M*/L gradients driven by IMF variation: Large impact on dynamical stellar mass estimates

M. Bernardi[1]⋆, R. K. Sheth[1], H. Dominguez-Sanchez[1], J.-L. Fischer[1], K.-H. Chae[1,2], M. Huertas-Company[1,3], F. Shankar[4]

[1] *Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[2] *Department of Astronomy and Space Science, Sejong University, 98 Gunja-dong Gwangjin-Gu, Seoul 143-747, Republic of Korea*
[3] *GEPI, Observatoire de Paris, CNRS, Univ. Paris Diderot; Place Jules Janssen, 92190 Meudon, France*
[4] *Department of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*

ABSTRACT

Within a galaxy the stellar mass-to-light ratio Υ_* is not constant. Recent studies of spatially resolved kinematics of nearby early-type galaxies suggest that allowing for a variable initial mass function (IMF) returns significantly larger Υ_* gradients than if the IMF is held fixed. We show that ignoring such IMF-driven Υ_* gradients can have dramatic effect on dynamical ($M_*^{\rm dyn}$), though stellar population ($M_*^{\rm SP}$) based estimates of early-type galaxy stellar masses are also affected. This is because $M_*^{\rm dyn}$ is usually calibrated using the velocity dispersion measured in the central regions (e.g. $R_e/8$) where stars are expected to dominate the mass (i.e. the dark matter fraction is small). On the other hand, $M_*^{\rm SP}$ is often computed from larger apertures (e.g. using a mean Υ_* estimated from colors). If Υ_* is greater in the central regions, then ignoring the gradient can overestimate $M_*^{\rm dyn}$ by as much as a factor of two for the most massive galaxies. Large Υ_*-gradients have four main consequences: First, $M_*^{\rm dyn}$ cannot be estimated independently of stellar population synthesis models. Second, if there is a lower limit to Υ_* and gradients are unknown, then requiring $M_*^{\rm dyn} = M_*^{\rm SP}$ constrains them. Third, if gradients are stronger in more massive galaxies, then accounting for this reduces the slope of the correlation between $M_*^{\rm dyn}/M_*^{\rm SP}$ of a galaxy with its velocity dispersion. In particular, IMF-driven gradients bring $M_*^{\rm dyn}$ and $M_*^{\rm SP}$ into agreement, not by shifting $M_*^{\rm SP}$ upwards by invoking constant bottom-heavy IMFs, as advocated by a number of recent studies, but by revising $M_*^{\rm dyn}$ estimates in the literature downwards. Fourth, accounting for Υ_* gradients changes the high-mass slope of the stellar mass function $\phi(M_*^{\rm dyn})$, and reduces the associated stellar mass density. These conclusions potentially impact estimates of the need for feedback and adiabatic contraction, so our results highlight the importance of measuring Υ_* gradients in larger samples.

Key words: galaxies: luminosity function, mass function – galaxies: structure – galaxies: fundamental parameters – galaxies: kinematics and dynamics

1 INTRODUCTION

A census of the stellar masses of galaxies is more useful than a census of their luminosity (e.g. Cole et al. 2001). There are two approaches to estimating the stellar mass of a galaxy. One fits the observed data – single or multi-band photometry and perhaps spectroscopy as well – to stellar population synthesis libraries (SP; e.g. Bruzual & Charlot 2003; Maraston 2005; Vazdekis et al. 2010; Conroy & Gunn 2010; Maraston & Strömbäck 2011). These include assumptions about the age, metallicity, star formation history, the amount of dust, of gas, and the initial mass function (IMF) (e.g. Worthey 1994; Thomas et al. 2011; Mendel et al. 2014; Villaume et al. 2017). The data constrain (some, usually degenerate, combination of) these parameters, and hence the stellar mass-to-light ratio, $M_*^{\rm SP}/L$. The stellar mass $M_*^{\rm SP}$

⋆ E-mail: bernardm@sas.upenn.edu

is then obtained by multiplying $M_*^{\rm SP}/L$ by an estimate of the total light. Note that here it is the total L which matters; the detailed shape of the light profile does not. Both $M_*^{\rm SP}/L$ and L carry significant uncertainties, although recent work suggests that the systematics associated with L are now subdominant (Bernardi et al. 2017a). The biggest systematic unknown for $M_*^{\rm SP}/L$ is the IMF.

The other is a dynamical estimate, $M_*^{\rm dyn}$, and comes in two flavors, both of which use the Jeans equation (see Cappellari 2016 for a review). These estimates almost always assume that $M_*^{\rm dyn}/L$ is constant throughout the galaxy, so that the observed shape of the light profile is used as an indicator of the mass profile. If the velocity dispersion is only measured on a single scale (e.g. the SDSS), then $M_*^{\rm dyn} \propto R_e \sigma^2/G$, where the size R_e and the constant of proportionality depend on the shape of the light profile – but the total L does not matter. In this case, the Jeans equation predicts the shape of the velocity dispersion profile, and the overall amplitude is got by matching the predicted shape to the observed σ (this is why L does not matter). If spatially resolved spectroscopy is available, so the profile $\sigma(r)$, rather than its value at only a single scale, is known, then the richer dataset allows one to constrain a richer family of models. Now the models must match σ over a range of scales. Typically, $\sigma(r)$ predicted by the light (Binney & Mamon 1982; Prugniel & Simien 1997) falls more steeply than observed (Jørgensen et al. 1995), so the difference is attributed to dark matter which increasingly dominates the mass in the outer regions. Importantly, in both cases, $M_*^{\rm dyn}$ does not depend on the details of the stellar population, so it is sometimes viewed as being less impacted by systematics than $M_*^{\rm SP}$.

It has been known for some time that, if one assumes that all galaxies have the same IMF, then $M_*^{\rm dyn}/M_*^{\rm SP}$ varies across the early-type population (Bender, Burstein & Faber 1992; Bernardi et al. 2003; Shankar & Bernardi 2009). Recent work has focussed on the fact that this ratio tends to increase with velocity dispersion (Cappellari et al. 2013 from ATLAS[3D]; Li et al. 2017 from MaNGA). This correlation is substantially reduced if the IMF is treated as a free parameter: galaxies with larger σ tend to have IMFs which result in larger $M_*^{\rm SP}/L$ (Conroy & van Dokkum 2012; Lyubenova et al. 2016; Tang & Worthey 2017). In this case, the difference between $M_*^{\rm dyn}$ and $M_*^{\rm SP}$ is explained by asserting that $M_*^{\rm SP}/L$ estimates are biased unless one accounts for IMF-related effects (Cappellari et al. 2013; Li et al. 2017; but see Clauwens et al. 2016).

While this is possible, it is potentially problematic for the following reason. Recall that $M_*^{\rm SP} = (M_*^{\rm SP}/L) \times L$. The improvements in L are most dramatic for the most massive galaxies (Bernardi et al. 2010; Bernardi et al. 2013; Meert et al. 2015), and have lead to a revision of the stellar mass density upwards by a factor of 2 (Bernardi et al. 2013, 2017a,b; Thanjavur et al. 2015; D'Souza et al. 2015). This weakens the role that must have been played by feedback in regulating star formation. If the $M_*^{\rm SP}$ estimates must be increased further because of IMF effects on $M_*^{\rm SP}/L$ (Bernardi et al. 2018), then the need for feedback will be further reduced. It is not obvious that this is reasonable.

We turn, therefore, to the possibility that some of the discrepancy between $M_*^{\rm dyn}$ and $M_*^{\rm SP}$ is driven by problems with $M_*^{\rm dyn}$. For this study, we will suppose that the prob-

lems are not due to shape of the light profile, but to the assumption that the stellar mass-to-light ratio Υ_* is constant within a galaxy. The constant Υ_* assumption is commonly made (e.g. the analysis of ATLAS[3D] in Cappellari et al. 2013) because it is convenient – it has no physical motivation. Galaxies have long been known to show gradients in color (photometry) and absorption line strength (spectroscopy). In the context of stellar population modelling, these indicate age and/or metallicity gradients. Typically, these gradients tend to increase the stellar population estimate of the stellar mass-to-light ratio in the central regions. Age and metallicity estimates which result from assuming a constant IMF imply a variation in Υ_* of about 50% within a galaxy (e.g. bottom row, second from left panel in Figure 10 of van Dokkum et al. 2017; also see Newman et al. 2015, who report $\Upsilon_* \propto R^{-0.15}$ in massive galaxies). When luminosity-weighted and averaged over the galaxy this is a smaller effect, which is why ignoring these gradients when estimating $M_*^{\rm dyn}$ may not be so problematic. Indeed, in their study of $M_*^{\rm dyn}$ in MaNGA, Li et al. (2017) report that the impact of such fixed-IMF Υ_*-gradients on $M_{\rm dyn}$ estimates is small (see their Figure 6).

However, more recent work suggests that IMF-sensitive line-indices also show gradients, which indicate that the IMF is more bottom-heavy in the center than it is beyond the half-light radius (Martín-Navarro et al. 2015; Lyubenova et al. 2016; van Dokkum et al. 2017; La Barbera et al. 2017; Li et al. 2017; Parikh et al. 2018), although there is not yet universal agreement (Alton et al. 2017; Vaughan et al. 2017). These tend to further increase the stellar population-based $M_*^{\rm SP}/L$ estimate in the central regions. The panel which is second from bottom right of Figure 10 of van Dokkum et al. (2017) suggests that the net effect can be a factor of 3 or more, meaning that IMF-gradients are the dominant contribution to gradients in Υ_*. Importantly, a factor of 3 variation is too large to be safely ignored when using the Jeans-equation.

van Dokkum et al.'s conclusions are based on only 6 objects. However, Parikh et al. (2018) study a much larger sample (by more than two orders of magnitude), drawn from the MaNGA survey, and they too see IMF-related gradients. The IMF gradients they report are slightly weaker than those in van Dokkum et al., and substantially weaker at smaller masses, but they are large enough that they should not be ignored.

