MNRAS **000**, 1–18 (2015) Preprint 5 January 2016 Compiled using MNRAS LATEX style file v3.0

A spiral galaxy's mass distribution uncovered through lensing and dynamics

Wilma H. Trick[1]*, Glenn van de Ven[1] and Aaron A. Dutton[1]
[1]*Max-Planck-Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany*

Accepted ???. Received ???; in original form ???

ABSTRACT

We investigate the matter distribution of a spiral galaxy with counter-rotating stellar core, SDSS J1331+3628 (J1331), independently with gravitational lensing and dynamical modelling. By fitting a gravitational potential model to a quadruplet of lensing images around J1331's bulge, we tightly constrain the mass inside the Einstein radius $R_{\rm ein} = (0.91 \pm 0.02)''(\simeq 1.83 \pm 0.04$ kpc) to within 4%: $M_{\rm ein} = (7.8 \pm 0.3) \cdot 10^{10} M_\odot$. We model observed long-slit major axis stellar kinematics in J1331's central regions by finding Multi-Gaussian Expansion (MGE) models for the stellar and dark matter distribution that solve the axisymmetric Jeans equations. The lens and dynamical model are independently derived, but in very good agreement with each other around $\sim R_{\rm ein}$. We find that J1331's center requires a steep total mass-to-light ratio gradient. A dynamical model including a NFW halo (with virial velocity $v_{200} \simeq 240 \pm 40$ km s^{-1} and concentration of $c_{200} \simeq 8 \pm 2$) and moderate tangential velocity anisotropy ($\beta_z \simeq -0.4 \pm 0.1$) can reproduce the signatures of J1331's counter-rotating core and predict the stellar and gas rotation curve at larger radii. However, our models do not agree with the observed velocity dispersion at large radii. We speculate that the reason could be a non-trivial re-distribution of matter due to a possible merger event in J1331's recent past.

Key words: gravitational lensing: strong – stars: kinematics and dynamics – galaxies: kinematics and dynamics – galaxies: photometry – galaxies: structure

1 INTRODUCTION

Gravitational lensing and dynamical modelling provide independent constraints on the mass distribution of galaxies. Combining them allows for valuable cross-checking opportunities to disentangle the stellar and dark matter (DM) content of galaxies.

Cosmological simulations suggest that cold DM forms cuspy halos following a Navarro-Frenk-White (NFW) profile (Navarro et al. 1996). However, the existence of central DM density cusps in massive galaxies depends strongly on the stellar mass-to-light ratio (e.g., Dutton et al. 2011a) and DM dominated dwarf galaxies even favour DM halos with cores (e.g., Moore 1994; de Blok et al. 2001). This discrepancy, known as the core-cusp problem, might be resolved by galaxy formation processes such as mergers and outflows (e.g., El-Zant et al. 2001; Pontzen & Governato 2012).

Determining the overall mass distribution in massive galaxies and separating the DM from the stellar mass components is therefore a crucial step in better understanding the structure and formation of galaxies and nature of DM.

Massive galaxies can act as gravitational lenses, deflect the light of background sources and give rise to multiple images. This strong gravitational lensing tightly constrains the projected total mass of the lens galaxy inside $\sim 1''$ (e.g., Treu 2010).

The mass profile at larger galactocentric radii can be probed by gas rotation curves that directly measure the galaxy's circular velocity profile (e.g., Rubin et al. 1980)[TO DO: Can I use this here?]. However, due to its dissipative nature, gas motions are very sensitive to disturbances by e.g., spiral arms and bars (e.g., Sellwood 2004).

Because stars are dissipationless dynamical tracers and present almost everywhere in the galaxy, stellar dynamical modelling can complement mass constraints from lensing at small and gas motions at large radii. As stellar motions are complex—a bulk rotation around one principal axis combined with random motions in all coordinate directions—, full dynamical modelling of rotation, dispersion and velocity anisotropies is needed to deduce the matter distribution.

The Sloan WFC Edge-on Late-type Lens Survey (SWELLS, WFC = Wide field camera) (Treu et al. 2011; Dutton et al. 2011b; Brewer et al. 2012; Barnabè et al. 2012; Dutton et al. 2013; Brewer et al. 2014) is dedicated

* E-mail: trick@mpia.de

to find and investigate spiral galaxies, which are (a) strong gravitational lenses and (b) seen almost edge-on, such that rotation curves can be easily measured. By combining lensing and dynamical modelling degeneracies inherent in both methods can be broken.

One of the SWELLS galaxies is the massive spiral SDSS J1331+3628, to which we refer as J1331 in the remainder of this work. It has bluish spiral arms and a large reddish bulge (see Figures 1a and 1b), which is superimposed by a quadruplet of extended bluish images approximately at a distance of 1″ from the galaxy center (see Figure 1c). The lensed object might be a star-forming blob of a background galaxy. J1331 stands out of the SWELLS sample because of its large counter-rotating core (see Figure 1d), which might be an indication that J1331 underwent a merger in its recent past.

Treu et al. (2011) confirmed that J1331 was a strong gravitational lens, measured its apparent brightness and estimated the stellar masses of disk and bulge. The lensing properties of J1331 were first analysed by Brewer et al. (2012). Dutton et al. (2013) measured the gas and stellar kinematics along the major axis and deduced J1331's mass profile from the gas rotation curve at large radii and total mass inside the Einstein radius from gravitational lensing, focusing mostly on the outer regions of J1331.

The goal of this work is to complement the previous work on J1331 by an in-depth analysis of the matter distribution in J1331's inner regions. We use stellar dynamical modelling in addition to lensing constraints, similar to a study of the Einstein Cross by van de Ven et al. (2010). We attempt to disentangle the stellar and DM components and test if employed axisymmetric Jeans models work also in the presence of a counter-rotating core. Ideally, this work on J1331 could also help understanding how mergers modify the mass distribution of a galaxy.

This paper is organized as follows: Section 2 summarizes the data, Section 3 gives an overview of the modelling techniques used in this work, and Section 4 presents our results on the surface photometry of J1331 using Multi-Gaussian expansions (Section 4.1), constraints from lensing (Section 4.2) and Jeans modelling based on the surface brightness only (Section 4.3) and including a NFW DM halo (Section 4.4). Finally Section 5 uses the result to discuss J1331's stellar mass-to-light ratio, possible merger history and starting points for future work.

2 DATA

Redshift and position. J1331 is located at right ascension = 202.91800° and a declination = 36.46999° (epoch J2000). Treu et al. (2011) found from SDSS spectra that J1331 has two redshifts inside 1″: J1331 itself has $z_d = 0.113$ and $z_s = 0.254$ is the redshift of the lensed background source (Brewer et al. 2012). According to the WMAP5 cosmology (Dunkley et al. 2009), J1331 has an angular diameter distance of 414 Mpc, which translates into a transverse scaling of $1'' \widehat{=} 2.01$ kpc.

HST imaging. We use HST imaging of J1331 by Treu et al. (2011). They performed high resolution imaging with the Hubble Space Telescope's (HST) Wide-Field Planetary Camera 2 (WFPC2) and its WF3 CCD chip. The images are a combination of four exposures with each an exposure time of 400 sec and were drizzled to a pixel scale of 1 pixel = 0.05″. In particular, we use the images in the filters F450W, to identify the positions of the bluish lensing images, and F814W (I-band) to create a surface brightness MGE model of the reddish bulge.

Stellar kinematics. For the dynamical modelling we use the stellar kinematics for J1331 measured by Dutton et al. (2013). They obtained long-slit spectra along J1331's major-axis with the Low Resolution Imaging Spectrograph (LRIS) on the Keck I 10m telescope. The width of the slit was 1″ and the seeing conditions had a FWHM of ∼ 1.1″. 1D Spectra for spatial bins of different widths along the major axis were extracted. They measured line-of-sight stellar rotation velocities ($v_{\rm rot}$) and stellar velocity dispersion (σ) from the spectrum containing the absorption lines Mgb (5177 Å) and FeII (5270, 5335 and 5406 Å) (analogous to Dutton et al. 2011b). Gas kinematics were extracted from Gaussian line profile fits to Hα (6563 Å) and [NII] (6583 and 6548 Å) emission lines, as tracers for ionized gas.

The stellar kinematics, $v_{\rm rot}$, σ and $v_{\rm rms}^2 = v_{\rm rot}^2 + \sigma^2$ are shown in Figure 1d. The rotation curve reveals a counter-rotating core within $2'' \simeq 4$ kpc. Outside of ∼ 3.5″ there is a steep drop in the dispersion. This could indicate the boundary between the pressure supported bulge and the rotationally supported disk, which appears around this radius in the F450W filter in Figure 1a. However, in the brighter F814W filter in Figure 1b the large reddish bulge extends out to ∼ 5″.

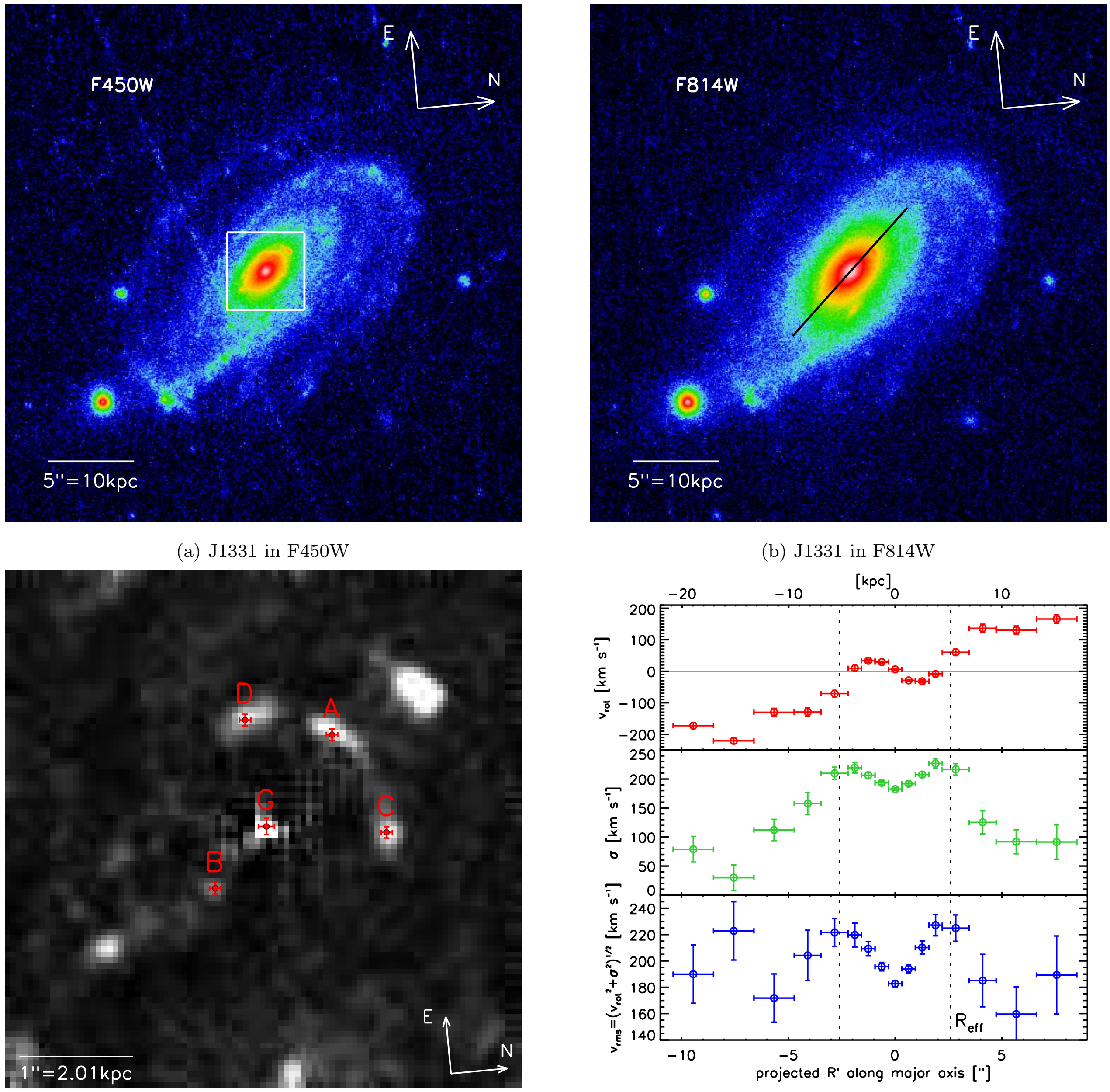
Inside of ∼ 4″, the data appears to be symmetric, outside of this the assumption of axisymmetry seems not to be valid anymore, considering the data. We add -2.3 km s^{-1} to the $v_{\rm rot}$ to ensure $v_{\rm rot}(R' = 0) \sim 0$ as a possible correction term for a misjudgement of the systemic velocity. We also symmetrize the data within 4″ and asign a minimum error of $\delta v_{\rm rms} > 5$ km s^{-1} to the $v_{\rm rms}$ data. In the JAM modelling, which is based on the assumption of axisymmetry, only stellar kinematics with $R' \lesssim 2.5''$ or $R' \lesssim 4''$ are used. Another reason to restrict to modelling on the bulge region is that our MGE in Table 3 is only a good representation of J1331's F814W light distribution inside ∼ 5″.

3 MODELLING

3.1 Surface brightness model

Multi-Gaussian Expansion (MGE). MGEs are used to parametrize the observed surface brightness (or projected total mass) of a galaxy as a sum of N 2-dimensional, elliptical Gaussians (Bendinelli 1991; Monnet et al. 1992; Emsellem et al. 1994, 1999). This work makes use of the algorithm and code[1] by Cappellari (2002). We assume all Gaussians to have the same center and position angle ϕ, i.e., the orientation of the Gaussians' major axis measured

[1] The IDL code package for fitting MGEs to images by Cappellari (2002) is available online at http://www-astro.physics.ox.ac.uk/~mxc/software. The version from June 2012 was used in this work.



