Author: Krause-Feldmeier, A.
Title: The assembly history of the milky way nuclear star cluster
Issue Date: 2016-09-13

3

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

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. (2014). The models are used to constrain the supermassive black hole mass $M_\bullet$, dynamical mass-to-light ratio Y, and the intrinsic shape of the cluster. Our best-fitting model has $M_\bullet = (3.0^{+1.1}_{-1.3}) \times 10^6\ M_\odot$, $Y = (0.90^{+0.76}_{-0.08})\ M_\odot/L_{\odot,4.5\mu m}$, 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 on the high-end of 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 semi-resolved 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 $M_{\mathrm{MWNSC}} = (3.1^{+2.6}_{-0.3}) \times 10^7\ M_\odot$ within a radius of $r = 2 \times r_{\mathrm{eff}} = 8.4$ pc. The best-fitting model is tangentially anisotropic in the central $r = 2$ 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.

A. Feldmeier-Krause, L. Zhu, N. Neumayer, G. van de Ven, P. T. de Zeeuw, R. Schödel

submitted to MNRAS on July 22, 2016

3.1 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 it 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 ± 0.6) $\times 10^6\ M_\odot$ (Ghez et al. 2008), (4.3 ± 0.39) $\times 10^6\ M_\odot$ (Gillessen et al. 2009b), or (4.02 ± 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 observational 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; Haller et al. 1996; Genzel et al. 1996). Also stellar proper motions were used to study the Galactic centre mass distribution (Schödel et al. 2009). Several studies combined radial velocity and proper motion data (Trippe et al. 2008; Do et al. 2013b; 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 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. 2014a) 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. 2013b; Chatzopoulos et al. 2015a), although the uncertainties are quite large (e.g. Do et al. 2013b). 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 halos. 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 2012) and a position angle $31\overset{\circ}{.}40$ East of North (J2000.0 coordinates, Reid & Brunthaler 2004) with respect to the Galactic plane. This chapter is organised as follows: We describe the kinematic and photometric data in Section 3.2. The dynamical models are introduced in Section 3.3. Section 3.4 discusses the results, and Section 3.5 summarizes the main conclusions.

3.2 Description of the data

3.2.1 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 $2\overset{''}{.}22\cdot$pixel^{-1}. We used the data cube that was cleaned from 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 PPXF (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.

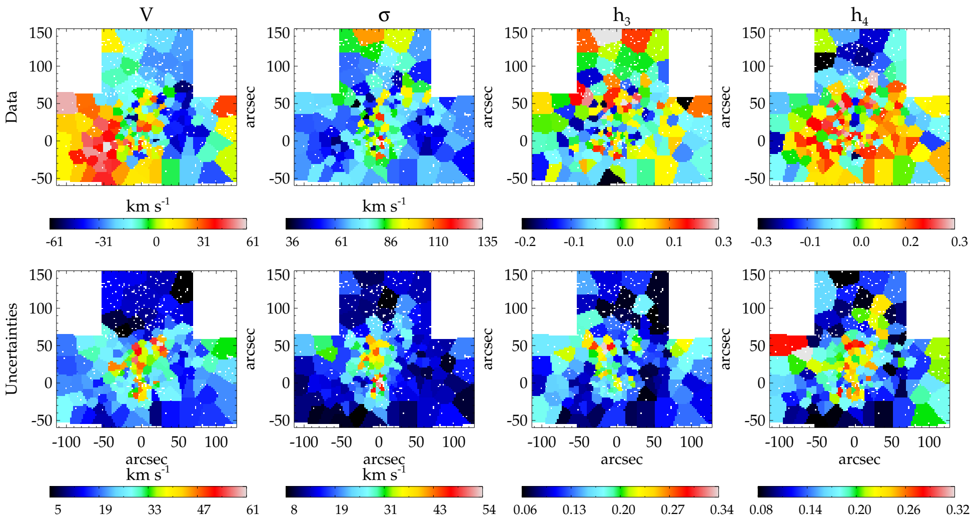


Figure 3.1: Kinematic data (top row) and respective uncertainties (bottom row). The columns denote velocity V, velocity dispersion σ, Gauss-Hermite moments h_3, and h_4. White pixels are due to excluded bright stars.

Since the Milky Way nuclear star cluster is at a distance of only 8 kpc, the spectroscopic observations are spatially semi-resolved. 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 $(\varepsilon_{V_i} + \varepsilon_{V_j})$. We do this for the uncertainty of velocity, velocity dispersion, h_3, and h_4, and find that it is required for about 68 per cent of the kinematic data uncertainties. Additionally, we point-symmetrise the kinematics using the procedure of van den Bosch & de Zeeuw (2010). The median uncertainties or V, σ, h_3, and h_4 are 24.6 km$\cdot$s^{-1}, 18.4 km$\cdot$s^{-1}, 0.15, and 0.17. 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$\cdot$s^{-1} and an increase in the velocity dispersion from about 65 km$\cdot$s^{-1} towards $\sigma_{\max}$=135 km$\cdot$s^{-1} at the centre. The kinematic maps are shown on the top row of Fig. 3.1, the uncertainties are shown on the bottom row.