In Section 2 we present the first analysis of how such IMF-driven M_*/L gradients impact Jeans-equation estimates of $M_*^{\rm dyn}$. In Section 3 we show how recalibrated masses modify the $M_*^{\rm dyn}/M_*^{\rm SP}$ scaling, as well as the stellar mass function. A final section summarizes and discuss consequences for estimates of dark matter fractions and evidence for/against adiabatic contraction. An Apppendix provides a fully analytic toy model which illustrates the main features of our results.

When necessary, we assume a spatially flat background cosmology with parameters $(\Omega_m, \Omega_\Lambda) = (0.3, 0.7)$, and a Hubble constant at the present time of $H_0 = 70$ km s^{-1}Mpc^{-1}, as these are the values adopted in most studies of the stellar mass function which we reference in our work. As we will be working at low z, all our conclusions are robust to small changes in these parameters.

2 EFFECT OF Υ_* GRADIENTS ON STELLAR MASS ESTIMATES

We study the effect of gradients in the observed (2d, projected) stellar mass-to-light ratio, Υ_*. We show that increasing Υ_* in the central regions has a much more dramatic effect on $M_*^{\rm dyn}$ than on $M_*^{\rm SP}$. Ignoring the gradient leads to an over-estimate of $M_*^{\rm dyn}$. Much of the analysis here is numerical. See Appendix A for a toy model which makes use of simple analytic approximations to the three-dimensional profiles of galaxies to illustrate these same points.

2.1 Observationally motivated scalings

We begin with the projected light distribution, $I(R)$, which is observed to follow a Sérsic (1963) profile:

$$I(R) = \frac{L}{R_e^2} \frac{b^{2n}}{2\pi\, n\Gamma(2n)}\, \mathrm{e}^{-b_n\, (R/R_e)^{1/n}}, \qquad (1)$$

with $b_n \approx 2n - 1/3 + 0.01/n$.

For the projected mass-to-light ratio we set

$$\Upsilon_*(R) = \Upsilon_{*0}\,(1 + \alpha - \beta\, R/R_e) \quad \text{if} \quad R/R_e \le \alpha/\beta, \quad (2)$$

with $\Upsilon_*(R) = \Upsilon_{*0}$ at larger R. The values $(\Upsilon_{*0}, \alpha, \beta) = (3, 2.33, 6)$ provide a good approximation to the scalings shown in the panel which is second from right in the bottom of Figure 10 of van Dokkum et al. (2017).

The total mass associated with this gradient is

$$M_\infty \equiv 2\pi \int_0^\infty \mathrm{d}R\, R\, J(R) = \Upsilon_{*0}L\left(1 + g(n, \alpha, \beta)\right), \quad (3)$$

where

$$J(R) \equiv I(R)\,\Upsilon_*(R). \qquad (4)$$

It is useful to think of the second term,

$$M_{\rm grad} \equiv \Upsilon_{*0}L\, g(n, \alpha, \beta), \qquad (5)$$

as the extra mass contributed by the gradient. For $\Upsilon_*(R)$ given by equation (2), $M_{\rm grad}$ can be written in terms of incomplete gamma functions. For $n = 6$ (Meert et al. 2015 show that this is typical for the massive galaxies, $M_*^{\rm SP} > 10^{11}\ {\rm M}_\odot$, of most interest in this paper), and the values of α and β given above, $M_{\rm grad} = 0.4\,\Upsilon_{*0}L$; the gradient term contributes an additional 40% to the total mass. The ratio $M_{\rm grad}/\Upsilon_{*0}L$ decreases for smaller n.

In what follows, we will illustrate our results using three choices for the gradient strength. These are summarized in Table 1. The smallest values are chosen to mimic the gradient associated with fixed-IMF (panel which is second from bottom left in Figure 10 of van Dokkum et al. 2017), and the largest are the values we just described (second from bottom right of Figure 10 in van Dokkum et al. 2017). We also study an intermediate case, in which the IMF in the central regions cannot become arbitrarily bottom heavy, but is capped at Salpeter (1995): one may think of this as a simple approximation to the gradients reported by Parikh et al. (2018), but we caution that they provide a number of different estimates of IMF-gradients which vary widely. While our three models have different slopes for Υ_*, the transition scale beyond which Υ_* becomes constant (i.e. Υ_{*0}) is the same: it equals $\alpha/\beta \sim 0.4R_e$. We have not explored models where this scale is also varied, simply because varying the

Table 1. Parameter values for M/L gradient strength (equation 2) driven by different assumptions about the IMF (fixed or variable) used in this study.

Model	IMF-var.	α	β	Source
fixed IMF	no	0.39	1.00	Chabrier
Salp$^{\rm IN}$-Chab$^{\rm OUT}$	yes	1.29	3.33	Salpeter-Chabrier
vD17	yes	2.33	6.00	van Dokkum et al.

slope alone allows us to illustrate all the important steps in the argument.

2.2 Stellar population estimate

The presence of gradients in Υ complicate interpretation of published $M_*^{\rm SP}$ values. To illustrate why, we consider various ways in which $M_*^{\rm SP}$ may have been estimated. These depend on the answer to two questions: Was the IMF treated as a free parameter? Was spatially resolved information used?

For the most common estimates, the answer to both questions is 'no'. For such fixed-IMF estimates, typically based on integrated multi-band photometry, what is returned is $\Upsilon_{*\rm fix}$, which is then multiplied by the luminosity to give $M_{\infty,\rm fix}$. In the current context, this would correspond to M_∞ of equation (3) with α, β values from the first row in Table 1 (i.e. Model = fixed IMF). The corresponding estimate when the IMF is allowed to vary is given by either of the other two pairs in the Table. In general, these will be larger than the one for a fixed IMF.

Often, however, Υ_* is estimated from spectroscopic measurements (line indices, etc.) which probe scales that are of order R_e or smaller. A particularly relevant example in the current context is Conroy & van Dokkum (2012); their IMF estimates were based on the light within $R_e/8$. If we use $R_{\rm obs}$ to denote this scale, then the reported $M_*^{\rm SP}$ is given by

$$M_*^{\rm SP} \approx \frac{2\pi \int_0^{R_{\rm obs}} \mathrm{d}R'\, R'\, J(R')}{2\pi \int_0^{R_{\rm obs}} \mathrm{d}R'\, R'\, I(R')} \times L, \qquad (6)$$

and will be greater than M_∞. Hence, if the IMF was a free parameter, then the estimate from within $R_{\rm obs}$ is an overestimate. On the other hand, if the IMF was held fixed, then the overestimate relative to $M_{\infty,\rm fix}$ might still be smaller than the true M_∞ (the one when IMF-gradients are allowed), because $M_{\infty,\rm fix} < M_\infty$. We show examples of this in Section 3.

2.3 Dynamical estimate

Dynamical mass estimates are also affected. For $M_*^{\rm dyn}$, the most relevant quantities are the magnitude of $M_{\rm grad}$ (equation 5) and its spatial distribution. Recall that, for the numbers given earlier (the model with the strongest gradient, i.e. Model = vD17 in Table 1), $M_{\rm grad}$ contributes an additional 40% to the total mass. However, this additional mass is entirely within $0.4R_e$. Absent gradients, the mass with $0.4R_e$ equals about $0.3\,\Upsilon_{*0}L$, so the gradient more than doubles the mass within $0.4R_e$. Since the dynamical mass is estimated by normalizing to σ measured within $R_e/8$ or so, the impact of this additional mass on the $M_{\rm dyn}$ estimate will be

much greater than 40%, and more like a factor of 2, as we now quantify.

Following, e.g., Binney & Mamon (1982), the three dimensional stellar mass profile is obtained by deprojecting $J(R) \equiv \Upsilon_*(R) I(R)$, the projected stellar mass profile:

$$\rho_*(r) = -\frac{1}{\pi} \int_r^\infty \mathrm{d}R \, \frac{\mathrm{d}J/\mathrm{d}R}{\sqrt{R^2 - r^2}}. \qquad (7)$$

With $\rho_*(r)$ in hand it is straightforward to obtain $M_*(< r)$, from which (integrating) the Jeans equation,

$$\frac{\mathrm{d}\,\rho_*(r)\sigma^2(r)}{\mathrm{d}r} + \frac{2\beta_\sigma(r)}{r}\,\rho_*(r)\sigma^2(r) = -\rho_*(r)\frac{GM(< r)}{r^2}, \quad (8)$$

yields $\sigma^2(r)$. The quantity $\beta_\sigma(r)$ (not to be confused with our parametrization of the Υ gradient) is the velocity anisotropy parameter: we set it to zero in what follows, and comment on this approximation at the end of this subsection. The enclosed mass $M(< r)$ is the sum of that in stars $M_*(< r)$ and in dark matter $M_{\mathrm{DM}}(< r)$. Although $M_{\mathrm{DM}}(< r)$ is not known, we expect $M(< r) \approx M_*(< r)$ on small enough scales. Light-weighting (rather than mass-weighting) and projecting $\sigma^2(r)$ gives $\sigma_p^2(R)$:

$$I(R)\,\sigma_p^2(R) = 2\int_R^\infty \frac{\mathrm{d}r\, r\,\rho_{\mathrm{L}}(r)\,\sigma^2(r)}{(r^2 - R^2)^{1/2}}\left[1 - \beta_\sigma(r)\frac{R^2}{r^2}\right], \quad (9)$$

where

$$\rho_{\mathrm{L}}(r) = -\frac{1}{\pi} \int_r^\infty \mathrm{d}R \, \frac{\mathrm{d}I/\mathrm{d}R}{\sqrt{R^2 - r^2}}. \qquad (10)$$

is the deprojected light profile. Often, it is the velocity dispersion within an aperture which is observed. This is

$$I(< R)\,\sigma_p^2(< R) \equiv 2\pi \int_0^R \mathrm{d}R\, R\, I(R)\,\sigma_p^2(R). \qquad (11)$$

The Jeans equation analysis treats Υ_{*0} – the constant amplitude factor which appears as the first term on the right hand side of equation (2) – as a free parameter which is fixed by matching the predicted $\sigma_p(< R)$ to that observed. Note that, unless a model for the dark matter is explicitly included, this matching must be done on small enough scales that neglecting dark matter is accurate. Once this has been done, then using $M_*(< r)$ in the Jeans equation even on larger scales yields an estimate of the contribution to the observed $\sigma_p(< r)$ which is due to the stars.

