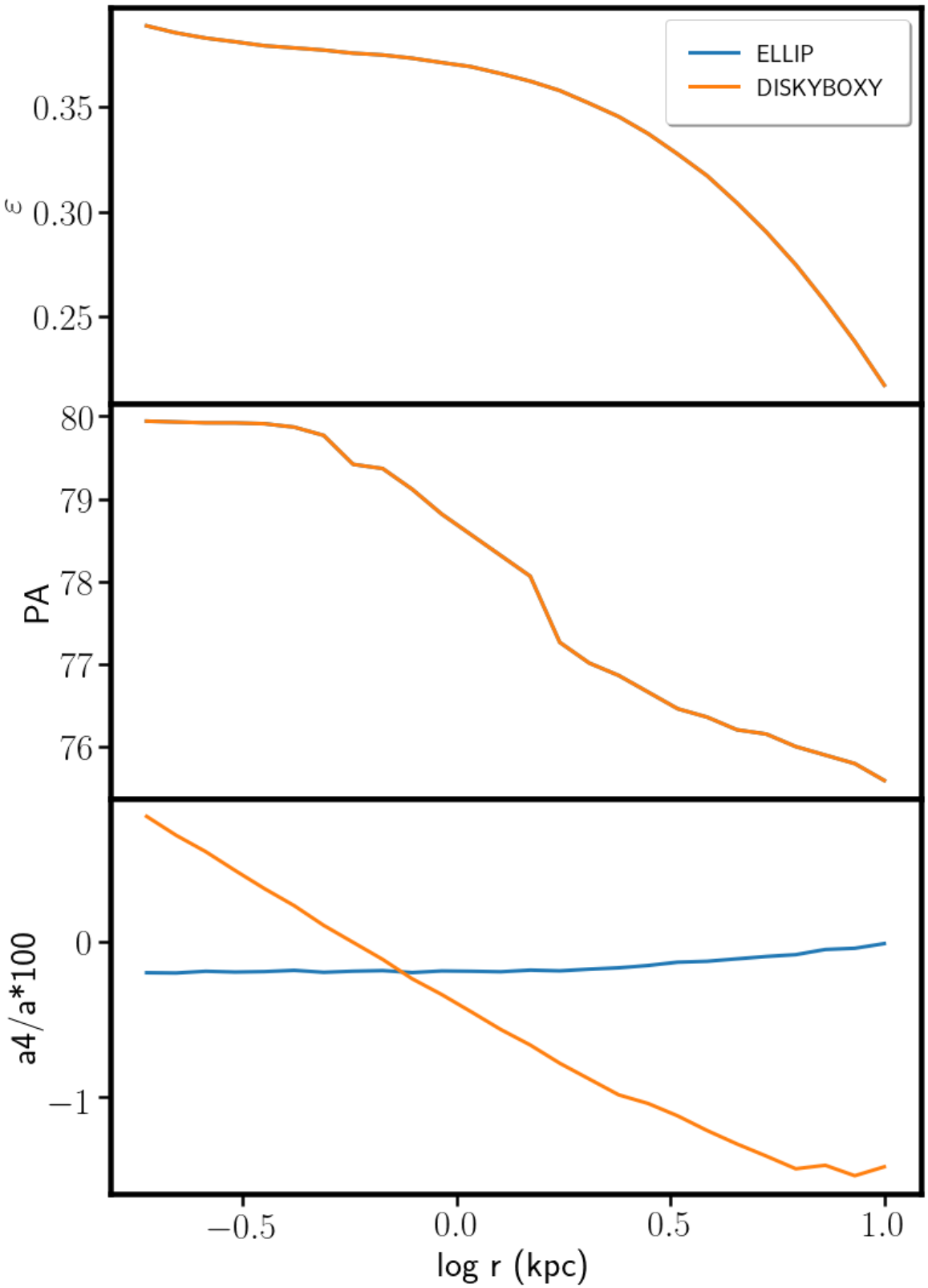


Figure 2.13: Radial profiles of the ellipticity ε (upper panel), PA (middle panel) and a_4 coefficient (lower panel) for *ELLIP* (blue line) and *DISCYBOXY* (orange line) projected using $\theta = \phi = \psi = 45°$. As expected, the ellipticity and PA profiles are identical while the *DISCYBOXY* model has discy isophotes ($a_4 > 0$) in the central regions and boxy ones ($a_4 < 0$) in the outer parts.

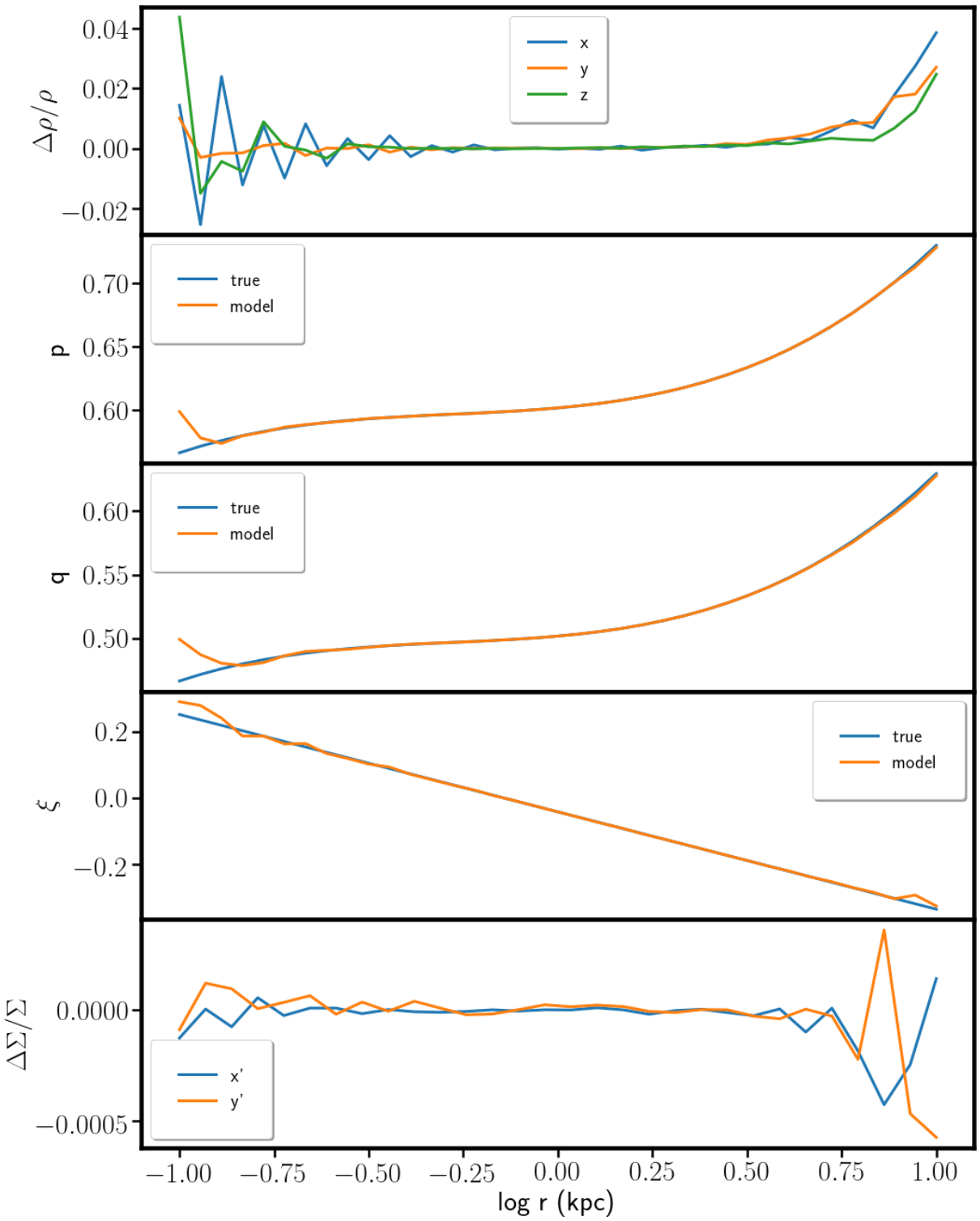


Figure 2.14: *From top to bottom:* relative difference between the true (blue) and the recovered (orange) intrinsic density along the principal axes x, y, z of model *DISCYBOXY* using the constrained shape deprojection algorithm; recovered p, q, and ξ profiles superimposed to the true ones ; percentage difference between the true and the recovered SB along the principal axes.

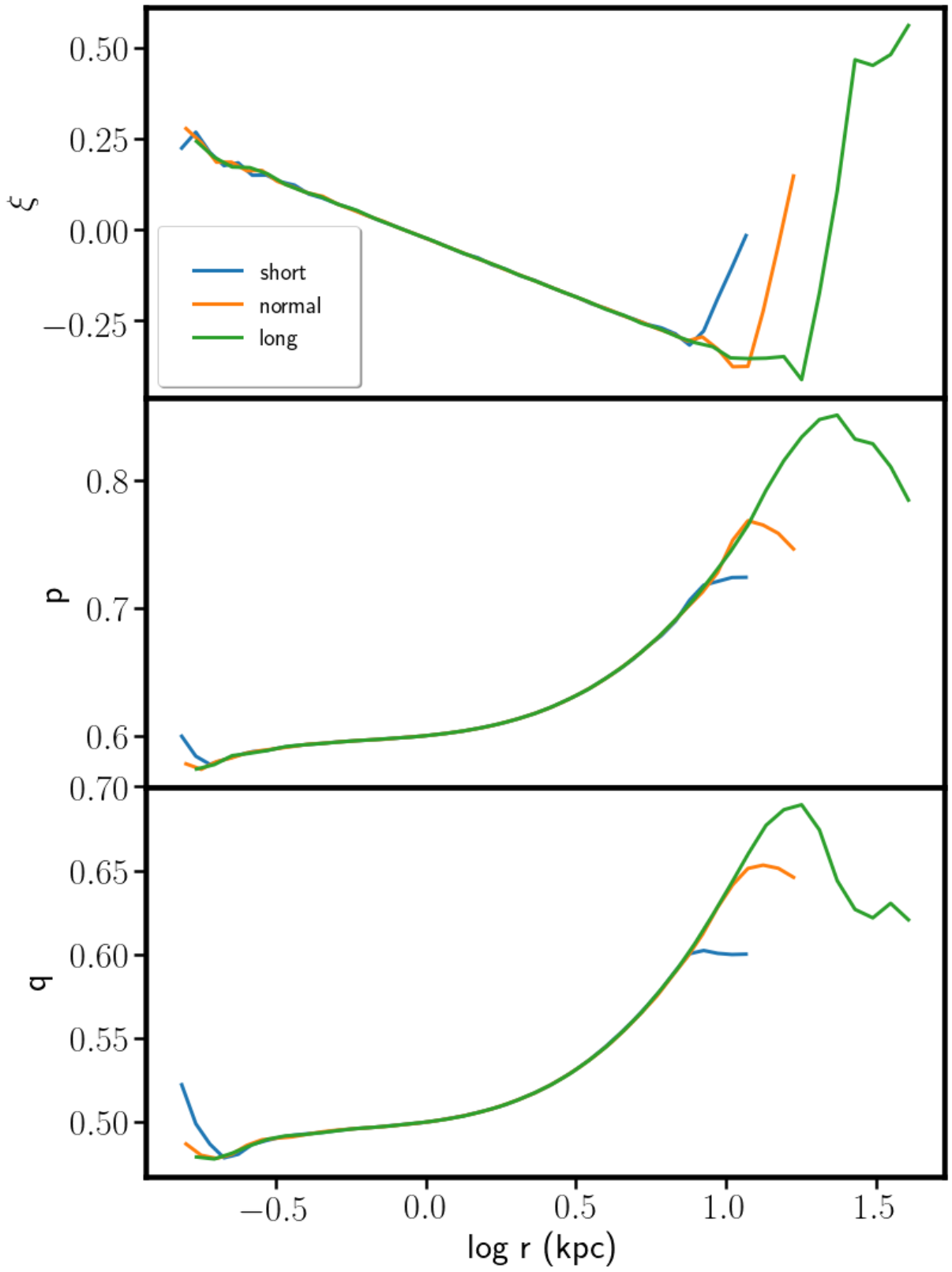


Figure 2.15: Recovery of p, q, and ξ profiles for model *DISCYBOXY* by assuming three ρ grids of different extension using the constrained shape deprojection algorithm. The more extended the grid, the farther out our code is able to recover the true profiles.

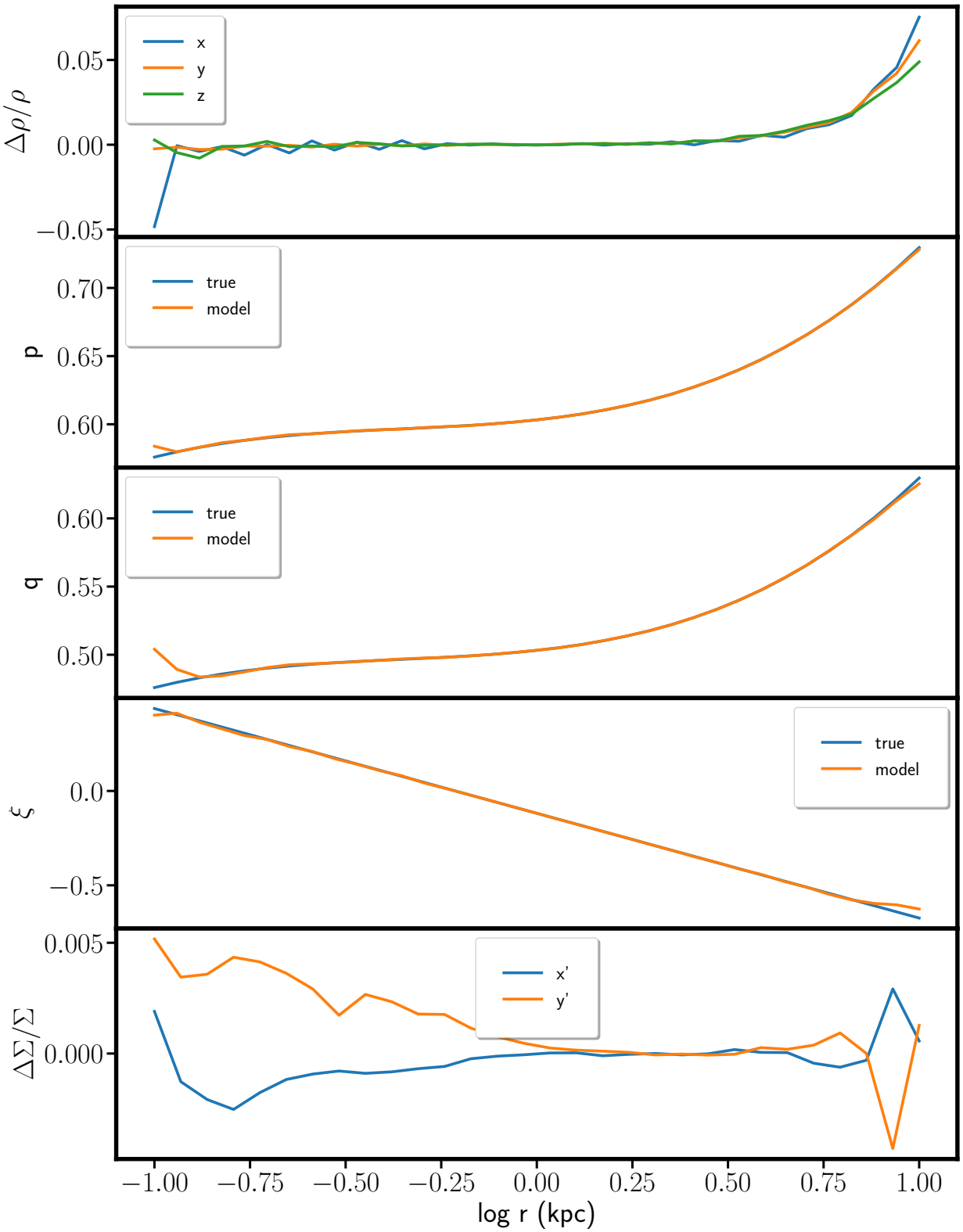
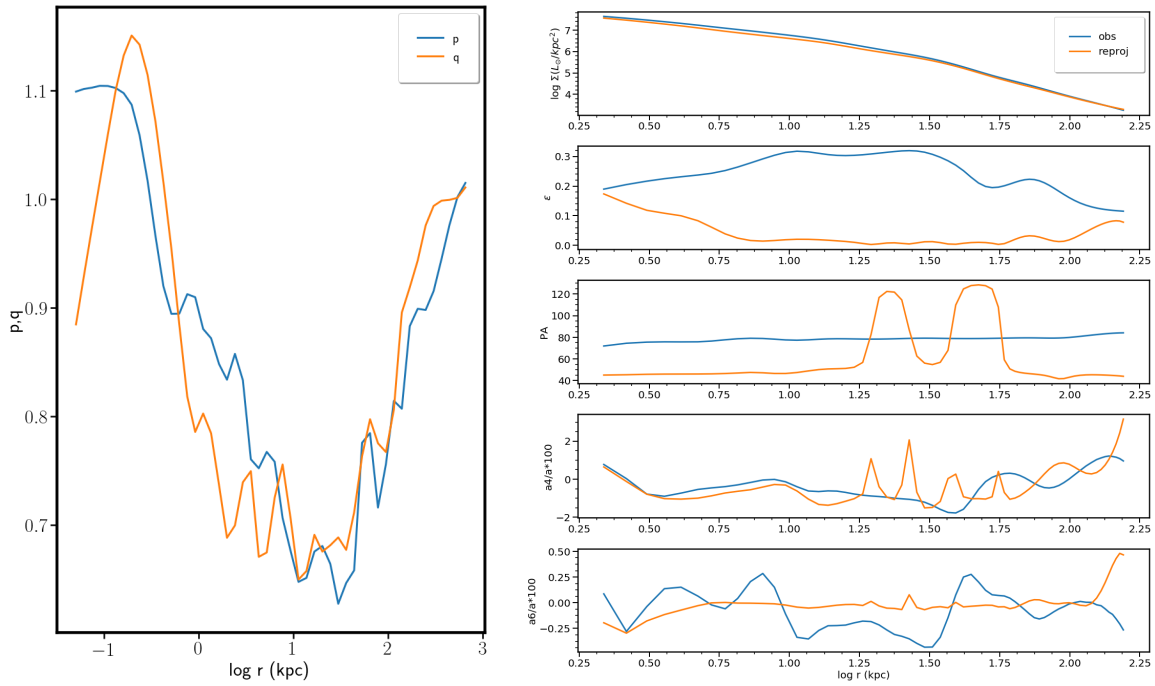


Figure 2.16: Same as Fig. 2.14 but using a Hernquist model which is less cuspy than the Jaffe model. Here too, the code fits both the intrinsic density and the projected SB very well.



(a) (b)

Figure 2.17: Discarding *NBODY* deprojections at wrong viewing angles.*Left:* p and q profiles, obtained for viewing angles $\theta = 40°$, $\phi = 50°$, and $\psi = 45°$. We formally discard this deprojection, because the profiles cross over. *Right:* Isophote shape analysis as a function of semi-major axis for the re-projection along the y-axis (orange) of the density obtained by deprojecting at $\theta = 30°$, $\phi = 60°$, and $\psi = 45°$. We discard this deprojection because of the unphysical jumps in the PA of the y-axis re-projection.

we achieve a very good recovery also when the density is stratified on *deformed* ellipsoids (2.29). This is not fully surprising, since the information available on the 4 planes in Fourier space discussed in Section 2.2 should be more than enough to constrain the four one-dimensional functions used in our procedure (assuming that the LOS is not parallel to a principal axes).

2.6 Reconstruction of viewing angles

The projection geometry of a galaxy is, of course unknown. In the axisymmetric case, it is completely described by the inclination angle i between the minor axis and the LOS, while in the triaxal case we need three angles, two (θ, ϕ) to specify the LOS direction and another one (ψ) to give a rotation around the LOS. However, it is unlikely that a given SB profile can be deprojected for every viewing geometry, which is something the fully non-parametric code is able to do by producing possibly unphysical densities. For example, only for a restricted set of viewing angle we can find an MGE (Cappellari, 2002) deprojection.

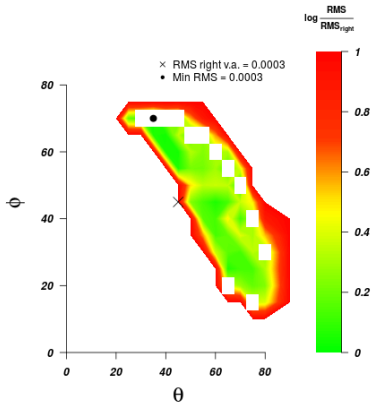
Here, we want to explore how significantly the range of allowed viewing angles can be narrowed using our constrained-shape deprojection approach. To this end we deprojected three different models for several *wrong* viewing angles and study the effects of such incorrect projection geometries. The models we consider are as follows.

- The ellipsoidal Jaffe model *ELLIP* (Figs. 2.12a & 2.13);
- The *DISCYBOXY* Jaffe model (discy in the centre, boxy towards the outer regions, see Figs. 2.12b, 2.12c & 2.13). The comparison between *ELLIP* and *DISCYBOXY* enables us to study the effects of deviations from perfect ellipticity;
- Finally, we consider a more realistic case, allowing us to study the effects of noise. To this end, we use an N-body model drawn from simulations of Rantala et al. (2018, 2019), which has an effective radius of 7 kpc and an assumed distance of 20 Mpc (model *NBODY*). Its projections are described in Neureiter et al., submitted to MNRAS, where it is used to test our newly developed triaxial Schwarzschild code. The semi-major axis of the innermost isophote is 0.5" (~ 0.05 kpc), while the outermost radius is 100 kpc. The grids onto which we place the SB and the intrinsic density have the same dimensions as those used for the Jaffe model. This gives a step of $\sim 0.18 \log$ kpc for the SB grid and of $\sim 0.12 \log$ kpc for the ρ grid. The SB grid has been chosen to be circular, while for the ρ grid we have taken flattenings of $P = 0.8$ & $Q = 0.7$.

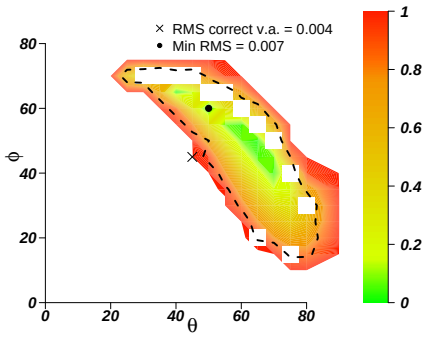
In all three cases, we project the true ρ using $\theta = \phi = \psi = 45°$. We first assume we knew the correct value of the angle ψ and then consider two different wrong values of $\psi = 30°$ and $\psi = 60°$. For each one of these cases, we deproject the models on a grid of θ, ϕ values linearly spaced from 0° to 90° with step of 5°. The code is free to search for the best-fitting p, q and ξ profiles. The RMS we obtain for *ELLIP* and *DISCYBOXY* at the correct viewing angles is about 0.03% in SB, 0.4% (for *ELLIP*) and 0.6% (for *DISCYBOXY*) in ρ. For *NBODY*, where noise is present, the RMS in SB is of the order of 0.8% while for ρ we get ∼14.4%. In this case, the RMS value is mostly driven by the noise rather than by poor extrapolation.

2.6.1 A recipe to compare deprojections obtained with different assumed viewing angles

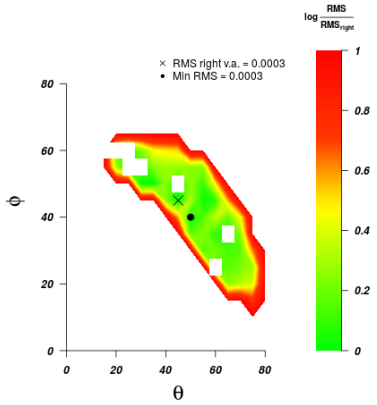
Reconstructed densities that fit well the given SB for a given choice of viewing angles could generate unrealistic SB profiles when projected to different viewing angles. In the following we adopt some criteria to find and eliminate these cases. These criteria incorporate observations of massive elliptical galaxies as a class, e.g. their observed ellipticity distributions, frequency and strength of isophote twists etc. in a *qualitative* way. We plan a more statistical analysis of this in a separate paper (de Nicola et al., in prep.). Depending on the class of galaxies considered, other criteria might be more useful. However, here we treat our mock SB data as if they were massive ellipticals to illustrate how well the viewing angles of these galaxies can be constrained photometrically.



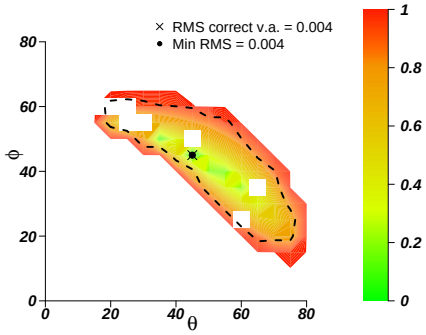
(a) SB comparison, $\psi = 30°$.



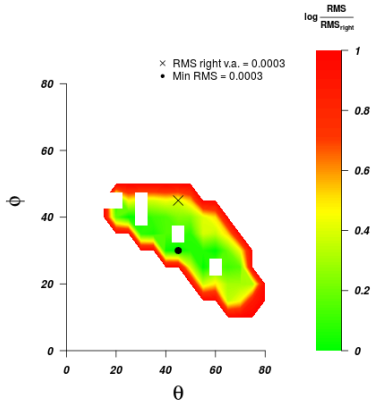
(b) ρ comparison, $\psi = 30°$.



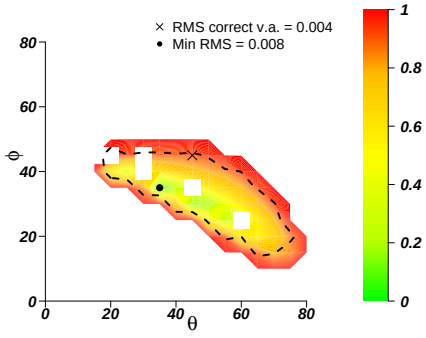
(c) SB comparison, $\psi = 45°$.



(d) ρ comparison, $\psi = 45°$.



(e) SB comparison, $\psi = 60°$.



(f) ρ comparison, $\psi = 60°$.

Figure 2.18: Logarithmic RMS errors in SB (left) and ρ (right) for model *ELLIP* scaled to those obtained when deprojecting at the correct viewing angles ($\theta = \phi = \psi = 45°$), obtained for constrained-shape deprojections at different assumed (wrong) viewing angles. The dashed contours on the right delimits the area inside which the RMS in SB is within twice the values for the correct viewing angles. The cross labels the correct (θ, ϕ), while the black dot is at the minimum RMS. Empty (white) squares depict regions discarded because of crossing p and q profiles.

Our main criterion to compare different deprojections is their relative likelihood, or the goodness of fit, respectively. We discard all those deprojections which have an RMS (in SB) larger than 0.1% for the Jaffe models *ELLIP* and *DISCYBOXY* and than 1% for *NBODY*. These values have been selected by checking the RMS (in SB) that we obtained when deprojecting the SB profiles for the true viewing angles (∼0.03% for *ELLIP* and *DISCYBOXY* and ∼0.8% for *NBODY*). The 1% threshold we use for *NBODY* is what we are likely to be using for real galaxies too.

Even those viewing directions that give an excellent fit to the observed SB can be ruled out if they happen to show p and q profiles which are not smooth, or have values which are either too low (≤ 0.2) or too high (≥ 5) with respect to the observed ellipticity distribution of elliptical galaxies. Finally, p and q profiles with interchanging principal axes are unlikely, since this would produce frequent and strong isophote twists (Fig. 2.11), which are not observed often in massive ellipticals (Goullaud et al., 2018). This means that we would accept a p or q profile which is always above unity but we would discard it in case it was above unity for some radii and below it for others[7].

Finally, we re-project the remaining densities along the principal axes and check the isophotal shapes, which is a technique already used in the axisymmetric case (e.g. Thomas et al. 2005). In fact, a plausible density for a giant elliptical galaxy is not expected to have too high (or too low) higher-order Fourier coefficients ($-5 \leq a_4 \leq 0.2$), too high ellipticity (≥ 0.6) or too severe twists ($\leq 40^\circ$). Examples of the second and third criteria are given in Fig. 2.17.

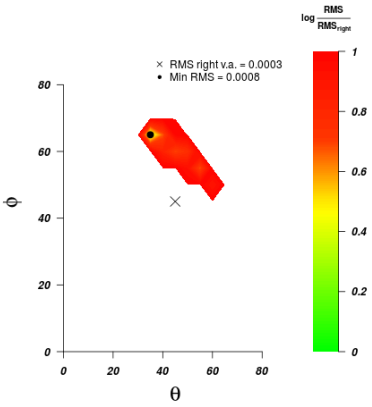
2.6.2 Results

In order to assess the results derived in the previous section, we plot in Figs. 2.18, 2.19, and 2.20 the RMS errors (both in SB and in ρ and scaled to the RMS for the correct viewing angles) as a function of (θ, ϕ) for the models *ELLIP*, *DISCYBOXY*, and *NBODY*, respectively. In the top panels $\psi = 45^\circ$ (correct value), while the middle and the lower panels are for $\psi = 30^\circ$ and $\psi = 60^\circ$ respectively. In all these plots, a cross shows the correct (θ, ϕ) and a block dot those corresponding to the minimum RMS. On the ρ plots (right panels), we also show as dashed curve the contour delimiting the area inside which the RMS in SB is within twice the value for the correct viewing angles. White quadratic holes are regions omitted because of implausible p or q profiles or re-projections.

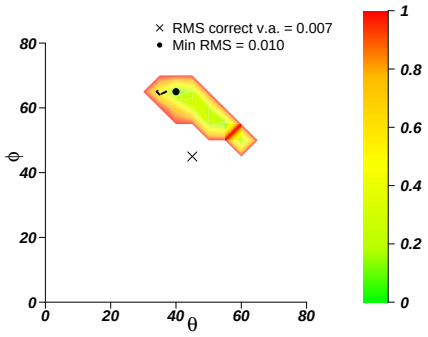
The main conclusions we can draw from these figures are as follows.

1. Regardless of the value of ψ, there are wrong viewing angles (θ, ϕ) for which the code can find a fit to the SB even slightly better than that at the true viewing angles. This does not happen for model *DISCYBOXY*, for which no solutions at wrong viewing angles are found, suggesting that the introduction of a variable ξ profile shrinks the region of acceptable deprojections. Larger acceptable regions are found for model *NBODY* (Fig. 2.20), where noise is present.

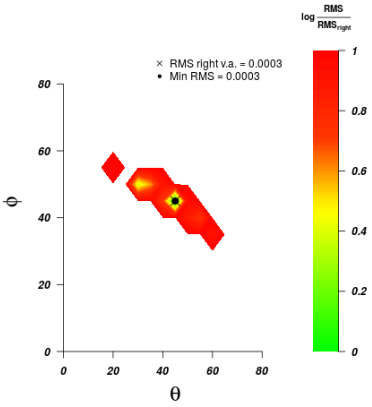
[7]These conditions can either be verified *a posteriori* or *a priori* by imposing constraints on p and q, both of which our code allows.



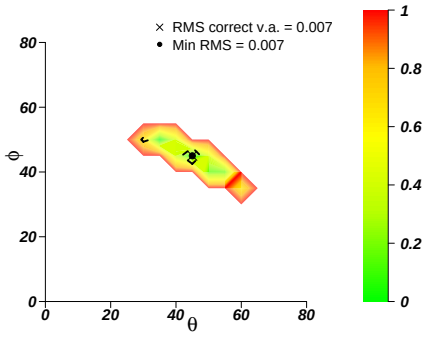
(a) SB comparison, $\psi = 30°$.



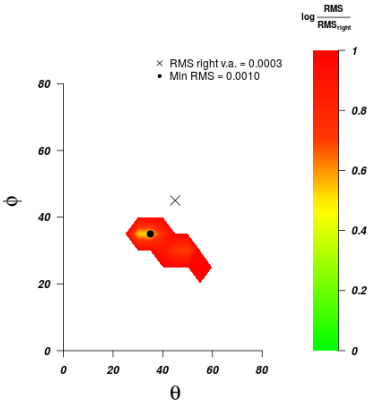
(b) ρ comparison, $\psi = 30°$.



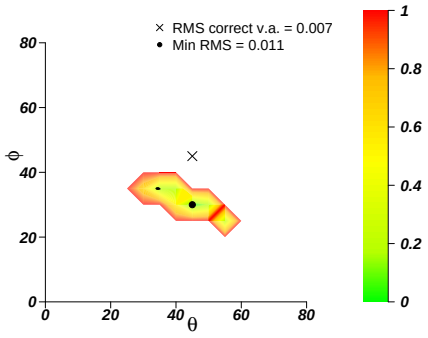
(c) SB comparison, $\psi = 45°$.



(d) ρ comparison, $\psi = 45°$.

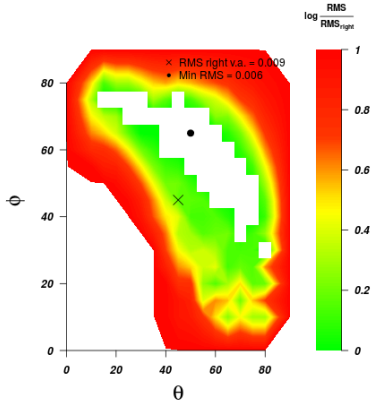


(e) SB comparison, $\psi = 60°$.

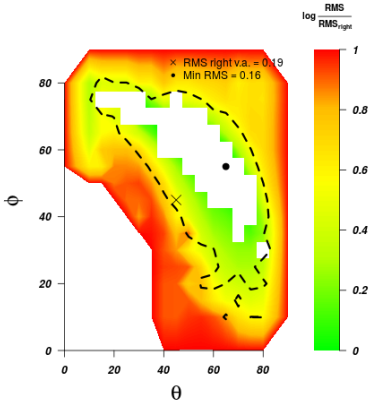


(f) ρ comparison, $\psi = 60°$.

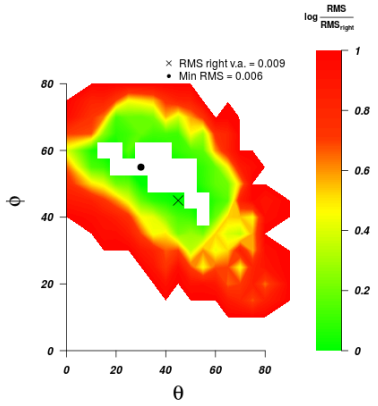
Figure 2.19: Same as Fig. 2.18 but for model *DISCYBOXY*. We do not plot the white squares since in this case we do not need them to constrain the viewing direction.



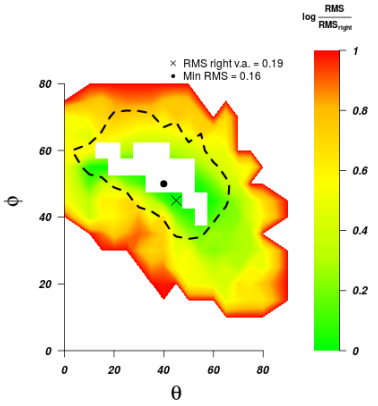
(a) SB comparison, $\psi = 30°$.



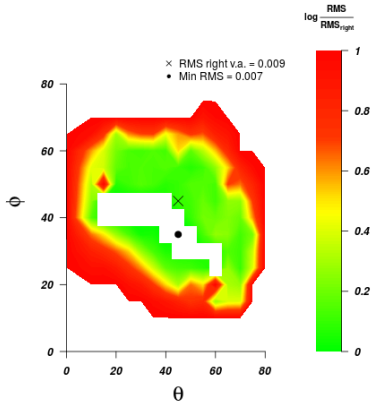
(b) ρ comparison, $\psi = 30°$.



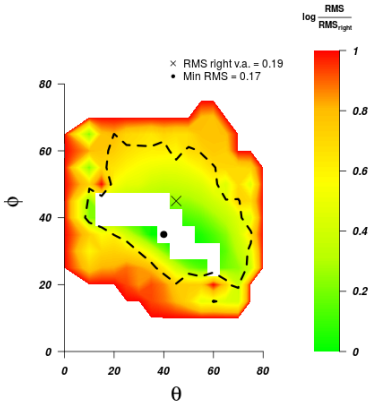
(c) SB comparison, $\psi = 45°$.



(d) ρ comparison, $\psi = 45°$.



(e) SB comparison, $\psi = 60°$.



(f) ρ comparison, $\psi = 60°$.

Figure 2.20: Same as Figs. 2.18 & 2.19 but for model *NBODY*.

2. For *ELLIP* and *DISCYBOXY* the RMS differences for both SB and ρ between the true and wrong viewing angles can be as high as three orders of magnitude (Figs. 2.18 & 2.19), but the noisy model *NBODY* allows only for a one order of magnitude range in RMS (Fig. 2.20). Furthermore, model *NBODY*, because of its noise, is the only one for which the viewing angles that give the best RMS *in* ρ are different from the true ones. However, these solutions have intersecting noisy p and q profiles, that according to our selection rules would be excluded (see Fig. 2.17).

3. For *ELLIP* and *NBODY*, not only we find wrong intrinsic densities that project to a very good fit to the observed SB (as it happens in the non-parametric case), but also intrinsic densities with low RMS's which do not project to an acceptable fit to the observed SB profile.

4. If ψ is wrong, then the $(\theta,\ \phi)$ pair that give the best RMS in SB is far off the true one. Moreover, the correct $(\theta,\ \phi)$ pair combined with the wrong ψ can result in an RMS an order of magnitude larger than for the correct projection. Since the observed ellipticity and twist of a given model depend both on $(\theta,\ \phi)$ and the p, q profiles, it is not immediately clear whether a set of wrong viewing angles cannot deliver a good solution, as the case discussed above for the two values of ψ shows. In Fig. 2.11 we have already seen that intersecting p and q profiles help in generating large observed twists.

5. The conditions we apply to the p, q profiles and the re-projections along the principal axes shrink the allowed range of viewing angles much more strongly for *NBODY* than for models *ELLIP* and *DISCYBOXY*. The presence of noise allows to generate deprojected intrinsic densities that are 'stranger' and therefore more easily eliminated than in a noise-free case. Since the SB profile of an ordinary massive elliptical are not as noisy as our *NBODY*, more deprojections are likely to survive these conditions when dealing with real galaxies.

We have shown in a qualitative manner that the statistical photometric properties of massive ellipticals (observed ellipticities, isophotal distortions and isophote twists) can be modelled with deformed ellipsoidal intrinsic density distributions. As long as the assumption of deformed-ellipsoidal density distributions holds, the range of possible deprojections shrinks considerably. In fact, since the deprojection becomes formally unique, comparing the fit quality of different deprojections at different assumed viewing angles can be used to narrow down the possible LOS of a massive galaxy just from photometric data. We plan to study in detail the intrinsic shape distribution of massive galaxies in a separate paper.

2.7 Comparison with the MGE approach

The Multi-Gaussian Expansion (Cappellari, 2002; van den Bosch et al., 2008) is a fast tool to deproject a SB profile, directly from a `FITS` file, assuming both the SB and the density profiles can be approximated as a sum of Gaussians. This analytic approach produces a fit that can be reduced to a small set of numbers, yields smooth solutions, is fast, is bound to deliver reasonable

Table 2.3: MGE fit to the Jaffe model detailed in this section. Luminosities are in counts/pixel/10^8, σ's in pixels and position angles (PA) in degrees. From these values one can compute the SB profile using equation (1) of Cappellari (2002).

L	σ	q_s	PA
1.80751	11.7388	0.72202	0.171955
1.11501	11.9416	0.485346	-1.85816
3.29186	23.4278	0.626633	-1.0591
2.82643	37.5102	0.630021	-0.217961
3.43627	54.6153	0.629184	-1.75377
4.63525	90.624	0.632035	0.102047
2.874	147.446	0.705932	-5
3.04795	210.449	0.59828	2.22027
6.79399	404.492	0.817577	-0.29154

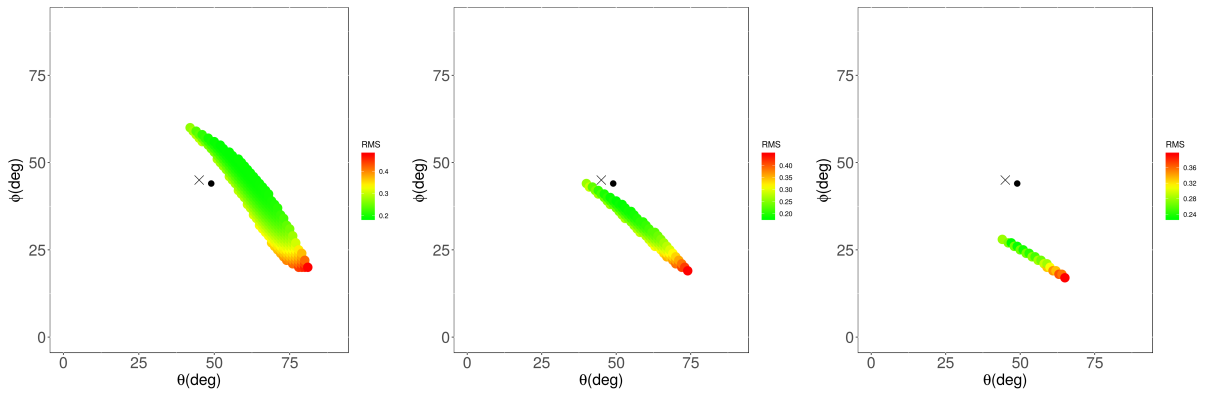
re-projected SB whatever viewing angles are considered, and delivers a unique deprojection for a set of allowed viewing angles. However, this set can be empty if the gaussians required to get a good fit span a large range of twists or have very low flattenings. Therefore, here we first apply it to the models considered in the previous sections *without* the additional flat components discussed in Appendix 2.B to assess its performances in terms of quality of the reproduced density and set of allowed viewing angles. Then, we compare its results with those obtained by our code for a real galaxy.

2.7.1 MGE performance on the Jaffe model

We make sure that our coordinate system on the plane of the sky is consistent with what the MGE assumes, namely that the major axis of the innermost Gaussian component is aligned with the x'-axis. Since the twist of our Jaffe model is nearly zero in the innermost regions (see Fig. 2.9a), we can assume that aligning the innermost isophote with the x'-axis is to a very good approximation the same as aligning the innermost Gaussian of the MGE fit. This is achieved by rotating our isophotes clockwise by the PA of the innermost isophote ($\sim 78°$) and adding this value to the $\psi = 45°$ we used above, giving a new $\psi_{\rm MGE} = 123°$.

We project the density of the *DISCYBOXY* model with viewing angles ($\theta = 45°, \phi = 45°, \psi_{\rm MGE} = 123°$) and generate the galaxy image in `FITS` format as an input to the code of Michele Cappellari[8] to produce the MGE fit. The procedure fits the image with a combination of $N = 8$ Gaussians which we report in Table 2.3. We tested several fits, each time by imposing different constraints on the flattenings q'_j and the twist $\Delta\psi_j$, allowing for up to 30 Gaussians in the fit. The constraints are needed because by letting the code run unconstrained we obtained a solution for which no possible viewing angles are found; we ended up using $q'_j \in [0.2, 1]$ and $\Delta\psi_j \in [-5, 5]°$. The RMS between the MGE SB and that we have on our grid (2.10) is 4.9%. We then compute the intrinsic densities corresponding to the allowed viewing angles (see

[8]http://www-astro.physics.ox.ac.uk/~mxc/software/



(a) $\psi = 108°$. (b) $\psi = 123°$. (c) $\psi = 138°$.

Figure 2.21: RMS between the true intrinsic density and that recovered by the MGE for the true value of $\psi = 123°$ (middle) and for two wrong values of $\psi = 108°$ (left) and $\psi = 138°$ (right) as a function of all possible θ, ϕ values compatible with this particular ψ. The cross and the black dot indicate the true (θ, ϕ) viewing angles and those at which the MGE deprojections give the least RMS, respectively.

