

# Axisymmetric orbit models of $N$ -body merger remnants: a dependency of reconstructed mass on viewing angle

J. Thomas,<sup>1,2★</sup> R. Jesseit,<sup>2★</sup> T. Naab,<sup>2</sup> R. P. Saglia,<sup>1</sup> A. Burkert<sup>2</sup> and R. Bender<sup>1,2</sup>

<sup>1</sup>Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse, D-85748 Garching, Germany

<sup>2</sup>Universitätssternwarte, Scheinerstr. 1, 81679 München, Germany

Accepted 2007 August 12. Received 2007 July 4; in original form 2007 January 19

## ABSTRACT

We model mock observations of collisionless  $N$ -body disc–disc mergers with the same axisymmetric orbit superposition program that has been used to model elliptical galaxies in Coma. The remnants sample representatively the shape distribution of disc–disc mergers, including the most extreme cases, like highly prolate, maximally triaxial and dominantly oblate objects. The aim of our study is to better understand how the assumption of axial symmetry affects reconstructed masses and stellar motions of systems which are intrinsically not axisymmetric, whether the axisymmetry assumption then leads to a bias and how such a potential bias can be recognized in models of real galaxies. The mass recovery at the half-light radius depends on viewing angle and intrinsic shape: edge-on views allow to reconstruct total masses with an accuracy between 20 per cent (triaxial/prolate remnants) and 3 per cent (oblate remnant). Masses of highly flattened, face-on systems are underestimated by up to 50 per cent. Deviations in local mass densities can be larger where remnants are strongly triaxial or prolate. Luminous mass-to-light ratios are sensitive to box orbits in the remnants. Box orbits cause the central value of the Gauss–Hermite parameter  $H_4$  to vary with viewing angle. Reconstructed luminous mass-to-light ratios, as well as reconstructed central masses, follow this variation. Luminous mass-to-light ratios are always underestimated (up to a factor of 2.5). Respective dark haloes in the models can be overestimated by about the same amount, depending again on viewing angle. Reconstructed velocity anisotropies  $\beta$  depend on viewing angle as well as on the orbital composition of the remnant and are mostly accurate to about  $\Delta\beta = 0.2$ . Larger deviations can occur towards the centre or the outer regions, respectively. We construct  $N$ -body realizations of the Schwarzschild models to discuss chaotic orbits and the virial equilibrium in our models. In this study we explore the extreme limits of axisymmetric models. Apparently flattened, rotating ellipticals of intermediate mass are likely close to both, axial symmetry and edge-on orientation. Our results imply that Schwarzschild models allow a reconstruction of their masses and stellar anisotropies with high accuracy.

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

## 1 INTRODUCTION

Subject of this paper is the reconstruction of *synthetic* dynamical systems –  $N$ -body merger remnants – with orbit models. The motivation behind is to better understand models of *real* dynamical systems, especially those of elliptical galaxies.

Elliptical galaxies are optically smooth stellar systems in approximate dynamical equilibrium. They can result from various kinds of merging processes – e.g. the merging of two discs (e.g. Toomre

& Toomre 1972) – or from some kind of monolithic collapse (e.g. Eggen, Lynden-Bell & Sandage 1962; Larson 1974). In a cosmological context an early-type can go through several distinct phases of the above prototypical forms (e.g. Naab et al. 2007).

Apart from following the cosmic evolution of potential progenitor systems the only way to determine elliptical galaxy evolutionary histories is to scan their present structure for characteristic fingerprints of different evolutionary events. This concerns both, scaling relations of ellipticals as a class and internal properties of individual systems. For example, the cold collapse of a stellar system results in a typical gradient from central isotropy to strong outer radial anisotropy in stellar orbits (van Albada 1982). Galaxy mergers, on

★E-mail: jthomas@mpe.mpg.de (JT); jesseit@usm.uni-muenchen.de (RJ)

the other hand, can produce a variety of dynamical systems. The final structure of disc–disc merger remnants depends, e.g. on progenitor properties (Barnes 1992; Hernquist 1992, 1993), the merging geometry (Weil & Hernquist 1996; Dubinski 1998) and on the mass ratio of the progenitors (Naab & Burkert 2003; Jesseit, Naab & Burkert 2007). Ellipticals as progenitors can be merged as well (Naab, Khochfar & Burkert 2006a).

The difficulty with real galaxies is that their intrinsic properties, like the intrinsic shape, the distribution of mass or the geometry of stellar orbits, are not directly observable. They have to be inferred from observations through dynamical modelling.

The state-of-the-art method for such modelling is Schwarzschild’s orbit superposition technique (Schwarzschild 1979). In very rough terms (1) the photometry of a galaxy is deprojected into the 3D internal light distribution; (2) the light distribution is multiplied with the stellar mass-to-light ratio ( $M/L$ ) and – depending on the specific application – a black hole or dark halo is added to obtain the composite mass distribution; (3) thousands of orbits are calculated in the resulting gravitational potential; (4) the orbits are added together to fit the kinematic and photometric observations of the galaxy. Thereby each orbit is weighted individually to optimize the match with the data.

Schwarzschild’s method can be read as a numerical implementation of Jeans’ theorem, which states that *stationary* distribution functions (DFs) (the density of stars in 6D phase space) of collisionless systems are *necessarily* functions of the integrals of motion (e.g. Binney & Tremaine 1987). In other words – since integrals of motion label orbits (and vice versa) – the phase-space density in a stationary system is constant along individual orbits. This explains why the fundamental building blocks of stationary dynamical systems are entire orbits and no density variation along individual orbits needs to be considered. In principle then, the only assumption underlying Schwarzschild modelling is that galaxies are stationary and collisionless. In practice, however, applications also assume a specific internal symmetry for each object under study. This is to reduce the degrees of freedom in the deprojection and to simplify the sampling of phase space with orbits. Axial symmetry is the simplest geometry to account for intrinsic flattening, inclination effects, rotation and the presence of disc-like subsystems in real galaxies. Several implementations of Schwarzschild’s method for axially symmetric potentials have been developed (Cretton et al. 1999; Gebhardt et al. 2000; Häfner et al. 2000; Thomas et al. 2004; Valluri, Merritt & Emsellem 2004; Cappellari et al. 2006) and have been used to analyse surveys of elliptical galaxy kinematics (Gebhardt et al. 2003; Cappellari et al. 2006; Thomas et al. 2007).

Comparison with phase-space DFs of synthetic galaxy models has proven that an accuracy level of better than 10 per cent can be achieved in the numerical construction of axisymmetric orbit superpositions involving all three integrals of motion (Cretton et al. 1999; Krajnovic et al. 2005; Thomas et al. 2004, 2005). Concerning applications to real galaxies, however, the distribution of apparent ellipticities, isophotal twists and/or minor-axis rotation indicate that ellipticals cannot be *exactly* axisymmetric (e.g. Bertola & Galletta 1979; Franx, Illingworth & Heckman 1989; Jedrzejewski & Schechter 1989; Tremblay & Merritt 1996). Up to now, it is not clear how such intrinsic deviations from rotational symmetry in real galaxies affect the results of axisymmetric dynamical modelling.

In this paper, we present first results of a project aimed to systematically survey the properties of axisymmetric Schwarzschild models that are applied to non-axisymmetric test objects. Specifically, we imitate realistic photometric and kinematical observations (realistic in terms of spatial coverage and resolution) of collisionless

$N$ -body merger remnants and fit them with the same Schwarzschild code that has been used for a study of Coma ellipticals (Thomas et al. 2007). We determine internal mass distributions and velocity anisotropies just as for real galaxies, but since we know the corresponding properties of our test objects, we can examine the models.

The final goal of our project is twofold. First, we want to explore possible systematic deviations that are caused by applying axisymmetric models to objects that do not respect any internal symmetry. Thereby, we want to understand how such deviations can be recognized when modelling a real galaxy, whose internal structure is not known a priori. Collisionless disc mergers are ideal for such a study, because they represent physically motivated dynamical systems that cover a large range of intrinsic shapes and dynamical structures.

By investigating how intrinsically non-axisymmetric systems are mapped on to axisymmetric models we also gain templates for the interpretation of real galaxy models. A second goal of our study is therefore to compare the resulting Schwarzschild models of merger remnants with Schwarzschild models of real galaxies. Since we use the same modelling code in both cases, differences are indicative for structural differences between galaxies and the analysed merger remnants. Knowing such differences allows a deeper understanding of the physical mechanisms involved in elliptical galaxy build-up.

The present paper focuses on the first part of the project. A detailed discussion of the results with respect to observations and models of real galaxies is planned for a future publication. We further plan to extend our survey of Schwarzschild models to samples of mergers involving gas physics and/or dynamical systems developing from cosmological initial conditions.

The paper is organized as follows. Section 2 describes the sample of merger remnants used for this work. Our implementation of Schwarzschild’s technique is reviewed in Section 3. Section 4 summarizes tests with a Hernquist sphere. The modelling results are detailed in Section 5 (general notes), Section 6 (reconstructed masses) and Section 7 (reconstructed velocity anisotropies). In Section 8 we discuss various modelling uncertainties. Section 9 deals with the viewing-angle dependency of the total mass recovery and Section 10 discusses the relation between reconstructed luminous  $M/L$  values and the central orbital structure of the merger remnants. Implications for models of real galaxies are briefly discussed in Section 11. The paper closes with a summary in Section 12.

## 2 MERGER SAMPLE

A careful selection of the sample of merger remnants is crucial for our study. The merger remnants for this paper are taken from the collisionless disc–disc merger sample of Naab & Burkert (2003), which is consistent with expectations from cosmological simulations (Khochfar & Burkert 2006). Their progenitor galaxies consist of exponential discs and Hernquist bulges with a bulge-to-disc ratio of 1:3, embedded in pseudo-isothermal dark matter haloes such that the overall circular velocity curve is approximately flat in the outer parts. With respect to their global kinematical and photometric properties, these merger remnants resemble giant ellipticals of intermediate mass (Naab & Trujillo 2006). For further details on the general properties and numerical details of how the merger remnants have formed we refer the reader to Naab & Burkert (2003).

Different merger remnants result from different merging geometries, but here we select them only according to their shape and do not care how they have formed. As the shape of the merger remnants

is very closely correlated to their orbital content (Jesseit et al. 2005), we know that sampling different shapes ensures that we explore a range of different orbital makeups as well.

## 2.1 Orbital composition and shape of remnants

According to rotational symmetry, all orbits in axisymmetric potentials conserve the  $z$  component  $L_z$  of angular momentum and are *minor-axis tubes*, or *Z tubes*. Such Z tubes can have various shapes between equatorial–radial, equatorial–circular, shell-like and polar–radial (e.g. Richstone 1982).

In triaxial dynamical systems we will expect more orbit classes (e.g. de Zeeuw 1985). In particular *box orbits* (most frequent in the centre) without net angular momentum, *boxlets* (resonant boxes found at larger radii) and *inner* and *outer major-axis tubes* (also *X tubes* in the following). Major-axis tubes have significant angular momentum around the long axis. As in axisymmetric potentials, also triaxial force fields support minor-axis tubes, which have a non-zero angular momentum with respect to the short axis. The abundances of different orbit classes will depend on the exact shape of the merger remnant.

The shape is determined by the ratio of the three principal axes of the inertial tensor calculated from the particle positions in the merger remnants. The main axes are denoted:  $X$  (long),  $Y$  (intermediate) and  $Z$  (short), respectively. The corresponding values of the inertial tensor are  $a$ ,  $b$  and  $c$ , respectively.

## 2.2 Sample selection

We choose six remnants that representatively sample the range of shapes realized by the collisionless mergers of Naab & Burkert (2003), including the most extreme cases: (1) a box orbit dominated remnant (TRIAX); (2) one with a high  $X$  tube fraction (PROLATE); (3) a nearly round object (ROUND); (4, 5) two very flattened remnants with different inner shape profiles (FLAT and ELONG); (6) one oblate remnant, dominated by  $Z$  tubes (OBLATE). Modelling of further remnants from this sample would bring little additional information.

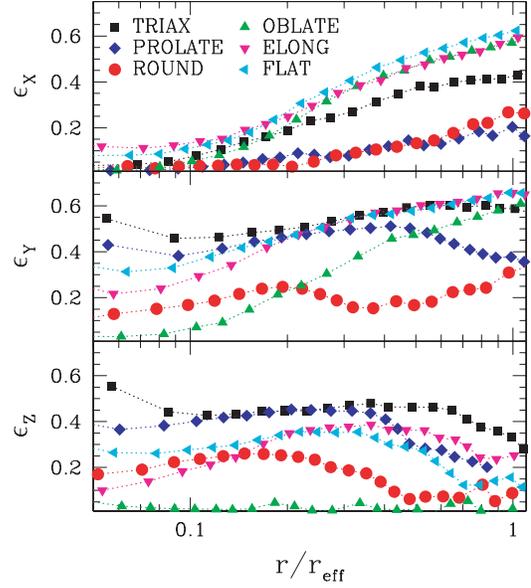
Table 1 summarizes orbital abundances. Respective ellipticity profiles  $\epsilon_X$ ,  $\epsilon_Y$  and  $\epsilon_Z$  that result from projecting the remnants along the three principal axes  $X$ ,  $Y$  and  $Z$  are shown in Fig. 1.

The Schwarzschild models considered in this work always assume oblate axial symmetry. Concerning projected ellipticities, axial symmetry implies either

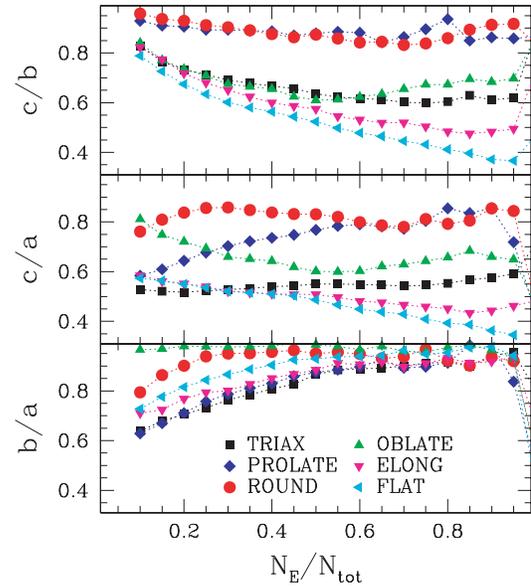
$$\epsilon_Z \equiv 0, \epsilon_X \equiv \epsilon_Y \text{ (oblate)} \quad (1)$$

**Table 1.** Selected merger remnants. (1) Merger remnant; (2) progenitor mass ratio; (3) merging geometry according to table 1 of Naab & Burkert (2003); (4–6) abundances of major orbit classes among the 40 per cent most bound particles (from Jesseit et al. 2005). Irregular orbits and orbits without classification are not included in the table.

Merger remnant			Box and boxlet	Z tube	Inner and outer X tube
(1)	(2)	(3)	(4)	(5)	(6)
TRIAX	1:1	5	0.57	0.24	0.06
PROLATE	1:1	7	0.40	0.21	0.29
ROUND	2:1	12	0.23	0.40	0.25
FLAT	2:1	17	0.47	0.35	0.04
ELONG	3:1	29	0.53	0.29	0.05
OBLATE	4:1	11	0.12	0.76	0.03



**Figure 1.** Projected ellipticities of merger remnants. From top to bottom:  $X$ ,  $Y$  and  $Z$  projections. Radii are scaled by the half-light radius  $r_{\text{eff}}$  of the respective projection.



**Figure 2.** Internal axis ratios of the merger remnants shown in Fig. 1 as function of the fraction  $N_E/N_{\text{tot}}$  of particles with the highest binding energy. From top to bottom:  $c/b$ ,  $c/a$  and  $b/a$ .

or

$$\epsilon_X \equiv 0, \epsilon_Y \equiv \epsilon_Z \text{ (prolate)}. \quad (2)$$

Fig. 1 reveals that one of the six modelled merger remnants is consistent with oblate axial symmetry (OBLATE), while two others are marginally consistent with prolate axial symmetry (ROUND, PROLATE).

Generally, the remnant sample of Naab & Burkert (2003) is deviant from oblate rotational symmetry in the inner regions. This can be inferred from Fig. 2, which shows profiles of internal axis ratios, calculated from the spatial distribution of the  $N_E$  most bound (luminous) particles ( $N_{\text{tot}}$  is the total number of luminous particles in

the remnant; cf. Jesseit et al. 2005). In terms of intrinsic axis ratios, oblate axial symmetry implies  $c < a$  and  $a \equiv b$ .

### 2.3 Mock observations

We will model the ‘observational’ data of the projections along the three principal axes for each remnant. Almost all observational properties, photometric or kinematic, will reach their maximum or minimum values at these projections. This is so because the principal axes are also the symmetry axes of the various orbit classes which means that particles will move perpendicular or parallel to one of the axes, depending on orbit class and projection. Consequently, the moments of the line-of-sight velocity distribution (LOSVD) will reach extreme values for the respective projections. Similar reasoning can be applied to the photometric properties, e.g. the isophotal shape parameter  $a_4$  (Bender & Möllenhoff 1987). Concerning  $Z$  tubes in triaxial potentials, e.g.  $a_4$  reaches its extreme values in the long-axis projection (most boxy) and intermediate-axis projection (most discy), respectively (cf. Jesseit et al. 2005 for detailed discussion). In summary: by modelling the principal projections, we are testing the extreme cases, where the influence of certain orbit classes on the observables is visible (or not). Influence of the viewing angle on our results will be discussed in detail in the following sections.

As we want to make meaningful statements about the recovery of real galaxy structure, we try to imitate observational conditions of a comparison sample of Coma ellipticals modelled by Thomas et al. (2007) with the same Schwarzschild code. This comparison sample consists of long-slit major and minor-axis spectra (Mehlert et al. 2000; Wegner et al. 2002). The photometry for the Coma galaxy models is constructed as a composite of *Hubble Space Telescope* (*HST*) (centre) and ground-based imaging (outer parts; cf. Thomas et al. 2005).

The photometry of the  $N$ -body mergers is obtained for two different resolutions as well. The coarse one is mimicking ground-based observations and extends to large radii. The one with higher resolution simulates *HST* data at small radii. We include seeing by smoothing the particle distribution with a Gaussian of width  $1/20r_{\text{eff}}$  (roughly the size of the numerical softening parameter) in the high-resolution case. In the coarse resolution case the seeing amounts to three quarters of the effective radius. Low- and high-resolution photometry are combined into one continuous profile, as described in Thomas et al. (2005).

To match the observational conditions of the Coma comparison sample the kinematic information is extracted at an intermediate resolution (smoothing width  $1.2 \text{ arcsec} \approx 1/6 r_{\text{eff}}$ ) along the apparent photometric major and minor axes (cf. Section 3.2). We have kinematic data out to about two effective radii (see Naab & Burkert 2003, for extensive discussion on how artificial observations are performed on  $N$ -body remnants).

## 3 SCHWARZSCHILD MODELLING

Our Schwarzschild models are described in Thomas et al. (2004) where a thorough discussion of the modelling implementation is given.

### 3.1 Model set-up

In the following, we briefly recall the basic steps of the Schwarzschild technique.

(i) The surface brightness of each remnant projection is deprojected at three inclination angles: the edge-on deprojection probes

the inclination where the Schwarzschild model is least flattened. The other extreme, the most flattened Schwarzschild model, is constructed at an inclination angle, for which the deprojection appears as an E7 galaxy (when seen from the side). Inclination angles resulting in intrinsically even flatter models are unreasonable because (1) ellipticals flatter than E7 are not observed and (2) remnants flatter than E7 are not in our merger remnant sample. As an intermediate case we also probe an inclination angle that leads to a Schwarzschild model resembling an E5 galaxy (when seen from the side).

The luminosity distribution  $\nu$  is assumed to be axisymmetric and we include surface brightness, ellipticity and  $a_4$  profiles in the deprojection, which is performed with the non-parametric program of Magorrian (1999).

(ii) Based on the deprojected luminosity profile  $\nu$  a mass distribution is constructed via

$$\rho = \Upsilon \nu + \rho_{\text{DM}}, \quad (3)$$

where  $\Upsilon$  determines the amount of mass that follows the light (we will denote  $\Upsilon$  the stellar mass-to-light ratio in the following). For the additional dark matter density  $\rho_{\text{DM}}$  we adopt a Navarro–Frenk–White (NFW) profile and the relation between concentration and mass given in Navarro, Frenk & White (1996). The dark haloes of the merger remnants do not follow these profiles exactly. The progenitor galaxies are embedded in pseudo-isothermal haloes with a flat central density core. After the merging the central dark matter slope steepens, but is still shallower than in NFW profiles (cf. Section 6.2).