3.2.2 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 et al. 2013), therefore we used near- and mid-infrared images.

For the central $40\farcs4 \times 40\farcs4$ (1.6 pc × 1.6 pc) we used the high-resolution NACO $H-$band mosaic of Schödel et al. (2009), which has a spatial scale of $0\farcs027 \cdot \mathrm{pixel}^{-1}$. 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 et al. 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 ($\sim 12.8''$, 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 PAH emission by Schödel et al. (2014a). We used the extinction and emission corrected 4.5 μm image to measure the light distribution. The image was smoothed to a scale of $5'' \cdot \mathrm{pixel}^{-1}$, and extends over $\sim$270 pc × 200 pc. We excluded a central circle with $r = 0.6$ pc ($\sim 15\farcs4$) to avoid contribution from ionised gas emission and young stars. In addition we masked out the young Quintuplet star cluster (Figer et al. 1999), and the dark 20-km$\cdot$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 model (Emsellem et al. 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. 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 ($16''-27\farcs8$). Then we measured the photometry on each image along 12 angular sectors, and converted the NACO photometry to the *Spitzer* flux. Assuming four-fold 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 $\sim 30''$. 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 $1'$. Fritz et al. (2016) used NACO $H-$ and K_S-band images in the central $r = 20''$. Out to $1\,000''$ ($\sim$39 pc) they used *Hubble Space Telescope* WFC3 data (M127 and M153 filters) and public VISTA Variables in the Via Lactea Survey images ($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

Table 3.1: The Multi-Gaussian Expansion (MGE) fit parameters for the 4.5 μm *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 is the axial ratio of the Gaussian components. I_{unscaled} is the peak surface brightness before scaling to Fritz et al. (2016).

I_{scaled} $\left[10^4\,\mathrm{L}_{\odot,4.5\mu\mathrm{m}}\,\mathrm{pc}^{-2}\right]$	σ_{MGE} [arcsec]	q	I_{unscaled} $\left[10^4\,\mathrm{L}_{\odot,4.5\mu\mathrm{m}}\,\mathrm{pc}^{-2}\right]$
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

et al. (2016). As a result the amplitudes of the inner Gaussians become smaller, but the outer Gaussians (σ_{MGE}>100″∼4 pc) are nearly unchanged. We list the components of the Multi-Gaussian Expansion in Table 3.1 and plot the profile in Fig. 3.2.

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 ionised 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 ionised gas emission and light from young stars are only important in the central parsec.

3.3 Dynamical models of the Milky Way nuclear star cluster

3.3.1 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

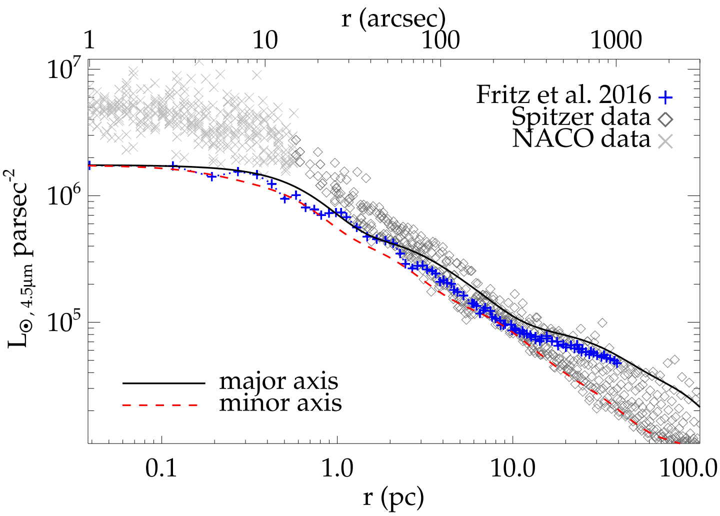


Figure 3.2: Surface brightness profile derived from a dust extinction and PAH emission corrected *Spitzer*/IRAC 4.5 μm image and NACO *H*-band mosaic for the centre, scaled to the measurements of Fritz et al. (2016, blue crosses). The black full line denotes the MGE fit along the major axis, and the red dashed line along the minor axis.

the intrinsic kinematical structure and the underlying gravitational potential (Schwarzschild 1979; van de Ven et al. 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.

