

Triaxial orbit-based modelling of the Milky Way nuclear star cluster

A. Feldmeier-Krause,^{1,2★} L. Zhu,³ N. Neumayer,³ G. van de Ven,³ P. T. de Zeeuw^{1,4}
and R. Schödel⁵

¹European Southern Observatory (ESO), Karl-Schwarzschild-Straße 2, D-85748 Garching, Germany

²The Department of Astronomy and Astrophysics, The University of Chicago, 5640 S. Ellis Ave, Chicago, IL 60637, USA

³Max-Planck-Institut für Astronomie, Königsstuhl 17, D-69117 Heidelberg, Germany

⁴Sterrewacht Leiden, Leiden University, Postbus 9513, NL-2300 RA Leiden, the Netherlands

⁵Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía s/n, E-18008 Granada, Spain

Accepted 2016 December 21. Received 2016 December 19; in original form 2016 July 22

ABSTRACT

We construct triaxial dynamical models for the Milky Way nuclear star cluster using Schwarzschild’s orbit superposition technique. We fit the stellar kinematic maps presented in Feldmeier et al. The models are used to constrain the supermassive black hole mass M_{\bullet} , dynamical mass-to-light ratio Υ and the intrinsic shape of the cluster. Our best-fitting model has $M_{\bullet} = (3.0_{-1.3}^{+1.1}) \times 10^6 M_{\odot}$, $\Upsilon = (0.90_{-0.08}^{+0.76}) M_{\odot}/L_{\odot,4.5\mu\text{m}}$ and a compression of the cluster along the line of sight. Our results are in agreement with the direct measurement of the supermassive black hole mass using the motion of stars on Keplerian orbits. The mass-to-light ratio is consistent with stellar population studies of other galaxies in the mid-infrared. It is possible that we underestimate M_{\bullet} and overestimate the cluster’s triaxiality due to observational effects. The spatially semiresolved kinematic data and extinction within the nuclear star cluster bias the observations to the near side of the cluster, and may appear as a compression of the nuclear star cluster along the line of sight. We derive a total dynamical mass for the Milky Way nuclear star cluster of $M_{\text{MWNSC}} = (2.1 \pm 0.7) \times 10^7 M_{\odot}$ within a sphere with radius $r = 2 \times r_{\text{eff}} = 8.4$ pc. The best-fitting model is tangentially anisotropic in the central $r = 0.5$ – 2 pc of the nuclear star cluster, but close to isotropic at larger radii. Our triaxial models are able to recover complex kinematic substructures in the velocity map.

Key words: Galaxy: centre – Galaxy: kinematics and dynamics.

1 INTRODUCTION

The Milky Way nuclear star cluster is the ideal object to study the dynamics of a stellar system around a supermassive black hole. At a distance of 8 kpc it is close enough to resolve the individual stars, and measure discrete velocities in three dimensions. Modelling the stellar kinematics can constrain the mass distribution of the star cluster, and reveal the presence of a central dark massive object. In the special case of our own Galaxy, it is possible to observe Keplerian orbits of stars around a dark, point-mass-like object in the Galactic Centre. These observations constrain this dark object to be a supermassive black hole with a mass of $(4.1 \pm 0.6) \times 10^6 M_{\odot}$ (Ghez et al. 2008), $(4.3 \pm 0.39) \times 10^6 M_{\odot}$ (Gillessen et al. 2009) or $(4.02 \pm 0.20) \times 10^6 M_{\odot}$ (Boehle et al. 2016). Unfortunately, similar high-resolution observations are not yet possible in other galaxies.

Already in the 1970s the requirement of a central supermassive black hole in the Galactic Centre was discussed to explain obser-

vatational data (e.g. Oort 1977). Several studies used stellar radial velocities to constrain the mass distribution in the Galactic Centre (e.g. Rieke & Rieke 1988; McGinn et al. 1989; Sellgren et al. 1990; Genzel et al. 1996; Haller et al. 1996). Also stellar proper motions were used to study the Galactic Centre mass distribution (Schödel, Merritt & Eckart 2009). Several studies combined radial velocity and proper motion data (Trippe et al. 2008; Do et al. 2013; Fritz et al. 2016). The mass distribution was derived using the spherical Jeans (1922) equations or the projected mass estimators of Bahcall & Tremaine (1981) for spherical systems. These studies found that a central dark mass of 2 – $5 \times 10^6 M_{\odot}$ is required to explain the observations.

Together with the increase of observational data, also the modelling became more advanced. Trippe et al. (2008) included the rotation of the nuclear star cluster in the modelling, although the rotation velocity of their data was too high (Schödel et al. 2009; Feldmeier et al. 2014). Feldmeier et al. (2014) and Chatzopoulos et al. (2015a) studied the Milky Way nuclear star cluster using axisymmetric Jeans models. Chatzopoulos et al. (2015a) showed the advantages of axisymmetric models over spherical Jeans models, which cannot explain the observed asymmetry of the velocity

* E-mail: afeldmei@uchicago.edu

dispersion of proper motions parallel and perpendicular to the Galactic plane. The nuclear star cluster appears to be flattened in its light distribution (Schödel et al. 2014) as well as in the kinematics (Chatzopoulos et al. 2015a). Most studies showed that the nuclear star cluster kinematics is in agreement with isotropy (Schödel et al. 2009; Do et al. 2013; Chatzopoulos et al. 2015a), although the uncertainties are quite large (e.g. Do et al. 2013). All these models assumed a constant mass-to-light ratio for the light distribution of the cluster.

In this study, we relax the assumption of axisymmetry and use triaxial orbit-based Schwarzschild (1979) models. Orbit-based models make no assumptions on the velocity anisotropy of the stellar motions, as Jeans models do. Further, the higher moments of the kinematics can also be included (Rix et al. 1997), which is important to break the degeneracy of mass and anisotropy in dynamical models.

Orbit-based models are commonly used to analyse line-of-sight velocity data of other galaxies (e.g. van der Marel et al. 1998; Gebhardt et al. 2000; Valluri et al. 2005; van den Bosch et al. 2008), and are an excellent tool to detect and measure the masses of supermassive black holes and dark matter haloes. For extragalactic systems, the data are usually obtained from integrated light observations. Each data point contains the accumulated kinematics of many stars, weighted by their respective brightness. However, modelling the dynamics of integrated light data may be prone to systematic uncertainties, and bias the results of the central black hole mass. Therefore, it is interesting to test dynamical models on systems for which we know the central black hole mass from other independent measurements. The Milky Way nuclear star cluster is a good object for this kind of test. Also megamaser disc galaxies are useful to validate stellar dynamical black hole measurements. Black hole mass measurements from megamasers are very precise with uncertainties of only about 10 per cent. However, there is currently only one megamaser disc galaxy with a stellar dynamical black hole mass measurement (van den Bosch et al. 2016), NGC 4258. Different dynamical studies found either a 15 per cent lower or a 25 per cent higher black hole mass than the maser measurement (Siopis et al. 2009; Drehmer et al. 2015).

We use the triaxial orbit-based code by van den Bosch et al. (2008) to model the light distribution and line-of-sight kinematics of the Milky Way nuclear star cluster. We use the spectroscopic data cube constructed by Feldmeier et al. (2014) for the kinematic data, and derive a surface brightness distribution using *Spitzer* 4.5 μm and NACO *H*-band images. We assume a galactocentric distance of 8 kpc (Malkin 2013) and a position angle $31^\circ 40'$ east of north (J2000.0 coordinates; Reid & Brunthaler 2004) with respect to the Galactic plane. This paper is organized as follows: We describe the kinematic and photometric data in Section 2. The dynamical models are introduced in Section 3. Section 4 discusses the results, and Section 5 summarizes the main conclusions.

2 DESCRIPTION OF THE DATA

2.1 Kinematic data

The line-of-sight velocity distribution (LOSVD) provides constraints on the dynamical structure of stellar systems. To extract this information, we used the near-infrared *K*-band spectroscopic data cube of Feldmeier et al. (2014), which has a pixel scale of 2.22 arcsec pixel⁻¹. We used the data cube that was cleaned of foreground stars and bright stars. The cleaned data cube contains the light of the old red giant star population.

We fitted the LOSVD as in Feldmeier et al. (2014), i.e. on the stellar CO absorption lines (2.2902–2.365 μm) with the IDL routine PPF (Cappellari & Emsellem 2004) and the high-resolution spectra of Wallace & Hinkle (1996) as template stars. We applied the same spatial binning as Feldmeier et al. (2014), resulting in 175 spatial bins. Feldmeier et al. (2014) fitted only the velocity V and velocity dispersion σ . However, we fitted in addition also higher moments of the LOSVD, in particular the Gauss–Hermite parameters h_3 and h_4 . We added noise to each of the 175 integrated light spectra in 100 Monte Carlo simulation runs and obtained a distribution for each moment of the LOSVD. The mean and standard deviation of the Monte Carlo distribution are taken as measurement and 1σ uncertainty of the kinematics.

Since the Milky Way nuclear star cluster is at a distance of only 8 kpc, the spectroscopic observations are spatially semiresolved. Bright stars can be resolved individually, and contribute a large fraction of the flux. For that reason we used the cleaned data cube of Feldmeier et al. (2014), where bright stars were excluded. However, the kinematic maps still show stochastic shot noise. As a consequence, the difference of the kinematics in adjacent bins can be higher than their uncertainties, which causes problems when we model the kinematics. The stochastic noise can be mistaken for signal, and this means the best fit will be achieved by modelling the shot noise. To prevent this, we increased our kinematic uncertainties ϵ_V such that the difference of the measurement in two adjacent bins ($V_i - V_j$) is less than the sum of their uncertainties ($\epsilon_{V_i} + \epsilon_{V_j}$). We did this for the velocity V , velocity dispersion σ , h_3 and h_4 , and find that it is required for about 68 per cent of the kinematic data uncertainties. Additionally, we point-symmetrized the kinematics using the procedure of van den Bosch & de Zeeuw (2010). The median uncertainties of V , σ , h_3 and h_4 are 24.6 km s⁻¹, 18.4 km s⁻¹, 0.15 and 0.17, respectively. Our resulting kinematic maps are consistent with the maps of Feldmeier et al. (2014). We find rotation in the velocity map of approximately 50 km s⁻¹ and an increase in the velocity dispersion from about 65 km s⁻¹ towards $\sigma_{\text{max}} = 135$ km s⁻¹ at the centre. The kinematic maps are shown on the top row of Fig. 1, the uncertainties are shown on the bottom row.

2.2 Imaging data and surface brightness distribution

The light distribution of the nuclear star cluster traces the stellar density. We require the two-dimensional light distribution of the red giant stars, which are our kinematic tracers. The extinction is high at optical wavelengths in the Galactic Centre ($A_V \sim 30$ mag; Scoville et al. 2003; Gao, Li & Jiang 2013), therefore we used near- and mid-infrared images.