Section 2.4.1 and equations (7-8) of Cappellari 2002). Clearly, while our code can produce a deprojection for each possible set of viewing angles, this is not possible for the MGE. Thus, to construct analogs of Fig. 2.18 we have isolated all solutions which have the true $\psi = 123°$ (Fig. 2.21b), then those at $\psi = 108°$ (Fig. 2.21a) and finally those at $\psi = 138°$ (Fig. 2.21c) and plot the RMS with respect to the true intrinsic density. The meaning of the black cross and dot are the same as in Fig. 2.18. The most significant findings are as follows.

- The quality of the MGE fit is poorer than the one achieved with the constrained-shape algorithm, delivering an RMS in SB of nearly 5%. This is not surprising, since a superposition of a series of densities localized in shells provides much more flexibility than a set of Gaussians.
- An MGE deprojection for the true viewing angles is possible.
- Of all possible MGE deprojections, the one that gives the best RMS in ρ ($\sim$ 18%) is obtained for viewing angles of $\theta = 49°$, $\phi = 44°$, $\psi = 119°$, different from the true values by a few degrees.
- The RMS in ρ yielded by the constrained-shape algorithm is significantly smaller (0.7%) than the one achievable with the MGE approach, even omitting the last 10 radial points, where the Gaussians have a sharp cut-off (see also Fig. 2.22).

2.7.2 Comparison using a real galaxy

It is now interesting to compare our code with MGE for a real galaxy, which can neither be described exactly by a sum of Gaussians nor has the form of eq. 2.29. We focus on the elliptical

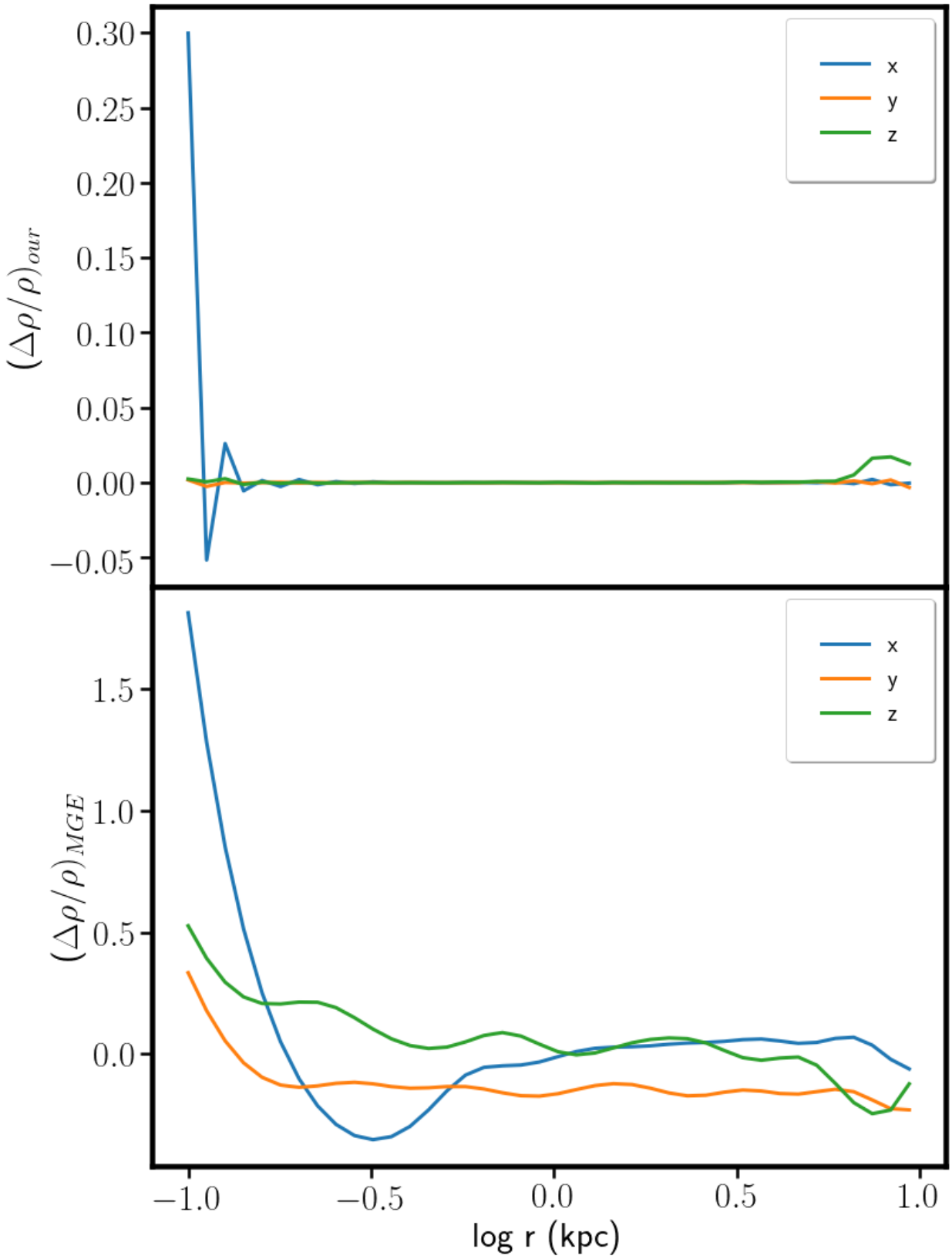
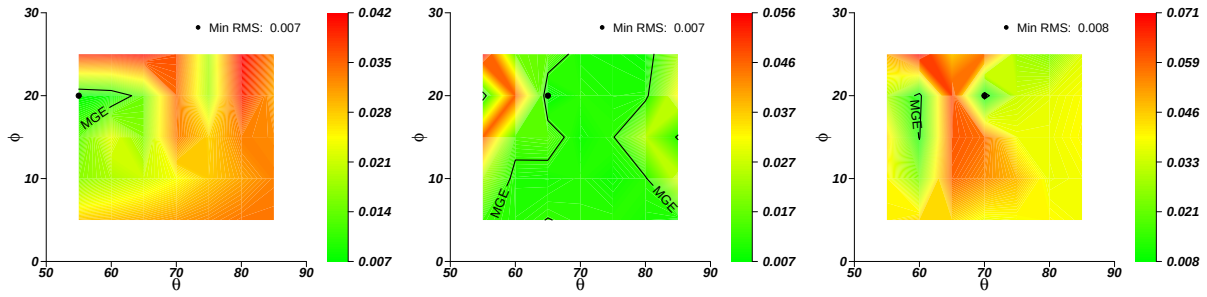


Figure 2.22: Comparison of the reconstructed intrinsic density of *DISCYBOXY* between our code (top) and MGE (bottom) on the three principal axes. For both cases, we took the deprojection at the viewing angles giving the best RMS in ρ, even though these are not exactly the true ones.



(a) $\psi = 30°$. (b) $\psi = 40°$. (c) $\psi = 50°$.

Figure 2.23: RMS between the true SB and that reconstructed by our code for the elliptical galaxy NGC 5831 for $\psi = 30°$, $40°$, and $50°$ as indicated. We sampled θ, ϕ in the intervals where the MGE method allows for a solution. The black dots indicate the respective (θ, ϕ) with the smallest RMS. The black contours bracket the regions where RMS $< 1.2\%$.

Table 2.4: Highest ellipticity ε and ranges for a_4 and a_6 found when re-projecting the intrinsic densities obtained for NGC 5831 along 60 random directions. No unphysical or unusual values for these coefficients are found.

ψ	ε	a_4	a_6
$30°$	0.496	$[-0.51, 1.43]$	$[-0.09, 0.78]$
$40°$	0.512	$[-0.78, 2.06]$	$[-0.32, 1.12]$
$50°$	0.478	$[-0.23, 1.65]$	$[-0.05, 0.91]$

galaxy NGC5831, which has a $\sim 35°$ isophote twist and is also used by Cappellari (2002) as an example for the performance of MGE in presence of isophote twist. The released MGE Python code fits the photometry with a sum of 11 Gaussians yielding an RMS of 4.7%; Cappellari (2002) quotes an even better RMS of 1.2%, that we adopt as a benchmark.

We computed all the (θ, ϕ, ψ) values compatible with the MGE fit, obtaining $\theta \in [55, 85]°$, $\phi \in [5, 25]°$, $\psi \in [106, 114]°$. Then, we used our near-ellipsoidal algorithm to deproject the surface density of the galaxy considering θ and ψ in the interval allowed by the MGE fit with a step of $5°$. Since the definition of ψ adopted in the MGE formalism differs from ours (see Section 2.7.1), we sampled $\psi \in [10, 180]°$ with a step of $10°$.

In Fig. 2.23, we plot the RMS error of the surface density as function of θ and ϕ for $\psi = 30°$, $40°$ and $50°$. **In all three cases, we find viewing directions at which we can fit the surface brightness better than the MGE**. In particular, for all three ψ values we find deprojections with RMS $< 1\%$, with the $\psi = 40°$ case having $\sim 70\%$ SBs below this threshold.

As outlined in Sec. 2.6.1, we reprojected the resulting intrinsic densities for 60 random viewing angles to ensure that the isophotes look reasonable. We show in Tab. 2.4 the largest ellipticity ε and the maximum/minimum a4/a6 values we found among all densities for a certain ψ. Here we do not apply any requirements on the twist angle τ since differently from most massive ellipticals this galaxy does have a strong twist in the outer regions. We see that the isophotes are never

particularly flat and do not show anomalous a_4 or a_6 values. It is interesting to note that we do not find any re-projection yielding significantly boxy isophotes.

The main conclusion is that, although the MGE naturally rules out unsmooth densities and is fast, it might bias the region of allowed angles θ, ϕ, ψ and deliver an SB fit and reconstructed density of relatively poor quality. We will investigate the impact of these shortcomings on dynamical modelling in a future paper.

2.8 Conclusions

We present two novel approaches to deproject elliptical galaxies under the assumption of triaxiality, the first fully non-parametric, the second stratified on deformed ellipsoidal shells. Both are able to deal with isophotal twist and can deproject systems that have the principal axes interchanging between them as a function of the distance from the centre.

The full non-parametric code can be used to explore a range of possible deprojections going beyond those allowed by current state-of-art algorithms, but at present does not allow for any control of the shape of the density. Our constrained-shape approach, on the other hand, allows for penalization towards discy/boxy shapes and controls the smoothness of the density along the major axis and of the density contours. Tests performed with benchmark Jaffe and Hernquist models of varying axis ratios and shape biases show that the intrinsic density can be recovered very well when the viewing angles are known and far enough from the principal axes, much better than what can be achieved with a Multiple Gaussian Expansion. When dealing with a noisy system such as an N-body simulation, the SB can be fitted with an RMS $\sim 1\%$, delivering a reconstructed density precise to 20%, when the viewing angles are known.

We are able to constrain the possible range of viewing angles by mapping the RMS of the fitted SB as a function of (θ, ϕ, ψ) and eliminating unphysical reconstructed densities, by examining their re-projected SB. This reduces the number of densities to be tested dynamically, which will be the subject of a forthcoming paper, towards the deprojections and dynamical modelling of real galaxies.

In this process we might discover that a number of galaxies appear similar to the *LARGEDISC* discussed in Appendix 2.B, where a massive disc component is present together with a triaxial bulge. For unfavourable viewing angles this component is invisible in projection. In Appendix 2.A we discuss a number of analytic descriptions of these possible *cloaked densities*, the triaxial extension of axisymmetric *conus densities*. In Appendix 2.B we show how one can flag these cases. We explore how well we can deproject triaxial bodies where flattened axisymmetric components are present. We find that the constrained-shape approach performs well if these (disc like) components do not exceed 15% of the total light. We develop a deformed-ellipsoidal shape plus axisymmetric component algorithm that is able to reconstruct well systems with nearly edge-on ($\approx 80^\circ$) discs, or flag the possible presence of important (i.e. contributing $\approx 50\%$ of the total light) disc components at unfavourable angles ($\leq 45^\circ$).

Appendix

2.A Constructing cloaked densities

We are now considering various ways to construct analytical models that project to nothing. Such *cloaked densities* can be added or subtracted to any model without changing its projection but potentially with drastic changes to its spatial shape.

2.A.1 Cloaked densities via differentiation

We are after triaxial functions whose Fourier transform vanishes on the four planes $\boldsymbol{\ell}_i \cdot \boldsymbol{k} = 0$. A simple such function is

$$\begin{aligned}(\boldsymbol{\ell}_1 \cdot \boldsymbol{k})(\boldsymbol{\ell}_2 \cdot \boldsymbol{k})(\boldsymbol{\ell}_3 \cdot \boldsymbol{k})(\boldsymbol{\ell}_4 \cdot \boldsymbol{k}) \\ = (\ell_x k_x)^4 + (\ell_y k_y)^4 + (\ell_z k_z)^4 - 2(\ell_x k_x)^2(\ell_y k_y)^2 \\ - 2(\ell_x k_x)^2(\ell_z k_z)^2 - 2(\ell_y k_y)^2(\ell_z k_z)^2, \quad (2.31)\end{aligned}$$

which is positive in each of the funnels around one fundamental axis and negative in the three-sided funnels in the middle of each octant (see Fig. 2.1). If we multiply the Fourier transform $\hat{f}(\boldsymbol{k})$ of some triaxial function $f(\boldsymbol{r})$ with (2.31), the corresponding density

$$\rho_0(\boldsymbol{r}) \equiv (\boldsymbol{\ell}_1 \cdot \boldsymbol{\nabla})(\boldsymbol{\ell}_2 \cdot \boldsymbol{\nabla})(\boldsymbol{\ell}_3 \cdot \boldsymbol{\nabla})(\boldsymbol{\ell}_4 \cdot \boldsymbol{\nabla}) f(\boldsymbol{r}) \quad (2.32)$$

is invisible when seen along any of the four LOS $\boldsymbol{\ell}_i$ – this is also obvious by doing the projection via integration by parts. Defining $\boldsymbol{R} = (X, Y, Z)^t \equiv (x/\ell_x, y/\ell_y, z/\ell_z)^t$, this can be expressed as

$$\rho_0(\boldsymbol{r}) = \left[\partial_X^4 + \partial_Y^4 + \partial_Z^4 - 2\partial_X^2\partial_Y^2 - 2\partial_X^2\partial_Z^2 - 2\partial_Y^2\partial_Z^2\right] f(\boldsymbol{r}). \quad (2.33)$$

Applying this procedure to a spherical Gaussian $f(\boldsymbol{r}) = G(\boldsymbol{r}) \equiv \exp(-\frac{1}{2}\boldsymbol{r}^2)$, we find $\rho_0 = h(\boldsymbol{r})G(\boldsymbol{r})$ with

$$\begin{aligned}h(\boldsymbol{r}) \;=\; & \ell_x^4(x^4 - 6x^2 + 3) + \ell_y^4(y^4 - 6y^2 + 3) + \ell_z^4(z^4 - 6z^2 + 3) \\ & -2\ell_x^2\ell_y^2(x^2 - 1)(y^2 - 1) - 2\ell_x^2\ell_z^2(x^2 - 1)(z^2 - 1) \\ & -2\ell_y^2\ell_z^2(y^2 - 1)(z^2 - 1). \quad (2.34)\end{aligned}$$

Table 2.A.1: **Functions used by ellipsoidal cloaked densities.** Analytical 3D spherical distributions $\varphi(r)$, normalised to $\varphi(0) = 1$, with Fourier transform $\hat{\varphi}(k)$ that vanishes for $k \geq 1$. Functions with $\forall r : \varphi(r) > 0$ have 3D Wendland (1995, 2005) functions as Fourier transform. σ_k^2 is the 1D variance of $\hat{\varphi}(k)$ and determines the width of φ since $\varphi(r \to 0) = 1 - \frac{1}{2}\sigma_k^2 r^2 + O(r^4)$.

$\hat{\varphi}(k<1)/\hat{\varphi}(0)$	$\varphi(r)$		σ_k^2	$\varphi > 0$
1	$3r^{-2}\left[\operatorname{sinc} r - \cos r\right]$	$\frac{1}{5}$	≈ 0.2	no
$1-k$	$12r^{-4}\left[2 - 2\cos r - r\sin r\right]$	$\frac{2}{15}$	≈ 0.1333	no
$(1-k)^2$	$60r^{-4}\left[2 + \cos r - 3\operatorname{sinc} r\right]$	$\frac{2}{21}$	≈ 0.0952	yes
$(1-k)^3(1+3k)$	$630r^{-7}\left[r(8+7\cos r) + (r^2-15)\sin r\right]$	$\frac{1}{12}$	≈ 0.0833	no
$(1-k)^4(1+4k)$	$5040r^{-8}\left[9r\sin r + (24-r^2)\cos r + 4r^2 - 24\right]$	$\frac{1}{15}$	≈ 0.0667	yes
$\frac{1}{2}(1+\cos k\pi)$	$\frac{3\pi^2}{\pi^2-6}\left[r^{-2}(\operatorname{sinc} r - \cos r) - \frac{1}{\pi^2-r^2}\left(\frac{\pi^2+r^2}{\pi^2-r^2}\operatorname{sinc} r + \cos r\right)\right]$	$\frac{\pi^4-20\pi^2+120}{5\pi^2(\pi^2-6)}$	≈ 0.1048	no

This function itself is not bounded from below, i.e. approaches $-\infty$ in certain directions. This means that $[1 + h(\boldsymbol{r})]G(\boldsymbol{r})$ is not a physical model, but $h(\boldsymbol{r})G(\boldsymbol{r})$ is, of course, bounded and can be added to another model such that the total is still non-negative.

When applying the recipe to a triaxial Gaussian $G(\boldsymbol{r}) = \exp(-\frac{1}{2}\boldsymbol{r}^t \cdot \mathbf{C}^{-1} \cdot \boldsymbol{r})$, then we again obtain $\rho_0 = h(\boldsymbol{r})G(\boldsymbol{r})$ with $h(\boldsymbol{r})$ as given in equation (2.34) after the replacements $\boldsymbol{\ell} \to \mathbf{C}^{-1/2} \cdot \boldsymbol{\ell}$ and $\boldsymbol{r} \to \mathbf{C}^{-1/2} \cdot \boldsymbol{r}$.

Another option is to pick $f(\boldsymbol{r}) = F(|\boldsymbol{R}|)$. Then

$$\begin{aligned}\rho_0(\boldsymbol{R}) = {} & 3\frac{F'}{R^3} - 3\frac{F''}{R^2} + 2\frac{F'''}{R} \\ & + Q(\boldsymbol{R})\left(-15\frac{F'}{R^3} + 15\frac{F''}{R^2} - 6\frac{F'''}{R} + F''''\right),\end{aligned} \tag{2.35}$$

where

$$Q(\boldsymbol{R}) \equiv R^{-4}\left[X^4 + Y^4 + Z^4 - 2X^2Y^2 - 2X^2Z^2 - 2Y^2Z^2\right], \tag{2.36}$$

which is maximal at $Q = 1$ on the axes, minimal at $Q = -1/3$ at $X^2 = Y^2 = Z^2$, and vanishes for $X \pm Y \pm Z = 0$ (with both signs independent), corresponding to $x/\ell_x \pm y/\ell_y \pm z/\ell_z = 0$, which holds on four planes, which in a sense are the reciprocal planes to those in $\boldsymbol{k}$ space where $\hat{\rho} = 0$.

2.A.2 Cloaked densities via Fourier transform of compact functions

When constructing a cloaked density via differentiation as in the previous sub-section, one has little control over the resulting shape. Here, we consider methods to construct cloaked densities with certain properties as Fourier transform of functions $\hat{f}(\boldsymbol{k})$ that vanish everywhere except for

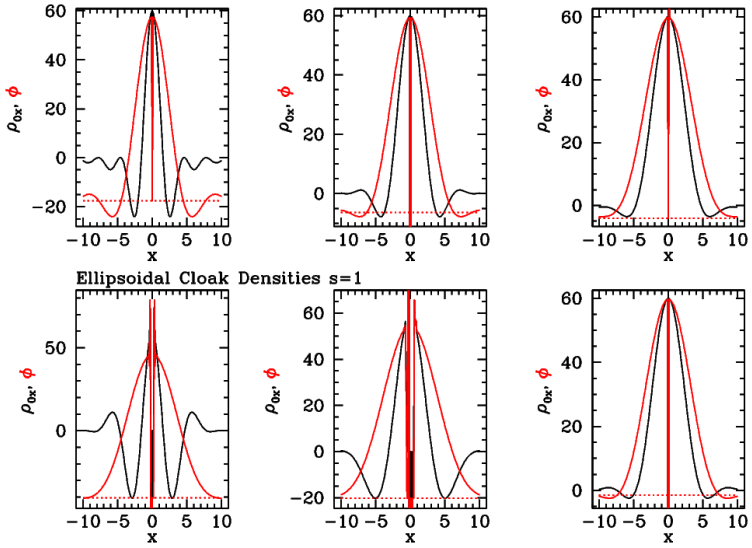


Figure 2.A.1: Plots of the six functions φ of Table 2.A.1 (red); the horizontal line shows the zero level) and the corresponding ρ_{0x} for $s = 1$ (black, see equation 2.39).

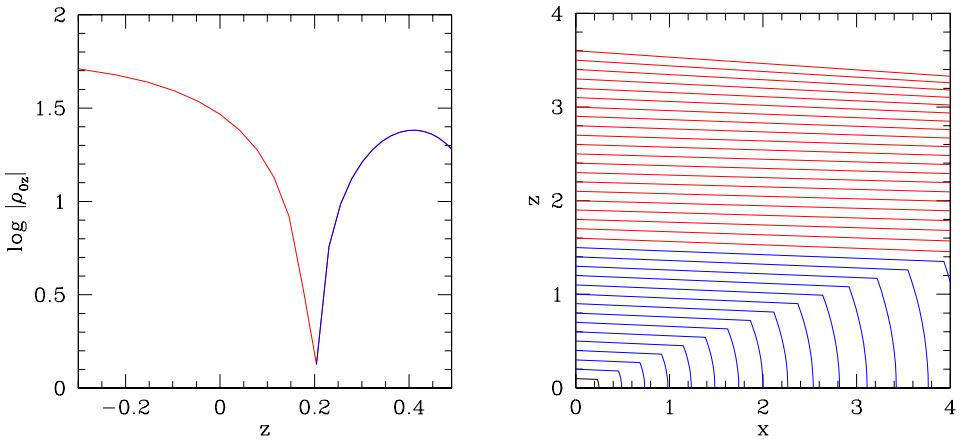


Figure 2.A.2: The equivalent of the bottom plot of Fig. 2.6 (left) and the middle plot of Fig. 2.7 for the density ρ_{0z} with $s = 1$ (see equation 2.39) and inverted colours (red for positive densities and blue for negative values).

a finite triaxial region. The simplest such functions are obtained by shifting a symmetric function of compact support by an amount $1/s$ along $\hat{\boldsymbol{x}}_i$ and superimpose it with a version shifted in the other direction:

$$\hat{\rho}(\boldsymbol{k}) = \tfrac{1}{2} \textstyle\sum_{\pm} \hat{f}(\boldsymbol{k} \pm \hat{\boldsymbol{x}}_i/s) \tag{2.37}$$

with density

$$\rho(\boldsymbol{r}) = \cos(\boldsymbol{r} \cdot \hat{\boldsymbol{x}}_i/s) f(\boldsymbol{r}). \tag{2.38}$$

This is a way to hide a disc perpendicular to the x_i axis, with typical scale height s, and with extent given by the typical scale of f.

Another possibility is to shift $\hat{f}$ along a direction in the three-sided funnels in the centre of each octant (see Fig. 2.1).

Ellipsoidal cloaked densities

One option is to take $\hat{f}(\boldsymbol{k})$ to be ellipsoidal. Let $\varphi(r)$ be a spherical function whose 3D Fourier transform $\hat{\varphi}(k)$ vanishes for $k > 1$. From such a function, we may construct an invisible model via the above recipe as

$$\rho_{0i} = \cos\left(\frac{\boldsymbol{r} \cdot \hat{\boldsymbol{x}}_i}{s}\right) \frac{\varphi(m)}{abc} \tag{2.39}$$

For this to be invisible its Fourier transform must not intersect the plane $\boldsymbol{\ell} \cdot \boldsymbol{k} = 0$ which requires

$$s < \ell_i \sigma_\ell, \qquad \sigma_\ell^{-2} = \boldsymbol{\ell} \cdot \mathbf{C}^{-1} \cdot \boldsymbol{\ell}. \tag{2.40}$$

So, not surprisingly it is easier to hide a disc that is near-perpendicular to the LOS (large ℓ_i) than other discs. Possible functions $\varphi(r)$ are listed in Table 2.A.1 and shown in Fig. 2.A.1. Fig. 2.A.2 shows the qualitative equivalent of the bottom plot of Fig. 2.6 (left) and the middle plot of Fig. 2.7 for the density ρ_{0z} with $s = 1$ (see equation 2.39).

Cuboidal cloaked densities

Instead of shifting ellipsoidal Fourier distributions, to generate cloaked densities, one may also use cuboidal distributions of the form

$$\hat{f}(\boldsymbol{k}) = \hat{h}_x(ak_x)\, \hat{h}_y(bk_y)\, \hat{h}_z(ck_z) \tag{2.41}$$

with $\hat{h}_i(k) \neq 0$ only for $|k| < 1$. For example the top-hat function and its n-fold self-convolution[9], which correspond to

$$h(x) = b_n(x) \equiv \mathrm{sinc}^n(x/n) \tag{2.42}$$

with

$$b_1(x) = \mathrm{sinc}\, x \equiv \frac{\sin x}{x}, \tag{2.43}$$

[9]These functions, also known as Schoenberg (1946) B-splines, are (modulo a scaling) identical to the Irwin (1927)-Hall (1927) probability density for the sum k of n independent variables, each drawn form a uniform distribution between $-1/n$ and $1/n$. The only difference to the common use of these functions is that we revert the role of the function and its Fourier transform so that the latter has compact support.

Table 2.A.2: **Functions for elliptic conus densities.** $\hat{\varphi}_1(x)$ is obtained from $\hat{\varphi}(k)$ via equation (2.50) (modulo a constant factor) and $\hat{\varphi}_1^{(2n)}(x)$ serves as vertical density profile for the conus density disc.

n	$\hat{\varphi}_1^{(2n)}(x<1)$	$\hat{\varphi}(k<1)$	comments
1	$1-6x^2+5x^4$	$(1-k^2)^{5/2}$	$\hat{\varphi}_1^{(2n)}$ discontinuous at $x=1$
1	$(1-7x^2)(1-x^2)$	$(1-k^2)^{7/2}$	
2	$3-30x^2+35x^4$	$(1-k^2)^{7/2}$	$\hat{\varphi}_1^{(2n)}$ discontinuous at $x=1$
2	$(3-42x^2+63x^4)(1-x^2)$	$(1-k^2)^{9/2}$	

which has as Fourier transform the top-hat function $\hat{b}_1 = \frac{1}{2}$ for $|x| < 1$ and 0 otherwise. The scaling of the argument by $1/n$ in (2.42) ensures that $\hat{h}_n(k) = 0$ for $|k| > 1$. Possible 3D densities are then

$$\rho_{i,\mathsf{n}}(\boldsymbol{r}) = \frac{1}{abc}\cos\left(\frac{r_i}{s}\right) b_{\mathsf{n}_x}\left(\frac{x}{a}\right) b_{\mathsf{n}_y}\left(\frac{y}{b}\right) b_{\mathsf{n}_z}\left(\frac{z}{c}\right) \tag{2.44}$$

with parameters $\mathbf{n} = (\mathsf{n}_x, \mathsf{n}_y, \mathsf{n}_z)$, $\boldsymbol{a} = (a, b, c)$, s and i. At large distances, these functions decay as $1/\boldsymbol{x}^{\mathbf{n}} \equiv 1/x^{\mathsf{n}_x} y^{\mathsf{n}_y} z^{\mathsf{n}_z}$. In order for this density to be invisible, its Fourier transform must not intersect the plane $\boldsymbol{\ell} \cdot \boldsymbol{k} = 0$, which requires that

$$\frac{\ell_i}{s} > \frac{\ell_x}{a} + \frac{\ell_y}{b} + \frac{\ell_z}{c}. \tag{2.45}$$

2.A.3 Cloaked conus densities

The method of the previous sub-section cannot create centrally diverging cloaked densities, because such distributions have power on all scales and their Fourier transform is not confined to a compact region. This is, however, only a shortcoming of this particular method and not inherent to cloaked densities: one may superpose many such models with ever smaller $\mathbf{C}$ and s to create a cuspy cloaked density.

Alternatively, we may construct a cloaked density from a Fourier transform that is defined everywhere inside a cone around one of the fundamental axes. Without loss of generality, we take this to be the z axis. Taking the cone to be elliptic, this gives the ansatz

$$\hat{\rho}(\boldsymbol{k}) = \hat{\varphi}\left(\sqrt{a^2k_x^2 + b^2k_y^2}\Big/ k_z\right)\hat{f}(|k_z|), \tag{2.46}$$

where as before $\hat{\varphi}(k)$ vanishes for $k > 1$, while $\hat{f}$ is as of yet unspecified. For this to be invisible

$$\ell_x^2/a^2 + \ell_y^2/b^2 > \ell_z^2. \tag{2.47}$$

Fourier transforming $\hat{\rho}(\boldsymbol{k})$ first in x and y and then in z gives

$$\begin{aligned}\rho(\boldsymbol{r}) &= \frac{1}{ab}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}k_z z}\; k_z^2 \hat{f}(k_z)\; \varphi(k_z\mu)\; \mathrm{d}k_z \\ &= \frac{1}{ab\mu}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\kappa z/\mu}\; \left(\frac{\kappa}{\mu}\right)^2 \hat{f}\left(\frac{\kappa}{\mu}\right)\; \varphi(\kappa)\; \mathrm{d}\kappa \end{aligned} \tag{2.48}$$

with $\mu^2 \equiv x^2/a^2 + y^2/b^2$. For this to result in a closed functional form, the freedom for the function $\hat{f}(k)$ must be exploited. If, for example, one takes $k^2\hat{f}(|k|) = (-\mathrm{i}k)^{2n}$ with[10] $n > 0$, then

$$\rho(\boldsymbol{r}) = \frac{1}{ab\mu^{2n+1}}\hat{\varphi}_1^{(2n)}\left(\frac{z}{\mu}\right), \tag{2.49}$$

where $\hat{\varphi}_1$ is the one-dimensional Fourier transform of $\varphi(r)$, which in turn was the two-dimensional Fourier transform of $\hat{\varphi}$ that vanishes at $k > 1$. By comparing their respective inverse Fourier transforms, one finds

$$\hat{\varphi}_1(x) = 2\int_0^{\sqrt{1-x^2}} \hat{\varphi}\left(\sqrt{x^2+k^2}\right)\mathrm{d}k = 2\int_x^1 \frac{\hat{\varphi}(r)\, r\; \mathrm{d}r}{\sqrt{r^2-x^2}}. \tag{2.50}$$

It follows that $\hat{\varphi}_1(x)$ also vanishes at $x > 1$, which implies that the density vanishes for $|z| > \mu$, i.e. $\rho(\boldsymbol{r})$ describes a flaring elliptic disc with vanishing column density and power-law mid-plane profile. Possible functions $\hat{\varphi}_1(r)$ are listed in Table 2.A.2.

2.A.4 Near-invisible densities

We now consider simple analytic density distributions with projections that do not vanish exactly, but are potentially very small. These may be useful in numerical work, for example as a perturbation to be added to another model as input for an iterative deprojection algorithm, or as a component of a superposition-based deprojection.

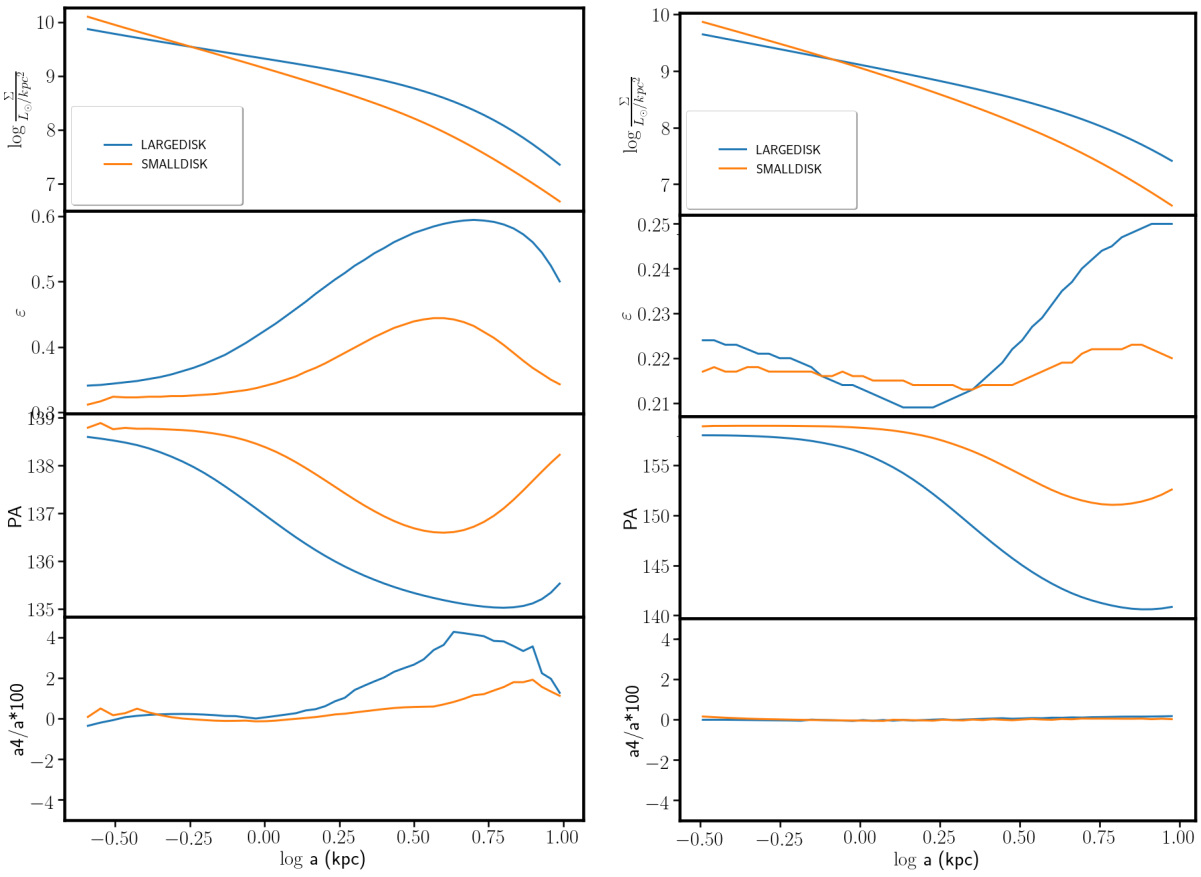
Near-invisible ellipsoidal models

The idea here is to replace the functions of compact support used in the previous sub-section with more general ellipsoidal models, i.e. use the recipe (2.38) with some model $f(\boldsymbol{r})$ whose Fourier transform $\hat{f}(\boldsymbol{k})$ may not vanish anywhere. Then, of course, the resulting $\hat{\rho}(\boldsymbol{k})$ will not vanish on the four planes $\boldsymbol{\ell}_i \cdot \boldsymbol{k} = 0$, but can be small on these planes if $\hat{f}(\boldsymbol{k})$ decays sufficiently fast and the scale s is sufficiently small, such that the projection Σ, though not vanishing, is hardly visible.

The simplest case is a near-invisible ellipsoidal Gaussian, when this recipe gives density

$$\rho(\boldsymbol{r}) = \cos\left(\frac{\boldsymbol{r}\cdot\hat{\boldsymbol{x}}_i}{s}\right)\frac{\exp\left(-\frac{1}{2}\boldsymbol{r}^t\cdot\mathbf{C}^{-1}\boldsymbol{r}\right)}{\sqrt{(2\pi)^3|\mathbf{C}|}}. \tag{2.51}$$

[10] Or $n > 1$ if $\hat{f}(0) = 0$ is required.



(a) $\theta = 80°$ (b) $\theta = 45°$

Figure 2.A.3: Surface brightess ellipticity ε, $a4$ and twist profiles as a function of semi-major axis when we project the *LARGEDISC* or the *SMALLDISC* model at $\theta = 80°$ (left) or $\theta = 45°$ (right), in both cases with $\phi = \psi = 45°$. High ellipticity and the positive $a4$, which are clear markers of a disc-like component, are only present for $\theta = 80°$ but not 45°.

Adding such a model generates a disc in the plane perpendicular to $\hat{\boldsymbol{x}}_i$ with projected surface density

$$\Sigma(\xi,\eta) = \exp\left(-\tfrac{1}{2}\ell_i^2\sigma_\ell^2/s^2\right)\ \cos\left[\frac{1}{s}\begin{pmatrix}\xi\\ \eta\end{pmatrix}\cdot\begin{pmatrix}\xi_i+\ell_i\sigma_\ell^2\sigma_{\xi\ell}^{-2}\\ \eta_i+\ell_i\sigma_\ell^2\sigma_{\eta\ell}^{-2}\end{pmatrix}\right]\ \Sigma_0(\xi,\eta), \tag{2.52}$$

where σ_ℓ was given in equation (2.40),

$$\sigma_{\xi\ell}^{-2} = \boldsymbol{\xi}^t\cdot\mathbf{C}^{-1}\cdot\boldsymbol{\ell}, \qquad \sigma_{\eta\ell}^{-2} = \boldsymbol{\eta}^t\cdot\mathbf{C}^{-1}\cdot\boldsymbol{\ell}, \tag{2.53}$$

while

$$\Sigma_0(\xi,\eta) = \frac{\exp\left(-\tfrac{1}{2}(\xi,\eta)^t\cdot\bar{\mathbf{C}}^{-1}\cdot(\xi,\eta)\right)}{\sqrt{(2\pi)^2|\bar{\mathbf{C}}|}} \tag{2.54}$$

is the the projected density of an ellipsoidal Gaussian. Thus, Σ differs from that of an ellipsoidal Gaussian by both a cosine modulation and suppression factor. For a substantial suppression $s \ll \ell_i\sigma_\ell$, which favours discs near-perpendicular to the LOS so that ℓ_i is large.

Near-invisible elliptical discs

We can also use a Gaussian for $\hat{\varphi}$ in the recipe of §2.A.3, i.e.

$$\hat{\rho}(\boldsymbol{k}) = k_z^2\mathrm{e}^{-\frac{1}{2}(a^2k_x^2+b^2k_y^2)/k_z^2}. \tag{2.55}$$

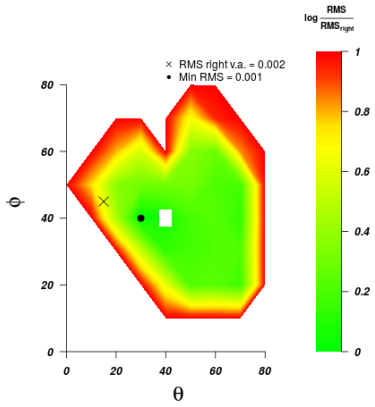
$$\rho(\boldsymbol{r}) = \frac{1}{(2\pi)^{3/2}ab}\frac{1}{\mu^5}\left(3-6\frac{z^2}{\mu^2}+\frac{z^4}{\mu^4}\right)\mathrm{e}^{-\frac{1}{2}z^2/\mu^2}. \tag{2.56}$$

2.B Probing the effects of hidden discs

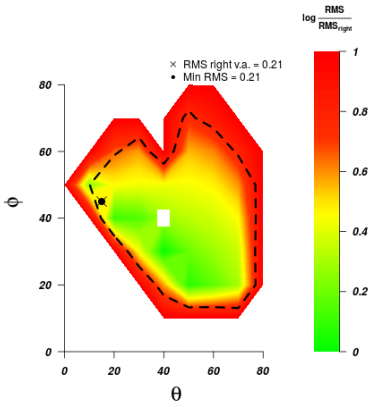
Massive elliptical galaxies have nearly elliptical isophotes and this justifies the assumption of the deformed ellipsoidal deprojection algorithm discussed in the previous sections. However, even these objects could harbour (possibly faint) disc components, possibly nearly invisible in projection (see discussion in Appendix 2.A). Here we explore the effects of hidden discs by considering a flat component whose intrinsic light density ρ_D is described by a double exponential profile, reminiscent of those observed for spiral galaxies:

$$\rho_D = N\mathrm{e}^{-\frac{\sqrt{x^2+y^2}}{h}}\mathrm{e}^{-\frac{z}{h_z}}. \tag{2.57}$$

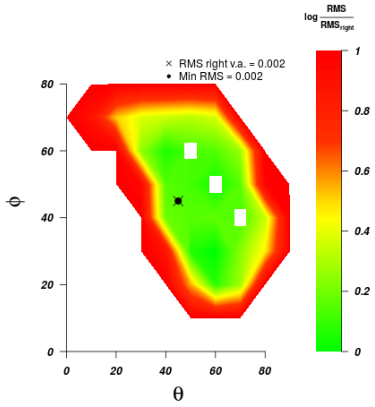
We choose the scale length and height to be $h = 0.5$ and $h_z = 0.1$, respectively, such that the half-light radius is similar to the one of the Jaffe model used above and the structure is flatter than the most flatten elliptical galaxies known. N is a normalization factor used to vary the disc mass. The density contours in the meridional plane are rhombi, i.e. quite different from the deformed ellipses of equation (2.29). We deproject the projection of ρ_D using our implementation of M99's



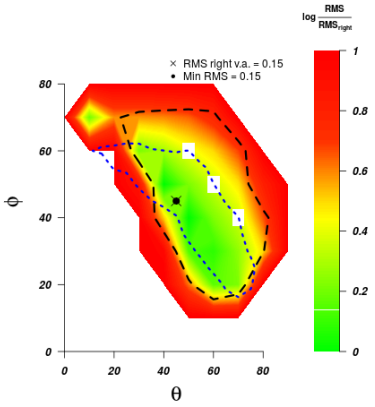
(a) SB comparison, $\theta = 15°$.



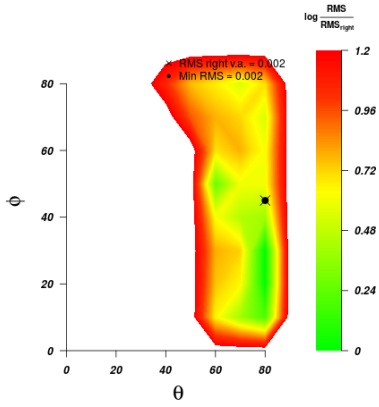
(b) ρ comparison, $\theta = 15°$.



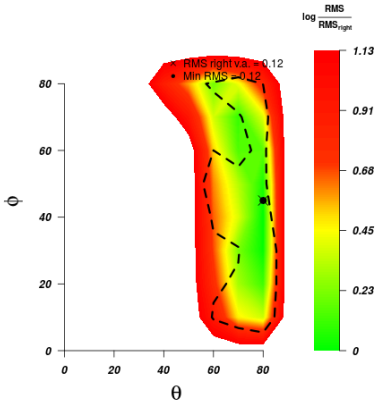
(c) SB comparison, $\theta = 45°$.



(d) ρ comparison, $\theta = 45°$.



(e) SB comparison, $\theta = 80°$.



(f) ρ comparison, $\theta = 80°$.

Figure 2.B.1: Same as Figs. 2.18-2.20 for model *SMALLDISC*. As in Fig. 2.18 (reproduced here by the blue dashed contour in the middle left panel), the area of good fits overlaps well with that where the intrinsic density matches the true one, but is larger. The true viewing angles are recovered well except for the small discrepancy at $\theta = 15°$.

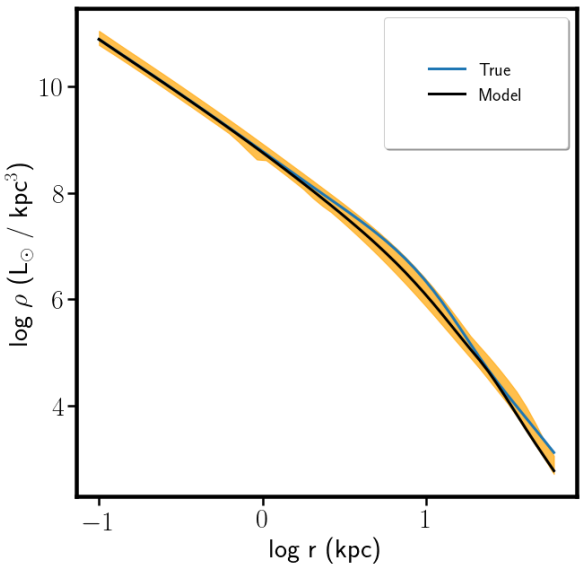


Figure 2.B.2: The intrinsic density along the major axis of model *SMALLDISC* (blue), and the range of densities recovered with the constrained-shape method for viewing angles compatible with the surface brightness obtained projecting at $\theta = \phi = \psi = 45°$. The black line shows the deprojection assuming these angles.

code, finding, as expected, that the deprojection is unique for $\theta = 90°$ and it can be tuned towards the true density by using the d_4 parameter of the code to obtain discy isophotes at lower θ angles.