3.3.1.1 Mass model

We calculated orbits in the combined gravitational potential of a supermassive black hole $\Phi_\bullet$ and the star cluster $\Phi_\star$, inferred from the imaging data. As we run triaxial models, there are three intrinsic shape parameters, p, q, and u, for the cluster. The shape parameters characterise the axis ratios for the long, intermediate and short axes x, y, and z. They are defined as $p = y/x$, $q = z/x$, and $u = x'/x$, where x' is the length of the longest axis x projected on the sky. Thus, u represents the compression of x due to projection on the sky. Each set of axis ratios refers to a set of viewing angles (θ, ϕ, ψ). The surface brightness distribution is deprojected given the intrinsic shape parameters p, q, u, and multiplied with the dynamical mass-to-light ratio Υ to get the intrinsic stellar mass density $\varrho_\star$. From Poisson's equation $\nabla^2\Phi_\star = 4\pi G\varrho_\star$ one calculates the gravitational potential. We do this for different values of the black hole mass $M_\bullet$, dynamical mass-to-light ratio Υ, and different shape parameters. In total our model has five free parameters, $M_\bullet$, Υ, p, q, 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 measure a dynamical mass-to-light-ratio, which combines the stellar mass-to-light-ratio with other components. These components are the young stars, ionised 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 $M_\odot$ (Feldmeier-Krause et al. 2015). However, the total enclosed stellar mass in the same region is $\sim 10^6\,M_\odot$ (Oh et al. 2009; Feldmeier et al. 2014), and the mass of the supermassive black hole is $4\times10^6\,M_\odot$. The mass of the young stars is therefore probably negligible. The hot ionised 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 circum-nuclear disc may contribute more mass, estimates range from $10^4\,M_\odot$ (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 et al. 2010). The circum-nuclear disc extends over a distance of about 1 pc to more than 5 pc from the centre. At 5 pc the total enclosed stellar mass is $\sim 10^7\,M_\odot$ (McGinn et al. 1989; Feldmeier et al. 2014). We decided to neglect the mass distribution of the circum-nuclear disc in our dynamical models, since it is very uncertain, and makes up only 0.1 to 10 per cent of the stellar mass. The contribution of dark matter to the nuclear star cluster mass is also neglected. Linden (2014) show 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 et al. (1996).

3.3.1.2 Orbit library

The orbit library should be as general as possible and representative for the potential. We assume 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 $\langle L_z\rangle$ 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 sample 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 sample $N_E = 35$ energies calculated from R_c in logarithmic steps ranging from $R_c = 10^{0.5}$ to $R_c = 10^{4.2}$, i.e. $3\farcs16$ to $4\fdg4$ or 0.12 pc to 616.5 pc. We note that the outer radius is about 3.5 times the outermost Gaussian $\sigma_{\rm MGE}$ 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 is 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 is integrated over 200 periods, and sampled on 100 000 points per orbit. For each orbit we store 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 are stored in 183 bins between $-7.4\,\sigma_{\max}$ and $+7.4\,\sigma_{\max}$. These numbers guarantee a proper sampling of the observed velocity profiles (Cretton et al. 1999).

3.3.1.3 Solving the orbital weight distribution

The model has to fit the kinematic data, the intrinsic, and the projected mass distribution. The fit is 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 (NNLS) logarithm 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 two per cent, which is the typical accuracy of the observed surface brightness (van den Bosch et al. 2008).

3.3.2 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. $M_\bullet \approx 2\times10^6\,M_\odot$). 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 $M_\odot/L_{\odot,4.5\mu\mathrm{m}}$). The starting model had $(p, q, u) = (0.84, 0.29, 0.99)$. We sampled different combinations of (p, q, u) in steps of (0.02, 0.01, 0.01), with $0.40 < p < 0.99$, $0.05 < q < 0.29$, and $0.70 < u \leq 0.99$. We found the best fit of the five parameters by calculating the χ^2 from the kinematic measurements. The number of observables is the number of kinematic bins times the number of kinematic moments, in our case $N = 175 \times 4 = 700$.

3.3.3 Modelling results

3.3.3.1 The best-fitting model

Our best-fitting parameters are $M_\bullet = 3.0\times10^6\,M_\odot$, $\Upsilon = 0.90$, $q = 0.28$, $p = 0.64$, $u = 0.99$. This corresponds to best-fitting viewing angles $\vartheta = 80°$, $\varphi = 79°$, $\psi = 91°$. We show the surface brightness map and the symmetrised kinematic maps in Fig. 3.3. 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 $\sim 20''$ (~ 0.8 pc) found by Feldmeier et al. (2014) are well reproduced in the model velocity map. The surface brightness map is reproduced within one per cent. The best fit has $\chi^2 = 290$. With $M = 5$ fitted parameters and $N = 700$ observational constraints, this means $\chi^2_{\mathrm{red}} = 0.42$. That χ^2_{red} 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. 3.4. We plot each combination of parameters. Red colours denote low χ^2, bluer, smaller symbols denote high χ^2.

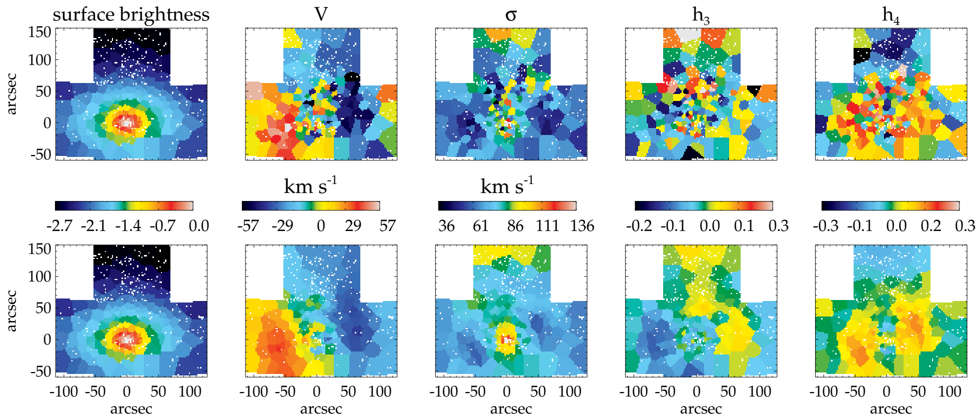


Figure 3.3: Comparison of the observed stellar surface brightness and kinematics (top row) and the best-fitting Schwarzschild model. The columns denote surface brightness, velocity V, velocity dispersion σ, Gauss-Hermite moments h_3, and h_4.

