The ATLAS[3D] project – XIX. Benchmark for early-type galaxies scaling relations from 260 dynamical models: mass-to-light ratio, dark matter, Fundamental Plane and Virial Plane

Michele Cappellari,[1*] Nicholas Scott,[1,2] Katherine Alatalo,[3] Leo Blitz,[3] Maxime Bois,[4] Frédéric Bournaud,[5] M. Bureau,[1] Alison F. Crocker,[6] Roger L. Davies,[1] Timothy A. Davis,[1,7] P. T. de Zeeuw,[7,8] Pierre-Alain Duc,[5] Sadegh Khochfar,[10] Davor Krajnović,[7] Harald Kuntschner,[7] Richard M. McDermid,[11] Raffaella Morganti,[12,13] Thorsten Naab,[14] Tom Oosterloo,[12,13] Marc Sarzi,[15] Paolo Serra,[12] Anne-Marie Weijmans[16] and Lisa M. Young[17]

[1] *Sub-department of Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH*
[2] *Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[3] *Department of Astronomy, Campbell Hall, University of California, Berkeley, CA 94720, USA*
[4] *Observatoire de Paris, LERMA and CNRS, 61 Av. de l'Observatoire, F-75014 Paris, France*
[5] *Laboratoire AIM Paris-Saclay, CEA/IRFU/SAp CNRS Université Paris Diderot, 91191 Gif-sur-Yvette Cedex, France*
[6] *Department of Astrophysics, University of Massachusetts, 710 North Pleasant Street, Amherst, MA 01003, USA*
[7] *European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany*
[8] *Sterrewacht Leiden, Leiden University, Postbus 9513, 2300 RA Leiden, the Netherlands*
[9] *Université Lyon 1, Observatoire de Lyon, Centre de Recherche Astrophysique de Lyon*
and Ecole Normale Supérieure de Lyon, 9 avenue Charles André, F-69230 Saint-Genis Laval, France
[10] *Max-Planck Institut für extraterrestrische Physik, PO Box 1312, D-85478 Garching, Germany*
[11] *Gemini Observatory, Northern Operations Centre, 670 N. A'ohoku Place, Hilo, HI 96720, USA*
[12] *Netherlands Institute for Radio Astronomy (ASTRON), Postbus 2, 7990 AA Dwingeloo, The Netherlands*
[13] *Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands*
[14] *Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany*
[15] *Centre for Astrophysics Research, University of Hertfordshire, Hatfield, Herts AL1 9AB, UK*
[16] *Dunlap Institute for Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada*
[17] *Physics Department, New Mexico Institute of Mining and Technology, Socorro, NM 87801, USA*

ABSTRACT

We study the volume-limited and nearly mass selected (stellar mass $M_{\rm stars} \gtrsim 6 \times 10^9$ M$_\odot$) ATLAS$^{\rm 3D}$ sample of 260 early-type galaxies (ETGs, ellipticals Es and lenticulars S0s). We construct detailed axisymmetric dynamical models, which allow for orbital anisotropy, include a dark matter halo, and reproduce in detail both the galaxy images and the high-quality integral-field stellar kinematics out to about $1R_{\rm e}$, the projected half-light radius. We derive accurate total mass-to-light ratios $(M/L)_e$ and dark matter fractions $f_{\rm DM}$, within a sphere of radius $r = R_{\rm e}$ centred on the galaxies. We infer masses $M_{\rm JAM} \equiv L \times (M/L)_e \approx 2 \times M_{1/2}$, where $M_{1/2}$ is the mass within a sphere enclosing half of the galaxy light. We also measure stellar $(M/L)_{\rm stars}$. We find that the thin two-dimensional subset spanned by galaxies in the $(M_{\rm JAM}, \sigma_e, R_e^{\rm max})$ coordinates system, which we call the Virial Plane (VP) has an observed rms scatter of 17%, which would imply an intrinsic one of just 4%. The VP satisfies the scalar virial relation $M_{\rm JAM} = 5.0 \times \sigma_e^2 R_e^{\rm max}/G$ within our tight errors. This is true when (i) the major axis $R_e^{\rm max}$ of the isophote containing half of the total projected galaxy light is used as galaxy radius, and (ii) σ_e is measured inside that isophote. This show that the larger scatter in the Fundamental Plane (FP) $(L, \sigma_e, R_e^{\rm max})$ is due to stellar population effects (including trends in the stellar Initial Mass Function [IMF]) and confirms that the deviation from the virial relation is due to a genuine $(M/L)_e$ variation. However, the details of how both $R_{\rm e}$ and σ_e are determined are critical in defining the precise deviation from this simple virial form. Differences in these details is the basic reason for a decades-old debate on the origin of the Fundamental Plane tilt. Even using excellent photometry, the main uncertainty in masses or M/L estimates using the scalar virial relation is in the measurement of $R_{\rm e}$: relative values are easy to obtain, but absolute normalizations are difficult to reproduce. This problem is already relevant for nearby galaxies and may cause significant biases in virial mass and size determinations at high-redshift. Dynamical model can eliminate these problems. We measure a median dark matter fractions of $f_{\rm DM} = 16\%$ in our sample, which implies the total density profile within $1R_{\rm e}$ is dominated by the stellar density $\rho_{\rm stars}$. When approximated by a power-law the latter has the 'isotermal' form $\rho_{\rm stars}(r) \propto r^{-2.0}$ with an intrinsic rms scatter of 0.2 for our sample. We revisit the $(M/L)_e - \sigma_e$ relation, which describes most of the deviations between the VP and the FP. The best-fitting relation is $(M/L)_e \propto \sigma_e^{0.69}$. Given that part of the $(M/L)_e$ increase with σ_e is already explained by systematic variations in age, metallicity (and possibly dark matter fraction), this relation provides an upper limit to any increase of the IMF mass normalization with σ_e. We find differences in the relation as a function of galaxy rotation and environment, with the correlation being slightly more shallow and with smaller scatter for slow rotating systems or for galaxies in Virgo. For the latter, when using the best distance estimates, we observe a scatter in $(M/L)_e$ of 10%, from which we infer an intrinsic one of just 6%. We perform an accurate empirical study of the relations between σ_e and the galaxies circular velocity within $1R_{\rm e}$ and find the two empirical relations $V_{\rm circ}(R_e^{\rm max}) \approx 1.51 \times \sigma_e$ and $\max(V_{\rm circ}) \approx 1.76 \times \sigma_e$, which are satisfied with small scatter and negligible dependence of σ_e for our entire sample. The accurate parameters described in this paper are used in the companion Paper XX of this series to explore the variation of global galaxy properties, including the IMF, on the projections of the VP.

Key words: galaxies: elliptical and lenticular, cD – galaxies: evolution – galaxies: formation – galaxies: structure – galaxies: kinematics and dynamics

1 INTRODUCTION

Scaling relations of early-type galaxies (ETGs, ellipticals E and lenticulars S0) have played a central role in our understanding of galaxy evolution, since the discovery that the stellar velocity dispersion σ (Faber & Jackson 1976) and the galaxy projected half-light radius $R_{\rm e}$ (Kormendy 1977) correlate with galaxy luminosity L. An important step forward was made with the discovery that these two relations are just projections of a relatively narrow plane, the Fundamental Plane (FP) (Faber et al. 1987; Dressler et al. 1987; Djorgovski & Davis 1987), relating the three variables $(L, \sigma_e, R_{\rm e})$.

When the plane is used as a distance indicator, as was especially the case at the time of its discovery, the luminosity can be replaced by the surface brightness within $R_{\rm e}$ as $\Sigma_e \equiv L/(2\pi R_{\rm e}^2)$ and the observed plane assumes the form

$$R_{\rm e} \propto \sigma^{1.33} \Sigma_e^{-0.82}, \qquad (1)$$

where the adopted parameters are the median of the 11 independent determinations tabulated in Bernardi et al. (2003).

It was immediately realized that the existence of the FP could be due to the galaxies being in virial equilibrium (e.g. Binney & Tremaine 2008) and that the deviation (tilt) of the coefficients from the virial predictions $R_{\rm e} \propto \sigma^2 \Sigma_e^{-1}$, could be explained by a smooth power-law variation of mass-to-light ratio M/L with

* E-mail: cappellari@astro.ox.ac.uk

mass (Faber et al. 1987). The FP showed that galaxies assemble via regular processes and that their properties are closely related to their mass. The tightness of the plane gives constraints on the variation of stellar population among galaxies of similar characteristics and on their dark matter content (Renzini & Ciotti 1993; Borriello et al. 2003). The regularity also allows one to use the FP to study galaxy evolution, by tracing its variations with redshift (van Dokkum & Franx 1996).

However, other reasons for the deviation of the coefficients are possible: the constant coefficients in the simple virial relation only rigorously apply if galaxies are spherical and homologous systems, with similar profiles and dark matter fraction. But both galaxies concentration (Caon et al. 1993) and the amount of random motions in their stars (Davies et al. 1983) were found to systematically increase with galaxy luminosity.

The uncertain origin of the tilt led to a large number of investigations about its origin, exploring the effects of (i) the systematic variation in the stellar population, or (ii) the non-homology in the surface brightness distribution or (iii) the kinematic, or (iv) the variation in the amount of dark matter, on the FP tilt and scatter (Renzini & Ciotti 1993; Prugniel & Simien 1994, 1996, 1997; Ciotti et al. 1996; Graham & Colless 1997; Forbes et al. 1998; Bertin et al. 2002; Borriello et al. 2003; Trujillo et al. 2004). Those works were all based on approximate galaxy spherical models, trying to test general hypotheses and not reproducing real galaxies in detail, which sometimes led to contrasting results. What became clear however was that various effects could potentially influence a major part of the FP tilt. Moreover it was found that the scatter in the FP implies a well regulated formation for ETGs.

The next step forward came with subsequent studies, which instead of testing general trends, used small samples of objects and tried to push to the limit the accuracy of measuring galaxy central masses, while reducing biases as much as possible. Those accurate total masses could be directly compared to the simple virial ones, testing for residual trends. Similar but independent studies were performed using two completely different techniques, either stellar dynamics (Cappellari et al. 2006) or strong gravitational lensing (Bolton et al. 2007, 2008; Auger et al. 2010a). The results from those efforts agree with each others, and showed that the tilt of the FP is almost entirely due to a genuine M/L variation.

In this paper we investigate once more the origin of the FP tilt. This new study is motivated by the dramatic increase in the size and quality of our galaxy sample, with respect to any previous similar study. We have in fact state-of-the-art SAURON (Bacon et al. 2001) stellar kinematics for all the 260 early-type galaxies of the ATLAS[3D] sample (Cappellari et al. 2011a, hereafter Paper I), which constitute a volume-limited and carefully selected sample of ETGs, down to a stellar mass of about $M_{stars} \gtrsim 6 \times 10^9 \ M_\odot$. This fact, combined with detailed dynamical models for the entire sample, allows us to test previous claims with unprecedented accuracy. The new models also include a dark matter halo and give constraints on the dark matter content in the centres of early-type galaxies. These measurements will be used in the companion Cappellari et al. (2012a, hereafter Paper XX) to provide a novel view of galaxy scaling relations.

In what follows, in Section 2 we present the sample and data, in Section 3 we describe the methods used to extract our quantities, in Section 4 we present our results on the FP tilt, dark matter and the $(M/L)\sigma$ relation, and finally we summarize our paper in Section 5.

2 SAMPLE AND DATA

2.1 Selection

The galaxies studied in this work are the 260 early-type galaxies which constitute the volume-limited and nearly mass-selected ATLAS[3D] sample (Paper I). The object were morphologically selected as early-type according to the classic criterion (Hubble 1936; de Vaucouleurs 1959; Sandage 1961) of not showing spiral arms or a disk-scale dust lane (when seen edge-on). The early-types are extracted from a parent sample of 871 galaxies of all morphological types brighter than $M_K = -21.5$ mag, using 2MASS photometry (Skrutskie et al. 2006), inside a local ($D < 42$ Mpc) volume of 1.16×10^5 Mpc^3 (see full details in Paper I).

2.2 Comparison to previous samples: dynamics and lensing

Our goal is to measure total masses, or equivalently mass-to-light ratios (M/L), in the central regions of galaxies. M/L of significant samples of individual ETGs have been previously obtained via dynamical modelling (e.g. van der Marel 1991 [37 ETGs]; Magorrian et al. 1998 [36 ETGs]; Gerhard et al. 2001 [21 ETGs]; Cappellari et al. 2006 [25 ETGs]; Thomas et al. 2007b [16 ETGs]; Williams et al. 2009 [14 ETGs]; Scott et al. 2009 [48 ETGs]) or strong gravitational lensing (e.g. Rusin et al. 2003 [22 ETGs]; Koopmans et al. 2006 [15 ETGs]; Bolton et al. 2008 [53 ETGs]; Auger et al. 2010a [73 ETGs]). An important, and perhaps not obvious, difference between the quantities obtained with the two techniques is that the dynamical models provide masses enclosed within a *spherical* radius, while strong lensing measures the mass inside a *cylinder* with axis parallel to the line-of-sight. Care has to be taken when comparing the two methods. A nice illustration of this fact is given in figure 1 of Dutton et al. (2011a).

An advantage of the strong lensing technique is that the recovered mass inside a cylinder with the radius of the Einstein ring is nearly insensitive to the mass distribution, and completely independent on the stellar dynamics. However, the requirement of a galaxy to act as a strong lens, necessarily imposes biases in the objects selection, and in particular limits mass measurements via strong lensing to the most massive nearby ETGs ($\sigma \gtrsim 200$ km s^{-1} in Auger et al. 2010a).

The dynamical modelling technique has the significant advantage that it can in principle be applied to any bound system made of stars. However, it requires a detailed treatment of the observed surface brightness and orbital distribution, in combination with integral-field data, for robust and accurate values (e.g. Cappellari et al. 2006).

In this paper we apply the dynamical modelling technique to the ATLAS[3D] sample of 260 early-type galaxies. This increases the sample size for which accurate total masses have been measured by a factor of four. Moreover the sample is volume-limited and statistically representative of the nearby galaxy population with stellar mass $M_{stars} \gtrsim 6 \times 10^9 \ M_\odot$ and in particular includes ETGs with velocity dispersion as low as $\sigma_e \approx 40$ km s^{-1} (see Paper I for an illustration of the characteristics of the sample).

2.3 Stellar kinematics and imaging

Various multi-wavelengths datasets are available for the sample galaxies (see a summary in Paper I). In this work we make use of the SAURON (Bacon et al. 2001) integral-field stellar kinematics within about one half-light radius R_e, which was introduced

in Emsellem et al. (2004), for the subset of 48 early-types in the SAURON survey (de Zeeuw et al. 2002), and in Paper I for the rest of the ATLAS^{3D} sample. Maps of the stellar velocity for all the 260 galaxies were presented in Krajnović et al. 2011, hereafter Paper II.

In this paper we are not interested in the shape of the stellar line-of-sight velocity distribution (LOSVD), but we want to approximate velocity moments which are predicted by the (Jeans 1922) equations. In Cappellari et al. (2007) we used semi-analytic models to compute a set of realistic galaxy LOSVDs with known velocity moments, using the Hunter & Qian (1993) formalism, as implemented in Emsellem et al. (1999). The models LOSVDs were used to broaden galaxy spectral templates and noise was subsequently added. The kinematics was then extracted from the synthetic spectra using pPXF Cappellari & Emsellem (2004) as done for the real galaxies. We found that $V_{\rm rms} \equiv \sqrt{V^2 + \sigma^2}$, where V and σ are the mean and standard deviation of the best fitting Gaussian provide a better empirical approximation to the velocity second moment $\langle v_{\rm los}^2 \rangle^{1/2}$ than an integral of a more general LOSVD described by the Gauss-Hermite parametrization (van der Marel & Franx 1993; Gerhard 1993). This is due to the large sensitivity of the moments to the wings of the LOSVD, which are observationally ill determined. For this reason all the kinematic quantities used in the paper are extracted using a simple Gaussian LOSVD in the pPXF software (keyword MOMENTS=2).

The photometry used in this work comes from the Sloan Digital Sky Survey (SDSS, York et al. 2000) data release eight (DR8 Aihara et al. 2011) for 225 galaxies and was supplemented by our own photometry taken at the 2.5-m Isaac Newton Telescope in the same set of filters and with comparable signal to noise for the rest of the sample galaxies (Scott et al. 2012, hereafter Paper XXI).

3   METHODS

3.1   Measuring galaxy enclosed masses

3.1.1   *Choosing the dynamical modelling approach*

Various dynamical modelling techniques have been developed in the past. They are all characterized by their ability to reproduce in detail, in a *non-parametric* way, the characteristics of the galaxy surface brightness. This contrasts with a more qualitative toy-model approach (e.g. Tortora et al. 2009; Treu et al. 2010) that assume a spherical shape and a simpler parametrization (e. g. Hernquist 1990 or Sersic 1968 profile) for the surface brightness of all galaxies. An accurate description of the galaxy surface brightness is a necessary requirement for quantitative and unbiased measurements of dynamical quantities as much of the kinematic information on real galaxies is contained in the photometry alone (Cappellari 2008). The state of the art in the field is currently represented by Schwarzschild (1979) orbit-superposition approach, which was originally developed to reproduce galaxy stellar densities and was later generalized to produce detailed fits to the stellar kinematics (Richstone & Tremaine 1988; Rix et al. 1997; van der Marel et al. 1998) and has been widely used for determinations of masses of supermassive black holes (e.g. van der Marel et al. 1997; Gebhardt et al. 2000a; Cappellari et al. 2002; Valluri et al. 2004; Houghton et al. 2006), for galaxy mass determinations (e.g. Cappellari et al. 2006; Thomas et al. 2007b) and to recover orbital distributions (e.g. Krajnović et al. 2005; Cappellari et al. 2007; van den Bosch et al. 2008; Thomas et al. 2009). A close contender technique, but not as widely used, is the particle-based made-to-measure method of (Syer & Tremaine

1996) as implemented to reproduce kinematical observables by various groups (de Lorenzi et al. 2007; Dehnen 2009; Long & Mao 2010). When the gravitational potential is assumed to be known, and the particles are chosen to fully sample all integrals of motion, the method effectively corresponds to a particle-based analogue of Schwarzschild's method, and is expected to provide similar results. However, the method may be very useful when the potential is derived from the particles in a self-consistent way. Not much however is known about the convergence and uniqueness of the solution in this case.

The sophistication and generality of the dynamical models has reached a level that exceeds the amount of information that the observations of external galaxies can provide. As a result the observations are unable to uniquely constrain all the model parameters, which suffer from degeneracies (Dejonghe & Merritt 1992; Gerhard et al. 1998; de Lorenzi et al. 2009; Morganti & Gerhard 2012). A key degeneracy is in the deprojection of the observed surface brightness into a three dimensional stellar mass distribution, which has been proved to be of mathematical nature (Rybicki 1987; Gerhard & Binney 1996) and applies even when the galaxy is assumed to be axisymmetric. However, similar degeneracies are likely to exists when higher (than zero) moments of the velocity are considered. This is expected from dimensional arguments: the current data provide at most a three-dimensional observable (an integral-field data cube), which is the minimum requirement to constrain the orbital distribution, which depends on three integrals of motion, for an assumed potential and known light distribution. It is unlikely for the data to contain enough information to constrain additional parameters, like the dark matter halo shape and the viewing angles (e.g. Valluri et al. 2004). Numerical experiments confirm that even with the best available integral-field stellar kinematics, and assuming the gravitational potential is known and axisymmetric, even the galaxy inclination cannot be reliably inferred from the data using general Schwarzschild models (Krajnović et al. 2005; Cappellari et al. 2006; van den Bosch & van de Ven 2009).

The situation becomes even more problematic when one considers the fact that the majority of early-type galaxies are likely to have bars. 30% have obvious bars (Paper II) in the ATLAS^{3D} sample, but more must be hidden by projection effects. Bars are characterized by figure rotation which is ignored by most popular modelling approaches. The treatment of bars could be included in the models as demonstrated in the two-dimensional limit by (Pfenniger 1984) and as done to models the Milky Way in three dimension (Zhao 1996; Häfner et al. 2000; Bissantz et al. 2004). However, no applications to external galaxies exists. This is due to the extra degeneracy that the addition of at least two extra model parameters, the bar pattern speed and position angle, will produce on an already degenerate problem. This combines with the dramatic increase in the non-uniqueness of the mass deprojection expected in a triaxial rather than axisymmetric distribution (Gerhard 1996) and in the additional unavoidable biases introduced by observational errors. All this is expected to further broaden the minima in the χ^2 distributions of the fits and increase the uncertainties and covariances in the recovered parameters.

