

# A catalogue of 2D photometric decompositions in the SDSS-DR7 spectroscopic main galaxy sample: preferred models and systematics

Alan Meert,<sup>★</sup> Vinu Vikram and Mariangela Bernardi

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

Accepted 2014 October 31. Received 2014 October 31; in original form 2014 June 15

## ABSTRACT

We present a catalogue of 2D, point spread function-corrected de Vaucouleurs, Sérsic, de Vaucouleurs+Exponential, and Sérsic+Exponential fits of  $\sim 7 \times 10^5$  spectroscopically selected galaxies drawn from the Sloan Digital Sky Survey (SDSS) Data Release 7. Fits are performed for the SDSS *r* band utilizing the fitting routine GALFIT and analysis pipeline PYMORPH. We compare these fits to prior catalogues. Fits are analysed using a physically motivated flagging system. The flags suggest that more than 90 per cent of two-component fits can be used for analysis. We show that the fits follow the expected behaviour for early and late galaxy types. The catalogues provide a robust set of structural and photometric parameters for future galaxy studies. We show that some biases remain in the measurements, e.g. the presence of bars significantly affect the bulge measurements although the bulge ellipticity may be used to separate barred and non-barred galaxies, and about 15 per cent of bulges of two-component fits are also affected by resolution. The catalogues are available in electronic format. We also provide an interface for generating postage stamp images of the 2D model and residual as well as the 1D profile. These images can be generated for a user-uploaded list of galaxies on demand.

**Key words:** galaxies: evolution – galaxies: fundamental parameters – galaxies: structure.

## 1 INTRODUCTION

The study of the structural components of galaxies has contributed substantially to the understanding of the formation and evolution of galaxies. The discovery of many scaling relations including the Faber–Jackson (Faber & Jackson 1976), Kormendy (Kormendy 1977), Tully–Fisher (Tully & Fisher 1977), the Fundamental Plane (Djorgovski & Davis 1987) and the morphology–density relation (Dressler 1980) refined models of galaxy formation and evolution. In addition, the structural components of galaxies in the local Universe trace morphological galaxy type and many other galaxy parameters related to both assembly and evolution of galaxies: colour, metallicity, gas fraction, central velocity dispersion (e.g. Bernardi et al. 2003a,b,c; Kauffmann et al. 2003; Tremonti et al. 2004). Properties may also trace halo size and galaxy environment and place constraints on a cold dark matter cosmology (e.g. Blanton et al. 2005; Bernardi 2009; Shankar et al. 2010a,b; Kravtsov, Vikhlinin & Meshcheryakov 2014). However, careful estimation of structural parameters for large numbers of galaxies is required to test different formation and evolution models.

The Sloan Digital Sky Survey (SDSS; York et al. 2000; Abazajian et al. 2009) has already provided a sample of many millions

of nearby galaxies. Future surveys like the Dark Energy Survey (The Dark Energy Survey Collaboration 2005) and Large Synoptic Survey Telescope (LSST; LSST Science Collaboration 2009) will produce larger data sets, both increasing the number and quality of galaxies available for analysis. At the same time, the growth of computing power makes it possible to analyse these data sets at a reasonable rate, making it possible to perform time-intensive analysis, like galaxy decompositions, on large data sets.

There has been much recent work on improving photometric decomposition of galaxies (e.g. Gadotti 2009; Simard et al. 2011; Kelvin et al. 2012; Lackner & Gunn 2012; Häußler et al. 2013). However, the accuracy of such fits is often questioned, particularly when multiple components are fitted (i.e. bulge+disc+bar etc.) and when fits are automated without individual inspection.

This paper presents a catalogue of 2D, point spread function (PSF)-corrected de Vaucouleurs, Sérsic, de Vaucouleurs+Exponential, and Sérsic+Exponential fits of  $\sim 7 \times 10^5$  spectroscopically selected galaxies drawn from the SDSS. Fits are presented for the SDSS *r* band utilizing the fitting routine GALFIT (Peng et al. 2002) and analysis pipeline PYMORPH (Vikram et al. 2010). This catalogue is one of the largest galaxy samples for which structural decompositions have been performed on SDSS galaxies.

The simulations presented in Meert, Vikram & Bernardi (2013, hereafter M13) are used as a benchmark for these fits. M13 used

\* E-mail: ameert@sas.upenn.edu

simulated galaxies drawn from galaxies in this work to test the accuracy of the fitting process. It established uncertainties on fitting parameters and showed that the choice of cutout size and background estimation were appropriate for the galaxies in this sample.

2D decompositions of SDSS galaxies that overlap with our catalogue have also been carried out by other groups. Simard et al. (2011, hereafter S11) presented decompositions of the photometric sample of SDSS ( $\approx 1.4$  million galaxies). Kelvin et al. (2012) limited to fitting a single Sérsic model to a subset of our catalogue. Lackner & Gunn (2012, hereafter LG12) fitted several models to galaxies at lower redshifts. We compare to S11, LG12 and other works in Sections 6 and 7.

This paper focuses on the spectroscopic sample, a subsample of SDSS and S11. We also present a method for identifying good and bad fits different from the statistical approach of S11 and the statistical and qualitative combination of LG12. Using a combination of comparisons between this work, S11, and LG12 (presented in Sections 6 and 7), we show that our catalogue improves on previous fits in many respects.

Several studies have already utilized this catalogue. Shankar et al. (2013) tested semi-analytical modelling of hierarchical formation. Huertas-Company et al. (2013) examined environmental effects on the size of galaxies. Bernardi et al. (2013) analysed the uncertainty in the bright end of the mass and luminosity functions (LF). Bernardi et al. (2014) also examined the biases automated decompositions imposed on the size-luminosity relation. Finally, Kravtsov et al. (2014) performed detailed fits of approximately 10 brightest cluster galaxies (BCGs) and found that their measurements agree more with our measurements than with the measurements of S11.

The paper is organized as follows: Section 2 describes the selection of the data, Section 3 briefly describes the PYMORPH fitting routine we used and the specifics of our fitting procedure including our choice of cutout size, background fitting, and neighbour fitting. Section 4 describes the flagging system used to identify poor fits and interpret fits as either bulge, disc, or two-component galaxies. Section 5 describes internal comparisons and consistency checks among the models we fit. Section 6 describes comparisons to measurements made by other groups including SDSS (Abazajian et al. 2009), S11, LG12, and Mendel et al. (2014). Section 7 describes the comparisons incorporating morphological information from previous visual or automated classifications. Section 8 presents the catalogue and describes how to use it. In addition, we describe the web page content associated with the catalogue. Finally, Section 9 concludes the paper with a summary of results and final remarks. The fits discussed in this paper and further recommendations for their use are available in electronic format as a public release.

## 2 THE DATA

### 2.1 SDSS CasJobs data

The data used in this analysis were drawn from the spectroscopic sample of the Legacy area<sup>1</sup> of the Sloan Digital Sky Survey Data Release 7 (DR7; Abazajian et al. 2009). The spectroscopic sample provides a well-established sample with well-defined and tested selection criteria. The criteria are presented in Strauss et al. (2002).

Galaxies listed in both the PhotoObj and SpecObj tables of the CasJobs DR7 data base that satisfy three main selection criteria were

<sup>1</sup> A list of fields in the Legacy survey is provided at <http://www.sdss.org/dr7/coverage/allrunsdr7db.par>

selected. Those were: (1) the extinction-corrected  $r$ -band Petrosian magnitude between magnitude 14 and 17.77; (2) the Photo pipeline identified the object as a galaxy (Type = 3); and (3) the spectrum was also identified as a galaxy (SpecClass = 2). We place a limit at the faint end of 17.77 mag in the  $r$  band because this is the lower limit for completeness of the SDSS Spectroscopic Survey (Strauss et al. 2002). The limit of 14 mag at the bright end is used to exclude large, nearby galaxies that are typically either too well resolved to be fitted with a standard smooth light profile or shredded into multiple smaller objects in the SDSS catalogue. These brightest galaxies may also be segmented over multiple images or so large that it prevents robust estimation of the background flux. Such galaxies require additional work to properly combine neighbouring images (see Blanton et al. 2011, for details).

The query used, omitting the names of selected data columns, is `SELECT p.objid ...FROM photoobj as p ...LEFT OUTER JOIN SpecObj as s on p.objID = s.BestObjID, segment g, field f, chunk c WHERE g.segmentID = f.segmentID and f.fieldID = p.fieldID and c.chunkID = g.chunkID and (p.petroMag_r - p.extinction_r) between 14.0 and 17.77 and p.type = 3 and s.specclass = 2 ORDER BY p.objid`, which produces 676 010 matches.

We apply additional cuts similar to Shen et al. (2003) and S11 to the data prior to fitting. We remove all galaxies with redshift  $< 0.005$  (1647 galaxies). These galaxies have large apparent sizes and resolved structure that make decomposition difficult. We also remove 20 galaxies with redshift  $> 1.0$ . Visual inspection reveals that these galaxies likely represent catastrophic failures in the redshift code.

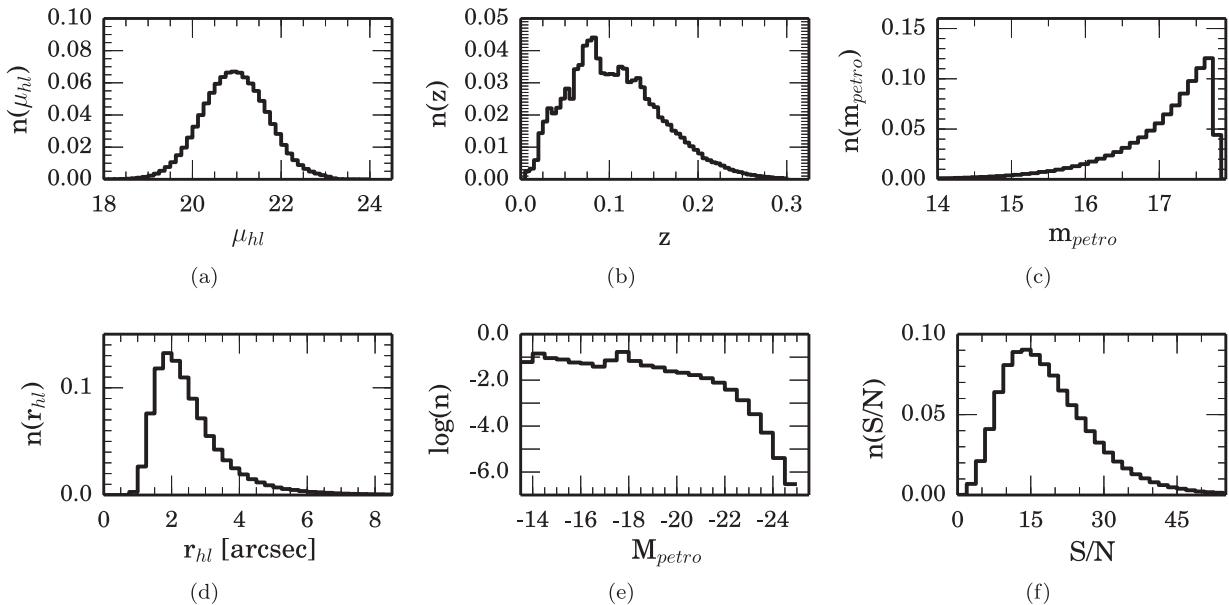
Galaxies with saturated pixels as indicated by the Photo flags are also removed from the sample (3207 galaxies). In addition, as discussed in Strauss et al. (2002), we apply a surface brightness cut on the mean surface brightness within the Petrosian half-light radius of  $\mu_{50,(r)} < 23.0$  mag, where

$$\mu_{50,(r)} = m_{\text{petro}, r} + 2.5 \log(2\pi r_{\text{petro}, 50}^2) \quad (1)$$

because there is incomplete spectroscopic target selection at brightnesses below this threshold. After applying all the cuts, 5529 galaxies (approximately 0.8 per cent of the sample) are removed and a sample of 670 722 galaxies remains. We identify this as our sample used throughout the paper. We consider the cuts described here in the completeness given in Section 8.

Fig. 1 shows the surface brightness distribution, redshift distribution, extinction-corrected  $r$ -band Petrosian magnitude,  $r$ -band Petrosian half-light radius,  $V_{\max}$ -weighted LF, and signal-to-noise ratio (S/N) distribution of the sample used in this paper. We define the S/N as the mean pixel flux within the half-light radius divided by the noise associated with that pixel, or

$$\begin{aligned} \left(\frac{S}{N}\right) &\equiv \frac{I_{\mu_{50}}}{N_{\text{avg}}} \\ N_{\text{avg}} &\equiv \sqrt{\frac{I_{\mu_{50}} + I_{\text{sky}}}{\text{gain}} + \text{dark variance}} \\ I_{\mu_{50}} &= 10^{-0.4(\mu_{50} - zp)} \times \text{plate scale}^{-2}, \end{aligned} \quad (2)$$



**Figure 1.** (a) The surface brightness distribution, (b) redshift distribution, (c) extinction-corrected  $r$ -band Petrosian magnitude, (d)  $r$ -band Petrosian half-light radius, (e)  $V_{\text{max}}$ -weighted LF, and (f) S/N distribution using the measurement of sky provided by the SDSS photometric pipeline of the sample used in this paper drawn from the DR7 SDSS spectroscopic galaxy sample. Bin counts are normalized so that each distribution integrates to 1.

where  $I_{\mu_{50}}$  is the source DN ('data numbers' or, equivalently, counts)<sup>2</sup> of the average surface brightness defined in equation (1). The zero-point,  $z_{\text{p}}$ , is calculated from the SDSS zero-point, extinction, and airmass terms associated with each image. The platescale is used to convert the surface brightness from counts per square arcsecond to counts per pixel.  $N_{\text{avg}}$  is the noise in a pixel using the SDSS background measurement as an estimate of the background flux and the average flux per pixel inside the Petrosian half-light radius as the galaxy flux.

We collect all of the identifying data as well as photometric measurements obtained from the SDSS CasJobs server into the table named the CasJobs Table distributed with the catalogue. For convenience, ID numbers were assigned to all galaxies contained in the catalogue to be used in place of the SDSS ObjID. These ID numbers are referred to as GALNUM and used throughout the available data products presented here. The GALNUM are used throughout the analysis as the unique identifier for each galaxy.

Using the information obtained from the SDSS CasJobs server, we download all necessary fpC images and PsField files from the SDSS. The PsField files provide the 2D reconstruction of the point spread function (PSF) necessary for fitting. The fpC images contain the galaxy and surrounding neighbourhood. We extract our own postage stamps from the fpC images for fitting rather than using the stamps provided through the atlas images produced by SDSS. This process is described in more detail in Section 3.1.

## 2.2 Additional parameters

In addition to the data provided by CasJobs, we collect parameters from a number of other studies. The Morphology catalogue of Huertas-Company et al. (2011, hereafter H2011) is an automated morphological classification that used a Bayesian SVM algorithm

<sup>2</sup> The counts are related to the number of photoelectrons collected by the detector through the gain of the detector amplifier. This distinction is important since the photoelectrons obey Poisson statistics.

to classify all galaxies in the spectroscopic sample based on data available as part of the SDSS DR7. Objects are matched to our catalogue based on the SDSS DR7 objID. The morphological parameters from the matching catalogue (H2011) for our sample is described in the electronic catalogue.

We also calculate  $K$ -corrections, distance modulus, angular diameter distance, and  $V_{\text{Max}}$  correction for each galaxy.  $K$ -corrections are calculated using version 4.2 of the  $K$ -correction code KCORRECT described in Blanton & Roweis (2007). To calculate the  $K$ -correction, the SDSS modelmag and modelmag\_err are used and data for all band passes ( $u, g, r, i, z$ ) are provided to the program. These terms are collected and provided with the data. We assume a cosmology with  $(H_0, \Omega_\Lambda, \Omega_m, \Omega_k) = (70 \text{ km s}^{-1}\text{Mpc}^{-1}, 0.7, 0.3, 0.0)$  when necessary.

## 3 THE FITTING PROCESS

In this section we describe the fitting process. Section 3.1 describes our choice of cutout size and the data used for fitting. Section 3.2 describes the profiles used during fitting. Section 3.3 briefly describes the PYMORPH pipeline used to fit the catalogue. Finally, Section 3.4 describes our masking and neighbour identification and focuses on how we verified the masking and simultaneous fitting in crowded fields where fitting is complicated and potentially biased by neighbouring objects.

### 3.1 Pre-processing of SDSS images

FpC images and psField files are the primary data used in the fitting procedure. The data were downloaded from the SDSS Data Archive Server. We used postage stamp images of galaxies during fitting. Postage stamp images of each source were extracted from the fpC image such that the stamp was 40 Petrosian half-light radii on a side ( $20 \times \text{petroR50\_r}$  from the centre of the image to the edge) and centred on the target source. The decision to cut at 20 half-light radii is justified in M13 based upon simulations and provides a large

number of background pixels (about 30 000–40 000 pixels for an average-sized image).

In addition, a minimum size of 80 pixels on each side was set to ensure that enough pixels were retained to properly determine the background. In reality, with such a large postage stamp size, this minimum of 80 pixels is rarely required.

A smaller cutout size could potentially be used when fitting the galaxy and would reduce the time needed to fit each galaxy. Section 3.5 and fig. 12 of M13 show that the effect of further reducing the cutout size is insignificant. The main driver of postage stamp size selection is to ensure that there are a sufficient number of background pixels for sky estimation. Section 3.6 of M13 shows that the estimate of background sky brightness is accurate to  $\sim 0.1$  per cent with a slight bias towards underestimating the background level using this stamp size. This sky bias does not noticeably bias the other fitted parameters.

We also extract a PSF from the PsField files using the READATLASIMAGES-V5\_4\_11 program distributed on the SDSS website.<sup>3</sup> The PSF provided by SDSS using the READATLASIMAGES-V5\_4\_11 program has a standard image size of 51 pixels on each side.

In addition, prior to fitting we remove the 1000 DN soft bias from the images and PSF. We create sigma images from the SDSS image cutout following the standard deviation calculation<sup>4</sup>

$$W_{i,j} = \sqrt{\frac{F_{i,j}}{\text{gain}} + \text{dark variance}}, \quad (3)$$

where  $W_{i,j}$  is the pixel sigma in DN,  $F_{i,j}$  is the pixel flux (again in DN), *gain* for the image as specified in SDSS CasJobs and used to account for the fact that the photoelectrons (rather than the DN) obey Poisson statistics, and ‘dark variance’ is the term used by SDSS to describe the contribution of the read noise and dark current to the image noise. Finally, we normalized the postage stamp and sigma images to a 1 s exposure prior to fitting.

### 3.2 The fitted profiles

The Sérsic model has been used extensively in galaxy studies since first being proposed by Sérsic (1963):

$$I(r) = I_e \exp\left(-b_n \left[\left(\frac{r}{R_e}\right)^{1/n} - 1\right]\right) \\ b_n = 1.9992n - 0.3271, \quad (4)$$

where Sérsic index ( $n$ ), half-light radius ( $R_e$ ), and surface brightness at  $R_e$  ( $I_e$ ) are the parameters used to define the profile.  $b_n$  uses the approximation from Capaccioli (1989) which is valid for  $0.5 < n < 10$ . When  $n = 4$ , the Sérsic model reduces to the de Vaucouleurs model (de Vaucouleurs 1948). For the fitting presented here, the Sérsic index is restricted to values less than or equal to 8.0. Higher values of the Sérsic index are not allowed. Although such galaxies may exist in nature (e.g. Graham & Driver 2007; Kormendy et al. 2009), we find that higher values of the Sérsic index are often associated with fitting problems in this sample.

For two-component models, a de Vaucouleurs or Sérsic model is used to model the light in the central part of the galaxy (often

associated with the bulge) and an exponential disc is added to model the portion of the galaxy farther from the centre (often associated with the rotational disc). The exponential model is defined by the scale radius ( $R_d$ ) and central surface brightness ( $I_d$ ). The disc is modelled using the function

$$I_{\text{Exp}}(r) = I_d \exp\left(\frac{-r}{R_d}\right). \quad (5)$$

The profiles defined in equations (4) and (5) are 1D profiles. The 1D profiles are used to generate 2D models by also fitting a centre on the image, position angle ( $\phi$ ), and axis ratio ( $b/a$ ) to each component.

Equations (4) and (5) are often interpreted as representing a bulge and a disc, respectively. However, fitting two-component models to galaxies does not guarantee that the two components measured are truly present. Many early-type galaxies show no signs of disc-like structures. Similarly, many late-type galaxies show little or no sign of a bulge in the central part of the galaxy. Also, when fitting multiple components, a significant second component may only indicate substantial departure from a single-component profile rather than the presence of a physically meaningful second component. For example, Gonzalez, Zabludoff & Zaritsky (2005), Donzelli, Muriel & Madrid (2011), and Huang et al. (2013) fit multiple components to Ellipticals (Ell) and BCGs without necessarily claiming the existence of additional physically distinct components. Also, the presence of a bar will affect fitting, changing the ellipticity and Sérsic index of the bulge component in the two-component models. We examine the effects of a bar component in Section 7.3.

We reserve judgement on the interpretation of the components until after discussion of the flagging system in Section 4. We merely comment here that there are many cases in which the components should not be interpreted as a physically meaningful bulge and disc.

We also note that the Sérsic model intended to represent the central region of two-component galaxies has broad wings and can have a large fraction of its total light at large radii. This is especially true at higher values of the Sérsic index and often causes the Sérsic model to dominate any exponential disc at large radii. We account for cases where the Sérsic component dominates at large radii during the flagging procedure described in Section 4 and Appendix A by focusing on the relative brightness of the two components only out to the radius where 90 per cent of the total light is enclosed. However, the reader should be aware that the Sérsic component may often dominate at radii beyond this point.

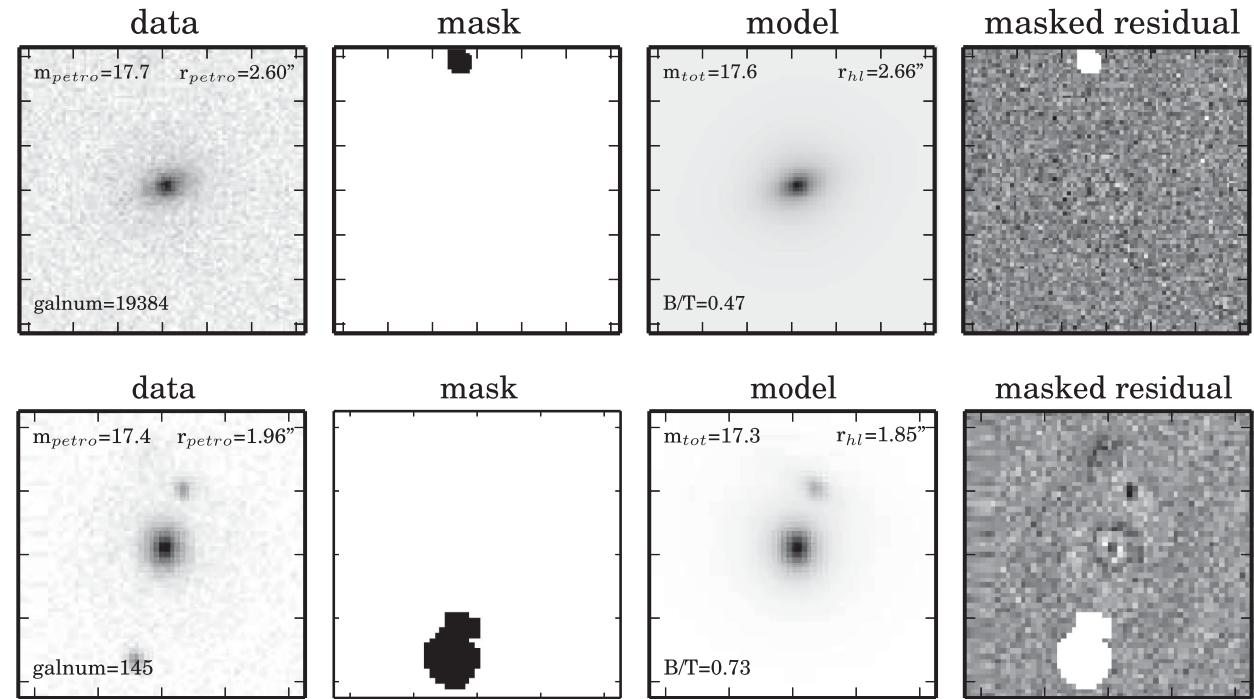
### 3.3 The fitting pipeline

We performed both one- and two-component fits to the sample described in Section 2 using PYMORPH (Vikram et al. 2010). PYMORPH is a PYTHON based automated software pipeline built on SExtractor (Bertin & Arnouts 1996) and the 2D fitting routine GALFIT (Peng et al. 2002). Both PYMORPH and GALFIT have been extensively tested (see Peng et al. 2002; Vikram et al. 2010, for more tests of GALFIT).