Although the overall normalization Υ_{*0} is a free parameter, the R-dependence of $\Upsilon_*(R)$, the steepness of the gradient, matters in what follows. This steepness will obviously affect the scale dependence of $\sigma_p(R)$. Therefore, anything else that changes the steepness of the predicted $\sigma_p(R)$ will also affect our results. For example, to illustrate our arguments, we set $\beta_\sigma = 0$, meaning we assume orbits are isotropic. Figure 7 in Ciotti & Lanzoni (1997) suggests that anistropies contribute less than 10% effects for the E+S0s of most interest here. We explore this briefly in the next subSection, but a detailed study of anisotropy is beyond the scope of this work.

2.3.1   Effect of gradients

We begin with a cartoon of the effect. The blue dots in Figure 1 represent measurements in data: they show the
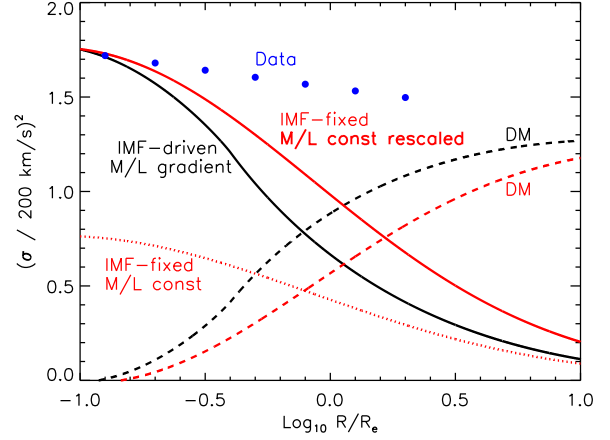
Figure 1. Effect of gradients in $\Upsilon_* = M_*^{\mathrm{SP}}/L$ (driven by IMF-gradients) on M_*^{SP} and M_*^{dyn} estimates. Filled blue circles represent the measured velocity dispersion (equation 12). Lower dotted red curve shows the predicted shape for a fixed IMF (e.g. Chabrier 2003) and a constant mass-to-light ratio when the light profile is Sérsic with $n = 6$; upper solid red curve shows the M_*^{SP}/L that is required for this shape to fit the observed σ^2 on small scales. The discrepancy between the solid red curve and the blue circles on larger scales (dashed red curve) is attributed to dark matter. Black solid and dashed curves show the corresponding results if M_*^{SP}/L increases towards the center as given by equation (2) with $(\alpha, \beta) = (2.33, 6)$. As the stellar mass is now more centrally concentrated, the associated velocity dispersion falls more steeply from the center, so that more dark matter is needed within R_e to explain why the observed σ^2 is relatively flat.

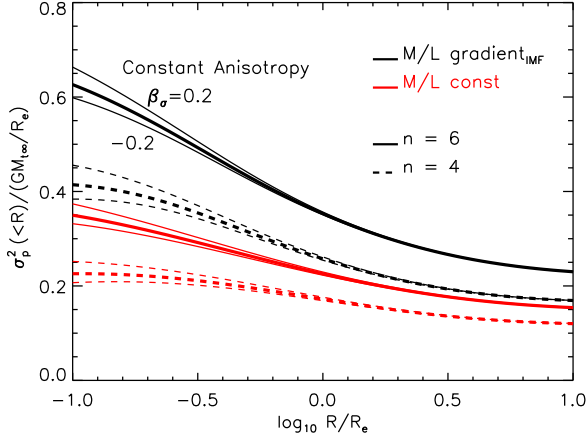mean light-weighted projected velocity dispersion within an aperture of projected size R_e:

$$\sigma_p(< R) = \sigma_p(< R_e)\,(R_e/R)^{0.06} \qquad (12)$$

(Jørgensen et al. 1995; Cappellari et al. 2006). The solid black curve shows $\sigma_p(< R)$ returned from our Jeans analysis. We scaled its height so that it matches the blue dots at $R = R_e/8$, as this is a common choice. (Normalizing on small R is necessary to be as immune to dark matter as possible.) The dynamical mass estimate follows from the procedure we described earlier. The black curve falls more steeply than the dots, showing that the mass in stars cannot account for the velocity dispersion observed on larger scales. The dashed curve shows the difference between the observed dispersion and that predicted by the stellar mass. This difference is usually attributed to dark matter. This additional contribution exceeds that from the stars beyond about $0.8 R_e$.

The red curve shows the corresponding estimate if we ignore gradients by setting $J(R) \propto I(R)$ and then follow all the same steps as before. When gradients are included, then $\rho_*(r)$ is steeper, and $\sigma_p(< R)$ is too. Therefore, a smaller total dynamical stellar mass can account for the small scale σ_p (because the mass is more centrally concentrated).

A crude estimate of the difference in the total stellar dynamical mass estimates (i.e., $M_{*\mathrm{dyn}}$, not the total dynamical stellar+dark matter mass) can be got as follows. Since $M_{*\mathrm{dyn}}(< R) \sim R\sigma^2(< R)/G$, the ratio of the red and black solid curves is approximately equal to the ratio of the two stellar mass estimates enclosed within R. At
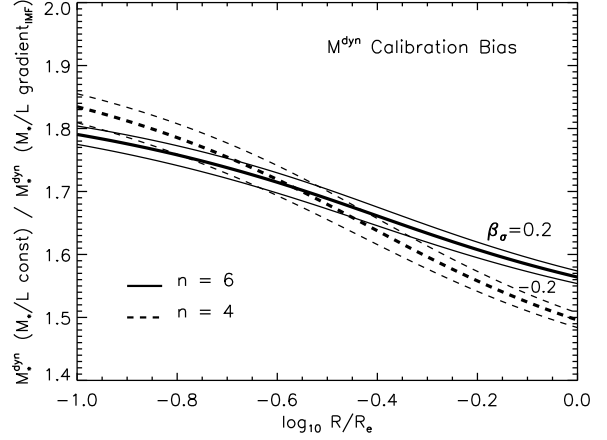
Figure 2. Effect of Υ_* gradients on the shape of the light-weighted velocity dispersion within a circular aperture of radius R, for three choices of the velocity anisotropy parameter $\beta_\sigma = (-0.2, 0, 0.2)$ when $n = 4$ (lower) and 6 (upper). Red curves assume that mass is proportional to light, and black curves include a gradient in Υ_* (equation 2 with $\alpha = 1.29$ and $\beta = 3.33$, i.e. Model = Salp$^{\mathrm{IN}}$-Chab$^{\mathrm{OUT}}$ in Table 1). The three sets of curves for each n show how a spatially constant velocity anisotropy $\beta_\sigma = (-0.2, 0, 0.2)$ affects the shape of the profile.

Figure 3. Bias in $M_{*\mathrm{dyn}}$ which results from ignoring Υ_* gradients, shown as a function of the scale on which the models are normalized to match the observed $\sigma_p^2(<R)$. This bias is given by the ratio of the black to corresponding red curves in the previous figure. When calibrating to the velocity dispersion measured inside $R_e/8$ – where dark matter is not expected to contribute significantly – the overestimate in M^{dyn} is ~ 1.8 for the Salp$^{\mathrm{IN}}$-Chab$^{\mathrm{OUT}}$ model (it is ~ 2 for the vD17 model). We do not show the ratio for $R > R_e$, because ignoring the effects of dark matter when calibrating to large scales is incorrect.

$R \ll R_e$, the two are similar, since both models are normalized to match the small scale velocity dispersion (the blue dots at $R \ll R_e$). However, the total stellar dynamical mass is $M_{*\mathrm{dyn}}(< R)$ as $R \to \infty$. So we must compare the red and black curves, not at $R = R_e/8$ where they are equal (by design), but at $R \gg R_e$. (At these large R, both curves lie well-below the blue dots because dark matter matters and these solid curves only show the stellar component.) There, the two $M_{*\mathrm{dyn}}(< R)$ estimates are quite different: the red is about $2\times$ larger than the black. This indicates that incorrectly ignoring gradients overestimates the total (i.e., the $R \to \infty$) $M_{*\mathrm{dyn}}$ estimate by a factor of about 2.

Moreover, since the solid red curve falls less steeply than the black one, it is closer to the observations over a wider range of scales. Therefore, the associated dark matter estimate (dashed red) is smaller: it only dominates the estimated contribution from the stars beyond about $1.8R_e$. Thus, ignoring gradients will lead one to systematically underestimate the actual dark matter contribution on small scales. The Appendix shows that these are generic trends.

Figure 2 shows how these trends depend on the Sérsic index n and on the whether or not the velocity dispersions are isotropic. In anticipation of our consideration of the data in the next section, we first truncate the projected Sérsic profile at $7.5R_e$. This reduces the total luminosity L by an n-dependent amount, which is of order 10%, and we use $M_{t\infty}$ to denote the total mass associated with this truncated profile. The figure shows two sets of black curves, and two sets of red: solid and dashed curves show results for $n = 4$ and 6. In all cases the curves show $\sigma_p^2(< R)/(GM_{t\infty}/R_e)$, where R_e is the projected half-light radius of the original (untruncated) profile. The lower (red) curves show the profiles if there is no Υ_* gradient: $(\alpha, \beta) = 0$. The upper (black) curves include a gradient, for which we used the intermediate set of values from Table 1: $(\alpha, \beta) = (1.29, 3.33)$. The central curve of each

set is thicker; it assumes the velocity dispersion is isotropic. The other two are for a constant anistropy of $\beta_\sigma = \pm 0.2$ at all r. (As an aside, comparison of our red $\beta_\sigma = 0$ curves with Figure 11 of Prugniel & Simien 1997 shows that truncating the light profile makes only a small difference.) While $\beta_\sigma \neq 0$ makes a visible difference, it is small compared to the effect of the Υ_* gradient. Moreover, as we argue below, the real quantity of interest is the ratio of each black curve to its corresponding red one: this removes most of the effect of $\beta_\sigma \neq 0$.