We chose a different approach. Rather than allowing for the full generality and degeneracies of the models, we adopt a modelling method that makes empirically-motivated assumptions to restrict the range of model solutions and improve the accuracy of the mass recovery. This is motivated by the finding that the kinematics of real fast-rotator early-type galaxies in the SAURON sample (de Zeeuw et al. 2002) is well approximated by models characterized by a remarkably simple and homogeneous dynamics,

characterized by a cylindrically-aligned and nearly oblate velocity ellipsoid $\sigma_\phi \approx \sigma_R \gtrsim \sigma_z$ (Cappellari 2008), as previously suggested by more general Schwarzschild's models (Cappellari et al. 2007; Thomas et al. 2009). The models are called Jeans Anisotropic MGE (JAM), where MGE stands for the Multi-Gaussian Expansion method of Emsellem et al. (1994), that is used to accurately describe the galaxy photometry. The JAM models can reproduce the full richness of the observed state-of-the-art SAURON integral-field kinematics of fast rotator ETGs using just two free parameters (Cappellari 2008; Scott et al. 2009; Cappellari et al. 2012b), providing a compact description of their dynamics. The JAM models are ideal for this work given that the nearly-axisymmetric fast rotator ETGs constitute the 86% of the ATLAS^{3D} sample (Paper II; Emsellem et al. 2011, hereafter Paper III). Moreover the JAM models only require the first two velocity moments (V and σ), and not the full LOSVD, which is not available for about half of the sample (see Paper I). The JAM models do not have the freedom to actually fit small-scale details of the kinematics, but they make a prediction based on an accurate description of the photometry and a couple of parameters. This constitutes an advantage in presence of noise and systematics in the data, as it makes spurious features easy to recognize and automatically exclude from the fit. Moreover the approach is at least three orders of magnitudes faster than Schwarzschild's approach.

Not all ETGs are well described by the JAM models however: some of the slow rotators in ATLAS^{3D} are likely nearly spherical in the region where we have stellar kinematics, but about 10% of the sample galaxies are weakly triaxial or out of equilibrium (Paper II). For those objects the modelling results should be treated with caution. Errors of up to 20% can arise when measuring masses of triaxial objects with axisymmetric models (Thomas et al. 2007a; van den Bosch & van de Ven 2009) and this should be kept in mind when interpreting our results. However, preliminary tests using real galaxies in the SAURON sample indicate excellent agreement between the M/L recovery using axisymmetric models and triaxial ones with identical data (van den Bosch 2008). Moreover, in what follows, unless explicitly mentioned, we verified that all conclusion are unchanged if we remove the slow rotator galaxies from the sample. Barred galaxies provide a further complication, which will be discussed in the next Section.

3.1.2 JAM models with dark halo

In practice the modelling approach we use in this paper starts by approximating the observed SDSS and INT r-band surface brightness distribution of the ATLAS^{3D} galaxies using the Multi-Gaussian Expansion (MGE) parametrization (Emsellem et al. 1994), with the fitting method and MGE_FIT_SECTORS software package of Cappellari (2002)[1]. The choice of the photometric band is a compromise between the need of using the reddest band, to reduce the contamination by dust, and the optimal signal-to-noise in the images. For barred galaxies the Gaussians of the MGE models are constrained to have the flattening of the outer disk, following Scott et al. (2009, their fig. 4). Full details of the fitting approach and illustrations of the quality of the resulting MGE fits are given in Paper XXI. The MGE models are used as input for the JAM method[1] (Cappellari 2008) which calculates a prediction of the line-of-sight second velocity moments $\langle v_{los}^2 \rangle$ for given model parameters and compare this to the observed V_{rms}.

---

[1] Available from http://purl.org/cappellari/idl

In Cappellari et al. (2006) it was shown that, when the surface brightness distribution is accurately reproduced and good quality integral-field data are available, simple two-integral Jeans models measure masses nearly as accurate as those of Schwarzschild's models, with errors of 6%. The agreement can be further improved by allowing for orbital anisotropy, in which case the two methods give equally accurate results (Cappellari 2008). We have run an extensive set of tests using JAM to determine the M/L of realistic numerical simulations (Lablanche et al. 2012, hereafter Paper XII). We found that for unbarred galaxies, even when the anisotropy is not accurately constant inside the region with kinematic data, the M/L can be recovered with maximum biases as small as 1.5%. The situation changes when the galaxies are barred. In this case biases of up to 15% can be expected for the typical bar strengths we find in ETGs.

The models we use here were already presented in Cappellari et al. (2012b), where they were used to uncover a systematic variation of the stellar IMF in ETGs. That paper (their table 1) describes six sets of JAM models for all the ATLAS^{3D} galaxies, making various assumptions on the dark matter halo. Given that the SAURON data are typically spatially limited to $1R_e$ one cannot expect to be able to robustly characterize the shape of the dark halo out to large radii from them (Mamon & Łokas 2005). However, as long as the radial profile and the flattening of the halos are not both the same as those of the stellar distribution, we can determine how much room the models allow for a dark matter halo, within the region constrained by the kinematics. The models were summarized in Cappellari et al. (2012b), but we describe them here in some more detail using the same lettering notation as that paper:

(A) Self-consistent JAM model: Here we assume that the mass distribution follows the light one as inferred from the deprojected MGE. In this case the model has three free parameters. Two parameters are non-linear: (i) the vertical anisotropy $\beta_z = 1 - \sigma_z^2/\sigma_R^2$ and (ii) the galaxy inclination i, which together uniquely specify the shape of the second velocity moment $\langle v_{los}^2 \rangle$, which is then linearly scaled by the $(M/L)_{JAM}$ to fit the two-dimensional V_{rms} data. We emphasize that, even though the models do not include a dark halo explicitly, $(M/L)_{JAM}$ does *not* represent the stellar M/L, as sometimes incorrectly assumed, but the *total* one, within a spherical region which has the projected size of our data (see discussion in Section 4.1.2). This set of models, like all others, has a central supermassive black hole with mass predicted by the $M_{BH} - \sigma$ correlation (Gebhardt et al. 2000b; Ferrarese & Merritt 2000), or a black holes mass as published, when available. The supermassive black hole has a minimal effect on M/L in nearly all cases, but we still exclude the central $R < 2''$ from the fits, for maximum robustness. Examples of mass-follows-light JAM models are shown in Fig. 1.

(B) JAM with NFW dark halo: This set of models adopted the approach introduced by Rix et al. (1997) to reduce the halo to a one-parameter family of models. This approach was already used with axisymmetric JAM models of disk galaxies, as done here, by Williams et al. (2009) and to construct spherical toy models of various stellar systems (Napolitano et al. 2005; Tollerud et al. 2011). We assume the halo is spherical and characterized by the two-parameters double power-law NFW profile (Navarro et al. 1996). We then adopt the halo mass-concentration $M_{200} - c_{200}$ relation (Navarro et al. 1996) as given by (Klypin et al. 2011) to make the halo profile a unique function of its mass M_{200}. The latter is not a critical
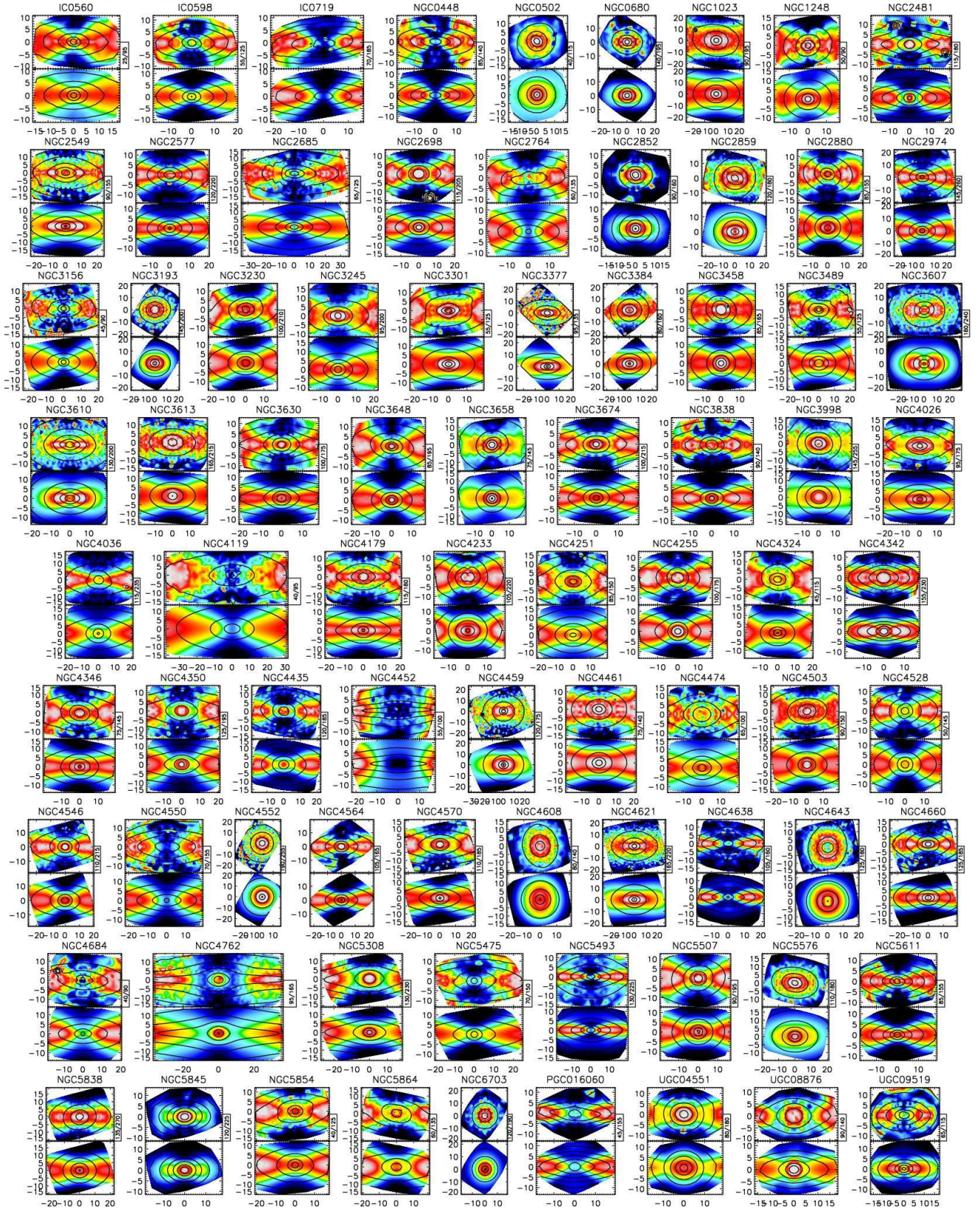
Figure 1. Mass-follows-light JAM model examples. In each panel, the top plot shows the by-symmetrized SAURON data, with the observed galaxy surface brightness overlaid, in steps of 0.5 mag. The bottom plot shows the best-fitting JAM model, with the MGE surface brightness assumed by the model. The models (A) have just two free non-linear parameters, the inclination and the global anisotropy (i, β_z), to reproduce the shape of the observed second velocity moments $V_{\rm rms} \equiv \sqrt{V^2 + \sigma^2}$, where V is the mean stellar velocity and σ is the stellar velocity dispersion. Yet, once the surface brightness is given, most of the variety in our maps can be reproduced. The most significant deviations between data and models are due to bars, recognizable from the asymmetries in the observed surface brightness. The predictive power of these simple JAM models qualitatively suggest that the assumed total potential may not be significantly in error, which would imply dark matter is unimportant (or accurately follows the light). The good fits shows that ETGs have a relatively simple dynamics within $1R_{\rm e}$.

assumption: our observations only sample a region well inside the predicted halo break radius, so that all our conclusion are unchanged if we describe the halo with a simple power law density profile $\rho(r) \propto r^{-1}$, as we numerically verified. The resulting JAM models have in this case four parameters: (i) The galaxy inclination i (ii) the anisotropy β_z, (iii) the *stellar* $(M/L)_{\rm stars}$, assumed spatially constant and (iv) the halo virial mass M_{200}, defined as the mass within the spherical radius r_{200} at which the average density is equal to 200 times the critical density of the Universe.

(C) JAM with contracted NFW dark halo: These models include a halo which is originally assumed to be of NFW form, with concentration specified by its mass via the $M_{200} - c_{200}$ relation as in (B). However, during the fitting process, for every choice of the model parameters, the halo is contracted according to the enclosed stellar mass distribution, which is defined by the (circularized) MGE and the corresponding $(M/L)_{\rm stars}$ parameter. For the contraction we used the prescription of Gnedin et al. (2011), which is an update of Gnedin et al. (2004). We verified that our IDL code produces the same output as the C language software CON-TRA by Gnedin et al. (2004), when the same input is given. The resulting JAM model has the same four free parameters $(i, \beta_z, (M/L)_{\rm stars}, M_{200})$ as in (B).

(D) JAM with general dark halo (gNFW): These models include a dark halo that generalizes the NFW profile (see also Barnabè et al. 2012)

$$\rho_{\rm DM}(r) = \rho_s \left(\frac{r}{r_s}\right)^{\gamma} \left(\frac{1}{2} + \frac{1}{2}\frac{r}{r_s}\right)^{-\gamma-3}. \qquad (2)$$

The density has the same asymptotic power-law slope $\beta = -3$ as the NFW halo, but it allows for a variable inner slope, which we constrained to the bounds $-1.6 < \gamma < 0$, by assigning zero probability to the prior $P(\rm model) = 0$ (Section 3.1.3) outside this parameters range. The ranges include a flat inner core $\gamma = 0$ and the NFW $\gamma = -1$. The upper bound was chosen as the nearly maximum slope we measured for all contracted halos in (C) (top panel of Fig. 2). However, recent simulations suggest that baryonic effects produce flatter halos than these predictions for a broad range of galaxy masses (Duffy et al. 2010; Governato et al. 2010; Inoue & Saitoh 2011; Pontzen & Governato 2012; Laporte et al. 2012; Macciò et al. 2012; Martizzi et al. 2012). Note that our adopted maximum halo slope is still generally more shallow than the typical 'isothermal' average power slope $\gamma' = 2.0$ the we measure for the stellar density alone within $1R_{\rm e}$ (bottom panel of Fig. 2). This fact is important to avoid model degeneracies between the stellar and halo densities. This model is the most general of all six and it includes any of the other five models as special cases. It has five free parameters: (i) the galaxy inclination, (ii) the anisotropy β_z, (iii) the stellar mass $M_{\rm stars}$, (iv) the halo inner slope γ and (v) the halo density ρ_s at r_s, which we parametrized using the dark matter fraction $f_{DM}(r = R_e)$ to reduce the strong correlation between ρ_s and γ during the parameter estimation. The break radius r_s of the halo was not included as a free parameter given that it is (in models E) generally 3–5 times larger than $R_{\rm e}$ and it is completely unconstrained by our data. We fixed $r_s = 20$ kpc, which is the median value for all models E, but we verified that nearly identical results are obtained if we describe the halo with a simple power-law density profile $\rho(r) \propto r^{-\gamma}$;

(E) JAM with fixed NFW dark halo: The halo has a NFW pro-

file without any free parameter. During the fitting process the halo mass M_{200} is determined from the enclosed stellar mass $M_{\rm stars}$, which is given by the total luminosity of the MGE model multiplied by its current $(M/L)_{\rm stars}$. This is done using the $M_{200} - M_{\rm stars}$ relation derived by Moster et al. (2010) (see also Moster et al. 2012), which matches the observed galaxy luminosity functions to the simulated halos mass function. However, negligible differences would have been obtained using e.g. the similar relations derived by Behroozi et al. (2010) or Guo et al. (2010). For a given halo mass, the concentration is specified by the $M_{200} - c_{200}$ relation as in (B). The only free model parameters are the three of the stellar component $(i, \beta_z, (M/L)_{\rm stars})$ as in (A). This fixed-halo assumption, in combination however with equally fixed spherical and isotropic Hernquist (1990) galaxy models, was also used by Auger et al. (2010b) and Deason et al. (2012).

(F) JAM with fixed contracted dark halo: The halo has a contracted profile without any free parameter. For a given stellar mass, the halo has initially the same NFW form as in (E), but the profile is contracted as in (C) using the prescription of Gnedin et al. (2011). The only free model parameters are the three of the stellar component $(i, \beta_z, (M/L)_{\rm stars})$ as in (A). This fixed-halo assumption, in combination however with equally fixed spherical and isotropic Hernquist (1990) galaxy models, was also used by Auger et al. (2010b).

3.1.3 *Bayesian inference of the JAM model parameters*

The determination of the JAM model parameters for the 260 ATLAS³ᴰ galaxies in Cappellari et al. (2012b) was done using Bayesian inference (Gelman et al. 2004). The same approach was adopted using JAM models in Barnabè et al. (2012) in combination with gravitational lensing. From Bayes theorem, the posterior probability distribution of a model, with a given set of parameters, given our data is

$$P(\rm model \,|\, data) \propto P(\rm data \,|\, model) \times P(\rm model). \qquad (3)$$

Here we make the rather common assumption of Gaussian errors, in which case the probability of the data, for a given model is given by

$$P(\rm data \,|\, model) \propto \exp\left(-\frac{\chi^2}{2}\right), \qquad (4)$$

with

$$\chi^2 = \sum_j \left(\frac{\langle v_{\rm los}^2 \rangle_j^{1/2} - V_{\rm rms,j}}{\Delta V_{\rm rms,j}}\right)^2. \qquad (5)$$

We further assume a constant *noninformative* prior $P(\rm model)$ for all variables within the given bounds.

The calculation of the posterior distribution is performed using the *adaptive* Metropolis et al. (1953) (AM) algorithm of Haario et al. (2001). The AM method adapts the multivariate Gaussian proposal distribution during the Markov chain Monte Carlo sampling, in such a way that the Gaussian proposal distribution has the same non-diagonal covariance matrix as the posterior distribution accumulated so far by the algorithm. This natural idea is similar to what is routinely done e.g. in the determination of cosmological parameters, where the covariance matrix of the posterior is calculated after a burn-in phase (e.g. Dunkley et al. 2005). However, the adaptive approach converges much more rapidly as the proposal
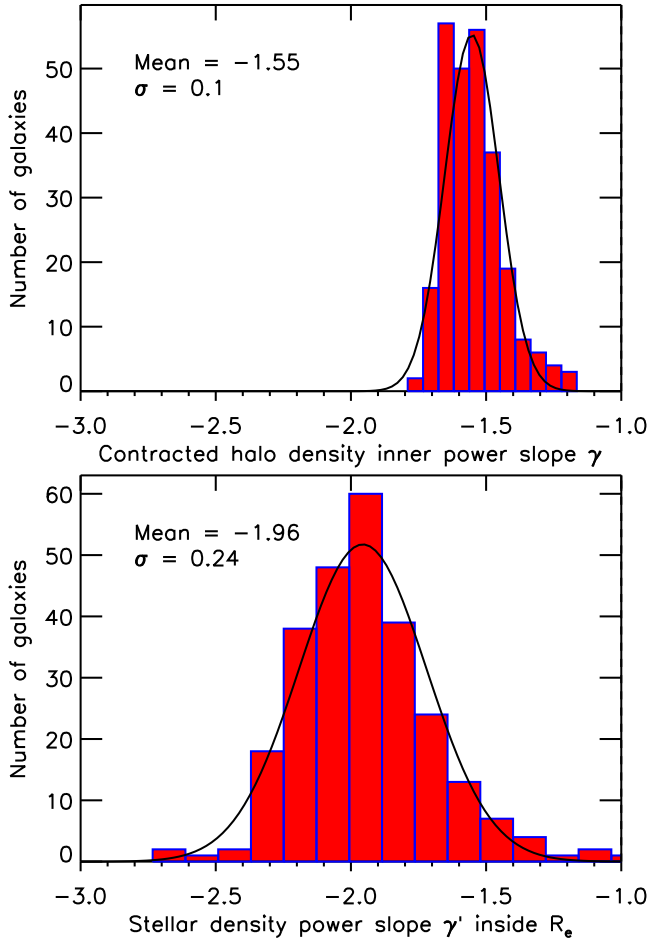
Figure 2. Inner slope of contracted dark halos and luminous matter. Top Panel: Histogram of the halo slope of contracted halos for all 260 ATLAS^{3D} galaxies in model (C). The slopes were determined by fitting a power law relation $\rho_{\rm DM}(r) \propto r^{\gamma}$ inside the radius $r < r_s/4$, where we verified the contracted halo profiles are accurately described by a power law. Bottom Panel: Histogram of the slope of the deprojected stellar mass density distribution from the MGE models. The slope was fitted inside a spherical radius $r = R_{\rm e}$. Although the stellar density $\rho_\star(r) \propto r^{\gamma'}$ inside that radius is not always accurately described by a power-law, on average the *stellar* slope peaks with high accuracy at at the 'isothermal' value $\gamma' \approx 2.0$, with an intrinsic scatter of just $\sigma = 0.24$ for our entire sample.

distribution starts approaching the posterior already after a few points have been sampled. We found the adaptive approach absolutely critical for the speed up of our calculation by orders of magnitudes, given the strong degeneracies between the model parameters producing inclined and narrow posterior distributions. Some examples of the posterior distributions obtained with our approach are shown in Fig. 3. Although the adaptive nature of the AM algorithm makes the resulting chain non-Markovian, their authors have proven that it has the correct ergodic properties (Haario et al. 2001) and for this reason it can be used to estimate the posterior distribution as in standard Markov chain Monte Carlo methods (Gilks et al. 1996).