For the central 40.4 arcsec \times 40.4 arcsec (1.6 pc \times 1.6 pc) we used the high-resolution NACO *H*-band mosaic of Schödel et al. (2009), which has a spatial scale of 0.027 arcsec pixel⁻¹. We preferred the *H* band over the *K* band in order to avoid light from gas emission lines in the *K* band (Br γ and He I; Paumard, Maillard & Morris 2004). Our kinematic tracers are cool late-type stars, but there are also more than 100 hot, young stars located in the centre of the cluster, within a projected radius $r = 0.5$ pc (~ 12.8 arcsec, Paumard et al. 2006). We masked out the young stars from the image with a 15 pixel radius. For the bright red supergiant IRS 7 we used a larger mask with a 30 pixel radius. Beyond the central 0.5 pc, the nuclear star cluster light is dominated by cool stars, and the contribution of young stars is negligible (Feldmeier-Krause et al. 2015).

For the large-scale light distribution, we used *Spitzer* IRAC images (Stolovy et al. 2006). These images were corrected for dust extinction and polycyclic aromatic hydrocarbon (PAH) emission by

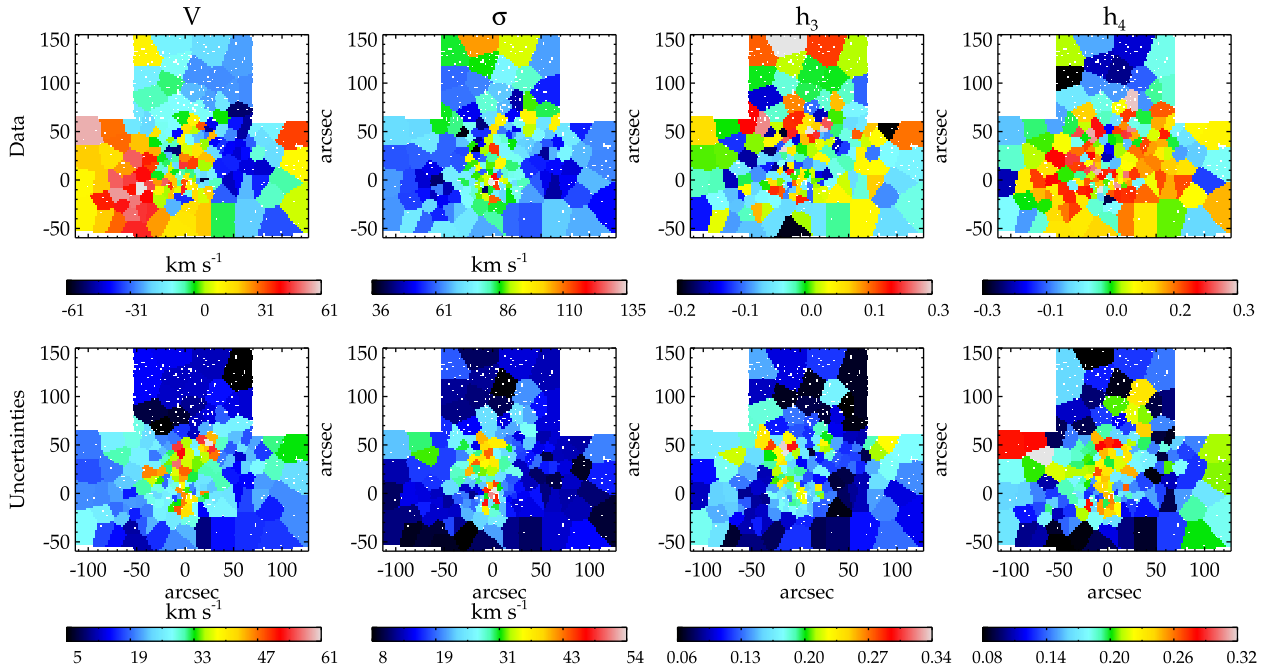


Figure 1. Kinematic data (top row) and respective uncertainties (bottom row). The columns denote velocity V , velocity dispersion σ , Gauss–Hermite moments h_3 and h_4 . White pixels are due to excluded bright stars.

Schödel et al. (2014). We used the extinction and emission corrected $4.5\ \mu\text{m}$ image to measure the light distribution. The image was smoothed to a scale of $5\ \text{arcsec}\ \text{pixel}^{-1}$, and extends over $\sim 270\ \text{pc} \times 200\ \text{pc}$. We excluded a central circle with $r = 0.6\ \text{pc}$ ($\sim 15.4\ \text{arcsec}$) to avoid contribution from ionized gas emission and young stars. In addition, we masked out the young Quintuplet star cluster (Figer, McLean & Morris 1999), and the dark $20\text{-km}\ \text{s}^{-1}$ cloud M-0.13-0.08 (García-Marín et al. 2011).

We used the `MGE_FIT_SECTORS` package (Cappellari 2002) to derive the surface brightness distribution. The Multi-Gaussian-Expansion (MGE) model (Emsellem, Monnet & Bacon 1994) has the advantage that it can be deprojected analytically. We measured the photometry of the two images along the major axis and the minor axis. We assumed that the cluster’s major axis is aligned along the Galactic plane, as found by Schödel et al. (2014), and constant. The centre is the position of Sgr A*, which is the radio source associated with the Galactic Centre supermassive black hole. We fitted a scale factor to match the photometry of the two images in the region where they overlap ($16\text{--}27.8\ \text{arcsec}$). Then we measured the photometry on each image along 12 angular sectors, and converted the NACO photometry to the *Spitzer* flux. Assuming fourfold symmetry, the measurements of four quadrants are averaged on elliptical annuli with constant ellipticity. Using the photometric measurements of the two images, we fitted a set of two-dimensional Gaussian functions, taking the point spread function (PSF) of the NACO image into account.

A comparison with the surface brightness profile of Fritz et al. (2016, their fig. 2) showed that our profile is steeper in the central $\sim 30\ \text{arcsec}$. A possible reason is the small overlap region of the *Spitzer* and NACO images, and that the *Spitzer* flux could be too high at the centre. Maybe the PAH emission correction of the *Spitzer* image was too low. The mid-infrared dust emission is significant out to almost $1\ \text{arcmin}$. Fritz et al. (2016) used NACO H - and K_S -band images in the central $r = 20\ \text{arcsec}$. Out to $1000\ \text{arcsec}$ ($\sim 39\ \text{pc}$) they used *Hubble Space Telescope* WFC3 data (M127

Table 1. The MGE fit parameters for the $4.5\ \mu\text{m}$ *Spitzer*/IRAC dust extinction and PAH emission corrected image in combination with the NACO H -band mosaic scaled to *Spitzer* flux. I_{scaled} is the peak surface brightness used in the dynamical modelling, σ_{MGE} is the standard deviation and q_{MGE} is the axial ratio of the Gaussian components. I_{unscaled} is the peak surface brightness before scaling to Fritz et al. (2016).

I_{scaled} ($10^4 L_{\odot, 4.5\ \mu\text{m}}\ \text{pc}^{-2}$)	σ_{MGE} (arcsec)	q_{MGE}	I_{unscaled} ($10^4 L_{\odot, 4.5\ \mu\text{m}}\ \text{pc}^{-2}$)
0.86	1.7	0.30	312
32.4	10.4	0.34	164
89.8	15.0	0.82	257
18.5	52.1	0.95	30.0
17.0	98	0.36	29.3
7.1	154	0.95	7.4
4.8	637	0.36	4.9
3.2	2020	0.30	3.2
1.3	4590	0.81	1.3

and M153 filters) and public VISTA Variables in the Via Lactea Survey images (H and K_S bands; Saito et al. 2012). We lowered the intensities of the central Gaussians by scaling our averaged profile to the one-dimensional flux density profile of Fritz et al. (2016). As a result the amplitudes of the inner Gaussians become smaller, but the outer Gaussians ($\sigma_{\text{MGE}} > 100\ \text{arcsec} \sim 4\ \text{pc}$) are nearly unchanged. We list the components of the MGE in Table 1 and plot the surface brightness profile in Fig. 2 (upper panel). We also show the projected axial ratio q_{proj} as a function of radius in the lower panel of Fig. 2. Out to the central $1\ \text{pc}$, q_{proj} is increasing from 0.4 to 0.7. Schödel et al. (2014) found a mean axial ratio of 0.71 ± 0.02 for the nuclear star cluster. This is in agreement with our maximum value of q_{proj} . However, q_{proj} decreases at larger radii, as the contribution from the nuclear stellar disc becomes more important and the light distribution therefore flatter.

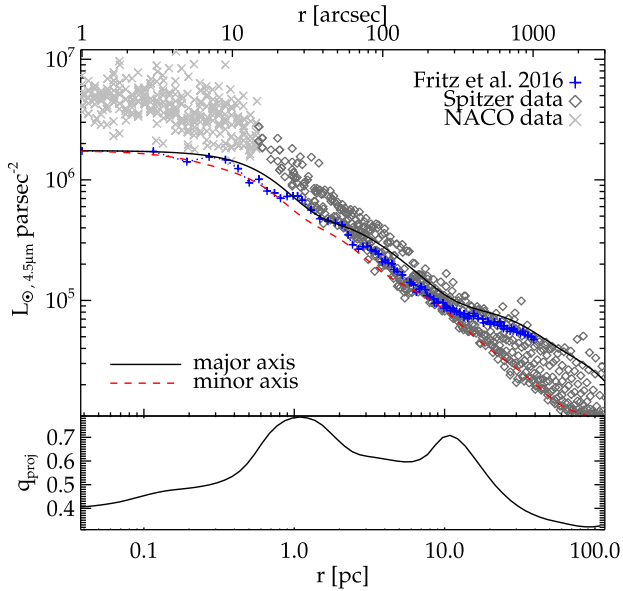


Figure 2. Upper panel: surface brightness profile derived from a dust extinction and PAH emission corrected *Spitzer*/IRAC 4.5 μm image and NACO *H*-band mosaic for the centre, scaled to the measurements of Fritz et al. (2016, blue crosses). The black full line denotes the MGE fit along the major axis, and the red dashed line along the minor axis. Lower panel: projected axial ratio q_{proj} as a function of r .

We note that there are three main differences with the surface brightness distribution derived by Feldmeier et al. (2014): (1) We used an *H*-band instead of a K_s -band NACO image to avoid ionized gas emission; (2) we masked young stars in the NACO image to match the distribution of stars used as kinematic tracers; and (3) we scaled the central photometry to the flux density data of Fritz et al. (2016) to avoid a possible overestimation of the central flux when scaled to the *Spitzer* image. All three changes influence only the central part of the surface brightness distribution, as ionized gas emission and light from young stars are only important in the central parsec.

The surface brightness profile is deprojected to obtain the three-dimensional spatial stellar light distribution. In general, the deprojection of a surface brightness profile is non-unique (e.g. Rybicki 1987; Franx 1988), and our deprojection is only one possible solution. The MGE deprojection produces smooth intrinsic densities, which are in agreement with the photometric observations (Cappellari 2002).

3 DYNAMICAL MODELS OF THE MILKY WAY NUCLEAR STAR CLUSTER

3.1 Schwarzschild’s method

Orbit-based models or Schwarzschild models are a useful tool to model the dynamics of stellar systems by orbit superposition. The first step of Schwarzschild’s method is to integrate the equations of motion for a representative library of stellar orbits in a gravitational potential Φ . Then the observables for each orbit are computed, considering projection, PSF convolution and pixel binning. The next step is to find orbital weights to combine the orbits such that they reproduce the observed data. Schwarzschild models are a powerful tool to recover the intrinsic kinematical structure and the underlying gravitational potential (Schwarzschild 1979; van de Ven, de Zeeuw

& van den Bosch 2008; van den Bosch & van de Ven 2009). We refer the reader for further details to van den Bosch et al. (2008) for implementation and van de Ven et al. (2008) for verification of the triaxial Schwarzschild code.