(a) J1331 in F450W

(b) J1331 in F814W

(c) The lensing images

(d) Stellar Kinematics by Dutton et al. (2013)

Figure 1. Hubble Space telescope (HST) images and stellar kinematics of the galaxy SDSS J1331+3628 (J1331), which has a large counter-rotating core and whose bulge acts as a strong lens for a bluish background source. *Panel (a) and (b):* HST/WFPC2/WFC3 images of J1331 by Treu et al. (2011) in two filters, F450W in panel (a) and F814W in panel (b). The black solid line in panel (b) shows the orientation of the major-axis. Its length is $10''$ and it indicates approximately where we carry out the dynamical modelling. *Panel (c):* Lensing images in the central region of J1331. An IRAF *ellipse* fit to the F450W surface brightness in panel (a) was subtracted from the image. The (smoothed) residuals within the white square in panel (a) are shown in panel (c). (The four bright blobs (A,B,C and D), that become visible, are arranged in a typical strong lensing configuration around the center of the galaxy (G). The configuration of the two additional blobs, that lie approximately on a line with A, B and G, does not suggest that they form a lensing doublet. They might be star forming regions of a background galaxy.) *Panel (d):* Stellar Kinematics along the galaxy's major axis as measured by Dutton et al. (2013), line-of-sight rotation velocity $v_{\rm rot}$, line-of-sight velocity dispersion σ and the rms-velocity $v_{\rm rms} = \sqrt{v_{\rm rot}^2 + \sigma^2}$. The dotted line in panel (b) indicates the galaxy's effective half-light radius (in the F814W filter), $R_{\rm eff} = 2.6" = 5.2$ kpc. The $v_{\rm rot}$ curve reveals that J1331 has a counter-rotating core within $R_{\rm eff}$.

counter-clockwise from the y'-axis of the coordinate system with polar coordinates (R', θ'). Then the surface brightness can be written as

$$\begin{aligned} I(R', \theta') &= \sum_{i=1}^{N} I_{0,i} \exp\left[-\frac{1}{2\sigma_i^2}\left(x'^2_i + \frac{y'^2_i}{q'^2_i}\right)\right] && (1) \\ \text{with } I_{0,i} &= \frac{L_i}{2\pi\sigma_i^2 q'_i} && (2) \\ \text{and } x'_i &= R' \cos(\theta' - \phi) \\ y'_i &= R' \sin(\theta' - \phi), \end{aligned}$$

where $I_{0,i}$ is the central surface brightness of each Gaussian, L_i its total luminosity, σ_i its dispersion along the major axis and q'_i the axis ratio between the elliptical Gaussian's major and minor axis.

We can also expand the telescope's point-spread function (PSF) as a sum of circular Gaussians,

$$\mathrm{PSF}(x', y') = \sum_j \frac{G_j}{2\pi\delta_j^2} \exp\left[-\frac{1}{2\delta_j^2}\left(x'^2 + y'^2\right)\right], \quad (3)$$

where $\sum_j G_j = 1$ and δ_j are in this case the dispersions of the circular PSF Gaussians.

The observed surface brightness distribution is a convolution of the intrinsic surface brightness in Equation (1) with the PSF in Equation (3): $(I * \mathrm{PSF})(x', y')$ is then again a sum of Gaussians and can be directly fitted to an image of the galaxy in question.

$I(R', \theta')$ describes the intrinsic, to 2D projected light distribution or surface density of the galaxy. Under the assumption that the galaxy is oblate and axisymmetric, and given the inclination angle i of the galaxy with respect to the observer, MGEs allow an analytic deprojection of the 2D MGE to get a 3D light distribution or density $\nu(R, z)$ for the galaxy,

$$\nu(R, z) = \sum_i \nu_{0,i} \exp\left[-\frac{1}{2\sigma_i}\left(R^2 + \frac{z^2}{q_i^2}\right)\right]. \quad (4)$$

The flattening of each axisymmetric 3D Gaussian q_i and its central density $\nu_{0,i}$ follow from the observed 2D axis ratio q'_i and surface density $I_{0,i}$ as

$$\begin{aligned} q_i^2 &= \frac{q'^2_i - \cos^2 i}{\sin^2 i} \\ \nu_{0,i} &= \frac{q'_i I_{0,i}}{q_i \sqrt{2\pi\sigma_i^2}}. \end{aligned}$$

3.2 Strong gravitational lens model

Lensing formalism. A gravitational lens is a mass distribution, whose gravitational potential Φ acts as a lens for light coming from a source positioned somewhere on a plane behind the lens. The angular diameter distance from the observer to the lens is D_d, to the source plane D_s, and the distance between the lens and source plane is D_{ds}. The deflection potential of the lens is its potential, projected along the line of sight z and rescaled to

$$\psi(\vec{\theta}) := \frac{D_{ds}}{D_d D_s} \frac{2}{c^2} \int \Phi(\vec{r} = D_d \vec{\theta}, z) \, \mathrm{d}z, \quad (5)$$

where $\vec{\theta}$ is a 2-dimensional vector on the plane of the sky. The light from the source at $\vec{\beta} = (\xi, \eta)$[2] is deflected according to the lens equation

$$\vec{\beta} = \vec{\theta}_i - \vec{\nabla}_\theta \psi(\vec{\theta})\Big|_{\vec{\theta}_i} \quad (6)$$

into an image $\vec{\theta}_i = (x_i, y_i)$. The gradient of the deflection potential $\vec{\nabla}_\theta \psi(\vec{\theta})$ is equal to the angle by which the light is deflected multiplied by D_{ds}/D_s.

The total time delay of an deflected light path through $\vec{\theta}$ with respect to the unperturbed light path is given by

$$\Delta t(\vec{\theta}) = \frac{(1 + z_d)}{c} \frac{D_d D_s}{D_{ds}} \left[\frac{1}{2}(\vec{\theta} - \vec{\beta})^2 - \psi(\vec{\theta})\right] \quad (7)$$

(Narayan & Bartelmann 1999). According to Fermat's principle the image positions will be observed at the extrema of $\Delta t(\vec{\theta})$.

The inverse magnification tensor

$$\mathscr{M}^{-1} \equiv \frac{\partial \vec{\beta}}{\partial \vec{\theta}} \overset{(6)}{=} \left(\delta_{ij} - \frac{\partial^2 \psi}{\partial \theta_i \partial \theta_j}\right) \quad (8)$$

describes how the source position changes with image position. It also describes the distortion of the image shape for an extended source and its magnification due to lensing according to

$$\mu \equiv \frac{\text{image area}}{\text{source area}} = \det \mathscr{M}.$$

Lines in the image plane for which the magnification becomes infinite, i.e., $\det \mathscr{M}^{-1} = 0$, are called critical curves. The corresponding lines in the source plane are called caustics. The position of the source with respect to the caustic detemines the number of images and their configuration and shape with respect to each other.

The Einstein mass M_{ein} and Einstein radius R_{ein} are defined via the relation

$$M_{\mathrm{ein}} \equiv M_{\mathrm{proj}}(< R_{\mathrm{ein}}) \overset{!}{=} \pi \Sigma_{\mathrm{crit}} R_{\mathrm{ein}}^2,$$

where $\Sigma_{\mathrm{crit}} \equiv \frac{c^2}{4\pi G} \frac{D_s}{D_d D_{ds}}$ is the critical density and $M_{\mathrm{proj}}(< R_{\mathrm{ein}})$ is the mass projected along the line-of-sight within R_{ein}. M_{ein} is similar to the projected mass within the critical curve M_{crit}.

Lens model. Following Evans & Witt (2003) we assume a scale-free model

$$\psi(R', \theta') = R'^{\alpha} F(\theta') \quad (9)$$

for the lensing potential, consisting of an angular part $F(\theta')$ and a power-law radial part, with (R', θ') being polar coordinates on the plane of the sky. The case $\alpha = 1$ corresponds to a flat rotation curve. We expand $F(\theta')$ into a Fourier series,

$$F(\theta') = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos(k\theta') + b_k \sin(k\theta')\right). \quad (10)$$

For this scale-free lens model the lens equation (6) becomes

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} R'_i \cos\theta'_i - R'^{\alpha-1}_i \left(\alpha \cos\theta'_i F(\theta'_i) - \sin\theta'_i F'(\theta'_i)\right) \\ R'_i \sin\theta'_i - R'^{\alpha-1}_i \left(\alpha \sin\theta'_i F(\theta'_i + \cos\theta'_i F'(\theta'_i)\right) \end{pmatrix} \quad (11)$$

[2] ξ and η are cartesian coordinates on the plane of the sky.

(Evans & Witt 2003), where $F'(\theta') \equiv \partial F(\theta')/\partial\theta'$. When we fix the slope α, then the lens equation is a purely linear problem and can be solved numerically for the source position (ξ, η) and the Fourier parameters (a_k, b_k) given one observed image at position $(x_i = R'_i \cos\theta'_i, y_i = R'_i \sin\theta'_i)$.

Model fitting. The free parameters of the lens model are: the source position (ξ, η), the radial slope α and Fourier parameters (a_k, b_k) of the lens mass distribution in Equations (9) and (10). We want to minimize the distance between the observed image positions, $\vec{\theta}_{oi}$, and those predicted by the lensing model, $\vec{\theta}_{pi}$. To avoid having to solve the lens equation (11) for $\vec{\theta}_{pi}$, we follow Kochanek (1991) and cast the calculation back to the source plane using the magnification tensor in Equation (8) to approximate $\vec{\theta} \simeq (\partial\vec{\theta}/\partial\vec{\beta})\vec{\beta} = \mathscr{M}\vec{\beta}$. The best fit lens model is then the one that minimizes

$$\begin{aligned}\chi^2_{\rm lens} &= \sum_i \left| \begin{pmatrix} \frac{1}{\Delta_x} & 0 \\ 0 & \frac{1}{\Delta_y} \end{pmatrix} \left(\vec{\theta}_{pi} - \vec{\theta}_{oi}\right) \right|^2 \\ &\simeq \sum_i \left| \begin{pmatrix} \frac{1}{\Delta_x} & 0 \\ 0 & \frac{1}{\Delta_y} \end{pmatrix} \mathscr{M}|_{\vec{\theta}=\vec{\theta}_{oi}} \begin{pmatrix} \xi - \tilde{\xi}_i \\ \eta - \tilde{\eta}_i \end{pmatrix} \right|^2, \end{aligned} \tag{12}$$

where (Δ_x, Δ_y) are the measurement errors of the image positions $\vec{\theta}_{oi}$. $\mathscr{M}|_{\vec{\theta}=\vec{\theta}_{oi}}$ is the magnification tensor and $(\tilde{\xi}_i, \tilde{\eta}_i)$ the source position according to the lens equation, both evaluated at the position of the i-th lensing image, $\vec{\theta}_{oi}$. Following van de Ven et al. (2010) we add a term

$$\chi^2_{\rm shape} = \lambda \sum_{k\geqslant 3} \frac{(a_k^2 + b_k^2)}{a_0^2} \tag{13}$$

which forces the shape of the mass distribution to be close to an ellipse. The total χ^2 to minimize is therefore

$$\chi^2 = \chi^2_{\rm lens} + \chi^2_{\rm shape} \tag{14}$$

We set $a_1 = b_1 = 0$, which corresponds to the choice of origin; in this case the center of the galaxy.

To be able to constrain the slope α, we would need flux ratios for the images as in van de Ven et al. (2010). But the extended quality of the images, possible dust obscuration and surface brightness fluctuations due to microlensing events, as well as the uncertainty in surface brightness subtraction, make flux determination too unreliable and we do not include them in the fitting.

3.3 Axisymmetric dynamical model

Jeans Anisotropic Models (JAM). JAM modelling assumes galaxies to be (a) collisionless (i.e., the collisionless Boltzmann equation for the distribution function $f(\vec{x}, \vec{v}, t)$ has to be satisfied, $\frac{{\rm d}f(\vec{x},\vec{v},t)}{{\rm d}t} = 0$), (b) in a steady state ($\frac{\partial}{\partial t} = 0$), (c) axisymmetric (best described in cylindrical coordinates (R, z, ϕ) and $\frac{\partial}{\partial\phi} = 0$). From this follow the axisymmetric Jeans equations as the vector-valued first moment of the Boltzmann equation, i.e.,

$$\int \vec{v}\frac{{\rm d}f}{{\rm d}t}\,{\rm d}^3v = 0.$$

To be able to solve the Jeans equations, additional assumptions about the velocity ellipsoid tensor $\langle v_i v_j \rangle$ have to be made. We follow Cappellari (2008) and assume firstly, that the galaxy's velocity ellipsoid is aligned with the cylindrical coordinate system, i.e., $\langle v_R v_z \rangle = \langle v_z v_\phi \rangle = \langle v_\phi v_R \rangle = 0$. Secondly, we assume a constant ratio between the radial and vertical second velocity moments,

$$\beta_z \equiv 1 - \langle v_z^2 \rangle / \langle v_R^2 \rangle. \tag{15}$$

This reduces the Jeans equations to two equations for $\langle v_z^2 \rangle$ and $\langle v_\phi^2 \rangle$, that can be solved by means of one integration,

$$n\langle v_z^2 \rangle(R,z) = \int_0^\infty n \frac{\partial\Phi}{\partial z}\,{\rm d}z \tag{16}$$

$$n\langle v_\phi^2 \rangle(R,z) = R\frac{\partial}{\partial R}\left(\frac{n\langle v_z^2 \rangle}{1-\beta_z}\right) + \frac{n\langle v_z^2 \rangle}{1-\beta_z} + Rn\frac{\partial\Phi}{\partial R} \tag{17}$$

where $n(\vec{x}) \equiv \int f(\vec{x}, \vec{v})\,{\rm d}^3v$ is the number density of (stellar) tracers and $\Phi(\vec{x})$ the galaxy's gravitational potential, generated by the mass density $\rho(\vec{x})$.