Our choice for NFW haloes is motivated by results of Monte Carlo simulations showing that one can always find an NFW halo among the above introduced family which mimics an (non-singular) isothermal distribution sufficiently well over the radial region considered here (Thomas et al. 2005). Moreover, for a few remnant projections we have also calculated cored logarithmic haloes, and the results do not change significantly (cf. Section 6.2). Then, since we do not lose generality, it is convenient to use the one-parameter family of NFW haloes. To explore possible effects of halo shapes we model each halo once with a spherical mass distribution and once with a flattening of the density distribution of  $c/b \equiv c/a = 0.7$ , where  $a$ ,  $b$  and  $c$  are the long, intermediate and short axes of the halo mass distribution, respectively. The haloes of the merger remnants are close to oblate-axial symmetry, with  $b/a > 0.9$  and  $0.7 \leq c/a \leq 1.0$ .

With the mass density fixed, the gravitational potential  $\Phi$  follows by solving Poisson’s equation.

(iii) In the gravitational potential  $\Phi$  a representative set of orbits is calculated. The orbit sampling is described in detail in Thomas et al. (2004).

(iv) In the final step the orbits are superposed to fit the photometric and kinematical constraints. The maximum entropy technique of Richstone & Tremaine (1988) is applied and the kinematic data are fitted by solving for the maximum of

$$S - \alpha \chi^2 \rightarrow \max. \quad (4)$$

$S$  is the entropy of the model and

$$\chi^2 \equiv \sum_{j=1}^{N_{\mathcal{L}}} \sum_{k=1}^{N_{\text{vel}}} \left( \frac{\mathcal{L}_{\text{mod}}^{jk} - \mathcal{L}_{\text{in}}^{jk}}{\Delta \mathcal{L}_{\text{in}}^{jk}} \right)^2 \quad (5)$$

measures the difference between input LOSVDs  $\mathcal{L}_{\text{in}}$  and model LOSVDs  $\mathcal{L}_{\text{mod}}$  (see Thomas et al. 2004, for more details about the calculation of  $S$  and  $\mathcal{L}$  in this context). Each LOSVD is binned into  $N_{\text{vel}}$  velocity bins and the input data consist of  $N_{\mathcal{L}}$  LOSVDs in total (cf. Section 3.2 for further details). The luminosity density is treated

as a boundary condition to equation (4). The regularization parameter  $\alpha$  in equation (4) allows to control the relative importance of  $\chi^2$  minimization (fit to data) and entropy maximization (smoothness of the distribution function).

In the following we will consider two cases for  $\alpha$ . First, models obtained with  $\alpha = 0$  will be called  $S_{\max}$  models, because for  $\alpha = 0$  the  $\chi^2$  term vanishes and the orbital weights are entirely determined by the maximization of  $S$  (under the boundary condition related to  $\nu$ ). In order to fit an orbit library to a given set of kinematical data,  $\alpha$  has to be positive. The larger  $\alpha$ , the better the fit will be. In case of real observations, very large  $\alpha$  can result in models that fit the noise in the data. Concerning our merger remnant fits, we assume that the input data are not affected significantly by noise (cf. Section 3.2). Therefore, as the second case for  $\alpha$ , we consider a value large enough such that the minimum of  $\chi^2$  is reached ( $\chi^2_{\min}$  models). This usually occurs around  $\alpha \approx 1$ .<sup>1</sup> Larger  $\alpha$  values do not change  $\chi^2$  or other model properties significantly.

### 3.2 Definition of $\chi^2$

To solve equation (4) one needs to evaluate the  $\chi^2$  term and, thus, to specify the  $\Delta\mathcal{L}_{\text{in}}^{jk}$  of equation (5). Insofar as the  $N$ -body simulations are viewed as a discrete  $N$ -particle realization of an underlying continuous phase-space DF, the mock observations should be interpreted to have some intrinsic Poisson scatter that decreases with increasing the number of particles. This describes the case of Section 4, where we test our modelling machinery with an  $N$ -body representation of a Hernquist sphere. It is also valid for the set-up of the progenitor systems. In both cases, the  $N$ -body system is an imperfect representation of an underlying continuous phase-space DF. The merger remnants, however, are not such an  $N$ -body sampling of some unknown DF. Instead, they just reflect the dynamical evolution of  $N$  particles from their particular initial conditions – irrespective of how these have been constructed. In this sense, after the relaxation induced by the merging, we treat the mock observations as ‘ideal’ observations of a discrete ( $N \approx 10^5$  particle) dynamical system, that we try to represent by Schwarzschild models. There is no obvious way to define  $\Delta\mathcal{L}_{\text{in}}^{jk}$  in this case, however.

For a statistical analysis the proper way to proceed is to add random fluctuations to the raw observations. The resulting noisy ‘data’ together with the ‘error bars’ from which the noise has been constructed provide a statistically consistent input to the models. However, our merger sample is small and it would be necessary to model several random realizations of the original raw data in order to avoid any influence of a particular noise pattern on the results. This is computationally too expansive as it means to model effectively dozens of data sets. Moreover, it is not the goal of this study to quantify uncertainties that originate from observational errors (which has been done elsewhere, e.g. Thomas et al. 2005). Instead, our aim is to explore possible *systematic* biases arising when treating non-axisymmetric objects with axisymmetric models. Therefore, we set up our model input as follows.

First, Gauss–Hermite moments  $v, \sigma, H_3$  and  $H_4$  (Gerhard 1993; van der Marel & Franx 1993) of the merger remnants are calculated as in Naab & Burkert (2001). The Gauss–Hermite moments are then used to calculate the LOSVDs  $\mathcal{L}_{\text{in}}$  at a set of radii typical for our comparison sample of Coma ellipticals. Corresponding observational errors of Coma galaxies at these radii are scaled to the

mock data.<sup>2</sup> The Gauss–Hermite ‘error bars’ are propagated into  $\Delta\mathcal{L}_{\text{in}}^{jk}$  by means of Monte Carlo simulations. The resulting LOSVDs  $\mathcal{L}_{\text{in}} \pm \Delta\mathcal{L}_{\text{in}}^{jk}$  are used as input for the Schwarzschild models without adding noise explicitly. Neglecting the noise makes uncertainties of derived model quantities (masses, internal velocity moments) unreliable. But for our purpose of identifying systematic trends it is only important to flag a best-fitting model in a similar way as a best-fitting model is determined for a real galaxy. The role of  $\Delta\mathcal{L}_{\text{in}}^{jk}$  is to specify the relative weight of different data points. The usage of error bars from real observations ensures that in our models data from different spatial regions are weighted similar as in models of real galaxies.

### 3.3 The best-fitting model

To obtain the best-fitting dynamical model we calculate Schwarzschild models on a grid in the 2D parameter space ( $\Upsilon, c$ ). Thereby we probe  $0.3 \leq \Upsilon \leq 1.3$  with  $\Delta\Upsilon = 0.1$  and  $2.5 \leq c \leq 30.0$  with  $\Delta c = 2.5$ . For each pair ( $\Upsilon, c$ ) on the grid one model is calculated with a spherical halo and another one with a halo flattening of  $q = 0.7$ . The procedure is repeated for up to three inclinations (cf. Section 3.1).

For this first set of models we use a coarse library set-up with  $2 \times 3500$  orbits, roughly half the number used to model Coma ellipticals (Thomas et al. 2007) and roughly twice the number that has been used by the Nuker team for models of galaxy centres (Gebhardt et al. 2000). Among the low-resolution models one, say with parameters ( $\Upsilon_f, c_f, q_f$ ), yields the lowest  $\chi^2$ . Around these parameter values we recalculate models with a larger number of orbits ( $2 \times 9000$  orbits as used for models of Coma galaxies by Thomas et al. 2007). The overall best-fitting model is chosen among these high-resolution fits according to the minimum of  $\chi^2$ .

For the high-resolution models we adjust the modelling strategy as follows. (1) As it will be discussed in Section 5.3, the best-fitting (low-resolution) model is always at an inclination  $i = 90^\circ$ . For the high-resolution case we therefore only consider edge-on geometries. (2) We examine the same grid for ( $\Upsilon, q$ ) as in the low-resolution case, but restrict concentrations around  $c_f$ , usually probing the region between  $c_f - 2\Delta c$  and  $c_f + 2\Delta c$ . If necessary we extend the concentration interval such that the best-fitting high-resolution model never occurs at the boundary of the sampled parameter space. When resampling with a larger number of orbits, we vary the halo concentrations in smaller steps of  $\Delta c = 1.0$ .

We do not find systematic differences between the models with  $2 \times 3500$  orbits and those with  $2 \times 9000$  orbits, respectively. For example, 12 out of 18 best-fitting luminous mass-to-light ratios  $\Upsilon_{\text{fit}}$  are the same in low-resolution and high-resolution models. In the remaining cases they change by  $\Delta\Upsilon_{\text{fit}} = 0.1$  (four models) and  $\Delta\Upsilon_{\text{fit}} = 0.2$  (two models), respectively. There is no preferred direction for the change  $\Delta\Upsilon_{\text{fit}}$ .

## 4 VALIDATION: A HERNQUIST SPHERE

To check all conversions from  $N$ -body systems to Schwarzschild models and back we first model a self-consistent Hernquist sphere (Hernquist 1990): we sample the isotropic Hernquist DF with

<sup>2</sup> We use fractional errors in  $v$  and  $\sigma$  but absolute errors for  $H_3$  and  $H_4$ . As template to create the error bars we use the observations of NGC 4807, which are prototypical for the Coma sample in terms of radial coverage and signal-to-noise ratio.

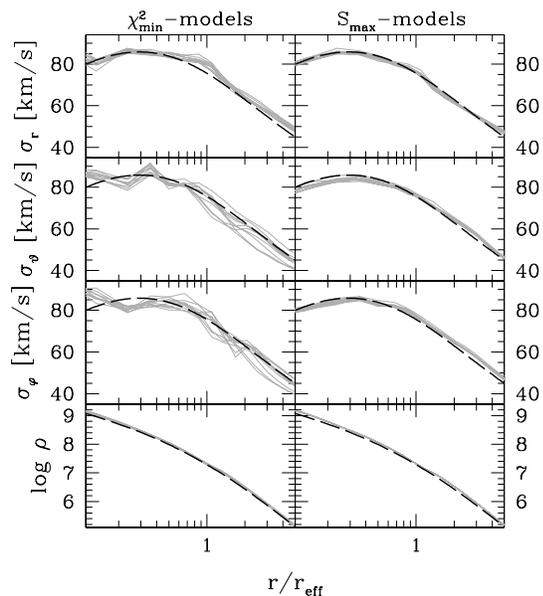
<sup>1</sup> The exact value we will use is  $\alpha = 0.9143$  and arises from the iterative solution of equation (4); see e.g. Thomas et al. (2005).

$N = 1.6 \times 10^5$  particles and ‘observe’ the resulting  $N$ -body realization in exactly the same way as the merger remnants. The number of particles  $N = 1.6 \times 10^5$  resembles the number of luminous particles in the 1:1 merger remnants, analysed later on. For the goal of verifying our machinery by reconstructing the Hernquist sphere we modify the modelling procedure as follows. (1) We only consider an inclination of  $i = 90^\circ$ , such that the deprojection is unique. (2) We only fit self-consistent Schwarzschild models, because the Hernquist sphere is set up self-consistently (without dark matter). Finally, in order to evaluate the influence of noise in the  $N$ -body representation, we combine Schwarzschild fits to 10 different Monte Carlo realizations of the Hernquist sphere. The only free mass parameter in this test run is the stellar mass-to-light ratio  $\Upsilon$ .

Application of our Schwarzschild models to the Hernquist sphere yields  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.993 \pm 0.037$ , where  $\Upsilon_{\text{fit}}$  and  $\Upsilon_{\text{in}}$  are the  $M/L$  of the Schwarzschild models (averaged over fits to 10 realizations of the Hernquist sphere) and the  $N$ -body input, respectively. The quoted uncertainties reflect the variance about the mean.

Internal velocity moments of Schwarzschild fits are shown in Fig. 3 together with the analytic profiles for the isotropic Hernquist sphere (Hernquist 1990). Results from Schwarzschild modelling are spherically averaged. For the  $\chi^2_{\text{min}}$  models in the left-hand panels of Fig. 3 a regularization parameter  $\alpha \approx 1$  has been used (cf. Section 3.1; the same value is used for the merger remnant fits). Apart from some noise in the Schwarzschild models, the overall agreement between analytic calculations and Schwarzschild fits is very good.

To understand whether the scatter in  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  and the internal moments originates from uncertainties in the Schwarzschild code or whether it comes from noise in the  $N$ -body realization, we also tried to reconstruct the Hernquist sphere by a method that is independent



**Figure 3.** Schwarzschild models for a Hernquist sphere. In each panel 10 solid grey lines represent models for 10 Monte Carlo samplings of a Hernquist sphere. The analytically derived reference profiles of the Hernquist sphere are shown by black, dashed lines. Three top rows: radial ( $\sigma_r$ , top), meridional ( $\sigma_\theta$ , middle) and azimuthal ( $\sigma_\phi$ , bottom) velocity dispersions (luminosity-weighted, spherical averages); bottom row: density. Left-hand column:  $\chi^2_{\text{min}}$  fits; right-hand column:  $S_{\text{max}}$  models (details in the text).

of noise in the  $N$ -body kinematics: the solution of equation (4) for  $\alpha = 0$ . As stated in Section 3.1, with  $\alpha = 0$  the  $\chi^2$  term vanishes and the orbit distribution is determined entirely by maximizing its entropy ( $S_{\text{max}}$  model). The idea behind considering  $S_{\text{max}}$  models here is the following: the maximization of  $S$  yields, in a sense, the smoothest DF for the given density profile. Assuming that this smoothing isotropizes stellar velocities then the  $S_{\text{max}}$  model would be identical to the (unique) isotropic DF, which is connected to any self-consistent spherical density profile. Maximizing  $S$  would therefore determine the orbital weights (and internal moments etc.) of our orbit representation of the Hernquist sphere without any fit to the kinematics.

Since the orbital weights in the  $S_{\text{max}}$  models are fixed, the only degree of freedom is the velocity scale  $\Upsilon$ . Results of the corresponding fits are shown in the right-hand panels of Fig. 3. As can be seen, the internal moments of the  $S_{\text{max}}$  models follow closely the analytic profiles, confirming the above speculations about the connection between entropy and isotropy in spherical systems. That the  $S_{\text{max}}$  models in fact match better with the analytical Hernquist profiles than the fits on the left-hand panels implies that the scatter in the fits is mainly caused by noise in the  $N$ -body LOSVDs. Uncertainties in the Schwarzschild code (finite number of orbits and finite numerical resolution) are instead negligible, as otherwise deviations between reference moments and orbit representation would be larger. Likewise, since the  $S_{\text{max}}$  models in the right-hand panels of Fig. 3 are based upon the deprojected  $N$ -body light profiles, noise in the  $N$ -body light profiles is also not the dominant driver for scatter in the left-hand panels.

Concerning  $M/L$  values we find  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 1.007 \pm 0.016$  in the mean over all 10  $S_{\text{max}}$  models. As stated above, the remaining scatter of about 1.5 per cent is due to noise in the  $N$ -body kinematics. We do not expect this scatter to have a significant influence on our results of fits to the merger remnants.

## 5 SCHWARZSCHILD FITS OF MERGER REMNANTS: GENERAL NOTES

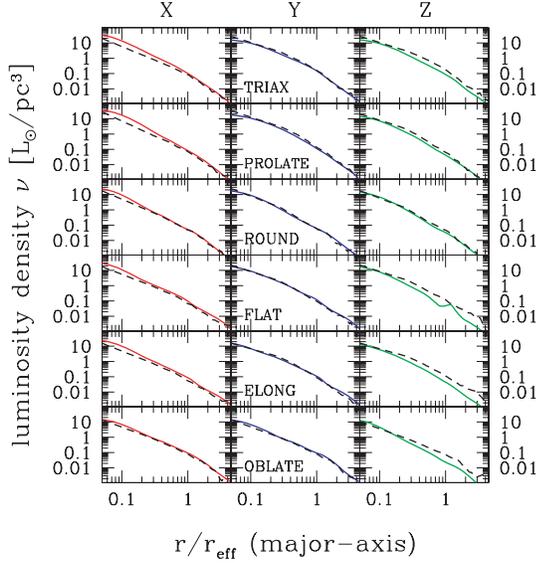
Now we discuss the models of simulated merger remnants. This section contains notes on general properties of the Schwarzschild fits and the deprojections.

### 5.1 Luminosity densities

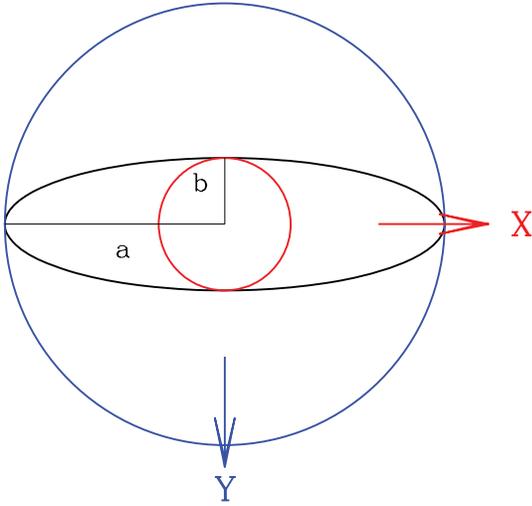
Fig. 4 compares the axisymmetric deprojections with the internal luminosity density profiles of the merger remnants. The figure only compares densities along the projected major axis. Results along other position angles are similar. For the merger remnants, the density is averaged over a plan-parallel wedge of size  $\Delta r \approx 0.05 r_{\text{eff}}$  along the major axis,  $\Delta z \approx 0.2 r_{\text{eff}}$  perpendicular to this axis (in the plane of the sky) and  $\Delta\phi = 45^\circ$  in the plane defined by the line of sight and the projected major axis.

If a remnant is seen along its long axis (left-hand panels), then the axisymmetric deprojection overestimates the density – especially near the centre. The opposite occurs if a merger is seen along the intermediate axis (middle panels): the axisymmetric deprojection of the  $Y$  projections underestimates the remnant density. Note that for the remnant in the bottom row (OBLATE)  $X$  and  $Y$  deprojections are almost equal, consistent with its oblate shape ( $b \approx a$ ).

Fig. 5 illustrates that the viewing-angle dependency of the deprojections reflects the intrinsic non-oblateness  $b/a \neq 1$  of most of



**Figure 4.** Luminosity density of merger remnants (black, dashed) and merger models (coloured, solid). From left- to right-hand panels: models of X, Y and Z projections. Densities are evaluated along the projected major axis.



**Figure 5.** Schematic view on to the  $(X, Y)$  plane of a prolate body ( $b/a < 1$ ). If the body is seen along the X-axis, then a deprojection assuming axial symmetry with the symmetry axis being perpendicular to the  $(X, Y)$  plane, overestimates the density inside  $r < b$  (small circle, red). Correspondingly, if the body is seen from the Y-axis, an axisymmetric deprojection underestimates the density inside  $r < a$  (large circle, blue).

our merger remnants.<sup>3</sup> The light inside an ellipse with  $b < a$ , if seen along the long axis, is quenched into the region  $r < b$  in the axisymmetric deprojection. Accordingly, the mean density of the deprojection inside  $b$  must be larger than the original density inside the same spatial region. Conversely, if the ellipse is viewed side-on, the axisymmetric deprojection stretches the light into the larger region  $r < a$  and, hence, underestimates the true density.

<sup>3</sup> We restrict the discussion to the edge-on case, since all our best-fitting Schwarzschild models have  $i = 90^\circ$  (cf. Section 5.3).

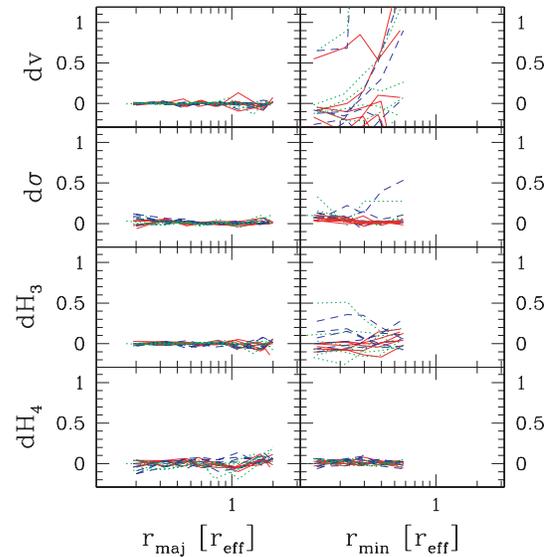
Concerning our merger remnants, deviations between deprojection and intrinsic light profile are the largest where  $b/a$  is the smallest (cf. Fig. 2) – in accordance with the above reasoning. At large radii, the intermediate-to-long axis ratio becomes  $b/a \approx 1$  and the deprojections of X and Y projections approach the luminosity profiles of the remnants.