M13 also tested fitting in SDSS conditions of S/N, plate scale, and seeing using simulated data. The test presented in M13 showed that estimates of the sky are expected to be accurate at the level of 0.05 per cent, and that fitted parameters are not significantly affected by sky errors at this level. In addition, fitting a Sérsic profile to a galaxy with two components will cause a bias in size and magnitude at the 5–10 per cent level. When the appropriate profile (i.e. one versus two component) is chosen, magnitude and half-light radius are expected to be accurate at the 5 per cent level (using  $1\sigma$  error bars). The flagging procedure presented in Section 4 is intended (among other

<sup>3</sup> The use of READATLASIMAGES-V5\_4\_11 for PSF extraction is described at [http://www.sdss.org/dr7/products/images/read\\_psf.html](http://www.sdss.org/dr7/products/images/read_psf.html)

<sup>4</sup> See the SDSS DR7 online documentation at <http://www.sdss.org/dr7/algorithms/fluxcal.html> for further discussion.



**Figure 2.** Two example fits output by the PYMORPH pipeline. Ser-Exp fits are shown. Each column contains the input image (far left), mask (centre left), fitted galaxy (centre right), and the residual image (far right). The first row shows a late-type galaxy. Some spiral structure is clearly evident in the residual. The second row shows an early-type galaxy with a smoother profile and a fitted neighbour.

goals) to reliably distinguish between one- and two-component fits, reducing any fitting bias resulting from fitting a one-component profile to a two-component galaxy (see M13 for more information on the simulations and results).

We use the PYMORPH pipeline to fit single Sérsic and single de Vaucouleurs fits (Ser and deV) as well as Sérsic + exponential disc (Ser-Exp) and de Vaucouleurs + exponential disc (deV-Exp) models to each galaxy in the SDSS *r* band. In addition to photometric decomposition, PYMORPH also supports measurement of several non-parametric structural parameters. The final catalogue reports the automagnitude (a Kron-like magnitude) and half-light radius measured by SExtractor. We also directly measure a total half-light radius and total axis ratio for all models using the image of the fitted model. These measurements allow direct comparison of the one- and two-component model half-light radii (See Section 5.3 and Fig. 12).

Two example fits and the corresponding residuals are presented in Fig. 2. Each row of the figure represents a different galaxy with the data, mask, fitted model, and residual presented in columns 1, 2, 3, and 4, respectively.

### 3.4 Masking and neighbour identification

In this section, we focus on the effects of neighbours on the fitting process. We discuss our choices for masking and simultaneously fitting neighbours. We justify these choices and describe how galaxies with fitting problems caused by neighbours are identified in the catalogue.

Fig. 3 summarizes the process we used to identify neighbour sources, decide whether to mask or simultaneously fit the neighbour, and verify that the choice is appropriate. This process is described throughout Section 3.4.

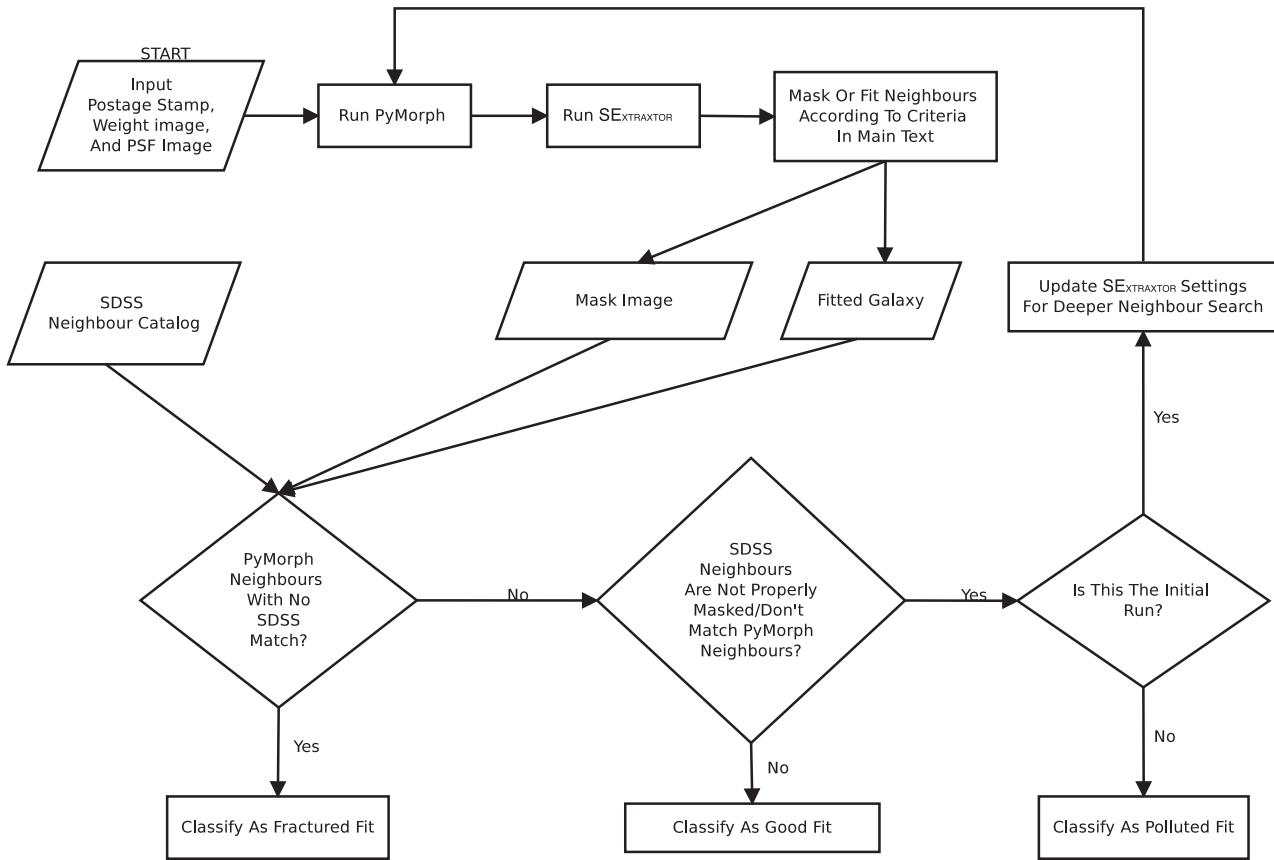
#### 3.4.1 The masking and neighbour identification process

PYMORPH performs image masking using the SExtractor program (Bertin & Arnouts 1996). Any sources identified by SExtractor are masked out for fitting unless the extra source is too close to the target galaxy to be properly masked. (The threshold for this case is set by the user.) Neighbours that are not simultaneously fit are masked according to the SExtractor segmentation image. Simultaneous fitting of the target source and neighbour source is performed in cases where the proximity of the sources makes masking ineffective. In this case, the extra source is simultaneously fitted with a single Sérsic profile while fitting the target galaxy.

PYMORPH detects all neighbouring sources in the fitted frame using SExtractor with settings BACK\_SIZE=64, ANALYSIS\_THRESHOLD=1.5 and DETECT\_MINAREA=6. We tested several values of these settings but found no effect on the final fits. Object detection is largely unaffected by varying the parameters.

After searching the frame with SExtractor, the SExtractor catalogue is used to decide whether to mask or fit neighbouring sources according to requirements set on the minimum fractional size of a neighbour relative to the target and the maximum separation between the neighbour and target in multiples of the sum of the half-light radii. When both conditions are satisfied simultaneously by a specific neighbour-target combination, the neighbour is simultaneously fitted with a Ser profile. Otherwise, the area occupied by the source as defined by the SExtractor segmentation image is masked out during fitting.

For the fits presented in this paper, neighbouring sources are simultaneously fit when the separation of the target and neighbour source is less than three times the sum of the two objects' semimajor half-light radii as measured by SExtractor. The neighbour source must be larger than 20 per cent of the area defined by the SExtractor radius of the target galaxy (i.e.  $r_{\text{SExtractor}, \text{neighbour}}^2 / r_{\text{SExtractor}, \text{target}}^2 > 0.2$ ).



**Figure 3.** A flowchart showing the process of masking and simultaneous fitting described in Section 3.4. The flowchart also includes the decision process used to identify ‘fractured’ or ‘polluted’ fits described in Section 3.4.2. This entire process is carried out for each galaxy individually.

We tested several masking methods, using circular and elliptical masks with sizes 2, 4, and 6 times the size of the neighbour sources. These masking configurations provided no reduction in the scatter of the recovered parameters when tested on simulations. Since no improvement was evident, we used the default configuration for analysis (i.e. masks are drawn according to the SExtractor segmentation image produced during fitting). While this likely leaves unmasked light from neighbouring sources, it has no effect on our determination of sky due to the large number of sky pixels. Different masking and handling of neighbouring sources was carried out by S11 as well as in Barden et al. (2012). Our masking generally covers a smaller area around each source when compared to these two works. The insensitivity of the sky estimation to this choice of masking is likely the result of the large cutout sizes used in our fitting, which reduces sensitivity of sky estimation to stray light from masked neighbours.

Fig. 4 shows some examples of galaxies fitted with Ser profiles. Each galaxy has masked or simultaneously fit neighbours. The decision of whether to mask or fit the neighbour was carried out as described in this section. Each row shows the input image, mask used during fitting, the final fitted image, and the residual of the fit. The half-light radius of the target galaxy and the neighbour jointly determine whether the neighbour is masked or simultaneously fit.

#### 3.4.2 Verifying masking and deblending conditions

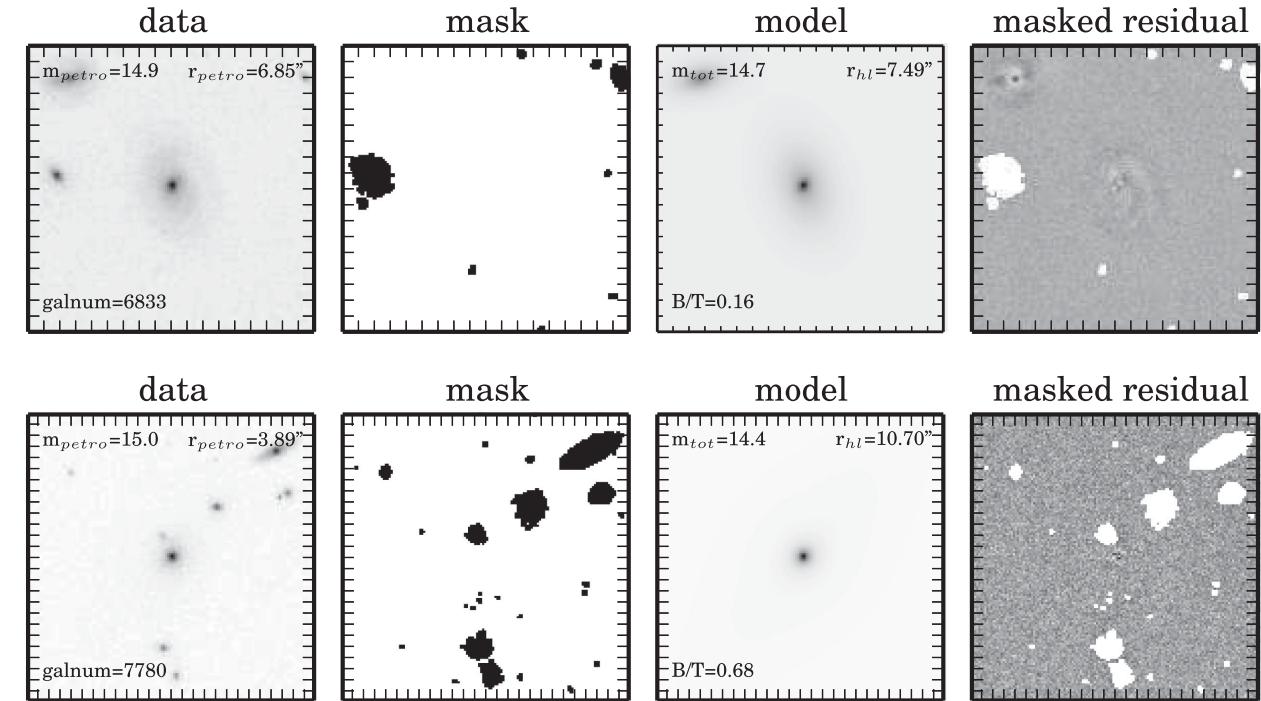
After fitting, we verify that neighbours are properly identified and masked or simultaneously fitted by comparing the catalogue

containing each targeted galaxy and any fitted neighbours against the five nearest Primary Photometric targets from the DR7 CasJobs PhotoPrimary table. We require that the CasJobs objects have Petrosian  $r$ -band magnitude brighter than 20 and be within 9 arcsec of the original galaxy fitted by PYMORPH. Objects farther or dimmer than these cuts are unlikely to cause fitting problems because they are more than  $\sim 5$  times dimmer and over 4 half-light radii away from the average galaxy in our catalogue.

There are  $\approx 78\,000$  potential neighbours in SDSS and  $\approx 100\,000$  simultaneously fitted PYMORPH neighbours (about 79 000 galaxies have one neighbour, about 8000 galaxies have two neighbours, and about 1500 galaxies have three or more simultaneously fitted neighbours, adding up to produce the  $\approx 100\,000$  neighbours).

Since the average number of PYMORPH neighbours per galaxy is small, there should be good agreement between the PYMORPH neighbours and the brightest neighbours found in SDSS. Therefore, we match our PYMORPH neighbours to SDSS neighbours by cross-matching the two catalogues with a search radius equal to the  $r$ -band Petrosian half-light radius of the SDSS neighbour (which is usually a few arcseconds). If a PYMORPH neighbour and SDSS neighbour are separated by less than this radius, we consider the two matched.

About 40 000 of the 100 000 fitted PYMORPH neighbours match SDSS objects. The remaining 60 000 unmatched PYMORPH neighbours and  $\approx 38\,000$  unmatched SDSS neighbours may cause problems during fitting. Following the procedure described in this section, we show that the number of galaxies with problems due to neighbours is only  $\sim 4000$  galaxies in the final sample.



**Figure 4.** Examples of two galaxies with neighbours fitted simultaneously (top row) or masked out (bottom row). From left to right, each row shows the input image, mask created after initial analysis of the image, fitted image, and residual. Regions that are masked out during fitting are shown in the mask image as black areas.

The two groups, unmatched SDSS neighbours and unmatched PYMORPH neighbours, possibly represent two different failures of the pipeline. Unmatched PYMORPH neighbours may be spurious neighbours fitted by PYMORPH after SExtractor improperly separates the target galaxy into several smaller fragments. Unmatched SDSS objects may be nearby neighbours that fail to be deblended from the target galaxy by SExtractor.

The vast majority of unmatched PYMORPH neighbours do not negatively affect fitting. These are objects that are dimmer than 20 mag in the  $r$  band or farther than 9 arcsec from the target galaxy, and do not have SDSS matches as a result of our original SDSS neighbour catalogue selection. Only a small number of unmatched PYMORPH neighbours are close to the target galaxy. In the largest and nearest galaxies, detailed galaxy structure (e.g. spiral arms or dust lanes) triggers improper separation and deblending. We call these cases ‘fractured’ galaxies and identify them by searching for the fitted galaxies where the  $r$ -band Petrosian magnitude is at least 0.5 mag brighter than the fitted magnitude and the unmatched fitted neighbour is less than 3.0 Petrosian half-light radii from the target galaxy. These cases are quite rare, representing only about 150 ‘fractured’ galaxies. The remainder of unmatched PYMORPH neighbours (almost all of the 60 000 PYMORPH neighbours) do not affect the fit because of their separation from the target and small magnitude.

Unmatched SDSS neighbours tend to happen when neighbours are superimposed on the target galaxy or very nearby. In principle, these situations can happen independent of the true physical separation of the neighbour and target (i.e. a star may be superimposed on a galaxy, the two of which should not be correlated in any physically meaningful way). The incidence of nearby neighbours may also be enhanced in dense environments (i.e. within clusters). This is a potentially important effect as it can bias our measurements of galaxies in clusters.

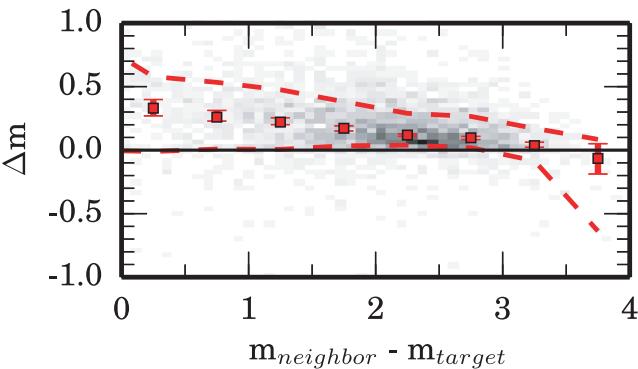
Many SDSS galaxies may be masked rather than simultaneously fit, causing it to appear that PYMORPH misses many neighbours. In order to examine the SDSS neighbours for contamination of our target galaxies, we examined the PYMORPH fitting masks to verify that these neighbours are not masked out. Any masked neighbours (for which at least 80 per cent of the pixels inside their half-light radii are masked) are removed from our set of potentially problematic SDSS neighbours. This removes approximately 4000 SDSS neighbours. We also remove SDSS neighbours that possess the DEBLEND\_NOPEAK  $r$ -band flag in the SDSS data (this indicates that no peak was found in the deblended source by the SDSS photopipeline, and the source is likely to be a spurious source).

Any SDSS neighbours that pass these cuts potentially corrupt our fits. We find empirically that a safe cut for considering these unmasked, unfit neighbours to be problems is if they are less than 3 half-light radii away and

$$m_n - m_t < 4.0 - r_d/r_{hl} \quad (6)$$

where  $r_d$  is the radial distance from the target galaxy to the neighbour,  $r_{hl}$  is the target galaxy half-light radius,  $m_n$  is the neighbour magnitude and  $m_t$  is the target magnitude. For example, a neighbour galaxy 4 arcsec away from an average galaxy in our sample (with half-light radius of 2 arcsec) may be no more than 2 mag dimmer than the target in order for us to be concerned that it was not chosen for simultaneous fitting. This cut naturally tapers, allowing brighter neighbours to be considered problematic out to larger radii relative to dimmer neighbour galaxies. After this cut, only about 7500 target galaxies potentially have contamination from unmatched SDSS neighbours.

Once the images with unfitted neighbours are identified, we perform an additional run of PYMORPH on the smaller sample. This run uses a different, deeper set of deblending settings for SExtractor as well as more generous settings of PYMORPH neighbour fitting. The



**Figure 5.** The change in magnitude (the fitted magnitude from the second pass minus the original fitted magnitude) of galaxies for which the fitting is rerun with deeper deblending and all the expected neighbours are detected. Individual galaxies are shown as grey scatter points. The median and 68 per cent contour are overplotted in red. The  $x$ -axis shows the difference in target and neighbour magnitudes. Galaxies tend towards dimmer magnitudes on the second fit which is consistent with a reduction in the contamination of the target galaxy by light from the neighbour source. Note that by a difference of 3 mag the median change in the parameters is consistent with zero. There is no need to look at unmatched SDSS sources dimmer than this.

SEXTRACTOR setting are changed to DEBLEND\_NTHRESH 64 rather than the original DEBLEND\_NTHRESH 32 and DEBLEND\_MINCONT 0.001 rather than the original DEBLEND\_MINCONT 0.005. This produces more fragmented sources with less contrast required to determine that a source is actually two blended objects. Changing these settings not only increases the likelihood of detecting the SDSS neighbours, but also of incorrectly ‘fracturing’ larger sources. We also reduce the neighbour source area requirement in PYMORPH for this fitting run. This makes PYMORPH more likely to perform a neighbour fit, requiring the neighbour to be only 10 per cent of the area of the target rather than the original setting of 20 per cent area.

After refitting we again test for unfit or unmasked neighbours. Galaxies still having neighbour problems at this point are flagged in the final catalogue as potentially being polluted by neighbour objects. This process improves the fits of  $\sim 3500$  galaxies, or 0.5 per cent of the total sample, for which deeper fitting identifies previously unmasked neighbours.

Fig. 5 shows the change in fitted Ser magnitude (the fitted magnitude from the second pass minus the original fitted magnitude) for galaxies where all neighbours were detected on a second pass of fitting using deeper deblending. Individual galaxies are shown as grey scatter points. The median and 68 per cent contour are overplotted in red. The  $x$ -axis shows the difference in target and neighbour magnitudes.

The plot shows that on average, the target galaxies get dimmer on a second pass as light from the neighbour is now properly associated with a different source (the neighbour) rather than being fitted as part of the target galaxy. A difference between target and neighbour of more than 3 mag causes little change in the fitted magnitude.

According to the 68 per cent contours on the plot, approximately 16 per cent of the refit galaxies become brighter during refitting. Since the original fits may include unmasked SDSS neighbours, the fits are usually poor. In some cases, the fit is not just contaminated by the neighbour, but the fit fails catastrophically, producing unreliable magnitudes and other fitting parameters. In these cases, the original fit would have been classified as a failed fit using our flagging

system. Such fits may produce a brighter magnitude when refit with proper identification of the neighbour target.

Similarly (although not shown here), separations of the source and target of more than 3 target half-light radii show little-to-no change in magnitude. The other fit parameters (i.e. Sérsic index, axis ratio, and radius) exhibit similar behaviours. Based on these observations, we believe our cutting criteria in equation (6) to be generous enough to capture the majority of cases where a true SDSS neighbour is causing fitting contamination.

We find about 4000 ‘polluted’ target galaxies for which the second pass with deeper deblending still fails to find the neighbour identified in SDSS. The ‘polluted’ and ‘fractured’ galaxies jointly comprise about 0.6 per cent of the sample. We mark galaxies that are suspected to have substantial fitting problems with a quality flag discussed in Section 4 and shown in Table A2. We also report the number of simultaneous fits for each galaxy in our catalogue since such fits substantially increase the number of free parameters during fitting and make it much more likely that the resulting fit has problems.

## 4 FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

While the fitting process is fairly straightforward, automated fitting routines tend to produce many poor fits. Non-physical ellipticity or sizes can easily be produced, especially if the sky is not properly estimated or neighbours are not properly accommodated. Inverted two-component profiles (where the component intended to fit the bulge fits the disc of a galaxy) are possible when the bulge has low Sérsic index, or when the S/N of the bulge is too low. Extreme cases of face-on or edge-on galaxies can also present difficulty. Overfitting (i.e. fitting components not truly present in the galaxy just to improve the  $\chi^2_v$ ) is easy to do, producing meaningless results for the components.

In this section, we discuss the flagging method used to separate good fits from poor fits in our catalogue. We attempt to provide a straightforward way of determining which fits to use and an interpretation of the fits that we claim to be good. Section 4.1 gives the motivation for our flagging system. Section 4.2 briefly describes the visually classified galaxies used to design and tune our flags. Section 4.3 describes the automated flags resulting from our visual inspection and gives a breakdown of the flags for the Ser-Exp catalogue (a more detailed description of the flags are available in the supplementary online material provided in Appendix A and example cases for each flag are presented in Appendix B).

### 4.1 The motivation for the flagging system

Independent inspection of the  $\chi^2$  values from each fitted model (deV, Ser, deV-Exp, and Ser-Exp) is a poor indicator of properly fitted models in our catalogue for some reasons. The most important reason is that while the  $\chi^2$  measures the ability of the model to appropriately fit the data, it makes no distinction between physical and unphysical models. Also, the number of degrees of freedom (DOF) is not well defined for non-linear models (see Andrae, Schulze-Hartung & Melchior 2010, for a more in-depth discussion), and the distribution of residual values do not approximate a normal distribution. Although the number of DOF’s and the resulting probability are not well defined, minimization of the  $\chi^2$  can still produce the ‘best’ fit by minimizing the  $\chi^2$  value. However, the statistical likelihood associated with the fitted  $\chi^2$  is not the best measure for determining physically meaningful fits.

**Table 1.** A breakdown of the automated flags characterized into categories useful for analysis. The first two groups can be used for analysis of total fits. When examining the subcomponents, consideration should be given as to exactly which groups of fits should be included.