The JAM modelling approach by Cappellari (2008) expresses the tracer and mass density in terms of MGEs (see also Emsellem et al. 1994). The tracer density $n(\vec{x})$ is assumed to be proportional to the observed and deprojected brightness distribution $\nu(R, z)$ in Equation (4). The mass density $\rho(R, z)$ can consist of several sets of MGEs, describing stellar and DM components. $\Phi(R, z)$ is generated from the mass density MGE by integrating the Poisson equation (Emsellem et al. 1994). Equations (16) and (17) together with (15) provide the velocity dispersion tensor $\langle v_i v_j \rangle = \langle v_i^2 \rangle$ (with $i, j \in \{R, \phi, z\}$). $\langle v_i^2 \rangle$ is then rotated by the inclination angle i to the coordinate system of the observer; (x', y') is the plane of the sky and z' the line-of-sight, where x' is alinged with the galaxy's major axis. Taking a light-weighted projection along the line-of-sight gives a model prediction for the line-of-sight velocity second moment, which is comparable to actual spectroscopic measurements of the second velocity moment. Details of the derivation using the MGE formalism are given in Cappellari (2008) and in the appendix of van de Ven et al. (2010). We therefore just give the result for the line-of-sight second velocity moment prediction from the Jeans equations,

$$\begin{aligned}&\left(I\langle v_{\rm los}^2 \rangle\right)(x', y') \\ &= 4\pi^{3/2} G \int_0^1 \sum_{k=1}^{N}\sum_{j=1}^{M} \nu_{0,k}\; q_j\; \rho_{0,j}\; u^2 \\ &\times \frac{\sigma_k^2\, q_k^2 \left(\cos^2 i + \frac{\sin^2 i}{1-\beta_{z,k}}\right) + \mathscr{D}\, x'^2 \sin^2 i}{(1-\mathscr{C}u^2)\sqrt{(\mathscr{A} + \mathscr{B}\cos^2 i)\left[1-(1-q_j^2)u^2\right]}} \\ &\times \exp\left\{-\mathscr{A}\left[x'^2 + \frac{(\mathscr{A}+\mathscr{B})\,y'^2}{\mathscr{A}+\mathscr{B}\cos^2 i}\right]\right\}{\rm d}u, \end{aligned} \tag{18}$$

with N Gaussians describing the tracer distribution and M Gaussians describing the mass distribution, $\rho_{0,j}$ being the j-th mass density Gaussian evaluated at $(R = 0, z = 0)$ and

$$\begin{aligned}\mathscr{A} &= \frac{1}{2}\left(\frac{u^2}{\sigma_j^2} + \frac{1}{\sigma_k^2}\right) \\ \mathscr{B} &= \frac{1}{2}\left\{\frac{1-q_k^2}{\sigma_k^2 q_k^2} + \frac{(1-q_j^2)u^4}{\sigma_j^2\left[1-(1-q_j^2)u^2\right]}\right\} \\ \mathscr{C} &= 1 - q_j^2 - \frac{\sigma_k^2 q_k^2}{\sigma_j^2} \\ \mathscr{D} &= 1 - \frac{q_k^2}{1-\beta_{z,k}} - \left[\left(1 - \frac{1}{1-\beta_{k,z}}\right)\mathscr{C} + \frac{1-q_j^2}{1-\beta_{z,k}}\right]u^2. \end{aligned}$$

The JAM modelling code by Cappellari (2008) evaluates Equation (18) numerically for a given luminous tracer and mass distribution MGE and a given inclination.

Data comparison. As data we use stellar line-of-sight rotation velocities $v_{\rm rot} \equiv \langle v_{\rm los}\rangle$ and velocity dispersions σ as described in Section 2. The JAM models give us a prediction for the second line-of-sight velocity moment $\langle v_{\rm los}^2\rangle$. The root mean square (rms) line-of-sight velocity $v_{\rm rms}$ allows a data-model comparison by relating theses velocities according to

$$v_{\rm rms}^2 = \langle v_{\rm los}^2\rangle = v_{\rm rot}^2 + \sigma^2.$$

The model in Equation (18) predicts the intrinsic $\langle v_{\rm los}^2\rangle$ at a given position on the sky, which needs then to be modified according to the mode of observation, to be comparable to the measurements. The measured $v_{\rm rms}$ is a light-weighted mean for a pixel along the long-slit of the spectrograph, with width $L_y = 1''$ (Dutton et al. 2013) and a certain given extent along the galaxy's major axis, L_x, i.e., for a rectangular aperture

$$\mathrm{AP}(x',y') \equiv \begin{cases} 1 & \text{for } -\frac{L_x}{2} \leqslant x' < +\frac{L_x}{2} \\ & \text{and } -\frac{L_y}{2} \leqslant y' \leqslant +\frac{L_y}{2} \\ 0 & \text{otherwise.} \end{cases}$$

The light arriving at the spectrograph itself was subject to seeing, i.e., a Gaussian with Full Width Half Maximum (FWHM) of $1.1''$ (Dutton et al. 2013),

$$\mathrm{PSF}(x',y') \equiv \mathcal{N}(0, \mathrm{FWHM}/2\sqrt{2\ln 2}).$$

The model predictions for $\langle v_{\rm los}^2\rangle$ have therefore to be convolved with the convolution kernel

$$\begin{aligned} K(x',y') &\equiv (\mathrm{PSF} * \mathrm{AP})(x',y') \\ &= \frac{1}{4}\prod_{u\in\{x',y'\}}\left[\mathrm{erf}\left(\frac{L_u/2-u}{\sqrt{2}\sigma_{\rm seeing}}\right) + \mathrm{erf}\left(\frac{L_u/2+u}{\sqrt{2}\sigma_{\rm seeing}}\right)\right], \end{aligned}$$

and weighted by the surface brightness $I(x',y')$, i.e.,

$$\begin{aligned} I_{\rm obs} &= I * K \\ \langle v_{\rm los}^2\rangle|_{\rm obs} &= \frac{(I\langle v_{\rm los}^2\rangle) * K}{I_{\rm obs}}. \end{aligned}$$

If provided with the convolution kernel, the JAM code[3] by Cappellari (2008) performs the convolution numerically. We set $L_x = 0.21''$ as the width of the model pixel, and get a prediction for the actual measurements in bins of width $0.63''$, $1.26''$ and $1.89''$ (Dutton et al. 2013) as light-weighted mean from each 3, 6 and 9 model pixels.

Rotation curve. The intrinsic rotation curve is the first velocity moment $\langle v_\phi\rangle = \sqrt{\langle v_\phi^2\rangle - \sigma_\phi^2}$. The observed rotation curve is the projection of the light-weighted contributions to $\langle v_\phi\rangle$ along the line-of-sight (Cappellari 2008),

$$I\langle v_{\rm los}\rangle = \int_{-\infty}^{+\infty} \nu\langle v_\phi\rangle \cos\phi \sin i \,\mathrm{d}z'.$$

[3] The IDL code package for Jeans Anisotropic Models (JAM) by Cappellari (2008) is available online at http://www-astro.physics.ox.ac.uk/~mxc/software. The version from June 2012 was used in this work.

The first velocity moments cannot be uniquely determined from the Jeans equations, which give only a prediction for the second velocity moments. Further assumptions are needed to separate the second velocity moments into ordered and random motion. Cappellari (2008) assumes that in a steady state there is no streaming velocity in R direction, i.e., $\langle v_R\rangle = 0$ and therefore $\sigma_R^2 = \langle v_R^2\rangle$. Then Cappellari (2008) relates the dispersions in R and ϕ direction such that

$$\langle v_\phi\rangle = \sqrt{\langle v_\phi^2\rangle - \sigma_\phi^2} \equiv \kappa\sqrt{\langle v_\phi^2\rangle - \langle v_R^2\rangle},$$

and the κ parameter quantifies the rotation: $\kappa = 0$ means no rotation at all and $|\kappa| = 1$ describes a velocity dispersion ellipsoid that is a circle in the R-ϕ plane (Cappellari 2008). The sign of κ determines the rotation direction. We can assign a constant κ_k to every Gaussian in the MGE formalism and

$$\nu\langle v_\phi\rangle = \left[\nu\sum_k \kappa_k^2\left([\nu\langle v_\phi^2\rangle]_k - [\nu\langle v_R^2\rangle]_k\right)\right]^{1/2}$$

is then the light-weighted circular velocity curve, given the second velocity moments found from the Jeans equations.

To model the counter-rotating core of J1331 with one free parameter, we employ the condition that the overall $\kappa(R)$ profile should smoothly and relatively steeply transit from $\kappa(R) = -\kappa' < 0$ at small R through $\kappa(R_0) = 0$ and increase to $\kappa(R) = \kappa' > 0$ at large R. Our imposed profile is

$$\kappa(R) = \kappa'\frac{R^2 - R_0^2}{R^2 + R_0^2}. \tag{19}$$

We find κ' by matching the model $\langle v_{\rm los}\rangle$ with the symmetrized $v_{\rm rot}$ data, where for a given κ' the κ_k are found from fitting the MGE generated profile $\kappa(R) = \sum_k \kappa_k\nu_k(r)/\sum_k\nu_k(r)$ to Equation (19). The observed zero point is at $R_0' \approx 2''$. In the deprojected galactic plane the radius of zero rotation would be at a $R_0 \gtrsim 2''$, and we choose it to be at $2.2''$.

Including a NFW halo. As mentioned above, JAM modelling allows to incorporate an invisible matter component in addition to the luminous matter in the form of an MGE. In Section 4.4 we will include a spherical Navarro-Frenk-White (NFW) DM halo (Navarro et al. 1996) in the dynamical model. The classical NFW profile has the form

$$\rho_{\rm NFW}(r) \propto r^{-1}(r + r_s)^{-2} \tag{20}$$

and two free parameters, the scale length r_s and a parameter describing the total mass of the halo. We will use v_{200}, which is the circular velocity at the radius r_{200} within which the mean density of the halo is 200 times the cosmological critical density $\rho_{\rm crit} \equiv (2H^2)/(8\pi G)$, i.e.

$$\begin{aligned} M_{200} &= M(< r_{200}) \\ \frac{M_{200}}{\frac{4}{3}\pi r_{200}^3} &= 200\rho_{\rm crit}(z=0) \\ v_{200} &= \sqrt{\frac{GM_{200}}{r_{200}}} \end{aligned}$$

with $\rho_{\rm crit}(z=0) = 1.43\cdot 10^{-7} M_\odot/\mathrm{pc}^3$ in the WMAP5 cosmology by Dunkley et al. (2009). How much the mass is concentrated in the center of the NFW halo is given by the concentration of the NFW halo defined by

$$c_{200} \equiv r_{200}/r_s. \tag{21}$$

Table 1. Galaxy Parameters of J1331.

redshift (Brewer et al. 2012)	z_d	0.113
angular diameter distance	D_d [Mpc]	414
scaling	1 kpc / 1 $''$	2.006
position angle from North	ϕ [degrees]	42.90°
average axis ratio	q'	0.598
average ellipticity	$\epsilon = 1 - q'$	0.402
apparent I-band magnitude	m_I [mag]	15.77
total I-band luminosity	$L_\mathrm{tot,I}$ [$10^{10} L_\odot$]	5.6
effective half-light radius	R_eff [$''$]	2.6
	R_eff [pc]	5.2

Table 2. F814W PSF MGE: Parameters of the circular four-Gaussian MGE in Equation (3) fitted to the radial profile of the synthetic HST/F814W PSF image.

j	G_j	δ_j [$''$]
1	0.184	0.038
2	0.485	0.085
3	0.222	0.169
4	0.109	0.487

There is a close relation between the concentration and halo mass in simulations (Navarro et al. 1996). Macciò et al. (2008) found this relation for the WMAP5 cosmology to be

$$\langle \log c_{200} \rangle (M_{200}) = 0.830 - 0.098 \log \left(h \frac{M_{200}}{10^{12} M_\odot} \right) \qquad (22)$$

(their equation 10), with a Gaussian scatter of $\sigma_{\log c_{200}} = 0.105$ (their Table A2).

4 RESULTS

4.1 Surface photometry for J1331 with MGEs

In this section we construct a model for the J1331's intrinsic light distribution in terms of MGEs (see Section 3.1). We use the HST image in the F814W (I-band) filter (Figure 1b) because J1331's central stellar component appears at longer wavelengths (i) smoother and more extended than in the F450W filter (Figure 1a), as it is less sensitive to young clumpy star-forming regions, (ii) much brighter than the bluish lensing images, and (iii) the imaging is less prone to extinction.

PSF for the HST/F814W filter. The one-dimensional MGE in Equation (3) is fitted to the radial profile of a synthetic image of the HST/F814W filter PSF, ignoring diffraction spikes and using the code by Cappellari (2002). The MGE parameters of the normalized PSF model are given in Table 2.

MGE for the inner regions. We fit a MGE to J1331's smooth central region within $\sim 5''$ from the HST/WFPC2/WF3/F814W image (Figure 1b). Bright objects close to the bulge (blobs possibly belonging to the background galaxy and parts of the foreground spiral arm) were masked during the fit. J1331's galaxy center, position angle (with respect to North through East) and average apparent ellipticity (see Table 1) are found from the images weighted first and second moment. The MGE fit splits the image in annuli with the given ellipticity and position angle and sectors of 5° width and fits an 5-Gaussian MGE of the form in Equation (1) convolved with the PSF MGE in Table 2 to it. The best fit MGE (PSF convolved) is compared to the data in Figure 2a and the corresponding parameters of the intrinsic surface brightness distribution are given in Table 3. The fit is a very good representation of the light distribution in the inner $5''$, but underestimates the light distribution outward.