The black cross denotes the best-fitting model. The value of q is constrained by the surface brightness profile. As the lowest value of q_{MGE} is 0.30, the deprojected q cannot be higher. Likewise, the value of $u = 0.99$ is the boundary value of the grid. The values of q, p, and u are averaged over the entire system, i.e. the nuclear stellar disc and the embedded nuclear star cluster. The upper left panel of Fig. 3.4 shows that for each value of u, the best-fitting Y is near to 0.90. A similar behaviour is found with p and q. There is only a slight increase of the best-fitting Y with higher values of p. At the same time, the best-fitting values of $M_\bullet$ do not vary strongly with q, p, or u (second row). 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 certainly contribute to the intrinsic shape fit. The supermassive black hole mass and the dynamical mass-to-light ratio are correlated. For higher values of Y, a lower $M_\bullet$ fits the data.

We show how χ^2 depends on the different parameters in Fig. 3.5. 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. $\sqrt{2(\mathrm{N}-\mathrm{M})} = 37.3$, where N = 700, and M = 5. This value was used as confidence limit by van den Bosch & van de Ven (2009). In Table 3.2 we list the 1σ and 3σ uncertainties.

3.3.3.2 Mass profile

We show the enclosed total mass as a function of the projected radius in Fig. 3.6, grey shaded contours are the 3σ uncertainty. The mass was computed within ellipses. We also plot the results of various other studies. Most studies assumed spherical symmetry, Feldmeier et al.

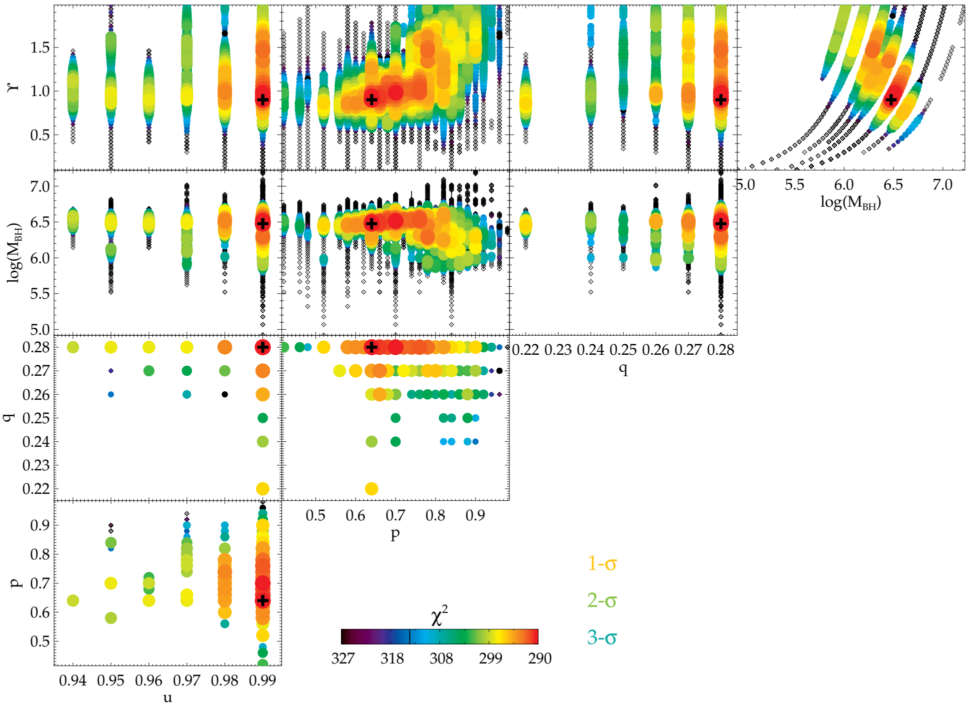


Figure 3.4: Illustration of the fitted parameter space. Each symbol denotes a model, the coloured symbols are models with $\Delta\chi^2 < \sigma_{\chi^2} = 37.3$, black diamonds are models with $\Delta\chi^2 > \sigma_{\chi^2}$. The 1σ, 2σ, and 3σ colours corresponding to $\Delta\chi^2 = 5.9$, 11.3 and 18.2, are denoted. The black cross denotes the best-fitting model.