Flag bit	Descriptive category	Per cent Ser-Exp	Example galaxy
–	<i>Good total and component magnitudes and sizes</i>	39.055	
10	Two-component galaxies	39.055	Figs B1, B2, and B3
11	No flags	18.095	Fig. B1
12	Good Ser, good Exp (some flags)	19.417	Fig. B2
13	Flip components	1.543	Fig. B3
–	<i>Good total magnitudes and sizes only</i>	54.945	
1	Bulge galaxies	18.964	Figs B4 and B5
2	No Exp component, $n_{\text{Ser}} > 2$	7.074	Fig. B4
3	Ser dominates always	11.889	Fig. B5
4	Disc galaxies	25.146	Figs B6, B7, B8, B9, and B10
5	No Ser component	16.876	Fig. B6
6	No Exp, $n_{\text{Ser}} < 2$ , flip components	0.551	Fig. B7
7	Ser dominates always, $n_{\text{Ser}} < 2$	0.103	Fig. B8
8	Exp dominates always	2.872	Fig. B9
9	Parallel components	4.745	Fig. B10
14	Problematic two-component galaxies	10.835	Figs B11, B12, B13, and B14
15	Ser outer only	7.504	Fig. B11
16	Exp inner only	0.425	Fig. B12
17	Good Ser, bad Exp, $B/T \geq 0.5$	0.017	Fig. B13
18	Bad Ser, Good Exp, $B/T < 0.5$	0.625	Fig. B14
19	Bulge is point	2.237	Fig. B15
20	Bad total magnitudes and sizes	6.000	

The focus of this work is not only on producing good fits, but on separating the cases where a second component is needed from those where the second component is not needed. We would also like to separate cases where the best fit is physically meaningful from those fits that are unphysical. For this, analysis beyond the  $\chi^2$  value is required. In place of the  $\chi^2$  test, we devise a series of physically motivated flags outlined in Table 1. We use these flags to determine the reliability of the various fits and the individual subcomponents. We also use the flags to mark poorly fitted galaxies.

The final goal of our flagging system is to identify galaxies as being in one of the following categories:

- (i) bulge-like galaxies,
- (ii) disc-like galaxies,
- (iii) two-component galaxies, and
- (iv) unknown type/poorly fitted/failed fits.

In the Ser and deV fits, these conditions simplify somewhat as we do not have to evaluate the appropriateness of the subcomponent parameters. Rather than focus on these broad (and perhaps vaguely defined) categories for classification, we develop a series of quantitative indicators based on the fitted parameters and assess the indicators to determine the quality of the fits.

Several methods of quality assessment have been used in previous works. Allen et al. (2006) separated galaxies based on the 1D radial profile, comparing the Ser and Exp components to separate galaxies into one of seven categories. More recently, Mendel et al. (2014) applied a similar criteria to distinguish one-component galaxies from galaxies that are better fit by two-component deV-Exp models. A different approach is to apply a statistical test similar to S11 and LG12, which perform an  $F$ -test to separate meaningful fits from galaxies that are likely overfit by the more complex models. Our approach is similar to the former groups rather than the latter. We separate the galaxies using a categorical description of the fits and show that these categories match the expected distributions of

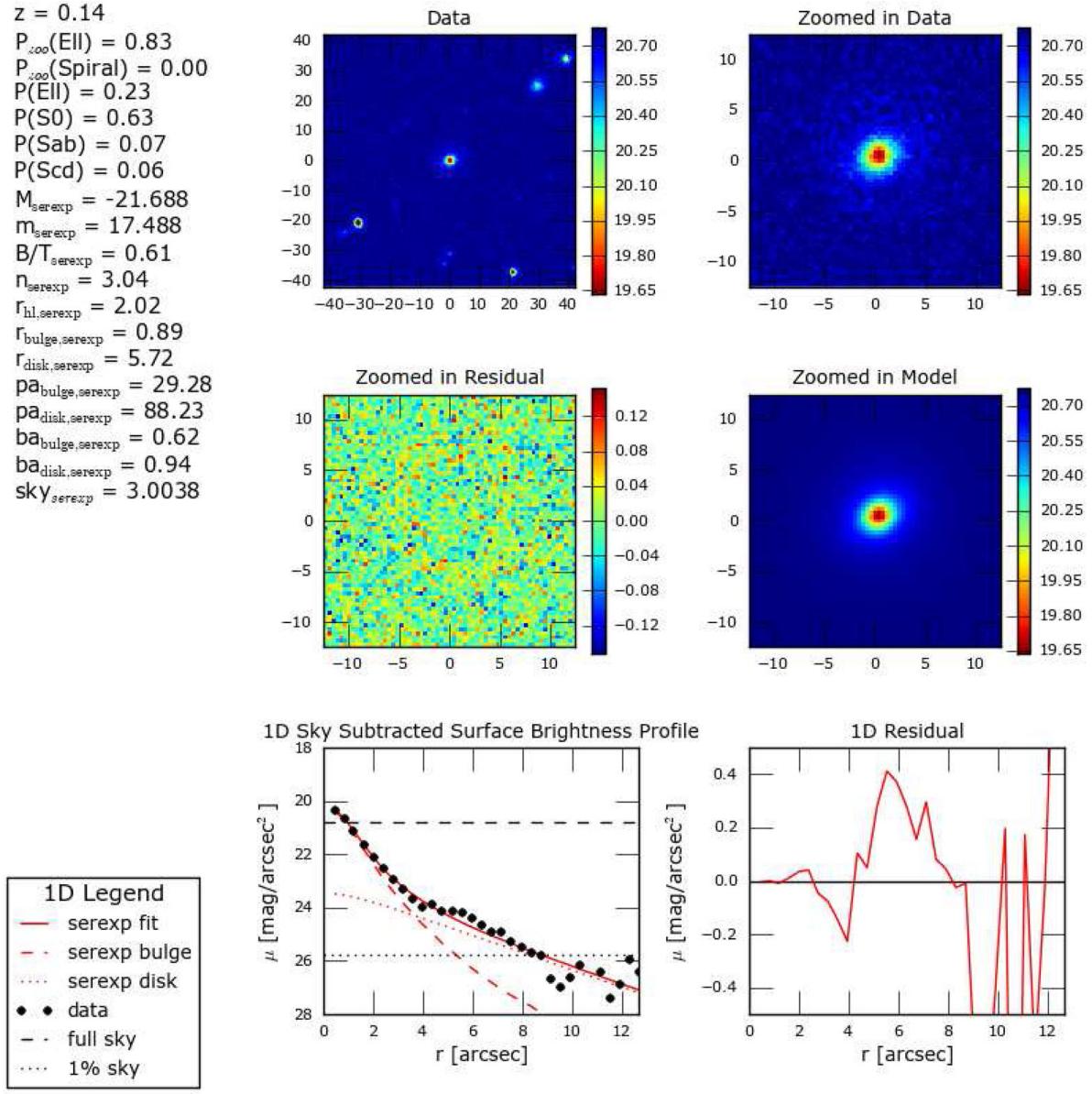
properties when compared to other observables like magnitude and radius. We compare the results of our method to the results of S11, LG12, and Mendel et al. (2014) in Sections 6 and 7.

## 4.2 The visual classification

No catalogue of  $10^5$  galaxies can be visually inspected in a reasonable amount of time without employing a large number of observers following a procedure similar to Galaxy Zoo (GZ; Lintott et al. 2008). We constructed a small training set of visually classified fits that were used to design our flagging criteria. For purposes of training and validation, we randomly selected a 1000 galaxy sample from the fitted galaxies for visual classification. The visual classification system has several categories including general fit problems and image characteristics. Each of the authors classified the sample by examining the fitted parameters, the 2D image, 2D model, 2D residual, the 1D profile, and the 1D residual.

Fig. 6 shows an example of the plots used to manually inspect the fits. Visual classification allowed us to enumerate the types of problems commonly found in the fits (i.e. effects of contamination or poorly defined bulges) and understand the relative importance of each common problem. The visual sample was selected randomly from the catalogue and fairly represents the full catalogue in photometric parameters (e.g. magnitude, radius, redshift, etc.). The incidence of problems associated with fitting in the visually classified sample is expected to be similar to that of the full catalogue. However, rarer problems are potentially missed in such a small sample.

After visual classification, the sample was randomly divided into two samples of approximately equal size, a training sample and a test sample. The training sample was used to define the automated flags described in Section 4.3. The test sample was set aside and used to test the reliability of the automated flags once they were defined.



Note: The 1d data is calculated using background-subtracted data. The 2d data is shown with background included.

**Figure 6.** An example panel used during visual classification showing the original stamp image, zoomed-in image of the galaxy, fitted galaxy, 2D residual, 1D galaxy profile, and 1D residual. In addition, many of the fitting parameters are displayed on the left-hand side of the image for comparison. When classifying the galaxy visually, we look for Ser components that dominate at large radii (which indicates a non-physical bulge), as well as misaligned components, large residuals, and other indications of bad fits.

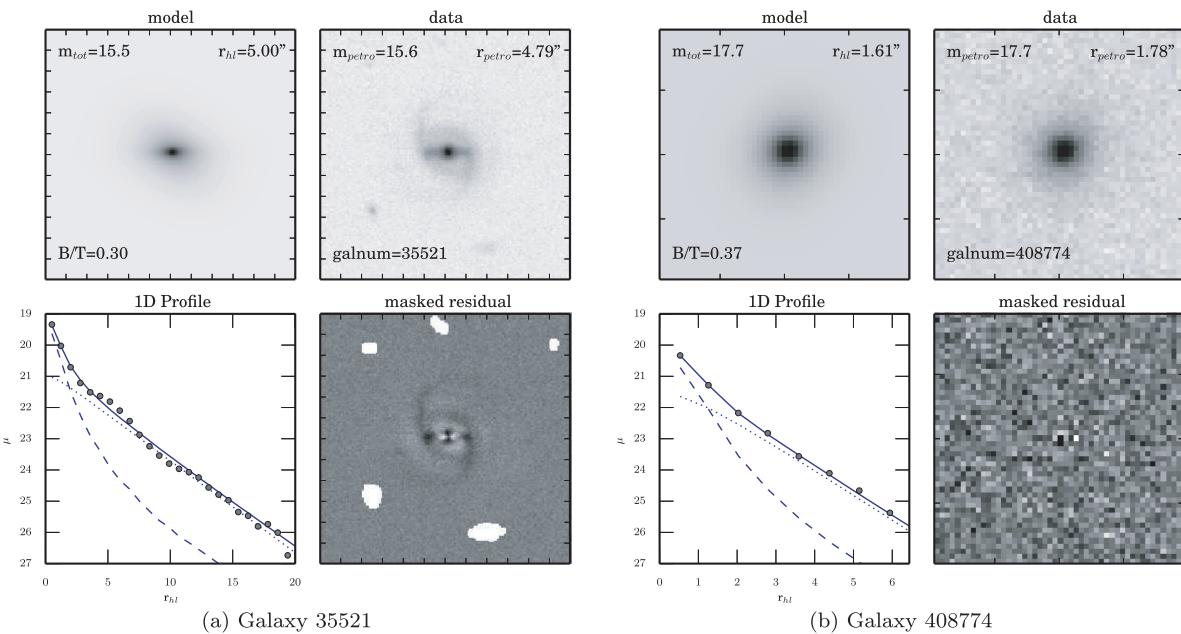
### 4.3 The automated flagging system

The automated bit flags are designed to accurately identify problems commonly observed during visual examination of the fits. Each automated flag has a tunable cutoff that is optimized using the training set of manually classified galaxies. Then the efficacy of the flag is evaluated using the test set of galaxies (which are also visually classified). Appendix A (available online) describes the flags in more detail including the decision criteria for the flags (see Table A1) and the per cent of galaxies possessing each flag (see Table A2).

Table 1 shows the flagging categories, the final percentage of Ser-Exp galaxies in each category, and references to example galaxies for each flag. Example galaxies selected to characterize

each flag are presented in Appendix B (available online). Fig. 7 shows an example of galaxies with bits 10 and 11 set. These panels and additional examples of other flag categories are available in Appendix B. We show two example galaxies for each flagging category. For each fit, we show the 2D data, fitted model, and residual. We also show the 1D radial data profile, bulge and disc component profiles, and the total fitted profile.

We use a tiered structure to describe the fits. The most general description is the accuracy of the total magnitude and radius. About 94 per cent of the Ser-Exp sample is classified as having an accurate measurement of the total magnitude and half-light radius. Within the first tier assessing the accuracy of the total magnitude and radius, we then separate galaxies into single-component galaxies (flag bits 1 and 4), two-component galaxies (flag bit 10), and problematic



**Figure 7.** Examples of Ser-Exp fits which we flag as good; i.e. the two components of the fit are real and without fitting problems. These galaxies are in category ‘No Flags’ in Table 1 and have flag bits 10 and 11 set.

two-component galaxies (flag bit 14) that have fit parameters that are difficult to interpret.

We recommend using the categories presented in Table 1 to select samples from the catalogue. In particular, we highlight the major categories of our flagging as follows.

(i) *Good two-component fits (bit 10 set)*. These are the galaxies we find to have two fitted components with intermediate bulge-to-total light ratio ( $B/T$ ) and reasonably well-behaved subcomponents. We recommend using both the subcomponents and the total magnitude and radius in any analysis. Bits 11–13, only one of which will be set, give a more specific description of the galaxy.

(ii) *Good bulge fits* (*bit 1 set*). These are the galaxies we find to have little or no evidence of an Exp second component. The  $B/T$  can be as low as 0.8, however, the Exp component is not trustworthy. We recommend using the *Ser* fit for these galaxies, and they should be treated as having  $B/T=1.0$  regardless of the fitted  $B/T$ . Bits 2 and 3, only one of which will be set, give a more specific description of the galaxy.

(iii) *Good disc fits (bit 4 set)*. These are the galaxies we find to have little or no evidence of a **Ser** second component. The  $B/T$  can be as high as 0.2, however, the **Ser** bulge component of the **Ser-Exp** fit is not trustworthy. We recommend using the **Ser** fit total magnitude and radius in any analysis. The **Ser** fit should be used for galaxies, and they should be treated as having  $B/T=0$ , regardless of the fitted  $B/T$ . Bits 5–9, only one of which will be set, give a more specific description of the galaxy.

(iv) *Problematic two-component fits (bit 14 set)*. These are the galaxies we find to have two fitted components with intermediate  $B/T$ , but at least one subcomponent has strange behaviour. These galaxies likely require additional investigation prior to including them in any analysis. Bits 15–19, only one of which will be set, give a more specific description of the galaxy.

(v) *Bad fits* (*bit 20 set*). These are the galaxies we find to have severe problems with the fit. They should not be included in any analysis without close examination. Even the total magnitude and total radius are believed to have significant errors. Bits 21–26, only

one of which will be set, give a more specific description of the galaxy.

For the single-component `Ser` and `deV` fits, the two-component categories listed above have no galaxies in them, but we retain the same flagging structure for all fits to ensure consistency and ease of use.

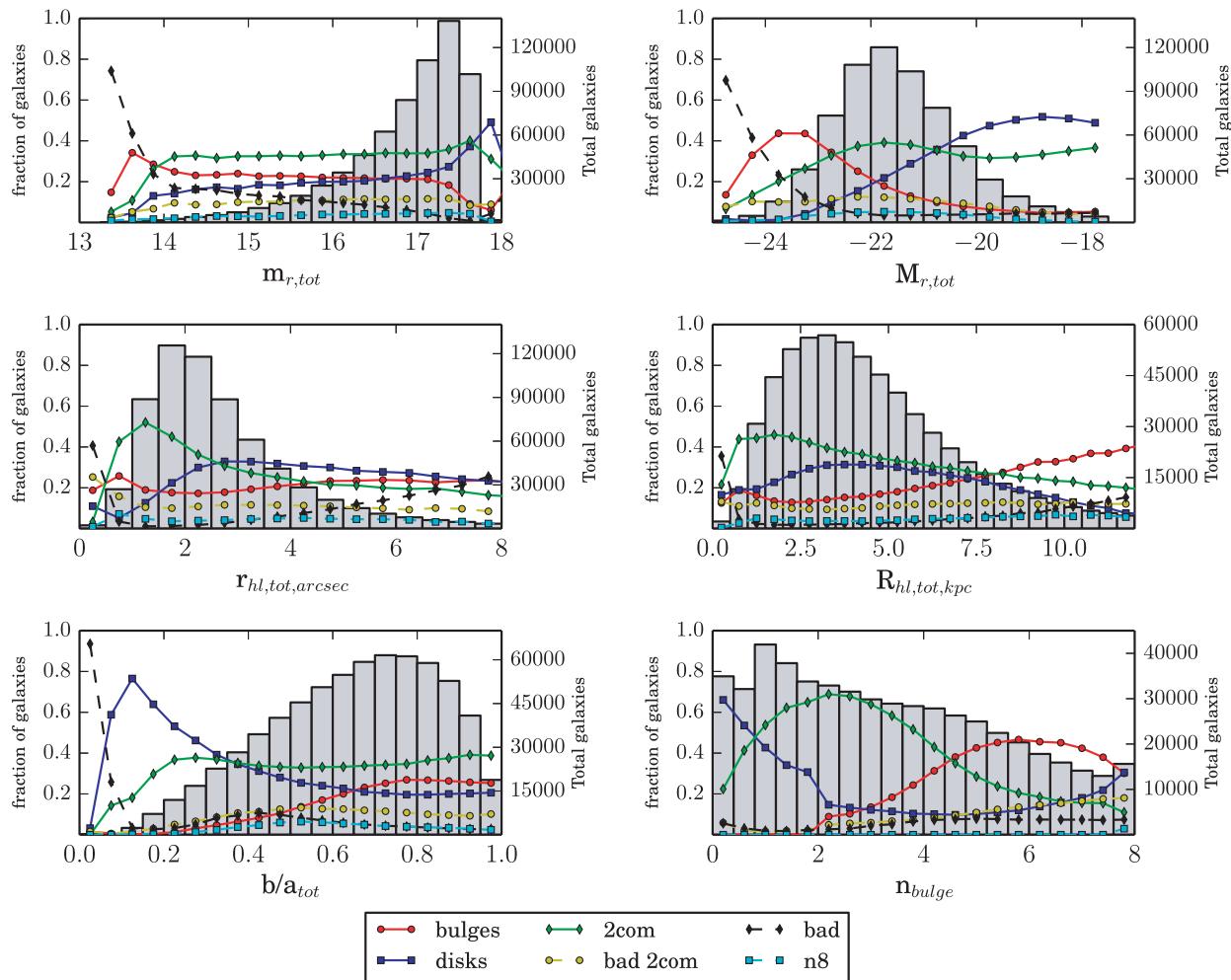
The flagging here focuses on identifying Ser versus Ser-Exp fits. While the Ser fit can reproduce the deV fit in cases where the Sérsic index is 4, it will be different for other values. We choose to carry out and report deV fits for comparison purposes. Similarly, the deV-Exp fit is included as a separate fit because the Ser-Exp fit is not necessarily equivalent to the deV-Exp fit.

## 5 INTERNAL COMPARISONS AND CONSISTENCY CHECKS

Before comparing the fits, in particular the **Ser-Exp** fits, to other groups, we perform some internal consistency checks. Section 5.1 examines the distribution of general classifications described in Section 4.3 for the **Ser-Exp** catalogue as a function of basic observables. Section 5.2 examines the bulge radii with respect to the PSF. Section 5.3 compares some of the fitted values of the **Ser** fit to the **Ser-Exp** fit.

### 5.1 Examination of flags with basic observables

Using the flags from Section 4.3, we should be able to reproduce sensible distributions of magnitude, radius, and ellipticity provided that our chosen flags can reliably separate good and bad fits within our sample. We also include the bulge Sérsic index, which is used during classification. In Figs 8 and 9, we show the percentage of two component (green points, labelled as ‘2com’, flag bit 10), bulge (red points, labelled as ‘bulges’, flag bit 1), disc (blue points, labelled as ‘discs’, flag bit 4), problematic two component (yellow points, labelled as ‘bad 2com’, flag bit 14), and failed galaxies (black points, labelled as ‘bad’, flag bit 20) in the Ser-Exp catalogue



**Figure 8.** The bin-by-bin per cent of galaxies of each type: disc (blue points, labelled as ‘discs’), bulge (red points, labelled as ‘bulges’), two component (green points, labelled as ‘2com’), problematic two component (yellow points, labelled as ‘bad 2com’), and failed galaxies (black points, labelled as ‘bad’), according to our categorical flags. We have also separated the good two-component models into those with Sérsic indices below 8 and those galaxies with acceptable fits but the Sérsic index of the bulge hits the  $n = 8$  boundary of the parameter space (cyan points, labelled as ‘n8’) to check for any bias resulting from the restriction on the fitted Sérsic index. In the background of the plot, we plot the total distribution of galaxies with respect to the parameter used to bin the data. For example, there are approximately 60 000 galaxies in our catalogue with apparent magnitudes between 16.5 and 16.75 (from the top-left panel). Of these 60 000 galaxies, about 37 per cent are good two-component fits, 20 per cent are bulges, 20 per cent are discs, and the remaining  $\sim 20$  per cent are a mixture of the remaining classes. When summing the percentages over all model types, each bin sums to 100 per cent. The percentage of two-component galaxies is mostly stable with respect to apparent size and magnitude. However, the data favour more two-component models at small half-light radii. This effect is examined in the text.

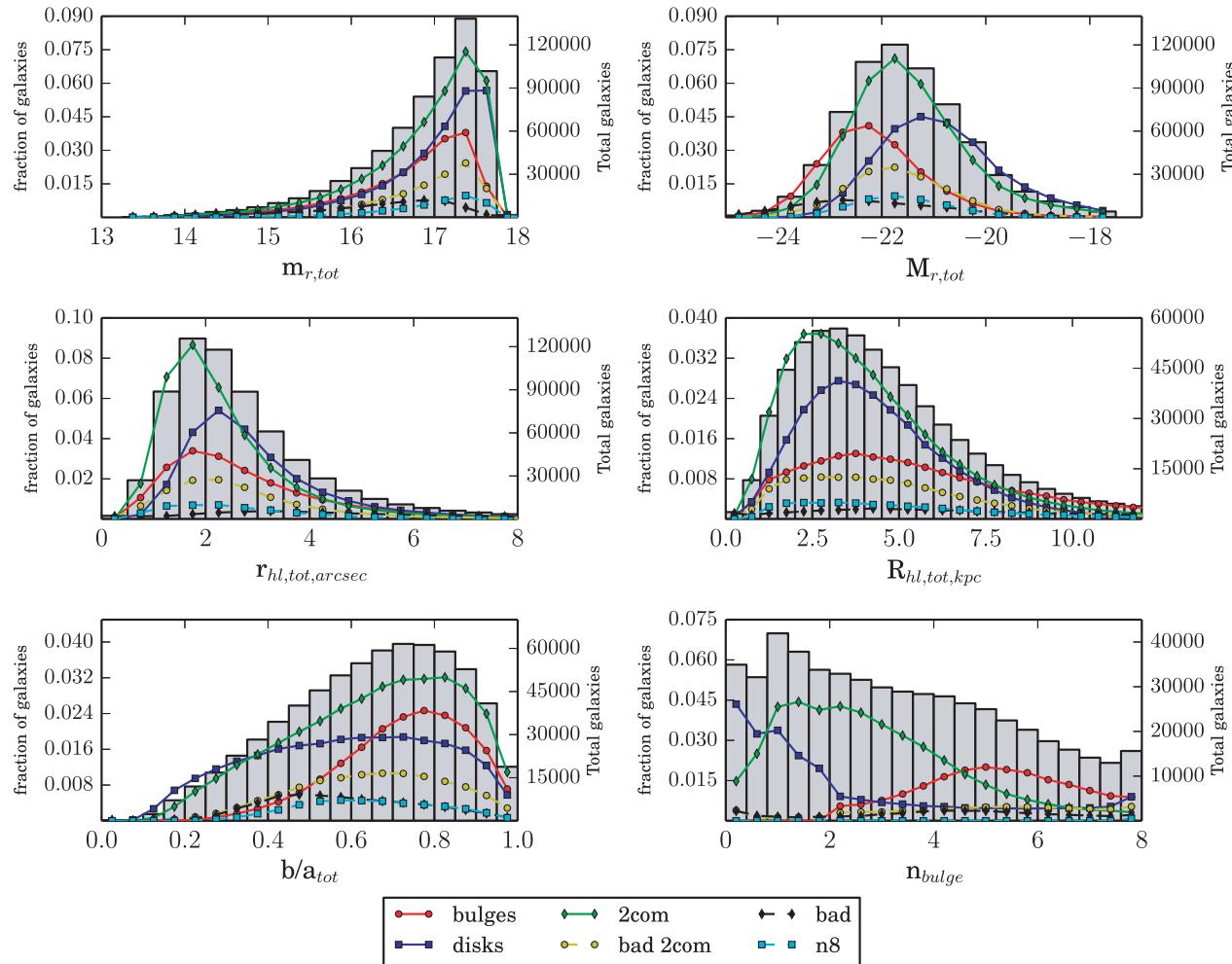
as a function of several parameters. The figures show the same distribution of the categories using different normalizations. Fig. 8 is normalized such that the sum of the fractions of all categories sums to 1 in each bin (i.e. the reported fractions for the blue line with square points show the fraction of galaxies at a given magnitude, size, etc. that are discs). Fig. 9 is normalized by the total number of galaxies in the sample (i.e. the reported fraction for the blue line with square points show the fraction of all galaxies in the sample galaxies that are discs and have a given magnitude, size, or other measured parameter).

In both figures, the total distribution of galaxies is plotted in the background in order to give the reader a sense of the number of galaxies in each bin. The left-hand axis denotes the fraction of galaxies in given category (e.g. bulge, disc, or other category) using the relevant normalization. The right-hand axis denotes the raw number of galaxies plotted in the background histogram relative to 670 722, the total number of galaxies in our sample.

