

**BULGES AND DISKS IN THE NEARBY UNIVERSE:**

**Applications to Evolution and Formation of Galaxies**

Alan Meert

A DISSERTATION

in

**Physics and Astronomy**

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the Degree of

**Doctor of Philosophy**

2015

**Supervisor of Dissertation**

---

Mariangela Bernardi

Associate Professor, Physics and Astronomy

**Graduate Group Chairperson**

---

Marija Drndić

Professor, Physics and Astronomy

**Dissertation Committee:**

Ravi Sheth, Professor, Physics and Astronomy

Evelyn Thomson, Associate Professor, Physics and Astronomy

Masao Sako, Associate Professor, Physics and Astronomy

Adam Lidz, Assistant Professor, Physics and Astronomy

# **BULGES AND DISKS IN THE NEARBY UNIVERSE:**

**Applications to Evolution and Formation of Galaxies**

**COPYRIGHT ©**

**2015**

**Alan Meert**

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 License.

To view a copy of this license, visit <http://creativecommons.org/licenses/>

*To my wife, Katy, whom I love without limits*

## Acknowledgements

I am very grateful to my advisor, Mariangela Bernardi, for her expertise, dedication, and help throughout my graduate career. She has been a great mentor, collaborator, and friend. Thank you to my committee members, Ravi Sheth, Evelyn Thomson, Masao Sako, and Adam Lidz for their service on my committee and the time they devoted to this process. Thank you to the University of Pennsylvania, the Department of Physics and Astronomy, and the Astro Group for their support throughout my career here at Penn.

Thank you and farewell to my collaborators at Penn, Mariangela Bernardi, Vinu Vikram, and Ravi Sheth as well as my friends and colleagues in the graduate group at Penn. Farewell to my fellow graduate students, especially Matt Malloy, Ben Schmitt, Joseph Clampitt, and David Moore, with whom I spent the last 6 years in pursuit of scientific knowledge. The friendships forged here will always be remembered fondly.

This work was supported in part by NASA grant ADP/NNX09AD02G, NSF-AST 0908241, 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 U.S. 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 Web Site is <http://www.sdss.org/>.

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 United States Naval Observatory, and the University of Washington.

I would like to thank my wife, Katy, for her patience and support. I would also like to thank my parents, Greg and Margaret, for their support and motivation throughout my life, including graduate school. Thanks to my siblings, grandparents, and other family members, on both my side and Katy's side of the family for their encouragement without which, this process would have been much less rewarding.

## ABSTRACT

Bulges and Disks in the Nearby Universe: Applications to Evolution and Formation of Galaxies

Alan Meert

Mariangela Bernardi

We present a catalog of 2D, PSF-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. Simulations are used to test the two-dimensional decompositions of SDSS galaxies. We compare these fits to prior catalogs and present a physically motivated flagging system which suggests that more than 90 per cent of two-component fits can be used for analysis. The catalogs provide a robust set of structural and photometric parameters for future galaxy studies. The catalog is then extended into the *g*-band and *i*-band. The data and fitted images are available online at [http://shalaowai.physics.upenn.edu/~ameert/fit\\_catalog/](http://shalaowai.physics.upenn.edu/~ameert/fit_catalog/).

Two applications of the catalog are presented. The sensitivity to the choice of light profile and its affect on the massive end of the stellar mass function is quantified. The total stellar mass density at  $z \sim 0.1$  is about 1.2 times larger than in previous analysis of the SDSS using the new catalog. The differences are most pronounced at the massive end, where the measured number density of objects having  $M_* \geq 6 \times 10^{11} M_\odot$  is approximately five times larger. The systematics of fitting different model profiles on the size-luminosity relation of galaxies in the SDSS (i.e. at  $z \sim 0.1$ ) are also quantified in the *r*-band. The net effect on the R-L relation is small, except for the most luminous tail. Neither the early- nor the late-type relations are pure power laws: both show significant curvature, which we quantify. This curvature confirms that two mass scales are special for both early- and late-type galaxies:  $M_* \sim 3 \times 10^{10}$  and  $2 \times 10^{11} M_\odot$ . The intrinsic scatter in the R-L relation is shown to decrease at large L and/or  $M_*$  and should provide additional constraints on models of how the most massive galaxies formed.

Future work applying the analysis to DES and LSST galaxies at higher redshifts is suggested to further constrain models of galaxy formation and evolution at higher redshifts.

# Contents

<b>Title</b>	i
<b>Copyright</b>	ii
<b>Dedication</b>	iii
<b>Acknowledgements</b>	iv
<b>Abstract</b>	vi
<b>Contents</b>	vii
<b>List of Tables</b>	xiii
<b>List of Figures</b>	xv
<b>1 Introduction</b>	1
1.1 Galaxy Types . . . . .	2
1.1.1 Elliptical . . . . .	2
1.1.2 Lenticular/S0 . . . . .	3
1.1.3 Spiral . . . . .	3
1.1.4 Galaxy classification . . . . .	4
1.2 The Luminosity function . . . . .	5

## CONTENTS

---

1.3	Scaling relations . . . . .	8
1.4	Constraining models of formation and evolution . . . . .	9
1.5	Testing models of galaxy formation and evolution . . . . .	13
<b>I</b>	<b>A catalog of DR7 spectroscopic galaxies</b>	<b>15</b>
<b>2</b>	<b>Catalog Selection</b>	<b>16</b>
2.1	The Sloan Digital Sky Survey . . . . .	16
2.2	SDSS CasJobs Data . . . . .	23
2.3	Additional Data . . . . .	28
2.3.1	Derived Parameters . . . . .	28
2.3.2	K-Corrections . . . . .	29
2.3.3	Rest-Frame Conversion Parameters . . . . .	29
2.3.4	$V_{\text{Max}}$ Correction for Selection Effects . . . . .	29
<b>3</b>	<b>Fitting Process</b>	<b>32</b>
3.1	Preprocessing of SDSS Images . . . . .	32
3.2	The Fitted Profiles . . . . .	35
3.3	The Fitting Pipeline . . . . .	37
3.3.1	PYMORPH . . . . .	37
3.3.2	SEXTRACTOR . . . . .	38
3.3.3	GALFIT . . . . .	39
3.3.4	Non-parametric measurements . . . . .	40
3.3.5	Post-fitting measurements . . . . .	41
3.4	Masking and neighbor identification . . . . .	41
3.4.1	The masking and neighbor identification process . . . . .	44
3.4.2	Failures of masking related to Nearby Neighbors and Source Fragmentation . . . . .	45

---

## CONTENTS

3.4.3	Verifying masking and deblending conditions . . . . .	49
<b>4</b>	<b>Simulations</b>	<b>58</b>
4.1	Creating the catalog simulations . . . . .	59
4.1.1	Selecting the simulation catalog . . . . .	59
4.1.2	Generating the images . . . . .	63
4.1.3	Creating the background . . . . .	65
4.1.4	Noise . . . . .	68
4.1.5	Final processing for fitting . . . . .	69
4.2	Testing PYMORPH image decompositions . . . . .	73
4.2.1	Noiseless images . . . . .	73
4.2.2	The effects of background, neighbor sources, and incorrect PSF extraction . . . . .	83
4.2.3	Testing with real images . . . . .	86
4.2.4	Varying the S/N and pixel size . . . . .	88
4.2.5	Effect of cutout size . . . . .	89
4.2.6	Masking and neighbor thresholding . . . . .	93
4.2.7	The effect of incorrect sky estimation . . . . .	95
4.3	Applicability of an F-test criteria . . . . .	99
4.4	Bias correction using simulations . . . . .	104
4.5	Extending the simulations to the <i>g</i> - and <i>i</i> -bands . . . . .	105
<b>5</b>	<b>Flagging and classification of good and bad fits</b>	<b>111</b>
5.1	The motivation for the flagging system . . . . .	111
5.2	The visual classification . . . . .	115
5.3	The automated flagging system . . . . .	119
5.3.1	Finding galaxies with no bulge or no disk . . . . .	122
5.3.2	Examples of galaxies possessing automated flags . . . . .	127

## CONTENTS

---

5.4	Final Flags . . . . .	133
5.4.1	Detailed descriptions of the final flags . . . . .	136
5.4.2	Examples of the final flags . . . . .	139
5.5	<i>g</i> - and <i>i</i> -band flag results . . . . .	156
<b>6</b>	<b>Internal comparisons and consistency checks</b>	<b>163</b>
6.1	Examination of flags with basic observables . . . . .	163
6.1.1	Behavior with apparent magnitude . . . . .	166
6.1.2	Behavior with absolute magnitude . . . . .	168
6.1.3	Behavior with apparent and absolute half-light radius . . . . .	169
6.1.4	Behavior with axis ratio . . . . .	172
6.1.5	Behavior with bulge Sérsic index . . . . .	172
6.1.6	Interpretation of the two-component galaxies . . . . .	172
6.2	<i>g</i> - and <i>i</i> -band results . . . . .	173
6.3	PSF effects on bulge radius . . . . .	178
6.4	Comparison of the <code>Ser</code> and <code>Ser-Exp</code> models . . . . .	182
<b>7</b>	<b>Comparisons to literature</b>	<b>187</b>
7.1	External catalogs used for comparison and analysis . . . . .	187
7.2	The <code>deV</code> fits . . . . .	190
7.3	The <code>Ser</code> fits . . . . .	197
7.4	The <code>deV-Exp</code> fits . . . . .	203
7.5	The <code>Ser-Exp</code> fits . . . . .	213
7.6	Comparing the preferred models . . . . .	217
<b>8</b>	<b>Comparisons using morphological information</b>	<b>223</b>
8.1	Preferred models as an indicator of morphological classification . . . . .	226
8.2	Model selection as a function of <i>T</i> -type . . . . .	228
8.3	Internal consistency checks using <i>T</i> -type . . . . .	234

---

## CONTENTS

8.3.1	Bulge behavior with $T$ -type . . . . .	234
8.3.2	Bulge and disk axis ratios . . . . .	235
8.4	Effects of bars . . . . .	238
8.5	Color-magnitude relations of the types . . . . .	241
<b>9</b>	<b>Discussion and use of the catalog</b>	<b>243</b>
 <b>II Applications of the catalog</b>		<b>249</b>
<b>10</b>	<b>Dependence of the bright end of the LF and MF on the fitted profile</b>	<b>250</b>
10.1	Introduction . . . . .	251
10.2	Luminosity and stellar mass functions . . . . .	252
10.2.1	The sample . . . . .	252
10.2.2	Dependence on assumed surface brightness profile . . . . .	254
10.2.3	Dependence on pipeline . . . . .	256
10.2.4	The luminosity function . . . . .	258
10.2.5	The stellar mass function . . . . .	259
10.3	Comparison with previous work . . . . .	263
10.4	Dependence on morphology . . . . .	266
10.5	Discussion . . . . .	267
<b>11</b>	<b>Systematic effects on the size-luminosity relations of early and late type galaxies: dependence on model fitting and morphology</b>	<b>272</b>
11.1	Introduction . . . . .	273
11.2	The SDSS DR7 sample . . . . .	276
11.2.1	Estimates of photometric parameters . . . . .	276
11.2.2	Morphologies . . . . .	277
11.3	The $R - L$ relation at $z \sim 0.1$ . . . . .	280

## CONTENTS

---

11.3.1 Single <b>Ser</b> fits and comparison with previous work . . . . .	282
11.3.2 Dependence on model fitting . . . . .	284
11.3.2.1 Early-types . . . . .	284
11.3.2.2 Comparison of early- and late-types . . . . .	289
11.4 Dependence on Bayesian Automated Classifier morphologies . . . . .	291
11.4.1 Small but statistically significant difference between Ellipticals and S0s	291
11.4.2 Scatter in log(size) around the mean relation for early-types . . . . .	293
11.5 Summary . . . . .	297
<b>III Future Work and Concluding Remarks</b>	<b>300</b>
<b>12 Future Work</b>	<b>301</b>
12.1 Further exploration of <i>g</i> - and <i>i</i> -bands . . . . .	301
12.2 Pushing to higher resolution and redshift . . . . .	303
12.3 Exploring the Green Valley . . . . .	303
<b>13 Conclusion</b>	<b>305</b>
<b>Glossary</b>	<b>308</b>
<b>References</b>	<b>310</b>
<b>Index</b>	<b>317</b>

# List of Tables

1.1	Hubble T-types, deV types and Hubble Types . . . . .	4
4.1	The bias and scatter of the fitted parameters from simulations . . . . .	87
5.1	Description of the visual flags . . . . .	118
5.2	Incidence of visual flags in the manually classified sample . . . . .	120
5.3	The mapping of visual flags to automated flags . . . . .	121
5.4	Automated flag numerical selection criteria . . . . .	123
5.5	The incidence of automated flags . . . . .	124
5.6	Final flagging breakdown for <i>r</i> -band <b>Ser-Exp</b> fits . . . . .	134
5.7	Description of the criteria used for the final flags . . . . .	137
5.8	Incidence the final flags in the <i>r</i> -band . . . . .	138
5.9	Noise properties of the <i>g</i> -, <i>r</i> -, and <i>i</i> -band images . . . . .	157
5.10	deV model table showing flagging breakdown for the <i>g</i> -, <i>r</i> -, and <i>i</i> -bands. . .	159
5.11	Ser model table showing flagging breakdown for the <i>g</i> -, <i>r</i> -, and <i>i</i> -bands. . .	160
5.12	deV-Exp model table showing flagging breakdown for the <i>g</i> -, <i>r</i> -, and <i>i</i> -bands. .	161
5.13	Ser-Exp model table showing flagging breakdown for the <i>g</i> -, <i>r</i> -, and <i>i</i> -bands. .	162
9.1	Description of columns in UPenn_PhotDec_nonParam_rband . . . . .	247
9.2	Description of columns in UPenn_PhotDec_Models_rband . . . . .	248

## LIST OF TABLES

---

10.1 PYMORPH, DR7, and S11 fitting parameters of LF and MF . . . . .	257
11.1 Eyeball morphological classifications from Fukugita et al. (2007) . . . . .	279
11.2 Eyeball morphological classifications from Nair and Abraham (2010) . . . . .	281
11.3 Luminosity-size relation . . . . .	287
11.4 Stellar mass-size relation . . . . .	288

# List of Figures

1.1	The Hubble Tuning Fork . . . . .	2
1.2	B/T as a function of type from Oohama et al. (2009) . . . . .	6
1.3	Schematic of the process of galaxy formation . . . . .	10
1.4	Predicted quantities of semi-analytical models . . . . .	11
1.5	Predicted B/T and elliptical fraction from Wilman et al. (2013) . . . . .	11
1.6	Predicted size-luminosity relations from Shankar et al. (2010a) . . . . .	12
2.1	The SDSS Telescope . . . . .	17
2.2	The SDSS focalplane . . . . .	18
2.3	SDSS filter transmission curves in vacuum . . . . .	19
2.4	A SDSS spectroscopic plate . . . . .	20
2.5	SDSS DR7 legacy survey photometric and spectroscopic sky coverage . . . . .	22
2.6	Distributions of observable values for catalog galaxies . . . . .	26
3.1	Flowchart of preprocessing . . . . .	34
3.2	Flowchart of PYMORPH . . . . .	37
3.3	Two example galaxy fits . . . . .	42
3.4	Flowchart of the Masking and Deblending process . . . . .	43
3.5	Example Masked and simultaneously fit neighbors . . . . .	46
3.6	An incorrectly fit close-pair of galaxies . . . . .	48

## LIST OF FIGURES

---

3.7	Example galaxy with fictitious CasJobs neighbor . . . . .	54
3.8	Change in magnitude of refitted galaxies with contaminating neighbors . . . . .	57
4.1	Comparison of simulated and original galaxy distributions . . . . .	62
4.2	Sky Brightness Distribution Used in Simulations . . . . .	66
4.3	Distribution of Real and Simulated Galaxy Radii Versus Apparent Magnitude	70
4.4	Distribution of Real and Simulated Galaxy S/N Versus Apparent Magnitude	71
4.5	Examples of Mock Galaxies and Background With and Without Noise . . . . .	72
4.6	<b>Ser</b> fit to <b>Ser</b> simulations . . . . .	77
4.7	<b>Ser-Exp</b> fit to <b>Ser</b> simulations . . . . .	78
4.8	<b>Ser</b> fit to <b>Ser-Exp</b> simulations . . . . .	79
4.9	Total fit parameters of <b>Ser-Exp</b> fit to <b>Ser-Exp</b> simulations . . . . .	80
4.10	Bulge fit parameters of <b>Ser-Exp</b> fit to <b>Ser-Exp</b> simulations . . . . .	81
4.11	Disk fit parameters of <b>Ser-Exp</b> fit to <b>Ser-Exp</b> simulations . . . . .	82
4.12	Error correlation matrix for simulated <b>Ser-Exp</b> galaxies . . . . .	85
4.13	<b>Ser</b> fit with <b>Ser</b> for resolution degraded by 2 . . . . .	90
4.14	<b>Ser</b> fit with <b>Ser-Exp</b> for resolution degraded by 2 . . . . .	90
4.15	<b>Ser-Exp</b> fit with <b>Ser</b> for resolution degraded by 2 . . . . .	90
4.16	<b>Ser-Exp</b> fit with <b>Ser-Exp</b> for resolution degraded by 2 . . . . .	91
4.17	<b>Ser-Exp</b> subcomponents fit with <b>Ser-Exp</b> for resolution degraded by 2 . . . . .	92
4.18	Magnitude, Radius, and Sky as a function of cutout size . . . . .	93
4.19	Testing fit stability as a function of mask shape and size . . . . .	96
4.20	SDSS Sky vs PYMORPH Estimation . . . . .	98
4.21	Effects of sky error on <b>Ser</b> model fit with <b>Ser</b> . . . . .	100
4.22	Effects of sky error on <b>Ser-Exp</b> model fit with <b>Ser-Exp</b> . . . . .	101
4.23	Magnitude and Half-light Radius of <b>Ser</b> galaxies with F-test . . . . .	103
4.24	Bias correction example using simulation results . . . . .	106
4.25	Simulations at S/N of <i>g</i> -band images . . . . .	109

---

## LIST OF FIGURES

5.1	The $\chi^2$ cdf for fits . . . . .	112
5.2	Example panel for visual classification . . . . .	117
5.3	Bulge and disk cuts versus B/T and magnitude . . . . .	126
5.4	Examples of initial automated flags for fitting problems . . . . .	128
5.5	Examples of automated flags suggesting a single component . . . . .	129
5.6	Examples of automated flags for Ser problems . . . . .	130
5.7	Examples of automated flags for Exp problems . . . . .	131
5.8	Examples of automated flags for large fitting inaccuracies . . . . .	132
5.9	Example Ser-Exp fits considered good using the final flags . . . . .	141
5.10	Additional example Ser-Exp fits considered good . . . . .	142
5.11	Example Ser-Exp fits with flipped components . . . . .	143
5.12	Example Ser-Exp fits with bulge-like profiles . . . . .	144
5.13	Additional example Ser-Exp fits with bulge-like profiles . . . . .	145
5.14	Example Ser-Exp fits with disk-like profiles . . . . .	146
5.15	Additional example Ser-Exp fits with disk-like profiles . . . . .	147
5.16	Additional example Ser-Exp fits with disk-like profiles . . . . .	148
5.17	Additional example Ser-Exp fits with disk-like profiles . . . . .	149
5.18	Additional example Ser-Exp fits with disk-like profiles . . . . .	150
5.19	Example Ser-Exp fits that are difficult to interpret . . . . .	151
5.20	Additional example Ser-Exp fits that are difficult to interpret . . . . .	152
5.21	Additional example Ser--Exp fits that are difficult to interpret . . . . .	153
5.22	Additional example Ser-Exp fits that are difficult to interpret . . . . .	154
5.23	Additional example Ser-Exp fits that are difficult to interpret . . . . .	155
6.1	Percentages of galaxies given as a percent of the bin . . . . .	164
6.2	Percentages of galaxies given as a percent of the total . . . . .	165
6.3	B/T as a function of half-light radius . . . . .	171
6.4	Magnitude distribution of galaxy types for each band . . . . .	174

## LIST OF FIGURES

---

6.5	Size distribution of galaxy types for each band . . . . .	176
6.6	Axis-ratio and bulge index distribution of galaxy types for each band . . . . .	177
6.7	Confusion matrix for bands and flagging categories . . . . .	179
6.8	Comparison of bulge circularized effective radius to the PSF HWHM . . . . .	180
6.9	Comparison of <b>Ser</b> fit to <b>Ser-Exp</b> fit . . . . .	183
6.10	Comparison of <b>Ser</b> fit to <b>Ser-Exp</b> fit, absolute quantities . . . . .	184
6.11	Comparison of <b>Ser</b> radius to the <b>Ser-Exp</b> bulge radius as a function of B/T	185
7.1	Limited comparison of fitted <b>deV</b> models . . . . .	191
7.2	Full comparison of fitted <b>deV</b> models, <i>g</i> -band . . . . .	192
7.3	Full comparison of fitted <b>deV</b> models, <i>r</i> -band . . . . .	193
7.4	Full comparison of fitted <b>deV</b> models, <i>i</i> -band . . . . .	194
7.5	Comparison of fitted <b>Ser</b> models . . . . .	198
7.6	Full comparison of fitted <b>deV</b> models, <i>g</i> -band . . . . .	199
7.7	Full comparison of fitted <b>Ser</b> models, <i>r</i> -band . . . . .	200
7.8	Full comparison of fitted <b>Ser</b> models, <i>i</i> -band . . . . .	201
7.9	The difference in <i>r</i> -band sky values . . . . .	201
7.10	Comparison of fitted <b>deV-Exp</b> models . . . . .	204
7.11	Comparison of global fit parameters for <b>deV-Exp</b> models in the <i>g</i> -band . . .	205
7.12	Comparison of bulge fit parameters for <b>deV-Exp</b> models in the <i>g</i> -band . . .	206
7.13	Comparison of disk fit parameters for <b>deV-Exp</b> models in the <i>g</i> -band . . .	207
7.14	Comparison of global fit parameters for <b>deV-Exp</b> models in the <i>r</i> -band . . .	208
7.15	Comparison of bulge fit parameters for <b>deV-Exp</b> models in the <i>r</i> -band . . .	209
7.16	Comparison of disk fit parameters for <b>deV-Exp</b> models in the <i>r</i> -band . . .	210
7.17	Comparison of global fit parameters for <b>deV-Exp</b> models in the <i>i</i> -band . . .	211
7.18	Comparison of component fit parameters for <b>deV-Exp</b> models in the <i>i</i> -band	212
7.19	Comparison of global fit parameters for <b>Ser-Exp</b> models in the <i>g</i> -band . . .	213
7.20	Comparison of component fit parameters for <b>Ser-Exp</b> models in the <i>g</i> -band	214

---

## LIST OF FIGURES

7.21 Comparison of global fit parameters for <b>Ser-Exp</b> models in the <i>r</i> -band . . . . .	215
7.22 Comparison of component fit parameters for <b>Ser-Exp</b> models in the <i>r</i> -band . . . . .	216
7.23 Comparisons of <i>g</i> -band SDSS cModel and Petrosian magnitude to models . . . . .	218
7.24 Comparisons of <i>r</i> -band SDSS cModel and Petrosian magnitude to models . . . . .	219
7.25 Comparisons of <i>i</i> -band SDSS cModel and Petrosian magnitude to models . . . . .	220
8.1 Type comparisons versus magnitude . . . . .	224
8.2 Type comparisons versus magnitude for all bands . . . . .	225
8.3 Type comparisons versus magnitude for all bands . . . . .	229
8.4 Comparison of model selection versus T-type . . . . .	230
8.5 <b>Ser-Exp</b> bulge Sérsic index, B/T, and bulge radius as a function of <i>T</i> -type . . . . .	234
8.6 <b>Ser-Exp</b> bulge axis-ratio separated by type . . . . .	236
8.7 <b>Ser-Exp</b> disk axis-ratio separated by type . . . . .	237
8.8 <b>Ser-Exp</b> fitting parameters for disc galaxies with bar components . . . . .	239
8.9 Color-magnitude plot . . . . .	242
9.1 The completeness of the catalog . . . . .	244
10.1 Difference between PyMORPH Sérsic fits, SDSS DR7, and S11 . . . . .	253
10.2 Comparison of <b>Ser</b> fits to other works . . . . .	255
10.3 SDSS Main galaxy LF based on several works . . . . .	260
10.4 SDSS Main galaxy MF based on several works . . . . .	261
10.6 Morphological dependence of the LF and MF . . . . .	264
11.1 The <i>r</i> -band single-Ser half-light radius versus total absolute magnitude . . . . .	283
11.2 Dependence of the size-luminosity and size-stellar mass on the profile . . . . .	285
11.3 The $R - M_*$ relation for <b>Ser-Exp</b> fits . . . . .	289
11.4 The $R - M_*$ relation for <b>Ser-Exp</b> fits using hard cuts on the sample . . . . .	290
11.5 BAC-weighted <b>Ser-Exp</b> based size-luminosity and size- $M_*$ relations . . . . .	292

## **LIST OF FIGURES**

---

11.6 Size-luminosity relation at fixed luminosity for Elliptical and S0 . . . . .	293
11.7 Size-luminosity relations at fixed velocity dispersion . . . . .	295
11.8 Observed and estimated scatter around size-luminosity relations . . . . .	298
12.1 Bulge and disk colors . . . . .	302

# Chapter 1

## Introduction

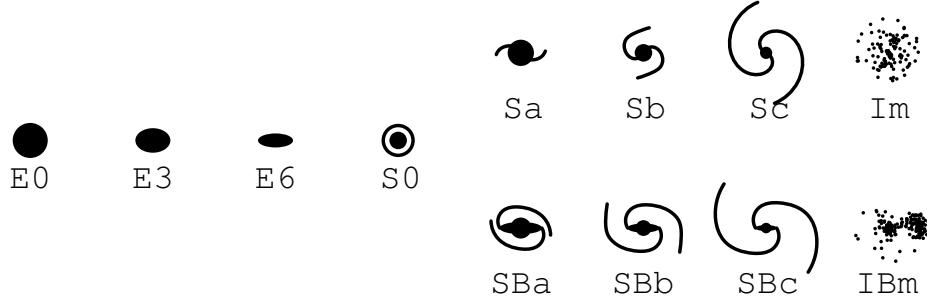
Convincing evidence of galaxies that exist beyond our home Galaxy, the Milky Way, was presented by Hubble (1929). Hubble used observations of Cepheid variable stars, which have a predictable relationship between absolute brightness and the period of the stars variability to show that the nearby nebula, Andromeda, was indeed much further away. It was so far, that there was no way the nebula could be part of the Milky Way.

Prior to this work Hubble constructed the “tuning fork” diagram for classifying galaxies (Hubble 1922, 1926) which has been refined since then. Figure 1.1 shows the tuning fork. The Hubble Tuning Fork as described in Abraham (1998). The fork arranges Elliptical galaxies (Early types) on the left and Spiral (Late types) on the right. The spirals are separated into barred and non-barred forms and the prominence of the bulge decreases toward the left. The winding of the spiral arms also becomes looser toward the right. The distinction of Early vs Late types are due to the incorrect belief that Elliptical galaxies evolved toward Spiral types. If anything, it is the opposite.

Since the 1920’s much work has been carried out cataloging, measuring, and interpreting galaxy properties and properties of the population of galaxies as a whole. Global galaxy properties (e.g. shape, size, total magnitude) were some of the first properties measured

## 1. INTRODUCTION

---



**Figure 1.1:** The Hubble Tuning Fork as described in Abraham (1998). The fork arranges Elliptical galaxies (Early types) on the left and Spiral (Late types) on the right. The spirals are separated into barred and non-barred forms and the prominence of the bulge decreases toward the left. The winding of the spiral arms also becomes looser toward the right.

for galaxies. The first galaxy catalogs (de Vaucouleurs 1948; Sérsic 1963b) focused on these properties because of the limitations of observers at the time and the lack of sufficient understanding of galaxy dynamics. These basic properties have since proved to be useful tests of evolutionary models for galaxies. In addition, properties like color and the sub-components contained in a galaxy have also become of interest to observers.

### 1.1 Galaxy Types

Galaxies can be roughly separated into several types. The types exhibit different colors, morphologies, and compositions.

#### 1.1.1 Elliptical

Elliptical galaxies appear smooth in structure. They vary in apparent ellipticity from circular (E0) to having an axis ratio (the ratio of the shortest axis of the ellipsoid to the longest) as high as 0.3 (E7). These galaxies tend to have older stellar populations with no young, blue stars. There is generally little cold gas in elliptical galaxies, and the star formation rates

## 1.1 Galaxy Types

---

are quite low ( $<<< 1M_{\odot}/\text{year}$ ) and are often found in clustered environments. The most massive galaxies (BCG's and cD's) are elliptical galaxies.

The radial light profile of Elliptical galaxies tend to follow a de Vaucouleurs profile (de Vaucouleurs 1948), or the modified version, the Sérsic profile (Sérsic 1963a):

$$I(r) = I_e \exp \left( -b_n \left[ \left( \frac{r}{R_e} \right)^{\frac{1}{n}} - 1 \right] \right) \quad (1.1)$$
$$b_n = 1.9992n - 0.3271,$$

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.  $b_n$  has several approximations, the approximation from Capaccioli (1989) is used here, which is valid for  $0.5 < n < 10$ . When  $n = 4$ , the Sérsic model reduces to the de Vacouleurs model. Sérsic index values tend to range between 4 and 6 for elliptical galaxies, although galaxies with higher index values may exist in nature (e.g. Graham and Driver 2007; Kormendy et al. 2009).

### 1.1.2 Lenticular/S0

S0 galaxies tend to share qualities with both the elliptical and spiral galaxies. They may have a thin, rotating disk as well as a dominant bulge component. They may also possess a bar.

### 1.1.3 Spiral

Spiral galaxies have a thin, rotating disk composed of cooler gas, dust, star-forming regions, and young blue stars. In addition, the disk often possesses spiral arms. The number, tightness of the winding, and continuity of the arms is variable. Spiral galaxies also may have a bar component as well as a redder bulge component at the center of the galaxy.

## 1. INTRODUCTION

---

<b>Hubble</b>	E	E-S0	S0	S0-Sa	Sa	Sa-b	Sb	Sb-c	Sc	Sc-Irr	Irr
<b>deV</b>	E	S0 <sup>-</sup>	S0 <sup>0</sup>	S0 <sup>+</sup>	Sa	Sab	Sb	Sbc	Scd	Sdm	Im
<b>T</b>	-5	-3	-2	0	1	2	3	4	6	8	10

**Table 1.1:** Comparison of the major classes of Hubble galaxies to the classification scheme of de Vaucouleurs.

Spiral disks are commonly treated as an exponential disk with a scale radius ( $R_d$ ) and central surface brightness ( $I_d$ ). The disk is modeled using the function

$$I_{Exp}(r) = I_d \exp\left(\frac{-r}{R_d}\right) \quad (1.2)$$

This is equivalent to setting the Sérsic index to 1 and redefining the radius and surface brightness parameters.

Since a bulge is also present, the bulge and disk can be simultaneously fitted using a Sérsic model for the bulge and an Exponential disk for the galactic disk. A simultaneous fit of these two components produces a measurement of the bulge and disk parameters. The relative contribution of each component (bulge and disk) are interesting because the latest type spirals appear to have little or no bulge and the earliest ellipticals have no disk. The relative contributions of bulge and disk are characterized by the amount of light contributed by the bulge relative to the total light of the galaxy. This is called the B/T ratio. With better resolution, additional galaxy components like bars, rings, etc. can be modeled and similar ratios can be constructed.

### 1.1.4 Galaxy classification

Over the last 50 years, other classification schemes have been produced. Another popular classification scheme is the de Vaucouleurs system, which uses a numerical T-type (de Vaucouleurs 1974) and refined the Hubble system. The relationship between the classifications are shown in Table 1.1.

## 1.2 The Luminosity function

---

These classifications are useful because they tend to follow trends in measurable properties of galaxies as well as fitting within models of galaxy formation. Merging morphological information with other measures of galaxies (e.g. luminosity, mass, color, etc.) can inform formation models.

For instance, the T-type has a broad relationship with B/T ratio for galaxies (Oohama et al. 2009; Simien and de Vaucouleurs 1986). Figure 1.2 shows the results taken from Oohama et al. (2009) for B/T vs. T-type. Oohama et al. (2009) fitted an  $n = 4$  bulge and an exponential disk to approximately 700 well-resolved, bright, nearby galaxies. They use a modified T-type, where T= 0 represents Elliptical, and T values of 1,2,3, 4, and 5 represent S0, Sa, Sb, Sc, and Sd galaxies, respectively. The B/T value decreases significantly with type, although there is large scatter. Such decompositions have only been carried out for small samples ( $\sim 1000$  galaxies), and only recently on larger samples (Lackner and Gunn 2012; Simard et al. 2011).

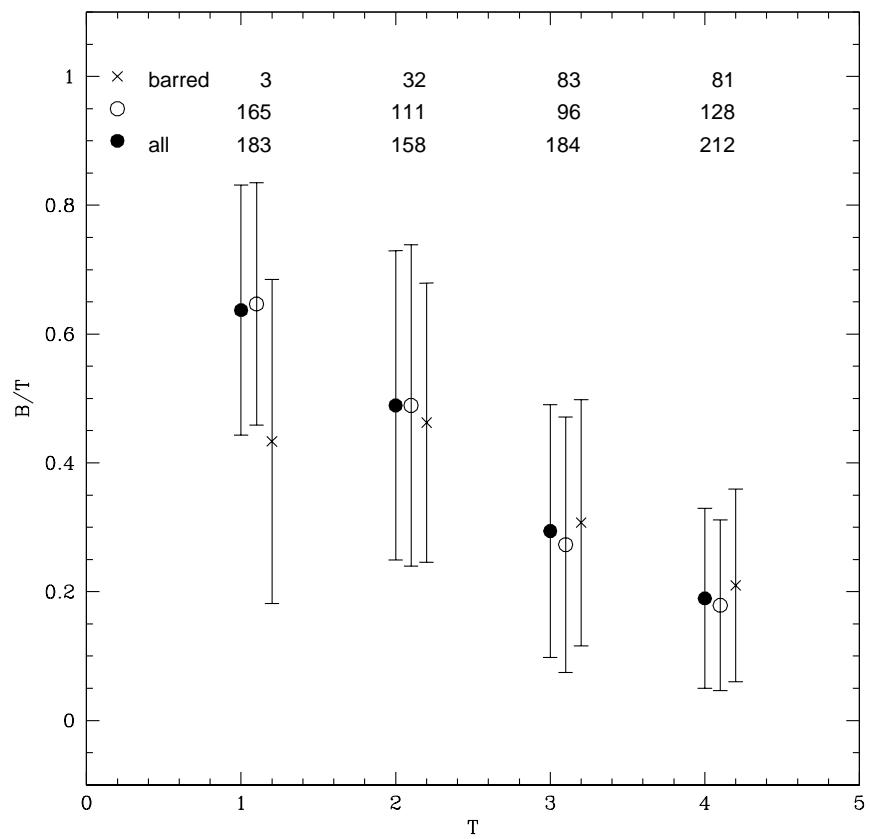
B/T measurements provide an automated proxy for the T-type, which has to be assigned by visual inspection. In addition, the B/T value is a continuous value, making it easier to incorporate or predict in numerical cosmological simulations. In addition, the bulge and disk components of galaxies trace many other galaxy parameters related to both assembly and evolution of galaxies (color, metallicity, gas fraction, central velocity dispersion) (Bernardi et al. 2003b,c,d). Properties may also trace halo size, galaxy environment, and place constraints on  $\Lambda$ CDM cosmology (Bernardi 2009; Blanton et al. 2005a; Shankar et al. 2010a,b). Therefore, accurate measurement of galaxy fit parameters (including B/T, radius, magnitude, mass, etc.) are useful for constraining formation models.

## 1.2 The Luminosity function

The Luminosity function (LF) describes the number density of galaxies at a given magnitude or luminosity. A common model for the luminosity function was developed in Press

## 1. INTRODUCTION

---



**Figure 1.2:** Results of Oohama et al. (2009) for  $B/T$  vs. T-type. A modified T-type is used, where  $T=0$  represents Elliptical, and  $T$  values of 1, 2, 3, 4, and 5 represent S0, Sa, Sb, Sc, and Sd galaxies, respectively.

## 1.2 The Luminosity function

---

and Schechter (1974) and Schechter (1976). It is written:

$$\Phi(L)dL = \phi^* \left( \frac{L}{L_*} \right)^\alpha \exp \left( \frac{-L}{L_*} \right) \frac{dL}{L_*} \quad (1.3)$$

where  $\phi^*$  a normalization defining the overall density of galaxies, often quoted in units of  $h^3 \text{ Mpc}^{-3}$ , and  $L_*$  is the characteristic luminosity above which the number of galaxies declined exponentially.  $\alpha$  represents the faint-end slope of the LF and is typically negative. This reflects the expectation that there are many more galaxies at low luminosities than at higher luminosities.

Equation 1.3 can also be written in magnitudes, requiring appropriate conversion of the luminosity. The measurement in magnitudes is closer to the explicitly measured quantity. Absolute calibration of telescope measurements is generally quite difficult, and using the magnitude system eliminates the explicit requirement for accurate absolute calibration.

There are several other important steps to accurate determination of the luminosity function. These include accurate estimations of the extinction (both due to our galaxy and to the galaxy being observed), accurate  $K$ -correction measurement, and accurate redshift estimation. In addition, assumptions about the cosmology of the universe are necessary in order to obtain a distance from the galaxy redshift.

Finally, there are many observational biases that can distort the LF. The most common is an apparent magnitude cut on the survey. Such cuts make the observable volume behave as a function of the observed magnitude. As a result, a  $V_{max}$  correction is applied, weighting each galaxy by the volume in which it can be observed (see Schmidt 1968, for details of the weighting and diagnostic tests).

Similar effects on the observable volume can arise as a result of redshift limits, size limits, or surface brightness limits, all of which are often seen in surveys (see Shen et al. 2003, for examples of these effects in SDSS data and methods for accounting for such effects).

Nonetheless, the LF can serve as a test of formation models by describing the statistical properties of the galaxy population, including the evolution of the population at different

## 1. INTRODUCTION

---

redshifts and the effects of environment on the galaxy population.

The LF can also be converted to a Mass function (MF) by determining the Mass-to-Light ratio ( $M/L$ ) . For galaxies, this requires modeling the stellar population of the galaxy, which is non-trivial, as well as determining the amount of gas, dust, and non-baryonic dark-matter in the galaxy. Another option is to use a lensing mass estimate for the galaxy, or some other probe that directly measures the galaxy mass (rotation curves, etc. ).

### 1.3 Scaling relations

The structural components of galaxies in the local Universe obey many scaling relations including the morphology-density relation (MDR; Dressler 1980). Elliptical galaxies and (to some extent, bulges of spiral galaxies) follow the Kormendy (Kormendy 1977), Faber-Jackson (Faber and Jackson 1976), and the Fundamental Plane (Djorgovski and Davis 1987). Disks of spiral galaxies also follow the Tully-Fisher (Tully and Fisher 1977) relation and spiral galaxy bulges follow the  $M_{BH} - \sigma$  relation.

The Fundamental Plane (Djorgovski and Davis 1987) arises from arguments related to the Virial theorem:

$$\frac{GM}{\langle R \rangle} = \langle v^2 \rangle \quad (1.4)$$

where  $G$  is Newton's gravitational constant,  $M$  is the mass of the object,  $\langle R \rangle$  is the average radius of the object, and  $\langle v^2 \rangle$  is the average squared speed. The left side is the mean potential energy per unit mass and the right side is the mean kinetic energy per unit mass. The Virial theorem relates the two with a factor of 2 (which has been omitted). The half-light radius,  $R_e$  can be related to  $\langle R \rangle$  up to a constant (i.e.,  $R_e = k_r \langle R \rangle$ ). Similarly, the central velocity dispersion is related to  $\langle v^2 \rangle$  (i.e.,  $\sigma_0 = k_v \langle v^2 \rangle$ ). The luminosity of the galaxy can be expressed as  $L = 2\pi \langle I \rangle_e R_e^2$  where  $\langle I \rangle_e$  is the average surface brightness within  $R_e$ .

Combining all of these terms yields:

$$R_e = K \sigma^2 \langle I \rangle_e^{-1} \left( \frac{M}{L} \right)^{-1} \quad (1.5)$$

## 1.4 Constraining models of formation and evolution

---

and taking the log of both sides yields

$$\log R_e = a \log \sigma_0 + b \log \langle I \rangle_e^{-1} + \text{constant} \quad (1.6)$$

which is the Fundamental Plane. The plane may show tilt or curvature if the  $M/L$  or the other constants are not truly constants, but have some dependence on the elliptical galaxies of interest.

The other relations (size-luminosity, Kormendy, Faber-Jackson) are the result of the projection of the fundamental plane into two dimensions.

## 1.4 Constraining models of formation and evolution

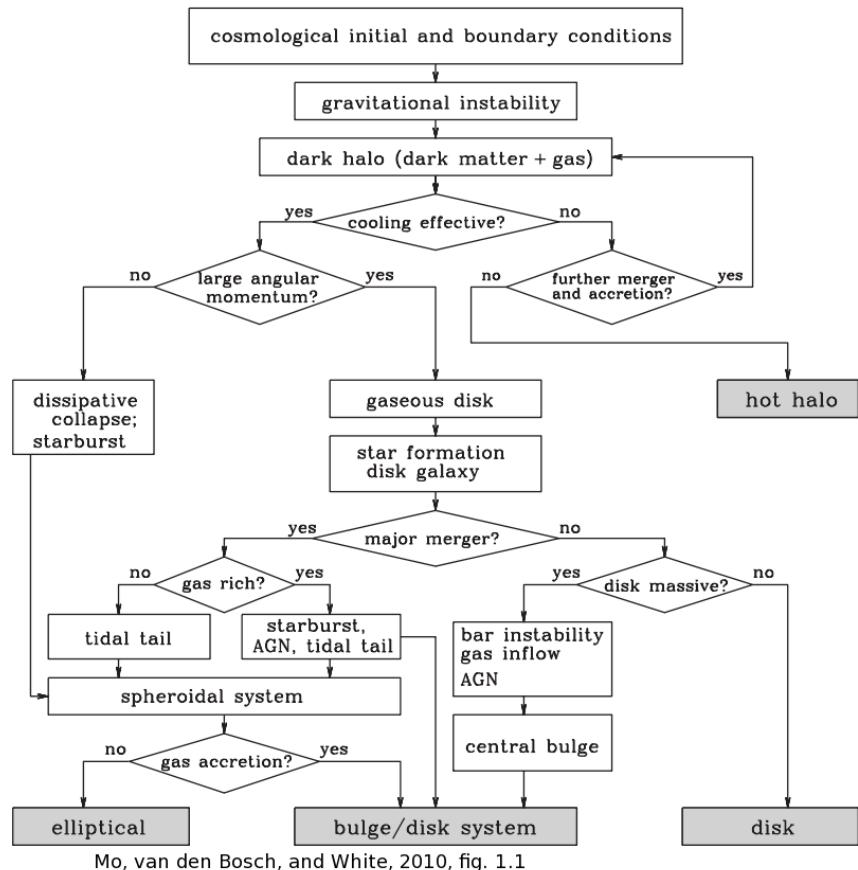
Figure 1.3 (Mo et al. 2010) shows a schematic of the process of galaxy formation. The process begins with slight overdensities in the early universe which eventually form galaxies. The white boxes in the flowchart represent the physical processes at work throughout the process of galaxy formation. The gray boxes show the final products (and what is ultimately observed today). By constraining the relative proportions of bulge, bulge+disk, and disk systems, we can set limits on the relative importance of the different pathways though the formation and evolution process.

Models of galaxy formation must agree with observations for the models to be valid. Thus, updating the local observations can lead to additional constraints on models. Figure 1.4 (Cole et al. 2000) shows many predicted quantities produced by semi-analytical models of galaxy formation. All of these measured quantities (perhaps with the exception of clustering) require accurate photometric measurements of the galaxy size, luminosity, and morphology. Therefore, accurate measurements are essential to improving these models.

Models also predict B/T values and fractions of elliptical galaxies (e.g. Wilman et al. 2013), and size-luminosity functions (e.g. Shankar et al. 2010a). So, accurate measurements of these properties in a large sample of low redshift galaxies could further constrain models of galaxy formation and evolution.

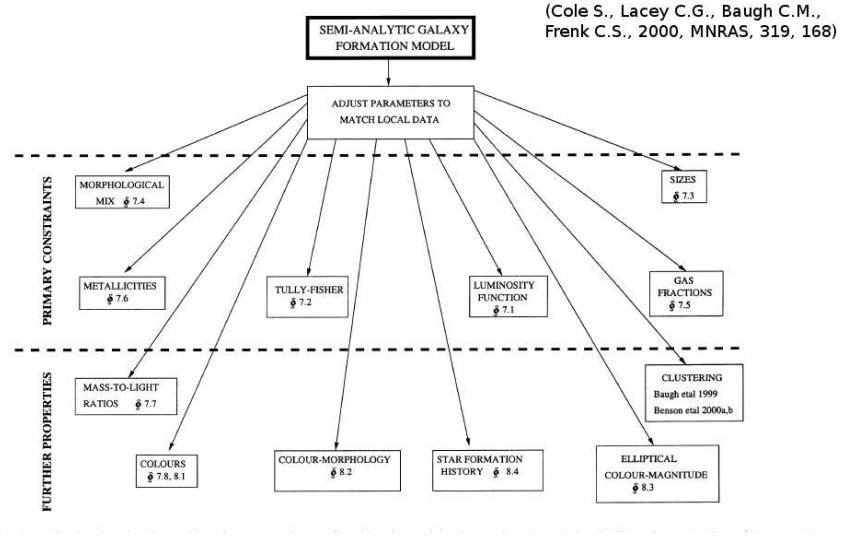
## 1. INTRODUCTION

---

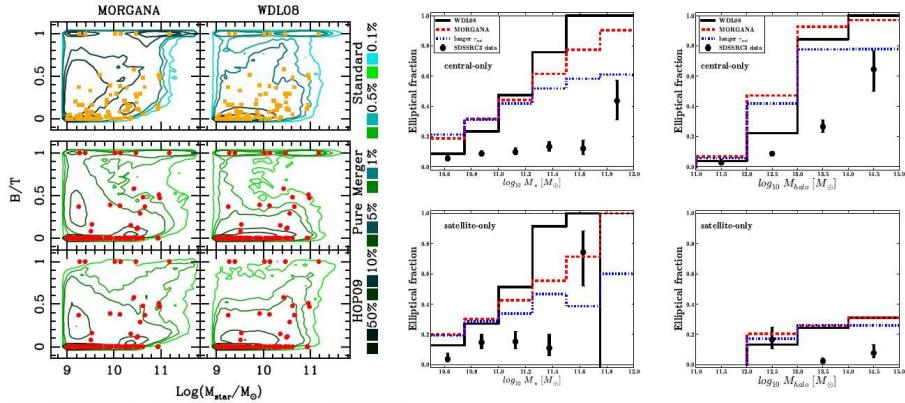


**Figure 1.3:** This figure is taken from Mo et al. (2010) and shows the schematic process of galaxy formation from the initial overdensities in the early universe through to present day.

## 1.4 Constraining models of formation and evolution



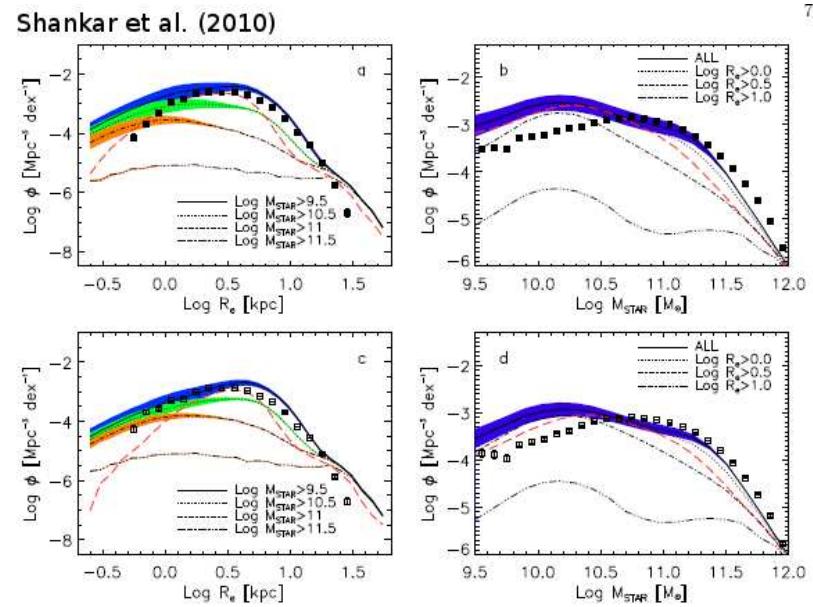
**Figure 1.4:** The predicted quantities of semi-analytical models of galaxy formation. The vast majority of these predictions are tested by making accurate measurements of photometric luminosity, size, and morphology, making such studies extremely important.



**Figure 1.5:** The predicted B/T and elliptical fraction from Wilman et al. (2013)

## 1. INTRODUCTION

---



7

**Figure 1.6:** Predicted size-luminosity relations from Shankar et al. (2010a).

## 1.5 Testing models of galaxy formation and evolution

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 and Jackson 1976), Kormendy (Kormendy 1977), Tully Fisher (Tully and Fisher 1977), the Fundamental Plane (Djorgovski and 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: color, metallicity, gas fraction, central velocity dispersion (e.g. Bernardi et al. 2003b,c,d; Kauffmann et al. 2003; Tremonti et al. 2004). Properties may also trace halo size and galaxy environment and place constraints on  $\Lambda$  cold dark matter cosmology (e.g. Bernardi 2009; Blanton et al. 2005a; Kravtsov et al. 2014; Shankar et al. 2010a,b). However, careful estimation of structural parameters for large numbers of galaxies is required to test different formation and evolution models. Testing galaxy models requires a large sample of galaxies containing photometric, spectroscopic, and morphological data.

The Sloan Digital Sky Survey (SDSS) has already provided a sample of many millions of nearby galaxies. Future surveys like the Dark Energy Survey (DES) and Large Synoptic Survey Telescope (LSST) 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 analyze these data sets at a reasonable rate, making it possible to perform time-intensive analysis, like galaxy decompositions, on large data sets. These decompositions can be used to make important comparisons between data and theory.

There has been recent work on improving photometric decomposition of nearby galaxies (Gadotti 2009; Häufler et al. 2013; Kelvin et al. 2012; Lackner and Gunn 2012; Simard et al. 2011). The approach is largely the same for these analyses. However, the different fitting processes continue to produce different results. The applicability of such fits is often questioned, particularly when multiple components are fitted (i.e., bulge+disk+bar etc.).

## 1. INTRODUCTION

---

This thesis presents a catalog of two-dimensional, 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. In Part I, fits are presented for the SDSS *g*-, *r*-, and *i*-bands utilizing the fitting routine GALFIT and analysis pipeline PYMORPH. This catalog is one of the largest galaxy samples for which structural decompositions have been performed on SDSS galaxies. Simulations are used to justify the choices made during fitting, and comparisons are made to previous works. Part II presents initial applications of the catalog to test and update important predictions of formation models, namely the luminosity function (Chapter 10) and the size-luminosity relation (Chapter 11). Part III briefly describes future work to be done with the catalog and some concluding remarks.

Parts of the catalog have been published in Meert et al. (2013) and Meert et al. (2015). The results of the catalog have been applied to several tests of galaxy formation and evolution. Shankar et al. (2013) tested semi-analytical modeling of hierarchical formation. Huertas-Company et al. (2013b) examined environmental effects on the size of galaxies. Bernardi et al. (2013) analyzed the uncertainty in the bright end of the mass and luminosity functions (LF). Bernardi et al. (2014) also examined the biases automated decompositions impose 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 other catalogs.

## Part I

# A catalog of DR7 spectroscopic galaxies

# Chapter 2

## Catalog Selection

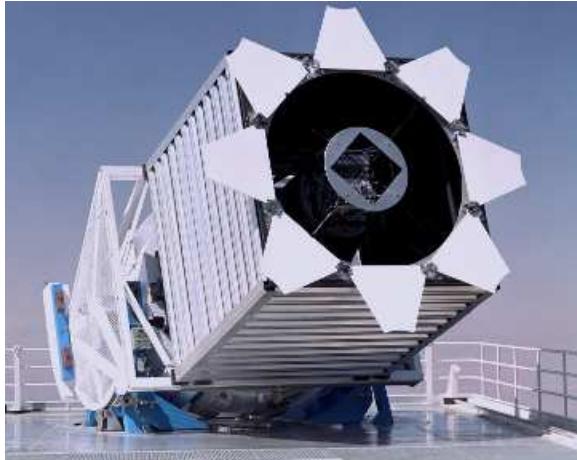
Using a well-understood sample is essential for testing evolutionary models of galaxy formation. It prevents biases from distorting the population of galaxies. If this is not possible then we should at least have a good estimate of the biases in our sample and some idea of how to correct for them.

### 2.1 The Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS) uses a single, dedicated 2.5 meter telescope located at Apache Point Observatory in New Mexico. The SDSS consists of two completed campaigns (SDSS-I and SDSS-II) as well as a current, ongoing campaign (SDSS-III). Figure 2.1 shows the telescope open and ready for observation, although the sky is light enough that such observations would not be possible. Gunn et al. (1998) describes the photometric camera in detail, and Gunn et al. (2006) describes the telescope itself.

The focalplane of the SDSS telescope is pictured in Figure 2.2. There are 6 columns (referred to as the “camCol”) in the imaging plane. Each column has an imaging CCD for each of the filter bands used in the survey. The telescope is operated in drift-scan mode for

## 2.1 The Sloan Digital Sky Survey



**Figure 2.1:** The SDSS telescope at Apache Point Observatory. Image Credit:Fermilab Visual Media Services and Sloan Digital Sky Survey

photometric observations. This means that targeted objects slowly drift across the imaging plane during observation. Software allows the telescope to compensate for this motion.

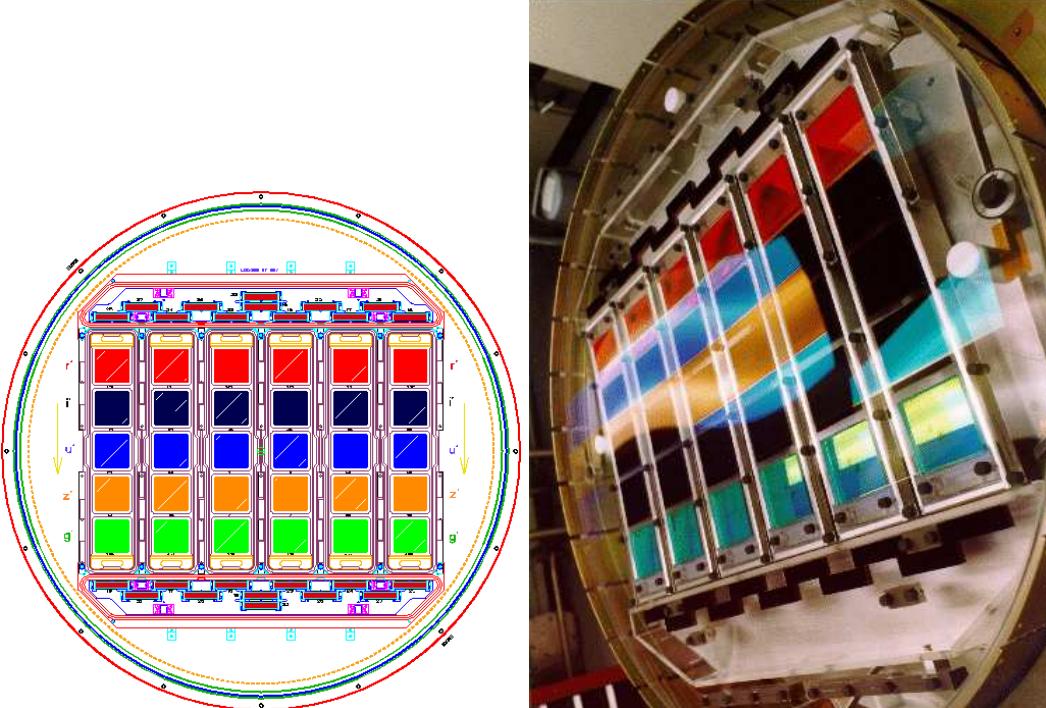
SDSS-I took place from 2000 through 2005 and is fully contained in the fifth data release (Adelman-McCarthy et al. 2007, DR5). The Early Data Release contains much useful information on the survey methods, design, and analysis pipelines (Stoughton et al. 2002, EDR). SDSS-I covered about 8000 sq. degrees of the sky, surveying millions of galaxies, hundreds of thousands of Quasars (QSOs) as well as many stars and other galactic objects. Survey data was taken in five bands, the  $u$ ,  $g$ ,  $r$ ,  $i$ , and  $z$  bands, spanning wavelengths from near-UV to near-IR. The images are 53.9 second exposures with platescale of 0.396 arcsec/pixel.

Figure 2.3 shows approximate transmission curves in vacuum for the SDSS filters.

Spectroscopic data was also obtained for up to  $\sim 600$  objects at a time using two fiber-fed spectrographs with moderate resolution ( $R$  1800 – 2200) and is detailed in Strauss et al. (2002). The spectroscopic survey includes galaxies, QSOs, and stars. For galaxies,

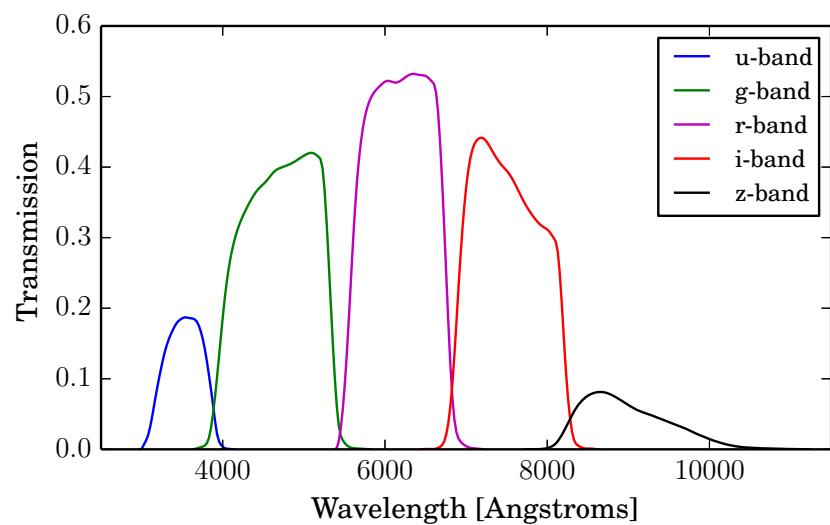
## 2. CATALOG SELECTION

---



**Figure 2.2:** An image and diagram of the SDSS focalplane showing the 6 camera columns (know as camCol). Each camCol contains a 2048x2048 chip for each of the  $u$ ,  $g$ ,  $r$ ,  $i$ , and  $z$  filters. Images are taken as objects drift across the focalplane, passing over each of the filtered chips in a camCol. Image Credit:Fermilab Visual Media Services and Sloan Digital Sky Survey

## 2.1 The Sloan Digital Sky Survey



**Figure 2.3:** The photometric SDSS filter transmission curves in vacuum.

## 2. CATALOG SELECTION

---



**Figure 2.4:** A spectroscopic plate used during the spectroscopic survey mounted into the telescope housing with drilled holes at observation locations and optical fibers inserted directing light to the spectrograph for observation. Each plate was drilled after spectroscopic targets selection and survey strategy were completed. Insertion of the fibers into these plates were done by-hand for each plate by workers on-site. Image Credit:Fermilab Visual Media Services and Sloan Digital Sky Survey

target selection is complete down to an r-band extinction corrected magnitude of 17.77 mags and a surface brightness of 23.0 mags/arcsec<sup>2</sup>. Figure 2.4 shows the spectroscopic plates with drilled hole and optical fibers inserted for observation. Each plate was drilled after spectroscopic targets selection and survey strategy were completed. Insertion of the fibers into these plates were done by-hand for each plate by workers on-site.

SDSS-II took place from 2005 through 2008 and is fully contained in the seventh data release (hereafter DR7; Abazajian et al. 2009). SDSS-II had three separate goals:

**The Sloan Legacy Survey:** The Sloan Legacy Survey completed the original SDSS imaging and spectroscopic goals. The final catalog contains 230 million celestial objects covering 8,400 square degrees of the sky and spectra of 930,000 galaxies, 120,000 quasars, and 225,000 stars.

**SEGUE:** The Sloan Extension for Galactic Understanding and Exploration (SEGUE)

---

## 2.1 The Sloan Digital Sky Survey

probed the structure and history of the Milky Way galaxy. It contains new imaging of 3500 sq. degrees and spectra of 240,000 stars covering a range of spectral classes and sizes.

**The Sloan Supernova Survey:** The Sloan Supernova Survey carried out repeated imaging of the 300 square degree southern equatorial stripe, referred to as “Stripe 82,” to discover and measure supernovae and other variable objects. In the course of three 3-month campaigns, the supernova survey discovered nearly 500 spectroscopically confirmed Type Ia supernovae, which are still being used to constrain cosmological models over the last 4 billion years.

Each section of the sky is identified by a run, rerun, and field number which uniquely identifies its location on the sky. The exact survey strategy used to collect the data is not relevant this study, so no more comments are made regarding it.

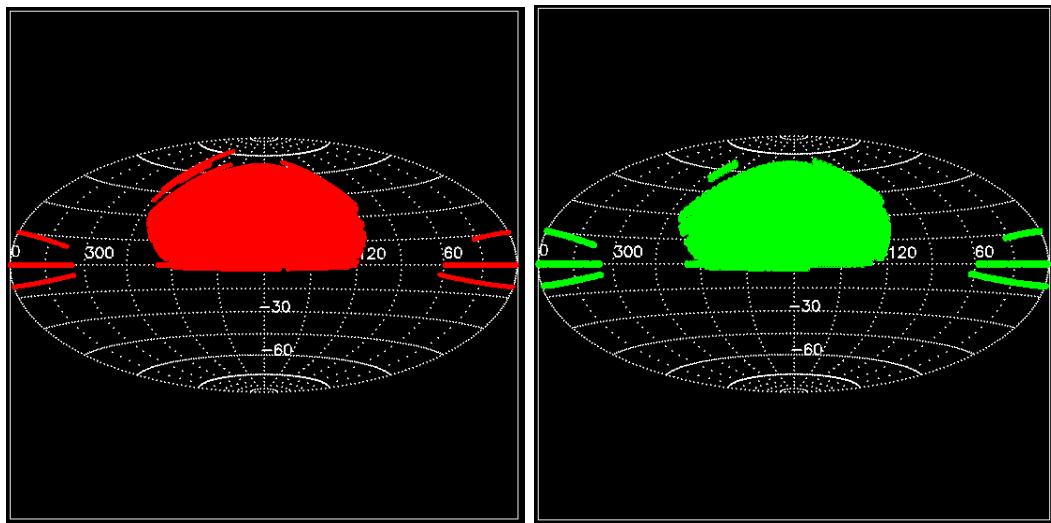
More information on SDSS-I and II are available through the original SDSS website:  
<http://www.sdss.org>.

SDSS-III is currently ongoing. It has several constituent surveys. The goals of these surveys are to constrain the Baryon Acoustic Oscillations using spectroscopic information on galaxies out to larger redshifts than the original SDSS Legacy survey. It also has several projects associated with Quasars, galactic objects, and Exoplanets. The current data release is DR10 (Ahn et al. 2014) with more data releases planned. More information can be found at <http://www.sdss3.org>.

SDSS was revolutionary for many reasons including the large sky coverage, the digital nature of its imaging, the public data releases, the size of the collaboration, and online data service made available to astronomers. This project uses the completed Legacy Survey from SDSS-II and DR7. Work is restricted to the  $g$ -,  $r$ -, and  $i$ -bands, which have the best signal-to-noise, and use galaxies from the spectroscopic survey. The details of this selection are described in the next section. Figure 2.5 shows the sky coverage for the legacy survey area in DR7 of the photometric (in red) and spectroscopic (in green) portions of the survey.

## 2. CATALOG SELECTION

---



**Figure 2.5:** The photometric (in red) and spectroscopic (in green) sky coverage of the SDSS DR7 legacy survey. The coordinates are given in right ascension (longitudinal) and declination (latitudinal) coordinates. Image Credit:Fermilab Visual Media Services and Sloan Digital Sky Survey

## 2.2 SDSS CasJobs Data

The data used in this analysis were drawn from the spectroscopic main galaxy sample of the DR7 Legacy Area. The spectroscopic sample provides a well-established sample with well-defined and tested selection criteria. The criteria are presented in Strauss et al. (2002).

The main galaxy sample consists of galaxies with  $r$ -band Petrosian magnitudes  $r \leq 17.77$  and  $r$ -band Petrosian half-light surface brightnesses  $\mu_{50} \leq 24.5 \text{ mag arcsec}^{-2}$ . The completeness of the sample exceeds 99 per cent. However, About 6 per cent of galaxies that satisfy the selection criteria are not spectroscopically observed because they have a neighboring target closer than 55 arcseconds, which is the minimum spectroscopic fiber separation.

Galaxies from the main sample have both spectroscopic and photometric information and are listed in both the `PhotoObj` and `SpecObj` tables of the CasJobs DR7 data base. Galaxies that satisfy three main selection criteria were selected. Those criteria were: (1) the extinction-corrected  $r$ -band Petrosian magnitude was 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`). The limit at the faint end (17.77 mag in the  $r$ -band) corresponds to the the lower limit for completeness of the SDSS Spectroscopic Survey (Strauss et al. 2002). The limit of at the bright end (14 mag) 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 catalog. 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 neighboring images (see Blanton et al. 2011, for details). In addition, this magnitude cut corresponds to the cut used in Simard et al. (2011, hereafter S11) which is used throughout this work as a benchmark for comparison.