3.1.1 Mass model

We calculated orbits in the combined gravitational potential of a supermassive black hole Φ_* and the star cluster Φ_* , inferred from the imaging data. As we run triaxial models, there are three intrinsic shape parameters, q , p and u , for the cluster. The shape parameters characterize the axial ratios for the long, intermediate and short axes a , b and c . They are defined as $q = c/a$, $p = b/a$ and $u = a'/a$, where a' is the length of the longest axis a projected on the sky. Thus, u represents the compression of a due to projection on the sky. Each set of axial ratios refers to a set of viewing angles (ϑ , ϕ , ψ , see also van den Bosch et al. 2008). The surface brightness distribution is deprojected given the intrinsic shape parameters q , p , u and multiplied with the dynamical mass-to-light ratio Υ to get the intrinsic stellar mass density ϱ_* . From Poisson’s equation $\nabla^2\Phi_* = 4\pi G\varrho_*$ one calculates the gravitational potential. We did this for different values of the black hole mass M_* , dynamical mass-to-light ratio Υ and different shape parameters. In total our model has five free parameters, M_* , Υ , q , p and u .

Besides the considered stellar population and the supermassive black hole, there are other components within the nuclear star cluster, which we neglected. We measured a dynamical mass-to-light ratio, which combines the stellar mass-to-light ratio with other components. These components are the young stars, ionized gas, neutral gas and dark matter. The young stars are at a distance of about 0.5 pc from the supermassive black hole. The lower limit of the total mass of young stars is $12\,000 M_\odot$. However, the total enclosed extended mass in the same region is $\sim 10^6 M_\odot$ (Oh, Kim & Figer 2009; Feldmeier et al. 2014), and the mass of the supermassive black hole is $4 \times 10^6 M_\odot$. The mass of the young stars is therefore probably negligible. The hot ionized gas has a mass of only a few $100 M_\odot$ (Ferrière 2012), and cannot influence the stellar dynamics significantly. The neutral gas in the circumnuclear disc may contribute more mass, estimates range from 10^4 (Etxaluze et al. 2011; Requena-Torres et al. 2012) to $10^6 M_\odot$ (Christopher et al. 2005), though this is probably the upper limit (Genzel, Eisenhauer & Gillessen 2010). The circumnuclear disc extends over a distance of about 1 pc to more than 5 pc from the centre. At 5 pc, the total enclosed mass is $\sim 10^7 M_\odot$ (McGinn et al. 1989; Feldmeier et al. 2014). We decided to neglect the mass distribution of the circumnuclear disc in our dynamical models, since it is very uncertain, and makes up only 0.1 to 10 per cent of the enclosed mass. The contribution of dark matter to the nuclear star cluster mass is also neglected. Linden (2014) showed that the fraction of dark matter in the central 100 pc of the Milky Way is about 6.6 per cent, assuming the traditional dark matter profile of Navarro, Frenk & White (1996).

3.1.2 Orbit library

The orbit library should be as general as possible and representative for the gravitational potential. We assumed that the orbits are regular and that three integrals of motion, E , I_2 and I_3 , are conserved. The orbit families consist of box orbits, which can cross the centre and have an average angular momentum of zero, and three types of tube orbits, which avoid the centre. The tube orbits are divided in short-axis tube orbits, which have non-zero mean angular momentum (L_z)

around the short axis, outer and inner long-axis tube orbits, which have non-zero mean angular momentum $\langle L_x \rangle$ around the long axis. The orbit grid should sample the entire phase space. It has to be dense enough to suppress discreteness noise, but integration has to be done in a reasonable amount of computing time.

We followed van den Bosch et al. (2008) and sampled the orbit energy E using a logarithmic grid in radius. Each energy E is linked to the radius R_c by calculating the potential at $(x, y, z) = (R_c, 0, 0)$. We sampled $N_E = 35$ energies calculated from R_c in logarithmic steps ranging from $R_c = 10^{0.5}$ to $10^{4.2}$, i.e. 3.16 arcsec to 4.4 or 0.12 pc to 616.5 pc. We note that the outer radius is about 3.5 times the outermost Gaussian σ_{MGE} of the MGE fit. We tested lower values of the inner radius but found consistent results. For each energy, the starting point of an orbit was selected from a linear grid over 14 values each. For details on the orbit sampling, we refer to van den Bosch et al. (2008). In total, we have $N_E \times N_{I_2} \times N_{I_3} = 35 \times 14 \times 14 = 6860$ orbits. Each orbit was integrated over 200 periods, and sampled on 100 000 points per orbit. For each orbit, we stored the intrinsic and projected properties. The projected orbits are stored in a (x', y', v_z) grid, with PSF convolution and pixel size of the observed data taken into account. The velocities were stored in 183 bins between $-7.4 \sigma_{\text{max}}$ and $+7.4 \sigma_{\text{max}}$. These numbers guarantee a proper sampling of the observed velocity profiles (Cretton et al. 1999).

3.1.3 Solving the orbital weight distribution

The model has to fit the kinematic data, the intrinsic and the projected mass distribution. The fit was done by finding a linear combination of the orbits, and solving for orbital weights γ_i . Each orbital weight corresponds to a mass on the respective orbit i , and the weights γ_i are therefore non-negative. We used the non-negative least-squares algorithm of Lawson & Hanson (1974), which was also used by Rix et al. (1997), van der Marel et al. (1998) and Cretton et al. (1999). One of the fitting constraints is to make sure that the model is self-consistent. It is required that the orbit superposition reproduces the intrinsic and projected aperture masses within 2 per cent, which is the typical accuracy of the observed surface brightness (van den Bosch et al. 2008).

3.2 Constraining the input parameters

We ran 4899 models with different parameter combinations of M_\bullet , Υ , q , p , u . The black hole mass M_\bullet was sampled in logarithmic steps of 0.2 from 5.5 to 7.5, starting with 6.3 (i.e. $M_\bullet \sim 2 \times 10^6 M_\odot$). The mass-to-light ratio Υ was linearly sampled between 0.1 and 2.0 with steps of 0.04, with a starting value of 0.6 (in units of $M_\odot/L_{\odot,4.5 \mu\text{m}}$). The starting model had $(q, p, u) = (0.29, 0.84, 0.99)$. We sampled different combinations of (q, p, u) with a step size of $(0.01, 0.02, 0.01)$, with the boundaries $0.05 < q < 0.29$, $0.40 < p < 0.99$ and $0.70 < u \leq 0.99$. When fitting the orbital weights, we already applied the self-consistency criteria to make sure that the photometry is fitted with a high accuracy of at least 2 per cent for each model. The best fit of the five parameters M_\bullet , Υ , q , p , u , however, was found by calculating the χ^2 of the kinematic moments. After the starting model, we computed different combinations of $(M_\bullet, \Upsilon, q, p, u)$ in the grid. We used a χ^2 analysis to find regions with good fits and started new models with a finer sampling at low χ^2 . This was done in several iterations, and it was not necessary to compute each point in the $(M_\bullet, \Upsilon, q, p, u)$ grid.

We measured four kinematic moments in 175 spatial bins of the spectroscopic map in Section 2.1. In each bin, we measured the

velocity v , velocity dispersion σ and the higher Gauss–Hermite moments h_3 and h_4 . Thus, the number of observables is the number of spatial bins times the number of kinematic moments, in our case, $N = 175 \times 4 = 700$. We minimized the function

$$\chi^2 = \sum_{i=1}^{175} \left(\frac{V_i - V_i^{\text{mod}}}{\epsilon_{V_i}} \right)^2 + \left(\frac{\sigma_i - \sigma_i^{\text{mod}}}{\epsilon_{\sigma_i}} \right)^2 + \left(\frac{h_{3,i} - h_{3,i}^{\text{mod}}}{\epsilon_{h_{3,i}}} \right)^2 + \left(\frac{h_{4,i} - h_{4,i}^{\text{mod}}}{\epsilon_{h_{4,i}}} \right)^2, \quad (1)$$

where v , σ , h_3 and h_4 denote the kinematic measurements, ϵ the respective measurement uncertainties, v^{mod} , σ^{mod} , h_3^{mod} and h_4^{mod} the model kinematics, to find the best-fitting model.

3.3 Modelling results

3.3.1 The best-fitting model

Our best-fitting parameters are $M_\bullet = 3.0 \times 10^6 M_\odot$, $\Upsilon = 0.90$, $q = 0.28$, $p = 0.64$, $u = 0.99$. This corresponds to best-fitting viewing angles $\vartheta = 80^\circ$, $\varphi = 79^\circ$, $\psi = 91^\circ$. We show the surface brightness map and the symmetrized kinematic maps in Fig. 3. The upper row are the data, the lower row are the maps of the best-fitting model. The misalignment of the kinematic rotation axis with respect to the photometry and the perpendicular rotating substructure at ~ 20 arcsec (~ 0.8 pc) found by Feldmeier et al. (2014) are well reproduced in the model velocity map. We also show the data and the best-fitting model in Fig. 4, the panels denote surface brightness, v , σ , h_3 and h_4 . The surface brightness map is reproduced within 1 per cent. The highest and lowest values of v , h_3 , h_4 , and the lowest values of σ in the data have higher absolute values than in the best-fitting model, but are consistent within their uncertainties. The best fit has $\chi^2 = 290$. With $M = 5$ fitted parameters and $N = 4 \times 175 = 700$ observational constraints, this means $\chi_{\text{red}}^2 = 0.42$. That χ_{red}^2 is less than one is partially due to the large uncertainties of the kinematics, and the fact that the kinematic measurements are correlated.