The estimated dynamical mass is

$$M_*^{\mathrm{dyn}} = k(R, n, \alpha, \beta)\, R_e\, \sigma_p^2(<R)/G, \qquad (13)$$

where $k(R, n, \alpha, \beta)$ is the inverse of the predicted $\sigma_p^2(< R)/(GM_{t\infty}/R_e)$ on scale R (e.g. Bernardi et al. 2018). Hence, the ratio of the black and red curves gives the amount by which one overestimates the total M_*^{dyn} if one ignores the gradient and normalizes to $\sigma_p(<R)$. Note that R appears in k to highlight the fact that k depends on the scale on which one chooses to normalize the model to the observed $\sigma_p(<R)$. (If there were no dark matter, and the stellar mass profile and velocity anistropies were correctly modelled, then this dependence would make the estimated M_*^{dyn} the same for all choices of R.)

Figure 3 shows this ratio for the various pairs of curves in Figure 2. If $R \sim R_e/8$, the ratio is nearly a factor of ~ 2. If, instead, one normalized at $R \sim R_e$, then this ratio is slightly smaller. Recall, however, that this ratio is only meaningful if dark matter makes a negligible contribution to $\sigma_p(< R)$. If there are no gradients, then dark matter already contributes significantly at $R \sim R_e$; if gradients matter, then the scale where dark matter can be ignored is even smaller (c.f. Figure 1). Therefore, only the $R \ll R_e$ values shown in Figure 3 are likely to be reliable.

The steepening of the stellar contribution to σ_p (because of Υ_* gradients) is what leads to a greater need for dark matter on small scales. However, anistropic velocity dispersions can also affect the shape of σ_p. So, if β_σ is not constrained by other observations, but is determined from the same Jeans equation analysis which determined Υ_{*0} and hence M_*^{dyn}, then there is some room for degeneracy between anisotropy and dark matter fraction on small scales. Previous work suggests this is a small effect (Ciotti & Lanzoni 1997). The similarity of the different curves for each n in Figure 3 suggests that the effect of including anisotropic velocities is indeed small compared to the factor of two coming from the IMF-driven Υ_*-gradient. I.e., anisotropic velocity dispersions, if present, will not change our main point that current estimates of the dark matter fraction must be revisited if gradients matter.

To summarize: observations suggest that the stellar mass-to-light ratio Υ_* increases towards the center of a galaxy (equation 2 and Table 1). If these gradients are ignored, then (i) stellar population-based mass estimates may be biased, and (ii) dynamical mass estimates will be more strongly biased towards overestimating the total stellar mass and underestimating the amount of dark matter on small scales as a result. In the next section, we use these results to illustrate the impact of gradients on the inferred stellar mass function.

3   POTENTIAL IMPACT OF Υ_* GRADIENTS IN SDSS

This section illustrates the potential impact of Υ_* gradients using the galaxies in the SDSS DR7 Main Galaxy sample (Abazajian et al. 2009).

3.1   The sample

We select the galaxies in the SDSS DR7 with r-band Petrosian magnitude limits $14 \leq m_r \leq 17.77$ mag (see Meert et al. 2015 for a detailed discussion of the sample selection), and we use the PyMorph photometry of Meert et al. (2015). The differences between PyMorph and SDSS pipeline photometry are significant for the most massive galaxies. See Bernardi et al. (2013), Meert et al. (2015), Fischer et al. (2017), and Bernardi et al. (2017b) for why PyMorph is preferred.

For single Sérsic fits, the relevant PyMorph parameters are the Sérsic index n, half-light radius R_e and total luminosity L of each object. The estimated total light L results from extrapolating the fitted (Sérsic) model to infinity. As a result, the single Sérsic fits are known to slightly overestimate the total light; integrating out to only $7.5R_e$ yields a more reliable luminosity estimate (Bernardi et al. 2017a; Fischer et al. 2017; Bernardi et al. 2018). In what follows, we will always use PyMorph truncated luminosities.

We will consider E+S0s separately from the full population. For this, we use the Bayesian Automated morphological classifications (hereafter BAC) of Huertas-Company et al. (2011), because they provide a probability $p(\mathrm{type})$ for each object. We can either weight by, or implement hard cuts in, this probability.

Fixed IMF (Chabrier 2003) $M_*^{\mathrm{SP-Chab}}$ estimates for all these galaxies are also available. These combine the dusty and dust-free $M_*^{\mathrm{SP-Chab}}/L$ estimates of Mendel et al. (2014), obtained from integrated multiband photometry, with the truncated Sérsic L of Meert et al. (2015). We correct these values for gradients using

$$M_*^{\mathrm{SP}} = M_*^{\mathrm{SP-Chab}} \frac{1 + g(n, \alpha, \beta)}{1 + g(n, 0.39, 1)}, \qquad (14)$$

where the values of α, β are taken from Table 1 as we discuss below. Note that the values of g here are for a Sérsic profile truncated at $7.5R_e$.

The SDSS pipeline also provides estimates of the velocity dispersion estimated within a circular fiber of radius $\theta_{\mathrm{fiber}} = 1.5$ arcsec. For a galaxy at redshift z with Sérsic index n and half light radius R_e, the discussion of the previous section suggests defining M_*^{dyn} using equation (13), with $R_{\mathrm{obs}}/R_e = d_{\mathrm{A}}(z)\theta_{\mathrm{fiber}}$. However, R_{obs} is approximately $R_e/2$ for the E+S0s in our sample. In view of our discussion of how gradients will tend to increase the need for dark matter (c.f. Figure 1), it is likely that M_*^{dyn} estimated using σ_{obs}^2 will be biased by dark matter. For this reason, we first aperture correct σ_{obs} to $R_e/8$ using equations (6) and (7) of Bernardi et al. (2018): $\sigma(<R)/\sigma(<R_e) = (R/R_e)^{-\gamma(n)}$. (This accounts for a weak dependence of the slope γ on the Sérsic index n instead of being fixed to a constant, 0.06, as in equation 12). We then use this corrected value to estimate

$$M_*^{\mathrm{dyn}} = k(R_e/8, n, \alpha, \beta) \left(\frac{R_{\mathrm{obs}}}{R_e/8}\right)^{\gamma(n)} \frac{R_e \sigma_{\mathrm{obs}}^2}{G}, \qquad (15)$$

where α and β are the same values we use in equation (14), and the entire analysis uses the truncated Sérsic profile. In this respect, our $(\alpha, \beta) = (0, 0)$ analysis differs from that in Bernardi et al. (2018) who used a truncated Sérsic for the light, but whose procedure for estimating M_*^{dyn} was more complex than ours (the net difference in M_*^{dyn} is small).

The aperture correction (i.e., γ) as well as the gradient correction (i.e., k) depend on galaxy type (i.e., on n). Whereas the aperture correction is a small effect, the gradient correction matters very much. Further work is necessary to quantify these in data, so here we will explore two simple modifications to the assumption that α and β are the same for all galaxies.

3.2   Type-dependent gradients

Our first model is motivated by Parikh et al. (2018), who suggest that gradients are smaller in lower mass galaxies. Therefore, for fixed-IMF M_*^{SP} values that are between 10.3 and 11.2 dex we multiply both α and β by

$$C_* = [\log_{10}(M_*^{\mathrm{SP-Chab}}/M_\odot) - 10.3]/(11.2 - 10.3), \quad (16)$$

with α and β equal to zero at smaller masses. This changes the slope of Υ_*, but keeps the transition scale beyond which $\Upsilon_*(R)$ becomes constant equal to $\alpha/\beta \sim 0.4R_e$ for all galaxies. We used the two mass scales identified by Bernardi et al. (2011) as being special: Various scaling relations change slope at these scales. We also study a model in which the gradient is weaker at small σ: For $\sigma_{e/8}$ between 100 and 250 km s^{-1} we multiply both α and β by

$$C_\sigma = (\sigma_{e/8}/\mathrm{kms}^{-1} - 100)/(250 - 100), \qquad (17)$$

with α and β equal to zero at smaller σ. The threshold values of $\sigma_{e/8}$ are those associated with our threshold values of $M_*^{\mathrm{SP-Chab}}$ based on the $\sigma_{e/8} - M_*^{\mathrm{SP-Chab}}$ relation. In what follows, we compare results using α and β from Table 1, with and without the mass- or σ-dependent rescalings we have just described.

Scaling with σ rather than M_* is motivated by the fact that IMF-indicators appear to be strongly correlated with σ (Conroy & van Dokkum 2012; Lyubenova et al. 2016). Moreover, a number of groups have reported that, when gradients are ignored, then $M_*^{\mathrm{dyn}}/M_*^{\mathrm{SP-Chab}}$ is large if σ is large (Cappellari et al. 2013; Li et al. 2017). This has fueled the argument that M_*^{dyn} estimates are robust, but fixed-IMF M_*^{SP} estimates are biased low if they have not accounted for (σ-dependent) IMF variations across the population. However, while van Dokkum et al. (2017) find a large scatter in Υ_* values within $R_e/8$, the range of values at R_e and beyond is much smaller. If there is a floor to the Υ_* value which is the same for all galaxies, then the ones with small σ, which Conroy & van Dokkum (2012) report have small Υ_* even at $R_e/8$, must have small gradients.

If we make a σ-dependent change to gradients, then we expect that accounting for them will impact the $M_*^{\mathrm{dyn}}/M_*^{\mathrm{SP-Chab}} - \sigma$ relation. How this will differ from scaling gradient strength with M_* is less obvious, which is why we believe it is interesting to compare both.