Moreover, to basically eliminate the burn-in phase of the AM method, we use the efficient and extremely robust DIRECT deterministic global optimization algorithm of Jones et al. (1993) to find the starting location without the risk for the Metropolis stage

to be stuck in a possible secondary minimum in multi-dimensional parameter space.

An important addition to the fitting process is an iterative sigma clipping of the kinematics, to remove spurious features in the data like stars or problematic bins at the edge of the SAURON field of view. This is important for a sample of the size of ATLAS^{3D}, where the quality of every Voronoi bin cannot be assessed manually for all galaxies. After an initial fit the few bins deviating more than 3σ of the local rms noise are excluded from the fit and a new fit is iteratively performed, until convergence.

3.2 Robust fitting of lines or planes to the data

3.2.1 Goodness of fit criteria

The apparently simple task of fitting linear relations or planes to a set of data with errors does not have a well defined and obvious solution and for this reason has continued to generate significant interest. A number of papers have discussed the solution of the corresponding least-squares problem (Isobe et al. 1986; Feigelson & Babu 1992; Akritas & Bershady 1996; Tremaine et al. 2002; Press et al. 2007), while more recent works have addressed the problem using Bayesian methods (Kelly 2007; Hogg et al. 2010). A popular method is the least-squares approach by Tremaine et al. (2002), which is an extension of the FITEXY procedure described in Press et al. (2007, section 15.3). The method defines the best fit of the linear relation $y = a + b(x - x_0)$ to a set of N pairs of quantities (x_j, y_j), with symmetric errors Δx_j and Δy_j, as the one that minimizes the quantity

$$\chi^2 = \sum_{j=1}^{N} \frac{[a + b(x_j - x_0) - y_j]^2}{(b\Delta x_j)^2 + (\Delta y_j)^2 + \epsilon_y^2}. \tag{6}$$

Here x_0 is an adopted reference value, close to the middle of the x_j values, adopted to reduce uncertainty in a and the covariance between the fitted values of a and b. While ϵ_y is the intrinsic scatter in the y coordinate, which is iteratively adjusted so that the χ^2 per degree of freedom $\nu = N - 2$ has the value of unity expected for a good fit. As recognized by Weiner et al. (2006), minimizing the above χ^2 corresponds to maximizing the likelihood of the data for an assumed intrinsic probability distribution of the observables described by the linear relation $y = a + b(x - x_0) + \epsilon_y$, where ϵ_y is the Gaussian scatter projected along the y coordinate, and one assumes a uniform prior in the x coordinate. equation (6) is only rigorously valid when the errors in x and y are Gaussian and uncorrelated (have zero covariances). A term $-2b \, {\rm Cov}(x_j, y_j)$ should be included in the denominator if the covariances are known and nonzero (e.g. Falcón-Barroso et al. 2011). The 1σ confidence interval in ϵ_y can be obtained by finding the values for which $\chi^2 = \nu \pm \sqrt{2\nu}$ as done by Novak et al. (2006). The apparent asymmetry of equation (6) with respect to the x and y variables does not imply we assume only the y variable has intrinsic scatter. In fact the assumed intrinsic distribution has a Gaussian cross section along any direction non parallel to the ridge line $y = a + b(x - x_0)$. The value ϵ_y merely specifies the dispersion along the arbitrary y direction. The formula would give completely equivalent results by interchanging the x and y variables if the distribution of x values was uniform and infinitely extended as assumed. Any difference in the fitting results when interchanging the x and y coordinates are due to the breaking of the uniformity assumptions.

equation (6) can be generalized to plane fitting by defining the best-fitting plane $z = a + b(x - x_0) + c(y - y_0)$ to a set of N
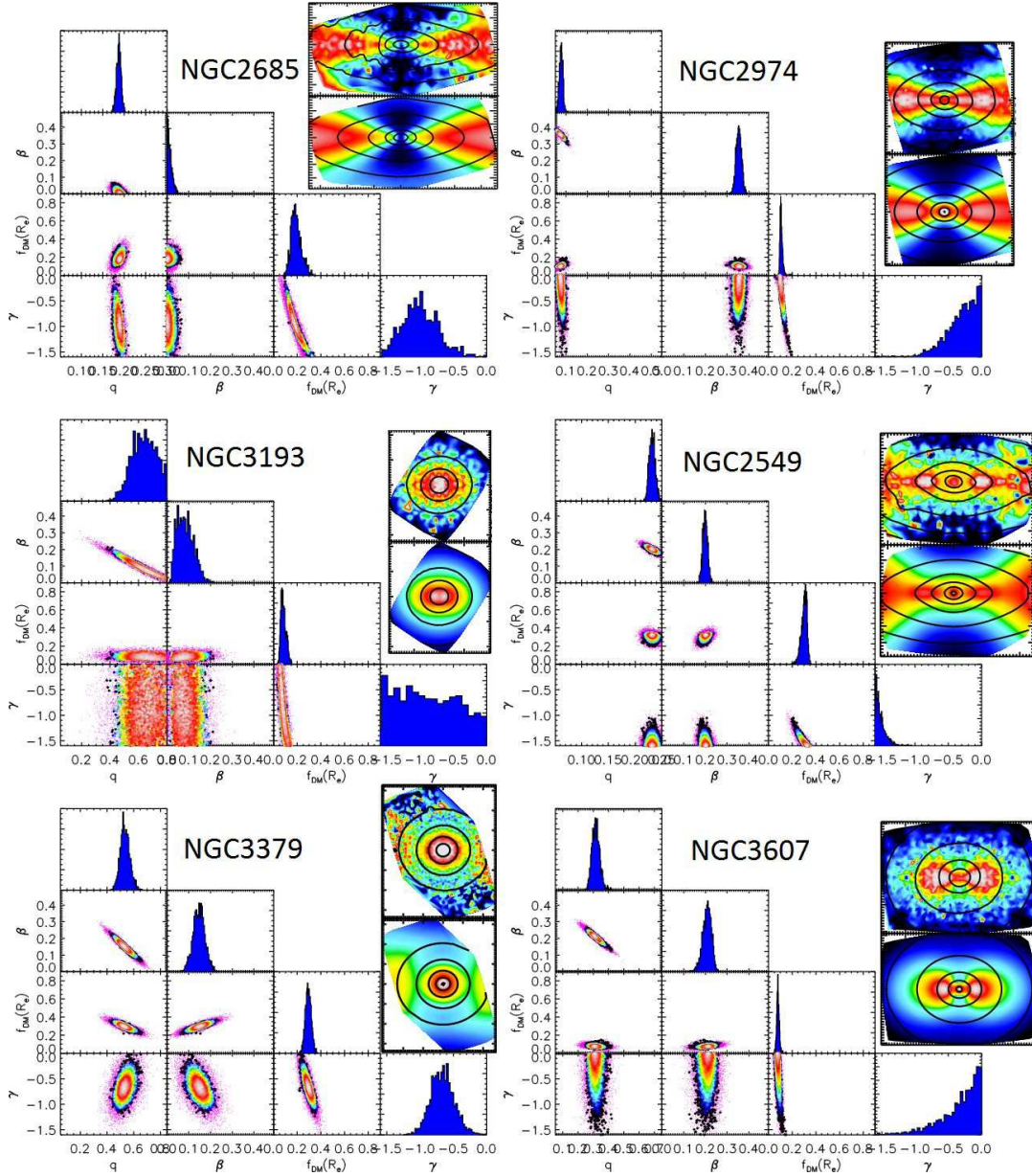
Figure 3. Examples of JAM dynamical modelling with general dark halo using Adaptive Metropolis. Each panel shows the posterior probability distribution for the parameters $(q, \beta_z, f_{\rm DM}, \gamma)$, using galaxy model (E) marginalized over two dimensions (colour contours) and one dimension (blue histograms). For every non-linear parameters combination the $(M/L)_{\rm stars}$ is linearly scaled to fit the data. We assumed ignorant (constant) priors on all model parameters. The name of the galaxies is written next to each panel. The symmetrized $V_{\rm rms}$ SAURON data, and the best-fitting model are shown on the right (as in Fig. 1). This plot illustrates a variety of situations and shapes of the kinematic field: (i) some models (NGC 2685 and NGC 3379) are well constrained within the explored parameters boundaries and have preferred slope $\gamma \approx -1$ like NFW; (ii) others (NGC 2774 and NGC3607) prefer a flat $\gamma \approx 0$ or even positive values; (iii) others (NGC 3193) have nearly unconstrained halo slope; (iv) others (NGC 2549) prefer steep halo slopes at the boundary $\gamma = -1.6$ of our allowed parameter range. In all cases there is a strong degeneracy in the halo slope, but the dark matter fraction is tightly constrained by the data to be small ($f_{\rm DM} \lesssim 30\%$ in these examples). Only allowing the inner halo slope to match the stellar density $\gamma' \approx -2.0$ could significant dark matter be included in some of the models.

triplets of quantities (x_j, y_j, z_j), with symmetric errors Δx_j, Δy_j and Δz_j, as the one that minimizes the quantity

$$\chi^2 = \sum_{j=1}^{N} \frac{[a + b(x_j - x_0) + c(y_j - y_0) - z_j]^2}{(b\Delta x_j)^2 + (c\Delta y_j)^2 + (\Delta z_j)^2 + \epsilon_z^2}, \quad (7)$$

Here x_0 and y_0 are adopted reference values, close to the middle of the x_j and y_j values respectively, adopted to reduce uncertainty in a and the covariance between the fitted values of a, b and c.

While ϵ_z is the intrinsic scatter in the z coordinate, which is iteratively adjusted so that the χ^2 per degrees of freedom $\nu = N - 3$ has the value of unity expected for a good fit. As in the two-dimensional case the minimization of equation (7) is equivalent to the maximization of the likelihood of the data, for an underlying probability distribution of the observables described by the relation $z = a + b(x - x_0) + c(y - y_0) + \epsilon_z$, where ϵ_z is the dispersion of the Gaussian intrinsic scatter in the plane, projected along the z coordinate, for a uniform prior in the x and y coordinates and assum-

ing uncorrelated and Gaussian errors in the x, y and z observables (zero covariances). equation (7) reduces to the so called orthogonal plane fit when the measurements errors are ignored and one simply assumes $\Delta x_j = \Delta y_j = \Delta z_j$. This latter form is the one generally used when fitting the Fundamental Plane (e.g. Jorgensen et al. 1996; Pahre et al. 1998; Bernardi et al. 2003). Contrary to the popular approach, equation (7) allows for intrinsic scatter in the relation, which is important to deriving unbiased parameters (Tremaine et al. 2002).

Recently Kelly (2007) proposed a Bayesian method to treat the linear regression of astronomical data in a statistically rigorous manner, allowing for intrinsic scatter, covariance between measurements and providing rigorous errors on the parameters in the form of random draws from the posteriori distribution (see also Hogg et al. 2010). He pointed out that the Tremaine et al. (2002) approach to linear fitting can lead to biased results in some circumstances. For this reason in all our fits we used both the results and errors derived from equation (6) and (7), and the corresponding results obtained with the Bayesian method and software by Kelly (2007), which was kindly made available as part of the IDL NASA Astronomy Library (Landsman 1993). In all cases differences between the two method where found to be insignificant, in both the fitted values and the errors, confirming the near conceptual equivalence of the two technically very different approaches.

3.2.2 Least Trimmed Squares robust fits

A general issue when fitting linear relations to data using least-squares methods is the presence of outliers, which can dominate the χ^2 and bias the parameter recovery. This is the reason why a number of previous studies have determined the parameters of the Fundamental Plane using the more robust method of minimizing absolute instead of squared deviations (e.g. Jorgensen et al. 1996; Pahre et al. 1998), at the expense of decreasing the statistical efficiency, namely larger errors on the fitted parameters. An alternative simple solution, which maintains the efficiency of the least-squares method for Gaussian distributions, consists of removing outliers by iteratively clipping points deviating more than 3σ from the currently best-fitting relation. A problem with the σ-clipping approach is that it is not guaranteed to converge to the desired solution in the presence of significant outliers. Alternative robust methods have been proposed (see Press et al. 2007, section 15.7). However, they complicate the error estimation and like the standard σ-clipping do not always converge.

After some experimentation with different robust approaches the only fully satisfactory solution we found is the Least Trimmed Squares (LTS) regression approach of Rousseeuw & Leroy (1987). The reason for its success is that the method, as opposed to other robust approaches, finds a *global* solution. The approach consists of finding the global minimum to

$$\chi_h^2 = \sum_{j=1}^{h} (r^2)_{j:N},\qquad(8)$$

where $(r^2)_{1:N} \leq (r^2)_{2:N} \leq \ldots \leq (r^2)_{N:N}$ are the ordered square residuals from the linear regression of a subset of $N/2 < h < N$ data points. In other words the LTS method consists of finding the subset of h data points providing the smallest χ_h^2, among *all* possible h-subsets. It's easy to realize that this approach is robust to the contamination of up to half of the data points, when $h \approx N/2$. This is a computational very expensive combinatorial problem for

which however a fast and nearly optimal solution (FAST-LTS) has recently been proposed by Rousseeuw & Van Driessen (2006).

In our implementations, which we called LTS_LINEFIT and LTS_PLANEFIT for the line and plane case respectively, we proceed as follows:

(i) We adopt as initial guess $\epsilon = 0$ for the intrinsic scatter in the y (for LTS_LINEFIT) or z coordinate (for LTS_PLANEFIT);

(ii) We start by default with $h = (N + p + 1)/2$, where p is the data dimension, and use the FAST-LTS algorithm to produce a least-squares fit[2]. to the set of points characterized by the smallest χ_h^2 (defined by equation 6 or 7);

(iii) We compute the standard deviation σ of the residuals for these h values and extend our selection to include all data point deviating no more than $2.6\,\sigma$ from the fitted relation (99% of the values for a Gaussian distribution);

(iv) We perform a new linear fit to the newly selected points;

(v) We iterate steps (iii)–(iv) until the set of selected points does not change any more;

(vi) We calculate the χ^2 for the fitted points;

(vii) The whole process (i)–(vi) is iterated varying ϵ using Brent's method (Press et al. 2007, section 9.3) until $\chi^2 = \nu$.

(viii) The errors on the coefficients are computed from the covariance matrix;

(ix) The error on ϵ is computed by increasing ϵ until $\chi^2 = \nu - \sqrt{2\nu}$ (we do not decrease it to avoid problems when $\epsilon \approx 0$).

This method was used to produce all fits in this paper and automatically exclude outliers. Note that although the approach may appear similar to the standard σ clipping one, and produces similar results in simple situations, the key difference is that in LTS_LINEFIT and LTS_PLANEFIT the clipping is done from the inside-out instead of the opposite. This was found to be the essential feature for the resulting extreme robustness, which was essential in particular to objectively select Virgo members in Fig. 15. Once the outliers are removed, the same set of points was used as input to Kelly (2007) Bayesian algorithm.

3.3 Measuring scaling relations parameters

3.3.1 Determination of L, R_e and $r_{1/2}$ from the MGE

Galaxy photometric parameters are generally determined using three main approaches: (i) fitting growth curves, where one constructs profiles of the enclosed light within circular annuli and extrapolates the outermost part of the galaxy profile to infinite radius, typically using the analytic growth curve of the $R^{1/4}$ (de Vaucouleurs 1948) profile (e.g. the Seven Samurai: Burstein et al. 1987 and Faber et al. 1989; the RC3: de Vaucouleurs et al. 1991 and Jorgensen et al. 1995a); (ii) fitting an $R^{1/n}$ (Sersic 1968) profile (e.g. Graham & Colless 1997), possibly including an exponential disk (e.g. Saglia et al. 1997), to the circularized profiles and finding the half-light from the models; (iii) fitting flattened two-dimensional models directly to the galaxy images, where the profile of the models is again parameterized by an $R^{1/4}$ (e.e. Bernardi et al. 2003), or by an $R^{1/4}$ bulge plus exponential disk (e.g. Gebhardt et al. 2003; Saglia et al. 2010; Bernardi et al. 2010).

---

[2] In all the nonlinear fits the minimization was performed with the IDL program MPFIT by Markwardt (2009), which is in an improved implementation of the MINPACK Levenberg-Marquardt nonlinear least-squares algorithm by Moré et al. (1980)

Here we have MGE photometric models for all the galaxies in the sample based on the SDSS+INT photometry (Paper XXI). Due to the large number of Gaussians used to fit the galaxy images, the MGE models provide a compact and essentially non-parametric description of the galaxies surface brightness, which reproduces the observations much more accurately than the simpler bulge and disk models, but more robustly than using the images directly. Our MGE fitting approach is in fact analogue to the popular GALFIT (Peng et al. 2002) software, when it is used to match every detail of a galaxy image using multiple components. Here we use the MGE models to measure the photometric parameters (L and R_{e}) in our scaling relations as done in Cappellari et al. (2009). A significant difference between this MGE approach and all the ones previously mentioned is that it does not extrapolate the galaxy light to infinite radii. Outside three times the dispersion $3\max(\sigma_j)$ of the largest MGE Gaussian, the flux of the model essentially drops to zero. No attempt is made to infer the amount of stellar light that we may have observed if we had much deeper photometry. For this reason this R_{e} must be necessarily smaller than the ones obtained via extrapolation to infinite radii. Earlier indications using deeper MegaCam photometry, which we have acquired for many of the galaxies in our sample (Duc et al. 2011)[hereafter Paper IX], seems to confirm that R_{e} determinations have to be used with caution and depend sensitively on the depth of the adopted photometry.

Our method has the advantage that it extracts R_{e} directly from the observations: it measures the radius enclosing half of the *observed* galaxy light. The extrapolation method depends on the assumed form of the unobservable galaxy profile out to infinite radii. This is perhaps sensible for genuine spheroidal and single-component elliptical galaxies, which are thought to be well described by Sersic (1968) profiles. But our volume-limited sample of ETGs is dominated by fast rotators (Paper II; Paper III), characterized by the presence of disks (Krajnović et al. 2012, hereafter Paper XVII) and closely linked to spiral galaxies (Cappellari et al. 2011b, hereafter Paper VII; Paper XX). Given the variety in the outer profiles of spiral galaxies (van der Kruit & Searle 1981; Pohlen & Trujillo 2006) it seems safer to make our results independent on the adopted profile extrapolation, but to use a quantity that can be directly inferred from the data. Neither the extrapolation method, nor our truncated one can claim to measure the 'true' R_{e} or the galaxies.

If the x-axis is aligned with the galaxy photometric major axis, and the coordinates are centered on the galaxy nucleus, the surface brightness of an MGE model at the position (x', y') on the plane of the sky, already analytically deconvolved for the atmospheric seeing effects, can be written as (Emsellem et al. 1994)

$$\Sigma(x', y') = \sum_{j=1}^{M} \Sigma_j \exp\left[-\frac{1}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{q_j'^2}\right)\right], \qquad (9)$$

where M is the number of the adopted Gaussian components, having peak surface brightness Σ_j, observed axial ratio $0 \le q_k' \le 1$ and dispersion σ_k along the major axis. The total luminosity of the MGE model is then:

$$L = \sum_{j=1}^{M} L_j = \sum_{j=1}^{M} 2\pi \Sigma_j \sigma_j^2 q_j', \qquad (10)$$

where L_j are the luminosities of the individual Gaussians.