Concerning the short axis,  $c/a$  quantifies the quenching of light along the line of sight as much as  $b/a$  quantifies it along the intermediate axis. Insofar, the Z projection is similar to the Y projection, which explains why Z deprojections underestimate luminosity densities of the mergers as well. A difference arises at large radii because  $b/a \rightarrow 1$ , whereas  $c/a$  stays roughly constant (e.g. FLAT, ELONG, OBLATE). Consequently, Z deprojections deviate over the whole radial range plotted in Fig. 4 and have a steeper slope than the luminosity profiles of the mergers.

## 5.2 Kinematic fits

Because the merger remnants do not obey oblate axial symmetry it is not clear whether their kinematics can be fitted by our models – which respect this symmetry – at all. Residuals in the kinematic fits are shown in Fig. 6. Except from minor-axis rotation  $v$  and asymmetry of the LOSVD  $H_3$  Schwarzschild models reproduce the data very well, to an accuracy of about a tenth of the assigned ‘error bars’. Since these ‘error bars’ are taken from observations, a comparable degree of triaxiality in real galaxies would be hardly recognisable in terms of a systematic offset between models and data.

Discrepancies between merger remnants and Schwarzschild models in minor-axis profiles of  $v$  and  $H_3$  are the result of oblate axial symmetry enforcing  $v \equiv H_3 \equiv 0$  in the models. Hence, the upper right-hand panel of Fig. 6 in fact shows the amount of minor-axis rotation in the remnants. Neglecting the latter in our fits implies that part of the kinetic energy of the merger remnants is missing in the Schwarzschild models. This could lead to an underestimation of the mass. However, the minor-axis rotation in Fig. 6 is of the order of the assigned error bars ( $d v \lesssim 1$ ), e.g. below 10 per cent of the



**Figure 6.** Residuals between Schwarzschild fits and remnant LOSVDs normalized to the assigned error bars:  $dv \equiv (v_{\text{fit}} - v_{\text{in}})/\Delta v_{\text{in}}$  (and analogously for  $\sigma$ ,  $H_3$  and  $H_4$ ). Left-hand panels: major axis; right-hand panels: minor axis; red/solid: X projections; blue/dashed: Y projections; green/dotted: Z projections.

kinetic energy in the dispersion (cf. radial profiles of  $v$ ,  $\sigma$ ,  $H_3$  and  $H_4$  and their assigned errors in Appendix B). We therefore do not expect that neglecting minor-axis rotation of the merger remnants has a dominant effect on our results.

### 5.3 Inclinations of best-fitting models

Although we probe models at three different inclinations for each remnant projection, the best-fitting model (with  $2 \times 3500$  orbits; cf. Section 3.3) always occurs at an inclination of  $i = 90^\circ$  (edge-on). This is not surprising for  $X$  and  $Y$  projections. However,  $Z$  projections could have been expected to be better represented by nearly face-on models, e.g. with  $i \approx 0^\circ$ .

However, according to the lower panel of Fig. 1 all remnants except the OBLATE one appear flattened when projected along the  $Z$ -axis ( $\epsilon_z > 0$ ). Axisymmetric models, on the other hand, are necessarily round when seen along the axis of symmetry. Thus, an axisymmetric  $i = 0^\circ$  model *cannot* fit the  $Z$  projection of most remnants.

Only one remnant (OBLATE) is close enough to axial symmetry that its  $Z$  projection is almost round. Why is the best-fitting model for this remnant again achieved for  $i = 90^\circ$ ? The main reason is probably the small rotation signal  $v \neq 0$  and  $H_3 \neq 0$  along the apparent major axis of its  $Z$  projection (the face-on view is not exactly round, cf. Fig. 1). At a viewing angle of  $i = 90^\circ$  the model can adjust the balance between prograde and retrograde orbits to fit  $v \neq 0$  and  $H_3 \neq 0$ . Instead, any rotation and asymmetric deviation from a Gaussian LOSVD disappear when looking at an axisymmetric system face-on:  $v \equiv H_3 \equiv 0$  (for all position angles). Thus, everything else being equal, a face-on model will necessarily have a larger  $\chi^2$  than an edge-on model. In fact 83 per cent of the  $\Delta\chi^2$  between the best-fitting edge-on and the best-fitting face-on model<sup>4</sup> of the OBLATE remnant, respectively, is due to differences in the fit to  $v$  and  $H_3$ . This is not a proof, but a strong indication that the residual rotation in the  $Z$  projection of the OBLATE remnant is the main driver for the best-fitting model to occur at an inclination of  $i = 90^\circ$ .

A triaxial dynamical system can exhibit various degrees of rotation in the  $Z$  projection. If this indeed causes the corresponding axisymmetric fit to prefer an inclination of  $i = 90^\circ$ , then the inclination mismatch is an unavoidable consequence of the false symmetry assumption. Concerning models of real galaxies, an additional complication enters through measurement errors: even for an exactly axisymmetric face-on galaxy one would determine  $v \neq 0$  and  $H_3 \neq 0$  due to measurement uncertainties. In such a case, a best-fitting axisymmetric inclination of  $90^\circ$  would be an artefact related to the ability of the modelling machinery to fit the noise in the data. Proper regularization could provide a way out of the inclination mismatch then. For the  $Z$  models of the OBLATE remnant we find indeed a best-fitting inclination  $i = 0^\circ$  for  $\alpha < 0.005$  (strong regularization).

A systematic investigation of the question whether noise in real data can bias axisymmetric models towards  $i = 90^\circ$  and whether this possible bias can be reduced by using proper regularization is out of the scope of this paper. For simplicity, we adopt the same regularization scheme to all merger remnants in the following. We expect this to significantly affect only the fits to the face-on projection of

the OBLATE remnant. Specifically, assuming the wrong inclination makes our  $Z$  model worse than it could possibly be with optimized regularization. In all other remnants the inclination mismatch is due to intrinsic non-axisymmetry.

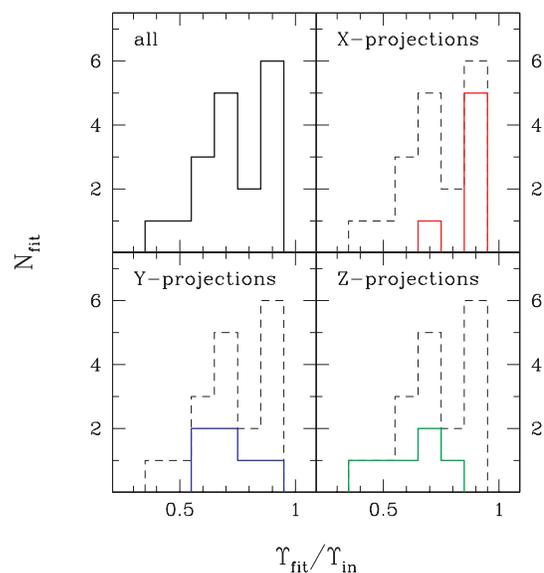
## 6 MASS DISTRIBUTION IN REMNANTS AND MODELS

Having discussed general features of the Schwarzschild models we now turn to the comparison of the mass distribution in models and the corresponding merger remnants.

### 6.1 Stellar mass-to-light ratio

Fig. 7 shows the distribution of (scaled)  $M/L$  values  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  obtained from our best-fitting Schwarzschild models. The reconstructed  $\Upsilon_{\text{fit}}$  systematically underestimate  $\Upsilon_{\text{in}}$ . 17 out of the 18 models have  $M/L$  values in the range  $0.5 \leq \Upsilon_{\text{fit}}/\Upsilon_{\text{in}} \leq 0.9$ , one model has  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.4$ . Separating the results according to the viewing angle yields that models of  $X$  projections (shortly  $X$  models below) recover the true  $M/L$  very well ( $\Upsilon_{\text{fit}} = 0.9 \Upsilon_{\text{in}}$  in all but one case; cf. upper right-hand panel of Fig. 7). In contrast, Schwarzschild models of  $Y$  and  $Z$  projections have  $M/L$  values distributed almost homogeneously in the range  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.5\text{--}0.9$  (bottom panels of Fig. 7). *Although* the luminosity density of the deprojection predicts less light in the corresponding models than there is in the merger. The reason for this behaviour will be further discussed in Sections 9 and 10.

It should be restated that our mock observations are not drawn from random projections. Therefore, Fig. 7 does not equal the distribution of  $M/L$  values that would result from modelling real galaxies (even if they would be structurally similar to the merger remnants). The most significant result here is that axisymmetric models tend to *underestimate* the mass fraction that follows the light. We have no proof for the generality of this result, but since we have modelled all three principal projections for each remnant we do not expect models from other viewing angles to deliver  $\Upsilon_{\text{fit}} > \Upsilon_{\text{in}}$ .



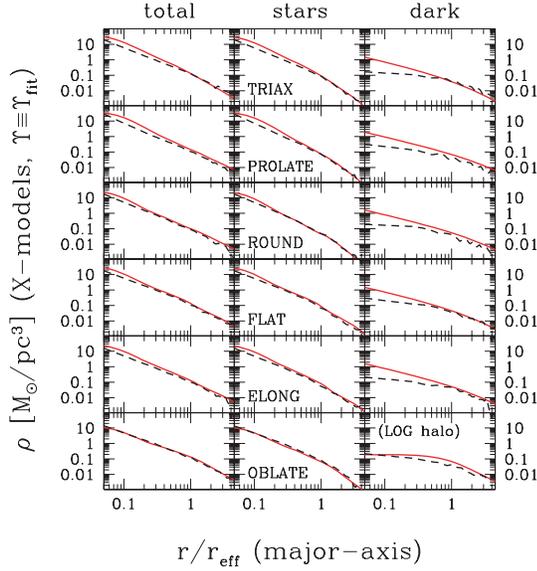
**Figure 7.** Distribution of best-fitting mass-to-light ratios  $\Upsilon_{\text{fit}}$  (scaled by the input value  $\Upsilon_{\text{in}}$ ). Top left-hand panel: whole sample; other panels: results for different projections (indicated in each panel; dashed lines: total distribution for comparison).

<sup>4</sup> The deprojection of axisymmetric bodies at  $i = 0^\circ$  is infinitely uncertain (Rybicki 1987; Gerhard & Binney 1996). For the above comparison a face-on model has been constructed using the ( $i = 90^\circ$ ) deprojection of the  $X$  projection. If the remnant would be exactly axisymmetric, then this deprojection would uniquely recover its intrinsic luminosity distribution.

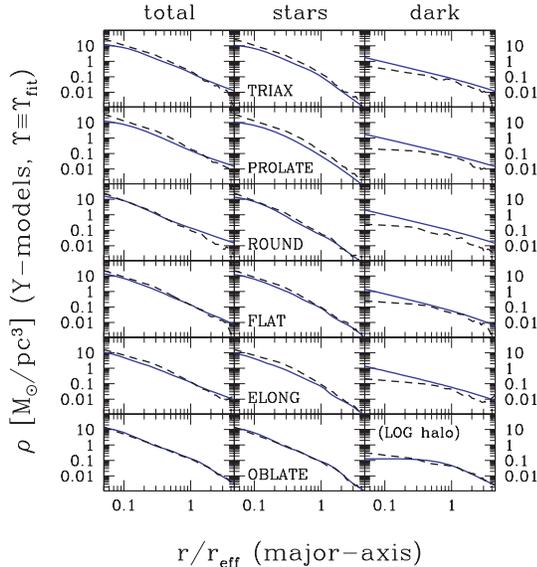
## 6.2 Mass densities

Our Schwarzschild models (and the merger remnants as well) contain both luminous as well as dark mass and  $\Upsilon$  only represents a fraction of the total mass. The next question is how well total and dark matter density profiles are represented in the Schwarzschild models. To explore this, Figs 8–10 survey radial density profiles of models and remnants separately for the three principal projections. The figures show intrinsic densities along the projected major axis. The middle panels (luminous mass density) differ from the  $\nu$  profiles of Fig. 4 only in the scaling (the stellar mass density equals  $\Upsilon_{\text{fit}} \nu$ ; cf. equation 3).

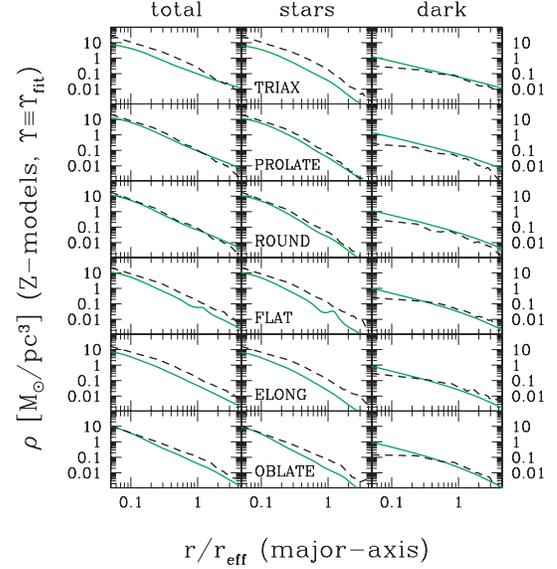
Evidently, in  $X$  models not only the luminosity density, but also the total mass in the inner regions is overestimated. Exceptional is the  $X$  model of the OBLATE remnant: because the remnant is close



**Figure 8.** Comparison of remnant (black/dashed) and Schwarzschild model (coloured/solid) mass density profiles (along projected major axis). Left-hand/middle/right-hand panels: total/luminous/dark mass. The figure surveys results from modelling  $X$  projections.



**Figure 9.** As Fig. 8, but for  $Y$  projections.



**Figure 10.** As Fig. 8, but for  $Z$  projections.

to axial symmetry, no overestimation of the central density occurs. In  $Y$  and  $Z$  models – parallel to the underestimation of the light – also the total mass density is underestimated. Again exceptional is the  $Y$  model of the OBLATE remnant: the total mass is well recovered. This reflects again the axial symmetry of the remnant, according to which  $X$  and  $Y$  projections are equivalent and both should allow a good reconstruction with our models.

The case of the OBLATE remnant also reveals a slight degeneracy in the mass recovery. The best-fitting  $X$  model has  $\Upsilon_{\text{fit}} = 0.7$ , while the best-fitting  $Y$  model is obtained with  $\Upsilon_{\text{fit}} = 0.9$ . Despite these different  $\Upsilon_{\text{fit}}$ , the total mass inside  $r_{\text{eff}}$  is recovered with high accuracy in both models: 2.8 per cent fractional accuracy in the  $X$  model and 0.4 per cent in the  $Y$  model, respectively. Thus, the total mass can be recovered with about the same accuracy, even if luminous masses differ by about 20 per cent.

Independent of projection, central dark matter densities are overestimated in all Schwarzschild models. Most likely, this reflects our choice of NFW profiles for the haloes of the models (cf. Section 3.1). In principle, an overestimation of the central dark matter density could cause an underestimation of the luminous mass for compensation. Near the centre, where the dark matter excess is most prominent, the luminous matter is, however, still a factor of 10 larger than the dark matter density (at  $0.1r_{\text{eff}}$ , for example). We therefore do not expect the central overprediction of dark matter to be important for the recovery of  $\Upsilon$ . Moreover, while the dark matter excess is projection *independent*, the underestimation of  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  is projection *dependent*.

Nevertheless we have additionally calculated a set of logarithmic (LOG) haloes for one merger remnant (OBLATE; the grid used to sample the haloes is described in Thomas et al. 2007). In case of the  $X$  and  $Y$  projections LOG haloes allow a slightly better fit than NFW haloes (cf. bottom right-hand panels of Figs 8 and 9). As will become clear from the discussions in the next sections, these models are in no respect systematically different from the models of other remnants which are calculated with NFW haloes.

We have also calculated logarithmic haloes for the  $X$  models of the TRIAX, PROLATE and ELONG remnants. In these cases as well as concerning the  $Z$  model of the OBLATE remnant, LOG haloes do not provide better fits. As a consequence, considering

LOG haloes does not change  $\Upsilon_{\text{fit}}$  of this models. It follows that the particular choice of the halo profile (between NFW and LOG) has little effect on our results. It merely influences the match to the dark matter component in a spatial region, where dark matter is a minor contributor to the total mass.

Towards the outer edge of the kinematical data ( $r_{\text{eff}} \lesssim r \lesssim 2r_{\text{eff}}$ ), mass densities of Schwarzschild models and merger remnants agree reasonably well. This holds for the total mass, as well as for luminous and dark components, separately. Around  $1-2r_{\text{eff}}$ , integrated total masses of Schwarzschild models are accurate to about 20 per cent. The  $Z$  models of the most flattened remnants (FLAT, ELONG, OBLATE) are deviant by up to 40–50 per cent.

## 7 VELOCITY ANISOTROPIES IN REMNANTS AND MODELS

We now consider in more detail the internal dynamical structure of the merger remnants and how it is represented by our Schwarzschild fits.

### 7.1 Anisotropy profiles

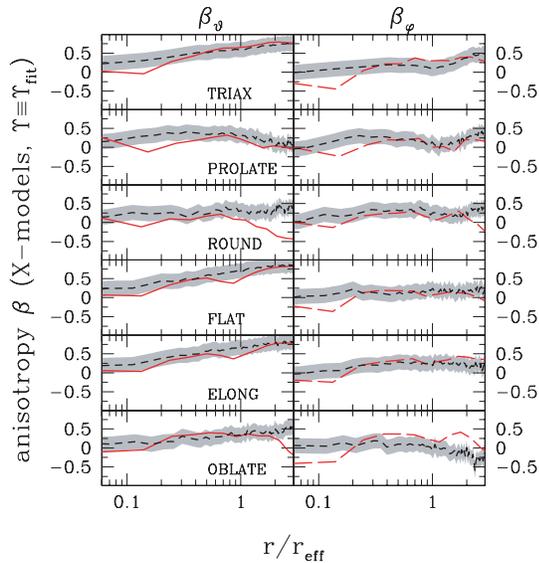
Figs 11–13 compare profiles of meridional anisotropy

$$\beta_{\vartheta} \equiv 1 - \frac{\sigma_{\vartheta}^2}{\sigma_r^2} \quad (6)$$

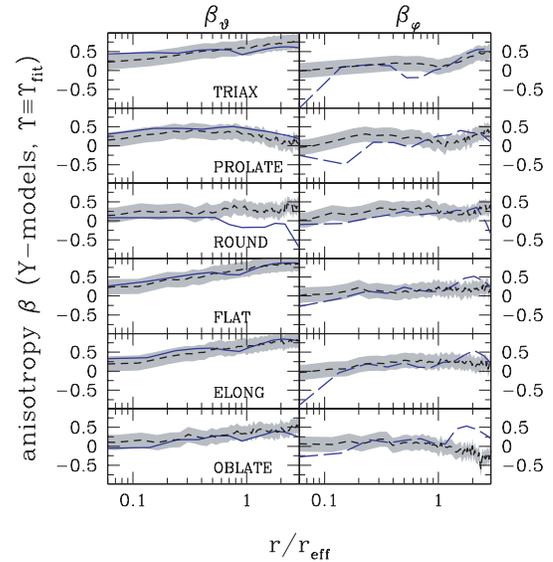
and azimuthal anisotropy

$$\beta_{\varphi} \equiv 1 - \frac{\sigma_{\varphi}^2}{\sigma_r^2} \quad (7)$$

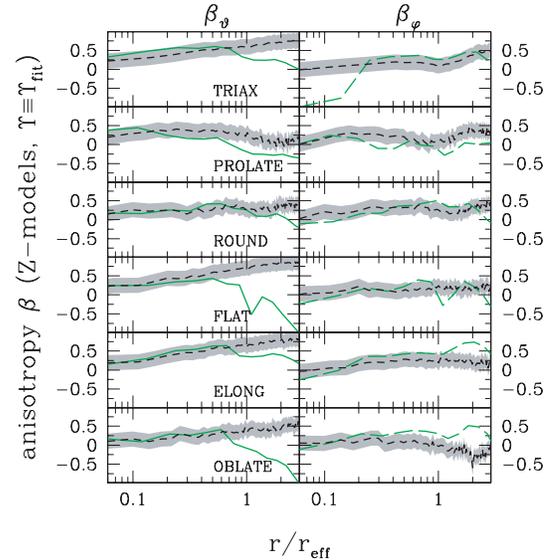
of Schwarzschild models and merger remnants. We use spherical coordinates  $r$ ,  $\vartheta$  and  $\varphi$ , oriented along the principal axes such that  $\varphi$  is the azimuth in the  $(X, Y)$  plane and  $\vartheta$  is the latitude. The velocity dispersions are luminosity-weighted spherical averages.