MGE for the outer regions. To get a handle on the light distribution also in the outer parts of J1331, where spiral arms dominate, we first fit a IRAF (Tody 1993) *ellipse* model to the F814W image (masking the brightest blobs in the spiral arms and outer regions). Only then we fit a 7-Gaussian MGE to the smooth ellipse model. The MGE does not perfectly reproduce the flatness of the ellipse model at every radius (see Figure 2b), but considering the spiral arm dominated outer regions of J1331, it is good enough for an approximate handling of the overall light distribution.

Transformation into physical units. To transform the MGE in units of counts into physical units, we apply a simplified version of the procedure described in Holtzman et al. (1995).

The scaling of the drizzled HST/WFC3 images is $S \equiv 0.05''$/pixel width and the total exposure time $T = 1600$ sec. Each Gaussian in the MGE has a total F814W luminosity L_i (in counts) and a central surface brightness (in counts per pixel) of

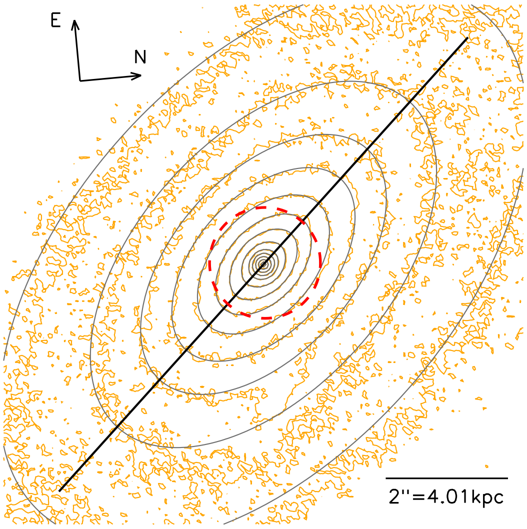
$$C_{0,i}[\text{counts/pixel}] = \frac{L_i[\text{counts}]}{2\pi\sigma[\text{pixel}]^2 q}.$$

This is then transformed into an I-band surface brightness (in mag $\times$ $(1'')^{-2}$) via

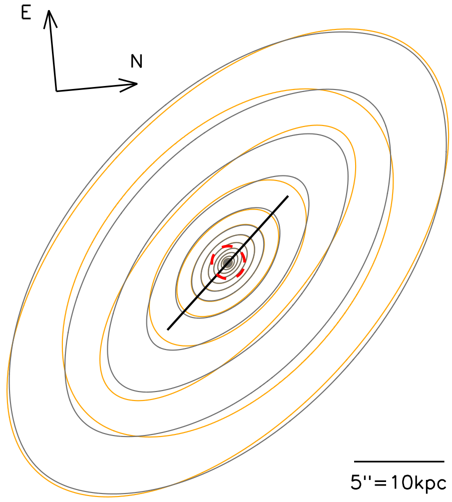
$$\mu_{I,0,i} \simeq -2.5 \log_{10} \left(\frac{C_{0,i}[\text{counts/pixel}]}{T[\text{sec}] \cdot S[''/\text{pixel width}]^2} \right) + Z + C + A_I, \qquad (23)$$

where $Z \simeq 21.62$ mag is a the zero-point from Holtzman et al. (1995), updated according to Dolphin (2000, 2008), for the photometric system of the HST/WFPC2 camera and the F814W filter, plus a correction for the difference in gain between calibration and observation. $C = 0.1$ mag corrects for the finite aperture of the WFPC2; and $A_I = 0.015$ mag is the extinction in the (Landolt) I-band towards J1331, according to the NASA/IPAC Extragalactic Database (NED)[4]. The color-dependent correction between the F814W filter and the I-band of the UBVRI photometric system is small (Holtzman et al. 1995) and we neglect it therefore. The last step is to transform the surface brightness $\mu_{I,0,i}$ (in mag) to

[4] The NASA/IPAC Extragalactic Database (NED, `https://ned.ipac.caltech.edu/`) is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The data for J1331 (SDSS J133140.33+362811.9) was retrieved in October 2013.



(a) MGE for J1331's inner regions.



(b) MGE for J1331's outer regions.

Figure 2. MGEs for J1331's surface brightness distribution. Comparison of contours with constant F814W surface brightness (orange lines) with the corresponding iso-brightness contours of the best fit MGE, convolved with the PSF in Table 3, (gray lines). The black line has a length of 10″ and it's orientation corresponds to the galaxy's position angle as found in Table 1. For comparison the Einstein radius as found in Table 5 is indicated as red dashed line. *Panel (a)* Central regions of J1331. The MGE model is a good representation of the galaxy's light distribution along the major axis within $\sim 5''$. Its parameters are given in Table 3. This MGE is used as model for the stellar tracer distribution in the dynamical Jeans modelling in Sections 4.3 and 4.4. *Panel (b)* Outer regions of J1331. The orange lines indicate here contours of a smooth IRAF *ellipse* fit to J1331 in the F814W filter; the gray lines are the corresponding best fit MGE. The green box corresponds to the image section shown in Panel (a). This MGE is not used for dynamical modelling, because the dynamics in the outer regions are strongly affected by non-axisymmetries (e.g., spiral arms). We use it, however, to estimate the galaxy's total luminosity and effective radius, and to get a prediction for the dynamics of the outer regions.

Table 3. Parameters of the best fit MGE to the F814W surface brightness of J1331 in Figure 2a. The fit is best inside an radius of 5″. The position angle is given in Table 1. This MGE is used in the dynamical modelling in Sections 4.3 and 4.4. The first column gives for each Gaussian the total F814W luminosity in Equation (2) in units of counts. The second column is the corresponding I-band peak surface brightness in Equation (1) in units of a luminosity surface density. The third and fourth column give the dispersion and the last column the axis ratio of the Gaussian in Equation (1).

i	total luminosity L_i [counts]	surface density $I_{0,i}$ [$L_\odot$/pc^2]	Gaussian dispersion σ_i [″]	σ_i [kpc]	axis ratio q'_i
1	9425.96	20768.	0.051	0.103	1.00
2	13173.0	3161.2	0.178	0.358	0.76
3	40235.0	1588.2	0.503	1.008	0.58
4	67755.2	502.25	1.180	2.368	0.56
5	203677.	136.51	3.891	7.805	0.57

the I-band surface density $I_{0,i}$ (in $L_\odot/\mathrm{pc}^2$) of the Gaussian, i.e.,

$$I_{0,i}[L_\odot \mathrm{pc}^{-2}] = (64800/\pi)^2\,(1+z_d)^4\,10^{0.4\left(M_{\odot,I}-\mu_{I,0,i}\right)},$$

where the term with z_d accounts for redshift dimming and $M_{\odot,I} = 4.08$ mag is the Sun's absolute I-band magnitude (Binney & Merrifield 1998). The luminosity L_i[counts] and the corresponding surface brightness density $I_{0,i}[L_\odot \mathrm{pc}^{-2}]$ of each Gaussian are given in Table 2a.

Inclination. To estimate the inclination of J1331 with respect to the observer, we use the observed axis ratio of the flattest ellipse in the IRAF *ellipse* model for J1331, which is $q' = 0.42$. This is similar to the disk axis ratio of $q' = 0.40$ found by Treu et al. (2011). If a typical thickness of an oblate disk is around $q_0 \sim 0.2$ (Holmberg 1958), the inclination follows from

$$\cos^2 i = \frac{q'^2 - q_0^2}{1-q_0^2}$$

and a correction of $+3°$ (Tully 1988). Our estimate for the inclination is therefore $i \approx 70°$. Given this inclination, the 2D MGE models can be deprojected into three dimensions (see Section 3.1).

Total luminosity and effective radius. J1331's total I-band luminosity is determined by summing up the luminosity contributions of all the MGE Gaussians for the outer regions (shown as gray lines in Figure 2b). We find $L_{\mathrm{tot,I}} \simeq 5.6 \cdot 10^{10}\ L_\odot$. This corresponds to an apparent magnitude of $m_I = 15.77$ mag. We determine the circularized effective radius R_{eff} of J1331 from the definition

	A	B	C	D	G
x_i [pixel]	12.1	-8.5	21.7	-3.3	$0.5 \pm\sqrt{2}$
y_i [pixel]	16.6	-10.4	-0.5	19.2	$0.5 \pm\sqrt{2}$

Table 4. Positions of the lensing images (A-D) and the galaxy center (G) in Figure 1c. The image positions were determined from the lens subtracted image for J1331 in Figure 4 of Brewer et al. (2012), rotated to the (x, y) coordinate system in Figure 1c. The pixel scale is 1 pixel $= 0.05''$ and the error of each image position is $\pm$ 1 pixel.

$L(< R_{\rm eff}) \equiv \frac{1}{2}L_{\rm tot}$ and the growth curve $L(< R')$ from the MGE model of the outer regions, where R' is the projected radius on the sky. We find the effective radius to be $R_{\rm eff} \simeq 2.6'' \hat{=} 5.2$ kpc. All values are summarized in Table 1.

4.2 Mass distribution from lensing

In this section we use the gravitational lensing formalism summarized in Section 3.2 to fit a scale-free galaxy mass model to the positions of the lensing images observed in J1331's central region.

Image positions. We determine the positions of the lensing images by first subtracting a smooth model for the galaxy's surface brightness from the original image. As models we use MGE fits and IRAF *ellipse* fits to J1331 in each the F450W and F814W filter. The lensing images become visible in the residuals (see Figure 1c). Because the images are extended, we use the position of the brightest pixel in each of the images. In addition, we consider the F450W-MGE subtracted residuals from Brewer et al. (2012). The lensing positions, as determined from the latter, are given in Table 4. The scatter of lensing positions, as determined from subtracting different brightness models from the galaxy in different filters, gives an error of ± 1 pixel on the image positions. To the galaxy center, which we assume to be the surface brightness peak in the F450W image, we apply an error of $\pm\sqrt{2}$ pixel.

Eight image position coordinates allow us to fit a lens mass model with only < 8 free parameters. We therefore do not fit Fourier components (a_k, b_k) with $k > 3$ in the lens mass model in Equations (9)-(10).

Even though the constraint from the image positions on α is very weak, we were however able to show that the image positions in Table 4 are consistent with a model with flat rotation curve. In the following we therefore set $\alpha = 1$.

Best fit lens model. We fit the lens mass model to the image positions in Table 4 by minimizing $\chi^2 = \chi^2_{\rm lens} + \chi^2_{\rm shape}$ (see Equations (12)-(14)). The best fit parameters are given in the first column of Table 5. Figure 3 shows the corresponding critical curve, caustic and Einstein radius, and the best fit source position. In this case, where $\alpha = 1$, the (tangential) critical curve is also an equidensity contour of the galaxy model (Evans & Witt 2003), which appears to have an elliptical mass distribution. The source is located close to a cusp of the diamond-shaped caustic: a lensing configuration for which we indeed expect four images. Figure 3 also shows the (smoothed) residuals from the F450W image subtracted by an IRAF *ellipse* brightness model and the contours of the best fit model's time delay surface. This demonstrates that, although we did not include any information about the shape of the lensing images in the fit, it is consistent with the predicted distortion for an extended source.

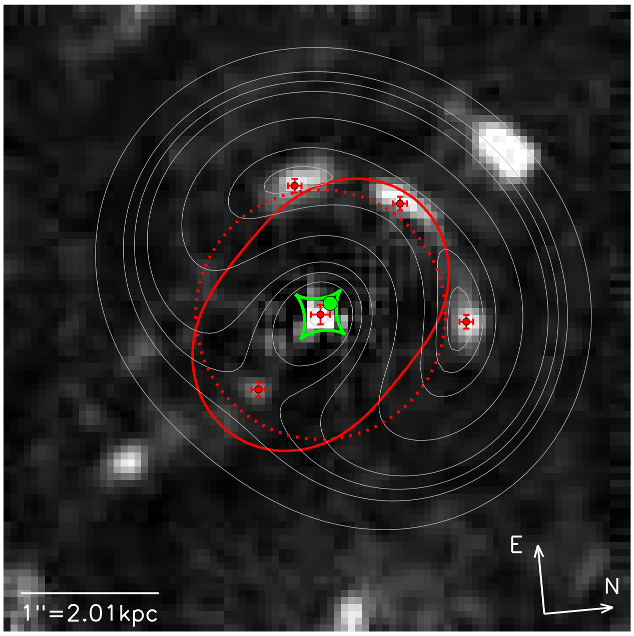
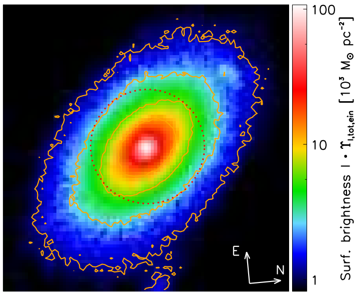


Figure 3. Lensing model (Table 5) found as best fit to the image positions. In the background we show the central region of J1331 in the F450W filter, subtracted by an IRAF *ellipse* model of the F450W surface brightness and smoothed to remove noise smaller than the PSF. The brightness peaks of the four lensing images and the galaxy center (Table 4) are marked as red dots. For the best fit lens model we show the Einstein radius (red dotted circle) and the critical curve (red solid line), which are located in the lens plane. We also show the caustic (green solid line) corresponding to the critical curve and best fit source position (green dot), which are located in the source plane. For $\alpha = 1$ the critical curve is a contour of constant surface density of the mass model. The grey lines show contours of the time delay surface given by Equation (7). Not only the position of the extrema, but also their shape is consistent with the observed, extended images, even though we did not use information about the image shape in the analysis.