In Cappellari et al. (2009) the effective radius of the MGE model was obtained by circularizing the individual Gaussians of the MGE, while keeping their peak surface brightness. This was achieved by replacing (σ_j, q_j') with $(\sigma\sqrt{q_j'}, 1)$. The luminosity of

the circularized MGE enclosed within a cylinder of projected radius R is then

$$L(R) = \sum_{j=1}^{M} L_j\left[1 - \exp\left(-\frac{R^2}{2\sigma_j^2 q_j'}\right)\right]. \qquad (11)$$

The circularized effective (half-light) radius R_{e} was found by solving $L(R) = L/2$, using a quick interpolation over a grid of $\log R$ values. When the MGE has constant axial ratio $q_j' = q'$ for all Gaussians, this approach finds the circularized radius $R_{\mathrm{e}} = \sqrt{ab} = a\sqrt{q'}$ of the elliptical isophote containing half of the analytic MGE light, where a is the major axis of the isophote. This is the quantity almost universally used for studies of scaling relations of ETGs. When the axial ratios of the different Gaussians are not all equal, the approach finds an excellent approximation for the radius $R_{\mathrm{e}} = \sqrt{A_e/\pi}$ of a circle having the same area A_e as the isophote containing half of the MGE light. In fact we verified that for all the MGE of the ATLAS$^{\mathrm{3D}}$ sample the two determinations agree with an rms scatter of just 0.17% and only for four of the flattest galaxies the difference is larger than 3%.

Hopkins et al. (2010) pointed out the usefulness of adopting as size parameter the major axis $R_e^{\max}$ of the half-light isophote instead of the circularized radius R_{e}, when analysing results of simulations. The motivation is that $R_e^{\max}$ is more physically robust and less dependent on inclination. Here we also calculate $R_e^{\max}$ for our observed galaxies as follows.

 (i) We construct a synthetic galaxy image from the MGE using equation (9), with size $\max(\sigma_j) \times \max(\sigma_j)$ (only one quadrant is needed for symmetry);

 (ii) We sample a grid of surface-brightness values $\mu_k = \mu(x_k, 0)$ along the MGE major axis, and for each value we calculate the light enclosed within the corresponding isophote;

 (iii) We find the surface brightness μ_e of the isophote containing half of the analytic MGE total light by solving $L(\mu) = L/2$ using linear interpolation;

 (iv) $R_e^{\max}$ is the maximum radius enclosed inside the isophote μ_e (the largest x coordinate).

We also calculate the circularized effective radius of the isophote $R_{\mathrm{e}} = \sqrt{A/\pi}$ of area A and the effective ellipticity ε_e of the MGE model inside that isophote as (Cappellari et al. 2007)

$$(1 - \varepsilon_e)^2 = q_e'^2 = \frac{\langle y^2 \rangle}{\langle x^2 \rangle} = \frac{\sum_{k=1}^{P} F_k\, y_k^2}{\sum_{k=1}^{P} F_k\, x_k^2}, \qquad (12)$$

where F_k is the flux inside the k-th image pixel, with coordinates (x_k, y_k) and the summation extends to the pixels inside the chose isophote. A similar quantity was calculated from the original galaxy images in Paper III, but we use here this new determination for maximum consistency between our ϵ_e and the ellipticity of the MGE models in the tests of Fig. 4.

We studied the dependence on inclination of the two definitions of effective radii using the photometry of real galaxies. For this we selected the 26 flattest galaxies in our sample, all having axial ratio $q' < 0.4$. These galaxies are likely to be close to edge-on. We assume they are exactly edge-on and we then use the MGE formalism (equations 9, 13 and 14) to deproject the surface brightness and calculate the intrinsic luminosity density. We then project it back on the sky plane at different inclinations, from edge on ($i = 90°$) to face on ($i = 0°$). At every inclination we calculate the two effective radii R_{e} and $R_e^{\max}$ (Fig. 5). The comparison shows that, as expected, the R_{e} of flattened objects can be much smaller when objects are edge-on than face-on, with a median decrease of
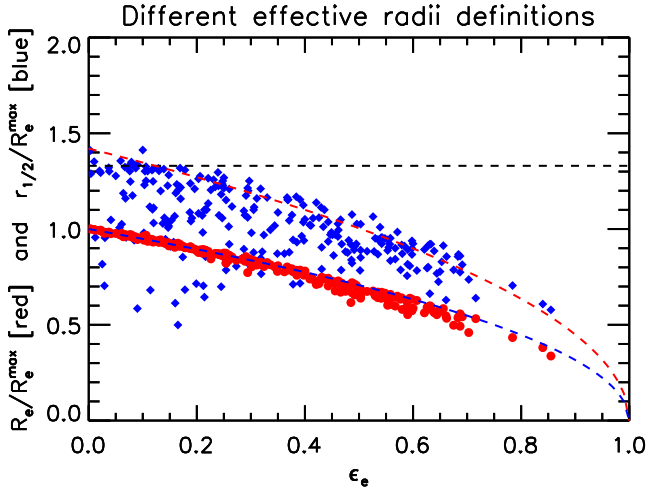
Figure 4. Different definitions of R_e as a function of the galaxy ellipticity. The red filled diamonds are the projected radii R_e of a cylinder with the same area of the half-light isophote. The blue filled circles are the radii $r_{1/2}$ of a sphere with the same volume as the half-light iso-surface. In both cases the radii are normalized to R_e^{max}, which is the projected semi-major axis of the half-light isophote, having ellipse of inertia of ellipticity ϵ_e. The red and blue dashed lines are the relations $f(\epsilon_e) = 1.42\sqrt{\epsilon_e}$ and $f(\epsilon_e) = \sqrt{\epsilon_e}$ respectively. The horizontal dashed line marks the theoretical value $4/3$, which approximately applies to a number of simple theoretical profiles.
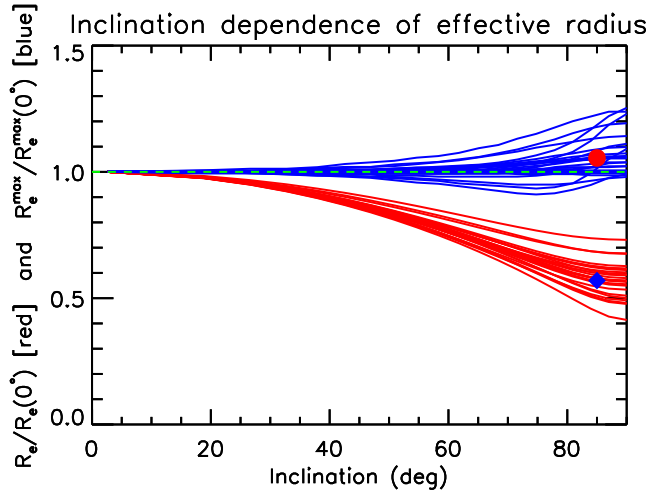
Figure 5. Inclination dependence for different definitions of the effective radius. The red lines show the change in the measured circularized R_e, normalized to the face-on value, when the inclination is changed from edge-on ($i = 90°$) to face-on, for the 26 flattest ATLAS³D galaxies. The blue diamond marks the median (43%) of the maximum variation. The blue lines show the same variation with inclination of the major axis R_e^{max} of the half-light isophote. The red circle is the median (5%) of the maximum variation.

43% (0.24 dex). The opposite is true for R_e^{max}, but the variations are dramatically smaller, with a median increase of 5% (0.02 dex). The two effective radii of course are the same for intrinsically spherical objects. The use of R_e^{max} instead of R_e is especially useful when one considers that 86% of the galaxies in ATLAS³D (and in the nearby Universe) are disk like (Paper II, III and VII).

In what follows we also need the radius $r_{1/2}$ of a sphere enclosing half of the galaxy light. For this we need to derive the intrinsic galaxy luminosity density from the MGE, assuming the best

fitting inclination of the JAM models. A possible deprojection of the observed MGE surface brightness can be derived analytically by deprojecting the individual Gaussians separately (Monnet et al. 1992). The solution is only unique when the galaxy is edge-on (Rybicki 1987). The deprojected luminosity density ν is given by

$$\nu(R, z) = \sum_{k=1}^{M} \frac{\Sigma_j q_j'}{\sqrt{2\pi}\,\sigma_j q_j} \exp\left[-\frac{1}{2\sigma_j^2}\left(R^2 + \frac{z^2}{q_j^2}\right)\right], \qquad (13)$$

where the individual components have the same dispersion σ_j as in the projected case (9), and the intrinsic axial ratio of each Gaussian becomes

$$q_j = \frac{\sqrt{q_j'^2 - \cos^2 i}}{\sin i}, \qquad (14)$$

where i is the galaxy inclination ($i = 90°$ being edge-on). To calculate $r_{1/2}$ from the intrinsic density of equation (13) one can proceed analogously to the approach used to measure the circularized R_e. This is done by making the three-dimensional MGE distribution spherical, while keeping the same total luminosity and peak luminosity density of each Gaussian. This is achieved by replacing (σ_j, q_j) with $(\sigma\, q_j^{1/3}, 1)$. The light of this new spherical MGE enclosed within a sphere of radius r is given by

$$L(r) = \sum_{j=1}^{M} L_j \left[\mathrm{erf}(h_j) - 2h_j \exp(-h_j^2)/\sqrt{\pi}\right], \qquad (15)$$

with $h_j = r/(\sqrt{2}\,\sigma_j\, q_j^{1/3})$ and erf the error function. And the half-light spherical radius $r_{1/2}$ is obtained by solving $L(r) = L/2$ by interpolation. As in the projected case, when all Gaussians have the same $q_j = q$, which means the density is stratified on similar oblate spheroids, the method gives the geometric radius $r_{1/2} = (abc)^{1/3} = a\,q^{1/3}$, where a is the semi-major axis of the spheroid. While when the q_j are different, this radius provides a very good approximation to the radius $r_{1/2} = [3V_e/(4\pi)]^{1/3}$ of a sphere that has the same volume V_e of the iso-surface enclosing half of the total galaxy light.

In Fig. 4 we compare the three definitions of R_e as a function of the observed effective ellipticity ϵ_e of the MGE, for all the galaxies in the ATLAS³D sample. Even though the galaxy isophotes are in most cases not well approximated by ellipses, and the galaxies are intrinsically not oblate spheroids, the ratio between R_e and R_e^{max} follows the relation for elliptical isophotes. When the galaxies are very close to circular on the sky R_e and R_e^{max} agree by definition. The situation is very different regarding the relation between $r_{1/2}$ and R_e^{max}. In this case, when the galaxy is edge-on, there is a simple ratio $r_{1/2}/R_e \approx 1.42$, but when the galaxies have lower inclinations, large variations in the ratio are possible, so that $r_{1/2}$ cannot be inferred from the observations, without the knowledge of the galaxy inclination, which generally require dynamical models. The situation is of course much simpler for spherical objects, in which case $r_{1/2}/R_e \approx 1.42$ as in the edge-on case. For comparison Hernquist (1990) found the theoretical value $r_{1/2}/R_e \approx 1.33$ for his spherical models, while Ciotti (1991) has shown that for a $R^{1/m}$ model the ratio is confined between 1.34–1.36, when $m = 2 - 10$, and the same applies to other simple profiles (Wolf et al. 2010). As expected our ratio is slightly larger, given that our models, like real galaxies, do not extend to infinite radii. For flatter models the cylindrical and spherical circularized radii are approximately related as $R_e/R_e^{max} = \sqrt{\epsilon_e}$, which one would expect for elliptical isophotes while the ratio $r_{1/2}/R_e$ remains approximately constant.

3.3.2 Comparing effective and gravitational radius

For an isolated spherical system in steady state one obtains from the scalar virial theorem (Binney & Tremaine 2008)

$$M = \frac{r_g \langle v^2 \rangle_\infty}{G}, \qquad (16)$$

where r_g is defined as the gravitational radius, which depends on the total and luminous mass distribution, M is the galaxy total luminous plus dark mass and $\langle v^2 \rangle_\infty$ is the mean-square speed of the galaxy stars, integrated over the full extent of the galaxy. In the spherical case $\langle v^2 \rangle_\infty = 3 \langle \sigma^2_{\mathrm{los}} \rangle_\infty$ and

$$M = 3 \frac{r_g \langle \sigma^2_{\mathrm{los}} \rangle_\infty}{G}. \qquad (17)$$

This formula is *rigorously* independent of anisotropy and only depends on the radial profiles of luminous and dark matter (Binney & Tremaine 2008, section 4.8.3).

When the spherical system is self-consistent ($L(r) \propto M(r)$) the gravitational radius can be easily calculated as

$$r_g = \frac{2L^2}{\int_0^\infty [L(r)/r]^2 \mathrm{d}r}. \qquad (18)$$

Here we evaluate this expression using a single numerical quadrature via equation (15), from the same spherical deprojected MGE we used in the previous Section to calculate $r_{1/2}$. The MGE is obtained by deprojecting the observed surface brightness at the JAM inclination and subsequently making the MGE spherical while keeping the same peak stellar density and luminosity of every Gaussian. In this way our calculation of r_g is rigorously accurate when the MGE is already spherical, while the formula provides a good approximation for flattened galaxies.

In Fig. 6 we plot the ratio $r_{1/2}/r_g$, for the full ATLAS[3D] sample as a function of the non-parametric Third Galaxy Concentration (TGC) defined in Trujillo et al. (2001) as the ratio between the light $L(R_\mathrm{e}) = L/2$ enclosed within an isophote of radius R_e and the one $L(R_\mathrm{e}/3)$ enclosed within an isophote with radius $R_\mathrm{e}/3$. Graham et al. (2001) have shown that this choice leads to a more robust measure of concentration than popular alternatives (e.g. Doi et al. 1993). We compute the TGC from the circularized MGE using equation (11), as done for R_e. We find a trend in the ratio for the galaxies in our sample that varies between $r_{1/2}/r_g \approx 0.3 - 0.4$ for the range of galaxy concentrations we observed. For comparison we also calculate the TGC and the corresponding r_g for spherical models described by the $R^{1/m}$ profile (Sersic 1968). This was done by constructing analytic profiles, truncating them to $R < 4R_\mathrm{e}$, to mimic the depth of the SDSS photometry, before fitting them with the one-dimensional MGE-fitting procedure of Cappellari (2002). Both TGC and $r_{1/2}/r_g$ span the ranges predicted for profiles with $m = 2 - 6$. Our trend in the ratio is more significant than the generally assumed near constancy around 0.40 ± 0.02, first reported by Spitzer (1969) for different polytropes, which agrees with the theoretical value $r_{1/2}/r_g = (1 + \sqrt{2})/6 \approx 0.402$ for a Hernquist (1990) profile (Mamon 2000; Łokas & Mamon 2001). However, the variation is indeed rather small, being only at the $\pm 15\%$ level around a median value of 0.35 in our sample.

The relatively small variations of the ratio between the gravitational and intrinsic $r_{1/2}$ or projected R_e half-light radii, explain the usefulness of the latter two parameters in measuring dynamical scaling relations of galaxies. This fact, combined with the rigorous independence on anisotropy, also explains the robustness of a mass estimator like
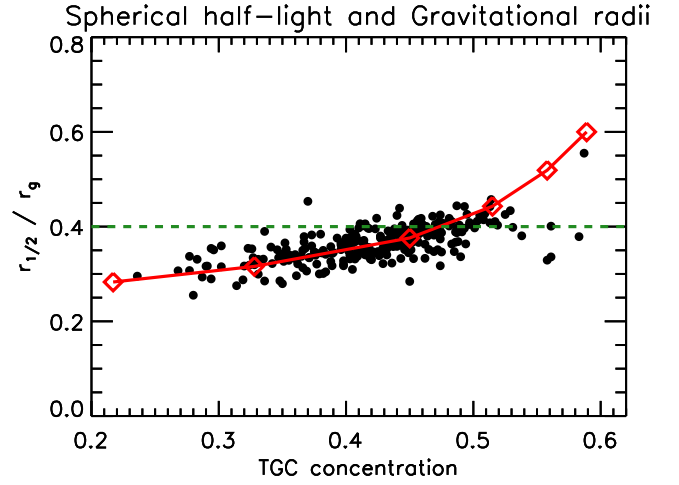
Figure 6. The black filled circles mark the ratio $r_{1/2}/r_g$ between the radius of the half-light sphere and the gravitational radius for all the galaxies in the sample. For comparison the solid red line indicates the same ratio for a spherical galaxy with an $R^{1/m}$ surface brightness profile. From left to right the red diamonds mark the locations $m = 1, 2, 4, 6, 8, 10$ respectively. The green dashed horizontal line indicates the theoretical value for a Hernquist (1990) profile.

$$M_{1/2} = k \frac{r_{1/2} \langle \sigma^2_{\mathrm{los}} \rangle_\infty}{G}, \qquad (19)$$

when the stellar systems can be assumed to be spherical and kinematics is available over the *entire extent* of the system, as pointed out by Wolf et al. (2010). Assuming the measured ratio $r_{1/2}/r_g \approx 0.4$ for galaxies with the approximate concentration of an $R^{1/4}$ profile, already in the self-consistent limit the expected coefficient is $k \approx 3/0.4/2 = 3.75$, which is close, but 25% larger than the corresponding coefficient $k = 3$ proposed by Wolf et al. (2010). However, the ratio $r_{1/2}/r_g$ we empirically measured on real galaxies, does not assume the outermost galaxy profiles are known and can be extrapolated to infinity, so it weakly depends on the depth of the photometry. For example, for a spherical galaxy that follows the $R^{1/4}$ profile to infinity, we obtain $r_{1/2}/r_g = 0.456$, which would imply $k = 3.29$ in the self-consistent limit. The remaining 10% difference from Wolf et al. (2010) is easily explained by the small increase of $\langle \sigma^2_{\mathrm{los}} \rangle_\infty$ due to the inclusion of a dark halo.

3.3.3 Determination of σ_e

Unfortunately the quantities $\langle v^2_{\mathrm{los}} \rangle_\infty$, or $\langle \sigma^2_{\mathrm{los}} \rangle_\infty$ are currently only observable via discrete tracers in objects like nearby dwarf spheroidal (dSph) galaxies (e.g. Walker et al. 2007), but it is still not a directly observable quantity in early-type galaxies. Nonetheless Cappellari et al. (2006) showed that in practice $\langle v^2_{\mathrm{los}} \rangle_e$, as approximated by σ_e, which can be empirically measured for large samples of galaxies, can still be used to derive robust central masses when applied to real, non-spherical ETGs, with kinematics extended to about $1R_\mathrm{e}$:

$$(M/L)[r = R_\mathrm{e}] \approx 5.0 \times \frac{R_\mathrm{e} \sigma^2_e}{GL}, \qquad (20)$$

where $(M/L)[r = R_\mathrm{e}]$ is estimated inside an iso-surface of volume $V = 4\pi R_\mathrm{e}^3/3$ (a sphere of radius R_e if the galaxy is spherical), and σ_e is the velocity dispersion calculated within a projected circular aperture of radius R_e. In this paper we improve

on the previous approach by measuring σ_e inside an effective ellipse instead of a circle. The ellipse has area $A = \pi R_e^2$ and ellipticity ε_e. The measurement is done by co-adding the luminosity-weighted spectra inside the elliptical aperture and measuring the σ of that effective spectrum using pPXF (Cappellari & Emsellem 2004). Due to the co-addition, the resulting spectrum has extremely high S/N (often above 300) and this makes the measurement robust and accurate. When the SAURON data do not fully cover R_e we correct the σ_e to $1R_e$ using equation (1) of Cappellari et al. (2006). σ_e has the big advantage over $\langle v_{\rm los}^2 \rangle_e$ that it can also be much more easily measured at high redshift, as it does not require spatially resolved kinematics. Integrated stellar velocity dispersions have started to become measurable up to redshift $z \approx 2$ (Cenarro & Trujillo 2009; Cappellari et al. 2009; van Dokkum et al. 2009; Onodera et al. 2010; van de Sande et al. 2011). Moreover the advantage of σ_e over the traditional central dispersion σ_c, is that it is empirically closer to the true second velocity moment $\langle v_{\rm los}^2 \rangle_\infty$ that appears in the virial equation (17) and is directly proportional to mass. Making the good approximation $(M/L)[r = R_e] \approx (M/L)[r = r_{1/2}]$, where $r_{1/2} \approx 1.33 R_e$, one can rewrite equation (20) in a form that is directly comparable to equation (19)

$$M_{1/2} \approx 2.5 \times \frac{R_e \sigma_e^2}{G} \approx 1.9 \times \frac{r_{1/2}\, \sigma_e^2}{G}. \qquad (21)$$

Note that the empirical coefficient 1.9 is significantly smaller than the value around 3.0 one predicts when using $\langle \sigma_{\rm los}^2 \rangle_\infty$ in equation (19) and we will come back to this point in Section 4.3.