3.3 Correlation between stellar population and dynamical masses

Figure 4 quantifies the impact of Υ_* gradients on the $M_*^{\mathrm{dyn}}/M_*^{\mathrm{SP-Chab}} - \sigma$ relation using the three pairs of α, β given in Table 1. Assuming the same gradients for all objects changes the amplitude but not the slope of the relation (small red dots in the left hand panels) compared to when one estimates M_*^{dyn} assuming a constant Υ_* (small green dots, same in all panels); reducing the gradient strength at small masses flattens and even reverses the correlation (small blue dots in the right hand panels). Decreasing the gradients for objects with small σ produces very similar results (small red dots in the right hand panels).

The bottom panels show that the change in amplitude is substantial if we use the largest values of (α, β) in Table 1 (i.e., Model = vD17); this yields M_*^{dyn} values that are too small compared to the corresponding $M_*^{\mathrm{SP-Chab}}$ values. Top panels show that even the much weaker gradients (Model = fixed IMF) produce noticable effects. In all cases, the panels on the right – in which gradients are weaker at low mass or velocity dispersion – show shallower or even reversed correlations. Note that the Figure only shows the effects of changing M_*^{dyn}; including the changes to M_*^{SP} as well yields qualitatively similar results, although the quantitative change to M_*^{SP} depends on whether the value of Υ_* which multiplies L was estimated only within the central regions or not. We illustrate the effect of changes to M_*^{SP} in Section 3.4.

This weakening of the correlation may account for conflicting claims in the literature about the veracity of the IMF-σ correlation and its relation to M_*^{dyn}. Conroy & van Dokkum (2012) report a relation which is steeper than most $M_*^{\mathrm{dyn}}/M_*^{\mathrm{SP}}$-$\sigma$ relations (e.g. Cappellari et al. 2013; Li et al. 2017). However, their analysis is really one of the stellar population estimate, based on spectra taken at $R_e/8$;

since this is smaller than $0.4R_e$, their results are potentially strongly impacted by gradients. If the IMF-gradients we have explored are realistic, then our results suggest that the M_*^{dyn} estimates of other groups should be revised downwards, rather than that the M_*^{SP} should be revised upwards. Of course, actual measurements of how the gradient correlates with σ (or M_*^{SP}) are needed to really settle the issue, so we hope that our results motivate further work quantifying such trends.

In the meantime, our results illustrate that requiring $M_*^{\mathrm{dyn}} = M_*^{\mathrm{SP}}$ may provide a useful constraint on gradient strength if there is a lower limit to Υ_{*0} of equation (2). If there is, and one uses it to estimate $M_*^{\mathrm{SP-min}}$ from the light, then adding the mass associated with the Υ_* gradient will increase this minimal SP mass estimate by M_{grad} of equation (5). On the other hand, if this additional mass steepens the mass profile compared to the light, it will decrease the dynamical mass estimate (Figure 1 and related discussion). As a result, if gradients are too strong, then it is possible that $M_*^{\mathrm{dyn}} < M_*^{\mathrm{SP-min}}$, which is unacceptable. The bottom left panel of Figure 4 shows a model for the Υ_*-gradient which results in $M_*^{\mathrm{dyn}} < M_*^{\mathrm{SP-Chab}}$ (by 0.2 dex on average). If $M_*^{\mathrm{SP-Chab}} \approx M_*^{\mathrm{SP-min}}$, then these gradients are unrealistic.

3.4 The stellar mass function

Including IMF variations across the galaxy population, but ignoring gradients within each galaxy, results in a substantial increase in the inferred stellar mass density that is locked-up in massive galaxies compared to if one assumed a Chabrier IMF for the full population (Bernardi et al. 2018). In effect, the IMF variation increases the stellar population estimates so that they come into good agreement with the dynamical estimates. However, the previous subsection showed that accounting for Υ_* gradients within each galaxy will reduce M_*^{dyn}; whether M_*^{SP} is also reduced, or must be increased slightly, depends on how it was estimated. Since these reduction factors can be different, it is not obvious that the two can remain in agreement. Figure 4 suggests that if gradients are too strong, then $M_*^{\mathrm{dyn}} < M_*^{\mathrm{SP}}$ which is obviously problematic (bottom panels). But for the intermediate strength gradients from our Table 1 (Model = Salp$^{\mathrm{IN}}$-Chab$^{\mathrm{OUT}}$), the two mass estimates can be in reasonable agreement, if gradients are weaker in lower mass galaxies (right middle panel).

Figures 5 and 6 show how these effects manifest in the mass function. They correspond to the models shown in the left and right hand panels of Figure 4: i.e. same gradient across the population, or σ-dependent gradients. In all panels, solid blue and hashed cyan regions show $\phi(M_*^{\mathrm{SP-Chab}})$ and $\phi(M_*^{\mathrm{dyn}})$ if gradients are ignored (from Bernardi et al. 2018); except for the top panel of Figure 5 we have normalized the results by a fiducial curve to more clearly illustrate the differences. The large differences between the blue and cyan regions, especially at large masses, are usually attributed to problems with $M_*^{\mathrm{SP-Chab}}$ rather than M_*^{dyn}, but our analysis suggests that accounting for gradients might lead to the opposite conclusion.