We illustrate the distribution of χ^2 for the 4899 models in Fig. 5. We plot each combination of parameters. Red colours denote low χ^2 , bluer, smaller symbols denote high χ^2 . The black cross denotes the best-fitting model. The observed projected flattening q_{MGE} of the surface brightness profile constrains the viewing angles and thus also the intrinsic shape parameters. In particular, a flat $q_{\text{MGE}} \ll 1$ means that the stellar system is observed along one of the principal planes (van den Bosch et al. 2008). We obtained as lowest value $q_{\text{MGE}} = 0.30$, this limits the possible projections and puts an upper limit on the parameter q . Likewise, the value of $u = 0.99$ is the boundary value of the grid. The values of q , p and u denote the respective minimum values of the radially varying intrinsic axial ratios q_{intr} , p_{intr} and u_{intr} over the entire radial range of the photometry, i.e. the nuclear stellar disc and the embedded nuclear star cluster. The upper left panel of Fig. 5 shows that for each value of u , the best-fitting Υ is approximately 0.90. A similar behaviour is found with q and p . There is only a slight increase of the best-fitting Υ with higher values of p . At the same time, the best-fitting values of M_\bullet do not show a strong dependence on q , p or u (second row) in the allowed parameter ranges. The intrinsic shape parameters do not influence our best fit for M_\bullet , as this measurement is mostly made from the inner bins and the outer bins contribute little. The outer bins, however, contribute to the intrinsic shape fit. The

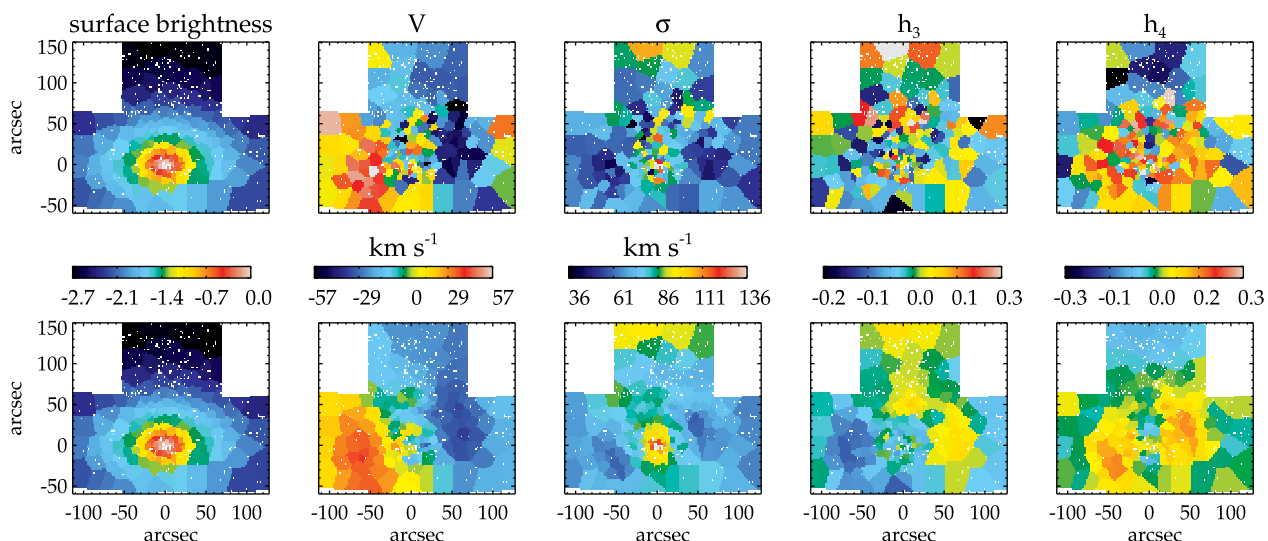


Figure 3. Comparison of the observed stellar surface brightness and kinematics (top row) and the best-fitting Schwarzschild model. The columns denote surface brightness, velocity v , velocity dispersion σ , Gauss–Hermite moments h_3 and h_4 .

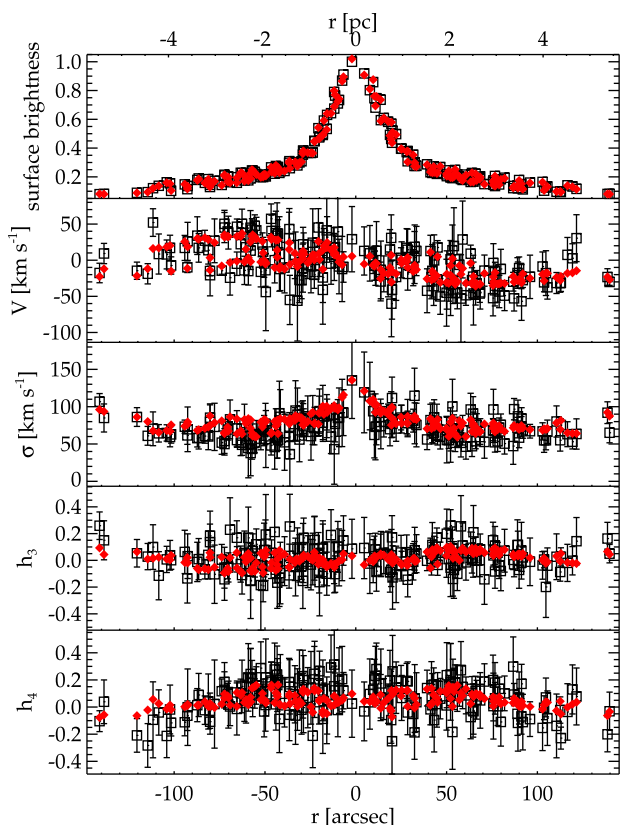


Figure 4. The surface brightness and kinematics of the best-fitting model (red filled diamond symbols) and the data (open black squares and error bars) as a function of radius. Data points in the Galactic east are plotted at negative radii, Galactic west at positive radii. From top to bottom: surface brightness, velocity v , velocity dispersion σ and higher Gauss–Hermite moments h_3 and h_4 .

supermassive black hole mass and the dynamical mass-to-light ratio are correlated. For higher values of Υ , a lower M_* fits the data.

We show how χ^2 depends on the different parameters in Fig. 6. The best-fitting model, which has the lowest χ^2 , is marked as blue

asterisk symbol. The blue lines denote the 1σ , 2σ and 3σ confidence limits, corresponding to $\Delta\chi^2 = 5.9$, 11.3 and 18.2. The red line illustrates the standard deviation of χ^2 itself, i.e. $\sqrt{2(N-M)} = 37.3$, where $N = 700$ and $M = 5$. This value was used as confidence limit by van den Bosch & van de Ven (2009). In Table 2, we list the 1σ and 3σ uncertainties.

3.3.2 Mass profile and intrinsic shape

We show the enclosed total mass as a function of the projected radius in Fig. 7, grey shaded contours are the 3σ uncertainties. The mass was computed within spherical shells. We also plot the results of various other studies. Most studies assumed a spherical shape of the cluster, Feldmeier et al. (2014) and Chatzopoulos et al. (2015a) assumed axisymmetry. Some of the studies used also different Galactocentric distances, so we scaled the masses using $R_0 = 8.0$ kpc. Our results are in agreement with several studies in the central 100 arcsec. However, we obtain a lower enclosed mass than Tripp et al. (2008) and Oh et al. (2009), who used the Jeans equation for a spherical system to obtain the enclosed mass. At larger radii $r \approx 400$ arcsec (~ 15.5 pc) beyond the reach of our kinematic data, we obtained a higher mass than Lindqvist et al. (1992), but we are in agreement at $r = 750$ –1600 arcsec. Their data extend to larger radii, but their assumption of spherical symmetry does no longer hold at such large radii.

The enclosed total mass within a sphere with the radius $r = 8.4$ pc, i.e. about two times the effective radius of the nuclear star cluster, is $M_{\text{MWNSC}} = (2.1 \pm 0.7) \times 10^7 M_\odot$, here we give the 3σ uncertainty.

The black hole influences the stellar kinematics only at the centre of the nuclear star cluster. Out to $r = 53$ arcsec (~ 2 pc), the best-fitting mass of the black hole ($M_* = 3.0 \times 10^6 M_\odot$) is higher than the enclosed stellar mass of our best-fitting model. Merritt (2004) defined the radius of influence of a black hole as the radius where the enclosed stellar mass equals two times the black hole mass. With this definition and a black hole mass of $4 \times 10^6 M_\odot$, we obtain $r_{\text{infl}} = (104^{+56}_{-29})$ arcsec, i.e. approximately $(4.0^{+2.2}_{-1.1})$ pc. This result is in agreement with Alexander (2005), who found $r_{\text{infl}} = 3$ pc. The kinematic measurements at larger radii have little influence on the black hole mass measurement, but are important to constrain the orbital structure and dynamical mass-to-light ratio.

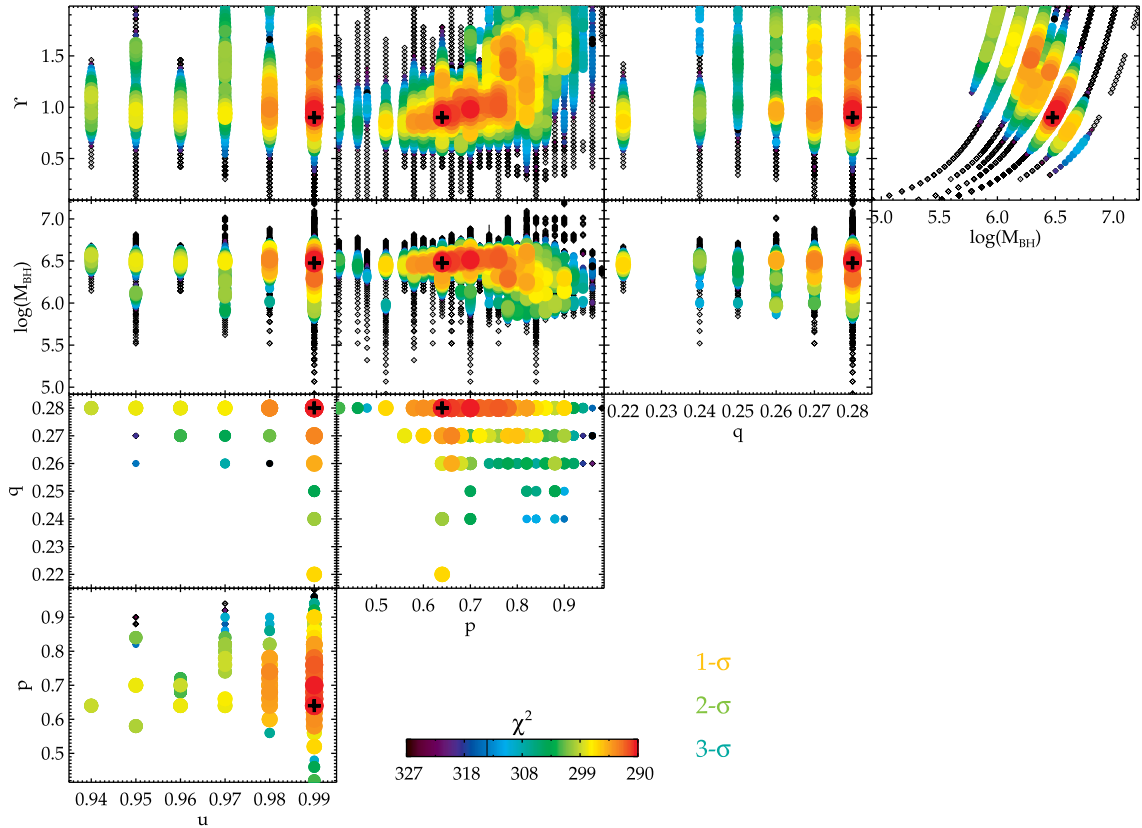


Figure 5. Illustration of the fitted parameter space, χ^2 was computed with equation (1). Each symbol denotes a model, the coloured symbols are models with $\Delta\chi^2 < \sigma_{\chi^2} = 37.3$, black diamonds are models with $\Delta\chi^2 > \sigma_{\chi^2}$. The 1σ , 2σ and 3σ colours corresponding to $\Delta\chi^2 = 5.9$, 11.3 and 18.2 , are denoted. The black cross denotes the best-fitting model.