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
```

## 2. CATALOG SELECTION

---

```
= 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.
```

Additional cuts similar to Shen et al. (2003) and S11 to the data prior to fitting. All galaxies with redshift < 0.005 (1647 galaxies) are removed. These galaxies have large apparent sizes and resolved structure that make decomposition difficult. In addition, the redshift may be contaminated by peculiar velocities, making the redshift-dependent distances unreliable.

Twenty galaxies with redshift > 1.0 were also removed. Visual inspection reveals that these galaxies likely represent catastrophic failures in the redshift code. Catastrophic redshift errors may exist at even lower redshift, and there are an additional 66 galaxies with  $0.4 < z < 1.0$  that we include out catalog. These galaxies may have the same problem.

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), a surface brightness cut on the mean surface brightness within the Petrosian half-light radius of  $\mu_{50, r} < 23.0$  mag, where

$$\mu_{50, r} = m_{\text{petro}, r} + 2.5 \log(2\pi r_{\text{petro}, 50}^2) \quad (2.1)$$

is necessary 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. This sample is used throughout the project and the cuts described here are considered in the sample completeness discussed in Chapter 9.

Figure 2.6 shows the surface brightness distribution, redshift distribution, extinction-corrected Petrosian magnitude, Petrosian half-light radius,  $V_{\max}$ -weighted luminosity function, and signal-to-noise distribution of the sample used in this work. The S/N is defined

## 2.2 SDSS CasJobs Data

---

as the mean pixel flux within the half-light radius divided by the noise associated with that pixel, or

$$\begin{aligned} \left(\frac{S}{N}\right) &\equiv \frac{I_{\mu_{50}}}{N_{avg}} \\ N_{avg} &\equiv \sqrt{\frac{I_{\mu_{50}} + I_{sky}}{\text{gain}} + \text{dark variance}} \\ I_{\mu_{50}} &= 10^{-0.4(\mu_{50} - zp)} \times \text{platescale}^{-2} \end{aligned} \quad (2.2)$$

where  $I_{\mu_{50}}$  is the source DN ('data numbers' or, equivalently, counts)<sup>1</sup> of the average surface brightness defined in Equation 2.1. The zeropoint,  $zp$ , is calculated from the SDSS zeropoint, 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.  $N_{avg}$  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.

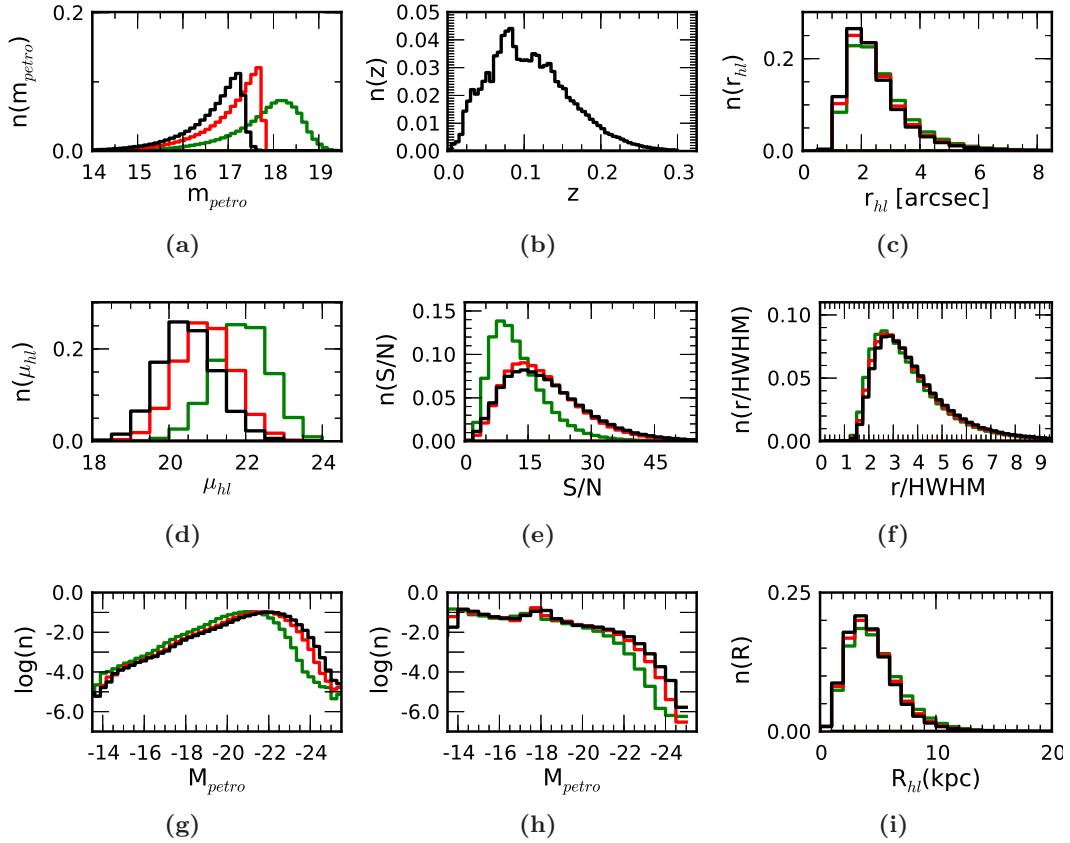
Figure 2.6 shows that the surface brightness (top left panel) is brightest in the *i*-band and dimmest in the *g*-band. The sizes of galaxies in each band (bottom row, left panel) are similar, although the *g*-band is slightly larger than the *r*-band on average. The *i*-band is slightly smaller than the *i*-band. This trend follows the same trend in average PSF width seen across the bands (1.47 vs 1.35 vs 1.28 arcsec for the *g*-, *r*-, and *i*-bands, respectively). The large difference in surface brightness is due to substantial differences in the total magnitudes of the objects (top row, right panel). However, the S/N is most relevant to fitting. The S/N (bottom row, right) shows that the *i*-band has S/N similar to, and perhaps slightly higher than, the *r*-band. Although the *i*-band galaxy images are brighter and the same size as the *r*-band galaxy images, which would imply a higher S/N. However, the sky brightness is also  $\approx 50\%$  brighter in the *i*-band. The *g*-band has comparably lower average S/N relative to the *r*-band. The average *g*-band sky is dimmer by 1 magnitude relative to the *r*-band,

---

<sup>1</sup>The counts are related to the number of photo electrons collected by the detector through the gain of the detector amplifier. This distinction is important since the photo electrons obey Poisson statistics.

## 2. CATALOG SELECTION

---



**Figure 2.6:** Distributions of (a) extinction-corrected Petrosian magnitude, (b) redshift, (c) Petrosian half-light radius, (d) surface brightness, (e) Signal-to-Noise using the the measurement of sky provided by the SDSS photometric pipeline, (f) PSF size, (g) absolute Petrosian magnitude, (h)  $V_{\text{max}}$ -weighted luminosity function, and (i) physical size, of the samples drawn from the DR7 spectroscopic galaxy sample. Bin counts are normalized to integrate to 1. The  $i$ - and  $r$ -bands are similar in many respects, although the S/N is larger for the  $i$ -band data. The  $g$ -band is different in many respects. It tends toward smaller sizes, lower S/N, and appears as complete rather than a magnitude-limited distribution of galaxies.

## 2.2 SDSS CasJobs Data

---

but the galaxy brightness is also dimmer on average. These differences, particularly in S/N must be accounted for when testing any algorithm intended to be used across all the fitted bands. Chapter 4.5 addressed this issue for the simulations used to test the work presented here, and Chapter 5.5 addresses the changes necessary to adapt the flagging developed in the *r*-band to work in the *i*- and *g*-bands. Differences in S/N are most concerning in the case of the **Ser-Exp** and **deV-Exp** fits where two-components are being simultaneously fit. Lowering the S/N of each component (because the total flux is now divided between the two fitted components) makes accurate measurement and decomposition difficult.

It is also potentially useful to use information about each band to inform the fits of neighboring bands. Indeed (Simard et al. 2011) attempted this by requiring many parameters (i.e., Sérsic index, radius, ellipticity) of the fitting model to be identical across the *g*- and *r*-bands, essentially using the two bands as a form of co-added data to increase the S/N. This increase of S/N comes at the expense of dis-allowing variation in the matched parameters, which may or may not be an appropriate assumption (i.e., in a two-component fit, we might expect the bulge size to change across bands, which is dis-allowed). Additionally, Häußler et al. (2013) enforced simple polynomial relationships in parameters across bands, using the neighboring bands to further constrain the acceptable parameter space to be searched by the fitting algorithm.

The most flexible method is to fit each band independently and examine the systematic effects of each band as necessary, making additional cross-band comparisons including color (for example, see Lackner and Gunn 2012). This is the method adopted for the data presented here.

All of the identifying data, photometric measurements, and spectroscopic measurements obtained from the SDSS CasJobs server were collected and stored in the table named the “CasJobs Table” distributed with this work. For convenience, ID numbers were assigned to all galaxies contained in the catalog to be used in place of the SDSS ObjID. These ID numbers are referred to as GALNUM and used throughout the available data products

## 2. CATALOG SELECTION

---

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, all necessary fpC images and PsField files from were downloaded 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 neighborhood. Individual postage stamps for each galaxy were extracted 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 Chapter 3.1.

### 2.3 Additional Data

In addition to the data provided by CasJobs, parameters were collected from a number of other studies. The Morphology catalog of Huertas-Company et al. (2011, hereafter H2011) is an automated morphological classification that used a Bayesian SVM algorithm to classify all galaxies in the spectroscopic sample based on data available as part of the SDSS DR7. Objects are matched to our catalog based on the SDSS DR7 `objID`. The morphological parameters from the matching catalog (H2011) for our sample is described in the electronic catalog.

#### 2.3.1 Derived Parameters

$K$ -corrections, distance modulus, angular diameter distance, and  $V_{\text{Max}}$  corrections are calculated for each galaxy in order to make account for observational effects on the data. These calculations assume a flat cosmology with  $(H_0, \Omega_\Lambda, \Omega_m) = (70 \text{ km s}^{-1}\text{Mpc}^{-1}, 0.7, 0.3)$ .

### 2.3.2 K-Corrections

$K$ -corrections are calculated using version 4.2 of the  $K$ -correction code KCORRECT described in Blanton and 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.

### 2.3.3 Rest-Frame Conversion Parameters

The distance modulus, DM, and angular diameter distance,  $D_A$ , are calculated in Mpc according to the usual equations:

$$D_M(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda}}, \quad (2.3a)$$

$$D_A(z) = \frac{D_M}{1+z}, \quad (2.3b)$$

$$D_L(z) = (1+z)D_M, \quad (2.3c)$$

$$\text{DM} = 5\log(D_L) + 25 \quad (2.3d)$$

using the notation given in Hogg (1999). The angular diameter distance is reported in the DERT in units of kpc per arcsec to allow for easy conversion between apparent and physical sizes of galaxies.

### 2.3.4 $V_{\text{Max}}$ Correction for Selection Effects

A  $V_{\text{Max}}$  term is also calculated for each galaxy. This represents the observable volume given the constraints on magnitude, surface brightness, and redshift used to construct the catalog. Volume corrections are calculated similar to that of Shen et al. (2003) (i.e., we account for all of the catalog cuts when calculating the observed volume for any galaxy in the catalog). As discussed in Chapter 2.2, the SDSS spectroscopic survey is complete down to a galactic extinction corrected Petrosian  $r$ -band magnitude of 17.77 and an extinction

## 2. CATALOG SELECTION

---

corrected Petrosian  $r$ -band surface brightness of 23.0 mag/arcsec $^2$ . Additionally, a minimum redshift cut of  $z_{\min} > 0.005$  and a maximum redshift cut of  $z_{\max} < 1.0$  were imposed. The upper redshift cut removes essentially no galaxies and changes the  $V_{\text{Max}}$  value of less than 10 galaxies. In addition, a minimum magnitude cut was included to exclude bright galaxies. Therefore, the volume in which any particular galaxy is observable must account for these effects.

The lower and upper bounds of redshift that limit this volume are determined by the conditions imposed on the data. The magnitude limits in the  $r$ -band,  $m_{r,\min} \leq m_{r,\text{obs}} \leq m_{r,\max}$  correspond to a lower and upper limit on the redshift at which the galaxy remains observable through the luminosity distance.

$$D_L(z_{\max,m}) = D_L(z)10^{-0.2(m_r - m_{r,\max})}, \quad (2.4a)$$

$$D_L(z_{\min,m}) = D_L(z)10^{-0.2(m_r - m_{r,\min})} \quad (2.4b)$$

The surface brightness limit constrains  $V_{\text{Max}}$  mainly through the  $(1+z)^4$  cosmological dimming effect. The maximum redshift from which a galaxy with mean surface brightness  $\mu_{50,r}$  at redshift  $z$  can be detected at the surface brightness limit of  $\mu_{50,r} = 23$  is given by

$$z_{\max,\mu} = (1+z)10^{(23.0 - \mu_{50,r})/10} - 1 \quad (2.5)$$

As is mentioned by S11, our selection criteria do not include a cut on size of the target galaxy, so the maximum redshifts may differ from that of Shen et al. (2003). The maximum and minimum redshifts for a galaxy in this work are then

$$z_{\min} = \max(z_{\min,m}, 0.005), \quad (2.6a)$$

$$z_{\max} = \min(z_{\max,m}, z_{\max,\mu}, 1.0) \quad (2.6b)$$

With the above limits on redshift, the comoving volume in which a target galaxy satisfies the sample selection criteria is:

$$V_{\text{Max}} = \int d\Omega f(\theta, \phi) \int_{z_{\min}(\theta, \phi)}^{z_{\max}(\theta, \phi)} \frac{(1+z)^2 D_A^2}{H(z)} dz \quad (2.7)$$

---

### 2.3 Additional Data

where  $H(z)$  is the Hubble parameter at redshift  $z$  and  $c$  is the speed of light.  $f(\theta, \phi)$  is the sampling fraction of point  $(\theta, \phi)$  on the sky, and  $\Omega$  is the solid angle. Following S11,  $f(\theta, \phi)$  is treated as constant. In the case of a flat universe, Equation 2.7 reduces to

$$V_{\text{Max}} = \frac{4\pi f_{\text{sky}}}{3} (D_M(z_{\text{max}})^3 - D_M(z_{\text{min}})^3) \quad (2.8)$$

where  $f_{\text{sky}} = 0.195$  is the fraction of the full sky covered by SDSS (8032 square degrees coverage for the SDSS DR7 Legacy survey out of 41253 square degrees in the sky). Calculated values of  $V_{\text{Max}}$  are provided as part of the data released with this work.

# Chapter 3

# Fitting Process

This chapter describes the fitting process. Section 3.1 describes the 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 and its components used to fit the catalog. Finally, Section 3.4 describes the masking and neighbor identification process used in this work.

## 3.1 Preprocessing 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 as described in Chapter 2.2. 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 center of the image to the edge) and centered on the target source.

The decision to cut at 20 half-light radii is justified in chapter 4.2.5 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. With

### 3.1 Preprocessing of SDSS Images

---

these settings, the estimate of background sky brightness is accurate to  $\sim 0.1\%$  with a slight bias toward underestimating the background level using this stamp size. This sky bias does not noticeably bias the other fitted parameters.

A smaller cutout size could potentially be used when fitting the galaxy and would reduce the time needed to fit each galaxy. Figure 4.18 and the related discussion show that the effect of further reducing the cutout size is insignificant to estimates of the total magnitude, half-light radius, and sky brightness. However, the original 20 half-light radii cutout sizes are used throughout the rest of this work.

The PSF is extracted from the PsField files using the READATLASIMAGES-v5\_4\_11 program distributed on the SDSS website<sup>2</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 the 1000 DN soft bias is removed from the images and PSF. Sigma images are created from the SDSS image cutout following the standard deviation calculation<sup>3</sup>

$$W_{i,j} = \sqrt{\frac{F_{i,j}}{\text{gain}}} + \text{dark variance} \quad (3.1)$$

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 photo electrons (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, the postage stamp and sigma images are normalized to a 1 s exposure prior to fitting. A simple diagram of this process is shown in Figure 3.1.

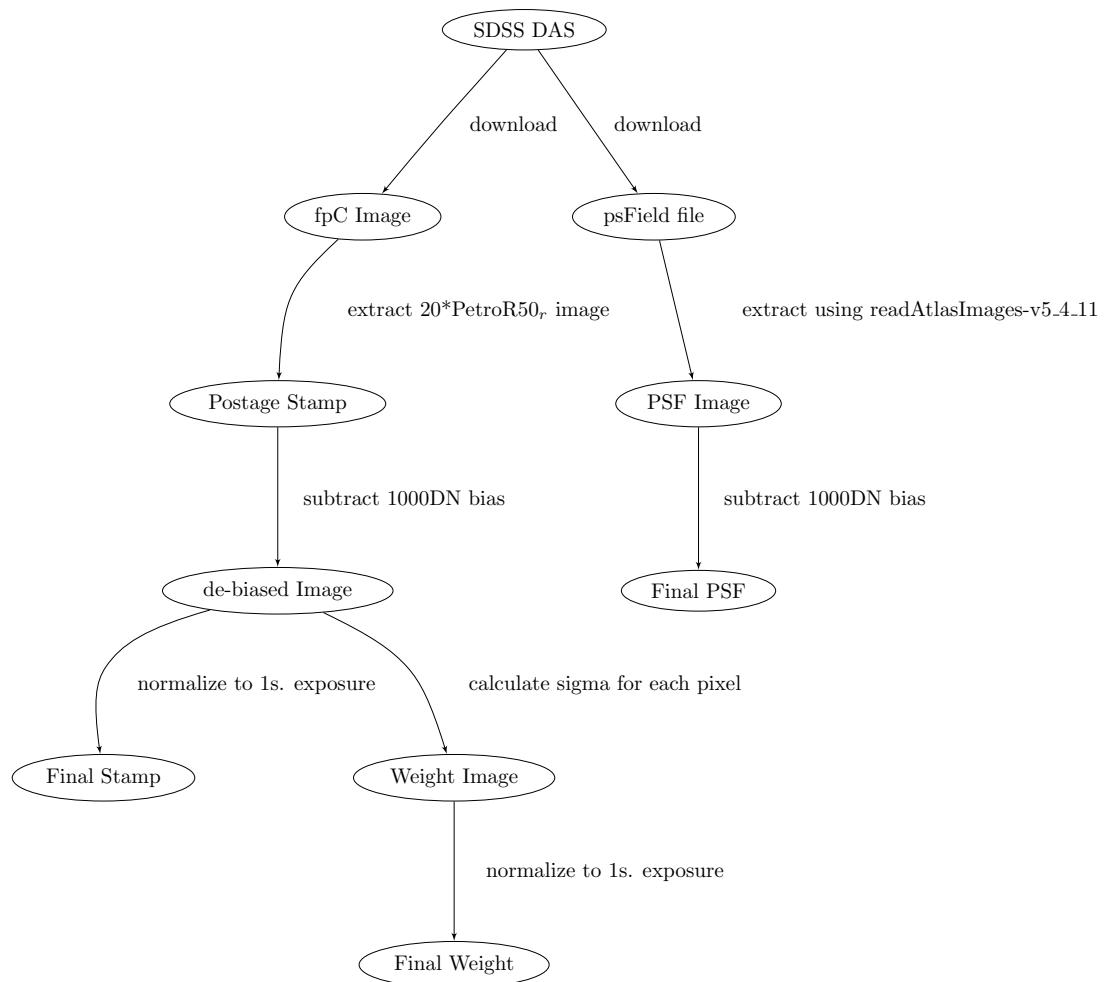
---

<sup>2</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>3</sup>See the SDSS DR7 online documentation at <http://www.sdss.org/dr7/algorithms/fluxcal.html> for further discussion.

### 3. FITTING PROCESS

---



**Figure 3.1:** The flowchart of the preprocessing done prior to fitting with PYMORPH.

## 3.2 The Fitted Profiles

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

$$I(r) = I_e \exp \left( -b_n \left[ \left( \frac{r}{R_e} \right)^{\frac{1}{n}} - 1 \right] \right) \quad (3.2)$$

$$b_n = 1.9992n - 0.3271,$$

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 and 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 center (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 modeled using the function

$$I_{Exp}(r) = I_d \exp \left( \frac{-r}{R_d} \right) \quad (3.3)$$

The profiles defined in equations 3.2 and 3.3 are 1D profiles. The 1D profiles are used to generate 2D models by also fitting a center on the image, position angle ( $\phi$ ), and axis ratio ( $b/a$ ) to each component.

Equations 3.2 and 3.3 are often interpreted as representing a bulge and a disc, respectively. However, fitting two-component models to galaxies does not guarantee that the two

### **3. FITTING PROCESS**

---

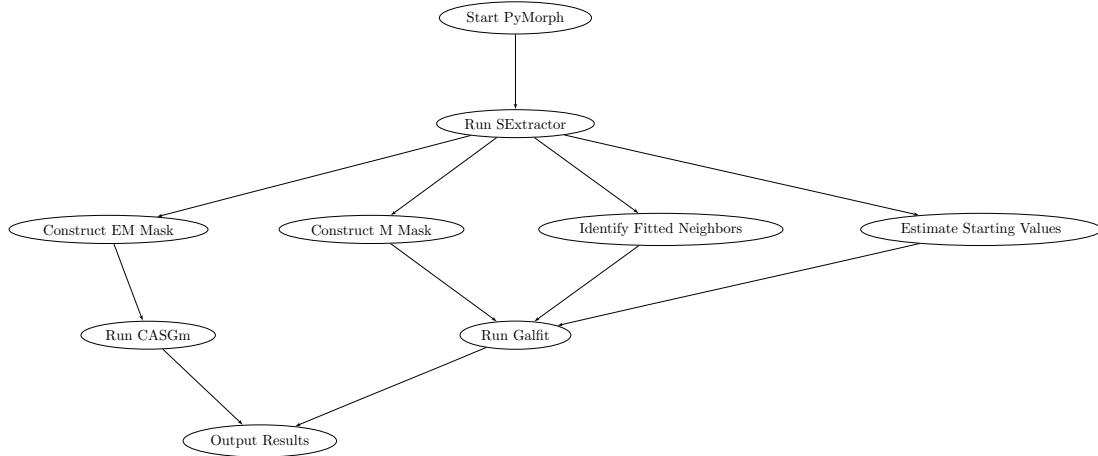
components measured are truly present. Many early-type galaxies show no signs of disc-like structures. Similarly, many late-type galaxies show little or no sign of a bulge in the central part of the galaxy. Also, when fitting multiple components, a significant second component may only indicate substantial departure from a single component profile rather than the presence of a physically meaningful second component. For example, Gonzalez et al. (2005), Donzelli et al. (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. The effects of bar presence is examined in Chapter 8.4.

Judgment on the physical interpretation of the components will be reserved until after discussion of the flagging system in Chapter 5. However, there are many cases in which the components should not be interpreted as a physically meaningful bulge and disc and this is an important consideration in this work.

The Sérsic model intended to represent the central region of two-component galaxies has broad wings but 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. Cases where the Sérsic component dominates at large radii are considered during the flagging procedure described in Chapter 5 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

---



**Figure 3.2:** The flowchart of the internal PYMORPH processes.

## 3.3 The Fitting Pipeline

This work fits both one- and two-component models to the sample described in Chapter 2.2 using PYMORPH (Vikram et al. 2010). PYMORPH is a Python based automated software pipeline built on SExtractor (Bertin and 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). After fitting, additional measurements used during flagging are performed.

### 3.3.1 PyMorph

PYMORPH was released in Vikram et al. (2010) and is written in PYTHON. The pipeline was developed to automate the process of target identification, masking, generation of GALFIT configuration files, running GALFIT, making non-parametric measurements, and collecting output data. A simple diagram of the internal PYMORPH process is shown in Figure 3.2.

### 3. FITTING PROCESS

---

PYMORPH can be run on full telescope images in a mode where objects are identified, selected and fit separately in a fully automated way. For this work, PYMORPH was provided with the postage stamp images of each galaxy, weight image, and PSF extracted as described in Chapter 3.1.

PYMORPH runs SExtractor and uses the output catalog to construct the image mask for fitting prior to running GALFIT on the postage stamp image.

#### 3.3.2 SExtractor

This work uses version 2.8.6 of SExtractor (Bertin and Arnouts 1996). The SExtractor code provides automated detection and photometry of sources in telescope images. The code can provide many measurements of photometry. The measurements reported by SExtractor are used as a starting point for the GALFIT code and need only be approximate. Because of this, the accuracy and appropriateness of specific SExtractor photometry choices will not be discussed here.

SExtractor works in a series of steps. First, it determines the background and separates pixels belonging to background from objects. Then it segments any area that is not background into separate objects and determines the properties of each object. The output of this process is written to a catalog. Optionally, the code can output images with object boundaries, pixel-by-pixel background estimation or other output. In addition to producing the list of objects with a series of measured parameters (ellipticity, size etc.) each object is classified into stars and galaxies (i.e., everything non-star-like) by a neural network.

For the background determination, all the pixels with intensity above a given threshold are treated as belonging to an object. Based on the structure of the intensity distribution and the settings for deblending, any given object may be split into two separate objects. Photometry is done on these by dividing up the intensity of the shared pixels.

Background is estimated using the BACK\_SIZE parameter. Within each area of size BACK\_SIZE the mean and standard deviation of the background is iteratively estimated

### 3.3 The Fitting Pipeline

---

with pixels deviating by more than  $3\sigma$  being pruned at each step until convergence.

DETECT\_THRESHOLD sets the minimum flux per pixel required for an object to be detected by SExtractor. In addition, all pixels of a detected object must be adjacent to each other and the total area must be as large or larger than DETECT\_MINAREA. For this work, DETECT\_MINAREA is set to 6 pixels and DETECT\_THRESHOLD is set to 1.5.

DEBLEND\_MINCONT and DEBLEND\_NTHRESH are the most relevant parameters for object deblending. SExtractor deblends objects by iteratively dividing contiguous groups of pixels that are above the background level into two objects brighter than a given threshold with a region below the threshold separating the two objects. The number of thresholds is set by DEBLEND\_NTHRESH (32 in this work), which are spaced exponentially between the DETECT\_THRESHOLD and the maximum pixel brightness in the object. In order to separate a single object into two objects at a given threshold, both objects must contain at least DEBLEND\_MINCONT or larger fraction of the light in the blended object. If either or both objects fail this criteria, then deblending does not split the object.

Using these settings, SExtractor identifies neighboring objects (including both stars and galaxies) and measures approximate values of the center, size, and brightness for each object as well as the brightness of the sky background at the location of the object. These measurements are used by PYMORPH to mask neighbors and configure GALFIT.

#### 3.3.3 Galfit

The PYMORPH pipeline uses GALFIT (Peng et al. 2002) version 2.0.3c 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.

GALFIT uses a non-linear least-squares minimizer to fit the selected profiles to each galaxy. The minimizer uses a Levenberg-Marquardt minimization algorithm (Press et al.

### 3. FITTING PROCESS

---

1992) and is a “greedy” algorithm, meaning that it uses certain heuristics to speed up convergence of the minimization. A result of this is that it is susceptible to finding local minimums and does not necessarily find the global minimum.

The measure of goodness-of-fit is the *reduced chi-squared*,  $\chi_{\nu}^2$  :

$$\chi_{\nu}^2 = \frac{1}{N_{DOF}} \sum_{x=1}^{nx} \sum_{y=1}^{ny} \frac{(f_{data}(x, y) - f_{model}(x, y))^2}{\sigma(x, y)^2} \quad (3.4)$$

The minimization proceeds iteratively because of the non-linear nature of the profiles and the need to convolve the image with the PSF each time before calculating the goodness-of-fit. While minimizing the  $\chi_{\nu}^2$  is appropriate for the fitting process, the value produced does not necessarily provide an accurate measure of the probability of an appropriate fit. The unsuitable nature of the  $\chi_{\nu}^2$  is a major driver of developing the flagging process presented as part of this work. The  $\chi_{\nu}^2$  is discussed further in Chapter 5.1.

#### 3.3.4 Non-parametric measurements

In addition to photometric decomposition, PyMorph also supports measurement of several non-parametric structural parameters described in Conselice (2003) and Lotz et al. (2004). These include the concentration index ( $C$ ) defined as the ratio of the radius that contains 80 per cent of the galaxy light to the radius that contains 20 per cent, or

$$C = 5 \log \left( \frac{r_{80}}{r_{20}} \right). \quad (3.5)$$

Several other definitions of this value exist, but all are related to the Sérsic index for a single-component galaxy.

The asymmetry ( $A$ ) is measured by taking the difference of the image and the image rotated 180 degrees normalized by the original pixel flux. The clumpyness ( $S$ ) is measured by summing the residual of the original image and a boxcar-smoothed version of the image. Gini coefficient and second-order moment of the galaxy are also measured by the code.

## 3.4 Masking and neighbor identification

---

These non-parametric measures of galaxy structure have been used in the past to differentiate galaxy types. The values for these measurements are reported in the final catalog. However, caution is necessary when using these parameters because PSF effects cannot be removed from the measurements. As a result, there may be strong resolution effects, particularly since much of the galaxy sample is similar in size to the image PSF.

### 3.3.5 Post-fitting measurements

We also directly measure a total half-light radius and total axis ratio for all models using the image of the fitted model. To do this, after fitting we generate an unconvolved image of the fitted model using GALFIT. Using the image moments, we calculate the position angle and ellipticity of the galaxy and measure in elliptical annuli extending out until 50 per cent of the total light of the image are enclosed. This numerical measurement is found to agree with the actual radius of the galaxy within 1 per cent for single-component galaxies where a direct comparison between the analytical and computed half-light radii can be compared. These measurements allow direct comparison of the one- and two-component model half-light radii (See Chapter 6.4 and Figure 6.9, for example).

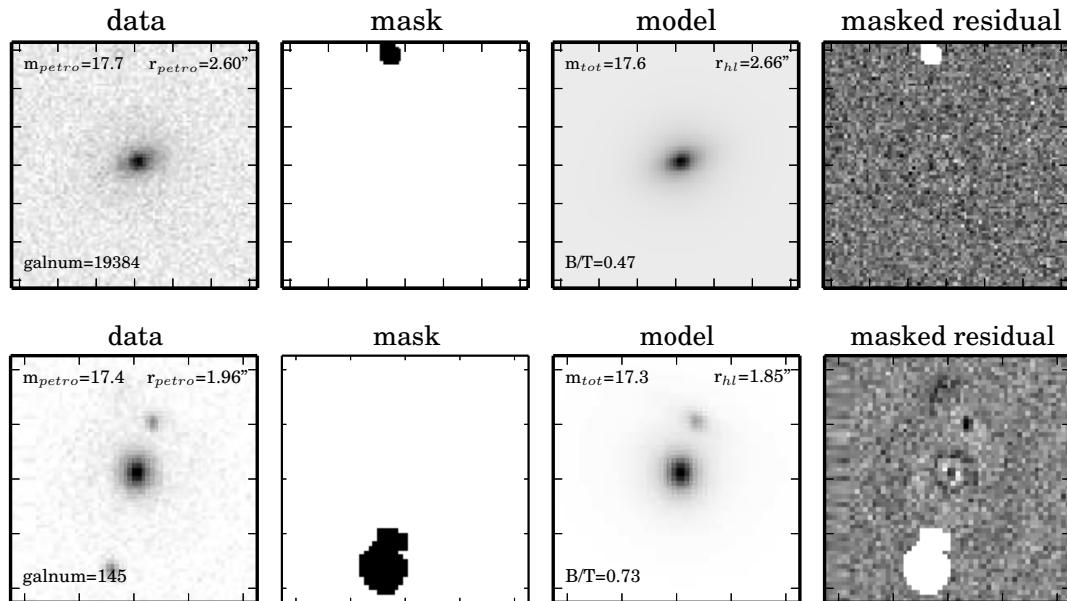
Figure 3.3 shows two example galaxy fits performed as part of this work using the PYMORPH pipeline. Ser-Exp fits are shown. Each column contains the input image (far left), mask (center left), fitted galaxy (center right), and the residual image (far right). The first row shows a well-fit galaxy. The second row shows an early-type galaxy with a smoother profile and a fitted neighbor. The three rightmost columns are all data produced as a result of running the pipeline and used in post-fitting analysis.

## 3.4 Masking and neighbor identification

In this chapter, the effects of masking and neighbors on fitting process. The choices for masking and simultaneously fitting neighbors are examined and justified. The chapter

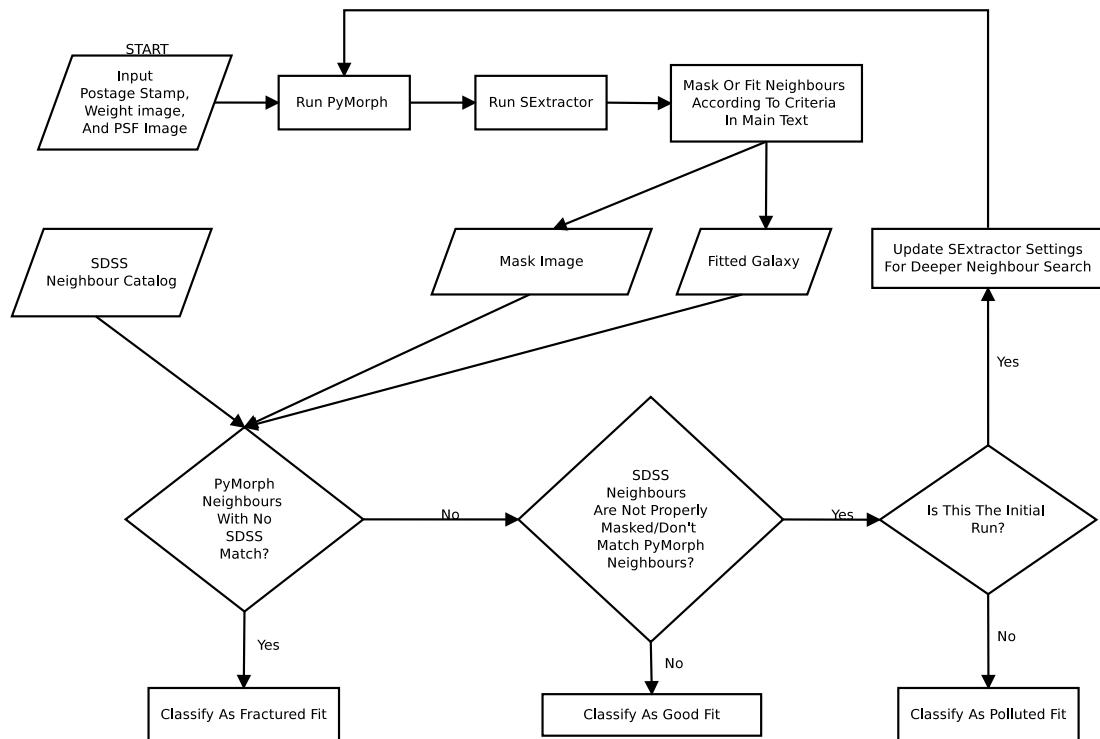
### 3. FITTING PROCESS

---



**Figure 3.3:** Two example fits output by the PyMORPH pipeline. Ser-Exp fits are shown. Each column contains the input image (far left), mask (center left), fitted galaxy (center right), and the residual image (far right). The first row shows a well-fit galaxy. The second row shows an early-type galaxy with a smoother profile and a fitted neighbor.

### 3.4 Masking and neighbor identification



**Figure 3.4:** A flowchart showing the process of masking and simultaneous fitting described in Chapter 3.4. The flowchart also includes the decision process used to identify ‘fractured’ or ‘polluted’ fits described in Chapter 3.4.3. This entire process is carried out for each galaxy individually.

### 3. FITTING PROCESS

---

also describes how galaxies with fitting problems caused by neighbors are identified in the catalog.

Figure 3.4 summarizes the process we used to identify neighbor sources, decide whether to mask or simultaneously fit the neighbor, and verify that the choice is appropriate. This process is described throughout Chapter 3.4.

#### 3.4.1 The masking and neighbor identification process

PYMORPH performs image masking using the SExtractor program (Bertin and 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.) Neighbors that are not simultaneously fit are masked according to the SExtractor segmentation image. Simultaneous fitting of the target source and neighbor 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 neighboring 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 unaffected by varying these parameters because the objects that require masking are large, bright neighbors with high S/N.

After searching the frame with SExtractor, the SExtractor catalog is used to decide whether to mask or fit neighboring sources according to requirements set on the minimum fractional size of a neighbor relative to the target and the maximum separation between the neighbor and target in multiples of the sum of the half-light radii. When both conditions are satisfied simultaneously by a specific neighbor-target combination, the neighbor 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.

---

### 3.4 Masking and neighbor identification

For the fits presented in this work, neighboring sources are simultaneously fit when the separation of the target and neighbor source is less than three times the sum of the two objects' semi-major half-light radii as measured by SExtractor. The neighbor source must be larger than 20 per cent of the area defined by the SExtractor radius of the target galaxy. Explicitly, these conditions are:

$$d_{tn} < 3(r_{hl,neighbor} + r_{hl,target}) \quad (3.6)$$

and

$$r_{hl,neighbor}^2 / r_{hl,target}^2 > 0.2 \quad (3.7)$$

Figure 3.5 shows some examples of galaxies fitted with `Ser` profiles. Each galaxy has masked or simultaneously fit neighbors. The decision of whether to mask or fit the neighbor was carried out as described in this chapter. 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 neighbor jointly determine whether the neighbor is masked or simultaneously fit.

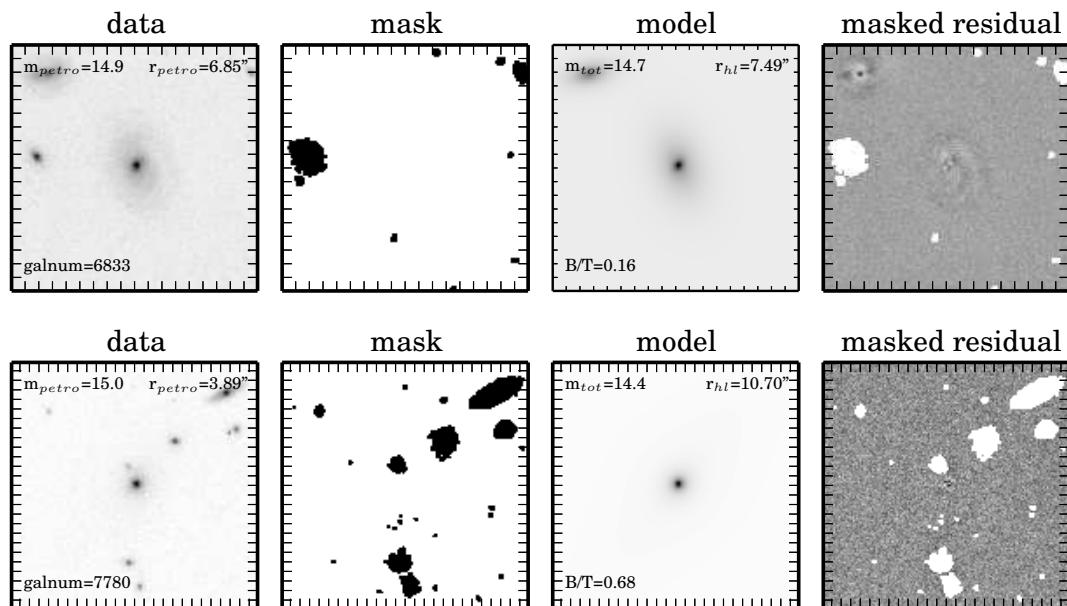
Several masking methods were tested prior to fitting, using circular and elliptical masks with sizes 2, 4, and 6 times the size of the neighbor sources (see Chapter 4.2.6). These masking configurations provided no reduction in the scatter of the recovered parameters when tested on simulations. Since 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).

#### 3.4.2 Failures of masking related to Nearby Neighbors and Source Fragmentation

A subset of catalog galaxies have nearby neighbors. In practice, neighbors far away from the target galaxy (i.e., larger than a few times the Petrosian half-light radius of the galaxy are

### 3. FITTING PROCESS

---



**Figure 3.5:** Examples of two galaxies with neighbors 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 the black areas.

### 3.4 Masking and neighbor identification

---

easily identified by PYMORPH and properly masked or simultaneously fit. However, there are cases where the galaxies are too close together and fitting will not properly proceed without human intervention.

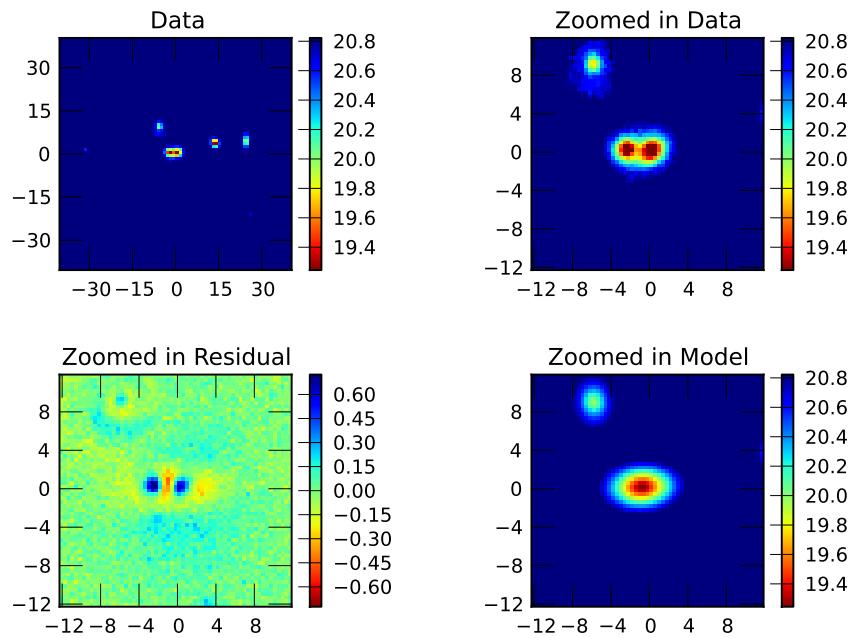
Figure 3.6 shows an example of two galaxies that sit very close together and are not properly identified as two separate sources. These galaxies are listed as two separate sources in the CasJobs database, but not separated by SExtractor during fitting. Problems such as this are not immediately noticeable using the catalog because the only galaxies selected for fitting are those galaxies in the spectroscopic catalog. Most neighbors fall below the spectroscopic sample cut (extinction-corrected  $m_r < 17.77$ ) and will not be included in the catalog.

Additionally, neighbors bright enough to pass the spectroscopic cuts may still be missing from the spectroscopic catalog due to fiber collisions at separations of 55 arcsec or less (Strauss et al. 2002). These collisions will occur only for galaxies that meet the criteria used for spectroscopic target selection. Galaxies that have an extinction-corrected Petrosian magnitude brighter than 17.77 will then be excluded from the sample if they caused a fiber collision during target selection. However, given the source density of these sources (of order 1 per fPc image), it is reasonable to expect very few overlapping bright sources. Fields with multiple visits can also avoid fiber collisions by measuring the colliding sources on separate occasions. This represents only a small number of catalog galaxies (less than 100). Instead, nearby sources dimmer than 17.77 magnitudes must also be considered.

The missed neighbors can often be detected by SExtractor using deeper deblending settings. However, increasing the sensitivity of the SExtractor deblending can lead to source fragmentation. If PYMORPH is too sensitive to unusual profile behavior, then a spurious source may be detected and fit. This fragments the source into multiple pieces, causing PYMORPH to under-report the brightness and size of the source as well as producing meaningless fitting parameters. Increasing the sensitivity of SExtractor only when the SDSS CASJobs data suggests that increased sensitivity is required balances these two concerns.

### 3. FITTING PROCESS

---



**Figure 3.6:** An example of two galaxies so close together that source separation is not done correctly and a single fit is applied.

### **3.4 Masking and neighbor identification**

---

This is determined by comparing the catalogs of SDSS and PYMORPH as described in the next section.

#### **3.4.3 Verifying masking and deblending conditions**

After an initial first run of the PYMORPH pipeline, verify that proper neighbor identification and masking or simultaneous fitting is verified by comparing the PYMORPH catalog containing each targeted galaxy and any fitted neighbors against the five nearest Primary Photometric targets from the DR7 CasJobs PhotoPrimary table. CasJobs objects are required to 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 the catalog.

There are  $\approx 78\,000$  potential neighbors in SDSS and  $\approx 100\,000$  simultaneously fitted PYMORPH neighbors (about 79 000 galaxies have one neighbor, about 8 000 galaxies have two neighbors, and about 1 500 galaxies have three or more simultaneously fitted neighbors, adding up to produce the  $\approx 100\,000$  neighbors).

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

Specifically, the matching conditions are:

### 3. FITTING PROCESS

---

$$d_X = (X_{target} - X_{neighbor}) - (colc_{r, target} - colc_{r, neighbor}) \quad (3.8)$$

$$d_Y = (Y_{target} - Y_{neighbor}) - (rowc_{r, target} - rowc_{r, neighbor}) \quad (3.9)$$

$$\sqrt{d_X^2 + d_Y^2} \leq r_{petro,hl,r} \quad (3.10)$$

where the separation is measured in pixels.  $X$  and  $Y$  are the coordinates on the cutouts used for fitting.  $colc$  and  $rowc$  are the coordinates on the fPc image. There are a small number of neighbors (about 2000 out of the 70,000 neighbor candidates) that come from different runs. For these objects, matching conditions are relaxed and any neighbor candidates that have the same angular separation from the target galaxy in both the SDSS and PYMORPH catalogs (these neighbors both lie in the same 1.5 arcsec annulus about the target galaxy) are considered matched. While this provides no sensitivity to direction, it is not a cause for concern, since galaxies have few fitted neighbors (most galaxies have 0 neighbors, and only about 10,000 have 2 or more neighbors as noted earlier in this section). Therefore, it is safe expect that the confusion rate in the neighbor-matching algorithm to be quite low.

Each fitted neighbor has a unique (GALNUM , neighbornum) pair assigned to it. The neighbornum of the matching CAST\_neighbors entry is reported as the neighbornum entry in the CAST\_neighbors table. The status of the fitted neighbor (1 for matched, 0 for not matched) is also recorded in the “ismatched” parameter of the r\_neighbors\_ser table. These tables are released with the catalog.

About 40 000 of the 100 000 fitted PYMORPH neighbors match SDSS objects. The remaining 60 000 unmatched PYMORPH neighbors and  $\approx$  38 000 unmatched SDSS neighbors may cause problems during fitting. Following the procedure described in this section, the number of galaxies with problems due to neighbors is only  $\approx$  4 000 galaxies in the final sample.

The two groups, unmatched SDSS neighbors and unmatched PYMORPH neighbors, possibly represent two different failures of the pipeline. Unmatched PYMORPH neighbors may

### 3.4 Masking and neighbor identification

---

be spurious neighbors fitted by PYMORPH after SExtractor improperly separates the target galaxy into several smaller fragments. Unmatched SDSS objects may be nearby neighbors that fail to be deblended from the target galaxy by SExtractor.

Unmatched SDSS neighbors tend to happen when neighbors are superimposed on the target galaxy or very nearby. In principle, these situations can happen independent of the true physical separation of the neighbor 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 neighbors may also be enhanced in dense environments (i.e., within clusters). This is a potentially important effect as it can bias measurements of galaxies in clusters.

However, SDSS has many suspicious entries at dimmer magnitudes. These sources are often the result of deblending errors. Figure 3.7 shows the fit to galaxy with a fictitious neighbor. According to CasJobs, there are two sources at this location with the same or similar magnitudes. These cases obviously do not constitute true neighbors and have no effect on the fitting.

As an additional constraint, we download the SDSS object flags<sup>4</sup> In particular, the DEBLEND\_NOPEAK flag in the band of interest (either *g,r,or i*) needs to be checked. This flag indicates that no peak was found in the deblended source, and the source is likely not a real source. Although not implemented here, one could have required that the objects have the BINNED1 flag set. The SDSS site discusses that sources should have either BINNED1, BINNED2, or BINNED4 to be detected. However, many of the BINNED2 sources are spurious detections of the outskirts of bright galaxies and scattered light from bright stars. In addition, SDSS says that very few objects flagged as BINNED4 are real astrophysical objects. We could also check for the NOPROFILE flag which indicates that no profile could be measured for the deblended source. Such sources are likely spurious, especially in the neighborhood of a

---

<sup>4</sup>For an explanation of the flags, see [http://www.sdss3.org/dr8/algorithms/flags\\_detail.php](http://www.sdss3.org/dr8/algorithms/flags_detail.php). See also Robert Lupton's page at [http://www.astro.princeton.edu/~rhl flags.html](http://www.astro.princeton.edu/~rhl	flags.html).

### 3. FITTING PROCESS

---

galaxy. These two additional flags were checked initially and found to provide no additional information for the sample of neighbors used in this work, so this flag was not used in the final matching.

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

SDSS sources that do not have the DEBLEND\_NOPEAK flag set and are not masked are considered to be polluting the target galaxy profile if they are unmatched to a fitted PYMORPH neighbor and the distance calculated as

$$\begin{aligned} x_d &= \cos(\phi) * (xn - xt) + (yn - yt) * \sin(\phi) \\ y_d &= -\sin(\phi) * (xn - xt) + (yn - yt) * \cos(\phi) \\ r_d &= \sqrt{ba * (x_d^2) + (y_d^2)/ba} * 0.396 \end{aligned} \tag{3.11}$$

where  $\phi$  and  $ba$  are the position angle and ratio of the semi-minor to semi-major axis of the fitted target galaxy.  $r_d$  is converted to arcsec by the pixel scale, 0.396 arcsec/pixel.  $r_d$  is compared to the maximum of the Petrosian half-light radius and the fitted half-light radius (this protects us in cases where the target shrinks to a point source, as can happen in cases poor fits).

Any SDSS neighbors that pass these cuts potentially corrupt the fits. An empirically verified safe cut for considering these unmasked, unfit neighbors to be problems is if they

---

### 3.4 Masking and neighbor identification

are less than 3 half-light radii away and

$$m_n - m_t < 4.0 - r_d/r_{hl} \quad (3.12)$$

where  $r_d$  is the radial distance from the target galaxy to the neighbor,  $r_{hl}$  is the target galaxy half-light radius,  $m_n$  is the neighbor magnitude and  $m_t$  is the target magnitude. For example, a neighbor galaxy 4 arcsec away from an average galaxy in the 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 neighbors to be considered problematic out to larger radii relative to dimmer neighbor galaxies. After this cut, only about 7500 target galaxies potentially have contamination from unmatched SDSS neighbors.

Another potential problem case are the so-called “fractured” galaxies. In the largest and nearest galaxies, detailed galaxy structure (e.g. spiral arms or dust lanes) triggers improper separation and deblending. These cases are called ‘fractured’ galaxies and identified by searching for the fitted galaxies where the Petrosian magnitude is at least 0.5 mag brighter than the fitted magnitude and the unmatched fitted neighbor is less than 3.0 Petrosian half-light radii from the target galaxy. Specifically,

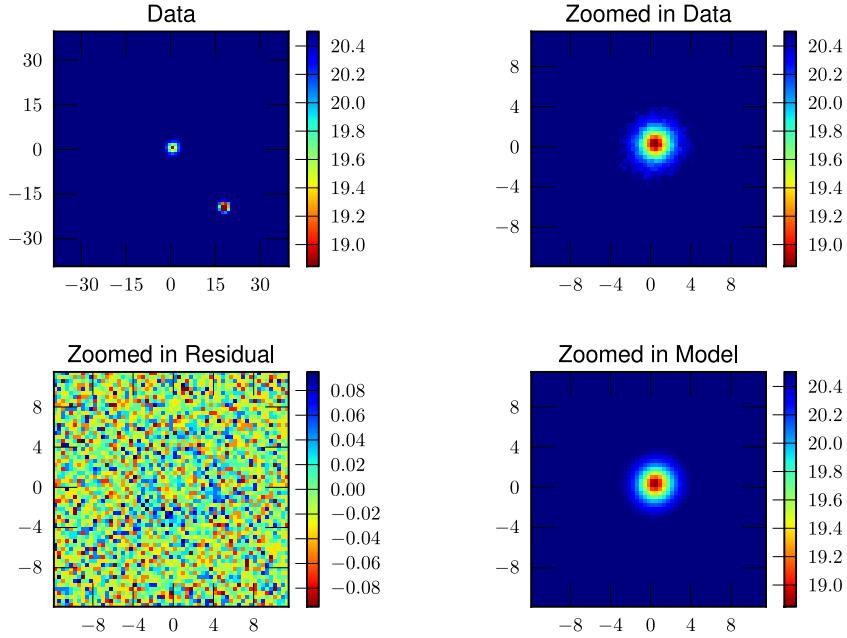
$$\begin{aligned} r_d &< 3.0 * r_{petro} \\ m_n - m_t &< 3.0 - dis/r_{petro} \\ r_d &< 9.0 \\ m_t - m_{petro} &> 0.5 \end{aligned} \quad (3.13)$$

These cases are quite rare, representing only about 150 ‘fractured’ galaxies in the entire catalog.

After all of these constraints are applied to the catalog, about 12,000 “polluted” sources and about 150 “fractured” galaxies are identified (which is about 2% of the sample). About 2,000 of these were initially failed fits. The remaining  $\sim 10000$  are refit using a deeper

### 3. FITTING PROCESS

---