The pink, purple and magenta regions (top to bottom) in Figure 5 show the effect of Υ_* gradients when they are the same across the population. If there are no IMF gradients,
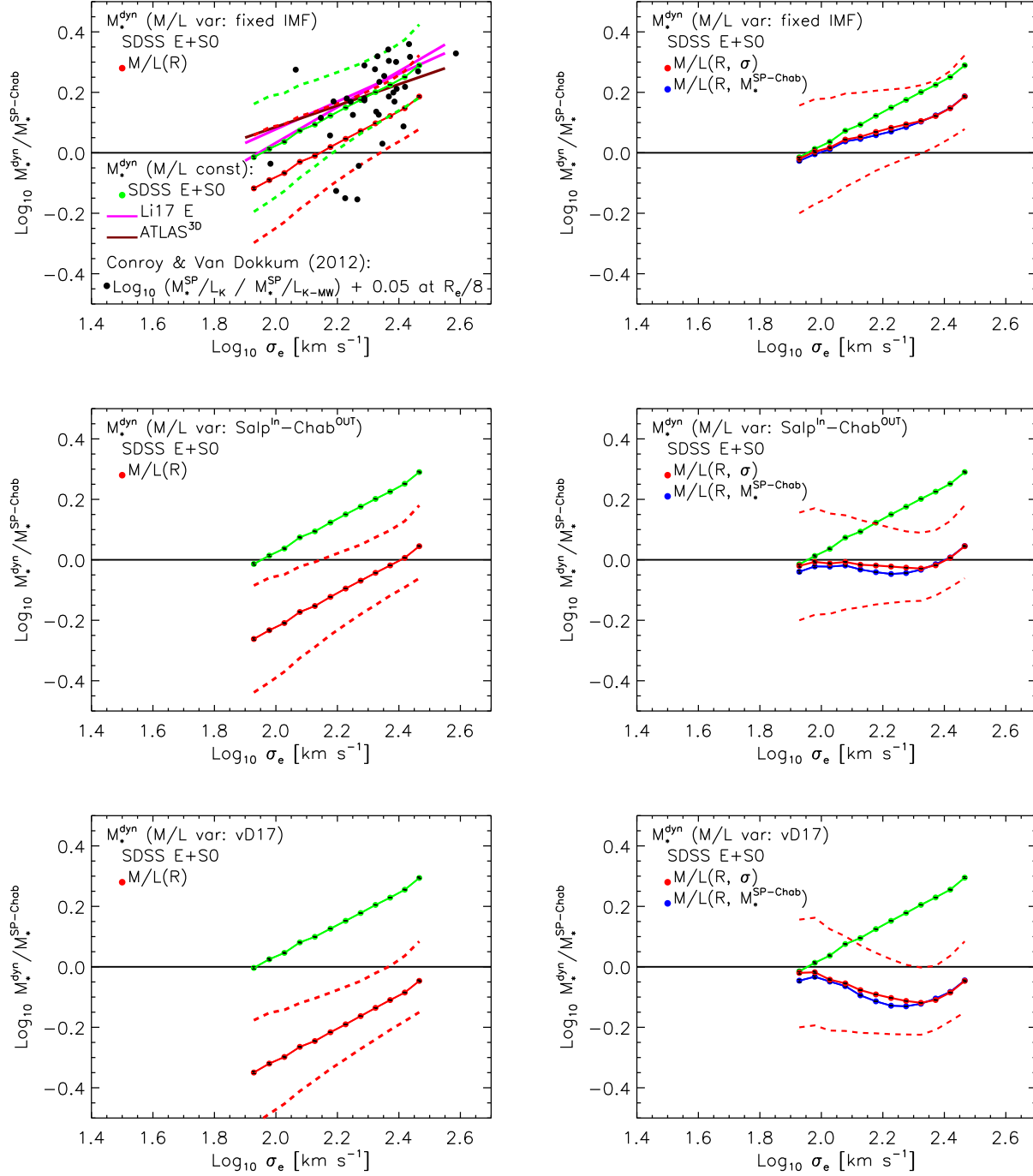
Figure 4. Effect of Υ_* gradients on the $M_*^{\mathrm{dyn}}/M_*^{\mathrm{SP}} - \sigma$ correlation. *Previous work, ignoring Υ_* gradients* (top left panel): Straight lines, same in each panel, show this correlation in the MaNGA (magenta lines; Li et al. 2017, from STARLIGHT and pPXF models) and ATLAS^{3D} (brown line; Cappellari et al. 2013) datasets if gradients are ignored, and M_*^{SP} estimates assume a Chabrier IMF for all galaxies. Black symbols by Conroy & van Dokkum (2012; see text for details) show the ratio between stellar mass estimates when the IMF is free to when it is fixed. (We have shifted their M_* estimate to account for the fact their Milky Way IMF is based on Kroupa 2001, which differs by 0.05 dex from Chabrier 2003.) Note, that these estimates were computed in the central regions (at $R_e/8$). Small green dots connected by a solid line, same in each panel, show the median $M_*^{\mathrm{dyn}}/M_* - \sigma$ relation in SDSS E+S0s if gradients are ignored; dashed lines show the region which encloses 68% of the objects at each σ_e (similar to Bernardi et al. 2018). *This work, accounting for Υ_* gradients* (all panels): Small red dots and associated curves result from accounting for gradients using the three models described in Table 1. Panels on the left assume the gradient strength is the same for all objects, and is due to a fixed (Chabrier) IMF (top – fixed IMF); to the IMF varying between Salpeter in the center and Chabrier beyond $0.4R_e$ (middle – Salp$^{\mathrm{IN}}$-Chab$^{\mathrm{OUT}}$); and to even larger gradients because the center is even more extreme than Salpeter (bottom – vD17). Stronger gradients result in greater offsets from the no-gradient case. Panels on the right show the same models, except that the gradient strength is assumed to decrease at lower masses (equation 16 – small blue dots) and lower velocity dispersions (equation 17 – small red dots). Both show flatter correlations; the slope in the panel on the bottom-most panel is even reversed.
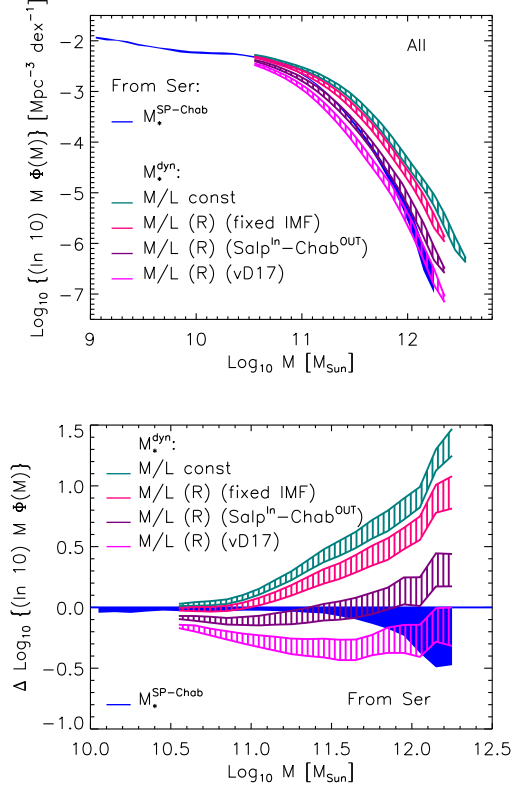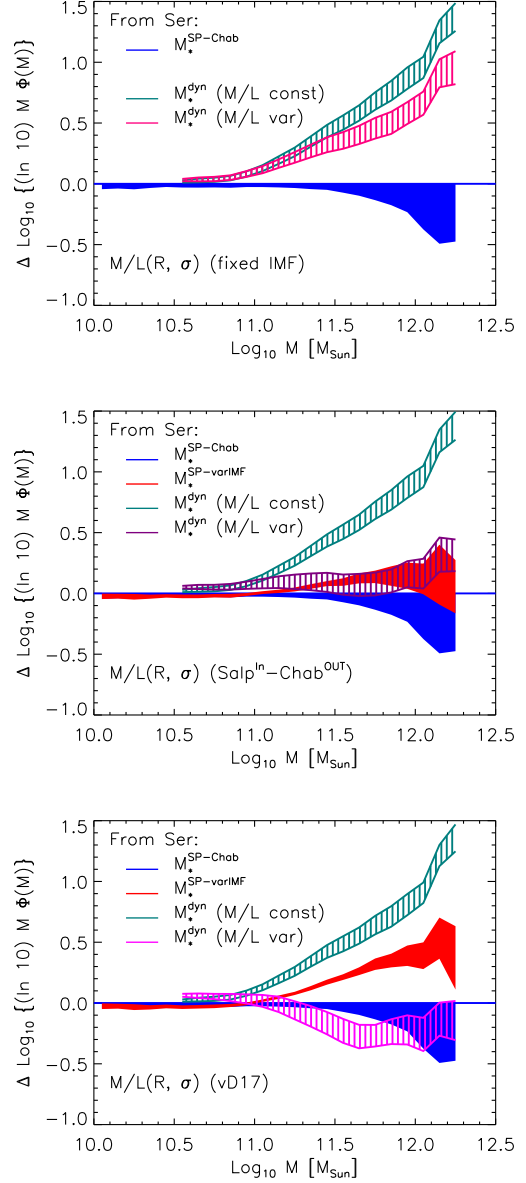
Figure 5. Stellar mass functions for the models shown in the left hand panels of Figure 4. Solid blue region shows the fixed-IMF estimate $\phi(M_*^{\mathrm{SP-Chab}})$ and hashed cyan region shows the corresponding $\phi(M_*^{\mathrm{dyn}})$ if gradients are ignored (from Bernardi et al. 2018). Hashed pink, purple and magenta regions (top to bottom) show the result of accounting for Υ_* gradients when estimating M_*^{dyn}, as labeled. Bottom panel shows the same results normalized by a fiducial curve to reduce the dynamic range. Clearly, accounting for Υ_* gradients brings $\phi(M_*^{\mathrm{dyn}})$ to better agreement with $\phi(M_*^{\mathrm{SP-Chab}})$ – this is because M_*^{dyn} has been reduced, rather than because M_*^{SP} has been increased.

so the only Υ_* gradients are those associated with a fixed IMF, then $M_*^{\mathrm{SP}} = M_*^{\mathrm{SP-Chab}}$ (see equation 14), and Υ_* gradients only modify M_*^{dyn}. Since they are weak, the decrease in M_*^{dyn} is small: this brings the hashed cyan region down to the hashed pink region. This is why such gradients are usually ignored. However, the stronger IMF-driven gradients produce more dramatic changes to $\phi(M_*^{\mathrm{dyn}})$.

The three panels in Figure 6 show how gradients impact the stellar mass function for the three models shown on the right hand side of Figure 4: these use the three sets of α, β given in Table 1, but scaled so that objects with small σ have smaller gradients (recall that scaling with mass gives similar results). The stellar population and dynamical mass results are similar only in the middle panel. When gradients are too weak (top panel), then $M_*^{\mathrm{SP}} < M_*^{\mathrm{dyn}}$, whereas the opposite is true when gradients are too strong (bottom panel). In the middle panel however, the two estimates are slightly larger than the fixed-IMF estimate $\phi(M_*^{\mathrm{SP-Chab}})$ which is currently in the literature (Bernardi et al. 2017a, 2018), but substantially smaller than the dynamical estimate associated with ignoring gradients. I.e., agreement between the

Figure 6. Stellar mass functions for the models shown in the right hand panels of Figure 4, divided by a fiducial curve to reduce dynamic range. Blue solid and hashed cyan regions, same in all panels, show fixed IMF $\phi(M_*^{\mathrm{SP-Chab}})$, and $\phi(M_*^{\mathrm{dyn}})$ if gradients are ignored. Red regions show $\phi(M_*^{\mathrm{SP}})$ if σ-dependent gradients have been accounted for. Pink, purple and magenta show the corresponding $\phi(M_*^{\mathrm{dyn}})$ estimates. Purple and red in middle panel agree because M_*^{dyn} has been reduced, rather than because M_*^{SP} has been increased. They disagree significantly in bottom panel, where gradients are strongest (also, $\phi(M_*^{\mathrm{dyn}} < \phi(M_*^{\mathrm{SP}})$, which is opposite compared to top panel).

two stellar mass functions is achieved by decreasing M_*^{dyn} so it matches $M_*^{\mathrm{SP-Chab}}$ rather than the other way round.

Clearly, an accurate census of the stellar mass density requires that gradients in Υ_* be well-quantified.

4   DISCUSSION

Observations suggest higher stellar mass-to-light ratios Υ_* in the central regions of galaxies. We illustrated a number of ways in which such gradients in Υ_* impact stellar population and dynamical mass estimates of galaxies.

Ignoring such gradients leads to overestimates of the dynamical stellar mass (Figure 1 and related discussion, as well as Figure A1 and equation A4), with $M_*^{\rm dyn}$ being more strongly biased (Figures 2 and 3). Since gradients are not well-quantified in the literature, we illustrated their effects using three models (equation 2 with Table 1). We noted that, if there is a minimum value to Υ_*, then requiring agreement between $M_*^{\rm dyn}$ and $M_*^{\rm SP}$ limits the allowed gradient strength (bottom left panel of Figure 4 and related discussion). We also explored the possiblitity that Υ_*-gradient strength depends on galaxy type – e.g. if gradients are larger at higher velocity dispersion or stellar mass. Gradients can modify the $M_*^{\rm dyn}/M_*^{\rm SP} - \sigma$ substantially (Figure 4).

Of the three simple models for gradients which we explored, only the one where gradients are driven by a transition from Salpeter in the inner regions to Chabrier beyond $0.4R_e$ produces agreement between $M_{\rm dyn}$ and $M_*^{\rm SP}$, and then, only if the gradient strength is weaker in less massive galaxies. In this model, IMF-driven gradients bring $M_*^{\rm dyn}$ and $M_*^{\rm SP}$ into agreement, not by shifting $M_*^{\rm SP}$ upwards by invoking constant bottom-heavy IMFs, as advocated by a number of recent studies (e.g. Li et al. 2017), but by revising $M_*^{\rm dyn}$ estimates in the literature (e.g. Cappellari et al. 2013) downwards.