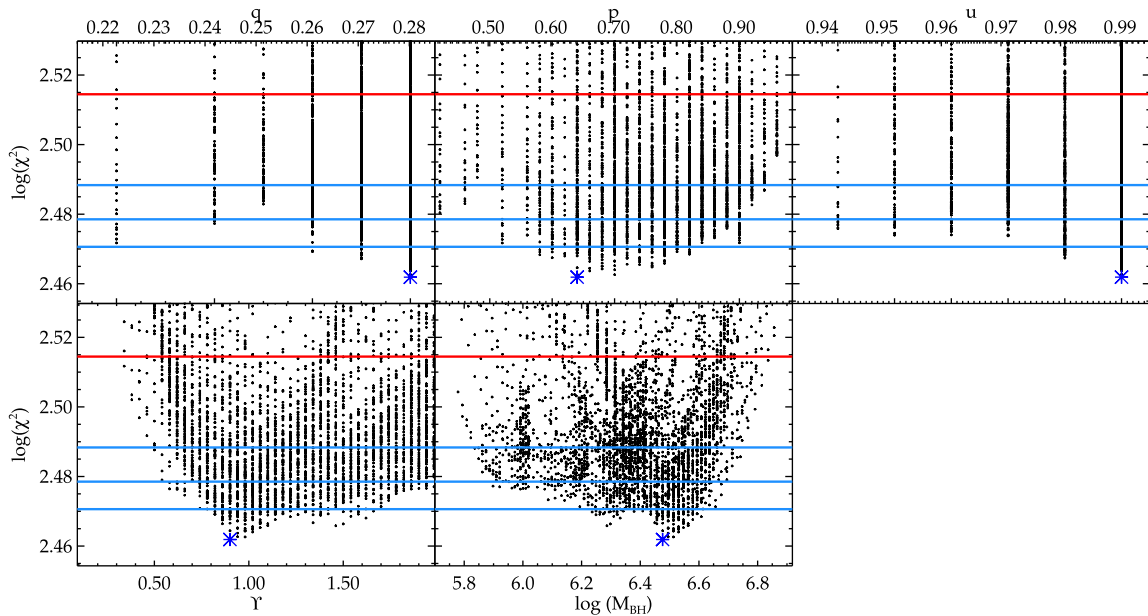


Figure 6. The χ^2 values plotted against the five free parameters (M_{\bullet} , Υ , q , p , u). χ^2 was computed with equation (1). The best-fitting model is denoted as blue asterisk, the 1σ , 2σ and 3σ confidence levels, corresponding to $\Delta\chi^2 = 5.9$, 11.3 and 18.2 , are denoted as blue lines. The red line denotes $\sigma_{\chi^2}^2 = 37.3$.

The shape of the nuclear star cluster is illustrated in Fig. 8. We show the intrinsic axial ratios q and p as a function of radius r . The axial ratio $q = c/a$ is low in the centre ($q = 0.3$), increases to $q = 0.8$ at $r \approx 35$ arcsec, and then decreases to $q = 0.6$ at $r = 150$ arcsec.

The best-fitting shape parameter, $q = 0.28$, is approximately the central value. The axial ratio $p = b/a$ is also low in the centre (~ 0.65), increases to $p = 0.9$ at $r = 40$ arcsec, and then decreases to $p = 0.75$ at $r = 150$ arcsec. Though, the uncertainty of p is rather

Table 2. The best-fitting model results and the 1σ and 3σ uncertainties, corresponding to $\Delta\chi^2 = 5.9$ and 18.2 .

Parameter	Best fit	1σ	3σ	Unit
M_\bullet	3.0	+1.1 -1.3	+2.4 -2.3	$\times 10^6 M_\odot$
Υ	0.90	+0.76 -0.08	+1.12 -0.32	$M_\odot/L_{\odot,4.5\mu\text{m}}$
q	0.28	+0.0 -0.02	+0.0 -0.06	—
p	0.64	+0.18 -0.06	+0.30 -0.22	—
u	0.99	+0.0 -0.01	+0.0 -0.05	—

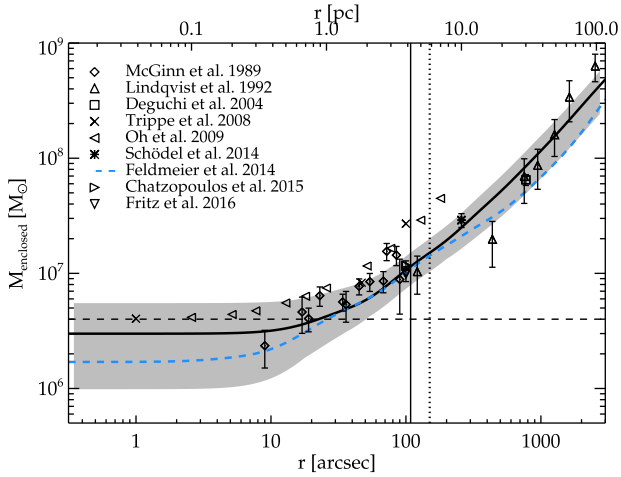


Figure 7. Enclosed total mass in spherical shells as a function of radius, in units of M_\odot and in logarithmic scaling. The black line denotes the enclosed mass with $\Upsilon = 0.90$ and $M_\bullet = 3.0 \times 10^6 M_\odot$, the grey shaded contours the 3σ uncertainty. The horizontal line denotes a supermassive black hole with the mass $M_\bullet = 4 \times 10^6 M_\odot$. The vertical, dotted line denotes the outer edge of the kinematic data, the vertical, solid line the effective radius. We also plot the results for the enclosed mass from previous studies. We scaled the masses to a distance $R_0 = 8.0$ kpc, if the study assumed a different Galactocentric distance: McGinn et al. (1989, diamonds, assumed $R_0 = 8.5$ kpc), Lindqvist, Häbing & Winnberg (1992, upward triangles, $R_0 = 8.5$ kpc), Deguchi et al. (2004, squares), Trippe et al. (2008, x-symbol), Oh et al. (2009, leftfacing triangles), Schödel et al. (2014, asterisk), Feldmeier et al. (2014, blue dashed line), Chatzopoulos et al. (2015a, rightfacing triangle, $R_0 = 8.3$ kpc) and Fritz et al. (2016, downward triangle, $R_0 = 8.2$ kpc).

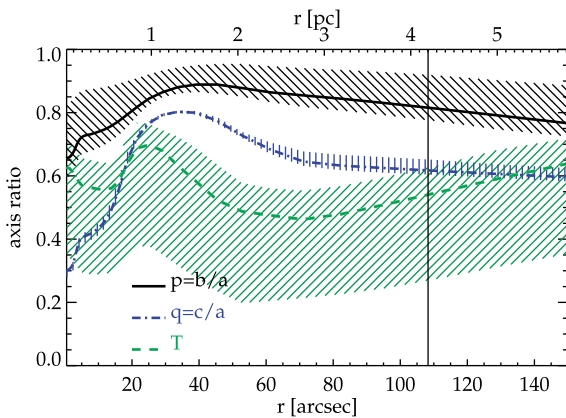


Figure 8. Intrinsic axial ratios as a function of radius r . The blue dot-dashed line denotes $q = c/a$, the black solid line $p = b/a$, the green dashed line the triaxiality $T = (1 - p^2)/(1 - q^2)$, the shaded regions the 1σ uncertainties, the vertical, solid line $r_{\text{eff}} = 4.2$ pc.

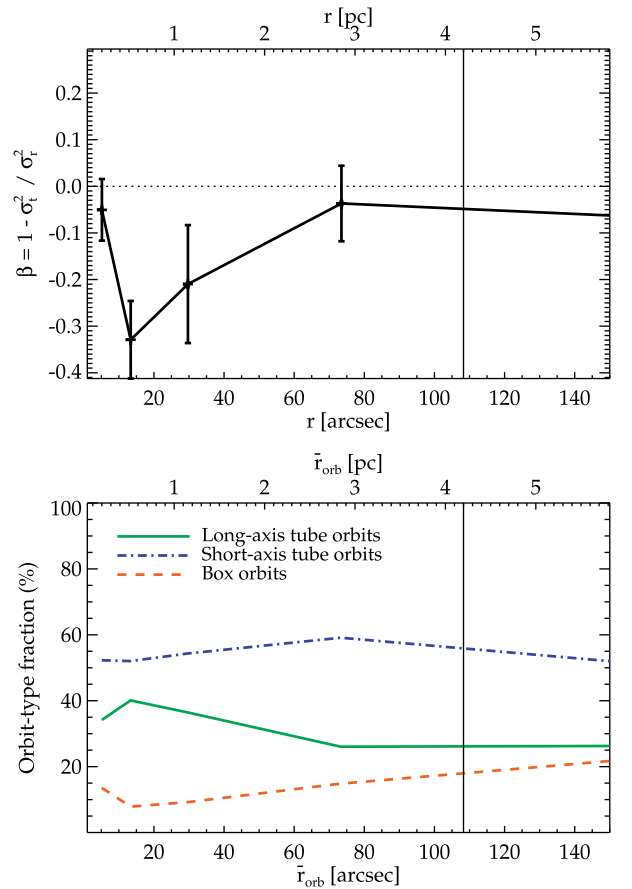


Figure 9. Top panel: anisotropy β as a function of radius r . Negative values denote tangential anisotropy, positive values radial anisotropy. Bottom panel: fraction of stellar mass per orbit type as a function of average orbit radius \bar{r}_{orb} . The green, solid line denotes long-axis tube orbits; the blue, dot-dashed line short-axis tube orbits; the red, dashed line box orbits; the vertical, solid line $r_{\text{eff}} = 4.2$ pc.

high, and the decrease not significant. Also the best-fitting axial ratio, $p = 0.64$, is close to the central value. We plot the triaxiality $T = (1 - p^2)/(1 - q^2)$, it varies between 0.45 and 0.7. The triaxiality increases from $r = 70$ arcsec to the outer radius $r = 150$ arcsec. This may be a signature of the increasing influence of the triaxial Galactic bulge at larger radii (Tsatsi et al. 2017). However, given the large uncertainty of T , this increase is not significant. A constant T or a decreasing T are not excluded by the large uncertainties.

3.3.3 Internal dynamics

The best-fitting model has tangential anisotropy in the centre of the cluster. The value of the anisotropy $\beta = 1 - \sigma_t^2/\sigma_r^2$ is negative, where σ_t is the tangential velocity dispersion and σ_r is the radial velocity dispersion. We show the anisotropy β as a function of radius in Fig. 9, top panel. We plot the mean anisotropy of the models within the 1σ uncertainty limit. The uncertainty of β is about 0.1. The plot extends to the outer edge of the kinematic data at 150 arcsec. The cluster kinematics becomes nearly isotropic at radii $r > 70$ arcsec.