4   RESULTS

4.1   Uncertainty in the scaling relations parameters

4.1.1   Errors in L, R_e and σ

In the study of galaxy scaling relations formal errors on L, R_e and σ are often adopted, as given in output by the program used for their extraction. These errors assume the uncertainties are of statistical nature. However, in many realistic situations the systematic errors are significant, but difficult to estimate. In this work, the availability of a significant sample of objects, with similar quantities measured via independent data or methods, allow for a direct comparison of quantities. This external comparison permits us to include systematic errors into our adopted errors, instead of just using formal or Monte Carlo errors.

In Paper XXI we compare the total magnitude M_r of the MGE model, as derived from the SDSS+INT r-band photometry to various other sources in the literature. We conclude that our total M_r are accurate at the 10% level, in the relative sense. This is the error we adopted in what follows. This accuracy is comparable to other state-of-the-art photometric surveys.

A comparison between the circularized half-light radii R_e of Paper I and the circularized R_e from the r-band MGE is shown in Fig. 7. In this case the scatter is of 0.058 dex, which would imply errors of 10% in the individual R_e. This must still be a firm upper limit to the errors, given that any relative variations, among galaxies, in the colour gradients in r and K_s will increase the scatter. Remarkably in this case our scatter between SDSS r-band and 2MASS K_s bands, for the entire sample, is as small as the best agreement (0.05 dex) reported by Chen et al. (2010), using the very same SDSS g-band photometry. We are not aware of other published independent R_e determinations from different data that agree
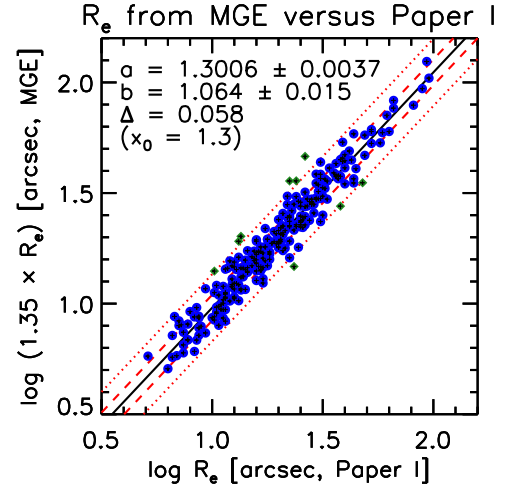
Figure 7. Testing the relative accuracy of size measurements. Comparison between the R_e from 2MASS plus RC3, matched to RC3 as described in Paper I, and the R_e from the MGEs. For a good match the MGE values have been *increased* by a significant factor 1.35. In what follows the effective radii will always already include this multiplicative factor. The coefficients of the best-fitting relation $y = a + b(x - x_0)$ and the corresponding observed scatter Δ in y are shown at the top left of the plot. The two red dashed and dotted lines mark the 1σ bands (enclosing 68% of the values for a Gaussian distribution) and 2.6σ (99%) respectively. The outliers automatically excluded from the fit by the LTS_LINEFIT procedure are shown as green diamonds.

with such a small scatter, and for such a large sample. The rms scatter we measure is twice smaller that their comparison in the same band between SDSS and ACSVCS. Our scatter is also twice smaller than a similar comparison we performed in Paper I between the R_e of 2MASS and RC3. We interpret the excellent reproducibility of our MGE R_e values, and the agreement with the values of Paper I, to the fact that in both 2MASS and our MGE models the total luminosities are *not* computed via a extrapolation of the profile to infinity, but truncated to the extent of the data. This result is a reminder of the fact that extrapolation is a dangerous practice, which should be avoided whenever possible. We argue that it is difficult to derive reproducible results on galaxy sizes, when one *assumes* the outer (unobservable) profile is accurately know. So for this work define R_e as the radius containing half of the *observed* light, not half of the ill-defined amount of total light we think the galaxy may have. Of course even our approach does not solve the problem of determining an absolute normalization of R_e, and our sizes appear well reproducible only in a relative sense.

However, a very important feature of Fig. 7 is the significant offset by a factor 1.35 between the MGE R_e and the values of Paper I, with the MGE values being smaller. In what follows all our MGE effective radii will always already include this multiplicative factor. The values of Paper I where determined from a combination of 2MASS (Skrutskie et al. 2006) and RC3 (de Vaucouleurs et al. 1991) R_e measures. But they were scaled to match on average the values of the RC3 catalogue, which were determined using growth curves extrapolated to infinity. The RC3 normalization agree within 5% with the SAURON determinations in (Cappellari et al. 2006; Kuntschner et al. 2006; Falcón-Barroso et al. 2011). Part of the 1.35 offset is simply due to the extrapolated light in an $r^{1/4}$ profile, outside the region where our galaxy extend on the SDSS or INT images. But the source of the remaining offset is unclear and confirms the difficulty of determining R_e. For comparison in Paper I we

showed that the 2MASS and RC3 values correlate well, but have an even more significant offset of a factor 1.7!

Various comparisons of the accuracy of kinematic quantities have been performed in the literature (e.g. Emsellem et al. 2004). The general finding is that the measurements of the galaxies velocity dispersion can be reproduced at best with an accuracy of $\approx 5\%$, mainly due to uncertainties in the stellar templates and various systematic effects that are difficult to control. Here in Fig. 8 we test the internal errors of our kinematic determination by comparing σ_e against the velocity dispersion measured withing a circular aperture of radius $R = 0.87$ kpc as adopted in Jorgensen et al. (1995b). We measure an rms scatter of $\Delta = 0.027$ dex between the two quantities, which corresponds to a 1σ error of 4.5% in each value. The two values do not measure the same quantity, as the two adopted apertures and fitted spectra are different, and for this reason both the actual velocity dispersion and the stellar population change in the two pPXF fits. For this reason the observed scatter provides a firm upper limit to the true internal uncertainties in σ_e. However, in what follows we still assume a conservative error of 5% in σ_e and $\sigma_{0.87}$, to account for possible systematics. The same choice was made e.g. in Tremaine et al. (2002) and Cappellari et al. (2006). We further compared our $\sigma_{0.87}$ values agaimts the literature σ compilation in the HyperLEDA database (Paturel et al. 2003), for 207 galaxies in common with our sample. A robust fit between the logarithm of the two quantities eliminating outliers with LTS_LINEFIT gives an observed rms scatter of 9% ($\Delta = 0.037$ dex), likely dominated by the heterogeneity of the HyperLEDA values, and no significant offset (2%) in the overall normaliztion. Apart from placing a very firm upper limit to our errors, this provides an external estimate of the typical uncertainties in the HyperLEDA values.

4.1.2 Errors in mass or M/L

To obtain an estimate of our mass and M/L errors for the full sample, we proceed similarly to Cappellari et al. (2006), namely we compare mass determinations using two significantly different modelling approaches. In Section 3.1.2 we described the six modelling approaches that were presented in Cappellari et al. (2012b) and we also use in this paper. For this test we compare the two very different set of models: the self-consistent model (A) and the models (B) which include a NFW halo with mass as free parameter. For the model with NFW halo we then compute the $(M/L)_e \equiv [M/L](r = R_e)$ by numerically integrating the luminous and dark matter of the models. The total M/L enclosed within an iso-surface of volume $V = 4\pi R_e^3/3$ is defined as follows

$$[M/L](r = R_e) \equiv \frac{L(R_e) \times (M/L)_{\rm stars} + M_{\rm DM}(R_e)}{L(R_e)}, \qquad (22)$$

where $M_{\rm DM}$ is the mass in the dark halo. This quantity is compared with the $(M/L)_{\rm JAM}$ of the self-consistent model in the top panel of Fig. 9. The agreement is excellent, with an rms scatter $\Delta = 0.030$ dex, consistent with errors of 5% in each quantity. This value is nearly the same as the value of 6% we estimated as modelling error in Cappellari et al. (2006) and confirms the original estimate of the random modelling uncertainties. There is no evidence for any significant trend or systematic offset.

Importantly this result clarifies two misconceptions regarding the use of self-consistent models to measure the M/L inside $r \approx R_e$ in galaxies. Self-consistent models, like the one used in Cappellari et al. (2006), do *not* significantly underestimate the total M/L as it is sometimes stated (e.g. Dutton et al. 2011b, section 3.7). Even though the model with dark halo has a total galaxy
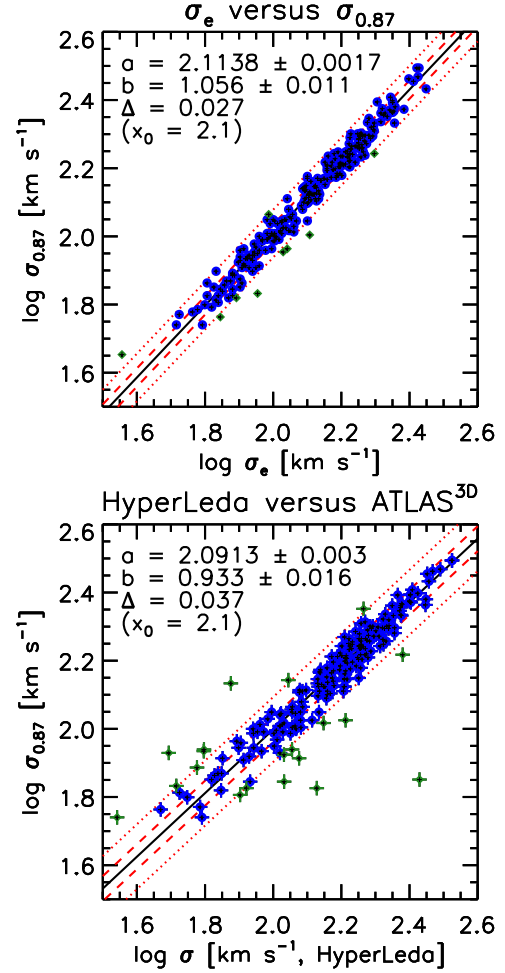
Figure 8. Testing the relative accuracy of σ_e determinations. *Top Panels:* Same as in Fig. 7 for the comparison between the dispersion σ_e, as measured with pPXF from the spectrum inside an elliptical aperture of area $A = \pi R_e^2$, and the quantity $\sigma_{0.87}$ measured within a spectrum inside a fixed circular aperture of radius $R = 0.87$ kpc as in Jorgensen et al. (1995b). *Bottom Panels:* Same as in the top panel, for the comparison between our $\sigma_{0.87}$ and the central σ provided by the HyperLeda database.

mass typically an order of magnitude larger inside the virial radius, and has a dramatically different mass profile at large radii, the model still measures an unbiased *total* M/L within a sphere of radius $r \approx R_e$, corresponding to the projected extent of the kinematical data. The robustness in the recovery of the enclosed total mass, in the region constrained by the data, even in the presence of degeneracies in the halo profile, was already pointed out by Thomas et al. (2005) and is demonstrated here with a much larger sample.

Of course the self-consistent $(M/L)_{\rm JAM}$ is larger than the purely stellar one $(M/L)_{\rm stars}$ if dark matter is present, according to the relation

$$(M/L)_{\rm JAM} \approx [M/L](r = R_e) = \frac{(M/L)_{\rm stars}}{1 - f_{\rm DM}(r = R_e)}, \qquad (23)$$

where the fraction of dark matter contained within an iso-density surface of mean radius R_e is defined as

$$f_{\rm DM}(r = R_e) \equiv \frac{M_{\rm DM}(R_e)}{L(R_e) \times (M/L)_{\rm stars} + M_{\rm DM}(R_e)}. \qquad (24)$$

The difference between $(M/L)_{\rm JAM}$ and the stellar M/L inferred from population models can then be used to give quantitative con-
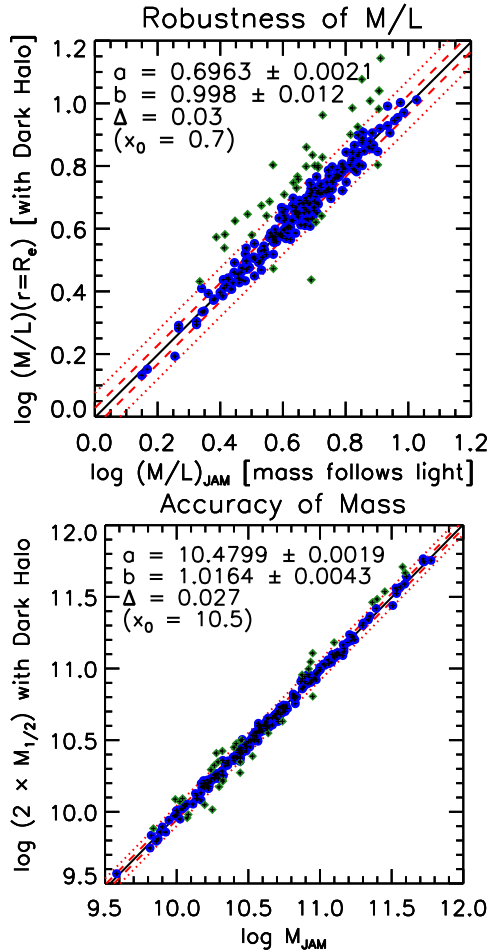
Figure 9. Accuracy of M/L and mass. *Top Panel:* Same as in Fig. 7 for the comparison between the $(M/L)_{\rm JAM}$ of the best-fitting self-consistent (total mass follows light) models, and the $(M/L)_e$, integrated within an iso-surface of volume $V = 4\pi R_e^3/3$ (for a spherical galaxy a sphere of radius $r = R_e$), including the contribution of both the stellar and the dark matter component. Except for some outliers due to inferior data, there is no bias between the two determinations, which are consistent with an intrinsic scatter of 5% in each quantity. *Bottom Panel:* Same as the top panel for the comparison between the total mass of the self-consistent JAM model and twice the mass within the half-light iso-surface, for the model with dark matter halo.

straints on the dark matter content and the form of the IMF, as done in Cappellari et al. (2006). Moreover the self-consistent models do not imply or require the dark mass to be negligible inside $r \approx R_e$ as sometimes stated (e.g. Thomas et al. 2011). Even for galaxies with inferred $f_{\rm DM} \approx 50\%$, due to inferior data (see later), the total (luminous plus dark) M/L within $1R_e$ is still accurately recovered by the simple self-consistent models. This makes the self-consistent models well suited to determine unbiased total M/L within $1R_e$ at high redshift (van der Marel & van Dokkum 2007; van der Wel & van der Marel 2008; Cappellari et al. 2009), where high-quality integral-field stellar kinematics still cannot be obtained and dark matter fractions cannot be extracted.

Using integral-field data the error in this measure of enclosed mass is as small as the one that can be obtained from strong lensing studies. The important difference between the two techniques is that the lensing results measure the total mass inside a projected cylinder (or elliptical cylinder), while the stellar kinematics gives

the total mass inside a spherical (or spheroidal) region. The lensing mass should be larger than the dynamical one if dark matter is present in the galaxy. The difference between these two quantities provides a measure of the dark matter content along the LOS and can be exploited to get some constraints on the dark matter profiles (Thomas et al. 2011; Dutton et al. 2011a).

For completeness we also show the accuracy in the recovery of the enclosed mass in the bottom panel of Fig. 9, as this is what is usually presented. This figure contains essentially the same amount of information as Fig. 9, given that the luminosities are identical on both axes, so the values still differ just by the M/L. This figure explicitly illustrates that with good accuracy

$$(M/L)_{\rm JAM} \approx 2 \times M_{1/2}. \tag{25}$$

The JAM models with dark halo additionally provide an estimate of the dark matter fraction $f_{\rm DM}$ (equation (24)) enclosed within the region constrained by the data $r = R_e$. For the galaxies where our kinematics does not cover $1R_e$, our $f_{\rm DM}$ will be more uncertain. The results is presented, as a function of galaxy stellar mass $M_{\rm stars}$ in Fig. 10 for the set of models (B), with a NFW halo, with mass as free parameter, and for the set of models (E), which have a cosmologically-motivated NFW halo, uniquely specified by $M_{\rm stars}$. We find a median dark matter fraction for the ATLAS$^{\rm 3D}$ sample of $f_{\rm DM} = 16\%$ for the full sample and $f_{\rm DM} = 12\%$ for the best models (B) and 17% with models (E). These value are broadly consistent, but on the lower limit, with numerous previous stellar dynamics determinations inside $1R_e$ from much smaller samples and larger uncertainties: Gerhard et al. (2001) found $f_{\rm DM} = 10 - 40\%$ from spherical dynamical modelling of 21 ETGs; Cappellari et al. (2006) inferred $f_{\rm DM} \approx 30\%$ by comparing dynamics and population masses of 25 ETGs, and assuming a universal IMF; Thomas et al. (2007b, 2011) measured $f_{\rm DM} = 23 \pm 17\%$ via axisymmetric dynamical models of 16 ETGs; Williams et al. (2009) measured a median fraction $f_{\rm DM} = 15\%$ with JAM models of 15 ETGs, as done here, but with more extended stellar kinematics to $\approx 2 - 3 \, R_e$; The results of Tortora et al. (2009) are not directly comparable, as they used spherical galaxy toy models and inhomogeneous literature data from various sources, however they are interesting because they explored a sample of 335 ETGs, comparable to ours, and report a typical $f_{\rm DM} = 30\%$ by comparison with stellar population.

The quite small $f_{\rm DM}$ that we measure is also consistent with the fact that the strong lensing analysis of the about 70 galaxies of the SLACS sample Bolton et al. (2006) finds a logarithmic slopes for the *total* (luminous plus dark matter) density close to isothermal. Subsequent re-analyses of their data all confirmed a trend $\rho_{\rm tot}(r) \propto r^{-2.0}$, with an intrinsic scatter of ≈ 0.2 (Koopmans et al. 2006, 2009; Auger et al. 2010a; Barnabè et al. 2011). In Fig. 2 we derive the same slope and intrinsic scatter for the stellar density alone, inside a sphere of radius $r = R_e$. This fact seems to suggest that dark matter does not play a significant role in galaxy centres and that the measured isothermal density slope is essentially due the stellar density distribution. Only a very steep dark matter slope close to isothermal $\rho_{\rm DM}(r) \propto r^{-2.0}$ like the average stellar distribution could allow for significant dark matter fractions, while still being consistent with these observations. We are not aware of any theoretical or empirical evidence for these very steep dark matter cusps in galaxies.
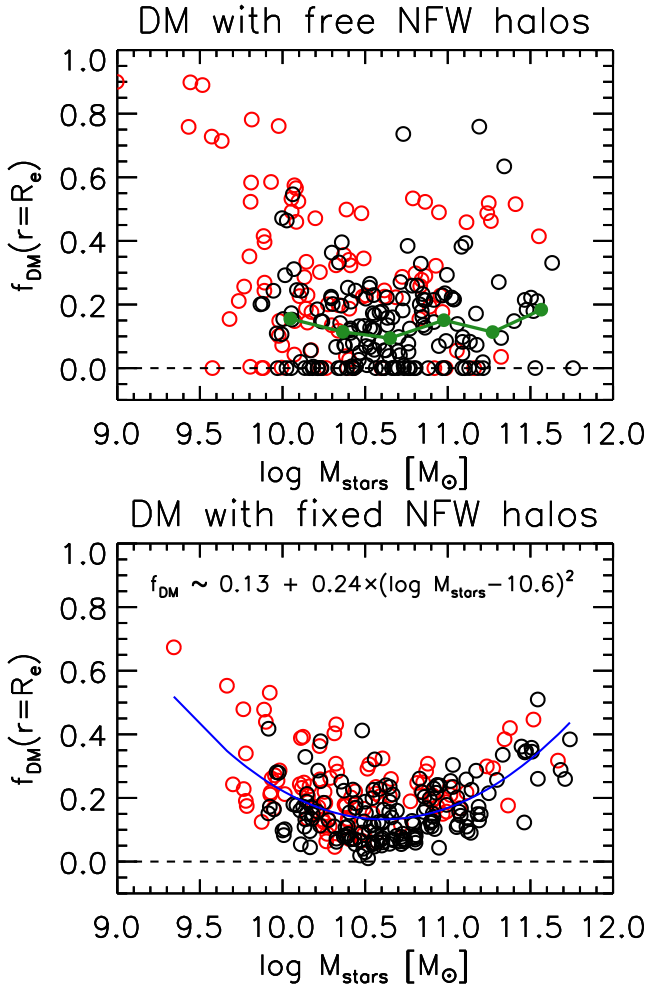
Figure 10. Dark matter fraction for ATLAS[3D] galaxies. The open circles indicate the fraction f_{DM} of dark matter enclosed within the iso-surface of volume $V = 4\pi R_e{}^3/3$ (in the spherical case within a sphere of radius R_e), for the best-fitting JAM models, as a function of the galaxy stellar mass M_{stars} inferred by the models. The black symbols are for the subset of 163 galaxies with the best models and data, while the red symbols indicate less impressive model fits (e.g. due to bars, interactions or low inclination) or inferior data. The *Top Panel* corresponds to the results for model (B), with a NFW halo having mass as free parameter. The median is $f_{DM} = 16\%$ for the full sample and $f_{DM} = 12\%$ for the best models. In a number of cases the model without dark matter is preferred. The solid green line indicates the median for six bins of mass. All significant f_{DM} values seems just spurious results of inferior data or modelling problems. The *Bottom Panel* is the same as the top one, for the set of models (E) which has a cosmologically-motivated NFW halo, uniquely determined by M_{stars} (see text for details). The median $f_{DM} = 17\%$. The blue line is a robust parabolic fit to all the data, which has best-fitting parameter written in the figure. The difference between these two panels likely illustrates our uncertainty in the individual dark matter fractions. The robust result is that dark matter fractions for halo slopes as steep as NFW or more shallow is small, with $f_{DM} < 29\%$ ($f_{DM} < 18\%$) in 90% (68% [1σ]) of the good models.