**Figure 11.** Comparison of velocity anisotropies. Left-hand panels: meridional  $\beta_{\vartheta}$  of Schwarzschild models (solid/coloured) and of merger remnants (short-dashed/black); right-hand panels: azimuthal  $\beta_{\varphi}$  of Schwarzschild models (long-dashed/coloured) and of merger remnants (short-dashed/black); grey:  $\pm 0.2$  absolute deviations from merger values for comparison. The figure shows results of  $X$  models.



**Figure 12.** As Fig. 11, but for  $Y$  models.

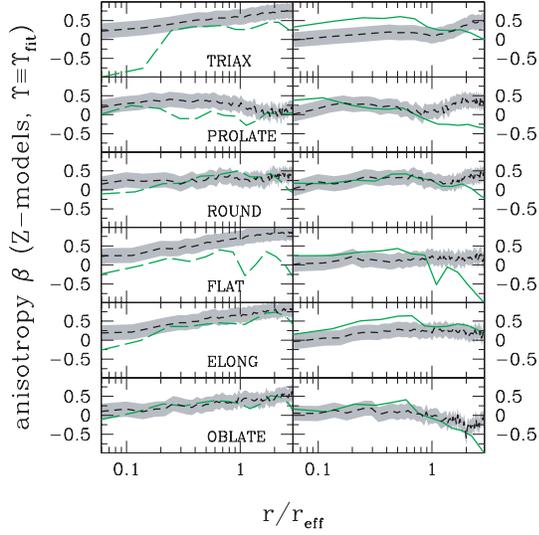


**Figure 13.** As Fig. 11, but for  $Z$  models.

In Figs 11–13  $\Delta\beta = \pm 0.2$  margins are highlighted. The choice of these margins is arbitrary, and is only to guide a quantification of deviations between mergers and models. In most cases these are smaller than  $\Delta\beta < 0.2$ . But there are some outliers (mostly among  $Z$  models). As a general rule,  $X$  and  $Y$  models fit better with the intrinsic properties of the merger remnants than  $Z$  models.

The mismatch of the  $Z$  models is partly due to the fact that the best-fitting Schwarzschild models are always achieved for an inclination  $i = 90^\circ$  (cf. Section 5.3). Apart from the related mismatch in the deprojection it raises a complication concerning the comparison of the internal moments: in the  $Z$  projection of a merger remnant, according to the above definitions, the azimuth  $\varphi$  appears as the angle in the plane of the sky. In the Schwarzschild models, however,  $\vartheta$  as defined above is the angle in the plane of the sky, as long as  $i = 90^\circ$ .

Much of the discrepancies between Schwarzschild models and merger remnants can be attributed to these different coordinate definitions. To show this Fig. 14 replicates the same profiles as



**Figure 14.** As Fig. 13, but Schwarzschild  $\beta_\phi$  (long-dashed) are compared to merger remnant  $\beta_\phi$  (left-hand panels) and Schwarzschild  $\beta_\phi$  (solid) are compared with  $\beta_\phi$  of the remnants (right-hand panels).

Fig. 13, but  $\beta_\phi$  of the Schwarzschild models is now compared to  $\beta_\phi$  of the merger remnants and vice versa. The differences between the mergers and the models are significantly smaller in Fig. 14 than in Fig. 13, especially among the most strongly flattened remnants.

## 7.2 Interpretation in terms of orbits

The remaining deviations between the anisotropy profiles of merger remnants and their corresponding Schwarzschild fits are most likely related to the different orbit families supported by  $N$ -body potentials on one side and axisymmetric potentials on the other.

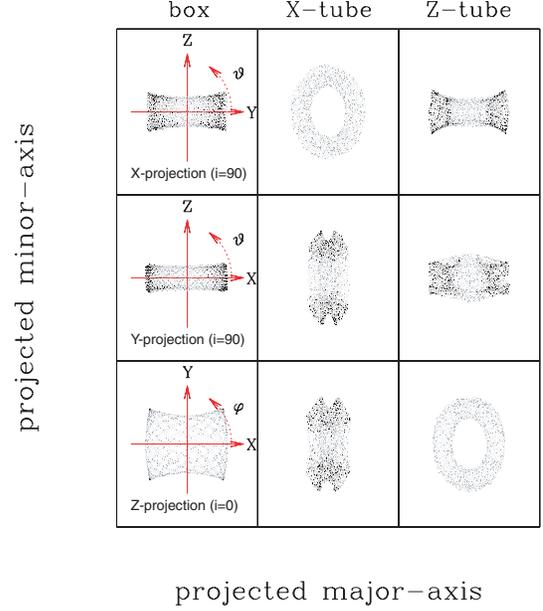
Figs 15 and 16 review principal projections of orbits numerically integrated in an  $N$ -body potential (Fig. 15) and in an axisymmetric potential (Fig. 16). Regions with  $v_{\text{los}} > v_\perp$  are plotted dark and regions with  $v_{\text{los}} < v_\perp$  are plotted grey. Thereby  $v_{\text{los}}$  is the absolute line-of-sight velocity in the given projection and  $v_\perp$  is the absolute magnitude of the velocity perpendicular to the line of sight. In dark areas most of the kinetic energy of an orbit is directed towards the observer, whereas in grey areas most of the kinetic energy is in motion perpendicular to the line of sight.

The tangential anisotropy of the X model for the PROLATE remnant can be explained by the dominance of X tubes in this remnant. According to Fig. 15 their round appearance in the X projection makes them most similar to the edge-on projection of axisymmetric shell orbits (cf. Fig. 16). The latter, in turn, have large  $\sigma_\phi$  and low  $\sigma_r$ , and cause the tangential anisotropy in the Schwarzschild model.

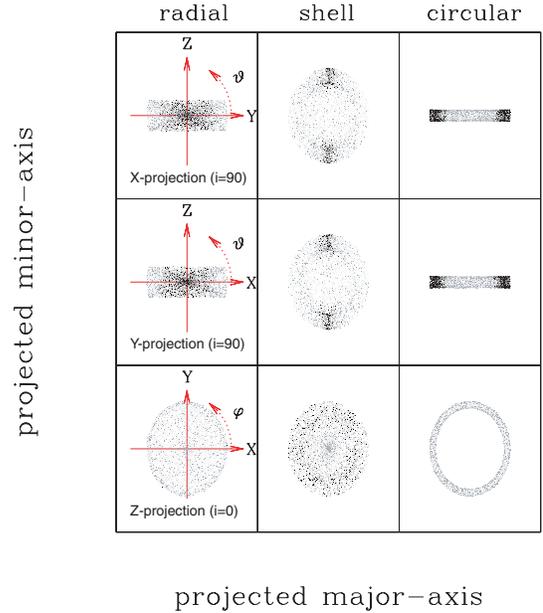
Likewise, the similarity of  $\beta_\phi$  in Z models with  $\beta_\phi$  of the merger remnants discussed at the end of Section 7.1 can be explained by the fact that the dominant orbits in the outer parts of merger remnants, Z tubes, appear nearly round when seen face-on. Again, they are likely mapped on to axisymmetric shell orbits, with the same consequence for the model's anisotropy as discussed for the X model of the PROLATE remnant.

## 8 MODELLING UNCERTAINTIES

Up to now we have presented the viewing-angle dependency of the masses and anisotropies which we reconstructed with our axisymmetric orbit models. The behaviour of the anisotropy could be



**Figure 15.** Numerically integrated orbits in an  $N$ -body potential. From left to right-hand panels: box orbit, X tube and Z tube. From top to bottom: X, Y and Z projections. Coordinate definitions are illustrated on the left-hand side. Black:  $v_{\text{los}} > v_\perp$ ; light grey:  $v_{\text{los}} < v_\perp$  (details in the text).



**Figure 16.** As Fig. 15, but for the case of an axisymmetric potential. From left- to right-hand panels: radial, shell-like and nearly circular orbit.

explained by the way in which projected properties of major remnant orbit families match with different axisymmetric orbits. The recovered masses are less easy to understand, in particular the low  $\Upsilon_{\text{fit}}$ . This section and Sections 9 and 10 are aimed to discuss the mass recovery in more detail. We start this discussion here by investigating whether the projection dependency of the mass recovery in our axisymmetric dynamical models is an artefact of the modelling machinery.

### 8.1 Stationarity assumption

As it has been stated in Section 1, the Schwarzschild method is based on Jeans' theorem and the assumption that the object to be modelled is stationary. Non-stationarity of the merger remnants can have a significant influence on the recovered masses. For example, if a remnant contracts because the ratio of its kinetic and potential energies is smaller than in virial equilibrium then a stationary model could deliver a mass smaller than the true one. Likewise, if a remnant expands then the recovered  $M/L$  could be too large.

Stationarity or virial equilibrium, respectively, implies that

$$2T_{ij} = -W_{ij}, \quad (8)$$

where  $T_{ij}$  denotes the kinetic energy tensor and  $W_{ij}$  denotes the potential energy tensor. Equation (8) holds for the luminous and the dark components separately, if both are stationary. In the following we only consider the luminous component. The calculation of its kinetic and potential energies is straightforward:

$$T_{ij} = \frac{1}{2} \sum_{\alpha=1}^{N_l} m_{\alpha} \dot{x}_i^{(\alpha)} \dot{x}_j^{(\alpha)}, \quad (9)$$

where the sum extends over all the  $N_l$  luminous particles of the merger remnant (with mass  $m_{\alpha}$  each) and  $\dot{x}_i^{(\alpha)}$  is the  $i$  component of the velocity of particle  $\alpha$ . The potential energy of the luminous component comprises the two contributions

$$W_{ij} = W_{ij}^{(ll)} + W_{ij}^{(ld)}, \quad (10)$$

where

$$W_{ij}^{(ll)} = -G \sum_{\alpha=1}^{N_l-1} m_{\alpha} \sum_{\beta=\alpha+1}^{N_l} m_{\beta} \frac{\xi_{ij}^{(\alpha\beta)}}{r_{\alpha\beta}^3} \quad (11)$$

and

$$\xi_{ij}^{(\alpha\beta)} \equiv (x_i^{(\alpha)} - x_i^{(\beta)}) (x_j^{(\alpha)} - x_j^{(\beta)}). \quad (12)$$

The sum extends over luminous particles only and  $r_{\alpha\beta}$  is the distance between particles  $\alpha$  and  $\beta$ , respectively. The contribution of dark matter comes in through

$$W_{ij}^{(ld)} = -G \sum_{\alpha=1}^{N_l} m_{\alpha} \sum_{\beta=1}^{N_d} \mathcal{M}_{\beta} \frac{\xi_{ij}^{(\alpha\beta)}}{r_{\alpha\beta}^3}. \quad (13)$$

In the last sum  $N_d$  denotes the total number of dark matter particles in the remnant (with mass  $\mathcal{M}_{\beta}$  each). In virial equilibrium total potential energy  $W = \sum W_{ii}$  and total kinetic energy  $T = \sum T_{ii}$  obey  $2T/|W| = 1$ .

The six modelled merger remnants have  $2T/|W| \in [0.960, 0.981]$  with the lowest value for the FLAT remnant and the largest value for the OBLATE one. Thus, the remnants are very close to virial equilibrium and we expect that the assumption of stationarity in the models should affect the models' masses at most at the 5 per cent level. Hence it is not the main driver for the low  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  in our models. In addition  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  is projection dependent whereas  $2T/|W|$  is projection independent.

### 8.2 Phase-space sampling

Another potential uncertainty in the modelling procedure is the difference in phase-space structure of merger remnants on one side and Schwarzschild models on the other: while the remnants are composed of a relatively large number of particles, each sampling a different orbit at one point, the Schwarzschild model is composed

of a relatively low number of orbits, each sampled very densely (we use about  $10^5$  time-steps for each orbit integration).

Concerning the sampling of the orbit (the time-step and total integration time used), our implementation of Schwarzschild's method has been successfully tested on continuous analytical dynamical models, like e.g. a Hernquist sphere (Thomas et al. 2004). To check whether a similarly good agreement can be achieved when modelling  $N$ -body systems, we have repeated the tests with discrete  $N$ -body realizations as modelling targets (cf. Section 4). The small uncertainties that we find imply that differences in phase-space structure are negligible.

### 8.3 Chaotic orbits

In the implementation of Schwarzschild's method applied here (as in most others) chaotic orbits are treated in the same way as regular orbits. This is not necessary in Schwarzschild models, but makes them computationally more efficient. A chaotic region in phase space at fixed  $E$  and  $L_z$  has to have a constant phase-space density according to Jeans' theorem. If such a region is represented by one (chaotic) orbit in the library, then the method works fine. However, it may happen that the (finite) integration time of the first orbit that is launched in the chaotic region is insufficient to cover the accessible phase-space volume entirely. Then the program will launch one or more other orbits to fill up the rest of the chaotic region. It is then likely (although not necessary) that these fractional orbits will have different phase-space densities in the final model. As a consequence, the model no longer satisfies Jeans' theorem. Several suggestions have been made to overcome this problem (e.g. Merritt & Fridman 1996; Häfner et al. 2000).

Since the main consequence of chaos in phase space is to break the stationarity of the Schwarzschild models, it should manifest itself in deviations from virial equilibrium and, thus, can be quantified by evaluating the virial equations of the Schwarzschild models. To calculate the kinetic energy tensors  $T_{ij}$  and  $W_{ij}$  defined in equations (9) and (10) we have constructed  $N$ -body realizations of each best-fitting Schwarzschild model as described in Appendix A.

For the obtained virial ratios we find  $2T/|W| \approx 1$  to within 15 per cent. This limits the amount of chaos in our orbit libraries. Deviations from  $2T/|W| = 1$  are not correlated with viewing angle but with halo concentration, which is an artefact of the  $N$ -body realization and further discussed in Appendix A. Thus, the margins for intrinsic non-stationarity are even smaller than the above quoted 15 per cent. The related uncertainties are not sufficient to explain the trends in the mass recovery.

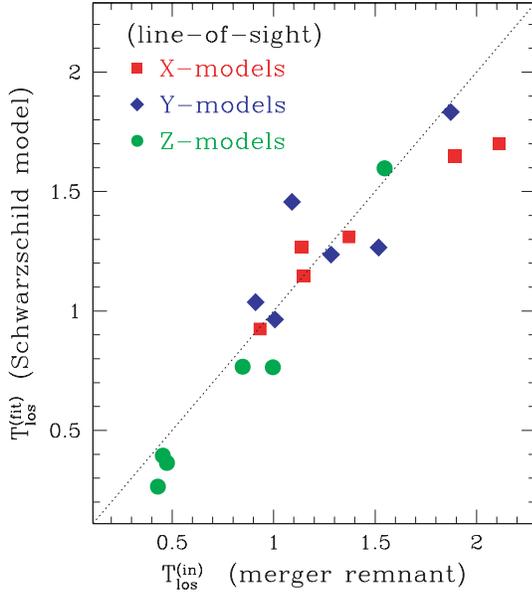
## 9 THE VIEWING ANGLE DEPENDENCY OF THE TOTAL MASS RECOVERY

The last section has ruled out modelling uncertainties as the main source for the magnitude and projection dependency of the mass recovery. We now investigate whether the different 3D shapes of models and merger remnants are the main driver of this dependency.

Globally, mass and kinetic energy are linked by the virial theorem,  $2T = |W| = \kappa M$ , where  $\kappa$  depends on the density profile. Accordingly, reconstructed masses  $M^{(\text{fit})}$  and input masses  $M^{(\text{in})}$  are related via

$$\frac{M^{(\text{fit})}}{M^{(\text{in})}} = \frac{\kappa^{(\text{in})} T^{(\text{fit})} + T_{\text{DM}}^{(\text{fit})}}{\kappa^{(\text{fit})} T^{(\text{in})} + T_{\text{DM}}^{(\text{in})}}, \quad (14)$$

with  $T^{(\text{fit})}$  and  $T_{\text{DM}}^{(\text{fit})}$  denoting the kinetic energy of luminous and dark matter in the Schwarzschild fit and  $T^{(\text{in})}$  and  $T_{\text{DM}}^{(\text{in})}$  being the analogue



**Figure 17.** Correlation between kinetic energies of merger remnants (in) and Schwarzschild models (fit; along the line of sight). Dotted line: one-to-one relation for comparison.

quantities of the merger remnant, respectively. Since it is basically  $T^{(\text{fit})}$  that is constrained by the LOSVD fits, it is instructive to study first the energy budget of the Schwarzschild models in comparison to the merger remnants. Then, equation (14) can be used to evaluate the implications on the reconstructed masses.

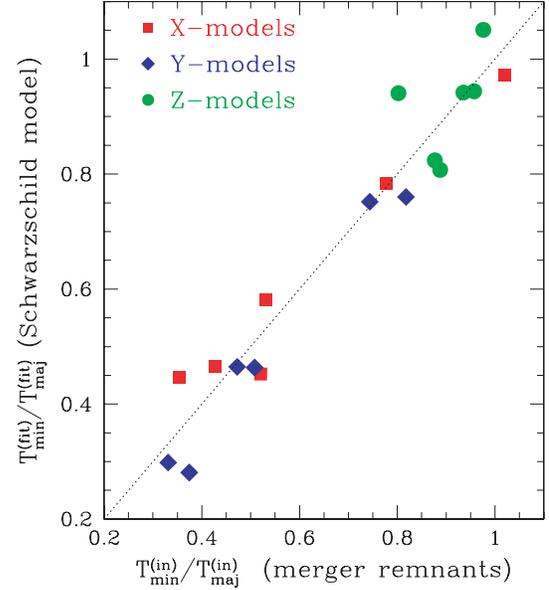
Because the virial theorem relates energies to *total* masses this section deals with the viewing-angle dependency of the total mass recovery. The  $M/L$  of the stellar component will be discussed separately in Section 10.

### 9.1 Energy budget of the Schwarzschild models

In the merger remnants, by definition of the axes (cf. Section 2)  $T_{xx} \geq T_{yy} \geq T_{zz}$ , whereas oblate axial symmetry implies  $T_{xx} \equiv T_{yy} \geq T_{zz}$  in the Schwarzschild fits.<sup>5</sup> In the following, it is convenient to switch from axis labels referring to the intrinsic shape of the remnant (e.g.  $X$ ,  $Y$  and  $Z$  as defined in Section 2) to projection-based labels: let us define the kinetic energy  $T_{\text{los}}$  as the energy parallel to the axis that points towards the observer,  $T_{\text{maj}}$  as the energy parallel to the axis that projects to the apparent major axis and  $T_{\text{min}}$  as the energy directed parallel to the apparent minor axis.

Figs 17 and 18 show that the line-of-sight energy  $T_{\text{los}}$  and the ratio  $T_{\text{min}}/T_{\text{maj}}$  of the two transversal energies are well recovered by the Schwarzschild models. This could have been expected since  $T_{\text{los}}$  is the energy mapped by the projected kinematics. A mismatch in  $T_{\text{los}}$  should manifest itself in the kinematic fits. Some scatter remains, however, because we do not assume full sky coverage with kinematic data. That the Schwarzschild models match also with  $T_{\text{min}}/T_{\text{maj}}$  of the remnants is plausible, because this energy ratio determines the shape (e.g. Binney & Tremaine 1987). And the shape is accounted

<sup>5</sup> We will only consider the diagonal elements  $T_{ii}$  in the following, because in the merger remnants as well as in the  $N$ -body realizations of the Schwarzschild models the other components are at least two orders of magnitude lower and, thus, energetically negligible.



**Figure 18.** Correlation between the ratio  $T_{\text{min}}/T_{\text{maj}}$  of the two transversal energies in merger remnants (in) and Schwarzschild models (fit), respectively. Dotted: one-to-one relation for comparison.

for in the Schwarzschild fits through the deprojected luminosity density, which is used as a boundary condition for our models.

The two relations revealed by Figs 17 and 18 have several important implications for the energy budget and, thus, the recovered masses of the Schwarzschild models.