We show the angular momentum distribution of the orbits in Fig. 10. The colours denote the density of orbits passing radius r with mean angular momentum $\langle \lambda_z \rangle$ (top panel) or $\langle \lambda_x \rangle$ (bottom panel). The plot of $\langle \lambda_z \rangle$ denotes rotation about the short z -axis. Orbits with

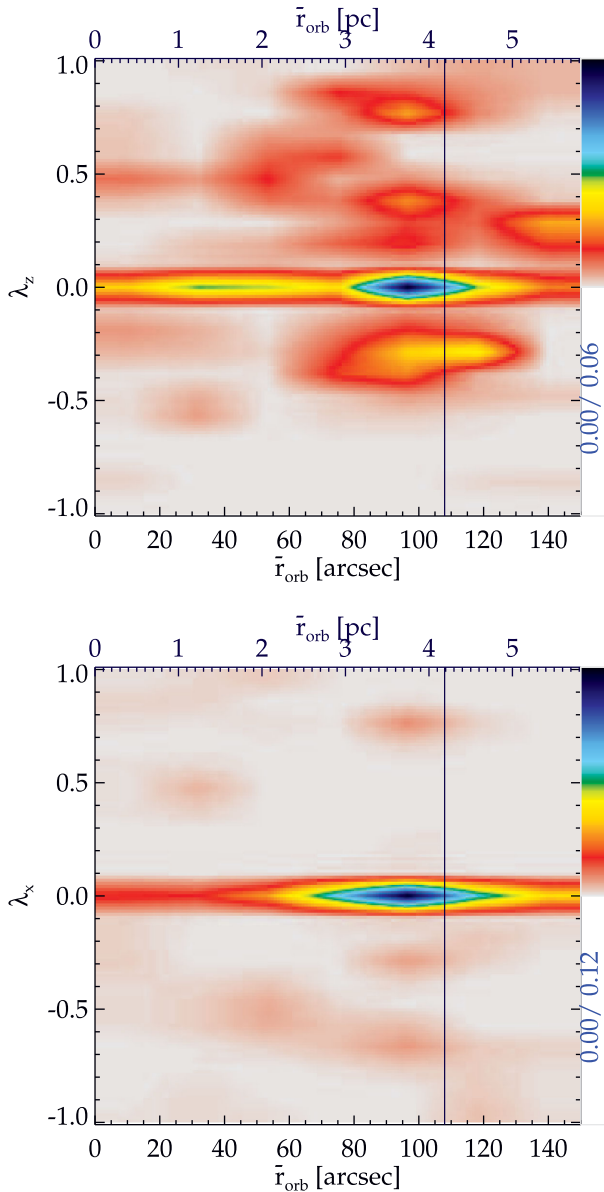


Figure 10. Orbit density with angular momentum λ_z (top panel), i.e. rotation around the short axis, and λ_x (bottom panel), i.e. rotation around the long axis as a function of average orbit radius \bar{r}_{orb} . Dark, blue colour indicates higher orbit density. The vertical line denotes $r_{\text{eff}} = 4.2$ pc.

$\langle \lambda_z \rangle \neq 0$ are contributed by short-axis tube orbits, while long-axis tube orbits have $\langle \lambda_z \rangle = 0$. On the other hand, $\langle \lambda_x \rangle$ denotes rotation about the long x -axis (bottom panel), and orbits with $\langle \lambda_x \rangle \neq 0$ are contributed by long-axis tube orbits. Short-axis tube orbits have $\langle \lambda_x \rangle = 0$. Long-axis tube orbits are most important in the central 20–60 arcsec and at larger radii $r \gtrsim 80$ arcsec. Short-axis tube orbits, which contribute in total more mass than long-axis tube orbits, are most important at $r = 60$ –140 arcsec. We illustrate the distribution of the stellar mass on the different orbit types also in Fig. 9 (bottom panel) as a function of radius. Most stars (>50 per cent) are on short-axis tube orbits, i.e. they orbit the minor axis. Long-axis tube orbits contribute about 40 per cent to the luminosity and thus stellar mass in the central 30 arcsec. They produce the perpendicular rotating substructure at $r \approx 20$ arcsec (~ 0.8 pc) found by Feldmeier et al.

(2014). At larger radii, long-axis tube orbits contribute only about 30 per cent to the stellar mass. Box orbits contribute little mass in the centre (<10 per cent), but their relative mass increases towards larger radii. At $r = 150$ arcsec (~ 5.8 pc), they contribute 20 per cent.

4 DISCUSSION

4.1 Difference of the resulting black hole mass

The currently best black hole mass estimate is $(4.02 \pm 0.20) \times 10^6 M_{\odot}$, derived by Boehle et al. (2016) from Keplerian stellar orbits around the supermassive black hole. Using axisymmetric Jeans models and the same spectroscopic data as our study, Feldmeier et al. (2014) found a lower value of $M_{\bullet} = (1.7_{-1.1}^{+1.4}) \times 10^6 M_{\odot}$. Our best fit using triaxial Schwarzschild models is $(3.0_{-1.3}^{+1.1}) \times 10^6 M_{\odot}$. This measurement is consistent with the direct measurements of Boehle et al. (2016) within the 1σ uncertainty limit. The result is also in agreement with the lower black hole mass of Feldmeier et al. (2014). We derived a 3σ lower limit for the black hole of $0.7 \times 10^6 M_{\odot}$, and an upper limit of $5.4 \times 10^6 M_{\odot}$. We briefly discuss the model degeneracies, possible reasons for the different black hole mass measurements and why our results are closer to the direct measurement than the black hole mass derived by Feldmeier et al. (2014).

4.1.1 Model degeneracies

Some model parameters seem to be correlated. This becomes clear when looking at Fig. 5. The best-fitting value of p apparently increases with increasing dynamical mass-to-light ratio Υ (second column of the first row). However, the value of p has little effect on M_{\bullet} , as can be seen in the second column of the second row in Fig. 5. With a higher value of p , the best-fitting M_{\bullet} increases slightly. At larger p , the χ^2 -contours of M_{\bullet} and Υ broaden. This means that for a more oblate axisymmetric cluster with p closer to one, M_{\bullet} and Υ are not as well constrained as with smaller values of p . The black hole mass $M_{\bullet} = 4.1 \times 10^6 M_{\odot}$ can be obtained with a higher value of $p = 0.82$ and $\Upsilon = 1.22$, this model is within the 1σ uncertainties.

The dynamical mass-to-light ratio Υ is inversely correlated with the black hole mass (fourth column of the first row in Fig. 5). The higher Υ , i.e. the more massive the cluster, the less massive is the black hole. This degeneracy is often obtained in dynamical models. Valluri, Merritt & Emsellem (2004) found that the degeneracy of M_{\bullet} depends on how well the black hole’s sphere of influence is resolved. The measurement of Υ is better constrained when the data extend to larger radii, provided that Υ is constant over the entire field. We have several kinematic data bins within the radius of influence of the supermassive black hole, and our data extend to one effective radius. This may not be sufficient to put strong constraints on Υ . To get agreement with the measurement of $(4.02 \pm 0.20) \times 10^6 M_{\odot}$ (Boehle et al. 2016), we would require a lower value of $\Upsilon \approx 0.75$, which is within the 1σ uncertainties.

4.1.2 Influence of the surface brightness profile

The shape of the surface brightness profile is important to estimate the mass of the supermassive black hole. The surface brightness profile has to represent the density of the kinematic tracer. We excluded young stars and ionized gas from the surface brightness profile, as these components contribute little mass compared to the cool, old stars we used as kinematic tracers. Excluding these components results in a lower surface brightness and stellar mass in

the centre compared to Feldmeier et al. (2014). The stellar mass we obtain at $r = 48$ arcsec (~ 1.8 pc) is $1.3 \times 10^6 M_{\odot}$ less. Our black hole mass is therefore higher, and closer to the direct measurement of $M_{\bullet} \approx 4 \times 10^6 M_{\odot}$. We ran the same axisymmetric models (Cappellari 2008), using the same kinematic data as Feldmeier et al. (2014), but our surface brightness distribution from Table 1. The best fit is obtained with $M_{\bullet} = (2.8^{+1.3}_{-0.8}) \times 10^6 M_{\odot}$, $\Upsilon = 0.89^{+0.12}_{-0.19}$ and a constant anisotropy of $\beta = -0.3$. This result is in agreement with the triaxial Schwarzschild models, and confirms that the surface brightness profile has a strong influence on the results of the black hole mass and dynamical mass-to-light ratio.

The deprojection of the surface brightness profile is non-unique and only one possible solution. We applied the MGE method, which produces smooth intrinsic densities (Cappellari 2002). We assumed that the deprojected density is smooth, as this is what we observe in other galaxies, where the clusters are observed from various viewing angles.

van den Bosch (1997) studied how much central density can be hidden in an oblate axisymmetric galaxy without effects on the projected surface brightness. They found that the percentage of hidden density in the centre of a Staekel potential is 0 per cent for inclination $i = 90^{\circ}$, and $\lesssim 10$ per cent for $i \gtrsim 80^{\circ}$ of the total galaxy mass. The percentage of hidden density can increase in steeper cusps. However, the effect on the dynamics is still negligible. For our models, a hidden density is equivalent to a spatially varying mass-to-light ratio.

4.1.3 Spatially varying mass-to-light ratio

We assumed a constant dynamical mass-to-light ratio Υ for the Schwarzschild models. We obtained $\Upsilon = 0.90^{+0.76}_{-0.08}$ (1σ uncertainty). The dynamical mass-to-light ratio combines the stellar mass-to-light ratio with other components, it is sensitive to the presence of gas or dark matter.

Our best-fitting value of $\Upsilon = 0.9$ is consistent with stellar-population studies. Norris et al. (2014) found $\Upsilon = 0.8\text{--}1.2$ at $4.6 \mu\text{m}$ for stellar populations with ages >7.0 Gyr and metallicities $[\text{Fe}/\text{H}]$ ranging from -1.0 to $+0.3$ dex. For populations with lower metallicity $[\text{Fe}/\text{H}] = -2.18$ dex and 13 Gyr age, Υ can be even 1.5. However, for younger stellar populations $\lesssim 5$ Gyr, Norris et al. (2014) obtained lower values $\Upsilon \approx 0.6$. Our measurement of Υ is averaged over the entire field of the kinematic data. We cannot exclude that the stellar age or metallicity changes over the range of the kinematic data. Stellar population studies of the red giant population were so far confined to the central 1–2 pc (e.g. Pfuhl et al. 2011; Do et al. 2015; Feldmeier-Krause et al. 2017). The stars in this region are mostly older than 5 Gyr and rather metal-rich. Our knowledge of the stellar population at the outer region of our field is based on only a few bright stars (e.g. Blum et al. 2003; Feldmeier et al. 2014). But these stars are brighter and probably younger than our kinematic tracer stars. However, the mass-to-light ratio for old stars in the mid-infrared varies modestly with age and metallicity in comparison to the optical mass-to-light ratio (Meidt et al. 2014). Therefore we do not expect a change of Υ by more than ~ 0.4 within the cluster. Should Υ vary with radius, our mass profile (Fig. 7) would have a different shape. For example, if Υ was lower in the centre than outside, this would increase M_{\bullet} , and there would be less mass in the stellar component.

However, the stellar mass-to-light ratio may also increase towards the central $r = 0.5$ pc, as massive stellar remnants may migrate to the centre. The mass and distribution of dark stellar remnants, i.e. stellar-mass black holes and neutron stars, in the central par-

sec of the nuclear star cluster is uncertain. For a top-heavy initial mass function, there could be $>1 \times 10^6 M_{\odot}$ in dark remnants (Morris 1993), though Löckmann, Baumgardt & Kroupa (2010) found a lower mass of about $1 \times 10^5 M_{\odot}$ for a canonical initial mass function.

In our models, we neglected the mass of molecular gas in the circumnuclear disc. The molecular gas may contribute $10^4\text{--}10^6 M_{\odot}$. The gas disc extends from $r \approx 1\text{--}7$ pc along the Galactic plane, but only to $r \approx 3$ pc along the minor axis (Ferrière 2012). Thus, the molecular gas is located in the central part of our spectroscopic field, but absent in the north. If the gas contributes significantly to the cluster mass, our assumption of spatially constant Υ would be violated, and Υ would be higher than for a stellar component alone. When we assume the maximum gas mass of $10^6 M_{\odot}$, the value of a constant Υ decreases to about 0.85, which is within our 1σ uncertainty limit.