Table 3.2: 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	$^{+0.76}_{-0.08}$	$^{+1.12}_{-0.32}$	$M_\odot/L_{\odot,4.5\mu m}$
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	$^{+0.0}_{-0.01}$	$^{+0.0}_{-0.05}$	

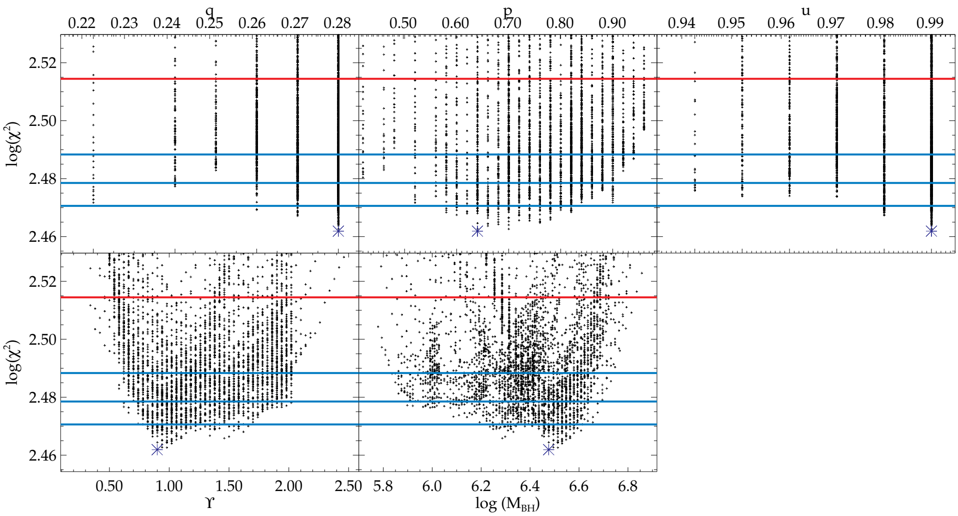


Figure 3.5: The χ^2 values plotted against the five free parameters ($M_\bullet$, Υ, q, p, u). The best-fitting model is denoted as blue asterisk, the 1σ, 2σ, and 3σ confidence level, corresponding to $\Delta\chi^2$ = 5.9, 11.3, and 18.2, are denoted as blue lines. The red line denotes $\sigma^2_\chi = 37.3$.

(2014) and Chatzopoulos et al. (2015a) assumed axisymmetry. Some of the studies used also different Galactocentric distances, so we scaled the masses to $R_0 = 8.0$ kpc. Our results are in agreement with other studies in the central 100″. At larger radii $r \approx 400''$ (~15.5 pc) beyond the reach of our kinematic data, we obtained a higher mass than Lindqvist et al. (1992a). Their data extend to larger radii, but their assumption of spherical symmetry does no longer hold at such large radii. Launhardt et al. (2002) took the flattening of the nuclear stellar disc into account and obtained $M_\star = (8.0 \pm 2) \times 10^8\, M_\odot$ within 120 pc, and in addition $M_{\mathrm{MWNSC}} = (3 \pm 1.5) \times 10^7\, M_\odot$ for the nuclear star cluster. Our best-fitting model has a total enclosed mass $M_\star = (8.8^{+7.4}_{-0.8}) \times 10^8\, M_\odot$ inside an ellipse with semi-major axis distance 120 pc, which is in agreement with Launhardt et al. (2002). The enclosed stellar mass at $r = 8.4$ pc, i.e. about two times the effective radius of the nuclear star cluster, is $M_{\mathrm{MWNSC}} = (3.1^{+2.8}_{-0.3}) \times 10^7\, M_\odot$. The uncertainty comes from the 1σ uncertainty of the mass-to-light ratio Υ.

The black hole influences the stellar kinematics only at the centre of the nuclear star cluster. Out to $r = 33''$ (~1.3 pc), the best-fitting mass of the black hole ($M_\bullet = 3.0 \times 10^6\, M_\odot$) is higher than the enclosed stellar mass of our best-fitting model. Assuming $M_\bullet = 4 \times 10^6\, M_\odot$, this radius increases to 41″ (~1.6 pc). 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 $r_{\mathrm{infl}} = 71''$ (~2.8 pc). This value of r_{infl} is higher than the result of Feldmeier et al. (2014, $(60^{+55}_{-17})''$), as our model has less stellar mass in the centre. We have excellent agreement with Alexander (2005), who found $r_{\mathrm{infl}} = 3$ pc. The kinematic measurements at larger radii have little influence on the black