We have also separated the good two-component models into those with Ser-Exp bulge Sérsic indices below 8 and those galaxies for which our flags indicate the fitted profiles are acceptable but the Ser-Exp bulge Sérsic index hits the  $n = 8$  boundary of the parameter space (cyan points, labelled as ‘n8’). Galaxies with the Ser-Exp bulge Sérsic index approaching the boundary of the parameter space may exhibit problems due to the inability of the code to reach the true minimum of the fit. We separate out these fits to check for any strange behaviour that would suggest substantial biases in the fitting due to this effect.

### 5.1.1 Behaviour with apparent magnitude

The original sample selected from SDSS is defined by a cut in extinction-corrected  $r$ -band apparent Petrosian magnitude removing all galaxies brighter than 14 or dimmer than the spectroscopic sample cut of 17.77. Examination of the type distribution with



**Figure 9.** The per cent of galaxies of each type, given as a percentage of the total sample. The colours and plot design are the same as Fig. 8. For example, approximately 3 per cent of all galaxies are ‘2com’ galaxies fitted with apparent magnitudes between 16.5 and 16.75 (from the top-left panel).

respect to total fitted apparent magnitude derived from the Ser-Exp fits by summing both components of the fit (top-left panel) reveals that the percentage of good two-component galaxies is independent of apparent magnitude across magnitudes 14–17.77 where the majority of our sample is located. Independence with apparent magnitude is expected unless there are observational biases in the flags. However, the endpoints (i.e. below 14 and above 18) show a large increase in the percentage of failed galaxies.

Since the Petrosian magnitudes are known to be dimmer at the brighter end of the magnitude distribution (Blanton et al. 2001; Bernardi et al. 2013) and dimmer for galaxies with higher Sérsic index (or greater concentration) (see fig. 7 of Graham & Driver 2005), galaxies with fitted magnitudes brighter than 14 (where our original selection cut is made; see Section 2) are not immediately identifiable as failed fits. Indeed, Petrosian magnitude is dimmer than fitted magnitude by approximately 0.2 for a Ser galaxy with Sérsic index of 4 (however this increases to about 0.55 mag for  $n = 8$  galaxies). Larger concentration (or equivalently, Sérsic index) is required to account for larger differences between Petrosian and fitted magnitude. Similarly, as concentration increases, the proportion of galaxies possessing at least that level of concentration decreases. So, a reasonable expectation is that the failure rate should increase

brighter than 14. In addition, fitted magnitudes should rarely be dimmer than the Petrosian magnitude unless there was a failure in deblending or some other photometry problem. Therefore, the majority of the galaxies outside the range of 14–17.77 mag should be flagged as failed cases because their magnitudes vary greatly from what is expected.

The increasingly high failure rate at the bright end of the apparent magnitude plotted above 14 and the increase in the failure rate at the dim end below 17.77 are an indicator of the ability of the flags to identify poorly fit galaxies. Indeed, the majority of the failed galaxies ( $\sim 800$  out of the  $\sim 1000$ ) in the magnitude bins brighter than 14 have either flags 21 or 22 set. These flags identify galaxies with large, extended components that are due to underestimating sky brightness or contamination from nearby neighbours. This makes sense as we expect galaxies with these bright magnitudes to have their brightness substantially overestimated.

Classification of bulge, disc, or 2com galaxies is also increasingly unreliable in galaxies outside of the 14–17.77 mag range. Since these galaxies are not selected in our original sample, we may not have an appropriate training sample for these galaxies. Therefore, careful consideration of any galaxies dimmer than 17.77 or brighter than 14 should be taken before including them in any analysis.

### 5.1.2 Behaviour with absolute magnitude

Figs 8 and 9 also show the behaviour of our component categories with respect to total fitted absolute magnitude derived from the Ser-Exp fits by summing both components of the fit (top-right panel). Bulge galaxies (i.e. elliptical) galaxies dominate at the brightest magnitudes while disc galaxies dominate at the dimmest magnitudes and two-component models dominate at the intermediate magnitudes. There is an increase in failed cases near the bright end of the distribution. Further inspection of this end of the distribution shows that it is a pile-up effect due to the preference of PYMORPH to overestimate the brightness of a galaxy when a fitting failure occurs. Large components (either Ser or Exp components; both occur at similar rates) can be wrongly used by PYMORPH to fit sky or neighbours. These large components will contribute to the brightness of the source, making it appear brighter than the true brightness. Therefore, the failed galaxies will tend to shift up the magnitude distribution to brighter magnitudes.

To test for this problem, we also examined the distribution of galaxy types using the model-independent Petrosian magnitudes. This shifts many of the failed cases back to the dimmer Petrosian magnitude bins and smooths out the distribution of failed galaxies at the bright end. As a result, the failure rate at the bright end is substantially lower when viewed in Petrosian magnitudes. The failure rate increases roughly linearly between  $-23$  and  $-25$  mag from about 10 to 30 per cent rather than increasing to 80 per cent as shown in Fig. 8. The percentage of failed cases in the magnitude range  $-23$  to  $-22$  also increases. This shows that a small failure rate in the dimmer bins (magnitudes between  $-23$  and  $-22$ ) is causing a substantial contribution to the number of galaxies in the brighter bins (brighter than  $-23$ ). However, these failed cases far outnumber the legitimate galaxies in the brightest bins causing the apparent failure rate to approach or exceed 50 per cent.

The occurrence of two-component models is also moderately higher in the magnitude bins brighter than  $-23$  when viewed in Petrosian magnitudes. The recent works of Mosleh, Williams & Franx (2013) and Davari et al. (2014) report improved fitting and recovery of magnitude and radius when using two-component models for nearby elliptical galaxies (as opposed to galaxies at  $z \sim 1$ ). Huang et al. (2013) also provide evidence for using three components when fitting well-resolved elliptical galaxies (requiring resolution substantially better than 1 kpc). The question of exactly how to interpret such an additional component is beyond the scope of this work, so we only comment on the trend here and caution against a simple bulge+disc interpretation of these galaxies. We will return to this issue in Section 7.

### 5.1.3 Behaviour with apparent and absolute half-light radius

Figs 8 and 9 also show the behaviour of the percentage of our component categories with respect to the total absolute half-light radius derived from the Ser-Exp fits by measuring the half-light radius of the total profile (second row, right). Here, we observe expected trends in physical size (larger physical size should be dominated by bulge galaxies). When observing in apparent size (second row, left), the incidence of two-component galaxies increases with smaller apparent size and peaks at 1.5 arcsec, above the half-width at half-maximum (HWHM) of the PSF (which is about 0.7 arcsec in the  $r$  band). At these sizes, the percentage of discs drops substantially while the percentage of bulges remains constant.

The shift in the percentage of galaxy types at small half-light radii is consistent with an interpretation that the observed shift to

two-component galaxies is due to observational effects of the magnitude limit on the distribution of galaxies rather than systematics in the fitting. For example, Fig. 8 shows that the peak of the pure disc sample is near 5 kpc in size. This is near the typical size of late-type discs for galaxies at  $-21$   $r$ -band Petrosian magnitude in the SDSS sample (see figs 5 and 6 of Shen et al. 2003). The  $-21$  magnitude also corresponds to the peak of the disc galaxy distribution in Fig. 9. When this size (5 kpc) is translated to an apparent size at  $z \approx 0.05$  the expected size of these disc galaxies would be nearly 4 arcsec, well above the sizes where this effect occurs. Even at  $z \approx 0.15$ , which is higher redshift than approximately 80 per cent of our sample, the expected size of these disc galaxies would be about 2 arcsec. Therefore, it is reasonable not to expect many disc galaxies below 2 arcsec where the drop in pure disc systems occurs.

Even if such a drop in disc galaxies was believed to be fitting bias, PSF effects have been shown to set in near the HWHM of the PSF (see Section 5.2 for justification and further discussion). The drop between 1 and 2 arcsec is above the HWHM of the PSF where sizes are potentially biased by the PSF.

A similar examination of bulge galaxies versus two-component galaxies is less conclusive. Both groups (bulge and two component) should be present at the smaller radii (below 2 arcsec).

Further examination of the  $B/T$  of the two-component and pure bulge galaxies is shown in Fig. 10. The median  $B/T$  as a function of the fitted total half-light radius for pure bulge galaxies (left) and two-component galaxies (right) is shown. Symbols show median values. 68 and 95 per cent contours are plotted as dashed and dot-dashed lines. Error bars on the median values represent 95 per cent confidence intervals (CI) obtained from bootstrap resampling.

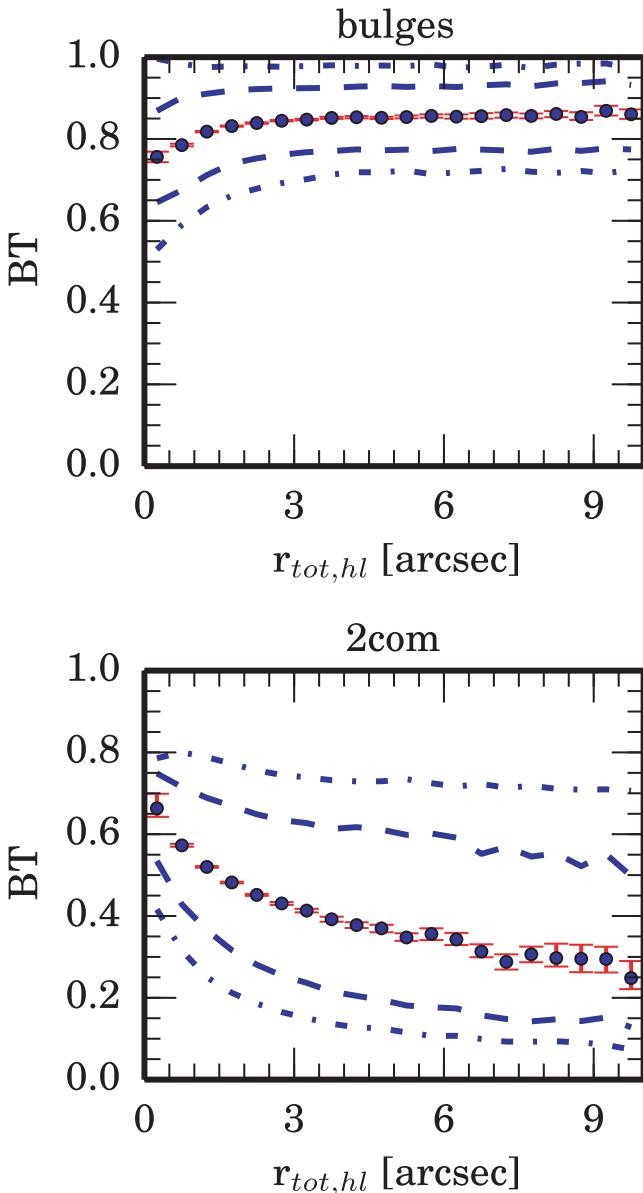
$B/T$  increases to a median value of 0.6 at small radii versus median values of 0.4 in the larger apparent radii bins. The fitted  $B/T$  of the Ser-Exp model for pure bulge galaxies also drops from  $\approx 0.85$  to  $\approx 0.75$  in the smallest radii bins. This indicates a possible bias in our catalogue of preferentially classifying apparently small galaxies as two-component galaxies with moderately high (0.6–0.8)  $B/T$  due to effects of the PSF. We examine bias due to the PSF further in Section 5.2. Here, we caution the reader that there may be some classification bias preferring two-component fits at small sizes.

### 5.1.4 Behaviour with axis ratio

Figs 8 and 9 also show the distribution of total axis ratios derived from the Ser-Exp fits by measuring the axis ratio of the total profile for the fitted galaxies in our catalogue (bottom left). If the flagging properly identifies pure bulge systems, we would expect to see pure bulge systems concentrated near axis ratios approaching 1 since early-type galaxies are ellipsoidal. Lower values of axis ratio should be dominated by pure discs and two component systems if these categories properly identify late-type galaxies. Fig. 8 shows this behaviour. Fig. 9 also shows that the peak of the bulge galaxy distribution is near 0.8 with few galaxies at small axis ratios.

### 5.1.5 Behaviour with bulge Sérsic index

Figs 8 and 9 also show the distribution of bulge Sérsic index derived from the Ser-Exp fits to galaxies in our catalogue (bottom right). Fig. 8 shows that two-component galaxies dominate the objects fitted with bulge indexes between 1 and 4. Galaxies with lower (higher) bulge indexes tend to be classified as pure disc (bulge) galaxies. Fig. 9 also shows similar behaviour.



**Figure 10.** Symbols show median  $B/T$  as a function of the fitted total half-light radius for pure bulge galaxies (left) and two-component galaxies (right). 68 and 95 per cent contours are plotted as dashed and dot-dashed lines. Error bars on the median values represent 95 per cent CI obtained from bootstrap resampling. The  $B/T$  of pure bulges drops as the apparent size of the galaxy decreases. At the same time, the  $B/T$  of the two-component galaxies increases, suggesting the ability to distinguish one- and two-component galaxies is reduced at small apparent size.

However, the reader should be cautioned that interpretation of these plots is complicated by the fact that the bulge Sérsic index is used during classification. Also, the  $B/T$  of these galaxies is not considered in this plot. Many galaxies, especially with low bulge index, have low  $B/T$ , making the observed bulge index unreliable.

#### 5.1.6 Interpretation of the two-component galaxies

The distribution of bulge and disc galaxies appear to make sense when the distributions are examined with respect to basic observables like magnitude, size, and axis ratio. Therefore, we consider it appropriate to associate our bulge and disc classes with early and

late-type galaxies, respectively. Further evidence for this claim is discussed in Section 7. In contrast, we have not yet explored the behaviour of the Ser-Exp bulge and disc components for galaxies that we claim to be two-component galaxies.

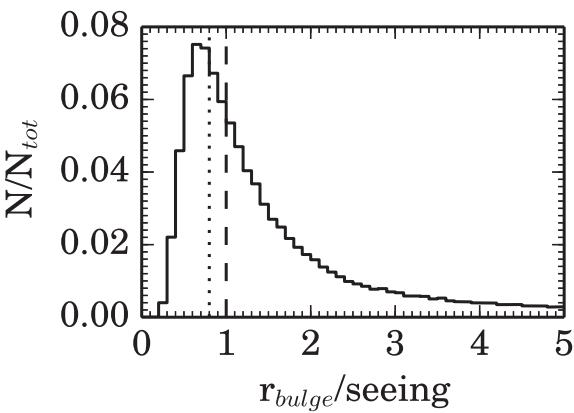
M13 showed that accurate measurement of the total size and magnitude can be accomplished without the components necessarily having physical interpretation. It also showed that PYMORPH can reliably recover Ser and Exp components of two-component models down to component magnitudes of roughly 19 in the  $r$  band. However, accuracy of the fitted components does not guarantee that the components represent physical bulges or discs. Indeed interpretation of the Ser-Exp subcomponents as true bulge and disc components is complicated. We continue to refer to the Ser and Exp components of two-component Ser-Exp fits as ‘bulge’ and ‘disc’ throughout the paper. However, we caution the reader that there are many cases where this simple interpretation does not make sense. We will return to this issue in Section 7.

#### 5.2 PSF effects on bulge radius

The PSF limits our ability to accurately recover component radii and Sérsic index when the PSF is larger than the component in question. Gadotti (2008) examined the effects of low resolution on galaxy parameters by performing detailed fitting of 17 nearby ( $z \sim 0.005$ ) nearly face-on ( $b/a > 0.9$ ) SDSS disc galaxies. The galaxies were then redshifted to  $z \sim 0.05$  and refit. Fig. 13 in Gadotti (2008) and the accompanying discussion shows that bulge sizes smaller than 80 per cent of the seeing radius (or equivalently 80 per cent of the PSF HWHM) can be biased high (by as much as 50 per cent). Bulge Sérsic indices can also be suppressed for values greater than 2 (by as much as 1), and  $B/T$  can be biased high (by as much as 0.1). However, the authors caution against extending these expectations to higher redshifts as the physical scale of the PSF grows substantially with redshift.

Gadotti (2009) studied a larger sample of galaxies compared to Gadotti (2008) (~3000 compared to 17 galaxies) with a wider range of types (including ellipticals). The decision of whether to fit a second component was based on individual inspection of the radial light profile. Fig. 7 in Gadotti (2009) shows the distribution of the quantity (bulge radius/PSF HWHM) for galaxies with a detected bulge component. The authors find that 3 per cent of their sample are bulges that are smaller than 80 per cent HWHM and 10 per cent are smaller than the HWHM. Bernardi et al. (2014) found similar results. However, excluding galaxies below the 80 per cent level or even the more conservative HWHM do not change the measurements of physical bulge sizes made by either group.

Fig. 11 shows the comparison of the bulge radius to PSF size for all Ser-Exp bulges in our catalogue. 15 per cent (24 per cent) of bulges lie below  $0.8 \times$  HWHM ( $1.0 \times$  HWHM). The percentage of poorly resolved bulges is somewhat higher than either Gadotti (2009) or Bernardi et al. (2014). For the pure bulge galaxies (i.e. flag bit 1 set), only 2 per cent (3 per cent) of pure bulge galaxies have radii smaller than  $0.8 \times$  HWHM ( $1.0 \times$  HWHM). Including pure bulge galaxies and two-component galaxies with  $B/T > 0.5$  (i.e. Ser-Exp fits with flag bit 1 set or with  $B/T > 0.5$  and flag bit 10 or 14 set), 7 per cent (12 per cent) of the bulges are smaller than  $0.8 \times$  HWHM ( $1.0 \times$  HWHM). These numbers are still higher, but in closer agreement with both Gadotti (2008) and the sample used in Bernardi et al. (2014). In contrast, 28 per cent (41 per cent) of two-component galaxies with  $B/T \leq 0.5$  have bulges smaller than  $0.8 \times$  HWHM ( $1.0 \times$  HWHM). The work of Gadotti (2009) suggests that a significant proportion of our Ser-Exp bulge components with



**Figure 11.** The comparison of Ser-Exp bulge circularized effective radius to the PSF HWHM. 15 per cent (24 per cent) of bulges lie below the cut of  $0.8 \times \text{HWHM}$  ( $1.0 \times \text{HWHM}$ ).

$B/T \leq 0.5$  are potentially biased to larger sizes and brighter magnitudes by poor resolution. This must be considered when looking at the bulge components of galaxies with significant discs. We include the PSF size in the catalogue so that this consideration can be made during future analysis.

M13 also showed that the effective bulge radius for bulges smaller than 1 arcsec is overestimated by  $\sim 5 \pm 20$  per cent. The simulations reflect a tendency to overestimate bulge radius as reported in Gadotti (2008). We do not correct for this effect here, but caution the user that small bulges (smaller than the HWHM) are likely biased larger and brighter. This effect likely contributes to the increase in two-component galaxies at small apparent sizes as discussed in Section 5.1.3.

### 5.3 Comparison of the Ser and Ser-Exp models

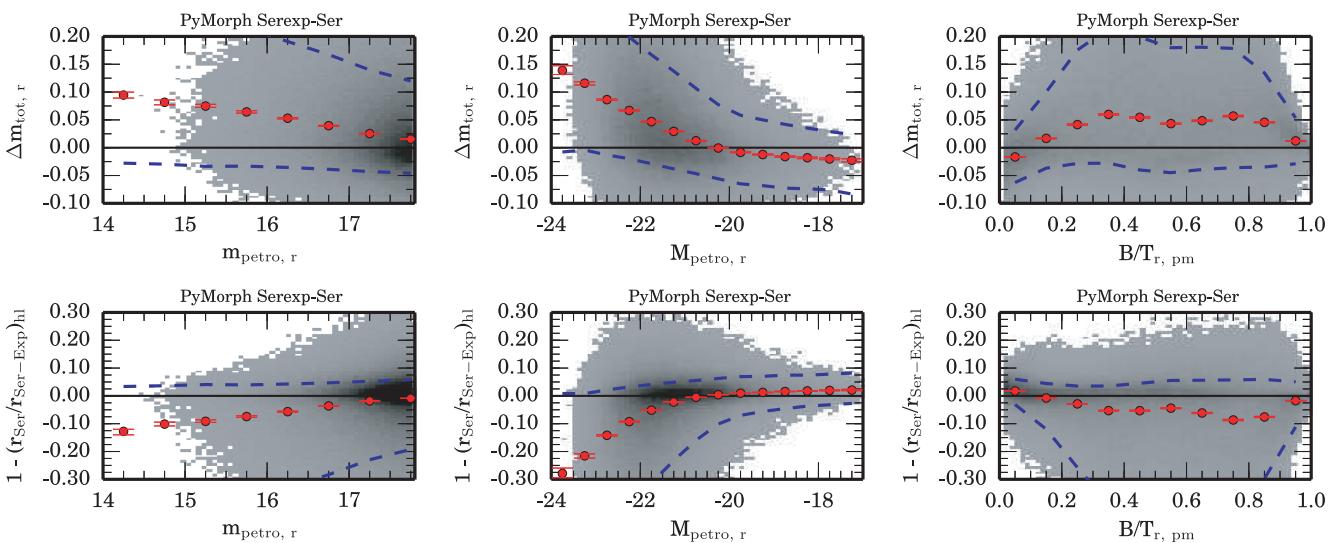
In this work, we favour the Ser-Exp model for both deciding on the structure of the galaxies (i.e. one versus two components) and for

estimating the total size and magnitude. Fig. 12 shows comparisons of fitted magnitude (top row) and half-light radius (bottom row) between the Ser and Ser-Exp fits. The top row shows comparisons of the fitted magnitude as a function of apparent Petrosian magnitude (left), absolute Petrosian magnitude (centre), and Ser-Exp  $B/T$  (right). Magnitude differences are reported as Ser-Exp – Ser magnitude, therefore a positive magnitude indicates a brighter Ser fit relative to the corresponding Ser-Exp fit. Median values for each bin are plotted in red with the error bars representing a 95 per cent bootstrap CI on the median. The 68 per cent contours of the data are plotted as blue dashed lines. The density of points is plotted in grey-scale with the low end of the density representing 1 per cent of the maximum density. Bins with a density below the minimum density are plotted in white.

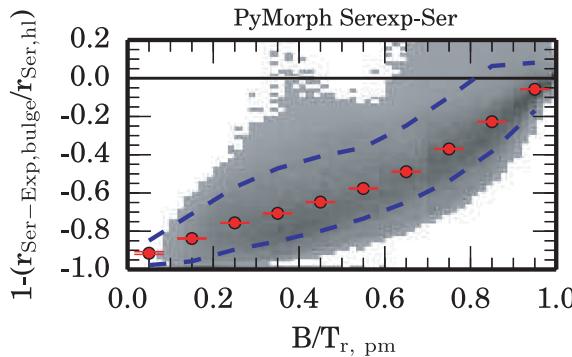
Analogous plots are presented for the half-light radius in the second row. In all the plots presented here, we bin galaxies by Petrosian magnitude in order to make consistent comparisons across the different models. The Petrosian magnitude provides a model-independent measure of the magnitude so that any comparison has the same distribution of galaxies among the bins.

We observe larger differences in the magnitude at brighter apparent magnitude, brighter absolute magnitude, and intermediate  $B/T$  values. These differences are consistent with the effects seen in the simulations of M13. The bias is caused by inappropriately fitting a one-component model to a more complicated light profile. This type of underfitting was shown to bias measurements of the Ser magnitude to be brighter than the combined Ser-Exp magnitude measured from the same galaxy. This effect is also reported by Mosleh et al. (2013) and Davari et al. (2014).

Fig. 1 of Bernardi et al. (2013) also shows the difference in magnitudes for the Ser and Ser-Exp fits of this work as a function of absolute magnitude. The figure also shows the comparison of the Ser fit presented in this work to the Ser fits of S11 and the SDSS magnitudes. The Ser magnitudes from this work are systematically brighter by up to 0.5 mag as you approach  $-24$ . This difference leads to large differences in the bright end of the LF. However, consistent difference of the Ser magnitudes is not an indication of bias in our fits because Kravtsov et al. (2014) showed that the



**Figure 12.** The comparison of Ser fit to Ser-Exp fit for galaxies in this work. The top row shows comparisons of magnitude (Ser-Exp – Ser) as a function of apparent Petrosian magnitude, absolute Petrosian magnitude, and  $B/T$ . The bottom row shows comparisons of half-light radius as a function of apparent Petrosian magnitude, absolute Petrosian magnitude and the fitted Ser-Exp  $B/T$ . At the low and high  $B/T$ , the bias is nearly 0 with smaller scatter. In the intermediate  $B/T$  values where a single-component model would be expected to fit poorly, the bias and scatter are larger.