The spatial distribution of dark matter in the Galactic Centre is uncertain. A classical cuspy Navarro et al. (1996) dark matter profile results in a dark matter fraction of about 6.6 per cent in the central 100 pc (Linden 2014). However, black hole accretion, dark matter annihilation and scattering alter the shape of the dark matter distribution in the Galactic Centre. Vasiliev & Zelnikov (2008) found that these effects produce a shallower dark matter profile in the central 2 pc than further out. The dark matter mass inferred from the classical cusp is reduced by up to 50 per cent in the central 2 pc. The contribution of dark matter to the nuclear star cluster mass should therefore be negligible. Although the dark matter distribution may be different from the luminous baryonic matter, and the dynamical mass-to-light ratio for that reason not spatially constant, the effect on the cluster mass distribution should be only minor.

4.2 Triaxial cluster shape

Our best-fitting model has axial ratios of $q = c/a = 0.28^{+0.0}_{-0.02}$, $p = b/a = 0.64^{+0.18}_{-0.06}$ and $u = a'/a = 0.99^{+0.0}_{-0.01}$. These axial ratios correspond to viewing angles $\vartheta = 80^{\circ}$, $\varphi = 79^{\circ}$ and $\psi = 91^{\circ}$. The angle ϑ denotes the polar viewing angle, φ the azimuthal viewing angle and ψ is the misalignment angle between photometric major axis and the projected intrinsic long axis (van den Bosch et al. 2008; van den Bosch & van de Ven 2009). For the best-fitting model, the angle α between the cluster's major axis and the line of sight is about 79° . The cluster's shape is illustrated in Fig. 8, the triaxiality parameter $T = (1 - p^2)/(1 - q^2)$ varies between 0.45 and 0.7, with an average at $T \sim 0.6$. An oblate axisymmetric system has $T = 0$, a prolate axisymmetric system has $T = 1$.

Also the Milky Way's bulge is triaxial, the axial ratios are $q = 0.26$ and $p = 0.63$ (Wegg & Gerhard 2013). The shape was derived from the density of red clump stars in the central 800 pc of the bulge. The Milky Way bulge is much larger than the nuclear star cluster, and extends out to about 2.5 kpc. At the outer edge of our data, at $r = 150$ arcsec, we obtain for the intrinsic shape parameters $q = 0.60$ and $p = 0.75$. Both q and p are higher than found in the bulge. However, they are also decreasing, though the decrease of p is not significant. The bulge has a peanut or X-shape (Nataf et al. 2010; McWilliam & Zoccali 2010). The angle α between the bulge major axis and the line of sight to the Galactic Centre is about 27° (Rattenbury et al. 2007; Wegg & Gerhard 2013), while we obtained 79° for the nuclear star cluster. There are also indications for another bar within the Galactic bulge from star count data in the inner $|l| \lesssim 1^{\circ} \approx 140$ pc (Alard 2001) or even $|l| \lesssim 4^{\circ} \approx 560$ pc (Nishiyama et al. 2005). Rodriguez-Fernandez & Combes (2008)

derived an angle $\alpha \approx 60^\circ\text{--}75^\circ$ for a thick triaxial nuclear bar with axial ratios $q \approx 0.55$ and $p \approx 0.75$.

One possible scenario for nuclear star cluster formation is that massive star clusters ($10^5\text{--}10^7 M_\odot$) formed in the galactic disc, migrated to the galaxy's centre and merged (Neumayer et al. 2011; Guillard, Emsellem & Renaud 2016). Simulations of multiple star cluster mergers and of star cluster accretion on a nuclear stellar component can produce triaxial nuclear star clusters (Bekki et al. 2004; Hartmann et al. 2011; Perets & Mastrobuono-Battisti 2014). However, so far no observational study was able to constrain the triaxial shape of nuclear star clusters in general. Hartmann et al. (2011) constrained the shape of two nuclear star clusters and found agreement with an axisymmetric shape. For the Milky Way nuclear star cluster, p may be as high as 0.94 within the 3σ uncertainties. Thus, a nearly axisymmetric shape is consistent with our data.

4.3 Caveats and considerations

4.3.1 Regime of semiresolved populations

We used integrated light spectroscopy to measure the stellar kinematics. This is the common approach for extragalactic systems, which have a distance of several megaparsec. The measured kinematics are weighted by the respective luminosities of different stars. As the Milky Way nuclear star cluster is only 8 kpc distant, we are in the regime of semiresolved populations. The brightest stars can be resolved individually, and these stars contribute a large fraction of the flux. In consequence, individual spatial bins can be dominated by a single star. Instead of measuring the spectrum of an ensemble of stars, one measures a spectrum in which a large percentage of the flux is contributed by one single star. This causes shot noise and high differences between neighbouring spatial bins. We accounted for this problem by excluding the brightest stars from the spectroscopic map. This method helps to significantly reduce the intrinsic scatter of the velocity dispersion (Lützgendorf et al. 2011; Bianchini et al. 2015). We further increased the kinematic uncertainties such that the data in two neighbouring bins have consistent values within their uncertainties. This helps to prevent that the models fit only stochastic shot noise. Due to the large kinematic uncertainties, the intrinsic shape parameters q , p , u and the dynamical mass-to-light ratio Υ are not very well constrained, and have large error margins.

At a distance of only 8 kpc, also the relative distances of the stars become more important. A star located on the near side of the nuclear star cluster, at a distance $d = 7.9$ kpc, contributes 1.05 times more flux than a star with the same absolute magnitude at the far side of the cluster, at $d = 8.1$ kpc. In an extragalactic system, the distance of a star at the near side and the distance of a star at the far side with respect to the observer are approximately the same, as the system is farther away. For a galaxy at $d = 5$ Mpc, a relative difference of 200 pc changes the flux only by a factor of 1.000 08. Even foreground stars that belong to the outer parts of the stellar system contribute roughly the same flux as a star with the same magnitude that is located in the galactic nucleus.

4.3.2 Interstellar extinction

Another observational complication is interstellar extinction in the Galactic Centre, which varies on arcsecond scales (Schödel et al. 2010). In particular, the field of view of the kinematic data contains the so-called 20-km s⁻¹ cloud (M-0.13-0.08, e.g. García-Marín et al. 2011) in the Galactic southwest. It lies at a projected distance of about 70 arcsec (~ 3 pc) from the centre, and probably about 5 pc in front of Sgr A* (Ferrière 2012). This cloud blocks

the light from stars of the nuclear star cluster. We cannot access the kinematics of stars behind this cloud. There is also interstellar dust within a projected distance of 20 arcsec (~ 0.8 pc) from the centre, i.e. within the radius of influence of the black hole. This dust causes extinction within the nuclear star cluster by up to 0.8 mag (Chatzopoulos et al. 2015b). As a consequence, the two effects of dimming by distance and by extinction add up and stars that lie on the far side of the nuclear star cluster appear even more faint than the stars on the near side.

4.3.3 Implications

Both the semiresolved stellar population and the inter-cluster extinction cause that our observations are biased to the near side of the nuclear star cluster. As a consequence, we measured a lower limit of the velocity dispersion. Feldmeier et al. (2014) found that the velocity dispersion in the projected radial range $6 \text{ arcsec} < r < 20 \text{ arcsec}$ is smaller compared to the velocity dispersion computed from proper motion data of Schödel et al. (2009), which is based on resolved stars. For resolved stars, the velocity dispersion is not weighted by the flux of the stars. An underestimated velocity dispersion means that the black hole mass measurement is biased to lower values.

This observational bias also influences the measurements of v , h_3 and h_4 . In particular, the cluster may appear compressed along the line of sight, and thus the value of $p = b/a = 0.64_{-0.06}^{+0.18}$ may be too low. As a consequence, $\Upsilon = 0.90_{-0.08}^{+0.76}$ would be underestimated (see second column of the first row in Fig. 5).

4.3.4 Influence of figure rotation

The Galaxy rotates, and with it the nuclear star cluster. In a non-axisymmetric, rotating system, centrifugal and Coriolis forces play a role. However, figure rotation and the resulting forces were not included in our triaxial models. Figure rotation influences the stellar orbits. The prograde and retrograde tube orbits no longer fill the same volumes, while the box orbits acquire net mean angular momentum (e.g. Heisler, Merritt & Schwarzschild 1982; Schwarzschild 1982; Sellwood & Wilkinson 1993; Skokos, Patsis & Athanassoula 2002). As a result, orbit-based, tumbling, triaxial models are computationally expensive. For that reason, only few such models were constructed that take kinematic data into account (Zhao 1996; Häfner et al. 2000; Wang et al. 2012, 2013). It is difficult to predict how our results would change in a rotating model. The inferred orbital structure will be affected (depending on the tumbling speed of the nuclear star cluster), but our results on the mass distribution are likely to be fairly robust, as the assumption of a constant mass-to-light ratio is probably more important.

5 SUMMARY AND OUTLOOK

We constructed for the first time triaxial orbit-based Schwarzschild models of the Milky Way nuclear star cluster. We used the spectroscopic integrated light maps of Feldmeier et al. (2014) to measure the cluster kinematics of the central 60 pc^2 of the Milky Way. As photometry we used *Spitzer* 4.5 μm and NACO *H*-band images, and measured a two-dimensional surface brightness distribution. We excluded young stars, avoided gas emission and dark clouds in the photometric data. Our triaxial models were based on the code by van den Bosch et al. (2008). Our best-fitting model contains a black hole of mass $M_\bullet = (3.0_{-1.3}^{+1.1}) \times 10^6 M_\odot$, a dynamical mass-to-light ratio of $\Upsilon = (0.90_{-0.08}^{+0.76}) M_\odot/L_{\odot,4.5 \mu\text{m}}$, and shape parameters $q = 0.28_{-0.02}^{+0.0}$, $p = 0.64_{-0.06}^{+0.18}$, and $u = 0.99_{-0.01}^{+0.0}$. Our black hole

mass measurement is in agreement with the direct measurement of $(4.02 \pm 0.20) \times 10^6 M_{\odot}$ by Boehle et al. (2016). We obtain a total cluster mass $M_{\text{MWNsc}} = (2.1 \pm 0.7) \times 10^7 M_{\odot}$ within a spherical shell with radius $r = 2 \times r_{\text{eff}} = 8.4$ pc. The best-fitting model is tangentially anisotropic in the central $r = 0.5$ – 2 pc of the nuclear star cluster, but close to isotropic at larger radii. The model is able to recover the long-axis rotation in the central $r = 0.8$ pc found by Feldmeier et al. (2014), and the misalignment of the kinematic rotation axis from the photometric minor axis.

There are several possible ways to extend the dynamical models in the future. One way is to include a component for the neutral gas disc inside the nuclear star cluster. If the gas mass is close to the upper limit of $10^6 M_{\odot}$, the dynamical mass-to-light ratio would probably decrease slightly, and in return would slightly increase the black hole mass. Modelling a spatially varying mass-to-light ratio may provide a better representation of the cluster's intrinsic properties. Further, proper motions can be included in combination with discrete line-of-sight velocities, as shown by van de Ven et al. (2006) and van den Bosch et al. (2006) for axisymmetric Schwarzschild models. Watkins et al. (2013) extended axisymmetric Jeans models and implemented discrete kinematic data without binning. Using discrete data means that the stars are not weighted by their luminosities. This prevents the previously discussed bias towards the near side of the cluster.