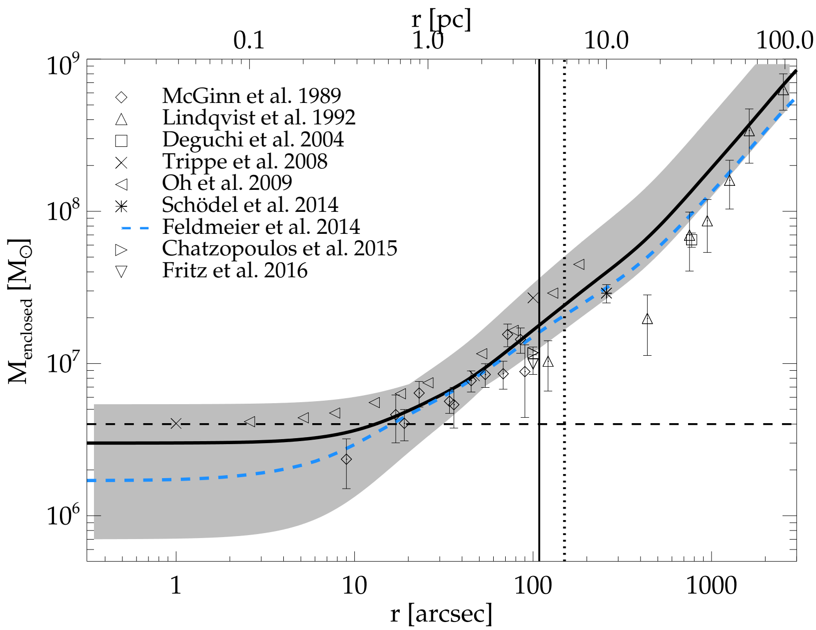


Figure 3.6: Enclosed total mass within a distance of 0.″3 to 50′ along the mean radius of the ellipses in units of $M_\odot$ and in logarithmic scaling. The black line denotes the enclosed mass with $\Upsilon = 0.90$ and $M_\bullet = 3.0 \times 10^6\, M_\odot$, the grey shaded contours are for $\Upsilon = 0.90\,^{+1.12}_{-0.32}$ and $M_\bullet = (3.0\,^{+2.4}_{-2.3}) \times 10^6\, M_\odot$. 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 denotes the effective radius. We also plot the results for the enclosed mass from previous studies. We scaled the masses to $R_0 = 8.0$ kpc if the study assumed a different Galactocentric distance: McGinn et al. (1989, diamonds, assumed R_0 = 8.5 kpc), Lindqvist et al. (1992a, 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. (2014a, 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).

hole mass measurement, but are important to constrain the orbital structure and dynamical mass-to-light ratio.

3.3.3.3 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. 3.7, top panel. We plot the mean anisotropy of the models within the 1σ uncertainty limit. The uncertainty of β is given by the standard deviation and is about 0.1. The plot extends to the outer edge of the kinematic data at 150″. The vertical, solid line denotes the photometric effective radius r_{eff} = 4.2 pc, the dashed line denotes the radius of influence r_{infl} = 71″ (~2.8 pc). The cluster kinematics becomes nearly isotropic at radii r >70″.

We show the angular momentum distribution of the orbits in Fig. 3.8. The colours denote the density of orbits passing radius r with mean angular momentum $\langle\lambda_z\rangle$ (left panel) or $\langle\lambda_x\rangle$ (right panel). The plot of $\langle\lambda_z\rangle$ denotes rotation about the short z-axis. Orbits with $\langle\lambda_z\rangle \neq 0$ 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″ and at larger radii $r \gtrsim 80''$. Short-axis-tube orbits, which contribute in total more mass than long-axis-tube orbits, are most important at r = 60–140″. We illustrate the distribution of the stellar mass on the different orbit types also in Fig. 3.7 (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 in the central 30″. They produce the perpendicular rotating substructure at $r \approx 20''$ (~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 fraction increases towards larger radii. At r = 150″ (~5.8 pc), they contribute 20 per cent.

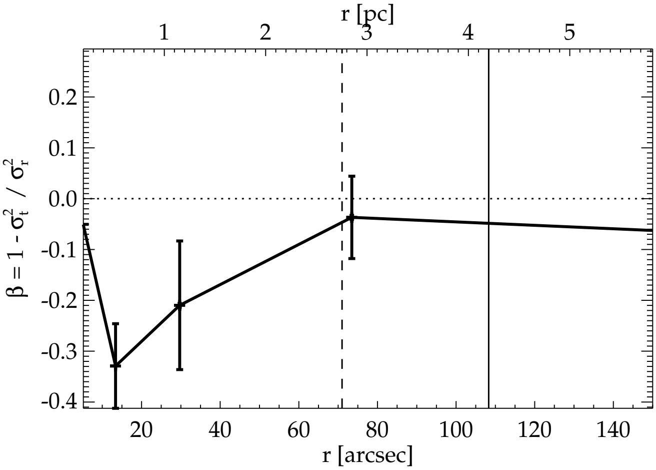
3.4 Discussion

3.4.1 Difference of the resulting black hole mass

The currently best black hole mass estimate is $(4.1 \pm 0.6)\times10^6\,M_\odot$ (Ghez et al. 2008), $(4.3 \pm 0.39)\times10^6\,M_\odot$ (Gillessen et al. 2009b) or $(4.02 \pm 0.20)\times10^6\,M_\odot$ (Boehle et al. 2016), derived from Keplerian stellar orbits around the supermassive black hole. Using axisymmetric Jeans models and the same spectroscopic data as this study, Feldmeier et al. (2014) found a lower value of $M_\bullet = (1.7^{+1.4}_{-1.1})\times10^6\,M_\odot$. The best fit using triaxial Schwarzschild models is $(3.0^{+1.1}_{-1.3})\times10^6\,M_\odot$. This measurement is consistent with the direct measurements of Ghez et al. (2008), Gillessen et al. (2009b), and 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 $0.7\times10^6\,M_\odot$, and an upper limit of $5.4\times10^6\,M_\odot$. 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).