To see this, let us assume for simplicity that

$$T_{\text{los}}^{(\text{fit})} \equiv T_{\text{los}}^{(\text{in})} \quad (15)$$

and

$$\frac{T_{\text{min}}^{(\text{fit})}}{T_{\text{maj}}^{(\text{fit})}} \equiv \frac{T_{\text{min}}^{(\text{in})}}{T_{\text{maj}}^{(\text{in})}} \quad (16)$$

(in other words, we replace the two approximate one-to-one correlations of Figs 17 and 18 by identities). By symmetry  $T_{\text{maj}}^{(\text{fit})} \equiv T_{\text{los}}^{(\text{fit})}$  in the Schwarzschild fits and it follows that

$$T_{\text{maj}}^{(\text{fit})} \equiv T_{\text{los}}^{(\text{in})}. \quad (17)$$

Moreover, according to equation (16)

$$T_{\text{min}}^{(\text{fit})} \equiv T_{\text{maj}}^{(\text{fit})} \frac{T_{\text{min}}^{(\text{in})}}{T_{\text{maj}}^{(\text{in})}}, \quad (18)$$

and by equation (17) we have

$$T_{\text{min}}^{(\text{fit})} \equiv T_{\text{los}}^{(\text{in})} \frac{T_{\text{min}}^{(\text{in})}}{T_{\text{maj}}^{(\text{in})}}. \quad (19)$$

Hence, the two relations (15) and (16) uniquely link the three relevant components of the kinetic energy tensor of the Schwarzschild model to the energy components of the merger remnant via equations (15), (17) and (19). Note that this holds only for edge-on models. If a model is calculated at  $i = 0^\circ$ , then  $T_{\text{min}}^{(\text{fit})} \equiv T_{\text{maj}}^{(\text{fit})}$  by symmetry. In this case, even the identity  $T_{\text{los}}^{(\text{fit})} = T_{\text{los}}^{(\text{in})}$  does not constrain the two ratios  $T_{\text{los}}^{(\text{fit})}/T_{\text{min}}^{(\text{fit})}$  and  $T_{\text{los}}^{(\text{fit})}/T_{\text{maj}}^{(\text{fit})}$ , respectively. This reflects the uncertainty in the flattening along the line of sight of the model.

To express the expected energy budget for a Schwarzschild fit of a given merger projection more quantitatively, it is necessary to figure out to which intrinsic axes  $T_{\text{los}}$ ,  $T_{\text{maj}}$  and  $T_{\text{min}}$  correspond.

**Table 2.** Ratios of kinetic energies of Schwarzschild fits and merger remnants according to equations (15), (17) and (19). (1) Projection; (2) line-of-sight energy; (3) projected long-axis energy; (4) projected short-axis energy; (5) total energy ( $T = T_{\text{los}} + T_{\text{maj}} + T_{\text{min}}$ ).

Projection	$\frac{T_{\text{los}}^{(\text{fit})}}{T_{\text{los}}^{(\text{in})}}$	$\frac{T_{\text{maj}}^{(\text{fit})}}{T_{\text{maj}}^{(\text{in})}}$	$\frac{T_{\text{min}}^{(\text{fit})}}{T_{\text{min}}^{(\text{in})}}$	$\frac{T^{(\text{fit})}}{T^{(\text{in})}}$
(1)	(2)	(3)	(4)	(5)
X	1	$\frac{T_{xx}^{(\text{in})}}{T_{yy}^{(\text{in})}} \geq 1$	$\frac{T_{xx}^{(\text{in})}}{T_{yy}^{(\text{in})}} \geq 1$	$\geq 1$
Y	1	$\frac{T_{yy}^{(\text{in})}}{T_{xx}^{(\text{in})}} \leq 1$	$\frac{T_{yy}^{(\text{in})}}{T_{xx}^{(\text{in})}} \leq 1$	$\leq 1$
Z	1	$\frac{T_{zz}^{(\text{in})}}{T_{xx}^{(\text{in})}} \leq 1$	$\frac{T_{zz}^{(\text{in})}}{T_{xx}^{(\text{in})}} \leq 1$	$\leq 1$

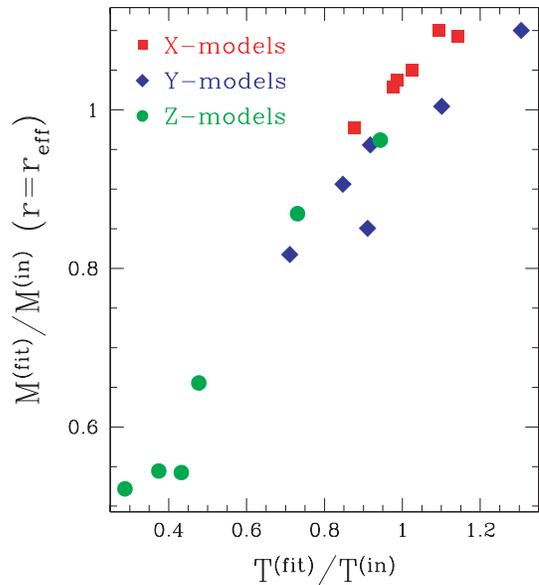
Simple algebra leads to Table 2, in which Schwarzschild model energies relative to remnant energies are given explicitly for each projection. The table allows us to draw some important conclusions. In fact, for  $X$  models the energy in the Schwarzschild model has to be larger than in the remnant,  $T^{(\text{fit})}/T^{(\text{in})} \geq 1$ , unless the remnant is oblate-axisymmetric. Contrary, in  $Y$  and  $Z$  models  $T^{(\text{fit})}/T^{(\text{in})} \leq 1$  and the energy in the Schwarzschild model has to be smaller than in the remnant.

## 9.2 Mass budget of the Schwarzschild models

What are the consequences for the masses of the Schwarzschild fits? For simplicity let us first assume that  $\kappa^{(\text{fit})} = \kappa^{(\text{in})}$ . Then, equation (14) predicts that the ratio  $M^{(\text{fit})}/M^{(\text{in})}$  of reconstructed and input masses equals the ratio of the corresponding total kinetic energies (luminous + dark). However, due to the lack of kinematic information about the constituents of the dark halo, the models have no access to the dark matter kinetic energy of the remnants and  $T_{\text{DM}}^{(\text{fit})}$  is not constrained. As a consequence, the results on the energy budget are significant for the comparison of reconstructed and input masses only inside a radius where luminous matter dominates, e.g. inside  $r_{\text{eff}}$ . There, the contribution of the dark matter kinetic energy is small. Assuming that its contribution is in fact negligible, then equation (14) and the last column of Table 2 imply  $M^{(\text{fit})}/M^{(\text{in})} \geq 1$  inside  $r_{\text{eff}}$  for  $X$  models and  $M^{(\text{fit})}/M^{(\text{in})} \leq 1$  for  $Y$  and  $Z$  models. The amount by which  $M^{(\text{fit})}$  exceeds  $M^{(\text{in})}$  in  $X$  models should be comparable to the amount by which  $M^{(\text{fit})}$  is reduced relative to  $M^{(\text{in})}$  in  $Y$  models and masses of  $Z$  models should be smaller than those of  $Y$  models. Fig. 19 relates the mass ratio  $M^{(\text{fit})}/M^{(\text{in})}$  (at the effective radius) to kinetic energies  $T^{(\text{fit})}/T^{(\text{in})}$  and globally confirms the just discussed trends between reconstructed masses and kinetic energies.

Hence, at the half-light radius, masses of our Schwarzschild fits are closely related to the energy budget of the models. The energy, in turn, derives from the fit to the kinematics and the shape of the modelling target via relations (15) and (16). For edge-on systems, the restrictions imposed by the assumption of axial symmetry together with these two relations already uniquely determine the luminous kinetic energy and, hence, the mass budget of the axisymmetric fits. Thereby it turns out that  $X$  models have to overestimate the true mass, while  $Y$  and  $Z$  models have to underestimate it.

We close this section with a few further comments on Fig. 19. According to equation (14), the relation between reconstructed masses and reconstructed energies can be tilted with respect to a one-to-one relation if the  $\kappa$  ratio varies systematically over the sample. In Section 5.1 we have discussed systematic variations of the deprojections with viewing angle which could cause the tilt with respect



**Figure 19.** Accuracy of reconstructed total (luminous + dark) mass inside  $r_{\text{eff}}$  versus ratio  $T^{(\text{fit})}/T^{(\text{in})}$  of total kinetic energies in Schwarzschild fits and merger remnants.

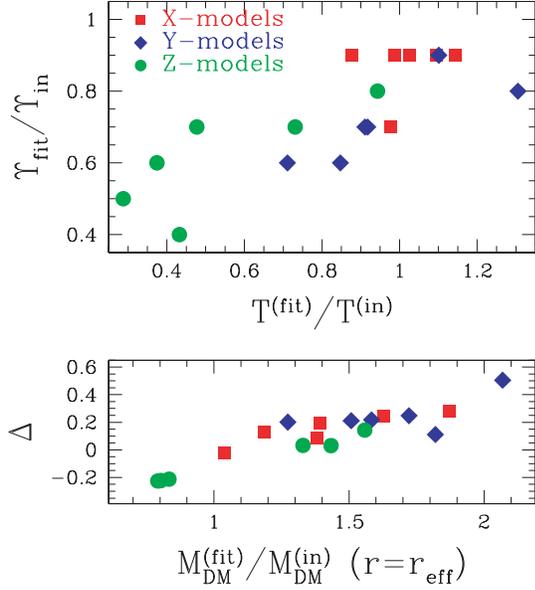
to the one-to-one case revealed by Fig. 19. In addition there is a correlation of the dark halo properties in the models with remnant projection, which could also contribute to the tilt in Fig. 19 (cf. Section 10.3). On top of that, the fact that we compare luminous kinetic energies with masses inside  $r_{\text{eff}}$  adds to the uncertainties in the step from  $T^{(\text{fit})}/T^{(\text{in})}$  to  $M^{(\text{fit})}/M^{(\text{in})}$  and can also tilt the relation or increase its scatter. Other sources of scatter are scatter in the relations (15) and (16). For example, the two  $Y$  models with  $M^{(\text{fit})}/M^{(\text{in})} > 1$  in Fig. 19 correspond to the two  $Y$  models above the one-to-one relation in Fig. 17.

## 10 CENTRAL REMNANT STRUCTURE AND THE LUMINOUS MASS-TO-LIGHT RATIO

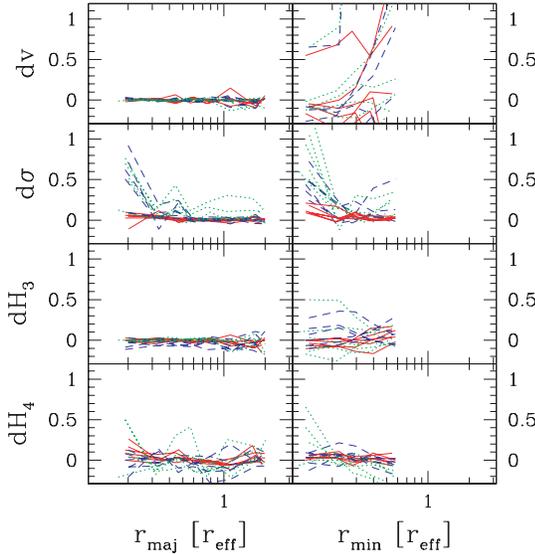
In this section we discuss the results for the modelled luminous  $M/L$  values. In contrast to the total masses the disagreement between models and mergers cannot be simply traced to the energy budget.

The upper panel of Fig. 20 shows  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  versus total kinetic energy  $T^{(\text{fit})}/T^{(\text{in})}$ . The scatter in Fig. 20 is much larger than in the corresponding Fig. 19 which deals with total masses. As expected from the tightness of the correlation in Fig. 19, deviations from a one-to-one correlation between  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  and  $T^{(\text{fit})}/T^{(\text{in})}$  are correlated with the dark matter content of the fits: where  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  is too low, the dark matter in the models overestimates the dark matter in the remnant and vice versa. A larger scatter in Fig. 20 than in Fig. 19 is not surprising because the reconstruction of the mass decomposition is less certain than the reconstruction of the total mass: there is some freedom in the modelling to shift mass from the luminous to the dark component (and vice versa) without changing the fit significantly. There are however two striking trends: (1)  $\Upsilon_{\text{fit}}$  is generally smaller than  $\Upsilon_{\text{in}}$  and (2) at a given value of  $T^{(\text{fit})}/T^{(\text{in})}$   $Y$  models suffer from a slightly stronger underestimation of  $\Upsilon$  than models of other projections.

The systematics in the lower panel of Fig. 20 indicate that  $\Upsilon_{\text{fit}}$ , unlike the total mass, is not merely set by the total kinetic energy but must instead depend on something else. To investigate this further,



**Figure 20.** Top: accuracy of reconstructed luminous mass-to-light ratio  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  versus ratio  $T^{(\text{fit})}/T^{(\text{in})}$  of total kinetic energies in Schwarzschild fits and merger remnants. Bottom: deviations  $\Delta \equiv T^{(\text{fit})}/T^{(\text{in})} - \Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  from a one-to-one correlation in the top panel versus accuracy of reconstructed dark mass (inside  $r_{\text{eff}}$ ).



**Figure 21.** As Fig. 6, but for Schwarzschild models with  $\Upsilon \equiv \Upsilon_{\text{in}}$ .

we have redetermined best-fitting Schwarzschild models (with  $2 \times 3500$  orbits; cf. Section 3.3) under the condition that the luminous  $M/L$  is fixed to the true value of the mergers,  $\Upsilon \equiv \Upsilon_{\text{in}}$ . The corresponding fits are illustrated in Fig. 21 in the same way as fits with optimized  $\Upsilon_{\text{fit}}$  are shown in Fig. 6. The most important result is that  $Y$  and  $Z$  models constrained to have the true  $M/L$  fail to fit the central kinematics: they predict too large central velocity dispersions.  $X$  models in Fig. 21 are only marginally different from those in Fig. 6, because most  $X$  models yield  $\Upsilon_{\text{fit}} \approx \Upsilon_{\text{in}}$  to within 10 per cent in any case.

## 10.1 The relation between flattening and central kinematics

That axisymmetric models of  $Y$  and  $Z$  projections with  $\Upsilon_{\text{fit}} = \Upsilon_{\text{in}}$  overpredict the velocity dispersion somewhere along the equator, can be qualitatively understood as follows. The flattening of an axisymmetric body always comes along with some excess of energy parallel to the  $(X, Y)$  plane. According to rotational symmetry, this excess energy is equally distributed in  $X$  and  $Y$ , respectively. Consequently, the projected kinetic energy in directions where the object looks most flattened is always relatively large. Only details of the radial distribution of the projected energy are not strongly constrained by the flattening alone. For example, if the flattening comes predominantly from near-circular orbits, then the central projected energy can be relatively low (circular orbits cross the central line of sight with zero line-of-sight velocity), while most energy resides in the outer regions. If radial equatorial orbits are responsible for the flattening, then the central projected energy (velocity dispersion) is relatively large, instead (e.g. Dehnen & Gerhard 1993).

The situation in triaxial bodies is different: again there is an excess of energy in the  $(X, Y)$  plane as soon as  $c < a$  and  $c < b$ . But this energy is no longer distributed equally between  $X$  and  $Y$ . If  $b < a$ , then there is more energy parallel to  $X$  than parallel to  $Y$ . It follows that the  $Y$  projection, in which the object is most flattened (one sees the shortest and longest axis in projection), has relatively low projected kinetic energy. In other words,  $Y$  projections of triaxial systems can be highly flattened in combination with a low specific projected energy (e.g. line-of-sight energy per mass). Now, since the flattening in axisymmetric systems is connected to a relatively large specific projected energy around the equatorial plane, a mismatch of projected dispersions somewhere around the equator is plausible (if the masses are equal).

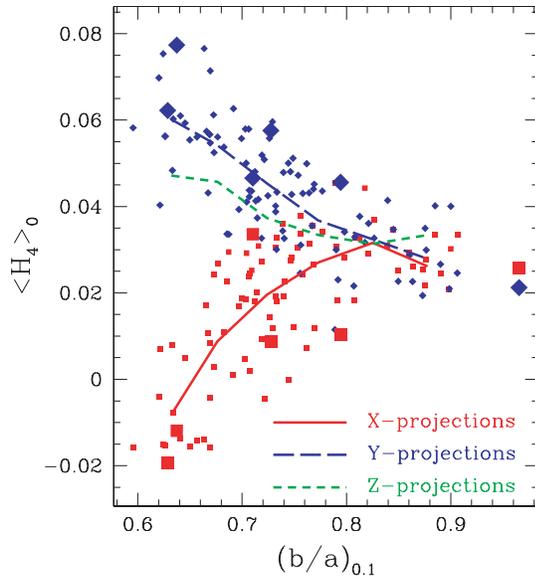
## 10.2 Flattening by box orbits and $\Upsilon_{\text{fit}}$

So far we have discussed the general situation. The fact that the specific dispersion of the axisymmetric models with  $\Upsilon_{\text{fit}} = \Upsilon_{\text{in}}$  is too large near the centre (cf. Fig. 21) is most likely connected to the specific structure of the merger remnants analysed here. They become prolate near the centre (cf. Section 2.2) and this inner prolateness is connected to particles moving preferentially on box orbits (Jesseit et al. 2005). Fig. 22 shows the interplay between intrinsic central shape and projected central kinematics. The former is quantified by the axis ratio  $b/a$  calculated from the spatial distribution of the 10 per cent most bound particles and the latter is expressed in terms of  $\langle H_4 \rangle_0$  (calculated inside an aperture of 2 arcsec – about  $r_{\text{eff}}/3$  or 1 kpc at the Coma distance of  $d = 100$  Mpc). The figure shows that high  $\langle H_4 \rangle_0$  occurs in  $Y$  projections (low line-of-sight velocities), while low or negative  $\langle H_4 \rangle_0$  appears when viewing the prolate centres end-on (large line-of-sight velocities). Differences between  $X$  and  $Y$  projections increase with decreasing  $b/a$ . Finally, as expected for nearly prolate systems,  $Z$  and  $Y$  projections are almost equivalent.

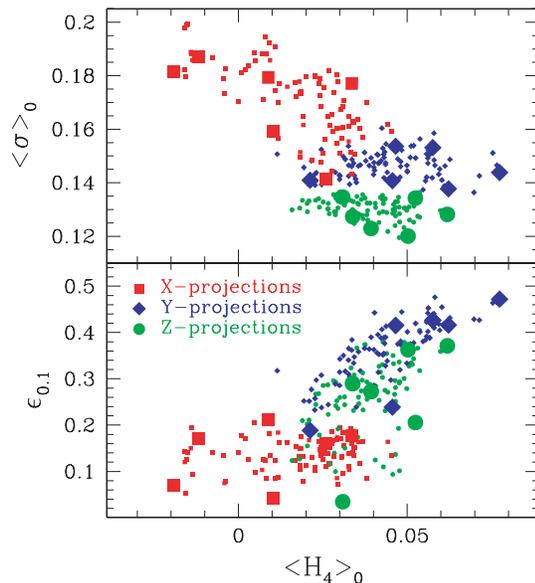
In principle, the projected central velocity dispersion should show analogue trends. However, it also depends on the total mass  $M$  and size  $R_h$  of a system and has to be normalized before different objects can be compared. One option is to use

$$\langle \sigma \rangle_0 \equiv \frac{\sigma_0}{\sqrt{GM/R_h}}, \quad (20)$$

where  $R_h$  is the half-mass radius of the light distribution and  $\sigma_0$  is the central velocity dispersion, measured in the same aperture as  $\langle H_4 \rangle_0$ . The top panel of Fig. 23 shows that  $\langle H_4 \rangle_0$  and  $\langle \sigma \rangle_0$  are closely correlated: high  $\langle H_4 \rangle_0$  comes along with low projected dispersions and vice versa, as expected from our above discussion. In



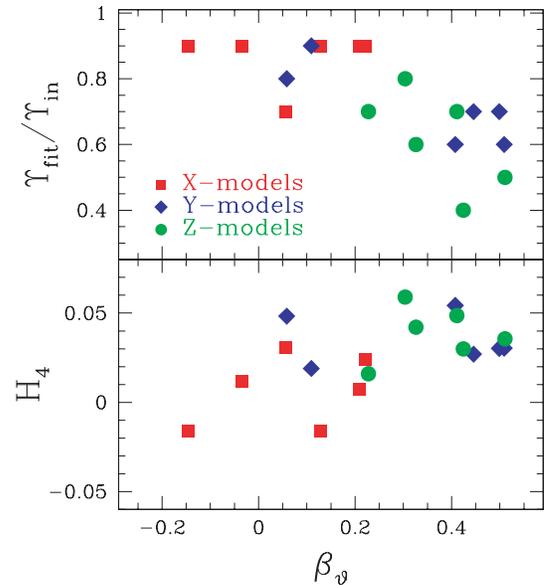
**Figure 22.** Central  $\langle H_4 \rangle_0$  in X and Y projections as function of  $b/a$  (10 per cent most bound particles). Small symbols: merger remnants of Naab & Burkert (2003); large symbols: subset of modelled merger remnants in this work; lines trace the mean of the distributions. For comparison also the mean of the Z projections is indicated.