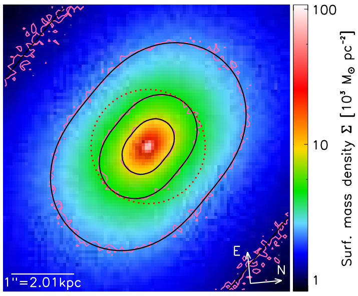
To estimate how the uncertainties in the determination of image positions and galaxy center affect the results, we sample random positions from two-dimensional Normal distributions with peaks and standard deviations according to Table 4. Model fits to many sampled image positions lead to probability distributions for the best fit shape parameters and Einstein quantities; peak and standard deviations are given in the second column of Table 5. We constrain the Einstein radius to within 2%, $R_{\rm ein} = (0.91 \pm 0.02)''$ and the projected mass within the critical curve with a relative error of 4%, $M_{\rm crit} = (7.9 \pm 0.3) \cdot 10^{10} M_\odot$. Our measurement of $R_{\rm ein}$ is consistent with that from Brewer et al. (2012), $R_{\rm ein,SWELLS} = (0.96 \pm 0.04)''$ (which used a singular isothermal ellipsoid as lens mass model and the intermediate-axis convention of the critical curve as the Einstein radius). The relative difference between our critical mass and that of

Table 5. Best fit lens model found from the peak image positions in Table 4 (first column) following the procedure described in Section 3.2 and assuming a flat rotation curve ($\alpha = 1$). The second column gives the corresponding best fit mean and standard deviation derived from Monte Carlo sampling of the Gaussian uncertainties around the image positions (column 2).

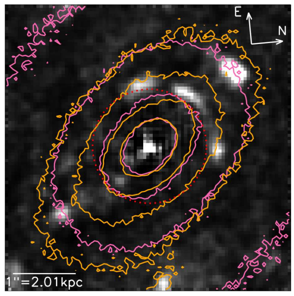
		lens model for peak image positions	lens model from Monte Carlo sampling of image positions			
Einstein radius	$R_{\rm ein}$ [$''$]	0.907	0.91	$\pm$	0.02	(2%)
Einstein mass	$M_{\rm ein}$ [$10^{10} M_\odot$]	7.72	7.8	$\pm$	0.3	(4%)
Critical mass	$M_{\rm crit}$ [$10^{10} M_\odot$]	7.87	7.9	$\pm$	0.3	(4%)
Source position	ξ [$''$]	0.095	0.09	$\pm$	0.03	(28%)
	η [$''$]	0.107	0.10	$\pm$	0.03	(27%)
Fourier coefficients	a_0	1.814	1.82	$\pm$	0.04	(2%)
	a_2	0.012	0.011	$\pm$	0.004	(35%)
	b_2	-0.057	-0.06	$\pm$	0.01	(25%)
	a_3	-0.0001	0.0000	$\pm$	0.0006	
	b_3	-0.0002	0.000	$\pm$	0.001	



(a) Observed light distribution.



(b) Predicted mass distribution.



(c) Comparison of mass and light.

Figure 4. Comparison of the observed F814W/I-band surface brightness distribution (Panel (a) and orange contours) and predicted mass distribution from lensing constraints (Panel (b) and pink contours). To allow for a qualitative comparison of the contours in Panel (c), the light distribution was turned into a mass distribution by multiplication with the total mass-to-light ratio inside the Einstein radius $\Upsilon^{\rm ein}_{\rm I,tot} = 5.56\Upsilon_{I,\odot}$. The Einstein radius is overplotted as red dotted circle. The uncertainties in the mass model in the second column of Table 5 were translated into random Monte Carlo noise in the mass contours. The smooth black contours correspond to the best fit model in the first column of Table 5. The background in Panel (c) shows again the surface brightness subtracted center of the galaxy to make the lensing images visible.

Brewer et al. (2012), $M_{\rm crit,SWELLS} = (8.86 \pm 0.61) \cdot 10^{10} M_\odot$, is 13%.

Comparison with light distribution. The surface mass distribution as predicted by the best fit lens model (Table 5) is shown in Figure 4b. We visualize the effect of the Fourier shape parameter uncertainties by introducing random noise to create a mock observation. From the mock image's second moment we find an average axis ratio for the lens mass model of $q_{\rm lens} \simeq 0.695$, which is consistent with the one found by Brewer et al. (2012), $q_{\rm lens,SWELLSIII} = 0.67 \pm 0.09$, while the light's average axis ratio is $q' = 0.598$ (see Table 1).

We estimate the total mass-to-light ratio within the Einstein radius $\Upsilon^{\rm ein}_{\rm I,tot} \equiv M_{\rm ein}/L_{\rm I,ein}$. For this, we integrate the MGE in Table 3 to get the total I-band luminosity within the Einstein radius $L_{\rm I,ein} = 1.40 \cdot 10^{10} L_\odot$. The corresponding Einstein mass-to-light ratio is therefore $\Upsilon^{\rm ein}_{\rm I,tot} = 5.56\Upsilon_{I,\odot}$. This is consistent with or slightly larger than the stellar mass-to-light ratio assuming a Salpeter initial mass function (IMF) $\Upsilon^{\rm sal}_{\rm I,*} = 4.7 \pm 1.2$ according to Treu et al. (2011) and Table 7 (see also discussion in Section 5.1).

We use $\Upsilon^{\rm ein}_{\rm I,tot}$ to transform the observed surface brightness in the F814W filter into a surface mass density (Figure 4a). Figure 4c then compares equidensity contours of both the predicted lens mass distribution and the observed surface brightness times $\Upsilon^{\rm ein}_{\rm I,tot}$.

Figure 4 leads to the following three findings: (1)The mass predicted from lensing and the observed light distribution are oriented in the same direction (i.e., have the same position angle). (2) Within and around the Einstein radius, mass and light distribution have a similar elliptical shape, while further out the mass distribution is slightly rounder. (3) The light distribution drops faster than the mass with increasing radius (which is—at least partly—because of the assumption of a flat rotation curve). However, the mass distribution is only constraint around the Einstein radius and otherwise an extrapolation.

4.3 JAM based on surface brightness

In this section we create dynamical models for J1331 following the procedure in Section 3.3. We use the deprojected surface brightness MGE from Table 3 for the tracer distribution $\nu(R, z)$ and to generate mass models $\rho(R, z)$. The only

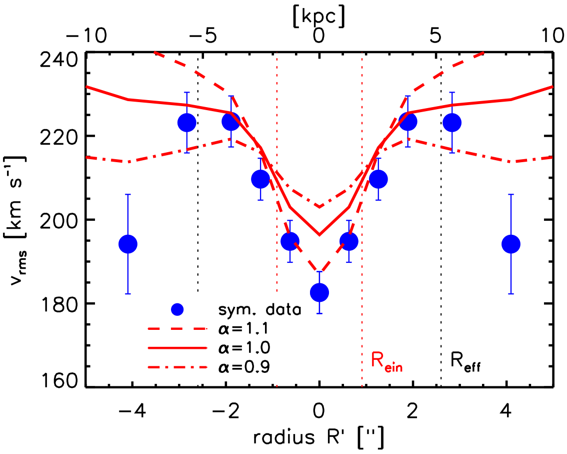


Figure 5. Comparison (not a fit!) of the symmetrized stellar $v_{\rm rms}$ data of J1331 (blue dots) with JAM models generated from mass distributions which were independently derived from lensing constraints in Section 4.2 (red lines). The red solid line corresponds to the lens model for a flat rotation curve ($\alpha = 1$) in Table 5; the red dashed and dash-dotted lines are best fit lens models found analogously from the image positions, but for a fixed rotation curve slope of $\alpha = 1.1$ and 0.9, respectively. For the JAM modelling a best fit MGE to the lens mass models were used, as well as the observed surface brightness MGE in Table 3, assuming velocity isotropy $\beta_z = 0$ and an inclination of $i = 70°$. The most reliable constraints are around the Einstein radius (vertical red dotted line); outside of it the JAM model is just an extrapolation.

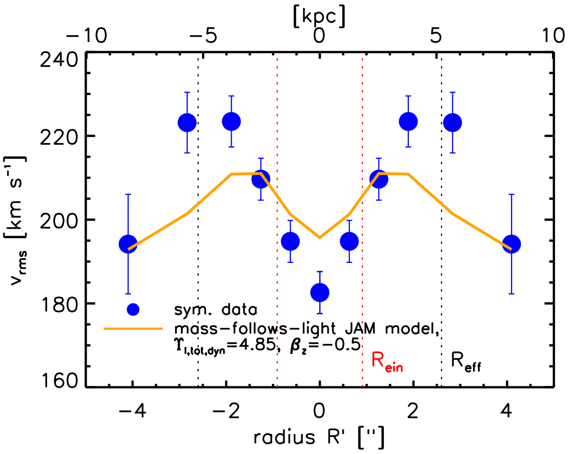


Figure 6. Comparison of the symmetrized $v_{\rm rms}$ data of J1331 (blue dots) with a best fit dynamical JAM model (orange line) assuming mass-follows-light and with two free parameters: $\Upsilon^{\rm dyn}_{\rm I,tot}$, the total I-band mass-to-light ratio found from dynamics, which converts the observed surface brightness in Table 3 into a mass distribution, and the velocity anisotropy parameter β_z. The "best" fit is $\Upsilon^{\rm dyn}_{\rm I,tot} = 4.8 \pm 0.1$ and $\beta_z = -0.5$, where the latter is however pegged at the lower limit of the allowed value range. This is obviously not a good model.

exception is the first test, where the mass model comes from lensing constraints.

JAM with lens mass model. We make an independent prediction for the $v_{\rm rms}$ curve by evaluating Equation 18 (with $\beta_z = 0$) for the lens mass model in Table 5 ($\alpha = 1$, flat rotation curve). In addition, we also calculate a $v_{\rm rms}$ curve for two lens models which were found analogously, but assumed a slightly rising (falling) rotation curve slope of $\alpha = 1.1$ ($\alpha = 0.9$). The predictions are compared with the data in Figure 5. While the most reliable constraint is around $R_{\rm ein}$, the agreement within $R' \sim 3''$ is still striking: The $\alpha > 1$ model recreates the observed central dip, while the $\alpha = 1$ model fits the wings around $R_{\rm eff}$. This is in concordance with observations in other galaxies. For $R' > R_{\rm eff}$ we would expect $\alpha < 1$—and the lensing model for $\alpha = 0.9$ has indeed a slightly dropping $v_{\rm rms}$ around $R_{\rm eff}$, like the data. A definite comparison in this regime is however difficult as the lens models are just extrapolations outside of $R_{\rm ein}$. While our lensing model assumes $\alpha(R') = $ const, Figure 5 suggests, that a model with variable $\alpha(R')$ could fit even better. Overall the lensing model is in very good agreement with the $v_{\rm rms}$ data within $R' \sim 3''$—even though it was derived completely independently.

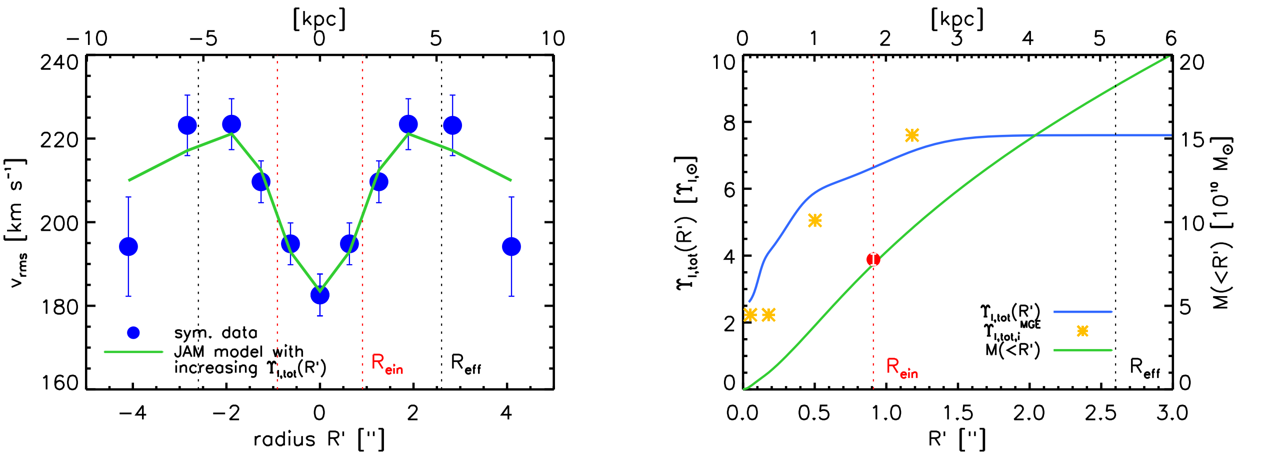
JAM with "mass-follows-light" and velocity anisotropy. The first JAM model that we fit to the observed $v_{\rm rms}$ within $R' < 5''$ is a mass-follows-light model. Mass-follows-light models are often used in dynamical JAM modelling (e.g., van de Ven et al. 2010; Cappellari et al. 2006) and generate a mass distribution by multiplying the intrinsic light distribution $\nu(R, z)$ by a constant total mass-to-light ratio $\Upsilon^{\rm dyn}_{\rm I,tot}$. This assumes that the DM is always a constant fraction of the total matter distribution within the region covered by the kinematics. This simplified mass model sometimes gives good representations of the inner parts of galaxies, where the stellar component dominates.

We also allow for a overall constant but non-zero velocity anisotropy β_z in the model. The model parameters $(\Upsilon^{\rm dyn}_{\rm I,tot}, \beta_z)$ that fit the $v_{\rm rms}$ data best are found using a χ^2-fit and are demonstrated in Figure 6.

For β_z we imposed the fitting limits $\beta_z \in [-0.5, +0.5]$. While the outer parts of galaxies often show radially biased velocity anisotropy up to ~ 0.5 (from dynamical modelling of observed elliptical galaxies, e.g., Kronawitter et al. (2000)) and cosmological simulations (e.g., Diemand et al. 2004; Fukushige & Makino 2001), the centers of galaxies are near-isotropic or have negative velocity anisotropy (Gebhardt et al. 2003). Only in extreme models (e.g., around in-spiralling supermassive black holes, e.g., Quinlan & Hernquist 1997) velocity anisotropies as low as ~ -1 have been found.