As a second step, we sum to the density of *ELLIP* the density ρ_D with normalisation N chosen such that the two components have mass ratios of 1 (*LARGEDISC*) or 5.67 (*SMALLDISC*, where the flattened component has 15% of the total mass). We project these densities for $\theta = 80°$, $45°$, and $15°$ with $\phi = \psi = 45°$. Decreasing θ makes it easier to hide the flattened component in projection. For *SMALLDISC* (and even more for *LARGEDISC*), the isophotes of the projected density at $\theta = 80°$ show a clear signature (high ellipticity and a_4 values, see Fig. 2.A.3, left). At $\theta = 45°$ the only possible signature for *LARGEDISC* is a $\sim 20°$ twist (Fig. 2.A.3, right), which lies just on the threshold of what we can observe in massive ellipticals (see Fig. 2.9a).

We are always able to deproject *SMALLDISC* using the constrained-shape method, matching well the projected surface brightness and with resonably good precision the intrinsic density, getting RMS in ρ of 12%, 15%, 20% at $\theta = 80°$, $45°$, $15°$, respectively. This corresponds to the range in density errors found when reconstructing the viewing angles for the Jaffe-only density (see Figs. 2.18 and 2.B.2). However, the region of allowed viewing angles in these cases is larger (Fig. 2.B.1).

For *LARGEDISC* the situation is more difficult. Given the strongly non-elliptical isophotes of the $\theta = 80°$ projection, the constrained-shape algorithm is unable to deliver projected densities matching the true ones. We cure this problem by modifying the deprojection algorithm: we add a non-parametric, axisymmetric, flattened component, that is added to the one with deformed ellipsoidal shape, and optimize it subject to regularization constraints together with the first component through the Metropolis procedure. With this code we are able to reproduce well the

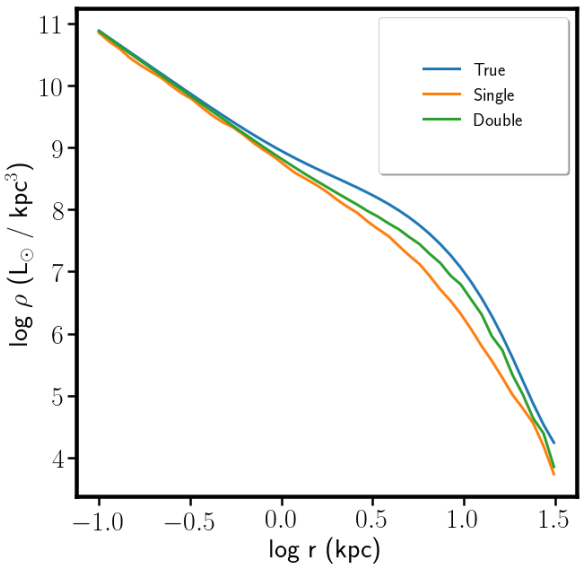


Figure 2.B.3: The intrinsic density along the major axis of *LARGEDISC* (blue), the density recovered with the constrained-shape method without (orange, 'single') and with a complementary axisymmetric model (green, 'double') assuming the true viewing angles $\theta = \phi = \psi = 45°$.

SB profile, recovering the intrinsic density with an RMS of less than 9%. Of course, since in this case the disc's signature can be seen in the photometry (Fig. 2.A.3a), we may also directly subtract it from the galaxy image as done by Scorza & Bender (1990).

When we project *LARGEDISC* at $\theta = 45°$ or 15°, the disc becomes impossible to spot from a photometric analysis alone (Fig. 2.A.3b) and the constrained-shape algorithm is able to reproduce the observed surface brightness very well. However, the intrinsic density can only be recovered up to an RMS of $\sim$ 36% (or even worse when $\theta = 15°$). Using the modified, constrained-shape-plus-axisymmetric-component algorithm we are able to reproduce the observed surface brightness to the same precision and the intrinsic density with an RMS of $\sim$ 22% (see Fig. 2.B.3). We *do not* see such a strong difference between the densities reconstructed with or without complementing the constrained-shape method with an axisymmetric component for models *SMALLDISC*, *ELLIP*, or *DISCYBOXY*.

This exploration can guide us when deprojecting the surface photometry of real elliptical galaxies that do not have clear signs for the presence of a disc component. If a disc component is present, we expect that the differences between intrinsic densities recovered with and without a complementary axisymmetric component exceed the variations observed as function of assumed viewing angles.

Chapter 3

Intrinsic shapes of Brightest Cluster Galaxies

3.1 Introduction

[1]Massive elliptical galaxies show, when observed projected on the plane of the sky, smooth elliptical contours with mild boxy deviations ($a_4 > 0$, Bender & Möllenhoff 1987), small twists (typically $\lesssim 10°$) and increasing ellipticity (ε) profiles towards outer radii (Goullaud et al., 2018). The average projected flattening is $< q' > \sim 0.8$ (Tremblay & Merritt, 1996; Weijmans et al., 2014; Chen et al., 2016; Ene et al., 2018), with $q' \equiv 1 - \varepsilon$. The presence of isophote twists, together with other factors such as the statistical distribution of ellipticity profiles provides evidence for the triaxiality of these objects (Illingworth & King, 1977; Bertola & Galletta, 1978; Vincent & Ryden, 2005).
Various works in the last 25 years have studied the average statistical distribution of intrinsic shapes for large galaxy samples at different redshifts z (Tremblay & Merritt, 1996; Vincent & Ryden, 2005; Weijmans et al., 2014; Chang et al., 2013; Chen et al., 2016; Ene et al., 2018; Li et al., 2018b). Deprojecting these distributions (e.g. Tremblay & Merritt 1996) allows to recover the *average* or *typical* intrinsic shape of the galaxies. These studies find that most massive objects are indeed triaxial, with a mean triaxiality parameter $T = \left(1 - p^2\right) / \left(1 - q^2\right)$ (Franx et al., 1991) in the range $[0.4, 0.8]$ (Vincent & Ryden, 2005), where $p \equiv b/a$, $q \equiv c/a$, and $a \geq b \geq c$ are the lengths of the three principal axes of the density ellipsoid. However, no study has yet attempted to *directly measure radially resolved intrinsic shapes of individual galaxies in large samples*.

In a recent paper, de Nicola et al. (2020) have presented a triaxial deprojection routine that fits the intrinsic shape of ellipsoidal galaxies and allows to constrain the viewing angles under which an object is seen by photometric data alone. This can be refined further in combination with the dynamical modeling of appropriate stellar kinematics (de Nicola et al., submitted), since the number of deprojections which need to be tested is drastically reduced, allowing the study of large samples of galaxies.

[1]Published paper: de Nicola S., Saglia R.P., Thomas J., Pulsoni C., Kluge M., Bender R., Valenzuela L., Remus R.-S. - Intrinsic shapes of Brightest Cluster Galaxies, ApJ, 2022, 933, 215

An interesting group of massive galaxies consists of the so-called Brightest Cluster Galaxies (BCGs). According to Kluge et al. (2020), a BCG is defined as the closest galaxy to the geometrical and the kinematical centre of a given galaxy cluster, although not necessarily the most luminous galaxy of the cluster itself. Lying deep in the potential well of the cluster, these giant ellipticals are able to increase their mass through processes such as galaxy mergers (Contini et al., 2018), cannibalism or tidal stripping (Mo et al., 2008). In a recent paper by Kluge et al. (2020), a sample of 170 BCGs was analyzed in great detail using extremely deep photometric observations, revealing that BCGs follow different scaling relations with respect to ordinary ETGs. BCGs are also interesting because their outer parts have probably grown predominantly by collision-less accretion and, hence, in a manner similar to the proposed growth of (collision-less) dark-matter halos.

The first goal of this paper is to constrain the intrinsic shapes and viewing angles of a representative subsample (56 objects) of this BCG catalogue with the deprojection method of de Nicola et al. (2020). Then, our second goal is to compare the recovered shapes to the ones of simulated massive galaxies and their dark matter halos. For this purpose we consider the IllustrisTNG (Nelson et al., 2018; Springel et al., 2018; Marinacci et al., 2018; Naiman et al., 2018; Pillepich et al., 2018) and Magneticum pathfinder (Hirschmann et al. 2014; Teklu et al. 2015[2]) simulations. These cosmological (magneto)-hydrodynamical model the formation and evolution of galaxies in a ΛCDM Universe including recipes for star formation and evolution, chemical enrichment of the inter-stellar medium, gas cooling and heating, black hole and supernova feedback. These simulations produce galaxy populations with properties in reasonable agreement with observations (Remus et al., 2017; Teklu et al., 2017; Genel et al., 2018; van de Sande et al., 2019; Rodriguez-Gomez et al., 2019; Pulsoni et al., 2020; Remus & Forbes, 2021).

The paper is structured as follows. Section 2 describes the galaxy sample used in this work. In Section 3 we explain the deprojection procedure. In Section 4 we present the results on the statistics of triaxial shapes and compare our findings with the TNG and Magneticum simulations. Finally, we draw our conclusions in Section 5. Throughout the paper we assume a flat cosmology with $H_0 = 69.6 kms^{-1}Mpc^{-1}$ and $\Omega_m = 0.286$.

3.2 The sample

The BCGs studied in this work come from a recently published sample (Tab. 1 of Kluge et al. 2020). Each BCG was observed in the g' band with the 2m Fraunhofer telescope at the Wendelstein Observatory (see Kluge et al. 2020 for technical details). The photometry has exactly the requirements needed for the present work, being extremely deep (down to $m_{g'} \sim 30$ mag) and reaching very large radii (typically well beyond 100 kpc). From the complete sample, we extract those objects for which supplementary F606W HST photometry (typical resolution ∼0.15") is available, excluding galaxies that are overall unrelaxed (see below). To combine these high-resolution data with those coming from Wendelstein observations, we first select the radii where we have data from both observation sets that are not affected by seeing (typically from 5

[2] www.magneticum.org

to 15 arcsec from the center), then we interpolate HST photometry at Wendelstein radii. Finally, we convert the HST data to the g' band, by determining the sky level and the scaling factor that minimize the differences between the two photometric sets. We combine the two sets by taking the HST values in the inner 10-15 arcsec, and the Wendelstein values at larger radii. In this way we have photometric data with both very high resolution in the center and also extending out to ∼100 kpc for the majority of the objects[3]. We complement this list with further 8 BCGs which we recently observed both in the H and/or Ks bands at the 8.4m Large Binocular Telescope (LBT) using Adaptive Optics, with typical resolution of ∼0.4". We combine the LBT photometry with the Wendelstein one using the same approach described above. In App. 3.A we show a comparison of the deprojection of two galaxies with and without high-resolution photometry to explore possible photometric effects, showing that they are small. Without high resolution data, the deprojections cannot probe the central regions of the galaxies, but reliable profiles are derived at larger radii.
This allows us to add 16 more BCGs with only Wendelstein data, for a total of 56 galaxies. The average isophotal flattening $< q' >$ is ~ 0.77, although almost every BCG becomes very flat ($q' \lesssim 0.4$) at large radii. In App. 3.B we show the ε and PA profiles for every BCG of the sample. Since BCGs often show signs of interactions with other neighbor galaxies of the cluster or AGN activity in the central regions (Kluge et al., 2020), and given that our triaxial code works under the assumption of (smooth) "deformed ellipsoids" (see eq. 29 of de Nicola et al. 2020), we omit the innermost/outermost isophotes from the deprojection when we find signs of incomplete relaxation (for example, in the form of bumpy εor PA profiles). This happens in the very center (e.g. AGNs, ongoing accretion) or in the very outer parts (where dynamical time scales are large). Notes on individual galaxies can be found in App. 3.D. Since we are interested in comparing our findings with simulations at large radii, we try, when possible, to extend the deprojection up to $2 - 4R_e$, with the values for half-light radii R_e taken from Tab. 4 of Kluge et al. (2020).

3.2.1 Selection effects

The full sample of Kluge et al. (2020) is drawn from the Abell–Corwin–Olowin catalog (ACO, Abell et al. 1989) by adopting redshift and volume-limiting constraints. It contains BCGs with redshift $z \lesssim 0.08$ (with 15 outliers) and has a slight Malmquist bias (see their Figure 3). However, a comparison of this sample with other large samples, also drawn from ACO, shows that they have about 80-90% of the objects in common.
In order to check the completeness of our sub-sample, we define R_N as the ratio of the number of galaxies in our sub-sample N_{sub} to the number of galaxies in the full sample N_{full}. We become progressively incomplete at larger redshifts: at $z \leq 0.04$, we have $0.5 \leq R_N \leq 1.0$, but at larger redshifts R_N is ≤ 0.2. Moreover, Fig. 3.1, left panel, shows that the mean redshift and redshift range covered by our subsample are smaller than the ones of the parent sample. This is expected, since both HST- and LBT-observed galaxies are at lower redshift than the average, and we picked them up to perform dynamical modeling.

[3]As shown by Kluge et al. (2020), at larger radii (typically as SB approaches 27-28 mag arcsec^{-2}) isophotal shape profiles become too noisy to be estimated reliably.

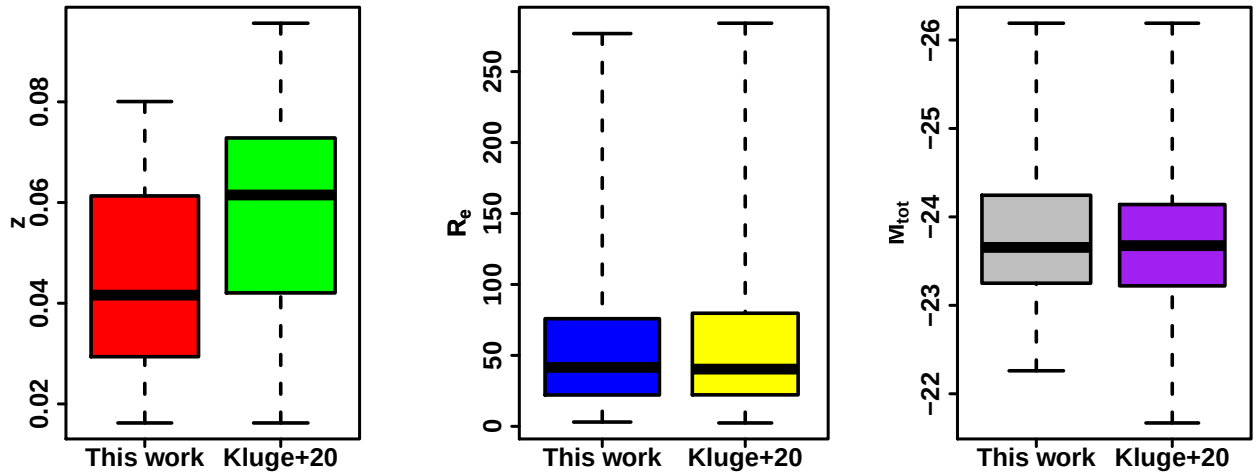


Figure 3.1: Boxplots showing redshifts (left panel), effective radii (central panel) and g' magnitudes (right panel) for Kluge et al. (2020)'s sample and our sub-sample. Our sub-sample is biased towards galaxies at lower redshift and also slightly towards galaxies with high R_e. Boxes are drawn from the first to the third quartiles, with the horizontal line in the middle of the boxes denoting the median. Whiskers span across the whole data range.

We do not find significant selection effects when considering the size (Fig. 3.1, middle panel) or the total absolute magnitude of the galaxies M_{tot}, (Fig. 3.1, right panel). The results are summarised in Tab. 3.1.

3.3 Deprojection procedure

In this section we describe the deprojection parameters used for the BCGs. An extensive description of the deprojection routine itself is given by de Nicola et al. (2020). In short, given the observed surface luminosity $I_{obs} = L/pc^2$ onto a polar elliptical grid, the code searches for the three-dimensional luminosity density ρ, placed onto an ellipsoidal grid, whose corresponding projected surface luminosity I_{fit} minimizes RMS= $\sqrt{< (\ln(I_{obs}/I_{fit}))^2 >}$. The algorithm works under the assumption that a galaxy can be described by what we call a "deformed ellipsoid", namely an ellipsoid whose radius is given by

$$m^{2-\xi(x)} = x^{2-\xi(x)} + \left[\frac{y}{p(x)}\right]^{2-\xi(x)} + \left[\frac{z}{q(x)}\right]^{2-\xi(x)} \tag{3.1}$$

where the exponent ξ can be used to generate disky
($\xi > 0$) or boxy ($\xi < 0$) bias. The three one-dimensional functions $p(x), q(x)$, and $\xi(x)$, along with the density on the x-axis $\rho_x(x)$, specify ρ at each point of the grid. Finally, the code uses

Variable	Intervals	R_N
	≤ 0.02	1.0
z	$[0.02, 0.05]$	~0.5
	≥ 0.05	~0.15
R_e	≥ 50	~0.25
M_{tot}	$[-22, -26]$	~0.25

Table 3.1: The ratio R_N of the number of galaxies in our sub-sample N_{sub} for a given interval to the number of galaxies in the full sample N_{full} for three variables of interest. Our sub-sample is biased towards low-redshift galaxies, while no significant bias in magnitude and effective radii is found.

a one-dimensional radial smoothing on $p(x)$, $q(x)$, $\xi(x)$ and $\rho_x(x)$ to penalize against unsmooth solutions.

3.3.1 Choice of parameters

Table 3.2 highlights for each BCG the parameters used for the deprojections. In the following we will give a detailed explanation of these parameters.

- *Grid sampling.* It is important to place both the observed I_{obs} and deprojected ρ on grids large enough to reproduce photometric information properly, but at the same time a very large grid would slow down the code significantly. Therefore, we start by interpolating the SB onto a finely sampled grid, and then gradually reduce both the number of radial ($n_{r'}$) and angular points ($n_{\theta'}$) as long as the comparison with the observations remains acceptable, that is, an error below 1% for every photometric variable (SB, ε, PA, a_4, a_6). Then, we set the number of radial points of the ρ-grid $n_r = n_{r'} + 20$, while the numbers of angular points n_θ and n_ϕ are typically the same as ($n_{\theta'}$) or slightly larger.

- *Grid extension in radius.* The innermost radii are the same for both grids, namely the semiminor axis of the innermost isophote. These have to be estimated for every BCG by taking into account the spatial resolution of the observations and by checking whether the galaxy shows central activity. In this last case, central regions are omitted. The outermost radii for the SB grid are estimated by making use of our software for isophotal fitting (Bender & Möllenhoff, 1987). We typically stop when we the isophotal shape profiles become noisy and have to be set to a constant values (typically at SB~27-28 mag asec^{-2}). The largest radii of the ρ-grid are then a few times those of the corresponding SB grid.

- *Grid flattenings.* The flattening of the SB grid can easily be estimated by considering the isophote PAs. For example, if a galaxy shows isophote structures with the major axes

aligned along the vertical axis on the plane of the sky y', it makes sense to use an elliptical polar grid flattened in the x'-direction, even with a 15-20° twist. As far as it concerns ρ, we first assume a spherical grid, then we re-deproject the galaxies at the best-fit inclination(s) using the recovered $< p(r) >$ and $< q(r) >$ as flattenings.

- *Smoothing.* As shown in de Nicola et al. (submitted) we can recover the true *intrinsic* density of a triaxial N-body simulation with an RMS of ∼10%. Since it is not entirely clear how to estimate the smoothing a priori, we take the four λ-values $\left[\lambda_\rho, \lambda_p, \lambda_q, \lambda_\xi\right]$ (cfr. eq. 30 of de Nicola et al. 2020) used with the simulation divided by a factor of 2, to take into account that our data are less noisy than the N-body simulation (the smoothing scales as λ^{-2}). Since the smoothing value affects the RMS one gets at the end (the higher the smoothing, the higher the RMS), we verified that for the best galaxies the RMS was comparable to the one we got for the simulation. The values we chose are [0.6,0.03,0.03,0.3]. A more rigorous implementation would be the minimization of the Akaike Information Criterion (AIC, Akaike 1974), as shown by Lipka & Thomas (2021) (see also Thomas & Lipka 2022). We defer this to a forthcoming paper.

- *Constraints on p, q.* Our code allows the possibility of deprojecting by imposing constraints on p, q. Since the code has shown excellent results in terms of recovering the right profiles, we only impose $p, q \geq 0.2$, to prevent too flat solutions which may give problems to the fit.

3.3.2 Viewing angles

Deprojections yield in general a non-unique solution (Gerhard, 1996; van den Bosch, 1997; de Nicola et al., 2020), unless one uses constraints about the galaxy shape. We impose the resulting deprojection to consist of a series of concentric perfect ellipsoids, i.e. setting $\xi(r) = 0$ at all radii[4], and without imposing biasing towards certain p, q profiles. Further parameters to be considered are the three viewing angles (θ, ϕ, ψ) required to identify the orientation in space of a triaxial galaxy. The first two give the orientation of the line-of-sight (LOS) in space, whereas the third one is a rotation about the LOS itself. If one could measure these angles, then the intrinsic axis ratios of the ellipsoid could be calculated analytically. Since this is not usually the case, we need to sample a grid of viewing angles and deproject every BCG trying out every set of viewing angles. The assumption of triaxiality, i.e. the galaxy has an 8-fold symmetry, allows us to sample the two viewing angles (θ, ϕ) in $[0, \pi/2]$, while ψ needs to be sampled in $[0, \pi]$. We sample each angle in 10°step, which gives us a total of 1800 inclinations to test for each BCG. As shown by de Nicola et al. (2020), although the viewing angle estimate through deprojections is not always perfect, the list of "good" solutions (see below) *always* includes the correct viewing angles.
In the same paper, the authors show that sampling one octant does not guarantee the "canonical" $1 \geq p \geq q$ order relationship (cfr. their Table 2). However, for a given deprojected density at a given set of viewing angles for which this does not happen, *it is always possible to find another*

[4]This assumption is justified by the fact that for the BCGs the deviations from elliptical shapes are small, i.e. $|a_4/a| < 5\%$ in the deprojected regions.

Galaxy	kpc/arcsec	r_{min}	r_{max}	I_{obs} grid	ρ grid	Photometry
2MASXJ0753	1.17	2.13	128.1	40×15	$60 \times 16 \times 16$	L + W
2MASXJ0900	1.426	2.35	104.2	50×15	$70 \times 15 \times 15$	W
2MASXJ1358	1.225	1.98	74.04	40×10	$60 \times 11 \times 11$	W
IC613	0.653	0.0940	116.1	50×10	$70 \times 11 \times 11$	H + W
IC664	0.679	0.0984	76.0	40×10	$60 \times 11 \times 11$	H + W
IC1101	1.48	1.06	70.2	40×12	$60 \times 13 \times 13$	H + W
IC1565	0.765	0.101	243.2	40×10	$60 \times 11 \times 11$	H + W
IC1634	1.336	1.58	156.8	50×10	$70 \times 11 \times 11$	W
IC1695	0.987	0.150	149.1	40×10	$60 \times 11 \times 11$	H + W
IC1733	0.714	0.528	33.8	30×12	$50 \times 13 \times 13$	W
IC2378	0.990	0.762	58.5	40×10	$60 \times 11 \times 11$	H + W
IC5338	1.10	4.91	178.3	40×10	$60 \times 11 \times 11$	H + W
LEDA1518	1.248	4.16	117.1	40×10	$60 \times 11 \times 11$	W
LEDA2098	1.467	2.82	127.3	40×10	$60 \times 11 \times 11$	W
MCG+01-60	1.16	0.622	38.4	30×12	$50 \times 13 \times 13$	W
MCG-02-02	1.083	3.44	173.7	40×10	$60 \times 11 \times 11$	W
MCG+02-04	0.869	0.213	184.3	50×12	$70 \times 13 \times 13$	H + W
MCG+02-27	0.653	0.114	103.4	40×10	$60 \times 11 \times 11$	H + W
MCG+02-58	1.52	1.69	100.1	40×10	$60 \times 11 \times 11$	H + W
MCG+03-04	1.375	2.76	66.34	40×10	$60 \times 11 \times 11$	W
MCG+03-38	0.886	0.165	87.6	40×10	$60 \times 11 \times 11$	H + W
MCG+04-28	2.53	1.15	299.7	40×12	$60 \times 13 \times 13$	H + W
MCG+05-32	1.44	1.19	312.9	40×10	$60 \times 11 \times 11$	H + W
MCG+05-33	1.23	2.64	67.1	40×10	$60 \times 11 \times 11$	H + W
MCG+09-13	1.362	3.78	103.1	40×10	$60 \times 11 \times 11$	W
MCG+09-20	1.296	2.49	133.4	40×10	$60 \times 11 \times 11$	W
NGC708	0.332	0.0551	61.2	40×15	$60 \times 15 \times 15$	H + W
NGC910	0.354	0.611	70.4	40×10	$60 \times 11 \times 11$	H + W
NGC1128	0.486	0.0895	31.7	40×10	$60 \times 11 \times 11$	H + W
NGC1129	0.361	0.160	98.0	60×12	$80 \times 12 \times 12$	H + W
NGC1275	0.359	0.739	81.9	40×10	$60 \times 11 \times 11$	W
NGC2329	0.396	0.534	43.3	50×12	$70 \times 13 \times 13$	H + W
NGC2804	0.559	1.58	35.6	40×10	$60 \times 11 \times 11$	W
NGC3550	0.703	0.505	99.9	40×12	$60 \times 13 \times 13$	L + W
NCG3551	0.640	0.382	33.4	30×12	$50 \times 13 \times 13$	W
NGC4104	0.577	0.970	60.2	40×12	$60 \times 13 \times 13$	L + W
NGC4874	0.469	0.0905	98.6	50×10	$70 \times 11 \times 11$	H + W
NGC6166	0.622	0.718	94.9	40×10	$60 \times 11 \times 11$	H + W
NGC6173	0.592	0.0613	109.2	40×10	$60 \times 11 \times 11$	H + W
NGC6338	0.552	0.334	132.4	50×10	$70 \times 11 \times 11$	H + W
NGC7647	0.818	0.323	97.5	50×10	$70 \times 11 \times 11$	H + W
NGC7649	0.835	0.659	94.1	40×10	$60 \times 11 \times 11$	H + W
NGC7720	0.611	0.720	153.1	40×10	$60 \times 11 \times 11$	H + W
NGC7768	0.545	0.307	102.3	40×10	$60 \times 11 \times 11$	H + W
SDSSJ0837	2.67	1.66	118.0	40×20	$60 \times 20 \times 20$	L + W
UGC716	1.19	0.832	175.7	40×10	$60 \times 11 \times 11$	L + W
UGC727	1.135	1.89	114.4	40×10	$60 \times 11 \times 11$	W
UGC1191	1.21	0.638	124.5	40×20	$60 \times 20 \times 20$	L + W
UGC2232	0.958	0.152	95.3	40×10	$60 \times 11 \times 11$	H + W
UGC2413	0.690	0.379	113.4	50×10	$70 \times 11 \times 11$	H + W
UGC4289	0.587	0.587	50.1	40×10	$60 \times 11 \times 11$	H + W
UGC6394	0.847	1.99	97.2	40×10	$60 \times 11 \times 11$	W
UGC9799	0.691	4.13	97.7	40×10	$60 \times 11 \times 11$	H + W
UGC10143	0.708	0.584	110.6	50×10	$70 \times 11 \times 11$	H + W
UGC10726	1.15	0.704	189.8	60×12	$80 \times 12 \times 12$	L + W
VV16IC	0.354	0.401	181.8	40×10	$60 \times 11 \times 11$	L + W

Table 3.2: *Col. 1*: Galaxy name. *Col. 2*: kpc/arcsec conversion factor. *Cols. 3-4*: Smallest and largest isophotal radii, in kpc. *Cols. 5-6*: I_{obs} and ρ-grid dimensions. *Col. 7*: The available photometry (W: Wendelstein, H: HST, L: LBT).

set of viewing angles for which the deprojection is equivalent but with $1 \geq p \geq q$, possibly in the octant with $\phi \in [\pi/2, \pi]$. In this case, if the density happens to be a "good" one in terms of the RMS (see below), we re-perform the deprojection for this new set of viewing angles such that the inequality $1 \geq p \geq q$ holds.
In order to find the "good" deprojections, we isolate for every galaxy g all viewing angles for which the deprojection has $\mathcal{R}_g \leq \delta \times \mathcal{R}_{\rm min,g}$, where $\mathcal{R}_g \equiv {\rm RMS}_g$ is the RMS and $\mathcal{R}_{\rm min,g}$ is the smallest RMS that we find among all viewing angles that we sampled for galaxy g. The factor δ determines how stringent the cut-off is. In another work aiming at modeling an N-body simulation (de Nicola et al., submitted), we adopt values in the range $\delta = 1.2 - 1.5$ studying the impact of the choice of this value on the number of plausible deprojections. Here, we try to be conservative adopting $\delta = 1.5$ The resulting light densities are those that we use to derive the results shown in Sec. 3.4.2.
Finally, we note that there may still be a degeneracy between the model and orientation parameters, but this is small and will not be taken into account throughout the remainder of the paper. In App. 3.A we examine the case of NGC7647, and derive profiles for the best-fit viewing angles, stopping the deprojection when the RMS reaches 1.5 ${\rm RMS}_{min}$ (which is the best value achievable). These p & q profiles are well within the range obtained by considering all possible viewing angles.

3.4 Results

Using the cut-off described in Sec. 3.3.2 we are able to reduce the number of inclinations from the initial value of 1800 by at least a factor of 3. The typical RMS values for the best-fit solutions are 0.01 to 0.03. The results are summarized in Tab. 3.3.

3.4.1 Reliability of the deprojections

As a first step we verify that our deprojections do reproduce the average photometry of the sample. First, we calculate for every galaxy the mean ε and the twist, defined as ΔPA = max(PA) - min(PA), both for the observed and the recovered photometry. Moreover, we reproject the best-fit densities ρ_g for every galaxy g at three different random viewing angles, computing the same averages as above. This is a good test to statistically verify that the recovered intrinsic shapes are compatible with the observed shape distribution. In Fig. 3.2 we show the histograms for ε (top row) and the twist (bottom row). A Kolmogoroff-Smirnov (KS, Kolmogorov 1933; Smirnov 1939) test returns p-values above the canonical 5% threshold[5] for both the $\varepsilon$and PA distributions, with this being valid for both the best-fit angles and the reprojections at random viewing angles. This confirms that the recovered photometric variables are statistically representative of the BCG sample.
A second step, we check the distribution of the two best-fit angles $< (\theta, \phi) >$, which specify the LOS position on the plane of the sky. In the upper panels of Fig. 3.3 we plot the two octants with the best-fit $< (\theta, \phi) >$ onto them. We see that there is a lack of solutions near the principal axes,

[5]This corresponds to values in the range 0.194-0.243 for the KS statistics D_n.

Galaxy	$10^2\times$ RMS$_{\rm best}$	Good Deprojections	best-fit angles	$< p >$	$< q >$	Δ_p	Δ_q
2MASXJ0753	5.1	96	(26,27,37)°	0.804	0.661	0.048	0.054
2MASXJ0900	2.2	114	(54,37,122)°	0.902	0.769	0.053	0.047
2MASXJ1358	5.0	62	(50,40,60)°	0.806	0.568	0.123	0.111
IC613	2.6	250	(70,50,90)°	0.879	0.782	0.047	0.040
IC664	3.2	204	(60,40,145)°	0.871	0.704	0.080	0.073
IC1101	3.0	28	(50,80,160)°	0.798	0.644	0.076	0.075
IC1565	2.4	312	(60,28,167)°	0.833	0.697	0.041	0.041
IC1634	2.8	378	(60,40,110)°	0.897	0.760	0.050	0.058
IC1695	2.9	376	(50,40,80)°	0.854	0.712	0.061	0.057
IC1733	0.73	380	(53,16,165)°	0.815	0.695	0.108	0.116
IC2378	1.6	118	(70,100,150)°	0.841	0.708	0.082	0.065
IC5338	3.2	58	(60,40,160)°	0.824	0.615	0.082	0.073
LEDA1518	4.1	146	(60,60,120)°	0.876	0.711	0.091	0.113
LEDA2098	2.3	64	(60,30,60)°	0.858	0.716	0.067	0.066
MCG+01-60	0.81	330	(36,34,110)°	0.909	0.824	0.084	0.122
MCG-02-02	2.2	94	(60,50,15)°	0.786	0.560	0.084	0.081
MCG+02-04	2.6	128	(70,130,90)°	0.876	0.747	0.033	0.033
MCG+02-27	3.0	184	(60,132,3)°	0.799	0.567	0.076	0.062
MCG+02-58	4.2	96	(60,100,130)°	0.763	0.632	0.082	0.101
MCG+03-04	2.3	114	(60,150,90)°	0.963	0.900	0.049	0.044
MCG+03-38	2.1	222	(60,140,145)°	0.903	0.813	0.075	0.060
MCG+04-28	12.7	144	(80,120,10)°	0.877	0.763	0.107	0.127
MCG+05-32	3.0	142	(56,143,3)°	0.841	0.639	0.056	0.062
MCG+05-33	2.6	112	(80,90,165)°	0.889	0.793	0.067	0.069
MCG+09-13	2.8	66	(60,40,75)°	0.822	0.611	0.083	0.088
MCG+09-20	5.3	76	(40,150,100)°	0.789	0.649	0.088	0.100
NGC708	2.7	162	(70,20,130)°	0.885	0.695	0.030	0.031
NGC910	3.2	142	(60,70,145)°	0.877	0.747	0.078	0.103
NGC1128	2.7	216	(80,130,145)°	0.941	0.884	0.015	0.021
NGC1129	4.6	494	(60,10,0)°	0.888	0.780	0.047	0.044
NGC1275	3.3	80	(50,140,80)°	0.780	0.599	0.087	0.083
NGC2329	1.8	140	(60,50,0)°	0.930	0.848	0.034	0.031
NGC2804	1.8	32	(61,132,93)°	0.897	0.772	0.068	0.046
NGC3550	6.8	422	(60,20,20)°	0.973	0.937	0.028	0.050
NCG3551	0.88	114	(48,48,151)°	0.856	0.687	0.090	0.093
NGC4104	1.5	74	(70,130,135)°	0.587	0.290	0.083	0.082
NGC4874	1.5	302	(60,30,150)°	0.925	0.818	0.025	0.025
NGC6166	2.5	80	(90,47,150)°	0.824	0.588	0.082	0.063
NGC6173	1.4	56	(60,120,30)°	0.724	0.427	0.088	0.096
NGC6338	2.1	66	(60,60,165)°	0.816	0.630	0.055	0.045
NGC7647	3.7	60	(56,53,18)°	0.773	0.623	0.033	0.032
NGC7649	3.0	96	(60,130,95)°	0.784	0.507	0.081	0.079
NGC7720	1.8	130	(70,60,150)°	0.753	0.502	0.076	0.050
NGC7768	2.9	114	(60,40,150)°	0.732	0.515	0.098	0.080
SDSSJ0837	2.6	364	(64,136,36)°	0.798	0.543	0.101	0.075
UGC716	3.2	130	(40,100,0)°	0.817	0.582	0.078	0.081
UGC727	2.6	102	(60,30,170)°	0.817	0.638	0.067	0.079
UGC1191	2.5	238	(60,60,40)°	0.850	0.691	0.061	0.049
UGC2232	2.2	198	(60,138,62)°	0.827	0.717	0.032	0.031
UGC2413	2.2	72	(60,50,130)°	0.788	0.538	0.078	0.058
UGC4289	2.3	94	(60,150,60)°	0.873	0.770	0.055	0.050
UGC6394	2.9	50	(54,127,102)°	0.855	0.689	0.091	0.078
UGC9799	3.0	66	(60,50,145)°	0.816	0.667	0.067	0.060
UGC10143	3.3	136	(70,150,165)°	0.782	0.569	0.073	0.061
UGC10726	1.7	134	(50,50,80)°	0.925	0.867	0.026	0.022
VV16IC	3.2	142	(70,50,0)°	0.738	0.514	0.066	0.046

Table 3.3: *Col. 1*: Galaxy name. *Col. 2*: Smallest RMS. *Col. 3*: Number of "good" deprojections. *Col. 4*: The best-fit viewing angles. *Cols 5-6*: Average p and q values. *Cols 7-8*: RMS on average p and q values.

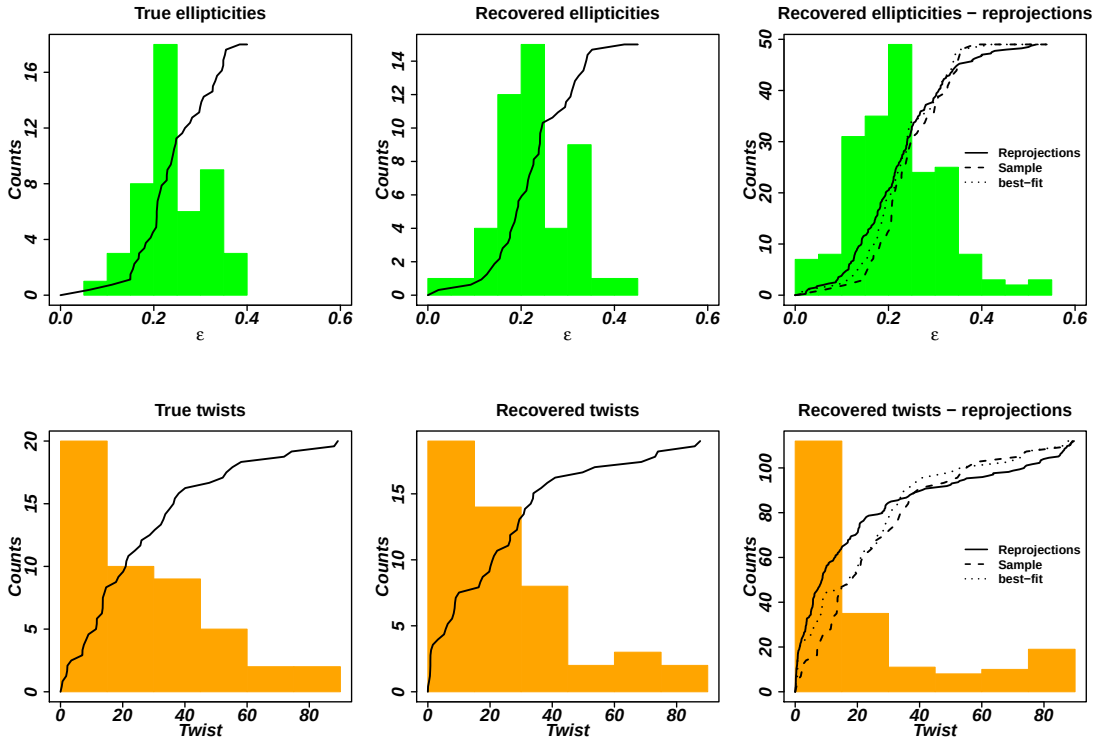


Figure 3.2: For both ε (top row) and PA (bottom row) we present three plots. The left panels show the observed mean values, the central panels the recovered values for the best-fit angles and the right panels the values obtained by re-projecting the best-fit solutions at random viewing angles. The solid lines are the cumulative distribution functions (cdfs), rescaled to the maximum count values. On the right panels we also show the cdf from the left and middle panels by dashed and dotted lines, respectively.

but this is only given by the fact that we are plotting *only the best-fit solution for each galaxy*. This clearly disfavours such viewing angles, because isophotal twists cannot occur along the principal axes of an ellipsoidal body. Hence, fits along an assumed LOS that coincides with one of the principal axes will deliver larger values of the RMS. In the bottom panel of Fig. 3.3 we show the entire distribution of deprojections over the two octants for the galaxy NGC708. Solutions on the principal axes are not excluded but lead on average to less good fits. Other examples are provided in the notes in App. 3.D.

3.4.2 Distribution of intrinsic axis ratios

We now present the *measured* shapes of the BCGs. Our deprojection algorithm *directly* yields the intrinsic axis ratios $p(r)$ and $q(r)$ as a function of the distance from the galaxy center. From these profiles, we compute the triaxiality parameter as a function of the radius $T(r)$. The profiles which we obtain by averaging over all good deprojections are shown in Fig. 3.C.1.

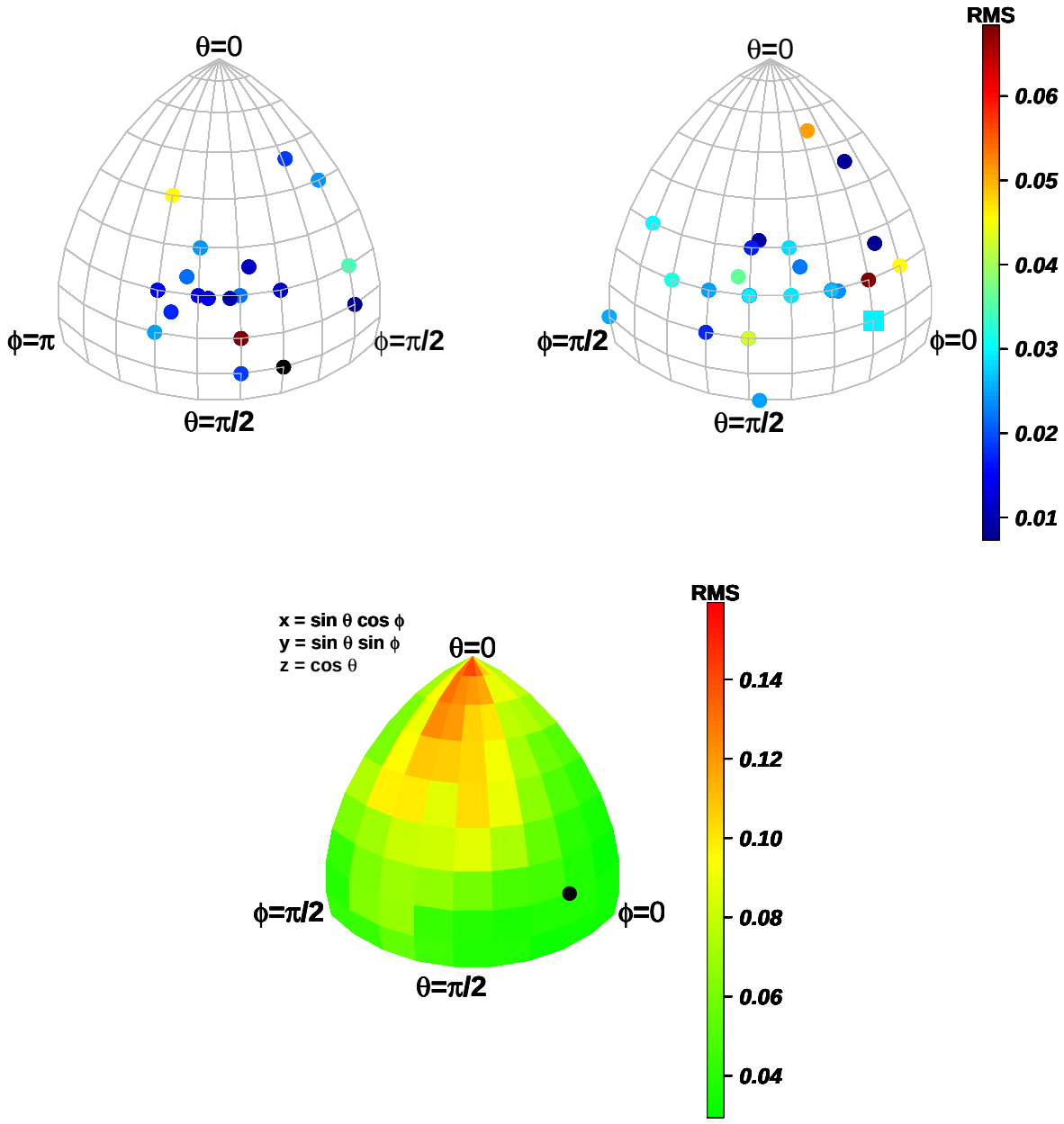


Figure 3.3: *Top:* Distribution of the best-fit angles on the two octants. The square at $(\theta, \phi) = (70, 20)^\circ$ labels the galaxy NGC708 (see below). The following galaxies appear superimposed: UGC 9799, UGC 2413, MCG-02-02 and NGC 2329 with $(\theta, \phi) = (60, 50)^\circ$; the galaxies MCG+09-13, IC1634, IC 5338 and IC 664 with $(\theta, \phi) = (60, 40)^\circ$; the galaxies 2MASXJ1358 and IC1695 with $(\theta, \phi) = (50, 40)^\circ$; the galaxies LEDA2098, NGC4874 and UGC727 with $(\theta, \phi) = (60, 30)^\circ$; the galaxies LEDA1518 and NGC6338 with $(\theta, \phi) = (60, 60)^\circ$; the galaxies UGC4289 and MCG+03-04 with $(\theta, \phi) = (60, 150)^\circ$; the galaxies MCG+02-04 and NGC 4104 with $(\theta, \phi) = (70, 130)^\circ$. *Bottom:* RMS distribution as a function of (θ, ϕ) for the galaxy NGC708, clearly showing the goodness of solutions close to the principal axes. The black point labels the best-fit solution, located at $(\theta, \phi) = (70, 20)^\circ$.