4.2 The classic Fundamental Plane

Since the discovery of the Fundamental Plane (FP) relation between luminosity, size, and velocity dispersion, in samples of local elliptical galaxies (Faber et al. 1987; Dressler et al. 1987; Djorgovski & Davis 1987), numerous studies have been devoted to the determination of the FP parameters either including fainter

galaxies (Nieto et al. 1990), fast rotating ones (Prugniel & Simien 1994), or lenticular galaxies (Jorgensen et al. 1996). The dependency of the FP parameters have been investigated as a function of the photometric band (Pahre et al. 1998; Scodeggio et al. 1998) or redshift (van Dokkum & Franx 1996). Moreover galaxy samples of more tha 10^4 galaxies have been studied (Bernardi et al. 2003; Graves et al. 2009; Hyde & Bernardi 2009). In this section, before presenting our result, we study the consistency of our FP parameters with previous studies.

Nearly all previous studies have used as variables the logarithm of the effective radius R_e, the effective surface brightness Σ_e and the (central) velocity dispersion σ. One of the reasons for this choice comes from the emphasis of the FP for distance determinations. Both Σ_e and σ are distance independent, so that all the distance dependence can be collected into the R_e coordinate by writing the FP as

$$\log R_e = a + b \log \sigma + c \log \Sigma_e. \tag{26}$$

In the top panel of Fig. 11 we present the edge-on view of our ATLAS[3D] FP, obtained with the LTS_PLANEFIT routine, where we use as velocity dispersion σ_e (Section 3.3.3) as done in Cappellari et al. (2006) and Falcón-Barroso et al. (2011), but here measure within an elliptical isophote. Our best-fitting parameters $b = 1.048 \pm 0.041$ and $c = -0.746 \pm 0.023$ are formally quite accurate, but significantly different from what is generally found by other studies: the median of the 11 determinations listed in table 4 of Bernardi et al. (2003) is $b = 1.33$ and $c = -0.82$, with an rms scatter in the values of $\sigma_b = 0.12$ and $\sigma_c = 0.03$. The observed scatter we measure $\Delta \approx 0.092$ in $\log R_e$ is very close to what has been found by other studies (e.g. Jorgensen et al. 1996 find 0.084).

To understand the possible reason of this disagreement we test the sensitivity of our estimate to the sample selection and the size of the kinematical aperture used for the σ determinations. For this we measure the velocity dispersion $\sigma_{0.87}$ inside a circular aperture with radius $R = 0.87$ kpc, as done by Jorgensen et al. (1995b). We also select the massive half of our sample by imposing a selection $\sigma_{0.87} > 130$ km s^{-1}. The resulting FP is shown in the middle panel of Fig. 11, and now both the fitted values and the observed scatter agree with previous values. For comparison we also show in the bottom panel of Fig. 11 the determination of the FP parameters, when using σ_e instead of $\sigma_{0.87}$, but keeping the same selection of the massive half of our ATLAS[3D] sample $\sigma_e > 130$ km s^{-1}. These values are also consistent with the literature. This illustrates the importance of sample-selection and σ extraction in the derivation of FP parameters. The increase of b as a function of the lower σ cut-off of the selection is fully consistent with the same finding by Gargiulo et al. (2009) and Hyde & Bernardi (2009) and we refer the reader to the latter paper for a more complete study of the possible biases in the FP parameters due to sample selection. The reason for the sensitivity of the FP parameters to the selection, is a result of the fact that the FP is not a plane, but a warped surface, as we demonstrate in Paper XX by studying the variation of the $(M/L)_{JAM}$ on the VP. So that the FP parameters depend on the region of the surface one includes in the fitting. This was also tentatively suggested by D'Onofrio et al. (2008).

Having shown that with our sample and method we can derive results that are consistent and at least as accurate as previous determinations, we now proceed to study the Virial Plane, by replacing the traditionally used stellar luminosity with the total dynamical mass. We call it in this way, because we will show it is entirely and accurately explained by the virial equilibrium condition. A similar study was performed by Bolton et al. (2007, 2008), and
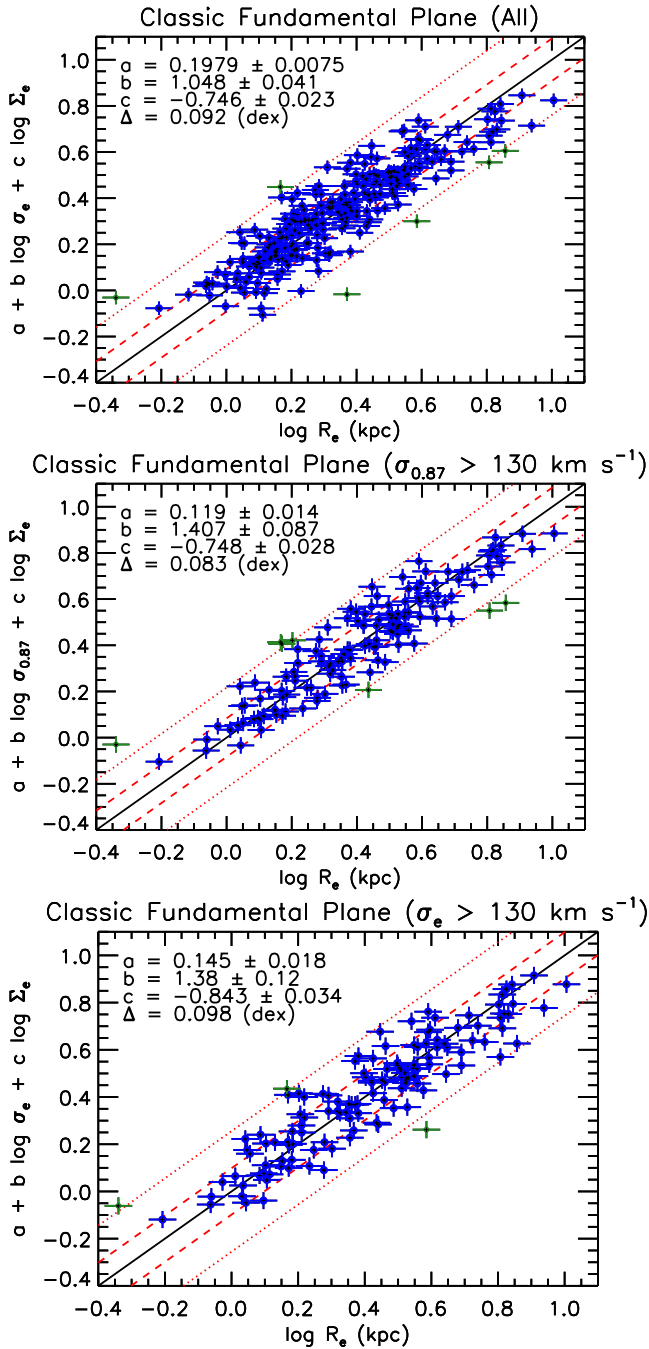
Figure 11. Classic fundamental plane. *Top Panel:* edge-on view of the FP for all the ATLAS^{3D} galaxies. The coefficients of the best-fitting plane $z = a + bx + cx$ and the corresponding observed scatter Δ are shown at the top left of the plot. The two dashed lines mark the 1σ bands (enclosing 68% of the values for a Gaussian distribution) and 2.6σ (99%). The outliers excluded from the fit by the LTS_PLANEFIT procedure are shown with green symbols. The errors are the projection of the observational errors, excluding intrinsic scatter. Middle Panel: Same as in the top panel, with $\sigma_{0.87}$ measured within a circle of 0.87 kpc. Only galaxies with $\sigma_{0.87} > 130$ km s^{-1} are included. Bottom Panel: Same as in the top panel using σ_e, but with Only including galaxies with $\sigma_e > 130$ km s^{-1}.

updated by Auger et al. (2010a), using masses derived from strong lensing analysis. They call their plane the "Mass Plane". Although our studies are closely related, we use a different term to emphasize the fact that, while the lensing masses are measured within a projected cylinder of radius $R = R_e/2$, parallel to the LOS, and for this reason they include a possible contribution of dark matter at large radii, our dynamical masses are measured within a sphere of radius $r = R_e$. This aspect was not made sufficiently clear in Cappellari et al. (2006) and we try to avoid possible confusion here.

4.3   From the Fundamental Plane to the Virial Plane

The classic form for the FP is ideal when the FP is used to determine distances. However, a different form seems more suited to studies where the FP is mainly used as a mass or M/L estimator. For this we rewrite the FP as

$$\log \left(\frac{L}{L_{\odot, r}} \right) = a + b \log \left(\frac{\sigma_e}{130 \, \text{km s}^{-1}} \right) + c \log \left(\frac{R_e}{2 \, \text{kpc}} \right) . \quad (27)$$

Here we normalized the σ_e and R_e values by the approximate median of the values for our sample, to reduce the covariance in the fitted parameters and the error in a. Using L instead of Σ_e has the advantage that it reduced the covariances between the pairs of observables (Σ_e, R_e). Here in fact, as opposed to when $\Sigma_e \equiv L/(2\pi R_e^2)$ is used, there is no explicit dependence between the three axes, which become independently measured quantities. The new fit to the FP is shown in the top panel of Fig. 12. In agreement with all previous authors the fitted parameters are very different from the values $b = 2$ and $c = 1$ expected in the case of the virial equation (20). The relation shows a negligible increase in the observed rms scatter, from $\Delta = 0.092$ dex (24%) to $\Delta = 0.10$ (26%). This may be due to the smaller covariances between our input measurements: the new scatter is now a better representation of the true scatter in the FP relation.

In the bottom panel of Fig. 12 we show for comparison the relation obtained by replacing the total galaxy luminosity with the dynamical mass $M_{1/2}$, which represent the *total* luminous plus dark matter enclosed within a iso-surface enclosing half of the galaxy light. In practice in the plots we show

$$M_{\text{JAM}} \equiv L \times (M/L)_{\text{JAM}} \approx 2 \times M_{1/2} \approx M_{\text{stars}}, \quad (28)$$

where $(M/L)_{\text{JAM}}$ is the total (luminous plus dark) dynamical M/L obtained using self-consistent JAM models (A), L is the total galaxy luminosity and $M_{1/2}$ is the total mass within a sphere of radius $r_{1/2}$ enclosing half of the total galaxy light, where $r_{1/2} \approx 1.33 R_e$ (Hernquist 1990; Ciotti 1991; Wolf et al. 2010; Fig. 4). The correctness of the $M_{\text{JAM}} \approx 2 \times M_{1/2}$ approximation is illustrated in the bottom panel of Fig. 9. While the $2 \times M_{1/2} \approx M_{\text{stars}}$ approximation is due to the relatively small amount of dark matter enclosed within $r = r_{1/2}$ (Fig. 10). This is only approximately true, but much larger errors are generally made when determining stellar masses from stellar population models, due the assumption of a universal IMF, which was recently shown not to represent real galaxies (van Dokkum & Conroy 2010; Cappellari et al. 2012b). None of our conclusions is affected by the last approximation, which only serves to allow for comparisons of our results to previous similar studies that use stellar mass as parameter.

Two features are obvious from the plot: (i) There is a dramatic reduction of the observed scatter from $\Delta = 0.10$ (26%) to $\Delta = 0.062$ (15%). This shows without doubt that a major part of the scatter in the FP is due to variations in the M/L, in agreement with independent results from strong lensing (Auger et al. 2010a);

(ii) The b coefficient substantially increase and is now much closer to the virial value $b = 2$, while the c coefficient remanins nearly unchanged. This confirms that much of the deviation of the FP from the virial predictions is due to a systematic variation in M/L along the FP, not to non-homology, also in agreement with previous dynamical (Cappellari et al. 2006) and strong lensing results (Bolton et al. 2008; Auger et al. 2010a).

The result of this exercise clearly shows that the existence of the FP is entirely due to the fact that galaxies can be remarkably well approximated by virialized stellar systems, with a relatively smooth variation of M/L. These facts have been clearly realized since the discovery of the FP (Faber et al. 1987) and have been generally assumed in most recent studies. The new findings on the tilt of the FP agree with a similar study of scaling relations in ETGs using accurate dynamical models and integral-field kinematics of a sample of just 25 galaxies (Cappellari et al. 2006) and with independent confirmations using strong gravitational lensing (Bolton et al. 2007, 2008; Auger et al. 2010a). Galaxy non-homology has a minor effect at best, when the determination of galaxy scaling parameters is pushed to the maximum accuracy and an attempt is made to remove the most important biases.

The level of accuracy at which the simple virial approximation holds is not entirely expected however, given the apparent complexity of galaxy photometry and kinematics. Of course the dynamical models assume equilibrium and rigorously satisfy the virial equations. One may think that a tight relation is a necessary feature of the approach. This is however not correct. It is true in fact that the models satisfy the scalar virial equation $2T + W = 0$ by construction, where T is the total kinetic energy and W is the total potential energy. However, given the complex multi-component nature of galaxies, the presence of bars, the importance of projection and the fact that the potential energy should include dark matter, it is far from obvious that one should be able to define any simple empirical measure of projected radius on the galaxy, and a measure of velocity dispersion within a limited region, so that the virial equation can be written in the simple form $M_{1/2} = k\,\sigma^2 R/G$ (designed for spherical homologous systems), with fixed exponents and nearly constant coefficient for the entire population!

In Fig. 13 we present a direct comparison between the new JAM M/L estimates within an iso-surface with volume $V = 4\pi R_e{}^3/3$ and the the simple virial estimate of equation (20) from Cappellari et al. (2006). Considering the modelling errors of 5% in M/L estimated in this paper, we infer an error of 15% in the virial estimates. This shows that, although the virial estimates do not suffer from strong biases, they provide errors about a factor 3 larger, even when using our good data.
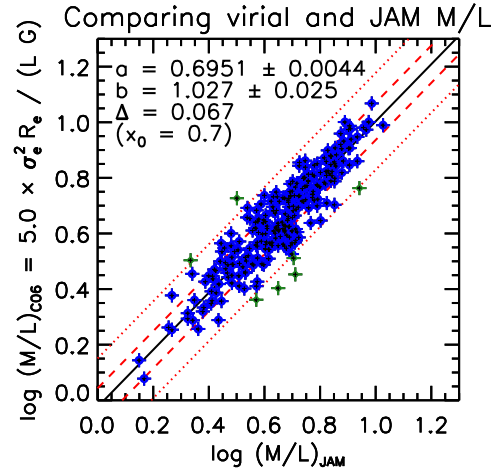
Our finding does not seem to agree with the small systematic offsets recently reported by Thomas et al. (2011). We suspect the disagreement may be an effect of small sample statistics and larger errors, given that they studied only 16 objects and did not use integral-field data. However, the difference they find may also be simply due to a systematic difference in their R_e determination, with respect to the SAURON ones. Our new strong empirical confirmation of the scaling of the coefficient in Cappellari et al. (2006), even in the presence of dark matter, also emphasizes the importance of using virial coefficients that are calibrated to the extent of the available kinematic data. The coefficient $k = 3.75$ given by Spitzer (1969) or $k = 3$ proposed by Wolf et al. (2010) for equation (19) should not be used to estimate central masses in early-type galaxies, where stellar kinematics out to at most a couple of R_e is available and the corresponding value $k \approx 1.9$ of equation (21) applies. The difference of the two coefficients is due to the fact that, while the es-

timator of Wolf et al. (2010) is a theoretical one, designed for spherical geometry, very extended kinematics, and assumes galaxy profiles are known to infinite radii, the one by Cappellari et al. (2006) is an empirical one, designed for quantitative measures of masses in the central regions of ETGs. Both estimators are useful in their own range of applicability, but they should not be used interchangeably, unless one can tolerate systematic biases of $\approx 60\%$ in the absolute mass normalization.

We stress here the importance and the difficulty of obtaining effective radii that are as accurate as the model calibration allows. Ultimately the general unreliability and poor reproducibility of effective radii determined from photometry of different quality is the main limiting factor to a quantitative use of the scalar virial relations to measure accurate masses or M/L, when a proper absolute normalization is essential, like in IMF studies of distant galaxies (Cappellari et al. 2009). If different methods or extrapolations, applied to different, but high-quality photometric data of local galaxies, can produce revisions in R_e by as much a s a factor of two (Kormendy et al. 2009; see also Chen et al. 2010), more significant biases should be expected when comparing local and high-redshift observations, as already pointed out by Mancini et al. (2010). When biases in R_e are present, only dynamical models can still provide robust central masses and M/L, due to the near insensitivity of the models to the shape of the outer mass and light profiles (van der Marel & van Dokkum 2007; van der Wel & van der Marel 2008; Cappellari et al. 2009).

Figure 13. Accuracy of the simple virial estimate. Comparison between the virial estimate of Cappellari et al. (2006) and the more accurate JAM values. The inferred rms errors in the estimation of M/L are 17%. Symbols and lines are as in Fig. 11.

Figure 12. From the Fundamental Plane to the Virial Plane. *Top Panel:* Edge on view of the FP. Symbols and lines are as in Fig. 11. *Bottom Panel:* Edge-on view of the MP. Note the decrease in the scatter, when making the substitution $L \rightarrow M$, and the variation in the coefficients, starting to approach the virial ones $b = 2$ and $c = 1$.

4.4 The $(M/L) - \sigma_e$ relation

In the previous sections we showed that the existence of the Fundamental Plane can be accurately explained by the virial relation combined with a smooth variation of the M/L. Here we study the previously reported correlation $(M/L) \propto \sigma_e^{0.8}$ (in the I-band) between the effective velocity dispersion and the dynamical M/L within a sphere of radius R_e (Cappellari et al. 2006; van der Marel & van Dokkum 2007). This relation was previously found to provide the tightest relation among other parameters of scaling relations (dynamical mass, luminosity or size), with an observed scatter of 18% and an inferred intrinsic one of just ∼13%, when using integral-field kinematics.

The $(M/L) - \sigma_e$ relation for the full ATLAS^{3D} sample is shown in the top-left panel of Fig. 14. Our new relation has an observed scatter of 29%, from which we infer an intrinsic scatter of 22%, when combining our 5% errors in the models with the distance errors for the various subsamples as described in section 2.2 of Paper I. We adopted as distance errors the median one for each given class of determinations reported in Paper I, instead of the individual errors, which are not easy to trust in every case, and that are likely dominated by systematics. The scatter is significantly larger than the previously reported one. The new relation has a formally accurate power slope of $b = 0.688 \pm 0.043$, which is a bit shallower than the previous one, based on a sample ten times smaller than the current one.

To understand the reason for the differences between our $(M/L) - \sigma_e$ slope and previous determinations, in the top-right panel Fig. 14 we plot the $(M/L) - \sigma_e$ relation for the subset of 78 galaxies with SBF distances from Tonry et al. (2001), as done in both Cappellari et al. (2006) and van der Marel & van Dokkum (2007). The relation for this subset now steepens and becomes fully consistent with the previous determinations. The reason for this is likely related to the fact that the Tonry et al. (2001) subsample is biased towards elliptical galaxies, which tend to be the brightest in our sample. A change in slope is then expected from the curvature of the $(M/L) - \sigma_e$ relation, which is not clearly visible in our range of σ values, but is implied by the deviations from our relation when other classes of objects with smaller of larger σ are considered (Zaritsky et al. 2006, 2008; Tollerud et al. 2011). A small but systematic increase in the slope is indeed visible when we select subsamples within different σ ranges from our ATLAS^{3D} sample. We conclude that the discrepancy between our newly fitted value and the previous works is due to the difference in the sample selection. The present sample is not only much large than the one used in previous studies, but also volume-limited so it provides a statistically representative view of the scaling relations in the nearby Universe.