**Figure 13.** The comparison of Ser radius to the Ser-Exp bulge radius as a function of  $B/T$  for galaxies in this work. As  $B/T$  increases, the Ser-Exp bulge radius approaches the Ser radius as expected. As  $B/T$  approaches 0, the ratio approaches 0 indicating that the bulge is shrinking in size with magnitude as would be expected if the bulges are properly fitting the central bulge of the galaxy. Median values for each bin are plotted in red with the error bars representing a 95 per cent bootstrap CI on the median. The 68 per cent contours of the data are plotted as blue dashed lines. The density of points is plotted in grey-scale with the low end of the density representing 1 per cent of the maximum density. Bins with a density below the minimum density are plotted in white.

magnitudes reported here are more consistent with their results than either SDSS or S11.

The bottom row of Fig. 12 shows comparisons of half-light radius between the Ser and Ser-Exp fits. We show the fractional difference in radii. For example, a value of 0 indicates no disagreement between the compared radii. A value of  $-1$  would indicate that the Ser model is 100 per cent larger compared to the Ser-Exp model. There is a difference of 5 per cent in radius and 0.05 mag at the intermediate  $B/T$  values. The lowest and highest  $B/T$  values show both agreement in the median difference between Ser and Ser-Exp models as well as decreasing scatter.

Fig. 13 also shows comparisons of bulge radius for the Ser and Ser-Exp fits. The Ser-Exp bulge radius gets smaller with decreasing  $B/T$  relative to the single-component Ser radius, as one might expect if the Ser component of the Ser-Exp model is fitting the central bulge of the galaxy. While this shows that the bulges are mostly compact, we discuss the correspondence between the Ser component of the Ser-Exp fit and the central bulge of galaxies further in Sections 6.5 and 7.

## 6 COMPARISONS TO LITERATURE

In this section, we show several comparisons to the literature. We also direct the reader to the work presented in M13, in which simulations were used to test and verify the accuracy of the fitting code. Although the flagging was not applied to these simulations, the simulations demonstrate the accuracy of the fitting algorithm, particularly for total magnitude and radius.

We describe the external catalogues used for comparison in Section 6.1. We compare the fits of SDSS, LG12, S11, and Mendel et al. (2014, hereafter Men14) to our fits where appropriate. Section 6.2 compares the deV fits to SDSS and LG12. Section 6.3 compares the Ser fits to S11 and LG12 Ser fits. Section 6.4 compares the deV-Exp fits to S11 and LG12 deV-Exp fits. Section 6.5 compares the Ser-Exp fits to S11 Ser-Exp fits. Section 6.6 presents a final comparison using the magnitudes measured using the preferred model of this work, Men14, and S11 to commonly used SDSS magnitudes (i.e. cModel and Petrosian).

### 6.1 External catalogues used for comparison and analysis

We use the fits of S11, SDSS DR7, and LG12 to make comparisons. S11 carried out fits of SDSS galaxies with Ser, deV-Exp, and Ser-Exp models using the GIM2D program (Simard et al. 2002). GIM2D uses the Metropolis search algorithm (Metropolis et al. 1953) to optimize the model parameters rather than a gradient descent algorithm similar to GALFIT. S11 also used SExtractor to mask out pixels dominated by neighbour galaxy light, which they refer to as ‘SEXTDEBL,’ or SExtractor deblending to separate the light of neighbouring objects from the target galaxy rather than the simultaneous fitting used in this work. They also chose a minimum number of pixels to use for sky estimation (20 000 pixels). The sky was fixed at this level during fitting rather than fitting the sky brightness as a free parameter. Fixing the sky was shown to provide the better fits than a fitted sky level for GIM2D (Häussler et al. 2007).

Additionally, S11 provide a statistical probability of a given galaxy being Ser, deV-Exp, or Ser-Exp based on an  $F$ -test. The  $F$ -test statistic is used to analyse if increasing the number of free parameters in the fit is statistically justified. This probability is used to select the preferred ‘best-fitting’ model from their data. S11 report two  $F$ -test probabilities, the  $P_{ps}$  probability, which is the probability that a one-component Ser model is preferred to a two-component deV-Exp model, and the  $P_{n4}$  probability, which is the probability that a two-component Ser-Exp model is preferred over a two-component deV-Exp model when attempting to explain the distribution of light in the observed galaxy.

Men14 re-analysed the S11 fits and classified the galaxies by a different method. Men14 used the radial light profile of the deV-Exp fit to separate galaxies into (1) bulge-dominated galaxies, (2) Exp dominated galaxies, (3) two-component galaxies, and (4) non-physical or unclear fits that do not fall into the previous categories.

LG12 provide an additional comparison to our data. The authors used a sample restricted to more nearby galaxies ( $0.003 < z < 0.05$ ) using SDSS DR8 (Aihara et al. 2011) data. The catalogue contains deV-Exp fits (referred to as ‘nb4’ in LG12), pseudo-bulge (exponential bulge + exponential disc, referred to as ‘nb1’ in LG12), deV (referred to as ‘dvc’ in LG12), Ser (referred to as ‘ser’ in LG12), and exponential disc or Exp (referred to as ‘exp’ in LG12) fits for 71 825 galaxies from the SDSS. We use the LG12 terminology for their fits throughout this paper to make it more clear which sample we are addressing.

In addition to fitting the models listed above, LG12 also give a classification of the ‘best-fitting’ model chosen from the five models they fit. LG12 assign a ‘best-fitting’ model using a combination of statistical and other metrics. For Exp models, statistical significance of the bulge and quality tests on the bulge magnitude, shape, and size is used to select Exp models rather than the Ser, nb1, or nb4 models. deV, or dvc, models are selected in a similar manner. Additional galaxies are selected as deV galaxies based on the colours and shapes of the disc in the nb4 fits. nb1 and nb4 galaxies are chosen from the remaining galaxies using the statistical significance of the bulge and quality tests on the fitting parameters intended to identify bad fits (i.e. tests on bulge ellipticity and bulge size relative to disc size). Remaining galaxies that do not satisfy any of these criteria are given the Ser model as the ‘best-fitting’ model.

LG12 report an absolute magnitude for each galaxy, which includes  $K$ -correction, extinction correction, and cosmological effects. While the assumed cosmology, the extinction correction, and the  $K$ -correction software used in this work are the same as those of LG12, the  $K$ -correction may be slightly different depending on the choice of input magnitudes (i.e. the Petrosian, SDSS model

magnitudes, or fitted magnitudes can be used to calculate a  $K$ -correction). Also the zero-point of the magnitudes may vary from the values used here due to the small calibration differences between SDSS DR7 and DR8. Differences in the zero-point calibration are expected because the calibration procedure (the ‘Ubercal’ algorithm; Padmanabhan et al. 2008) is a global algorithm, using all of the imaging data to determine the overall calibration rather than just a single frame. Since the volume of imaging data increased between DR7 and DR8, this can cause slight differences in the calibration.

The overlap of LG12 and this work contain galaxies at the low-redshift end of the galaxy distribution in our catalogue. These galaxies generally have better resolution and are brighter than the full sample of our catalogue, so agreement between LG12 and this work only provide a lower bound on the bias and scatter of our full catalogue. However, this comparison provides a test of the most optimal fitting conditions where resolution effects are less of a concern.

As in the last section, we bin all the plots presented here by Petrosian magnitude in order to make consistent comparisons across all the works. The Petrosian magnitude provides a model-independent measure of the magnitude so that any comparison has the same distribution of galaxies.

For the analysis presented in this section, we treat the ‘best-fitting’ models given by S11 and LG12 as the most appropriate models to use in comparison. Therefore, we compare our fits to the best fits of the external works (e.g. we compare our deV-Exp fits to galaxies identified as deV-Exp by S11 or to galaxies identified as ‘nb4’ by LG12). This will reduce the bias introduced by fitting an incorrect model to the galaxy since we will be comparing the fits with the highest confidence of being correct.

Sections 6.6 and 7 compare the S11, Men14, and LG12 classifications to the choice of best model based on the flagging presented in Section 4. In these sections, we examine the agreement between the various ‘best-fitting’ models of these different works.

## 6.2 The deV fits

The SDSS pipeline computes PSF-convolved 2D deV fits in addition to other parametric and non-parametric measurements. The SDSS fits are truncated at  $7 r_{\text{eff}}$  to go smoothly to zero beyond  $8 r_{\text{eff}}$  and also employ some softening of the profile within  $r = r_{\text{eff}}/50$  (Stoughton et al. 2002). No truncation of the profile is imposed upon the fits presented in this paper. For a deV profile, 94 per cent of the light is contained within  $8 r_{\text{eff}}$  and 93 per cent of the light is contained within  $7 r_{\text{eff}}$ . So, an offset between the deV model fit by SDSS and the deV fit presented here is expected. If different pipelines recover the same fitting parameters for sky, radius, axis ratio, etc. the expected offset due to the profile truncation is 0.0716 mag, assuming that half the light between 7 and  $8 r_{\text{eff}}$ , where the softening takes place, is also truncated.

For purposes of a fair comparison, we correct SDSS magnitudes by making the deV magnitudes of SDSS brighter by 0.07 mag. LG12 also truncate the deV profile following the same prescription as SDSS. However, LG12 do not soften the centre of the profile within  $r = r_{\text{eff}}/50$ . A similar offset is expected for the LG12 ‘dvc’ fits, so we apply the same correction of 0.07 mag. No modification is made to the radii of the fits or any other fitting parameters.

Fig. 14 shows a comparison of fitted deV magnitude from this work, SDSS, and LG12 as a function of apparent Petrosian magnitude (top row) and absolute Petrosian magnitude (bottom). SDSS deV fits and PYMORPH deV fits are compared for the 21 7131 galaxies in SDSS that are best fit by the deV profile as identified by `fracdev_r >= 0.8`. LG12 comparisons use the  $\sim 9000$  galaxies

identified as ‘dvc’ galaxies using the ‘best model’ parameter of LG12. We also remove all galaxies with flag bit 20 set (i.e. galaxies with fitting problems). This removes 2.82 per cent of the deV sample.

Magnitude comparisons are shown as the first model minus the second model. For example, the top-left panel compares PYMORPH with SDSS, so magnitude differences are quoted as PYMORPH – SDSS and negative values indicate that PYMORPH produces a brighter magnitude. Median values for magnitude bins are plotted in red with the error bars representing a 95 per cent bootstrap CI on the median. The 68 per cent contours of the data are plotted as blue dashed lines. The density of points is plotted in grey-scale with the low end of the density representing 1 per cent of the maximum density. Bins with a density below the minimum density are plotted in white.

In the left-hand column of Fig. 14, PYMORPH deV magnitudes agree with SDSS with a scatter of 0.05 (using 68 per cent contours) across most of the magnitude range but show a systematic increase to about 0.07 mag brighter than SDSS values at the bright end of the apparent magnitude range (top row, left-hand column). While SDSS and this work differ at the bright end, the majority of the galaxies show no systematic bias. In contrast, LG12 have an offset of  $-0.03$  to  $-0.05$  across the entire magnitude range (centre column of Fig. 14). LG12 comment on this systematic difference in their paper noting a  $-0.025$  mag offset in their deV fits to SDSS galaxies with `fracdev_r > 0.5` as well as a  $\sim 9$  per cent difference in the fitted radii. We investigated the source of this offset, but can find no reason for it. Although the zero-point of DR8 and DR7 vary up to 0.2 mag, correcting for this effect does not reduce the offset.

Median differences in the measured radii are below 5 per cent [see Appendix C1 (available online) for the radius comparisons] when comparing LG12 and this work. The difference in radii agrees with the observed magnitude differences, suggesting that the PYMORPH fits presented here are larger and brighter compared to the LG12 fits.

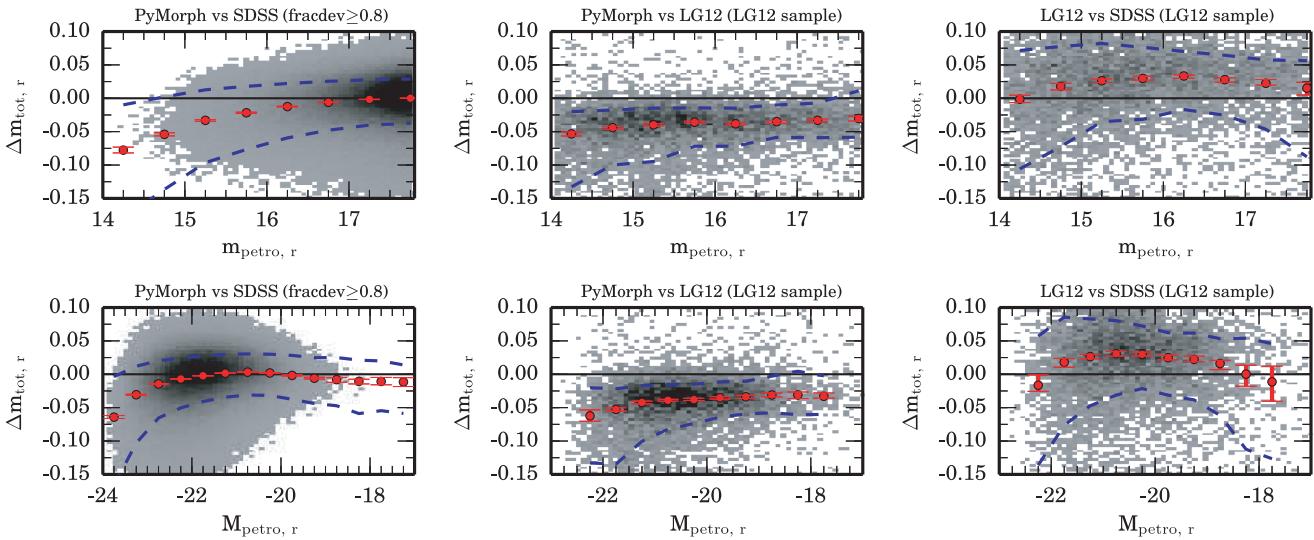
The radii in SDSS are 5–10 per cent smaller when compared to PYMORPH. LG12 comment on this difference as well, attributing it to effects of the softening of the fitted profile in the centre (inside  $r_{\text{eff}}/50$ ) which suppresses the half-light radius in SDSS. This cannot, however, be the source of the disagreement between LG12 and this work since neither work implements such softening.

The level of the sky in our fitting is found to be, on average, 0.25 per cent dimmer when compared to SDSS (see also Fig. 16 and related discussion). Sky level and the fitted magnitude have been shown to be correlated (see M13, and references therein for a full discussion). Bias in the sky level may explain the slight differences in magnitudes. In M13, overestimates of sky at the level of 0.5 per cent are shown to suppress fitted Sérsic magnitude by  $\approx 0.1$  mag. We discuss the sky brightness further in Section 6.3.

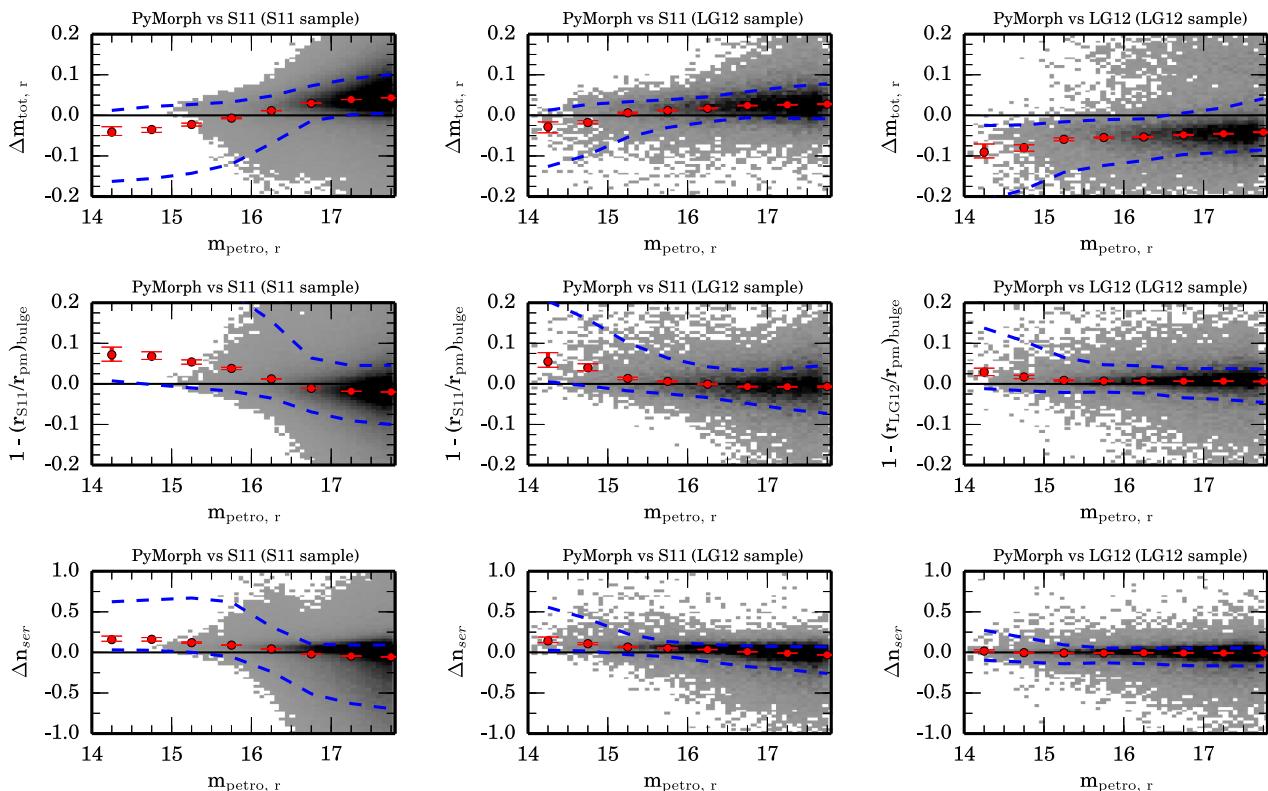
## 6.3 The Sérsic fits

We also compare our Sérsic fits to those of LG12 and S11. Fig. 15 shows the difference between magnitude (first row), half-light radius (second row), and Sérsic index (third row). We compare PYMORPH and S11 for the full sample (left-hand column), PYMORPH and S11 for the LG12 sample (centre column), and PYMORPH and LG12 (right-hand column). Plots are in a format similar to that of Fig. 14. For the Sérsic fits we also examine the differences in Sérsic index,  $n_{\text{Sérsic}}$ . For the full comparison of the Sérsic fits see Appendix C2 (available online).

For the plots presented in Fig. 15, we select galaxies in both S11 and LG12 that are ‘best fit’ by a Sérsic model according to S11. We



**Figure 14.** A comparison of fitted deV magnitude from this work, SDSS, and **LG12** as a function of apparent Petrosian magnitude (top row) and absolute Petrosian magnitude (bottom). SDSS deV fits and PYMORPH deV fits are compared for the 217 131 galaxies in SDSS that are best fitted by the deV profile as identified by  $\text{fracdev}_r \geq 0.8$ . **LG12** comparisons use the ~9000 galaxies identified as ‘dvc’ galaxies using the ‘best model’ parameter of **LG12**. We also remove all poorly fitted galaxies with flag bit 20 set (i.e. galaxies with fitting problems). This removes 2.82 per cent of the deV sample. Magnitude comparisons are shown as the first model – the second model. For example, the top-left panel compares PYMORPH with SDSS, so magnitude differences are quoted as PYMORPH – SDSS and negative values indicate that PYMORPH produces a brighter magnitude. Median values for magnitude bins are plotted in red with the error bars representing a 95 per cent bootstrap CI on the median. The 68 per cent contours of the data are plotted as blue dashed lines. The density of points is plotted in grey-scale with the low end of the density representing 1 per cent of the maximum density. Bins with a density below the minimum density are plotted in white. SDSS and PYMORPH agree within 0.05 across the magnitude range, except at the bright end where we are consistently brighter than SDSS. **LG12** exhibit an offset to both SDSS and our fits of up to 0.05. See Appendix C1 (available online) for additional comparisons.



**Figure 15.** The difference between the Ser magnitude (first row), radius (second row), and Sérsic index (third row). The left-hand column shows the comparison of this work and **S11** for the full sample. The centre column shows a comparison of this work and **S11** for the galaxies of the **LG12**, low-redshift sample. The right-hand column show comparisons for this work and **LG12**. **S11** and **LG12** galaxies are identified as best fit ‘ser’ models in **S11**. The format of the plot is the same as Fig. 14. The **LG12** fits exhibit an offset in magnitude similar to the offset seen in the deV fits. Differences between this work and **LG12** for bright galaxies are reduced in comparison to the **S11**–PYMORPH comparison. See Appendix C2 (available online) for additional comparisons.

use the **S11** classification in this case because **LG12** use the **Ser** category as a default category for fits that are not well described by the other models they fit. The authors comment that the **LG12 Ser** galaxies are primarily low Sérsic index. **S11** sample a broader range of Sérsic indices because they do not have separate **deV** or **Exp** categories. **LG12** do not mention any truncation of the **Ser** profile, so we do not apply a correction here. The top-left panel of Fig. 15 shows that the **LG12** magnitudes are offset again relative to PYMORPH. Comparison with **S11** (middle row, left) also show a zero-point offset, but in the opposite direction (i.e. **S11** is brighter than this work while **LG12** is fainter). In addition, there is an overall trend in **S11** across the magnitude range.

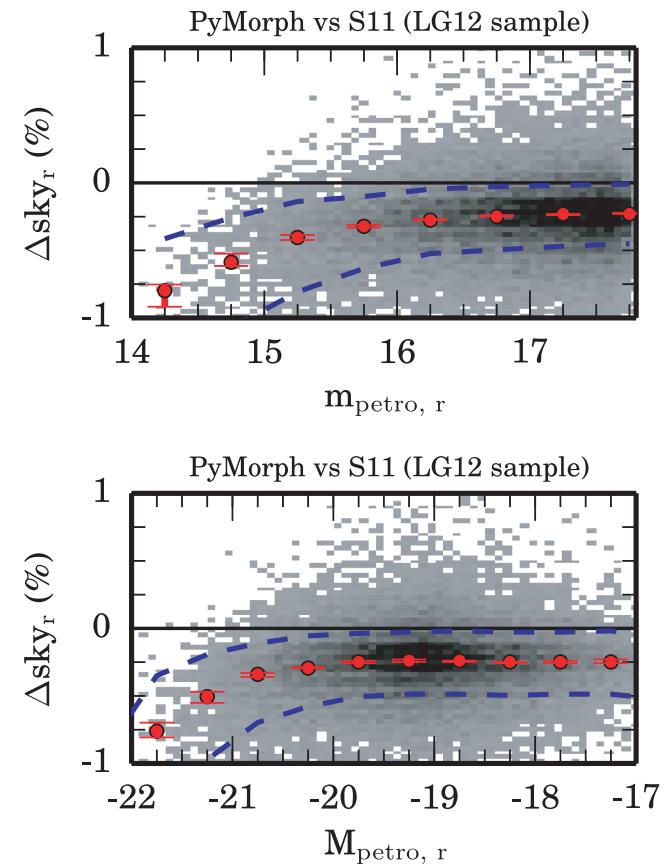
The radii of **S11** are also smaller at the bright end of the apparent magnitude range when compared to this work (bottom row, centre column). This trend is reduced in the **LG12** comparison (top row, centre). The Sérsic index also displays a trend in the **S11** data (bottom row, right) that does not appear in the comparison of **LG12** and this work (top row, right). Although this trend appears to be insignificant, there are larger biases in the Sérsic index when compared as a function of absolute magnitude (see Appendix C2, available online). This trend was also discussed in Bernardi et al. (2014), where the authors argue that the behaviour of the Sérsic index is consistent with systematic errors in the Sérsic index due to oversubtracting sky.

We used the **S11** sky values (Simard, private communication) to examine the effect of the sky on the **Ser** fits. PYMORPH prefers sky levels about 0.25 per cent lower than that of **S11**. **S11** sky brightness is similar to SDSS. Fig. 16 shows a comparison of **S11** and PYMORPH sky levels as a function of apparent (top) and absolute (bottom) Petrosian magnitude. There is an offset in sky level of approximately 0.25 per cent, similar to the SDSS. However, this difference increases in the **Ser** fits with PYMORPH sky levels appearing up to 1 per cent dimmer in the brightest apparent and absolute magnitude bin.

The simulations in M13 showed that PYMORPH estimates the sky with a bias  $\sim 0.1$  per cent which is a factor of  $\approx 2$  smaller than the observed difference seen for the SDSS measurements. Furthermore, the underestimate of sky brightness observed in M13 was not large enough to cause a measurable bias in recovered magnitude for the simulations. Since we expect an offset of 0.1 per cent in sky brightness if SDSS sky measurements are accurate, the observed offset of 0.25 per cent indicates that SDSS sky levels are likely slight overestimates (about 0.15 per cent of sky brightness). Since **S11** sky levels have a similar offset, we expect that they are overestimates as well. Although Häussler et al. (2007) argue that such an overestimate of sky improves GIM2D performance, offsetting the inefficient masking of neighbouring galaxies.