The left and central panels of Fig. 3.4 show the histograms of the average over all radial bins and over all acceptable deprojections $p(r)$ and $q(r)$ for every galaxy of the sample, i.e. averaged on every "good" deprojection. We get $< p(r) >= 0.84$ and $< q(r) >= 0.68$, with scatters of ∼0.1 (see Tab. 3.3 for the values we get for each BCG). For comparison, Ene et al. (2018) used ellipticity distribution and found $< p(r) >= 0.88$ and $< q(r) >= 0.65$ for a sample of slow rotators. The histogram of the mean triaxiality parameter, presented in the right panel of Fig. 3.4, shows that although BCGs follow different scaling relations from ordinary ETGs, they have $0.39 \leq < T > \leq 0.72$, in agreement with the findings of Vincent & Ryden (2005) for a sample consisting only of ordinary ETGs. The conclusion here is that the triaxiality is extremely high for every object of the sample, with no object showing a mean triaxiality outside of the $[0.39 - 0.72]$ interval.
We do not detect correlations between $< p(r) >$, $< q(r) >$ or $< T(r) >$ and the size. A weak correlation is seen with absolute magnitudes: bright BCGs appear rounder than fainter ones, having approximately the same triaxiality. The trend is more clearly seen when considering the radial profiles (see Figs. 3.5 & 3.6).

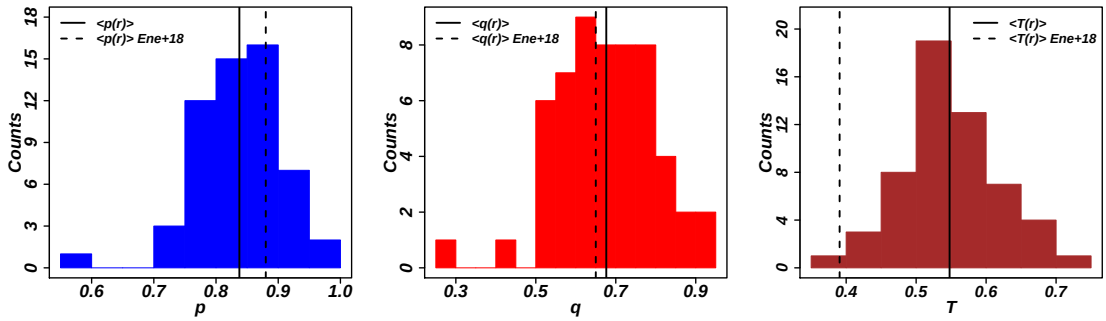


Figure 3.4: Histograms of the mean intrinsic axis ratios $< p(r) >$ (left panel) and $< q(r) >$ (middle panel) and of the corresponding triaxiality parameter $< T(r) >$ (right panel) for every galaxy of our sample. We also compare the mean values from our histograms with the findings of Ene et al. (2018) from the MASSIVE survey for slow rotators.

3.4.3 Comparison with the TNG and Magneticum simulations

In order to compare the recovered shape profiles to the shape profiles of simulated galaxies, we use the IllustrisTNG and Magneticum simulations. We consider the 110.7^3 Mpc^3 and 68^3 Mpc^3 cosmological volumes, respectively, as a good compromise between resolution and number of massive galaxies simulated. In TNG100, the mean mass of the stellar particles is 1.4×10^6 M$_\odot$ while the dark matter particles have masses 7.5×10^6 M$_\odot$. The Plummer equivalent gravitational softening length for both stars and dark matter at redshift $z = 0$ is $r_{soft} = 0.74$ kpc. Instead, Magneticum has stellar particles with masses of 2.6×10^6 M$_\odot$ and DM particles with masses of 5.1×10^7 M$_\odot$, while $r_{soft} = 2$ kpc for DM and 1 kpc for stars, respectively.
We select simulated galaxies with total mass larger than $10^{13} M_\odot$ that are the most massive members of their group (so-called 'central'). We divide these galaxies into 2 mass bins, with the

$\Delta \log (M_{tot}/M_\odot)$	TNG	Magneticum
13-13.35	116	31
$\geq$ 13.35	86	20

Table 3.4: The number of galaxies for every total mass bin from the TNG100 (second column) and Magneticum (third column) simulations.

number of objects in each mass bin summarized in Tab 3.4. From these, we derive $p(r)$ and $q(r)$ profiles for the dark matter and the stellar component separately. This is done by diagonalizing the inertia tensor

$$I_{ij} = \frac{\sum_n m_n x_{n,i} x_{n,j}}{\sum_n m_n}, \tag{3.2}$$

where $x_{n,i}$ are the coordinates of the stellar particles and m_n their mass, calculated in ellipsoidal shells (Zemp et al., 2011). We choose 10 radial bins logarithmically spaced along the intrinsic major axis of the galaxies from 3 to 100 kpc. In each step, the iterative procedure adjusts the flattening of the ellipsoidal shell and the direction of principal axes to the iso-density contours, until it converges within 1% in both p and q. We verified that the variation in the direction of the principal axes are generally within 5 degrees between 3 and 100 kpc and that fixing their position to a mean direction (for example, measured within 1 effective radius) slightly overestimates the axis ratios by a few percent, up to a median $\sim$ 3% in p and 7% in q at 100 kpc. This allows a comparison with the shape profiles derived for our BCGs with our deprojection code, which keeps the direction of principal axes in the 3D deprojected model fixed.
For each radial bin we compare the average profiles with those derived for our BCGs with our deprojection code, doing the same for the triaxiality parameter $T(r)$. We split the BCG sample into a bright one with $M_{tot} < -23.7$ and a faint one with $M_{tot} > -23.7$, each of which with 21 galaxies. Assuming a M/L ratio of 6, this corresponds to a stellar mass of $2.2\times10^{12} M_\odot$. Similarly, we split the simulated galaxies in two samples considering a total mass cut of $2.2 \times 10^{13} M_\odot$.
In Figs. 3.5 and 3.6 we show the comparison between the average profiles for stars and DM respectively, plotting the BCGs using lines, the simulated TNG100 galaxies using squares and the simulated Magneticum galaxies using triangles. For the BCGs we also show the RMS in each radial bin as error bar, with a typical value of ~0.1. This implies a typical error on the mean value of ~0.02. The RMS for the simulated galaxies is of the same order. The left panels show the faint BCG sample together with the less massive simulated galaxies; the right panels show the bright BCG sample together with the more massive simulated galaxies (see Tab. 3.4).
Bright BCGs appear slightly rounder than faint BCGs by $\Delta p \sim 0.04$ and $\Delta q \sim 0.08$, but with the same triaxiality. This trend is not obvious when looking at simulated galaxies.
The comparison of the profiles of BCGs and simulated galaxies shows that there is a strong disagreement in the inner regions, especially when the simulated stellar component is considered. The disagreement is less pronounced for the simulated DM halos. In particular, p and q of TNG100 galaxies have values down to 0.2-0.3, which would imply the presence of squashed

structures in almost all galaxies. The Magneticum galaxies are generally rounder, but still flatter than the observed ones. This shortcoming of simulations in reproducing the correct distribution of the ellipticity of massive (slowly rotating) systems is well documented: different sets of simulations predict a population of flat slow rotators with ellipticities as high as 0.55-0.6 (Li et al. 2018a using Illustris[6], Schulze et al. 2018 using Magneticum, Naab & Burkert 2003 using collision-less N-body simulations). In contrast, the observed ellipticity profiles on the sky (see top panels of Fig. 3.B.1) demonstrate that most BCGs are round near the center, and it is statistically impossible that all of them are axisymmetric systems viewed close to face-on or pole-on. Turning to the outermost regions, we find a much better agreement between the profiles of BCGs and of simulated galaxies, for both stars and DM. In particular, the BCGs and DM $p(r)$ and $q(r)$ profiles from TNG (and Magneticum at the high-mass end) follow a similar decreasing trend with a slight offset.

The average profile $T(r)$ of our BCGs is almost flat with r at a value of ≈ 0.55 with an RMS scatter of about 0.08, showing that these objects are overall triaxial. The TNG100 simulation generates objects which tend to be prolate in the center and as triaxial as our BCGs in the outer parts. The Magneticum simulation almost matches the observed average profile at intermediate masses when looking at the dark component, but produces more oblate/prolate profiles in the lower/larger mass bins.

Assuming that the flattening of the stellar component in the simulations compared to the observations is due to implementation in the hydrodynamics scheme and that the dark matter is unaffected by this, the similarity between the observed and simulated DM properties and the fact that the observations show similar triaxiality in the outskirts make it plausible that the light distribution of the outer regions of BCGs is tracing the underlying DM halos and may allow to even probe the nature of dark matter. In particular, recent N-body simulations (Robertson et al. 2019; Fischer et al. 2022) that study mergers of galaxy clusters show that the shape of dark matter subhaloes depends on their physical properties: self-interacting dark matter produces rounder halos than classical λCDM.

3.5 Conclusions

We have deprojected the photometry of a representative sample of 56 BCGs covering a large radial range with good resolution, from the innermost to the outermost regions probing into the intracluster light. The deprojection algorithm is able to generate SB profiles which are representative of the observed photometry. Moreover, the results show that the BCGs are consistent with random orientations in space. For the first time, we have measured radial profiles $p(r), q(r)$ and $T(r)$. The recovered shapes point to strongly triaxial galaxies, rounder at the centre and flatter at large radii. A comparison with the results of the TNG100 and Magneticum simulations shows that BCGs at large radii are a tracer of the DM halo they are embedded in, possibly probing the nature of dark matter. Extending this analysis to galaxies at higher redshifts can probe the

[6]For these simulations, values as high as 0.8 are found (Pulsoni et al., 2020).

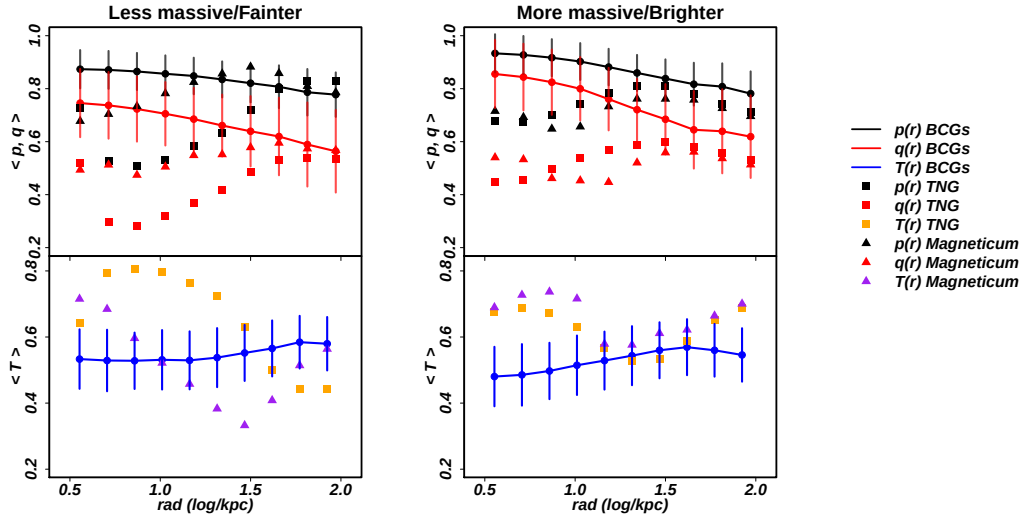


Figure 3.5: Comparison between the $p(r), q(r)$ average profiles (top panels) and $T(r)$ (bottom panels) of our BCGs and of the stellar component of TNG100 and Magneticum simulations objects. BCGs are rendered using lines, while squares and triangles are used for TNG- and Magneticum-simulated galaxies, respectively. For the BCGs we also compute the RMS in each radial bin, showing it as error bar, with a typical value of $\sim 0.08 - 0.1$.

formation history of such objects, although getting SB profiles with high enough signal-to-noise ratio at large values of z certainly represents a challenge for present-day facilities.
The extremely strong triaxiality of these objects stresses the need for triaxial dynamical modeling of the stellar kinematics (e.g. Neureiter et al. 2020) in order to recover unbiased BH mass and M/L estimates, reconstruct the anisotropy profiles of these galaxies and evaluate the effects of the different number of DOF. We will address these issues in two forthcoming papers (de Nicola et. al, submitted; Neureiter et. al, submitted).

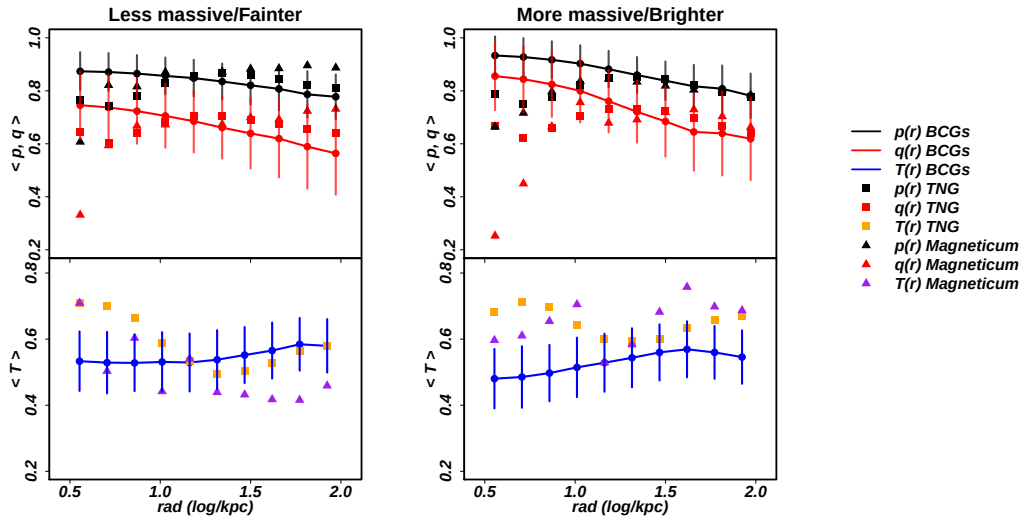


Figure 3.6: Similar to Fig. 3.5, but showing the $p(r), q(r)$ and $T(r)$ average profiles of the dark halo component of TNG100 and Magneticum simulations objects.

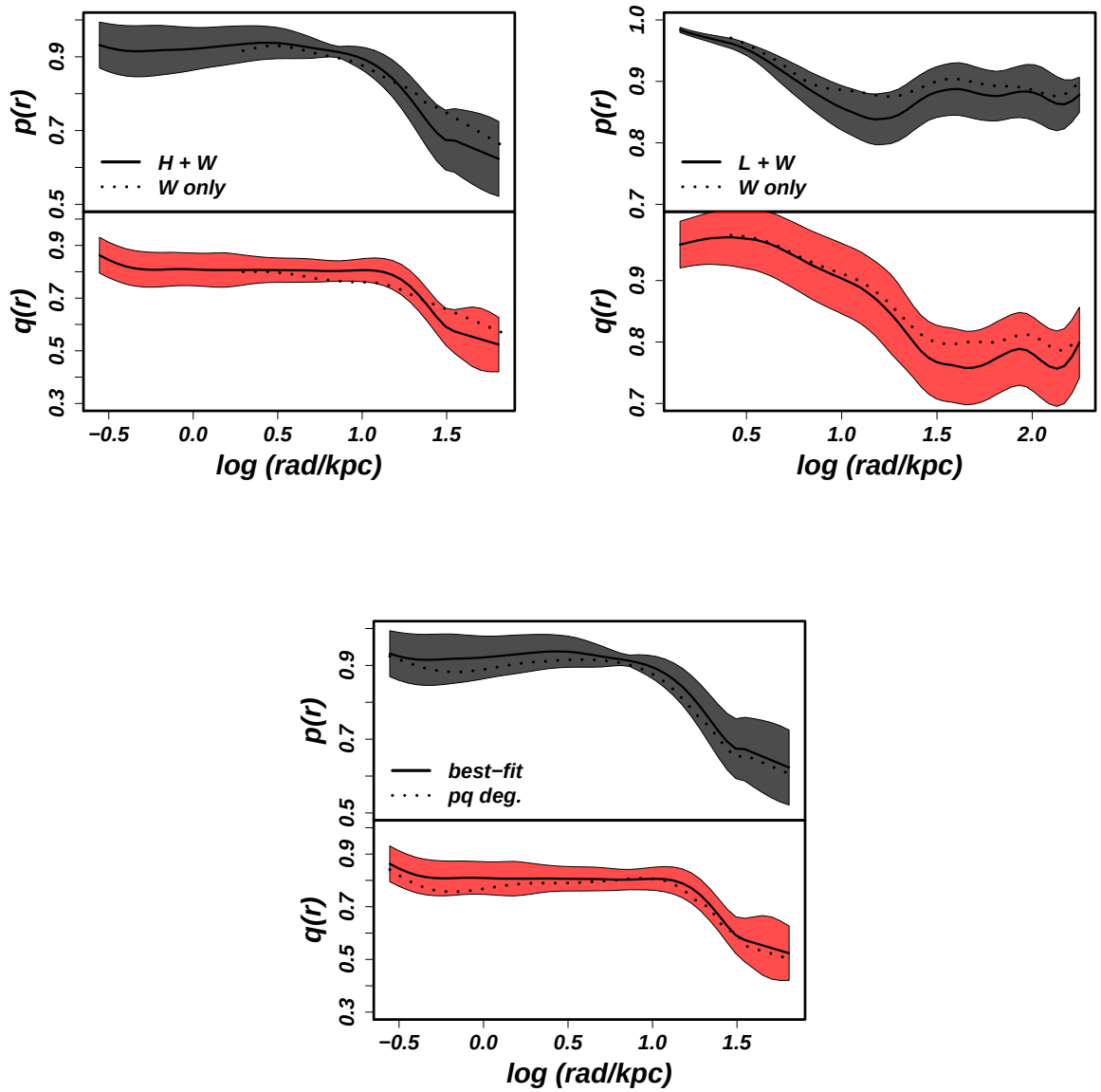


Figure 3.7: *Top panels:* Comparison of the deprojections obtained at the best-fit viewing angles with or without HST (left, galaxy NGC7647) and LBT photometry (right, galaxy UGC10726). The solid lines are the best-fit profiles with both high-resolution and Wendelstein photometry, while the dotted lines show the solutions with Wendelstein photometry only. The coloured regions are given by the RMS values. *Bottom panel:* For the galaxy NGC7647, we perform a deprojection stopping it when the RMS reaches $1.5 \times \mathrm{RMS}_{\mathrm{min}}$ to judge the effects of possible degeneracies between the model and the viewing angles. The resulting profiles are shown as dotted lines, solid lines and coloured regions are as above.

Appendix

3.A Resolution and degeneracy effects

We analyze here the systematic effects stemming from the lack of high-resolution HST or LBT data with Wendelstein observations as well as from the residual degeneracy between the p & q profiles and the viewing angles.
The first point can easily be investigated by deprojecting galaxies with and without high-resolution data, verifying how much the deprojection differs between the two cases. We choose two galaxies, NGC7647 (for the HST case) and UGC10726 (for the LBT case), and re-perform the deprojection using Wendelstein data only for the best-fit viewing angles. These two galaxies represent stringent tests given the very low scatter for both p and q (see Tab. 3.3) among different solutions at different viewing angles and the code yielding a low best-fit RMS value. Moreover, the central regions of these galaxies are relaxed, meaning that we can exploit HST and LBT data up to the innermost radii.
In the two top panels of Fig. 3.7 we show the $p(r)$, $q(r)$ profiles from the HST(LBT)+Wendelstein case as solid lines, while the Wendelstein-only profiles are shown as dotted lines for p and q. The Wendelstein-only deprojection cannot probe the inner region of the galaxy, but remains within the region delimited by the RMS (shown as coloured area in the figures) at larger radii.
As a second test, we take the galaxy NGC7647 and deproject it at the best-fit viewing angles (with HST photometry) stopping the deprojection as soon as the RMS reaches $1.5 \times \mathrm{RMS}_{\mathrm{min}}$. In the bottom panel of Fig. 3.7 we compare this solution to the best-fit one. Also in this case the deviations from the best-fit solution are smaller than the scatter due to the different viewing angles for which an acceptable deprojection is found.
Thus, we conclude the lack of the high-resolution photometry does not change the conclusions reported in this paper. Moreover, considering only the best-fit solution (in terms of the RMS) for a given set of viewing angles probes the range of acceptable p & q profiles.

3.B Ellipticity and PA profiles

In Fig. 3.B.1 we show the ε and the PA profiles for the BCGs of our sample. Omitted points (see also notes in App. 3.D) are not shown.

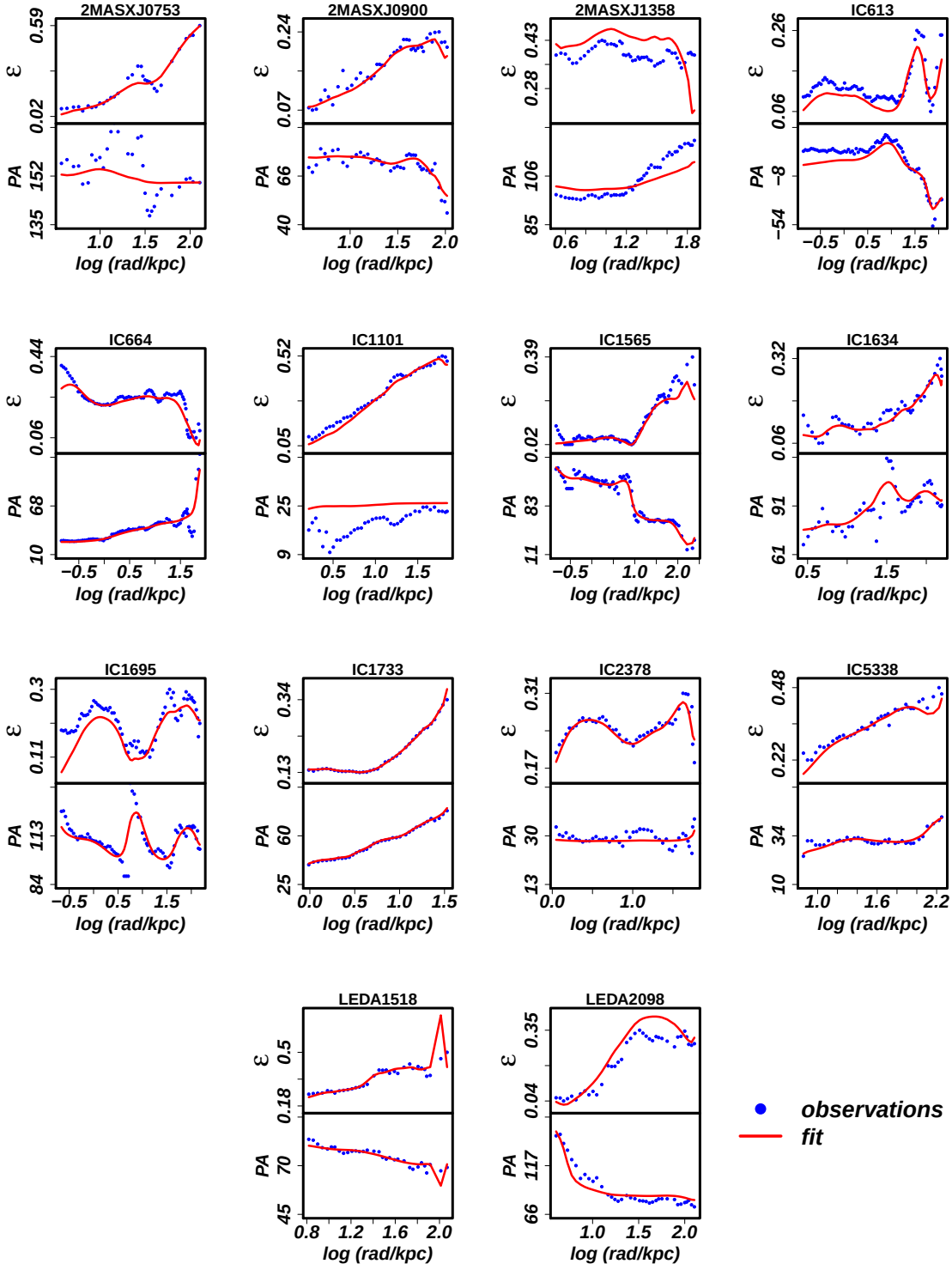


Figure 3.B.1: ε (top panels) and PA (bottom panels) profiles of every BCG considered in this work. Blue points are the original photometry, whereas the red lines show our fits. The radii are given in $\log_{10}$ kpc.

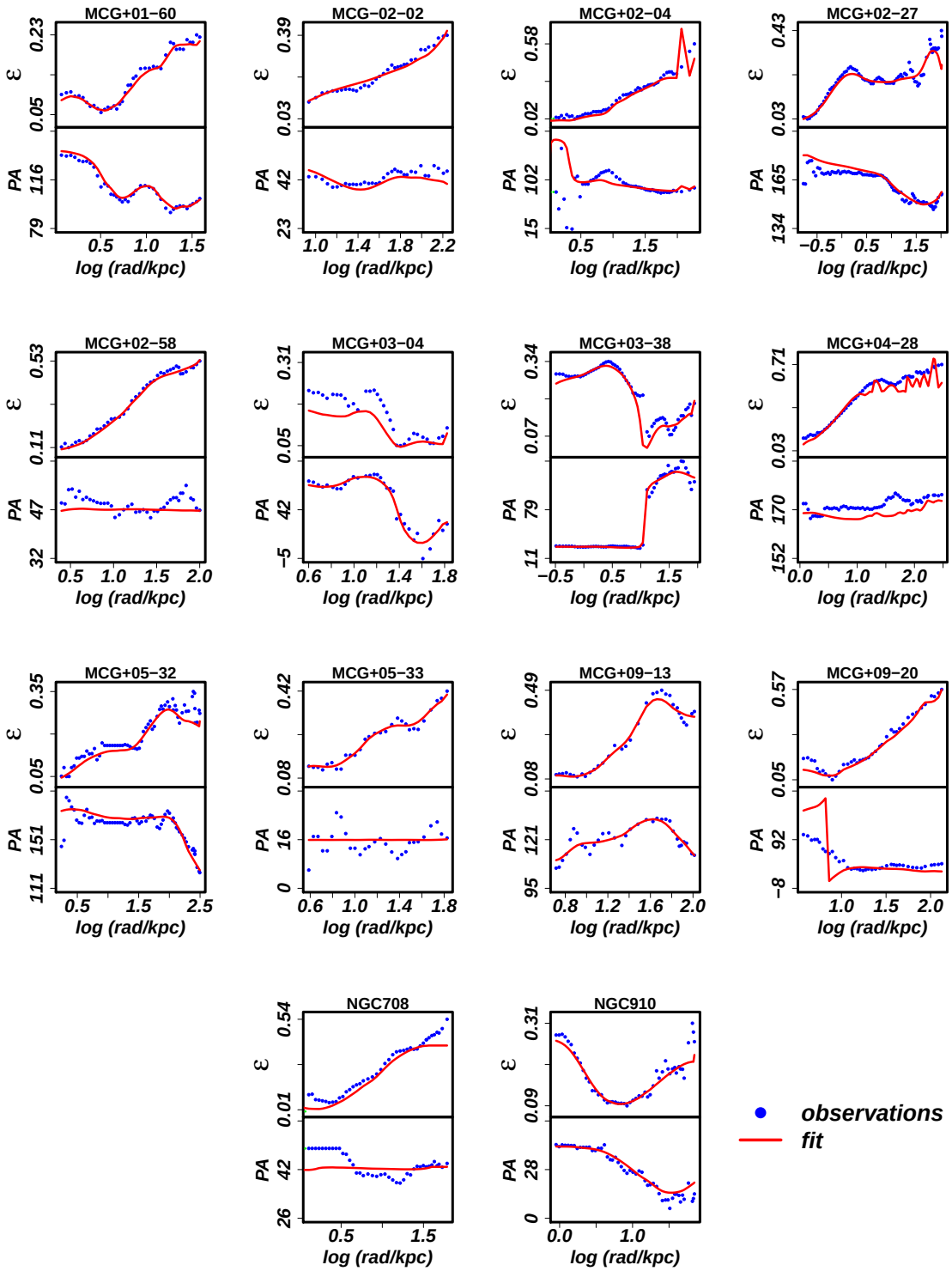


Figure 3.B.1 *(continued)*

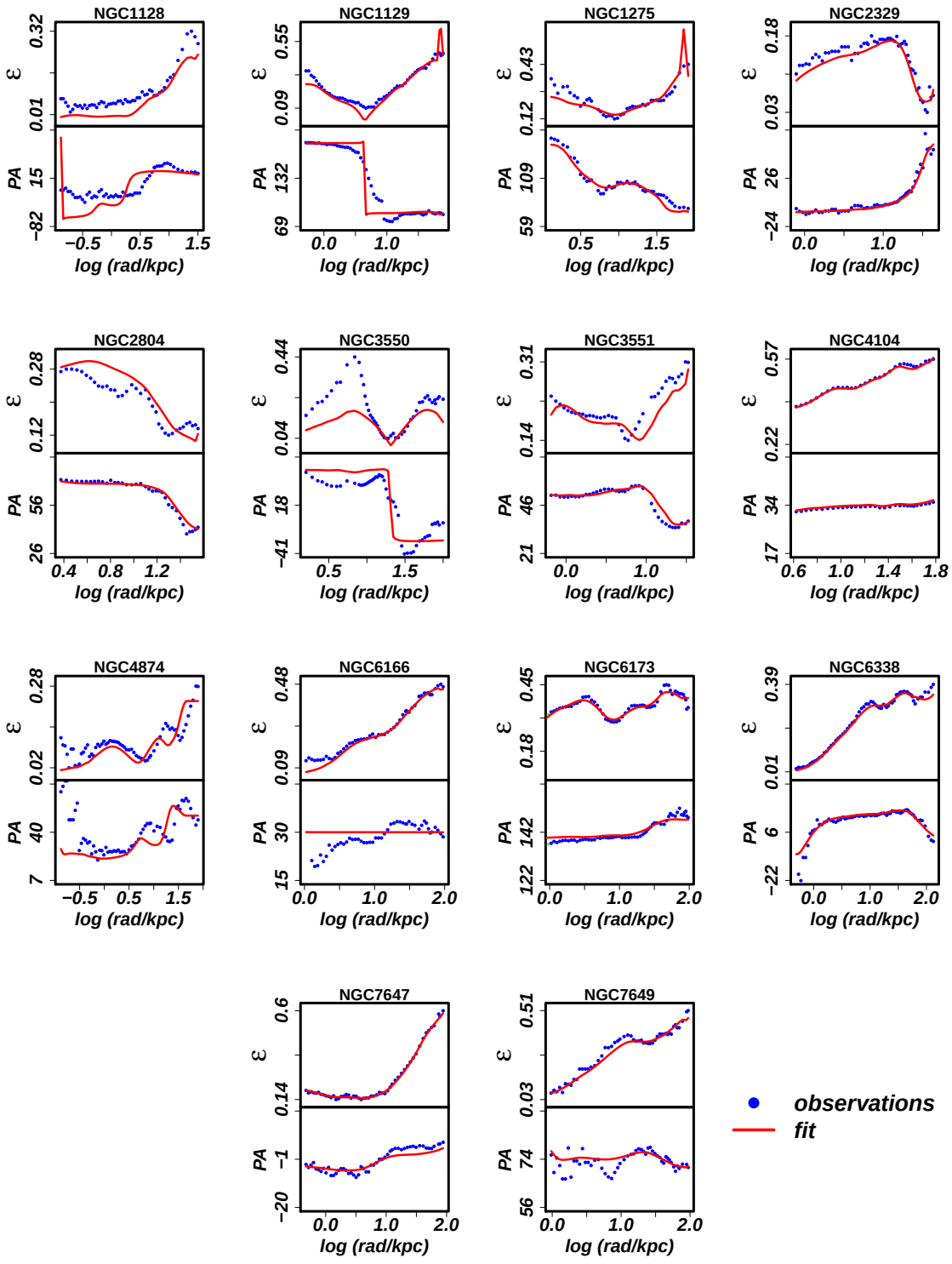


Figure 3.B.1 *(continued)*

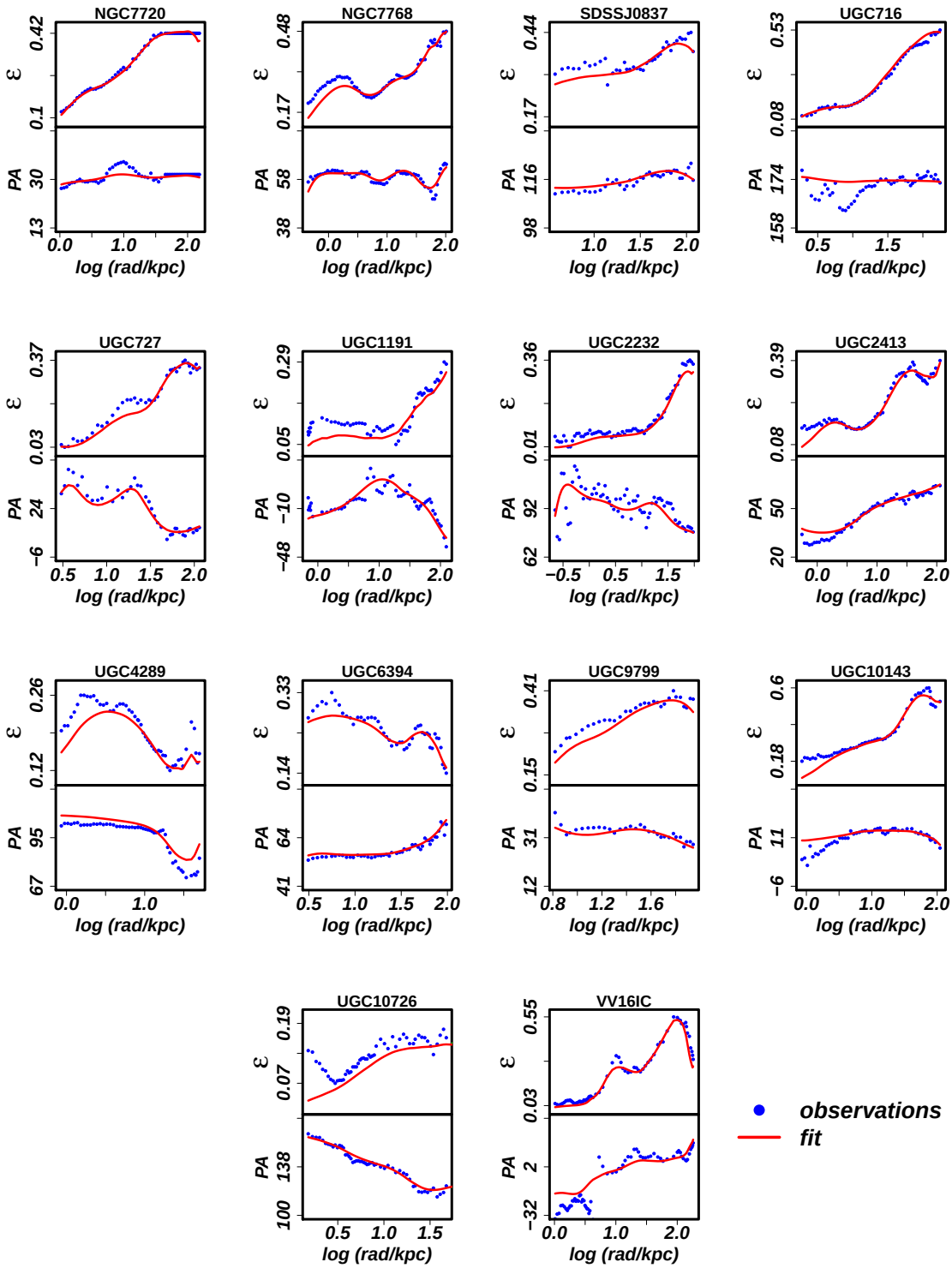


Figure 3.B.1 *(continued)*

3.C Intrinsic shape profiles

In Fig. 3.C.1 we show intrinsic axis ratio profiles, along with the corresponding triaxiality profiles, for every BCG of the sample. The profiles are computed by averaging over all acceptable deprojections.

3.D Notes on individual galaxies

- **2MASXJ0753:** This is one of the galaxies observed at LBT. The PSF effects are clearly visible way beyond the 0.4" seeing value estimated during the observations, and therefore all the affected points are not taken into account. The galaxy shows bumpy/noisy $\varepsilon$and PA profiles which cannot be described accurately, despite the fact that the viewing angles are not close to the principal axes.

- **2MASXJ0900:** This well-fitted galaxy has a sudden 25° twist in the outermost regions, which is well reproduced. This may be due to the ICL given that in the innermost the regions there is no significant twist.

- **2MASXJ1358:** This is an example of a galaxy whose εdoes not change much as a function of the radius. Our best-fit slightly overestimates it, while the twist is underestimated. We stop the deprojection at 75 kpc since the isophotal parameters cannot be adequately fitted anymore beyond this radius.

- **IC613:** The galaxy is very round ($\varepsilon \lesssim 0.1$ until beyond 10 kpc). The outermost radii hint at something not in equilibrium, which the code does fit well. The PA at the centre is not well reproduced, but since $\varepsilon$is small in the central region, this is not a serious issue.

- **IC664:** This is a rare example of a galaxy which is flat in the center and round in the outskirts. This low εgenerates an unrealistic twist at large radii, but the galaxy still shows a nice constant PA at lower radii which is well reproduced by the code.

- **IC1101:** The best-fit angles (θ, ϕ, ψ)= (50,80,160)° almost lie on the (y, z) plane, which does not allow for a precise recovery of the twist. Possibly these angles are not the correct ones, because one of the random re-projections does produce a better fit to the observed twist. The somewhat bumpy $\varepsilon$and PA profiles point to the presence of not fully relaxed structures.

- **IC1565:** The galaxy is well fitted with the exception of the outermost εpoints which likely belong to the intracluster light. The huge twist shows some oscillations, hinting at a not completely relaxed galaxy.

- **IC1634:** The galaxy has very noisy profiles, which hint at a not yet relaxed galaxy. However, the code reproduces the trends very well.

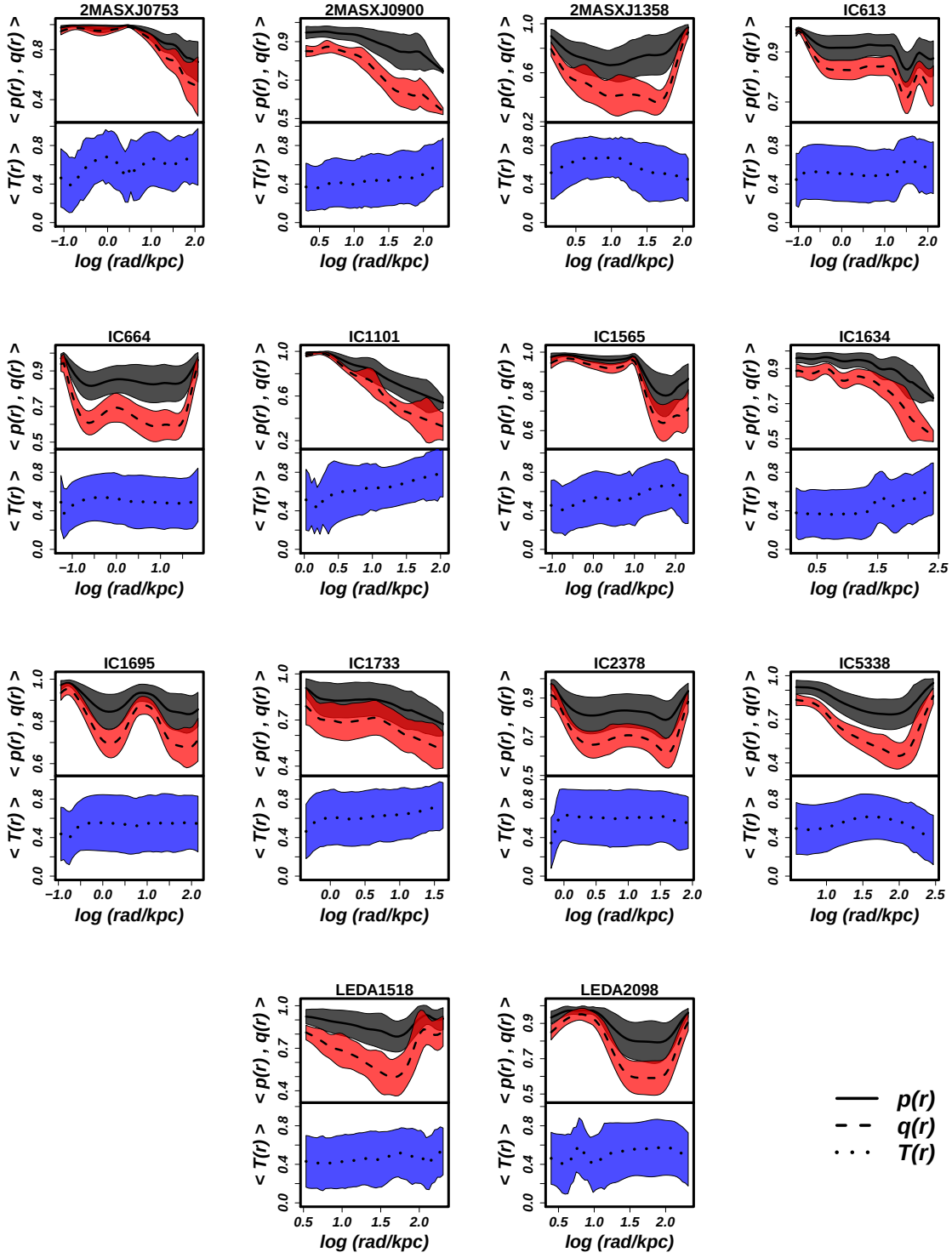


Figure 3.C.1: Intrinsic axis ratios p, q (top panels) and triaxiality T (bottom panels) profiles of every BCG considered in this work. The solid, dashed and dotted lines are the average among all good profiles (see Sec. 3.3.2), while the coloured regions mark the RMS values. The radii are given in $\log_{10}$ kpc.