In the middle-left panel of Fig. 14 we show the $(M/L) - \sigma_e$ of the 36 slow rotator ETGs defined in Paper III. We confirm a detectable offset in the relation, with the slow rotators having slightly larger M/L, as previously reported in Cappellari et al. (2006). However, the difference is just at the 9% level. There is also a change in the slope, with the slow rotators defining a more shallow relation that the full population. We also confirm the smaller scatter in the relation, as reported by Falcón-Barroso et al. (2011) for the colour-σ and FP relations. The slow rotators have an observed scatter of 22%, and an inferred intrinsic one of 15% in the $(M/L) - \sigma_e$ relation. This is likely due to the fact that significant amounts of cold gas and star formation, which affect the M/L but not σ, are in fast rotators (Paper IV, McDermid et al. in preparation). The relation for the fast rotators (middle-right panel) agrees

with the global one, as expected from the fact that they dominate the ATLAS^{3D} sample.

The dependence of the slope and zero point of the $(M/L) - \sigma_e$ relation on environment effects is shown in the bottom panels of Fig. 14. As discussed in Paper VII, most of the environmental differences in the ATLAS^{3D} sample can be characterized by whether a galaxy belongs to the Virgo cluster or not. The left panel shows the 58 ATLAS^{3D} galaxies in Virgo. They follow the same shallow relation as the slow rotators, but with the zero point of the global relation. The observed scatter decreases to just 14%, in part due to the accurate distances from ACSVCS (Mei et al. 2007). However, the intrinsic scatter $\Delta(M/L)$ also further decreases to just 10%. This is consistent with the intrinsic scatter measured by Cappellari et al. (2006), using a radically different set of models and different distance estimates (no ACSVCS), but on a sample that, contrary to the ATLAS^{3D} sample, was dominated by Virgo galaxies. The decrease in the scatter must be related to the decrease in the fraction of young objects in Virgo (Kuntschner et al. 2010; McDermid et al. in preparation). It again confirms that the scatter of the $(M/L) - \sigma_e$ relation is dominated by stellar population effects, as previously demonstrated for the FP. The two results are two ways of looking at the same thing, given that the $(M/L) - \sigma_e$ relation is the projection of the differences between the FP and VP along the σ_e axis. For completeness we also show in the bottom-right the relation for non-Virgo galaxies, which dominate the sample and again are consistent, albeit a bit steeper, than the global relation.

In the top panel of Fig. 15 we show how the tightness of the $(M/L) - \sigma_e$ relation can be used to cleanly select galaxies belonging to the Virgo cluster. Here we selected all ATLAS^{3D} galaxies contained within a cylinder of radius of $R = 12°$ centred on the Virgo cluster (approximately at the location of the galaxy M87) and assigned to all of them the cluster distance of $D = 16.5$ Mpc from Mei et al. (2007). We then used the LTS_LINEFIT routine to fit a line. Even in the presence of 20 dramatic outliers out of 79 objects, the method is able to robustly converge to a clean relation.[3] The method selects 59 galaxies within the 99% (2.6σ) confidence bands from the best-fitting relation. The plot reveals a tight sequence in the $(M/L) - \sigma_e$, which corresponds to galaxies in the Virgo cluster, with an observed scatter of $\Delta(M/L) = 0.063$ (16%). It is reassuring to see that this relation, which uses no individual distance information for the galaxies, agrees both in the slope and zeropoint with the ones for all ATLAS^{3D} galaxies, even though it has smaller scatter. Galaxies above the relation lie in the background of Virgo, and their difference in distance modulus from Virgo is $2.5\times$ the difference in $\log(M/L)$ from the best-fitting relations. In this fit we assume that the distance error are due to the 1σ depth of the Virgo cluster. Adopting the value of $\sigma_D = 0.6 \pm 0.1$ from Mei et al. (2007) we derive an intrinsic scatter in M/L of $\epsilon_{M/L} = 0.039$ dex (9%).

When we select only the galaxies with SBF distances from the ACSVCS (Mei et al. 2007) (bottom panel of Fig. 15), we find a relation with the same slope, but a decreased observed scatter of $\Delta(M/L) = 0.047$ (11%). For this relatively small, but still statistically significant sample of 32 galaxies, the inferred intrinsic scatter in M/L would be a mere 8%! Considering that ETGs appear to have very small fractions of dark matter in their central region (Fig. 10), a small scatter in dynamical M/L should be expected

---

[3] Other robust method like (i) minimizing the absolute deviation, (ii) using iterated biweight estimates or (iii) M-estimates (Press et al. 2007, section 15.7), failed to provide a sensible solution to this problem.
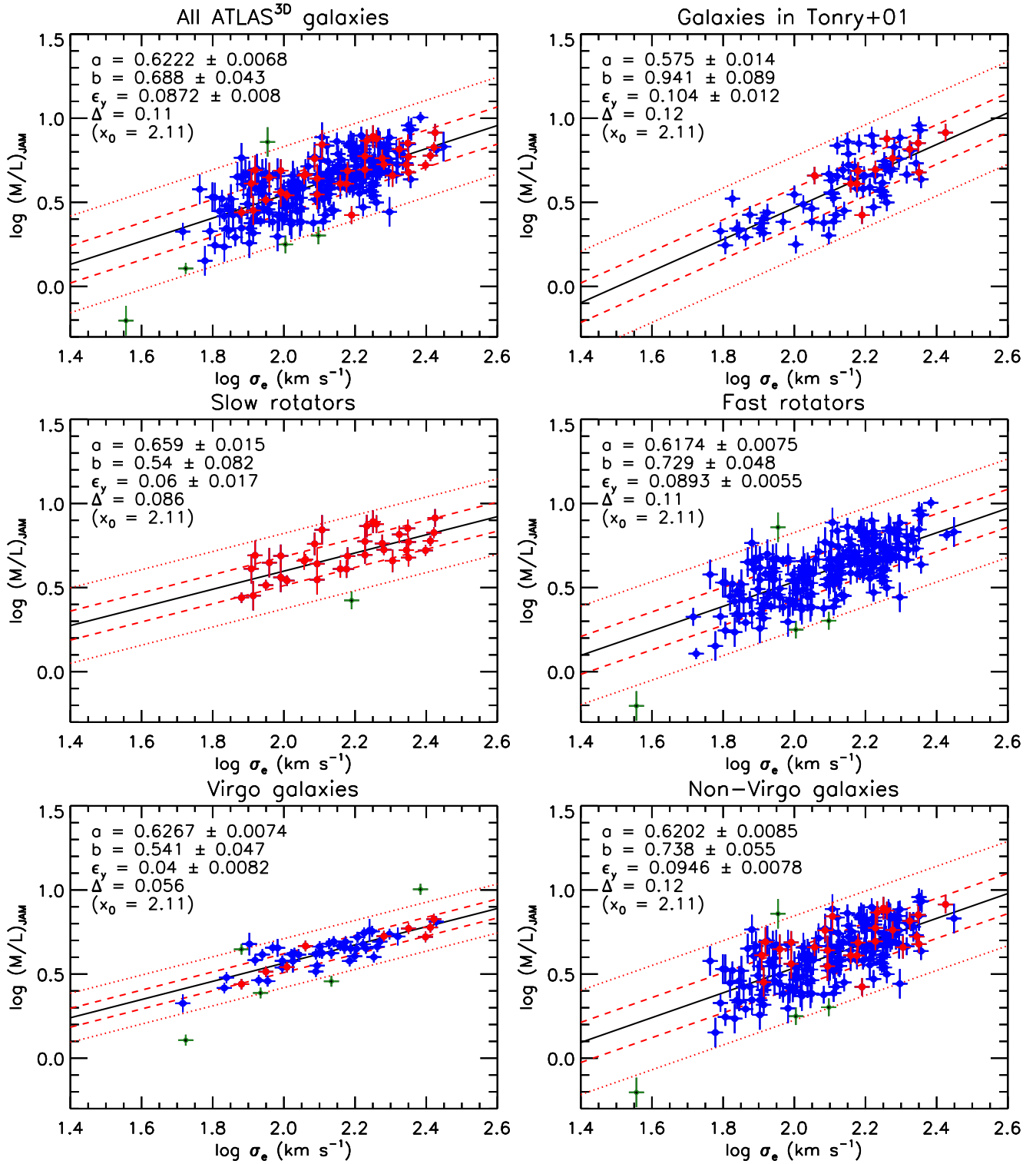
Figure 14. The $(M/L)_e - \sigma_e$ relation. From left to right and from top to bottom the relation is shown (i) for all ATLAS[3D] galaxies; (ii) for the subset in Tonry et al. (2001); (iii) for the subset of slow rotators (from Paper III); (iv) for the subset of fast rotators (from Paper III); (v) for the subset of galaxies in the Virgo cluster; (vi) for subset not in the Virgo cluster. In all plots the blue symbols are fast rotators , while red symbols are slow rotators. Green symbols are outliers excluded from the fit by LTS_LINEFIT.
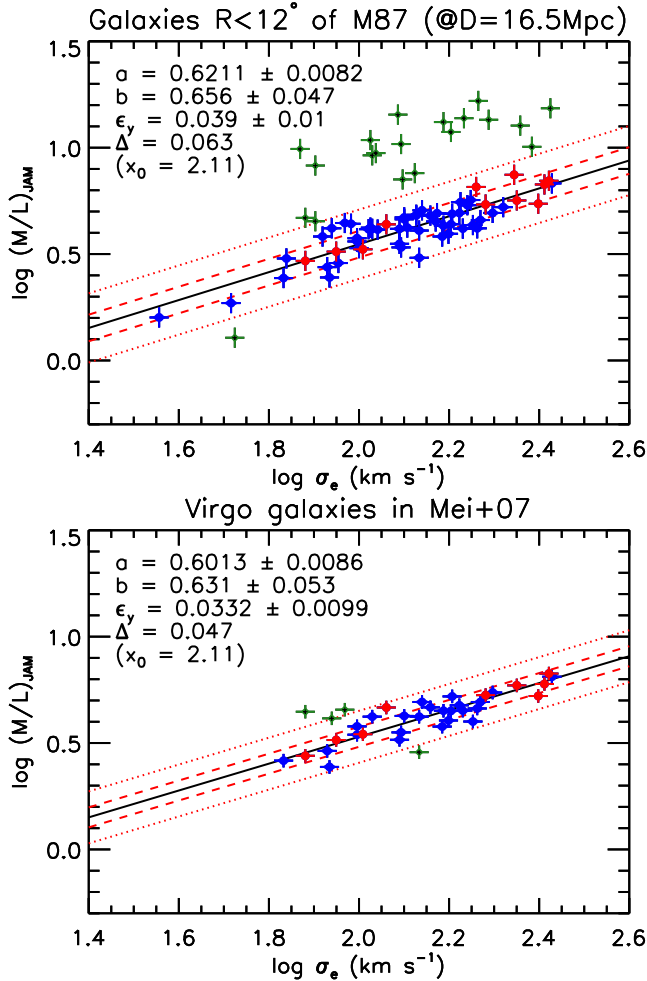
Figure 15. Scatter in the $(M/L)_e - \sigma_e$ relation in the Virgo galaxy cluster. *Top Panel:* All ATLAS³ᴰ galaxies within 12° of the center of the Virgo cluster have been assigned a fixed distance of $D = 16.5$ Mpc. The measured M/L naturally defines a clean $(M/L)_e - \sigma_e$ relation for galaxies belonging to the cluster. The scatter in this relation is due to a combination of the cluster depth and the intrinsic scatter in the relation. *Bottom Panel:* $(M/L)_e - \sigma_e$ relation for the galaxies in Mei et al. (2007). The accurate distances produce a quite significant decrease in the scatter, down to just 11%, indicating that both the $(M/L)_{\rm JAM}$ and the SBF distances are significantly more accurate than this value and confirming that the SBF distances are able to resolve the spatial structure of Virgo, along the LOS, as claimed.

from the extreme tightness of the colour-magnitude relation in clusters (Bower et al. 1992) and specifically for the ACSVCS galaxies (Chen et al. 2010), given that colour is a direct tracer of the M/L of the stellar population (Bell & de Jong 2001). Our small scatter finding confirms the remarkable accuracy of the ACSVCS SBF distances and their ability to resolve the cluster structure as claimed. It shows that the intrinsic $(M/L) - \sigma_e$ relation is extremely tight, but its study is limited in our sample by the distance errors. It would be valuable to perform a similar analysis as in the top panel of Fig. 15, with integral-field data and accurate models, in a cluster like Coma, sufficiently close that good stellar kinematics can be obtained, but sufficiently far that errors in the distance can be virtually ignored. The smaller intrinsic scatter inferred for this sample, with respect to the one in the top panel, suggests that, either they are not drawn from the same population, or the ACSVCS sample in Mei et al. (2007) spans a slightly smaller set of distances within

the Virgo cluster, than the ATLAS³ᴰ Virgo sample. The tightness of this correlation also places tight constraints on the possible intrinsic scatter on the IMF $- \sigma$ trend that we discuss in Paper XX.

4.5 Relation between σ_e and the maximum circular velocity

Previous studies (Zaritsky et al. 2006, 2008; McGaugh et al. 2010; Dutton et al. 2011b) have tried to unify dynamical scaling relations of spiral galaxies and early-type galaxies. For spirals one can measure the rotation velocity of the gas, which appears in the Tully & Fisher (1977) relation between galaxy luminosity (or mass) and its maximum (asymptotic) circular velocity $\max(V_{\rm circ})$, typically measured from the kinematics of the neutral gas at large radii. For early-type galaxies one can measure the velocity dispersion, which enters the Faber & Jackson (1976) and Fundamental Plane relations. Unification of the scaling relations is done by converting velocity dispersion into the circular velocity $V_{\rm circ}(R_e^{\rm max})$ at the half-light radius or into the maximum one $\max(V_{\rm circ})$ adopting constant factors.

Typical conversion factors for $V_{\rm circ}(R_e^{\rm max})$ used in the literature range from $\sqrt{2}$ to $\sqrt{3}$ (Courteau et al. 2007). For example Padmanabhan et al. (2004) estimates $k \approx 1.65$. While Schulz et al. (2010) adopts $k \approx 1.7$ and Dutton et al. (2011b) uses $k \approx 1.54$.

Our dataset provides accurate σ_e for all galaxies, together with circular velocities from our dynamical models. This allows for a robust empirical calibration of the relation. The correlation between σ_e and $V_{\rm circ}(R_e^{\rm max})$ is shown in Fig. 16 and the best-fitting relation has the form

$$V_{\rm circ}(R_e^{\rm max}) \approx 1.51 \times \sigma_e. \tag{29}$$

Considering the variety of photometric profile and galaxy flattening in our complete sample of ETGs, it is remarkable that the relation has a scatter of just 8%, with a weak dependency on σ_e.

Even slightly tighter is the correlation between σ_e and $\max(V_{\rm circ})$, which has the form

$$\max(V_{\rm circ}) \approx 1.76 \times \sigma_e, \tag{30}$$

and an observed scatter of 7%. Importantly this relation show essentially no trend with σ_e. It is worth noting that the $\max(V_{\rm circ})$ defined here is the peak in the rotation curve within the region where we have stellar kinematics, which is generally within $1R_e$. This value should not be confused with the asymptotic value of the circular velocity at large radii, which is generally used in the Tully & Fisher (1977) relation (but see Davis et al. 2011). Although the the so-called bulge-halo conspiracy van Albada & Sancisi (1986) generally tend to make the two values similar (e.g. see Williams et al. 2009), this fact has never robustly established for a significant sample of ETGs.

As shown in Fig. 17, the maximum in the circular velocity is almost always reached well inside $1R_e$, with 85% of the peak $V_{\rm circ}$ happening at a radius within $R_e/2$ and a median radius of just $R_e/5$.

5 SUMMARY

We construct detailed dynamical models (JAM), based on the Jeans equations and allowing for orbital anisotropy, for the volume-limited and essentially mass-selected ATLAS³ᴰ sample of early-type galaxies. The models fit in detail the two-dimensional galaxy images and reproduce in detail the integral-field stellar kinematics
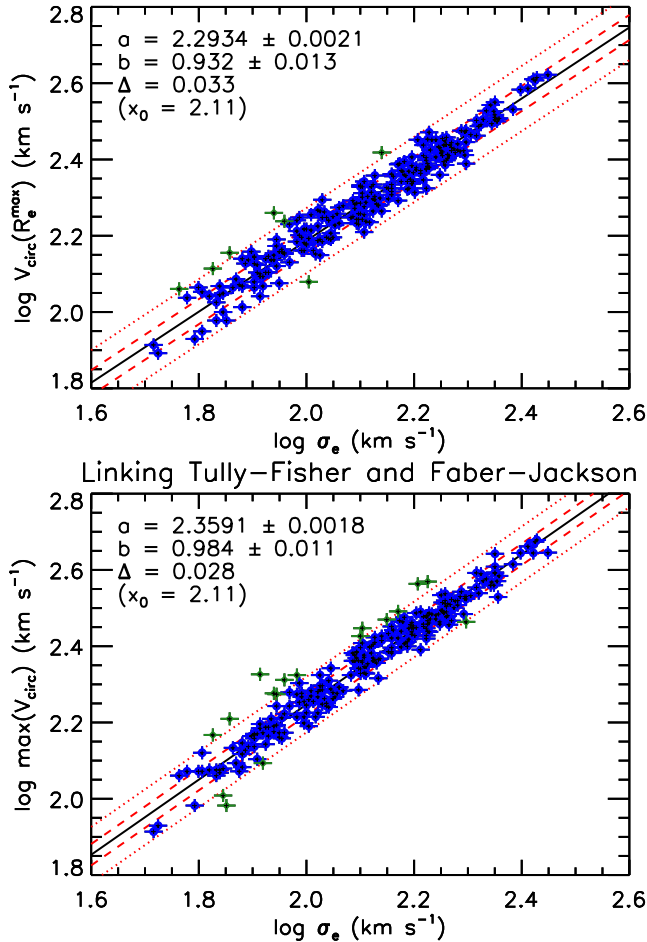
Figure 16. Circular velocity $V_{\rm circ}$ versus σ_e. *Top Panel:* correlation between the circular velocity $V_{\rm circ}(1R_e)$ inferred from our models at $1R_e$, and σ_e. *Bottom Panel:* correlation between the peak circular velocity $\max(V_{\rm circ})$ (within $1R_e$) and σ_e.
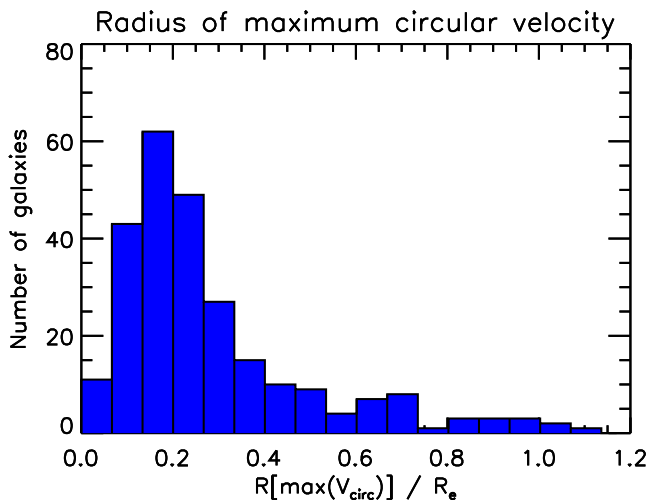
Figure 17. Histogram for the distribution of the radius R/R_e at which the maximum circular $\max(V_{\rm circ})$ is reached, as a fraction of the galaxy effective radius R_e.

obtained with SAURON out to about $1R_e$, the projected half-light radius. We derive accurate total mass-to-light ratios $(M/L)_e$ and dark matter fractions $f_{\rm DM}$, within a sphere of radius $r = R_e$ centred on the galaxies. We infer masses $M_{\rm JAM} \equiv L \times (M/L)_e \approx 2 \times M_{1/2}$, where $M_{1/2}$ is the mass within a sphere enclosing half of the galaxy light. We also measure stellar $(M/L)_{\rm stars}$.

We test the accuracy of our mass determinations by running models with and without dark matter and we find that the enclosed total $(M/L)_e$ is a robust quantity, independent of the inclusion of a dark-matter halo, with an rms accuracy of 5% and negligible bias. In other words, even using simple mass-follow-light models, one recovers the total enclosed $(M/L)_e$ with good accuracy and small bias. We illustrate the tecniques we use to measure radii and global kinematical quantities from our data, and to robustly fit linear relations or planes to the data, even in the presence of outliers and significant intrinsic scatter. We stress the difficulty of measuring absolutely calibrated effective radii R_e, and we argue againt extrapolation in the profiles, for more reproducible results. Systematic offsets in R_e determinations are the main limitation for the use of the scalar virial relation for mass estimates, and may affect size comparisons as a function of redshift.