**Figure 3.7:** An example of a galaxy that has a neighbor with the same magnitude centered less than 2 pixels. This is a case where there is no need for a neighbor fit

deblending setting on SExtractor and stricter masking conditions that make it more likely for a neighbor to be detected and simultaneously fit rather than masked or never detected.

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

### 3.4 Masking and neighbor identification

---

not only increases the likelihood of detecting the SDSS neighbors, but also of incorrectly ‘fracturing’ larger sources. We also reduce the neighbor source area requirement in PY-MORPH for this fitting run. This makes PYMORPH more likely to perform a neighbor fit, requiring the neighbor to be only 10 per cent of the area of the target rather than the original setting of 20 per cent area.

After refitting the test for unfit or unmasked neighbors described in this section is repeated. Galaxies still having neighbor problems at this point are flagged in the final catalog as potentially being polluted by neighbor 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 neighbors.

Figure 3.8 shows the change in fitted `Ser` magnitude (the fitted magnitude from the second pass minus the original fitted magnitude) for galaxies where all neighbors were detected on a second pass of fitting using deeper deblending. Individual galaxies are shown as gray scatter points. The median and 68 per cent contour are over-plotted in red. The  $x$ -axis shows the difference in target and neighbor magnitudes.

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

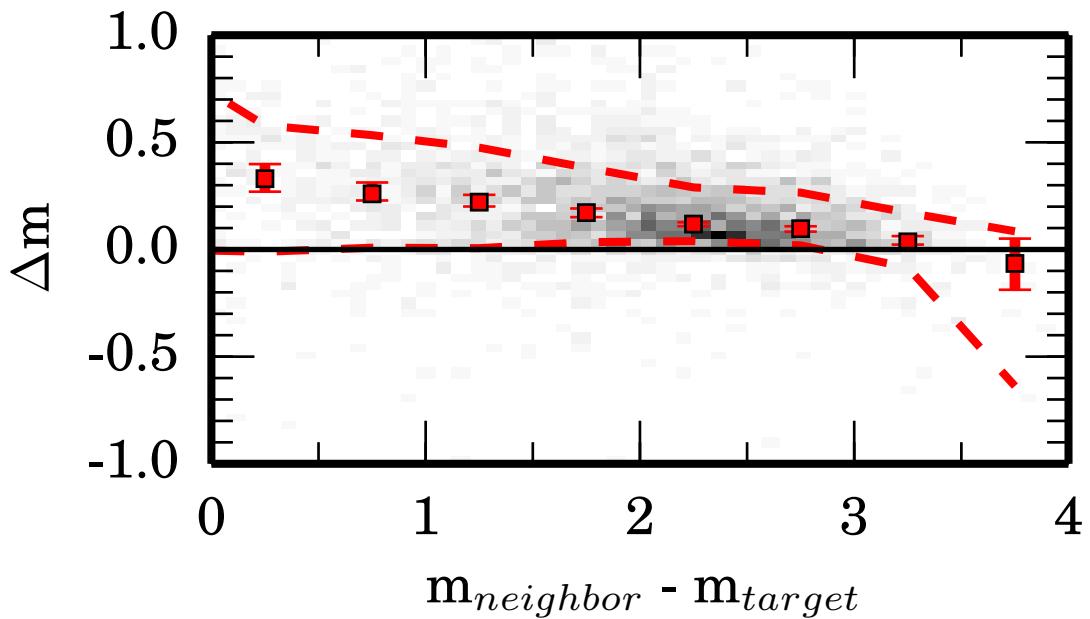
According to the 68 per cent contours on the plot, approximately 16 per cent of the refit galaxies become brighter during refitting. Since the original fits may include unmasked SDSS neighbors, the fits are usually poor. In some cases, the fit is not just contaminated by the neighbor, but the fit fails 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 neighbor target.

### **3. FITTING PROCESS**

---

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 behaviors. Based on these observations, we believe our cutting criteria in Equation 3.12 to be generous enough to capture the majority of cases where a true SDSS neighbor is causing fitting contamination.

About 4 000 ‘polluted’ target galaxies still fail during the second pass with deeper de-blending. The ‘polluted’ and ‘fractured’ galaxies jointly comprise about 0.6 per cent of the sample. Galaxies that are suspected to have substantial fitting problems are marked with a quality flag discussed in Chapter 5 and shown in Table 5.8. The number of simultaneous fits for each galaxy in our catalog is also reported since such fits substantially increase the number of free parameters during fitting and make it much more likely that the resulting fit has problems.



**Figure 3.8:** 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 neighbors are detected. Individual galaxies are shown as gray scatter points. The median and 68 per cent contour are over-plotted in red. The  $x$ -axis shows the difference in target and neighbor 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 neighbor 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.

# Chapter 4

## Simulations

In order to quantify the systematics and robustness of the fits of  $g$ -,  $r$ -, and  $i$ -band SDSS spectroscopic galaxies presented throughout this work, simulations of single and two-component galaxies are used to test the fitting process. These simulations are referred to as “mocks,” and are generated using an independent code. The simulations are fit using the same PY-MORPH pipeline discussed in Chapter 3.3 and used for the photometric decompositions of the DR7 galaxies presented in this work.

Following several detailed studies which have used simulations to test the robustness of different fitting algorithms (e.g. Häussler et al. 2007; Lackner and Gunn 2012; Simard et al. 2002), the main goal is to test the robustness of PYMORPH pipeline software on SDSS galaxies. These simulations test the effects of increased signal-to-noise as well as increased resolution, PSF errors, and sky determination. The simulations are specifically applicable to the SDSS galaxies in this catalog.

Unbiased samples selected to accurately represent the distribution of SDSS galaxies are used to estimate and correct the systematic error on recovered parameters as well as estimate reasonable uncertainties on fit parameters.

A description of the simulation process is presented in Chapter 4.1. This includes con-

---

## 4.1 Creating the catalog simulations

structing a catalog of realistic galaxy parameters (Chapter 4.1.1); generating galaxy surface brightness profiles based on these parameters (Chapter 4.1.2); generating sky and noise (Chapters 4.1.3 and 4.1.4); and including seeing effects in the final image. The completed simulations are run through the fitting pipeline, and the fits are analyzed in Chapter 4.2. Examination of the recovery of structural parameters in noise-free images is presented in Chapter 4.2.1. Parameter recovery in realistic observing conditions including both neighboring sources and the effects of incorrect PSF estimation is examined in Chapter 4.2.2.

`PYMORPH` is further tested by inserting mocks into real SDSS images to test the dependence on local density in Chapter 4.2.3. Dependence of the fits on resolution and signal-to-noise are tested in Chapter 4.2.4. This is particularly interesting since the pipeline may be applied to DES and future LSST surveys.

The effect of changing the fitted cutout size is presented in Chapter 4.2.5, and the effect of incorrect background estimation is also examined in Chapter 4.2.7.

Single-component Sérsic galaxy models (`Ser`), de Vaucouleurs bulge with an exponential disk models (`deV-Exp`) and Sérsic bulge with exponential disk models (`Ser-Exp`) are all used to test the fitting. The fitting bias of the single-component `deV` model is not examined. Although `deV` fits are performed on the SDSS catalog and released as part of this work, the `deV` model is not preferred for any work presented here. As a result, there is no reason to examine the bias present.

## 4.1 Creating the catalog simulations

### 4.1.1 Selecting the simulation catalog

The set of mocks used to test the fitting code are drawn from an initial run of the photometric decompositions of the catalog. These galaxy parameters represent the *r*-band image decompositions of a complete sample of the SDSS spectroscopic catalog containing all galaxies with spectroscopic information in SDSS DR7 (Abazajian et al. 2009).

## 4. SIMULATIONS

---

For each model (**Ser**, **deV-Exp**, and **Ser-Exp**), a representative sub-sample of physically meaningful photometric decompositions is randomly drawn from the catalog. In order to ensure that the galaxies are representative of the full catalog, the distributions of basic observational parameters of SDSS galaxies (surface brightness, redshift, apparent Petrosian magnitude, Petrosian half-light radius, and absolute magnitude) were examined.

Some restrictions on fit parameters are necessary to ensure that outliers are removed from the parameter space used to generate the simulations. Galaxies that do not satisfy these basic cuts are removed to ensure that the parameters used to generate the images are physically motivated. The cuts do not significantly bias the galaxy distributions shown in Figure 4.1. The cuts are:

1. Any Sérsic components must have Sérsic index less than 8.
2. Half-light radius of any Sérsic component must be less than 40 kpc.
3. In the two-component fits, the ratio of the bulge half-light radius to disk scale radius should be less than 1, or the galaxy should be bulge dominated ( $B/T > 0.5$ ).

Conditions (i) and (ii) are used to prevent selection of **Ser** models with extended profiles that are likely the result of incorrect sky estimation during the fitting process. Condition (iii) ensures that any disk dominated galaxies have a bulge component that is smaller than the disk.

After enforcing the cuts on the sample, 10 000 fitted galaxy profiles for each of the **Ser**, **deV-Exp**, and **Ser-Exp** models are selected at random without regard to the morphological classification of the original galaxy. The fitted parameters of these sample galaxies are used to generate the mocks used in testing the pipeline.

Selecting galaxy samples independent of galaxy morphology allows the **deV-Exp** and **Ser-Exp** samples to contain some galaxies that do not truly possess a second component. Using the flagging presented in Chapter 5 more of these galaxies would be removed. The flagging was not available at the time of the simulation tests.

## 4.1 Creating the catalog simulations

---

Additionally, there will be some truly two-component galaxies (i.e., both bulge and disk components are present) that are misrepresented by a single Sérsic fit. However, this sampling method will not invalidate the results of the tests. Testing the ability to recover simulated galaxy parameters only requires a realistic sample of galaxy profiles. These samples satisfy this requirement. Single Sérsic galaxies in the original sample, simulated as mock **Ser** galaxies and fit with **Ser** models, test the ability to recover Sérsic parameters. Similarly, **Ser** mocks with **Ser-Exp** models, show bias resulting from over-fitting a galaxy. Fitting the **Ser-Exp** mocks with a **Ser** model shows the bias due to under-fitting.

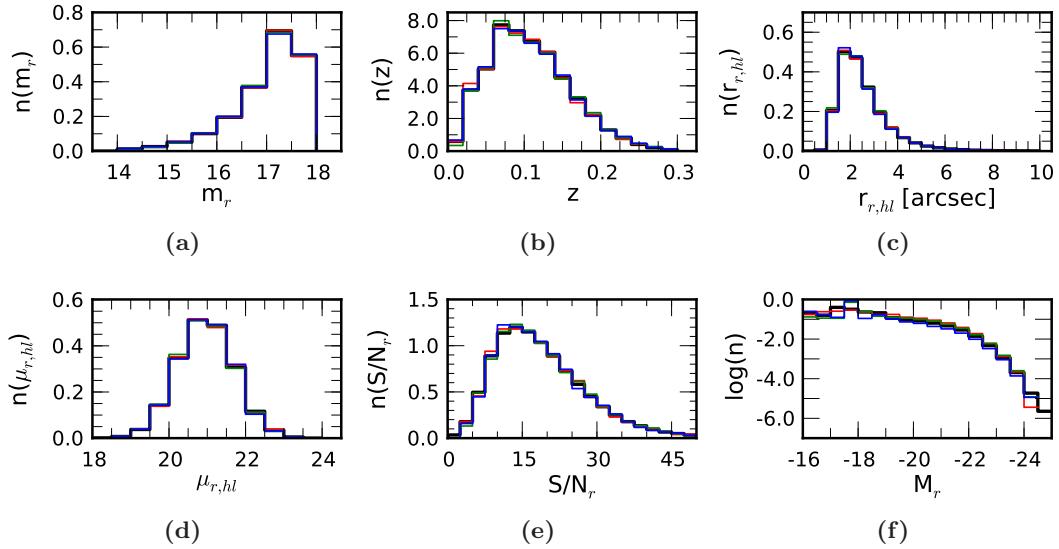
Fitting a single-component model regardless of galaxy structure or morphology is a common practice (e.g. Blanton et al. 2005b; Häussler et al. 2007; Simard et al. 2011). Figure 4.8b shows that bias of 0.05 mags and 5 per cent of the half-light radius result from fitting a two-component galaxy with a single component and that this bias increases to 0.1 mags and 10 per cent of the half-light radius for brighter galaxies. These biases are important in analyzing the results of a single-component fitting catalog. For example, Bernardi et al. (2013) shows that intermediate B/T galaxies can often be fit by Sérsic models with large Sérsic indexes, which can lead to misclassification if cuts similar to Shen et al. (2003) are used.

Figure 4.1 shows the distributions of surface brightness, redshift, extinction-corrected  $r$ -band Petrosian magnitude,  $r$ -band Petrosian half-light radius, and absolute magnitude of all SDSS spectroscopic galaxies (in black) and simulation samples: **Ser** (red), **deV-Exp** (green), and **Ser-Exp** (blue). The distribution of mock galaxies reproduces the observed distribution for all three samples for each observational parameter as verified by a KS 2-sample test.

Figure 4.1 also presents the signal-to-noise (S/N) of the mock samples as compared to the parent distribution. The S/N of the images is a limiting factor in the fitting process, so care must be taken to ensure that the S/N is not artificially increased in the simulations when compared to true SDSS galaxies. This S/N is calculated using the  $r$ -band Petrosian magnitude and  $r$ -band Petrosian half-light radius. Petrosian quantities are used to make a fairer

## 4. SIMULATIONS

---



**Figure 4.1:** Distributions of (a) extinction-corrected  $r$ -band Petrosian magnitude, (b) redshift, (c)  $r$ -band Petrosian half-light radius, (d) surface brightness, (e) signal-to-noise, and (f)  $V_{\max}$ -weighted luminosity function of the samples in the simulations drawn from the DR7 spectroscopic galaxy sample. The distribution of all DR7 spectroscopic galaxies is shown in black. Distributions of the **Ser**, **deV-Exp**, and **Ser-Exp** mocks are shown in red, green, and blue, respectively. Bin counts are normalized to integrate to 1. The distributions of the mocks are representative of the full catalog fitted in this work and are appropriate to compare to the DR7 spectroscopic sample as verified by a Kolmogorov-Smirnov 2-sample test. The signal-to-noise ( $S/N$ ) will be discussed further in Chapter 4.1.4. In calculating this  $S/N$ , the measurement of sky provided by the PYMORPH pipeline rather is used rather than DR7 to identify and separate target counts from sky counts. PYMORPH sky estimation is shown to be more accurate than the SDSS estimation provided in the DR7 catalog.

## 4.1 Creating the catalog simulations

---

comparison among all the samples. Because the Petrosian quantities are non-parametric, they avoid the complications that arise in assessing the possible biases introduced during fitting. Any differences in S/N are not large enough to significantly bias the distributions as verified by the KS 2-sample test. Therefore, the samples are fair representations of the underlying distribution of SDSS spectroscopic galaxies. The S/N is discussed further in Chapter 4.1.4.

Testing the accuracy of the PYMORPH fitting routine does not necessarily require an unbiased parameter distribution. In reality, all that is required is a sample with sufficient coverage of the parameter space represented by the data. The simulations use smooth profiles, simplifications of the true galaxies that are observed in SDSS. Examination of the results of fitting these simplified models and comparison to fits of true observed galaxies can potentially yield useful information regarding galaxy structure. In Bernardi et al. (2013), the simulations are used together with the decompositions of the SDSS spectroscopic sample to characterize the scatter in the size-luminosity relation as well as examine possible biases. In order to make these comparisons, an unbiased sample is required. The distributions shown in Figure 4.1 show that the simulations are appropriate to use for this purpose.

### 4.1.2 Generating the images

Two-dimensional normalized photon distributions are generated from the one-dimensional Sérsic profiles and the one-dimensional exponential profiles of each bulge and disk component. Disk components are only simulated where required, as is the case for two-component fits. When multiple components are to be simulated, each component's normalized photon distribution is generated separately and combined prior to generating the simulated exposure.

Two-dimensional galaxy profiles are treated as azimuthally symmetric one-dimensional galaxy light profiles that are deformed according to an observed ellipticity. The galaxy

## 4. SIMULATIONS

---

profiles are generated using the structural parameters generated from photometric decompositions as described in the previous section. Single-component galaxy profiles and the bulges of two-component galaxies are generated according to the Sérsic profile

$$I(r) = I_e \exp \left( -b_n \left[ \left( \frac{r}{R_e} \right)^{\frac{1}{n}} - 1 \right] \right) \quad (4.1)$$

$$b_n = 1.9992n - 0.3271$$

where Sérsic index ( $n$ ), half-light radius ( $R_e$ ), and surface brightness at  $R_e$  ( $I_e$ ) are selected simultaneously from the catalog described in the previous section.

For the **deV-Exp** and **Ser-Exp** cases, an exponential disk (Equation 4.1 with  $n = 1$ ) is added to the Sérsic component to model the disk component of the galaxies. The disk is modeled using a slightly modified version of Equation 4.1. This model requires input parameters scale radius ( $R_d$ ) and central surface brightness ( $I_d$ ).

$$I_{Exp}(r) = I_d \exp \left( \frac{-r}{R_d} \right). \quad (4.2)$$

After generating the two-dimensional profile, the image is pixelated by integrating over each pixel area. The details of this integration are largely unimportant. However, the simulation must take careful account of the integration in the central pixels, where the profile can vary greatly over a single pixel. Various oversampling methods have been devised to properly correct this common problem (e.g. Häussler et al. 2007; Peng et al. 2002). The simulations in this paper have been tested to ensure that the pixel-by-pixel integration is accurate to  $\approx 3$  per cent of the corresponding Poisson noise in a given pixel. Therefore, the simulations are treated as exact calculations of the galaxy photon distributions since any noise from the integration contributes only a small amount to the total noise budget.

The pixelated galaxy is numerically convolved with a PSF extracted from SDSS DR7 data using `read_PSF` program distributed by SDSS<sup>5</sup>. The choice of this PSF is discussed in Chapter 4.2.2.

---

<sup>5</sup>`read_PSF` is part of the `readAtlasImages-v5_4_11` package available at [http://www.sdss.org/dr7/products/images/read\\_psf.html](http://www.sdss.org/dr7/products/images/read_psf.html)

### 4.1.3 Creating the background

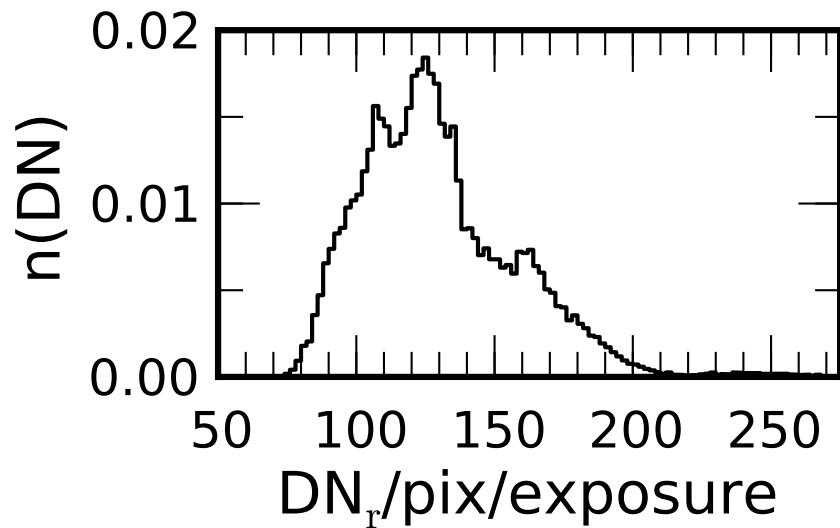
Two hundred background images, each equal in size to an SDSS fpC image, are also simulated for testing purposes. These images contain constant background and a randomly selected field of galaxies taken from an SDSS fpC image. The SDSS catalog provides rudimentary photometric decompositions of each star and galaxy. Galaxies are fit with an exponential disk and a de Vaucouleurs ( $n = 4$ ) bulge independently. The best fit is reported as a linear combination of the two fits using the `fracdev` parameter to express the ratio of the de Vaucouleurs model to the total light in the galaxy.

For the simulated background used in this paper, each galaxy is generated using the combined profile of the two fits. The de Vaucouleurs bulge and exponential disk component are separately simulated according to the magnitude, radius, ellipticity, and position angle reported in SDSS. Each component is simulated using the method described in Chapter 4.1.2. The background galaxy is constructed by adding the two components using the `fracdev` parameter. The galaxy is then inserted into the fpC image. Any foreground stars are also simulated as point sources and inserted into the image.

The mean sky of all SDSS observations as given in the SDSS `photoobj` table by the `sky_r` parameter is used for the background sky counts in the image. The distribution of the sky flux is plotted in Figure 4.2 in units of counts (or DN) per pixel per exposure. The median and mean values for a 54 second SDSS exposure are  $\approx 125$  and  $\approx 130$  counts per pixel, respectively. The mean value of 130 counts per pixel is used as the background in the simulations. This sky background is applied to the entire chip as a constant background; no gradient is simulated across the image. Background gradients should be approximately constant across a single galaxy. This assumption is verified by inserting the simulated galaxies into real SDSS fpC images near known clusters, where the sky contribution should be higher and gradients are more likely. In Chapter 4.2.3, the simulations show that there is little change in the behavior of the fits in these types of environments.

#### 4. SIMULATIONS

---



**Figure 4.2:** The distribution of sky values for data in the SDSS CASJOBS catalog. These data are drawn from the `sky_r` CASJOBS parameter and are converted into counts (DN) per pixel per standard SDSS image exposure of 54 seconds. This distribution is used to determine the sky value used in the simulations. As an approximation, the mean value of 130 counts/pixel/exposure is selected.

## 4.1 Creating the catalog simulations

---

Previous work has improved the measurements of sky background (see Blanton et al. 2011). However, these corrections tend to focus on areas of large, bright galaxies or on making the sky subtraction stable for purposes of tiling fpC images together. Since the focus is on maintaining the proper S/N for the simulations, the sky levels provided in the SDSS database are sufficient, provided that they maintain the correct S/N. The S/N distribution of the simulations and the original SDSS galaxy sample are discussed in Chapter 4.1.4.

Diffraction spikes and other image artifacts are not directly simulated. However, the SDSS photo pipeline often misidentifies additional phantom sources along an observed diffraction spike. These phantom sources are modeled in the background, and so these effects are approximately modeled. It is reasonable to expect that the diffraction effects should not have a large effect on the fitting process, as their elongated straight structure does not mimic galaxy structure. The dominant effect produced by the bright stars in the field is bias in the background estimation in the nearby neighborhood of a star.

After simulation of the background images, and prior to adding noise, each background image is convolved with a random SDSS PSF selected from original fpC image upon which the individual image is based. Selecting PSFs from original SDSS images introduces a variation in PSF size between mock galaxies inserted into images and the background galaxies. However, this variation is not of concern for us in the fitting process because the vast majority of galaxies (over 90 per cent of all galaxies) do not have neighbors near enough to the target galaxy to require simultaneous fitting. For these galaxies, the PSF applied to neighboring galaxies is of no interest in the fitting process because the sources are masked out. The details of this masking are not discussed in the remainder of the paper. Modifying the masking conditions produce no noticeable difference in the fitted values for the simulations. For the remaining 10 per cent of galaxies, there may be some variation in the fit due to differing PSFs. PSF sizes can differ between target and neighboring galaxies by up to a factor of 2. However in practice, this happens for less than 1 per cent of galaxies of the galaxies with neighbors. Furthermore, incorrect PSF tends to only cause effects at

## 4. SIMULATIONS

---

the centers of galaxies. So although using a PSF that is different from the background PSF will affect the recovered parameters of the neighbor, it will not affect the target galaxy.

### 4.1.4 Noise

After generating a target galaxy and inserting it into a background, Poisson noise is added using the average inverse gain of an SDSS image ( $4.7 \text{ e}^-/\text{DN}$ ) and the average contribution of the dark current and read noise, referred to as the “dark variance,” ( $1.17 \text{ DN}^2$ ), to determine the standard deviation for each pixel. Specifically,

$$F_{i,j} \equiv I_{i,j} + \text{bkrd}_{i,j} \quad (4.3)$$

is the total flux in pixel  $(i,j)$  (i.e., the sum of the source and background fluxes in the pixel), and

$$\sigma_{i,j} = \sqrt{\frac{F_{i,j}}{\text{gain}} + \text{dark variance}} \quad (4.4)$$

so

$$\left(\frac{\text{S}}{\text{N}}\right)_{i,j} \equiv \frac{I_{i,j}}{\sigma_{i,j}}, \quad (4.5)$$

for a single pixel.

Since the fitting pipeline is dependent on the S/N, it is essential that the simulated S/N is comparable to SDSS. The distribution of the average S/N per pixel within the half-light radius for the simulations and the original galaxies is plotted in Figure 4.1e. The S/N distribution of simulations and the SDSS spectroscopic galaxies agree as verified by a KS 2-sample test, therefore the simulations appropriately approximate the S/N of SDSS galaxies contained in the catalog.

An unbiased selection in the previously mentioned parameters is not sufficient to guarantee fair sampling of the S/N with respect to magnitude, nor does it prevent fictitious correlations among multiple fit parameters. In fact, correlations among fit parameters are to be expected if the PYMORPH pipeline is robustly measuring properties of the target

## 4.1 Creating the catalog simulations

---

galaxies (many correlations exist among physical parameters). It is difficult, and largely unnecessary, to examine every possible relationship for correlations introduced by biases in the sample selection process.

Examining the S/N and the half-light radius versus apparent magnitude help to ensure the appropriateness of the simulation. Systematic differences in radius will lead to systematic variation in the S/N of the sample. The scatter in recovered fitting parameters is also examined as a function of magnitude. Therefore, the S/N as a function of apparent magnitude should appropriately reflect that of the parent sample from SDSS.

Figure 4.3 presents the half-light radius versus apparent magnitude, and Figure 4.4 presents the S/N versus apparent magnitude. The points shown in red and blue correspond to the `Ser` and `Ser-Exp` mocks, respectively. The underlying SDSS parent distribution is shown in black. Figure 4.3 shows that the `Ser` and `Ser-Exp` models are in close agreement with the full SDSS sample. The `Ser` and `Ser-Exp` model radii agree across the magnitude range. The S/N agrees with the full SDSS sample or is slightly below that of SDSS. The lower signal-to-noise, although not exactly that of SDSS, will not bias the tests toward better results, so these samples are acceptable for testing. The `deV-Exp` sample, which is not shown here, tends to have smaller radii and higher S/N at brighter magnitude. The results of tests using the `deV-Exp` model are not discussed in the remainder of this paper. They can be found in M2014.

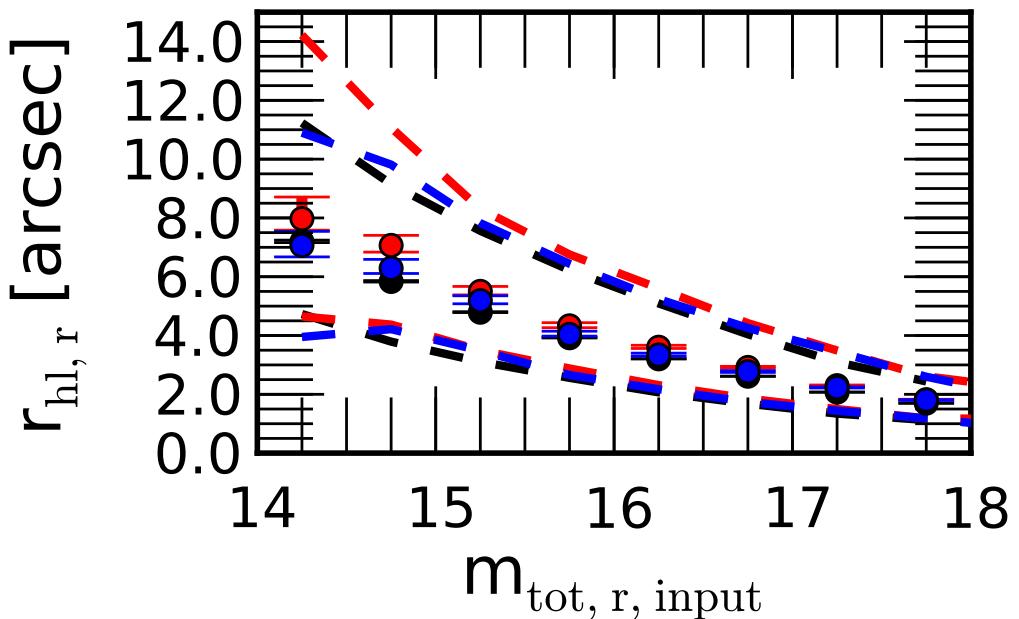
### 4.1.5 Final processing for fitting

For each mock galaxy, a weight image of the  $\sigma_{i,j}$  values is generated according to Equation 4.4. This image is supplied along with the input image to the pipeline in order to calculate the  $\chi^2$  value for the fit.

Figure 4.5 shows some examples of mock galaxies throughout the simulation process. This includes the noiseless mock galaxy, the noiseless simulated background, the composite image of galaxy and background, and the composite image after adding Poisson noise with

#### 4. SIMULATIONS

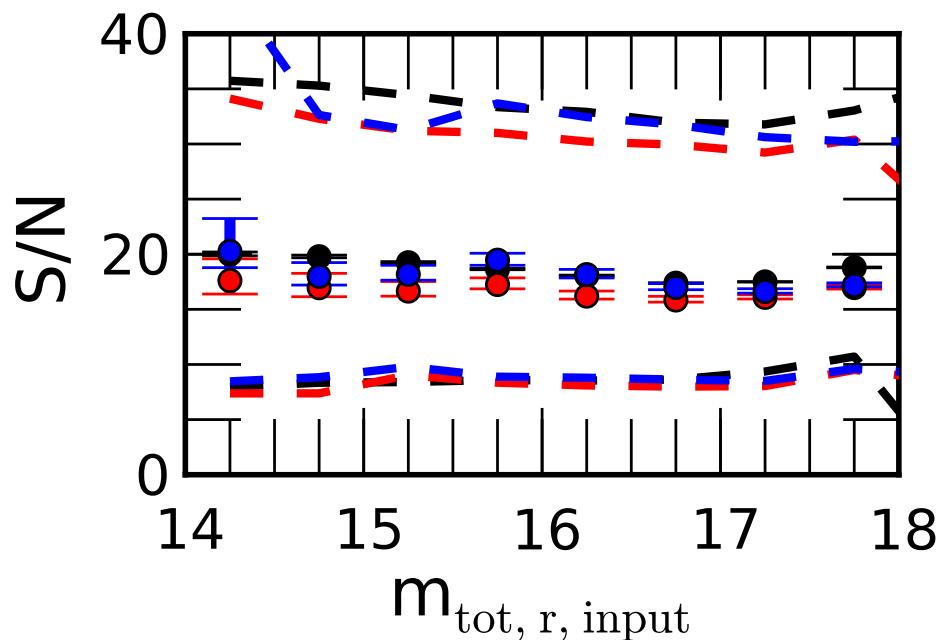
---



**Figure 4.3:** The distribution of galaxy radii as a function of apparent magnitude for the parent SDSS sample in black, the **Ser** model in red and the **Ser-Exp** model in blue. The median in each bin is shown with error bars representing the 95 per cent CI on the median. Corresponding dashed lines show the extent of the middle 68 per cent of data. The **Ser-Exp** model is in close agreement across the entire magnitude range while the **Ser** model begins to diverge at brighter magnitudes.

---

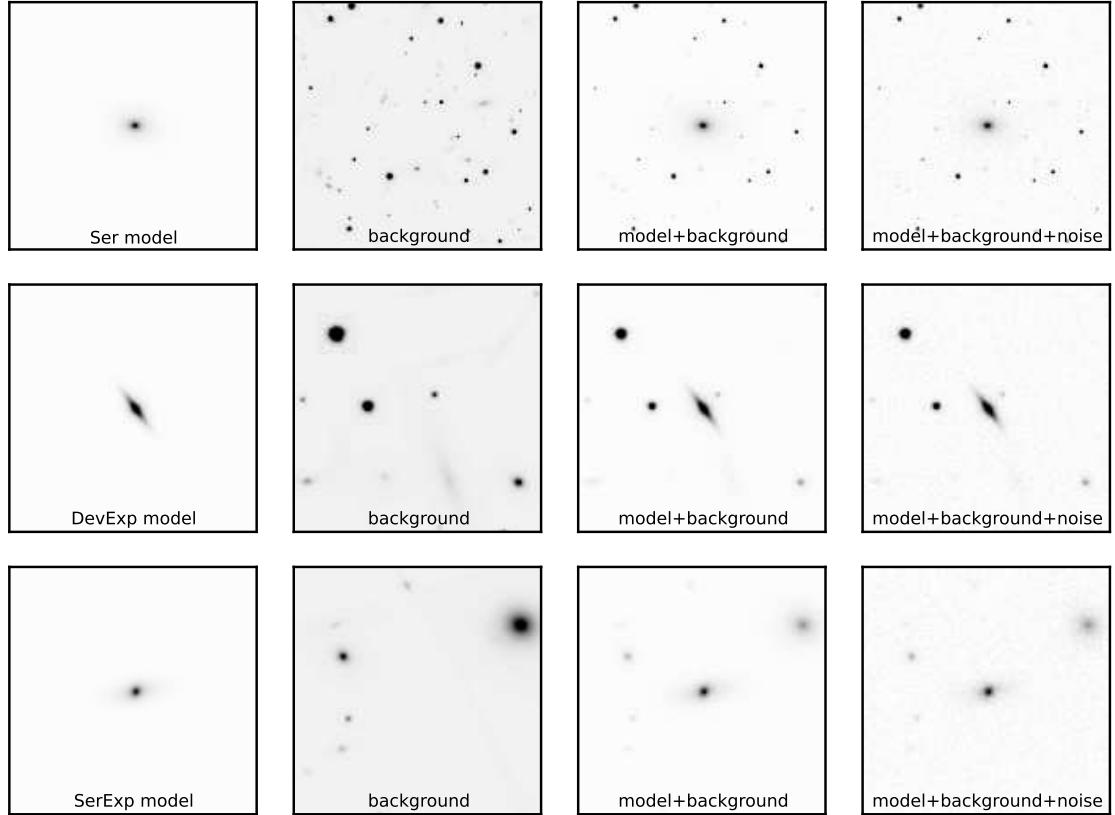
#### 4.1 Creating the catalog simulations



**Figure 4.4:** The distribution of galaxy S/N as a function of apparent magnitude is presented in the same format as Figure 4.3. The **Ser-Exp** and **Ser** models are in close agreement with the full sample across the entire magnitude range.

#### 4. SIMULATIONS

---



**Figure 4.5:** Examples of mock galaxies and background shown before and after adding Poisson noise. Top, middle, and bottom rows show randomly selected sample **Ser**, **deV-Exp**, and **Ser-Exp** profiles, respectively. From left to right, the columns show the mock galaxy, simulated background, background+galaxy, and final image with noise.

---

## 4.2 Testing PyMorph image decompositions

$\sigma_{i,j}$  defined in Equation 4.4. The final image size used for fitting is 20 times the Petrosian  $r$ -band half-light radius. A discussion of this choice of stamp size is presented in Chapter 4.2.5.

## 4.2 Testing PyMorph image decompositions

In order to test the parameter recovery of the PYMORPH pipeline on SDSS spectroscopic galaxies, the PYMORPH pipeline is applied mocks described in Chapter 4.1. The PYMORPH pipeline uses GALFIT to fit smooth profiles to the the mock galaxies. The pipeline is applied to several different realizations of the mock galaxies. These realizations increase in complexity from a noiseless image to an image with real noise and (possibly clustered) neighboring sources. The ability of PYMORPH to reliably recover model parameters is limited by both the S/N and the resolution of the mock galaxy. Understanding the systematic effects of S/N and resolution is useful in interpreting the data presented in Chapter 5. It may also be used to correct biases in the data as described later in Chapter 4.4.

### 4.2.1 Noiseless images

As an initial test, the pipeline is applied to simulations prior to adding noise, background counts, or neighboring sources. This produces the minimum scatter in the data, serves to verify that the simulations are correct, and shows that PYMORPH is properly functioning.

The total apparent magnitude, half-light radius, and additional fit parameters recovered by fitting the noiseless images of the **Ser** and **Ser-Exp** models are presented in Figures 4.6a, 4.7a, 4.8a, 4.9a, 4.10a, and 4.11a. The plots show the difference in simulated and fitted values (fitted value - input value). The difference is shown versus the input magnitude as well as the input value of the respective fit parameter. The gray-scale shows the density of points in each plane with red points showing the median value. Error bars on the median value are the 95 per cent confidence interval on the median obtained from bootstrapping. Blue dashed lines show the regions which contain 68 per cent of the objects.

## 4. SIMULATIONS

---

Figures 4.6a and 4.9a show the corresponding fit is well constrained (**Ser** fit with **Ser**, and **Ser-Exp** with **Ser-Exp**). The total magnitude and half-light radius are both constrained well within 1 per cent error on the flux or radius ( $\sigma_{\text{total mag}} \approx 0.01$  mag and  $\sigma_{\text{radius}} \approx 1$  per cent). However, the scatter increases somewhat for the sub-components of the **Ser-Exp** fit (see Figures 4.10a and 4.11a). As the components of the **Ser-Exp** model become dim (bulge/disk magnitude approaches 18.5), the component contribution to the total light becomes small. The origin of the magnitude limit is merely an artifact of the selection criteria requiring that all galaxies have total magnitude brighter than 17.77. This implies that components with magnitude of  $\approx 18.5$  or dimmer are necessarily sub-dominant components and contribute at most  $\approx 50$  per cent of the light to the total profile. On average, components dimmer than 18.5 magnitudes contribute about 25 per cent of the total light to a typical galaxy in this sample, and this contribution drops rapidly to about 10 per cent by 19 magnitudes. In these cases, the sub-dominant component will be much less apparent in the image and, therefore, less important to the overall  $\chi^2$  of the fit, allowing for greater error in the parameters of that component. In addition, once Poisson noise is considered, these dimmer components suffer from much lower S/N. Later tests (Chapter 4.2.4) show substantial error on these components due to the low flux and resulting low S/N.

Additionally, sub-dominant components (in particular, bulges) may be much smaller than the overall size of the galaxy. This makes bulge parameter recovery susceptible to resolution effects. These effects are also explored in Chapter 4.2.4.

The magnitude and half-light radius are also well constrained when a **Ser** galaxy is fit with a **Ser-Exp** profile (Figure 4.7a). However, a **Ser-Exp** galaxy fit with a **Ser** profile produces large biases in the magnitude and half-light radius (Figure 4.8a).

As already mentioned, the total magnitude and half-light radius are well constrained ( $\sigma_{\text{total mag}} \approx 0.01$  mag and  $\sigma_{\text{radius}} \approx 1$  per cent) in cases where the correct model is applied to the mocks (i.e., **Ser** mock fit with a **Ser** model). This is not always the case when the wrong model is applied (i.e., **Ser-Exp** mock fit with a **Ser** model). When attempting to

## 4.2 Testing PyMorph image decompositions

---

fit the simulated **Ser-Exp** mocks with a **Ser** model, there is a measurable bias of order .01 magnitudes in total magnitude. The scatter of both the size and magnitude is increased by an order of magnitude. This bias and increased scatter becomes even larger in later tests. It is obvious that a single-component galaxy cannot properly model a two-component galaxy in general, and therefore, significant problems would be expected when attempting to fit a single-component profile to a two-component galaxy. Nevertheless, this type of fit is often performed on real data at low to moderate resolution and S/N where it is unlikely to recover a robust two-component fit. An important observation is that the **Ser-Exp** fit provides the most stable estimate of the half-light radius and total magnitude regardless of the true simulated galaxy model (**Ser**, **deV-Exp**, or **Ser-Exp**). The additional freedom in the **Ser-Exp** model and the fact that the **Ser** and **deV-Exp** models are special cases of the **Ser-Exp** model would lead us to expect this result. Therefore, it is advisable to always use a **Ser-Exp** fit in the course of fitting SDSS galaxies unless there is specific evidence to the contrary.

One systematic effect in the pipeline that has been noted by other groups (e.g. Blanton et al. 2005b; Guo et al. 2009), is the underestimate of Sérsic index at larger Sérsic indexes. At Sérsic indexes of  $n \approx 4$ , the Sérsic index is underestimated by less than 1 per cent. However, this underestimate increases in the later tests. The data suggest that a substantial component of this error is due to the resolution limits of the SDSS sample. At larger Sérsic index, a high sampling rate at the center of the galaxy is useful in distinguishing the preferred value of the Sérsic index. The effect of image resolution is further explored in Chapter 4.2.4.

Since no Poisson noise is added to these images, the scatter apparent in these fits is a combination of the limitations of the SDSS data (in particular resolution), systematics inherent in the PYMORPH routine (as well as the GALFIT routine used by PYMORPH), and any parameter degeneracies inherent in the models.

## 4. SIMULATIONS

---

GALFIT uses the Levenberg-Marquardt minimization method (Press et al. 1992) to find the minimum of the  $\chi^2$  distribution of the fit. The Levenberg-Marquardt method is not a global search algorithm but rather follows the steepest decent to a local minimum. As the parameter space becomes more complicated, GALFIT has more trouble accurately recovering parameters. Adding components to the fit (i.e., going from a one-component to two-component fit or going from a fixed Sérsic index component to one with a free Sérsic index) will not only complicate the  $\chi^2$  surface, making convergence less likely, but may introduce true degeneracies in the parameter space.

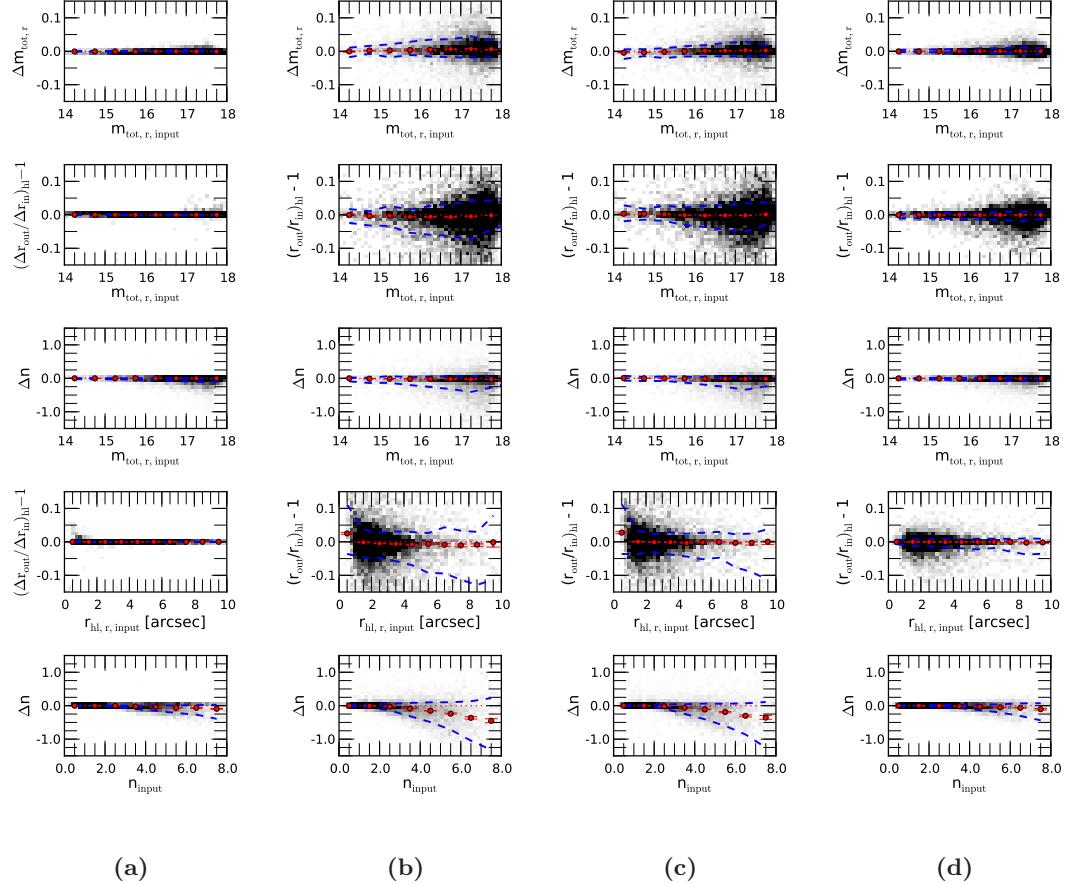
For instance, the **Ser-Exp** fit of a galaxy of very late type often suffers from over-fitting. The bulge component will tend to fit the disk of the galaxy as a second disk component with  $n_{\text{bulge}} \approx 1$ . This is obviously an unintended solution to the fitting but one that is equally valid from an  $\chi^2$  perspective. In practice, it is difficult to prevent this type of convergence without artificially constraining the fitting routine. Such constraints are generally discouraged and can lead to other negative effects including convergence to a non-optimal solution. The best solution to the parameter degeneracy is close examination of any two-component fits in cases where  $n_{\text{bulge}} \approx 1$ , or  $B/T \approx 0$  or 1.

In addition, PYMORPH reports statistical error estimates on the fitted parameters as returned from GALFIT. These errors are found to be an underestimate of the true error in the fits by as much as an order of magnitude. This gross underestimation of the error is also reported by Häussler et al. (2007) as well as being discussed in the GALFIT user notes<sup>6</sup>. Following Häussler et al. (2007), the ratio of the uncertainty reported by GALFIT to the deviation of the measured parameters (referred to as  $\sigma/\Delta$ ) should be greater than 1 for approximately 68 per cent of the data if the estimated uncertainty is appropriate. However, this is not the case for any of the parameters in the fits. We discuss a simple method for correcting the systematic bias and estimating the uncertainty in Chapter 4.4.

---

<sup>6</sup>See the technical FAQs at <http://users.obs.carnegiescience.edu/peng/work/galfit/TFAQ.html>

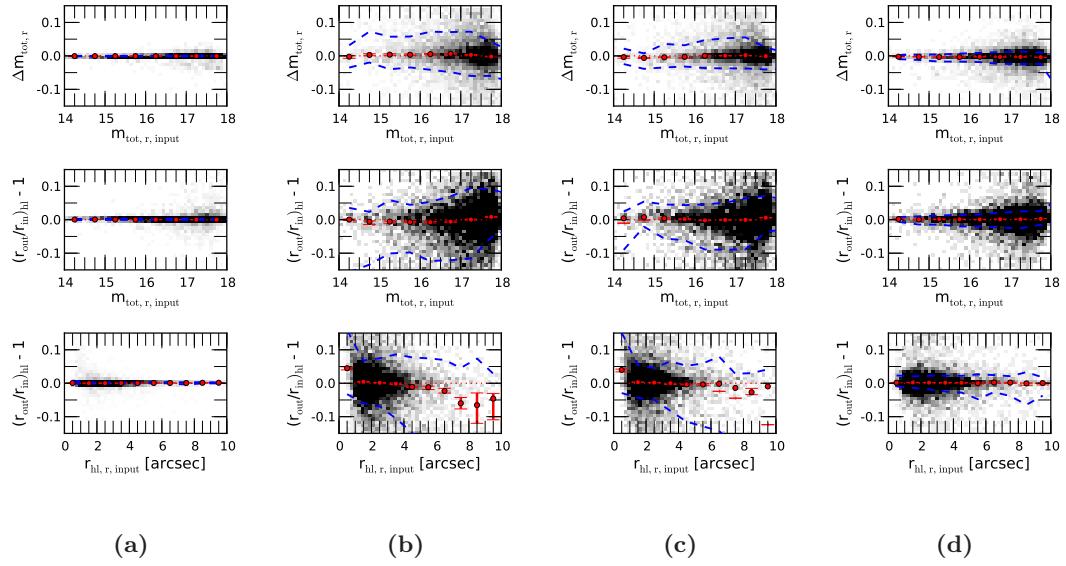
## 4.2 Testing PyMorph image decompositions



**Figure 4.6:** The simulated and recovered apparent magnitude, half-light radius, and Sérsic index for a **Ser** galaxy fit with a **Ser** model in four cases: **(a)** the image prior to adding Poisson noise, **(b)** the fiducial case containing simulated sky, Poisson noise, PSF errors, and neighboring sources, **(c)** the fiducial case with S/N increased by a factor of 4, and **(d)** the fiducial case with resolution increased by a factor of 2. Over-plotted are the bias (red points) in the fitted values. All plots show the 68 per cent (dashed line) scatter in blue. The density of points is plotted in gray-scale. The Sérsic index shows increasing underestimate up to  $\approx 0.5$  (or  $\approx 6$  per cent) at the largest Sérsic indexes.

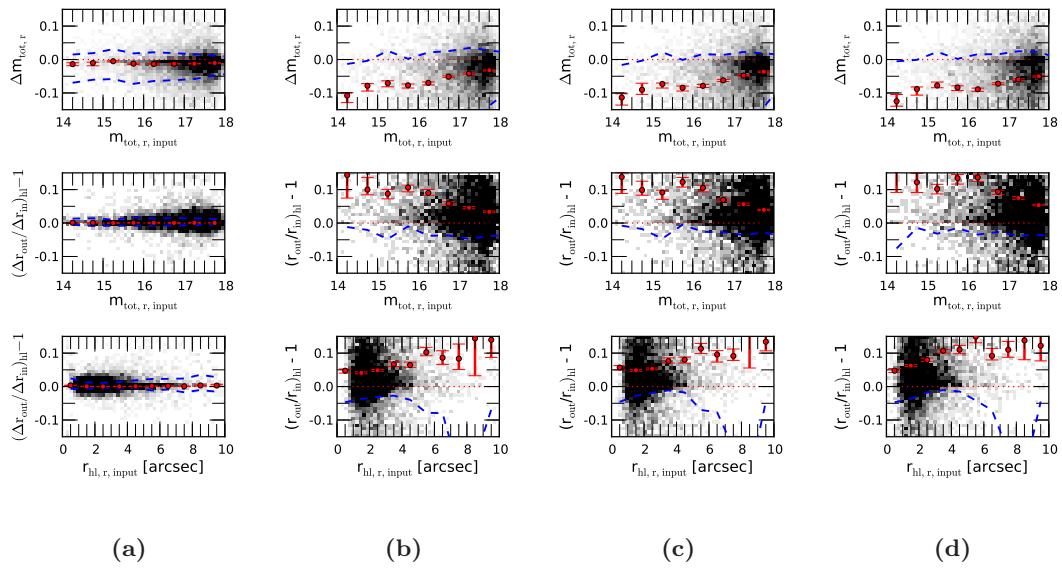
## 4. SIMULATIONS

---



**Figure 4.7:** The simulated and recovered apparent magnitude and half-light radius for a `Ser` galaxy fit with a `Ser-Exp` model in four cases: (a) the image prior to adding Poisson noise, (b) the fiducial case containing simulated sky, Poisson noise, PSF errors, and neighboring sources, (c) the fiducial case with S/N increased by a factor of 4, and (d) the fiducial case with resolution increased by a factor of 2. Over-plotted are the bias (red points) in the fitted values. All plots show the 68 per cent (dashed line) scatter in blue. The density of points is plotted in gray-scale.

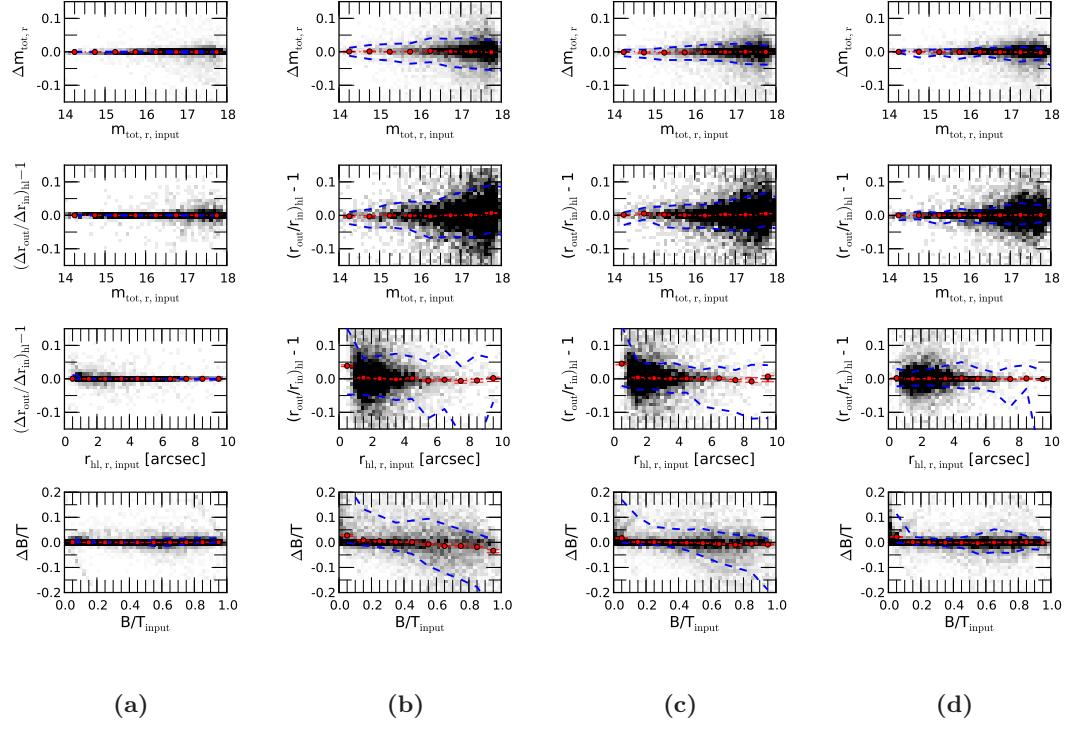
## 4.2 Testing PyMorph image decompositions



**Figure 4.8:** The simulated and recovered apparent magnitude and half-light radius for a **Ser**-**Exp** galaxy fit with a **Ser** model in four cases: **(a)** the image prior to adding Poisson noise, **(b)** the fiducial case containing simulated sky, Poisson noise, PSF errors, and neighboring sources, **(c)** the fiducial case with S/N increased by a factor of 4, and **(d)** the fiducial case with resolution increased by a factor of 2. Over-plotted are the bias (red points) in the fitted values. All plots show the 68 per cent (dashed line) scatter in blue. The density of points is plotted in gray-scale. The inability of the **Ser** profile to accurately model a **Ser**-**Exp** galaxy is clearly evident. Errors in magnitude and half-light radius are correlated and the error in radius is largely driven by errors in the largest, brightest objects. However, systematic errors occur even at the dimmer magnitudes. **Ser** fits tend toward recovering larger, brighter objects when applied to a true two component galaxy.

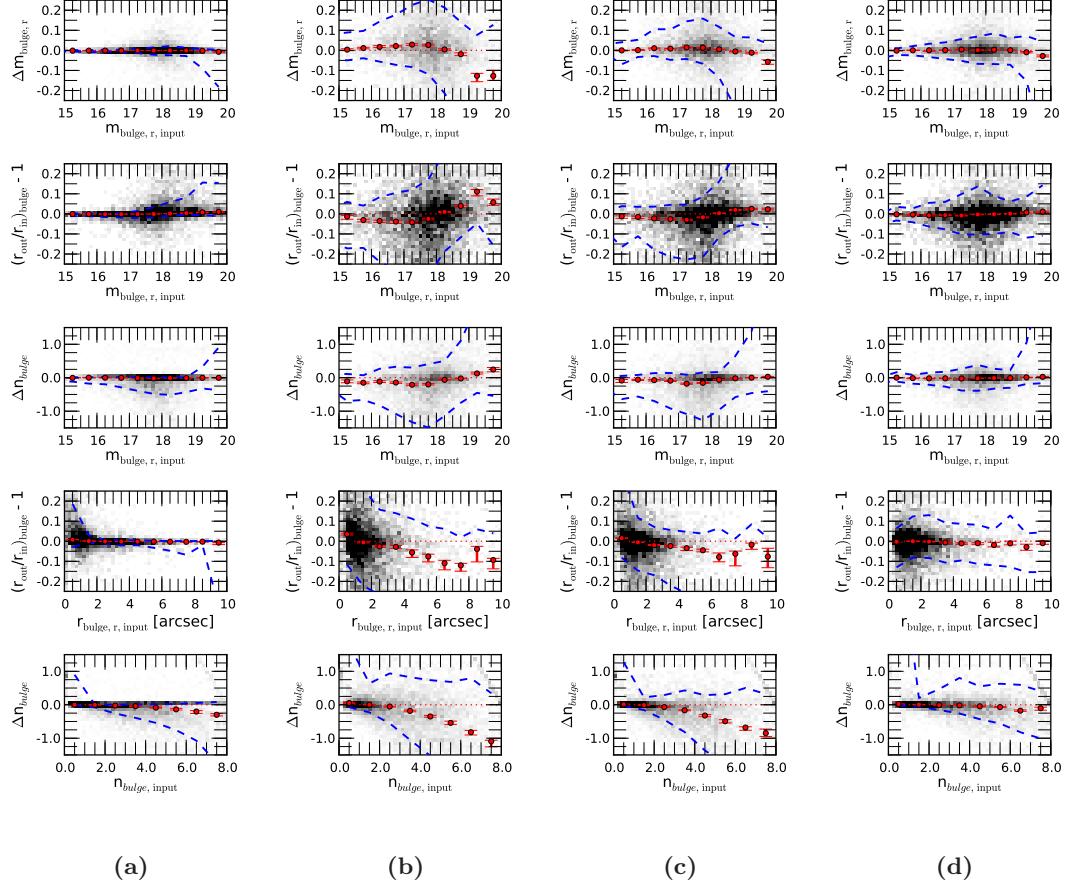
## 4. SIMULATIONS

---



**Figure 4.9:** The simulated and recovered apparent magnitude, half-light radius, and bulge-to-total light ratio for a Ser-Exp galaxy fit with a Ser-Exp model in four cases: **(a)** the image prior to adding Poisson noise, **(b)** the fiducial case containing simulated sky, Poisson noise, PSF errors, and neighboring sources, **(c)** the fiducial case with S/N increased by a factor of 4, and **(d)** the fiducial case with resolution increased by a factor of 2. Over-plotted are the bias (red points) in the fitted values. All plots show the 68 per cent (dashed line) scatter in blue. The density of points is plotted in gray-scale. The apparent trend in B/T from overestimation at lower B/T values to underestimation at higher B/T values is largely due to the boundaries on the parameter space forcing the scatter to be asymmetric (e.g. a galaxy with true B/T=1 cannot be estimated to have B/T>1).

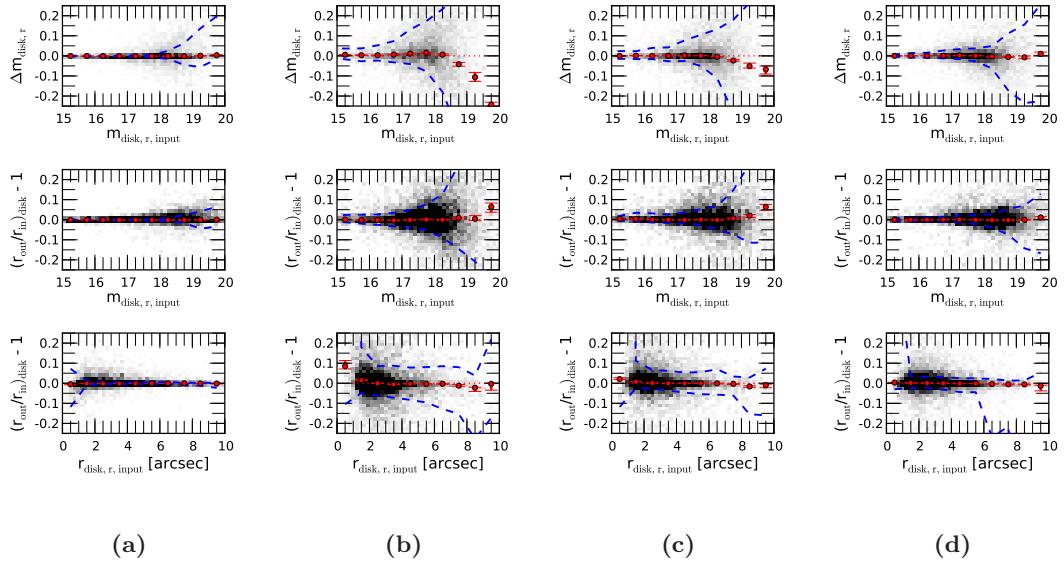
## 4.2 Testing PyMorph image decompositions



**Figure 4.10:** The simulated and recovered bulge component parameters for a Ser-Exp galaxy fit with a Ser-Exp model in four cases: (a) the image prior to adding Poisson noise, (b) the fiducial case containing simulated sky, Poisson noise, PSF errors, and neighboring sources, (c) the fiducial case with S/N increased by a factor of 4, and (d) the fiducial case with resolution increased by a factor of 2. Over-plotted are the bias (red points) in the fitted values. All plots show the 68 per cent (dashed line) scatter in blue. The density of points is plotted in gray-scale.

## 4. SIMULATIONS

---