We propose a simple explanation for this effect in the **S11** fits. **S11** used at least 20 000 pixels nearest to each galaxy to estimate the sky (the nearest 20 000 pixels that are classified as neither source, nor neighbour pixels). This is, in general, a large number of pixels that sample the sky at many different radii. However, as the size of the target galaxy grows, the annuli that form the perimeter of the galaxy grow as well. This leads to a systematic sampling of the sky for the brighter and larger galaxies. For the extended objects studied here, this will lead to an overestimate of the sky and, as a consequence, a suppression of galaxy size and brightness. Using an image cutout that instead scales with the galaxy radius, as is used here, ensures that the same range of half-light radii are sampled for sky estimation and prevents this systematic effect.

We note that two works have already shown that our measurements should be preferred to those of **S11**. First, Kravtsov et al.

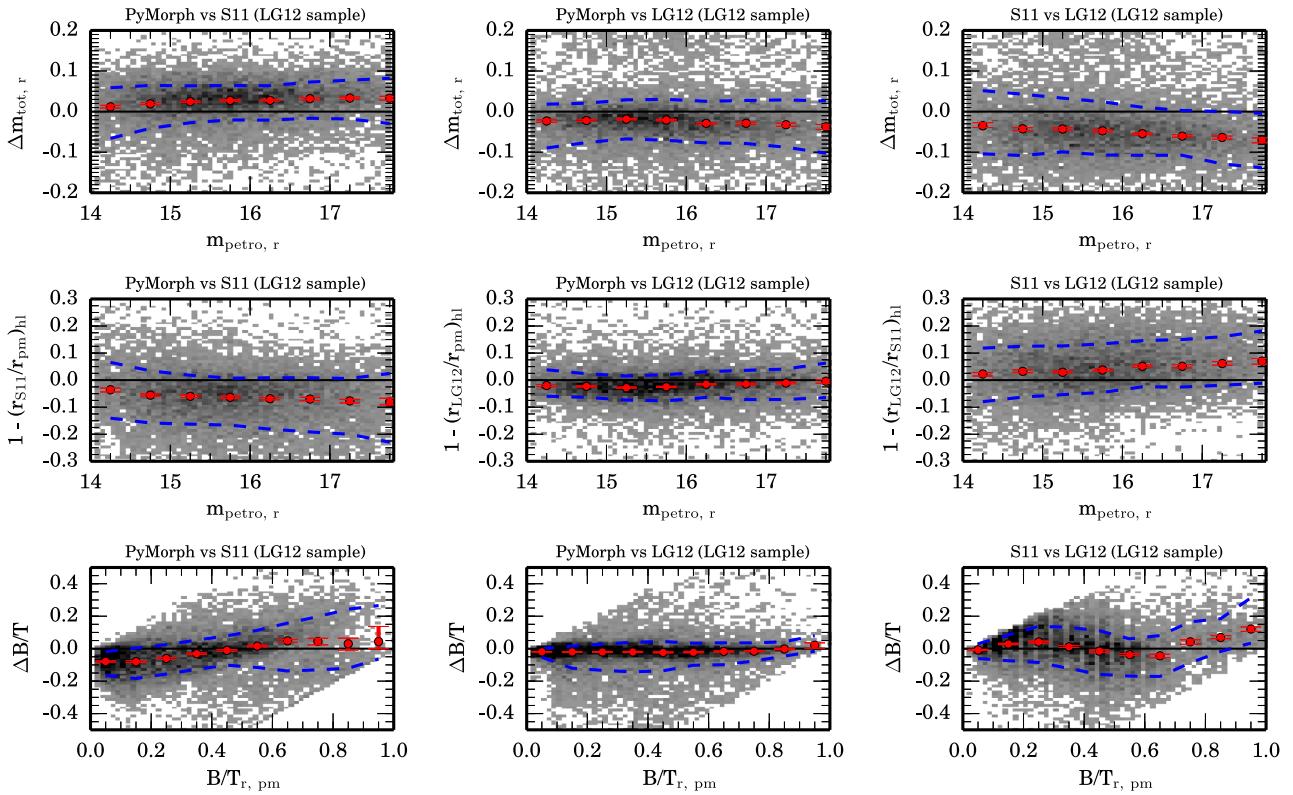


**Figure 16.** The difference in the  $r$ -band sky values of **S11** **Ser** fits and our PYMORPH **Ser** fits shown as PYMORPH – **S11** (negative values indicate that PYMORPH produces a dimmer sky measurement compared to **S11**). The top row shows the difference between the **S11** and PYMORPH **Ser** sky brightness as a function of the apparent Petrosian magnitude. The second row shows the difference in sky versus absolute Petrosian magnitude. An offset in sky brightness is observed. This difference increases with increasing brightness, suggesting that the **S11** sky values are systematically biased brighter for brighter and larger galaxies in both the observed and absolute frame, causing a systematic bias in the  $r$ -band magnitudes of **S11**.

(2014) performed detailed fits of approximately 10 BCGs and found that their measurements agree more with our measurements than with the measurements of **S11**. Secondly, we direct the reader to appendix of Bernardi et al. (2014) which shows an unexpected redshift evolution of the Sérsic index in the **S11** data. The trend in Sérsic index is not observed in the fits of this work. Bernardi et al. (2014) argue that this is another reason to prefer the fits of this work to those of **S11**. Therefore, we recommend that caution should be exercised when using the **Ser** fits from different works to properly account for such differences in fitting.

#### 6.4 The **deV**–**Exp** fits

The **deV**–**Exp** model is the final model that allows for a direct comparison of both the **S11** and **LG12** fits with our own. We see broad agreement between PYMORPH, **S11**, and **LG12** for the total half-light radius and total magnitude. Fig. 17 shows comparisons of total magnitude (top row), half-light radius (middle row), and  $B/T$  (bottom row) for this work compared to **S11** (left-hand column), this work compared to **LG12** (centre column) and **S11** compared to **LG12** (right-hand column). The PYMORPH and **LG12** magnitude and



**Figure 17.** The difference between the PYMORPH deV-Exp magnitude, half-light radius, and  $B/T$  and those of LG12 deV-Exp and S11 deV-Exp fits for galaxies in the LG12 sample. The top row shows the comparison of magnitude between S11 and PYMORPH (left), PYMORPH and LG12 (centre), and S11 and LG12 (right). The second row shows a comparison of the total half-light radius as a function of apparent magnitude. The third row shows a comparison of  $B/T$  as a function of PYMORPH  $B/T$ . S11 and LG12 galaxies are identified as best-fitting ‘nb4’ models in LG12. The format of the plot is the same as Fig. 14. S11  $B/T$  has larger scatter with respect to either PYMORPH or LG12. See Appendix C3 (available online) for additional comparisons.

half-light radius agree with smaller scatter than either comparison to S11.  $B/T$  values of LG12 and this work are also in better agreement (bottom row). S11 tend to overestimate  $B/T$  relative to this work by about 0.1 at the low  $B/T$  while the LG12 fits are in closer agreement to this work across the  $B/T$  range.

The scatter in the subcomponents is quite broad making any conclusions difficult. When comparing this work to S11, bulge magnitudes have a scatter of  $\sim 0.3$  mag, bulges tend to be dimmer in this work compared to S11 by as much as 0.1 mag, and bulges are larger by almost 10 per cent. There is better agreement in the disc magnitudes, but PYMORPH discs tend to be 5 per cent larger when compared to S11. The scatter decreases when comparing this work with LG12, although a difference in bulge and disc parameters is still observed. We refer the reader to Appendix C3 (available online) which shows the full comparison of the subcomponents.

## 6.5 The Ser-Exp fits

While S11 provide a Ser-Exp model fits, they expect less than 10 per cent of galaxies to support such a model according to the  $\chi^2$  arguments in their analysis. Also, they apply a prior on the Sérsic index of the bulge during fitting that favours a traditional  $n = 4$  bulge. LG12 do not attempt to fit such a model to the low- $z$  sample. As a result, our ability to compare the Ser-Exp fits is somewhat limited.

Fig. 18 shows a comparison of the total magnitude, radius and  $B/T$  values for galaxies selected as Ser-Exp according to S11. These galaxies have  $F$ -test probabilities below 0.32 for both the  $P_{\text{ps}}$  and  $P_{n4}$

probabilities, indicating that the Ser-Exp fit is significantly better than either the deV-Exp or Ser fits. The additional comparisons of the Ser-Exp model to S11 are presented in Section C4 for completeness.

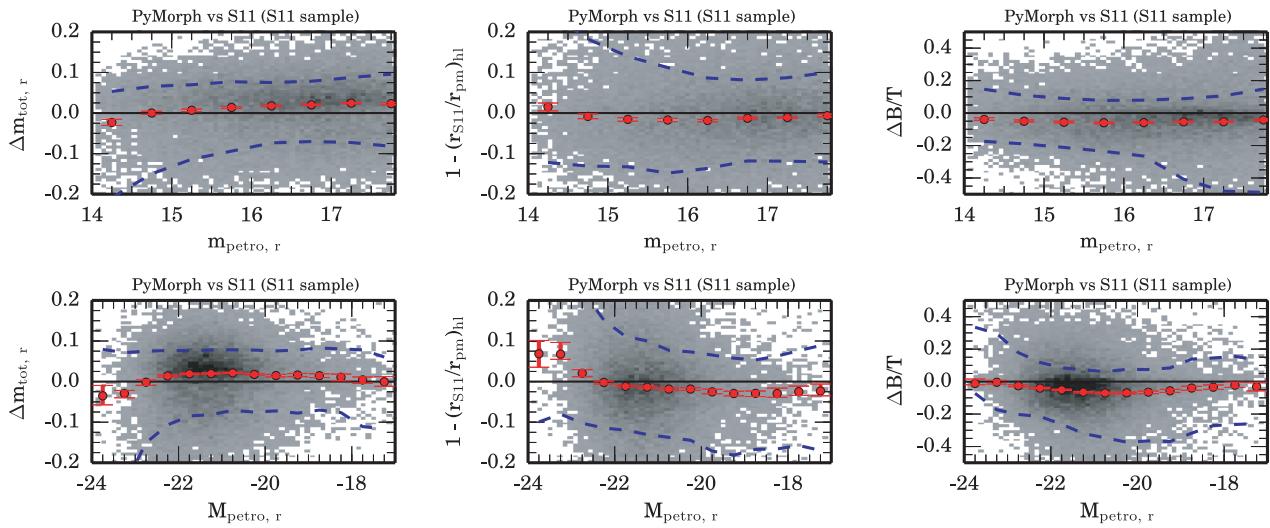
The total values are in agreement with wider scatter compared to that seen in Figs 15 and 17. However, the bright end of both the apparent magnitude and absolute magnitude distributions show trends similar to the Ser fits. The S11 fits use the same fixed sky level for all fitted models. Since we see evidence of bias in the sky level for the Ser galaxies of S11, it is likely that the same problems exist in the Ser-Exp fits. The components of the Ser-Exp fits have wider scatter than the analogous parameters in the deV-Exp fits.

More on the general comparisons with the S11 measurements is presented in Section 7 which support the accuracy of our Ser-Exp fits.

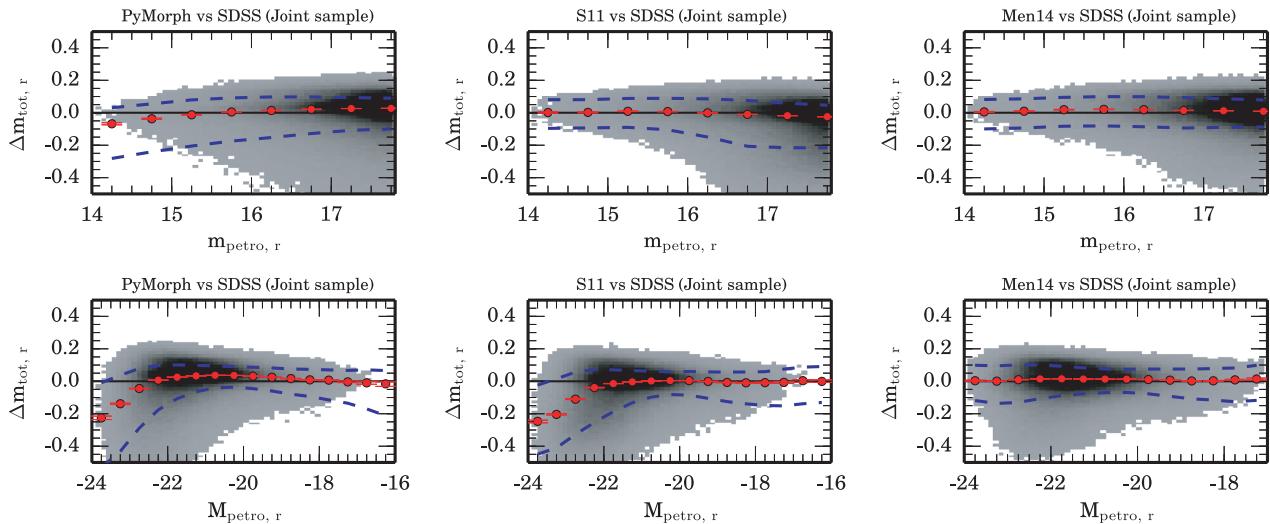
## 6.6 Comparing the preferred models

We conclude this section by comparing the ‘best model’ magnitudes of this work, S11, and the Men14 classification of S11 fits to commonly used SDSS magnitudes. Bernardi et al. (2013) showed that the choice of fitted model can substantially alter the bright end of the LF. Here, we examine whether these differences can be eliminated by the use of the ‘best-fitting’ magnitude rather than using magnitudes measured using only a single fitted model.

For the fits of this work, we construct a ‘best fit’ from the combination of the Ser and Ser-Exp catalogue as follows: first, galaxies with flag bit 10 or 14 set (galaxies shown as ‘2com’ and ‘prob



**Figure 18.** The difference between the PYMORPH Ser-Exp magnitude, half-light radius, and  $B/T$  for **S11** Ser-Exp fits in the full catalogue. The top row shows the comparison of magnitude (left), half-light radius (centre), and  $B/T$  (right) between **S11** and PYMORPH as a function of apparent Petrosian magnitude. The second row shows the same comparisons as a function of absolute Petrosian magnitude. Galaxies identified as best fit by Ser-Exp models in **S11** are used in this plot. The format of the plot is the same as Fig. 14. The Ser-Exp fits have wider scatter than the analogous parameters in the deV-Exp fits. See Appendix C4 (available online) for additional comparisons.



**Figure 19.** The difference between SDSS cModel magnitude and the magnitude fit by PYMORPH, **S11**, and **Men14** for the galaxies appearing in all three catalogues. Galaxies classified as an unknown profile in **Men14** (`Proftype=4`) are excluded from the plots. This removes approximately 10 per cent of our original catalogue. We also exclude any failed fits from our catalogue.

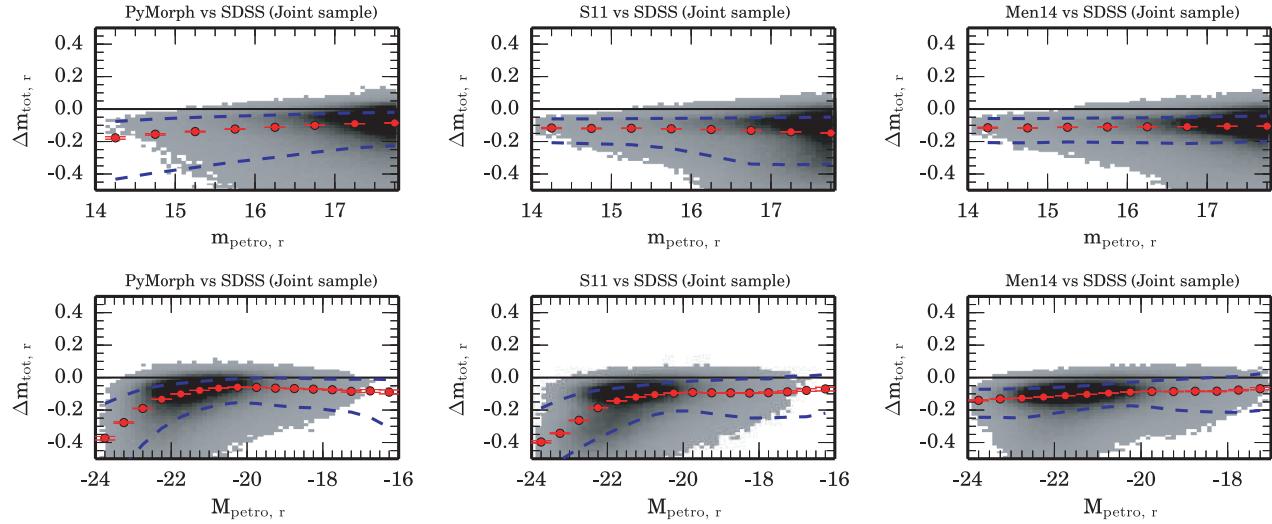
2com' in Fig. 21) are given the total magnitude measured by fitting the Ser-Exp model. Galaxies with flag bits 1 or 4 set (i.e. 'bulge' or pure 'disc' in Fig. 21) are given the total magnitude of the Ser fit. Finally, any galaxies with flag 20 set (i.e. 'bad' fits) for the Ser-Exp fit are given the fitted Ser magnitude if the Ser fit does not have flag bit 20 set.

For **S11** fits, we assign fitted Ser magnitudes to galaxies with  $F$ -test probabilities  $P_{ps} \geq 0.32$ . deV-Exp fit magnitudes are used for galaxies with  $F$ -test probabilities  $P_{ps} < 0.32$  and  $P_{n4} \geq 0.32$ . Galaxies with  $F$ -test probabilities  $P_{ps} < 0.32$  and  $P_{n4} < 0.32$  are given the reported Ser-Exp magnitudes.

For the **Men14** classification of **S11** fits, galaxies are assigned Ser magnitudes if they have `Proftype=1` or `2` (i.e. the bulge category or disc category, respectively) in the **Men14** data. Galaxies are assigned the deV-Exp fit magnitude if they have `Proftype=3`

(i.e. the two-component category). `Proftype=4` (i.e. the problematic fits) are ignored. After assigning the galaxies to these categories, we select only those galaxies present in all three data sets for comparison.

Fig. 19 shows the comparison of PYMORPH, **S11**, and **Men14** to the SDSS cModel magnitudes for the full sample. The cModel magnitudes (defined in Bernardi et al. 2010) are calculated from the linear combination of the independently fit SDSS deV and Exp models that best fits the galaxy. The parameter that sets the fraction of flux contributed by the deV fit is defined as `fracdev`. A `fracdev=1` galaxy is best fit by the deV model, while a `fracdev=0` galaxy is best fit by the Exp model. The `fracdev` parameter provides a very crude estimate similar to the  $B/T$  measured during simultaneous fitting. The cModel magnitudes are corrected for the offset due to profile truncation as was previously done in Fig. 14.



**Figure 20.** The difference between SDSS Petrosian magnitude and the magnitude fit by PYMORPH, S11, and Men14 for the galaxies appearing in all three catalogues. Galaxies classified as an unknown profile in Men14 (`Proftype=4`) are excluded from the plots. This removes approximately 10 per cent of our original catalogue. We also exclude any failed fits from our catalogue.

The difference between cModel magnitudes and S11 or this work increases at brighter magnitudes. This effect is not present using the Men14 selection of S11 (right-hand column). Bernardi et al. (2013) showed that the systematic effects on the bright end of the luminosity and stellar mass functions can cause a substantial underestimate of the bright end of the LF. The lowest estimates of the bright end of the LF occur in the SDSS Petrosian and cModel measurements. Since Men14 agree more closely with SDSS at the bright end, this suggests that Men14 are selecting models that underestimate the brightest galaxies by a substantial amount.

The large difference between Men14 and S11 in the brightest magnitude bins may be the result of Men14 strongly favouring the two-component deV-Exp fit for all galaxies as opposed to the preference of this work and S11 for single-component Ser fits at these magnitudes. The brightest, largest galaxies tend to have higher Sérsic index, which requires more light in the wings of the profile in order to account for the concentration observed at the centre of the galaxy. Similar concentrations can be reproduced with two lower concentration components (deV bulge and Exp disc), but when integrated to infinity, these two-component galaxies have less total light compared to the broader Ser profile. Fitting of a two-component deV-Exp model appears to produce a dimmer estimate of the total magnitude relative to the Ser model.

Fig. 20 shows the similar comparison of PYMORPH, S11, and Men14 to the SDSS Petrosian magnitudes. Differences approaching 0.5 mag in the brightest bins are observed in S11 and this work. These differences will increase the number of bright galaxies in the brightest bins of the LF. While Bernardi et al. (2013) only explored cases of pure fits (i.e. fitting the entire sample with Ser fits only or fitting the entire sample with Ser-Exp fits only), the large differences in the LF reported there likely persist based on the differences observed here where the ‘best-fitting’ profile is used.

## 7 COMPARISONS USING MORPHOLOGICAL INFORMATION

Finally, we examine our selection of bulge, disc and two-component galaxies using morphological information from other catalogues.

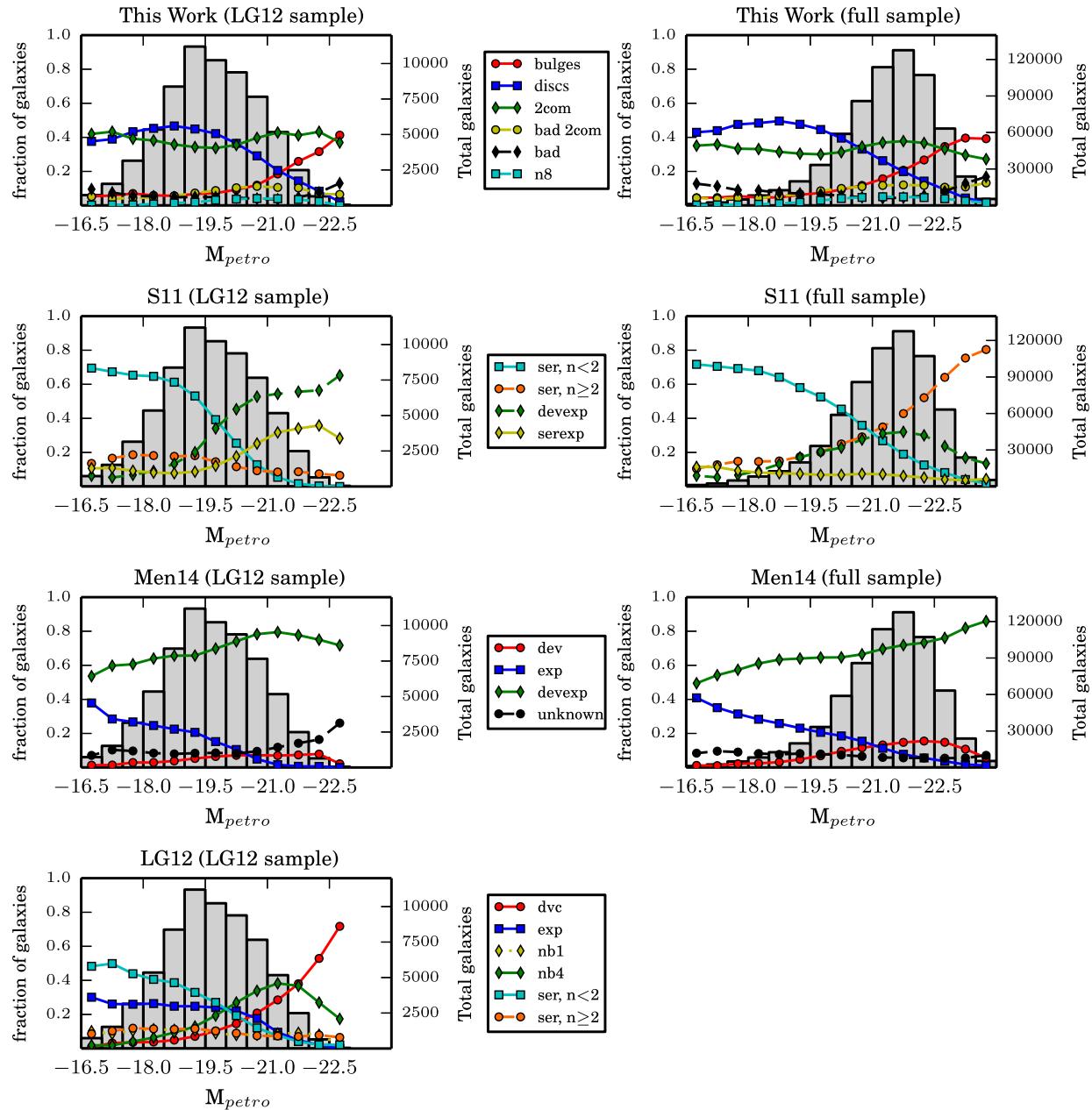
The work of Nair & Abraham (2010) gives a detailed visual morphological classification of a set of about 10 000 nearby galaxies. The GZ2 project (Lintott et al. 2008, 2011) provides another morphological classification of nearly half of our sample. These catalogues give morphological classifications not dependent on fitting and provide another test of our profile-based flagging. Section 7.1 compares the preferred models of LG12, S11 and Men14 to our selection with respect to magnitude and *T*-types calculated using the work of Nair & Abraham (2010). Section 7.2 examines a few internal checks, including *B/T* and axis ratios separated by *T*-type. Section 7.3 discusses possible effects of bars on the fitting based on barred galaxies identified in GZ2.