**Figure 23.** Central  $\langle H_4 \rangle_0$  versus  $\langle \sigma \rangle_0$  (top) and projected ellipticity  $\epsilon_{0.1}$  of the 10 per cent most bound particles (bottom). Small symbols: all remnants; large symbols: remnants with Schwarzschild models.

addition, high  $\langle H_4 \rangle_0$  (and low  $\langle \sigma \rangle_0$ , respectively) are connected to large projected ellipticities, as illustrated in the lower panel of the figure. Concluding, inner box orbits in the merger remnants indeed cause a situation as described in Section 1.0.1: Y and Z projections of the remnant centres have high flattening in combination with low  $\langle \sigma \rangle_0$  and high  $\langle H_4 \rangle_0$ .

How are the remnant centres mapped on to the axisymmetric Schwarzschild models? Particles on box orbits stream perpendicular to the line of sight of Y and Z projections. The respective high  $H_4$  values provoke radial anisotropy in the inner regions of Y and Z



**Figure 24.** Meridional anisotropy  $\beta_\vartheta$  at  $r = 0.2 r_{\text{eff}}$  versus  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  (top) and versus  $H_4$  at same radius (bottom).

models (lower panel of Fig. 24). It has been mentioned above that axisymmetric systems which are flattened by equatorial radial orbits have large central velocity dispersions (e.g. Dehnen & Gerhard 1993, see also the edge-on projection of the axisymmetric radial orbit in Fig. 16). However, the Y and Z projections of the remnants are characterized by the opposite: low central velocity dispersions. Consequently, Y and Z models constrained to have  $\Upsilon \equiv \Upsilon_{\text{in}}$  predict a too large central velocity dispersion (cf. Fig. 21).

Assuming that the anisotropy in the models is fixed by the constraints imposed through  $H_4$ , then there is only one way to match the low central dispersion of the remnants: to reduce the inner mass. The inner mass, in turn, depends on only one parameter,  $\Upsilon$ , because both, mergers and models, are virtually free of dark matter in their centres. Accordingly, if it is indeed the flattening by box orbits in the merger remnants, e.g. the combination of high ellipticity, positive  $\langle H_4 \rangle_0$  but low  $\langle \sigma \rangle_0$  that causes the low  $\Upsilon_{\text{fit}}$  in Y and Z models, then we would expect the following behaviour of the models: the larger  $\langle H_4 \rangle_0$ , the larger the radial anisotropy in the Schwarzschild fits. The larger the anisotropy, in turn, the larger  $\langle \sigma \rangle_0$  for  $\Upsilon = \Upsilon_{\text{in}}$ . Thus, we expect that models which are forced to be more radially anisotropic will have a lower  $\Upsilon_{\text{fit}}$  for compensation. The upper panel of Fig. 24 confirms that indeed the lowest  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  appears in models that are most strongly radially anisotropic near the centre (at  $0.2 r_{\text{eff}}$ ).

Hence, the systematically low  $\Upsilon_{\text{fit}}$  of Y and Z models in Fig. 20 most likely reflect a lack of appropriate counterparts of box orbits in axisymmetric potentials that can support a high flattening in combination with a low central velocity dispersion (per mass).

The underestimation of  $\Upsilon$  in Y and Z models is at first glance similar to the underestimation of the total mass in these models (cf. Section 9.2). However, models of X projections also deliver  $\Upsilon_{\text{fit}} < \Upsilon_{\text{in}}$ . This is different from the reconstruction of total masses, which are instead overestimated in X models. On one hand, it should be noted that – irrespective of  $\Upsilon_{\text{fit}} < \Upsilon_{\text{in}}$  – the total mass of X models, even close to the centre, is often larger than in the remnants because of the overestimation of light in the deprojection (cf. Section 5.1). On the other hand, the discussion of the Y and Z models has revealed that  $\Upsilon_{\text{fit}}$  depends primarily on the central orbital structure and not,

as the total mass, on the global energy budget. Therefore the reconstruction of  $\Upsilon$  is different from the reconstruction of the total mass. In particular, it most likely depends on the specific structure of the merger remnants considered here. We cannot rule out that systems exist in which, say,  $\langle H_4 \rangle_0$  and the inner flattening combine in a way such that  $X$  models would be forced to have  $\Upsilon_{\text{fit}} > \Upsilon_{\text{in}}$ . This needs further exploration of a broader sample of modelling targets. The  $X$  projections of the merger remnants analysed here apparently do not require  $\Upsilon_{\text{fit}} > \Upsilon_{\text{in}}$  to be modelled adequately.

### 10.3 The role played by dark matter

In Fig. 24  $Y$  and  $Z$  models behave similarly (low  $\Upsilon_{\text{fit}}$ , large radial anisotropy) as expected from the central box orbits in the merger remnants. In Fig. 20, however,  $Y$  models are different from  $Z$  models in their dark haloes: the underestimation of luminous mass in  $Y$  models is compensated for by relatively massive haloes, while this is apparently not the case in  $Z$  models. According to Section 9 higher projected kinetic energies in  $Y$  projections are responsible for their higher total masses (compared to  $Z$  models). The difference between  $Y$  and  $Z$  models could be related to the different views on minor-axis tubes, that dominate the outer parts of the remnants (Jesseit et al. 2005).  $Y$  projections map them edge-on such that they contribute significantly to the total line-of-sight energy. In  $Z$  projection they appear face-on and most of their kinetic energy is perpendicular to the line of sight. The transition from side-on inner box orbits to edge-on outer  $Z$  tubes may be the origin for the local maximum in some dispersion profiles of  $Y$  projections (TRIAX, FLAT; cf. Appendix B).  $Z$  projections lack a similar maximum, as expected if it is caused by the edge-on view on  $Z$  tubes. The increase of  $\sigma$  at the transition from box orbits to  $Z$  tubes could explain why the low  $\Upsilon_{\text{fit}}$  of  $Y$  models (set by the fit to the central remnant kinematics) have to be compensated for with massive haloes at larger radii.  $Z$  models do not need such compensating halo components, because  $\sigma$  drops smoothly.

It appears that for our modelling the inclusion of dark matter in the fits has two main effects: (1) it allows to trace the true mass structure of the modelling targets in the outer parts (because the remnants contain dark matter); (2) in some cases it can improve the fit to the *central* kinematics by allowing for an artificial ( $M/L$ ) gradient over the spatial region dominated by luminous mass. In particular it offers to combine low central  $\Upsilon$  with larger outer  $M/L$ . This second issue related to dark matter in the models raises the question whether fitting self-consistent triaxial systems with self-consistent axisymmetric models might be different: the mass-to-light ratio  $\Upsilon$  of the fits is then constrained by the inner as well as the outer kinematics. Accordingly, the quality of the overall fit could be less good. For example, in the self-consistent case  $\Upsilon$  cannot be reduced arbitrarily to match low central  $\langle \sigma \rangle_0$  such as those arising in specific  $Y$  projections, as otherwise the mass in the outer parts would be insufficient to fit the corresponding kinematics there. Consequently, the scatter in the relation shown in Fig. 17 could increase. This, in turn, would affect the conclusions drawn about the projection dependency of the mass budget of the axisymmetric fits, because they are partly based on this relation.

In summary, the luminous  $M/L$ , because it is the only parameter that controls the central mass, depends not primarily on the total projected energy (like the total mass). Instead it is more sensitive to the central orbital structure of the merger remnants. In particular the low  $\Upsilon_{\text{fit}}$  of  $Y$  and  $Z$  models results from the lack of orbits that resemble side-on views of box orbits, i.e. which appear highly flattened, show large transversal motions and low central line-of-sight

dispersions simultaneously. The radial anisotropy induced in the axisymmetric models by the transversal streaming of box orbits in the remnants requires a lowering of  $\Upsilon_{\text{fit}}$  to keep the central dispersion low. In  $Y$  models massive haloes partly compensate for the low outer luminous masses that result from the low  $\Upsilon_{\text{fit}}$  required for the central fit. These massive haloes are needed to fit side-on views of outer  $Z$  tubes and the related relatively large velocity dispersions.  $Z$  models do not require massive haloes, because when viewed face-on the outer  $Z$  tubes produce relatively low dispersions that do not need particularly large masses to be fitted.

## 11 IMPLICATIONS FOR MODELS OF REAL GALAXIES

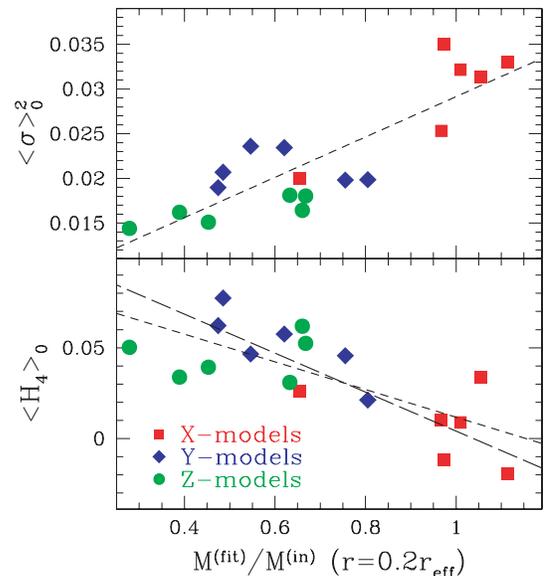
In Sections 9 and 10 we have followed the projection trends in the recovery of total and luminous masses back to the restrictions imposed by axial symmetry. We now discuss very briefly some implications for models of real galaxies.

### 11.1 A possible bias in the reconstruction of central masses

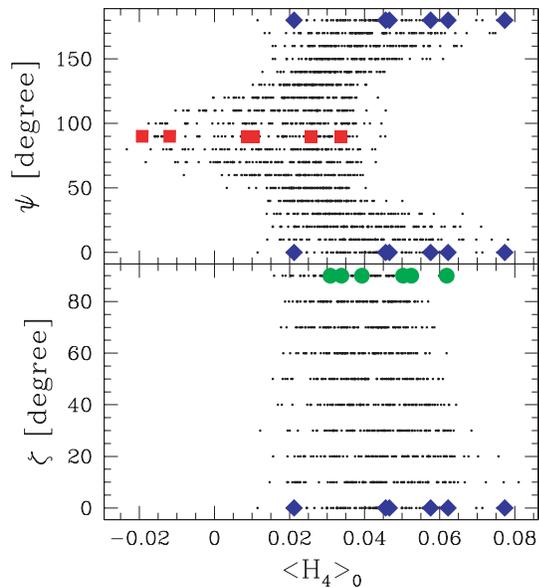
The dependency of  $\Upsilon_{\text{fit}}$  on the central kinematics of the merger remnants discussed in Section 10 implies that  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  is connected to  $\langle H_4 \rangle_0$  and  $\langle \sigma \rangle_0$ . Insofar as models and remnants are dominated by luminous mass near their centres, one would expect that  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  resembles the ratio  $M^{(\text{fit})}/M^{(\text{in})}$  between reconstructed and input mass, evaluated at some radius near the centre. Consequently,  $M^{(\text{fit})}/M^{(\text{in})}$  should be connected to  $\langle H_4 \rangle_0$  and  $\langle \sigma \rangle_0$  as well.

The relationships between  $M^{(\text{fit})}/M^{(\text{in})}$  (evaluated at  $0.2 r_{\text{eff}}$ ) and the central kinematical parameters  $\langle H_4 \rangle_0$  and  $\langle \sigma \rangle_0$  are shown in Fig. 25. Because masses globally scale with velocities squared, we have chosen  $\langle \sigma \rangle_0^2$  as the ordinate in the top panel. Indeed the inner  $M^{(\text{fit})}/M^{(\text{in})}$  is closely correlated to  $\langle H_4 \rangle_0$  and  $\langle \sigma \rangle_0^2$ . Linear fits yield

$$\frac{M^{(\text{fit})}}{M^{(\text{in})}} \Big|_{0.2} = 44.33 \langle \sigma \rangle_0^2 - 0.29 \quad (21)$$



**Figure 25.** Central velocity dispersion  $\langle \sigma \rangle_0^2$  (top) and  $\langle H_4 \rangle_0$  (bottom) versus  $M^{(\text{fit})}/M^{(\text{in})}$  (evaluated at  $0.2 r_{\text{eff}}$ ). Short-dashed lines: linear fits. The long-dashed line in the bottom panel shows the relation expected according to the fit from the top panel and the correlation between  $\langle \sigma \rangle_0$  and  $\langle H_4 \rangle_0$  shown in the top panel of Fig. 23. Symbols as labelled in the bottom panel.



**Figure 26.** Central  $\langle H_4 \rangle_0$  versus viewing angle. Top: azimuth  $\Psi$  in  $(X, Y)$  plane ( $\zeta \equiv 0$ ); bottom: latitude  $\zeta$  in  $(Y, Z)$  plane ( $\Psi \equiv 0$ ); large symbols (as in Fig. 23): modelled merger remnants.

and

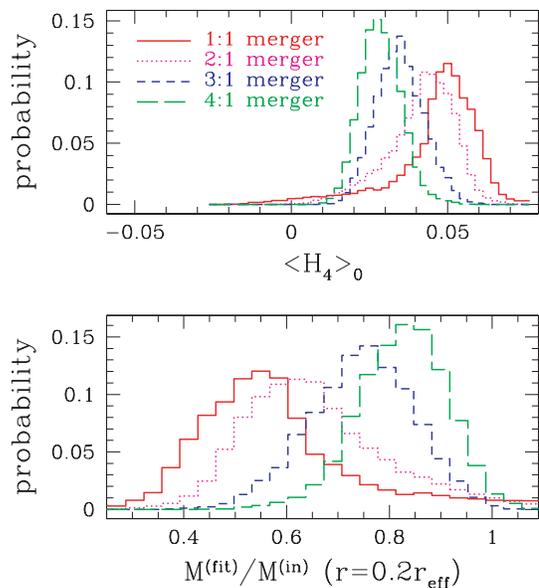
$$\left. \frac{M^{(\text{fit})}}{M^{(\text{in})}} \right|_{0.2} = 1.2 - 13.1 \langle H_4 \rangle_0, \quad (22)$$

respectively (short-dashed lines in Fig. 25). The two relations should be almost equivalent, because according to Fig. 23  $\langle H_4 \rangle_0$  and  $\langle \sigma \rangle_0$  are closely correlated in the merger remnants. In fact, the predicted relation between  $M^{(\text{fit})}/M^{(\text{in})}$  and  $\langle H_4 \rangle_0$  that derives from a linear fit to the top panel of Fig. 23 in combination with equation (21) is consistent with the actual relation between the mass ratio and  $\langle H_4 \rangle_0$  that is shown in the lower panel of Fig. 25. For comparison, this relation is indicated by the long-dashed line.

Both quantities,  $\langle \sigma \rangle_0$  and  $\langle H_4 \rangle_0$  are viewing-angle dependent. The case of the principal axes has already been discussed in Section 1.0.2. The full viewing-angle dependency of  $\langle H_4 \rangle_0$  for all merger remnants of Naab & Burkert (2003) is shown in Fig. 26. Thereby  $\Psi$  is defined as the azimuth in the  $(X, Y)$  plane and  $\zeta$  is the latitude. In accordance with our previous discussion  $\langle H_4 \rangle_0$  peaks in  $Y$  and  $Z$  projections and varies little with  $\zeta$  between these projections (prolateness). In the  $(X, Y)$  plane  $\langle H_4 \rangle_0$  decreases smoothly when approaching the long-axis projection. A similar behaviour of  $H_4$  with viewing angle has been observed in  $N$ -body binary mergers of discs with massive bulges, but without dark matter (Heyl, Hernquist & Spiegel 1995).

If the merger remnants studied here would be seen at random projections on the sky, then the viewing-angle dependency of  $\langle H_4 \rangle_0$  would give rise to the frequency distribution shown in the top panel of Fig. 27. Equal-mass mergers would have on average the highest  $\langle H_4 \rangle_0$ , because they are nearly prolate and the  $\langle H_4 \rangle_0$  distribution is dominated by the positive values around  $Y$  and  $Z$  projections, respectively. Towards 4:1 mergers the average  $\langle H_4 \rangle_0$  decreases slightly.

Assuming that relation (22) holds for all viewing angles and merger remnants, then the frequency distribution of  $\langle H_4 \rangle_0$  can be used to predict the distribution of reconstructed central  $M^{(\text{fit})}/M^{(\text{in})}$ . The latter is plotted in the lower panel of Fig. 27: axisymmetric models of these merger remnants would be always biased towards too low central masses. The bias would be strongest for equal-mass



**Figure 27.** Top: central  $\langle H_4 \rangle_0$  from random projections of the merger remnant sample of Naab & Burkert (2003) (results are separated for different progenitor mass ratios); bottom: resulting central  $M^{(\text{fit})}/M^{(\text{in})}$  according to relation (22).

mergers, while masses of the more axisymmetric 4:1 mergers would be recovered better.

Note, however, that for the construction of the lower panel of Fig. 27 we have assumed that the relation (22) holds for models of all merger remnants and at all viewing angles. This needs to be verified on the basis of a broader sample of models. Likewise it is not clear whether the connection between  $\langle H_4 \rangle_0$  and the central  $M^{(\text{fit})}/M^{(\text{in})}$  also holds for non-axisymmetric targets of more general shapes.

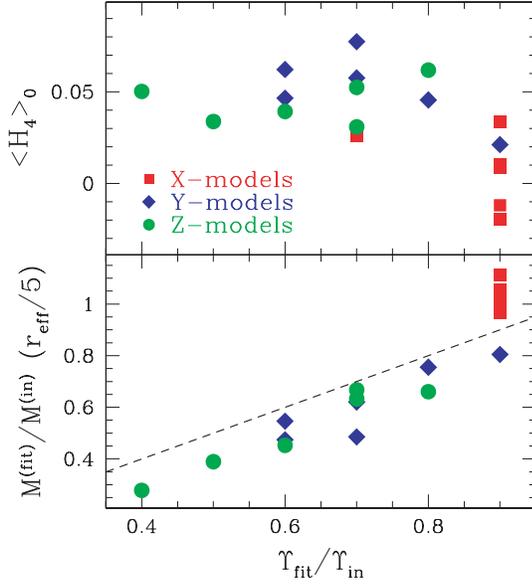
## 11.2 A possible bias in the reconstruction of luminous mass-to-light ratios

As stated above, if models and merger remnants are void of dark matter in their central regions, then  $M^{(\text{fit})}/M^{(\text{in})} \approx \Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$ . Accordingly, if the merger remnants analysed here would be seen at random projections on the sky, then  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  would be subject to a similar bias as the central  $M^{(\text{fit})}/M^{(\text{in})}$ . However, the relation between  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  and  $\langle H_4 \rangle_0$ , which determines the bias, is more scattering than the one with the inner  $M^{(\text{fit})}/M^{(\text{in})}$  and it cannot be described by a straight line (upper panel of Fig. 28). The reason is that the total mass near the centre, say inside  $0.2 r_{\text{eff}}$ , is determined by the product of the central luminosity with  $\Upsilon_{\text{fit}}$  (and a negligible amount of central dark matter). The central luminosity however, is overestimated in  $X$  models, but underestimated in  $Y$  and  $Z$  models (cf. Section 5.1). Consequently,  $X$  models with a specific  $\Upsilon$  have larger  $M^{(\text{fit})}/M^{(\text{in})}$  than  $Y$  or  $Z$  models with the same  $\Upsilon$ . This is illustrated in the lower panel of Fig. 28.

Regardless of the complicated shape of the relation between  $\langle H_4 \rangle_0$  and  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  it is clear that reconstructed luminous  $M/L$  values of the merger remnants analysed here would be always underestimated.

## 11.3 The mass recovery in models with $\Upsilon \equiv \Upsilon_{\text{in}}$

If one is interested in the recovery of luminous  $M/L$  values, one could use the relationship between  $\langle H_4 \rangle_0$  and  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  (upper panel of Fig. 28) to correct dynamically derived  $M/L$  values (roughly) for



**Figure 28.** Top:  $\langle H_4 \rangle_0$  versus  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$ ; bottom: connection between the parameter ratio  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  and the actual ratio  $M^{(\text{fit})}/M^{(\text{in})}$  of integrated masses (luminous + dark) inside  $0.2 r_{\text{eff}}$ . The dashed line marks the identity relation.