**Figure 4.11:** The simulated and recovered disk component parameters for a Ser-Exp galaxy fit with a Ser-Exp model in four cases: (a) the image prior to adding Poisson noise, (b) the fiducial case containing simulated sky, Poisson noise, PSF errors, and neighboring sources, (c) the fiducial case with S/N increased by a factor of 4, and (d) the fiducial case with resolution increased by a factor of 2. Over-plotted are the bias (red points) in the fitted values. All plots show the 68 per cent (dashed line) scatter in blue. The density of points is plotted in gray-scale.

---

## 4.2 Testing PyMorph image decompositions

### 4.2.2 The effects of background, neighbor sources, and incorrect PSF extraction

When analyzing real data, it is not possible to extract the PSF at the target galaxy to arbitrary accuracy. Interpolation is required and generally performed on a network of the nearest stars to the target galaxy. This effect is tested through extraction of a neighboring PSF to be used during fitting in place of the PSF used to generate the image.

The neighbor PSF used in fitting is randomly selected from a location within a 200 pixel box surrounding the source. This provides approximately even sampling of distances from nearly 0 to about 170 pixels in separation from the source which corresponds to a separation of  $\approx 0$  to  $\approx 67.32$  arcseconds between the target galaxy and the location used for PSF extraction. This inserts some PSF error into the process of fitting as would be expected in the case of real data. However, it also retains the similarity between the PSF used for simulation and the PSF used for fitting. A strong similarity between the two would be expected since the PSF generally will not vary greatly over the area of a single fpC image.

Target galaxies are randomly inserted into the simulated fpC images described in Chapter 4.1.3. The simulated fpC images contain sky as well as neighboring sources. The PSF of the neighboring sources will have a different PSF than the target galaxy. This effect is not of concern in this work.

Prior to fitting, a new cutout is extracted from the total image (containing the target galaxy and background) ensuring that the target galaxy is at the center of the stamp image. By constructing new postage stamp images in this manner, there is sufficient variation in the background and fitting the incorrect galaxy can be prevented.

These fits (containing error in PSF reconstruction, neighboring sources, and noise) are the closest simulation to actual observing conditions analyzed in this work. Therefore the fits and the resulting measures of scatter and bias are adopted as the fiducial estimates of scatter and bias when using the pipeline.

## 4. SIMULATIONS

---

Figures 4.6b, 4.7b, 4.8b, 4.9b, 4.10b, and 4.11b show that the input values are recovered with marginal scatter. The total magnitude and half-light radius remain well constrained ( $\sigma_{\text{totalmag}} \approx 0.05$  mag and  $\sigma_{\text{radius}} \approx 5$  per cent) in cases where the correct model is fit to the mock galaxy. However, this scatter becomes larger when the wrong model is fit. The underestimate of the Sérsic index, particularly at large values, persists.

Further examination of the two-component fits show that the pipeline has difficulty extracting dim components (bulge or disk magnitude dimmer than  $\approx 18.5$ ). In these ranges, the components are observed at lower S/N and the pipeline loses sensitivity to the model parameters. The **Ser-Exp** fit shows an underestimate of Sérsic index, which is even stronger than in the single-component case, and an underestimate of bulge radius. However, the disk parameters remain unbiased with an increase in scatter of the model parameters. The increased stability of the disk parameters relative to the bulge parameters was also noted in Simard et al. (2011). In their paper, the authors comment that this may be due to the fixed profile shape (due to the fixed Sérsic index,  $n = 1$ ) or to the fact that on average bulges are more compact than disks leading to a resolution effect. This stability is the result of the increased resolution as disk sizes in the sample are roughly 3 times the FWHM of the PSF while bulges are smaller, on average approximately equal to the FWHM of the PSF in size. Further discussion is in Chapter 4.2.4.

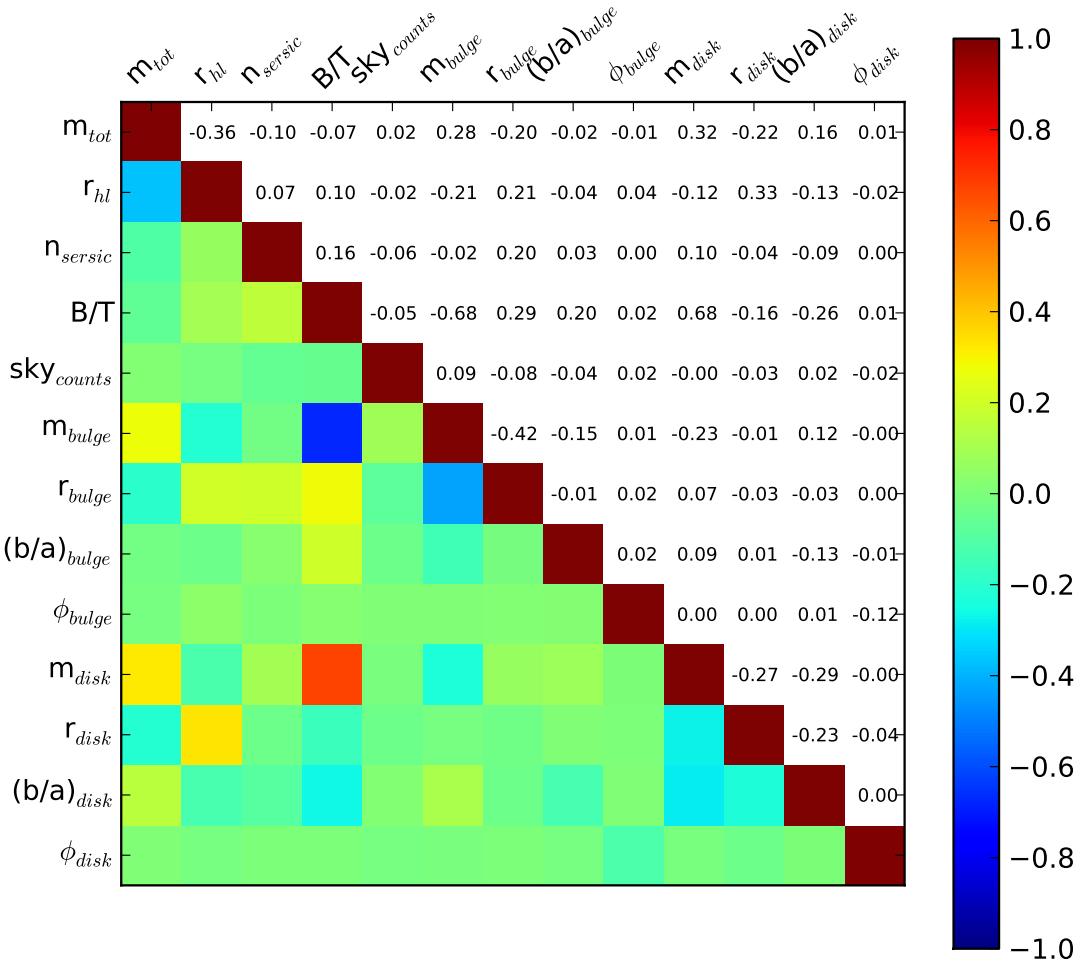
In general, the **Ser-Exp** fits are problematic and require much care when analyzing individual components. However, as was already shown, total magnitude and half-light radius are still tightly constrained.

Table 4.1 summarizes the bias and scatter in the fits; they exhibit trends with both the input value of the parameter and the input magnitude of the galaxy. This behavior is not properly encapsulated in the overall measure of bias, so these values are useful only as an example of the relative scale of bias and scatter for each parameter.

Errors can be correlated across many fit parameters. Figure 4.12 shows the error correlation matrix for the **Ser-Exp** mocks fit with a **Ser-Exp** model. The strong correlations

## 4.2 Testing PyMorph image decompositions

---



**Figure 4.12:** The correlation matrix for a mock Ser-Exp galaxy fit with a Ser-Exp model.

## 4. SIMULATIONS

---

between bulge-to-light ratio and the bulge and disk magnitudes are expected as well as the correlation among the radii of the bulge component with the Sérsic index. While the correlation matrix suggests that there is little correlation between sky estimation error and the fitted parameters, there is indeed a strong correlation in model errors with sky estimation error.

The apparent lack of correlation of sky error with the other fitting parameters is somewhat surprising. However, Figures 4.21 and 4.22 suggest a possible explanation for the apparent lack of correlation. Correlation of parameter errors with sky errors is non-linear and asymmetric with respect to over- or underestimating the sky. The fits discussed in this section are shown on Figures 4.21 and 4.22 in red. These points lie in a region where sky error does not significantly bias most parameters. In addition, the scatter of the sky values is quite small. This small scatter prevents us from sampling the broader covariance of the sky. If, for example, the recovered sky value was an underestimate of 0.5 per cent, then there would be a measurable covariance of fitting parameters with sky due to the steepness of the parameter bias with respect to sky level. The sky estimation is discussed further in Chapter 4.2.7.

### 4.2.3 Testing with real images

To verify the validity of the simulated background and to test the fitting pipeline in clustered environments, the mock galaxies were inserted into real SDSS fpC images. The fpC images are selected from SDSS DR7 images containing spectroscopic galaxy targets.

Plots of the fitted values are omitted here, because the scatter and the bias in the fits remain unchanged, suggesting that the sky background and neighboring sources common to an SDSS spectroscopic galaxy are properly modeled.

Dense environments provide an additional test for the pipeline. The GMBCG catalog (Hao et al. 2011) provides a selection of fpC images that contain dense environments. Brightest cluster galaxies (BCGs) with galaxies in the original catalog to select fpC images

## 4.2 Testing PyMorph image decompositions

---

Simulated model	Parameter	fitted model			
		Ser		SerExp	
		bias	1- $\sigma$	bias	1- $\sigma$
Ser	$m_{tot}$ [mag]	$0.00 \pm 0.09$	$[-0.02, 0.02]$	$-0.02 \pm 0.18$	$[-0.11, 0.02]$
	$r_{hl}$ [arcsec]	$-0.01 \pm 0.68$	$[-0.14, 0.01]$	$0.11 \pm 1.53$	$[-0.11, 0.04]$
	sky [ per cent]	$-0.05 \pm 0.14$	$[-0.11, -0.05]$	$-0.05 \pm 0.14$	$[-0.11, -0.05]$
	Sérsic Index	$-0.08 \pm 0.57$	$[-0.44, 0.01]$	–	–
SerExp	$m_{tot}$ [mag]	$-0.07 \pm 0.18$	$[-0.22, -0.01]$	$-0.02 \pm 0.15$	$[-0.09, 0.01]$
	$r_{hl}$ [arcsec]	$0.49 \pm 1.86$	$[-0.02, 0.30]$	$0.07 \pm 1.17$	$[-0.07, 0.04]$
	sky [ per cent]	$-0.08 \pm 0.15$	$[-0.16, -0.07]$	$-0.06 \pm 0.13$	$[-0.11, -0.05]$
	B/T	–	–	$0.00 \pm 0.15$	$[-0.07, 0.03]$
	$m_{bulge}$ [mag]	–	–	$-0.14 \pm 0.71$	$[-0.73, 0.06]$
	$m_{disk}$ [mag]	–	–	$-0.04 \pm 0.50$	$[-0.41, 0.05]$
	$r_{bulge}$ [arcsec]	–	–	$0.08 \pm 0.97$	$[-0.27, 0.11]$
	$r_{disk}$ [arcsec]	–	–	$0.07 \pm 0.82$	$[-0.14, 0.08]$
	Sérsic Index	–	–	$0.06 \pm 1.98$	$[-0.90, 0.14]$

**Table 4.1:** The bias and scatter of the fitted parameters of the simulated images with background and PSF effects. These values are provided for illustrative purposes only. There is much underlying structure in the errors when compared to their respective input values or the magnitude of the component.

with cluster members including the BCG. The mock galaxies are then inserted into the image which is run through the pipeline. In the previous simulations, intracluster light and gradients in the sky were not modeled. These tests show what the effects may be. Once again, the errors remain unchanged, showing that no environmental correction is necessary when using the fits from the pipeline.

Placing mock galaxies near cluster members tests for systematic effects in crowded fields. However, further examination of BCG galaxies is necessary before properly modeling them for this purpose. For example, the curvature at the bright end observed in the size-luminosity relation of early-type galaxies (see Bernardi et al. 2013) appears to be due to an increasing incidence of BCGs, which define steeper relations than the bulk of the early-type population (e.g. Bernardi et al. 2007, 2013). However, the curvature could also be due to intracluster light (e.g. Bernardi 2009). The ability to test the systematic effects associated with BCGs using the method outlined above is severely limited due to the existence of a BCG at the preferred location for placing the test galaxy (i.e., the center of the cluster).

## 4. SIMULATIONS

---

Therefore, the stability of recovered fit parameters with respect to environment cannot be assumed to extend to BCGs based on the tests presented here alone. Further tests for the largest, brightest galaxies are needed to explore this possibility.

### 4.2.4 Varying the S/N and pixel size

The effects of the S/N and image resolution on the bias and scatter in the fits can be quantified. Figures 4.6c, 4.7c, 4.8c, 4.9c, 4.10c, and 4.11c show the effect of increasing the S/N by a factor of 4 while holding all other parameters fixed. Similarly, Figures 4.6d, 4.7d, 4.8d, 4.9d, 4.10d, and 4.11d show the effect of increasing resolution by a factor of 2 while holding S/N constant. Corresponding decrements in these parameters were performed, although they are not presented in this paper.

Improving the resolution by a factor of two substantially improves the ability to recover the radius and Sérsic index with reduced bias. For instance, the Sérsic index bias is reduced to  $\approx 0.1$  at the larger values. Additionally, the bulge parameters of the **Ser-Exp** fit improve substantially with improved resolution. Corresponding changes in the S/N reduce the scatter, but by a small amount relative to the effect of the resolution change. In addition, changing the S/N does not remove the observed bias in Sérsic index or bulge size. This leads us to conclude that the limitations of the resolution of SDSS are the leading factor in causing systematic offsets in the half-light radius, Sérsic index, and other fitting parameters (including the bulges of the **Ser-Exp** fits). While increasing the S/N will reduce the scatter in the fits, increased resolution is necessary to properly recover unbiased values.

Lackner and Gunn (2012) also examined the effects of changing S/N and resolution on SDSS galaxies (see Figures 5-11 of their paper). The authors found that decreased resolution and S/N increases the relative error in the Sérsic index and radius. They recommended that **Ser** galaxies (and the bulge and disk sub-components of two-component galaxies) have radii,  $R_{hl} \gtrsim 0.5 \times \text{FWHM}$ . This cut removes  $\approx 1$  per cent of the **Ser** mocks and  $\approx 22$  per

## 4.2 Testing PyMorph image decompositions

---

cent of the **Ser-Exp** mocks from the simulated samples with a preference toward galaxies above  $z = 0.05$ .

While this condition is sufficient to keep the relative error in the half-light radius and Sérsic index comparable to the error in the magnitude, this condition fails to remove the bias in the galaxy samples. Figure 4.6b shows that the underestimate of Sérsic index occurs at larger values. These galaxies tend to exhibit radii larger than the PSF. Given that the average FWHM of PSFs in the sample is  $\approx 1.3''$ , it is not possible to remove the bias in Sérsic index. Clearly, reliable measurements are dependent on both the Sérsic index of the object and its radius relative to the resolution. Both parameters must be accounted for when deciding on an appropriate resolution cut.

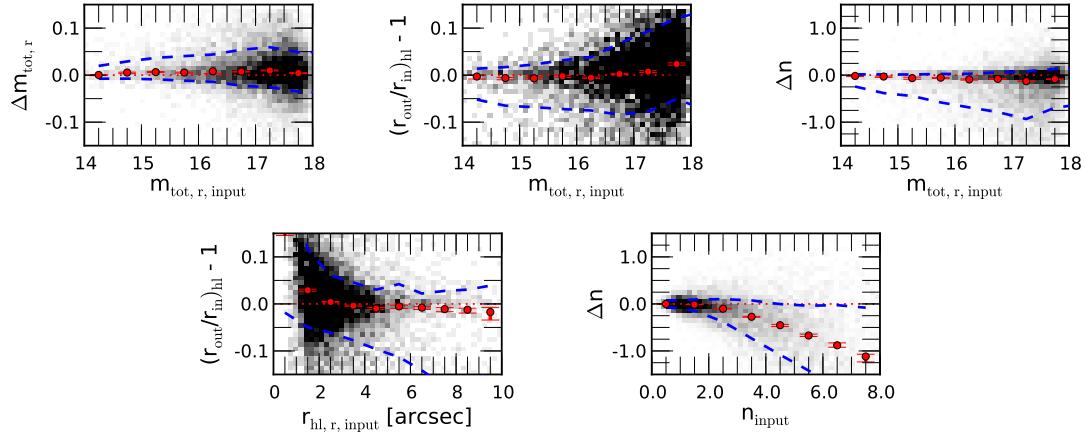
Extending the Lackner and Gunn (2012) recommendation to include a Sérsic index dependent term is sufficient to provide for recovery of Sérsic index  $> 4$  with bias  $\approx 0.1$  or  $\approx 1$  per cent. Galaxies should have circularized half-light radii  $R_{hl} \gtrsim 0.5 * \text{FWHM} \times n$ . This removes nearly 75 per cent of the sample. While such large cuts are sufficient to remove the bias in radius and Sérsic index for the **Ser** fits, the data are certainly biased relative to the original catalog after the cuts. Rather than remove these galaxies, it is possible to correct for the bias following a simple statistical argument presented in Chapter 4.4.

### 4.2.5 Effect of cutout size

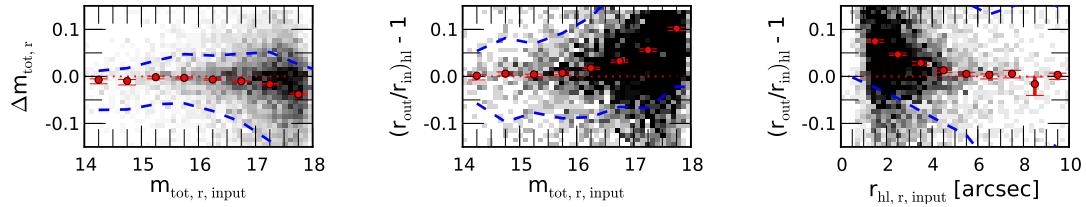
It is important to select a cutout size for fitting that does not significantly bias the fits produced by **PYMORPH**. The most important consideration is to provide enough sky pixels to allow the **PYMORPH** program to properly determine the sky level in the images. Reducing cutout size may cause overestimation of background and corresponding errors in the other fit parameters. However, the **PYMORPH** pipeline and **GALFIT** fit a constant background to the galaxy image. Since a larger image could make sky gradients more significant, this could bias the fits when a larger cutout is used. Minimizing error when estimating the sky

#### 4. SIMULATIONS

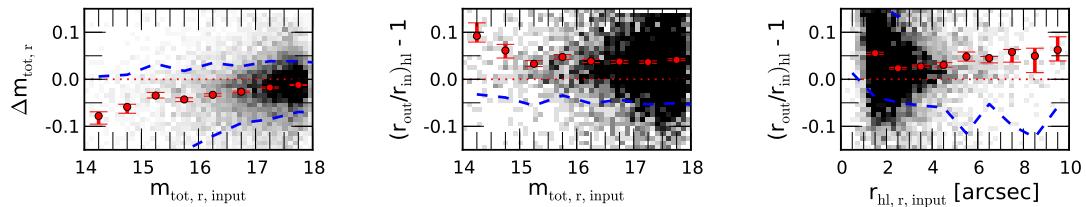
---



**Figure 4.13:** The `Ser` galaxies fit with a `Ser` model for resolution degraded by a factor of 2.

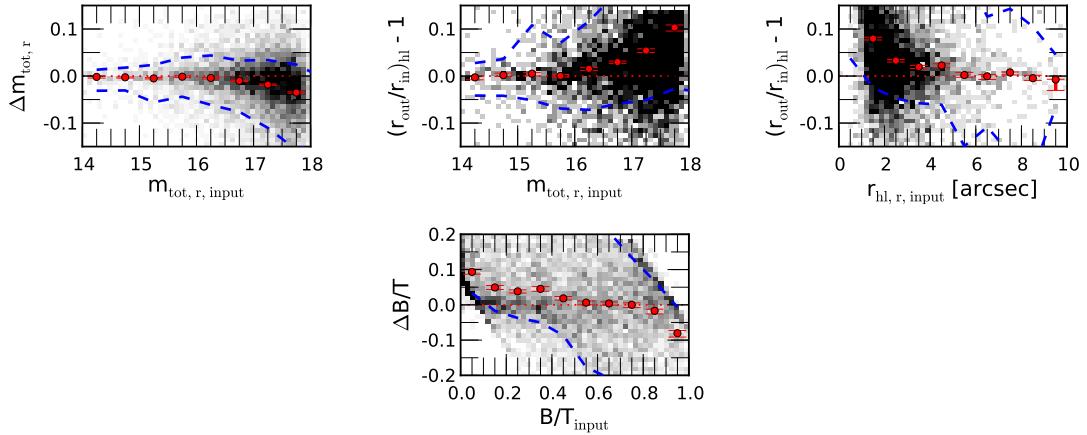


**Figure 4.14:** The `Ser` galaxies fit with a `Ser-Exp` model for resolution degraded by a factor of 2.



**Figure 4.15:** The `Ser-Exp` galaxies fit with a `Ser` model for resolution degraded by a factor of 2.

## 4.2 Testing PyMorph image decompositions



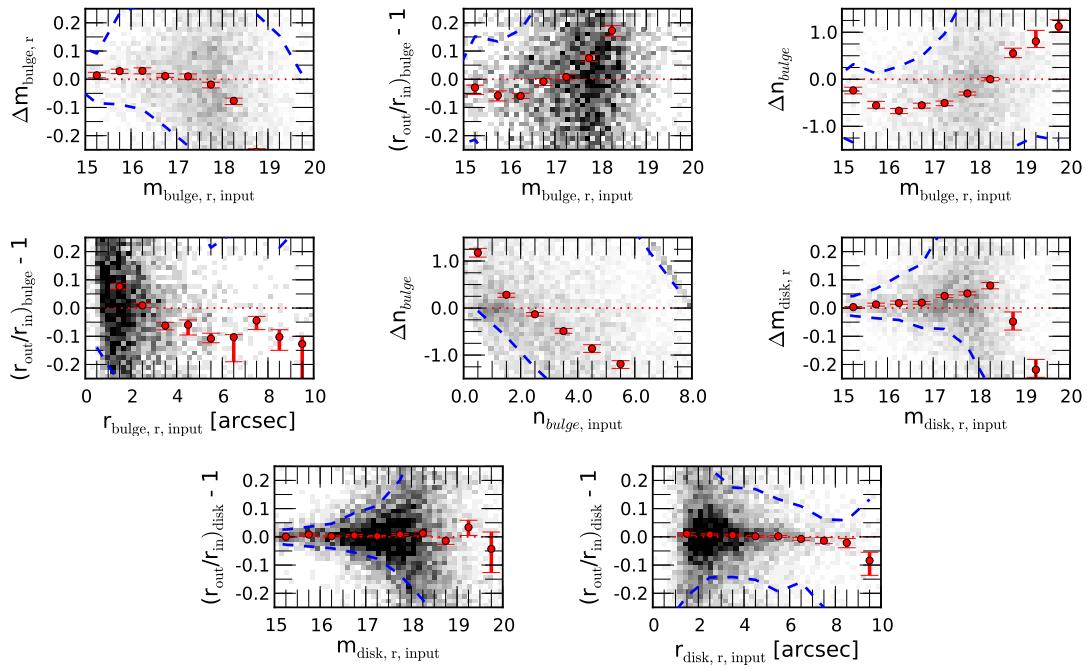
**Figure 4.16:** The Ser-Exp galaxies fit with a Ser-Exp model for resolution degraded by a factor of 2.

level without introducing a gradient term and further complicating the fitting process is preferred.

To test for optimal cutout size, mocks were fit with cutout sizes between 10 and 25 Petrosian half-light radii. In Figure 4.18 the error and  $1-\sigma$  scatter in the error on the total magnitude, half-light radius and sky (showing SExtractor sky in blue and estimates from this work in red) are plotted as a function of cutout size. Smaller sizes clearly bias sky estimates made by SExtractor, but only minor improvement in the scatter of any parameters is achieved by using cutout sizes above 20 half-light radii. Since the SExtractor sky is the starting point for the fitting, a size of 20 half-light radii is used for the images. The sky estimates of SExtractor improve substantially. However, GALFIT sky estimation is stable over these sizes. Because GALFIT sky estimation is largely independent of the initial starting SExtractor value (which is expected if GALFIT is truly finding the best fit to the galaxy), it is likely the case that cutout sizes smaller than even 10 half-light radii could be used for analysis.

#### 4. SIMULATIONS

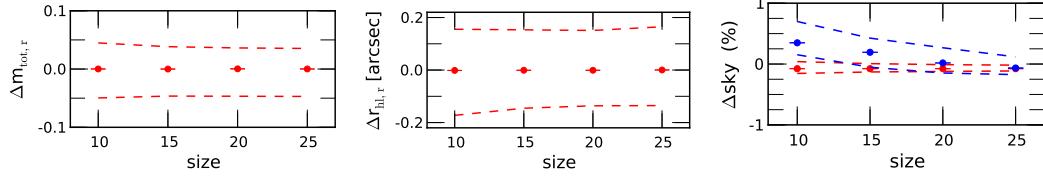
---



**Figure 4.17:** The sub-components of Ser-Exp galaxies fit with a Ser-Exp model for resolution degraded by a factor of 2.

## 4.2 Testing PyMorph image decompositions

---



**Figure 4.18:** The mean difference of the total magnitude (left column), PSF-corrected half-light radius (center column), and sky estimation (right column) as a function of cutout size for **Ser-Exp** mocks fitted with a **Ser-Exp** model. Other simulated models behave similarly. For sky estimation, the sky measured by **GALFIT** is plotted in red. **SEXTRACTOR** sky measurements are plotted in blue for reference. One- $\sigma$  scatter in the fits is plotted as a dashed line. Improvement in scatter when fitting for cutout sizes above 20 Petrosian radii is limited, so a 20 half-light radii cutout size is used for all images. Fit parameters seem to have no sensitivity to cutout size in this range, suggesting that it may even be possible to use smaller cutouts.

Additional plots of other parameters are omitted in this section. The other fitted parameters show little or no sensitivity to cutout size in the range of cutout sizes used. However, as previously discussed, the bias and scatter may not be equally affected across all model parameters. The effects may be concentrated in a small part of the parameter space.

### 4.2.6 Masking and neighbor thresholding

Different settings of the masking configuration were also tested. These different settings provided no reduction in the scatter of the recovered parameters when tested on simulations. Since no improvement was evident, the default configuration was selected for analysis (i. e., masks are drawn according to the **SEXTRACTOR** segmentation image produced during fitting). While this likely leaves unmasked light from neighboring sources, it has no apparent effect on the determination of sky. The insensitivity of the sky estimation to this choice of masking is likely the result of the large cutout sizes used during fitting, which reduces sensitivity of sky estimation to stray light from masked neighbors. The large number of sky pixels used during fitting prevents skewed fitting results due to the small amount of

## 4. SIMULATIONS

---

stray light from masked neighbors. Different masking and handling of neighboring sources was carried out by S11 as well as in Barden et al. (2012). The masking generally covers a smaller area around each source when compared to these two works.

SExtractor uses a flexible elliptical aperture around every detected object and measures the aperture flux in the background-subtracted image. This elliptical aperture is generated from moments of the light distribution in pixels above a given threshold (called the DETECT\_THRESHOLD). Therefore, the detection threshold will change the auto-mag value by changing the aperture size. Of course any change to the SExtractor settings that influences the level of the sky will also cause a corresponding shift in the auto-mag value.

For example, the GAMA survey (Barden et al. 2012) chose the DETECT\_THRESHOLD to be 1.7. This decision is made by choosing the threshold that minimizes the difference in number of detected sources between GAMA SExtractor and SDSS Photo Pipeline (see Hill et al. 2011, for all the details)). The DETECT\_THRESHOLD was set at 1.5. This threshold represents the minimum excess of flux relative to the sky sigma necessary for a pixel to be considered part of the target galaxy. Therefore, lowering the threshold will make target sources appear larger and brighter. In practice, this has little effect on the fitted models, as the SExtractor values are only used as input. PYMORPH rarely converges to the values given by SExtractor. However, the SExtractor radii are used to generate the masks. Therefore, the masking will be slightly more conservative as compared to GAMA (PYMORPH will mask out a slightly larger region around sources). This difference in masking is small enough to have no effect.

There are other possible settings that could bias the SExtractor magnitude, but these appear to be the same in this work and GAMA fitting. For instance, both catalogs use  $2.5 \times$ Kron radius as the fixed aperture size for magnitude measurements. The PHOT\_APERTURES setting, which sets the minimum aperture used for photometry can also influence the measurement of magnitude by artificially inflating the value of the mag-

## 4.2 Testing PyMorph image decompositions

---

nitude for small objects. In practice, this has no effect since the galaxies in these in this work are not smaller than 5 pixels (the default value of PHOT\_APERTURES).

Figure 4.19 shows the scatter in `Ser` models fit with `Ser` profiles for several different masking settings. The first row shows the original masking (based on the segmentation image produced by SExtractor) adopted throughout this work. This mask is produced by masking the segmentation image for all sources identified by SExtractor, except for the target. The second, third and fourth rows show the result of masking an ellipse of radius 2, 4, and  $6 * \text{semi-major axis}$  for each neighboring source, respectively. The columns show the results of reducing the threshold for neighboring sources being simultaneously fit. The first column is the default settings of `threshold=3` and `threshold_area=0.2` used throughout this work and shown in Equations 3.6 and 3.7. The second column uses `threshold=4` and `threshold_area=0.2`. The third column uses `threshold=4` and `threshold_area=0.1`. Neighbor fits result in 10 per cent, 20 per cent, and 30 per cent of fits for the columns 1, 2, and 3, respectively. Magnitude is effectively independent of the thresholding and masking conditions for the total sample.

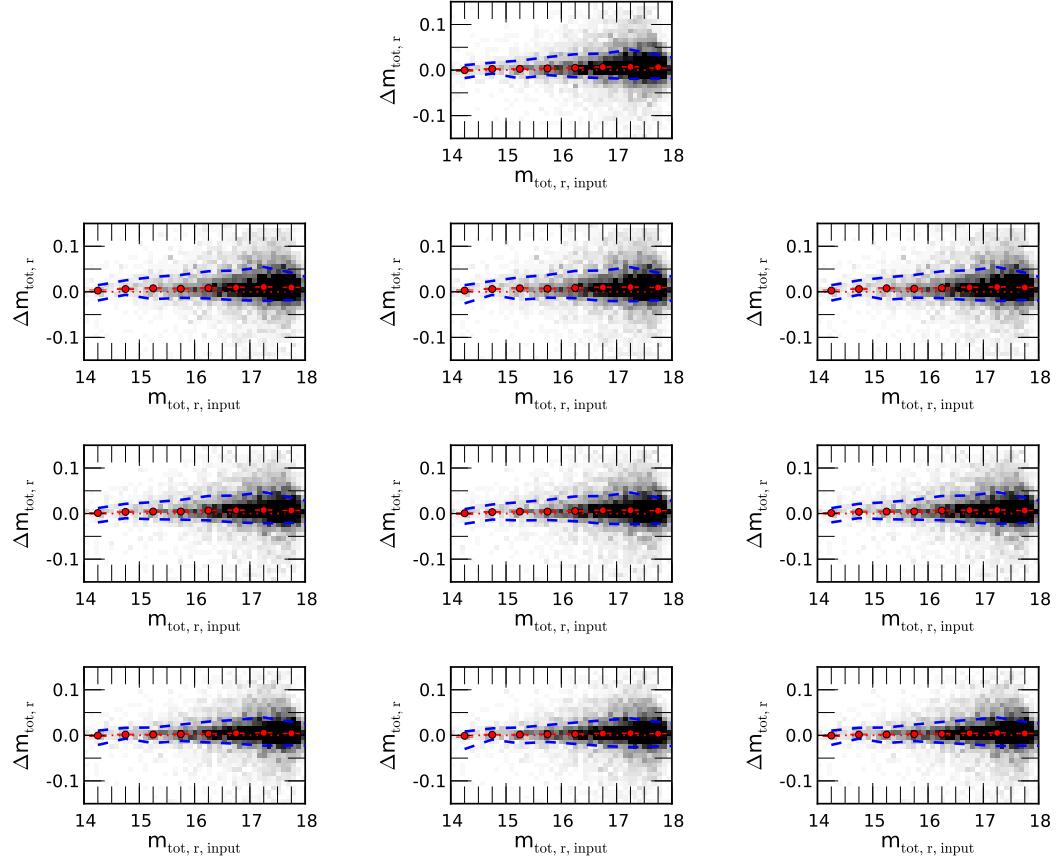
### 4.2.7 The effect of incorrect sky estimation

Estimation of the sky in the vicinity of the target galaxy has a high level of uncertainty. Indeed, accurate sky determination is likely not even a solvable problem as discussed briefly in Blanton et al. (2011). To determine the bias introduced by the sky estimation, the fitting pipeline was tested in cases of both underestimation and overestimation of the sky. The sky was fitted at the simulated sky level, as well as at simulated sky level  $\pm 0.5$  per cent and  $\pm 1.0$  per cent. These ranges were chosen to represent the range of differences between the sky estimations and those provided in the CASJOBS database for the SDSS spectroscopic sample.

Figure 4.20 shows a comparison of sky estimates using PYMORPH to those provided from the SDSS photometric data pipeline. This comparison is performed on data from the

## 4. SIMULATIONS

---



**Figure 4.19:** The scatter in `Ser` models fit with `Ser` profiles. The first row shows the original masking (based on the segmentation image produced by SExtractor) adopted throughout this work. The second, third and fourth rows show the result of masking an ellipse of radius 2, 4, and  $6 * \text{semi-major axis}$  for each neighboring source, respectively. The columns show the results of reducing the threshold for neighboring sources being simultaneously fit. The first column is the default settings of `threshold=3` and `threshold_area=0.2` as was done throughout this work. The second column uses `threshold=4` and `threshold_area=0.2`. The third column uses `threshold=4` and `threshold_area=0.1`. Neighbor fits result in 10, 20, and 30 per cent of fits for the columns 1, 2, and 3, respectively. Magnitude is effectively independent of the thresholding and masking conditions for the total sample.

## 4.2 Testing PyMorph image decompositions

---

catalog presented in M2014. The Figure shows the normalized distribution of differences in sky estimation in bins of 0.1 per cent. A negative difference indicates that the sky measured by PYMORPH is lower than that reported by SDSS. The vertical red solid line indicates the median of the distribution. The red dashed, dot-dashed, and dotted lines indicate the 68-95-99 per cent ranges of the data, respectively. The 95 per cent range of sky values is approximately between  $\pm 1$  per cent difference. This range is adopted as the range to test for sky variation.

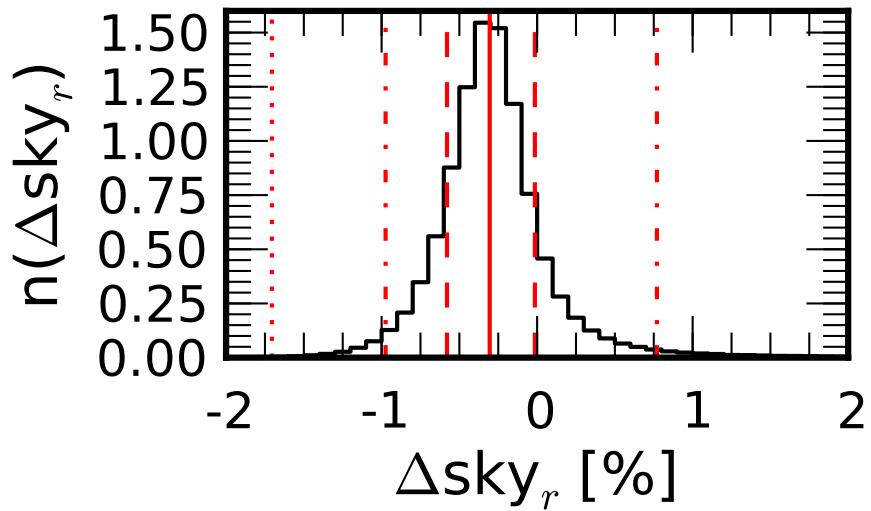
The results of incorrectly estimating the sky are shown in Figures 4.21 and 4.22. The results of fitting galaxies using the standard PYMORPH pipeline, treating sky level as a free parameter in the fit, are shown in red. PYMORPH systematically underestimates the sky at the 0.1 per cent level. However, the scatter is very tight as indicated by the vertical dashed red lines. The fitting results at fixed sky levels of the correct value and  $\pm 0.5$  per cent and  $\pm 1.0$  per cent are shown in black. Errors approaching 0.5 per cent clearly introduce a large bias in the fits. The 0.5 per cent level is an important level because it is the approximate level of overestimation shown in the preceding section (Chapter 4.2.5) found by SExtractor.

Note the asymmetry of the effects of incorrect sky estimation on fitting parameters. In particular, an underestimate of sky is much more detrimental to the fit than the corresponding overestimate. The reason for this asymmetry is due to changes in the perceived “flatness” of the profile at large radii. When the sky is overestimated, the galaxy profile tends to 0 flux too early. This causes a decrease in the Sérsic index and a decrease in the radius. However, when the sky is underestimated, there will be an extended, approximately constant brightness profile at larger radii. The only way to model such a profile is for Sérsic index to diverge to larger values which produce flat, extended profiles at large radii.

Guo et al. (2009) examined the effects of sky uncertainties in regards to the covariance between magnitude and both Sérsic index and half-light radius. They randomly sampled sky estimates from a distribution contained mostly within  $\pm 1$  per cent. They found similar variation of Sérsic index (varying by 2 or more in some cases of underestimating the sky

#### 4. SIMULATIONS

---



**Figure 4.20:** The percent difference between the sky estimate of PYMORPH for SDSS galaxies and the sky estimated by the SDSS photometric pipeline for those same galaxies. The normalized distribution of differences is shown in bins of 0.1 per cent. A negative difference indicates that the sky measured by PYMORPH is lower than that reported by SDSS. The vertical red solid line indicates the median of the distribution. The red dashed, dot-dashed, and dotted lines indicate the 68-95-99 per cent ranges of the data, respectively. The 95 per cent range of sky values is approximately between  $\pm 1$  per cent difference, this is adopted as the range used to test the effects of improper sky estimation.

### 4.3 Applicability of an F-test criteria

---

and varying by less than 1 in the case of overestimation). The asymmetry in bias due to incorrect sky estimation is apparent in Figure 5 of Guo et al. (2009), but not explicitly commented upon.

Figure 4.20 shows that PYMORPH consistently estimates the sky  $\approx 0.25$  per cent lower than that of the SDSS pipeline. Figures 4.18, 4.21, and 4.22 show that PYMORPH has a systematic underestimate of the true sky at the  $\approx 0.1$  per cent level. This bias is much smaller than the bias associated with using the SExtractor sky estimate as shown in Figure 4.20 (especially for smaller cutout size), which suggests that the sky values in SDSS are slightly overestimated.

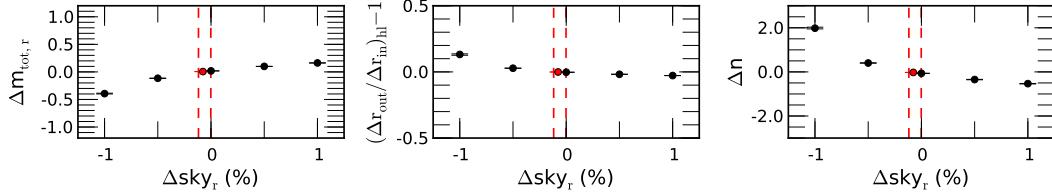
Ser-Exp disk components are remarkably robust to the errors in sky estimation, while bulge parameters suffer greatly, especially when the sky is underestimated. Upon further examination of Figure 4.22, the bulge parameters of the model are more accurately estimated when the sky is treated as a free parameter in the fit rather than when the sky is fixed at the correct value. However, this improvement does not suggest that underestimate of the sky is the preferred fitting outcome. It merely reflects the fact that the systematic effects due to underestimation of the sky are opposite to the underlying biases in half-light radius and Sérsic index. If the PYMORPH pipeline we applied to an image with higher S/N and increased resolution, the correct estimate of the sky would be preferred to prevent systematic overestimate of these parameters.

## 4.3 Applicability of an F-test criteria

The preceding sections have shown the covariance, bias, and scatter in the parameter estimation for the Ser and Ser-Exp models. In reality, the effects above will combine to yield a total scatter, covariance, and bias that should approach those shown in 4.2.2. the simulations give us an idea of the behavior of the PYMORPH pipeline when fitting SDSS galaxies as presented in M2014.

## 4. SIMULATIONS

---



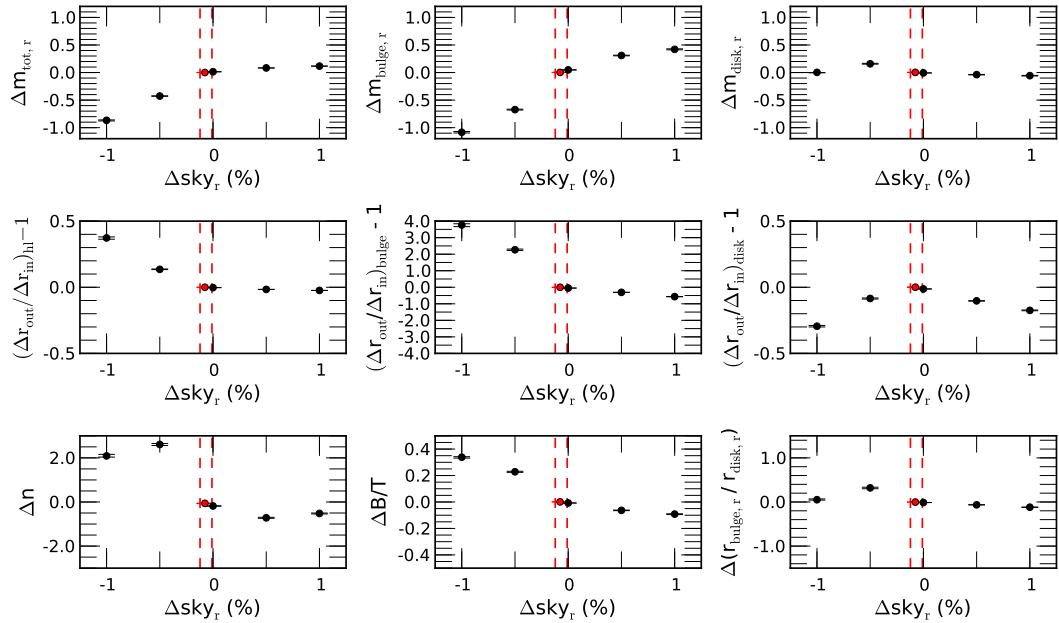
**Figure 4.21:** The simulated and recovered apparent magnitude (left), half-light radius (center), and Sérsic index (right) for a **Ser** galaxy fit with a **Ser** model. The residuals are plotted as a function of the sky level. Points plotted in black are from fits performed with fixed sky. The over-plotted points in red are the result of fitting with sky level as a free parameter in the fit. The vertical dashed red lines mark the 68 per cent scatter of the free sky determination. The fits are slightly biased low, and this contributes to a small overall bias in fit parameters.

The simulations show that the recovery of global fitting parameters (total magnitude and half-light radius) in the case of SDSS galaxies is remarkably robust, even in the case of the **Ser-Exp** fits. Two-component fits present a more difficult test for the pipeline. Both the bulge and disk components exhibit increased scatter relative to the scatter of the global parameters. In addition, the bulge component exhibits a systematic underestimation of the radius, Sérsic index, and magnitude, particularly for bulges with larger radii or higher Sérsic index.

The galaxies fit in this work have a median size roughly equivalent to the average PSF of SDSS. For most galaxies, the resolution necessary to accurately resolve bulge substructure is not present. As shown in Chapter 4.2.4, the ability to recover small bulges is improved by a factor of 2 increase in resolution. Finer resolution in central regions of the galaxy is also necessary to fully recover larger Sérsic indexes without bias. Even with these systematics, the two-component fits are still necessary to recover unbiased global parameters and can provide insight into the structure of galaxies.

The use of two-component models is potentially ill-advised for many SDSS galaxies as the respective sub-components may be too small to be well-resolved. This is suggested by

### 4.3 Applicability of an F-test criteria



**Figure 4.22:** The simulated and recovered parameters of a simulated Ser-Exp galaxy fit with a Ser-Exp profile. The residuals are plotted as a function of the sky level. Points plotted in black are from fits performed with fixed sky. The over-plotted points in red are the result of fitting with sky as a free parameter of the model. The vertical dashed red lines mark the 68 per cent scatter of the GALFIT sky determination. Notice that disk parameters are relatively robust while bulge parameters suffer from incorrect sky estimation. Underestimates of sky level have a particularly strong effect on the bulge.

## 4. SIMULATIONS

---

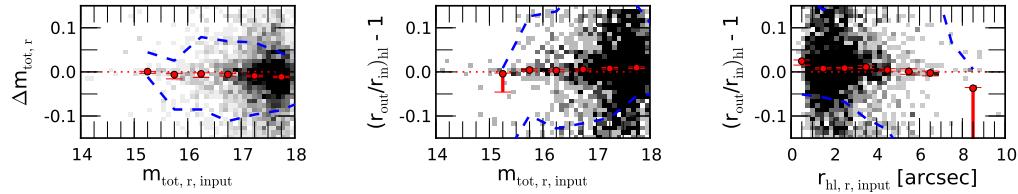
Simard et al. (2011) as well as Lackner and Gunn (2012) (if one uses the suggested resolution cut based on the PSF FWHM). However, the data show that this recommendation should be conditional on the galaxy parameters of interest. While it may be true that bulge parameters of the **Ser-Exp** fit become unreliable at small radii, using only the **Ser** fit radius will bias a sample of SDSS galaxies containing both single and two-component profiles (see Figure 4.8b). However, there are no cases where the **Ser-Exp** fit introduces bias. It is advisable to use the **Ser-Exp** half-light radius and magnitude as the total magnitude of the galaxy when examining a sample such as this.

The F-test offers a potentially powerful way to distinguish when it is necessary to use a more complicated two-component model. The F-test can compare the  $\chi^2$  values among nested linear models with Gaussian errors (Lupton 1993). Although the models are not linear and the error distribution is not strictly Gaussian, one can apply the F-test to the fits. Following Simard et al. (2011), an F-test probability of 0.32 is adopted as the cutoff indicating a more complicated model is required. When a low F-test probability is found,  $P_{correct} < 0.32$ , the more complicated model (i.e., going from a one-component to two-component fit, or allowing the Sérsic index of the bulge to vary) provides a better fit to the observed profile. In cases where a **Ser** fit is used rather than a **Ser-Exp** fit, the improvement in fitting is large enough to justify using a model with more free parameters. The improved fit is not merely the result of using a more flexible model. A similar test was performed by Lackner and Gunn (2012) to select among a pure disk or disk+bulge model.

If the selection based on the F-test is correct, then the resulting measurements of total magnitude and half-light radius will be unbiased. Using the **Ser-Exp** mocks fit with each of the **Ser** and **Ser-Exp** models, F-test comparing the **Ser** and **Ser-Exp** fits was preformed. The preferred fit (either **Ser** or **Ser-Exp**) of the **Ser-Exp** mocks is then used to assess the bias in the half-light radius and magnitude.

By examining the subset of **Ser-Exp** mocks for which the F-test determines the **Ser** model to be the appropriate fit, one can test the ability of the F-test to select galaxies that

### 4.3 Applicability of an F-test criteria



**Figure 4.23:** Total magnitude and half-light radius of **Ser** fits of **Ser-Exp** mocks shown by F-test to be sufficiently well fit by **Ser** models. The fits are unbiased, but the scatter in the recovered values are approximately twice as wide in half-light radius and magnitude as compared to the **Ser-Exp** fits in Figure 4.9b.

are correctly represented by **Ser** models. Figure 4.23 shows the resulting distribution of total magnitude and half-light radius of this subset of **Ser-Exp** mocks fit with **Ser** models. The bias originally observed in Figure 4.8b is not evident. However, the scatter in the recovered values are approximately twice as wide as in Figure 4.9b, indicating that while the fits are unbiased, some sensitivity is lost by using the simpler (and ultimately incorrect) model. The remaining **Ser-Exp** mocks, for which the **Ser-Exp** fit is determined by F-test to be most appropriate, are also unbiased in total magnitude and half-light radius. From this test, using the F-test to determine the most appropriate fitted model allows for unbiased measurement of the half-light radius and total magnitude.

Using the **Ser** mocks, the false positive rate (**Ser** mocks classified as needing a **Ser-Exp** fit according to the F-test) for the F-test with a significance level of 0.32 is 5 per cent, suggesting that there is a low level of contamination in a two-component sample selected using the F-test. Using **Ser-Exp** mocks with  $0.2 < B/T < 0.8$  and  $n_{bulge} > 2$ , which are considered true two-component galaxies, the false negative rate (**Ser-Exp** mocks classified as needing only a **Ser** fit according to the F-test) is 34 per cent, missing a substantial fraction of the galaxies with two components. While selection using the F-test is sufficient to remove the measured bias in global fitting parameters and is able to select a relatively pure sample of two-component galaxies, it does not select a complete sample of two-component galaxies.

## 4. SIMULATIONS

---

Clearly caution is necessary when using the F-test to select two-component galaxies from fitting routines. However, the F-test can indicate when the global parameters of a **Ser** model are likely unbiased regardless of the underlying galaxy type.

Following Simard et al. (2011), one can also select the fitted model based on a tiered approach, first performing the F-test on the **Ser** and **deV-Exp** fits. Galaxies for which the **deV-Exp** fit gives a statistically significant improvement are then tested again to determine whether the **Ser-Exp** fit is preferable to the **deV-Exp** fit. The preferred fit (either **Ser**, **deV-Exp**, or **Ser-Exp**) of the **Ser-Exp** mocks is then used to assess the bias in the half-light radius and magnitude. This approach did not significantly alter the results.

Many galaxies exhibit more complex structure than a single- or two-component structure. Even the case of a two-component model often oversimplifies galaxy structure. Bars, rings, central sources, clumpyness, or asymmetry cannot be effectively modeled in the simulations. This limits the process to only determining a lower-bound on the uncertainty in the parameter estimates. However, correcting fits using this lower bound improves the fit of the observed galaxy.

### 4.4 Bias correction using simulations

A simple example of bias correction can be applied following the procedure outlined in Simard et al. (2002). Given the simulated and fitted values of the Sérsic index for the **Ser** model, the output value represents the measured value in real data. The simulated value represents the true underlying value of the galaxy Sérsic index. The average bias and uncertainty in the bias, labeled as *Bias* and  $\Delta\text{Bias}$ , as a function of output Sérsic index can be measured. Additionally, random error in the fits from the width of the bias distribution as a function of Sérsic index, labeled as  $\Delta\text{Random}$ , can be measured. Then the corrected

---

## 4.5 Extending the simulations to the *g*- and *i*-bands

---

Sérsic index and uncertainty on the corrected index is

$$\begin{aligned} n_{corrected} &= n_{fitted} - Bias(n_{fitted}) \\ \Delta n &= \sqrt{\Delta_{galfit}^2 + \Delta Bias^2 + \Delta Random^2} \end{aligned} \quad (4.6)$$

Applying this correction allows us to correct bias as a function of both simulated and fitted Sérsic index for the sample of galaxies used in this work. The results of this process are shown in Figure 4.24.

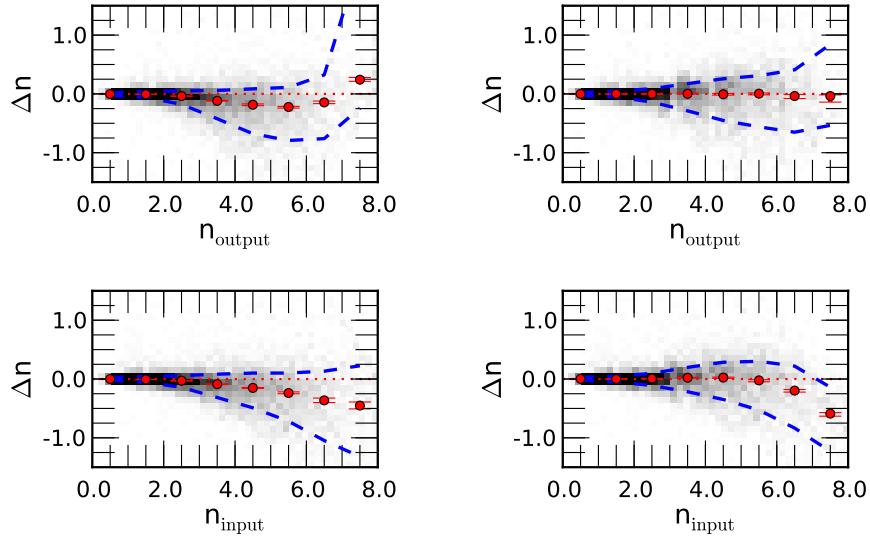
The statistical correction for the bias in the sample in both the simulated and fitted bases works for most values of the Sérsic index. However, there is an under-correction at high simulated Sérsic value. This effect appears to be due to the boundaries of the parameter space that PYMORPH is allowed to search for the best fit model. By restricting PYMORPH to values of  $n < 8$ , galaxies simulated with Sérsic index of 8 will be preferentially underestimated. However, the highest bins of fitted Sérsic index contain many more galaxies with over-estimated Sérsic index. Therefore the net correction will be negative and not appropriate for the highest bins. The error correction at higher bins could be improved by allowing GALFIT to explore larger values of the Sérsic index. Additional corrections may also be considered (i.e., divide in both magnitude and Sérsic index prior to computing the bias correction) depending on the specifics of a given study.

## 4.5 Extending the simulations to the *g*- and *i*-bands

the tests were performed on *r*-band data from SDSS. The performance of the pipeline can change when observing in different bands. This change is primarily dependent on the change in the S/N and resolution between bands (due to the changing brightness of the sky, color of the galaxy, and size relative to the PSF) and on the different galactic structures to which neighboring SDSS filter bands are sensitive. In principle, these effects could be measured from the simulations presented in this paper by adjusting the S/N and background level.

## 4. SIMULATIONS

---



**Figure 4.24:** An example of the bias correction of the Sérsic index of the `Ser` model. The error in the Sérsic index ( $n_{output} - n_{input}$ ) versus output value is presented before (top left panel) and after (top right panel) correction. The same correction is shown in the bottom row versus the simulated value of the Sérsic index. The correction is applied in the  $n_{output}$  basis. This appropriately corrects the bias as a function of  $n_{input}$  except at high  $n$  where the correction fails. The reason for this failure is due to the boundaries of the allowed  $n$  parameter space. Galaxies in the highest bins of output Sérsic index are a combination of poorly fit, low Sérsic index galaxies that are artificially constrained to fall in the high bins, and correctly fit, high Sérsic index galaxies. The result of this mixture is a net negative correction on galaxies with high Sérsic index.

## 4.5 Extending the simulations to the *g*- and *i*-bands

---