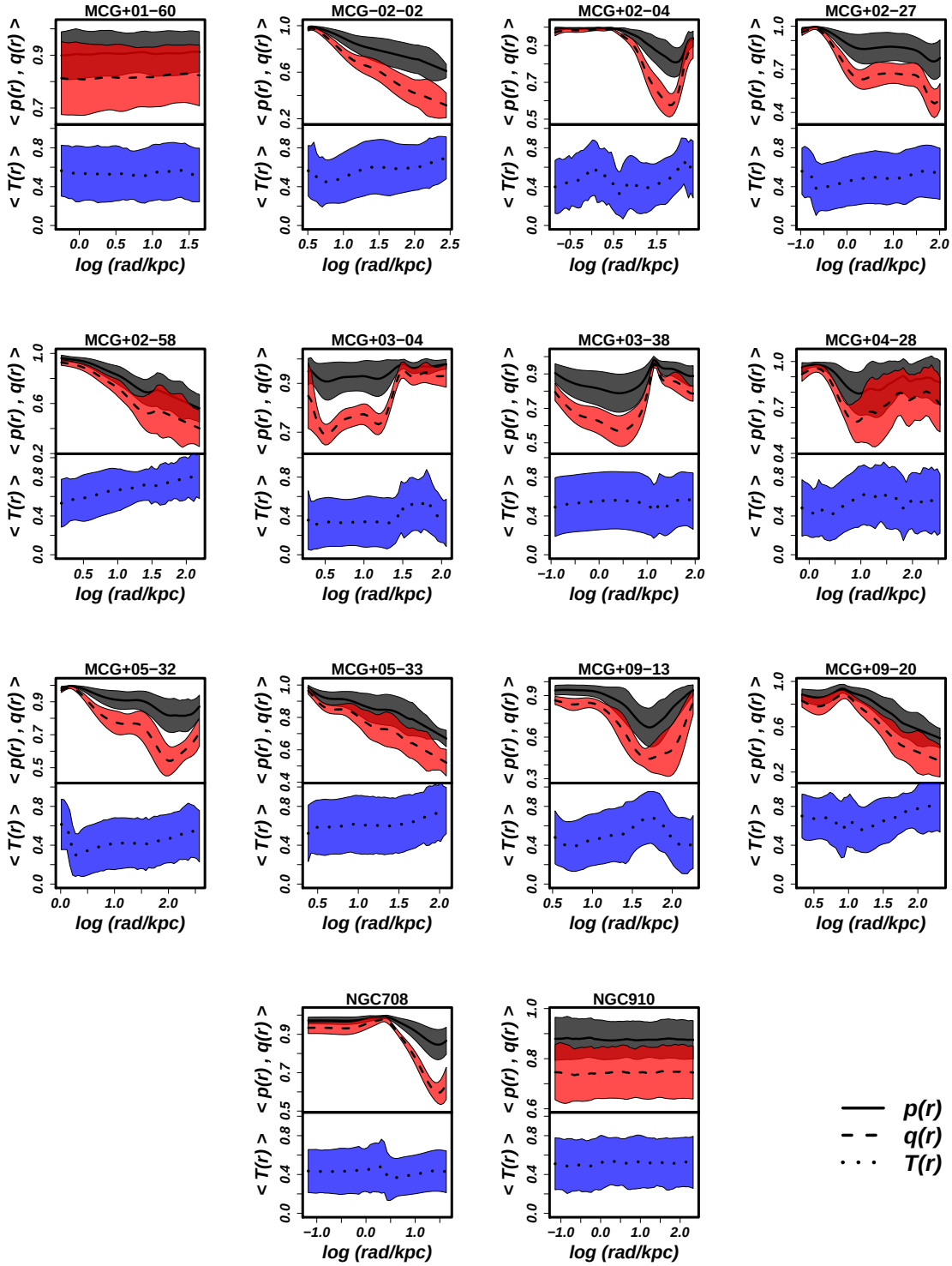


Figure 3.C.1 *(continued)*

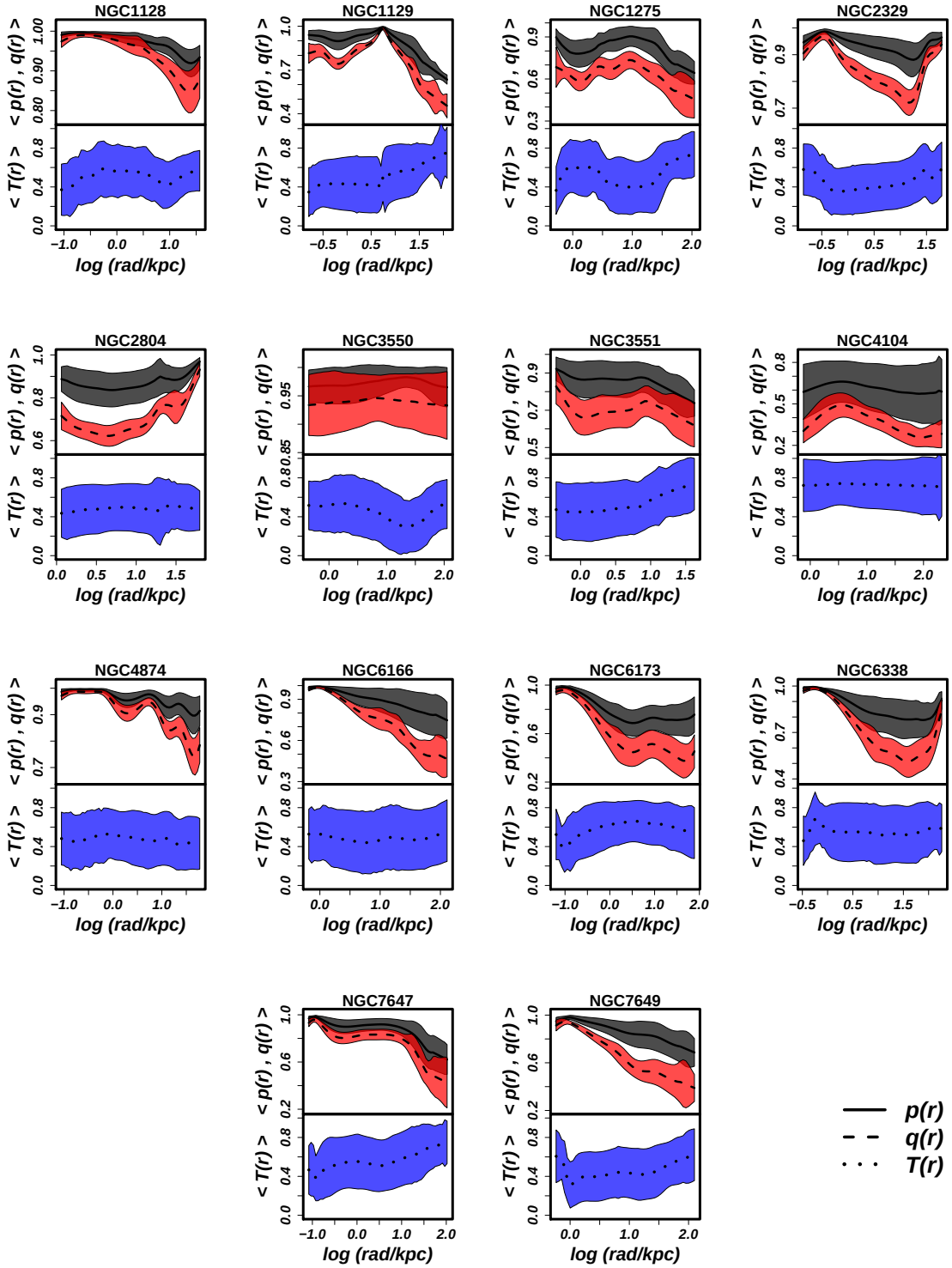


Figure 3.C.1 *(continued)*

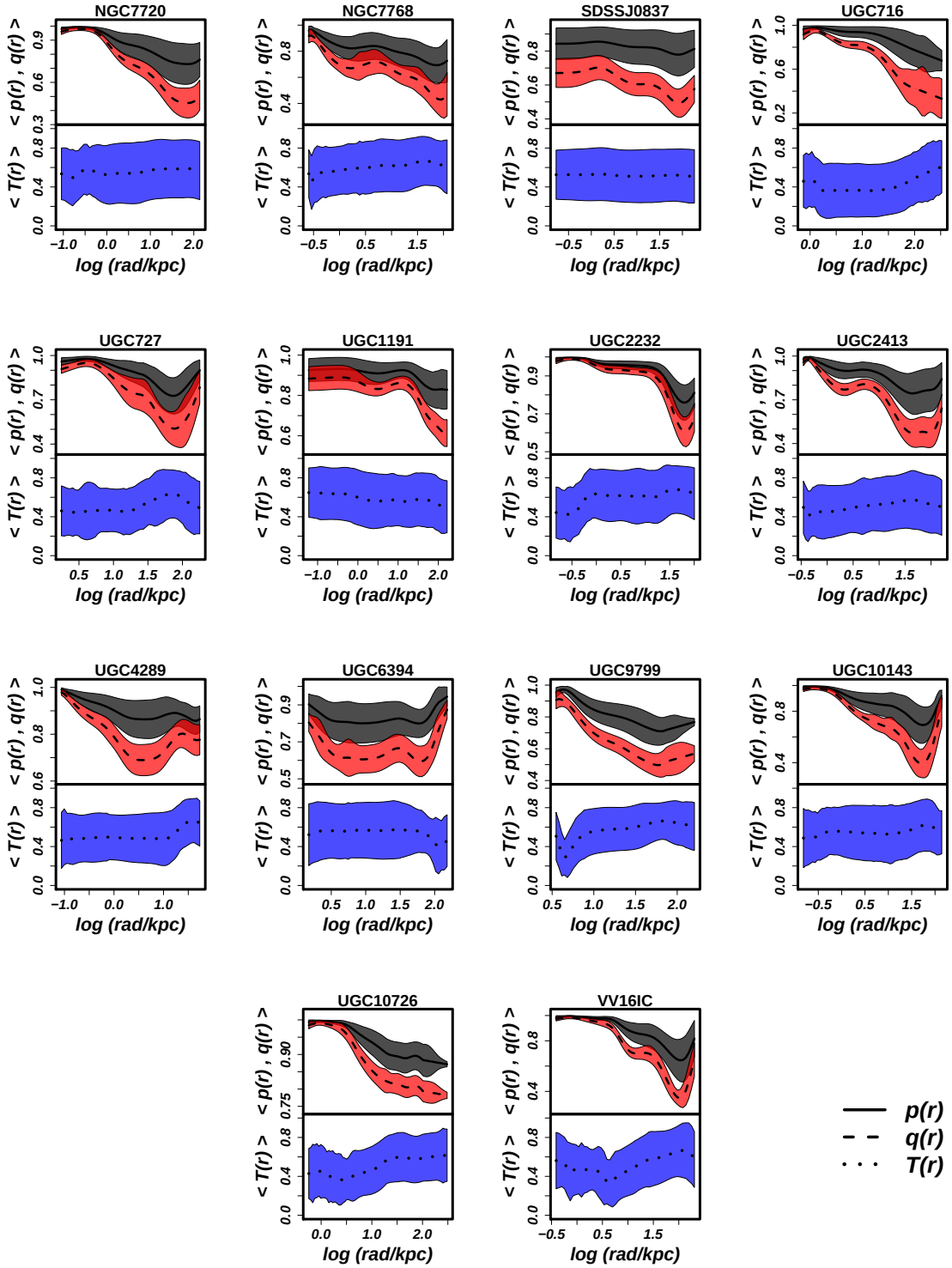


Figure 3.C.1 *(continued)*

- **IC1695:** Although both the $\varepsilon$and the PA profiles are complex, the code reproduces these profiles well, with the exception of the major bump in the PA, which is underestimated by a factor of 2.

- **IC1733:** We stop the deprojection at 35 kpc because of possible unrelaxed structures at larger radii. It is the best-fitted galaxy of the sample, with an RMS of 0.0073.

- **IC2378:** The best-fit angles (θ, ϕ, ψ)= (70,10,150)° are close to the x-axis. The somewhat noisy but almost constant PA profile allows a good fit with low RMS and indicates that the galaxy might be oriented along one of the principal axes.

- **IC5338:** We start the deprojection at 4.9 kpc because of possible AGN contamination. The last points do give a suspicious twist increase probably linked to the intracluster light; however, this is nicely reproduced.

- **LEDA1518:** The galaxy has a 20° twist and a smooth ellipticity profile, although we note that ε> 0.2 in the central regions already. The deprojection reproduces the profiles very well except for the outermost points.

- **LEDA2098:** The galaxy has a huge 80° twist, but this is mostly given by the very low ellipticity in the central regions. The code reproduces these profiles well, slightly overestimating εat the outer radii.

- **MCG+01-60:** We start the deprojection at 1.16 kpc, since for this galaxy only Wendelstein data are available. This is another galaxy with $\mathrm{RMS}_{\mathrm{best}} \leq 0.01$.

- **MCG-02-02:** This well relaxed galaxy has a typical ellipticity profile rising steadily and a small 10° twist. No solutions compatible with v.a. along the principal axes are found, but we get acceptable deprojections at $\theta = 10°$.

- **MCG+02-04:** The twist is well reproduced with the exception of a bump around 10 kpc. In the central regions we suspect the galaxy to be not fully relaxed, because of an unrealistic ∼150°twist at small radii. Therefore, we start the deprojection at 1.3 kpc.

- **MCG+02-27:** The 30° twist present in the central regions is not reproduced very well; however since $\varepsilon$is very low this does not significantly affect the goodness of the fit. The bumpy εprofile suggests that relaxation is not complete.

- **MCG+02-58, MCG+05-33:** The best-fit angles (θ, ϕ, ψ)= (80,0,165)° and (θ, ϕ, ψ)= (60,10,130)° almost lie along the x-axis and on the (x, z) plane, respectively. This does not allow for a good twist recovery; however, the true PA profiles oscillate around 15° (which would indeed give $\psi = 165°$ if the galaxy were along x) for MCG+05-33 and around 50° for MCG+02-58 (for which $\psi = 130°$ would be the right value). We measure twists oscillating around 20°and 9°, respectively. This indicates that the two galaxies are oriented along one of the principal axes, but possibly not fully relaxed yet.

- **MCG+03-04:** This interesting galaxy is flat in the central regions, where our code slightly underestimates the ellipticity, and gets rounder in the outskirts. The PA profile is tricky, since the twist is small in the central regions before getting significantly bigger at large radii, where $\varepsilon$is small. The fact that the twist is small in the central regions enables the code to obtain good fits close to the principal axes, as for IC2378.

- **MCG+03-38:** Another flat galaxy in the central regions. Both $\varepsilon$and PA jump wildly at ∼10-15 kpc, as if a decoupled structure were there in the centre. Nevertheless, the code returns a good fit to the observed photometry.

- **MCG+04-28:** The high RMS (0.122) indicates that the fit is not satisfactory beyond ∼10 kpc, where the galaxy becomes extremely flat. The somewhat bumpy twist is overall small, therefore the systematic offset between model and data is not worrying.

- **MCG+05-32:** The galaxy is well fitted. The photometry shows an unrealistic twist in the first ∼10 isophotes, which is probably the result of the low ellipticity.

- **MCG+09-13, MCG+09-20:** These are both galaxies showing typical isophotes of massive ellipticals, although with some bumps. The twist in the central regions for MCG+09-20 is due to εalmost going to 0.

- **NGC708:** Although the best-fit viewing angles are not exactly on the principal axes, there are several good solutions compatible with such inclinations, as shown in Fig. 3.3. The photometry has not been deprojected within the first 1.2 kpc because of a dust lane. The scale of the plot in Fig. 3.B.1 might give the wrong impression of a poorly recovered twist, which is not the case.

- **NGC910:** Like NGC1129, εgoes down and then up again. The 30° twist is well fitted.

- **NGC1128:** We do not include the galaxy in the twist histogram, as $\varepsilon$is almost always below 0.1, except for the outermost radii.

- **NGC1129:** The best-fit angles (θ, ϕ, ψ)= (60,10,0)° almost lie on the (x, z) plane. εgoes down to 0 and then increases again. Given that the twist is roughly 90°, this could be a galaxy compatible with viewing angles along the principal axes (as suggested by the best-fit viewing angles) *and* with intersecting p, q profiles.

- **NGC1275:** This peculiar galaxy shows a high ellipticity both in the innermost and in the outermost regions, with a dip in between. The 80° twist is very well recovered.

- **NGC2329:** We omit the innermost points because of an unrealistic bump in ε.

- **NGC2804:** εdecreases towards the outermost regions. In the first 10 kpc, the twist is completely absent. We stop the deprojection at 65 kpc because the isophotal parameters cannot be measured anymore beyond this radius.

- **NGC3550:** This galaxy was observed at LBT under poor seeing conditions. Moreover, both the $\varepsilon$and the PA profiles hint at a not fully relaxed galaxy.
- **NCG3551:** We omit the innermost three isophotes because of resolution problems when deriving the isophotal parameters (only Wendelstein images are available for this galaxy). The deprojection beyond 35 kpc also becomes unfeasible since the galaxy shows signs of non-equilibrium, however the deprojection yields RMS ≤ 0.01.
- **NGC4104:** The first 1.5" arcseconds must be discarded because of poor seeing. Also the outermost points (from 60 kpc) are omitted due to contamination from a neighbor galaxy. It is one of the flattest galaxies of the sample, with εalways between 0.4-0.6.
- **NGC4874:** Very round galaxy with noisy profiles. We include it in the histogram, although the only region where εstabilizes above 0.1 is beyond 10 kpc.
- **NGC6166:** The best-fit angles (θ, ϕ, ψ)= (40,0,60)° lie on the (x, z) plane. The low RMS might be explained by the fact that with the exception of the innermost radii (where most of the twist occurs but $\varepsilon$is low) the true PA oscillates around the constant PA recovered by the code.
- **NGC6173:** We dropped the poorly fitted central region of the galaxy.
- **NGC6338:** The deprojection starts at ∼0.6" because of possible AGN activity.
- **NGC7647:** Nothing to signal here.
- **NGC7649, NGC7720, NGC7768:** It is not entirely clear how relaxed the galaxies are. The PA profiles are somewhat noisy with very small twists, while the εprofiles increase smoothly with radius (with the exception of the central regions of NGC7768) with minor dips.
- **SDSSJ0837:** The same considerations about the observations made for 2MASXJ0837 also apply for this galaxy. However, this galaxy does not show signs of non-equilibrium.
- **UGC716:** The best-fit angles (θ, ϕ, ψ)= (40,0,60)° almost lie on the (y, z) plane. This is another galaxy which might indeed be close to the principal axes despite the small ∼10° twist near the round center.
- **UGC727:** See comments of MCG+09-13 and MCG+09-20.
- **UGC1191:** We start the deprojection at ∼0.6-0.7 kpc because of PSF effects. The galaxy has a large twist (∼40°) which is well reproduced.
- **UGC2232:** Nothing to signal here.
- **UGC2413:** We note a slight offset in the central regions between the true photometry and the recovered one, probably because of resolution effects given by the spherical ρ-grid.

- **UGC4289:** The galaxy shows a somewhat noisy εprofile along with a PA profile with an abrupt $\sim 30^\circ$ twist starting from $\sim$20 kpc. We stop the deprojection at 50 kpc because of possible contamination from neighbour galaxies.
- **UGC6394:** The ellipticity decreases as a function of radius and shows some bumps, while the PA profile is much smoother and very well recovered. For this galaxy we also obtain prolate deprojections compatible with the observed photometry.
- **UGC9799:** We start the deprojection at 4 kpc to avoid the center affected by probable AGN contamination.
- **UGC10143:** The same considerations made for UGC2413 also apply to this galaxy. We start the deprojection at 0.6 kpc because of a sudden 100° twist in the innermost regions.
- **UGC10726:** The same considerations made for UGC1191 also apply to this galaxy. We omit the outermost isophotes.
- **VV16IC:** See comments of UGC9799.

Chapter 4

Accuracy and precision of triaxial orbit models II: Viewing angles, shape and orbital structure

4.1 Introduction

[1] The recovery of the intrinsic shape of a galaxy as well as the reconstruction of the three-dimensional stellar dynamics rely on projected quantities that we see on the plane of the sky. By using 2D-images of the galaxy on the plane of the sky, it is possible to reconstruct the intrinsic 3D luminosity density (hereafter ρ) that projects to the observed image (or isophotes) for a certain galaxy inclination. The shape of this 3D-density, measured in terms of the axis ratios $p \equiv y/x$, $q \equiv z/x$ and of the triaxiality parameter $T = (1 - p^2)/(1 - q^2)$ (Franx et al., 1991), allows us to make inferences about the galaxy formation history. A particularly relevant example concerns the merging history that leads to the formation of elliptical galaxies: the most massive, triaxial, rounder galaxies, form through dry mergers, while fast-rotating, flatter galaxies, form through wet mergers (Bender 1988; Barnes & Hernquist 1992; Bender et al. 1992; Kormendy & Bender 1996; Kormendy et al. 2009; Bois et al. 2011; Khochfar et al. 2011; Naab et al. 2014, see Cappellari 2016 for a review). Moreover, the light density itself is used as a constraint to calculate the dynamics of stars found around the galaxy center such that the resulting Line-of-Sight Velocity Distribution (LOSVD) matches the observed one (Schwarzschild, 1979; Cretton et al., 1999, 2000; Gebhardt et al., 2000; Verolme et al., 2002; Valluri et al., 2004; Thomas et al., 2004; Valluri et al., 2005; Thomas et al., 2005; van den Bosch et al., 2008; Vasiliev & Valluri, 2020; Neureiter et al., 2021). With the stellar dynamics in hand, one can not only measure BH masses or mass-to-light (Υ) ratios, but also estimate the anisotropy profile $\beta(r)$. This is of particular relevance for core-elliptical galaxies, whose central cores are believed to be generated by BH-scouring and, hence, they should show a tangential bias in the innermost regions (Thomas et al., 2014; Rantala et al., 2018, 2019). While dynamical models indeed suggested tangential anisotropy in the centers of

[1]Published paper: de Nicola S., Neureiter, B., Thomas J., Saglia R.P., Bender R. - Accuracy and precision of triaxial orbit models II: Viewing angles, shape and orbital structure, MNRAS, 2022, 517, 3445

elliptical galaxies of different kinds – cored or non-cored (e.g. Gebhardt et al. 2003; Cappellari & McDermid 2005; Schulze & Gebhardt 2011; McConnell et al. 2012a; Thomas et al. 2014) – the structure of galaxies with depleted stellar cores is special: their anisotropy is remarkably uniform and changes from inner tangential anisotropy to outer radial anisotropy at the core radius as predicted (e.g. Thomas et al. 2014, 2016; Mehrgan et al. 2019).
Generally, one uses both photometrical and kinematical data to measure the galaxy shapes. Nevertheless, regardless of the approach one chooses to tackle down the deprojection/dynamical modeling, knowing the galaxy inclination is a key ingredient. Unfortunately, the inclination cannot be measured in general, and the issue gets particularly severe when dealing with massive ellipticals, which are typically disk-less galaxies. Moreover, observational evidences such as isophotal twists or misalignment between kinematic and photometric axes show that these galaxies are not axisymmetric, but rather triaxial (Vincent & Ryden, 2005; Ene et al., 2018), meaning that one needs to specify three viewing angles (θ, ϕ, ψ) instead of only one angle i needed in the axisymmetric case (Binney, 1985; de Zeeuw & Franx, 1989). Assuming the wrong viewing angles when deprojecting a galaxy will almost always yield the wrong shape (de Nicola et al., 2020), which will also likely lead to a wrong estimate of the anisotropy profile $\beta(r)$ of the galaxy and to biased estimates of the mass-to-light ratio Υ and the black hole mass $M_{\rm BH}$ when dynamically modeling the galaxy.
In principle, one could deproject a galaxy assuming a large number of possible viewing angles and then dynamically model all these three-dimensional luminosity densities to obtain an estimate of the galaxy viewing angles. In practice, this is not feasible because dynamical models are both very computing time- and memory-consuming. It is thus important to develop a deprojection tool which allows for a significant reduction of the number of possible viewing angles, generating physically plausible densities and keeping the degeneracy arising from the deprojection under control (Gerhard, 1996; Kochanek & Rybicki, 1996; Magorrian, 1999; de Nicola et al., 2020). One commonly used routine, the Multi-Gaussian-Expansion (MGE, Cappellari 2002) does not allow to explore different deprojections for a given set of viewing angles and could in principle not yield a single possible viewing angle if either the isophotes are very flattened or the twist is large. On the dynamical modeling side, available triaxial Schwarzschild codes (e.g. van den Bosch et al. 2008) deliver mass-to-light ratios and Dark Matter (DM) fractions of simulated galaxies deviating by 15-25% from the true values (see Jin et al. 2019).
Motivated by such arguments, we have developed two codes aimed at filling these gaps:

- In de Nicola et al. (2020) (hereafter dN20), we presented our novel triaxial semi-parametric deprojection code SHAPE3D, that finds the best- fit light density ρ projecting to a certain surface brightness under the assumption of being strati- fied onto "deformed" ellipsoids (see Sec. 3.1 of the paper). It also allows to bias the solution towards a certain degree of boxiness or diskiness and/or to certain $p(r)$ and $q(r)$ profiles. Unless we observe a triaxial galaxy exactly along one of the principal axes, the deprojection is unique if the density is a (deformed) ellipsoidal. Moreover, our code tackles the degeneracy problem, allowing for an exploration of possible density distributions that project to nothing. Finally, we show how the possible viewing angles of a galaxy can be significantly reduced from photometry alone. This not only helps the dynamical modeling, but also allows us to directly estimate

galaxy shapes from photometry alone (de Nicola et al., 2022b).

- In Neureiter et al. (2021) (hereafter BN21), we presented our novel triaxial Schwarzschild code SMART, which extends the axisymmetric code of Thomas et al. (2004). It exploits several advanced features: it fits the full non-parametric LOSVDs rather than Gauss-Hermite parametrisations, uses a 5-dimensional starting space for a better orbit sampling in the central regions and adopts a novel model selection technique to prevent overfitting, optimise the smoothing and deal with the different number of Degrees-of-Freedom (DOFs) for a given model (Lipka & Thomas, 2021; Thomas & Lipka, 2022). The code was tested using a sophisticated, state-of-the-art N-body simulation aimed at reproducing the formation history of giant ellipticals (Rantala et al., 2018, 2019) showing that if the intrinsic 3D distribution of the stars is known, the code shows excellent recoveries of $M_{\rm BH}$, Υ , the normalization of the DM halo as well as the internal velocity moments with an accuracy never achieved by any pre-existing Schwarzschild codes.

Here we combine the two algorithms and use both photometry and kinematics to constrain the galaxy viewing angles even better, and to recover the correct galaxy shape and anisotropy profile. Moreover, we quantify how large are the errors on the recovered mass parameters coming from the deprojections. To our knowledge, only in van den Bosch & van de Ven (2009) such study has been attempted. The paper is structured as follows. In Sec. 2 we briefly report the main features of the N-body simulation and our two codes. Sec. 3 describes our procedure, whose results are presented in Sec. 4. Finally, we draw our conclusions in Sec. 5. Results on the recovery of the mass parameter are presented in a companion paper (Neureiter et al. 2022, submitted to MNRAS, hereafter Paper I).

4.2 Data and Code

We test our strategy using an N-body simulation extensively described in Rantala et al. (2018, 2019). The simulation follows the merging process of two gas-free elliptical galaxies with supermassive black holes and was originally performed to study the formation and the evolution of the so-called core elliptical galaxies, whose light-deficient central regions are thought to be generated by (multiple) BH scouring events (Faber et al., 1997; Merritt, 2006a; Kormendy & Bender, 2009; Thomas et al., 2016; Rantala et al., 2018, 2019; Mehrgan et al., 2019).
The N-body simulation shows features commonly observed in massive early-type galaxies and closely resembles NGC1600 (Thomas et al., 2016). Projected on the sky, it is relatively round in the central regions and becomes more flattened when moving out to large radii (see Fig. 4.1). In terms of the internal shapes, the galaxy is close to spherical in the central core and then becomes triaxial at large radii (see Fig. 4.2). The anisotropy β, defined as

$$\beta = 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2}, \tag{4.1}$$

is negative in the central regions. Here $\sigma_r, \sigma_\theta, \sigma_\phi$ are the components of the velocity dispersion in spherical coordinates. This tangential bias is observed in several massive ellipticals (Thomas

et al., 2014, 2016; Mehrgan et al., 2019). It is likely due to the core scouring mechanism, which leads to an ejection of stars with radial orbits in the central regions (Rantala et al., 2018, 2019; Frigo et al., 2021).
We tested four different projections of the simulation at four different viewing angles (see Fig. 4.3). Two projections are along the intermediate (INTERM) and the minor (MINOR) axis of the galaxy[2], respectively, one lying exactly in between (MIDDLE) and finally another one at random viewing angles (RAND). These are summarized in Tab. 4.1, while their isophotes are shown in Fig. 4.1. The projections along the principal axes are the only ones where the deprojection is degenerate even if the surface brightness profile would *exactly* be the projection of a "deformed ellipsoid" model (see Sec. 4.2.1 below), since at least one of the two intrinsic shape parameters $p(x)$ or $q(x)$ is not constrained when the LOS coincides with one of the principal axes of the galaxy. MIDDLE is a case we already considered in dN20, while RAND happens to be the only projection showing a significant isophote twist ($\sim 20°$), although the isophotes only show a weak ellipticity $\varepsilon \leq 0.15$.

4.2.1 Deprojection

Our triaxial deprojection algorithm SHAPE3D is extensively described in dN20. Here we just report its main features for the sake of the reader's convenience.

- The surface brightness and ρ are placed onto polar elliptical and ellipsoidal grids, respectively;
- The algorithm works under the assumption that a galaxy can be described by what we call a "deformed ellipsoid", namely an ellipsoid whose radius is given by

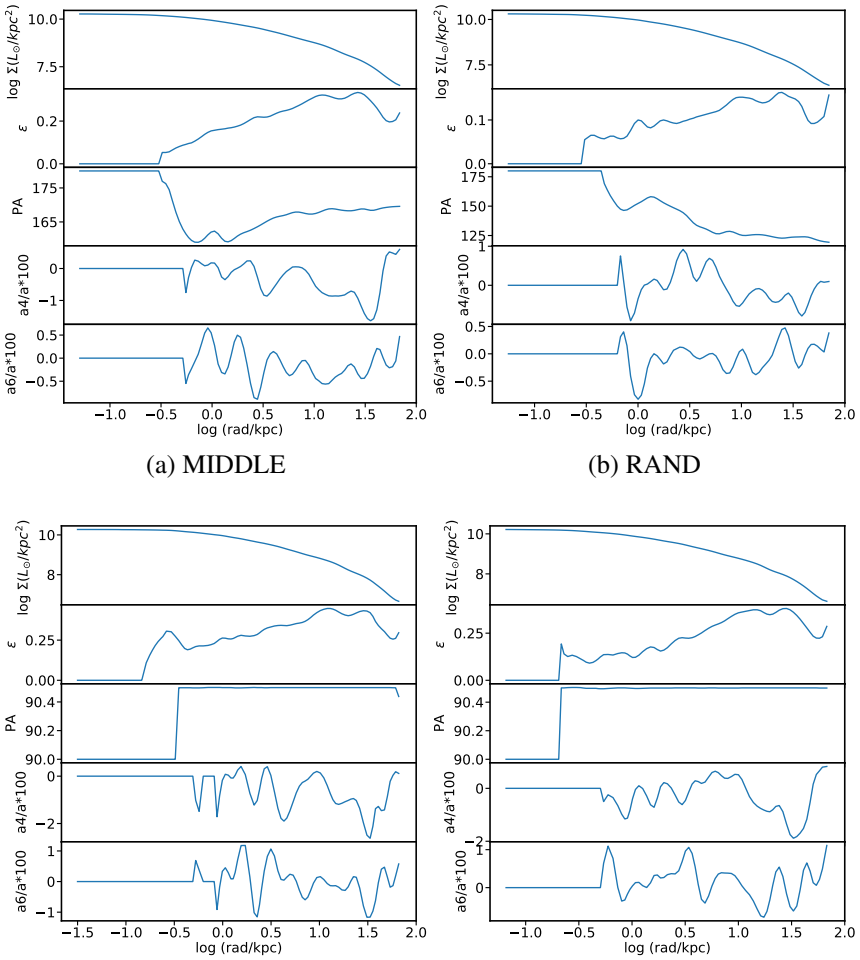
$$m^{2-\xi(x)} = x^{2-\xi(x)} + \left[\frac{y}{p(x)}\right]^{2-\xi(x)} + \left[\frac{z}{q(x)}\right]^{2-\xi(x)} \tag{4.2}$$

  where the exponent ξ can be used to generate disky ($\xi > 0$) or boxy ($\xi < 0$) iso-density surfaces. The three one-dimensional functions $p(x)$, $q(x)$, and $\xi(x)$, along with the density on the x-axis $\rho_x(x)$, specify ρ at each point of the grid.
- Since the algorithm is not a fully parametric method, it requires regularization. What the code minimizes is the likelihood $L = -\frac{\chi^2}{2} + P$ where χ^2 compares the differences between the observed and the modeled surface brightness and P is a penalty function used to disfavour unsmooth solutions.

4.2.2 Dynamical Modeling

`SMART`, the code that we use to compute dynamical models is the triaxial sibling of the Schwarzschild axisymmetric routine of Thomas et al. (2004). The code was presented in BN21. Schwarzschild

[2]We do not discuss the projection along the major axis, see Paper I.



(a) MIDDLE (b) RAND

(c) INTERM (d) MINOR

Figure 4.1: Isophotes of the four projections (Tab. 4.1) of the N-body simulation considered throughout this paper. For each projection, the four panels show (from top to bottom) the logarithmic surface brightness Σ, the ellipticity ε, the position angle PA, and higher-order distortions a_4 and a_6 of the isophotes as a function of the semi-major axis length.

Projection	$(\theta, \phi, \psi)^\circ$	Remarks
MIDDLE	(45,45,45)°	-
RAND	(60.4,162.3,7.5)°	Viewing angles have been drawn randomly
INTERM	(90,90,90)°	Projection along y. $p(r)$ photometrically unconstrained
MINOR	(0,90,90)°	Projection along z. $q(r)$ photometrically unconstrained

Table 4.1: The four projections of the N-body simulation studied in this paper. *Col 1:* The projection name. *Col 2:* The projection angles. *Col 3:* Notes on the individual projection.

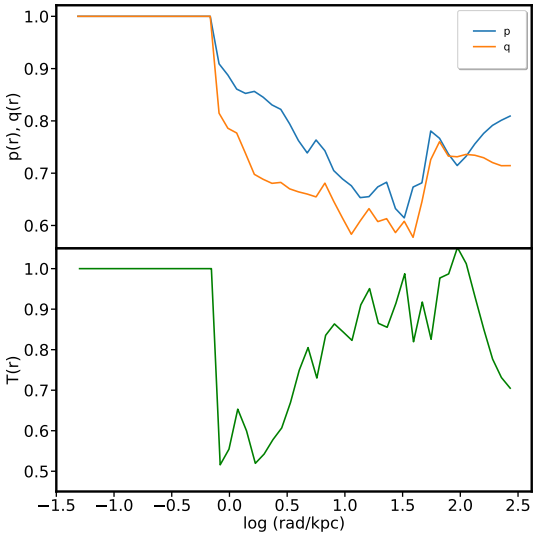


Figure 4.2: $p(r), q(r)$ (top panel) and $T(r)$ (bottom panel) profiles derived for the N-body simulation. The galaxy is spherical in the central regions, reaches the maximum triaxiality at ~3 kpc and then becomes prolate in the outskirts.

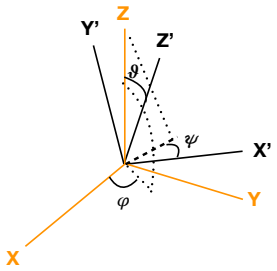


Figure 4.3: Geometric meaning of the viewing angles θ and ϕ, which determine the LOS direction Z', and ψ, which is a rotation around the LOS itself.

models are very flexible and do not require *a priori* assumptions on the shape and/or the anisotropy of the galaxy. As done above for the deprojection routine, we report here the most important features of the code.

1. The triaxial densities ρ derived from the deprojection of the surface brightness for a wide range of orientation parameters are considered;

2. A trial total density is constructed as:

$$\rho_{\rm TOT} = M_{\rm BH} \times \delta(r) + \Upsilon \times \rho + \rho_{\rm DM} \tag{4.3}$$

   where the first term is the point-like Keplerian potential coming from the black hole, the second term yields the stars' contribution through Υ and the deprojected density ρ and the third term is the DM density, which we compute assuming a modified gNFW (Navarro et al., 1997) profile (see below);

3. Poisson's equation is solved to obtain the potential which allows the computation of a representative time-averaged orbit library;

4. The orbital weights are computed such that the difference between the modeled and the observed LOSVDs is minimized for the assumed orientation and mass profile, with the stellar density as a Lagrangian constraint.

Steps (ii)-(iv) are then repeated for different $M_{\rm BH}$, Υ , orientation parameters and dark halo parameters to find the best model. Like our axisymmetric Schwarzschild code, `SMART` uses the entire LOSVD as a constraint for the orbit model rather than just Gauss-Hermite moments. Like most Schwarzschild codes we use a penalised maximum-likelihood approach to deal with the large number of formal model variables. Specifically, our code uses a maximum-entropy technique and maximises

$$\hat{S} = S - \alpha\chi^2, \tag{4.4}$$

where S is an entropy function. For the current study we use an entropy term related to the Shannon entropy (see Paper I). The strength of the entropy penalty is controlled by the regularisation parameter α and χ^2 compares the observed LOSVDs to the fitted ones. The optimal choice of the smoothing strength in penalised models is important to avoid both overfitting and oversmoothing. Thomas & Lipka (2022) have derived a generalised information criterion AIC_p for penalised models and demonstrated how it can be used to optimise smoothing strengths in a purely data-driven way. To find the correct value of α that prevents the code from finding solutions which fit the data well but are too noisy or others which are too entropy-biased we minimise

$$\mathrm{AIC}_p = \chi^2 + 2\mathrm{m}_{\rm eff}. \tag{4.5}$$

It generalises the classical Akaike Information Criterion (AIC) by using the concept of effective free parameters ($m_{\rm eff}$, c.f. Lipka & Thomas 2021) rather than a count of the number of variables[3].

[3] $m_{\rm eff}$ is calculated by running bootstrap simulations and evaluating the covariance between different noise patterns added to the best-fit model and the response of the fit to this noise.

The connection between the smoothing α and $m_{\rm eff}$ can be intuitively understood since the higher the smoothing, the less flexible the model will be, thus resulting in a smaller number of effective free parameters. A new feature of our code is that we individually optimise the smoothing for each trial mass model (see Paper I).

Lipka & Thomas (2021) have shown that determining the best-fit model for a galaxy is a model selection problem rather than a classical parameter-estimation problem. This is due to the number of effective parameters $m_{\rm eff}$ that does not only vary for different smoothings but also from one mass model to another (Lipka & Thomas, 2021). Hence we also use model selection based on AIC_p to determine the best-fit mass and orientation parameters (see Paper I).

4.3 Methodology

Our strategy is similar for the four N-body projections highlighted in Tab. 4.1. We simulate realistic observational conditions by combining both the deprojection and the dynamical model as we do for observations of real galaxies. We first run several deprojections in one octant to shrink the region of possible viewing angles (Sec. 4.3.1). Then, we dynamically model all plausible densities (Sec. 4.3.2) in order to find the viewing angles that give the best agreement with the observed kinematics and analyze the resulting shapes and anisotropy profiles.

4.3.1 Reducing the number of viewing direction with photometry only

We want to reduce of the number of viewing angles compatible with a specific photometric data set. In dN20, we showed that SHAPE3D can deal with this task and – moreover – can recover the correct intrinsic 3D density ρ (for the correct viewing angles). This is true as long as the object under study can be approximated by a nearly ellipsoidal shape and overcomes the degeneracy that is inherent to deprojections in general (see App. A of dN20)[4]. Our approach is described in detail below.

1. We generate the four galaxy images corresponding to each of the four N-body projections (Tab. 4.1) by projecting the intrinsic density calculated from the particles. Differently from what we did in dN20, here we choose not go through isophotal fits and directly fit the projected image[5].

2. We define a (θ, ϕ, ψ) grid, with (θ, ϕ) going from 0° to 90° and ψ going from 0° to 170°, each angle with a step of 10°. This is justified as long as we can assume triaxial symmetry, i.e. that the galaxy appears identical when viewed from different octants. For the tests in this paper this is guaranteed because we average the particle distributions in different octants (see also eq. 10 and Fig. 1 of dN20). The assumption works also well for the most massive ellipticals which do not show significant disky features or bars.

[4]Even for ellipsoidal bodies, the uniqueness of the deprojection is only true if we do not project along one of the principal axes. In Sec. 4.4.1 we discuss a possible solution for this case.

[5]If we had chosen to go through isophotal fits, then the level of noise in the images would be smoothed out since it would be impossible to reproduce the noisy isophotes properly using Fourier coefficients.

3. We deproject the surface brightness images for every possible viewing direction that we have defined on our (θ, ϕ, ψ) grid. Each deprojection is carried out using a 30×12 grid for the surface brightness and a $50 \times 13 \times 13$ grid for ρ. The smallest radius of both grids is at 0.05 kpc (for the simulation 1 kpc $\sim$ 10 asec), whereas the largest radii are at 68 kpc for surface brightness and at 270 kpc for ρ.

4. We select a threshold for the RMS, $\mathrm{RMS_{thr}}$, above which a viewing direction is discarded. The values we choose are 0.01 mag for RAND, 0.013 mag for MIDDLE and 0.02 mag for MINOR and INTERM. These thresholds allow us to discard ~90% of the deprojected densities.

5. We then discard all densities showing $p(r)$ & $q(r)$ profiles such that the order relation between the principal axes is not conserved at all radii, i.e. profiles intersecting with each other. We do this because, as shown in dN20, such solutions are likely to generate twists > 40°, which are not observed in (relaxed) massive ellipticals. All viewing directions fulfilling this requirement and yielding an RMS smaller than $\mathrm{RMS_{thr}}$ are dynamically modeled. As one can see from Tab. 4.2, the deprojection cut-off does a very good job in reducing the number of viewing angles compatible with the photometry, given that we typically need to sample only ~5% of the deprojected light distributions.

6. Finally, when we look at a galaxy along one of the principal axes[6] ρ cannot be uniquely constrained by the projected surface brightness alone. In order to test to which extent we can probe the full range of allowed densities, we take the MINOR case as an example and consider various different values for the unconstrained shape parameter (in this case $q(r)$) and run additional deprojections. This is only needed when the LOS lies on one of the principal axes, meaning that we need to perform this exercise at $(\theta, \phi) = (0, 90)^\circ$. For each one of these light densities, we need to launch a separate set of dynamical models keeping the viewing angles fixed but varying the 3D density. For this exercise, we sample $q(r)$ from 0.5 to 0.8.

The typical RMS = $\sqrt{< (\ln(I_{\mathrm{obs}}/I_{\mathrm{fit}}))^2 >}$ between the true and the recovered surface luminosity for the best-fit solutions are ~0.009 for MIDDLE and RAND and ~0.012 for the two projections along the principal axes.
From Fig. 4.5 we also see that the correct LOS, specified by θ and ϕ, is always included between the densities which we model with `SMART`. Before moving on to the dynamical model, it is worth making a few comments about the different N-body projections:

- The RMS= $\sqrt{< (\ln(\rho_{\mathrm{true}}/\rho_{\mathrm{fit}}))^2 >}$ between the true ρ from the N-body simulation is 0.095 and 0.097 for MIDDLE and RAND, respectively. Along the principal axes the RMS is a bit worse (~0.15), as expected since either $p(r)$ or $q(r)$ cannot be uniquely recovered anymore. For each of the four projections, we show in Fig. 4.4 the comparison between the

[6]A clue about a galaxy orientation along the principal is e.g. the lack of isophotal twist, which cannot occur in this case.

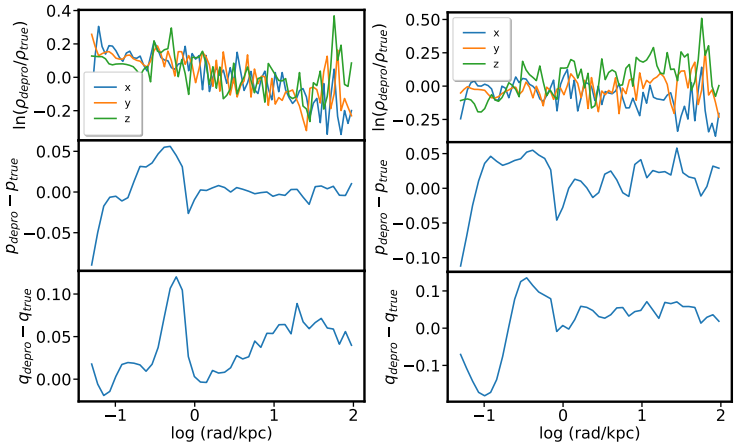
observed and the recovered density, as well as between the observed and the recovered $p(r)$ and $q(r)$. The largest deviations are, as expected, observed for the INTERM projection along the y-axis (and thus for $p(r)$) as well as the MINOR projection along the z-axis (and thus for $q(r)$), since in these cases information about $p(r)$ or $q(r)$ cannot be recovered from the photometry given the position of the LOS.