### 7.1 Preferred models as an indicator of morphological classification

#### 7.1.1 Magnitude distribution of preferred models

In this section, we compare our selection of bulge, disc, and two-component galaxies presented in Section 4 with S11, LG12, and Men14, all of which include a ‘preferred’ model in their catalogues. Fig. 21 shows the per cent of model types as a function of absolute Petrosian magnitude for two samples. The left-hand column shows the results for the low-*z* sample of LG12. The right-hand column shows the results for the entire spectroscopic catalogue used in this work. The panels show the models of this work (top row), S11 (second row), Men14 (third row), and LG12 (fourth row). Model types for each group are plotted as lines with symbols. The fraction of galaxies reported is the fraction within the respective bin. For each panel, the sum of all models in any given bin is 1. The total distribution of galaxies as a function of magnitude is plotted in the background histogram and the histogram scale is on the right-hand axis of the plot.

The fits of this work show expected trends with almost no bulges (only a few per cent) and a mixture of disc and two-component galaxies at low magnitudes (below about −19.5). Two-component fits are the dominant model between −20 and −22 and bulges dominate at magnitudes brighter than −22. This behaviour is visible in both samples and shows that the flag-based model



**Figure 21.** Type comparisons as a function of absolute Petrosian magnitude for this work (top row), S11 (second row), Men14 (third row), and LG12 (fourth row). The left-hand column shows the classifications for the smaller, low- $z$  LG12 sample. The right-hand column shows the distribution for the full sample presented in this work. Model types for each group are plotted. The distribution of galaxies as a function of magnitude is plotted in the background histogram and the histogram scale is on the right-hand axis of the plot. The fraction of galaxies reported is the fraction within the respective bin. For each panel, the sum of all models in any given bin is 1. Men14 are largely insensitive to the absolute magnitude of the galaxy. S11 (second row) have a large shift from fitting galaxies brighter than -21 with two-component models in the low- $z$  sample to predominantly fitting Ser models to the same galaxies in the full sample. This indicates that S11 are more sensitive to resolution than to the actual morphology of the galaxy. LG12 find much higher proportions of one-component models at the low and high magnitudes compared to this work ( $\sim 80$  per cent for LG12 compared to  $\sim 40$  per cent for this work).

selection is reasonably independent of apparent magnitude and resolution effects. Such effects would likely cause a different appearance in the model selection for the low- $z$  LG12 catalogue (left-hand column) when compared to the full sample (right-hand column).

In contrast, S11 (second row) have a large shift from fitting galaxies brighter than -21 with two-component models in the low- $z$  sample (about 80 per cent of galaxies brighter than -21 in the LG12 sample are two component) to predominantly fitting Ser models to

the galaxies with similar absolute magnitude in the full sample. This indicates that S11 are potentially more sensitive to observational effects (e.g. resolution or S/N) than to the actual morphology of the galaxy. However, the general behaviour in the full sample is similar to the fits presented in this work. The single-component, disc-like fits dominate at faint magnitudes. Two-component fits are more prevalent at intermediate magnitudes (in the neighbourhood of -21) and single-component bulges make up a large fraction of galaxies at the brightest magnitudes.

[Men14](#) (third row) are less sensitive to the absolute magnitude of the galaxy, with high percentages (between 60 and 80 per cent) of [deV](#)-Exp galaxies across the entire magnitude range in the [LG12](#) sample. [Men14](#) do identify slightly more discs in the full sample, but the [deV](#)-Exp model is still dominant across the magnitude range.

[LG12](#) (fourth row) cannot be compared to the full sample. In the [LG12](#) sample, the behaviour of both [LG12](#) and this work is similar over the range of magnitudes examined here. Bulges dominate at the brightest magnitudes, discs are much more frequent at dimmer magnitudes, and two-component fits dominate in the region surrounding  $-20$  mag. However, [LG12](#) find higher proportions of one-component models at the low and high magnitudes compared to this work ( $\sim 80$  per cent for [LG12](#) compared to  $\sim 40$  per cent for this work). The authors discuss in their paper that an initial identification of bulges using fitting parameters and  $\chi^2_\nu$  statistics does not produce a high enough percentage of bulges, so many galaxies are chosen to be ‘dvc’ (i.e. [deV](#) fits) based on colour information rather than the fitting parameters. Many of these galaxies have large diffuse components that are fitted by the Exp component of the two-component models.

As was briefly mentioned in Section 3.2, we find similar behaviour in our fits (the incidence of two-component fits is higher than expected at the brightest magnitudes). We choose not to force these galaxies to be fit by a single component in order to avoid the magnitude bias reported in [M13](#). Instead, we choose to relax our definition of the fitted models and caution the user that a significant Exp component may be an indication of an extended halo component rather than a classical disc in early-type galaxies.

### 7.1.2 Model selection as a function of $T$ -type

We assign a  $T$ -type to each galaxy in the catalogue using the type probabilities (Ell, S0, Sab, and Scd) provided by the BAC ([H2011](#)). The BAC of [H2011](#) used a Bayesian approach to assign probabilities of being one of four broad galaxy types (Ell, S0, Sab, or Scd) to each galaxy using colour, total axis ratio, and concentration as measured by the SDSS. An SVM algorithm was used to produce the probabilities, using the Fukugita et al. (2007) sample as the training set.

We calibrate the probabilities from [H2011](#) to the  $T$ -types using a simple linear model

$$T = -4.6 \times P(\text{Ell}) - 2.4 \times P(\text{S0}) + 2.5 \times P(\text{Sab}) + 6.1 \times P(\text{Scd}). \quad (7)$$

The coefficients of the equation are calibrated to the visually classified galaxies of Nair & Abraham (2010) by an unweighted linear regression. A similar comparison was used in Willett et al. (2013), although a symbolic regression was used rather than a linear regression. In Willett et al. (2013), the result of the symbolic regression was a linear model in parameters relating only to the dominance of the bulge. The regression fit in this work estimates  $T$ -types between  $-5$  and  $4$  with median bias of  $0$  in  $T$ -type and  $68$  per cent of the estimates within  $\pm 2$  in  $T$ -type. Although the scatter is relatively broad, we can reliably separate early and late types (the difference in  $T$ -type for these galaxies is more than  $4$ ). Since the Nair catalogue is quite small (only about  $10\,000$  galaxies from this work are present in Nair), we use this extension to estimate  $T$ -types for the entire sample common to [S11](#), [LG12](#), and this work. Since the parameters used in the [H2011](#) model are not specific to any particular fitting method or fitting code, we expect them to be unbiased across the studies examined here.

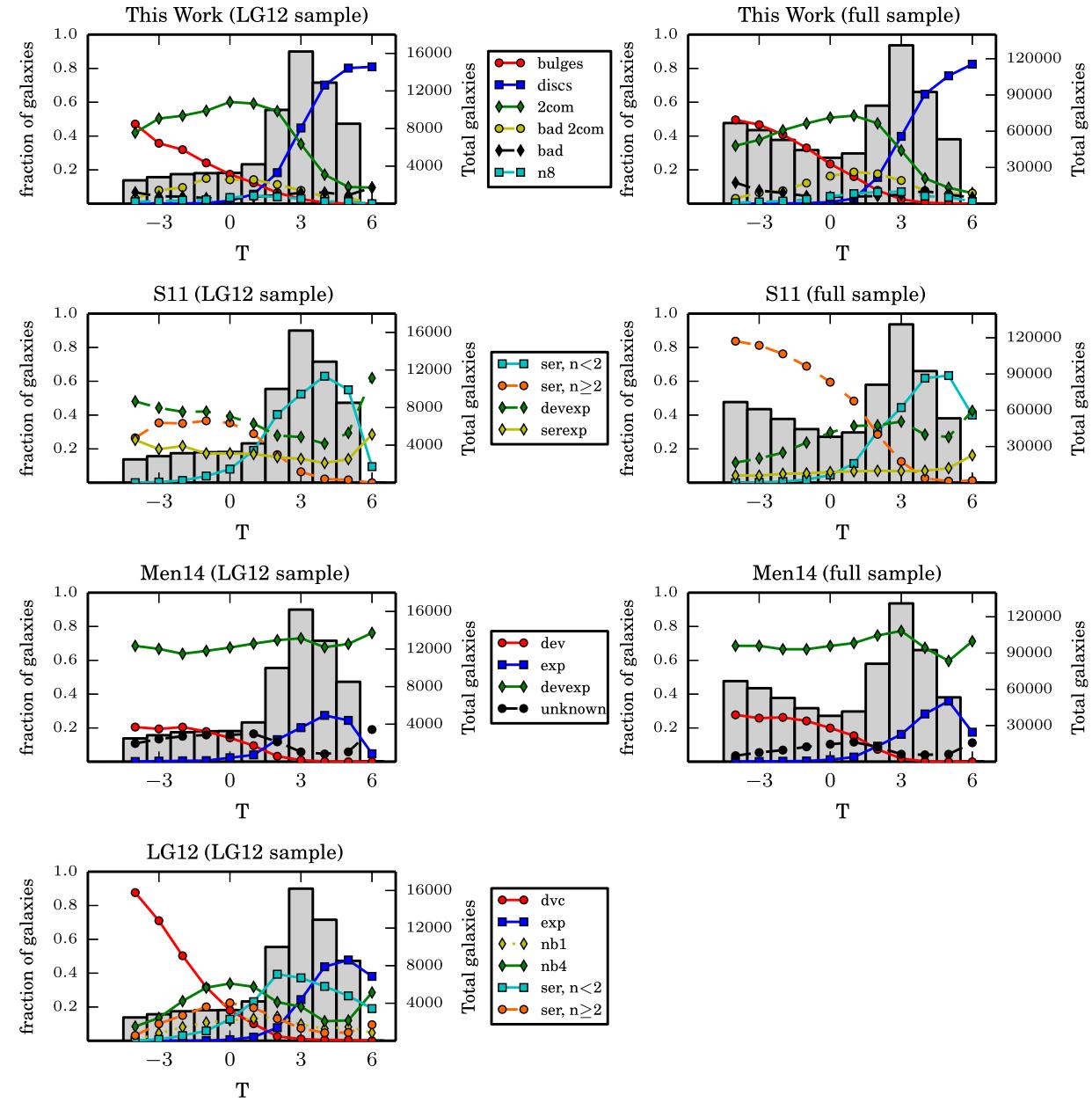
Fig. 22 shows the distributions of the preferred models of [Men14](#), [S11](#), [LG12](#), and our Ser-Exp categories of bulge, disc, and two-component fits as a function of  $T$ -type for galaxies in the [LG12](#) sample. Each plot shows the percentage of galaxies for each preferred model as a function of  $T$ -type. Ser models are divided into two categories,  $n < 2$  and  $n \geq 2$ , to better understand whether the preferred models are more disc like or bulge like. Although this is the same Sérsic index cut used in this work, we also use other criteria (based on the  $B/T$ , bulge and disc axis ratio, etc.) in addition to Sérsic index to classify pure bulges and pure discs (see Section 6.6).

For this work, we plot the distribution of our three models (bulge, bulge+disc (2com), and disc) as well as the failed galaxies (called ‘bad’), the problematic two-component fits (called ‘bad 2com’) and the two-component galaxies with  $n = 8$  bulges (called ‘n8’). For the [S11](#) fits, the [deV](#)-Exp, Ser-Exp, and Ser models are plotted. For the [LG12](#) galaxies, the five different models are shown with ‘nb1’ referring to two-component galaxies with  $n = 1$  bulges and ‘nb4’ referring to galaxies with  $n = 4$  bulges. The [Men14](#) fits are separated into [deV](#), Exp, [deV](#)-Exp, and unknown following the Proftype provided by the authors and used earlier in this paper (see Section 6.6).

The classification used in this work (top row) performs largely as expected over the range of  $T$ -types. Late  $T$ -types (above  $T = 4$ ) are  $70$ – $80$  per cent discs and the remainder is approximately equal parts two-component and failed fits. Pure bulges approach  $0$  per cent of the sample at this end. The early end (below  $T = -2$ ) has an increasing percentage of pure bulge systems, but has a large contribution from two-component models. This is discussed further in Section 7.2.2. These Exp components are not indicative of the presence of a true disc but rather a departure from a pure Sérsic profile. Finally, the failed fits (in black) and the  $n = 8$  bulges (in cyan) are uniformly distributed in  $T$ -type. This is expected if the failure rate is not correlated with  $T$ -type. A lack of correlation with  $T$ -type is preferable for evolutionary or environmental studies because it reduces the likelihood of introducing a bias by excluding these categories.

The [S11](#) classification (second row) again shows substantial differences in bulge preference between the [LG12](#) and full samples. This difference occurs for the early types ( $T < 0$ ) similar to Fig. 21 where the shift from two-component to one-component galaxies occurs at brighter absolute magnitude. When the Ser-Exp and [deV](#)-Exp categories are combined into a class of two-component systems,  $46$  per cent of the entire [LG12](#) sample are represented by this class. In contrast, the full [S11](#) sample for has  $35$  per cent two-component galaxies ( $26$  and  $9$  per cent are [deV](#)-Exp and Ser-Exp, respectively). This shows a possible dependence on the image quality of the fitted galaxies (see second row, left). Nearby galaxies are more likely to be considered for two-component fits using the [S11](#) criteria. This is expected since nearby galaxies are better resolved and more likely to have resolved structure. However, we would expect the well-resolved ellipticals to not be as strongly affected as the intermediate types where more structure is expected. These results suggest that while [S11](#) fits are similar to this work (see Section 6) [S11](#) model selection (i.e. Ser, [deV](#)-Exp, or Ser-Exp) is possibly affected by observational effects and may not provide an accurate indication of galaxy type.

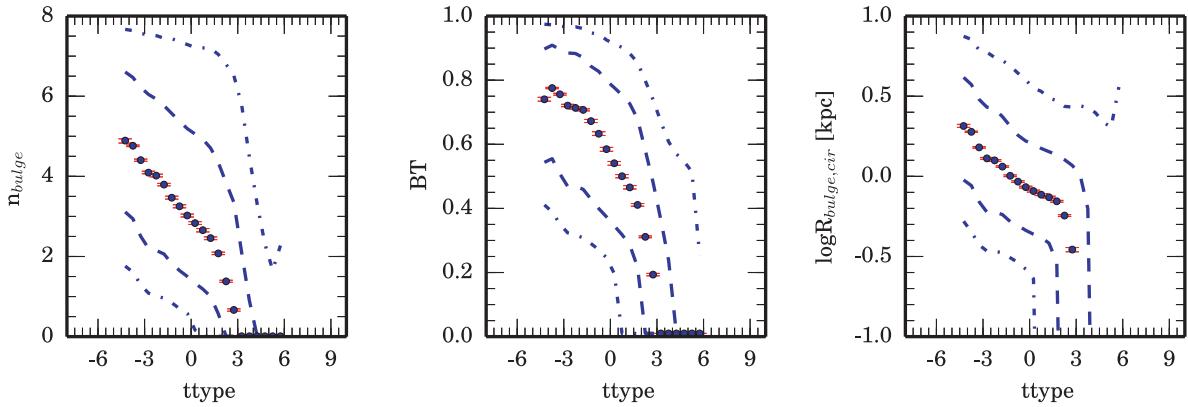
The [Men14](#) classification (third row) is largely independent of  $T$ -type, similar to the behaviour with absolute magnitude discussed in Section 7.1.1. The dominant one-component model transitions from [deV](#) to Exp across the  $T$ -types, but makes up nearly  $20$  per cent of the sample across the range.



**Figure 22.** Comparisons of the inferred  $T$ -type for this work (top row), S11 (second row), Men14 (third row), and LG12 (fourth row). The left-hand column shows the classifications for the smaller, low- $z$  LG12 sample. The right-hand column shows the distribution for the full sample presented in this work. Model types for each group are plotted. The distribution of galaxies as a function of  $T$ -type is plotted in the background histogram and the histogram scale is on the right-hand axis of the plot. The fraction of galaxies reported is the fraction within the respective bin. For each panel, the sum of all models in any given bin is 1. Men14 are again largely insensitive to the  $T$ -type of the galaxy. S11 (second row) have a large shift from fitting early-type galaxies with two-component models in the low- $z$  sample to predominantly fitting Ser models to the same galaxies in the full sample. This indicates that S11 are more sensitive to resolution than to the actual morphology of the galaxy. The bin normalization and background histogram are the same as in Fig. 21.

The LG12 classification (bottom row) shows a large percentage of deV fits at the early end of the  $T$ -types. This is achieved by using a colour cut on the data. Forcing galaxies to be deV profiles based on the colour without considering the fitted models is a qualitatively different process compared to using the profiles to select preferred models, as is done in this work. There is dependence of Sérsic index across type, and we show in Section 7.2.1 that Ell galaxies (i.e. the most negative  $T$ -types) are dominant at higher values of Sérsic index. Choosing the deV model for these galaxies may then impose a bias on the fitted magnitudes for the brightest galaxies in the LG12 fits. LG12 have a similar percentage of pure disc systems at the late

end compared to the fits of this work. Combining the Ser galaxies with  $n < 2$  and the Exp galaxies in LG12, LG12 pure discs at  $T = 4\text{--}6$  account for approximately 70 per cent of galaxies compared to 80 per cent in this work. Also, the Ser models with  $n \geq 2$  (orange points) and the nb4 (green points) models appear to be related by a simple constant fraction (i.e. they are scaled versions of each other). The similar shape of the green and orange curves suggests that the two samples may identify the same type of galaxy, with the distinction between one- and two-component models being the result of observational limitations rather than intrinsic differences in the galaxies.



**Figure 23.** The Ser-Exp bulge Sérsic index,  $B/T$ , and bulge radius for all galaxies as a function of  $T$ -type. Pure discs as denoted by our flags have been given  $B/T = 0$  and the radius and Sérsic index of the bulge are set to 0. The values trend towards zero at later  $T$ -types. Median values are plotted as the points. 68 and 95 per cent contours are plotted as dashed and dot-dashed lines. Error bars on the median values represent 95 per cent CI obtained from bootstrap resampling.

## 7.2 Internal consistency checks using $T$ -type

### 7.2.1 Bulge behaviour with $T$ -type

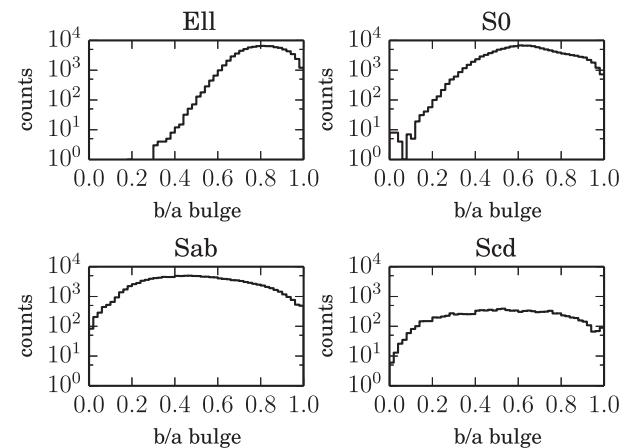
We can examine a few internal checks on the component parameters of bulges and discs in our two-component models. Fig. 23 shows the median Ser-Exp bulge Sérsic index,  $B/T$ , and bulge radius for all galaxies in our catalogue as a function of  $T$ -type calculated using equation (7). 68 and 95 per cent contours are plotted as dashed and dot-dashed lines. Error bars on the median values represent 95 per cent CI obtained from bootstrap resampling. Pure discs, as denoted by our flags, have been given  $B/T = 0$  and the radius and Sérsic index of the bulge are set to 0, but not excluded from the sample.

$B/T$ , bulge radius, and bulge Sérsic index all decrease with increasing  $T$ -type. The median bulge Sérsic index (left-hand panel) for the earliest  $T$ -types is approximately  $5 \pm 1.5$ , using the 68 per cent contours. Median Ser-Exp bulge Sérsic index decreases to 2 by  $T$ -type of 2 before dropping rapidly to zero due to the increased presence of pure disc systems. Median  $B/T$  decreases from 0.8 to 0.2 over the same range while median bulge size also decreases. Since we do not consider the  $T$ -type at all during the flagging or fitting, and we also do not use the colour or any other source of morphological information during fitting, the behaviour of the bulges in our sample is good evidence for proper fitting and effective flagging.

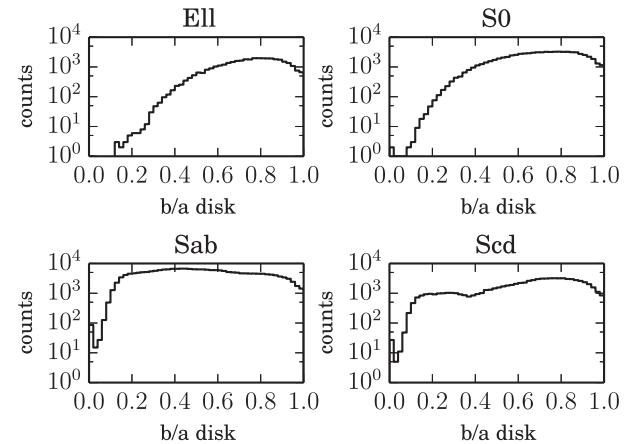
### 7.2.2 Bulge and disc axis ratios

Disc and bulge axis ratios are expected to have different distributions if we are truly measuring bulges and discs as opposed to fitting unphysical components. For instance, bulges are expected to be rounder with lower ellipticities while discs should be more evenly distributed across the range of axis ratios. This distribution for discs is expected for thin discs with uniformly distributed disc inclination. Figs 24 and 25 show the bulge and disc axial ratios for galaxies considered to be (1) two-component or pure bulge for the bulge axis ratio and (2) two-component or pure disc for the disc axis ratio as defined by our flags. We divide the galaxies using the  $T$ -type defined in equation (7). We then bin galaxies as either Ell, S0, Sab or Scd using equation (8).

$$\text{Type} = \begin{cases} \text{Ell} & : T \leq -3 \\ \text{S0} & : -3 < T \leq 0.5 \\ \text{Sab} & : 0.5 < T \leq 4 \\ \text{Scd} & : 4 < T \end{cases} \quad (8)$$

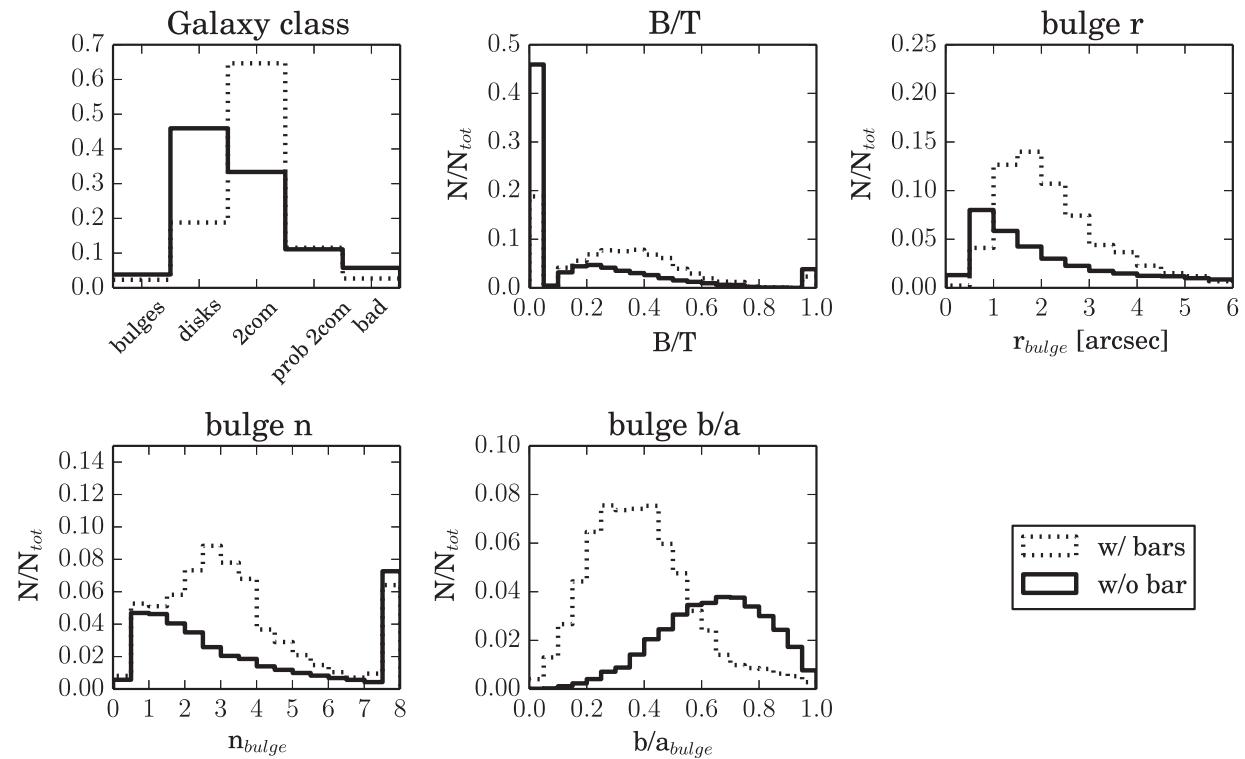


**Figure 24.** The Ser-Exp bulge axis ratio for all good two-component or pure bulge galaxies separated by type.



**Figure 25.** The Ser-Exp disc axis ratio for all good two-component or pure disc galaxies separated by type.

The axis ratio for early types is peaked near 0.8 for the Ell and 0.6 for the S0 galaxies. The distributions get progressively flatter as later types are considered. The increase of bulges with lower  $b/a$  seen in the S0, Sab, and Scd samples is, in part, explained by poor fitting due to contamination of a bar. We address this in Section 7.3. Overall, the distributions shown here are flatter than LG12 but similar to those



**Figure 26.** The distribution of Ser-Exp fitting parameters for disc galaxies with bar components and those that do not have a bar, as identified by selecting on the bar versus no\_bar Galaxy Zoo 2 t03 classification. Preferred models using our Ser-Exp flags (top left), fitted Ser-Exp  $B/T$  (top centre), Ser-Exp semimajor bulge half-light radius (top right), Ser-Exp bulge Sérsic index (bottom left), and Ser-Exp bulge axis ratio (bottom centre) are shown.

reported in S11 for two-component galaxies. The distributions of disc axial ratios are flatter for the later types. We expect a flat distribution in disc axis ratio if galaxies have thin discs due to random orientation of galaxies with respect to the observer.