The reduction in $M_*^{\rm dyn}$ has the effect of shifting the cutoff in $\phi(M_*^{\rm dyn})$ to masses that are approximately $2\times$ smaller (Figures 5 and 6). In our Salp$^{\rm IN}$-Chab$^{\rm OUT}$ model, agreement between $\phi(M_*^{\rm SP})$ and $\phi(M_*^{\rm dyn})$ is achieved by decreasing $M_*^{\rm dyn}$ so it matches $M_*^{\rm SP-Chab}$ rather than the other way round. As a result, the shape of the mass function is close to $\phi(M_*^{\rm SP-Chab})$ which is provided in tabular form by Bernardi et al. (2018). Therefore, an accurate census of the stellar mass density requires gradients in Υ_* to be well-quantified.

Smaller dynamical stellar masses (compared to previous $M_*^{\rm dyn}$ estimates, such as those provided by the JAM-analyses of Cappellari et al. 2013 and Li et al. 2017 for the ATLAS$^{\rm 3D}$ and MaNGA samples, respectively) also mean that dark matter must play a more dominant role than previously thought (Figure 1). Although the details will also depend on velocity anisotropy, our analysis of the simplest case suggest these will be sub-dominant (Figures 2 and 3). In turn, this impacts discussions of whether or not dark matter halos must have contracted adiabatically because of the baryons which cooled and formed stars in their centers (Newman et al. 2015; Shankar et al. 2017, 2018).

More centrally concentrated baryonic mass also impacts inferences about the black hole sphere of inference, especially if the Sérsic index n is small. Without gradients, σ_p is predicted to decrease at small r, so if an increase is observed, it may be attributed to a central black hole. Anisotropic velocity dispersions complicate the discussion (Binney & Mamon 1982; van der Marel 1994), but since Υ_*-gradients can also change the small scale slope, they are an additional complication. However, if black hole mass fractions are as small as Shankar et al. (2016) suggest, then they only matter on much smaller scales than those considered here (equation A7 and related discussion).

Finally, more centrally concentrated baryons will change estimates of the radial acceleration relation which is sometimes used to constrain modifications to standard gravity (Chae et al. 2014; Tortora et al. 2014; Chae & Gong 2015; Lelli et al. 2016; Posacki et al. 2016; Tortora et al. 2018; Chae et al. 2018; Appendix A2 here). For all these reasons, we hope that our work stimulates efforts to quantify how gradients in the stellar mass-to-light ratio depend on galaxy type.

Our final point is a philosophical one. If Υ_* gradients can be ignored, then $M_*^{\rm dyn}$ provides an estimate of the stellar mass that is independent of stellar population modeling. This independence is sometimes used to argue that the extra work required to estimate $M_*^{\rm dyn}$ is well-worth the trouble, since $M_*^{\rm dyn}$ is insensitive to the details of the stellar population which can lead to systematic biases in $M_*^{\rm SP}$. However, at the present time, Υ_* gradients are typically estimated from stellar population modeling, so the need to include gradients when estimating $M_*^{\rm dyn}$ removes this independence. As independent checks on mass-estimates are desirable, our work provides strong motivation for alternative probes of the IMF, such as the microlensing method of Schechter et al. (2014).

REFERENCES

Abazajian, K. N., Adelman-McCarthy, J. K., Ageros, M. A., et al. 2009, ApJS, 182, 543

Alton P. D., Smith, R. J. & Lucey, J. R. 2017, MNRAS, 468, 1594

Bell E. F., de Jong R. S., 2001, ApJ, 550, 212

Bender R., Burstein D., Faber S. M., 1992, ApJ, 399, 462

Bernardi, M., Sheth, R. K., Nichol, R. C., Schneider, D. P., Brinkmann, J. 2005, AJ, 129, 61

Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, MNRAS, 404, 2087

Bernardi, M., Roche, N., Shankar, F. & Sheth, R. K. 2011, MNRAS, 412, L6

Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697

Bernardi M., Meert A., Sheth R. K., Fischer J.-L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017a, MNRAS, 467, 2217

Bernardi M., Meert A., Sheth R. K., Fischer J.-L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017b, MNRAS, 468, 2569

Bernardi M., Sheth R. K., Fischer J.-L., Meert A., Chae K.-H., Dominguez-Sanchez H., Huertas-Company M., Shankar F., Vikram, V., 2018, MNRAS, in press (arXiv:1710.07296)

Binney J. J., Mamón G. A, 1982, MNRAS, 200, 361

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Cappellari M., et al., 2006, MNRAS, 366, 1126

Cappellari M., et al., 2013, MNRAS, 432, 1862

Cappellari, M. 2016, ARA&A, 54, 597

Chabrier G., 2003, ApJ, 586, 133

Chae K.-H., Bernardi M., Kravtsov A. V., 2014, MNRAS, 437, 3670

Chae K.-H., Gong I.-T., 2015, MNRAS, 451, 1719

Chae K.-H., Bernardi M., Sheth R. K., 2018, PRL, submitted (arXiv:1707.08280)

Ciotti L., Lanzoni, B., 1997, A&A, 321, 724

Clauwens B., Schaye J., Franx M. 2016, MNRAS, 462, 2832

Cole, S. et al. 2001, MNRAS, 326, 255

Conroy C., Gunn J. E., 2010, ApJ, 712, 833

Conroy C., van Dokkum P. G., 2012, ApJ, 760, 71

D'Souza R., Vegetti S., Kauffmann G. A. M., 2015, MNRAS, 454, 4027

Fischer J.-L.,. Bernardi M., Meert A., 2017, MNRAS, 467, 490

Hernquist L., 1990, ApJ, 356, 359

Huertas-Company M., Aguerri J. A. L, Bernardi M., Mei S., Sánchez Almeida J., 2011, A&A, 525, A157

Jørgensen I., Franx M., Kjaergaard P., 1995, MNRAS, 273, 1097

Kroupa P., 2001, MNRAS, 322, 231

La Barbera F., Vazdekis A., Ferreras I., et al., 2017, MNRAS, 457, 1468

Lelli F., McGaugh S. S., Schombert J. M., Pawlowski M. S. 2016, ApJ, 827, 19

Li, H., et al., 2017, ApJL, 838, 77

Lyubenova M., Martín-Navarro I., van de Ven G., Falcón-Barroso J., Galbany L., Gallazzi A., García-Benito R., González Delgado R., et al., 2016, MNRAS, 463, 3220

Maraston C., 2005, MNRAS, 362, 799

Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785

Martín-Navarro I., La Barbera F., Vazdekis A., Falcón-Barroso J., Ferreras I., 2015, MNRAS, 447, 1033

Meert A., Vikram V., Bernardi M., 2015, MNRAS, 446, 3943

Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, ApJS, 210

Newman A. B., Ellis R. S., Treu T., 2015, ApJ, 814, 26

Parikh T., Thomas D., Maraston C., Westfall K.B., Goddard D., Lian J., Meneses-Goytia S., Jones A., Bershady M., Bizyaev D., Bundy K., Roman-Lopes A., Zheng Z., MNRAS, 2018, submitted

Peletier R. F., Valentijn E. A., Jameson R. F., 1990, A&A, 233, 62

Posacki, S., Cappellari, M., Treu, T., Pellegrini, S., Ciotti, L., 2015, MNRAS, 446, 493

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

Salpeter E. E., 1955, ApJ, 121, 161

Shankar F., Bernardi M., 2009, MNRAS, 396L, 76

Schechter P. L., Pooley D., Blackburne J. A., Wambsganss J., 2014, ApJ, 793, 96

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6

Shankar F., Bernardi M., Sheth R. K., Ferrarese L., Graham A. W., Savorgnan G., Allevato V., Marconi A., Läsker R., Lapi A. 2016, MNRAS, 460, 3119

Shankar F., et al., 2017, ApJ, 840, 34

Shankar F., et al., 2018, MNRAS, in press (arXiv:1711.07986)

Tang B., Worthey G., 2017, MNRAS, 467, 647

Thanjavur K., Simard L., Bluck A. F. L., Mendel T., 2016, MNRAS, 459, 44

Thomas D., Maraston C., Johansson J., 2011, MNRAS, 412, 2183

Tortora C., Napolitano N. R., Romanowsky A. J., Jetzer Ph., Cardone V. F., Capaccioli M., 2011, MNRAS, 418, 1557

Tortora C., Romanowsky A. J., Cardone V. F., Napolitano N. R., Jetzer Ph., 2014, MNRAS, 438, L46

Tortora C., Koopmans L. V. E., Napolitano N. R., Valentijn E. A. 2018, MNRAS, 473, 2324

van der Marel R., 1994, MNRAS, 270, 271

van Dokkum, P., Conroy, C., Villaume, A., Brodie, J., Romanowsky, A., 2017, ApJ, 841, 68

Vaughan S. P., Houghton R. C. W., Davies R. L. & Zieleniewski, S. 2017, MNRAS, revised (arXiv:1612.00364)

Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, MNRAS, 404, 1639

Villaume A., Conroy C., Johnson B., Rayner J., Mann A. W., van Dokkum P., 2017, ApJS, 230, 23

Worthey G., 1994, ApJS, 95, 107

Wu H., Shao Z., Mo H. J., Xia X., Deng Z., 2005, ApJ, 622, 244

APPENDIX A: SIMPLE MODEL FOR EFFECTS OF M_*/L GRADIENTS AND A HERNQUIST PROFILE

We use Hernquist's (1990) profile to illustrate the effect of gradients in the stellar mass-to-light profile on the estimated stellar population and dynamical mass.