We find that the thin two-dimensional subset spanned by galaxies in the $(M_{\rm JAM}, \sigma_e, R_e^{\rm max})$ coordinates system, which we call the Virial Plane (VP) has an observed rms scatter of 17%, which would imply an intrinsic one of just 4%. The VP satisfies the scalar virial relation $M_{\rm JAM} = 5.0 \times \sigma_e^2 R_e^{\rm max}/G$ within our tight errors. However, this is only true if one pays special attention to the methodology employed to determine the galaxy global parameters and in particular, (i) one uses as scale radius the major axis $R_e^{\rm max}$ of the 'effective' isophote enclosing half of the total projected galaxy light (without extrapolating the profile beyond the data), and (ii) one measures the velocity dispersion σ_e (which includes rotation and random motions) from a spectrum derived inside that effective isophote. This confirms with unprecedented accuracy previous claims (Cappellari et al. 2006; Bolton et al. 2008) that galaxies accurately satisfy the virial relations and that the existence of the FP is entirely explained by virial equilibrium plus some intrinsic variations in the total $(M/L)_e$.

We revisit the $(M/L)_e - \sigma$ relation and measure a shallower observed slope than previously reported. The difference is due to the selection of the sample of galaxies previously used to fit the relations. We find that the correlation depends both on galaxy rotation and environment, in the sense that both for the subsamples of the galaxies in Virgo, or for the subsample of slow rotators, the relation is more shallow and has a reduced scatter. In the best case, when the most accurate distances are used, the observed scatter drops to 10% and the intrinsic one is estimated to be a mere 6%.

We study the correlation between σ_e and the circular velocity from the dynamical models. We find that $V_{\rm circ}(R_e^{\rm max}) \approx 1.51 \times \sigma_e$ and $\max(V_{\rm circ}) \approx 1.76 \times \sigma_e$. The relations have an observed scatter of 7–8% and show very little dependence on σ_e.

The accurate global dynamical scaling parameters for the ETGs in the ATLAS$^{\rm 3D}$ sample are used in the companion Paper XX to explore different projection of the Virial Plane and the variation of galaxy physical parameters.

REFERENCES

Aihara H., Allende Prieto C., An D., et al., 2011, ApJS, 193, 29
Akritas M. G., Bershady M. A., 1996, ApJ, 470, 706
Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Moustakas L. A., Burles S., 2010a, ApJ, 724, 511
Auger M. W., Treu T., Gavazzi R., Bolton A. S., Koopmans L. V. E., Marshall P. J., 2010b, ApJ, 721, L163
Bacon R. et al., 2001, MNRAS, 326, 23
Barnabè M., Czoske O., Koopmans L. V. E., Treu T., Bolton A. S., 2011, MNRAS, 415, 2215
Barnabè M. et al., 2012, MNRAS, 423, 1073
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Bell E. F., de Jong R. S., 2001, ApJ, 550, 212
Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, MNRAS, 404, 2087
Bernardi M. et al., 2003, AJ, 125, 1866
Bertin G., Ciotti L., Del Principe M., 2002, A&A, 386, 149
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press
Bissantz N., Debattista V. P., Gerhard O., 2004, ApJ, 601, L155
Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, ApJ, 638, 703
Bolton A. S., Burles S., Treu T., Koopmans L. V. E., Moustakas L. A., 2007, ApJ, 665, L105
Bolton A. S., Treu T., Koopmans L. V. E., Gavazzi R., Moustakas L. A., Burles S., Schlegel D. J., Wayth R., 2008, ApJ, 684, 248

Borriello A., Salucci P., Danese L., 2003, MNRAS, 341, 1109
Bower R. G., Lucey J. R., Ellis R. S., 1992, MNRAS, 254, 601
Burstein D., Davies R. L., Dressler A., Faber S. M., Stone R. P. S., Lynden-Bell D., Terlevich R. J., Wegner G., 1987, ApJS, 64, 601
Caon N., Capaccioli M., D'Onofrio M., 1993, MNRAS, 265, 1013
Cappellari M., 2002, MNRAS, 333, 400
Cappellari M., 2008, MNRAS, 390, 71
Cappellari M. et al., 2006, MNRAS, 366, 1126
Cappellari M. et al., 2012a, MNRAS, submitted (Paper XX) arXiv
Cappellari M. et al., 2009, ApJ, 704, L34
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cappellari M. et al., 2007, MNRAS, 379, 418
Cappellari M. et al., 2011a, MNRAS, 413, 813 (Paper I)
Cappellari M. et al., 2011b, MNRAS, 416, 1680 (Paper VII)
Cappellari M. et al., 2012b, Nature, 484, 485
Cappellari M., Verolme E. K., van der Marel R. P., Kleijn G. A. V., Illingworth G. D., Franx M., Carollo C. M., de Zeeuw P. T., 2002, ApJ, 578, 787
Cenarro A. J., Trujillo I., 2009, ApJ, 696, L43
Chen C., Côté P., West A. A., Peng E. W., Ferrarese L., 2010, ApJS, 191, 1
Ciotti L., 1991, A&A, 249, 99
Ciotti L., Lanzoni B., Renzini A., 1996, MNRAS, 282, 1
Courteau S., McDonald M., Widrow L. M., Holtzman J., 2007, ApJ, 655, L21
Davies R. L., Efstathiou G., Fall S. M., Illingworth G., Schechter P. L., 1983, ApJ, 266, 41
Davis T. A., Bureau M., Young L. M., et al., 2011, MNRAS, 414, 968 (Paper V)
de Lorenzi F., Debattista V. P., Gerhard O., Sambhus N., 2007, MNRAS, 376, 71
de Lorenzi F. et al., 2009, MNRAS, 395, 76
de Vaucouleurs G., 1948, Annales d'Astrophysique, 11, 247
de Vaucouleurs G., 1959, Handbuch der Physik, 53, 311
de Vaucouleurs G., de Vaucouleurs A., Corwin, Jr. H. G., Buta R. J., Paturel G., Fouque P., 1991, Third Reference Catalogue of Bright Galaxies. Volume 1-3, XII, 2069 pp. 7 figs.. Springer-Verlag Berlin Heidelberg New York
de Zeeuw P. T. et al., 2002, MNRAS, 329, 513
Deason A. J., Belokurov V., Evans N. W., McCarthy I. G., 2012, ApJ, 748, 2
Dehnen W., 2009, MNRAS, 395, 1079
Dejonghe H., Merritt D., 1992, ApJ, 391, 531
Djorgovski S., Davis M., 1987, ApJ, 313, 59
Doi M., Fukugita M., Okamura S., 1993, MNRAS, 264, 832
D'Onofrio M. et al., 2008, ApJ, 685, 875
Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, ApJ, 313, 42
Duc P.-A. et al., 2011, MNRAS, 417, 863 (Paper IX)
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161
Dunkley J., Bucher M., Ferreira P. G., Moodley K., Skordis C., 2005, MNRAS, 356, 925
Dutton A. A. et al., 2011a, MNRAS, 417, 1621
Dutton A. A. et al., 2011b, MNRAS, 416, 322
Emsellem E. et al., 2011, MNRAS, 414, 888 (Paper III)
Emsellem E. et al., 2004, MNRAS, 352, 721
Emsellem E., Dejonghe H., Bacon R., 1999, MNRAS, 303, 495
Emsellem E., Monnet G., Bacon R., 1994, A&A, 285, 723
Faber S. M., Dressler A., Davies R. L., Burstein D., Lynden-Bell D., 1987, in Nearly Normal Galaxies. From the Planck Time to the Present, Faber S. M., ed., pp. 175–183

Faber S. M., Jackson R. E., 1976, ApJ, 204, 668

Faber S. M., Wegner G., Burstein D., Davies R. L., Dressler A., Lynden-Bell D., Terlevich R. J., 1989, ApJS, 69, 763

Falcón-Barroso J. et al., 2011, MNRAS, 417, 1787

Feigelson E. D., Babu G. J., 1992, ApJ, 397, 55

Ferrarese L., Merritt D., 2000, ApJ, 539, L9

Forbes D. A., Ponman T. J., Brown R. J. N., 1998, ApJ, 508, L43

Gargiulo A. et al., 2009, MNRAS, 397, 75

Gebhardt K. et al., 2000a, AJ, 119, 1157

Gebhardt K. et al., 2000b, ApJ, 539, L13

Gebhardt K. et al., 2003, ApJ, 597, 239

Gelman A., Carlin J., Stern H., Rubin D., 2004, Bayesian data analysis. CRC press

Gerhard O., 1996, in Spiral Galaxies in the Near-IR, Rix D. M. . H.-W., ed., p. 138

Gerhard O., Jeske G., Saglia R. P., Bender R., 1998, MNRAS, 295, 197

Gerhard O., Kronawitter A., Saglia R. P., Bender R., 2001, AJ, 121, 1936

Gerhard O. E., 1993, MNRAS, 265, 213

Gerhard O. E., Binney J. J., 1996, MNRAS, 279, 993

Gilks W., Richardson S., Spiegelhalter D., 1996, Markov chain Monte Carlo in practice. Chapman & Hall/CRC

Gnedin O. Y., Ceverino D., Gnedin N. Y., Klypin A. A., Kravtsov A. V., Levine R., Nagai D., Yepes G., 2011, arXiv:1108.5736

Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16

Governato F. et al., 2010, Nature, 463, 203

Graham A., Colless M., 1997, MNRAS, 287, 221

Graham A. W., Trujillo I., Caon N., 2001, AJ, 122, 1707

Graves G. J., Faber S. M., Schiavon R. P., 2009, ApJ, 698, 1590

Grillo C., Gobat R., Lombardi M., Rosati P., 2009, A&A, 501, 461

Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111

Haario H., Saksman E., Tamminen J., 2001, Bernoulli, 223

Häfner R., Evans N. W., Dehnen W., Binney J., 2000, MNRAS, 314, 433

Hernquist L., 1990, ApJ, 356, 359

Hogg D. W., Bovy J., Lang D., 2010, arXiv:1008.4686

Hopkins P. F., Bundy K., Hernquist L., Wuyts S., Cox T. J., 2010, MNRAS, 401, 1099

Houghton R. C. W., Magorrian J., Sarzi M., Thatte N., Davies R. L., Krajnović D., 2006, MNRAS, 367, 2

Hubble E. P., 1936, Realm of the Nebulae. Yale Univ. Press, New Haven

Hunter C., Qian E., 1993, MNRAS, 262, 401

Hyde J. B., Bernardi M., 2009, MNRAS, 396, 1171

Inoue S., Saitoh T. R., 2011, MNRAS, 418, 2527

Isobe T., Feigelson E. D., Nelson P. I., 1986, ApJ, 306, 490

Jeans J. H., 1922, MNRAS, 82, 122

Jones D. R., Perttunen C. D., Stuckman B. E., 1993, Journal of Optimization Theory and Applications, 79, 157, 10.1007/BF00941892

Jorgensen I., Franx M., Kjaergaard P., 1995a, MNRAS, 273, 1097

Jorgensen I., Franx M., Kjaergaard P., 1995b, MNRAS, 276, 1341

Jorgensen I., Franx M., Kjaergaard P., 1996, MNRAS, 280, 167

Kelly B. C., 2007, ApJ, 665, 1489

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

Koopmans L. V. E. et al., 2009, ApJ, 703, L51

Koopmans L. V. E., Treu T., Bolton A. S., Burles S., Moustakas L. A., 2006, ApJ, 649, 599

Kormendy J., 1977, ApJ, 218, 333

Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJS, 182, 216

Krajnović D., Cappellari M., Emsellem E., McDermid R. M., de Zeeuw P. T., 2005, MNRAS, 357, 1113

Krajnović D. et al., 2011, MNRAS, 414, 2923 (Paper II)

Krajnović D. et al., 2012, MNRAS, submitted (Paper XVII)

Kuntschner H. et al., 2006, MNRAS, 369, 497

Kuntschner H., Emsellem E., et al., 2010, MNRAS, 408, 97

Lablanche P.-Y. et al., 2012, MNRAS, 424, 1495 (Paper XII)

Landsman W. B., 1993, in Astronomical Society of the Pacific Conference Series, Vol. 52, Astronomical Data Analysis Software and Systems II, R. J. Hanisch R. J. V. Brissenden . J. B., ed., pp. 246–+

Laporte C. F. P., White S. D. M., Naab T., Ruszkowski M., Springel V., 2012, MNRAS, 424, 747

Łokas E. L., Mamon G. A., 2001, MNRAS, 321, 155

Long R. J., Mao S., 2010, MNRAS, 405, 301

Macciò A. V., Stinson G., Brook C. B., Wadsley J., Couchman H. M. P., Shen S., Gibson B. K., Quinn T., 2012, ApJ, 744, L9

Magorrian J. et al., 1998, AJ, 115, 2285

Mamon G. A., 2000, in Astronomical Society of the Pacific Conference Series, Vol. 197, Dynamics of Galaxies: from the Early Universe to the Present, F. Combes G. A. Mamon . V. C., ed., pp. 377–+

Mamon G. A., Łokas E. L., 2005, MNRAS, 363, 705

Mancini C. et al., 2010, MNRAS, 401, 933

Markwardt C. B., 2009, in Astronomical Society of the Pacific Conference Series, Vol. 411, Astronomical Society of the Pacific Conference Series, D. A. Bohlender D. Durand . P. D., ed., pp. 251–+

Martizzi D., Teyssier R., Moore B., Wentz T., 2012, MNRAS, 422, 3081

McGaugh S. S., Schombert J. M., de Blok W. J. G., Zagursky M. J., 2010, ApJ, 708, L14

Mei S. et al., 2007, ApJ, 655, 144

Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., Teller E., 1953, The journal of chemical physics, 21, 1087

Monnet G., Bacon R., Emsellem E., 1992, A&A, 253, 366

Moré J., Garbow B., Hillstrom K., 1980, User guide for MINPACK-1. Argonne National Laboratory Argonne, IL

Morganti L., Gerhard O., 2012, MNRAS, 422, 1571

Moster B. P., Naab T., White S. D. M., 2012, arXiv:1205.5807

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

Napolitano N. R. et al., 2005, MNRAS, 357, 691

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

Nieto J.-L., Davoust E., Bender R., Prugniel P., 1990, A&A, 230, L17

Novak G. S., Faber S. M., Dekel A., 2006, ApJ, 637, 96

Onodera M. et al., 2010, ApJ, 715, L6

Padmanabhan N. et al., 2004, New Ast., 9, 329

Pahre M. A., Djorgovski S. G., de Carvalho R. R., 1998, AJ, 116, 1591

Paturel G., Petit C., Prugniel P., Theureau G., Rousseau J., Brouty M., Dubois P., Cambrésy L., 2003, A&A, 412, 45

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

Pfenniger D., 1984, A&A, 141, 171

Pohlen M., Trujillo I., 2006, A&A, 454, 759

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

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2007, Numerical recipes: The art of scientific computing, 3rd edn. Cambridge Univ. Press, Cambridge

Prugniel P., Simien F., 1994, A&A, 282, L1

Prugniel P., Simien F., 1996, A&A, 309, 749

Prugniel P., Simien F., 1997, A&A, 321, 111

Renzini A., Ciotti L., 1993, ApJ, 416, L49+

Richstone D. O., Tremaine S., 1988, ApJ, 327, 82

Rix H.-W., de Zeeuw P. T., Cretton N., van der Marel R. P., Carollo C. M., 1997, ApJ, 488, 702

Rousseeuw P., Van Driessen K., 2006, Data Mining and Knowledge Discovery, 12, 29, 10.1007/s10618-005-0024-4

Rousseeuw P. J., Leroy A. M., 1987, Robust regression and outlier detection. John Wiley & Sons, Inc., New York, NY, USA

Rusin D., Kochanek C. S., Keeton C. R., 2003, ApJ, 595, 29

Rybicki G. B., 1987, in IAU Symposium, Vol. 127, Structure and Dynamics of Elliptical Galaxies, de Zeeuw P. T., ed., p. 397

Saglia R. P., Bertschinger E., Baggley G., Burstein D., Colless M., Davies R. L., McMahan, Jr. R. K., Wegner G., 1997, ApJS, 109, 79

Saglia R. P. et al., 2010, A&A, 524, A6+

Sandage A., 1961, The Hubble Atlas. Carnegie Institution, Washington

Schulz A. E., Mandelbaum R., Padmanabhan N., 2010, MNRAS, 408, 1463

Schwarzschild M., 1979, ApJ, 232, 236

Scodeggio M., Gavazzi G., Belsole E., Pierini D., Boselli A., 1998, MNRAS, 301, 1001

Scott N. et al., 2009, MNRAS, 398, 1835

Scott N. et al., 2012, MNRAS, submitted (Paper XXI)

Sersic J. L., 1968, Atlas de galaxias australes. Cordoba, Argentina: Observatorio Astronomico, 1968

Skrutskie M. F. et al., 2006, AJ, 131, 1163

Spitzer, Jr. L., 1969, ApJ, 158, L139+

Syer D., Tremaine S., 1996, MNRAS, 282, 223

Thomas J., Jesseit R., Naab T., Saglia R. P., Burkert A., Bender R., 2007a, MNRAS, 381, 1672

Thomas J. et al., 2009, MNRAS, 393, 641

Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Corsini E. M., Wegner G., 2005, MNRAS, 360, 1355

Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Corsini E. M., Wegner G., 2007b, MNRAS, 382, 657

Thomas J. et al., 2011, MNRAS, 415, 545

Tollerud E. J., Bullock J. S., Graves G. J., Wolf J., 2011, ApJ, 726, 108

Tonry J. L., Dressler A., Blakeslee J. P., Ajhar E. A., Fletcher A. B., Luppino G. A., Metzger M. R., Moore C. B., 2001, ApJ, 546, 681

Tortora C., Napolitano N. R., Romanowsky A. J., Capaccioli M., Covone G., 2009, MNRAS, 396, 1132

Tremaine S. et al., 2002, ApJ, 574, 740

Treu T., Auger M. W., Koopmans L. V. E., Gavazzi R., Marshall P. J., Bolton A. S., 2010, ApJ, 709, 1195

Trujillo I., Burkert A., Bell E. F., 2004, ApJ, 600, L39

Trujillo I., Graham A. W., Caon N., 2001, MNRAS, 326, 869

Tully R. B., Fisher J. R., 1977, A&A, 54, 661

Valluri M., Merritt D., Emsellem E., 2004, ApJ, 602, 66

van Albada T. S., Sancisi R., 1986, Royal Society of London Philosophical Transactions Series A, 320, 447

van de Sande J. et al., 2011, ApJ, 736, L9+

van den Bosch R. C. E., 2008, PhD thesis, Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands

van den Bosch R. C. E., van de Ven G., 2009, MNRAS, 398, 1117

van den Bosch R. C. E., van de Ven G., Verolme E. K., Cappellari M., de Zeeuw P. T., 2008, MNRAS, 385, 647

van der Kruit P. C., Searle L., 1981, A&A, 95, 105

van der Marel R. P., 1991, MNRAS, 253, 710

van der Marel R. P., Cretton N., de Zeeuw P. T., Rix H.-W., 1998, ApJ, 493, 613

van der Marel R. P., de Zeeuw P. T., Rix H.-W., Quinlan G. D., 1997, Nature, 385, 610

van der Marel R. P., Franx M., 1993, ApJ, 407, 525

van der Marel R. P., van Dokkum P. G., 2007, ApJ, 668, 756

van der Wel A., van der Marel R. P., 2008, ApJ, 684, 260

van Dokkum P. G., Conroy C., 2010, Nature, 468, 940

van Dokkum P. G., Franx M., 1996, MNRAS, 281, 985

van Dokkum P. G., Kriek M., Franx M., 2009, Nature, 460, 717

Walker M. G., Mateo M., Olszewski E. W., Gnedin O. Y., Wang X., Sen B., Woodroofe M., 2007, ApJ, 667, L53

Weiner B. J. et al., 2006, ApJ, 653, 1049

Williams M. J., Bureau M., Cappellari M., 2009, MNRAS, 400, 1665

Wolf J., Martinez G. D., Bullock J. S., Kaplinghat M., Geha M., Muñoz R. R., Simon J. D., Avedo F. F., 2010, MNRAS, 406, 1220

York D. G., Adelman J., Anderson, Jr. J. E., et al., 2000, AJ, 120, 1579

Zaritsky D., Gonzalez A. H., Zabludoff A. I., 2006, ApJ, 638, 725

Zaritsky D., Zabludoff A. I., Gonzalez A. H., 2008, ApJ, 682, 68

Zhao H., 1996, MNRAS, 283, 149