3.4.1.1 Model degeneracies

Some model parameters seem to be correlated. This becomes clear when looking at Fig. 3.4. 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. 3.4 . With a lower value of p, the



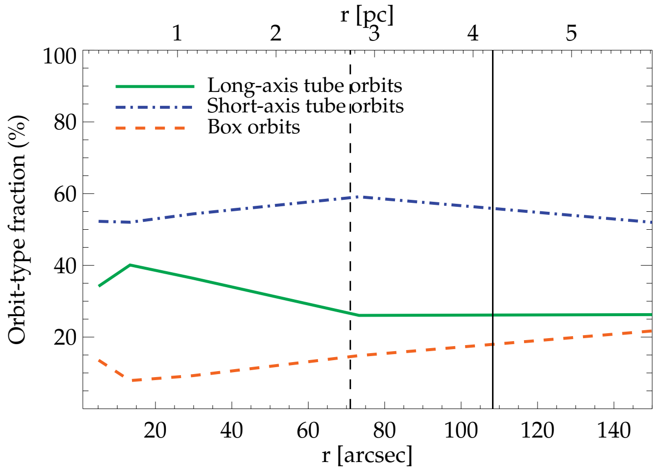


Figure 3.7: Top: Anisotropy β as a function of radius r. Negative values denote tangential anisotropy, positive values radial anisotropy. Bottom: Orbital structure of the Milky Way nuclear star cluster as a function of radius. The green, solid line denotes long-axis-tube orbits; the blue, dot-dashed line short-axis-tube orbits; the red, dashed line box orbits. The vertical, solid line denotes $r_{\text{eff}} = 4.2$ pc (Schödel et al. 2014a); the vertical, dashed line $r_{\text{infl}} = 71''$.

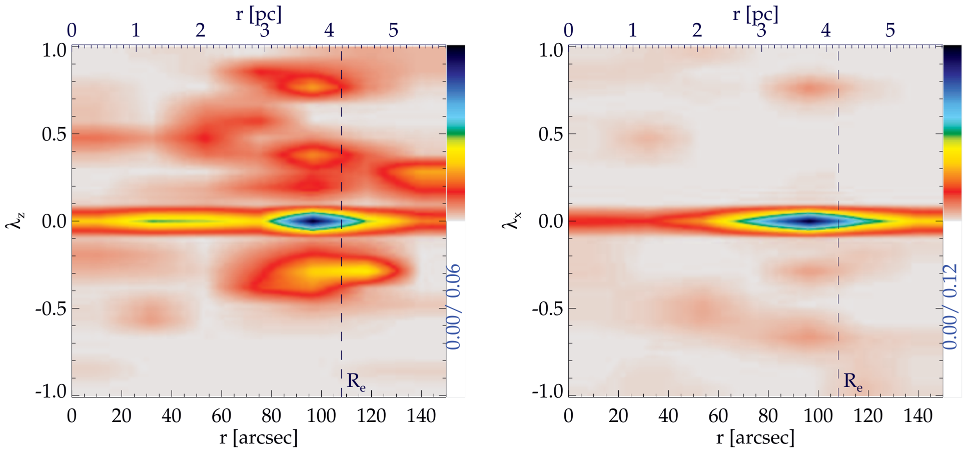


Figure 3.8: Orbit density with angular momentum λ_z (left), i.e. rotation around the short axis, and λ_x (right), i.e. rotation around the long axis. Dark, blue colour indicates higher orbit density. The vertical, dashed line denotes $r_{\rm eff}$ = 4.2 pc (Schödel et al. 2014a).

best-fitting M$_\bullet$ decreases only slightly. At larger p, the χ^2-contours of M$_\bullet$ and Y broaden. This means that for a more oblate axisymmetric cluster with p closer to one, M$_\bullet$ and Y are not as well constrained as with smaller values of p.

The dynamical mass-to-light ratio Y is inversely correlated with the black hole mass (fourth column of the first row in Fig. 3.4). The higher Y, i.e. the more massive the cluster, the less massive is the black hole. This degeneracy is often obtained in dynamical models. Valluri et al. (2004) found that the degeneracy of M$_\bullet$ depends on how well the black hole's sphere of influence is resolved, whereas the measurement of Y is better constrained when the data extend to larger radii. 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 Y. To get agreement with the measurement of (4.1 ± 0.6) $\times10^6$ $M_\odot$ (Ghez et al. 2008), (4.3 ± 0.39) $\times10^6$ $M_\odot$ (Gillessen et al. 2009b), and (4.02 ± 0.20) $\times10^6$ $M_\odot$ (Boehle et al. 2016), we would require a lower value of Y$\approx$ 0.75.

3.4.1.2 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 ionised 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 $r = 32''$ ($\sim$1.2 pc) is 1×10^6 $M_\odot$ less. Our black hole mass is therefore higher, and closer to the direct measurement of M$_\bullet$ $\approx 4\times10^6$ $M_\odot$. 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 3.1. 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.