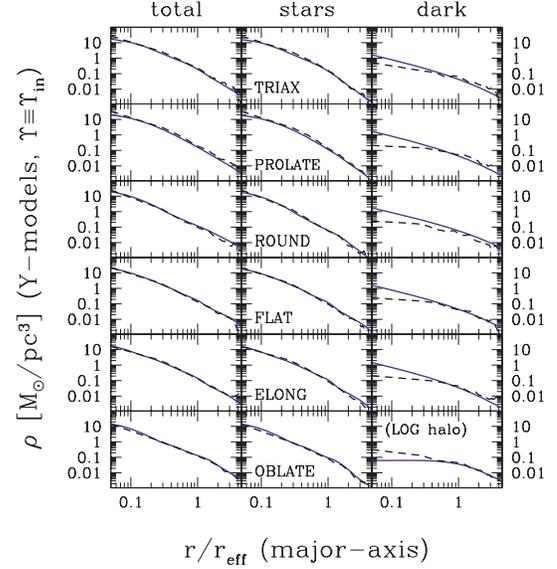
the effects of intrinsic non-axisymmetry. However, the accuracy of this step is limited by the systematics of the deprojections (lower panel of Fig. 28). Another way to obtain unbiased luminous  $M/L$  values would be to apply stellar population models to line indices, although this requires knowledge of the initial stellar mass function. In any case it is interesting to ask, whether knowledge of the true  $\Upsilon$  could help to improve other aspects of the dynamical modelling, e.g. the recovery of the intrinsic anisotropy. To investigate this, we now compare Schwarzschild fits obtained under the condition that the luminous  $M/L$  is fixed to its true value,  $\Upsilon \equiv \Upsilon_{\text{in}}$ , with the merger remnants. Since most best-fitting  $X$  models with variable  $\Upsilon$  already have  $\Upsilon_{\text{fit}} \approx \Upsilon_{\text{in}}$  to within 10 per cent (cf. Section 6.1), fixing  $\Upsilon \equiv \Upsilon_{\text{in}}$  does not change  $X$  models significantly. We therefore skip  $X$  models in the remainder of this section.

Fig. 29 compares the intrinsic mass densities of  $Y$  models with  $\Upsilon \equiv \Upsilon_{\text{in}}$  to the merger remnants in a similar fashion as models with variable  $\Upsilon$  were compared in Fig. 9. The figure clearly shows an improvement in the mass recovery when  $\Upsilon$  is known. This holds for both, total as well as dark mass densities. Fig. 30 covers the case of  $Z$  models. Here, although again the mass recovery improves, the discrepancy between the density profiles of models and remnants in the outer parts remains. As it has already been discussed in Section 5.1, the light profile of  $Z$  models differs from the remnants mainly in having the wrong *slope*. Knowing just the true scaling  $\Upsilon_{\text{in}}$  cannot remove this mismatch.

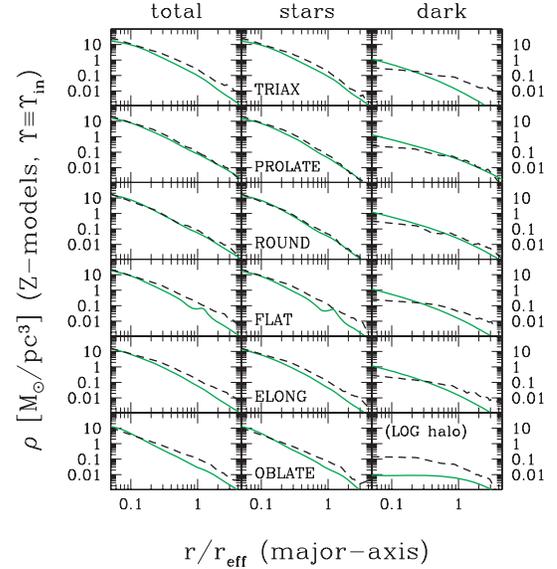
#### 11.4 Anisotropy in models with $\Upsilon \equiv \Upsilon_{\text{in}}$

Figs 31 and 32 compare anisotropy profiles of  $Y$  and  $Z$  models with  $\Upsilon \equiv \Upsilon_{\text{in}}$  to the merger remnants. The general trend is that models with  $\Upsilon \equiv \Upsilon_{\text{in}}$  become strongly tangentially anisotropic ( $\beta_\theta < 0$ ,  $\beta_\varphi < 0$ ) in the outer regions ( $r \gtrsim 0.5 r_{\text{eff}}$ ), especially  $Z$  models. Towards the centres most of the models shown in Figs 31 and 32 become radially anisotropic ( $\beta_\varphi > 0$ ), with a local peak around  $0.1\text{--}0.3 r_{\text{eff}}$ .

All in all then, fixing  $\Upsilon \equiv \Upsilon_{\text{in}}$  improves the reconstruction of the intrinsic mass structure, but deviations in internal velocity moments increase.



**Figure 29.** As Fig. 9, but for models with  $\Upsilon \equiv \Upsilon_{\text{in}}$ .

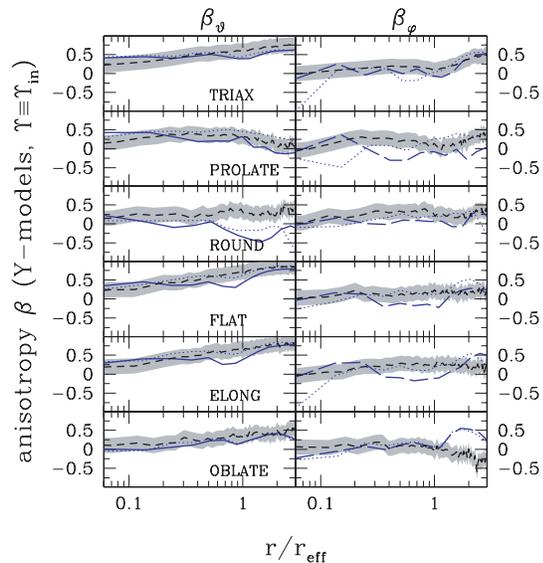


**Figure 30.** As Fig. 10, but for models with  $\Upsilon \equiv \Upsilon_{\text{in}}$ .

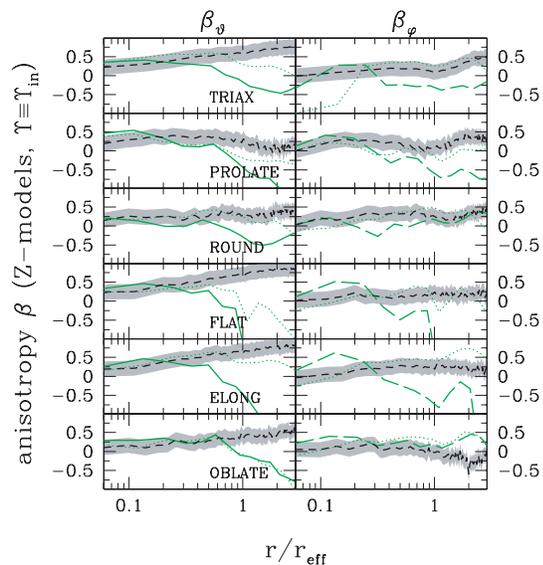
## 12 SUMMARY

We have modelled a set of collisionless disc–disc mergers with exactly the same axisymmetric orbit superposition program that has been used to model a sample of medium bright giant Coma ellipticals. The models assume a constant  $M/L$  for the luminous matter and a dark halo of the NFW type (Navarro et al. 1996). The remnants we model result from the collisionless merger of progenitor systems composed of a bulge, a disc (both with the same  $M/L$ ) and a dark halo. They are chosen to cover representatively the range of intrinsic shapes and dynamical structures of the Naab & Burkert (2003) merger sample, including the most extreme cases.

Intrinsic triaxiality causes a strong viewing-angle dependency of projected properties of the remnants. As in axisymmetric models many viewing angles are equivalent this must result in a



**Figure 31.** As Fig. 12, but for models with  $\Upsilon \equiv \Upsilon_{\text{in}}$ . Profiles of the best-fitting models shown in Fig. 12 are repeated for comparison (dotted).



**Figure 32.** As Fig. 31, but for Z models.

corresponding dependency of the fits on the viewing angle of the remnant. The goal of this study is to investigate this dependency.

Some projected properties of triaxial systems, e.g. minor-axis rotation or isophotal twists, are obviously incompatible with oblate axial symmetry. Apart from these, we do not find any obvious mismatch between our Schwarzschild models and the merger remnants: residuals in the kinematic fits are smaller than typical observational errors.

### 12.1 Remnant shapes and the mass recovery

We find that the reconstruction of the total (luminous + dark) mass at the effective radius depends primarily on viewing angle and not on the orbital structure of the merger remnants. This is so, because the global mass budget of the axisymmetric models is fixed by two constraints. First, the total line-of-sight energies of the luminous

components of models and mergers match, which is a consequence of the fit to the projected kinematics. Secondly, although the two transversal kinetic energies are not constrained in the models, the ratio of both has to be the same in models and mergers, respectively: the ratio determines the flattening and is constrained by the fit to the luminosity density. For edge-on models the restrictions imposed through axial symmetry then already fix the total mass budget of the Schwarzschild models. Thereby models of  $X$  projections overestimate the true mass, models of  $Y$  projections underestimate it on about the same level and models of  $Z$  projections have masses even lower than those of  $Y$  models. The exact amount of overestimation or underestimation depends on the intrinsic shape of the merger remnants.

In the merger remnants analysed here deviations of cumulative masses inside  $1-2 r_{\text{eff}}$  are mostly below 20 per cent. Extreme values of underestimations are larger than those of overestimations. The strongest underestimations occur among intrinsically very flattened, face-on systems, where the mass can be underestimated by up to 50 per cent. The underestimation is due to a wrong inclination of the corresponding models, which arises from the fact that most remnants appear either flattened or show some residual rotation along the minor axis, respectively, when viewed along their short axis. Both phenomena exclude face-on ( $i = 0^\circ$ ) oblate axisymmetric models, cause an underestimation of the luminous kinetic energy and, hence, an underestimation of the mass inside the effective radius.

The luminous  $M/L$  is always underestimated,  $\Upsilon_{\text{fit}} < \Upsilon_{\text{in}}$ . Unlike the total mass, it does not derive primarily from the total kinetic energy of the Schwarzschild model. Although  $\Upsilon_{\text{fit}}$  varies similarly with viewing angle as the total mass, this variation is mediated by the central kinematics of the merger remnants. Box orbits cause a combination of high projected flattening and low line-of-sight motions (high central  $H_4$ , low central  $\sigma$ ) in  $Y$  and  $Z$  projections. Box orbits are mapped on to radial orbits in the Schwarzschild fits. The corresponding increase of the central velocity dispersion in the model then requires a lowering of the central mass to achieve a good match to both, the high central  $H_4$  and the low  $\sigma$ . As the luminous mass-to-light ratio  $\Upsilon_{\text{fit}}$  is the only parameter that controls the central mass,  $\Upsilon_{\text{fit}} < \Upsilon_{\text{in}}$  in  $Y$  and  $Z$  models (up to a factor of 2.5). Models of long-axis projections yield the best approximations to the mergers, in all but one case  $\Upsilon_{\text{fit}} \approx \Upsilon_{\text{in}}$  to within 10 per cent.

The asymmetric motion of particles, especially those on inner box orbits, correlates central kinematical parameters like  $\langle \sigma \rangle_0$  and  $\langle H_4 \rangle_0$  with viewing angle, a result that has already been found in  $N$ -body simulations without dark matter (Heyl et al. 1995). The link between box orbit kinematics and  $\Upsilon_{\text{fit}}$  also gives rise to a correlation of the reconstructed central mass (relative to the merger mass)  $M^{(\text{fit})}/M^{(\text{in})}$  with  $\langle H_4 \rangle_0$  – and viewing angle.

The deficit of luminous mass in models of projections along the intermediate axis is compensated for by dark haloes that are more massive than in the remnants. Such massive haloes are necessary to fit the relatively large outer dispersions of these projections. Models of short-axis projections do not have massive dark haloes: their outer dispersions are low. The different outer dispersions of projections along the intermediate and short axis, respectively, arise due to different views on minor-axis tubes dominating the outskirts of the remnants.

If the luminous  $M/L$  is fixed to its true value, then the mass recovery of  $Y$  models improves (on average 6 per cent accuracy at  $2 r_{\text{eff}}$ ). Models of  $Z$  projections suffer from a wrong slope in the deprojection arising from the mismatch in the inclination already discussed above. As the inclination is wrong, mere knowledge of the true  $\Upsilon$  does not improve the mass reconstruction of  $Z$  models much.

In general, improvements of the mass recovery are to the expense of strong outer tangential anisotropy in the Schwarzschild models, which weaken the match with the remnants' dynamics.

## 12.2 Anisotropies in Schwarzschild models of mergers

The viewing-angle dependency of merger projections induces a viewing-angle dependency of anisotropies in corresponding Schwarzschild models. For example,  $X$  tubes, when seen end-on, are nearly round and are represented by shell orbits in axisymmetric models, resulting in strong tangential anisotropy in the Schwarzschild fit. The same  $X$  tubes seen side-on appear radially extended and are represented by radially extended orbits in the models, increasing their radial anisotropy. Because different orbit families give rise to different viewing-angle dependencies (e.g. Jesseit et al. 2005) the anisotropy of Schwarzschild fits depends on projection as well as on the orbital makeup.

$Z$  models of intrinsically flat, disc-like merger remnants are dominated by meridional motions ( $\sigma_\theta > \sigma_r$ ). This can be explained as an inclination effect: as stated above, the flattening of most remnants requires an inclination  $i > 0^\circ$  in the axisymmetric models. In this case,  $Z$  tubes dominating the outskirts of the remnant project to nearly round shapes and are mapped on to axisymmetric shell orbits of the same shape – causing a meridional anisotropy.

As found for the mass reconstruction, modelling the long-axis projection of a merger remnant yields a better match to the intrinsic structure, while the largest deviations between remnants and models appear among the short-axis models. In any case, deviations  $\Delta\beta$  between Schwarzschild fits and merger remnants are below  $\Delta\beta < 0.2$  at most radii. Towards the centre and/or towards the outer regions deviations can be larger, however.

## 12.3 Real galaxies

We have tested our axisymmetric orbit superposition code on a sample of rather extreme merger remnants, covering a wide range of non-axisymmetric as well as axisymmetric, but highly flattened, dynamical systems. The aim was to probe the limits of the method. If real ellipticals would resemble the merger remnants considered here in terms of their orbital structure, then random viewing angles would provoke scatter in anisotropies of axisymmetric dynamical models. Furthermore, dynamically derived stellar  $M/L$  values would be on average underestimated and the amount of underestimation would be correlated with the central value of  $H_4$ . We plan a detailed comparison of the models of merger remnants described here with models of Coma ellipticals in a future publication. If models of real ellipticals show less scatter in anisotropies or if there is no sign for a systematic underestimation of stellar  $M/L$  values, then this would be an indication for their intrinsic shapes to be closer to axial symmetry than in our  $N$ -body merger sample.

This would have two consequences. First, axisymmetric galaxies would be recovered with higher accuracy, even better than the merger remnants analysed here whose masses and anisotropies at the effective radius are mostly well matched. This holds especially for flattened and rotating systems, which are known to be close to edge-on. The case of round and non-rotating galaxies (even if they are axisymmetric) is more ambiguous as the inclination of these systems is a priori unknown.

Secondly, knowing that galaxies are close to axisymmetry provides clues about their formation: dissipation during the formation can change elliptical galaxy properties significantly (e.g. Cox et al. 2006; Robertson et al. 2006). Particularly, it can drive the final object towards axial symmetry (e.g. Barnes & Hernquist 1996; Naab,

Jesseit & Burkert 2006b). We therefore also plan to extend our study to binary disc mergers including gas physics, star formation and feedback from a central black hole as well as on simulations from cosmological initial conditions.

## ACKNOWLEDGMENTS

We thank Scott Tremaine, Ortwin Gerhard and the anonymous referee for comments and suggestions that helped to improve the manuscript. This work was supported by DFG Sonderforschungsbereich 375 'Astro-Teilchenphysik' and DFG priority programme 1177.

## REFERENCES

- Barnes J. E., 1992, *ApJ*, 393, 484  
 Barnes J. E., Hernquist L., 1996, *ApJ*, 471, 115  
 Bender R., Möllenhoff C., 1987, *A&A*, 177, 71  
 Bertola F., Galletta G., 1979, *A&A*, 77, 363  
 Binney J., Tremaine S., 1987, *Galactic Dynamics*. Princeton Univ. Press, Princeton, NJ  
 Cappellari M. et al., 2006, *MNRAS*, 366, 1126  
 Cox T. J., Jonsson P., Primack J. R., Somerville R. S., 2006, *MNRAS*, 373, 1013  
 Cretton N., de Zeeuw P. T., van der Marel R. P., Rix H. W., 1999, *ApJS*, 124, 383  
 Dehnen W., Gerhard O. E., 1993, *MNRAS*, 261, 311  
 de Zeeuw P. T., 1985, *MNRAS*, 216, 273  
 Dubinski J., 1998, *ApJ*, 502, 141  
 Eggen D. J., Lynden-Bell D., Sandage A. R., 1962, *ApJ*, 136, 748  
 Franx M., Illingworth G., Heckman T., 1989, *ApJ*, 344, 613  
 Gebhardt K. et al., 2000, *AJ*, 119, 1157  
 Gebhardt K. et al., 2003, *ApJ*, 583, 92  
 Gerhard O. E., 1993, *MNRAS*, 265, 213  
 Gerhard O. E., Binney J., 1996, *MNRAS*, 279, 993  
 Häfner R., Evans N. W., Dehnen W., Binney J., 2000, *MNRAS*, 314, 433  
 Hernquist L., 1990, *ApJ*, 356, 359  
 Hernquist L., 1992, *ApJ*, 409, 548  
 Hernquist L., 1993, *ApJ*, 400, 460  
 Heyl J. S., Hernquist L., Spergel D. N., 1995, *ApJ*, 448, 64  
 Jedrzejewski R., Schechter P. S., 1989, *AJ*, 98, 147  
 Jesseit R., Naab T., Burkert A., 2005, *MNRAS*, 360, 1185  
 Jesseit R., Naab T., Peletier R., Burkert A., 2007, *MNRAS*, 376, 997  
 Khochfar S., Burkert A., 2006, *A&A*, 445, 403  
 Krajnović D., Cappellari M., Emsellem E., McDermid R. M., de Zeeuw P. T., 2005, *MNRAS*, 357, 1113  
 Larson R. B., 1974, *MNRAS*, 169, 229  
 Magorrian J., 1999, *MNRAS*, 302, 530  
 Mehlert D., Saglia R. P., Bender R., Wegner G., 2000, *A&AS*, 141, 449  
 Merritt D., Fridman T., 1996, *ApJ*, 460, 138  
 Naab T., Burkert A., 2001, *ApJ*, 555, L91  
 Naab T., Burkert A., 2003, *ApJ*, 597, 893  
 Naab T., Trujillo I., 2006, *MNRAS*, 369, 625  
 Naab T., Khochfar S., Burkert A., 2006a, *ApJ*, 636, L81  
 Naab T., Jesseit R., Burkert A., 2006b, *MNRAS*, 372, 839  
 Naab T., Johansson P. H., Ostriker J. P., Efstathiou G., 2007, *ApJ*, 658, 710  
 Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563  
 Richstone D. O., 1982, *ApJ*, 252, 496  
 Richstone D. O., Tremaine S., 1988, *ApJ*, 327, 82R  
 Robertson B., Cox T. J., Hernquist L., Franx M., Hopkins P. F., Martini P., Springel V., 2006, *ApJ*, 641, 21  
 Rybicki G., 1987, in de Zeeuw T., ed., *Proc. IAU Symp. 127, Structure and Dynamics of Elliptical Galaxies*. Reidel, Dordrecht, p. 397  
 Schwarzschild M., 1979, *ApJ*, 232, 236  
 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, in press  
 Toomre A., Toomre J., 1972, ApJ, 178, 623  
 Tremblay B., Merritt D., 1996, AJ, 111, 2243  
 Valluri M., Merritt D., Emsellem E., 2004, ApJ, 602, 66  
 van Albada T. S., 1982, MNRAS, 201, 939  
 van der Marel R. P., Franx M., 1993, ApJ, 407, 525  
 Weil M., Hernquist L., 1996, ApJ, 457, 51  
 Wegner G., Corsini E. M., Saglia R. P., Bender R., Merkl D., Thomas D., Thomas J., Mehlert D., 2002, A&A, 395, 753

## APPENDIX A: $N$ -BODY REALIZATIONS OF SCHWARZSCHILD MODELS

In the following we describe the construction of  $N$ -body realizations of Schwarzschild models. Thereby we assume that all particles have the same mass, e.g. that there is a global constant of proportionality linking the number density of particles in phase space to the phase-space mass density. To simplify the notation, we also assume  $\Upsilon \equiv 1$ . The generalization to particle masses varying from orbit to orbit or to systems with  $\Upsilon \neq 1$  is straightforward.