Figure 2.6 in Chapter 2 showed that the S/N of the *i*-band is similar to that of the *r*-band fits. Therefore, the performance of the flagging in the *i*-band is expected to be similar to that of the *r*-band. However, the *g*-band fits were found to be lower in S/N with the peak of the distribution approximately a factor of 2 smaller than the peak in the *r*-band.

The simulations of this chapter were primarily based on assuming the S/N of the *r*-band data, so further investigation may be necessary prior to accepting the *g*-band fits. When the appropriate profile (i.e., one vs. two-component) is chosen, magnitude and half-light radius are expected to be accurate at the 5 per cent level (using  $1\sigma$  errorbars). However, these tests presumed the average S/N of an *r*-band image. This is not a concern for the *i*-band, which has both better seeing and brighter galaxies. In the *i*-band the sky is also brighter offsetting any gains in S/N and making the distribution S/N very similar to the *r*-band, if not skewed slightly higher. In the *g*-band, the sky is dimmer, but the galaxies are also dimmer and the PSF wider such that the average S/N is substantially lower. This is expected to affect the measurements in a few ways. First, the fitting parameters, particularly in the case of the `deV-Exp` and `Ser-Exp` fits may be biased or have very large uncertainty. Second, (and partially as a result of the first issue) the choice of bulge/two-component/disk based on the flagging may be biased.

Since the resolution is only moderately different between bands (less than 10 per cent) the effects are expected to be somewhat smaller. However, prior to fitting the *g*-band, the mock galaxies are resimulated with S/N reduced by a factor of 2 to explore the effects of reduced S/N in the *g*-band. Figure 4.25 shows a comparison of the original fiducial simulations of `Ser-Exp` fits and the results of lowering the S/N by a factor of 2, similar to what is observed in the *g*-band images. Figure 4.25 compares the original fiducial estimates from Chapter 4.2.2 (i.e., the scatter and bias in fits at the same S/N as the *r*-band) in the top row to the lower S/N fits (S/N a factor of 2 smaller, similar to *g*-band measurements) in the bottom row. Comparisons are shown for the bulge half-light radius (left-hand column), disk half-light radius (center column), and B/T (right-hand column). The plots show that

## 4. SIMULATIONS

---

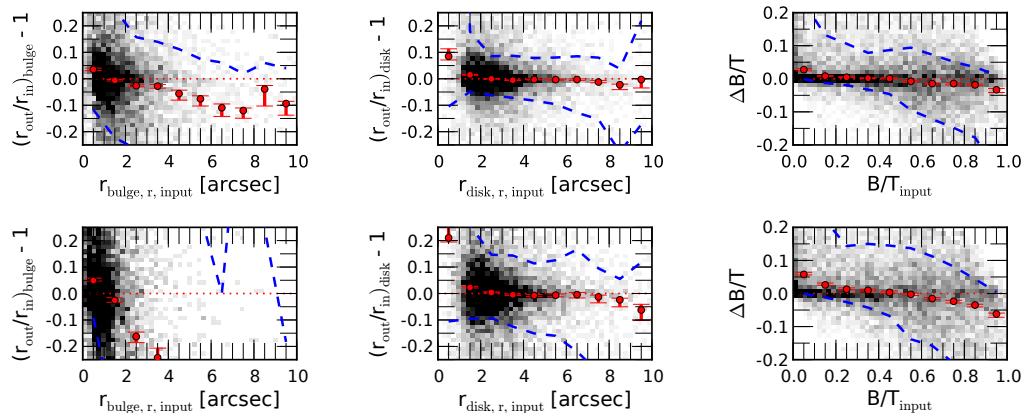
the bulge half-light radius is much more susceptible to underestimation, with the median bulge size underestimated by more than 20 per cent at radii larger than 2 arcsec. The 68 percent scatter in bulge size estimation also increases substantially. Disk size is more stable, showing little or no bias, but the 68 percent scatter increases by roughly a factor of 2. The scatter in B/T also gets broader, although the overall trend remains approximately the same (with slight overestimates of B/T at small B/T values to underestimates of B/T at high B/T values approaching 0.1. However, this trend is partially due to the intrinsic limits on B/T which only allow for underestimates(overestimates) at large(small) B/T because B/T is bounded between 0 and 1.

The behavior of total magnitude, total half-light radius, bulge and disk magnitude and bulge Sérsic index are similar to the effects shown on disk half-light radius. There is growth in the scatter of the estimated parameters by as much as a factor of two, but no additional bias is introduced. The *g*-band estimates of magnitude and half-light radius are expected to be accurate at the 10 per cent level (using  $1-\sigma$  errorbars) compared to 5 per cent accuracy in the *r*- and *i*-bands.

A smaller number of pure bulge galaxies may also be expected as a result of the larger underestimate of bulge size and decrease in the median B/T seen above 0.8. These effects may cause or magnify any shift away from pure bulge galaxies in the *g*-band.

Additional, size differences may alter the results of the simulations somewhat. However the distributions of sizes do not change drastically over the three bands fitted in this work. However, these simulations become an increasingly poor estimate of error in bluer bands where the photometry becomes more sensitive to star forming regions. These regions tend to be clumpier and, therefore, less well represented by a smooth profile. Therefore, *g*-band fits may present more scatter than the *r*- or *i*-band data. These clumpy regions are difficult to model with the smooth models presented here. One could attempt a hybrid approach to generating simulated data whereby one isolates clumpyness in nearby galaxies and use this as a template to add clumpyness to smaller SDSS galaxies. However, the details of

## 4.5 Extending the simulations to the *g*- and *i*-bands



**Figure 4.25:** Comparison of the original fiducial simulations of Ser-Exp fits and the results of lowering the S/N by a factor of 2 (bottom row), similar to what is observed in the *g*-band images. Comparisons are shown for the bulge half-light radius (left-hand column), disk half-light radius (center column), and B/T (right-hand column). The median value is plotted in red with 95 per cent CI errorbars shown. The 68 percent scatter of the data is shown as the dashed contours. The plots show that the bulge half-light radius is much more susceptible to underestimation, with the median bulge size underestimated by more than 20 per cent at radii larger than 2 arcsec. The 68 per cent scatter in bulge size estimation also increases substantially. Other fitted parameters primarily show increases in the 68 per cent scatter of approximately a factor of 2.

#### **4. SIMULATIONS**

---

this process are not addressed here because the galaxies used in this work are at the limit of resolution for detecting the bulges of galaxies. It is unlikely (except for the few largest galaxies) that resolution of star-forming regions are even possible.

# Chapter 5

## Flagging and classification of good and bad fits

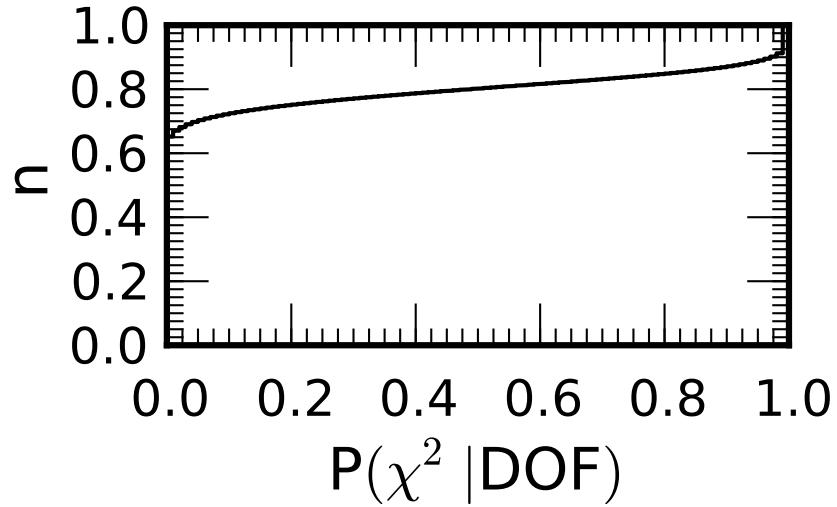
This chapter discusses the flagging method used to separate good fits from poor fits in the catalog. The flagging system provides a straight-forward way of determining which fits to use and an interpretation of the good fits. Chapter 5.1 gives the motivation for the flagging system. Chapter 5.2 briefly describes the visually classified galaxies used to design and tune our flags. Chapter 5.3 describes the automated flags resulting from our visual inspection and gives a breakdown of the flags for the **Ser-Exp** catalogue. A detailed description of the flags are also provided in and example cases for each flag are presented.

### 5.1 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 the catalog 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

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.1:** The cumulative distribution of  $P(\chi^2|\text{DOF})$  for the **Ser** fits. The probabilities are primarily confined to the endpoints of the probability range. Over 70% of galaxies have **Ser** fits with  $P(\chi^2|\text{DOF}) < 0.05$ . Only 20% of galaxies have  $P(\chi^2|\text{DOF}) > 0.5$ . It would be unreasonable to reject such a large proportion of the sample, so we examine other methods of distinguishing good fits from poor fits rather than using the  $\chi^2$  value.

---

## 5.1 The motivation for the flagging system

et al. 2010, for a more in-depth discussion), and the distribution of residual values do not approximate a normal distribution.

A simple cut on  $P(\chi^2|\text{DOF})$  could lead to rejecting many poorly fitted galaxies, but would miss many of these unphysical models. For example, Figure 5.1 shows the cumulative distribution of  $P(\chi^2|\text{DOF})$  for the `Ser` fits. A cut of  $P(\chi^2|\text{DOF}) < 0.05$ , a commonly used level of statistical significance, would lead us to reject 70% of the fitted galaxies as poor fits. Rejecting such a large portion of galaxies would make use of such a catalog useless. 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.

Of course, poor performance is not an acceptable reason to relax constraints on the significance of fitted  $\chi^2$ . Chapter 4 showed that successful fits of galaxies occur well in excess of 30 per cent of galaxies, and the percentage of successful fits was closer to 90 per cent. Based on the simulations of Chapter 4, the  $\chi^2$  is not a reliable predictor of the appropriateness of the fit. Additionally, for non-linear models, the number of degrees of freedom is not well defined (see Andrae et al. (2010) for a more in-depth discussion), and tests of the  $\chi^2$  value do not produce reliable statistics. Some groups have argued that the DOF should be reduced further in proportion to the size of the PSF (e.g. Lackner and Gunn 2012). Although the number of DOF's and the resulting probability is not well defined, minimization of the  $\chi^2$  can still produce the “best” fit. However, the statistical likelihood associated with the fitted  $\chi^2$  is not reliable.

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. Reliable identification of physically meaningful and unphysical fits is also desired. 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 5.6. These flags are used to determine the

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

reliability of the various fits and the individual subcomponents as well as mark poorly fitted galaxies.

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

1. bulge-like galaxies,
2. disk-like galaxies,
3. two-component galaxies, and
4. unknown type/poorly fitted/failed fits.

In the **Ser** and **deV** fits, these conditions simplify somewhat since there are no subcomponents. Rather than focus on these broad (and perhaps vaguely defined) categories for classification, a series of quantitative indicators based on the fitted parameters is defined and assessed 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. The approach in this work is similar to the former groups rather than the latter. Galaxies are separated using a categorical description of the fits. The results of the flagging are compared to the results of S11, LG12, and Mendel et al. (2014) in Chapters 7 and 8.

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 neighbors 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  $\chi^2_\nu$ ) is easy to do, producing meaningless results for the components.

Chapter 4 also tested fitting in SDSS conditions of S/N, platescale, and seeing using simulated data. The test presented in Chapter 4 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 **Ser** 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 components) is chosen, magnitude and half-light radius are expected to be accurate at the 5 per cent level (using  $1\sigma$  errorbars). The flagging procedure presented in Section 5 is intended (among other 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 Chapter 4 for more information on the simulations and results).

## 5.2 The visual classification

No catalog 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). A small training set of visually classified fits was used to design the flagging criteria. For purposes of training and validation, 1 000 galaxies were randomly selected from the fitted galaxies for visual classification. The visual classification system has several categories including general fit problems and image characteristics. The sample was classified by examining the fitted parameters, the 2D image, 2D model, 2D residual, the 1D profile and, the 1D residual. These panels are generated on the fly and available to

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

the reader upon request from the catalog website:

[http://shalaowai.physics.upenn.edu/~ameert/fit\\_catalog/view\\_galaxies/](http://shalaowai.physics.upenn.edu/~ameert/fit_catalog/view_galaxies/).

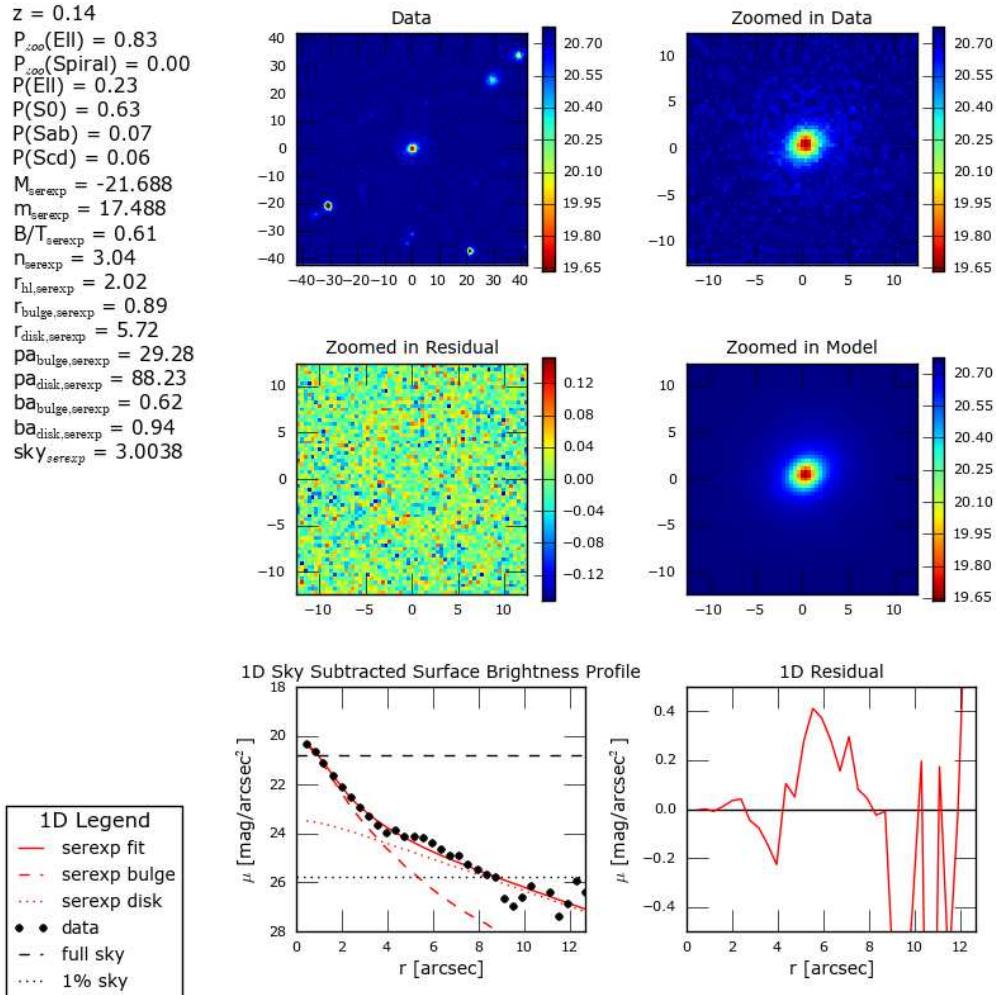
Figure 5.2 shows an example of the plots used to manually inspect the fits. Visual classification makes it possible to enumerate the types of problems commonly found in the fits (i.e., effects of contamination or poorly defined bulges) and understand the relative importance of each problem. The visual sample was selected randomly from the catalog 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 catalog. 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 5.3. The test sample was set aside and used to test the reliability of the automated flags once they were defined.

Classifications were carried out for each galaxy. After viewing each fit, we assign a series of flags intended to give useful information about the appropriateness of the total fit and the physical or unphysical nature of the components. Each fitted model (the **deV**, **Ser**, **deV-Exp**, and **Ser-Exp**) is assigned an independent set of flags. However, our focus here and throughout the text is on classifying the **Ser-Exp** fits.

The visual classification system has several categories including general fit problems and image characteristics. Table 5.1 shows the description of these visual flags. The visual flags contain a component focusing on the fitted model and a component focusing on the photometric qualities of the galaxy and its neighborhood. For instance, we expect that strong bar components may affect the fit of the **Ser** component to the galaxy. This would increase the ellipticity of the **Ser** component and potentially alter its radius. The ability to distinguish these bar cases would be useful. The effects of bar presence are discussed further in Chapter 8.4.

## 5.2 The visual classification



**Figure 5.2:** An example panel used during visual classification showing the original image, fitted galaxy, 2-d residual, 1-d galaxy profile, and 1-d residual. Many fitting parameters are displayed on the left side of the image for comparison. During visual classification, checks include looking for **Ser** components that dominate at large radii, mis-aligned components, large residuals, and other indications of bad fits.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

Flag bit	Fit Flag	Flag Description
0	Bad Central Fit	The fit residuals are large in the range $r < 2r_{hl}$
1	Bad Middle Fit	The fit residuals are large in the range $2r_{hl} \leq r \leq 4r_{hl}$
2	Bad Outer Fit	The fit residuals are large in the range $r > 4r_{hl}$
3	Ser Fitting Outer Part of the Galaxy	A bad Ser component that dominates the fit at large radii is observed
4	Exp fitting Inner Part of the Galaxy	A bad Exp component that dominates the fit at the central radii is observed
5	Centering Problems	The center of the model is offset from the data, usually due to neighbor contamination
6	Sky Subtraction Problems	The sky-brightness is incorrect and affects the total magnitude of the fitted model
7	Ser/Exp Position Angle Difference	The position angles of the Ser and Exp components differ substantially
8	High Ellipticity Exp Component	The Exp component has high ellipticity
9	High Ellipticity Ser Component	The Ser component has high ellipticity
10	Large ExpRadius/ExpFitting Sky	Large radius of Exp component indicating fitting sky
11	Bad Fit of Sub-Dominant Second Component	The dimmer component is incorrectly fit due to contamination or some other issue

Flag bit	Image Flags
0	contamination from nearby sources
1	Background problems
2	crowded field
3	Evidence of past or present interaction
4	Asymmetry
5	Low Surface Brightness components
6	Bar component evident

**Table 5.1:** The visual classification system used to build a training and validation set for the machine classifier. The choice of classification was made on the basis of the model 2d image, 2d residual, 1d profile and 1d residual. These data were compiled in a single panel for each model. An example of such a panel is shown in Figure 5.2. The first section of the table shows the possible visual flags available to the human classifier and a description of the flag. The second section of the table includes additional observations available to the human during classification. Although these additional observations do not necessarily indicate a poor fit, they are potentially correlated with common fitting problems (e.g. a nearby neighbor contaminating the fit can cause the fitted galaxy to align with the neighbor).

### **5.3 The automated flagging system**

---

The incidence of flags in the visual classification of the **Ser-Exp** fit is shown in Table 5.2. Only the **Ser-Exp** fit is shown since it is the focus of the flagging and the least well-constrained of the fits. A flagging procedure that is effective on the **Ser-Exp** fits should do equally well (or perhaps even better) on the simpler models.

The visual classifications are used to tune the automated flags used in the main text. The mapping used to compare automated flags and visual flags is given in Table 5.3. The mapping shows the simple automated flags in the first column and the visual analog in the second column. Chapter 5.3 gives a full description of the simple automated flags. We rely upon the simple automated flags to classify galaxies in the catalog into the refined categories presented in the main text.

## **5.3 The automated flagging system**

The automated flags consist of two tiers. The first tier consist of simple, quantitative flags that attempt to replicate many of the problems observed during manual examination of the fits. The second tier consists of the final flags used to separate galaxies into pure bulge/pure disk/two-component galaxies. The second tier of flags is more interpretive, attempting to explain each fit in a more useful way.

The tier-1 automated flags are explained in Table 5.4. These flags are chosen to approximately map to the visual classification described in Chapter 5.2 and contain definitive numerical cuts. Each individual flag has a tunable cutoff that is optimized using the training set of manually classified galaxies. The efficacy of the flag is evaluated using the test set of galaxies (which are also visually classified). A list of the automated flags and the percentage of the test set flagged are presented in Table 5.5. The visually classified sample is randomly divided into a training set ( $\approx 50\%$  of the data) used to train the classifier and the test set (the remaining  $\approx 50\%$  of the data).

These flags are binary classifiers. Each classifier is evaluated using the Matthews Correlation Coefficient (hereafter MCC; see Baldi et al. 2000, for a discussion of the MCC). The

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

Flag bit	Fit Flag	Flag Incidence out of 1000 galaxies
0	Bad Central Fit	45
1	Bad Middle Fit	47
2	Bad Outer Fit	35
3	Ser Fitting Outer Part of the Galaxy	115
4	Exp Fitting Inner Part of the Galaxy	46
5	Centering Problems	13
6	Sky Subtraction Problems	12
7	Ser/Exp Position Angle Difference	61
8	High Ellipticity ExpComponent	108
9	High Ellipticity SerComponent	141
10	Large ExpRadius/ExpFitting Sky	22
11	Bad Fit of Sub-Dominant Second Component	212

Flag bit	Image Flag	Flag Incidence out of 1000 galaxies
0	Contamination from nearby source	146
1	Background Problems	7
2	Crowded field	110
3	Evidence of past or present interaction	46
4	Asymmetry	103
5	Low Surface Brightness	1
6	Bar Component Evident	47

**Table 5.2:** The results of visual classification for the Ser-Exp models. Many flags are sparsely populated. While this is encouraging with respect to the fitting (the low incidence of flags suggests that the fitting is performing relatively well), it is problematic for training and tuning purposes. After dividing the sample into a training sample and test sample, the incidence of flags is further reduced by a factor of two. This can lead to tuning problems as we don't have enough of a signal for the automated flags to rely on.

### 5.3 The automated flagging system

---

Automated Flag	Equivalent Visual Flag
<i>Centering</i>	<i>Centering Problems</i>
<i>Ser is Sky</i>	<i>Sky Subtraction Problems</i> AND NOT <i>large Exp radius/Exp fitting sky</i> AND $b/a_{bulge} > 0.4$ AND $r_{bulge,cir}/r_{sex} > 3.0$
<i>Exp is Sky</i>	<i>large Exp radius/Exp fitting sky</i> AND $b/a_{disk} > 0.4$
<i>No Ser Likely</i>	NONE
<i>Ser Fitting Outer</i>	<i>Ser fitting outer part of the galaxy</i>
<i>High e Ser</i>	<i>high ellipticity Sercomponent</i>
<i>Ser is Disk</i>	<i>Ser fitting outer part of the galaxy</i> AND <i>Exp fitting inner part of galaxy</i> AND $n_{bulge} < 2$
<i>Ser PA Problem</i>	<i>Ser/Exp position angle difference</i> AND $B/T < 0.5$
<i>Ser Contaminated</i>	NOT <i>large Expradius/Expfitting sky</i> & $b/a_{bulge} < 0.4$ & $r_{bulge,cir}/r_{sex} > 3.0$
<i>Low Sersic index Ser</i>	NONE
<i>Ser Dominates Always</i>	NONE
<i>Exp Fitting Inner</i>	<i>Exp fitting inner part of galaxy</i>
<i>Exp Contaminated</i>	<i>large Expradius/Expfitting sky</i> & $b/a_{disk} < 0.4$ & $B/T > 0.3$
<i>High e Exp</i>	<i>high ellipticity Expcomponent</i>
<i>No Exp Likely</i>	NONE
<i>Exp PA Problem</i>	<i>Ser/Exp position angle difference</i> AND $B/T > 0.5$
<i>Exp Dominates Always</i>	NONE
<i>Parallel Components</i>	NONE
<i>Flip Components</i>	NONE
<i>Galfit Failure</i>	NONE

**Table 5.3:** The mapping of the visual classification system to the automated flags. Most flags have a visual analog, the flags that do not have a visual analog can be divided into two categories. The first category of flags without visual analogs contains flags that can be objectively defined and don't require tuning. This category includes *Ser Dominates Always*, *Exp Dominates Always*, *Low Sersic index Ser*, and *Galfit Failure*. The remainder, including *Flip Components*, *Parallel Components*, *No Exp Likely*, and *No Ser Likely* are defined according following the discussion in Chapter 5.3.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

MCC is useful in binary classification problems because it takes into account all terms in the confusion matrix. The MCC considers the relative proportions of True Negative cases in addition to all other possibilities. Other common performance metrics like precision, recall, and the f-score do not consider the true negative rate and are not optimal for binary classification. The MCC returns a value between -1 and 1 for each test; 1 being a perfect predictor, 0 being random, and -1 corresponding to predictions that are exactly opposite of the true value. The MCC is also proportional to the  $\chi^2$  with 3 degrees of freedom, so the significance of the choice of flags can be based on the  $\chi^2$  probability of identification vs. the null hypothesis that the flags are not accurate predictors of fit problems identified by manual classification.

Many combinations of flags can exist in the sample. Interpretation of these flag combinations is necessary for the flagging to be useful during fitting. The catalog is separated into useful groups in Table 5.6, and the table shows the relative percentages of the catalog for each group. These are the categories that are referred to as “tier-2” flags and are the flags most useful for analysis purposes.

To define the final flags, combinations of “tier-1” flags are manually explored and assigned to descriptive categories.

### 5.3.1 Finding galaxies with no bulge or no disk

The decision on where to cut bulges and disks, setting the line between one- and two-component galaxies is somewhat arbitrary. There are a number of cases where the galaxy has only a single component. The most definitive of these is when the second component is too dim relative to the brighter component.

According to the simulations shown in previous sections, scatter in the fitted values of the bulge and disk sub-components increases substantially when the sub-component is dimmer than 19 magnitudes. Large disparities in the magnitudes of the components make the smaller component less important and more likely to be fitting noise or substructure

### 5.3 The automated flagging system

---

Flags	Flagging Criteria
<b>Centering</b>	(galaxy centroid-sdss galaxy centroid) $>0.7 * r_{sex}$
<b>Parallel Components</b>	$RMS(B/T - \mu(B/T)) : 0 < r < r(0.9L_{tot}) \& n_{bulge} < 2.0$
<b>No Ser Likely</b>	$1000 * (0.2 \cdot B/T)^3 + (m_{bulge} - 19) > 0.5$
<b>No Exp Likely</b>	$1000 * (B/T - 0.8)^3 + (m_{disk} - 19) > 0.5$
<b>Ser Component Flags</b>	
<i>Ser Contaminated</i>	$r_{bulge,cir}/r_{sex} > 4.0 \text{ AND } b/a_{bulge} \leq 0.6$
<i>Ser is Sky</i>	$r_{bulge,cir}/r_{sex} > 4.0 \text{ AND } b/a_{bulge} > 0.6$
<i>High e Ser</i>	$b/a_{bulge} < 0.4 \text{ AND } B/T < 0.25$
<i>Ser PA Problem</i>	$\Delta\phi > 45 \text{ AND } B/T < 0.5 \text{ AND } b/a_{bulge} < 0.75 \text{ AND } b/a_{disk} < 0.75$
<i>Ser Fitting Outer</i>	$B/T < 0.7 \text{ AND } B/T(r(0.9L_{tot})) > 0.5$
<i>Ser is Disk</i>	$r_b/r_d > 1.5 \text{ AND } n_{bulge} < 2$
<i>Ser Dominates Always</i>	$B/T(r) > 0.5 : 0 < r < r(0.95L_{tot})$
<i>Low Sersic index Ser</i>	$r_b/r_d \leq 1.5 \text{ AND } n_{bulge} < 2$
<b>Exp Component Flags</b>	
<i>Exp is Sky</i>	$r_{disk,cir}/r_{sex} > 3.0 \text{ AND } b/a_{disk} > 0.6$
<i>Exp Contaminated</i>	$r_{disk,cir}/r_{sex} > 3.0 \text{ AND } b/a_{disk} \leq 0.6$
<i>High e Exp</i>	$b/a_{disk} < 0.4 \text{ AND } B/T > 0.75$
<i>Exp PA Problem</i>	$\Delta\phi > 45 \text{ AND } B/T > 0.5 \text{ AND } b/a_{bulge} < 0.75 \text{ AND } b/a_{disk} < 0.75$
<i>Exp Fitting Inner</i>	$B/T(0) < 0.5 \text{ AND } B/T(r: r < 2.7r_{hl}) > 0.5$
<i>Exp Dominates Always</i>	$B/T(r) < 0.5 : 0 < r < r(0.95L_{tot})$
<b>Galfit Failure</b>	Galfit fails to converge and crashes

**Table 5.4:** The selection criteria used to define the automated flags. The fist column gives the flag name, the second column gives the numerical criteria corresponding to that flag.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

Flags	% of Test Sample	% of Total Catalog
<b>No Flags</b>	21.88	21.84
<i>Centering</i>	1.04	1.54
<b>Parallel Components</b>	5.00	4.85
<b>No Ser Likely</b>	16.25	17.43
<b>No Exp Likely</b>	8.13	7.84
<b>Ser Component Flags</b>		
<i>Ser Contaminated</i>	1.87	1.50
<i>Ser is Sky</i>	1.25	0.78
<i>High e Ser</i>	4.37	4.91
<i>Ser PA Problem</i>	5.83	5.40
<i>Ser Fitting Outer</i>	13.54	10.07
<i>Ser is Disk</i>	2.50	2.14
<i>Ser Dominates Always</i>	13.33	14.62
<i>Low Sersic index Ser</i>	20.00	19.50
<b>Exp Component Flags</b>		
<i>Exp is Sky</i>	0.42	0.48
<i>Exp Contaminated</i>	1.46	1.86
<i>High e Exp</i>	3.12	3.09
<i>Exp PA Problem</i>	5.21	4.45
<i>Exp Fitting Inner</i>	4.79	3.56
<i>Exp Dominates Always</i>	5.83	5.73
<b>Galfit Failure</b>		

**Table 5.5:** The automated flags applied to the test sample and to the full catalog. Quantities are shown as a percentage of the respective sample. Columns 1, 2, and 3 are the flag name, the percentage of galaxies possessing the given flag from the test sample, and the percentage of the full sample, respectively. Multiple flags can be assigned to a single galaxy, so percentages do not add to 100%. However, the galaxies possessing “no flag” cannot appear in any other flag categories. Galaxies flagged as having *No Ser* (*No Exp*) are excluded from the other *Ser*(*Exp*)-related flags. These flags are less useful for analysis and difficult to interpret.

### 5.3 The automated flagging system

---

that is not of interest. For example, a B/T value of 0.05 is very small and indicates that, while there may be a small component at the center of the target galaxy, it is not likely to be a bulge. This may be a small, point source near the center, or the result of the disk not perfectly fitting the galaxy even though it is correct across the image. Similar problems are present in disks.

Since the extremes in magnitude and B/T are strong indicators of a galaxy best described by a single component, a cubic polynomial cut line in B/T vs magnitude space allows identification of these differences. The cuts are

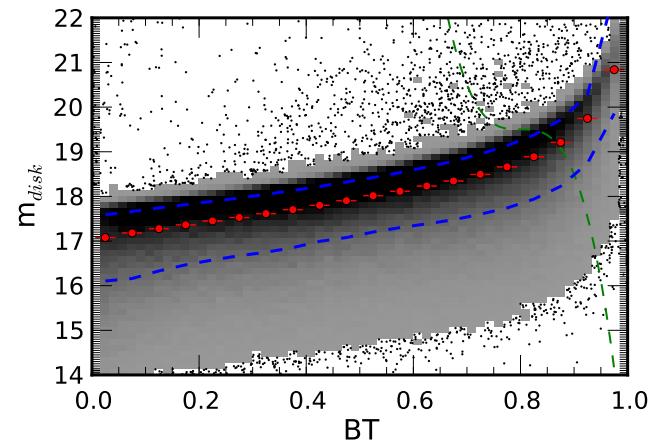
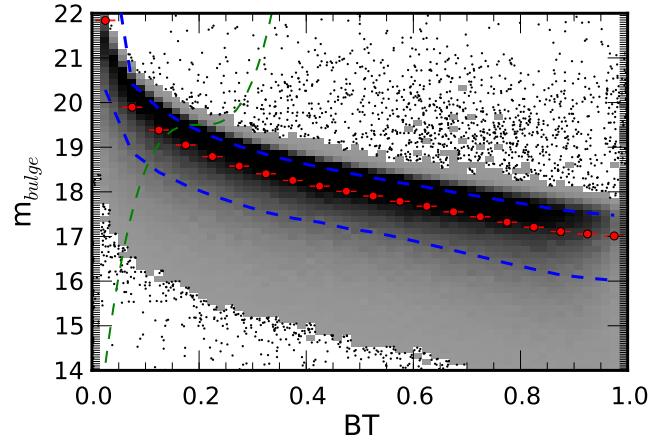
$$\begin{aligned} 1000 * (0.2 - B/T)^3 + (m_{bulge} - 19) &> 0.5 \Rightarrow \text{No Ser component} \\ 1000 * (B/T - 0.8)^3 + (m_{disk} - 19) &> 0.5 \Rightarrow \text{No Exp component.} \end{aligned} \tag{5.1}$$

Cases of “No Exp component” are further separated into Bulge and Disk galaxies based on the Sérsic index ( $n \geq 2 \Rightarrow$ Bulge galaxy,  $n < 2 \Rightarrow$ Disk galaxy). These cuts remove any galaxy with the secondary component dimmer than 19.5 mag if it represents less than 20% of the total light of the galaxy and cuts at progressively brighter magnitudes. Using this cut, secondary components that represent less than 5% of the total light are considered to be a single component regardless of the component magnitude. This is consistent with the visual classification that was previously performed. The cuts capture 17% of bulges and 7.8% of disks in our sample.

Figure 5.3 shows the entire sample of galaxies in the m vs B/T plane with the cut drawn as a green dashed line for each of the bulge and disk cuts. The choice of a cubic term in B/T, while arbitrary, gives the desired behavior. As the secondary component becomes brighter, it has higher signal-to-noise and should then be more reliable, even when it represents a smaller component of the total light. Likewise, if the component is dimmer, then it should represent much more of the total light for it to be considered trustworthy.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.3:** The no bulge cut and no disk cut drawn as a green dashed line superimposed over the total distribution of galaxies from the catalog. Note that the cut becomes more generous in B/T for brighter magnitudes and more restrictive for dimmer magnitudes.

## **5.3 The automated flagging system**

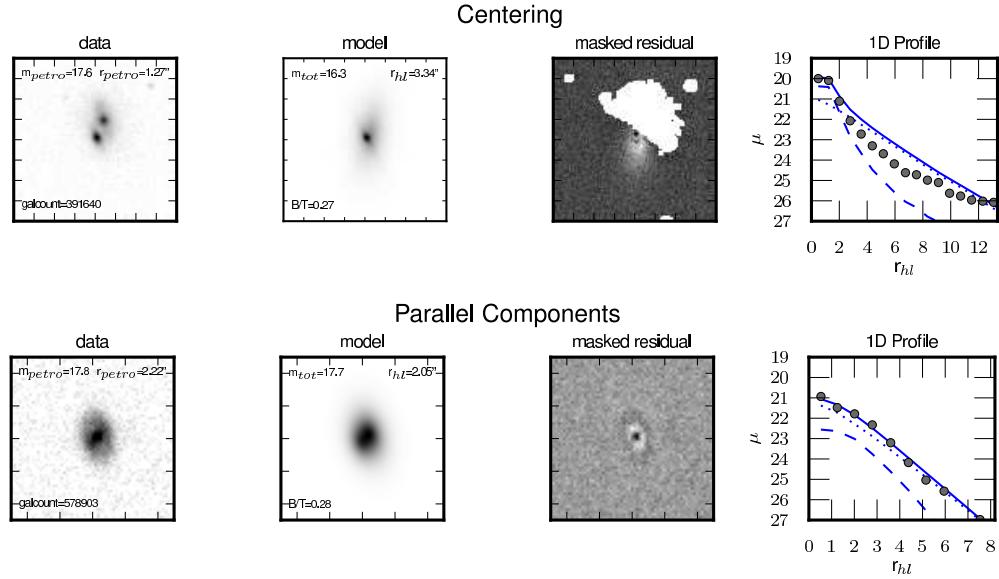
---

### **5.3.2 Examples of galaxies possessing automated flags**

This section shows examples of galaxies flagged with the various automated flags. Figures 5.4-5.8 show different flags.

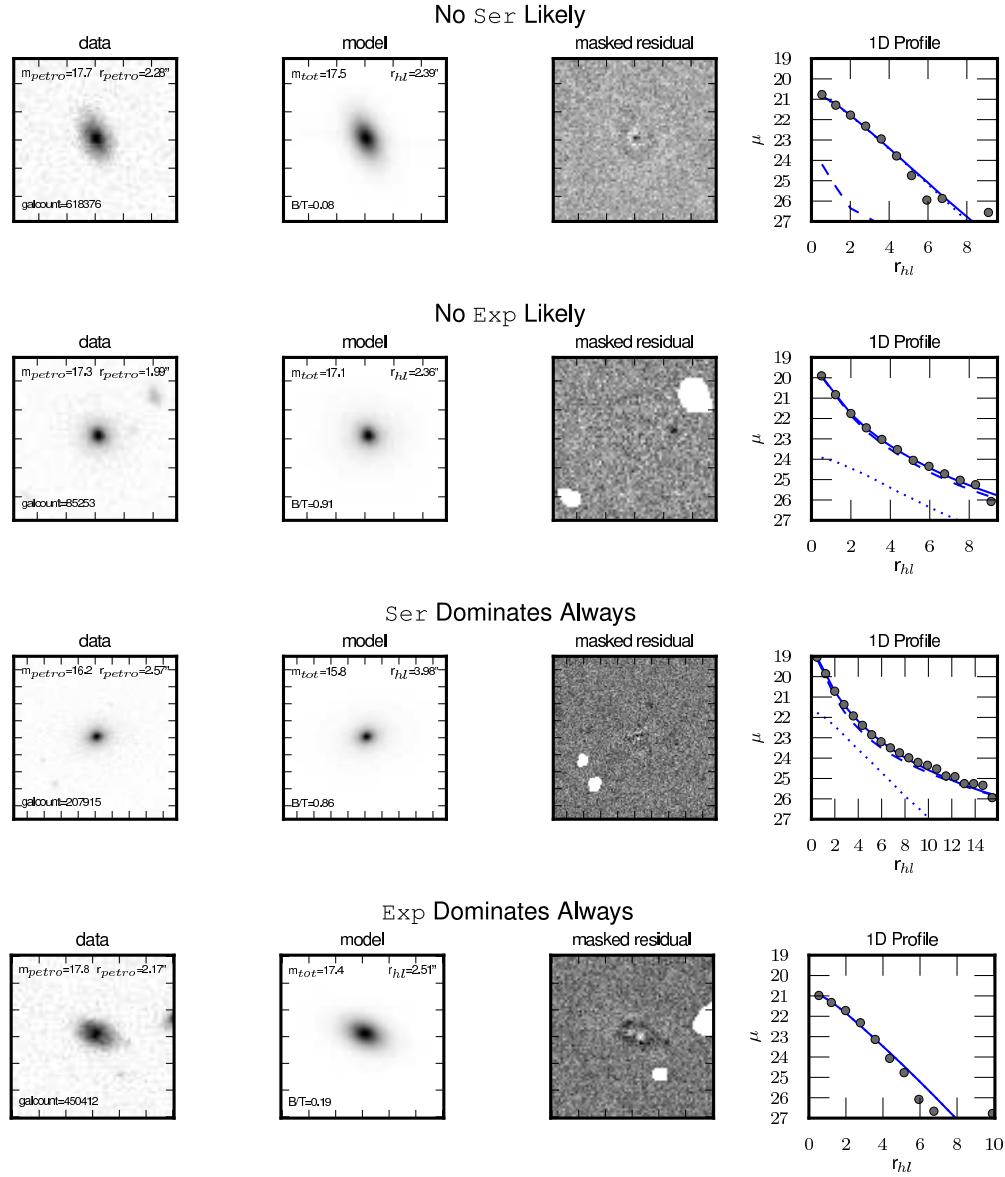
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.4:** Example Ser-Exp fits possessing tier-1 automated flags. “Centering” and “Parallel Components” are shown. “Centering” suggests a poor fit because the fitted object is not at the location of the target galaxy. “Parallel Components” suggests that both the **Ser** and **Exp** components behave similarly and should be considered a single component. The flag is denoted above each row. Each row shows the data image, model image, and masked residual image. The data image has the Petrosian magnitude and radius as well as the unique id for our catalog. The model image displays the fitted magnitude and radius for comparison as well as the fitted B/T ratio. On the far right e show the one-dimensional surface brightness profile as a function of the half-light radius. The dots represent the data while the blue solid, dashed, and dotted lines represent the total, **Ser**, and **Exp** component surface brightness from the fitted model, respectively.

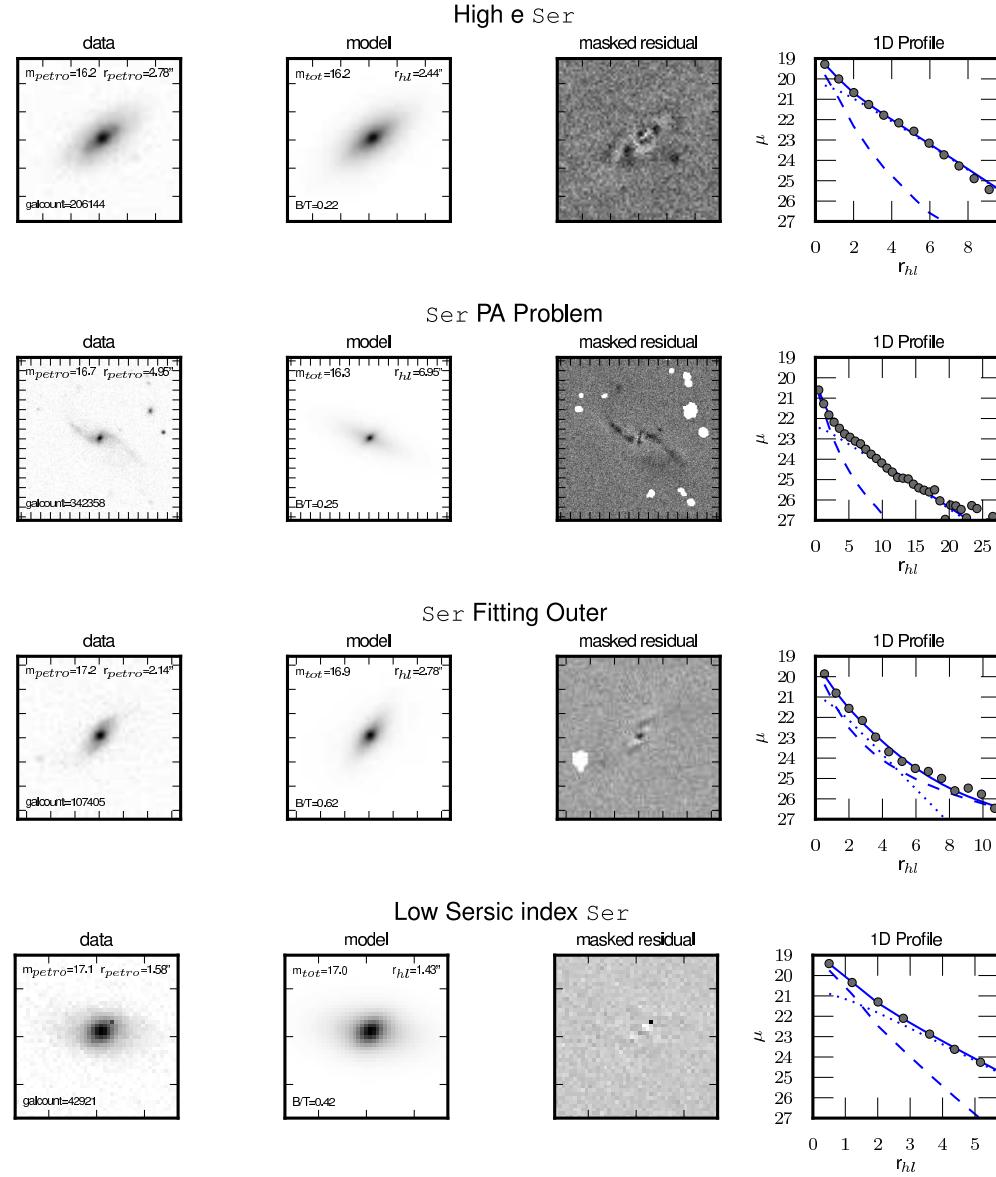
### 5.3 The automated flagging system



**Figure 5.5:** Example Ser-Exp fits possessing tier-1 automated flags that suggest the presence of only one component. The plots are shown in the same format as Figure 5.4.

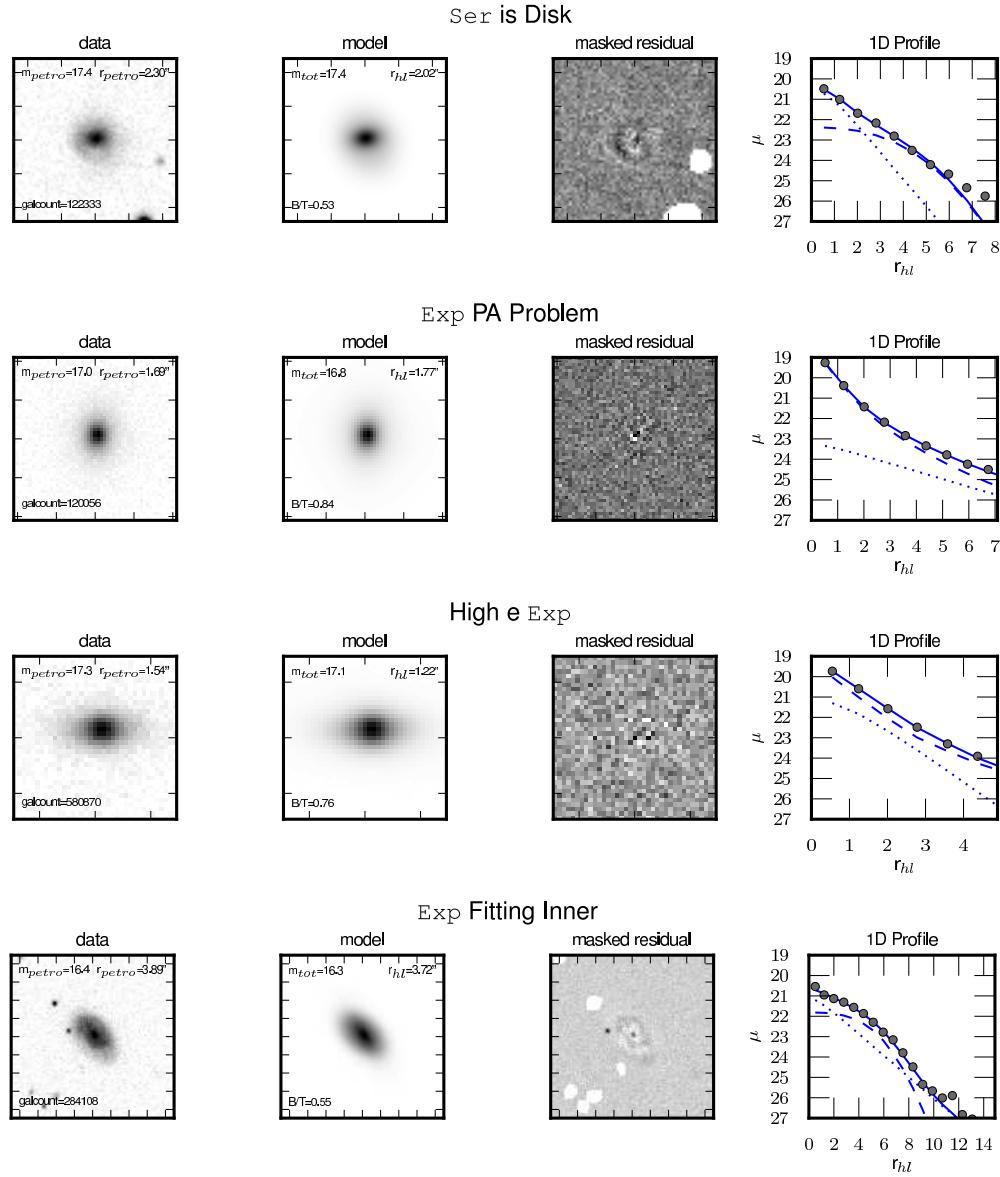
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.6:** Example Ser-Exp fits possessing tier-1 automated flags that suggest problems with the **Ser** component. The plots are shown in the same format as Figure 5.4.

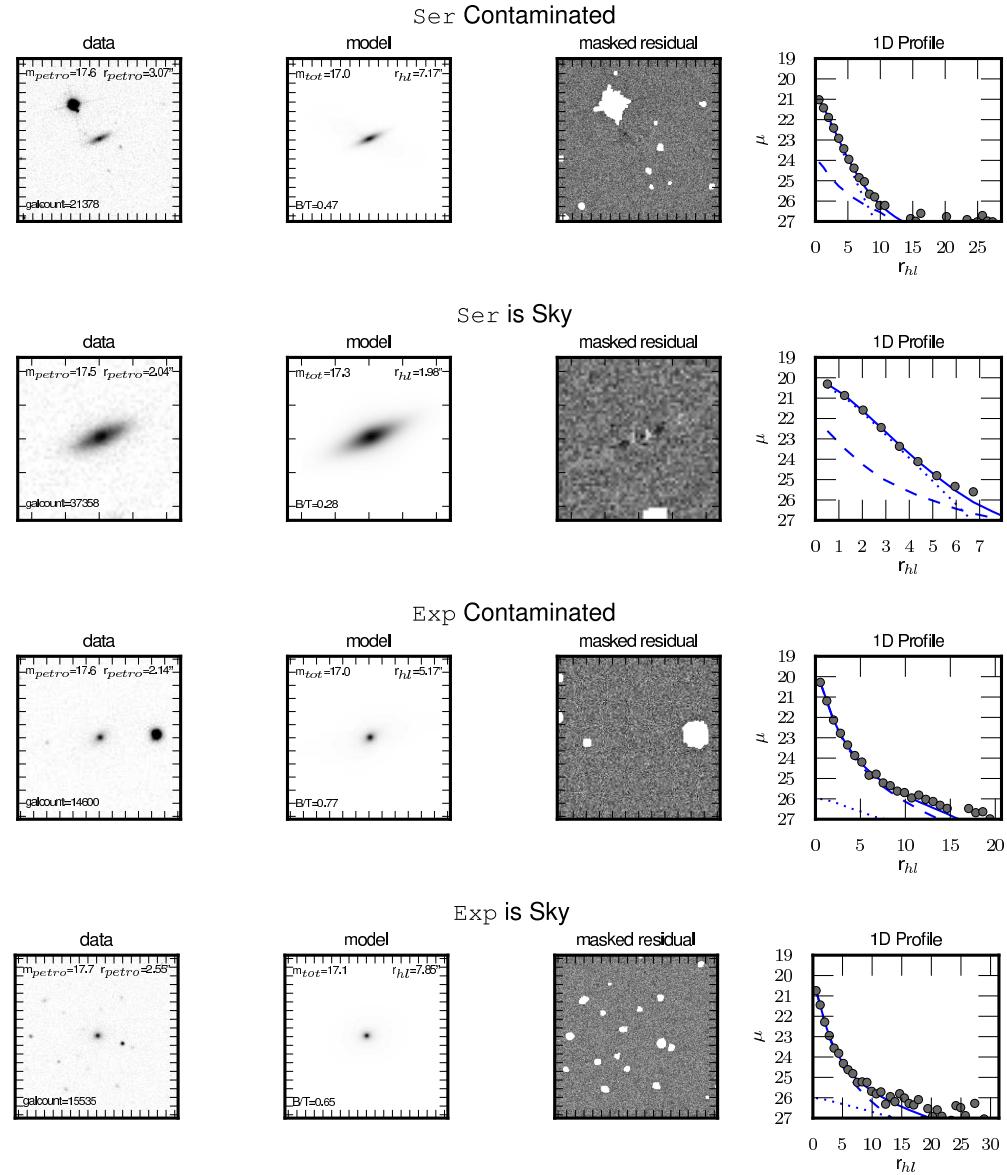
### 5.3 The automated flagging system



**Figure 5.7:** Example Ser-Exp fits possessing tier-1 automated flags that suggest problems with the Exp component. The plots are shown in the same format as Figure 5.4.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.8:** Example Ser-Exp fits possessing tier-1 automated flags that suggest large inaccuracies in fitting. The plots are shown in the same format as Figure 5.4.

## 5.4 Final Flags

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).

Table 5.6 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 Chapter 5.4.2. Two example galaxies are shown for each flagging category. For each fit, the 2D data, fitted model, and residual as well as the 1D radial data profile, bulge and disc component profiles, and the total fitted profile are shown.

A tiered structure is used 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, the galaxies are separated into single-component galaxies (flag bits 1 and 4), two-component galaxies (flag bit 10), and problematic two-component galaxies (flag bit 14) that have fit parameters that are difficult to interpret.

The categories presented in Table 5.6 are the recommended flags used to select samples from the catalog. In particular, major categories of the flagging are as follows.

- **Good Two-Component Fits (bit 10 set)**—

These are the galaxies found to have two fitted components with intermediate bulge-to-total light ratio (B/T) and reasonably well behaved subcomponents. Both the subcomponents and the total magnitude and radius can be used in any analysis. Bits 11-13, only one of which will be set, give a more specific description of the galaxy.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

Flag bit	Description	Per cent Ser-Exp	Example galaxy
-	Good total and component magnitudes and sizes	39.055	
10	Two-component galaxies	39.055	Figure 5.9, 5.10, 5.11
11	No flags	18.095	Figure 5.9
12	Good Ser, good Exp (some flags)	19.417	Figure 5.10
13	Flip components	1.543	Figure 5.11
-	Good total magnitudes and sizes only	54.945	
1	Bulge galaxies	18.964	Figure 5.12, 5.13
2	No Exp component, $n_{Ser} > 2$	7.074	Figure 5.12
3	Ser dominates always	11.889	Figure 5.13
4	Disk galaxies	25.146	Figure 5.14, 5.15, 5.16, 5.17, 5.18
5	No Ser component	16.876	Figure 5.14
6	No Exp, $n_{Ser} < 2$ , Flip Components	0.551	Figure 5.15
7	Ser dominates always, $n_{Ser} < 2$	0.103	Figure 5.16
8	Exp dominates always	2.872	Figure 5.17
9	Parallel components	4.745	Figure 5.18
14	Problematic two-component galaxies	10.835	Figure 5.19, 5.20, 5.21, 5.22
15	Ser outer only	7.504	Figure 5.19
16	Exp inner only	0.425	Figure 5.20
17	Good Ser, bad Exp, $B/T \geq 0.5$	0.017	Figure 5.21
18	Bad Ser, good Exp, $B/T < 0.5$	0.625	Figure 5.22
19	Bulge is point	2.237	Figure 5.23
20	Bad total magnitudes and sizes	6.000	

**Table 5.6:** 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 sub-components, consideration should be given as to exactly which groups of fits should be included.

---

## 5.4 Final Flags

- **Good Bulge Fits (bit 1 set)**–

These are the galaxies found 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. The `Ser` fit is recommended 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.

- **Good Disk Fits (bit 4 set)**–

These are the galaxies found 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. Using the `Ser` fit is recommended for 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.

- **Problematic Two-Component Fits (bit 14 set)**–

These are the galaxies with two fitted components with intermediate B/T, but at least one sub-component has strange behavior. 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.

- **Bad Fits (bit 20 set)**–

These are the galaxies found 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 the same flagging structure is used for all fits to ensure consistency and ease of use.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

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. The `deV` fits are reported 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.

### 5.4.1 Detailed descriptions of the final flags

This chapter provides a more detailed descriptions of the flags used in this catalog. Table 5.7 gives a short description of the fitting parameters used to set the value of each flag. The majority of the flags use the B/T, Sérsic index, and radial light profiles of the galaxy to set the flags. The flags also consider the axis ratio and the difference in position angle of the bulge and disc because small axis ratios combined with large differences in position angle of the two components is an unphysical fit for most cases.  $B/T(r)$  in flags 3, 7, 8, etc. refers to the radial B/T profile. For the radial profile, the local B/T ratio (i.e., the ratio of light contributed by the bulge at the annulus of width 1 arcsec centred on radius  $r$  to the light contributed by the disc) is plotted.  $r_{sextractor}$  is the half-light radius reported by SExtractor during fitting. This radius is used to set the scale of the galaxy during fitting and to test whether components have become unphysically large after fitting.

Flag bit 9 is used to determine whether the bulge and disc are fitting the same component of the galaxy. This is most common in very late-type galaxies (i.e.,  $T$ -types greater than 4) that have no visible bulge. The flag criteria reference  $r(0.9L_{tot})$ , which is the radius enclosing 90 per cent of the total light of the galaxy. The flag also measures the quantity  $RMS(B/T - \mu(B/T)) : 0 < r < r(0.9L_{tot})$ . The quantity  $RMS(B/T - \mu(B/T))$  is the standard deviation of B/T about the mean B/T value. The average is taken over the radii interior to the radius enclosing 90 per cent of the light. If the standard deviation is less than 0.1, then the B/T ratio is approximately the same at all radii and the components are parallel. Two parallel components, each with Sérsic index  $< 2$ , are indistinguishable from one brighter exponential component, so these galaxies are flagged as discs and the

## 5.4 Final Flags

---

Flag bit	Analysis flag	Flag criteria
<b>Good total and component magnitudes and sizes</b>		
10	<b>Two-component galaxies</b>	
11	No flags	No flags are present
12	Good Ser, Good Exp (some flags)	Some minor automated flags
13	Flip components	Exp component fitting the inner and Ser component fitting the outer part of the profile
<b>Good total magnitudes and sizes only</b>		
1	<b>Bulge galaxies</b>	
2	No Exp component, $n_{\text{Ser}} > 2$	$1000(B/T - 0.8)^3 + (m_{\text{bulge}} - 19) > 0.5 \text{ AND } n_{\text{Ser}} \geq 2$
3	Ser dominates always	$B/T(r) \geq 0.5 \text{ for all } r \text{ AND } n_{\text{Ser}} \geq 2$
4	<b>Disc galaxies</b>	
5	No Ser component	$1000(0.2 - B/T)^3 + (m_{\text{bulge}} - 19) > 0.5$
6	No Exp, $n_{\text{ser}} < 2$ , Flip Components	$1000(B/T - 0.8)^3 + (m_{\text{disc}} - 19) > 0.5 \text{ AND } n_{\text{Ser}} < 2$
7	Ser dominates always, $n_{\text{Ser}} < 2$	$B/T(r) \geq 0.5 \text{ for all } r \text{ AND } n_{\text{Ser}} < 2$
8	Exp dominates always	$B/T(r) < 0.5 \text{ for all } r$
9	Parallel components	$RMS [(B/T - \mu(B/T)) : 0 < r < r(0.9L_{\text{tot}})] < 0.1 \text{ AND } n_{\text{bulge}} < 2.0$
14	<b>Problematic two-component galaxies</b>	
15	Ser outer only	$B/T < 0.7 \text{ AND } B/T(r(0.9L_{\text{tot}})) > 0.5 \text{ AND } n_{\text{bulge}} \geq 2.0 \text{ AND } B/T(0) > 0.5$
16	Exp inner only	$B/T(0) < 0.5 \text{ AND } B/T(r: r < 2.7r_{hl}) > 0.5 \text{ AND } n_{\text{bulge}} \geq 2.0 \text{ AND NOT Ser Outer Only}$
17	Good Ser, bad Exp, $B/T \geq 0.5$	$B/T > 0.75 \text{ AND } \Delta\phi > 45 \text{ AND } b/a_{\text{bulge}} < 0.75 \text{ AND } b/a_{\text{disc}} < 0.4$
18	Bad Ser, good Exp, $B/T < 0.5$	$B/T < 0.25 \text{ AND } \Delta\phi > 45 \text{ AND } b/a_{\text{bulge}} < 0.4 \text{ AND } b/a_{\text{disc}} < 0.75$
19	Bulge is point	Circularized bulge radius $< 0.188 \text{ arcsec}$
20	<b>Bad total magnitudes and sizes</b>	
21	Centering problems	$(\text{galaxy centroid-SDSS galaxy centroid}) > 0.7 \times r_{\text{sextractor}}$
22	Ser component contamination by neighbours or sky	$r_{\text{bulge,cir}}/r_{\text{sextractor}} > 4.0$
23	Exp component contamination by neighbours or sky	$r_{\text{bulge,cir}}/r_{\text{sextractor}} > 4.0$
24	Bad Ser and bad Exp components	Failure of measurements
25	GALFIT failure	GALFIT fails to converge or other failure of the pipeline
26	Polluted or fractured	CasJobs neighbours not properly masked/masked or target is separated into 2 or more objects

**Table 5.7:** The description of our categories as described in the main text. The major flag bits used to select different catalogs are flag bits 10 (good two-component fits), 1 (good bulge galaxy), 4 (good disc galaxy), 14 (problematic two-component fit), and 20 (bad fit). The use of these flags is described in Chapter 5.4.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

Flag bit	Descriptive Category	Per cent SerExp	Per cent Test