3.4.1.3 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 higher than expected from stellar-population studies at $3.6\,\mu$m, which found $\Upsilon = 0.4 - 0.75$ (McGaugh & Schombert 2013; Meidt et al. 2014). At $4.5\,\mu$m, Υ is rather less than at $3.6\,\mu$m (Oh et al. 2008). 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 mostly confined to the central 1 pc. 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 ~0.3 within the cluster. Should Υ vary with radius, our mass profile (Fig. 3.6) could 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 parsec 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 et al. (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 circum-nuclear disc. The molecular gas may contribute $10^4 - 10^6\, M_\odot$. The gas disc extends from $r \approx 1{-}7$ 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. Further, our result of Υ would be higher than expected from stellar population studies. 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.

3.4.2 Triaxial cluster shape

Our best-fitting model has axis ratios of $p = y/x = 0.64\,^{+0.18}_{-0.06}$, $q = z/x = 0.28\,^{+0.0}_{-0.02}$, and $u = x'/x = 0.99\,^{+0.0}_{-0.01}$. These axis ratios correspond to viewing angles $\vartheta = 80°$, $\varphi = 79°$, and $\psi = 91°$. 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 strongly triaxial, with a triaxiality parameter $T = (1 - p^2)/(1 - q^2) = 0.64$. An oblate axisymmetric system has $T = 0$, a prolate axisymmetric system has $T = 1$.

Also the Milky Way's bulge is triaxial, the axis ratios are $p = 0.63$ and $q = 0.26$ (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. Intriguingly, the intrinsic shape parameters p and q of the Galactic bulge agree with our best-fit results for the nuclear star cluster within the error bars. 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.

One possible scenario for nuclear star cluster formation is that massive star clusters (10^5–$10^7\,M_\odot$) formed in the galactic disc, migrated to the galaxy's centre and merged (Neumayer et al. 2011; Guillard et al. 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 systematical 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.

3.4.3 Caveats and considerations

3.4.3.1 Regime of semi-resolved 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 Mpc. 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 semi-resolved populations. The brightest stars can be resolved individually, and these stars contribute a large fraction of the flux. A consequence of this is that 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 p, q, 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 1.00008. 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.

3.4.3.2 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″ (~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″ (~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.

3.4.3.3 Implications

Both the semi-resolved 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 $6'' < r < 20''$ 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 $p = y/x = 0.64^{+0.18}_{-0.06}$ may be too low. As a consequence, $\Upsilon = 0.90^{+0.76}_{-0.08}$ would be underestimated (see second column of the first row in Fig. 3.4). However, our best-fitting result of Υ is already higher than what we expect from stellar population studies (McGaugh & Schombert 2013; Meidt et al. 2014), and also higher than the result of Feldmeier et al. (2014), who found $\Upsilon = 0.56^{+0.22}_{-0.26}$. They assumed axisymmetry with $p = 1$, and thus $y = x$, i.e. the intermediate and long axis have the same length. With $p < 1$, the system extends less along the intermediate axis than in the oblate axisymmetric case.

3.4.3.4 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, and the progratde and retrograde tube orbits no longer fill the same volumes, while the box orbits acquire net mean angular momentum (e.g. Heisler et al. 1982; Schwarzschild 1982; Sellwood & Wilkinson 1993; Skokos et al. 2002). As a result, orbit-based tumbling triaxial models are computationally expensive. Other than an early attempt by Zhao (1996) no such models have been constructed that take into account kinematic data. 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.

3.5 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 by Feldmeier et al. (2014) to measure the cluster kinematics of the central 60 pc^2 of the Milky Way. As photometry we used *Spitzer* 4.5μm and NACO $H-$band images, and measured a two-dimensional surface brightness distribution. We excluded young stars, avoided gas emission and dark clouds. 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 $M_\bullet = (3.0^{+1.1}_{-1.3}) \times 10^6\, M_\odot$, a dynamical mass-to-light ratio of $\Upsilon = (0.90^{+0.76}_{-0.08})\, M_\odot/L_{\odot,4.5\mu\mathrm{m}}$, and shape parameters $p = 0.64^{+0.18}_{-0.06}$, $q = 0.28^{+0.0}_{-0.02}$, and $u = 0.99^{+0.0}_{-0.01}$. Our black hole mass measurement is in agreement with the direct measurement of $M_\bullet = (4.1 \pm 0.6) \times 10^6\, M_\odot$ (Ghez et al. 2008), $(4.3 \pm 0.39) \times 10^6\, M_\odot$ (Gillessen et al. 2009b), and $(4.02 \pm 0.20) \times 10^6\, M_\odot$ (Boehle et al. 2016). We obtain a cluster mass $M_{\mathrm{MWNSC}} = (3.1^{+2.8}_{-0.3}) \times 10^7\, M_\odot$ within $r = 2 \times r_{\mathrm{eff}} = 8.4$ pc. The best-fitting model is tangentially anisotropic in the central $r = 2$ pc of the nuclear star cluster, but is 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.

Acknowledgments

R.S. 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.

Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile (289.B-5010(A)).