A1 Newtonian gravity

Suppose that the three dimensional distribution of the light follows a Hernquist (1990) profile

$$\rho_{\rm L}(r) = \frac{L/R_H^3}{2\pi\,(r/R_H)(1+r/R_H)^3} \equiv \frac{L/R_H^3}{2\pi\,r_H(1+r_H)^3}, \quad \text{(A1)}$$

where L is the total luminosity, R_H is a characteristic radius and $r_H \equiv r/R_H$. The luminosity within r is

$$L(<r) = L\,\frac{r_H^2}{(1+r_H)^2}. \quad \text{(A2)}$$

If M_*/L increases towards the center as $\Upsilon_{*0}\,(1+f/r_H)$, so the mass density profile is

$$\rho_{\rm M}(r) = \Upsilon_{*0}\,\frac{f+r_H}{r_H}\,\rho_{\rm L}(r), \quad \text{(A3)}$$

then

$$M(<r) = \Upsilon_{*0}\,L(<r)\,(1+f+2f/r_H). \quad \text{(A4)}$$

Thus $M_*(<r)/L(<r) = \Upsilon_{*0}\,(1+3f)$ at $r = R_H$ and $\Upsilon_{*0}\,(1+f)$ at $r \gg R_H$. Therefore, the total mass is $1+f$ times larger than ΥL, the value if there were no gradient.

There are two natural proxies for the mass estimated by stellar population models when gradients are ignored. One is the light-weighted mass-to-light ratio: $M_\infty = \Upsilon_{*0}L(1+f)$. However, if $M_*^{\rm SP}/L$ was estimated from the light within the projected half light radius $R_e \approx 1.8153R_{\rm H}$, then it might be more natural to use $2\times$ the mass projected within R_e, which is $\Upsilon_{*0}L(1+1.816f)$. This will generically overestimate M_∞.

We turn now to dynamical mass estimates. For these, if we assume the velocity dispersion is isotropic, then the Jeans equation implies that

$$\rho_M(r)\,\sigma^2(r) = \int_r^\infty \frac{{\rm d}r}{r}\,\rho_M(r)\,\frac{GM(<r)}{r}. \quad \text{(A5)}$$

The integral can be done analytically, and yields

$$\rho_{\rm M}(r)\frac{\sigma^2(r)}{GM_\infty/R_{\rm H}} = \int_r^\infty \frac{{\rm d}r}{r}\,\rho_M(r)\,\frac{M(<r)/M_\infty}{r/R_{\rm H}}$$
$$= -\frac{\rho_{\rm M}(r)/(1+f)}{12(f+r_{\rm H})(r_{\rm H}+1)}\Big(f^2\big(5r_{\rm H}(5r_{\rm H}A+12)-12\big)$$
$$- 12\big(f\,(25f-14)+1\big)\,r_{\rm H}^2\,(r_{\rm H}+1)^4\,\ln(1/r_{\rm H}+1)$$
$$- 2fr_{\rm H}(7r_{\rm H}A+18)+r_{\rm H}^2A\Big), \quad \text{(A6)}$$
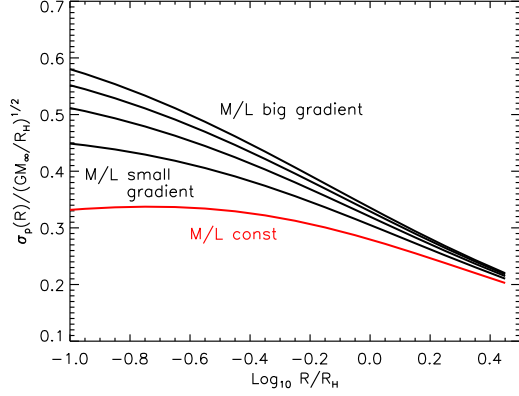
Figure A1. Effect of gradients in the mass-to-light ratio on the light-weighted projected velocity dispersion if the light distribution follows a Hernquist profile with scale radius R_H and the mass-to-light ratio increases towards the center as $1 + fR_H/r$. Curves, from bottom to top, show σ_p for $f = (0, 0.25, 0.5, 0.75, 1)$: gradients steepen the σ_p profile. When normalized to the dispersion observed on scale R, the stellar dynamical mass estimate is proportional to the inverse of the square of the quantity shown on the y-axis.

where $M_\infty \equiv \Upsilon_{*0}L(1 + f)$ and $A = 2r_H(3r_H(2r_H + 7) + 26) + 25$. If we add a central black hole of mass $M_\bullet$, then

$$\frac{M_\bullet}{M_\infty} \frac{\rho_M(r)}{6r_H(f + r_H)/(1 + r_H)}$$

$$\left(2(3 - 5f)(r_H + 1)^3 r_H^2 \ln(1 + 1/r_H) \right.$$

$$\left. + f(5r_H(2r_H + 1)B + 2) - 3r_H(2r_H + 1)B\right), \quad (A7)$$

where $B = 6r_H(r_H + 1) - 1$, must be added to the expression above.

The light-weighted projected dispersion at projected radius R is given by equation (9) in the main text. Models with a constant mass to light ratio have $\rho_M(r) = \Upsilon_{*0}\rho_L(r)$, so the shape of the light profile determines the shape of σ_p. The dynamical mass is then estimated by finding that Υ_{*0} which gives the observed value of σ_p: hence, the value of Υ_{*0} from stellar population modeling is irrelevant.

However, when there are gradients, then stellar population modelling is *required* to get ρ_M/ρ_L, so the dynamical mass estimate depends on the SP modeling (upto an overall normalization factor). I.e., the SP modelling is needed to estimate the shape, after which the normalization is given, as before, by fitting to data. For example, suppose we normalize using the observed σ_p on some scale R_{obs}. Then,

$$M_{dyn} = k(R_{obs}, f) R_H \sigma_{obs}^2/G, \quad (A8)$$

where $k(R_{obs}, f) = (GM_\infty/R_H)/\sigma_p^2(R_{obs})$ is the inverse of the (square of the) value of the curve with the relevant value of f shown in Figure A1.

If we assume there are no gradients when, in fact, there are, then we will use the red curve to match the observations instead of the correct one. This will have two consequences. First, the resulting stellar mass estimate will be too high by a factor of $k(R_{obs}, f)/k(R_{obs}, 0)$. This can be a significant overestimate if $R_{obs} < R_H$. For example, if $R_{obs} = R_e/8 \approx 0.2R_H$ and $f = 0.5$ then this factor is $(0.475/0.33)^2 \approx 2$.

The second effect has to do with the fact that the curves with larger f are steeper. In practice, observed $\sigma(R)$ profiles are shallower than the ones predicted by a Jeans equation analysis of the light (+ constant mass to light ratio). The difference is attributed to dark matter. If we were to slide the red curve upward so it matched a black one at some $R = R_{obs}$, then it would lie above the black beyond beyond R_{obs}. This would reduce the amount of dark matter which is required on scales beyond R_{obs}. As a result, dynamical mass estimates which ignore gradients generically bias one towards underestimating the amount of dark matter on scales where the light is seen. This will lead one to underestimate the amount of adiabatic contraction of the dark matter that the baryons in the central regions caused.

A2   Modified Newtonian dynamics

For MOND, if $a_N \equiv GM(< r)/r^2$ then the acceleration becomes $a_N(r) f(a_N/a_0)$. There are a number of parametrizations of this scaling, but to illustrate our results, we will use $f_{Bek}(x) = 1 + 1/\sqrt{x}$. For a Hernquist light profile,

$$a_N(r) = \frac{G \Upsilon_{*0}L/R_H^2}{(1 + r/R_H)^2} \equiv \frac{a_H}{(1 + r_H)^2}, \quad (A9)$$

so $f_{Bek} = 1 + (1 + r_H)\sqrt{a_0/a_H}$. Thus, the three dimensional velocity dispersion satisfies

$$\rho_L(r)\sigma^2(r) = R_H a_H \frac{L/R_H^3}{2\pi} \int_{r_H}^\infty \frac{dr}{r} \left[\frac{1}{(1 + r)^5} + \frac{\sqrt{a_0/a_H}}{(1 + r)^4}\right]. \quad (A10)$$

The term proportional to a_0/a_H is a correction to the Newtonian expression given earlier (equation A6 with $f = 0$); it only matters when $a_H \ll a_0$. The integral for this correction factor can be done analytically, yielding $(G \Upsilon_{*0}L/R_H)(L/R_H^3)/2\pi$ times

$$\sqrt{\frac{a_0}{a_H}}\left(\ln(1/r_H + 1) - \frac{3r_H(2r_H + 5) + 11}{6(1 + r_H)^3}\right).$$

The presence of $a_H \propto \Upsilon_*$ under the square root sign means that $\sigma(r)$ is not linearly proportional to Υ_*. Therefore, in contrast to the Newtonian case, as one changes the normalization factor Υ_*, the shape of $\sigma(r)$ also changes. This is because, in the Newtonian case, one treats the discrepancy between what the light predicts and what is observed by add an independent term, $M_{dm}(< r)$ to the total mass. In MOND, what matters is the ratio a_0/a_H, so the amount to be 'added' depends on what is present.

With gradients,

$$a_N(r) \equiv a_H \frac{(1 + f + 2f/r_H)}{(1 + r_H)^2}, \quad (A11)$$

so $f_{Bek} = 1 + \sqrt{a_0/a_H}(1 + r_H)/\sqrt{1 + f + 2f/r_H}$. This complicates the integral for $\sigma(r)$, but not the fact that the shape of $\sigma(r)$ depends on Υ_*.

For example, when $f = 1$ then $\rho_M(r) = \rho_L(r)(1 + r_H)/r_H$ so $M(< r) = M_\infty r_H/(1 + r_H)$ with $M_\infty = 2\Upsilon_{*0}L$,

and

$$\frac{\sigma^2(r)}{GM_\infty/R_\mathrm{H}} = 6\,r_\mathrm{H}^2\,(1+r_\mathrm{H})^2\,\ln(1+1/r_\mathrm{H})$$

$$-\,\frac{(1+2r_\mathrm{H})\left[6r_\mathrm{H}(1+r_\mathrm{H})-1\right]}{2} \qquad\text{(A12)}$$

$$+\,\sqrt{\frac{a_0}{2a_\mathrm{H}}}\,\frac{1-2r_\mathrm{H}\left[4r_\mathrm{H}\left(2r_\mathrm{H}-2D+3\right)-8D+3\right]}{3/2D},$$

where $D = \sqrt{r_\mathrm{H}(1+r_\mathrm{H})}$.