The best fit in Figure 6 however strives to very negative β_z to be able to reproduce the deep central dip in the $v_{\rm rms}$ curve.[5] But $\beta_z = -0.5$ is not even a remotely agreeable fit and lower anisotropies are not to be expected or realistic. We also tested radial profiles for $\beta_z(R)$ of the form proposed by Baes & van Hese (2007), which was however equally unable to reproduce the data. We conclude, that this is due to the well-known degeneracy between anisotropy and mass profile [TO DO: REF] and the mass-follows-light model is *not* a good representation of the mass distribution in J1331's inner regions.

[5] Without limiting the fitting range, the best fit would be a unrealistically low $\beta_z \sim -2$.



(a) Comparison of $v_{\rm rms}$ data and best fit model.

(b) Projected local mass-to-light ratio profile and enclosed mass.

Figure 7. JAM model found by fitting an increasing total mass-to-light ratio $\Upsilon_{\rm I,tot}(R')$ profile used to generate a mass model from the light distribution. This is done by assigning a different mass-to-light ratio to each Gaussian in the MGE in Table 3. *Panel (a):* Comparison between the stellar $v_{\rm rms}$ data (blue points) and the best fit model (green line). *Panel (b):* Projected mass-to-light profile $\Upsilon_{\rm I,tot}(R')$ along the major axis (blue line, left axis) of the best fit model. The best fit mass-to-light ratios of the first four Gaussians are plotted against each Gaussians σ (yellow stars). The two Gaussians with the largest σ (the fifth is not shown) have the same best fit mass-to-light ratio. Shown is also the enclosed mass inside the projected radius R' on the sky (green line, right axis). The enclosed mass curve is overplotted with the independent finding for the Einstein mass $\pm 4\%$ in Table 5 (red dot) at the Einstein radius (red dotted line). Shown is also the effective half-light ratio $R_{\rm eff}$ (black dotted line). [TO DO: Mention second axis only after 1st axis is explained.]

[TO DO: Make sure that we explain somewhere, why we model only in the inner regions]

JAM with increasing mass-to-light ratio. In Section 4.2 we found from lensing constraints, that the light distribution might drop faster with radius than the mass distribution. This could correspond to a radially increasing total mass-to-light ratio. As velocity anisotropy alone cannot explain the observed kinematics in a simple a mass-follows-light model, we now allow for a mass-to-light ratio gradient in the JAM modelling. We therefore generate a mass model from the light distribution in Table 3 by assigning each of the five Gaussians in the MGE its own total mass-to-light ratio $\Upsilon_{{\rm I,tot},i}$ and replace the total luminosity in Equation (2) L_i with the Gaussians total Mass $M_i = \Upsilon_{{\rm I,tot},i} L_i$. We treat the five $\Upsilon_{{\rm I,tot},i}$ as free fit parameters and only require that $\Upsilon_{{\rm I,tot},j} \geqslant \Upsilon_{{\rm I,tot},i}$ when the corresponding $\sigma_j \geqslant \sigma_i$ to ensure that the overall mass-to-light ratio is increasing with radius.

Figure 7b shows the best fit (projected local) mass-to-light ratio profile, which rises from $\Upsilon_{\rm I,tot} = 2.53$ in the center and approaches a value of $\Upsilon_{\rm I,tot} = 7.60$ outside of the fitted region at $R' \gtrsim 3''$. The central mass-to-light ratio is in agreement with the mass-to-light ratio $\Upsilon_{\rm I,*}^{\rm chab} = 2.5 \pm 0.6$, given in Table 7 based on the results of Treu et al. (2011) assuming a stellar population with a Chabrier (2003) IMF (see also Section 5.1). When assuming that galaxy bulges are in general older and redder in the center [TO DO: REF], i.e., $\Upsilon_{\rm I,*}$ is more likely to drop with radius than to rise, the strong increase of $\Upsilon_{\rm I,tot}(R')$ might be due to a strong contribution of dark matter in J1331.

Figure 7a shows that the best fit model nicely reproduces the central dip in the $v_{\rm rms}$ curve, even though it has difficulties fitting the drop around $R' \sim 4''$. The latter might be because we only allowed the $\Upsilon_{\rm I,tot}(R')$ to rise. A slight drop could be expected when the reddish bulge turns into the bluish disk and the contribution of the stellar component becomes less due to a lower $\Upsilon_{\rm I,*}$ for younger and bluer populations.

In Figure 7b we overplot the enclosed mass profile with the Einstein mass $M_{\rm ein} = (7.77 \pm 0.33) \cdot 10^{10} M_\odot$ at the Einstein radius found from lensing in Table 5. The agreement between the Einstein mass and the independently found $M(< R_{\rm ein}) = 7.49 \cdot 10^{10} M_\odot$ from dynamical modelling is striking.

4.4 JAM with a NFW dark matter halo

The modelling attempts in the previous sections suggest that J1331's mass distribution could be more roundish in the inner regions and more massive at larger radii than expected from the distribution of stars alone. A DM halo in addition to the stellar component could explain these findings.

Including a DM halo. We therefore proceed by modelling the mass distribution with (a) a stellar component, which we get from the light MGE in Table 3 (deprojected to the intrinsic $\nu(R, z)$) times a constant stellar mass-to-light ratio $\Upsilon_{\rm I,*}$, and (b) a spherical NFW DM component (see Section 3.3) with halo scale length r_s and circular velocity at the virial radius v_{200} as free parameters. In the JAM modelling we use a 10-Gaussian MGE fit to the classical NFW profile in Equation (20).

Modelling and priors. The full set of fit parameters is $(\Upsilon_{\rm I,*}, r_s, v_{200}, \beta_z)$, where β_z is again the constant velocity anisotropy parameter. We will investigate this parameter

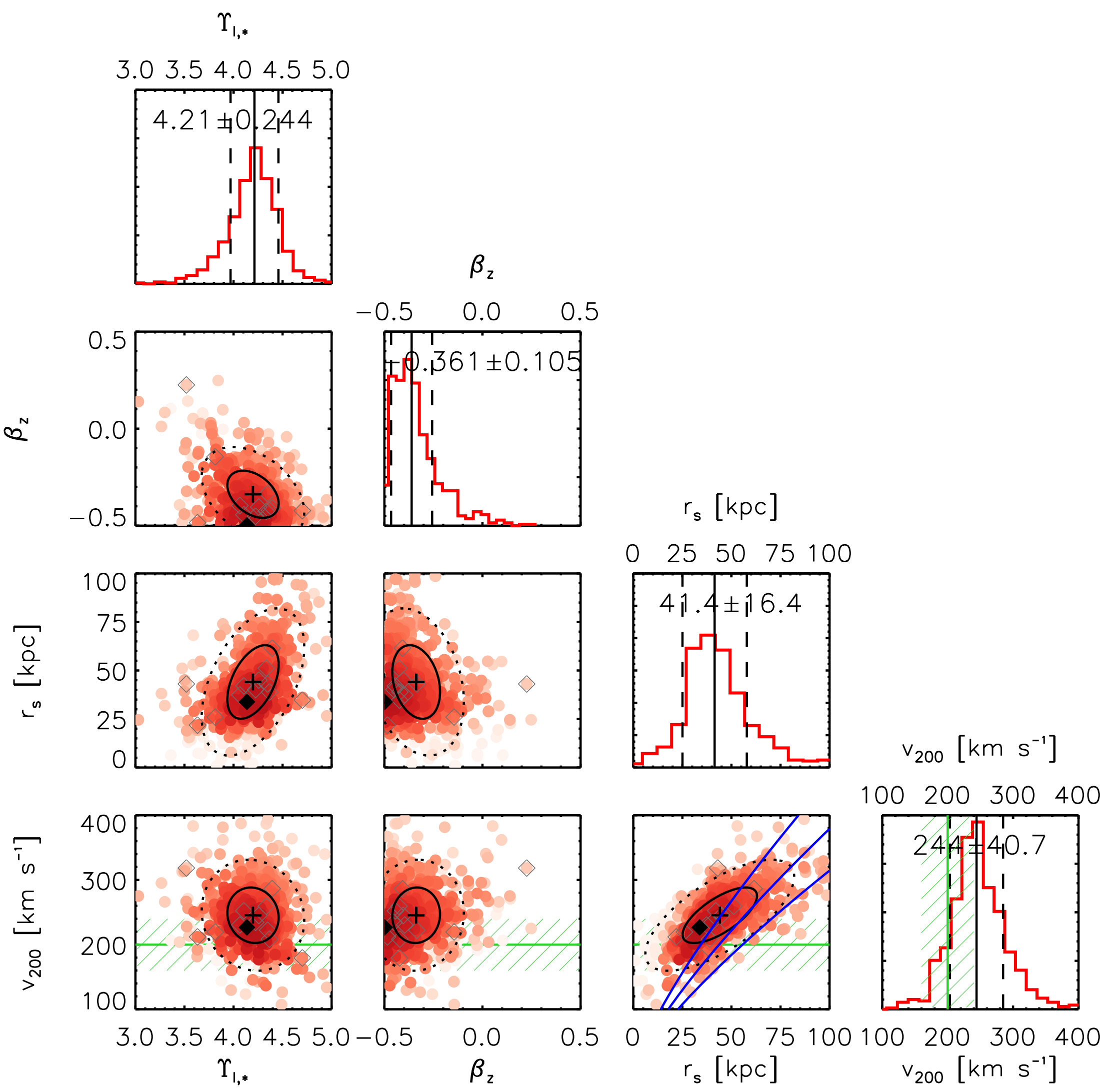


Figure 8. Posterior probability distribution sampled with MCMC (red dots and histograms) for a JAM model with NFW halo, parametrized by r_s and v_{200}, a stellar mass distribution generated from the I-band MGE in Table 3 and a constant stellar mass-to-light ratio $\Upsilon_{\mathrm{I},*}$, and constant velocity anisotropy β_z. Shown are also the priors used for J1331's NFW halo, $\mathcal{N}(200\ \mathrm{km\ s}^{-1}, 40\ \mathrm{km\ s}^{-1})$ (green) and the concentration vs. halo mass relation by Macciò et al. (2008) from Equation (22) in terms of v_{200} vs. r_s with 1σ scatter (blue). The MCMC samples are color coded according to their probability (darker red for higher probability); the sample point with the highest probability is marked by a black diamond. The black cross is the mean of the distribution and the ellipses are derived from the covariance of matrix of the sample set and correspond approximately to 1σ (black solid ellipse) and 2σ (black dotted ellipse). The histograms of the marginalized 1D distributions are overplotted by the mean (black solid lines) and 1σ error (black dashed lines), whose values are also quoted in the figure and in Table 6. The grey diamonds mark a random sub-selection of 12 samples; the corresponding models are shown in Figure 9. [TO DO: Maybe mention also: There are slight covariances between $\Upsilon_{\mathrm{I},*}$, r_s and β_z: The smaller the DM contribution in the center (i.e., the larger $\Upsilon_{\mathrm{I},*}$), the less concentrated is the halo (i.e., the larger r_s), the more velocity anisotropy is needed to reproduce the central dip. As the effect of v_{200} is mostly at larger radii, this parameter doesn't show any covariances, but is also mostly constrained by the prior.]

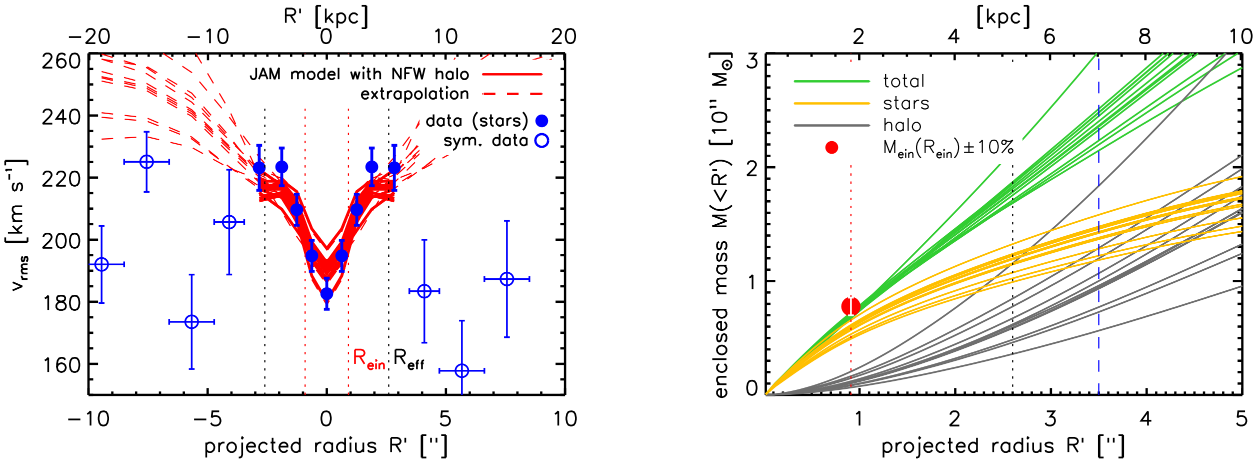
space with a MCMC[6] (Foreman-Mackey et al. 2013) and use priors for the halo parameters to guide the fit to a realistic NFW halo shape.

Dutton et al. (2010) give a relation for halo vs. stellar mass for late-type galaxies. Using the stellar mass estimate for J1331 from Treu et al. (2011), $m_* = (1.06 \pm 0.25) \cdot 10^{11} M_\odot$ for the Chabrier IMF estimate, we find

[6] The Python code package for *emcee*, a Monte Carlo Markov Chain implementation by Foreman-Mackey et al. (2013) is available online at `http://dan.iel.fm/emcee/current/`. The version from October 2013 was used in this work.

stellar I-band mass-to-light ratio	$\Upsilon_{\mathrm{I},*}$	4.2	$\pm$	0.2
velocity anisotropy	β_z	-0.4	$\pm$	0.1
NFW halo scale length	r_s [kpc]	40	$\pm$	20
NFW halo virial velocity	v_{200} [km s^{-1}]	240	$\pm$	40
NFW halo concentration	c_{200}	8	$\pm$	2
NFW halo mass	M_{200} [$10^{12}M_\odot$]	5	$\pm$	2

Table 6. Summary of the best fit parameters of the JAM model with NFW halo from the MCMC exploration in Figure 8. The halo mass and concentration are calculated from the the best fit r_s and v_{200}.