—	Trust total and component magnitudes and sizes	39.055	39.167
10	Two-component galaxies	39.055	39.167
11	No flags	18.095	17.917
12	Good Ser, good Exp (some flags)	19.417	18.333
13	Flip components	1.543	2.917
—	Trust total magnitudes and sizes only	54.945	54.375
1	Bulge galaxies	18.964	18.958
2	No Exp component, $n_{Ser} > 2$	7.074	7.917
3	Ser dominates always	11.889	11.042
4	Disc galaxies	25.146	23.958
5	No Ser component	16.876	15.625
6	No Exp, $n_{Ser} < 2$ , flip components	0.551	0.208
7	Ser dominates always, $n_{Ser} < 2$	0.103	0.625
8	Exp dominates always	2.872	2.917
9	Parallel components	4.745	4.583
14	Problematic two-component galaxies	10.835	11.458
15	Ser outer only	7.504	8.125
16	Exp inner only	0.425	0.625
17	Good Ser, bad Exp, $B/T \geq 0.5$	0.017	0.000
18	Bad Ser, good Exp, $B/T < 0.5$	0.652	0.625
26	Bulge is point	2.237	2.083
19	Bad total magnitudes and sizes	6.000	6.458
20	Centering problems	0.557	0.625
21	Ser component contamination by neighbors or sky	2.129	3.333
22	Exp component contamination by neighbors or sky	2.392	2.083
23	Bad Ser and bad Exp components	0.239	0.417
24	GALFIT failure	0.355	0.208
25	Polluted or fractured	0.676	0.833

**Table 5.8:** A breakdown of the descriptive categories useful for analysis. The percentages of the total catalog are shown for each of the fitted models and our visually classified test set. The one component models (**Ser** and **deV**) can not be classified as two-component models, by definition. For the **Ser** and **deV** models, many of the categories are empty. The major distinction for the one-component models is whether the fits have major problems (i. e., flag bit 20 is set).

## 5.4 Final Flags

---

components are considered unreliable. The **Ser** model is the most appropriate model in this case, similar to all the bulge and disc flag categories.

Flag bit 24 (bad **Ser** and bad **Exp**) also includes failures that indicate strange fitting behavior. This category includes galaxies where the  $B/T > 0.5$  and the measured  $B/T$  is lower than 0.5, or vice versa. These cases occur when the bulge or disc parameters get very unusual (radii approaching 0 or  $b/a$  approaching 0).

The remaining flag bits are self-explanatory, so they are not described here. Table 5.8 shows the percentage of galaxies in each flag type for the **Ser-Exp** model and the visually classified test sample. The percentages for the **deV**, **Ser**, **deV-Exp**, and **Ser-Exp** models are shown in Table 5.10, 5.11, 5.12, and 5.13, respectively. The one-component models (**Ser** and **deV**) cannot be classified as two-component models, by definition, so many of the categories are empty. The major distinction for the one-component models is whether the fits have major problems (i. e., flag bit 20 is set). The failure rate (flag bit 20 set) increases with the complexity of the fits, but stays below 7 per cent for the **Ser-Exp** model and 8 per cent for the test sample. An additional  $\sim 10$  per cent of the two-component fits have problematic interpretations. The remaining  $\sim 80 - 85$  per cent of two-component galaxies appear to have good fits that can be used in analysis. As discussed previously, this does not guarantee that the fits can be interpreted as a physical bulge and a physical disc model. Many caveats to this interpretation still exist.

### 5.4.2 Examples of the final flags

This chapter, provides sample **Ser-Exp** fits to galaxies for each of the flagging categories described in Table 5.6. Although examples are shown for only for the **Ser-Exp** fits, the same flags are applied to the **deV-Exp** fit. For the **deV** and **Ser** fits, a subset of these flags is used that include only the flags applicable to one-component fits (i. e., the ‘bad’ flags and pure bulge or pure disc flags).

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

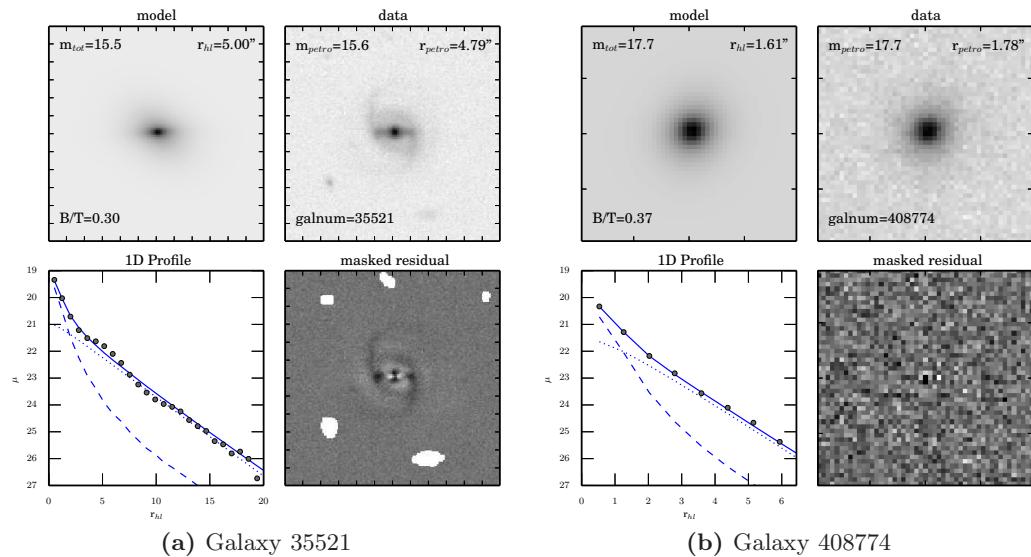
---

Each figure in this section shows two example galaxies that have the flag described in the figure caption. Each example galaxy shows the 2D fitted model and 2D data (top left and right, respectively). It also shows the 1D fitted profile and 2D masked residual image (bottom left and right, respectively). The total magnitude, B/T, and fitted, PSF-corrected half-light radius in arcseconds are shown in the 2D model image. The 2D data image has the GALNUM (our unique galaxy identification number), apparent Petrosian magnitude, and Petrosian half-light radius in arcseconds shown for comparison.

The 1D fitted profiles show the measured 1D profile using elliptical annuli as black dots. The total profile is plotted as the solid blue line. The **Ser-Exp** bulge is plotted as the dashed blue line and the **Ser-Exp** disc is plotted as the dotted blue line. The radial profiles are plotted in mag arcsec<sup>-2</sup> and the x-axis shows the radius in units of the fitted half-light radius.

Figures 5.9, 5.10, and 5.11 show various types of good two-component fits. Figures 5.12 and 5.13 show bulge fits. Figures 5.14, 5.15, 5.16, 5.17, and 5.18 show disc galaxies. Finally, Figures 5.19, 5.20, 5.21, 5.22, and 5.23 show two-component fits where the components are difficult to interpret.

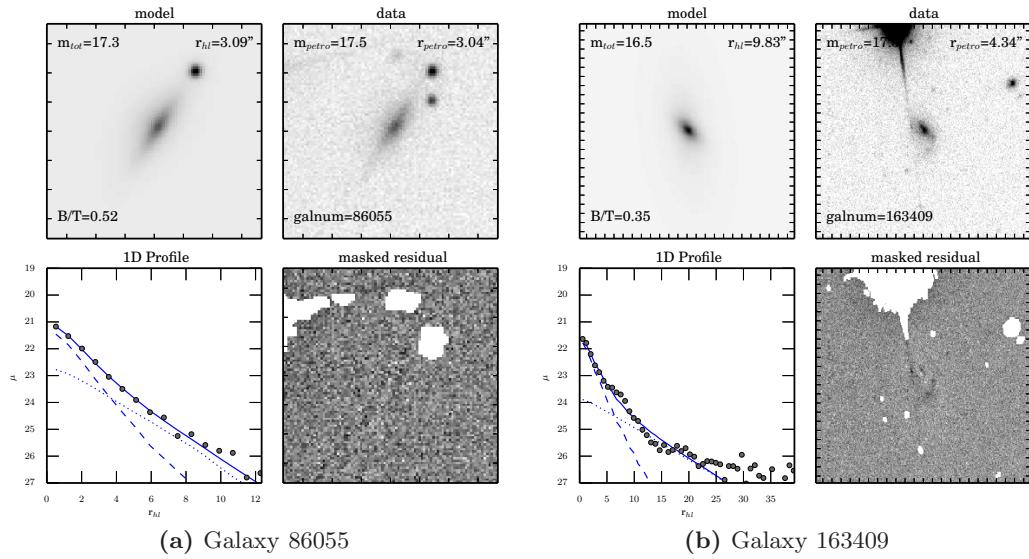
## 5.4 Final Flags



**Figure 5.9:** Example Ser-Exp fits considered good using our flagging criteria. For these galaxies, we claim that the two components of the fit are real and without fitting problems. These galaxies are in category ‘no flags’ in Table 5.6 and have flag bits 10 and 11 set.

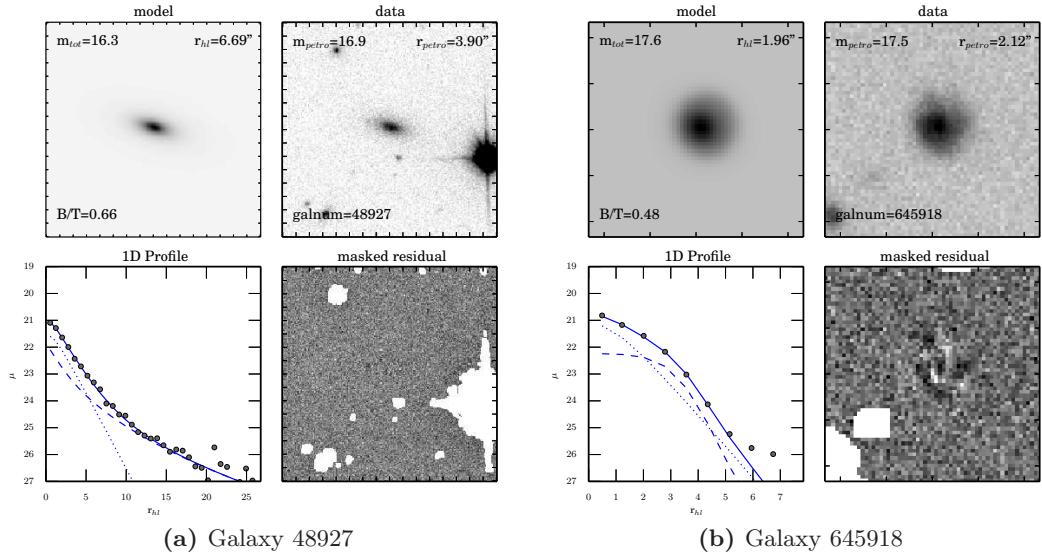
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.10:** Example Ser-Exp fits considered good using our flagging criteria. For these galaxies, we claim that the two components of the fit are real and largely free of fitting problems. However, there can be some isolated problems with the fits. For example, galaxy 86055 (left) has a high ellipticity bulge ( $b/a \approx 0.2$ ) and galaxy 163409 (right) has a bulge with low Sérsic index ( $n_{bulge} \approx 1.1$ ) and the bulge and disc position angles differ by a large amount ( $\approx 45^\circ$ ). However, we do not find enough reason to remove the fit as unphysical. These galaxies are in category ‘good Ser, good Exp (some flags)’ in Table 5.6 and have flag bits 10 and 12 set.

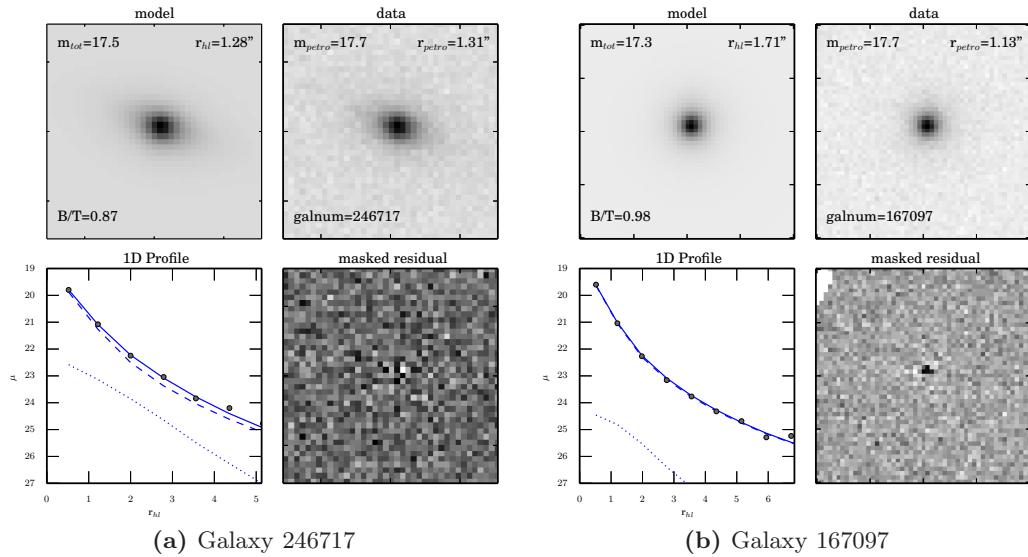
## 5.4 Final Flags



**Figure 5.11:** Example Ser-Exp fits considered good using our flagging criteria. For these galaxies, we claim that the two components of the fit are real. However, the components need to be interchanged to fit with the standard interpretation of the fits that we use throughout the paper. These galaxies are in category ‘flip components,  $n_{Ser} < 2$ ’ in Table 5.6 and have flag bits 10 and 13 set.

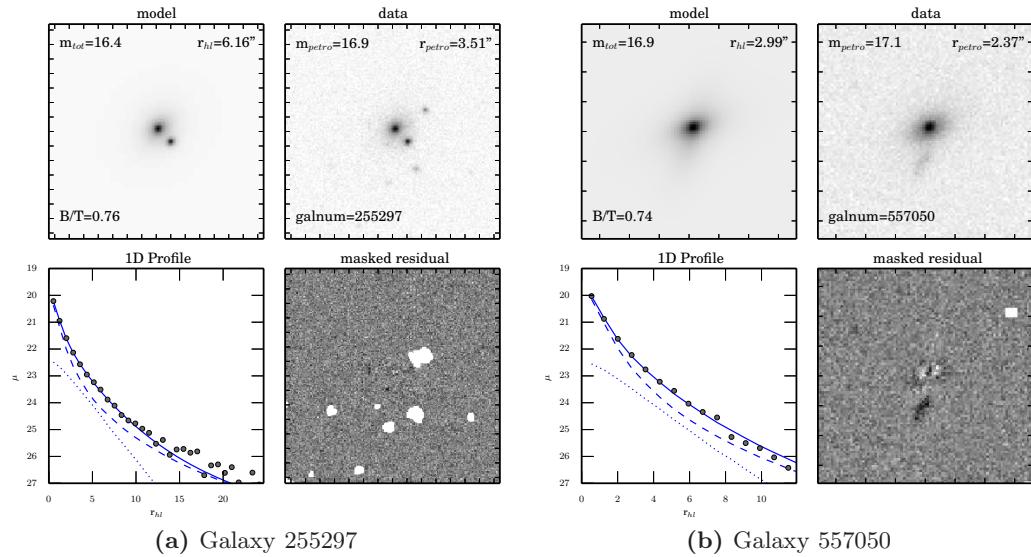
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.12:** Example Ser-Exp galaxies considered single-component galaxies with bulge-like profiles. The profiles have little or no Exp component contributing light to the galaxy. These galaxies are in category ‘no Exp component,  $n_{\text{Ser}} > 2$ ’ in Table 5.6 and have flag bits 1 and 2 set.

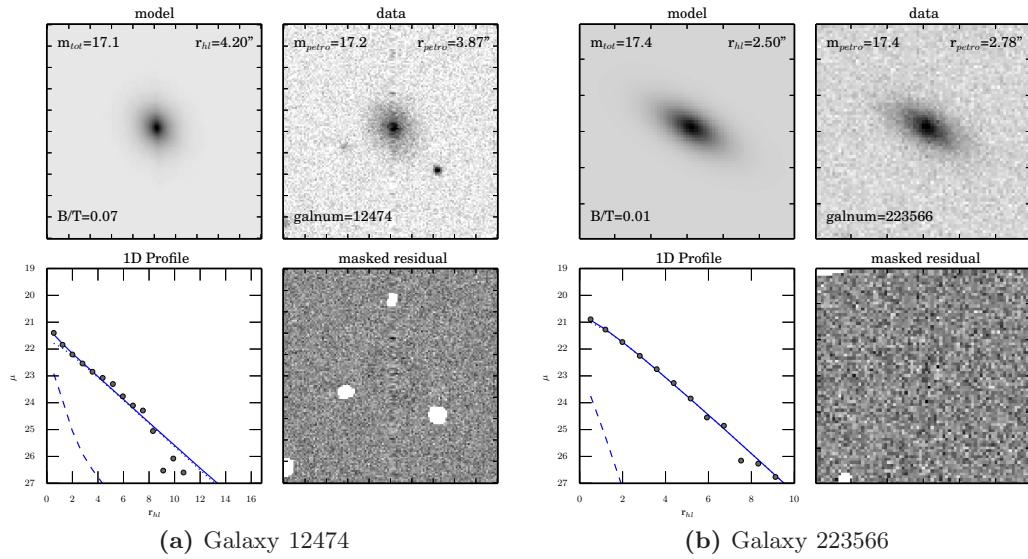
## 5.4 Final Flags



**Figure 5.13:** Example **Ser-Exp** galaxies considered single-component galaxies with bulge-like profiles. The profiles tend to have some small Exp component, but the **Ser** component dominates at all radii. These galaxies are in category ‘**Ser** dominates always’ in Table 5.6 and have flag bits 1 and 3 set.

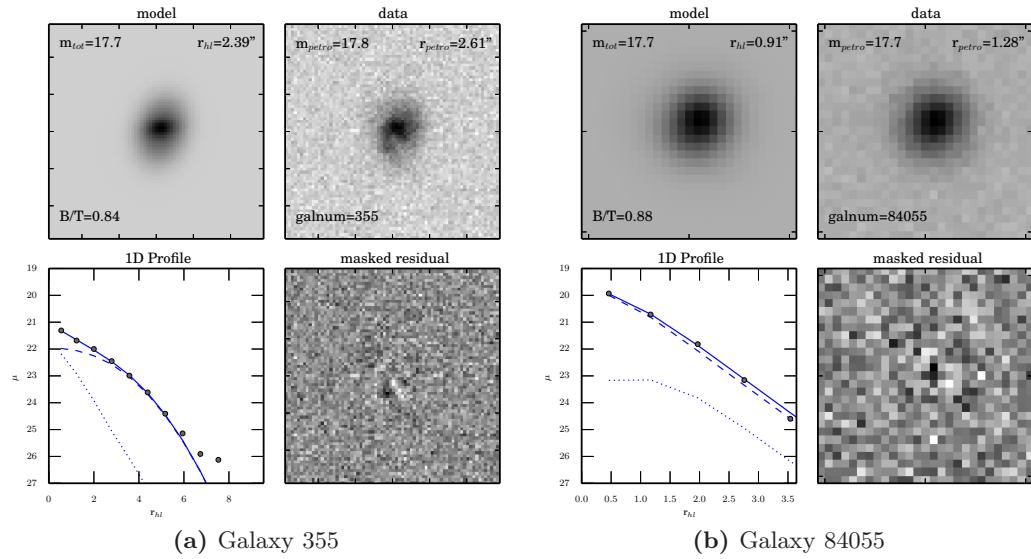
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.14:** Example Ser-Exp galaxies considered single-component galaxies with disc-like profiles. The profiles tend to have little or no contribution to the total light by the Ser component. These galaxies are in category ‘no Ser component’ in Table 5.6 and have flag bits 4 and 5 set.

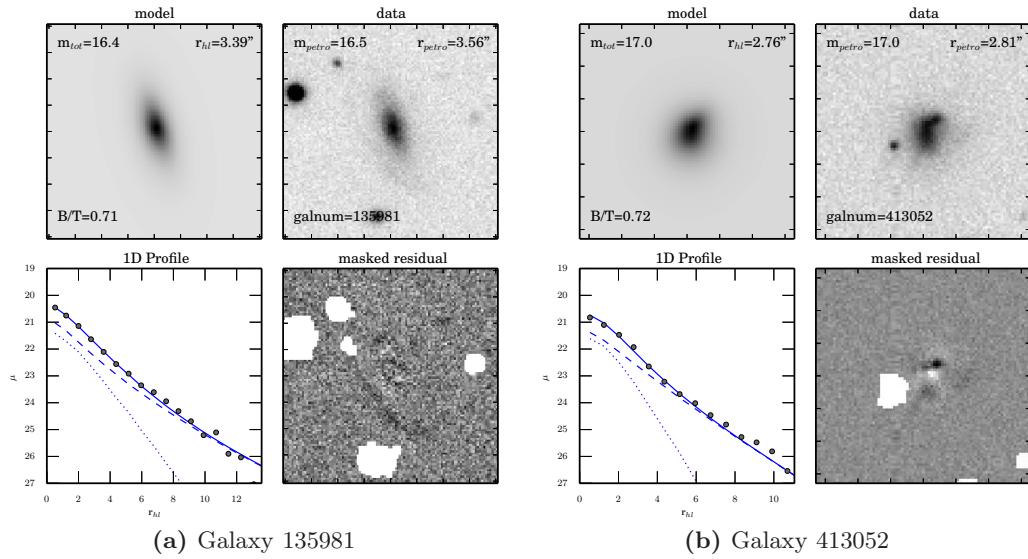
## 5.4 Final Flags



**Figure 5.15:** Example **Ser-Exp** galaxies considered single-component galaxies with disc-like profiles. The profiles tend to have little or no Exp component, but the **Ser** component is disc-like with a Sérsic index less than 2. These galaxies are in category ‘no Exp,  $n_{Ser} < 2$ , flip components’ in Table 5.6 and have flag bits 4 and 6 set.

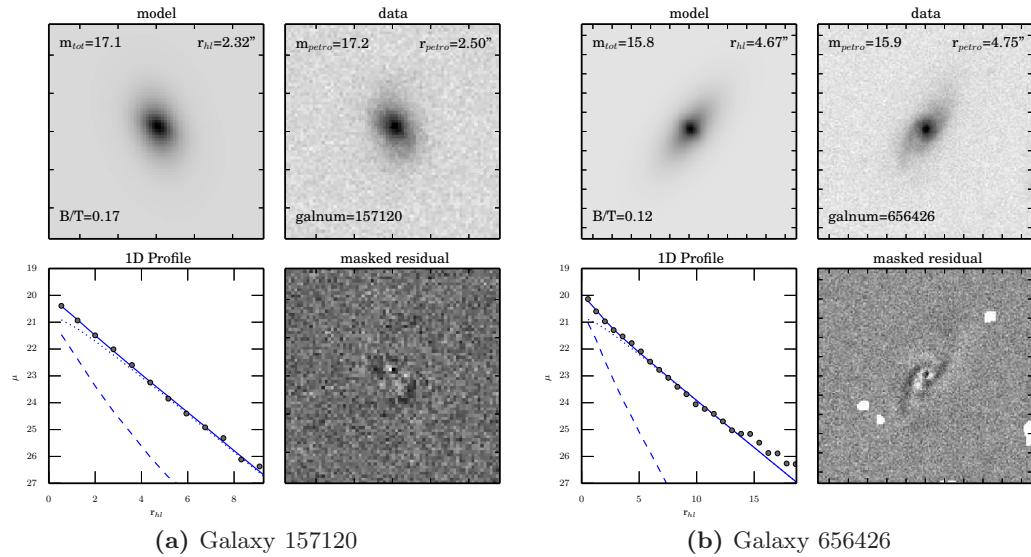
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.16:** Example **Ser-Exp** galaxies considered single-component galaxies with disc-like profiles. The profiles tend to have some small Exp component, but the **Ser** component dominates at all radii and is disc like with a Sérsic index less than 2. These galaxies are in category ‘**Ser** dominates always,  $n_{\text{Ser}} < 2$ ’ in Table 5.6 and have flag bits 4 and 7 set.

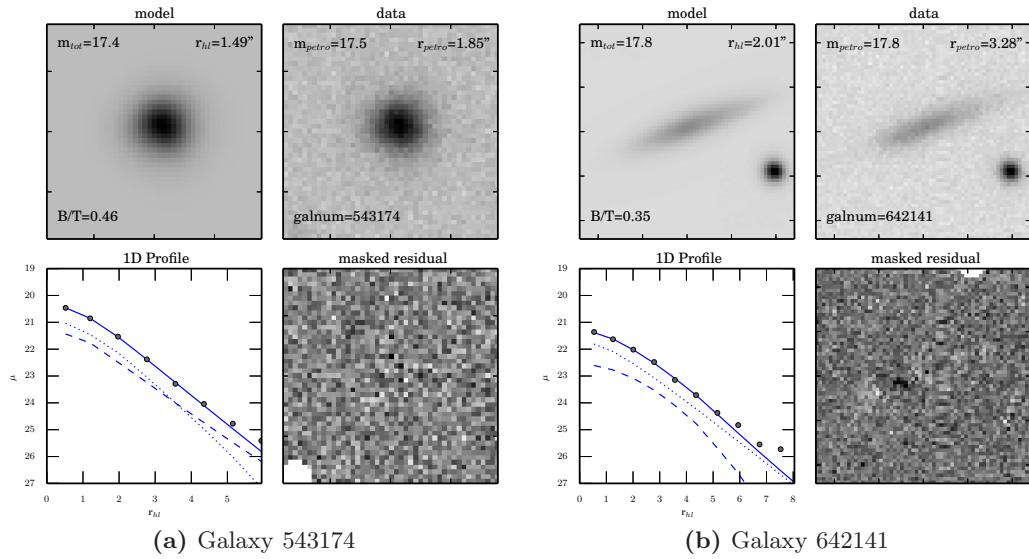
## 5.4 Final Flags



**Figure 5.17:** Example **Ser-Exp** galaxies considered single-component galaxies with disc-like profiles. The profiles tend to have some small **Ser** component, but the **Exp** component dominates at all radii. These galaxies are in category ‘**Exp** dominates always’ in Table 5.6 and have flag bits 4 and 8 set.

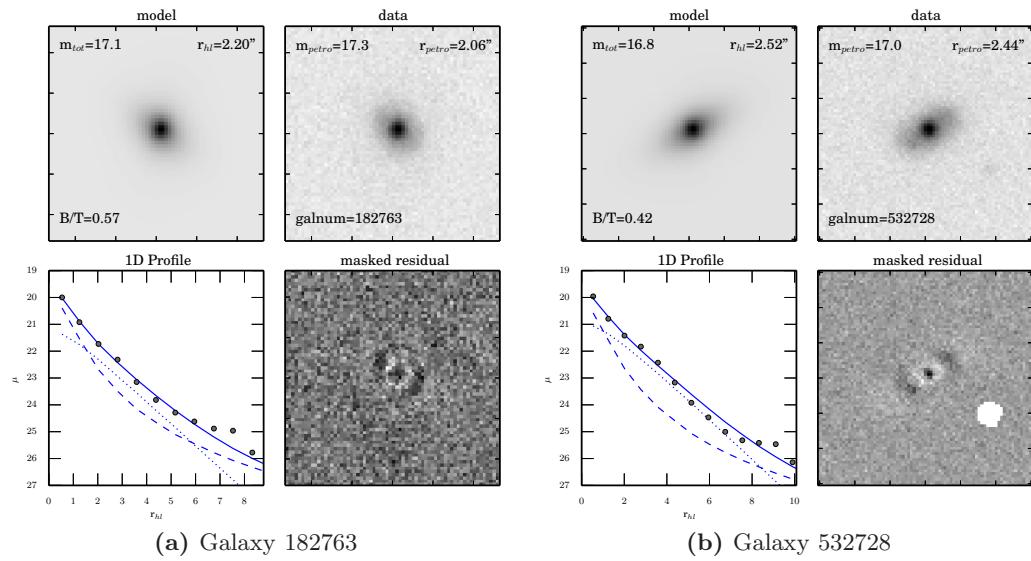
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.18:** Example **Ser-Exp** galaxies considered single-component galaxies with disc-like profiles. The profiles tend to have both **Ser** and **Exp** components, but the **Ser** component is disc-like with Sérsic index less than 2 and components contribute similar amounts of light to the profile across a range of radii (they appear to parallel each other in the 1D fit). These galaxies are in category ‘parallel components’ in Table 5.6 and have flag bits 4 and 9 set.

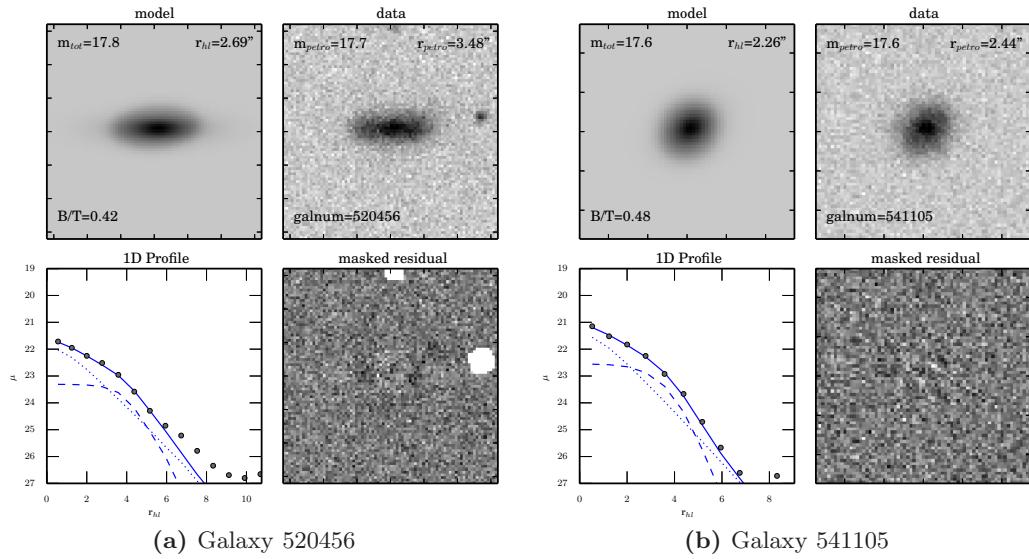
## 5.4 Final Flags



**Figure 5.19:** Example **Ser**-**Exp** galaxies considered two-component galaxies although the components may have problematic interpretations. The profiles tend to have both **Ser** and **Exp** components, but the **Ser** component dominates at large radii. This can be related to sky problems and artificially increase the bulge radius and magnitude. These galaxies are in category ‘**Ser** outer only’ in Table 5.6 and have flag bits 14 and 15 set.

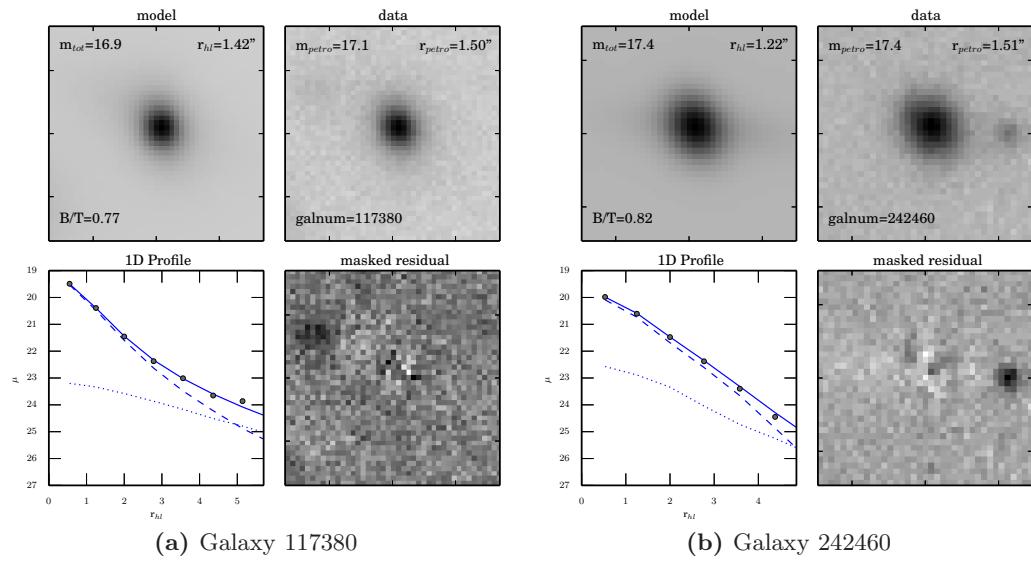
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.20:** Example **Ser-Exp** galaxies considered two-component galaxies although the components may have problematic interpretations. The profiles tend to have both **Ser** and **Exp** components, but the **Exp** component dominates at small radii and the **Ser** component is not necessarily disc like. These galaxies are in category ‘**Exp inner only**’ in Table 5.6 and have flag bits 14 and 16 set.

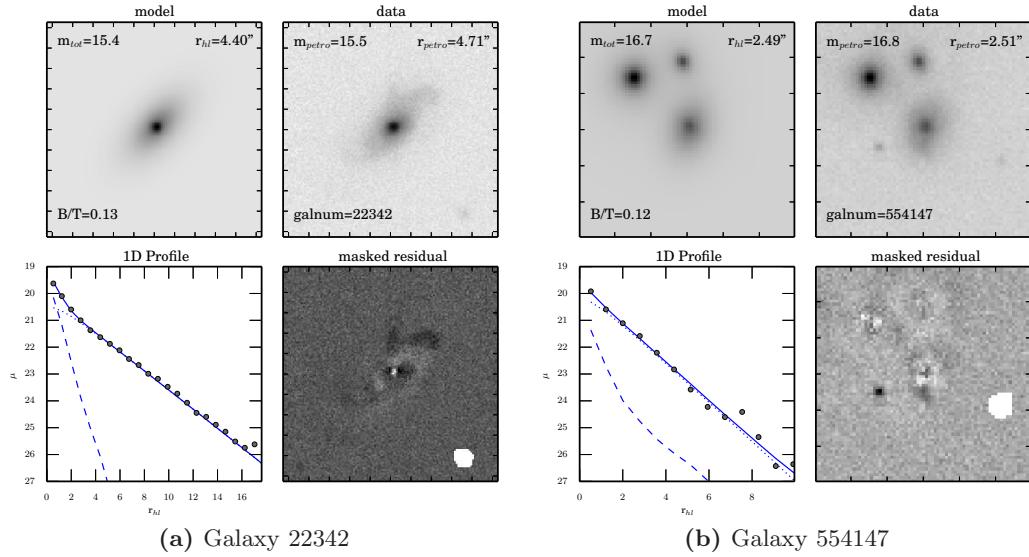
## 5.4 Final Flags



**Figure 5.21:** Example **Ser-Exp** galaxies considered two-component galaxies although the components may have problematic interpretations. The profiles tend to have both **Ser** and **Exp** components, but the **Exp** component is highly elliptical and misaligned with the galaxy. These galaxies are in category ‘good **Ser**, bad **Exp**,  $B/T \geq 0.5$ ’ in Table 5.6 and have flag bits 14 and 17 set.

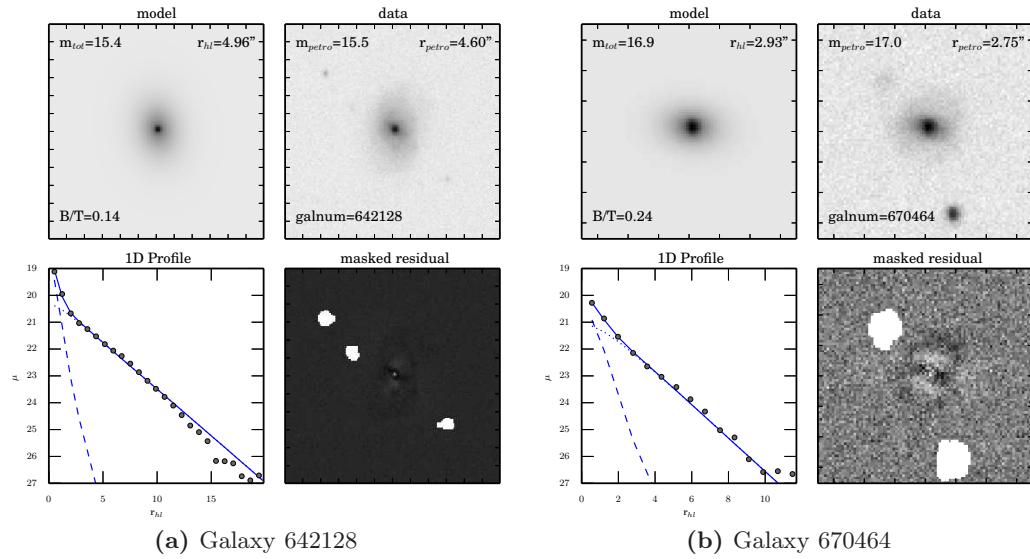
## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---



**Figure 5.22:** Example **Ser-Exp** galaxies considered two-component galaxies although the components may have problematic interpretations. The profiles tend to have both **Ser** and **Exp** components, but the **Ser** component is highly elliptical and misaligned with the galaxy. These galaxies are in category ‘bad **Ser**, good **Exp**,  $B/T<0.5$ ’ in Table 5.6 and have flag bits 14 and 18 set.

## 5.4 Final Flags



**Figure 5.23:** Example **Ser**-**Exp** galaxies considered two-component galaxies although the components may have problematic interpretations. The profiles tend to have both **Ser** and **Exp** components, but the **Ser** component has shrunk to a point source. These galaxies are in category ‘bulge is point’ in Table 5.6 and have flag bits 14 and 19 set.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

### 5.5 *g*- and *i*-band flag results

The flagging procedure presented in this chapter is intended (among other 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. The majority of the flags are constructed using relative measurements that should not require adjustment to accommodate measurements in bands other than the *r*-band. For instance, determining that a galaxy has a contaminated bulge or disk component (flag bits 22 and 23) requires comparing the size of the bulge and disk components to the Petrosian radius in the corresponding band. Since this measurement is not band-specific, it should change automatically when different bands are used.

However, the bounds for determining when a bulge or disk component is not present (i.e., flag bits 1 and 4) are partially determined by a cut on magnitude calibrated to the *r*-band fits using a training set of visually classified galaxies. Since the magnitudes and S/N are different in the *g*- and *i*-bands, it is possible that these cuts need to be adjusted to work in the neighboring bands.

In particular, the no bulge condition was

$$1000(B/T - 0.8)^3 + (m_{disc} - 19) > 0.5 \quad (5.2)$$

and the no disc condition was

$$1000(0.2 - B/T)^3 + (m_{bulge} - 19) > 0.5 \quad (5.3)$$

These cuts were set based on visual classification of fitted galaxies by the authors and the fact that scatter in the bulge or disc components grows substantially when the component was dimmer than 19. This cut of 19 was based on the assumption of S/N present in the *r*-band images. In particular, the average sky brightness and noise properties in the *g*-, *r*-, and *i*-bands differ as well as the average magnitudes.

## 5.5 *g*- and *i*-band flag results

---

<b>Band</b>	<b>r<sub>arcsec</sub></b>	<b>zeropoint</b>	$\mu_{\text{sky}}$ [mag arcsec <sup>-2</sup> ]	<b>Gain</b>	<b>Dark</b>	<b>Variance</b>	<b>S/N</b>	<b>m</b>
g	1.20	24.2	22.0	3.85	1.44	1.34	1.34	19.0
r	1.12	23.9	21.0	4.70	1.00	1.34	1.34	19.0
i	1.00	23.6	20.3	4.86	5.75	1.34	1.34	18.2

**Table 5.9:** Median values of radius, zeropoint, sky brightness, gain, and dark variance for the *g*-, *r*-, and *i*- bands. The S/N and magnitude are also shown. The *r*-band flags use a cut of 19 magnitudes as the basis for determining that a bulge or disk component is not reliable. Here, we show the median S/N of an object in the *r*-band, then using median values for the *g*- and *i*-bands, we calculate the equivalent magnitude of an object with the same S/N (within 1 per cent) as the *r*-band value. This band-dependent magnitude is used for bulge and disk flagging in the two-component *g*- and *i*-band fits and accounts for differences in the sky brightness and zeropoints between the bands.

Using Equation 2.2, the equivalent magnitude in the *g*-(and *i*)-bands of a component with the same S/N as an *r*-band magnitude 19.0 object with median sky brightness, median bulge(or disc) half-light radius, and median noise properties were calculated. Table 5.9 summarizes the median values used in each band and the equivalent magnitude for the *g*- and *i*-bands calculated using these values. The *g*-band requires no adjustment. The *i*-band requires more adjustment due to the increasing sky brightness and dark variance in the band. We use these magnitudes as the basis for the flag bits 1 and 4 (i.e., “no bulge” or “no disk”) flags. These flags are the only flags that use an absolute cut on magnitude, so the remaining flags require no such adjustment.

Tables 5.10, 5.11, 5.12, and 5.13 show the flagging categories, the final percentage of galaxies in each flagging category for the *g*-, *r*-, and *i*-bands (left, center, and right columns) for the **deV**, **Ser**, **deV-Exp**, and **Ser-Exp** models, respectively.

About 94% of the *r*-band **Ser-Exp** sample is classified as having an accurate measurement of the total magnitude and half-light radius. This percentage is reduced in the *g*- and *i*-bands to approximately 92 percent. The increase in failed fits can be mostly attributed to increasing neighbor contamination, which is reflected in the flagging as an increase of the

## **5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS**

---

“**Ser** component contamination by Neighbors or Sky” and the “**Exp** component contamination by Neighbors or Sky” flags which contain an additional  $\approx 1.5$  per cent of the total sample compared to the original *r*-band fits.

In addition to the increase of failed galaxies, there is a decrease in the percentage of two-component fits from  $\approx 39$  percent in the *r*-band 33 percent in the *g*-band and 37 per cent in the *i*-band. At the same time, the number of pure bulge galaxies drops in the bluer filters from 20 percent in the *i*-band to 17 per cent in the *g*-band and the number of disks increases substantially from 24 per cent in the *i*-band to 31 per cent in the *g*-band). These observations agree with a simple bulge+disk model of galaxies where the disk tends to be bluer and the bulge is redder. It may also partially explain the increase in failed cases due to contamination or sky since the *g*-(*i*-) band will feature the disk(bulge) more prominently and the second component will be dimmer or not present due to the different color. With dim or low S/N component, there is a more likely possibility that the fitting routine will settle into a local minima that improperly uses this second component since it is not well constrained by the observations. A poorly constrained component may be used to fit stray light from neighboring galaxies, increasing the failure rate as observed here.

## 5.5 *g*- and *i*-band flag results

---

Descriptive Category	% dev, <i>g</i>	% dev, <i>r</i>	% dev, <i>i</i>
<b>Trust Total and Component Magnitudes and Sizes</b>	—	—	—
<b>Two-Component Galaxies</b>	—	—	—
No Flags	—	—	—
Good <b>Ser</b> , Good <b>Exp</b> (Some Flags)	—	—	—
Flip Components, $n_{Ser} < 2$	—	—	—
<b>Trust Total Magnitudes and Sizes Only</b>	96.623	97.653	96.958
<b>Bulge Galaxies</b>	96.623	97.653	96.958
No <b>Exp</b> Component, $n_{Ser} > 2$	96.623	97.653	96.958
<b>Ser</b> Dominates Always	—	—	—
<b>Disk Galaxies</b>	—	—	—
No <b>Ser</b> Component	—	—	—
No <b>Exp</b> , $n_{Ser} < 2$ , Flip Components	—	—	—
<b>Ser</b> Dominates Always, $n_{Ser} < 2$	—	—	—
<b>Exp</b> Dominates Always	—	—	—
Parallel Components	—	—	—
<b>Problematic Two-Component Galaxies</b>	—	—	—
<b>Ser</b> Outer Only	—	—	—
<b>Exp</b> Inner Only	—	—	—
Good <b>Ser</b> , Bad <b>Exp</b> , $B/T \geq 0.5$	—	—	—
Bad <b>Ser</b> , Good <b>Exp</b> , $B/T < 0.5$	—	—	—
Tiny Bulge, otherwise good	—	—	—
<b>Bad Total Magnitudes and Sizes</b>	3.377	2.347	3.042
Centering Problems	0.292	0.338	0.311
<b>Ser</b> Component Contamination by Neighbors or Sky	1.865	1.302	1.651
<b>Exp</b> Component Contamination by Neighbors or Sky	—	—	—
Bad <b>Ser</b> and Bad <b>Exp</b> Components	—	—	—
Galfit Failure	0.386	0.187	0.286
Polluted or Fractured	1.041	0.679	0.947

**Table 5.10:** deV model table showing flagging breakdown for the *g*-, *r*-, and *i*-bands.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

Descriptive Category	% ser, <i>g</i>	% ser, <i>r</i>	% ser, <i>i</i>
<b>Trust Total and Component Magnitudes and Sizes</b>	—	—	—
<b>Two-Component Galaxies</b>	—	—	—
No Flags	—	—	—
Good <b>Ser</b> , Good <b>Exp</b> (Some Flags)	—	—	—
Flip Components, $n_{Ser} < 2$	—	—	—
<b>Trust Total Magnitudes and Sizes Only</b>	95.502	97.380	96.612
<b>Bulge Galaxies</b>	52.177	58.378	60.972
No Exp Component, $n_{Ser} > 2$	52.177	58.378	60.972
<b>Ser</b> Dominates Always	—	—	—
<b>Disk Galaxies</b>	43.325	39.001	35.640
No <b>Ser</b> Component	—	—	—
No Exp, $n_{Ser} < 2$ , Flip Components	43.325	39.001	35.640
<b>Ser</b> Dominates Always, $n_{Ser} < 2$	—	—	—
Exp Dominates Always	—	—	—
Parallel Components	—	—	—
<b>Problematic Two-Component Galaxies</b>	—	—	—
<b>Ser</b> Outer Only	—	—	—
Exp Inner Only	—	—	—
Good <b>Ser</b> , Bad <b>Exp</b> , $B/T \geq 0.5$	—	—	—
Bad <b>Ser</b> , Good <b>Exp</b> , $B/T < 0.5$	—	—	—
Tiny Bulge, otherwise good	—	—	—
<b>Bad Total Magnitudes and Sizes</b>	4.498	2.620	3.388
Centering Problems	0.320	0.357	0.322
<b>Ser</b> Component Contamination by Neighbors or Sky	3.012	1.618	2.036
<b>Exp</b> Component Contamination by Neighbors or Sky	—	—	—
Bad <b>Ser</b> and Bad <b>Exp</b> Components	—	—	—
Galfit Failure	0.327	0.124	0.239
Polluted or Fractured	1.046	0.681	0.948

**Table 5.11:** Ser model table showing flagging breakdown for the *g*-, *r*-, and *i*-bands.

## 5.5 *g*- and *i*-band flag results

---

Descriptive Category	% devexp, <i>g</i>	% devexp, <i>r</i>	% devexp, <i>i</i>
<b>Trust Total and Component Magnitudes and Sizes</b>	33.077	42.223	44.524
<b>Two-Component Galaxies</b>	33.077	42.223	44.524
No Flags	22.918	30.152	31.952
Good Ser, Good Exp (Some Flags)	8.526	10.701	11.180
Flip Components, $n_{Ser} < 2$	1.633	1.369	1.391
<b>Trust Total Magnitudes and Sizes Only</b>	59.362	52.444	48.390
<b>Bulge Galaxies</b>	15.043	14.636	14.211
No Exp Component, $n_{Ser} > 2$	9.610	4.646	3.398
Ser Dominates Always	5.433	9.989	10.813
<b>Disk Galaxies</b>	36.291	28.462	24.225
No Ser Component	34.502	25.041	20.070
No Exp, $n_{Ser} < 2$ , Flip Components	—	—	—
Ser Dominates Always, $n_{Ser} < 2$	—	—	—
Exp Dominates Always	1.789	3.421	4.156
Parallel Components	—	—	—
<b>Problematic Two-Component Galaxies</b>	8.029	9.346	9.954
Ser Outer Only	4.906	5.289	5.494
Exp Inner Only	0.632	0.514	0.534
Good Ser, Bad Exp, $B/T \geq 0.5$	0.003	0.011	0.010
Bad Ser, Good Exp, $B/T < 0.5$	0.620	0.884	1.042
Tiny Bulge, otherwise good	1.868	2.648	2.874
<b>Bad Total Magnitudes and Sizes</b>	7.560	5.334	7.086
Centering Problems	0.762	0.599	0.769
Ser Component Contamination by Neighbors or Sky	1.816	1.251	1.608
Exp Component Contamination by Neighbors or Sky	3.887	2.788	3.868
Bad Ser and Bad Exp Components	0.634	0.177	0.148
Galfit Failure	0.389	0.249	0.290
Polluted or Fractured	1.038	0.677	0.945

**Table 5.12:** deV-Exp model table showing flagging breakdown for the *g*-, *r*-, and *i*-bands.

## 5. FLAGGING AND CLASSIFICATION OF GOOD AND BAD FITS

---

Descriptive Category	% serexp, <i>g</i>	% serexp, <i>r</i>	% serexp, <i>i</i>
<b>Trust Total and Component Magnitudes and Sizes</b>	33.290	39.055	37.049
<b>Two-Component Galaxies</b>	33.290	39.055	37.049
No Flags	14.266	18.095	18.088
Good <b>Ser</b> , Good <b>Exp</b> (Some Flags)	16.662	19.417	17.704
Flip Components, $n_{\text{Ser}} < 2$	2.362	1.543	1.258
<b>Trust Total Magnitudes and Sizes Only</b>	58.616	54.945	54.926
<b>Bulge Galaxies</b>	17.245	18.964	20.104
No Exp Component, $n_{\text{Ser}} > 2$	11.255	7.074	10.058
Ser Dominates Always	5.990	11.890	10.046
<b>Disk Galaxies</b>	31.197	25.146	24.193
No Ser Component	21.438	16.876	18.453
No Exp, $n_{\text{Ser}} < 2$ , Flip Components	1.455	0.551	0.744
Ser Dominates Always, $n_{\text{Ser}} < 2$	0.121	0.103	0.074
Exp Dominates Always	2.013	2.872	1.607
Parallel Components	6.170	4.745	3.316
<b>Problematic Two-Component Galaxies</b>	10.174	10.835	10.629
Ser Outer Only	7.218	7.504	8.383
Exp Inner Only	0.943	0.425	0.325
Good Ser, Bad Exp, $B/T \geq 0.5$	0.008	0.017	0.012
Bad Ser, Good Exp, $B/T < 0.5$	0.512	0.652	0.552
Tiny Bulge, otherwise good	1.493	2.237	1.357
<b>Bad Total Magnitudes and Sizes</b>	8.094	6.000	8.025
Centering Problems	0.619	0.557	0.679
Ser Component Contamination by Neighbors or Sky	3.096	2.129	2.848
Exp Component Contamination by Neighbors or Sky	2.992	2.393	3.354
Bad Ser and Bad Exp Components	0.824	0.239	0.347
Galfit Failure	0.418	0.355	0.311
Polluted or Fractured	1.038	0.676	0.946

**Table 5.13:** A breakdown of the flags now shown for the g-, r-, and i-band fits. The prevalence of galaxies with two-component fits decreases toward the bluer end from  $\approx 39$  per cent in the i-band to 33 per cent in the g-band. The prevalence of pure bulge galaxies also decreases toward bluer bands while the prevalence of disk galaxies increases. This is as expected using a simple bulge+disk model with a bluer disk and redder bulge.

# Chapter 6

## Internal comparisons and consistency checks

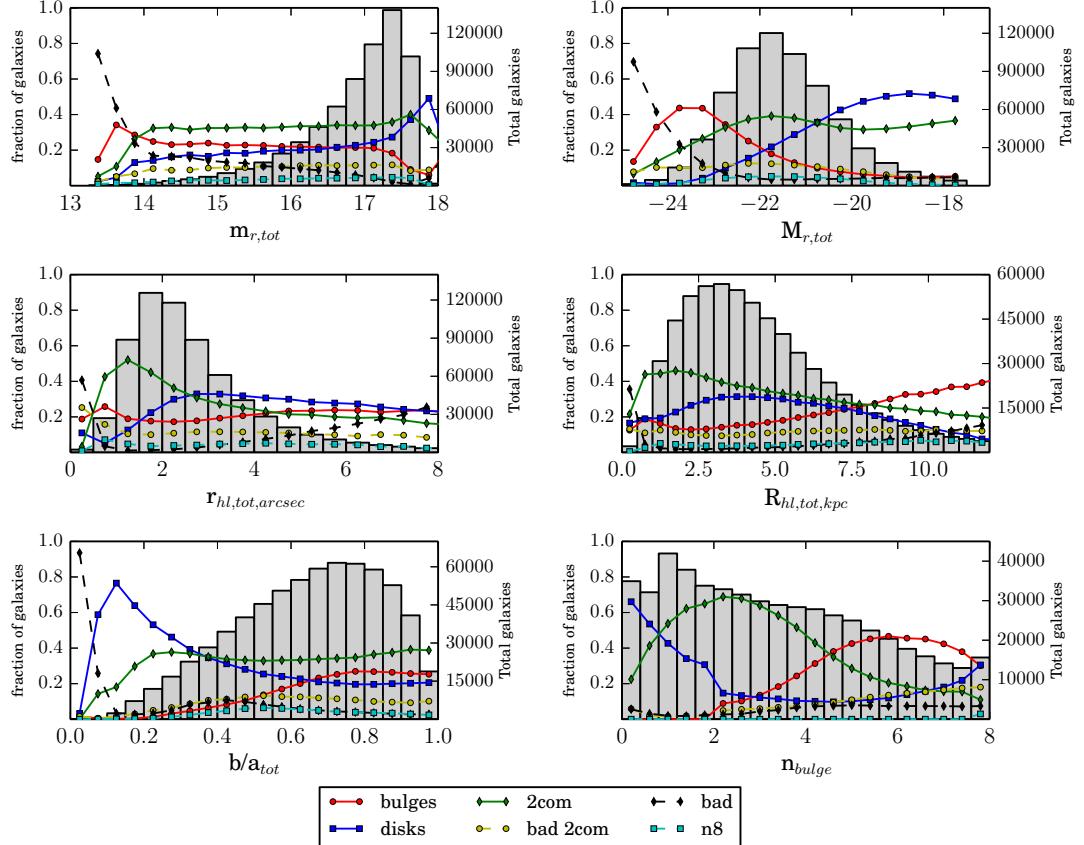
Before comparing the fits, in particular the **Ser-Exp** fits, to other groups, some internal consistency checks are performed. Chapter 6.1 examines the distribution of general classifications described in Chapter 5.3 for the **Ser-Exp** catalog as a function of basic observables. Chapter 6.3 examines the bulge radii with respect to the PSF. Chapter 6.4 compares some of the fitted values of the **Ser** fit to the **Ser-Exp** fit.

### 6.1 Examination of flags with basic observables

Using the flags from Chapter 5.3, one should be able to reproduce sensible distributions of magnitude, radius, and ellipticity provided that the chosen flags can reliably separate good and bad fits within the sample. The bulge Sérsic index is also included, which is used during classification. Figure 6.1 and Figure 6.2 show the percentage of two-component (green points, labeled as ‘2com’, flag bit 10), bulge (red points, labeled as ‘bulges’, flag bit 1), disc (blue points, labeled as ‘discs’, flag bit 4), problematic two-component (yellow points, labeled as ‘bad 2com’, flag bit 14), and failed galaxies (black points, labeled as

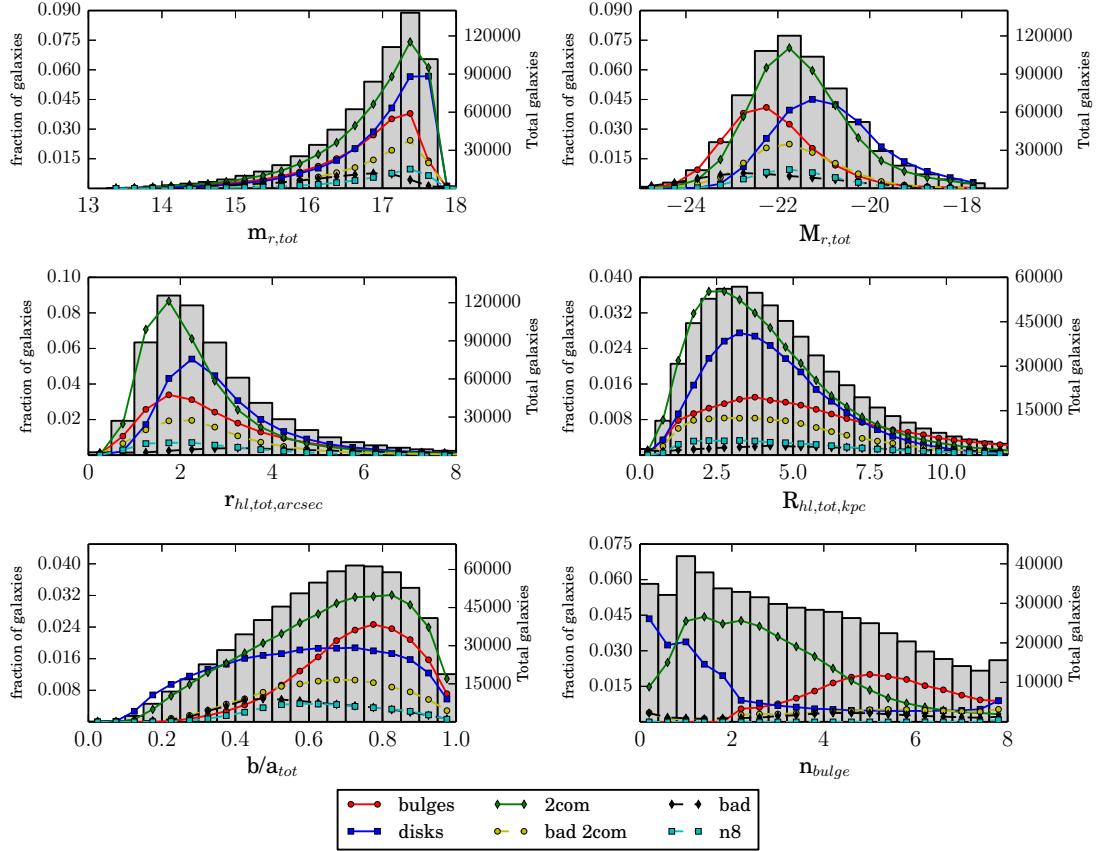
## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---



**Figure 6.1:** The bin-by-bin percent of galaxies of each type: disk (blue points, labeled as ‘discs’), bulge (red points, labeled as ‘bulges’), two-component (green points, labeled as ‘2com’), problematic two-component (yellow points, labeled as ‘bad 2com’), and failed galaxies (black points, labeled as ‘bad’), according to the categorical flags. The good two-component models are separated into those with Sérsic indexes 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, labeled as ‘n8’) to check for any bias resulting from the restriction on the fitted Sérsic index. In the background of the plot, the total distribution of galaxies with respect to the parameter used to bin the data is plotted.

## 6.1 Examination of flags with basic observables



**Figure 6.2:** The percent of galaxies of each type, given as a percentage of the total sample. The colors and plot design are the same as Figure 6.1. 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)

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---

‘bad’, flag bit 20) in the **Ser-Exp** catalog as a function of several parameters. The figures show the same distribution of the categories using different normalizations. Figure 6.1 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). Figure 6.2 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 indexes 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, labeled 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 behavior that would suggest substantial biases in the fitting due to this effect.

### 6.1.1 Behavior 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 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

## 6.1 Examination of flags with basic observables

---

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 (Bernardi et al. 2013; Blanton et al. 2001) and dimmer for galaxies with higher Sérsic index (or greater concentration) (see Figure 7 of Graham and Driver 2005), galaxies with fitted magnitudes brighter than 14 (where our original selection cut is made; see Chapter 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 neighbors. This makes sense as we expect galaxies with these bright magnitudes to have their brightness substantially overestimated.

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---

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, the training set may not be appropriate 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.

### 6.1.2 Behavior with absolute magnitude

Figures 6.1 and 6.2 also show the behavior of the 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 over estimate 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 neighbors. 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, the distribution of galaxy types was examined 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 Figure 6.1. 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

---

## 6.1 Examination of flags with basic observables

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 et al. (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 more difficult, caution is advised against a simple bulge+disc interpretation of these galaxies. Chapter 8 further discusses this issue.

### 6.1.3 Behavior with apparent and absolute half-light radius

Figure 6.1 and 6.2 also show the behavior 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, the expected trends in physical size (larger physical size should be dominated by bulge galaxies) are observed. 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, Figure 6.1 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

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---

magnitude in the SDSS sample (see Figures 5 and 6 of Shen et al. 2003). The  $-21$  magnitude also corresponds to the peak of the disc galaxy distribution in Figure 6.2. 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 Chapter 6.3 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.

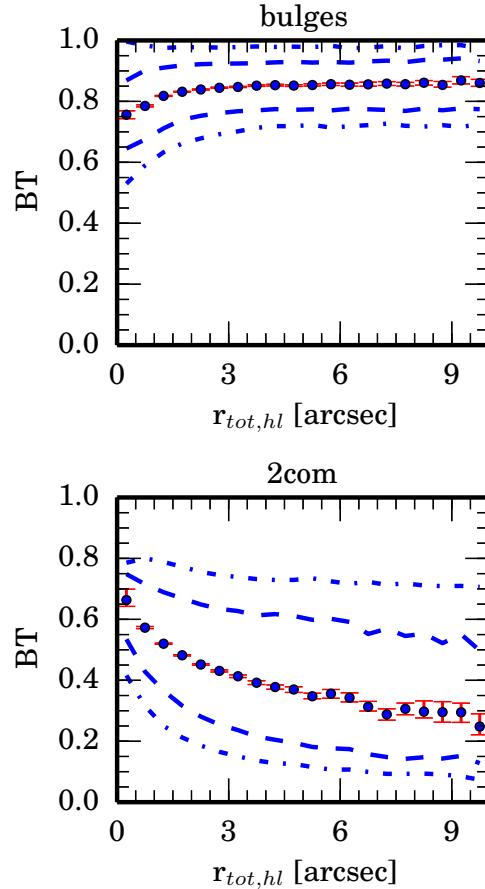
A similar examination of bulge galaxies versus two-component galaxies is less conclusive. Both groups (bulge and two-component) should be present at the smaller radii (below 2 arcsec).

Further examination of the B/T of the two-component and pure bulge galaxies is shown in Figure 6.3. 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. Errorbars 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 catalog 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 Chapter 6.3. Here, we caution the reader that there may be some classification bias preferring two-component fits at small sizes.

## 6.1 Examination of flags with basic observables

---



**Figure 6.3:** 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. Errorbars 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.

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---

### 6.1.4 Behavior with axis ratio

Figure 6.1 and 6.2 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 catalog (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. Figure 6.1 shows this behavior. Figure 6.2 also shows that the peak of the bulge galaxy distribution is near 0.8 with few galaxies at small axis ratios.

### 6.1.5 Behavior with bulge Sérsic index

Figure 6.1 and 6.2 also show the distribution of bulge Sérsic index derived from the Ser-Exp fits to galaxies in our catalog (bottom right). Figure 6.1 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. Figure 6.2 also shows similar behavior.

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.

### 6.1.6 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, it may be appropriate to associate our bulge and disc classes with early and late-type galaxies, respectively. Further evidence for this claim is discussed in Chapter 8.

Chapter 4 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. The **Ser** and **Exp** components of two-component **Ser-Exp** fits will be referred to as ‘bulge’ and ‘disk’ throughout this work. The reader should be cautioned that there are many cases where this simple interpretation does not make sense. Additional discussion is in Chapter 8.

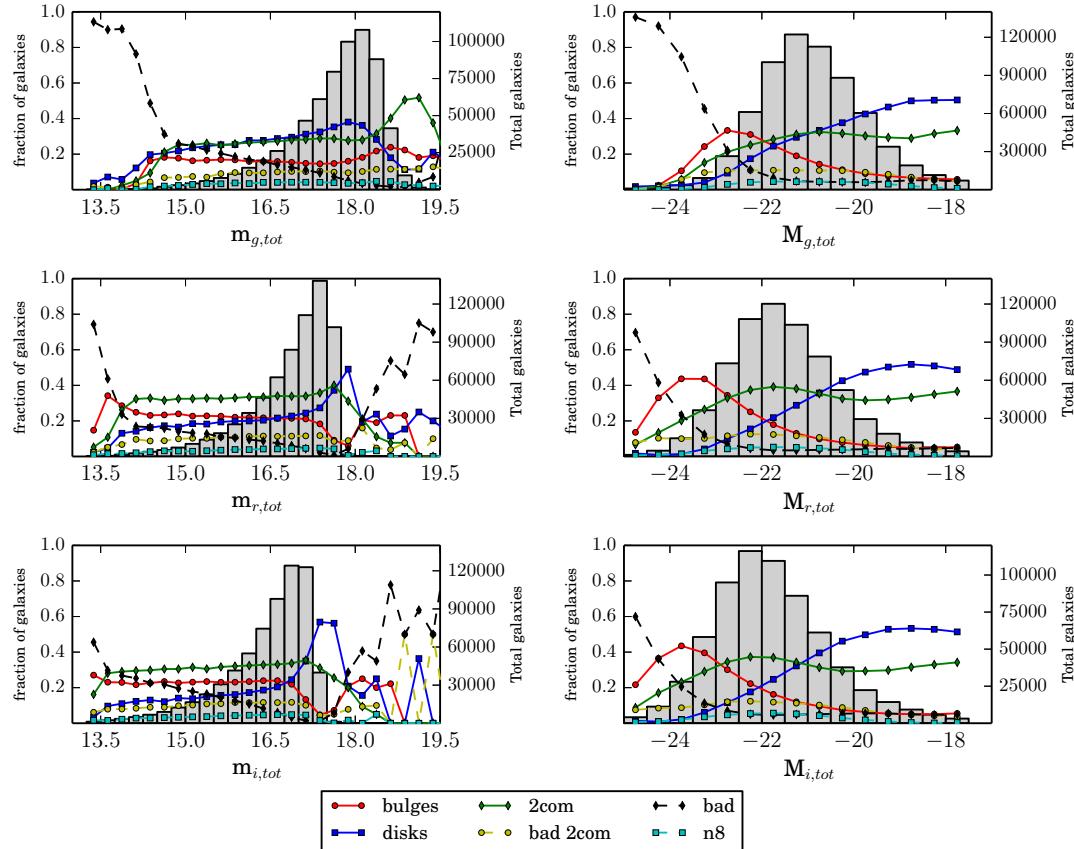
## 6.2 *g*- and *i*-band results

Figure 6.4 shows the apparent magnitude (left-hand side) and absolute magnitude distribution (right-hand side) of the *g*- (top row), *r*- (middle row), and *i*-band (bottom row) for galaxies in the catalog. The figure shows the bin-by-bin percent of galaxies of each type; Disk (blue points, labeled as “disks”), Bulge (red points, labeled as “bulges”), Two-Component (green points, labeled as “2com”), Problematic Two-Component (yellow points, labeled as “bad 2com”), Failed Galaxies (black points, labeled as “bad”), according to the categorical flags. The good two-component models are also separated into those with Sérsic indexes 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, labeled as “n8”) to check for any bias resulting from the restriction on the fitted Sérsic index. In the background of the plot, the total distribution of galaxies is plotted with respect to the parameter used to bin the data.

The *r*- and *i*-bands appear similar in behavior, while the *g*-band shows increased contributions from pure disk galaxies at the peak of the apparent magnitude distribution. This increased contribution is also reflected in the absolute magnitude distribution where there

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---



**Figure 6.4:** The magnitude distribution of bin-by-bin percent of galaxies of each type; Disk (blue points, labeled as “disks”), Bulge (red points, labeled as “bulges”), Two-Component (green points, labeled as “2com”), Problematic Two-Component (yellow points, labeled as “bad 2com”), Failed Galaxies (black points, labeled as “bad”), according to our categorical flags. We have also separated the good two-component models into those with Sérsic indexes 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, labeled 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.

## 6.2 *g*- and *i*-band results

---

is a smaller region with 2-component models dominating and the range of magnitudes with majority pure disk population is larger. At 18.5 or dimmer in the *g*-band, two-component models begin to dominate. A similar trend is not seen in the dimmer *r*- or *i*-band fits. This may be a flagging or fitting systematic. However, it may also be real signal. Galaxies with 18.5 or dimmer *g*-band magnitude are definitely brighter in the redder bands (if this were not true, then the galaxy would not have been selected as part of the sample). These redder galaxies are more likely to contain bulge components and bulge components are generally **Ser**-like rather than **Exp**-type profiles. So it is reasonable to expect that the dimmer *g*-band objects would show an increasing fraction of profiles containing bulges. The absolute magnitude distribution of the *g*-band supports the argument made here that there is not a negative systematic effect caused by the change in proportions of galaxy types at the dimmer *g*-band apparent magnitudes.

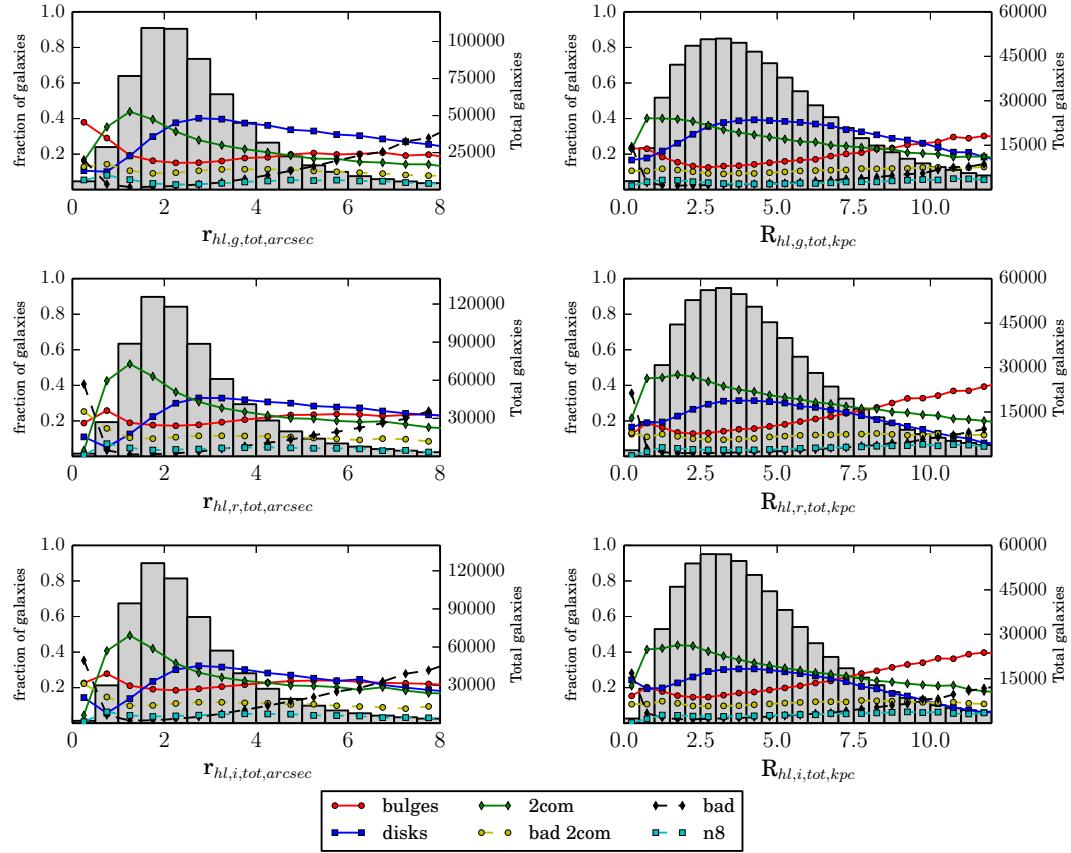
Figure 6.5 shows the apparent half-light radius (left-hand side) and physical half-light radius distribution (right-hand side) of the *g*- (top row), *r*- (middle row), and *i*-band (bottom row) for galaxies in the catalog. The figure is presented in the same way as Figure 6.4. All the apparent radii exhibit similar behavior with smaller radii being dominated by two-component models. This is possibly a systematic due to the sizes approaching the PSF size which is discussed in Chapter 6.3. The physical half-light radii are also similar in behavior across the bands. The bulges, disks, and 2com models appear to have the same behavior across the bands with only the overall relative scaling changing between bands.

Figure 6.6 shows the axis-ratio (left-hand side) and bulge Sérsic index (right-hand side) of the *g*- (top row), *r*- (middle row), and *i*-band (bottom row) for galaxies in the catalog. The figure is presented in the same way as Figure 6.4. Again, all bands show similar behavior.

Figures 6.4, 6.5, and 6.6 show that the fitting is quite similar in distribution across the bands with a smooth transition from having more disks in the *g*-band to fewer disks in the *g*-band. However, in addition to stability across the entire sample, single galaxies should show strong correlation across the three fits.

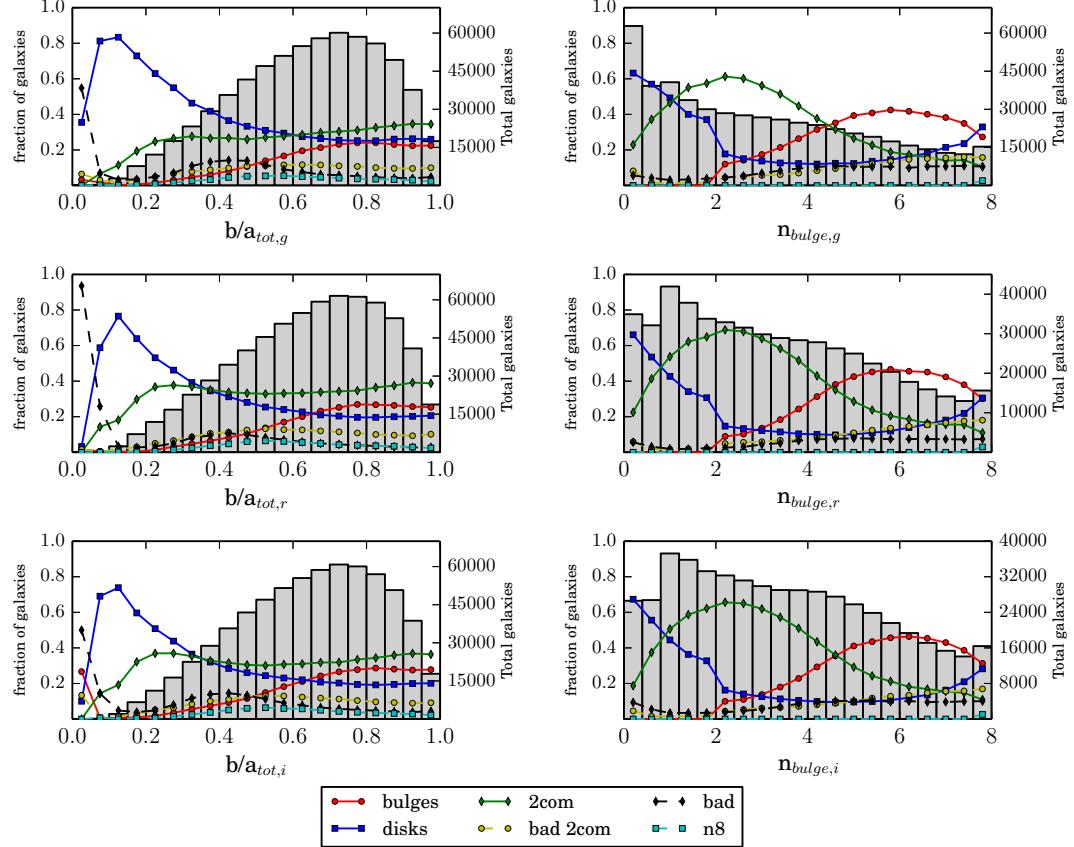
## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---



**Figure 6.5:** The size distribution of bin-by-bin percent of galaxies of each type; Disk (blue points, labeled as “disks”), Bulge (red points, labeled as “bulges”), Two-Component (green points, labeled as “2com”), Problematic Two-Component (yellow points, labeled as “bad 2com”), Failed Galaxies (black points, labeled as “bad”), according to the categorical flags. Good two-component models are separated into those with Sérsic indexes 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, labeled as “n8”) to check for any bias resulting from the restriction on the fitted Sérsic index. In the background of the plot, the total distribution of galaxies is plotted with respect to the parameter used to bin the data.

## 6.2 *g*- and *i*-band results



**Figure 6.6:** The total axis-ratio and bulge Sérsic index distribution of bin-by-bin percent of galaxies of each type; Disk (blue points, labeled as “disks”), Bulge (red points, labeled as “bulges”), Two-Component (green points, labeled as “2com”), Problematic Two-Component (yellow points, labeled as “bad 2com”), Failed Galaxies (black points, labeled as “bad”), according to the categorical flags. The good two-component models are also separated into those with Sérsic indexes 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, labeled as “n8”) to check for any bias resulting from the restriction on the fitted Sérsic index. In the background of the plot, the total distribution of galaxies is plotted with respect to the parameter used to bin the data.

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---

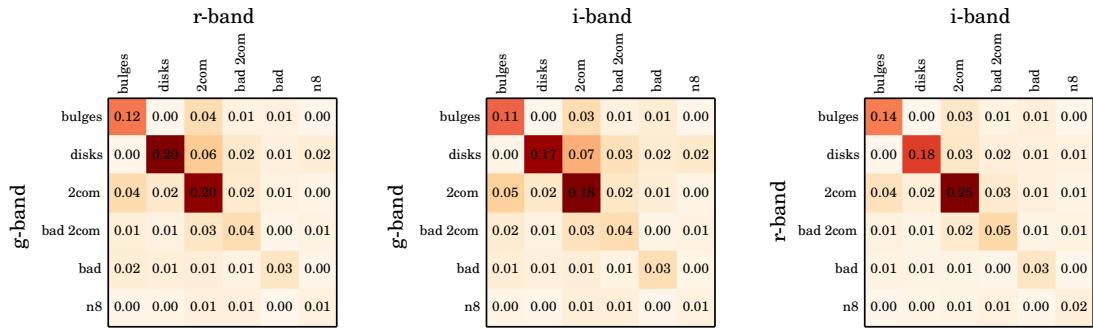
Figure 6.7 shows the breakdown of the fitting categories across two bands at a time. Each cell of the matrix shows the fraction of the entire catalog that is classified in the combination of bands and fitting categories denoted by the axes of the matrix (e.g. the top left cell of the first matrix shows that 12 percent of the sample is classified as “pure bulge” in both the  $g$ - and  $r$ -band fits). The majority of galaxies ( $\approx 65 - 70$  per cent) are in the top left  $3 \times 3$  quadrant of each matrix, indicating that the majority of galaxies are similarly identified across the bands. Also, the bulge-disk and disk-bulge cells are below 0.5 per cent (indicated by the 0.00 value) showing that few or no galaxies switch directly from bulge to disk categories, which would be unexpected. Additionally, the diagonal cells contain between 50 and 60 percent of the sample, further indicating that there is stability in the flagging classification used in this work.

### 6.3 PSF effects on bulge radius

The PSF limits the 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. Figure 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 indexes 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) ( $\sim 3000$  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. Figure 7 in Gadotti (2009) shows the distribution of the quantity (bulge radius/PSF

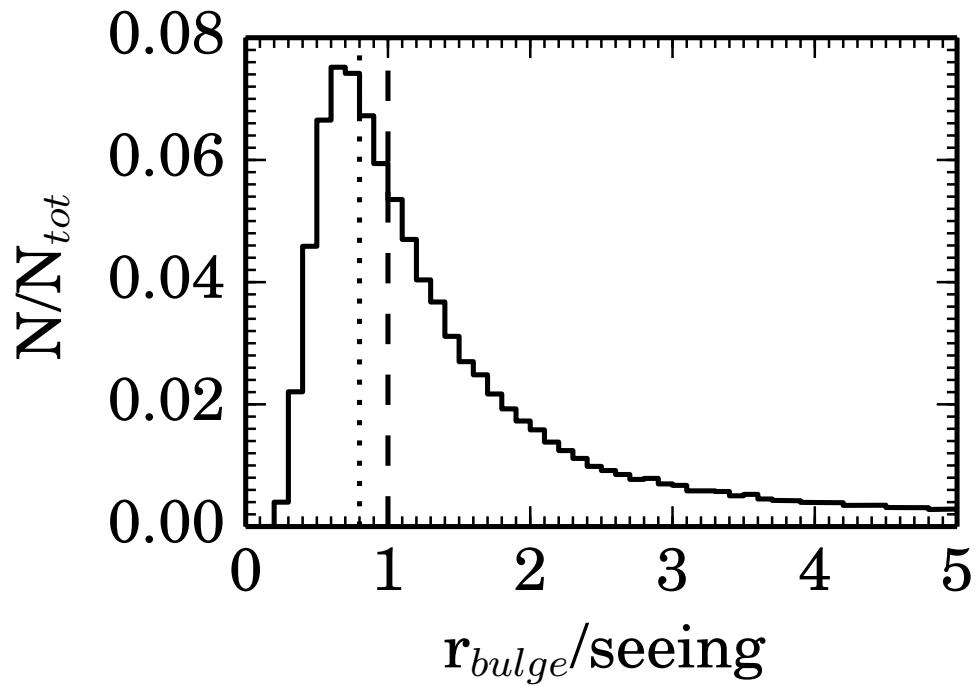
### 6.3 PSF effects on bulge radius



**Figure 6.7:** The confusion matrix for each combination of bands and flagging categories. The number in each cell denotes the fraction of the catalog classified by the combination of categories and bands denoted on the axes. For example, the top left cell of the first matrix shows that 12 percent of the sample is classified as “pure bulge” in both the g- and r-band fits. All possible combinations of bands are shown. The majority of galaxies ( $\approx 65 - 70$  per cent) are in the top left  $3 \times 3$  quadrant of each matrix, indicating that the majority of galaxies are similarly identified across the bands. Also, the bulge-disk and disk-bulge cells are below 0.5 per cent (indicated by the 0.00 value) showing that few or no galaxies switch directly from bulge to disk categories, which would be unexpected.

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---



**Figure 6.8:** 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×HWHM(1.0×HWHM).

### 6.3 PSF effects on bulge radius

---

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 percent 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.

Figure 6.8 shows the comparison of the bulge radius to PSF size for all **Ser-Exp** bulges in our catalog. 15 per cent (24 per cent) of bulges lie below  $0.8 \times \text{HWHM}$  ( $1.0 \times \text{HWHM}$ ). 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  $0.8 \times \text{HWHM}$  ( $1.0 \times \text{HWHM}$ ). 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  $0.8 \times \text{HWHM}$  ( $1.0 \times \text{HWHM}$ ). 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  $0.8 \times \text{HWHM}$  ( $1.0 \times \text{HWHM}$ ). The work of Gadotti (2009) suggests that a significant proportion of our **Ser-Exp** bulge components with  $B/T \leq 0.5$  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 catalog so that this consideration can be made during future analysis.