- The roundest projection RAND shows the largest number of possible viewing directions (although the threshold is only a factor ∼1.1 above the best-fit RMS), which is expected since the isophotes are very round, while the other three projections are flatter (Fig. 4.1). On the other hand, INTERM and MINOR are much better constrained, with no solutions when ψ is wrong. This is intriguing since their ellipticity profile is similar to the MIDDLE case, for which a much larger variety of solutions is found, *and* with a larger threshold compared to the respective best-fit RMS values. This might be due to the twist being as small as ∼0.5°, since numeric uncertainties in SHAPE3D might generate solutions with small twists and a profile without twist is particularly simple to fit assuming an inclination along the principal axes.

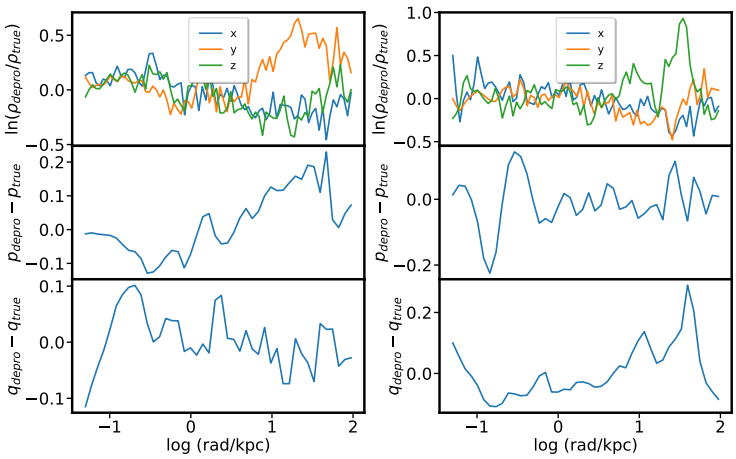
4.3.2 Dynamical modeling

Having significantly reduced the number of possible viewing angles using photometry only, we now turn to the dynamical modeling in order to further constrain the viewing angles and therefore the galaxy shape. Technical details about how we derive the kinematics derived from the simulation, such as the resolution and the number of fitted bins for each projection, the noise level etc. are discussed in Paper I.

Our strategy consists of taking each density distribution, which has survived the deprojection cut-off, and its corresponding set of viewing angles and use it in eq. 4.3 along with different $M_{\rm BH}$, Υ values and the DM halo density. We sample 10 $M_{\rm BH}$ values linearly spaced in the range $\left[1\times10^{10}, 3\times10^{10}\right] M_{\odot}$, while for Υ we use $[0.6, 1.4]$ as interval, again sampled with 10 values. For the DM halo we assume a profile similar to a Zhao (1996) profile (with $\alpha = 1$) in the scaled ellipsoidal radius r/r_s. The free parameters are the inner slope γ, the scale radius r_s and the density normalisation C. Since the N-body simulation started with progenitor halos approximated by a Hernquist sphere, we adjust β such that the outer logarithmic slope of our halo model equals $(\beta-\gamma) = 4.5$ rather than the canonical NFW value of $(\beta-\gamma) = 3$. We measured the asymptotic slope of the simulated halos directly from the DM particles of the simulation at large radii. Because the DM halo derived from the simulation is triaxial we have the two intrinsic shape parameters $p_{\rm DM}$, $q_{\rm DM}$ of the DM halo as additional free parameters. Thus, we have a total of 10 parameters to be fitted. For this, we use the `NOMAD` software (Audet & Dennis, 2006; Le Digabel, 2011; Audet & Warren, 2017) which automatically explores the parameter space, looking for the model minimizing $\mathrm{AIC_p}$. The best-fit model parameters are the ones where $\mathrm{AIC_p}$ is minimized. In principle, the $\mathrm{AIC_p}$ distribution – just like χ^2 – could be used to estimate the uncertainty of the models (masses, viewing angles etc.). We will illustrate this below for the anisotropy. However, in general, we prefer to estimate model uncertainties in a different way for two reasons. Firstly,



(a) MIDDLE (b) RAND



(c) INTERM (d) MINOR

Figure 4.4: Comparison between recovered (at the correct viewing angles) and true quantities for the four projections of the N-body simulation (Tab 4.1) considered throughout the paper. In the top panels we show the ratios of the light densities along the three principal axes, whereas in the bottom panels we show the differences for the corresponding $p(r)$, $q(r)$ profiles.

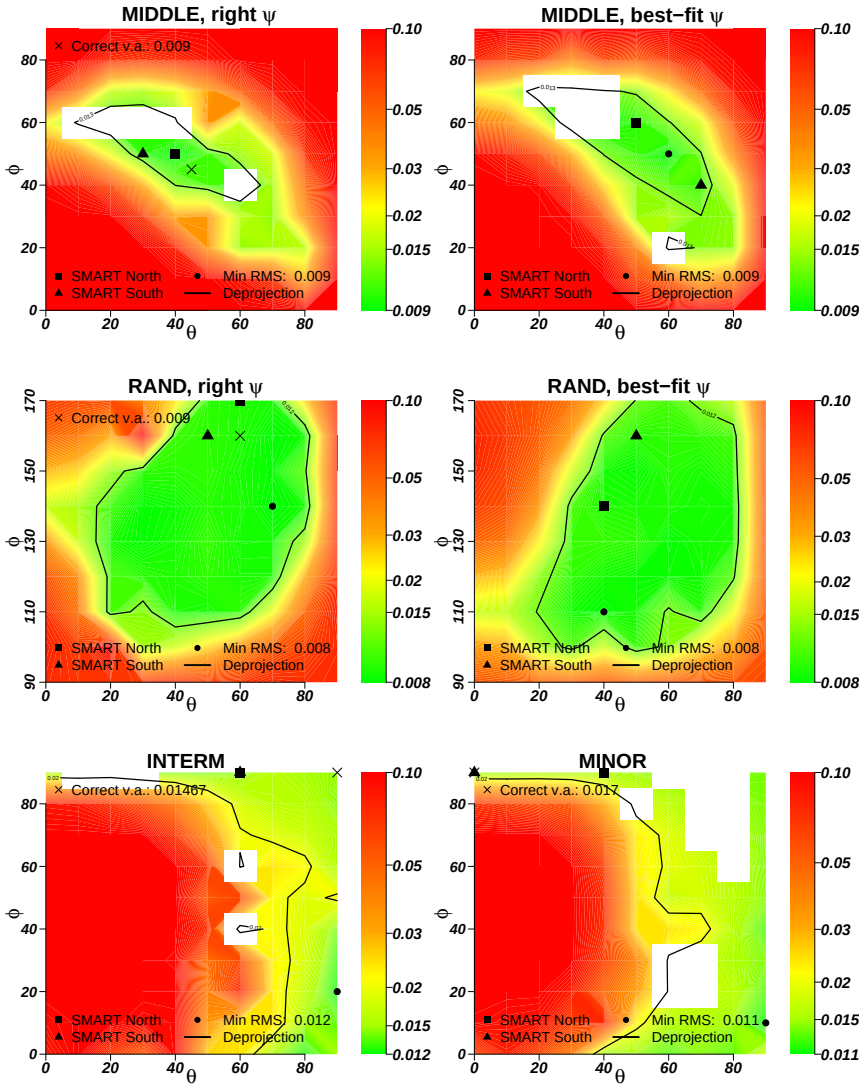


Figure 4.5: Logarithmic RMS errors in surface brightness $\Delta \log \Sigma$ (left) for the four N-body projections (Tab. 4.1), obtained for constrained-shape deprojections at different assumed θ, ϕ viewing angles for the correct values of ψ and for the best-fit ψ from the dynamical models. For INTERM and MINOR these two values coincide. The cross labels the correct (θ, ϕ). This is not shown in the "best-fit" plots because in this case ψ is not the correct one. The black dot is at the best-fit solution from the deprojection, while the black square and triangle indicate the two best-fit solutions from the dynamical modeling. The solid contour delimits the area inside which the RMS in surface brightness is within twice the minimum value we find on the grid. Finally, empty (white) squares depict regions discarded because of crossing p and q profiles.

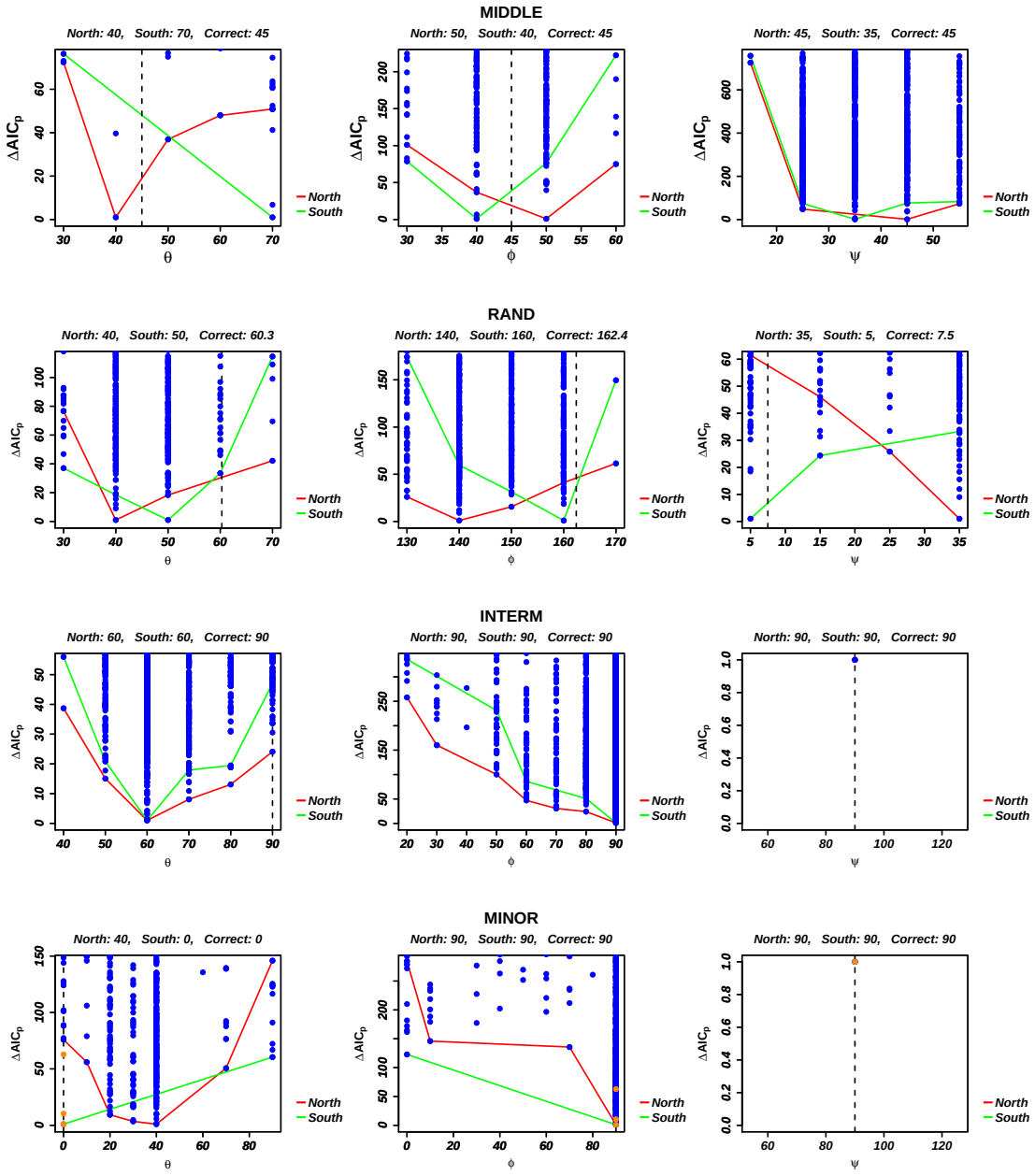


Figure 4.6: 1D-plots of $\mathrm{AIC_p}$ (eq. 4.5) against the viewing angles (from left to right: (θ, ϕ, ψ)) for the four N-body projections (From top to bottom row: MIDDLE, RAND, INTERM, MINOR) considered in this work. In each plot, the blue points are individual model runs (though we tested many more models outside the plotted range) while the dashed black line labels the true viewing angles. The red and green lines follow the lower envelope of the best-fit models for each (θ, ϕ, ψ) value and refer to the two halves of the galaxy ("North" and "South"), which we model separately. For the MINOR projection, we show the models along the principal axes when testing different deprojections as orange points.

Table 4.2: Summary of the discarding process of possible viewing angles using photometry only. *Col.1:* The particular model (see also Tab 4.1). *Col.2:* Total number of deprojections (i.e. different viewing angles) that we tested. *Col.3:* Number of viewing angles left over after the RMS cut-off. *Col.4:* Number of densities that we dynamically modeled. The difference between *Col.3* and *Col.2* yields the number of viewing angles omitted through the RMS, whereas the difference between *Col.4* and *Col.3* yields the number of viewing angles omitted because of implausible $p(r)$ and/or $q(r)$ profiles.

Model	Number of deprojections	After depro cut-off	Dynamical modeling
MIDDLE	1800	61	10
RAND	1800	75	9
INTERM	1800	25	6
MINOR	1800	25	6

when dealing with real observations the noise level in the data is often not exactly known but has to be estimated – which makes $\mathrm{AIC_p}$ or χ^2 distributions uncertain by themselves. Secondly, reliable errors from $\mathrm{AIC_p}$ or χ^2 require to sample the respective distributions in all parameters *completely*. In order to gain efficiency (we only seek to find the *optimal* model) and to get independent from the noise estimate for the data, we split the available kinematical data into two sets, which we label "North" and "South". The data set to which each bin is assigned to gets determined by the bin position with respect to the galaxy's apparent minor axis. In this way, each of the two data sets has roughly the same amount of bins. We model these two data sets independently and estimate the uncertainty of all relevant model parameters from the variance between these two fitting runs. Below we will also show the 1D functions $\mathrm{AIC_p}\,(\theta)$, $\mathrm{AIC_p}\,(\phi)$, $\mathrm{AIC_p}\,(\psi)$ obtained by minimising $\mathrm{AIC_p}$ over all other parameters to illustrate how well individual parameters can be estimated. For the simulations this is justified since we know the noise level in the "data" exactly. We summarize in Tab. 4.2 the number of tested viewing angles, those that survived the RMS cut-off and those for which $\mathrm{AIC_p}$ coming from the dynamical model is within 50 + $\mathrm{AIC_{p,min}}$ ("good dynamical models")[7].

4.4 Results and Discussion

We now turn to the analysis of the results and discuss them in detail, focusing on the viewing angles, anisotropy and shape recovery. The results on M_{BH} and Υ are discussed in BN21; for completeness, we report them in Tab. 4.3.

[7]This AIC threshold allows us to perform a very conservative comparison, since typically a diffence of 10-20 in $\mathrm{AIC_p}$ is already considered significant.

	$\lvert M_{\rm BH}\rvert$	$\lvert\Upsilon\rvert$
MIDDLE	1.67 ± 0.07	1.04 ± 0.05
RAND	1.67 ± 0.22	1.09 ± 0.13
INTERM	1.78 ± 0.11	1.05 ± 0.08
MINOR	1.56 ± 0.11	1.05 ± 0.08

Table 4.3: Recovery precision of $M_{\rm BH}$ and Υ for the four projections described in this paper. The $M_{\rm BH}$ values are given in units of $10^{10} M_\odot$. The deviations from the true values $M_{\rm BH} = 1.7 \times 10^{10} M_\odot$ and $\Upsilon = 1.0$ are less than 5%. Finally, the standard deviations are given by modeling the two sides of each projection.

4.4.1 Viewing angles recovery

dN20 showed that photometric constraints alone can shrink the range of possible viewing angles significantly because deprojections assuming the wrong viewing angles fit the observations not as good as deprojections near the true viewing angles. Hence, first testing deprojections in one octant and then using a cut-off in the RMS achieved between the photometric data and the deprojection model allows to select the best intrinsic densities. In Fig. 4.5 we show how well the photometric data of the merger simulation can be deprojected as a function of the assumed orientation of the line-of-sight. The panels with the correct value of ψ are closely analogous to Fig. 20 of dN20. They show that the deprojection alone helps in reducing dramatically the range of viewing angles that need to be tested when dynamically modeling the galaxy.

The naive expectation would be that the additional constraints from the observed kinematics will improve the viewing angle recovery. The results of the dynamical modelling are summarized in Fig. 4.6, where we show $\mathrm{AIC_p}$ (eq. 4.5) as a function of the three viewing angles (Fig. 4.3) for the four N-body projections. Away from the principal axes (i.e. MIDDLE and RAND) the viewing angles are well constrained (within 15°), while for the two cases where the LOS coincides with one of the principal axes (INTERM and MINOR) this is only true for ϕ and ψ (which are recovered correctly). The third angle θ shows a slightly larger offset (30° for INTERM, 20° for MINOR). For these two principal-axis projections the angle ψ was already fixed from the photometric constraints alone.

To compare the dynamical and photometric results we have also included the best-fit dynamical viewing angles in Fig. 4.5. In some cases the dynamically determined best-fit ψ differs from the correct value. We included additional panels for these best-fit ψ if necessary. In general, for the *correct* value of ψ, our Schwarzschild code can constrain the LOS position with less scatter than the deprojection alone – as expected – although with no significant improvement on the best-fit viewing angles.

4.4.2 The primary importance of the deprojected shape

As already noted, the deprojection of a triaxial body is generically degenerate. In particular, there exist infinitely many deprojections even at a given viewing angle. For example, the flattening $q(r)$ is completely unconstrained when the LOS coincides with the intrinsic minor axis. The reason

is that the intrinsic density distribution along the z-axis (the LOS in this case) is photometrically simply not accessible in this case (and similar for the other principal axes). However, our deprojection code SHAPE3D allows to probe deprojections with different intrinsic shapes at the same viewing angle. So, for the MINOR projection we constructed a set of deprojections all assuming the minor-axis as LOS but with different intrinsic $q(r)$. In Fig. 4.7 we show for each constant $q(r)$ value (see Sec. 4.3.1) in the range $[0.5, 0.8]$, the corresponding best-fit $\Delta\mathrm{AIC_p}$ values for the two modeled galaxy halves. The minima are located at $q = 0.65$ and $q = 0.6$, in agreement with the true profile in the region where the galaxy becomes triaxial (see Fig. 4.2). In other words, the kinematic constraints are sufficient to identify the correct intrinsic shape out of the set of deprojections with different $q(r)$. In fact, the constraints on the shape are quite strong: the $\Delta\mathrm{AIC_p}$ changes by more than $\Delta\mathrm{AIC_p} > 40$ over the sampled shape interval. For comparison, in the case of the MINOR projection, a change of $\Delta\mathrm{AIC_p} \sim 40$ corresponds to a change of $\theta \sim 40°$. The constraints on the shape are therefore quite significant, in particular for the viewing angle recovery.

In Fig. 4.6 the improvement in $\mathrm{AIC_p}$ due to the inclusion of the above described deprojections with different $q(r)$ is indicated by the orange points. As expected, including additional deprojections with different intrinsic flattenings for this orientation significantly improves the results of the viewing angle recovery. In fact, they allow to recover the correct galaxy viewing angles in the "southern" part of the galaxy.

These results imply that the constraints from modern kinematic data may in many cases be sufficient to discriminate between different deprojections at the same viewing angle. For a successfull viewing angle recovery flexible deprojection tools that allow to probe different intrinsic shapes at a given viewing angle are therefore vital. In particular, our results suggest, that if we would try to also optimise the deprojected shapes at other LOS orientations, the viewing angle recovery with the dynamical models would improve. However, the precision that we can achieve with our fiducial deprojections in terms of the mass recovery (see Paper I) and in terms of the recovery of the intrinsic shapes and orbital anisotropies (see below) suggest that this is not necessary and that the viewing angles themselves are only of secondary importance for the dynamical modelling.

4.4.3 Shape recovery

The orientation parameters and the intrinsic shape are intimately connected. However, there is no generic simple connection for *triaxial* objects in general – due to the high degree of degeneracy in the deprojection (cf. last Sec. 4.4.2). In many currently used deprojection methods a close connection between shape and orientation is nevertheless imprinted through additional assumptions made upon the intrinsic structure of galaxies. Now we test how well we can recover the intrinsic shape of the merger simulation from the four projected mock data sets.

Our main results from the dynamical modelling are:

- Our results are accurate: we get Δp and $\Delta q \lesssim 0.1$ for each of the four projections;
- The kinematical data help in reducing the scatter of the recovered profiles;

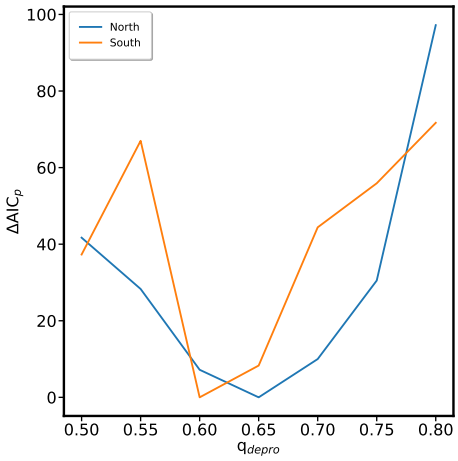


Figure 4.7: Best-fit AIC_p values plotted against the corresponding q value used to generate ρ. The results refer to the MINOR projection.

- The results are best when the LOS is not one of the principal axes.

This is summarized in Fig. 4.8. For the purely photometric results, we compare the shape *profiles* yielded by the deprojections using a similar approach as in de Nicola et al. (2022b), i.e. we consider every deprojection that is qualified as good according to the criteria given in de Nicola et al. (2022b) and calculate the average $< p(r) >, < q(r) >$ among them. The range encompassed by these 'good' deprojections is shown by light-coloured regions in Fig. 4.8, the average profiles are shown as dashed lines. We calculate the differences between $< p(r) >, < q(r) >$ and the true profiles at all radii and then average these, reporting the results in Tab. 4.4. We observe that the shape is best recovered for MIDDLE, with $\Delta p, \Delta q < 0.06$. Similarly good is the recovery of the rounder RAND projection. This is somewhat expected, since when the viewing angles are located between the principal axes, SHAPE3D performs optimally in recovering the intrinsic shape if the viewing angles are known (de Nicola et al., 2020).

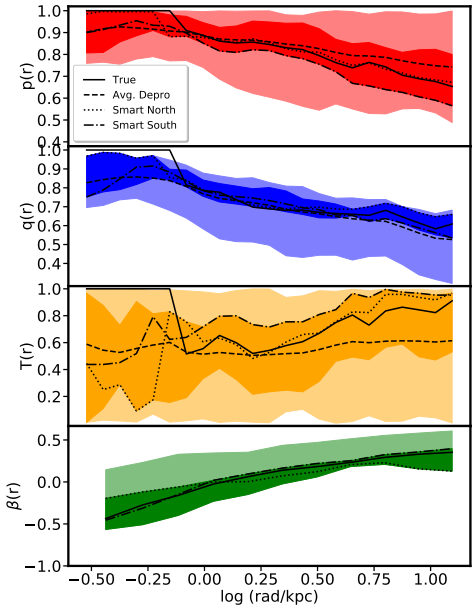
The results for the principal-axis projections INTERM and MINOR are slightly worse, which is again expected given the fact that the intrinsic shapes are less constrained[8]. Nevertheless, the intrinsic shapes can be reconstructed photometrically with an accuracy of ~ 0.1.

For all four projections, the best-fit triaxiality *profiles* deviate from the truth in the central regions (where the simulation is spherical), but approach the value of one in the outskirts, where the simulation becomes prolate.

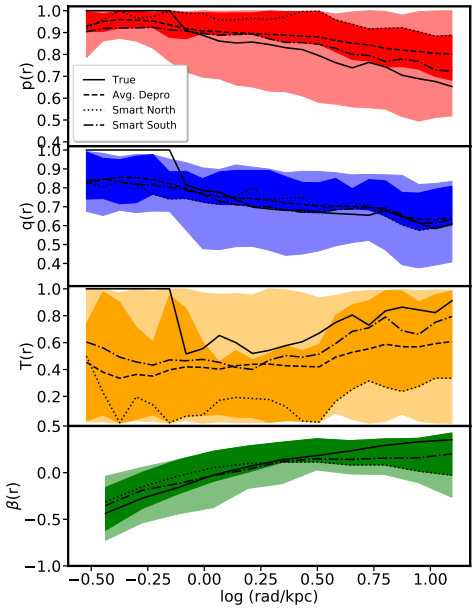
The dynamical modelling improves the shape recovery significantly. This is shown in Fig. 4.8 by the dark-coloured regions which encompass all dynamical models within $AIC_p \leq AIC_{p,min}+50$. The figure also shows the intrinsic shape profiles of the actual best-fit dynamical model for each projection and each modelled half as dotted and dash-dotted lines, respectively.

The darker regions from the dynamical models are in all cases within the larger uncertainty regions derived from the photometric constraints alone (light colors). To quantify this for the

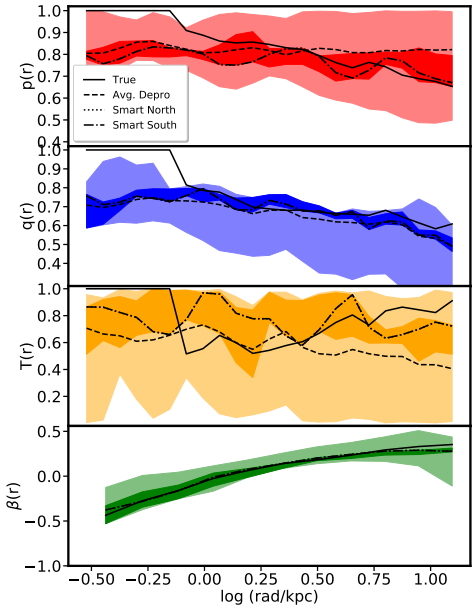
[8]For example, along the intermediate axis, the fact that we choose $P = 0.8$ for the ρ-grid implies that deprojections close to the intermediate axis will be biased towards $p(r) \sim 0.8$ at each r.



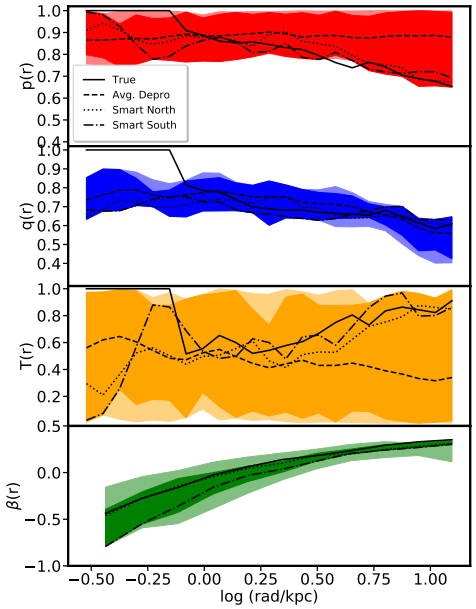
(a) MIDDLE



(b) RAND



(c) INTERM



(d) MINOR

Figure 4.8: For each of the four projections we show, from top to bottom panels, $p(r), q(r), T(r)$ and $\beta(r)$ intervals that we get considering every acceptable deprojection (lighter color) or only those deprojections for which the dynamical model yields an $\mathrm{AIC_p} \leq \mathrm{AIC_{p,min}} + 50$ (darker color). The solid lines show the correct profiles, whereas the dashed lines show the average profile that we get among all good deprojections. Finally, the dotted and dash-dotted lines are the best-fit profiles from the dynamical models.

shape parameters, we average $p(r)$ and $q(r)$ over all models within $\mathrm{AIC_p} \leq \mathrm{AIC_{p,min}} + 50$. The differences between these *average* profiles and the true ones and also the difference between the profiles of the single best-fit models and the true profiles are quoted in Tab. 4.4. As one can see, adding kinematical data to the analysis does improve the precision of the estimates, but not their accuracy. Even by choosing a very high threshold $\mathrm{AIC_{p,min}} + 50$ to select the favoured dynamical models, the interval embedding our 'good deprojections' is narrowed down in most cases, delivering smaller scatter. The true profiles are found within these intervals or very close to them, with the only exception being p for INTERM and q for MINOR in the central regions. This is expected, since p (for the INTERM projection) and q (for the MINOR projection) are hidden by the LOS (cf. Sec. 4.4.2).

4.4.4 Anisotropy recovery

The 3D intrinsic shape measured in terms of the $p(r)$ and $q(r)$ profiles and the orbital structure – the anisotropy profile $\beta(r)$ of a galaxy are related to each other through the Tensor Virial Theorem, although the shape does not uniquely determine the orbital anisotropy. Therefore, even the good shape recovery that we discussed in the last Sec. 4.4.3 does not guarantee that the orbital structure is well recovered. We quantify the anisotropy from the internal velocity moments yielded by our triaxial Schwarzschild code. They are computed as the quadratic mean of σ_r, σ_θ, σ_ϕ in spherical shells.

The main results from the dynamical modelling are very similar as for the shape recovery: our results are accurate and we get $\Delta\beta \lesssim 0.1$ for each of the four projections (Fig. 4.8 and Tab. 4.5). In particular, we recover the expected tangential bias in the central regions for all four projections very well together with the general trend towards radial anisotropy at larger radii. The largest deviations occur at the smallest and the largest radii, respectively, which is expected[9]

4.4.5 Model advancements

The analysis performed in this paper shows that the combination of our triaxial deprojection and dynamical modeling codes enables us to recover the intrinsic shape and anisotropy of the galaxy with good accuracy, meaning that intrinsic degeneracies in both the deprojection problem and in the determination of the orbital dynamics do not play a significant role (see Sec. 4.4.6 below). This is true at least for the setup that we have chosen (integral-field-type data coverage and usage of the entire LOSVDs as constraints), for the given simulation (with a realistic formation scenario and intrinsic shape/orbital structure) and for our newly developed codes. An important contribution comes from SHAPE3D, which allows to narrow down the range of possible viewing angles significantly. Within the variation of the intrinsic shapes among the remaining viewing angles, this already allows to determine the shape of the simulation with good accuracy. A posteriori, this verifies that SHAPE3D alone can be used to make inferences about the intrinsic

[9]At large radii, because near the edge of the field of view (FoV) the constraints on the orbital structure become weaker since ever more orbits have apocentres outside the region constrained by kinematical data. The same is true at small radii, where the finite resolution (mostly of the simulation) weakens the constraints on the orbit model.

shapes of real galaxies based on photometric data alone – at least in a statistical sense (de Nicola et al., 2022b).

BN21 showed that with noiseless data and without the uncertainty introduced by the deprojection step the anisotropy of the same N-body simulation can be reconstructed within $\Delta\beta = 0.05$ with our triaxial orbit code. The larger errors here come from the realistic amount of noise in the kinematic data and from the uncertainties intrinsic to the deprojection. However, the anisotropy recovery is still very good. The fact that we have chosen to go up to $\mathrm{AIC}_{\mathrm{p,min}} + 50$, which is very conservative, and that we are able to recover the correct $p(r)$, $q(r)$ with an accuracy typically smaller than 0.085 and reproduce the correct trend of $\beta(r)$ regardless of where the correct viewing angles are, provides a further indication of the stability of our dynamical models and the negligible role of degeneracies when using a setup as described here (see also Sec. 4.4.6).

Lipka & Thomas (2021) have shown that the optimisation of the mass and orientation parameters of Schwarzschild fits is a model selection problem rather than a simple parameter optimisation. Intimately connected to this is the fact that the effective number of parameters $\mathrm{m_{eff}}$ that a Schwarzschild Fit consists of is variable from model to model. The different number of DOFs, which depends on the particular model (in our case the chosen potential), can bias the results if one uses a χ^2 optimization method. We discuss the importance of the correct model selection for the mass models in Paper I. In a similar way, the optimisation of any smoothing penalty can also be performed using a model selection (Thomas & Lipka, 2022). To demonstrate the importance of the smoothing optimisation, we show in Fig. 4.9 for the MIDDLE projection the true β(r) profile, along with various profiles that we get for the southern half of the merger by just varying the smoothing strength α yet keeping the mass distribution and orientation fixed. We see that even in the same mass distribution different α values may yield anisotropy profiles with deviations of up to 30% from the true one. And this even though we only show β profiles that lead to formally acceptable fits to the data (i.e. $\chi^2/N_{\mathrm{data}} < 1$). An optimal choice of the smoothing is hence necessary to reach the accuracy that we report here.

4.4.6 Summary and Discussion

To summarize, the combination of our deprojection and dynamical modeling recovers the correct shape and anisotropy of the simulated galaxy with deviations $\lesssim 0.1$. As we show in Paper I the mass recovery has a similar accuracy of about 10 percent. This is not surprising as the masses can only be recovered with high accuracy when the orbital structure is correct and vice versa. The viewing angles turn out to be the most uncertain properties with an accuracy of about $\sim 30°$.

BN21 have shown that the anisotropy can be recovered very robustly from kinematic data similar to the one used here (full non-parametric LOSVDs and two-dimensional spatial coverage) when the LOS is given. In Sec. 4.4.2 we have seen that the shape can be recovered very robustly when the LOS is given. Moreover, since in these models the anisotropy was not held fixed and because we know that the best-fit models have the correct anisotropy it is quite straightforward to conclude that kinematic data of the kind used here contain enough information to constrain shape and anisotropy (and mass) together at a given LOS. A little more surprising is the fact that the full modelling of the N-body reveals that both, anisotropy and shape, can even be recovered robustly when the orientation of the LOS is a bit uncertain. One interpretation would be that

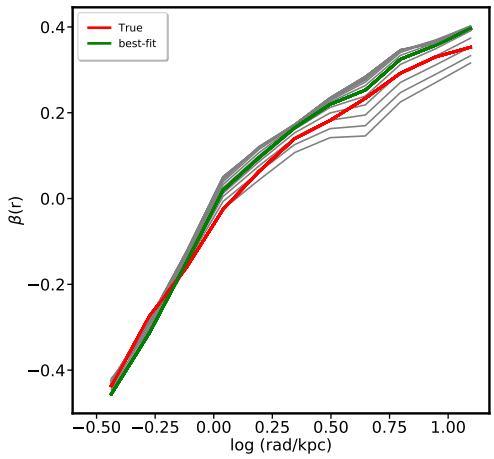


Figure 4.9: For the best-fit mass model and orientation of the southern half of the MIDDLE projection we show all β(r) profiles for every smoothing strength α that leads to an acceptable kinematic fit ($\chi^2/N_{\rm data} < 1$). The red line is the true profile from the simulation, while the green line is our fiducial model derived from the smoothing optimisation using AIC$_p$. All other profiles are plotted as grey lines. We see that the smoothing optimisation is an important factor for the improved accuracy of dynamical models that we report here.

the constraints on the viewing angles, the shape and the anisotropy do not interfere strongly with each other. And an observation in favour of the possibility that at least shape and anisotropy are not strongly entangled is the fact that the predicted uniform central anisotropy structure in massive (triaxial) ellipticals, i.e. the systematic change from central tangential anisotropy to outer radial anisotropy around the core radius, has already been demonstrated with axisymmetric models (Thomas et al., 2014). To a certain degree our tests suggest that something similar is true for shape and viewing angles: that the shape constraints force the model towards the correct deprojection even if the viewing angles are not correct. However, we have also seen that this critically depends on whether or not a deprojection with an appropriate shape is among the candidates or not. Flexible deprojection tools like SHAPE3D are therefore important. All in all it has emerged that for the recovery of the masses, intrinsic shapes and orbital structure the correct viewing angles are only of secondary importance. This is probably related to the fact that the merger simulation studied here and massive elliptical galaxies as well are not very strongly flattened.

4.4.7 Bias vs. scatter

A very important point that can now be addressed is whether or not one needs to go through the dynamical modeling in order to obtain acceptable estimates of the galaxy shape profiles. From Tab. 4.4 we see that the *bias* from the true shape of the simulation which we obtain can only be barely improved (if at all) using kinematical information. Therefore, our conclusion is that the

photometric information suffices to obtain a robust estimate (within 0.1) of the correct galaxy shape when considering the average over all orientations as best-fit guess. This approach is exactly what de Nicola et al. (2022b) used to derive shapes of Brightest Cluster Galaxies (BCGs). Nevertheless, the kinematical information helps in reducing the *scatter*, as it can be seen from both Fig. 4.8 (the darker regions are narrower) and Tab. 4.4. Therefore, the conclusion here is that if one simply wants to obtain an estimate of the galaxy shape, then simply deprojecting the surface brightness profile is enough, but in order to make these estimates more robust, it is still preferable to go through the dynamical modeling.
Finally, our findings suggest a possible approach for the dynamical modeling of the galaxy, which is needed to determine the orbit distribution along with the mass parameters. In fact, one could model *only one* light density, the one with p, q profiles closest to $< p >, < q >$ and only use this density for the dynamical model.

4.4.8 Comparison with previous studies

Another study focusing on the recovery of the intrinsic shape of a galaxy using both photometric and kinematic information is van den Bosch & van de Ven (2009). In this work, the authors show that the shape of a triaxial Abel model with constant $p(r) = 0.9$ and $q(r) = 0.77$ projected at $(\theta, \phi) = (60, 60)^\circ$ (thus at similar viewing angles compared to our MIDDLE and RANDOM projections) can be well constrained only if the galaxy shows significant rotation both in the central and in the outer regions. Our work shows that for the slowly rotating N-body simulation, $\Delta p, \Delta q \lesssim 0.1$ regardless of the photometry and the correct viewing angles. Another improvement with respect to van den Bosch & van de Ven (2009) is that in their case when a round, slow-rotator is considered, at almost every viewing angle a solution with recovered p, q $\sim$0.1 away from the true value (thus very good) can be found, while in our case even for the roundest projection (RANDOM) as well as for those without twists (INTERM and MINOR) our estimates are more accurate and we are able to exclude most of the viewing angles.
The triaxial code used by van den Bosch & van de Ven (2009) has been used in Jin et al. (2019) to recover the intrinsic shapes of nine simulated early-type galaxies from the Illustris simulations (three of which are triaxial). These galaxies are more similar to our N-body simulation since they have $p(r)$ and $q(r)$ profiles which are not constant as a function of radius. Here, average deviations of 0.07 in p and 0.14 in q (but with deviations as large as 0.25 in q) are found. Nevertheless, only four (for p) and one (for q) of the nine galaxies they consider show deviations smaller than 0.1, and the anisotropy profiles of these galaxies are also recovered with a lower accuracy (read their Figure 12). Moreover, Quenneville et al. (2022) have recently shown that the triaxial code used for the analysis (van den Bosch et al., 2008) did not project orbits correctly, which may lead to a substantial bias in mass and shape parameters (but see Thater et al. 2022).

Clearly, an exact comparison between these different works is not possible for a variety of reasons (e.g. van den Bosch & van de Ven 2009 and Jin et al. 2019 use the MGE as deprojection code), but given the results presented here our methodology appears superior.

Projection	Profile	$< \Delta p > \pm \sigma_{\Delta p}$	$< \Delta q > \pm \sigma_{\Delta q}$
MIDDLE	AVG ORIENT PHOT	0.046 ± 0.084	0.059 ± 0.071
	AVG ORIENT PHOT + KIN	0.033 ± 0.057	0.045 ± 0.042
	BF SMART	0.038 ± 0.034	0.044 ± 0.042
RAND	AVG ORIENT PHOT	0.069 ± 0.082	0.062 ± 0.089
	AVG ORIENT PHOT + KIN	0.091 ± 0.048	0.098 ± 0.057
	BF SMART	0.088 ± 0.055	0.070 ± 0.022
INTERM	AVG ORIENT PHOT	0.091 ± 0.102	0.106 ± 0.067
	AVG ORIENT PHOT + KIN	0.088 ± 0.035	0.086 ± 0.025
	BF SMART	0.088 ± 0.000	0.095 ± 0.000
MINOR	AVG ORIENT PHOT	0.105 ± 0.086	0.085 ± 0.055
	AVG ORIENT PHOT + KIN	0.086 ± 0.082	0.084 ± 0.050
	BF SMART	0.053 ± 0.023	0.101 ± 0.023

Table 4.4: Estimates of the mean deviations of the recovered $p(r), q(r)$ profiles from the correct ones from the simulations along with their RMS. *Col. 1:* The projection. *Col. 2:* Whether we consider the average among all deprojections (AVG DEPRO PHOT, light region of Fig. 4.8), among all deprojection for which $\mathrm{AIC_p} \leq \mathrm{AIC_{p,min}} + 50$ (AVG DEPRO PHOT + KIN, dark region of Fig. 4.8) or the best-fit solution from the dynamical modeling averaged over the two galaxy halves (BF SMART). *Cols. 3-4:* Differences between recovered and correct profiles, averaged on all radii for which we have kinematical data. For the INTERM projection the two best-fit solutions yielded by `SMART` are the same.

	MIDDLE	RAND	INTERM	MINOR
$< \Delta\beta > \pm \sigma_{\Delta\beta}$	0.067 ± 0.077	0.114 ± 0.048	0.023 ± 0.000	0.078 ± 0.076

Table 4.5: Similar as Tab. 4.4 but with the recovered $\beta(r)$ profiles from the true ones considering the two best-fit models for each galaxy half.

4.5 Conclusions

We have investigated how well the viewing angles, intrinsic shape and orbital structure of triaxial galaxies can be recovered by employing a novel approach to an N-body simulation with high resolution. For the first time we combine our newly developed codes for the modelling of triaxial galaxies: (i) our new semi-parametric triaxial deprojection routine SHAPE3D (de Nicola et al., 2020) which allows to probe degeneracies of deprojections at the same viewing angle and to shrink the region of possible orientations of a galaxy purely based on photometric data; (ii) our new triaxial orbit superposition code `SMART` (Neureiter et al., 2021) which exploits the entire kinematic information contained in non-parametrically sampled LOSVDs and uses a 5D orbital sampling to represent all orbital shapes in galaxy centers; (iii) our new model selection methods which allow to adaptively optimise the smoothing for each trial mass model/orientation and overcomes potential biases in χ^2-based approaches.

We tested projections along four representative viewing directions of this galaxy. We exploit the uniqueness of our deprojections for a given set of viewing angles and show that the region of possible viewing directions can be significantly reduced relying solely on the deprojections. Using the recovered luminosity densities as input for our triaxial Schwarzschild code, we determine the correct galaxy viewing angles to within 15° for MIDDLE and RANDOM, while for INTERM and MINOR θ is 30° and 20° off, respectively, but ϕ and ψ are perfectly recovered. This translates to robust estimates of the galaxy intrinsic shape profiles $p(r)$ and $q(r)$. In two cases where the LOS lies far away from the principal axes (MIDDLE and RAND), SHAPE3D provides the correct galaxy shape profiles within 0.1. The same is found for $p(r)$ for the MINOR projection and for $q(r)$ for the INTERM projection. For the MINOR case, where the $q(r)$ profile cannot be recovered since hidden by the LOS, we tested different deprojections with different intrinsic shapes, improving the $\mathrm{AIC_p}$ values and recovering the correct viewing angles. Thus, even if the best-fit angles do not change, this exercise leads to an improvement in the quality of the fit anyway, and should be repeated every time one finds a galaxy whose photometry is compatible with a deprojection along one of the principal axes. Moreover, in Paper I we have shown that the best-fit models also yield the correct BH mass and Υ parameters.
The anisotropy parameter β shows for each single projection the tangential bias in the central regions, expected to be generated from SMBH core scouring. It is significant that this is true even along the principal axes where either $p(r)$ or $q(r)$ cannot be recovered, showing the robustness of our Schwarzschild code in recovering the correct velocity moments and the correct orbital structure of the simulation. On the other side, the fact that even if the angles are not exactly correct the $p(r)$ and $q(r)$ profiles are well recovered (and so are the mass parameters, see paper I) hints at the possibility of needing a very low number of deprojections for the dynamical models, therefore reducing the parameter space to be sampled. All that is needed would be to analyze the favoured deprojections, select representative $p(r)$ and $q(r)$ profiles and pick up only one deprojection for each pair of profiles. Finally, our results point out that the known intrinsic photometric and kinematic degeneracies do not prohibit a precise and accurate reconstruction of the intrinsic structure of a triaxial galaxy. In our models, key ingredients are the non-parametric analysis of the photometric and kinematic data and advancements in the orbit sampling and model selection. All these novel improvements will be used in forthcoming works when we will

dynamically model real massive galaxies.