ACKNOWLEDGEMENTS

This paper is based on observations collected at the European Organization for Astronomical Research in the Southern hemisphere, Chile (289.B-5010(A)). GvdV acknowledges partial support from Sonderforschungsbereich SFB 881 ‘The Milky Way System’ (sub-project A7 and A8) funded by the German Research Foundation. RS acknowledges funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n. 614922. We thank Remco van den Bosch for helpful discussions about the project. We finally thank the anonymous reviewer for useful comments and suggestions.

REFERENCES

Alard C., 2001, *A&A*, 379, L44
 Alexander T., 2005, *Phys. Rep.*, 419, 65
 Bahcall J. N., Tremaine S., 1981, *ApJ*, 244, 805
 Bekki K., Couch W. J., Drinkwater M. J., Shioya Y., 2004, *ApJ*, 610, L13
 Bianchini P., Norris M. A., van de Ven G., Schinnerer E., 2015, *MNRAS*, 453, 365
 Blum R. D., Ramírez S. V., Sellgren K., Olsen K., 2003, *ApJ*, 597, 323
 Boehle A. et al., 2016, *ApJ*, 830, 17
 Cappellari M., 2002, *MNRAS*, 333, 400
 Cappellari M., 2008, *MNRAS*, 390, 71
 Cappellari M., Emsellem E., 2004, *PASP*, 116, 138
 Chatzopoulos S., Fritz T. K., Gerhard O., Gillessen S., Wegg C., Genzel R., Pfuhl O., 2015a, *MNRAS*, 447, 948
 Chatzopoulos S., Gerhard O., Fritz T. K., Wegg C., Gillessen S., Pfuhl O., Eisenhauer F., 2015b, *MNRAS*, 453, 939
 Christopher M. H., Scoville N. Z., Stolovy S. R., Yun M. S., 2005, *ApJ*, 622, 346
 Cretton N., de Zeeuw P. T., van der Marel R. P., Rix H.-W., 1999, *ApJS*, 124, 383
 Deguchi S. et al., 2004, *PASJ*, 56, 261
 Do T. et al., 2013, *ApJ*, 779, L6
 Do T., Kerzendorf W., Winsor N., Støstad M., Morris M. R., Lu J. R., Ghez A. M., 2015, *ApJ*, 809, 143

Drehmer D. A., Storch-Bergmann T., Ferrari F., Cappellari M., Riffel R. A., 2015, *MNRAS*, 450, 128
 Emsellem E., Monnet G., Bacon R., 1994, *A&A*, 285, 723
 Etxaluze M., Smith H. A., Tolls V., Stark A. A., González-Alfonso E., 2011, *AJ*, 142, 134
 Feldmeier A. et al., 2014, *A&A*, 570, A2
 Feldmeier-Krause A. et al., 2015, *A&A*, 584, A2
 Feldmeier-Krause A., Kerzendorf W., Neumayer N., Schödel R., Nogueras-Lara F., Do T., de Zeeuw P. T., Kuntschner H., 2017, *MNRAS*, 464, 194
 Ferrière K., 2012, *A&A*, 540, A50
 Figer D. F., McLean I. S., Morris M., 1999, *ApJ*, 514, 202
 Franx M., 1988, *MNRAS*, 231, 285
 Fritz T. K. et al., 2016, *ApJ*, 821, 44
 Gao J., Li A., Jiang B. W., 2013, *Earth Planets Space*, 65, 1127
 García-Marín M. et al., 2011, *ApJ*, 738, 158
 Gebhardt K. et al., 2000, *AJ*, 119, 1157
 Genzel R., Thatte N., Krabbe A., Kroker H., Tacconi-Garman L. E., 1996, *ApJ*, 472, 153
 Genzel R., Eisenhauer F., Gillessen S., 2010, *Rev. Mod. Phys.*, 82, 3121
 Ghez A. M. et al., 2008, *ApJ*, 689, 1044
 Gillessen S., Eisenhauer F., Trippe S., Alexander T., Genzel R., Martins F., Ott T., 2009, *ApJ*, 692, 1075
 Guillard N., Emsellem E., Renaud F., 2016, *MNRAS*, 461, 3620
 Häfner R., Evans N. W., Dehnen W., Binney J., 2000, *MNRAS*, 314, 433
 Haller J. W., Rieke M. J., Rieke G. H., Tamblyn P., Close L., Melia F., 1996, *ApJ*, 456, 194
 Hartmann M., Debattista V. P., Seth A., Cappellari M., Quinn T. R., 2011, *MNRAS*, 418, 2697
 Heisler J., Merritt D., Schwarzschild M., 1982, *ApJ*, 258, 490
 Jeans J. H., 1922, *MNRAS*, 82, 122
 Lawson C. L., Hanson R. J., 1974, *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs
 Linden T., 2014, in de Grijs R., ed., *Proc. IAU Symp. 303, The Galactic Center: Feeding and Feedback in a Normal Galactic Nucleus*. Kluwer, Dordrecht, p. 403
 Lindqvist M., Habing H. J., Winnberg A., 1992, *A&A*, 259, 118
 Löckmann U., Baumgardt H., Kroupa P., 2010, *MNRAS*, 402, 519
 Lützgendorf N., Kissler-Patig M., Noyola E., Jalali B., de Zeeuw P. T., Gebhardt K., Baumgardt H., 2011, *A&A*, 533, A36
 McGinn M. T., Sellgren K., Becklin E. E., Hall D. N. B., 1989, *ApJ*, 338, 824
 McWilliam A., Zoccali M., 2010, *ApJ*, 724, 1491
 Malkin Z., 2013, in de Grijs R., ed., *Proc. IAU Symp. 289, Statistical analysis of the determinations of the Sun’s Galactocentric distance*. Astron. Soc. Pac., San Francisco, p. 406
 Meidt S. E. et al., 2014, *ApJ*, 788, 144
 Merritt D., 2004, *Coevolution of Black Holes and Galaxies*. Cambridge Univ. Press, Cambridge, p. 263
 Morris M., 1993, *ApJ*, 408, 496
 Nataf D. M., Udalski A., Gould A., Fouqué P., Stanek K. Z., 2010, *ApJ*, 721, L28
 Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
 Neumayer N., Walcher C. J., Andersen D., Sánchez S. F., Böker T., Rix H.-W., 2011, *MNRAS*, 413, 1875
 Nishiyama S. et al., 2005, *ApJ*, 621, L105
 Norris M. A., Meidt S., Van de Ven G., Schinnerer E., Groves B., Querejeta M., 2014, *ApJ*, 797, 55
 Oh S., Kim S. S., Figer D. F., 2009, *J. Korean Astron. Soc.*, 42, 17
 Oort J. H., 1977, *ARA&A*, 15, 295
 Paumard T., Maillard J.-P., Morris M., 2004, *A&A*, 426, 81
 Paumard T. et al., 2006, *ApJ*, 643, 1011
 Perets H. B., Mastrobuono-Battisti A., 2014, *ApJ*, 784, L44
 Pfuhl O. et al., 2011, *ApJ*, 741, 108
 Rattenbury N. J., Mao S., Sumi T., Smith M. C., 2007, *MNRAS*, 378, 1064
 Reid M. J., Brunthaler A., 2004, *ApJ*, 616, 872
 Requena-Torres M. A. et al., 2012, *A&A*, 542, L21
 Rieke G. H., Rieke M. J., 1988, *ApJ*, 330, L33

- Rix H.-W., de Zeeuw P. T., Cretton N., van der Marel R. P., Carollo C. M., 1997, *ApJ*, 488, 702
- Rodriguez-Fernandez N. J., Combes F., 2008, *A&A*, 489, 115
- Rybicki G. B., 1987, in de Zeeuw P. T., ed., *Proc. IAU Symp. 127, Structure and Dynamics of Elliptical Galaxies*. Reidel, Dordrecht, p. 397
- Saito R. K. et al., 2012, *A&A*, 537, A107
- Schödel R., Merritt D., Eckart A., 2009, *A&A*, 502, 91
- Schödel R., Najarro F., Muzic K., Eckart A., 2010, *A&A*, 511, A18
- Schödel R., Feldmeier A., Kunneriath D., Stolovy S., Neumayer N., Amaro-Seoane P., Nishiyama S., 2014, *A&A*, 566, A47
- Schwarzschild M., 1982, *ApJ*, 263, 599
- Schwarzschild M., 1979, *ApJ*, 232, 236
- Scoville N. Z., Stolovy S. R., Rieke M., Christopher M., Yusef-Zadeh F., 2003, *ApJ*, 594, 294
- Sellgren K., McGinn M. T., Becklin E. E., Hall D. N., 1990, *ApJ*, 359, 112
- Sellwood J. A., Wilkinson A., 1993, *Rep. Prog. Phys.*, 56, 173
- Siopis C. et al., 2009, *ApJ*, 693, 946
- Skokos C., Patsis P. A., Athanassoula E., 2002, *MNRAS*, 333, 847
- Stolovy S. et al., 2006, *J. Phys. Conf. Ser.*, 54, 176
- Trippe S. et al., 2008, *A&A*, 492, 419
- Tsatsi A., Mastrobuono-Battisti A., van de Ven G., Perets H. B., Bianchini P., Neumayer N., 2017, *MNRAS*, 464, 3720
- Valluri M., Merritt D., Emsellem E., 2004, *ApJ*, 602, 66
- Valluri M., Ferrarese L., Merritt D., Joseph C. L., 2005, *ApJ*, 628, 137
- van de Ven G., van den Bosch R. C. E., Verolme E. K., de Zeeuw P. T., 2006, *A&A*, 445, 513
- van de Ven G., de Zeeuw P. T., van den Bosch R. C. E., 2008, *MNRAS*, 385, 614
- van den Bosch F. C., 1997, *MNRAS*, 287, 543
- van den Bosch R. C. E., de Zeeuw P. T., 2010, *MNRAS*, 401, 1770
- van den Bosch R. C. E., van de Ven G., 2009, *MNRAS*, 398, 1117
- van den Bosch R., de Zeeuw T., Gebhardt K., Noyola E., van de Ven G., 2006, *ApJ*, 641, 852
- van den Bosch R. C. E., van de Ven G., Verolme E. K., Cappellari M., de Zeeuw P. T., 2008, *MNRAS*, 385, 647
- van den Bosch R. C. E., Greene J. E., Braatz J. A., Constantin A., Kuo C.-Y., 2016, *ApJ*, 819, 11
- van der Marel R. P., Cretton N., de Zeeuw P. T., Rix H.-W., 1998, *ApJ*, 493, 613
- Vasiliev E., Zelnikov M., 2008, *Phys. Rev. D*, 78, 083506
- Wallace L., Hinkle K., 1996, *ApJS*, 107, 312
- Wang Y., Zhao H., Mao S., Rich R. M., 2012, *MNRAS*, 427, 1429
- Wang Y., Mao S., Long R. J., Shen J., 2013, *MNRAS*, 435, 3437
- Watkins L. L., van de Ven G., den Brok M., van den Bosch R. C. E., 2013, *MNRAS*, 436, 2598
- Wegg C., Gerhard O., 2013, *MNRAS*, 435, 1874
- Zhao H., 1996, *MNRAS*, 283, 149

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