Chapter 4 also showed that the effective bulge radius for bulges smaller than 1 arcsec is overestimated by  $\sim 5$  per cent  $\pm 20$  per cent. The simulations reflect a tendency to overestimate bulge radius as reported in Gadotti (2008). There is no correction for this effect here, but user should be aware 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 Chapter 6.1.3.

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---

### 6.4 Comparison of the **Ser** and **Ser-Exp** models

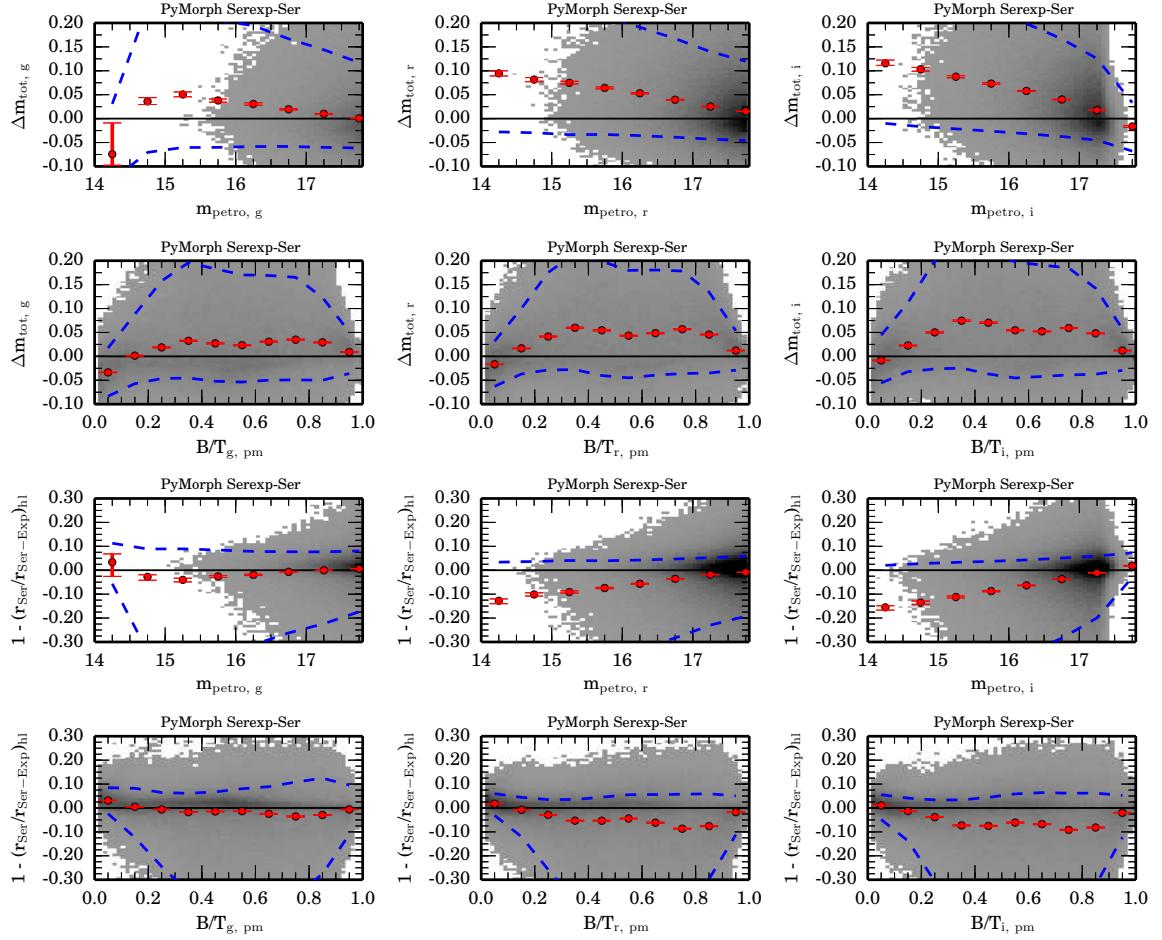
This work favors the **Ser-Exp** model for both deciding on the structure of the galaxies (i. e., one versus two components) and for estimating the total size and magnitude. Figure 6.9 shows comparisons of fitted magnitude (top row) and half-light radius (bottom row) between the **Ser** and **Ser-Exp** 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 errorbars 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 gray-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 in Chapter 4. The bias is caused by inappropriately fitting a one-component model to a more complicated light profile. This type of under-fitting 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).

Figure 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

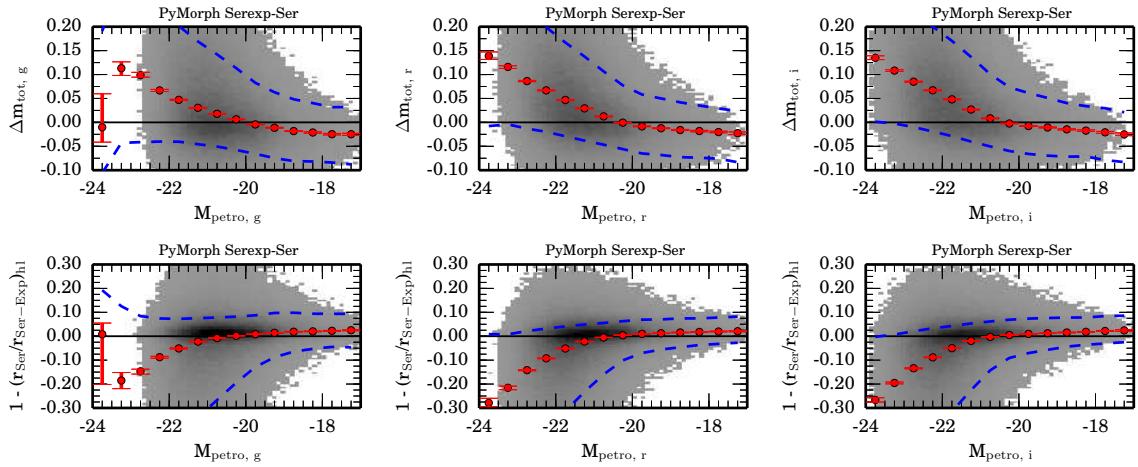
## 6.4 Comparison of the Ser and Ser-Exp models



**Figure 6.9:** 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.

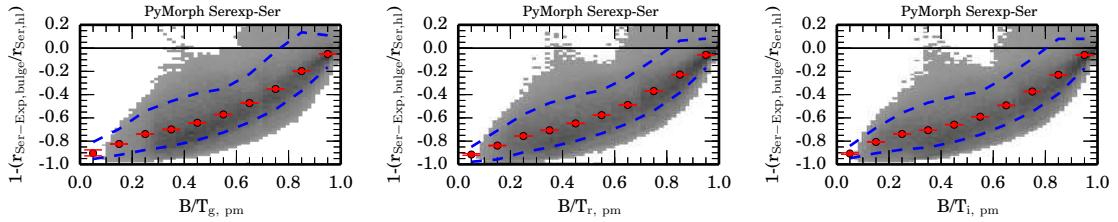
## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---



**Figure 6.10:** 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.

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



**Figure 6.11:** 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 errorbars 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 gray-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.

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

The bottom row of Figure 6.9 shows comparisons of half-light radius between the **Ser** and **Ser-Exp** fits. The plot shows 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.

Figure 6.11 also shows comparisons of bulge radius for the **Ser** and **Ser-Exp** fits. The

## 6. INTERNAL COMPARISONS AND CONSISTENCY CHECKS

---

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 Chapters 7.5 and 8.

# Chapter 7

## Comparisons to literature

In this section, we show several comparisons to the literature. We also direct the reader to the work presented in Chapter 4, 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 catalogs used for comparison in Chapter 7.1. We compare the fits of SDSS, LG12, S11 and Mendel et al. (2014, hereafter Men14) to our fits where appropriate. Chapter 7.2 compares the `deV` fits to SDSS and LG12. Chapter 7.3 compares the `Ser` fits to S11 and LG12 `Ser` fits. Chapter 7.4 compares the `deV-Exp` fits to S11 and LG12 `deV-Exp` fits. Chapter 7.5 compares the `Ser-Exp` fits to S11 `Ser-Exp` fits. Chapter 7.6 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).

### 7.1 External catalogs 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

## 7. COMPARISONS TO LITERATURE

---

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 neighbor galaxy light, which they refer to as ‘SEXTDEBL,’ or SExtractor deblending to separate the light of neighboring 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 analyze 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-analyzed 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 this work. The authors used a sample restricted to more nearby galaxies ( $0.003 < z < 0.05$ ) using SDSS DR8 (Aihara et al. 2011) data. The catalog 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’

---

## 7.1 External catalogs used for comparison and analysis

---

in LG12) fits for 71 825 galaxies from the SDSS. The LG12 terminology is used 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 insignificance 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 colors 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 zeropoint of the magnitudes may vary from the values used here due to the small calibration differences between SDSS DR7 and DR8. Differences in the zeropoint 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 catalog. These galaxies generally have better resolution and are brighter than the full sample of our catalog, so agreement between LG12 and this work only provide

## 7. COMPARISONS TO LITERATURE

---

a lower bound on the bias and scatter of our full catalog. However, this comparison provides a test of the most optimal fitting conditions where resolution effects are less of a concern.

All plots presented here are binned 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, the ‘best-fitting’ models given by S11 and LG12 are treated as the most appropriate models to use in comparison. Therefore, the fits of this work are compared to the best fits of the external works (e.g. `deV-Exp` fits from this work are compared 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.

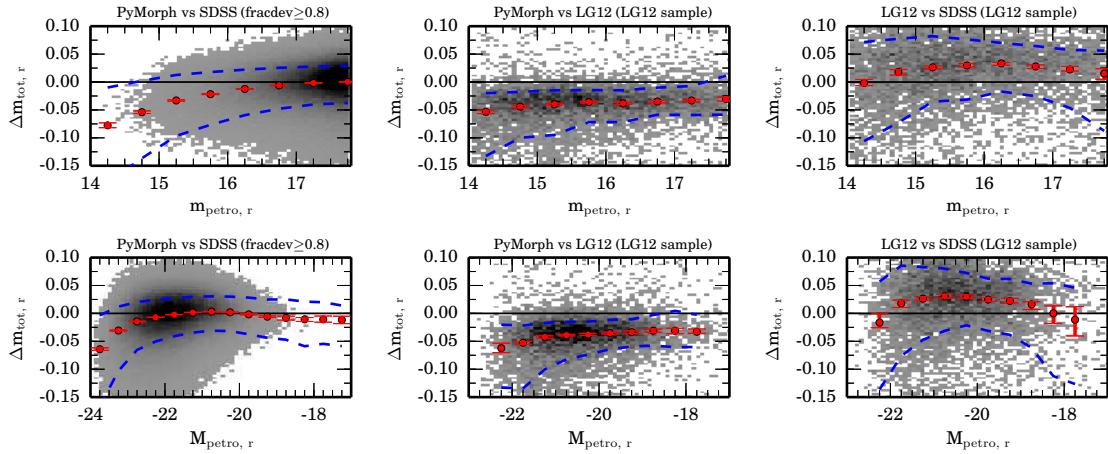
Chapters 7.6 and 8 compare the S11, Men14, and LG12 classifications to the choice of best model based on the flagging presented in Chapter 5.

### 7.2 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  $7 r_{eff}$  to go smoothly to zero beyond  $8 r_{eff}$  and also employ some softening of the profile within  $r = r_{eff}/50$  (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  $8 r_{eff}$  and 93 per cent of the light is contained within  $7 r_{eff}$ . 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  $8 r_{eff}$ , 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

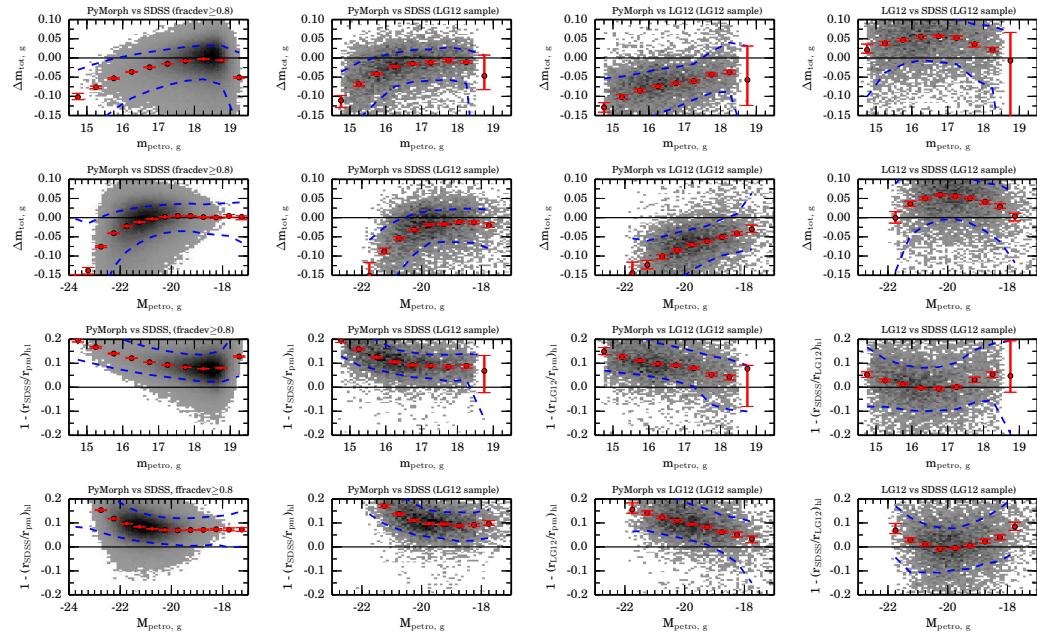
## 7.2 The deV fits



**Figure 7.1:** 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 217131 galaxies in SDSS that are best fitted by the deV profile as identified by  $\text{fracdev}_r \geq 0.8$ . LG12 comparisons use the  $\sim 9000$  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 errorbars 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 gray-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.

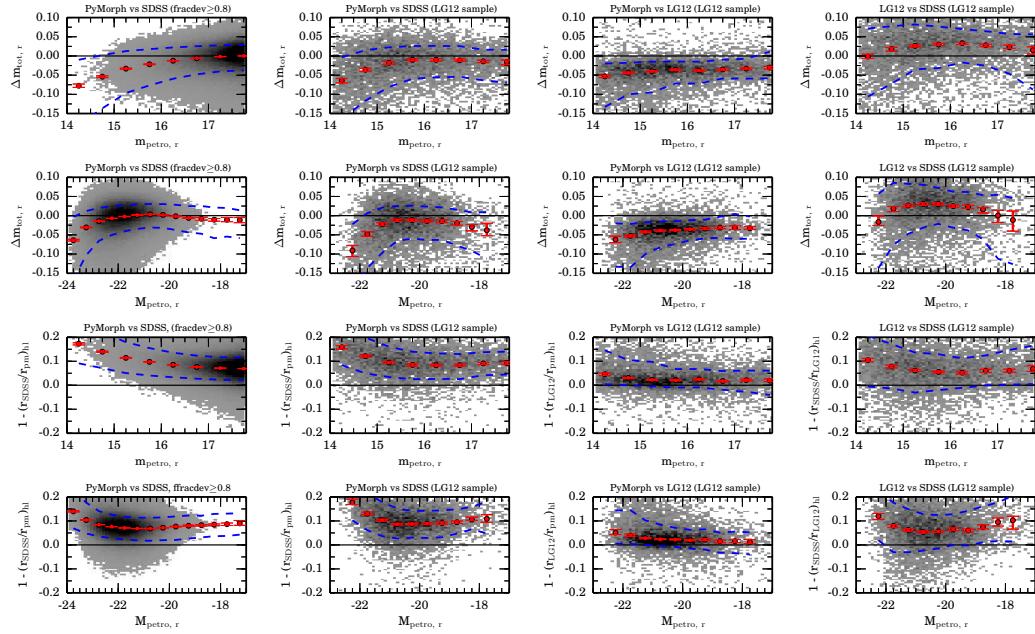
## 7. COMPARISONS TO LITERATURE

---



**Figure 7.2:** Similar to Figure 7.1, but now also including the comparison of SDSS with LG12 galaxies (far right column). The panels show the difference in magnitude and radius as a function of apparent or absolute Petrosian magnitude. The offsets are discussed in the main text.

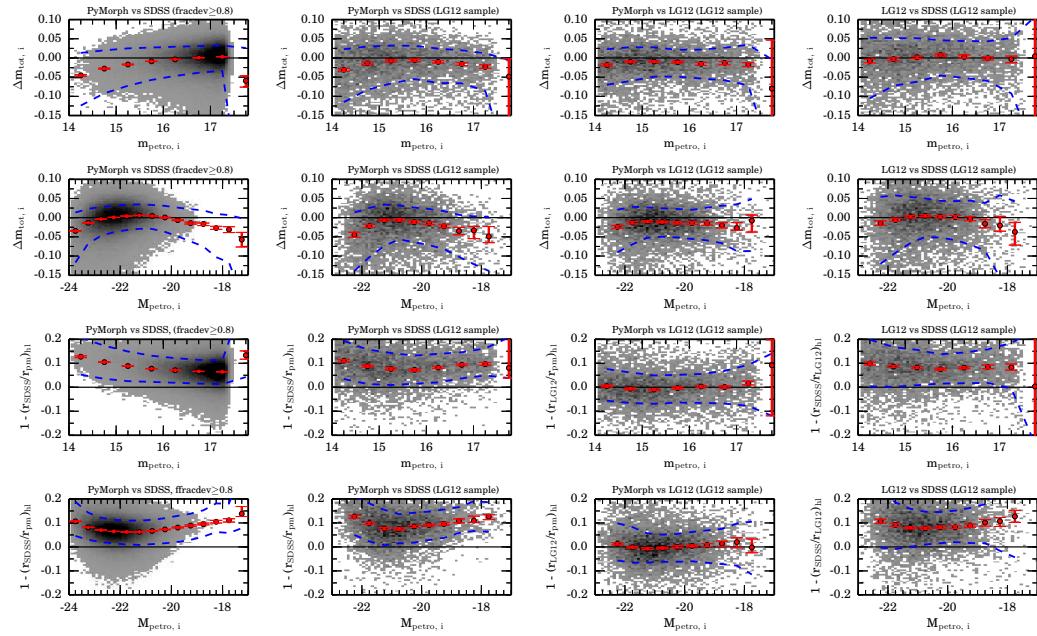
## 7.2 The deV fits



**Figure 7.3:** Similar to Figure 7.1, but now also including the comparison of SDSS with LG12 galaxies (far right column). The panels show the difference in magnitude and radius as a function of apparent or absolute Petrosian magnitude. The offsets are discussed in the main text.

## 7. COMPARISONS TO LITERATURE

---



**Figure 7.4:** Similar to Figure 7.1, but now also including the comparison of SDSS with LG12 galaxies (far right column). The panels show the difference in magnitude and radius as a function of apparent or absolute Petrosian magnitude. The offsets are discussed in the main text.

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.

Figure 7.1 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 217131 galaxies in SDSS that are best fit by the `deV` profile as identified by `fracdev_r` $\geq 0.8$ . 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 errorbars 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 gray-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 Figure 7.1, 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 Figure 7.1). 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

## 7. COMPARISONS TO LITERATURE

---

in the fitted radii. We investigated the source of this offset, but can find no reason for it. Although the zeropoint 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 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.

Figures 7.2, 7.3, and 7.4 show expanded comparisons of  $g$ -,  $r$ -, and  $i$ -bands, respectively. Looking at the third column of each figure, the comparison of LG12 with PYMORPH shows better agreement in the  $i$ -band with larger discrepancies in the bluer  $g$ -band. This effect is explained by the choice of fitting used in LG12. The authors only fitted profiles to the  $r$ -band data. After fitting, all parameters were fixed except for the magnitude of the profile and then the profile was fit to the  $g$ -band and  $i$ -band data. This process biases the results if the half-light radius or other parameters (e.g. ellipticity, Sérsic index, etc.) change as a function of the observed band. The PYMORPH fitting allows for these changes by independently fitting each band. As a result, the PYMORPH parameters change across the bands while the LG12 fitting parameters remain fixed. This effect appears as a trend from over- to under-estimation as is observed in these plots.

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 Figure 7.9 and related discussion). Sky level and the fitted magnitude have been shown to be correlated (see Chapter 4, and references therein for a full discussion). Bias in the sky level may explain the slight differences in magnitudes. In

Chapter 4, overestimates of sky at the level of 0.5 per cent are shown to suppress fitted **Ser** magnitude by  $\approx 0.1$  mag. We discuss the sky brightness further in Chapter 7.3.

### 7.3 The **Ser** fits

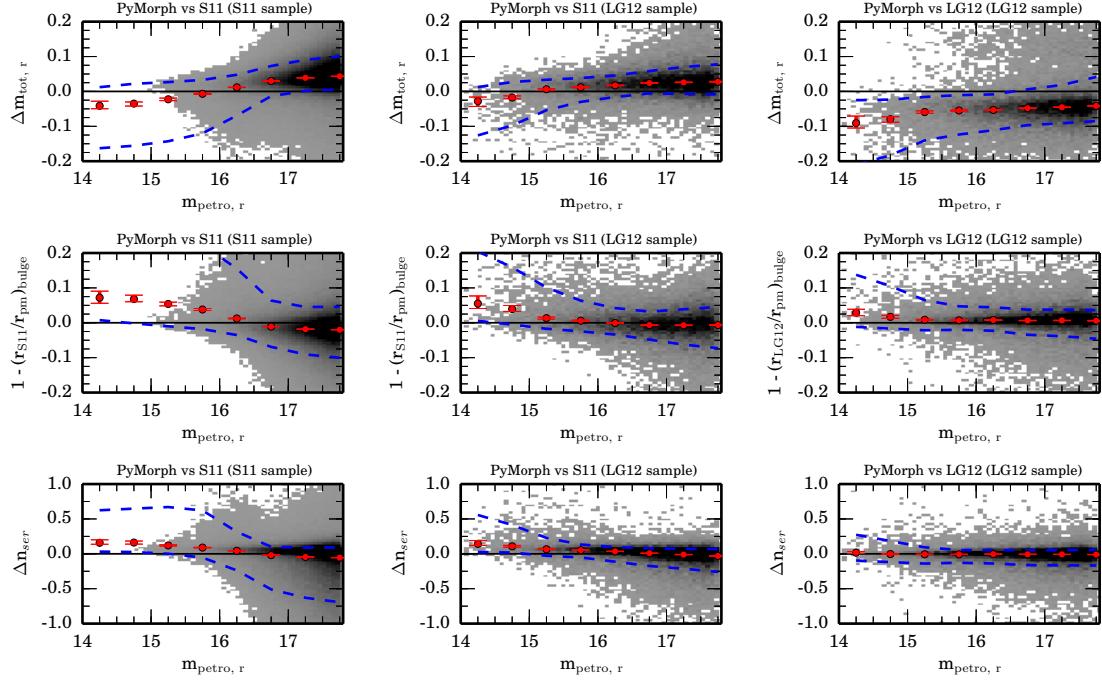
We also compare our **Ser** fits to those of LG12 and S11. Figure 7.5 shows the difference between magnitude (first row), half-light radius (second row), and Sérsic index (third row). We compare PYMORPH and S11 for the full sample (left-hand column), PYMORPH and S11 for the LG12 sample (centre column), and PYMORPH and LG12 (right-hand column) Plots are in a format similar to that of Figure 7.1. For the **Ser** fits we also examine the differences in Sérsic index,  $n_{ser}$ .

For the plots presented in Figure 7.5, we select galaxies in both S11 and LG12 that are ‘best fit’ by a **Ser** model according to S11. We 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 indexes 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 Figure 7.5 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. This trend was also discussed

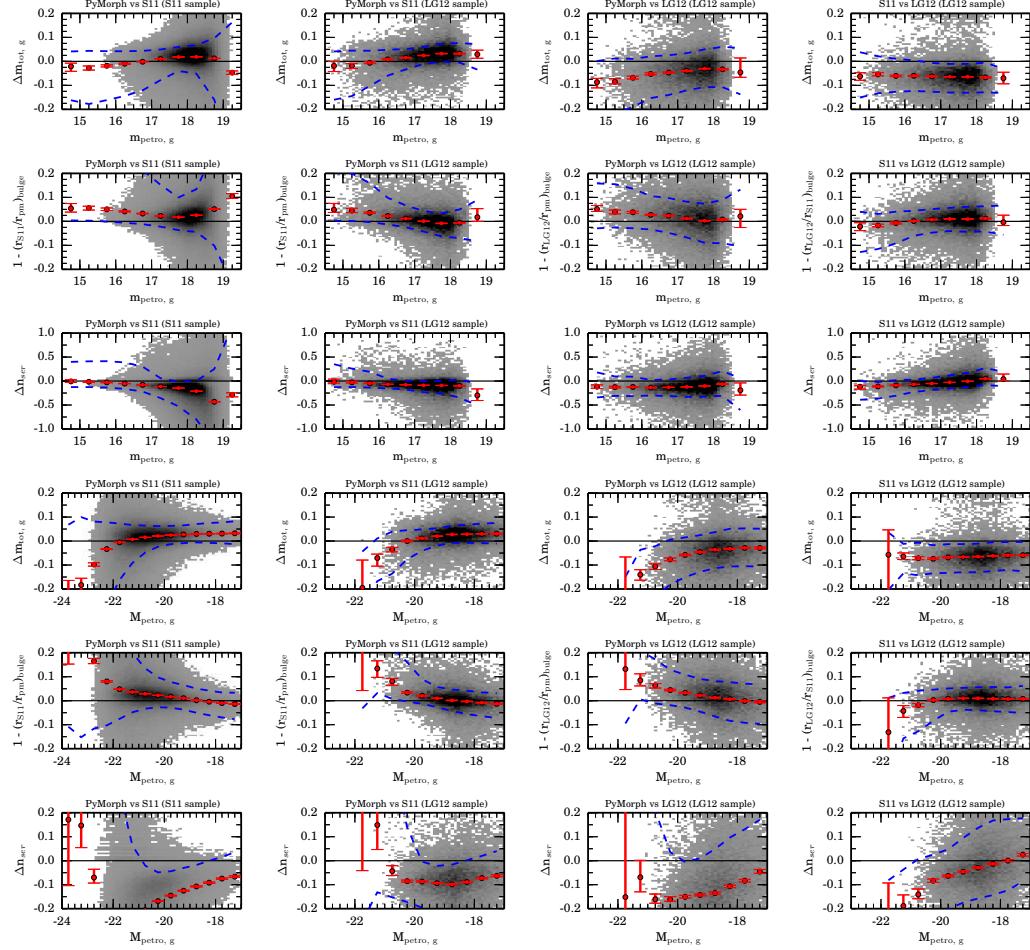
## 7. COMPARISONS TO LITERATURE

---



**Figure 7.5:** 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 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 Figure 7.1. 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.

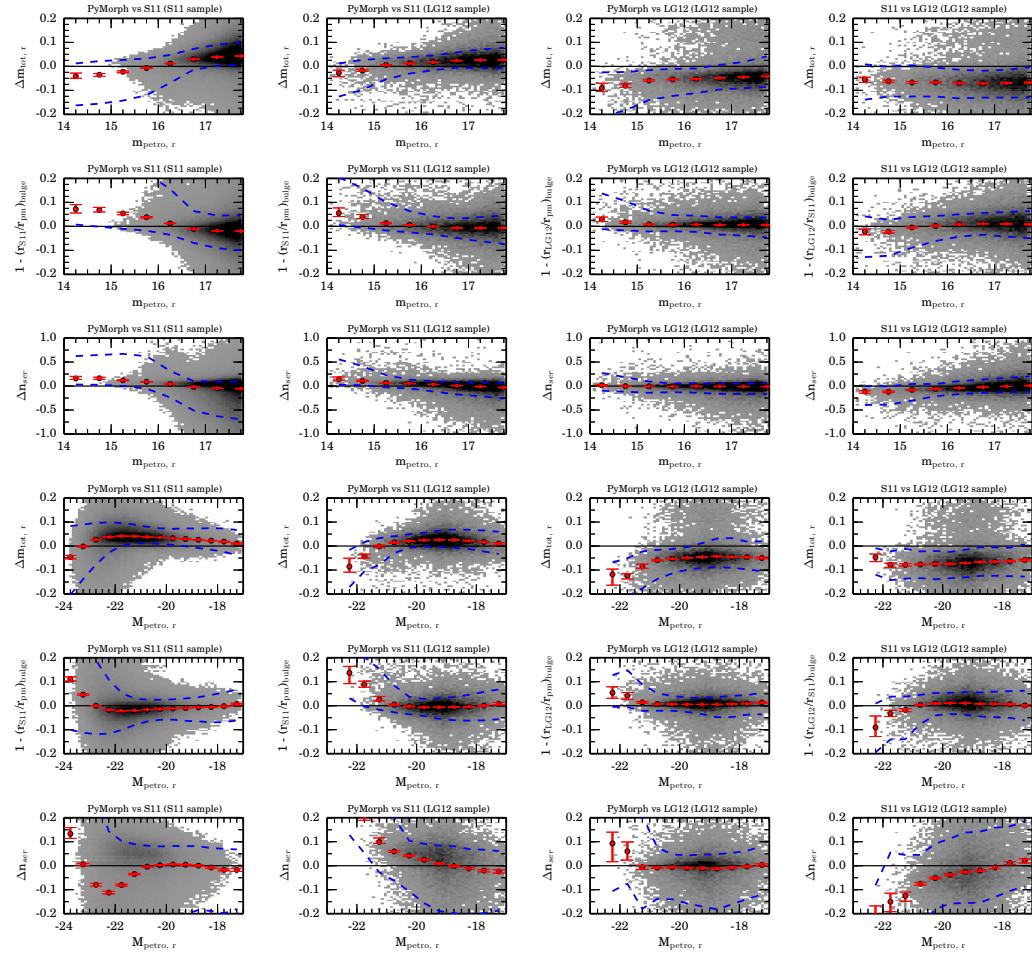
### 7.3 The Ser fits



**Figure 7.6:** Similar to Figure 7.5, but now also including the comparison of S11 with LG12 galaxies (far right column). The panels show the difference in magnitude, half-light radius, and Sérsic index as a function of apparent or absolute Petrosian magnitude. Note the trends in S11 when compared to LG12 or this work. Similar trends in Sérsic index, magnitude, and radius are not present in the comparison of this work with LG12.

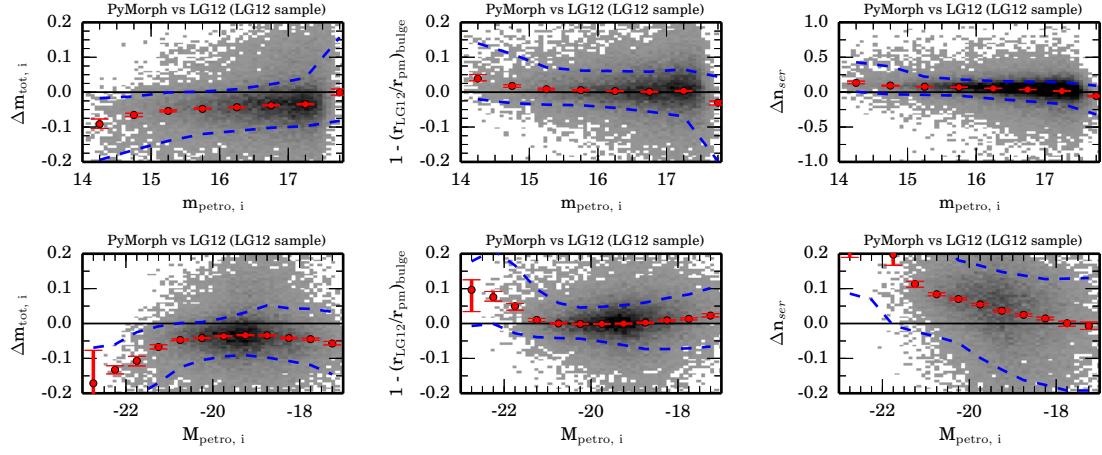
## 7. COMPARISONS TO LITERATURE

---

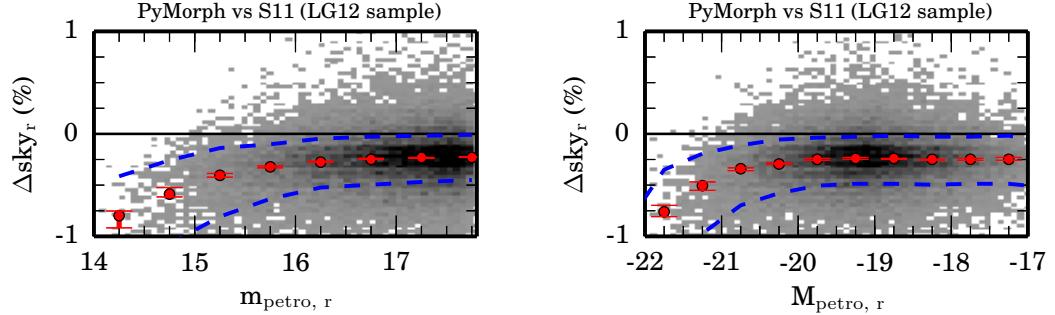


**Figure 7.7:** Similar to Figure 7.5, but now also including the comparison of S11 with LG12 galaxies (far right column). The panels show the difference in magnitude, half-light radius, and Sérsic index as a function of apparent or absolute Petrosian magnitude. Note the trends in S11 when compared to LG12 or this work. Similar trends in Sérsic index, magnitude, and radius are not present in the comparison of this work with LG12.

### 7.3 The Ser fits



**Figure 7.8:** Similar to Figure 7.5, but including only comparisons the LG12 fits. The panels show the difference in magnitude, half-light radius, and Sérsic index as a function of apparent or absolute Petrosian magnitude. S11 *i*-band fits were not carried out by the authors.



**Figure 7.9:** 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.

## 7. COMPARISONS TO LITERATURE

---

in Bernardi et al. (2014), where the authors argue that the behavior of the Sérsic Index is consistent with systematic errors in the Sérsic index due to over subtracting sky.

Figures 7.6, 7.7, and 7.8 show expanded comparisons of *g*-, *r*-, and *i*-bands, respectively. Figure 7.8, showing the *i*-band comparisons is presented in slightly different format because the S11 catalog does not provide *i*-band fits and no comparisons can be made as a result. This figure is equivalent to the third column of the *g*- and *r*-band plots. Looking at the third column of each figure, the comparison of LG12 with PYMORPH shows similar systematic trends across the bands as the **Ser** fits. The trend is most apparent in the Sérsic index where better agreement is apparent in the *r*-band with larger discrepancies in the bluer *g*-band and redder *i*-band.

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. Figure 7.9 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 Chapter 4 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 Chapter 4 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 neighboring 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 neighbor 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.

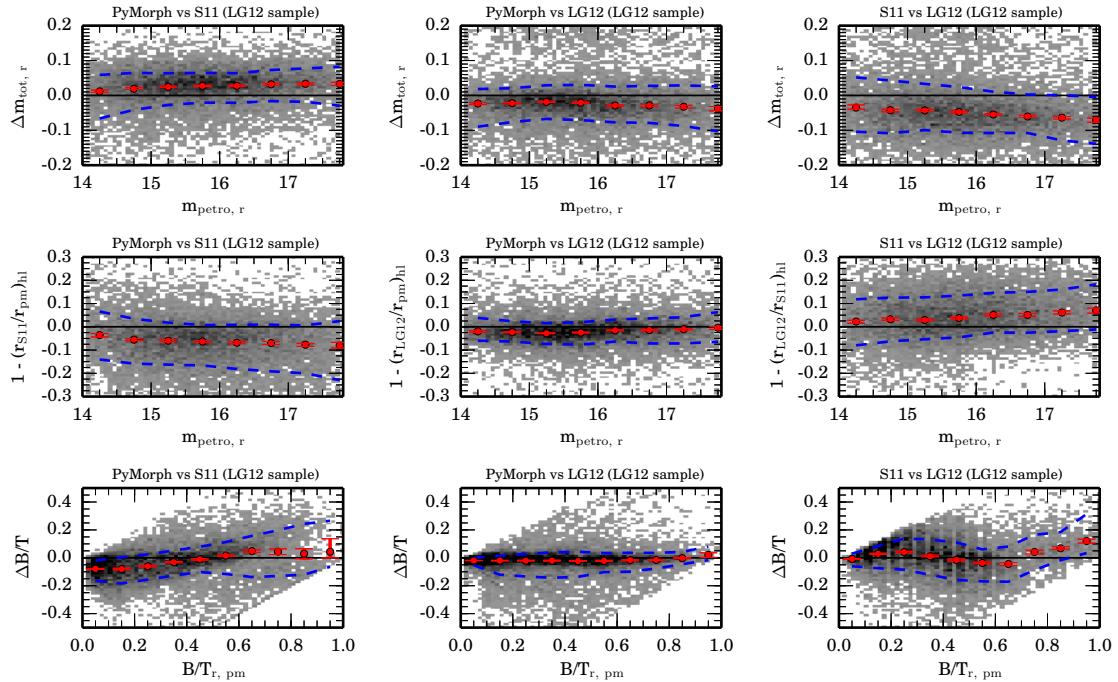
We note that two works have already shown that our measurements should be preferred to those of S11. First, Kravtsov et al. (2014) performed detailed fits of approximately 10 BCGs and found that their measurements agree more with our measurements than with the measurements of S11. Secondly, we direct the reader to appendix A 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.

## 7.4 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. Figure 7.10 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 half

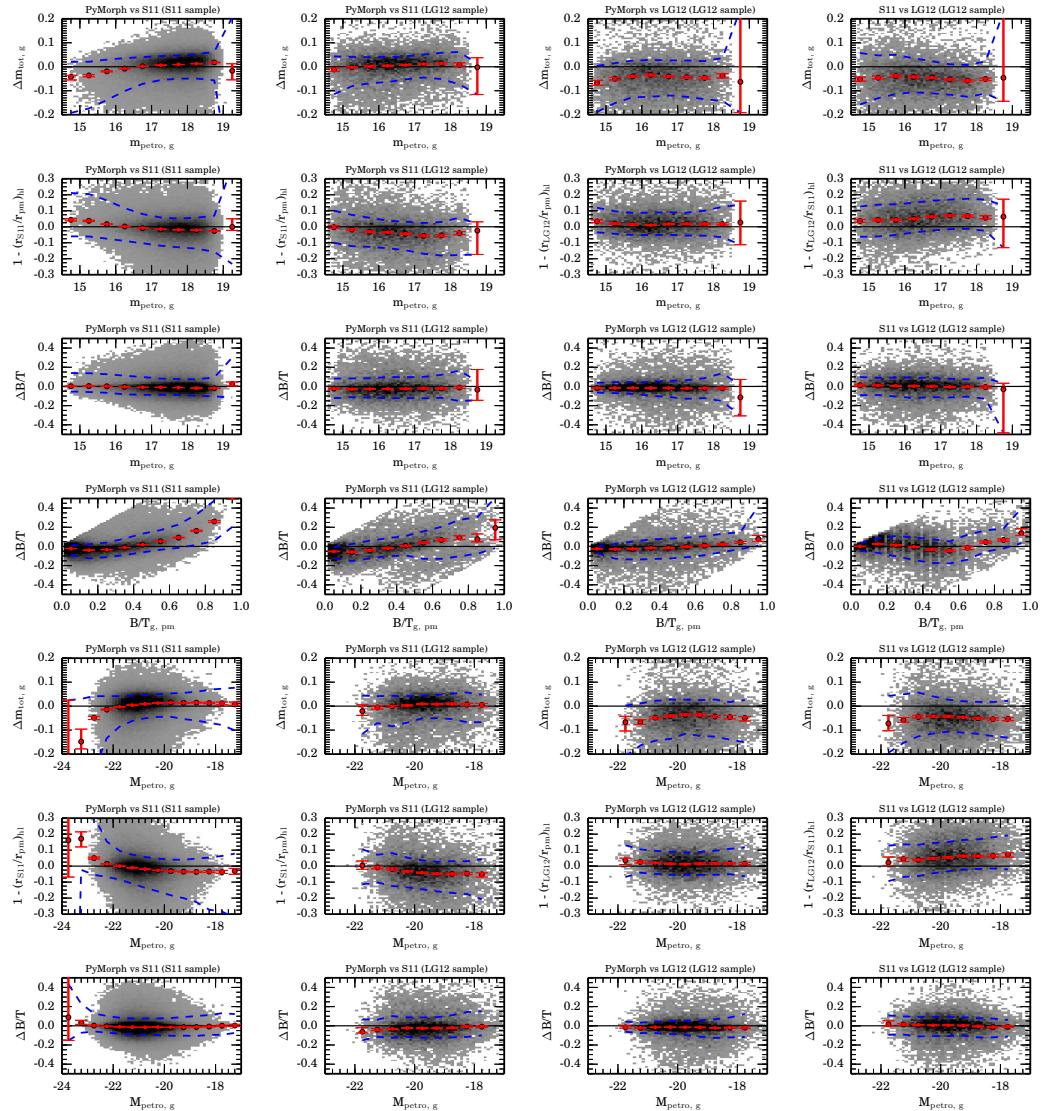
## 7. COMPARISONS TO LITERATURE

---



**Figure 7.10:** The difference between the PYMORPH  $\text{deV-Exp}$  magnitude, half-light radius, and B/T and those of LG12  $\text{deV-Exp}$  and S11  $\text{deV-Exp}$  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 Figure 7.1. S11 B/T has larger scatter with respect to either PYMORPH or LG12.

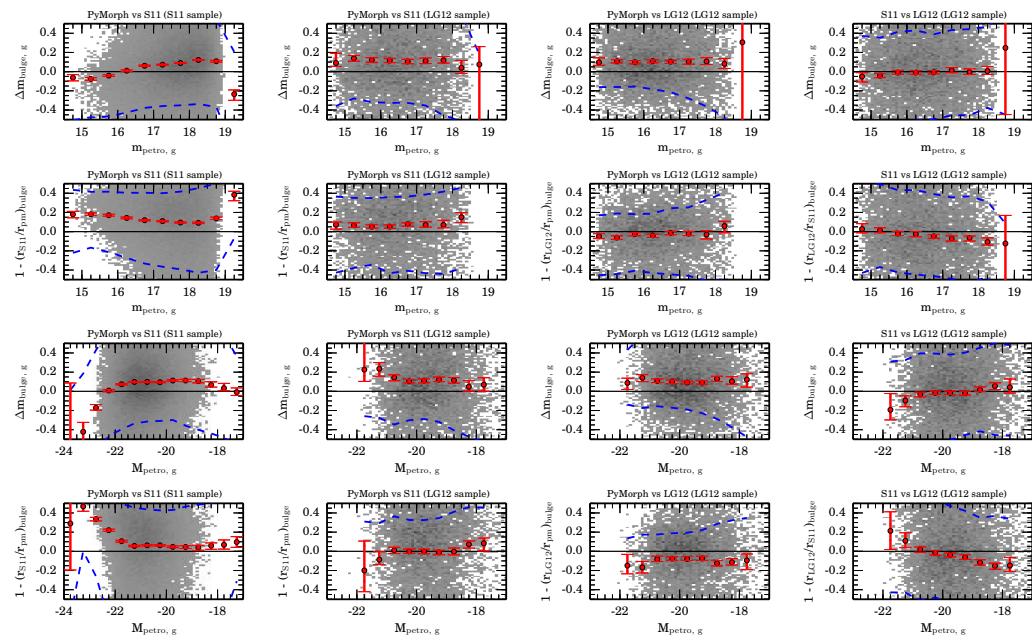
## 7.4 The deV-Exp fits



**Figure 7.11:** Comparison of the  $g$ -band deV-Exp total magnitude, half-light radius, and B/T versus apparent Petrosian magnitude (top 3 rows), B/T (row 4), and absolute Petrosian magnitude (bottom 3 rows) for this work, S11, and LG12. Plot format is the same as Figure 7.10.

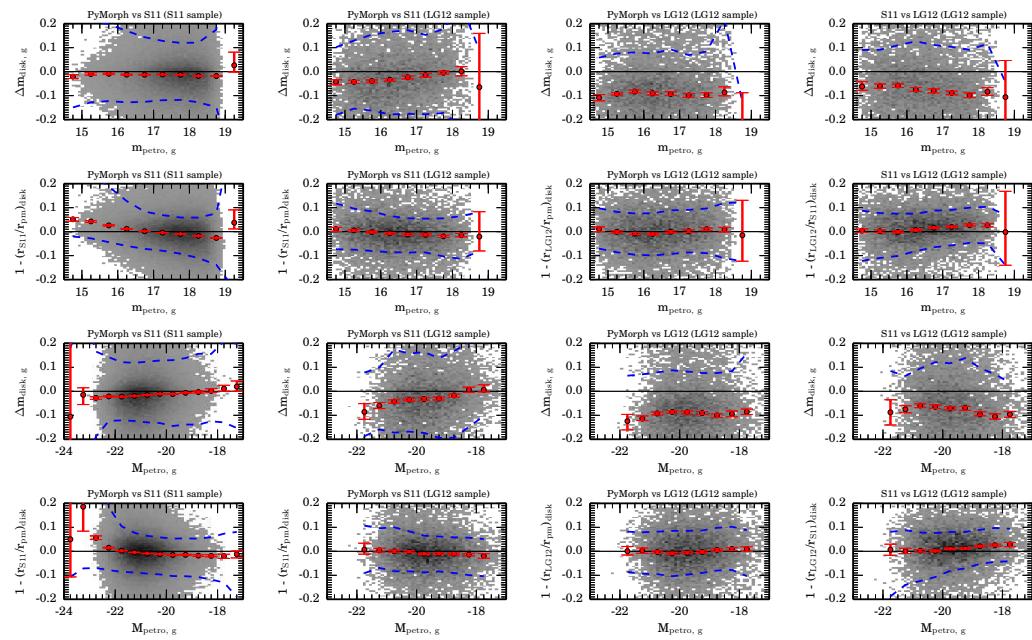
## 7. COMPARISONS TO LITERATURE

---



**Figure 7.12:** Comparison of the deV-Exp bulge component fits of this work, S11, and LG12 in the *g*-band.

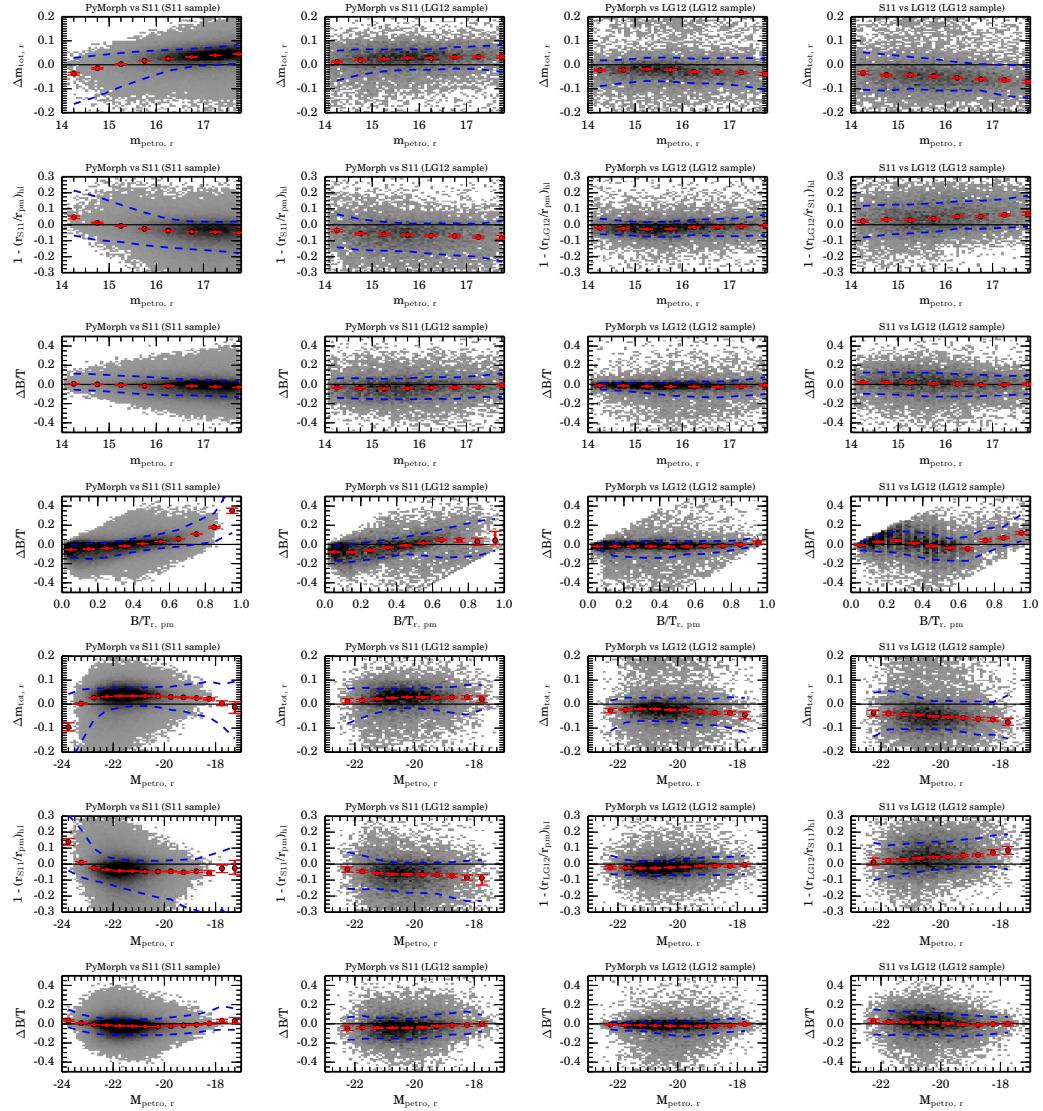
## 7.4 The deV-Exp fits



**Figure 7.13:** Comparison of the deV-Exp disk component fits of this work, S11, and LG12 in the *g*-band.

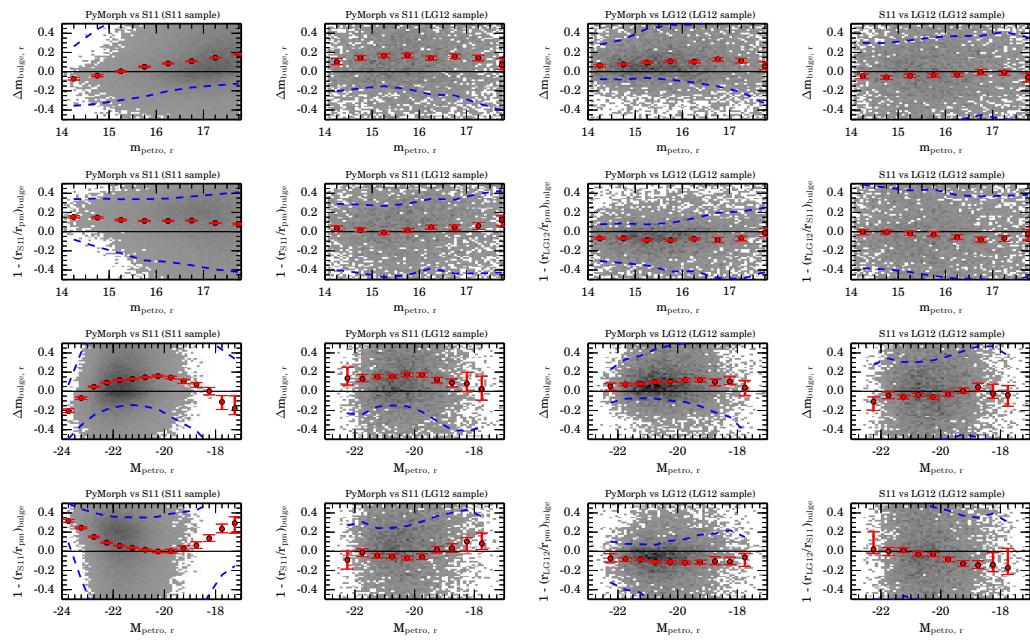
## 7. COMPARISONS TO LITERATURE

---



**Figure 7.14:** Comparison of the  $r$ -band deV-Exp total magnitude, half-light radius, and B/T versus apparent Petrosian magnitude (top 3 rows), B/T (row 4), and absolute Petrosian magnitude (bottom 3 rows) for this work, S11, and LG12. Plot format is the same as Figure 7.10.

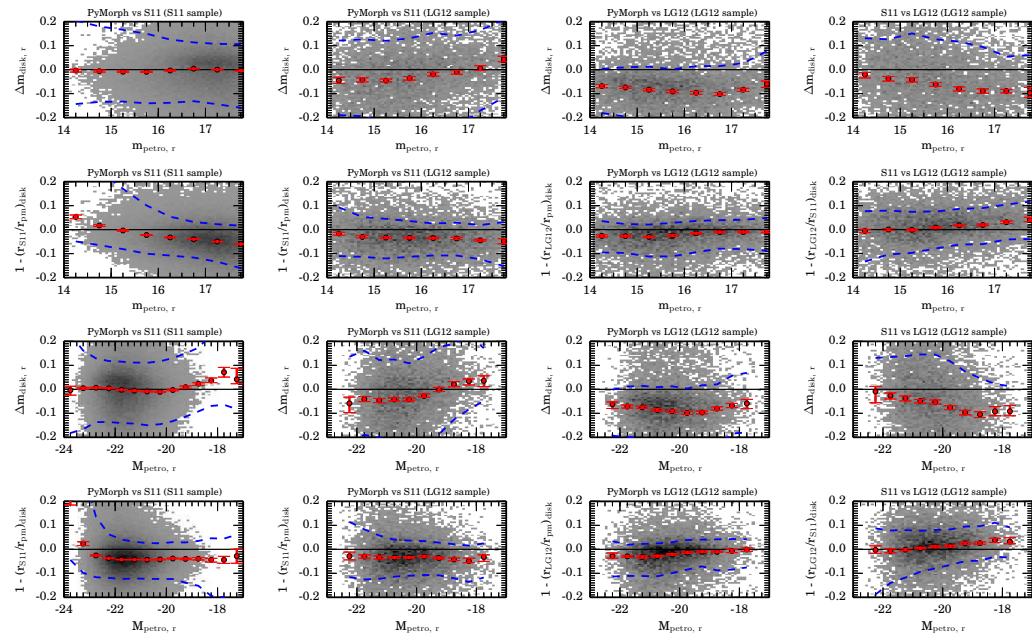
## 7.4 The deV-Exp fits



**Figure 7.15:** Comparison of the deV-Exp bulge component fits of this work, S11, and LG12 in the  $r$ -band.

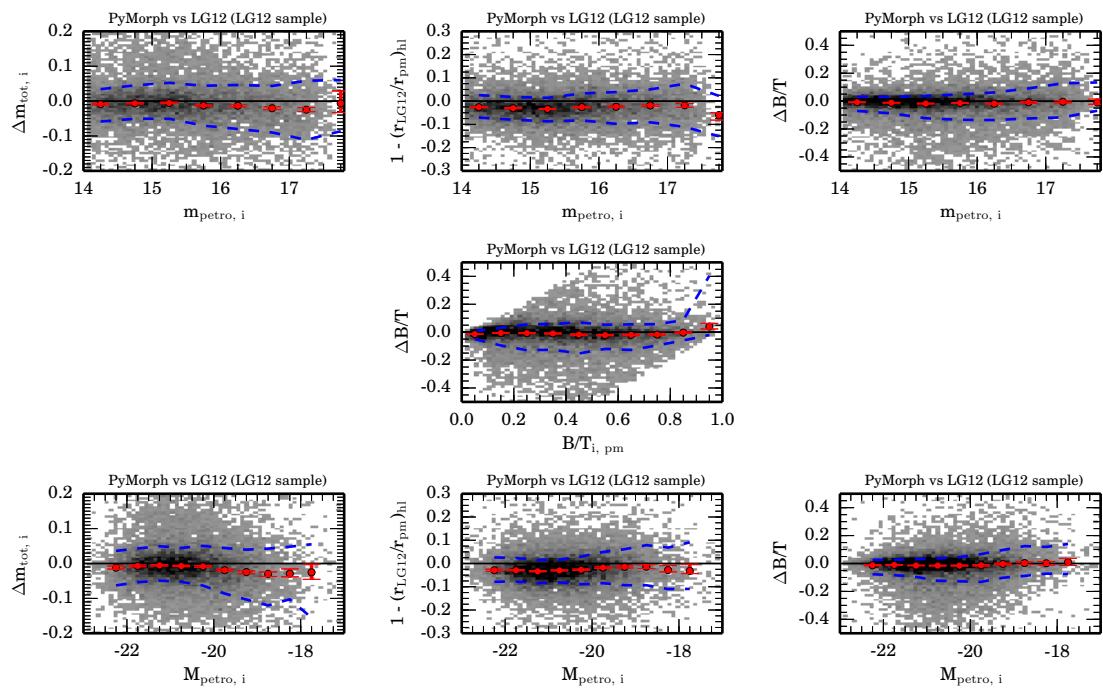
## 7. COMPARISONS TO LITERATURE

---



**Figure 7.16:** Comparison of the deV-Exp disk component fits of this work, S11, and LG12 in the  $r$ -band.

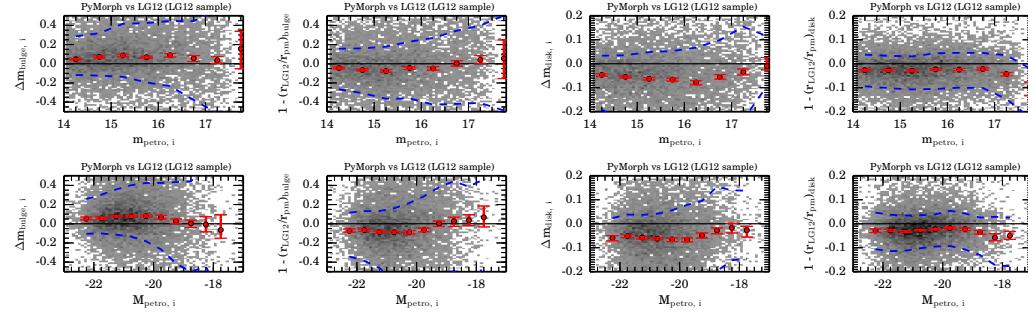
## 7.4 The deV-Exp fits



**Figure 7.17:** Comparison of the  $i$ -band deV-Exp total magnitude, half-light radius, and B/T versus apparent Petrosian magnitude (top row), B/T (middle row), and absolute Petrosian magnitude (bottom row) for this work and LG12. Plot format is the same as Figure 7.10.

## 7. COMPARISONS TO LITERATURE

---



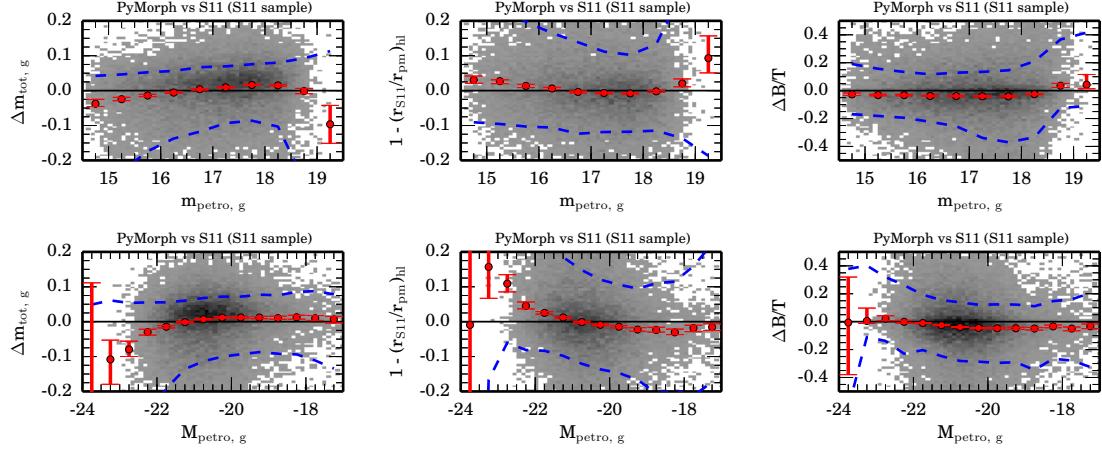
**Figure 7.18:** Comparison of the deV-Exp component fits of this work and LG12 in the  $i$ -band. The format of the plot is the similar to as Figure 7.10. The panels show the difference in magnitude and half-light radius as a function of apparent and absolute Petrosian magnitude for the bulge, and disc. The scatter is quite broad, making it difficult to draw any conclusions. Again the  $i$ -band comparisons to S11 are omitted since no S11 fits were available in this band.

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 sub components 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.

Figures 7.11, 7.14, and 7.17 show expanded comparisons of the total magnitude, half-light radius, and B/T in the  $g$ -,  $r$ -, and  $i$ -bands, respectively. Figures 7.12, 7.13, 7.15, 7.16, and 7.18 show expanded comparisons of the bulge and disk magnitude and half-light radius in the  $g$ -,  $r$ -, and  $i$ -bands.

## 7.5 The Ser-Exp fits



**Figure 7.19:** Comparison of the **Ser-Exp** fits of this work with S11 in the *g*-band. The format of the plot is similar to Figure 7.21. The panels show the difference in total magnitude, half-light radius, and B/T (left, center, and right columns, respectively) as a function of apparent Petrosian magnitude (row 1) and as a function of absolute Petrosian magnitude (row 2).

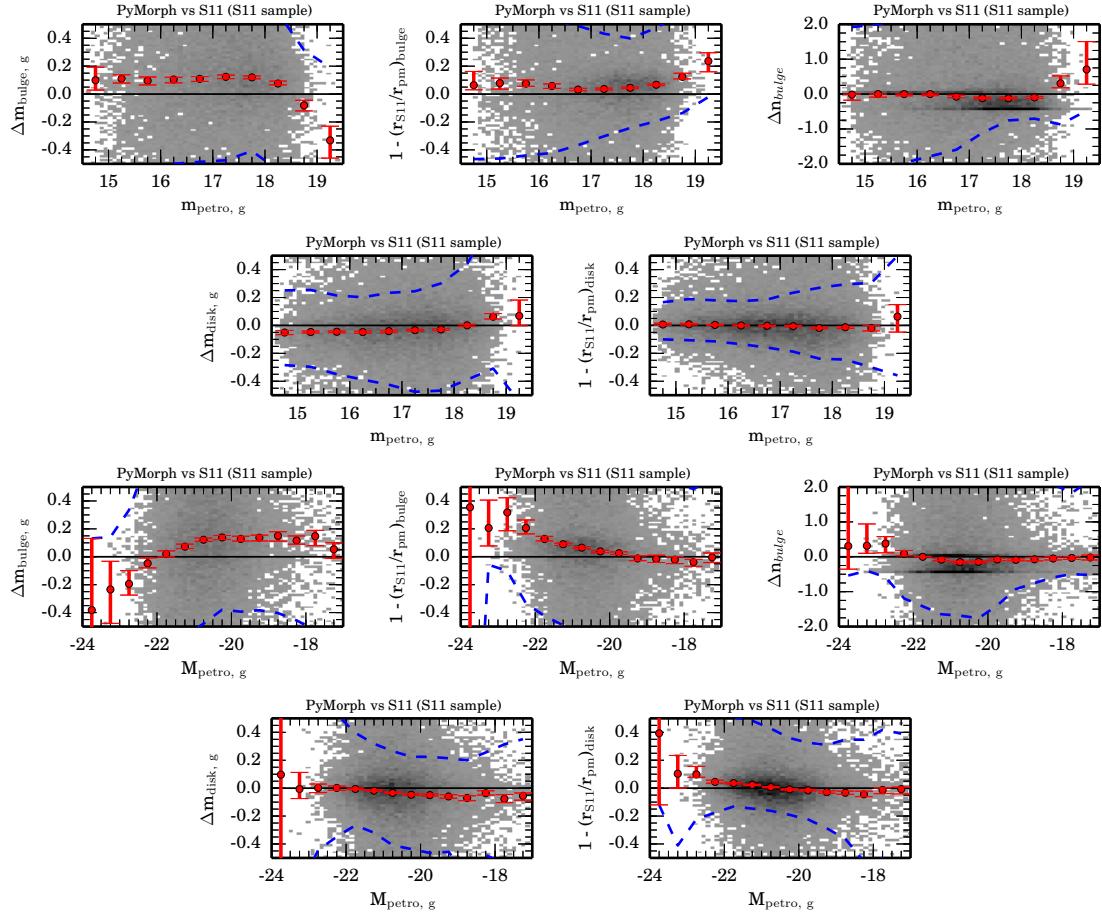
Figure 7.8, showing the *i*-band comparisons is presented in slightly different format because the S11 catalog does not provide *i*-band fits and no comparisons can be made as a result. This figure is equivalent to the third column of the *g*- and *r*-band plots. Looking at the third column of each figure, the comparison of LG12 with PYMORPH shows similar systematic trends across the bands as the **Ser** fits. The trend is most apparent in the Sérsic index where better agreement is apparent in the *r*-band with larger discrepancies in the bluer *g*-band and redder *i*-band.

## 7.5 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 favors a traditional  $n = 4$  bulge.

## 7. COMPARISONS TO LITERATURE

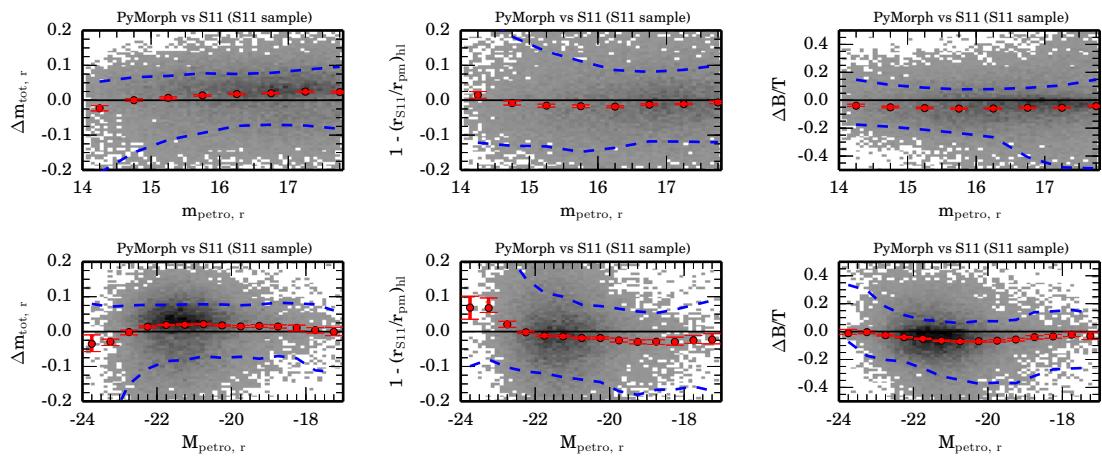
---



**Figure 7.20:** Comparison of the Ser-Exp component fits of this work and S11 in the  $g$ -band. The panels show the difference in magnitude, half-light radius, and (bulge) Sérsic index as a function of apparent and absolute Petrosian magnitude for the bulge, and disc. The scatter is quite broad, making it difficult to draw any conclusions.

## 7.5 The Ser-Exp fits

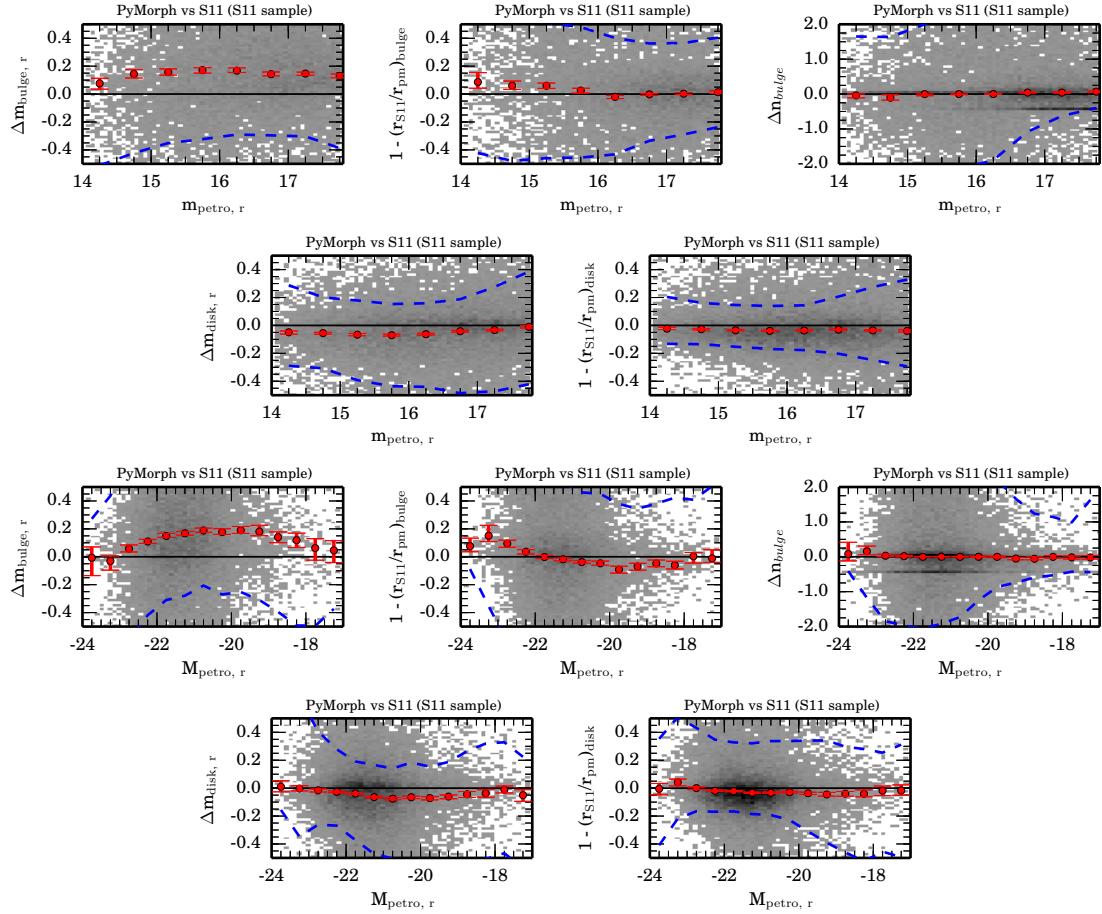
---



**Figure 7.21:** Comparison of the Ser-Exp fits of this work with S11 in the  $r$ -band. The format of the plot is similar to Figure 7.1. The panels show the difference in total magnitude, half-light radius, and B/T (left, center, and right columns, respectively) as a function of apparent Petrosian magnitude (row 1) and as a function of absolute Petrosian magnitude (row 2). The Ser-Exp fits have wider scatter than the analogous parameters in the deV-Exp fits.

## 7. COMPARISONS TO LITERATURE

---



**Figure 7.22:** Comparison of the Ser-Exp component fits of this work and S11 in the  $r$ -band. The panels show the difference in magnitude, half-light radius, and (bulge) Sérsic index as a function of apparent and absolute Petrosian magnitude for the bulge, and disc. The scatter is quite broad, making it difficult to draw any conclusions.

## 7.6 Comparing the preferred models

---

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.

Figures 7.19 and 7.21 show 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  $P_{pS}$  and  $P_{n4}$  probabilities, indicating that the **Ser-Exp** fit is significantly better than either the **deV-Exp** or **Ser** fits. Figures 7.20 and 7.22 show a comparison of the component fitting parameters.

The total values are in agreement with wider scatter compared to that seen in Figures 7.5 and 7.10. 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 Chapter 8 which support the accuracy of our **Ser-Exp** fits.

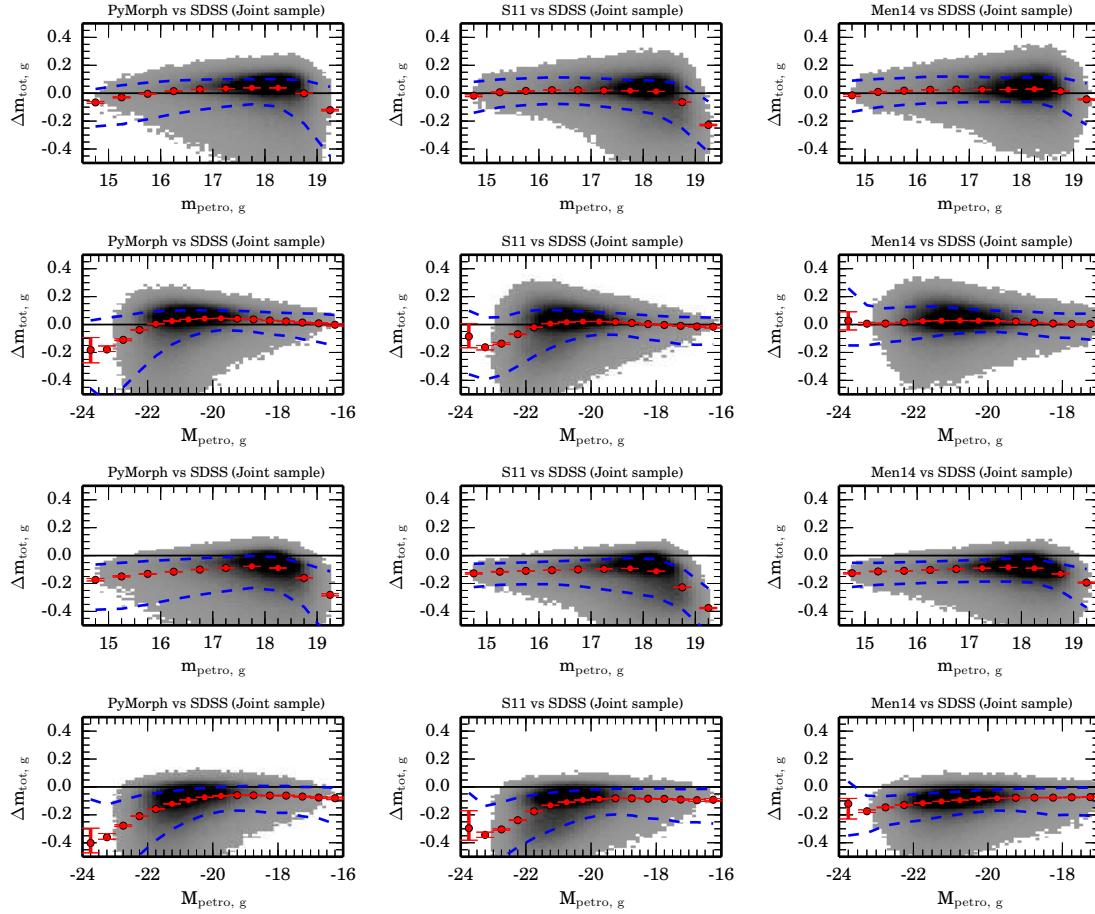
## 7.6 Comparing the preferred models

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** catalog as follows: first, galaxies with flag bit 10 or 14 set (galaxies shown as ‘2com’ and ‘prob 2com’ in Figure 8.1) are given the total magnitude measured by fitting the

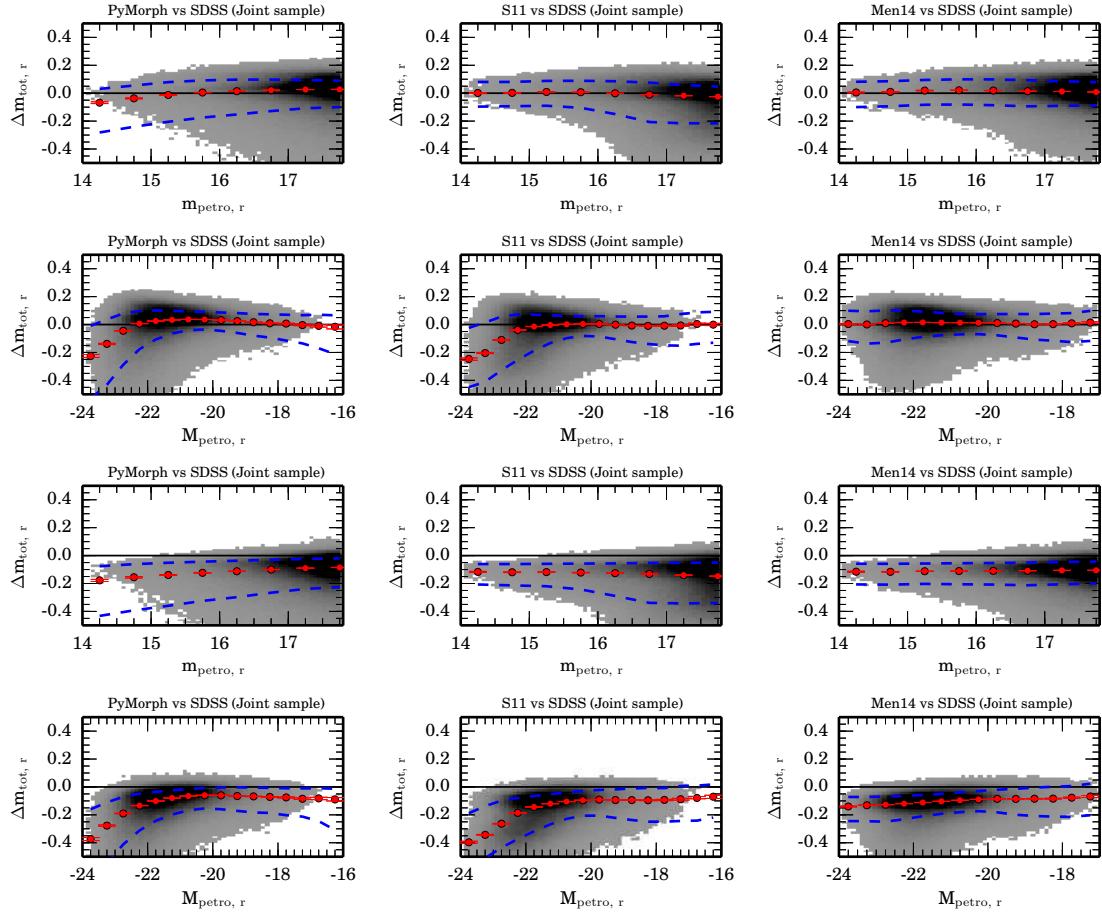
## 7. COMPARISONS TO LITERATURE

---



**Figure 7.23:** The difference between  $g$ -band SDSS cModel magnitude (first two rows)/Petrosian magnitude (last two rows) compared to the magnitude fit by PYMORPH, S11, and Men14 for the galaxies appearing in all three catalogs. Galaxies classified as an unknown profile in Men14 (`Proftype=4`) are excluded from the plots. This removes approximately 10 per cent of our original catalog. We also exclude any failed fits from our catalog.

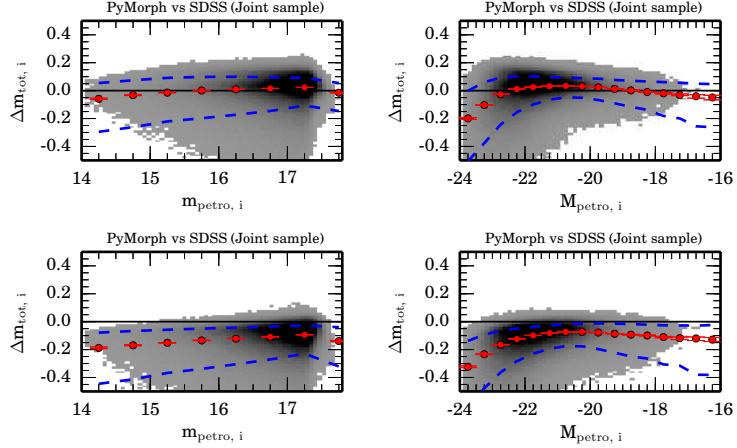
## 7.6 Comparing the preferred models



**Figure 7.24:** The difference between  $r$ -band SDSS cModel magnitude (first two rows)/Petrosian magnitude (last two rows) compared to the magnitude fit by PYMORPH, S11, and Men14 for the galaxies appearing in all three catalogs. Galaxies classified as an unknown profile in Men14 (`Proftype=4`) are excluded from the plots. This removes approximately 10 per cent of our original catalog. We also exclude any failed fits from our catalog.

## 7. COMPARISONS TO LITERATURE

---



**Figure 7.25:** The difference between  $i$ -band SDSS cModel magnitude (first row)/Petrosian magnitude (second row) compared to the magnitude fit by PYMORPH. S11 and Men14 are not compared since they do not do  $i$ -band fits. We exclude any failed fits from our catalog.

**Ser-Exp** model. Galaxies with flag bits 1 or 4 set (i.e., ‘bulge’ or pure ‘disc’ in Figure 8.1) 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

## 7.6 Comparing the preferred models

---

data sets for comparison.

Figure 7.24 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 `CModel` magnitudes are a simplified estimate for the total light of a galaxy surface brightness distribution for galaxies that follow neither a pure exponential disk (`Exp`) nor a deVaucouleur's (`deV`) profile (Bernardi et al. 2007). The `CModel` fitting proceeds by separately fitting a `Exp` profile and a `deV` profile to the total light profile of the galaxy. These separately generated fits are then jointly fit, adjusting the relative brightness contributed by each model but keeping the other parameters (i.e., radius, axis-ratio, position angle, and center) fixed. This is achieved by fitting the parameter `fracdev` in the equation:

$$F_{tot} = \text{fracdev} F_{\text{deV}}(r, \phi) + (1 - \text{fracdev}) F_{\text{Exp}}(r, \phi) \quad (7.1)$$

where  $F_{\text{deV}}(r, \phi)$  and  $F_{\text{Exp}}(r, \phi)$  are the flux at a given distance ( $r$ ) and position angle ( $\phi$ ) for each of the previously fit `deV` and `Exp` profiles. For intermediate types (i.e., not pure `deV` or pure `Exp` galaxies) the separate `CModel` components (`deV` and `Exp`) will not be directly related to the bulge and disk components of the galaxy, but will look more like the size and shape of the total galaxy profile.

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 Figure 7.1.

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

## 7. COMPARISONS TO LITERATURE

---

and stellar mass functions can cause a substantial underestimate of the bright end of the LF (see also Chapter 10). 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 favoring 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.

Figure 7.24 also 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.

Figures 7.23 and 7.25 show similar comparisons of the *g*- and *i*-band fits. Behavior of the PYMORPH fits is largely the same as in the *r*-band data.

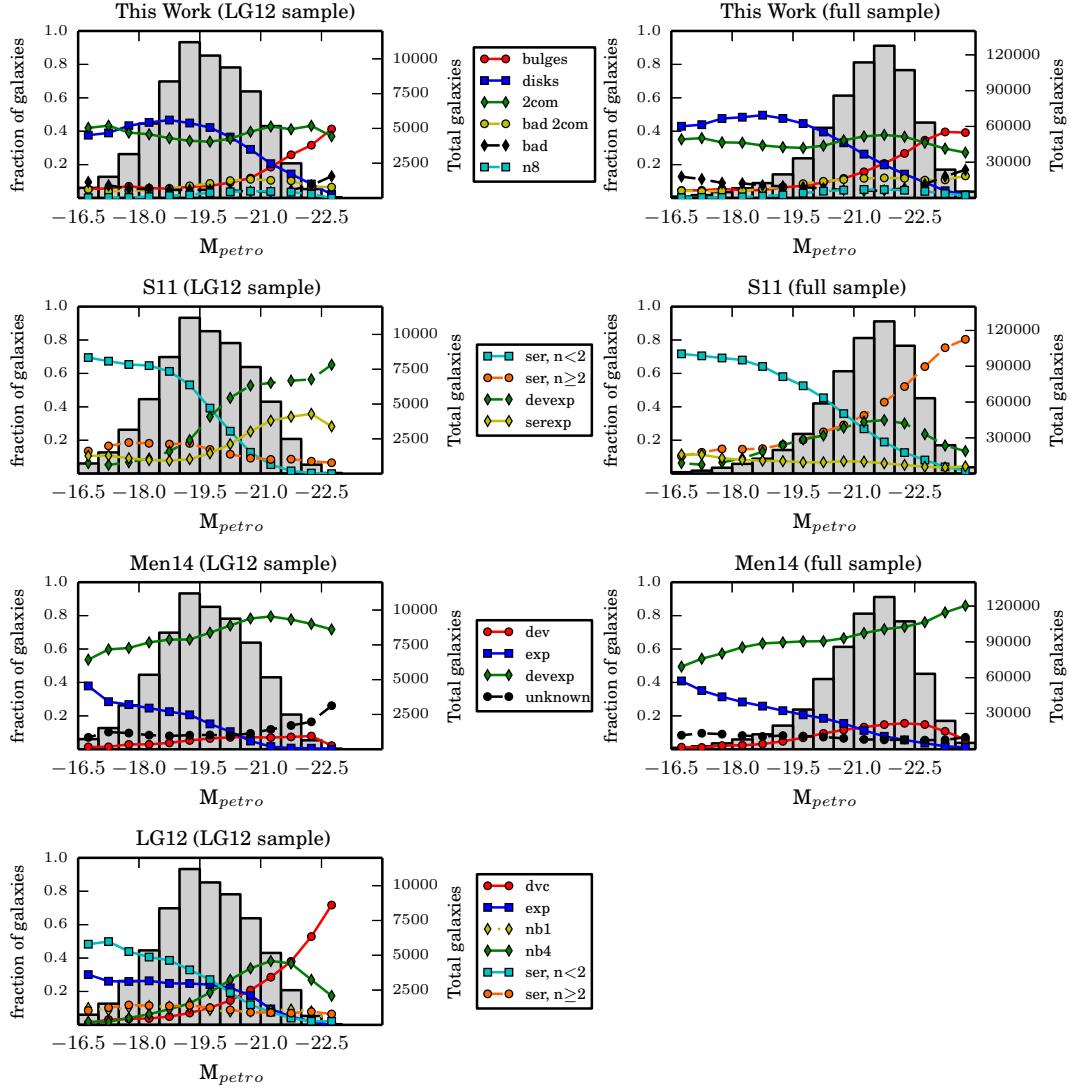
## Chapter 8

# Comparisons using morphological information

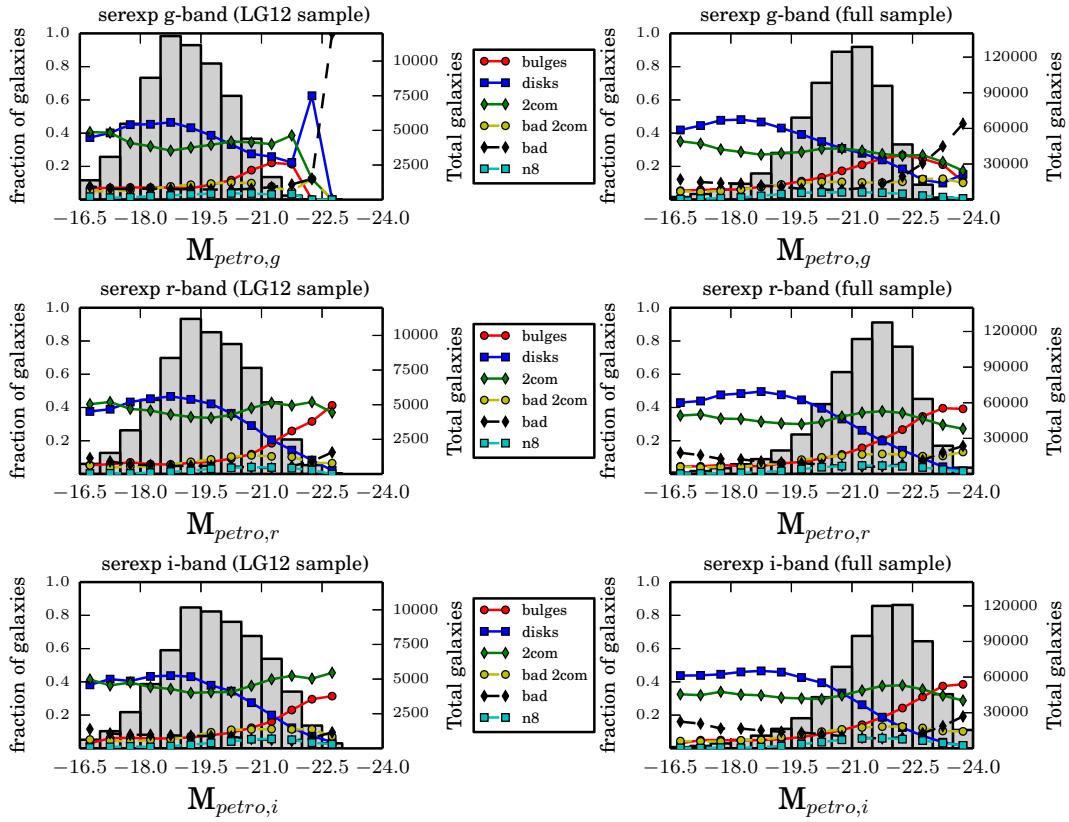
Finally, we examine our selection of bulge, disc and two-component galaxies using morphological information from other catalogs. The work of Nair and Abraham (2010) gives a detailed visual morphological classification of a set of about 10 000 nearby galaxies. The GZ2 project (Lintott et al. 2011, 2008) provides another morphological classification of nearly half of our sample. These catalogs give morphological classifications not dependent on fitting and provide another test of our profile-based flagging. Chapter 8.1 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 and Abraham (2010). Chapter 8.3 examines a few internal checks, including B/T and axis ratios separated by  $T$ -type. Chapter 8.4 discusses possible effects of bars on the fitting based on barred galaxies identified in GZ2.

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---



**Figure 8.1:** Type comparisons as a function of absolute Petrosian magnitude for this work (top row), S11 (second row), Men14 (third row), and LG12 (fourth row) for the smaller, low- $z$  LG12 sample (left) and the full sample (right). Plot design is the same as Figure 6.1.



**Figure 8.2:** Type comparisons as a function of absolute Petrosian magnitude for the g-, r-, and i-bands of this work. The left column shows the classifications for the smaller, low-z LG12. The right 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 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.

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---

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

In this section, we compare our selection of bulge, disc, and two-component galaxies presented in Chapter 5 with S11, LG12, and Men14, all of which include a ‘preferred’ model in their catalogs. Figure 8.1 shows the percent 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 catalog 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 percent) 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 behavior is visible in both samples and shows that the flag-based model 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 catalog (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

## 8.1 Preferred models as an indicator of morphological classification

---

actual morphology of the galaxy. However, the general behavior 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 neighborhood 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 behavior 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 color 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 Chapter 3.2, we find similar behavior 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 Chapter 4. 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.

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---

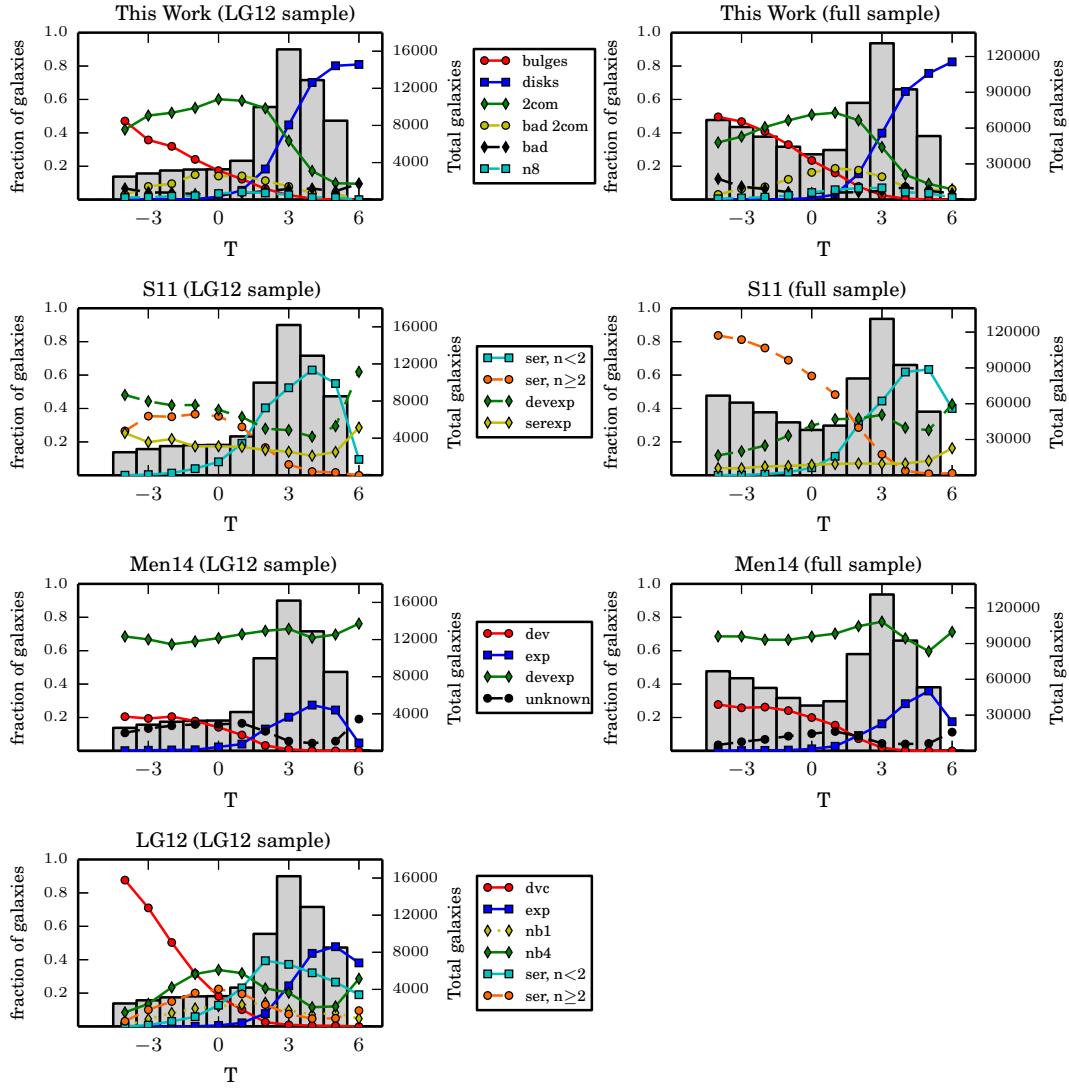
Figure 8.2 shows the percent of model types as a function of absolute Petrosian magnitude for two samples. The left column shows the results for the low-z sample of LG12. The right column shows the results for the entire spectroscopic catalog used in this work. The panels show the models fit to the g-band (top row), r-band (second row), and i-band (third row). Model types 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 axis of the plot.

The fits of this work show expected trends with almost no bulges (only a few percent) and a mixture of disk 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. In the *g*-band, the failure rate is substantially higher at the brightest magnitudes. However, the entire *g*-band sample is dimmer and galaxies as bright as -24 in the *g*-band are extremely rare, making it more likely that galaxies as bright as -24 are indeed incorrectly fit galaxies. The plots again demonstrate the stability of the flagging across the bands. Effects of S/N or resolution would likely cause a different appearance in the model selection for the different bands.

### 8.2 Model selection as a function of *T*-type

We assign a *T*-type to each galaxy in the catalog 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 color, 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.

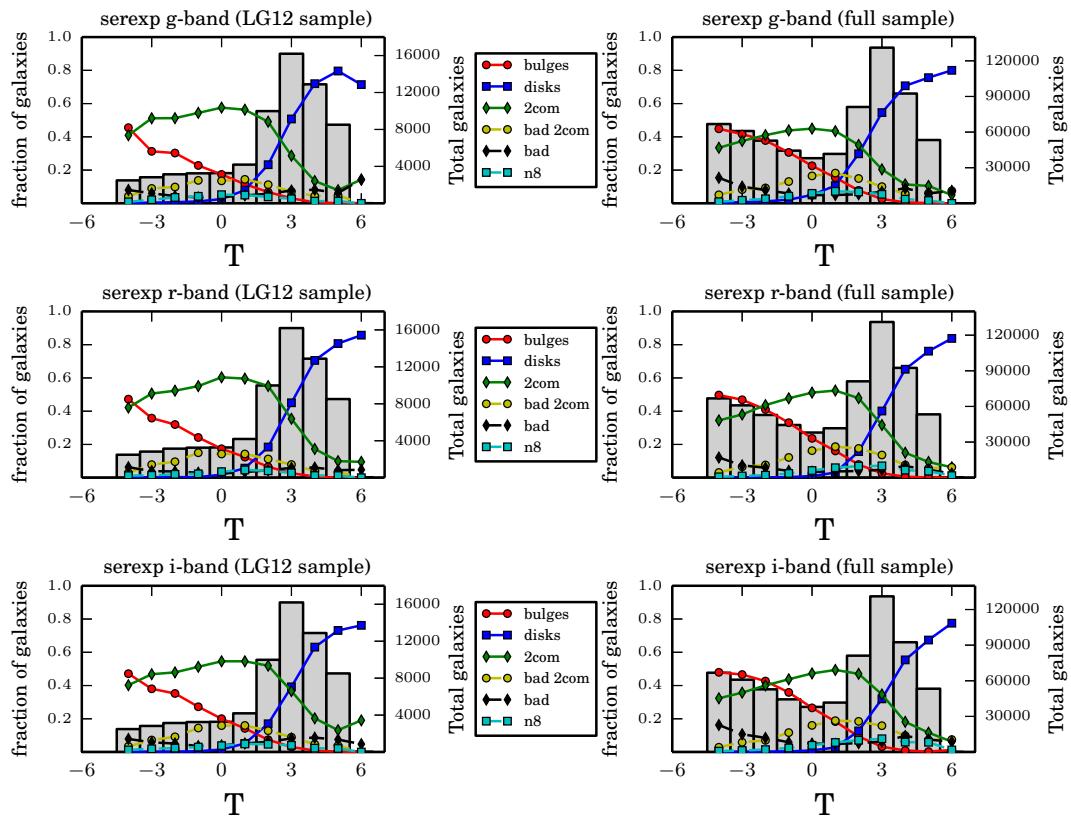
## 8.2 Model selection as a function of $T$ -type



**Figure 8.3:** Comparisons of the inferred  $T$ -type for this work (top row), S11 (second row), Men14 (third row), and LG12 (fourth row) for the smaller, low- $z$  LG12 sample (left) and the full sample (right). Plot design is the same as Figure 6.1.

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---



**Figure 8.4:** Comparisons of the inferred T-type for this work. Plot design is the same as Figure 8.2.

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

---

We calibrate the probabilities from H2011 to the  $T$ -types using a simple linear model:

$$\begin{aligned} T = & -4.6 \times P(Ell) - 2.4 \times P(S0) \\ & + 2.5 \times P(Sab) + 6.1 \times P(Scd) \end{aligned} \quad (8.1)$$

The coefficients of the equation are calibrated to the visually classified galaxies of Nair and 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 catalog 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.

Figure 8.3 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 Chapter 7.6).

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

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---

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 Chapter 7.6).

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 percent 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 Chapter 8.3.2. 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 Figure 8.1 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

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

---

the intermediate types where more structure is expected. These results suggest that while S11 fits are similar to this work (see Chapter 7) 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.

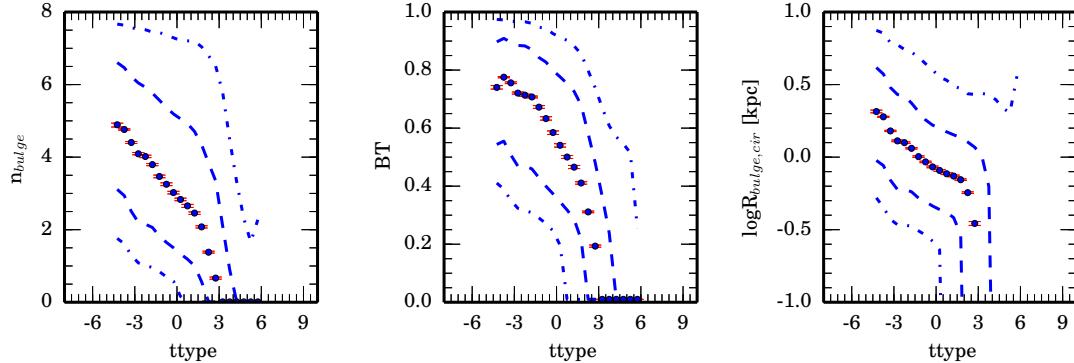
The Men14 classification (third row) is largely independent of  $T$ -type, similar to the behavior with absolute magnitude discussed in Chapter 8.1. The dominant one-component model transitions from **deV** to **Exp** across the  $T$ -types, but makes up nearly 20 per cent of the sample across the range.

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 color cut on the data. Forcing galaxies to be **deV** profiles based on the color 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 Chapter 8.3.1 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 8.4 shows the distribution of flags for the g- (top row), r- (middle), and i-band (bottom row) using the T-types as a proxy for morphology. The classification used in this work is virtually independent of band, but differences are observable in transitional ranges

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---



**Figure 8.5:** 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. Errorbars on the median values represent 95 per cent CI obtained from bootstrap resampling.

of  $T$ -type. Near  $T$ -type= 3 where bulges should become less prominent, the fraction of “pure disk” galaxies steadily drops toward the redder bands and the “2com” proportion increases, as would be expected for redder dimmer bulges.

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

#### 8.3.1 Bulge behavior with $T$ -type

We can examine a few internal checks on the component parameters of bulges and discs in our two-component models. Figure 8.5 shows the median **Ser-Exp** bulge Sérsic index, B/T, and bulge radius for all galaxies in our catalog as a function of  $T$ -type calculated using Equation 8.1. 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

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

---

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 color or any other source of morphological information during fitting, the behavior of the bulges in our sample is good evidence for proper fitting and effective flagging.

#### 8.3.2 Bulge and disk axis ratios

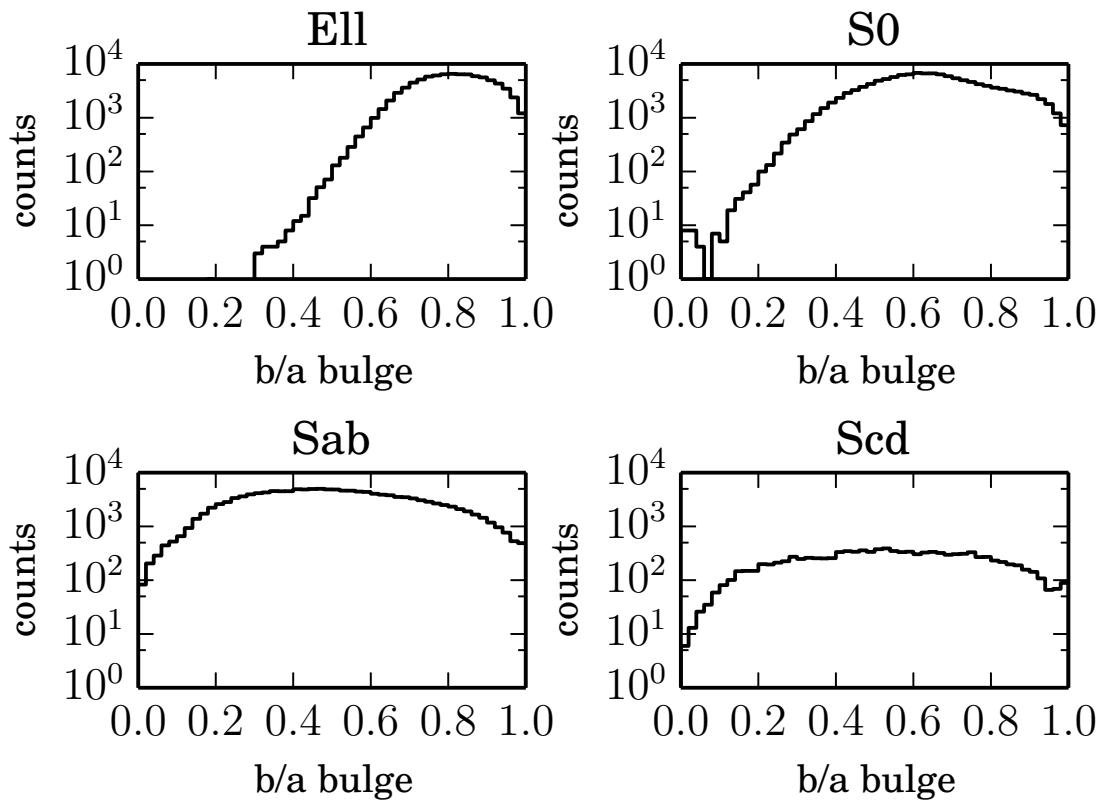
Disk and bulge axis ratios are expected to have different distributions if we are truly measuring bulges and discs as opposed to fitting unphysical components. For instance, bulges are expected to be rounder with lower ellipticities while discs should be more evenly distributed across the range of axis ratios. This distribution for discs is expected for thin discs with uniformly distributed disc inclination. Figures 8.6 and 8.7 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 8.1. We then bin galaxies as either Ell, S0, Sab or Scd using Equation 8.2.

$$\text{Type} = \begin{cases} \text{Ell} & : T \leq -3 \\ \text{S0} & : -3 < T \leq 0.5 \\ \text{Sab} & : 0.5 < T \leq 4 \\ \text{Scd} & : 4 < T \end{cases} \quad (8.2)$$

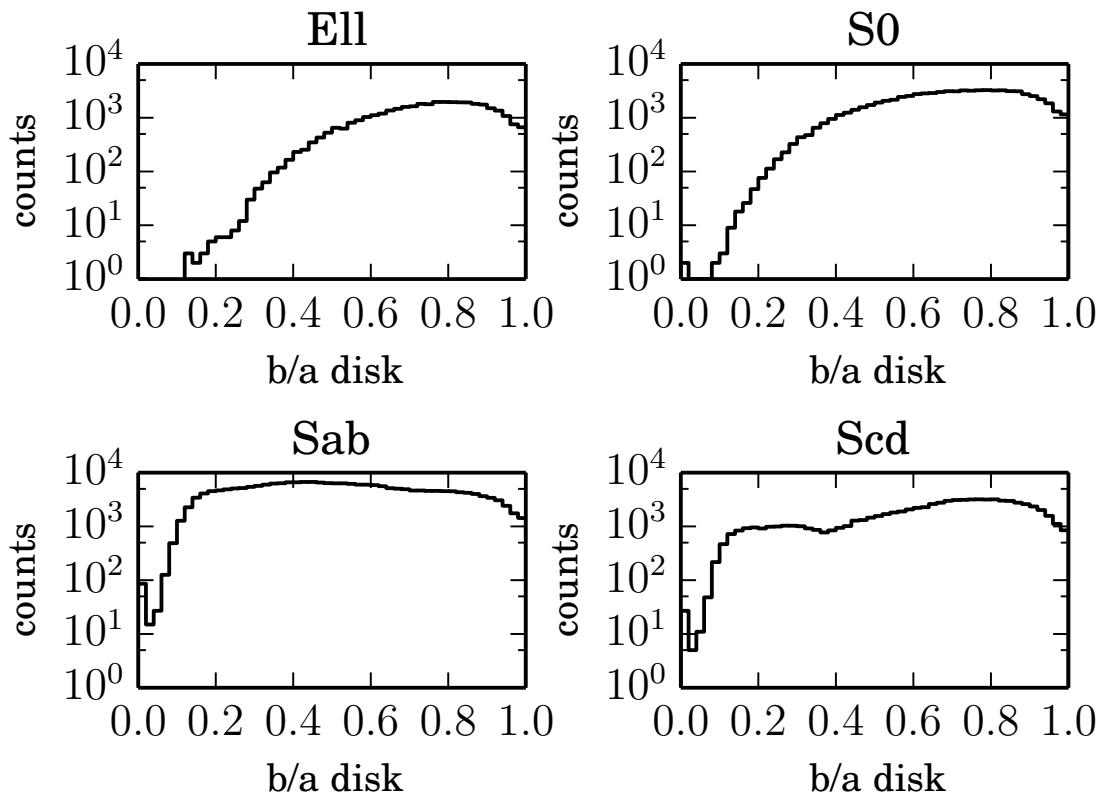
The axis ratio for early types is peaked near 0.8 for the Ell and 0.6 for the S0 galaxies. The distributions get progressively flatter as later types are considered. The increase of

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---



**Figure 8.6:** The Ser-Exp bulge axis ratio for all good two-component or pure bulge galaxies separated by type.



**Figure 8.7:** The Ser-Exp disc axis ratio for all good two-component or pure disc galaxies separated by type.

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---

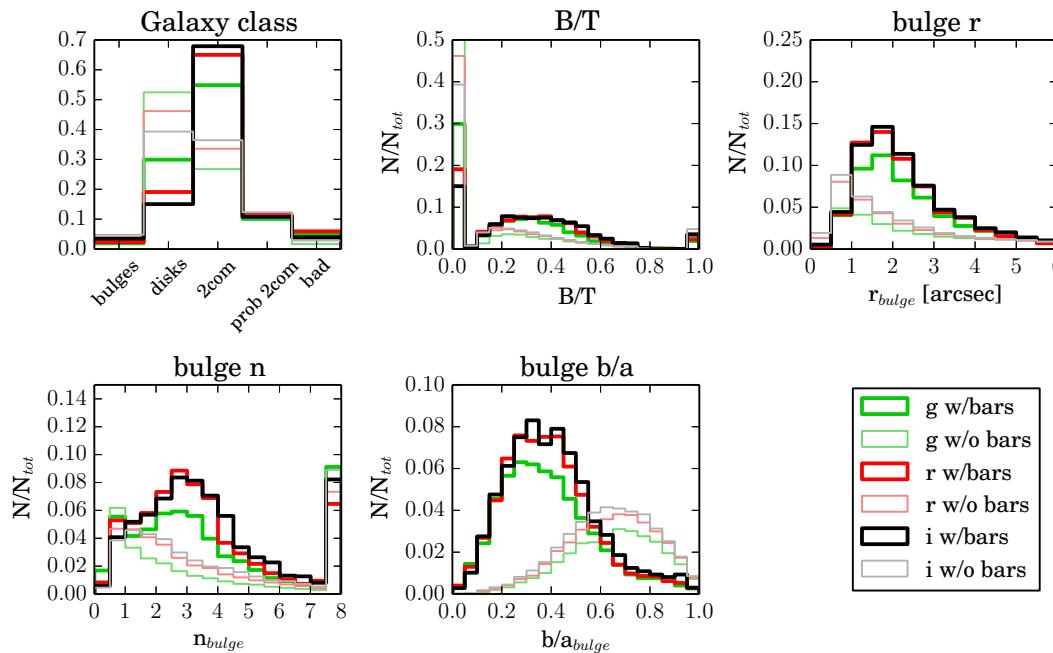
bulges with lower b/a seen in the S0, Sab, and Scd samples is, in part, explained by poor fitting due to contamination of a bar. We address this in Chapter 8.4. Overall, the distributions shown here are flatter than LG12 but similar to those reported in S11 for two-component galaxies. The distributions 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.

### 8.4 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.



**Figure 8.8:** The distribution of Ser-Exp fitting parameters for disk galaxies with bar components and those that do not have a bar, as identified by selecting on the bar vs no\_bar GZ2 t03 classification. Preferred models using our Ser-Exp flags (top left), fitted Ser-Exp B/T (top center), Ser-Exp semi-major bulge half-light radius (top right), Ser-Exp bulge Sérsic index (bottom left), and Ser-Exp bulge axis ratio (bottom center) are shown. Black lines show the results for the i-band. Red and green show the r- and g-bands, respectively.

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---

Figure 8.8 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 GZ2 t03 classification). Preferred models using our **Ser-Exp** flags (top left), fitted **Ser-Exp** B/T (top center), **Ser-Exp** semi major bulge half-light radius (top right), **Ser-Exp** bulge Sérsic index (bottom left), and **Ser-Exp** bulge axis ratio (bottom center) for these galaxies are shown.

Figure 8.8 shows that 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 semi major 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

## 8.5 Color-magnitude relations of the types

---

relative flatness of the bulge axis ratio distribution in Figure 8.6. Proper modeling of the distribution may provide a constraint on the bulge fraction in the sample.

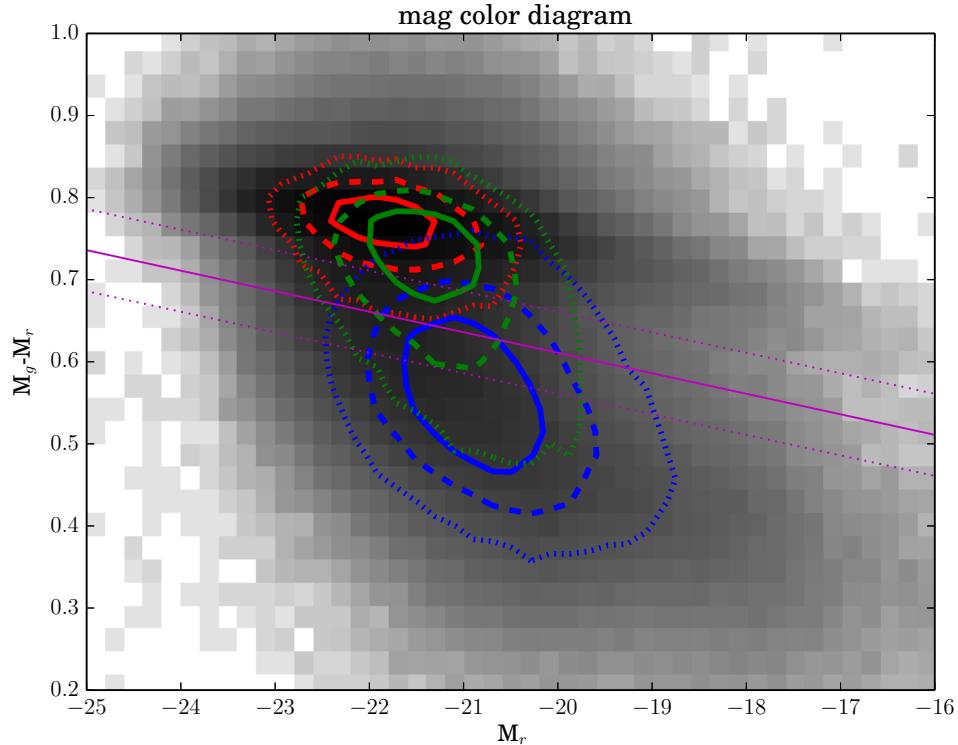
### 8.5 Color-magnitude relations of the types

As a final comparison, Figure 8.9 shows the color-magnitude plot using the  $g-r$  color and  $r$ -band magnitudes of galaxies in our catalog. The x-axis uses  $r$ -band Cmodel magnitude of SDSS and the y-axis uses the  $g-r$  color of the SDSS Modelmag. This combination of measurements is found to be most accurate by Stoughton et al. (2002). The catalog is divided into several classes based on the confusion matrices presented in Figure 6.7. Figure 8.9 shows galaxies classified as pure bulge in all bands in red contours ( $\approx 62\,000$  galaxies), pure disk in all bands in blue contours ( $\approx 109\,000$  galaxies), and two-components in all bands in green contours ( $\approx 116\,000$  galaxies). The solid, dashed, and dotted contours contain 25, 50, and 75 per cent of each group, respectively. The purple solid line is the definition of the centre of the “green valley” adopted by LG12. The dotted purple lines above and below mark the edges of the “green valley” used by LG12.

The contours show that there is good separation in the flagging with pure bulges centered on the red sequence and pure disks covering the blue cloud. Two-component galaxies appear to strongly favor the red sequence, but have a center closer to the “green valley” and more overlap with the “green valley” than either the pure bulge or pure disk sample. However, more analysis is required. Deeper analysis of the green valley galaxies in this sample could further inform recent claims that the “green valley” is not a single evolutionary pathway for galaxies, but contains several pathways from actively star-forming to passive galaxies (Schawinski et al. 2014). This is a goal for future analysis and is discussed in Chapter 12.3.

## 8. COMPARISONS USING MORPHOLOGICAL INFORMATION

---



**Figure 8.9:** Color-magnitude plot using the  $g-r$  color and  $r$ -band magnitudes of galaxies in our catalog. The x-axis uses  $r$ -band Cmodel magnitude of SDSS and the y-axis uses the  $g-r$  color of the SDSS Modelmag. The catalog is divided into several classes based on the confusion matrices presented in Figure 6.7. Galaxies classified as pure bulge in all bands in red contours ( $\approx 62\,000$  galaxies), pure disk in all bands in blue contours ( $\approx 109\,000$  galaxies), and two-components in all bands in green contours ( $\approx 116\,000$  galaxies). The solid, dashed, and dotted contours contain 25, 50, and 75 per cent of each group, respectively. The purple solid line is the definition of the centre of the “green valley” adopted by LG12. The dotted purple lines above and below mark the edges of the “green valley” used by LG12.

## Chapter 9

# Discussion and use of the catalog

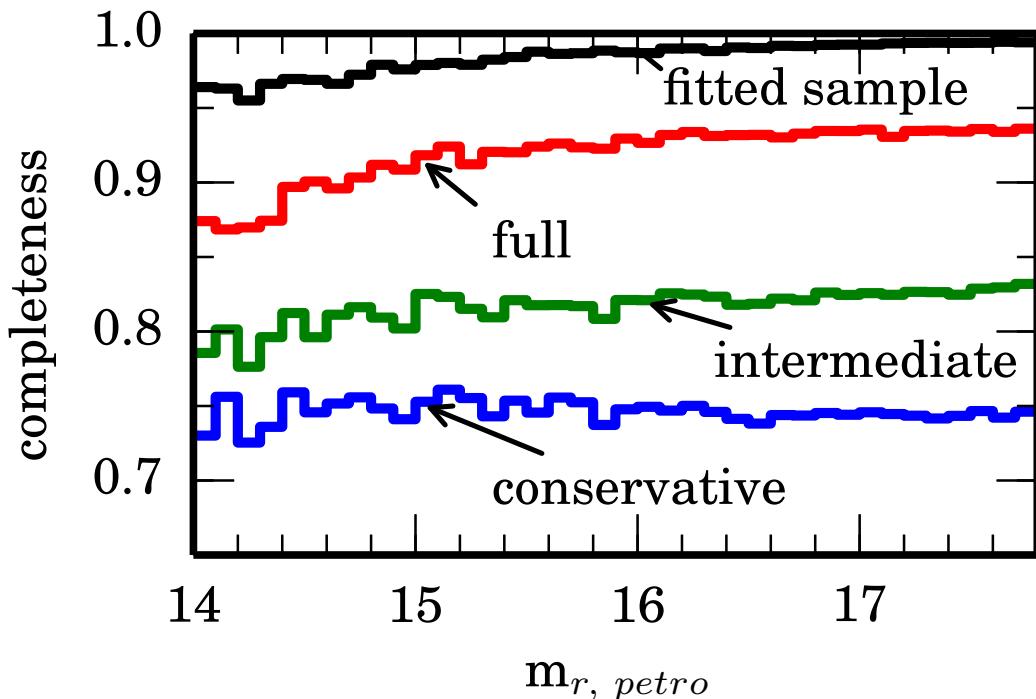
When using the catalog, it is recommended to remove galaxies flagged as bad (flag 20) as these galaxies have catastrophically bad estimates of total magnitude and radius. Additional galaxies may be removed depending on how conservative the user seeks to be. The problematic two-component fits (flag 14) 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 the bulge and disc components of galaxies with bit 6, 7, or 13 set (which were flagged as inverted profiles in the **Ser-Exp** fit) have been swapped. 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 catalogs.

The fits and flags are provided for each of the four models (**deV**, **Ser**, **deV-Exp**, and **Ser-Exp**). In addition, the preferred catalog used in Sections 7.6 and 8 is provided for public use. This preferred catalog 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 8 are

## 9. DISCUSSION AND USE OF THE CATALOG

---



**Figure 9.1:** The completeness of the samples described in Section 9. 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.

---

not included in these tables, but can be calculated directly using Equation 8.1.

The suggested composite samples drawn from the preferred fit catalog described in the previous paragraph are:

**The conservative catalog** 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 catalog.

**The intermediate catalog** Use the catalog above plus all galaxies with final flag bit 10 set and bulge Sérsic index  $=8$ .

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

Figure 9.1 shows the completeness of the three samples described above. The ‘fitted sample’ represents our selection after the cuts made in Chapter 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 catalog 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, there may be cases in which these models are preferred for analysis, especially when comparing to prior works. 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. All flagging is carried

## 9. DISCUSSION AND USE OF THE CATALOG

---

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 catalog are available online at [http://www.physics.upenn.edu/~ameert/SDSS\\_PhotDec/](http://www.physics.upenn.edu/~ameert/SDSS_PhotDec/). The website provides an interface for generating panels similar to Figure 5.2, 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 9.1 and 9.2 describe the format of the data tables released as part of this work. The data are distributed 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 neighbors, etc.). The following extensions contain the ‘best model’ (the combination of **Ser** and **Ser-Exp** fits described in Chapters 7.6 and 8), 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 5.3 (labeled 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 catalog and external works.