Chapter 5

Vacuum-cleaning the dusty NGC708: black hole mass, intrinsic shape and orbital structure

5.1 Introduction

NGC708 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 et al., 2005). This stellar envelope probably traces the cluster potential and is typically referred to as ICL. While many BCGs have a Surface Brightness (SB) profile which is better fitted by a double Sersic component rather than a single one, Kluge et al. (2020) report that for this galaxy a single Sersic profile with Sersic index $n = 2.96$ is sufficient.

HST images of the galaxy show the presence of a prominent, edge-on dust lane, extending ∼3" in the southern direction and ∼6" in the northern direction (Wegner et al., 2012). Its origin probably lies in a merger, a commonly observed phenomenon in BCGs. In fact, BCGs lie at the center 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 Sec. 5.2.1).

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 ∼3.1" (∼1 kpc, see Sec. 5.2.1). 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 et al., 1991; Faber et al., 1997; Merritt, 2006b; Milosavljević & Merritt, 2001; Thomas et al., 2014, 2016; Mehrgan et al., 2019). 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 to the SB of the central core itself (Mehrgan et al., 2019) have been discovered. For NGC708, 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.

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 do this, mass-to-light ratios Υ calculated by dynamically modeling these galaxies are compared to estimates from stellar population analyses (Thomas et al., 2003; Maraston, 2005; Maraston & Strömbäck, 2011; Conroy et al., 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 ($\Upsilon_{\rm SSP}$) and dynamicals models (Υ_{*}) for NGC708 are published, finding evidence for a bottom-heavy IMF. However, in that work the authors did not include a central BH in the models.

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 NGC708 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 even larger radii. In this case, an axisymmetric approximation can possibly yield biased black hole mass $M_{\rm BH}$ estimates (van den Bosch & de Zeeuw, 2010) as well as Υ estimates biased by 50% (Thomas et al., 2007).

As shown in Chapter 4 (see also Neureiter et al. 2022, submitted), a new sophisticated triaxial machinery which combines the semi-parametric deprojection algorithm `SHAPE3D` (see Chapter 2), the non-parametric code for kinematics extraction `WINGFIT` (Thomas et al., in prep.) and the triaxial Schwarzschild code `SMART` (Neureiter et al., 2021) was recently tested. The main goal of this chapter is to present the preliminary results of the application of this triaxial machinery to NGC708. A particular source of interest is given by $M_{\rm BH}$, whose value has never been determined. The chapter is structured as follows. Sec. 5.2 describes the photometric data and the deprojections, while Sec. 5.3 describes the spectroscopic data and kinematics extraction. Sec. 5.4 presents the results of the dynamical modeling, which are discussed in Sec. 5.5. Finally, we draw our conclusions in Sec. 5.6.

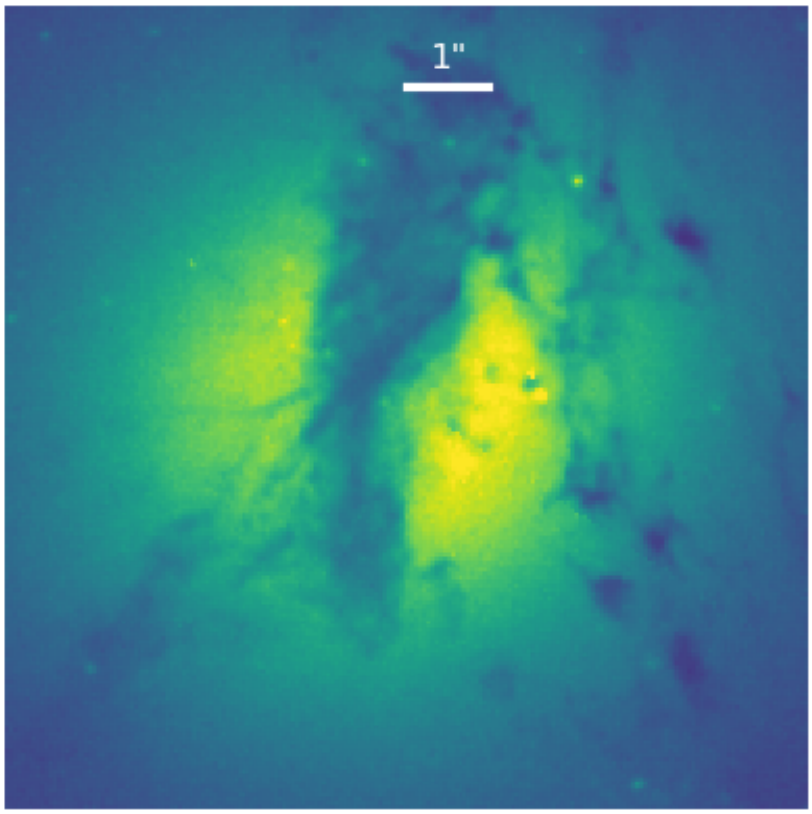


Figure 5.1: HST image of NGC708, BCG of A262, with its prominent dust lane. North is up and East to the left. Image provided by Matthias Kluge (private communication).

5.2 Photometry

The photometry used in this work comes from two different image sources. The first is a g'-band image obtained with the Fraunhofer Telescope at the Wendelstein observatory using the Wendelstein Wide Field Imager (WWFI, Kosyra et al. 2014), with typical seeing of FWHM = 1.2 ± 0.2 and pixel size of 0.2". The total field of view (FOV) of 27.6×28.9' 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 Sec. 3 of Kluge et al. (2020). A second image comes from high-resolution HST observations (see Fig. 5.1), with typical resolution of 0.1". These were carried out using the Wide Field Planetary Camera 2 (WFPC2) using the filter F622W. The camera consists of a grid with dimensions 800×800 pixels, with pixelsize 0.0455", yielding a FOV of 800×800" (Wegner et al., 2012).
The galaxy belongs to a subset of 170 Brightest Cluster Galaxies (BCGs), observed down to a SB $\sim$ 30 mag asec^{-1} (Kluge et al., 2020). This allows to reach the galaxy regions where the interaction with the intra-cluster light (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). Instead, 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 $M_{\rm BH}$ estimates.
Because the two 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 $r < 15$", and WWFI data at $r > 15$". In this way, we have a very high resolution at the centre, which is what we need for a robust estimate of $M_{\rm BH}$, and extending out to large radii ($\geq$ 100 kpc). The resulting photometry is be used to perform the triaxial deprojection.

5.2.1 Isophote features

We use the method of Bender & Möllenhoff (1987) to extract the isophote parameters from the galaxy image. These include the SB value, 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 image sources before combining the isophotes as described above. The resulting isophotes are shown in Fig. 5.2. A few key-points to mention are:

- There are wiggles in both the SB and the ε. These are probably signs of a recent merger;
- The ε and PA profiles show at small radii huge bumps due to the dust lane. This also generates wiggles in the SB profile at small radii;

[1]The camera itself consists of 4 CCD, each one having dimensions 4096×4109 pixels.

- With the exception of the dust-dominated central region, where the dust lane dominates, all variables follow similar trends as those observed in ordinary ETGs (Bender et al., 1988, 1992; Kormendy & Bender, 1996): ε is low at small radii before increasing at large radii, there is only a weak twist $< 10°$ and the Fourier coefficients remain ≤ 1.5;

- The core of the galaxy can be described in terms of the cusp-radius r_γ, defined as the radius where dlogΣ/dlogr = -1/2, where $\Sigma = 10^{-0.4\times SB}$. Adopting this definition, we measure a core of $\sim$ 1 kpc. According to the M_{BH}-core size relation (Thomas et al., 2016), this would imply a central BH with mass $\sim 10^{10} M_\odot$.

- 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", i.e. farther out than the core size.

In practice, we are dealing with a (possibly still ongoing) merger with dust at the galaxy centre. Given that we carry out all our analysis assuming the galaxy to be triaxial (see e.g. Sec. 5.2.2 below), we set ε and the PA to constant values in the innermost regions (see Fig. 5.2).

5.2.2 Deprojection parameters

`SHAPE3D`, the code which we use to deproject the SB profile of NGC708, has been extensively described in Chapters. 2, 3 and 4. Therefore, a detailed presentation of the code features will not be repeated here; we only highlight the selected deprojection parameters.
We place the SB on a circular 40×10 grid, with innermost radius at $\sim$1.02" and outermost radius at $\sim$120.4", 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 bet 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 $RMS_{\log\Sigma} = 0.006$, $RMS_\varepsilon = 0.008$ and $RMS_{PA} = 0.602°$. For the intrinsic density ρ we choose an ellipsoidal $60 \times 11 \times 11$ grid. The flattenings P, Q of this grid are determined for each set of viewing angles calculating the expected values for a perfect ellipsoid (see App.A of de Zeeuw & Franx 1989 or eqs. 2-3 of Cappellari 2002) and averaging these, unless the deprojection is along the principal axes. For example, along the y-axis, we determine Q as described above while for P we try different values.
We sample the viewing angles as described in Chapter 3. From all deprojections we select those for which $RMS \leq 1.2 \times RMS_{min}$, where $RMS_{min} = 0.027$ is the best-fit RMS. The viewing angles that provide the best deprojection within the RMS interval are $(\theta, \phi, \psi) = (70,20,130)°$. The resulting light densities are those we use for the dynamical modeling 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 observed isophotes and those of the recovered ones, finding $RMS_{\log SB} = 0.017$, $RMS_\varepsilon = 0.027$ and $RMS_{PA} = 1.84°$. The isophotal fits to the best-fit SB, superimposed to the observed photometry, are shown in Fig. 5.2.

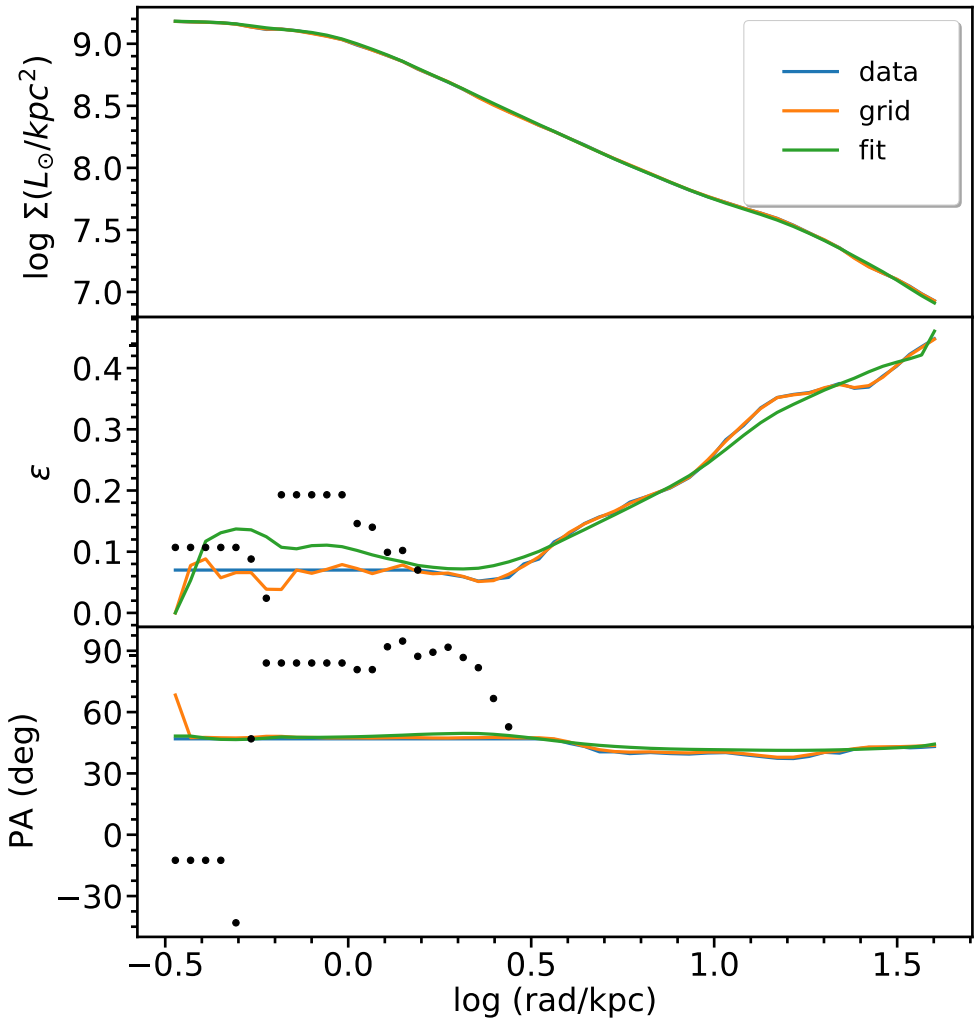


Figure 5.2: Isophotes (from top to bottom: SB, ellipticity, PA) of the galaxy NGC708. The blue lines represent the observed photometry, while the orange lines are computed by placing the SB on the grid and performing isophotal fits. The green lines come from the best-fit SB, found at (θ, ϕ, ψ)= (70,30,120)°. Finally, the black points show the original photometry, which is contaminated by the dust lane.

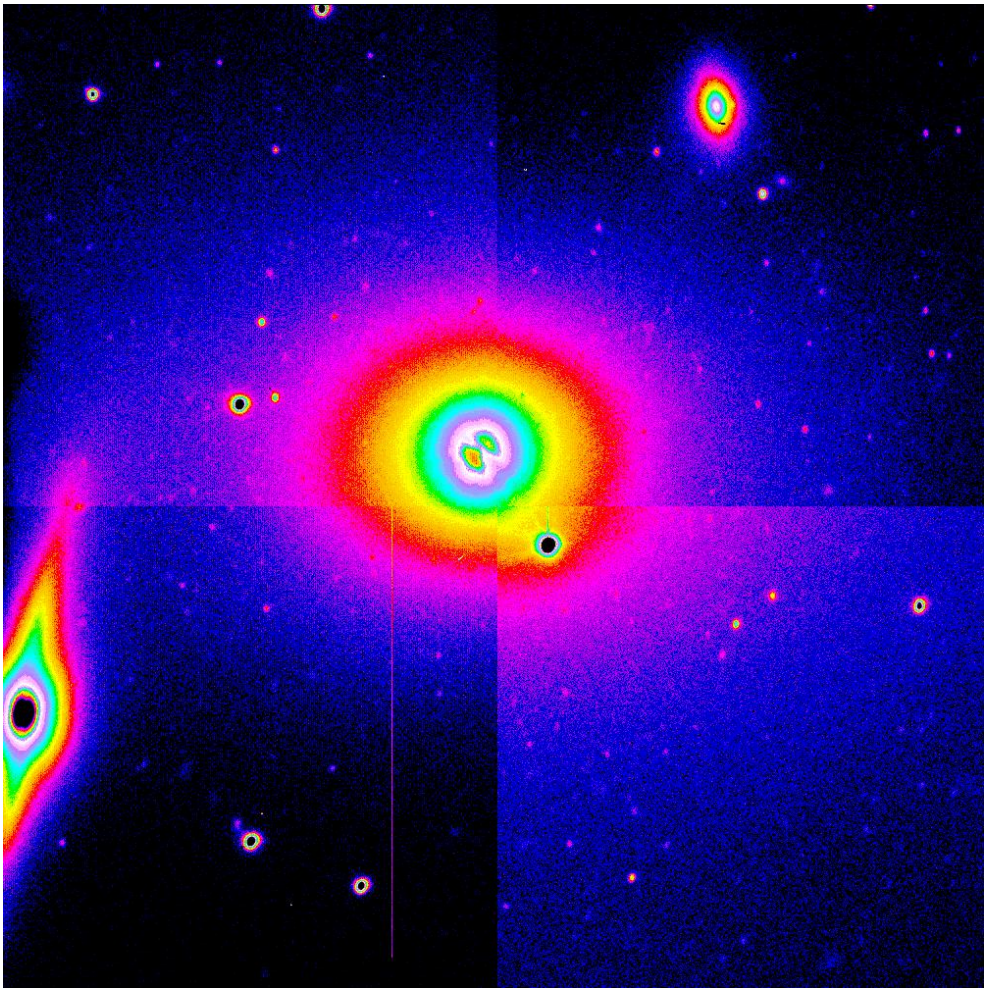


Figure 5.3: LBT acquisition image of NGC708 before starting the spectroscopy. In the center of the galaxy there is a prominent "double center": this effect is generated by the dust lane.

Setup	PA	Center	Number of bins
MAJOR	35	(0,0)	16
MINOR	125	(0.7369, 2.5913)	22
MAJOR+45	80	(0,0)	15
MINOR+45	170	(0.7369, 2.5913)	23

Table 5.1: Technical details about the four slit configurations that we used to observe NGC708. *Col. 1:* Setup name. *Col. 2:* PA of the slit, measured from North to East. *Col. 3:* Center of the slit in a NW frame of reference centered 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).

5.3 Spectroscopy

5.3.1 MODS Observations

Long-slit spectroscopic data for NGC708 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) to acquire the galaxy spectra. 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', works in the range [3200-10000] Å and has spectral resolution $\lambda/\Delta\lambda = 5000/3.2 \sim 1500$ with slit width 0.8" in the blue. All science images were corrected for bias, dark, flat fields and wavelength calibrated. Moreover, we acquired one sky image after each setup to allow for background subtraction. The pixel scale of each image is ~ 0.12 arcsec/pixel.
The galaxy was observed using two different configurations. During the October 2019 run, the two slits, each one with width of 0.8", 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). 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 $\sim$ 1.4" (FWHM). The relevant pieces of information of the four setups for the blue part of the spectrum are reported in Tab. 5.1.
Because of the dust lane, the galaxy shows two nuclei (see Fig. 5.3). While MAJOR and MAJOR+45 were centered on the brightest peak, MINOR and MINOR+45 were centered on the fainter one. In a NW coordinate system centered on the brightest peak, the fainter has coordinates (0.7369, 2.5913) arcseconds.
For each one of the four setups we spatially binned the data (see Tab. 5.1) along the slit. In each spatial bin we extracted the spectrum and measured the kinematics.

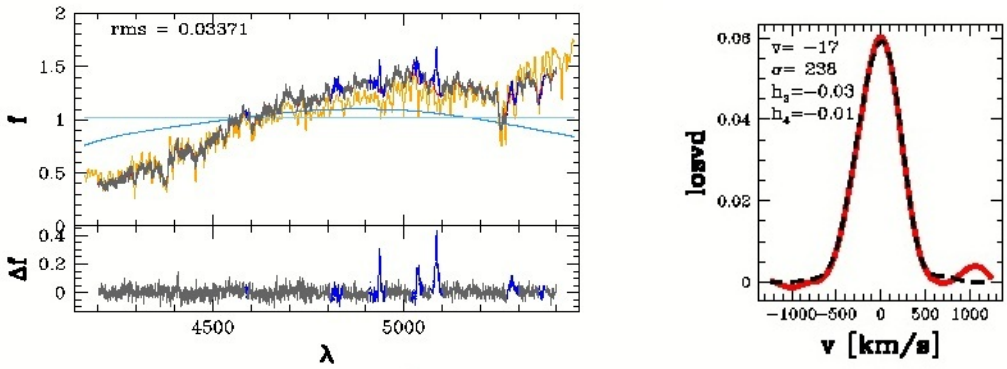


Figure 5.4: *Left:* Example of fitted spectrum by `WINGFIT`. The top panel shows the observed spectrum (in gray), while the yellow line is the stellar template before it is broadened to account for the different resolution the stellar templates have with respect to the galaxy. The lighter and darker cyan lines are the additive and multiplicative polynomials, while the parts of the spectrum highlighted in blue are masked and omitted from the fit. Finally, the bottom panel shows the fit residuals. *Right:* Corresponding non-parametric LOSVD recovered by `WINGFIT` (red line) along with the best fit parametric LOSVD (dashed line), whose coefficients are also printed in the plot window.

5.3.2 Extracting the kinematics

The extracted spectrum has prominent emission lines: in the blue part, [OIII] 5007 Å, [NI] 5199 Å, Hβ are clearly visible. The red part of the spectrum is dominated by emission lines such as Hα, [NII] 6583 Åand [SII] 6730 Å.

To extract the kinematic variables from the spectrum we use the non-parametric LOSVD fitting routine `WINGFIT` (Thomas et al. 2022, in prep.). `WINGFIT` uses the novel model selection technique described in Chapter 4 (Lipka & Thomas, 2021; Thomas & Lipka, 2022) to optimize the smoothing of the LOSVDs.

`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. (2022, submitted), their usage can generate artificially enhanced wings. Therefore, we limited ourselves to 4-th order multiplicative polynomials.
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. 5.4.
The recession velocity of the galaxy barycentre needs to be subtracted from the actual estimate (see Cappellari 2017) to set the zero-point of the 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, is shown in Fig. 5.5. 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 Δv ~100 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 Sec. 5.3.1). 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. Finally, the velocity dispersion σ hits a ceiling at ~250 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 $\sim 10^{10} M_\odot$, the galaxy is a potential catastrophic outlier in the $M_{\rm BH} - \sigma$ relation: using the $M_{\rm BH} - \sigma$ relation of Saglia et al. (2016), we infer a black hole mass smaller than 10^9 M$_\odot$.

5.4 Dynamical modeling

We now put together the results of the previous two sections and turn to the dynamical modeling of NGC708 under the assumption of triaxiality. The galaxy has already been modeled 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 Chapter 3), 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.

5.4.1 Our code

To compute the dynamical models of NGC708 we use our recently developed triaxial Schwarzschild code SMART. This is described in Chapter 4 and, more extensively, in 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).

[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.

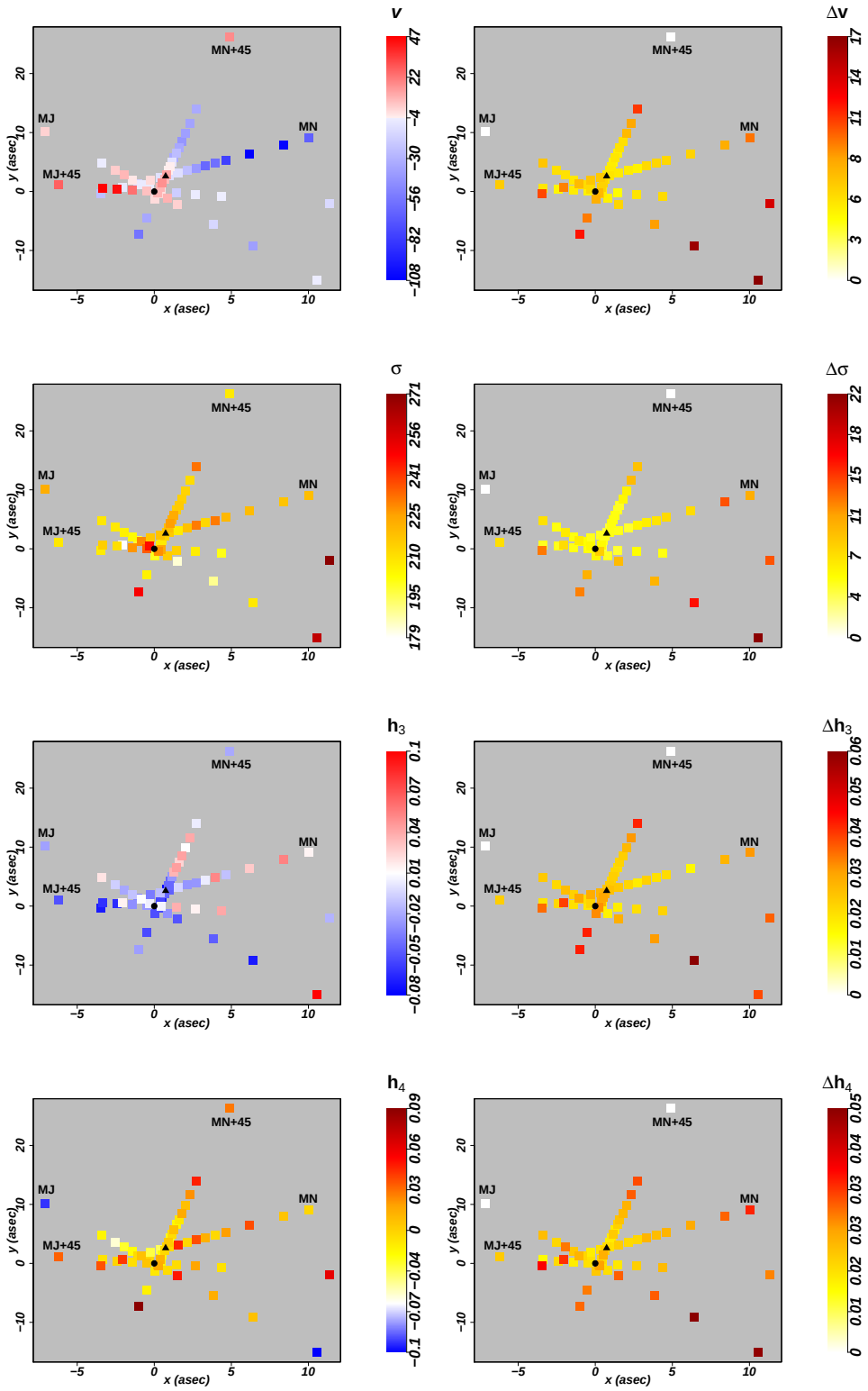


Figure 5.5: Kinematics of NGC708, measured with our code `WINGFIT` using MODS observations. The first column shows the kinematical variables, whereas the second column shows the derived uncertainties. The black points label the slit centers (circle MAJOR and MAJOR+45, triangle MINOR and MINOR+45). The four slits are labeled in the plots. From top to bottom: v, σ, h_3, h_4.

Variable	Range	No. Values	Step size
$M_{\rm BH}$	[4 $\times 10^9$ - 1.2 $\times 10^{10}$] $M_\odot$	5	4 $\times 10^9$ $M_\odot$
Υ	[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
$p_{\rm DM}$	[0.8 - 1.0]	3	0.1
$q_{\rm DM}$	[0.8 - 1.0]	3	0.1

Table 5.2: Sampling the 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.

5.4.2 Modeling strategy

The parameters needed to build the trial gravitational potential (defined in eq. 4.3) are: $M_{\rm BH}$, Υ , 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 $p_{\rm DM}$ and $q_{\rm DM}$. 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 β = 4.5. The value of β differs from the canonical value of 3 of the gNFW models: as we did in Chapter 4, we assume that the halo progenitors can be modeled 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 modeling strategy can be summarized as follows:

1. We fix the viewing angles at best-fit orientation (θ, ϕ, ψ) = (70,20,130)° and launch a NOMAD run. The other 6 parameters are (log-)linearly sampled as reported in Tab. 5.2. With the exception of $M_{\rm BH}$, 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 of 1.4", measured from our spectroscopic observations. Given that the size of the central core (∼3.1") is expected to approximate the sphere-of-influence radius $r_{\rm SOI}$ well and given that we model with a DM halo, we conclude that we can reliably estimate $M_{\rm BH}$ (see discussion in Rusli et al. 2013a). We verify this assumption below.

2. 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 Sec. 5.4.3).

3. 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.

In Chapter 4 we have shown that by selecting deprojections using the approach described in Sec. 5.2.2 and computing the average shape profiles $p(r)$ and $q(r)$ among all these, the resulting

profiles approximate the correct ones very well. The modeling 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.

5.4.3 Results

The shape of the best-fit density that comes out from the second NOMAD run is shown in Fig. 5.6. In the very central regions, the density is triaxial, then becomes slightly oblate before becoming triaxial again from 2 kpc. The viewing angles for this density are $(\theta, \phi, \psi) = (80, 90, 135)°$: for this orientation, close to the y-axis, we sampled three different deprojections using as starting values $p(r) = 1.0, 0.95, 0.9$ at all radii[3]. The last value provided the best-fit density. This agrees well with the findings of Chapter 3: the $p(r)$, $q(r)$, $T(r)$ intervals shown in Fig. 3.C.1 overlap well with those found in this work. It should be noted that the best-fit solution does not appear in Chapter 3 because there only one deprojection[4] per orientation was tested.
The results of the third and final NOMAD run are shown in Fig. 5.7, for a total of 1617 models. For each variable, we marginalize the 6-dimensional $\mathrm{AIC_p}$ distribution to recover the six one-dimensional functions $\mathrm{AIC_p}(M_{\mathrm{BH}})$, $\mathrm{AIC_p}(\Upsilon)$, $\mathrm{AIC_p}(\rho_0)$, $\mathrm{AIC_p}(\gamma)$, $\mathrm{AIC_p}(\mathrm{p_{DM}}$ and $\mathrm{AIC_p}(\mathrm{q_{DM}})$. The fit to the kinematics is shown in the left panel of Fig. 5.8, where we plot the residuals between the input kinematics and the fitted values by SMART. Here, we see that the code is able to recover all four kinematic variables well[5]; in particular, the minor-axis rotation is well reproduced. Moreover, in the right panel of Fig. 5.8 we show an example of a fit to an individual LOSVD, showing that SMART reproduces it very well: in this case, we have $\chi^2/\mathrm{N}_{data} = 0.73$.
The most relevant result is the detection of a $10^{10}\mathrm{M}_\odot$ SMBH in the center of NGC708. We calculate the size of its SOI in three ways: using the central velocity dispersion σ_0 as $\mathrm{r_{SOI}} = GM_{\mathrm{BH}}/\sigma_0^2$ and using the stellar mass derived from our models as $M_{\mathrm{tot}}(\mathrm{r_{SOI}}) = 2\,M_{\mathrm{BH}}$ (Merritt, 2004) and as $M_{\mathrm{tot}}(\mathrm{r_{SOI}}) = M_{\mathrm{BH}}$ (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 $\mathrm{r_{SOI}} = 0.73$, 1.83 and 1.21 kpc, respectively[6]. The strong inconsistency points out that the SIS model does not work well for this galaxy.
The derived $\Upsilon_* = 3.4$ agrees with the results of Wegner et al. (2012), who found 4.17 ± 1.05 in the Kron-Cousins R-band for a Kroupa (2001) IMF using SSP models (Maraston, 2005). The DM halo has a very high normlization $\rho_0 = 10^{8.06}\ \mathrm{M}_\odot/\mathrm{kpc}^3$ 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 ($\mathrm{p_{DM}} = 1.0$, $\mathrm{q_{DM}} = 0.8$), while the light density has $T(r) > 0.8$ only in the central core.

[3]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.

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

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

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

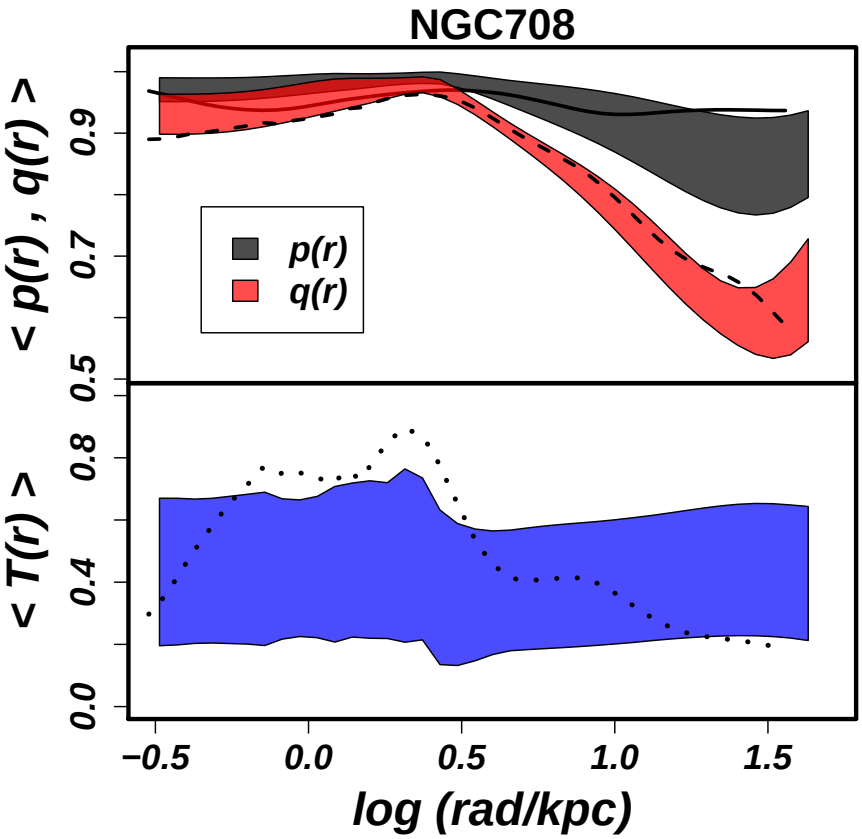


Figure 5.6: *Top:* The recovered intrinsic shape (solid line $p(r)$, dashed line $q(r)$) for the best-fit solution, found at $(\theta, \phi, \psi) = (80, 90, 135)^\circ$, superimposed to the intervals plotted in Fig. 3.C.1. This best-fit solution was not included in the analysis presented in Chapter. 3. $p(r)$ cannot be fitted well due to the unfavourable orientation; $q(r)$ becomes quite low (< 0.6) at large radii. *Bottom:* Same for the triaxiality parameter $T(r)$. In the very central regions the galaxy is triaxial, then becomes oblate and then from $r \gtrsim 2.05$ kpc becomes increasingly triaxial.

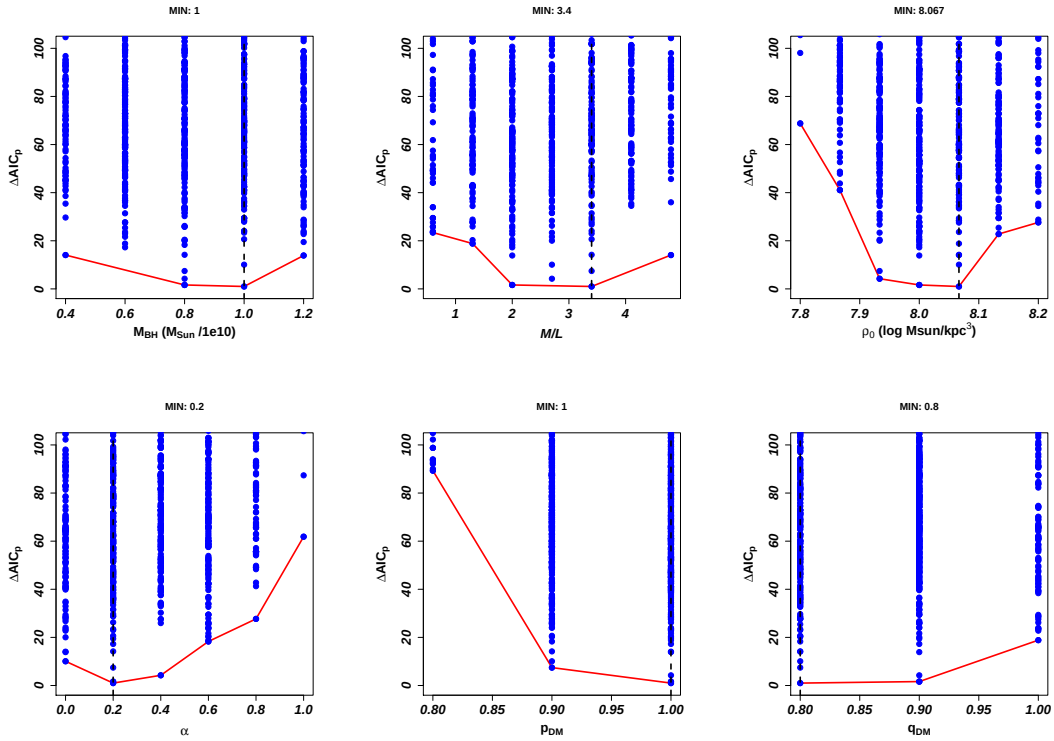
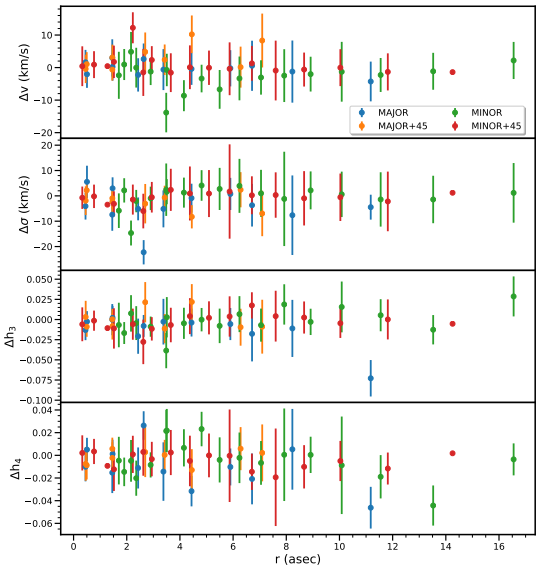


Figure 5.7: AIC_p values, calculated following eq. 4.5, plotted against the 6 variables fitted in our final NOMAD run. The blue points are the individual models. The majority of them is not shown as it falls outside the plotted range. The black dashed line labels the best-fit value, while the red line follows the best-fit models for each tested value. *Left to right, top to bottom:* $M_{\rm BH}$, Υ , ρ_0, γ, $p_{\rm DM}$, $q_{\rm DM}$.



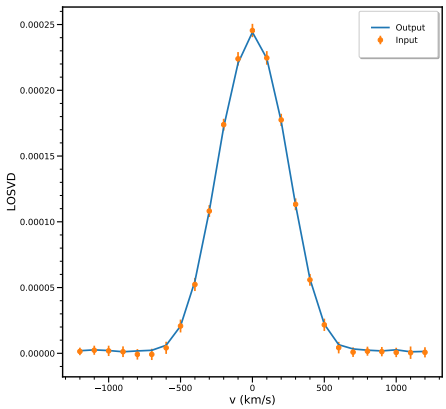


Figure 5.8: *Top:* Residuals between the input kinematics for our triaxial Schwarzschild models and the modeled parameters. Different colors are used for the four slits. *Bottom:* Example of fit to a single LOSVD. The orange points are the input velocity bins, while the light blue line is the recovered LOSVD.

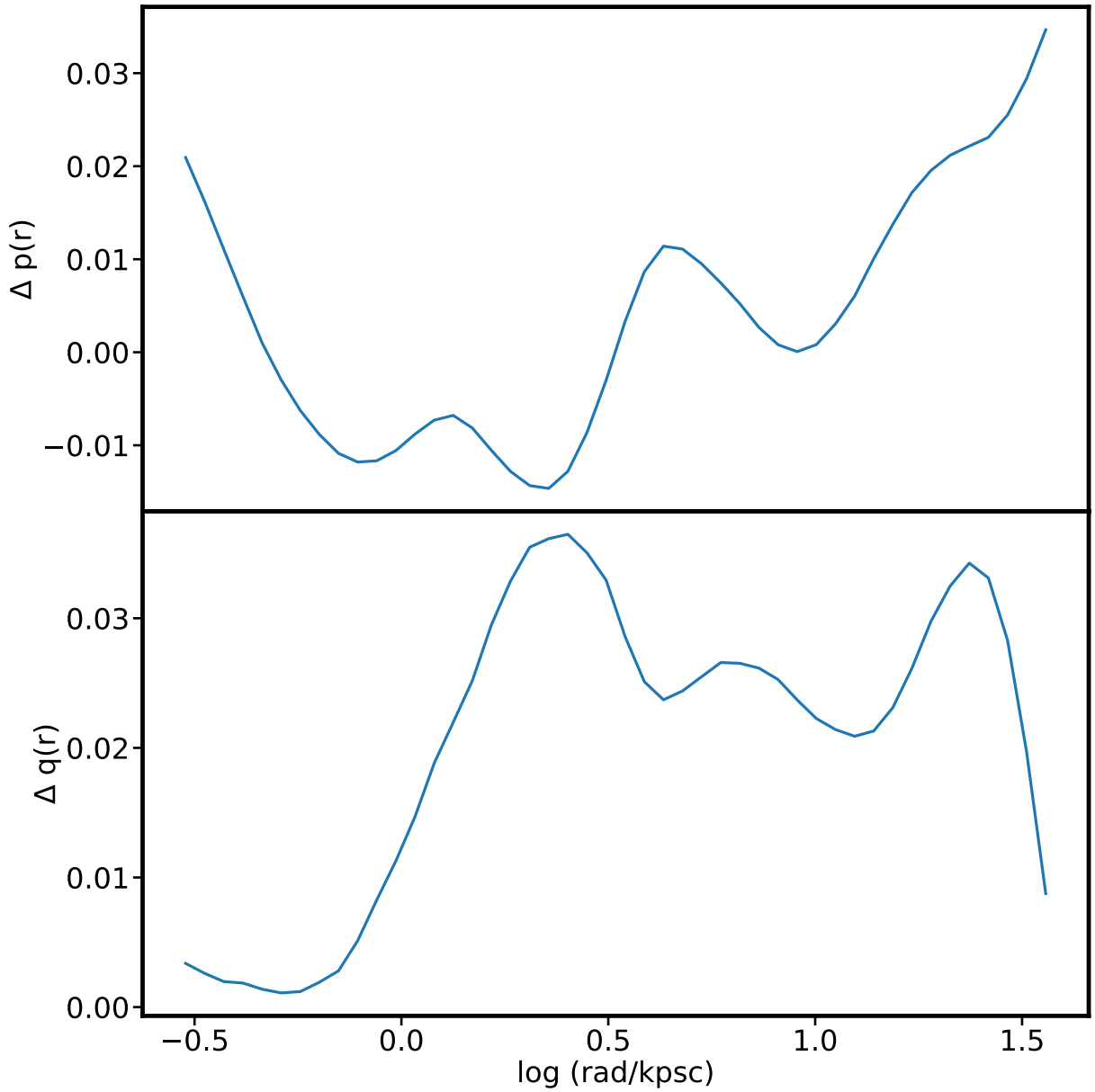


Figure 5.9: Differences between the intrinsic shape profiles $p(r)$ (top) and $q(r)$ (bottom) for the best-fit light density ρ found by NOMAD and the average shape profiles $< p(r) >$ and $< q(r) >$ among all deprojections selected for the dynamical modeling. The very small residuals confirm the findings of Chapter 4: photometry alone suffices to constrain the intrinsic shape of a massive, triaxial ETG.

5.5 Discussion

We have presented the preliminary results of the dynamical modeling of NGC708, BGC of A262. This is the first BCG which we modeled using our full non-parametric triaxial pipeline (SHAPE3D + WINGFIT + SMART): another work taking on dynamical modeling of NGC5419, a massive ETG, is Neureiter et al. (2022, in prep.). In what follows we discuss our results in more detail.

5.5.1 Intrinsic shape

As explained in Chapter 4, 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[7] which we select for the dynamical modeling and look for the deprojection which is closest to it. Instead, the dynamical modeling is needed to reduce the scatter of this estimate.
Here, we tested this approach with a real galaxy for the first time: 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 $(\theta, \phi, \psi) = (80, 90, 135)^\circ$, is indeed remarkably close to the average $< p(r) >$, $< q(r) >$ profiles (see Fig. 5.9). 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. 5.2, reaching $q(r) < 0.6$ at large radii. The corresponding triaxiality profile can be divided into three regions. In the very innermost radii the galaxy is triaxial, but the triaxiality parameter is extremely sensitive to p and q variations if these are close to 1. Then, the galaxy becomes increasingly oblate, reaching the maximum value $T_{max} = 0.95$ at $r = 2.04$ kpc. From this point, the triaxiality parameter decreases quickly, reaching the maximum triaxiality $T = 0.5$ at ~10 kpc. Averaging over all radii, we find $< T >= (1- < p(r) >^2)/(1- < q(r) >^2) = 0.455$. The two main conclusions here are (i) that only one density may be used for the dynamical models and (ii) that NGC708 is triaxial.

5.5.2 The Black Hole: scaling relations, anisotropy and orbital structure

Postponing the discussion of error bars to our final publication, perhaps the most important finding of this chapter is the discovery of a 10^{10} $M_\odot$ SMBH. Even using the smallest r_{SOI} value among the three we report in Sec. 5.4.3, 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 robust. 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 $< 10^{10}$ $M_\odot$. Measurements in this region include NGC4889 ($M_{BH} = 2.1 \times 10^{10}\, M_\odot$, McConnell et al. 2012b) and NGC1600 ($M_{BH} = 1.7 \times 10^{10}\, M_\odot$, Thomas et al. 2016). At the very high-mass end, the scaling relation between SMBHs and the velocity dispersion of the host

[7]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.