(a) Comparison of the v_{rms} data with the best fit JAM models.

(b) Enclosed mass profile.

Figure 9. [TO DO: Rewrite caption] Best fit JAM model including a NFW halo and velocity anisotropy with parameters given in Table 6. The 12 lines shown correspond to the 12 models randomly drawn from the posterior probability distribution and marked as grey diamonds in Figure 8. Overplotted are the Einstein radius (red dotted line) and the effective half-light radius (black dotted line). The blue dashed line marks the radius within which the data and model were fitted. *Panel a)* shows the comparison of the symmetrized v_{rms} data (solid blue points) with the best fit JAM models including a NFW halo (red solid lines). Also shown is the non-symmetrized data at larger radii (open blue dots) and an extrapolation of the best fit models, using the same model parameters but the I-band surface brightness MGE for the outer regions of J1331 derived from the Ellipse model (red dashed lines). *Panel b)* shows the circular velocity curve of the total mass (green), and separately the contribution of the stellar mass (yellow, again generated from the MGE in Table 3) and DM (grey). *Panel c)* shows the corresponding enclosed mass profile. Overplotted is also the Einstein mass at the Einstein radius with a 10% error, which was used in the fit as an additional constraint.

$v_{200} = (202^{+44}_{-33})^{+12}_{-13}$. The first error is due to the 2σ scatter in the relation by Dutton et al. (2010). The second error is the propagated error due to the uncertainty in the stellar mass. We use this as a rough estimate for the halo of J1331 and as Gaussian prior on v_{200},

$$p(v_{200}) = \mathcal{N}(200\ \mathrm{km\ s}^{-1}, 40\ \mathrm{km\ s}^{-1}).$$

We also use the concentration vs. halo mass relation by Macciò et al. (2008) in Equation (22) as a prior on the concentration, i.e.

$$p(\log c_{200} \mid v_{200}) = \mathcal{N}\left(\langle \log c_{200}\rangle (M_{200}) \mid 0.105\right).$$

For the velocity anisotropy parameter β_z we will again employ a uniform prior

$$p(\beta) = \mathcal{U}(-0.5, +0.5)$$

to exclude very unrealistic anisotropies. The full prior used is then

$$p(\Upsilon_{\mathrm{I},*}, r_s, v_{200}, \beta) = \frac{1}{\ln(10 r_s)} p(\log c_{200} \mid v_{200}) \cdot p(v_{200}) \cdot p(\beta),$$

where the factor $1/\ln(10r_s)$ is the Jacobian of the transformation from the halo parameters $(v_{200}, \log c_{200})$ to (v_{200}, r_s).

We restrict the v_{rms} fit to a region $R' \lesssim 3.5''$, approximately within the effective half-light radius $R_{\mathrm{eff}} = 2.6''$. We also include the Einstein mass $M(< R_{\mathrm{ein}})$ with a 10% error as an additional constraint in the fit.

Result. Figure 8 shows the posterior probability distribution (pdf) of the fit sampled with an MCMC. Overplotted are also the priors used to constrain the NFW halo. The mean parameters are summarized in Table 6. We find that the best fit NFW halo strives to be more massive and with a higher concentration (due to a smaller r_s) than proposed by the priors. The model also prefers very negative velocity anisotropies. Both, the high halo concentration and low β_z, are needed to reproduce the central dip of the v_{rms} curve. Figure 9 illustrates the range of best fit models according to the extent of the pdf. The models fit the v_{rms} data in the inner regions quite well (Figure 9a) and are also consistent with the Einstein mass (see Figure 9b). The extrapolation of the model to larger radii however overestimates the data, does not exhibit a drop around $\sim 6''$ at all and seems to be therefore overall too massive.

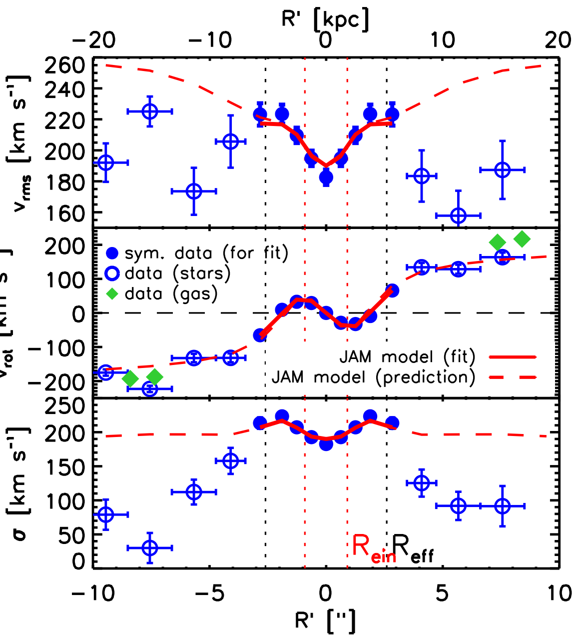


Figure 10. Generating the rotation curve from the JAM model with NFW halo and constant velocity anisotropy with the mean parameters in Table 6 and with the best fit rotation parameter $\kappa' = 0.76$ in Equation (19) (red solid lines). The second velocity moment in the first panel and the first velocity moment in the second panel with the additional fit parameter κ' were fitted to the symmetrized data (solid blue points). The velocity dispersion is simply $\sigma = \sqrt{v_{\rm rms}^2 - v_{\rm rot}^2}$. At larger radii we compare the unsymmetrized data (open blue dots), the gas kinematics from Dutton et al. (2013) (green diamonds) and a JAM model using the same model parameters but the light distribution MGE generated from the Ellipse model in Figure 2b as a prediction for the outer regions of J1331 (red dashed lines). The central regions are very well reproduced and we can also nicely predict the rotation curve at larger radii. Only at larger radii the $v_{\rm rms}$ and $v_{\rm rot}$ overestimate the measurements, probably due to a too massive NFW halo.

We also fitted a model with a cored logarithmic DM halo. The cored halo models are in general slightly less massive than the NFW halo and therefore fit the outer regions of J1331 better. However, the density profile of the cored halo as well as the I-band light distribution within the plane drop as $\rho(r) \propto r^{-2}$. There is therefore a strong degeneracy between the stellar mass and the DM. Overall, we were not able to obtain tight constraints on the cored logarithmic halo.

Rotation curve. We generate a rotation curve from the best fit mean model in Table 6, whose $v_{\rm rms}$ curve is shown in the first panel of Figure 10. Following the procedure in Section 3.3, we find the rotation curve by fitting the rotation parameter κ' to the symmetrized $v_{\rm rot}$ data within $R' = 3.5''$. The best fit with $\kappa' = 0.76$ is given in the second panel of Figure 10. The third panel shows the dispersion that follows from $\sigma = \sqrt{v_{\rm rms}^2 - v_{\rm rot}^2}$. Our assumptions for $\kappa(R)$ nicely reproduce a $v_{\rm rot}$ model with counter-rotating core. Although we fitted $v_{\rm rot}$ only to the inner regions, the extrapolation to large radii fits the data also very well.

While the dispersion σ in the center fits by construction quite good, the predicted dispersion is much larger than the data. We would expect the disk rotationally supported and therefore have a low velocity dispersion; especially dispersions as high as ~ 200 km s^{-1} are more likely to be observed in the pressure supported bulges of galaxies. There might be something unexpected with the σ measurements around $\sim 5''$, but at large radii the the best fit model NFW halo is simply too massive.

5 DISCUSSION AND CONCLUSION

We've presented different dynamical models for the central region of J1331. Some of them capture the observed kinematics, but none of them work at both small and large radii. In the following we discuss possible reasons, also by comparing our results to previous work.

5.1 On J1331's central stellar mass-to-light ratio

Some of the ambiguities in recovering J1331's matter distribution could be resolved by learning more about stellar populations with different IMFs in J1331. In particular, a sophisticated guess for the stellar mass-to-light ratio in the bulge could be compared to our very reliable measurement of the total mass-to-light ratio inside the Einstein radius $\Upsilon_{\rm I,tot}^{\rm ein} = 5.56\Upsilon_{I,\odot}$. This would then either support or contradict the presence of a significant amount of DM in the bulge.

Traditional choices for the IMF are the bottom-heavy IMF by Salpeter (1955),

$$\xi(m) \propto m^{-x}, x = 2.35,$$

where $\xi(m)\,{\rm d}m$ is the number of stars with mass m in $[m, m + {\rm d}m]$, and the IMFs by Kroupa (2002) and Chabrier (2003), which are in agreement with each other and predict less low-mass stars.

Ferreras et al. (2013) found a relation between the central stellar velocity dispersion σ_0 in early-type galaxies and the IMF slope x; a higher σ_0 suggests a more bottom-heavy IMF. For a unimodal (Salpeter-like) IMF and $\sigma_0 \simeq 200$ km s^{-1} in J1331 (see Figure 1d) they predict $x \approx 2.33$, which is close to the standard Salpeter slope, also supported by Spiniello et al. (2014). When assuming a bi-modal (Kroupa-equivalent-like) IMF, Ferreras et al. (2013) predict $x \approx 2.85$ for J1331's central velocity dispersion. This is more bottom-heavy than the standard Kroupa (2002) IMF. Overall, the central velocity dispersion suggests a rather bottom-heavy IMF in J1331's bulge and therefore large stellar mass-to-light ratio.

Treu et al. (2011) estimated J1331's stellar bulge mass given a Salpeter IMF and measured the I-band AB magnitude of the bulge. Transformed to a stellar I-band mass-to-light ratio, their results would correspond to $\Upsilon_{\rm I,*}^{\rm sal} = 4.7 \pm 1.2$ (see Table 7). This is not too far from $\Upsilon_{\rm I,*} = 4.2 \pm 0.2$ (see Table 6), which we found when including a NFW halo in the JAM modelling.

When Treu et al. (2011) assumed a Chabrier IMF, their result translates to $\Upsilon_{\rm I,*}^{\rm chab} = 2.5 \pm 0.6$ (see Table 7). In Section 4.3 we created a dynamical model from only the surface brightness distribution and an increasing mass-to-light ratio profile without additional DM halo. We found that such a

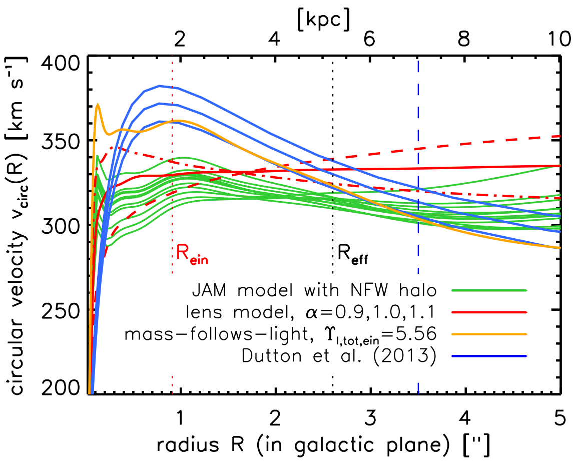


Figure 11. Comparison of the circular velocity curve of J1331's inner regions for different models: (a) JAM model with NFW DM halo from Section 4.4 and Figure 9 (green). (b) Lens model from Section 4.2 and Figure 5 with $\alpha = 1.1$ (red dashed line), $\alpha = 1$ (red solid line) and $\alpha = 0.9$ (red dash-dotted line). (c) Mass-follows-light model, which uses the F814W surface brightness in Table 3 and the mass-to-light ratio in the Einstein radius, $\Upsilon_{\mathrm{I,tot}}^{\mathrm{ein}} = 5.56$, to generate a mass distribution, as in Figure 4a (orange line). (d) Model from gas kinematics and Einstein mass found by Dutton et al. (2013) (their Figure 2, best model with 68% confidence region) (blue lines).

model would be perfectly consistent with the Einstein mass, predict a total $\Upsilon_{\mathrm{I,tot}}(R' \sim 0) = 2.53$—being consistent with the Chabrier IMF estimate by Treu et al. (2011)—and rise quickly to $\Upsilon_{\mathrm{I,tot}}(R' \gtrsim R_{\mathrm{ein}}) \gtrsim 6$.

We also compare our results from Section 4.4 with the study by Dutton et al. (2013). They found that the bulge of J1331 has an IMF *more* bottom-heavy than the Salpeter IMF from fitting a NFW halo to (1) the Einstein mass and (2) gas kinematics at larger radii $\gtrsim 8''$. In Section we fitted a mass model with NFW halo to (1) the Einstein mass and (2) stellar kinematics within $\sim 3.5''$. Our best fit $\Upsilon_{\mathrm{I},*} = 4.2 \pm 0.2$ (see Table 6) indicates a *less* bottom-heavy IMF than the Salpeter IMF. Dutton et al. (2013) found systematically lower NFW halo masses ($v_{\mathrm{circ,halo}}(5'') \sim 120$ km s^{-1} according to their Figure 2) than we did ($v_{\mathrm{circ,halo}}(5'') \sim 200$ km s^{-1}).

5.2 On J1331's central kinematics

In Figure 11 we compare the circular velocity curve found by Dutton et al. (2013) with a mass-follows-light model scaled to fit our Einstein mass (by multiplying the light distribution in Table 3 with $\Upsilon_{\mathrm{I,tot}}^{\mathrm{ein}} = 5.56$). Within $R < 5''$ they agree with each other.

The models in this work used more than just the Einstein mass to constrain the matter distribution at small radii: The lens mass model constrained also the shape of the mass distribution within the lensing image configuration at $R_{\mathrm{ein}} \sim 1''$. The dynamical models used stellar kinematics inside $R' \simeq 3.5''$ [TO DO: Check]. We compare the lens mass model's v_{circ} (for $\alpha = 1.0 \pm 0.1$) with the NFW JAM model (Table 6) in Figure 11 as well. Within and around R_{ein} they are consistent with each other, even though they where independently derived. They do not agree with the result by Dutton et al. (2013) and in Section 4.3 we showed, that "mass-follows-light" is not a good model for J1331.