Also, for the early-type galaxies, the distribution of disc axis ratios is flatter but tends to follow the bulge distribution. This is an additional indication that the second component fit by the disc is not a true disc (i.e. a rotationally flattened disc) but an extended component similar to the bulge. A detailed study on this will be done in the future.

### 7.3 Effects of bars

The simple models fit in this paper and the other works of LG12 and S11 neglect many often observed components of galaxies. The effect of these components is a concern whenever the fitted values are used to test models of formation and evolution. Gadotti (2008) showed that the effect of neglecting a bar or AGN component can still be substantial at lower resolution. A significant bar can increase  $B/T$  and the bulge effective radius by 20 per cent or more. Gadotti (2008) also showed that AGN components had little effect on the bulge Sérsic index (this is most likely due to the loss of resolution which suppresses the Sérsic index, acting oppositely to the effects of the AGN).

Comparing galaxies with known bars to galaxies without bars can establish how strong the effect of bars is in our sample. GZ2 (Willett et al. 2013) presents detailed morphologies for more than 300 000 of the largest and brightest SDSS galaxies ( $m_r < 17$ ). The galaxies were visually classified by thousands of citizen-scientists and corrected for any classification bias introduced by the citizen-scientists using spectroscopic information. GZ2 presents a debiased classification

as well as a simple binary flag that is intended to select a pure, but possibly incomplete sample of many morphological classifications.

Fig. 26 presents the distribution of fitting parameters for disc galaxies with bar components and those that do not have a bar, as identified by GZ (by selecting on the bar versus no\_bar Galaxy Zoo 2 t03 classification). Preferred models using our Ser-Exp flags (top left), fitted Ser-Exp  $B/T$  (top centre), Ser-Exp semimajor bulge half-light radius (top right), Ser-Exp bulge Sérsic index (bottom left), and Ser-Exp bulge axis ratio (bottom centre) for these galaxies are shown.

Fig. 26 shows that we find approximately 50 per cent of the galaxies without bars to be classified as pure discs (i.e. no measurable bulge component). However, when a bar is present, about 65 per cent of barred galaxies are fitted with two components and the percentage of pure disc galaxies drops to  $\sim 20$  per cent. In addition, the  $B/T$  and Sérsic index of the Ser-Exp bulge in barred galaxies is increased.

The distribution of semimajor bulge half-light radii is skewed towards smaller values for bar-free galaxies compared to galaxies that have bar component. This is especially true when considering that almost 50 per cent of galaxies without bars are classified as pure disc versus only 20 per cent of galaxies with bars. These galaxies are included in the normalization of the plots, but do not add to the visible part of this distribution. Cases with the presence of a bar also complicate the interpretation of the bulge component, since the relative brightness of the bar and bulge must be considered when interpreting the radius of this component.

The axis ratio of galaxies with bars is much more bar-like (close to a median of 0.4) than bulge like. This is clear evidence that Ser-Exp bulges of late-type galaxies can be strongly affected by bar contributions and caution should be taken to separate bulges and bars in

the Ser-Exp fits. The distribution of Ser-Exp bulge axis ratio does offer a potential method of separating galaxies with bars from the non-barred galaxies. The distribution of Ser-Exp bulge axis ratios of the two samples are distinct in their median  $b/a$ , although the overlap is quite large. Galaxies without a visually confirmed bar have round bulges while barred galaxies have Ser-Exp bulges of much higher ellipticity. This may also explain the relative flatness of the bulge axis ratio distribution in Fig. 24. Proper modelling of the distribution may provide a constraint on the bulge fraction in our sample. However, this is beyond the scope of this work.

## 8 DISCUSSION AND USE OF THE CATALOGUE

When using the catalogue, we recommend removing galaxies flagged as bad (flag 20) as these galaxies have catastrophically bad estimates of total magnitude and radius. Additional galaxies may be removed depending on how conservative the user seeks to be. The problematic two-component fits (flag 14) or the two-component fits with bulge Sérsic index  $n = 8$  may be used for total magnitude and radius measurements, but the subcomponents are not reliable.

The user should also be aware that we have swapped the bulge and disc components of galaxies with bit 13 set (which were flagged as inverted profiles in the Ser-Exp fit). These galaxies have  $B/T$  inverted and the components reversed relative to the Ser-Exp fit. Therefore, no additional alterations must be made to account for the inverted nature of the profile. However, using the ‘raw’ fit produced prior to flagging requires swapping bulge and disc parameters and inverting the  $B/T$ . This alteration has been done in all catalogues.

We provide our fits and flags for each of the four models (deV, Ser, deV-Exp, and Ser-Exp). In addition, we provide the preferred catalogue used in Sections 6.6 and 7 for public use. This preferred catalogue has Ser fits for galaxies flagged as pure bulge or pure disc. The remaining galaxies have Ser-Exp fits. The derived  $T$ -types used in Section 7 are not included in these tables, but can be calculated directly using equation (7).

We suggest one of these composite samples drawn from the preferred fit catalogue described in the previous paragraph.

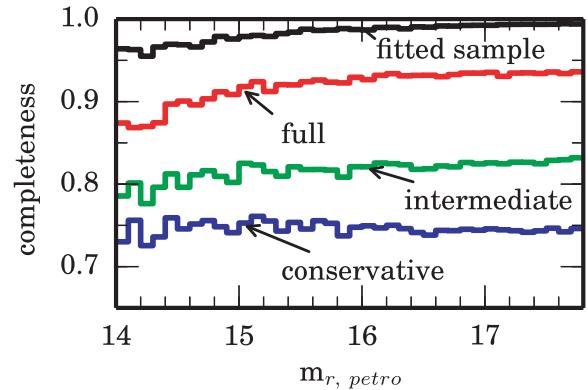
(i) *The conservative catalogue.* Select all galaxies with final flag bits 11, 12, or 13 set and bulge Sérsic index  $<8$ . In addition, the user should select galaxies with Ser-Exp final flag bits 1 or 4 set. These galaxies will have  $B/T$  of 1 (for bulges; final flag bit 1 set) or a  $B/T$  of 0 (for discs; final flag bit 4 set) and the relevant Ser parameters are reported in the catalogue.

(ii) *The intermediate catalogue.* Use the catalogue above plus all galaxies with final flag bit 10 set and bulge Sérsic index = 8.

(iii) *The full catalogue.* Use the catalogue above plus all galaxies with final flag bit 14 set. This is the least restrictive version of the catalogue but may include galaxies with strange, difficult-to-interpret fit parameters.

Fig. 27 shows the completeness of the three samples described above. The ‘fitted sample’ represents our selection after the cuts made in Section 2. All completeness calculations are relative to the original magnitude-limited galaxy sample downloaded from SDSS DR7. Not surprisingly, the completeness drops with more conservative catalogue choices. However, the completeness is largely flat across the magnitude range with a slight decrease of the order of 0.05 at magnitudes brighter than 14.5 extinction-corrected Petrosian  $r$ -band magnitude.

While the deV and deV-Exp models are not a focus of this work, we recognize that there may be cases in which these models are preferred for analysis, especially when comparing to prior works.



**Figure 27.** The completeness of the samples described in Section 8. The ‘fitted sample’ represents our selection after the cuts used in Section 2. All completeness calculations are relative to the original magnitude-limited galaxy sample downloaded from SDSS DR7.

The breakdown of flags for these models is shown in the online Appendix A. If the deV-Exp model is desired for two-component galaxies rather than the Ser-Exp used here, the same criteria used to draw the composite samples described in this section may be applied to the deV-Exp fits. This may be done at the user’s discretion. We point out here that all flagging is carried out in the same manner for the Ser-Exp and deV-Exp models, except where the fits do not allow such treatment (e.g. there can be no cases of bulges appearing disc like in the deV-Exp sample since the Sérsic index is fixed at 4 in all deV-Exp fits). The single-component deV and Ser fits also receive the same treatment during flagging.

The data files for this catalogue are available online at [http://www.physics.upenn.edu/~ameert/SDSS\\_PhotDec/](http://www.physics.upenn.edu/~ameert/SDSS_PhotDec/). We also provide an interface for generating panels similar to Fig. 6, i.e. postage stamp images of the 2D model and residual as well as the 1D profile. These panels can be generated for a user-uploaded list of galaxies on demand.

Tables 2 and 3 describe the format of the data tables released as part of this work. We distribute the data as a binary table using the FITS standard. The first binary extension contains the model-independent measurements for each galaxy (e.g. SExtractor measurements, the number of fitted neighbours, etc.). The following extensions contain the ‘best model’ (the combination of Ser and Ser-Exp fits described in Sections 6.6 and 7), the deV model, the Ser model, the deV-Exp model, and the Ser-Exp model in that order. These extensions include the fitted values for magnitude, radius, and other parameters as well as the flags described in Section 4.3 (labelled as `finalflag`, in column 34). A separate table containing RA/Dec./ $z$  information and other identifying information is also available to allow matching between this catalogue and external works.

## 9 CONCLUSIONS

A catalogue of deV, Ser, deV-Exp, and Ser-Exp galaxies was constructed for the SDSS DR7 spectroscopic sample. We used the PYMORPH pipeline, including the SExtractor and GALFIT programs, to perform 2D decompositions (see Sections 2 and 3). We developed a physically motivated flagging system that removes poor fits and accurately identifies pure bulge, pure disc, and two-component systems (Sections 4.3 and 7.1). After applying the flagging system to the Ser-Exp fit, we identified about 94 per cent of our fitted sample as having reliable total magnitude and half-light

**Table 2.** Description of columns in the electronic table UPenn\_PhotDec\_nonParam\_rband. The data are model-independent measurements fitted by PYMORPH. Problematic data or parameters are replaced with –999.

Column number	Column name	Explanation	Data type
0	SexMag	The SExtractor magnitude (mag)	Float
1	SexMagErr	The SExtractor magnitude error (mag)	Float
2	SexHrad	The SExtractor half-light radii (arcsec)	Float
3	SexSky	The SExtractor sky brightness (mag arcsec <sup>-2</sup> )	Float
4	num_targets	The number of targets	Int
5	num_neighorfit	The number neighbour sources fitted with Ser profiles	Int
6	C	The concentration	Float
7	C_err	The concentration error	Float
8	A	The asymmetry	Float
9	A_err	The asymmetry error	Float
10	S	The smoothness	Float
11	S_err	The smoothness error	Float
12	G	The Gini coefficient	Float
13	M20	The M <sub>20</sub> value	Float
14	extinction	The SDSS-provided galactic extinction (magnitude)	Float
15	dismod	The calculated distance modulus	Float
16	kpc_per_arcsec	The angular scale (kpc/arcsec)	Float
17	Vmax	The volume used for Vmax corrections (Mpc <sup>3</sup> )	Float
18	SN	The average S/N per pixel inside the half-light radius	Float
19	kcorr	k-correction calculated using the SDSS model magnitudes	Float

measurements. About 39 per cent of the sample are two-component galaxies with well-behaved components. An additional ≈11 per cent may be two-component fits, but with difficult-to-interpret components. The remaining 44 per cent are pure bulge and disc galaxies.

We compared the fits to the low-redshift sample of [LG12](#) and the larger samples from [S11](#), [Men14](#), and SDSS. We showed that some measurable systematic differences exist in sky brightness as well as per cent-level systematic differences in size and magnitude. Sections 6.2 and 6.3 discuss some of these differences on a catalogue-by-catalogue basis. This catalogue consistently fits brighter magnitudes to the brightest galaxies (in absolute magnitude) when compared to previous works (e.g. SDSS, [S11](#), [LG12](#), and [Men14](#)) (Section 6.1). Recent work suggests that the magnitudes reported in this work are more accurate than previously measured magnitudes ([Bernardi et al. 2014](#); [Kravtsov et al. 2014](#)).

Throughout this work, we have focused on the Ser-Exp model. Simulations (e.g. [M13](#); [Mosleh et al. 2013](#); [Davari et al. 2014](#)) have shown that there are cases where the Ser-Exp fit produces more reliable measurements of the total magnitude and half-light radius compared to the Ser fit. Examination of the Ser-Exp fits and the flagged categories show that the morphological classes assigned in this catalogue using the flags (i.e. bulge, disc, or two component) also correlate well with estimated *T*-types (Section 7.1.2). Our fits also behave appropriately in *B/T*, bulge size, and bulge Sérsic index with respect to *T*-type (Section 7.2.1). The different methods of model selection used by [S11](#), [LG12](#), [Men14](#), and this work show some differences in the breakdown of preferred models, particularly at brighter magnitudes and for nearby galaxies. However, the judgement of which model selection method is ‘best’ is dependent on individual science goals, and each selection method may be optimal for different studies.

The two-component fits are difficult to interpret in many circumstances. We examined several potential fitting systematics. We make several observations of potential bias in our catalogue.

(i) Bias due to resolution effects for the Ser bulge components of the Ser-Exp fits is likely present in the catalogue. The majority

of galaxies affected by this are the small, low *B/T* galaxies. Galaxies with *B/T* > 0.5, including pure bulge galaxies have a size distribution similar to [Gadotti \(2009\)](#), and we expect the resolution to have little effect on the determination of bulge size for the pure bulge and *B/T* > 0.5 as a result (Section 5.2).

(ii) Bars can strongly alter the measured parameters for the bulges of galaxies with lower *B/T* values, although using the *b/a* distribution can help to separate barred and non-barred galaxies (Section 7.3).

(iii) At the bright end, two-component models should not be interpreted as traditional bulge+disc systems, even though the Ser-Exp fit provides a more accurate measurement of galaxy half-light radius and magnitude. The axis ratio of discs in the early type (see Section 7.2.2 and Figs 24 and 25) suggest that these components are similar to the bulge component and may represent departures from a single Ser profile or an extended halo around the galaxy.

This catalogue will be extended to *g* and *i* bands in the near future to allow for multiwavelength analysis.

## ACKNOWLEDGEMENTS

We would like to thank our referee for many helpful comments and suggestions. These have greatly improved the paper.

We would also like to thank Luc Simard for providing sky brightness values for the fits published in their paper ([Simard et al. 2011](#)). We would also like to thank Mike Jarvis and Joseph Clampitt for many helpful discussions.

This work was supported in part by NASA grant ADP/NX09AD02G and NSF/0908242.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS website is <http://www.sdss.org/>.

**Table 3.** Description of columns in the electronic table UPenn\_PhotDec\_Models\_rband. The data are the ‘best model’, deV, Ser, deV-Exp, and Ser-Exp model fit parameters fitted by PYMORPH. Unfit parameters or missing data are replaced with values of –999.

Column number	Column name	Explanation	Data type
0	m_tot	Total fitted apparent magnitude	Float
1	BT	The $B/T$ (bulge-to-total light ratio) of the fit	Float
2	r_tot	The half-light radius (arcsec) of the total fit	Float
3	ba_tot	The axis ratio (semiminor/semimajor of the total fit)	Float
4	xctr_bulge	The bulge $x$ centre (pixels)	Float
5	xctr_bulge_err	The bulge $x$ centre error (pixels)	Float
6	yctr_bulge	The bulge $y$ centre (pixels)	Float
7	yctr_bulge_err	The bulge $y$ centre error (pixels)	Float
8	m_bulge	The bulge magnitude	Float
9	m_bulge_err	The bulge magnitude error	Float
10	r_bulge	The bulge radius (arcsec)	Float
11	r_bulge_err	The bulge radius error (arcsec)	Float
12	n_bulge	The Sérsic index	Float
13	n_bulge_err	The Sérsic index error	Float
14	ba_bulge	The bulge $b/a$	Float
15	ba_bulge_err	The bulge $b/a$ error	Float
16	pa_bulge	The bulge position angle (degrees)	Float
17	pa_bulge_err	The bulge position angle error (degrees)	Float
18	xctr_disk	The disc $x$ centre (pixels)	Float
19	xctr_disk_err	The disc $x$ centre error (pixels)	Float
20	yctr_disk	The disc $y$ centre (pixels)	Float
21	yctr_disk_err	The disc $y$ centre error (pixels)	Float
22	m_disk	The disc magnitude	Float
23	m_disk_err	The disc magnitude error	Float
24	r_disk	The disc radius (arcsec)	Float
25	r_disk_err	The disc radius error (arcsec)	Float
26	ba_disk	The disc $b/a$	Float
27	ba_disk_err	The disc $b/a$ error	Float
28	pa_disk	The disc position angle (degrees)	Float
29	pa_disk_err	The disc position angle error (degrees)	Float
30	GalSky	The PYMORPH sky brightness (mag/arcsec $^2$ )	Float
31	GalSky_err	The PYMORPH sky brightness (mag/arcsec $^2$ )	Float
32	chi2nu	The $\chi^2/\text{DOF}$	Float
34	finalflag	The primary quality flag described in this work	Float
36	autoflag	The intermediate, visually calibrated, automated flag described in this work	Float
37	pyflag	The PYMORPH run flag	Float
38	pyfitflag	The PYMORPH fit flag	Float

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the US Naval Observatory, and the University of Washington.

## REFERENCES

- Abazajian K. N. et al., 2009, ApJS, 182, 543  
 Aihara H. et al., 2011, ApJS, 193, 29  
 Allen P. D., Driver S. P., Graham A. W., Cameron E., Liske J., de Propris R., 2006, MNRAS, 371, 2  
 Andrae R., Schulze-Hartung T., Melchior P., 2010, preprint ([arXiv:1012.3754](https://arxiv.org/abs/1012.3754))  
 Barden M., Häubler B., Peng C. Y., McIntosh D. H., Guo Y., 2012, MNRAS, 422, 449  
 Bernardi M., 2009, MNRAS, 395, 1491  
 Bernardi M. et al., 2003a, AJ, 125, 1849  
 Bernardi M. et al., 2003b, AJ, 125, 1866  
 Bernardi M. et al., 2003c, AJ, 125, 1882  
 Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, MNRAS, 404, 2087  
 Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697  
 Bernardi M., Meert A., Vikram V., Huertas-Company M., Mei S., Shankar F., Sheth R. K., 2014, MNRAS, 443, 874  
 Bertin E., Arnouts S., 1996, A&AS, 117, 393  
 Blanton M. R., Roweis S., 2007, AJ, 133, 734  
 Blanton M. R. et al., 2001, AJ, 121, 2358  
 Blanton M. R., Eisenstein D., Hogg D. W., Schlegel D. J., Brinkmann J., 2005, ApJ, 629, 149  
 Blanton M. R., Kazin E., Muna D., Weaver B. A., Price-Whelan A., 2011, AJ, 142, 31  
 Capaccioli M., 1989, in Corwin H. G., Jr Bottinelli L., eds, World of Galaxies (Le Monde des Galaxies). Springer-Verlag, Berlin, p. 208  
 Davari R., Ho L. C., Peng C. Y., Huang S., 2014, ApJ, 787, 69

- de Vaucouleurs G., 1948, Ann. Astrophys., 11, 247  
 Djorgovski S., Davis M., 1987, ApJ, 313, 59  
 Donzelli C. J., Muriel H., Madrid J. P., 2011, ApJS, 195, 15  
 Dressler A., 1980, ApJ, 236, 351  
 Faber S. M., Jackson R. E., 1976, ApJ, 204, 668  
 Fukugita M. et al., 2007, AJ, 134, 579  
 Gadotti D. A., 2008, MNRAS, 384, 420  
 Gadotti D. A., 2009, MNRAS, 393, 1531  
 Gonzalez A. H., Zabludoff A. I., Zaritsky D., 2005, ApJ, 618, 195  
 Graham A. W., Driver S. P., 2005, PASA, 22, 118  
 Graham A. W., Driver S. P., 2007, ApJ, 655, 77  
 Häussler B. et al., 2007, ApJS, 172, 615  
 Häußler B. et al., 2013, MNRAS, 430, 330  
 Huang S., Ho L. C., Peng C. Y., Li Z.-Y., Barth A. J., 2013, ApJ, 766, 47  
 Huertas-Company M., Aguerri J. A. L., Bernardi M., Mei S., Sánchez Almeida J., 2011, A&A, 525, A157 (H2011)  
 Huertas-Company M., Shankar F., Mei S., Bernardi M., Aguerri J. A. L., Meert A., Vikram V., 2013, ApJ, 779, 29  
 Kauffmann G. et al., 2003, MNRAS, 341, 33  
 Kelvin L. S. et al., 2012, MNRAS, 421, 1007  
 Kormendy J., 1977, ApJ, 218, 333  
 Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJS, 182, 216  
 Kravtsov A., Vikhlinin A., Meshcheryakov A., 2014, preprint ([arXiv:1401.7329](https://arxiv.org/abs/1401.7329))  
 Lackner C. N., Gunn J. E., 2012, MNRAS, 421, 2277 (LG12)  
 Lintott C. J. et al., 2008, MNRAS, 389, 1179  
 Lintott C. et al., 2011, MNRAS, 410, 166  
 LSST Science Collaboration 2009, preprint ([arXiv:0912.0201](https://arxiv.org/abs/0912.0201))  
 Meert A., Vikram V., Bernardi M., 2013, MNRAS, 433, 1344 (M13)  
 Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, ApJS, 210, 3 (Men14)  
 Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, J. Chem. Phys., 21, 1087  
 Mosleh M., Williams R. J., Franx M., 2013, ApJ, 777, 117  
 Nair P. B., Abraham R. G., 2010, ApJS, 186, 427  
 Padmanabhan N. et al., 2008, ApJ, 674, 1217  
 Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266  
 Sérsic J. L., 1963, Bol. Asociacion Argentina Astron. La Plata Argentina, 6, 41  
 Shankar F., Marulli F., Bernardi M., Dai X., Hyde J. B., Sheth R. K., 2010a, MNRAS, 403, 117  
 Shankar F., Marulli F., Bernardi M., Boylan-Kolchin M., Dai X., Khochfar S., 2010b, MNRAS, 405, 948  
 Shankar F., Marulli F., Bernardi M., Mei S., Meert A., Vikram V., 2013, MNRAS, 428, 109  
 Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978  
 Simard L. et al., 2002, ApJS, 142, 1  
 Simard L., Mendel J. T., Patton D. R., Ellison S. L., McConnachie A. W., 2011, ApJS, 196, 11 (S11)  
 Stoughton C. et al., 2002, AJ, 123, 485  
 Strauss M. A. et al., 2002, AJ, 124, 1810  
 The Dark Energy Survey Collaboration 2005, preprint ([arXiv:astro-ph/0510346](https://arxiv.org/abs/astro-ph/0510346))  
 Tremonti C. A. et al., 2004, ApJ, 613, 898  
 Tully R. B., Fisher J. R., 1977, A&A, 54, 661  
 Vikram V., Wadadekar Y., Kembhavi A. K., Vijayagovindan G. V., 2010, MNRAS, 409, 1379  
 Willett K. W. et al., 2013, MNRAS, 435, 2835  
 York D. G. et al., 2000, AJ, 120, 1579

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**APPENDIX A. AUTOMATED FLAGGING TABLES.**

**APPENDIX B. EXAMPLES OF FLAGGED GALAXIES.**

**APPENDIX C. ADDITIONAL FIT COMPARISONS.**

(<http://mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stu2333/-DC1>).

Please note: Oxford University Press are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

This paper has been typeset from a TeX/LaTeX file prepared by the author.