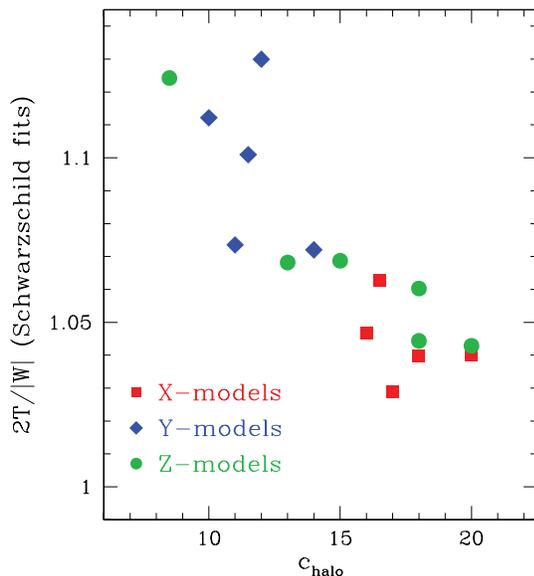
The orbit library of any Schwarzschild model provides a natural partition of phase space into cells  $dV(i, j)$ , where  $dV(i, j)$  is the phase-space volume of the cell that orbit  $i$  covers during time-step  $dt_i^{(j)}$ . The probability  $p(i, j)$  to find a particle in cell  $dV(i, j)$  is proportional to

$$p(i, j) \propto f_i dV(i, j), \quad (\text{A1})$$

where  $f_i$  is the phase-space density along orbit  $i$ . The latter is related to the weight  $w_i$  of the corresponding orbit and its total phase-space volume  $V_i$  via  $f_i = w_i/V_i$  (e.g. Thomas et al. 2004). According to the time-average theorem

$$\frac{dV(i, j)}{V_i} \propto \frac{dt_i^{(j)}}{\tau_i}, \quad (\text{A2})$$

where  $\tau_i$  is the total integration time of orbit  $i$ . Hence, the probability to find a particle in phase-space cell  $dV(i, j)$  is



**Figure A1.** Correlation between the virial ratio  $2T/|W|$  and the halo concentration in the Schwarzschild models.

$$p(i, j) = w_i \frac{dt_i^{(j)}}{\tau_i}. \quad (\text{A3})$$

If the luminosity is normalized such that  $\sum w_i \equiv 1$  then the  $p(i, j)$  of equation (A3) provide a complete partition of the interval  $[0, 1]$ :

$$[0, 1] = \bigcup_{i=1}^{N_{\text{orb}}} \bigcup_{j=1}^{N_t(i)} \mathcal{P}(i, j), \quad (\text{A4})$$

where  $N_{\text{orb}}$  is the number of orbits in the library,  $N_t(i)$  is the number of time-steps in the integration of orbit  $i$  and

$$\mathcal{P}(i, j) = \begin{cases} [a_{ij}, b_{ij}] : 1 \leq i \leq N_{\text{orb}}, 1 \leq j \leq N_t(i), \\ \quad (i, j) \neq (N_{\text{orb}}, N_t(N_{\text{orb}})) \\ [a_{ij}, b_{ij}] : i = N_{\text{orb}}, j = N_t(N_{\text{orb}}) \end{cases} \quad (\text{A5})$$

with

$$a_{ij} \equiv \sum_{k=0}^{i-1} w_k + \sum_{m=0}^{j-1} w_i \frac{dt_i^{(m)}}{\tau_i} \quad (\text{A6})$$

and

$$b_{ij} \equiv \sum_{k=0}^{i-1} w_k + \sum_{m=0}^j w_i \frac{dt_i^{(m)}}{\tau_i} \quad (\text{A7})$$

( $w_0 \equiv 0$  and  $dt_i^{(0)} \equiv 0$ ,  $1 \leq i \leq N_{\text{orb}}$ ). The length of each subinterval equals the probability to find a particle in phase-space cell  $dV(i, j)$ :

$$b_{ij} - a_{ij} = w_i \frac{dt_i^{(j)}}{\tau_i} = p(i, j). \quad (\text{A8})$$

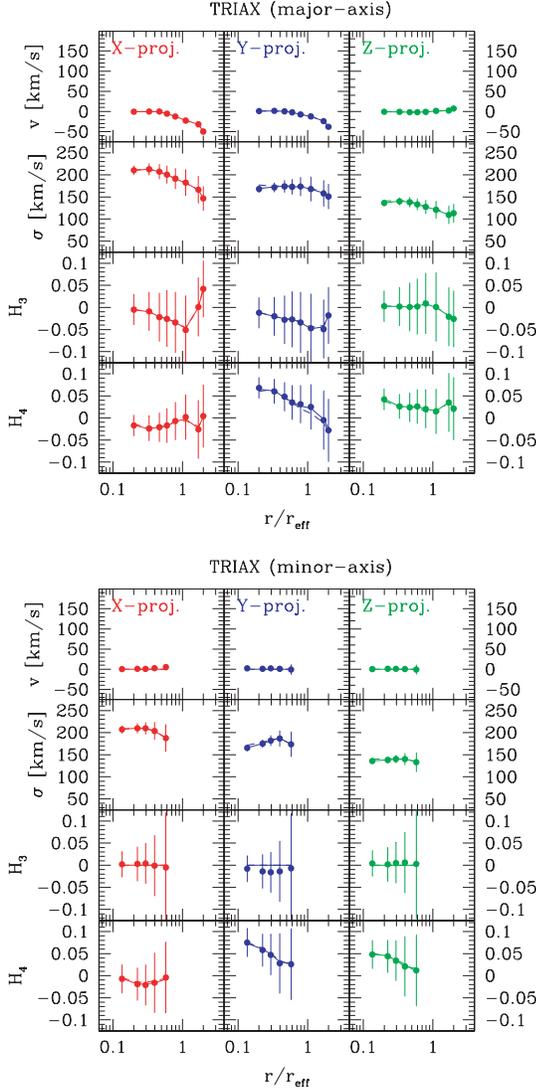
An  $N$ -body realization with  $N_l$  particles of, say, the luminous component can now be constructed by choosing  $N_l$  random numbers  $k \in [0, 1]$ . Each  $k$  falls into one subinterval  $\mathcal{P}(i_k, j_k)$ . Accordingly, particle  $k$  has to be dropped on orbit  $i_k$  during time-step  $j_k$ . The  $N$ -body realization of the dark halo can be constructed in a similar way, provided corresponding orbital weights are given. For the discussion in this paper we only need an  $N$ -body representation of the dark halo density profile, not of the halo kinematics. Thus, we can choose any DF for the halo that supports its density profile. The  $N$ -body realizations of the haloes in this work have been calculated from orbit weights that maximize the entropy of the dark matter DF. Their calculation is described in Thomas et al. (2007). We use  $N_l = \mathcal{N}_d = 50\,000$  particles to sample the Schwarzschild models up to  $10 r_{\text{eff}}$ .

Cutting the  $N$ -body realization beyond  $10 r_{\text{eff}}$  introduces a spurious correlation between the virial ratio  $2T/|W|$  and the halo concentration  $c_{\text{halo}}$  in the Schwarzschild fits. This is shown in Fig. A1: the less concentrated the halo, the larger  $2T/|W|$ . The systematic offset  $2T/|W| > 1$  of the  $N$ -body realization is thereby most likely caused by the fact that the halo is undersampled if the concentration is low ( $c_{\text{halo}} \lesssim 13$ ) and the halo scaling radius is correspondingly large. Then, the potential well is too shallow and  $|W|$  is underestimated. For the deviations from virial equilibrium discussed in Section 8.3 it follows that they are mostly an artefact of the  $N$ -body realization and not due to intrinsic non-stationarity of the models.

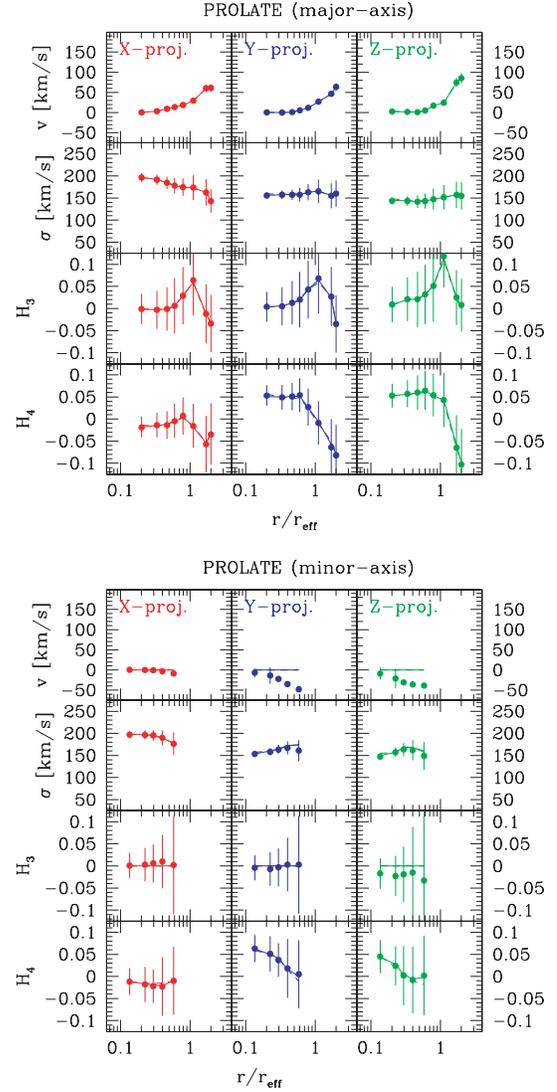
**APPENDIX B: DATA AND MODELS**

Figs B1–B6 survey the kinematics of the individual targets (dots with error bars) and the model fits (lines). For each target the best-fitting

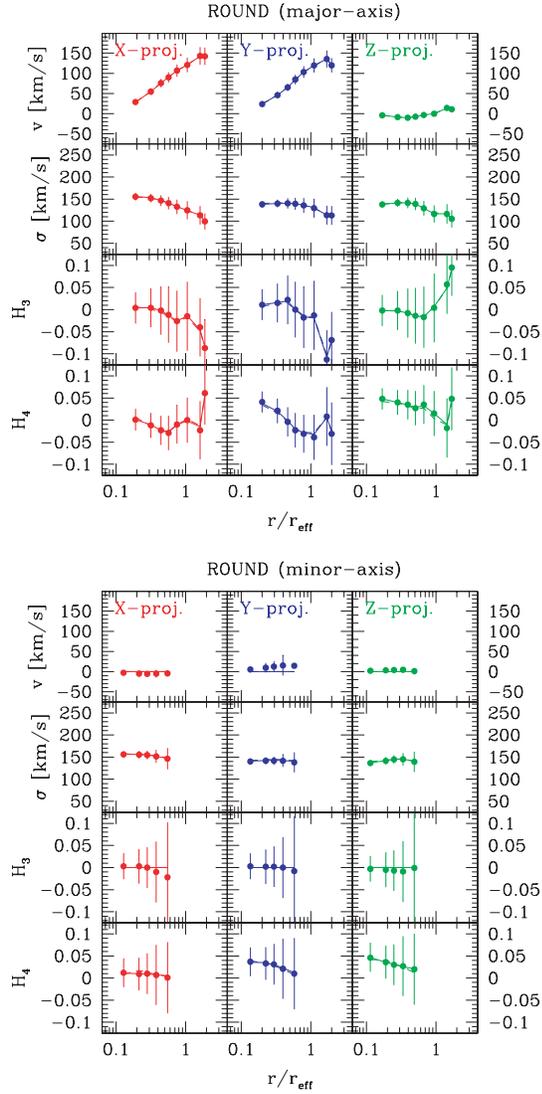
model (solid; best-fitting luminous mass-to-light ratios  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  are given in the captions and summarized in Table B1) as well as the best fit with  $\Upsilon \equiv \Upsilon_{\text{in}}$  (dashed) are shown. Table B1 also lists the accuracy of reconstructed total masses at  $r_{\text{eff}}$ .



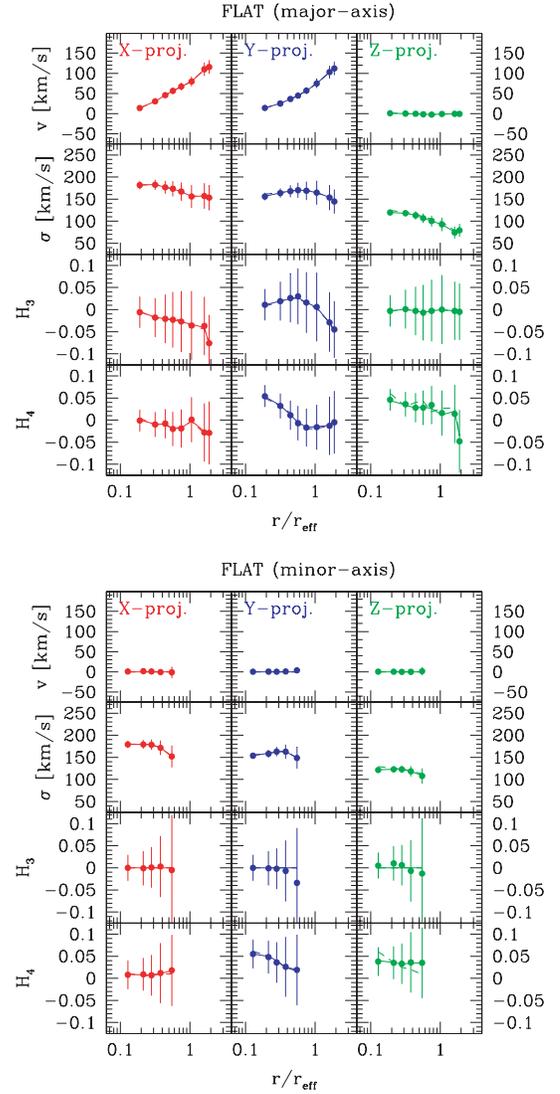
**Figure B1.** Comparison of model kinematics (solid lines: best fit; dashed lines  $\Upsilon_{\text{fit}} \equiv \Upsilon_{\text{in}}$ ) and input data (dots with error bars) for the TRIAX target. Top panel: major axis; bottom panel: minor axis. In each panel, from top to bottom  $v$ ,  $\sigma$ ,  $H_3$  and  $H_4$ ; left-hand panels:  $X$  projection; middle panels:  $Y$  projection; right-hand panels:  $Z$  projection. The luminous  $M/L$  values of the Schwarzschild fits are  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.9, 0.7, 0.4$  ( $X$ ,  $Y$  and  $Z$  models, respectively).



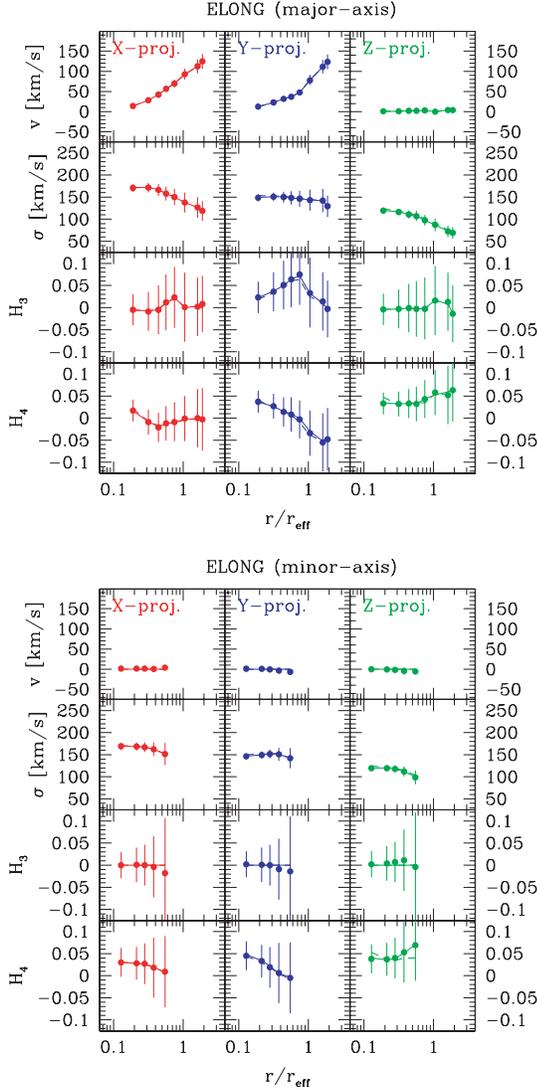
**Figure B2.** As Fig. B1, but for the PROLATE target.  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.9, 0.6, 0.8$  ( $X$ ,  $Y$  and  $Z$  models, respectively).



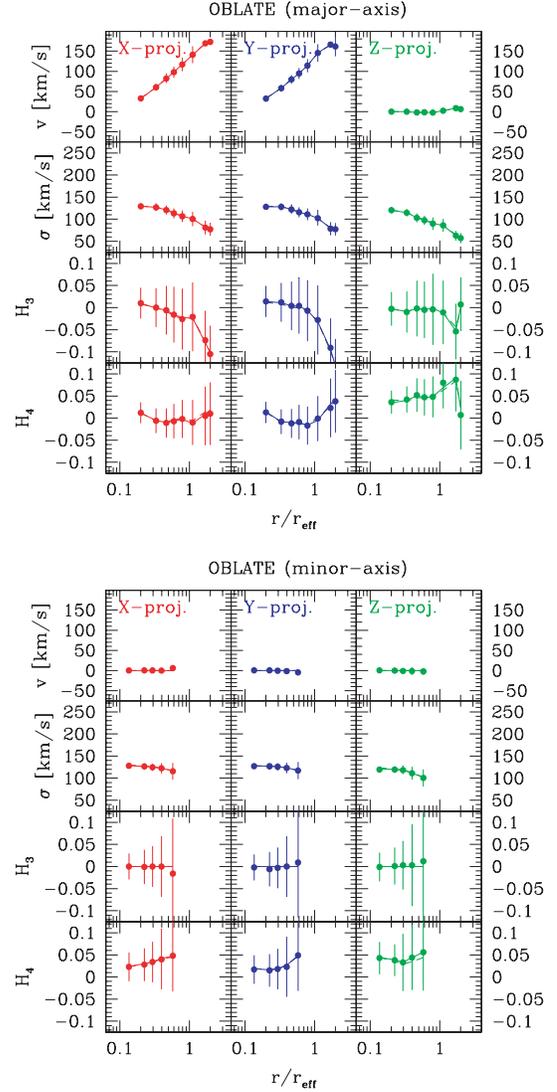
**Figure B3.** As Fig. B1, but for the ROUND target.  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.9, 0.8, 0.7$  (X, Y and Z models, respectively).



**Figure B4.** As Fig. B1, but for the FLAT target.  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.9, 0.7, 0.6$  (X, Y and Z models, respectively).



**Figure B5.** As Fig. B1, but for the ELONG target.  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.9, 0.6, 0.5$  (X, Y and Z models, respectively).



**Figure B6.** As Fig. B1, but for the OBLATE target.  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}} = 0.7, 0.9, 0.7$  (X, Y and Z models, respectively).

**Table B1.** Accuracy of reconstructed masses. (1) Merger remnant; (2–4) reconstructed  $\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$  for X, Y and Z models, respectively; (5–7) fractional error  $\Delta M \equiv (M^{(\text{fit})} - M^{(\text{in})})/M^{(\text{in})}$  of reconstructed (luminous + dark) mass inside  $r_{\text{eff}}$  for X, Y and Z models, respectively.

Remnant	$\Upsilon_{\text{fit}}/\Upsilon_{\text{in}}$			$\Delta M$ (per cent)		
	X	Y	Z	X	Y	Z
(1)	(2)	(3)	(4)	(5)	(6)	(7)
TRIAX	0.9	0.7	0.4	-2.3	-14.9	-45.8
PROLATE	0.9	0.6	0.8	3.7	-18.2	-3.8
ROUND	0.9	0.8	0.7	9.3	10.0	-13.1
FLAT	0.9	0.7	0.6	5.1	-4.4	-45.6
ELONG	0.9	0.7	0.5	10.0	-9.4	-47.8
OBLATE	0.7	0.9	0.7	2.8	0.4	-34.5

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.