Overall, we were not able to find a consistent explanation for J1331's central v_{rms} dip. It cannot be due to tangential velocity anisotropy in the center alone (see Section 4.3, Figure 6). It also cannot be explained by a strong contribution of DM inside the bulge together with a moderate tagential velocity anisotropy, because the corresponding DM halos would still be too massive to fit the data in the outer regions (see Section 4.4, Figure 9a]).e

Another feature in the stellar kinematics, that none of our JAM models was able to reproduce in the slightest and which we therefore excluded in the modelling, was the dip in v_{rms} around $\sim 6''$, which occurs around the transition from bulge to disk (see Figures 1b and 1d).

5.3 On J1331's possible merger history and modelling failures

J1331 has a large counter-rotating stellar core within $\sim 2''$. This suggests a process in J1331's past in which two components with angular momenta oriented in opposite directions were involved.

Accretion of gas on retrograde orbits and subsequent star formation could lead to a younger and counter-rotating stellar population [TO DO: REF]. However, to form enough stars such that the net rotation of the large and massive core is retrograde, a very large amount of gas would have had to be accreted by J1331—which is not very likely [TO DO: REF].

Another scenario are galaxy mergers. Major mergers including large amounts of gas can form kinematically decoupled cores (KDCs), see e.g., [TO DO: REF: Lauren Hofman et al 2010 - read paper how this fits with the reverse color gradient]. During a minor merger, the dense nucleus of a satellite galaxy on a retrograde orbit could survive the dissipationless accretion and spiral to the core due to tidal friction (Kormendy 1984).

Usually ellipticals and the bulges of massive spirals appear reddest in their center and getting increasingly bluer with larger radii. Mergers can reverse this behaviour in the remnant's core. A star formation burst triggered by the major merger could lead to a new young stellar population in the galaxy's bulge [TO DO: Check, if I can claim that]. In case of a minor merger the satellite nucleus now residing in the merger remnant's very center is also much younger than the bulge stars of the massive progenitor, because star formation in low-mass satellite galaxies occurs over a longer period and later than in massive galaxies [TO DO: REF]. The different stellar populations in a merger remnant can be associated with different Υ_* and sometimes even show up as a reverse colour gradient within the bulge in photometry [TO DO: REF].

Even though investigation of the photometry of J1331 did not reveal a distinct blue core in J1331, we cannot fully exclude the possibility that J1331 has Υ_* gradient. And as discussed [TO DO: Where???] an increasing Υ_* could explain the observed central v_{rms} dip. While Υ_* gradients could be easily included in JAM modelling, it would add much more complexity and degeneracies. We would need either more kinematic data and/or a sophisticated guess for the Υ_* gradient to constrain the model parameters reliably.

	L [$10^{10} L_{\odot}$]	Chabrier IMF M_* [$10^{10} M_{\odot}$]	$\Upsilon^{\rm chab}_{\rm I,*}$	Salpeter IMF M_* [$10^{10} M_{\odot}$]	$\Upsilon^{\rm sal}_{\rm I,*}$
bulge	3.10 ± 0.15	7.8 ± 1.8	2.5 ± 0.6	14.5 ± 3.7	4.7 ± 1.2
disk	2.35 ± 0.11	2.9 ± 0.7	1.2 ± 0.3	5.2 ± 1.1	2.2 ± 0.5
total	5.45 ± 0.19	10.6 ± 1.9		19.7 ± 3.9	

Table 7. Total I-band luminosity, stellar mass and mass-to-light ratio, calculated from the I-band AB magnitudes and stellar masses found for J133's bulge and disk by Treu et al. (2011) (their table 2) for comparison with this work. The transformation from AB magnitudes to the Johnson-Cousins I-Band used the relation $I[\text{mag}] = I[\text{ABmag}] - 0.309$ from Frei & Gunn (1994) (their table 2). For the conversion from apparent magnitude to total luminosity the redshift $z = 0.113$ Brewer et al. (2012) was turned into a luminosity distance using the cosmology by Dunkley et al. (2009).

Another way how major mergers can modify the structure of galaxies is a kinematic twist [TO DO: REF] or misalignment (warp) of kinematic and photometric major axis [TO DO: REF]. A kinematic twist shows up as a misalignment of the kinematic major axis of bulge and disk in projection [TO DO: No idea if it works like this.].

In both cases—kinematic twist and warp in J1331—the assumption of axisymmetry in our dynamical modelling would not be satisfied anymore. If such a kinematic twist is present in J1331 cannot be identified from kinematics along the photometric major axis alone, but should be immediately obvious in 2D kinematic maps. We note that if the kinematic major axis was off from the photometric major axis in the outer regions of J1331, the measured $v_{\rm rot}$ in the disk [TO DO: some sinus i] would be much lower than the actual maximum $v_{\rm circ}$. This could maybe also explain the so far inexplicable dips in $v_{\rm rms}$ at the transition from bulge to disk. It is therefore not completely ruled out, that J1331 might have a more massive DM halo than Dutton et al. (2013) found.

Mergers also could have changed the 3D shape of the DM halo and the NFW halo would be therefore not a good model for J1331's DM halo. We also tried to model J1331 with a cored logarithmic halo. However, due to degeneracies in the modelling, we were not able to either constrain the profile for a cored logarithmic halo, or to rule it out.

[TO DO: include somewhere "satellite core has lower velocity dispersion."]

5.4 Future work

Standard JAM modelling approaches seem not to work for J1331. A JAM model for J1331 would need to allow for stellar mass-to-light ratio gradients within the galaxy, velocity anisotropy and a dark matter halo. Because of degeneracies between stellar and dark mass, and matter distribution and anisotropy profile, such a dynamical model would not lead to very tight constraints on the model parameters. A search for colour gradients in J1331 and/or investigation of absorption line indices could support or contradict the suspicion of the existence of stellar mass-to-light ratio gradients in J1331. Subsequently detailed stellar population analyses of the spectra taken along J1331's major axis should be conducted to constrain the mass-to-light ratio reliably. [TO DO: Include in this discussion the findings from the color profile.]

In addition, the dynamical modelling should use more of the available information on J1331 and fit dynamics (stellar and gas kinematics from Dutton et al. (2013)) simultaneously with the gravitational lensing (image positions, shape and even flux ratios) in a similar fashion to Barnabè et al. (2012). To also model the extent, shape and flux of the lensing images, the method by Treu & Koopmans (2004); Warren & Dye (2003) could be employed, which models the surface brightness distribution of the images and source on a pixelated grid. However, for this to work a good model for the galactic extinction would be needed.

High-resolution integral-field spectroscopy could help with this, allow for spatially resolved stellar population analysis and dynamical modelling in two dimensions. [TO DO: Include in this discussion that this could also help test, if the vrms dip is maybe because of twist, i.e., the measurements where not done along the major axis in the disk] This would lead to much better understanding of J1331's structure and mass distribution and therefore answer questions on how minor mergers might modify spiral galaxies.

[TO DO: Include in discussion: High resolution IFU: high spatial resolution –¿ to resolve small features. High spectral resolution –¿ because the velocity dispersion in the outer regions of the galaxy are very low and to measure them accurately, need high spectral resolution.]

[TO DO: Include the following discssion somwhere: The above discussion motivates the following speculation: In the absence of a strong DM contribution in the center, the overall stellar-mass-to-light ratio within the Einstein radius indicates a bottom-heavy IMF close to the Salpeter IMF, consistent with estimates from the velocity dispersion. At the same time the central dip can then be only explained, if there was an increase in stellar mass-to-light ratio with radius *within* the bulge. *If* there was a central stellar population with an IMF close to the Chabrier (2003) IMF, surrounded by a more bottom-heavy population, the central stellar kinematics would be well explained and be fully consistent with the lensing results. (According to Figure 4c the lensing result might not predict a mass to light ratio gradient inside $R_{\rm ein}$, but then again the mass slope $alpha = 1$ was only weakly constrained.) The disk of J1331 has a lower $\Upsilon_{\rm I,*}$ than the bulge (see Table 7 according to Treu et al. (2011)). Such a drop in $\Upsilon_{\rm I,*}$ at the transition region from bulge to disk around $\sim 5''$ could lead to the observed drop in $v_{\rm rms}$, while an increasing contribution of a lower mass DM halo at larger radii as found by Dutton et al. (2013), would recover the kinematics in the out regions of J1331's disk.]

[TO DO: Why are disk and bulge mass in SWELLS I different to the one in SWELLS IV??]

5.5 Summary

[TO DO: Ask Glenn and Aaron how much of the above should find its way into the summary.]

REFERENCES

Baes M., van Hese E., 2007, Astronomy and Astrophysics, 471, 419

Barnabè M., et al., 2012, Monthly Notices of the RAS, 423, 1073

Bendinelli O., 1991, Astrophysical Journal, 366, 599

Binney J., Merrifield M., 1998, Galactic Astronomy

Brewer B. J., et al., 2012, Monthly Notices of the RAS, 422, 3574

Brewer B. J., Marshall P. J., Auger M. W., Treu T., Dutton A. A., Barnabè M., 2014, Monthly Notices of the RAS, 437, 1950

Cappellari M., 2002, Monthly Notices of the RAS, 333, 400

Cappellari M., 2008, Monthly Notices of the RAS, 390, 71

Cappellari M., et al., 2006, Monthly Notices of the RAS, 366, 1126

Chabrier G., 2003, Publications of the ASP, 115, 763

Diemand J., Moore B., Stadel J., 2004, Monthly Notices of the RAS, 352, 535

Dolphin A. E., 2000, Publications of the ASP, 112, 1397

Dolphin A. E., 2008, Zero Points relative to Holtzman et al. (1995)

Dunkley J., et al., 2009, Astrophysical Journal, Supplement, 180, 306

Dutton A. A., Conroy C., van den Bosch F. C., Prada F., More S., 2010, Monthly Notices of the RAS, 407, 2

Dutton A. A., et al., 2011a, MNRAS, 416, 322

Dutton A. A., et al., 2011b, Monthly Notices of the RAS, 417, 1621

Dutton A. A., et al., 2013, Monthly Notices of the RAS, 428, 3183

El-Zant A., Shlosman I., Hoffman Y., 2001, Astrophysical Journal, 560, 636

Emsellem E., Monnet G., Bacon R., 1994, Astronomy and Astrophysics, 285, 723

Emsellem E., Dejonghe H., Bacon R., 1999, Nonthly Notices of the RAS, 303, 495

Evans N. W., Witt H. J., 2003, Monthly Notices of the RAS, 345, 1351

Ferreras I., La Barbera F., de la Rosa I. G., Vazdekis A., de Carvalho R. R., Falcón-Barroso J., Ricciardelli E., 2013, Monthly Notices of the RAS, 429, L15

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publications of the ASP, 125, 306

Frei Z., Gunn J. E., 1994, Astronomical Journal, 108, 1476

Fukushige T., Makino J., 2001, Astrophysical Journal, 557, 533

Gebhardt K., et al., 2003, Astrophysical Journal, 583, 92

Holmberg E., 1958, Meddelanden fran Lunds Astronomiska Observatorium Serie II, 136, 1

Holtzman J. A., Burrows C. J., Casertano S., Hester J. J., Trauger J. T., Watson A. M., Worthey G., 1995, Publications of the ASP, 107, 1065

Kochanek C. S., 1991, Astrophysical Journal, 373, 354

Kormendy J., 1984, Astrophysical Journal, 287, 577

Kronawitter A., Saglia R. P., Gerhard O., Bender R., 2000, Astronomy and Astrophysics, Supplement, 144, 53

Kroupa P., 2002, Science, 295, 82

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, Monthly Notices of the RAS, 391, 1940

Monnet G., Bacon R., Emsellem E., 1992, Astronomy and Astrophysics, 253, 366

Moore B., 1994, Nature, 370, 629

Narayan R., Bartelmann M., 1999, in Dekel A., Ostriker J., eds, Formation of Structure in the Universe, Proceedings of the 1995 Jerusalem Winter School. Cambridge University Press, pp 360–432

Navarro J. F., Frenk C. S., White S. D. M., 1996, Astrophysical Journal, 462, 563

Pontzen A., Governato F., 2012, MNRAS, 421, 3464

Quinlan G. D., Hernquist L., 1997, New Astronomy, 2, 533

Rubin V. C., Ford W. K. J., . Thonnard N., 1980, ApJ, 238, 471

Salpeter E. E., 1955, Astrophysical Journal, 121, 161

Sellwood J. A., 2004, Dynamics of Astrophysical Discs

Spiniello C., Trager S., Koopmans L. V. E., Conroy C., 2014, Monthly Notices of the RAS, 438, 1483

Tody D., 1993, in Hanisch R. J., Brissenden R. J. V., Barnes J., eds, Astronomical Society of the Pacific Conference Series Vol. 52, Astronomical Data Analysis Software and Systems II. p. 173

Treu T., 2010, Annual Review of Astron and Astrophys, 48, 87

Treu T., Koopmans L. V. E., 2004, Astrophysical Journal, 611, 739

Treu T., Dutton A. A., Auger M. W., Marshall P. J., Bolton A. S., Brewer B. J., Koo D. C., Koopmans L. V. E., 2011, Monthly Notices of the RAS, 417, 1601

Tully R. B., 1988, Nearby galaxies catalog

Warren S. J., Dye S., 2003, Astrophysical Journal, 590, 673

de Blok W. J. G., McGaugh S. S., Bosma A., Rubin V. C., 2001, ApJ, 552, L23

van de Ven G., Falcón-Barroso J., McDermid R. M., Cappellari M., Miller B. W., de Zeeuw P. T., 2010, Astrophysical Journal, 719, 1481