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 et al., 2009) and, hence, generate $M_{\rm BH}$ values which are higher than the prediction of the canonical $M_{\rm BH} - \sigma$ relation, for which typically $M_{\rm BH} \propto \sigma^{5.2\div5.4}$ (Saglia et al., 2016; van den Bosch, 2016). This is expected: SMBHs correlate with the bulge parameters (Saglia et al., 2016; de Nicola et al., 2019), which are 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 $M_{\rm BH} - \sigma$ we assume, NGC708 is an outlier. The coefficients found by Saglia et al. (2016) omitting pseudobulges[8] predict $M_{\rm BH} = 5.75 \times 10^8$ M$_\odot$, making the galaxy an outlier *by almost a factor of 20*. We note that NGC708 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 $M_{\rm BH}$-M$_{\rm bul}$ relation (Magorrian et al., 1998). As estimate of M$_{\rm bul}$ 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$_{\rm bul} = 2.8 \times 10^{11}M_\odot$. 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 et al. (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].
The commonly proposed formation mechanism for these central cores is the gravitational slingshot caused by SMBHs lying at the center 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; Thomas et al. 2014), and in dry mergers there is no gas that can replenish the center. Therefore, scaling relations linking $M_{\rm BH}$ 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$_\gamma$ (Rusli et al., 2013b; Thomas et al., 2016) and with the SB of the core itself (Mehrgan et al., 2019). Possibly NGC708 is a mild outlier also in the $M_{\rm BH}$-core galaxies scaling relations: the coefficients of Thomas et al. (2016) for the $M_{\rm BH}$-r$_\gamma$ and those of Mehrgan et al. (2019) for the $M_{\rm BH}$-SB relations[10] predict a BH with mass $\sim 1.5 \times 10^{10}$.
The core scouring mechanism generates a tangential anisotropy ($\beta < 0$, eq. 1.26) in the central regions, because stars on radial orbits come closer to the SMBH and are more likely to be ejected. Radial anisotropy is expected at larger radii (Milosavljević & Merritt, 2001). In Fig. 5.10 we plot the anisotropy profile $\beta(r)$ for the best-fit model. Indeed, the profile shows the characteristics described above, although less pronounced than in other core-galaxies. In the future we will compare the derived $\beta(r)$-profile with simulations which reproduce the formation of cores in ETGs (Rantala et al., 2019; Frigo et al., 2021).

[8]Pseudobulges correlate weakly - if at all - with $M_{\rm BH}$ (Kormendy & Kennicutt, 2004; Kormendy & Ho, 2013) because their BHs are still accreting mass through infalling gas. In fact, these systems often host bars (Erwin et al., 2015b).

[9]Note that all scaling relations have intrinsic scatters. Using the (large) intrinsic scatter of the $M_{\rm BH}$-M$_{\rm bul}$ relation $\epsilon = 0.61$ from Bogdán et al. (2018), NGC708 remains an outlier by a factor of 2.6.

[10]This relation was derived in the V-band, which is close to the g'-band photometry used in this work nonetheless.

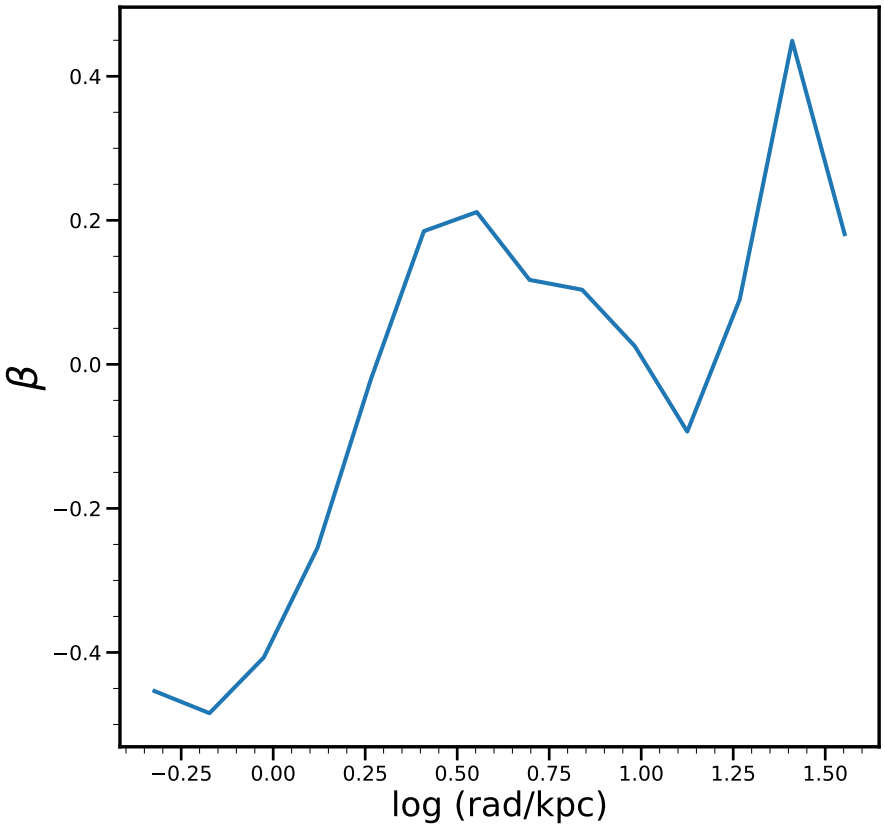


Figure 5.10: The recovered anisotropy of NGC708 for the best-fit model. We find tangential anisotropy in the central regions. This is a fingerprint of BH core scouring and therefore provides evidence for past merger(s). The fact that the tangential bias is not so pronounced as in most core-galaxy might indicate that the progenitors of NGC708 were core-galaxies themselves. At larger radii, the galaxy shows a small radial anisotropy.

5.6 Summary and conclusions

We have produced dynamical models of NGC708, 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 extraction `WINGFIT` and our triaxial Schwarzschild code `SMART`.
The galaxy shows several interesting features. We detect a SMBH with mass 10^{10} $M_{\odot}$, one of the few measurements in this mass range, which makes the galaxy a strong outlier in the $M_{\rm BH} - \sigma$ and $M_{\rm BH} - M_{\rm bul}$ relations. 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 has been observed for Holm15A, which is also an outlier in the scaling relations.
The intrinsic shape of the galaxy is, on average, triaxial. We find that the galaxy is observed close to the intermediate axis. In particular, the galaxy is consistent with an oblate geometry in the central regions, but becomes increasingly triaxial after ~2 kpc. The triaxiality could be already be inferred given the rotation along the minor axis in the kinematics, but no strong clues about it are visible in the photometry, stressing the need of using triaxial models even if the photometry might be compatible with an axisymmetric geometry.
The results presented in this chapter are still somewhat preliminary and need further improvements, before being presented in the final publication. Dust effects need to be taken into account for a robust estimation of r_{γ}. Test on simulated data taking into account the slit setups used here need to be performed, similarly to what described in Chapter 4, to assess the uncertainties on the mass parameters. A discussion of the orbit distribution and mass structure implied by our best-fit modeling is required, including a comparison to simulated dry mergers of cored ellipticals.

Chapter 6

Summary and conclusions

This thesis focuses on the reconstruction of galaxy shapes in the case of triaxial geometry and its applications both as stand-alone method and in combination with dynamical modeling of a galaxy. To achieve this, a robust deprojection algorithm that is required. Such an algorithm should be able to recover the correct intrinsic density when the viewing angles are known, explore the range of light distributions whose projection along the LOS vanishes, making them invisible but modifying the intrinsic structure of a galaxy (Rybicki, 1987; Gerhard, 1996; van den Bosch, 1997; Magorrian, 1999; de Nicola et al., 2020) and reduce the number of possible orientations compatible with the photometry. The widely used parametric code MGE (Cappellari, 2002) provides only one solution per orientation and, for every allowed orientation, delivers a recovered SB with always the same error, thus not allowing for a further restriction of the range of viewing angles. In Chapter 2 I presented two new codes to deproject SB profiles: the first one non-parametrically and the second one (`SHAPE3D`) assuming the galaxy to be stratified on concentric ellipsoids but recovering the shape profiles non-parametrically. The fully non-parametric code is only the second code of this kind (see Bissantz & Gerhard 2002) and successfully deprojects every possible SB profile obtained by projecting a triaxial galaxy. Unfortunately, this approach does not well suit our needs in practice because:

- Degeneracy cannot be kept under control: the correct intrinsic light distribution cannot be recovered even if the viewing angles are known;
- A solution for every possible orientation is found: given that dynamical models are very time consuming, sampling a large number of different light distribution would be unfeasible. By the time this thesis was written, the common belief was that photometric information alone could not help in constraining the viewing angles.

Despite being less general, the second approach proved to be the winning choice and has been used to derive every result presented in this thesis. In fact, `SHAPE3D`is able to:

- Recover the true intrinsic density if the viewing angles are known. In de Nicola et al. (2020), this is shown both analytically and through several examples using toy models as well as an N-body simulation which reproduces the real massive ETG NGC1600 (Rantala et al., 2018);

- Find different solutions at the *same* orientation, which allows to probe the degeneracy;
- Significantly restrict the range of possible orientations using on photometric information only. This is done using an RMS cut-off, where the RMS is used to take into account the differences between the observed and the modeled SB at each point of the grid.

A smaller number of viewing angles that need to be tested when computing the dynamical models reduces the computation time hugely. In Chapter 4 I show that *only one* deprojection may be enough for the dynamical modeling.

Given that the true intrinsic density can be derived, in the worst case, with an RMS of 10% for the N-body simulation, and that the orientation is very well constrained by photometry only, in Chapter 3 I exploit these findings and use my code to deproject a sample of 56 SB profiles of BCGs, presenting *radially resolved* measurements of intrinsic shape profiles *based on photometry alone*. The data come from Wendelstein observations: they allow us to probe the outermost galaxy regions, where the effects of ICL take over and where the DM contribution is significant. For most BCGs high-resolution photometry (either HST or LBT) in the central regions is also available, in view of future dynamical BH mass measurements. My deprojection code successfully reproduces the statistics of observed ellipticites and position angles, a robust check of the goodness of the deprojection. The number of orientations typically decreases by a factor of 3 for the roundest galaxies up to more than 100 for the flattest galaxies when the RMS cut-off is applied. I derive an estimate of the best-fit intrinsic shape profiles by averaging over all acceptable deprojections, showing that BCGs are *extremely* triaxial, even more than ordinary ETGs. In particular, the higher triaxiality is due to a lower $q \equiv c/a$ value, while $p \equiv b/a$ agrees with ETGs (see Vincent & Ryden 2005; Ene et al. 2018). The axisymmetric approximation used on a triaxial galaxy can lead to substantial biased BH mass (van den Bosch & de Zeeuw, 2010) and Υ gradient (Thomas et al., 2007, 2014) estimates, and should be avoided in these cases. Finally, cosmological simulations such as IllustrisTNG (Nelson et al., 2018; Pillepich et al., 2018; Springel et al., 2018; Marinacci et al., 2018) and Magneticum (Hirschmann et al., 2014; Teklu et al., 2015) aim at reproducing the formation and the evolution history of galaxies and can be used to compare simulated galaxy and DM halo intrinsic shapes with those recovered by my code. I show that simulations do not reproduce the correct flattenings in the central regions well, systematically producing too flat objects, while in the galaxy outskirts the agreement is much better. In particular, DM shape profiles agree very well at ~100 kpc, thus implying that by deprojecting BCG SB profiles, one can make inferences about the halo shapes. DM simulations probing different halo models (e.g. ΛCDM and fuzzy DM with different particle cross sections, Robertson et al. 2019; Fischer et al. 2022) show that the halo shape depends on the DM model one assumes, meaning that we could probe the physics behind the halo formation.

The triaxiality of the BCGs emphasizes the need for triaxial dynamical models: therefore, in Chapter 4 I combine my deprojection code with our new triaxial Schwarzschild code SMART (Neureiter et al., 2021), to investigate how well we can constrain the shape and the orbit distribution with a fully non-parametric triaxial machinery. The kinematic data needed for the dynamical

models are fitted non-parametrically with our code WINGFIT (Thomas et al. 2022, in prep.) exploiting the full information contained in the LOSVD. The non-parametric LOSVDs and light densities are then used to construct dynamical models. Both SMART and WINGFIT use a novel model selection technique (Lipka & Thomas, 2021) which automatically estimates the best model smoothing and avoids biases linked to the different number of DOFs each model has. This is efficiently implemented as described in Thomas & Lipka (2022).
I consider the same N-body simulation (Rantala et al., 2018) used in Chapter 2, projected at four different orientations: two along the principal axes, one exactly in the middle and one at randomly drawn viewing angles. I deproject the galaxy and reduce the number of light densities that need to be sampled by more than a factor of ∼30. At the correct viewing angles, the intrinsic density is recovered with ∼15% accuracy along the principal axes, ∼10% otherwise. To estimate how well the deprojection *alone* recovers the galaxy intrinsic shape, I calculate the average shape profiles among all good deprojections, before moving on to dynamical modeling the galaxy.
The results of the dynamical modelling point out that:

- The true viewing angles are recovered with a typical accuracy of 15°;
- If I consider the average differences between the true and the recovered shape among all orientations, I barely find any improvements when the dynamical modeling is used to further constrain the galaxy orientation. Given that the differences are always ≤ 0.1, I conclude that photometry alone suffices in recovering the correct intrinsic shape of a galaxy;
- Even if it does not improve the (already very low) *bias*, the dynamical modelling reduces the *scatter*. In fact, several more orientations can be ruled out adding kinematic information rather than with photometric information only;
- The dynamical modeling also allows to calculate the anisotropy profile of the galaxy, and thus the orbit distribution. I find that the anisotropy profile $\beta(r)$ is recovered within 10%, confirming the findings of Neureiter et al. (2021) where no deprojected light density was used. Particularly relevant is that the tangentially-biased anisotropy found in the galaxy innermost regions is very well recovered.

These results are valid regardless of the chosen projection. The same accuracy can be reached when estimating the mass parameters and the mass distribution (Neureiter et al. 2022, submitted).

A final application consists in tacking on the dynamical modeling of a real BCG. This is what I do in Chapter 5, where I present the preliminary results of the triaxial modeling of the BCG NGC708. I show that even in case of a particularly challenging galaxy, showing signs of an ongoing merger as well as emission lines in its spectrum, it is possible to robustly estimate $M_{\rm BH}$, Υ and the orbit distribution.

This work explores in detail a fully triaxial dynamical modeling algorithm, focusing on the photometrical aspect. One of the most important findings of this work is how much information we can get by simply deprojecting galaxies. This is often underestimated. Degeneracy needs to

be dealt with and reduced, but at the same time an algorithm delivering only one solution per orientation is not ideal. My code can deal with both issues: the ellipsoidal approximation allows to recover the correct intrinsic density, but the non-parametric fit to the shape profiles makes it suitable to explore more than one light profile compatible with the observed photometry for a given orientation. With the aid of deep-photometry data, this work directly shows that BCGs are extremely triaxial and trace the shape of the underlying DM halos, allowing to investigate the physics behind halo evolution. Moreover, photometric data can recover the galaxy shape with $\Delta p, \Delta q \leq 0.1$. All this can be achieved with photometry only, provided that the deprojection is accurately performed.
Nevertheless, the importance of accurate dynamical models must not be underestimated. This is another very important conclusion of this work: the accuracy of triaxial models of which SHAPE3D is an integral part is far higher than previously thought and published by e.g. van den Bosch et al. (2008); Jin et al. (2019); Vasiliev & Valluri (2020). A recovery of the correct orbital distribution is possible with $\Delta\beta < 0.1$, as well as bringing the scatter of the estimated galaxy intrinsic shape down to 0.05. The ingredients are a deprojection technique able to recover the right galaxy shape and restrict the number of orientations compatible with a given photometry, a non-parametric LOSVD fitting routine to exploit the full information contained in the LOSVD, an accurate triaxial dynamical modeling algorithm able to sample a 5D orbital space and a model selection technique which allows to find the best, *unbiased* solution by taking into account the different number of free parameters each model has.

Given the promising results which I obtained, this new technique will be systematically applied in the future to massive ETGs with the following goals:

- Derive unbiased triaxial $M_{\rm BH}$ estimations. This is particularly relevant to reduce the scatter of BH-host galaxies scaling relations. In particular, it is possible with the aid of high-resolution photometry to derive the size and SB of core radii, which are shown to strongly correlate with the central BH.
- Derive unbiased Υ estimations to investigate how much the IMF varies among massive ETGs and to reduce the bias with estimations from stellar population models;
- Study the distribution of different orbit families in triaxial galaxies.
- Finally, the uniqueness of the deprojection along with the fact that the intrinsic shapes of individual galaxies can be measured using photometry alone opens the window for many new applications. In particular, individual deprojections of large galaxy sample can be performed.

Appendix A

Code specifications

SHAPE3D is a Python3 (Van Rossum & Drake, 2009) code written exploiting to the object-oriented paradigm. In Python3, this is particularly convenient, given that anything defined within the code (variables, methods, arguments, etc.) is treated as an object, and can be conveniently interchanged between the various classes making up the code. It uses standard Python libraries such as NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), AstroPy (Astropy Collaboration et al., 2013, 2018) and Matplotlib (Hunter, 2007), making it very easy to use on a nowadays commonly used processor. The newest version of each library with which the code has successfully been tested is 1.21.2 for NumPy, 1.7.0 for Scipy, 4.2.0 for Astropy and 3.5.1 for Matplotlib, with `Python 3.8` as software version.
Python's flexibility allows for different kinds of supported input files. In case of the deprojection, SHAPE3D is able to deal with both isophote tables, regardless of the file format, and with FITS file containing galaxy images. Output files are typically in `.txt` or `.dat` format for easy visual inspection, while images are saved in `.png` format (even if I typically use the `.pdf` format for publication-quality images). An exception is given by the best-fit intrinsic density and its corresponding SB, which are stored as `.npy` binary files. These can be easily opened using the `numpy.load` function of NumPy, and allow to save and read arrays of arbitrary dimensionality.
There are four classes making up the code:

- `Ellipsoidal_triaxial_depro`: this is the most important class of the whole code. It is the only one that needs to be instanced to run a deprojection, since it inherits the methods from all other classes. The methods defined within the class set up the deprojection by computing the starting guess for the intrinsic density and perform it using the Metropolis algorithm described in chapter 2. The penalty function are also defined here. Its `__init__` method reads in the input, places it on the grid and perform the deprojection according to the user-specified parameters. Finally, it saves the results and takes care of generating the plots.

- `Proj`: this class can be used to project a certain intrinsic density at a given orientation, along with evaluating the goodness of fit with a target SB, possibly convolved with a PSF. This class is never instanced in the code, with `Ellipsoidal_triaxial_depro` simply using its methods.

- `Sbgrid`: this class takes care of transforming the SB from the isophotal table or `FITS` file to the grid-based one. Therefore, it gets instanced immediately by `Ellipsoidal_triaxial_depro`. I typically use this to check how the SB on the grid looks before starting the deprojection.
- `Utils`: This class contains a bunch of functions used for several purposes (e.g. fitting, interpolations, integration, generating plots). Its methods are inherited by all other classes of the code and can also be used for diagnostics.

Each class instance can be used as stand-alone to perform the tasks described above by supplying the required methods.
In addition, I also developed a `Nonparametric_triaxial_depro` class, which can be used in combination with the other classes exactly as
`Ellipsoidal_triaxial_depro`. The only difference lies in the Metropolis algorithm, since in this case the deprojection is fully non-parametric.
Finally, a `multiple_depro_analysis.py` file contains a pipeline which I use to generate useful plots when dealing with a large number of deprojections at the same time. This is the case when deprojecting a real galaxy, for which several orientations must be tested.

Appendix B

Theorem proofs

B.1 Proof of the Fourier Slice Theorem

This section shows the mathematical proof of the Fourier Slice Theorem, which was introduced in Sec. 1.2.2. The theorem states that given a certain function, say $f(\boldsymbol{x})$, and its Fourier transform $\hat{f}(\boldsymbol{k})$, then the projection of $f(\boldsymbol{x})$ along the LOS has as Fourier transform the slice of $\hat{f}(\boldsymbol{k})$ along the projection line. In other words, the following two operations are equivalent:

- Project $f(\boldsymbol{x})$ along the LOS and Fourier-transform its projection;
- Fourier-transform $f(\boldsymbol{x})$ and then slice the resulting $\hat{f}(\boldsymbol{k})$ along the projection line.

Proof. The projection of a n-dimensional function $f(\boldsymbol{x})$ on a m-dimensional linear sub-manifold is

$$p(x_1, .., x_m) = \int f(\boldsymbol{x})\, d^{n-m}\boldsymbol{x} \tag{B.1}$$

with $m < n$. Instead, the Fourier transform of f reads

$$\hat{f}(\boldsymbol{k}) = \int f(\boldsymbol{x}) e^{-2\pi i(\boldsymbol{k}\cdot\boldsymbol{x})} d\boldsymbol{x}. \tag{B.2}$$

Thus, the slice of $\hat{f}(\boldsymbol{k})$ parallel to the projection sub-manifold is

$$\begin{aligned} s(k_1, .., k_m) = \hat{f}(k_1, .., k_m, 0, .., 0) &= \int f(\boldsymbol{x}) e^{-2\pi i \sum_i^m k_i x_i} d\boldsymbol{x} \\ &= \int \left[\int f(\boldsymbol{x}) d^{n-m}\boldsymbol{x}\right] e^{-2\pi i \sum_i^m k_i x_i} d^m\boldsymbol{x} \\ &= \int p(x_1, .., x_m) e^{-2\pi i \sum_i^m k_i x_i} d^m\boldsymbol{x} \end{aligned} \tag{B.3}$$

where in the last passage we have used eq. B.1. Thus, $s(k_1, .., k_m)$ is the Fourier transform of $p(x_1, .., x_m)$, which proves the theorem.

B.2 Proof of the Jeans Theorem

This section reports the proof of the Jeans Theorem, according to which any steady-state solution of the CBE depends on the phase-space coordinates only through integrals of motion, and any function of the integrals of motion is a steady-state solution of the CBE.

Proof. If a distribution function f solves the CBE, then it must be an integral of motion given the form of eq. 1.21, which proves the first part of the theorem. Instead, let us assume $F_1, F_2, ..., F_n$ be n integral of motions. Then, if f is a function of these n integrals of motions, we can write

$$\frac{d}{dt} f\left[F_1(\mathbf{x},\mathbf{v}), F_2(\mathbf{x},\mathbf{v}), ..., F_n(\mathbf{x},\mathbf{v}),\right] = \sum_{i=1}^{n} \frac{\partial f}{\partial F_i}\frac{dF_i}{dt} = 0 \tag{B.4}$$

where in the first equality we have exploited the fact that F_i is an integral of motion and therefore its derivatives with respect to $\mathbf{x}, \mathbf{v}$ vanish. This concludes the proof of the theorem.

Appendix C

Shape of an ideal ellipsoid

This appendix contains formulae linking projected and intrinsic variables and are valid for an ellipsoidal body. As already assumed in the main body of the thesis, the ellipsoid has axis lengths $a \geq b \geq c$. The axis ratios are $p \equiv b/a$ and $q \equiv c/a$ and the ellipsoid has an orientation specified by the viewing angles (θ, ϕ, ψ) defined in Sec. 1.2.1, with projection equations defined in eqs. 2.12 & 2.13. The projection of an ellipsoidal body yields elliptic contours, whose axis ratio $q' \equiv 1 - \varepsilon$ depends on p, q and on the angles (θ, ϕ):

$$q' = \sqrt{\frac{A + C - \sqrt{(A-C)^2 + B^2}}{A + C + \sqrt{(A-C)^2 + B^2}}} \tag{C.1}$$

where

$$\begin{aligned} A &= \frac{\cos^2\theta}{q^2}\left(\sin^2\phi + \frac{\cos^2\phi}{p^2}\right) + \frac{\sin^2\theta}{p^2} \\ B &= \cos\theta \sin 2\phi \left(1 - \frac{1}{p^2}\right)\frac{1}{q^2} \\ C &= \left(\frac{\sin^2\phi}{p^2} + \cos^2\phi\right)\frac{1}{q^2} \end{aligned} \tag{C.2}$$

as reported in e.g. Gerhard (1994). A derivation of this formula can be found in Binney (1985). On the other side, for a given set of viewing angles, ellipticity and PA, the intrinsic axis ratios p, q can be calculated as (de Zeeuw & Franx, 1989)

$$1 - q^2 = \frac{\delta' \left[2\cos 2\Psi^* + \sin 2\Psi^* \left(\sec\theta \cot\phi - \cos\theta \tan\phi\right)\right]}{2\sin^2\theta \left[\delta' \cos\Psi^* \left(\cos\Psi^* + \cot\phi \sec\theta \sin\Psi^*\right) - 1\right]} \tag{C.3a}$$

$$p^2 - q^2 = \frac{\delta' \left[2\cos 2\Psi^* + \sin 2\Psi^* \left(\cos\theta \cot\phi - \sec\theta \tan\phi\right)\right]}{2\sin^2\theta \left[\delta' \cos\Psi^* \left(\cos\Psi^* + \cot\phi \sec\theta \sin\Psi^*\right) - 1\right]} \tag{C.3b}$$

where $\delta' = 1 - q'^2$. Here, Ψ^* is the angle between the projection of the z-axis on the plane of the sky and the semi-major axis of the projected ellipses (Fig. 2b of de Zeeuw & Franx 1989). Therefore, ***this coincides with the viewing angle Ψ^* if and only if the x′-axis is aligned with***

the isophote major axis. For a NE-aligned coordinate system, such as that used by SHAPE3D, the relation between Ψ^* and ψ is

$$\Psi^* = \psi + \mathrm{PA} - 90. \tag{C.4}$$

From eqs. C.3 it is clear that ellipticity gradients or twists imply that $p(r)$, $q(r)$ cannot be constant values at all radii. Moreover, there is no guarantee that a (physically plausible) solution can be found for all viewing angles, which makes ellipsoidal models suitable for removing viewing directions which are incompatible with a given photometry. We see that these equations are undetermined if the LOS lies on the principal axes, in agreement with the fact that at least one of p and q cannot be measured in this case. For example, along the y-axis there is no information about p, since the length of the intermediate axis b is hidden by the LOS, implying that the photometry is compatible with a prolate ($b = c$), oblate ($b = a$) or triaxial ($a > b > c$) case, making the deprojection more complicated.
The axisymmetric case is significantly easier. Here, only one viewing angle is required, and the relation between projected and intrinsic axis ratio is simply

$$q^2 \sin^2 i + \cos^2 i = q'^{\alpha} \tag{C.5}$$

where $\alpha = 2, -2$ for the oblate ($p = 1$) and prolate ($p = q$) case, respectively.

Bibliography

Abell G. O., Corwin Harold G. J., Olowin R. P., 1989, ApJS, 70, 1

Abramowitz M., Stegun I. A., 1972, Handbook of Mathematical Functions

Akaike H., 1974, IEEE Transactions on Automatic Control, 19, 716

Andrae R., Schulze-Hartung T., Melchior P., 2010, arXiv e-prints, p. arXiv:1012.3754

Arnold J. A., et al., 2014, ApJ, 791, 80

Astropy Collaboration et al., 2013, A&A, 558, A33

Astropy Collaboration et al., 2018, AJ, 156, 123

Audet C., Dennis J. E., 2006, SIAM Journal on Optimization, 17, 188

Audet C., Warren H., 2017, Derivative-Free and Blackbox Optimization. Springer

Barnes J. E., Hernquist L., 1992, Nature, 360, 715

Bender R., 1988, A&A, 202, L5

Bender R., 1990, A&A, 229, 441

Bender R., Möllenhoff C., 1987, A&A, 177, 71

Bender R., Döbereiner S., Möllenhoff C., 1988, A&AS, 74, 385

Bender R., Surma P., Döbereiner S., Möllenhoff C., Madejsky R., 1989, A&A, 217, 35

Bender R., Burstein D., Faber S. M., 1992, ApJ, 399, 462

Bender R., Saglia R. P., Gerhard O. E., 1994, MNRAS, 269, 785

Bendinelli O., 1991, ApJ, 366, 599

Bernardi M., et al., 2003a, AJ, 125, 1849

Bernardi M., et al., 2003b, AJ, 125, 1866

Bertola F., Galletta G., 1978, ApJL, 226, L115

Binggeli B., 1980, A&A, 82, 289

Binney J., 1978a, Comments on Astrophysics, 8, 27

Binney J., 1978b, MNRAS, 183, 501

Binney J., 1985, MNRAS, 212, 767

Binney J., 2005, MNRAS, 363, 937

Binney J., Mamon G. A., 1982, MNRAS, 200, 361

Binney J., Merrifield M., 1998, Galactic Astronomy

Binney J. J., Tremaine S., 2008, Galactic dynamics. 2nd ed. Princeton, NJ, Princeton University Press

Bissantz N., Gerhard O., 2002, MNRAS, 330, 591

Bogdán Á., Lovisari L., Volonteri M., Dubois Y., 2018, ApJ, 852, 131

Bois M., et al., 2011, MNRAS, 416, 1654

Bryant J. J., et al., 2015, MNRAS, 447, 2857

Caon N., Capaccioli M., D'Onofrio M., 1993, MNRAS, 265, 1013

Cappellari M., 2002, MNRAS, 333, 400

Cappellari M., 2008, MNRAS, 390, 71

Cappellari M., 2016, ARAA, 54, 597

Cappellari M., 2017, MNRAS, 466, 798

Cappellari M., Emsellem E., 2004, PASP, 116, 138

Cappellari M., McDermid R. M., 2005, Classical and Quantum Gravity, 22, S347

Cappellari M., et al., 2006, MNRAS, 366, 1126

Cappellari M., et al., 2011a, MNRAS, 413, 813

Cappellari M., et al., 2011b, MNRAS, 416, 1680

Cappellari M., et al., 2012, Nature, 484, 485

Cattaneo A., et al., 2009, Nature, 460, 213

Chang Y.-Y., et al., 2013, ApJ, 773, 149

Chen C.-Y., Hwang C.-Y., Ko C.-M., 2016, The Astrophysical Journal, 830, 123

Cimatti A., Daddi E., Renzini A., 2006, A&A, 453, L29

Ciotti L., Lanzoni B., Renzini A., 1996, MNRAS, 282, 1

Conroy C., van Dokkum P. G., Villaume A., 2017, ApJ, 837, 166

Contini E., Yi S. K., Kang X., 2018, MNRAS, 479, 932

Contopoulos G., 1956, ZA, 39, 126

Cooper A. P., Parry O. H., Lowing B., Cole S., Frenk C., 2015, MNRAS, 454, 3185

Côté P., et al., 2007, ApJ, 671, 1456

Cretton N., de Zeeuw P. T., van der Marel R. P., Rix H.-W., 1999, ApJS, 124, 383

Cretton N., Rix H.-W., de Zeeuw P. T., 2000, ApJ, 536, 319

Davies R. L., Birkinshaw M., 1986, ApJL, 303, L45

Davies R. L., Birkinshaw M., 1988, ApJS, 68, 409

Davies R. L., Efstathiou G., Fall S. M., Illingworth G., Schechter P. L., 1983, ApJ, 266, 41

Dehnen W., 1993, MNRAS, 265, 250

Dehnen W., 2009, MNRAS, 395, 1079

Dejonghe H., 1987, MNRAS, 224, 13

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

Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, ApJ, 313, 42

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

Emsellem E., Monnet G., Bacon R., 1994, A&A, 285, 723

Emsellem E., et al., 2007, MNRAS, 379, 401

Emsellem E., et al., 2011, MNRAS, 414, 888

Ene I., et al., 2018, MNRAS, 479, 2810

Erwin P., et al., 2015a, MNRAS, 446, 4039

Erwin P., et al., 2015b, MNRAS, 446, 4039

Event Horizon Telescope Collaboration et al., 2019, ApJL, 875, L4

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

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

Fabian A. C., 2012, ARAA, 50, 455

Falcón-Barroso J., Martig M., 2021, A&A, 646, A31

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

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189

Fall S. M., Frenk C. S., 1983, AJ, 88, 1626

Ferrarese L., et al., 2006, ApJS, 164, 334

Fischer M. S., Brüggen M., Schmidt-Hoberg K., Dolag K., Kahlhoefer F., Ragagnin A., Robertson A., 2022, MNRAS,

Förster Schreiber N. M., et al., 2006, ApJ, 645, 1062

Foster C., et al., 2017, MNRAS, 472, 966

Franx M., Illingworth G. D., 1988, ApJL, 327, L55

Franx M., Illingworth G., de Zeeuw T., 1991, ApJ, 383, 112

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

Gallagher John S. I., Ostriker J. P., 1972, AJ, 77, 288

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

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

Genel S., et al., 2018, MNRAS, 474, 3976

Genzel R., et al., 2006, Nature, 442, 786

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

Gerhard O. E., 1994, in Contopoulos G., Spyrou N. K., Vlahos L., eds, , Vol. 433, Galactic Dynamics and N-Body Simulations. pp 191–274, doi:10.1007/3-540-57983-4_21

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

Gerhard O. E., Binney J., 1985, MNRAS, 216, 467

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

Gerhard O., Kronawitter A., Saglia R. P., Bender R., 2001, AJ, 121, 1936

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

Goullaud C. F., Jensen J. B., Blakeslee J. P., Ma C.-P., Greene J. E., Thomas J., 2018, ApJ, 856, 11

Graham A. W., Guzmán R., 2003, AJ, 125, 2936

Gunn J. E., Gott J. Richard I., 1972, ApJ, 176, 1

Hall P., 1927, Biometrika, 19, 240

Harris C. R., et al., 2020, Nature, 585, 357

Hausman M. A., Ostriker J. P., 1978, ApJ, 224, 320

Hernquist L., 1990, ApJ, 356, 359

Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304

Hubble E. P., 1927, The Observatory, 50, 276

Hubble E. P., 1936, Realm of the Nebulae

Hunter J. D., 2007, Computing in Science & Engineering, 9, 90

Hunter C., de Zeeuw P. T., 1992, ApJ, 389, 79

Illingworth G., King I. R., 1977, ApJL, 218, L109

Irwin J. O., 1927, Biometrika, 19, 225

Jaffe W., 1983, MNRAS, 202, 995

Jeans J. H., 1915, MNRAS, 76, 70

Jedrzejewski R., Schechter P. L., 1989, AJ, 98, 147

Jesseit R., Cappellari M., Naab T., Emsellem E., Burkert A., 2009, MNRAS, 397, 1202

Jin Y., Zhu L., Long R. J., Mao S., Xu D., Li H., van de Ven G., 2019, MNRAS, 486, 4753

Khochfar S., et al., 2011, MNRAS, 417, 845

King A., Pounds K., 2015, ARAA, 53, 115

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

Kochanek C. S., Rybicki G. B., 1996, MNRAS, 280, 1257

Kolmogorov A., 1933, Inst. Ital. Attuari, Giorn., 4, 83

Kondratev B. P., Ozernoi L. M., 1979, Soviet Astronomy Letters, 5, 37

Kormendy J., 1977, ApJ, 218, 333

Kormendy J., 1984, ApJ, 287, 577

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

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

Kormendy J., Ho L. C., 2013, ARAA, 51, 511

Kormendy J., Kennicutt Jr. R. C., 2004, ARAA, 42, 603

Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJS, 182, 216

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

Kroupa P., 2001, MNRAS, 322, 231

Lauer T. R., et al., 1995, AJ, 110, 2622

Le Digabel S., 2011, ACM Transactions on Mathematical Software

Li H., Mao S., Emsellem E., Xu D., Springel V., Krajnović D., 2018a, MNRAS, 473, 1489

Li H., Mao S., Cappellari M., Graham M. T., Emsellem E., Long R. J., 2018b, ApJL, 863, L19

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

Lynden-Bell D., 1962, MNRAS, 124, 1

Ma C.-P., Greene J. E., McConnell N., Janish R., Blakeslee J. P., Thomas J., Murphy J. D., 2014, ApJ, 795, 158

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

Marinacci F., et al., 2018, MNRAS, 480, 5113

Matthews T. A., Morgan W. W., Schmidt M., 1964, ApJ, 140, 35

Mazzalay X., Thomas J., Saglia R. P., Wegner G. A., Bender R., Erwin P., Fabricius M. H., Rusli S. P., 2016, MNRAS, 462, 2847

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., 2012a, ApJ, 756, 179

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

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

Merritt D., 2004, in Ho L. C., ed., Coevolution of Black Holes and Galaxies. p. 263 (arXiv:astro-ph/0301257)

Merritt D., 2006a, ApJ, 648, 976

Merritt D., 2006b, ApJ, 648, 976

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

Mihos J. C., Hernquist L., 1994, ApJL, 437, L47

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

Mo H., Van Den Bosch F., White S., 2008, Galactic dynamics. 2nd ed. Princeton, NJ, Princeton University Press

Monnet G., Bacon R., Emsellem E., 1992, A&A, 253, 366

Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613

Moore B., Lake G., Katz N., 1998, ApJ, 495, 139

Morgan W. W., Lesh J. R., 1965, ApJ, 142, 1364

Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569

Naab T., Burkert A., 2003, ApJ, 597, 893

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

Naab T., et al., 2014, MNRAS, 444, 3357

Naiman J. P., et al., 2018, MNRAS, 477, 1206

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Nelson D., et al., 2018, MNRAS, 475, 624

Neureiter B., et al., 2020, MNRAS, 500, 1437

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

Nipoti C., Londrillo P., Ciotti L., 2003, MNRAS, 342, 501

Ostriker J. P., Tremaine S. D., 1975, ApJL, 202, L113

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

Pasquali A., van den Bosch F. C., Rix H. W., 2007, ApJ, 664, 738

Peebles P. J. E., 1972, ApJ, 178, 371

Pillepich A., et al., 2018, MNRAS, 475, 648

Pogge R. W., et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. p. 77350A, doi:10.1117/12.857215

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

Pounds K. A., Reeves J. N., King A. R., Page K. L., O'Brien P. T., Turner M. J. L., 2003a, MNRAS, 345, 705

Pounds K. A., King A. R., Page K. L., O'Brien P. T., 2003b, MNRAS, 346, 1025

Power C., Zubovas K., Nayakshin S., King A. R., 2011, MNRAS, 413, L110

Pulsoni C., Gerhard O., Arnaboldi M., Pillepich A., Nelson D., Hernquist L., Springel V., 2020, A&A, 641, A60

Pulsoni C., Gerhard O., Arnaboldi M., Pillepich A., Rodriguez-Gomez V., Nelson D., Hernquist L., Springel V., 2021, A&A, 647, A95

Quenneville M. E., Liepold C. M., Ma C.-P., 2022, ApJ, 926, 30

Quinlan G. D., Hernquist L., Sigurdsson S., 1995, ApJ, 440, 554

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, ApJL, 872, L17

Reeves J. N., O'Brien P. T., Ward M. J., 2003, ApJL, 593, L65

Remus R.-S., Forbes D. A., 2021, arXiv e-prints, p. arXiv:2101.12216

Remus R.-S., Dolag K., Naab T., Burkert A., Hirschmann M., Hoffmann T. L., Johansson P. H., 2017, MNRAS, 464, 3742

Richstone D. O., 1976, ApJ, 204, 642

Richstone D. O., Tremaine S., 1984, ApJ, 286, 27

Rix H.-W., White S. D. M., 1990, ApJ, 362, 52

Rix H.-W., White S. D. M., 1992, MNRAS, 254, 389

Rix H.-W., de Zeeuw P. T., Cretton N., van der Marel R. P., Carollo C. M., 1997, ApJ, 488, 702

Roberts P. H., 1962, ApJ, 136, 1108

Robertson A., Harvey D., Massey R., Eke V., McCarthy I. G., Jauzac M., Li B., Schaye J., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3646–3662

Rodriguez-Gomez V., et al., 2019, MNRAS, 483, 4140

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 Symposium Vol. 127, Structure and Dynamics of Elliptical Galaxies. p. 397

Saglia R. P., Colless M., Burstein D., Davies R. L., McMahan R. K., Wegner G., 2001, MNRAS, 324, 389

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

Sahu N., Graham A. W., Davis B. L., 2022, ApJ, 927, 67

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

Sargent W. L. W., Schechter P. L., Boksenberg A., Shortridge K., 1977, ApJ, 212, 326

Sarzi M., et al., 2010, MNRAS, 402, 2187

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

Schoenberg I. J., 1946, Q. Appl. Math., 4, 45

Schulze A., Gebhardt K., 2011, ApJ, 729, 21

Schulze F., Remus R.-S., Dolag K., Burkert A., Emsellem E., van de Ven G., 2018, MNRAS, 480, 4636

Schwarzschild M., 1979, ApJ, 232, 236

Schwarzschild M., 1993, ApJ, 409, 563

Scorza C., Bender R., 1990, A&A, 235, 49

Scorza C., Bender R., 1995, A&A, 293, 20

Sersic J. L., 1968, Atlas de Galaxias Australes

Smirnov N. V., 1939, Bull. Math. Univ. Moscou, 2, 3

Springel V., et al., 2018, MNRAS, 475, 676

Stark A. A., 1977, ApJ, 213, 368

Syer D., Tremaine S., 1996a, MNRAS, 282, 223

Syer D., Tremaine S., 1996b, MNRAS, 282, 223

Teklu A. F., Remus R.-S., Dolag K., Beck A. M., Burkert A., Schmidt A. S., Schulze F., Steinborn L. K., 2015, ApJ, 812, 29

Teklu A. F., Remus R.-S., Dolag K., Burkert A., 2017, MNRAS, 472, 4769

Thater S., et al., 2022, arXiv e-prints, p. arXiv:2205.04165

Thomas J., Lipka M., 2022, MNRAS, 514, 6203

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

Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Richstone D., 2004, MNRAS, 353, 391

Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Corsini E. M., Wegner G., 2005, MNRAS, 360, 1355

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

Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Corsini E. M., Wegner G., 2009, ApJ, 691, 770

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, Vol. 311, Galaxy Masses as Constraints of Formation Models. pp 36–39

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

Toomre A., 1977, in Tinsley B. M., Larson Richard B. Gehret D. C., eds, Evolution of Galaxies and Stellar Populations. p. 401

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

Tremaine S., Richstone D. O., Byun Y.-I., Dressler A., Faber S. M., Grillmair C., Kormendy J., Lauer T. R., 1994, AJ, 107, 634

Tremblay B., Merritt D., 1995, AJ, 110, 1039

Tremblay B., Merritt D., 1996, AJ, 111, 2243

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

Valluri M., Merritt D., Emsellem E., 2004, ApJ, 602, 66

Valluri M., Ferrarese L., Merritt D., Joseph C. L., 2005, ApJ, 628, 137

Van Rossum G., Drake F. L., 2009, Python 3 Reference Manual. CreateSpace, Scotts Valley, CA

Vasiliev E., Valluri M., 2020, ApJ, 889, 39

Verolme E. K., et al., 2002, MNRAS, 335, 517

Verro K., et al., 2022, A&A, 660, A34

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

Virtanen P., et al., 2020, Nature Methods, 17, 261

Wagner S. J., Bender R., Moellenhoff C., 1988, A&A, 195, L5

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

Weijmans A.-M., et al., 2014, MNRAS, 444, 3340

Wendland H., 1995, Adv. Comp. Math., 4, 389

Wendland H., 2005, Scattered Data Approximation. Cambridge University Press, Cambridge, UK

Werner N., McNamara B. R., Churazov E., Scannapieco E., 2019, Space Science Reviews, 215, 5

Zemp M., Gnedin O. Y., Gnedin N. Y., Kravtsov A. V., 2011, ApJS, 197, 30

Zhao H., 1996, MNRAS, 278, 488

de Lorenzi F., Debattista V. P., Gerhard O., Sambhus N., 2007, MNRAS, 376, 71

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 Vaucouleurs G., 1948, Annales d'Astrophysique, 11, 247

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

de Zeeuw T., Merritt D., 1983, ApJ, 267, 571

van Dokkum P. G., et al., 2008, ApJL, 677, L5

van Gorkom J. H., 2004, in Mulchaey J. S., Dressler A., Oemler A., eds, Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution. p. 305 (`arXiv:astro-ph/0308209`)

van de Sande J., et al., 2019, MNRAS, 484, 869

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 den Bosch R. C. E., van de Ven G., 2009, MNRAS, 398, 1117

van den Bosch R. C. E., van de Ven G., Verolme E. K., Cappellari M., de Zeeuw P. T., 2008, MNRAS, 385, 647

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

van der Marel R. P., Cretton N., de Zeeuw P. T., Rix H.-W., 1998, ApJ, 493, 613