---

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_neighborfit	the number neighbor 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

**Table 9.1:** 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.

## 9. DISCUSSION AND USE OF THE CATALOG

---

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 (semi-minor/semi-major 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 half-light radius (arcsec)	Float
11	r_bulge_err	The bulge radius error (arcsec)	Float
12	n_bulge	The bulge Sérsic index	Float
13	n_bulge_err	The bulge 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_disc	The disc x centre (pixels)	Float
19	xctr_disc_err	The disc x centre error (pixels)	Float
20	yctr_disc	The disc y centre (pixels)	Float
21	yctr_disc_err	The disc y centre error (pixels)	Float
22	m_disc	The disc magnitude	Float
23	m_disc_err	The disc magnitude error	Float
24	r_disc	The disc half-light radius (arcsec)	Float
25	r_disc_err	The disc radius error (arcsec)	Float
26	n_disc	The disc Sérsic index	Float
27	n_disc_err	The disc Sérsic index error	Float
28	ba_disc	The disc b/a	Float
29	ba_disc_err	The disc b/a error	Float
30	pa_disc	The disc position angle (degrees)	Float
31	pa_disc_err	The disc position angle error (degrees)	Float
32	GalSky	The PyMORPH sky brightness (mag/arcsec <sup>2</sup> )	Float
33	GalSky_err	The PyMORPH sky brightness (mag/arcsec <sup>2</sup> )	Float
34	chi2nu	The $\chi^2/\text{DOF}$	Float
35	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

**Table 9.2:** 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.

## **Part II**

# **Applications of the catalog**

## Chapter 10

# Dependence of the bright end of the LF and MF on the fitted profile

This chapter compares the luminosity and stellar mass functions based on SDSS CModel magnitudes and PYMORPH Ser and Ser-Exp models of galaxies in the SDSS. This work was completed with Mariangela Bernardi, Ravi Sheth, Vinu Vikram, Marc Huertas-Company, and Francesco Shankar and the results were published in Bernardi et al. (2013).

Chapter 7.6 showed that PYMORPH fits return more light compared to the CModel fits. The predicted masses are larger than CModel-based masses as a result. The differences are most pronounced at the massive end, where the measured number density of objects having  $M_* \geq 6 \times 10^{11} M_\odot$  is  $\sim 5 \times$  larger. Overall, the total stellar mass density as  $z \sim 0.1$  is a factor of about 1.2 larger than the analysis of Bernardi et al. (2010). The differences with respect to Petrosian-like magnitudes, are even more dramatic. In addition to these effects, different studies have large differences in the assumed  $M_*/L$ , which can exist even after scaling to the same IMF. While these allow one to quantify the differences which arise because of the assumed light profile, the most appropriate profile for high mass galaxies is still debated.

## 10.1 Introduction

Halo Model based abundance matching techniques or extreme value statistics require a good parametrization of galaxy abundance at the bright, massive end of the LF in order to understand the origin of the most luminous galaxies (e.g. Paranjape and Sheth 2012). Bernardi et al. (2010) noted that the most luminous galaxies were more abundant than expected from the most commonly used parameterizations of the LF. They also showed that this mis-match was important for constraining assembly of the brightest galaxies via major or minor mergers. In particular, the stellar mass function affects models which use the observed abundance and its evolution to constrain formation models. They also showed that the conversion from  $\phi(L)$  to  $\phi(M_*)$  is sensitive to the assumed stellar mass-to-light ratio, for which, there is still no consensus.

In addition to the large systematic differences arising from assumptions about the stellar mass-to-light ratio, the massive end of the stellar mass function is sensitive to the exact method used to fit light profiles to the most luminous galaxies. Bernardi et al. (2010) used luminosities estimated from the **CModel** magnitudes output by the DR7 (Abazajian et al. 2009). These tend to return more light than the more commonly used estimates based on the Petrosian radius defined by the SDSS, especially for the brightest objects (see Chapter 7.6). However, some of this difference was due to sky subtraction problems in the SDSS. Bernardi et al. (2010) applied a crude correction for this to the **CModel** magnitudes, but not to the Petrosian magnitudes output by the SDSS pipelines, primarily because essentially all previous work with Petrosian magnitudes made no such correction.

The **CModel** magnitudes are defined in Equation 7.1 and discussed in Chapter 7.6. The fits presented as part of this thesis, in particular the **Ser** and **Ser-Exp** fits, typically return even more light than the **CModel** magnitudes (see Figure 7.24, also Bernardi et al. 2014), in part because of the improved treatment of the sky, but also because differences in the model which is fitted to the observed light profile matter. The **Ser** and **Ser-Exp** models are integrated to infinitely large radius, while the **CModel** and other SDSS magnitudes are

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---

truncated at several half-light radii. A correction was applied to account for the profile truncation. However, a difference in magnitude is still present.

These differences impact estimates of the luminosity and stellar mass functions at the bright end. Chapter 10.2 describes the sample used to examine the effects on the LF and MF, shows the luminosity and stellar mass functions, quantifies how they depend on the fit to the light profile and provides simple fitting formula which quantify the results. While these results allow one to easily account for the dependence on the light profile (e.g. using `Ser` instead of SDSS `CModel` or Petrosian magnitudes), which  $M_*/L$  estimate is most appropriate is still an open question, and deserves further study. For reasons described in Bernardi et al. (2010), all  $M_*/L$  estimates assume a Chabrier (2003) IMF. Chapter 10.3 shows that even though a number of recent works have made this same choice for the IMF (Baldry et al. 2012; Moustakas et al. 2013), they still have  $M_*/L$  values which are very different from the values presented as part of this work (Bernardi et al. 2010, i.e., ), from one another, and from earlier work (Bell et al. 2003). The fact that differences in  $M_*/L$  arise even when the same IMF is assumed is not generally considered and is a substantial sign of the intrinsic uncertainty in these measurements. Chapter 10.4 shows how the luminosity and stellar mass functions depend on morphological type, where the type is determined by the Bayesian Automated Classification of Huertas-Company et al. (2011).

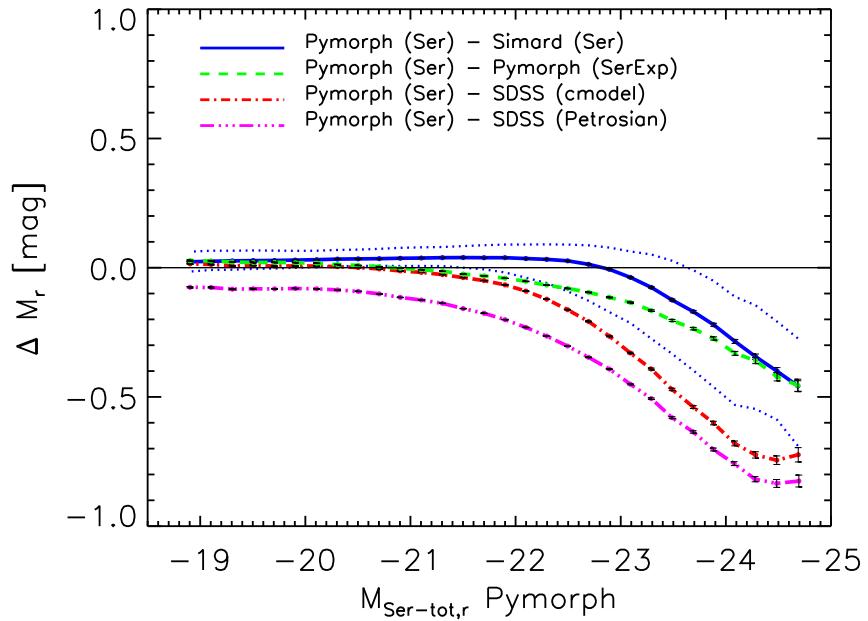
### 10.2 Luminosity and stellar mass functions

#### 10.2.1 The sample

To provide a direct comparison with previous work, the sample of Bernardi et al. (2010) was used. This sample consists of about 260,000 SDSS galaxies having  $14.5 \leq m_{r\text{Pet}} \leq 17.7$ . DR7 Petrosian and `CModel` estimates of the total light for each of these objects were used in the analysis. These are known to suffer from sky-subtraction and crowded-field/masking problems (Bernardi et al. 2010; Meert et al. 2013, 2015). For this analysis, the `CModel`

## 10.2 Luminosity and stellar mass functions

---



**Figure 10.1:** Difference between PYMORPH Sérsic fits and SDSS DR7 Petrosian, SDSS CModel, PYMORPH SerExp, and Sérsic fits from Simard et al. (2011) (bottom to top), for galaxies in the sample selected by Bernardi et al. (2010). Petrosian magnitudes are always the faintest, whereas single Sérsic-based magnitudes tend to be the brightest. Dotted lines around PYMORPH (Ser)–Simard (Ser) show the 16th and 84th percentiles of the distribution; these are similar to the scatter around the median for the other curves.

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---

magnitudes are crudely corrected for the SDSS sky subtraction problems as described in Bernardi et al. (2010). Analogous corrections to the Petrosian magnitudes are rarely made, so no such correction is made here. This allows easier comparison with previous work (this is discussed further in the context of Figure 10.1). Finally,  $k$ - and evolution corrections were computed for each object following Bernardi et al. (2010).

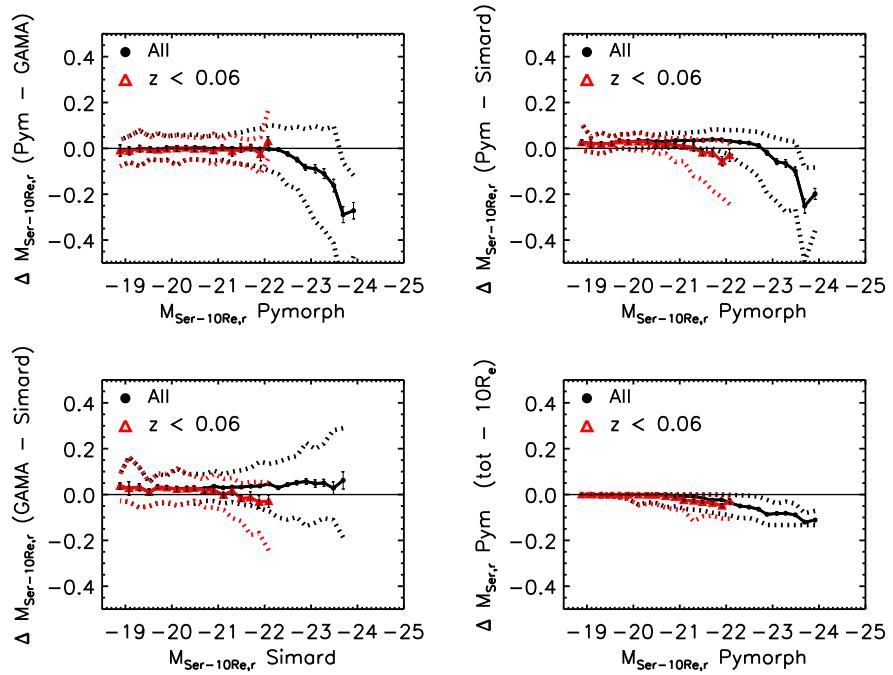
### 10.2.2 Dependence on assumed surface brightness profile

Chapter 4.2.1 and 4.2.2 showed that the magnitudes and half-light radii output by PY-MORPH depend on the model which is fit. For example, fitting what is really a two-component image with a `deV` profile will generally underestimate the total light. On the other hand, the total light associated with the best-fit `Ser` or a `Ser-Exp`, is less-biased from its true value (Bernardi et al. 2007, 2014, e.g. ). Chapter 4.2.1 and 4.2.2 as well as Bernardi et al. (2014) have also shown that in objects brighter than  $L_*$ , fitting a two-component `Ser-Exp` model to what is really just a `Ser` results in a noisier, but unbiased, recovery of the input parameters. In contrast, fitting a `Ser` to what is truly a two-component system results in significant biases.

Although the `Ser-Exp` model is more accurate, the `Ser` fit is often performed on real data when the resolution and S/N are make it unlikely to recover a robust two-component fit. However, either of these models are expected to be more realistic than a `deV` model.

The estimates of the total light from a `Ser` or `Ser-Exp` model are generally larger than those based on the `CModel` magnitudes output by the SDSS pipelines (Bernardi et al. 2007, 2014; Hill et al. 2011; Mosleh et al. 2013, e.g. ), and both are larger than the SDSS DR7 Petrosian magnitudes. Figure 10.1 illustrates that this difference can be large. Some of this is due to the difference in the treatment of the sky, and some to the differences between the fitted models. For example, the offset at the faint end between the Petrosian and the other models is almost entirely due to the fact that the SDSS DR7 pipeline tended to overestimate

## 10.2 Luminosity and stellar mass functions



**Figure 10.2:** Comparison of Ser fits for the SDSS galaxies in common to Simard et al. (2011), Kelvin et al. (2012), and PYMORPH (our notation  $M_{\text{Ser}-10R_e}$  reflects the fact that Kelvin et al. truncate the profile at  $10R_e$ ). Red triangles show a similar analysis if one restricts to objects with  $z < 0.06$  as done in Baldry et al. (2012); this shallower volume does not probe the highest luminosities that are of most interest here. Dotted lines show the 16th and 84th percentiles of the distribution.

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---

the contribution from the sky, thus making the SDSS Petrosian magnitudes about 0.05 mags too faint.

After this study was complete, He et al. (2013) quantified the effects of sky-subtraction and masking problems on the SDSS DR7 Petrosian values. Accounting for these effects makes the Petrosian magnitudes 0.05 mags brighter at the faint end, and 0.2 mags brighter at the bright end. (Our own reanalysis, based on PYMORPH sky-estimates suggests this difference is slightly smaller: about 0.1 mags at  $M_r \leq -23.5$ .) As a result, He et al. (2013) Petrosian magnitudes are slightly brighter than our CModel magnitudes at the bright end (recall that the CModel magnitudes include only a crude correction for the SDSS sky subtraction problems), but they are generally fainter than our Ser or Ser-Exp values at the bright end. There are several systematic effects on Petrosian magnitudes. Petrosian magnitudes underestimate the total light when the light profile has extended wings (extended wings are manifested as high Sérsic index in the fits). This is particularly important at the bright end (Andreon and Cuillandre 2002; Binggeli and Cameron 1991; Blanton et al. 2001; Brown et al. 2003; Graham et al. 2005; Trujillo et al. 2001, e.g. ). At the bright end, this leads to an underestimate of order  $\sim 0.3$  mags or more (i.e., , this matters more than the sky-subtraction problems), which is similar to the difference between the CModel and Ser-Exp magnitudes. This expected difference is consistent with the He et al. (2013) finding that even the revised Petrosian magnitudes are systematically fainter than aperture fluxes based on deeper photometry which reaches to 1% of the sky.

Although the dependence on the assumed light profile is what has motivated this study, it is reasonable to ask if these differences are indeed larger than those associated with different pipelines which fit the same model.

### 10.2.3 Dependence on pipeline

As a check of reductions, we have also used luminosities from the Ser photometric reductions of Simard et al. (2011). Figure 10.1 shows that these are in good agreement with PYMORPH

## 10.2 Luminosity and stellar mass functions

---

Fit	$\phi_*$ $10^{-2}\text{Mpc}^{-3}$	$L_*$ $10^9 L_\odot$	$\alpha$	$\beta$	$\phi_\gamma$ $10^{-2}\text{Mpc}^{-3}$	$L_\gamma$ $10^9 L_\odot$	$\gamma$	$\rho_L$ $10^9 L_\odot \text{Mpc}^{-3}$
cmodel	0.928	0.3077	1.918	0.433	0.964	1.8763	0.470	0.136
Sérsic	1.343	0.0187	1.678	0.300	0.843	0.8722	1.058	0.150
SerExp	1.348	0.3223	1.297	0.398	0.820	0.9081	1.131	0.146
Sérsic (Simard)	1.920	6.2456	0.497	0.589	0.530	0.8263	1.260	0.152

Fit	$\phi_*$ $10^{-2}\text{Mpc}^{-3}$	$M_*$ $10^9 M_\odot$	$\alpha$	$\beta$	$\phi_\gamma$ $10^{-2}\text{Mpc}^{-3}$	$M_\gamma$ $10^9 M_\odot$	$\gamma$	$\rho_{M_s}$ $10^9 M_\odot \text{Mpc}^{-3}$
cmodel	0.766	0.4103	1.764	0.384	0.557	4.7802	0.053	0.276
Sérsic	1.040	0.0094	1.665	0.255	0.675	2.7031	0.296	0.344
SerExp	0.892	0.0014	2.330	0.239	0.738	3.2324	0.305	0.330
Sérsic (Simard)	0.820	0.0847	1.755	0.310	0.539	5.2204	0.072	0.349

**Table 10.1:** Parameters of  $\phi(L_r)$  (top rows) and  $\phi(M_*)$  (bottom rows) derived from fitting equations (10.1) to the observed counts based on different magnitudes.

except for a small offset ( $\sim 0.05$  mags), although the differences become large at the bright end. Chapter 7.3 shows more detailed comparisons of S11 and this work. Figures A1 and A2 in Bernardi et al. (2014) and discussion on sky estimates in Chapter 4.2.7 and 7.3 suggest that the PyMorph estimates are less biased compared to the S11 fits. Nevertheless, these differences are small compared to PyMorph–CModel.

The PyMorph and S11 luminosities come from integrating the fitted profile to infinity while other authors truncate the profile, typically at some multiple of the half-light radius. For example, the analysis of galaxies in the GAMA survey (Galaxy And Mass Assembly survey Kelvin et al. 2012) truncates the fits at  $10R_e$ . Figure 10.2 compares PyMorph, Simard, and GAMA values for the **Ser** magnitude. PyMorph, S11, and GAMA DR1 values (Driver et al. 2011) are compared for the  $\sim 7335$  galaxies for which all three reductions are available. This sample is set by the fact that the GAMA DR1 covers 100 sq.deg. of the SDSS. GAMA has 10750 matches with the DR7 SDSS spectroscopic galaxy sample, of which 7335 galaxies are in the Bernardi et al. (2010) sample used here. The bottom right panel shows that the truncation matters at the level of 0.05 mags only at  $M_r < -22$ . But otherwise, if truncated similarly, then GAMA and PyMorph are in good agreement at

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---

$M_r > -22$ , whereas PYMORPH returns significantly more light than the other two at the bright end. Differences at the bright end are likely similar in origin (i.e., sky subtraction issues) to those with respect to S11 (Bernardi et al. 2014; Meert et al. 2013, 2015, see).

### 10.2.4 The luminosity function

For each of the estimates of the total light shown in Figure 10.1, the luminosity function is estimated as in Bernardi et al. (2010) using the  $V_{\max}$  method of Schmidt (1968). Each galaxy is weighted using  $1/V_{\max}(L_{rPet})$ , where  $V_{\max}$  is the maximum comoving volume within which the object could have been included in the sample, accounting for both the bright and faint magnitude limits.

Figure 10.3 shows the luminosity functions for the SDSS Petrosian and CModel magnitudes (corrected for the SDSS sky subtraction problems as described in Bernardi et al. (2010)), and Ser-Exp and Ser magnitudes (from PYMORPH). Although the difference between the Petrosian and CModel magnitudes has been known for some time, the fact that Ser-based counts lie substantially above those based on the SDSS outputs has only recently begun to attract attention. For example, the GAMA based results of Hill et al. (2011) point to this difference, but because GAMA covers a substantially smaller volume than the SDSS, it does not probe the high luminosity end which is of most interest here. Our PYMORPH reductions, which are in good agreement with Hill et al. (2011) at  $M_r > -23$ , show that at  $M_r < -23$  the difference with respect to CModel counts is dramatic.

The PYMORPH-based counts are in good agreement with those which use the Simard et al. (2011) Ser reductions, except at luminosities brighter than  $M_r \sim -24$ , where PYMORPH tends to be brighter, so the PYMORPH luminosity function shows more high luminosity objects. This agreement illustrates that our finding that single-Sérsic fits return substantially more objects in the high luminosity tail than do CModel magnitudes is robust to changes in the reduction pipeline.

## 10.2 Luminosity and stellar mass functions

---

The solid curves show the result of fitting

$$X\phi(X) = \phi_\alpha \beta \left( \frac{X}{X_*} \right)^\alpha \frac{e^{-(X/X_*)^\beta}}{\Gamma(\alpha/\beta)} + \phi_\gamma \left( \frac{X}{X_\gamma} \right)^\gamma e^{-(X/X_\gamma)} \quad (10.1)$$

with  $X = L$  to the counts. The associated luminosity density is  $\rho_X = \phi_\alpha X_* \Gamma[(1+\alpha)/\beta]/\Gamma[\alpha/\beta] + \phi_\gamma X_\gamma \Gamma[1+\gamma]$ . The first term in equation (10.1) is the same functional form as that used by Bernardi et al. (2010); the second is required to fit the slight bump at the faint-end. The parameters which yield the best-fit are given in Table 10.1. Note that the value of  $X_*$  is not as intuitive as is its mean value  $X_* \Gamma[(1+\alpha)/\beta]/\Gamma[\alpha/\beta]$ .

The observed distributions shown here have been broadened slightly by measurement errors. Bernardi et al. (2010) showed how to modify the analog of Equation 10.1 to estimate the parameters of the intrinsic distribution. In practice, the difference between the intrinsic and observed broadened distributions is small; much smaller than the difference between the PYMORPH and CModel counts, so here the results of the observed distribution are shown rather than the intrinsic one.

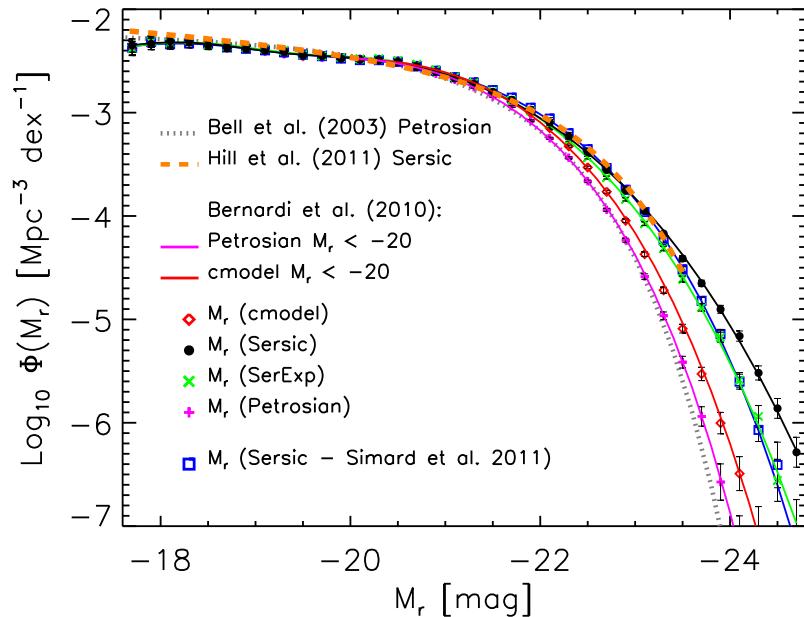
### 10.2.5 The stellar mass function

Figure 10.4 shows the associated stellar mass functions. In all cases,  $M_*$  was estimated from the luminosity and the CModel  $g - r$  color assuming the Chabrier (2003) IMF as described in Bernardi et al. (2010). CModel color is used here because the main goal of this paper is to study the effect on  $\phi(M_*)$  from changes in  $L$ . Using CModel colors ensure that  $M_*/L$  estimates for each object are the same as in Bernardi et al. (2010). However, the  $L$  estimate for each object differs ( $\text{Petrosian} \neq \text{CModel} \neq \text{PYMORPH}$ ). Additionally, if the stellar population models used to estimate  $M_*/L$  are incorrect, or if the IMF is mass-dependent, then this will modify the results. Chapter 10.3 shows comparison with other work.

The estimate from Baldry et al. (2012) lies below all the others. This is remarkable because it is based on the GAMA-Ser reductions, and the associated  $\phi(L)$  (Hill et al. 2011, from) is in good agreement with PYMORPH-based estimate. Therefore, the difference in

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

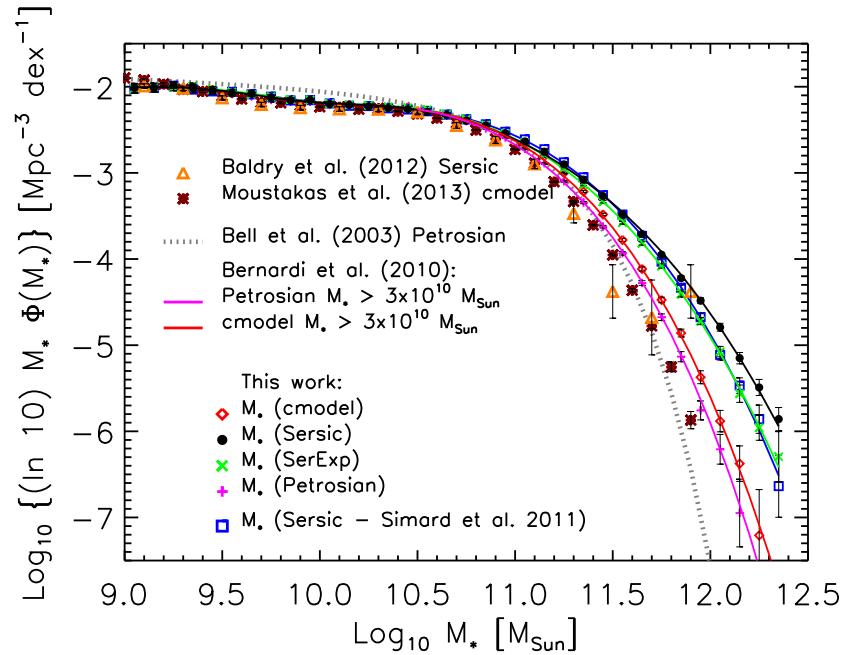
---



**Figure 10.3:** SDSS Main galaxy luminosity function based on Petrosian, CModel, Ser from S11 and PYMORPH Ser-Exp and Ser magnitudes (bottom to top at  $M_r = -24$ ). Smooth curves show the result of fitting equation (10.1) to the counts; associated best-fit parameter values are given in Table 10.1. For the Petrosian and CModel magnitudes, the curve shown is that reported by Bernardi et al. (2010) on the basis of fitting to  $M_r < -20$ . The Petrosian and Ser based fits of Bell et al. (2003) and Hill et al. (2011), respectively, are also shown for comparison.

## 10.2 Luminosity and stellar mass functions

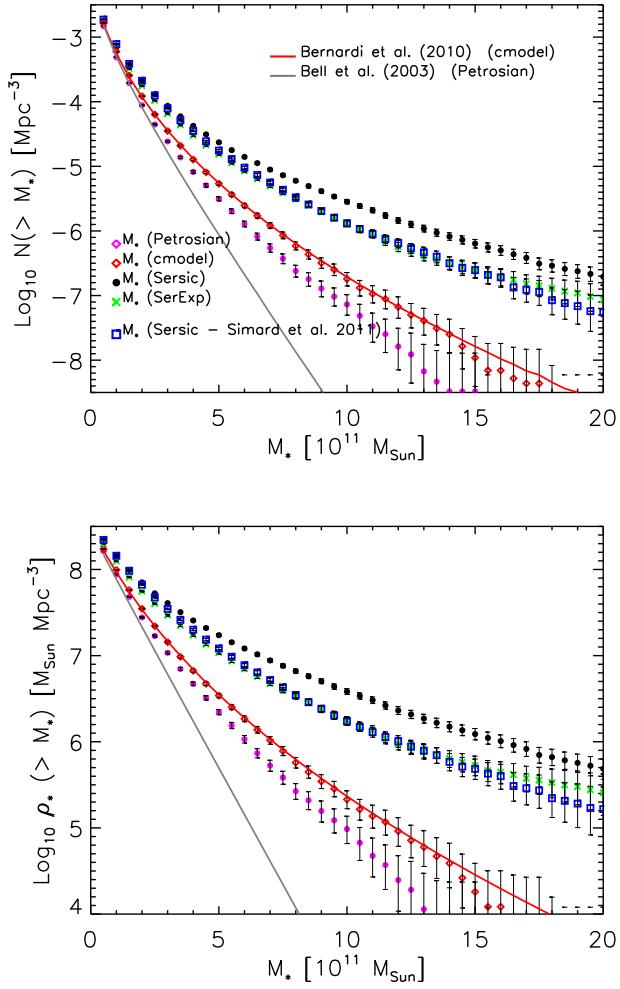
---



**Figure 10.4:** Same as Figure 10.3, but now for the associated stellar mass functions. Recent stellar mass functions from Baldry et al. (2012, ;based on Sérsic magnitudes) and Moustakas et al. (2013, ;based on CModel magnitudes) are also shown. All stellar masses assume a Chabrier IMF.

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---



**Figure 10.5:** Similar to Figure 10.4, but now showing cumulative rather than differential counts. Top and bottom panels show number and stellar mass density respectively. To facilitate comparison with previous work we show the fit of Bell et al. (2003).

### 10.3 Comparison with previous work

---

$\phi(M_*)$  must be entirely due to  $M_*/L$ , even though Baldry et al. (2012) also assume a Chabrier IMF. This is discussed further in Chapter 10.3.

Figure 10.5 shows cumulative (rather than differential) counts, both for number and stellar mass-weighted density. The number counts at the mass scale above which the number density of objects is  $10^{-6}\text{Mpc}^{-3}$  is larger by a factor of  $\sim 2$  compared to the CModel-based counts (a factor of  $\sim 3$  compared to Bell et al. (2003)). Alternatively, at  $M_* = 6 \times 10^{11} M_\odot$ , the PYMORPH counts lie a factor of  $\sim 8$  above those based on CModel magnitudes (a much larger factor above Bell et al. (2003)). For the mass-weighted counts the corresponding discrepancies at  $10^6 M_\odot \text{Mpc}^{-3}$  or  $6 \times 10^{11} M_\odot$  are similar or slightly larger.

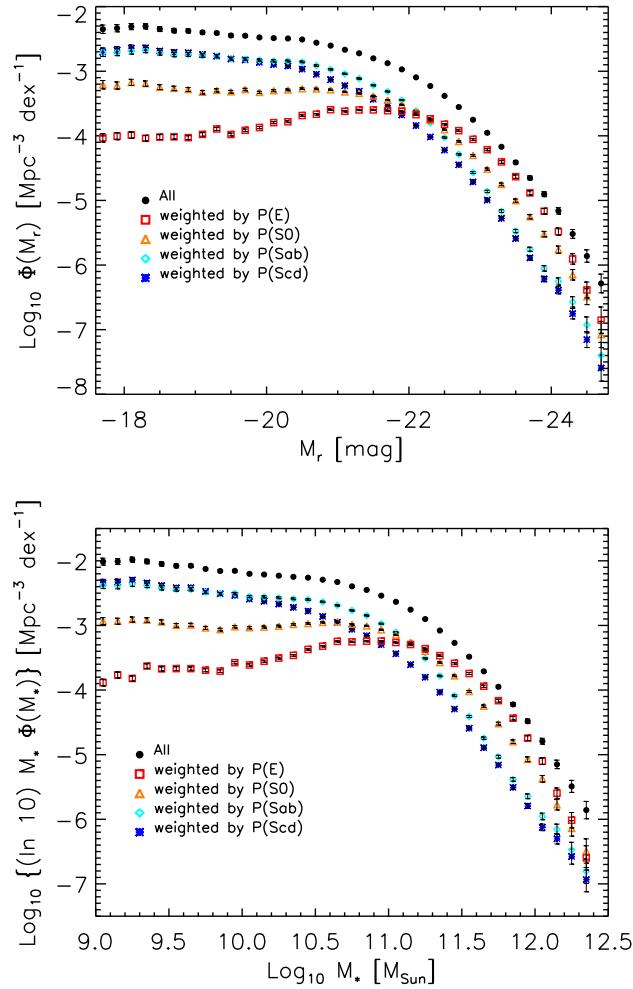
## 10.3 Comparison with previous work

Recent focus has concentrated more on  $\phi(M_*)$  than  $\phi(L)$ . Unfortunately, this combines uncertainty associated in the total light with uncertainty in the stellar mass-to-light ratio  $M_*/L$  (e.g. mismatched stellar templates, mass-dependence of the IMF, etc.). To illustrate this, the results of PYMORPH were compared to two of the most recent determinations of  $\phi(M_*)$ : the results of Baldry et al. (2012) and Moustakas et al. (2013). Although all three assume a Chabrier IMF, the estimates of  $\phi(M_*)$  made in Baldry et al. (2012) and Moustakas et al. (2013) are more similar to one another than they are to PYMORPH. However, this implies large differences in their  $M_*/L$  values since they used different luminosities to get  $M_*$ .

The Baldry et al. (2012) analysis is based on **Ser** fits to the light profiles from galaxies with  $z < 0.06$  in the GAMA survey reported by Kelvin et al. (2012). Of the 7335 GAMA matches in the Bernardi et al. (2010) sample which are studied here, only 1612 have  $z < 0.06$ . The red triangles in Figure 10.2 show that, for these 1612 galaxies, the **Ser** PYMORPH estimates of the total light are in good agreement with those derived by Kelvin et al. (2012), and used by Baldry et al. (2012). However, the redshift cut eliminates most of the high luminosity objects which are of most interest in this work.

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---



**Figure 10.6:** Morphological dependence of the luminosity (top) and stellar mass (bottom) functions. In all cases, the luminosities are based on **Ser** magnitudes, stellar masses assume a Chabrier IMF, and morphologies are based on the BAC method of Huertas-Company et al. (2011).

### 10.3 Comparison with previous work

---

GAMA-based luminosity function for these  $z < 0.06$  objects is in good agreement with that based on PYMORPH **Ser** reductions when restricted to  $z < 0.06$ . This is in good agreement with that for the full GAMA sample of Hill et al. (2011) shown in Figure 10.3. The smaller volume means the error bars are large, and the comparison is effectively limited to abundances greater than about  $10^{-4}$  Mpc $^{-3}$  dex $^{-1}$ . However, PYMORPH-based  $\phi(M_*)$  estimate lies well above that of Baldry et al. (2012), so  $M_*/L$  values for PYMORPH must be larger. The Baldry et al. (2012) estimate is much closer to, but slightly below, the estimate of Bell et al. (2003, scaled to a Chabrier IMF). However, since the Baldry et al. (2012) **Ser**-based  $\phi(L)$  estimate agrees with PYMORPH, and this lies well above the  $\phi(L)$  associated with the SDSS Petrosian magnitudes which were used by Bell et al. (2003) (see Figure 10.3), the Baldry et al. (2012)  $M_*/L$  values must also be smaller than those of Bell et al. (2003).

Moustakas et al. (2013) use **CModel** magnitudes. Since their sample is essentially the same SDSS sample as the Bernardi et al. (2010) sample, the luminosity function should agree with that of Bernardi et al. (2010). However, Figure 10.4 shows that although their stellar mass function is similar to the **CModel**-derived  $\phi(M_*)$  at low masses, it is different at the high end, indicating that the Moustakas et al. (2013)  $M_*/L$  ratios are smaller than those used here. This may be related to the choice of template used to estimate  $M_*/L$  at higher masses, for reasons given in Fig. 22 and associated discussion of Bernardi et al. (2010). On the other hand, despite differences at lower masses, the Moustakas et al. (2013)  $\phi(M_*)$  is reasonably well approximated by the Bell et al. (2003) estimate at higher masses. Since the **CModel** luminosity function is different from the Petrosian one at these high masses, the Moustakas et al. (2013)  $M_*/L$  values must be smaller than those of Bell et al. (2003), and different again from those of Baldry et al. (2012) (who used **Ser** rather than **CModel** magnitudes).

Therefore,

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---

- by comparing the PYMORPH **Ser**-based  $\phi(M_*)$  with that based on S11, Figure 10.4 quantifies how differences due to the uncertainties in a given light profile fit affect  $\phi(M_*)$ ;
- by comparing the **Ser-Exp** and **Ser**-based  $\phi(M_*)$  fits, Figure 10.4 quantifies the effect of fitting different light profile models;
- by comparing the **Petrosian** based  $\phi(M_*)$  to the Bell et al. (2003) fit, the **CModel**-based estimate to that from Moustakas et al. (2013), or **Ser** (or S11) based  $\phi(M_*)$  to Baldry et al. (2012), Figure 10.4 quantifies how systematic differences in  $M_*/L$  affect  $\phi(M_*)$ .

This shows that the effects on  $\phi(M_*)$  of using the total luminosity computed from different fits to the light profile are dramatic; it is important to specify how the light profile was fit when reporting a luminosity or stellar mass function.

### 10.4 Dependence on morphology

Combining the PYMORPH **Ser-Exp** reductions with the Bayesian Automated morphological Classifier of Huertas-Company et al. (2011) as was carried out in Chapter 8, can determine how the luminosity and stellar mass functions depend on morphology. Since the BAC returns the probability that an object is one of four types, the sum of the four probabilities is unity for each object. Therefore, to estimate the luminosity function, each object,  $j$ , is weighted by  $p_j(\text{type})/V_{\max}(L_j^{\text{Pet}})$ .

In practice, there are a number of faint objects for which  $p_j(E) \leq 0.15$ . These can dominate over the counts from similarly faint objects for which  $p_j(E) > 0.85$ . If these low values of  $p$  are simply the result of errors in the BAC algorithm, then these will wrongly boost the luminosity function at the faint end. To check the magnitude of the effect, all values of  $p \leq 0.15$  are set to zero, and the weights are renormalized so that the sum over

the four  $p$  values (some of which are now zero) is still unity. This reduces the counts of faint ellipticals and luminous spirals by a factor of about two. The modified counts are shown in Figure 10.6. Figure 10.6 shows that early-types (E and S0) dominate the high-mass end and later-types (Sa to Sd) dominate at lower masses as is expected.

These estimates were compared to Nair and Abraham (2010), who used the T-Type classification using the modified RC3 (Third Reference Catalogue of Bright Galaxies) classifiers. Setting E galaxies to  $T = -5$  and  $T = -4$ , S0 galaxies to  $T = -3$ ,  $T = -2$  and  $T = -1$ , Sa galaxies to  $T = 0$ ,  $T = 1$  and  $T = 2$ , Sb galaxies to  $T = 3$  and  $T = 4$ , and Scd galaxies to  $T = 5$ ,  $T = 6$  and  $T = 7$  in place of the BAC probabilities, PYMORPH E and S0 counts are in quite good agreement. However, at the bright end Nair and Abraham (2010) tend to classify objects as E rather than S0 when BAC divides its weights between E and S0. Differences are slightly more pronounced for later types. PYMORPH Scd counts are in good agreement at the faint end, but lie substantially above Nair and Abraham (2010) at high luminosities (where the counts are falling exponentially). On the other hand, PYMORPH Sab counts lie about a factor of two below Nair and Abraham (2010) at high  $L$ , suggesting that BAC assigns some weight to the extreme Scd classification when Nair and Abraham (2010) choose the extremes.

## 10.5 Discussion

PYMORPH Ser or Ser-Exp-based estimates of the total light of a galaxy are larger than those based on SDSS pipeline outputs (Petrosian or CModel; Figure 10.1). As a result our PYMORPH-based luminosity and stellar mass functions are different from previous work. They have more light and mass at the bright, massive end (Figures 10.3 and 10.4).

Petrosian magnitudes have been popular in recent years. However, the SDSS Petrosian-based luminosity functions presented here, which are similar to those in the literature, lie well-below any of the others. This is consistent with a number of recent studies which agree

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---

that Petrosian, and **CModel** magnitudes, underestimate the total light (Bernardi et al. 2014; Hall et al. 2012; Hill et al. 2011; Kelvin et al. 2012; Meert et al. 2015; Simard et al. 2011). Some of this is due to sky-subtraction problems which affect the SDSS pipelines (Bernardi et al. 2010) more than **PYMORPH** (Meert et al. 2015). While Petrosian magnitudes with better sky-subtraction are brighter (He et al. 2013, e.g. ), thus reducing the difference with respect to **PYMORPH** magnitudes, the results presented here suggest that Petrosian values should only be used after applying the sorts of corrections advocated by Graham et al. (2005) which partially account for the fact that Petrosian magnitudes underestimate the total light when the light profile has extended wings (Andreon and Cuillandre 2002; Blanton et al. 2001; Trujillo et al. 2001, e.g. ). For the brightest galaxies, even such corrected Petrosian magnitudes are fainter than **PYMORPH** magnitudes. In addition, Petrosian-derived quantities are not seeing corrected; this impacts studies of scaling relations (Hyde and Bernardi 2009, e.g. see Appendix A in) more than studies of the mass function. For these reasons, fits to **Ser** or **Ser-Exp** profiles are to be preferred, especially in higher redshift datasets where seeing is an issue.

The **Ser**-based luminosity functions from **PYMORPH** and S11 are very different from the **Ser**-based analysis in Blanton et al. (2003b). However, the **Ser** parameters used by Blanton et al. (2003b) were estimated from a 1-dimensional radial surface brightness profile, measured in  $\sim 5 - 10$  azimuthally averaged annuli. This procedure is expected to be significantly less accurate than the 2-dimensional fits to the whole galaxy image performed by **PYMORPH** and S11. In addition, these more recent analyses include a more careful treatment of the background sky, especially in crowded fields, so these more recent **Ser** reductions (as well as those of Kelvin et al. (2012) for the subset of objects in the GAMA survey) supersede those of Blanton et al. (2003b).

The fits presented here indicate that the luminosity density at  $z \sim 0.1$  is about 10% larger than previous work with **CModel** magnitudes (Table 10.1), and about a factor of two larger than when based on Petrosian SDSS DR7 magnitudes (Figure 10.5). This differ-

ence is driven by the most luminous objects which are predominantly quiescent early-type galaxies (Figure 10.6). Since a number of authors now agree that SDSS pipeline magnitudes underestimate the true luminosity, and these more recent algorithms are in reasonably good agreement with the PYMORPH reductions used here (Figure 10.2), there is now good agreement that the bright end of the luminosity function may be substantially brighter than the SDSS DR7 Petrosian magnitudes suggest.

As one might expect given this analysis of the luminosity function, the (Chabrier IMF-based) stellar mass densities at  $z \sim 0.1$  are about 25% larger than previous work (Bernardi et al. 2010) with CModel magnitudes (Table 10.1), which was itself considerably larger than when based on the SDSS DR7 Petrosian magnitudes often used for this purpose (Figure 10.5). As a result, the estimates presented here have implications for studies of the evolution of the star formation rate, the growth of stellar mass in galaxies, the processes by which this mass was assembled, and Halo Model analyses of the  $M_* - M_{\text{halo}}$  relation. For example:

- Higher stellar mass density at  $z \sim 0$  resolves the tension with respect to the total mass density inferred from the integrated star formation rate (SFR), as noted in Bernardi et al. (2010).
- A higher number density of massive galaxies in the local Universe allows for a higher incidence of major (in addition to minor) mergers in driving the stellar mass growth of the most massive central galaxies at late times (Bernardi et al. 2011a,b; Shankar et al. 2013, e.g.). This conclusion rests, of course, on the quality of the determination of the high redshift stellar mass function, a task which is expected to be even more challenging than in the local Universe. In this respect, it is interesting that the  $z \sim 1$  counts at  $M_* \geq 10^{11.5} M_\odot$  of Carollo et al. (2013) are in rather good agreement with our  $z \sim 0$  (Ser-based) estimate, strongly limiting merger rates at the high mass end.

## 10. DEPENDENCE OF THE BRIGHT END OF THE LF AND MF ON THE FITTED PROFILE

---

- A higher number density at high masses would better match the stellar and dynamical mass functions, possibly reducing the need for a strong mass-dependent variation of the IMF (see, e.g. Fig. 23 in Bernardi et al. 2010), although the true extent of the latter statement relies on accurate dynamical mass measurements with appropriate effective radii and structure constants.
- The stellar mass function in the local Universe is one of the fundamental ingredients in popular semi-empirical models for populating halos with galaxies, such as the halo occupation and abundance matching techniques (Berlind and Weinberg 2002; Cooray and Sheth 2002; Leauthaud et al. 2012; Moster et al. 2013; Shankar et al. 2006; Vale and Ostriker 2004; Zehavi et al. 2011, e.g. ). If the more massive galaxies are more abundant, then Halo Model analyses will assign them to lower mass halos. Since lower mass halos are less strongly clustered, the most massive galaxies will be less strongly clustered than current models assume.
- This would also imply that the median baryon fraction at the high mass end may be significantly higher than previously thought. This may pose serious questions about the impact of feedback from active galactic nuclei, in the quasar and/or radio modes (Croton et al. 2006; Granato et al. 2004; Silk and Mamon 2012, e.g. ); even more so, if one considers that the baryon fraction of ellipticals at their formation epoch ( $z > 1$ ) must have been even higher.
- More stellar mass at the high mass end directly impacts studies of how the stellar fraction compares with that in the gas detected by X-ray and Sunyaev-Zeldovich experiments. e.g. , by decreasing the halo mass one should associate with a given stellar mass, it potentially reduces the discrepancy shown in Figure 9 of Planck Collaboration et al. (2013).
- These effects on the stellar to halo mass mapping go in the same direction as those from plausible changes to the IMF at the massive end (Bernardi et al. 2010, and

references therein), so the data may be indicating that a major revision of the results on the galaxy-halo mapping at the high mass end, is called-for.

The estimates of the amount of stellar mass in the most massive objects used in this study are at least  $2\times$  larger than other recent determinations (e.g. Baldry et al. 2012; Moustakas et al. 2013). Even though the estimates of the luminosity function are in much better agreement than this  $\phi(M_*)$ -based number suggests. Chapter 10.3 shows that the  $M_*$  values are similar only because they assume very different  $L$  and  $M_*/L$  values from one another. Therefore, until the field converges on what it believes to be reliable  $M_*/L$  estimates for the highest mass objects (see Mitchell et al. 2013, for ongoing discussion of this point), these results argue against calibrating models to published stellar mass functions: calibrating to the luminosity function is more robust, as it does not combine differences in  $L$  and  $M_*/L$  into a single number  $M_*$ .

When using these results, it is important to bear in mind that the PYMORPH estimates assume that the galaxy is either a **Ser**, or a **Ser-Exp** system. These both allow for substantial light beyond the core of the image, so there is some question as to just what it is that the profiles are fitting. The most massive objects tend to be the brightest cluster galaxies (BCGs), so some of the excess light returned by **Ser** and/or **Ser-Exp** fits may contain intracluster light (ICL). If so, then these estimates of the total stellar mass are appropriate for models which associate the ICL with the central galaxy. This may even be physically reasonable, since most of the accretion and stripping which occurred as the cluster assembled likely happened during accretion onto what is now the central object.

## Chapter 11

# Systematic effects on the size-luminosity relations of early and late type galaxies: dependence on model fitting and morphology

This chapter quantifies the systematics in the size-luminosity relation of galaxies in the SDSS main sample (i.e., at  $z \sim 0.1$ ) which arise from fitting different one- and two-component model profiles to the  $r$ -band images. This work was completed with Mariangela Bernardi, Vinu Vikram, Marc Huertas-Company, Simona Mei, Francesco Shankar and Ravi Sheth and the results were published in Bernardi et al. (2014).

Chapter 4.2.1 and 4.2.2 have shown that in objects brighter than  $L_*$ , fitting a two-component **Ser-Exp** model to what is really just a **Ser** results in a noisier, but unbiased, recovery of the input parameters. In contrast, fitting a **Ser** to what is truly a two-component system results in significant biases. For objects brighter than  $L_*$ , fitting a **Ser** profile to what is really a two-component **Ser-Exp** system leads to biases in the half-light radius, which is

increasingly overestimated as Sérsic index,  $n$ , of the fitted single component increases. The half-light radius is also overestimated at intermediate B/T( $\sim 0.6$ ) where a single component cannot effectively account for the shape of the profile created by two distinct components. However, the net effect on the  $R - L$  relation is small, except for the most luminous tail.

The dependence on morphology of incorrect fitting on the  $R - L$  relation is also examined. This analysis is one of the first to use Bayesian-classifier derived weights, rather than hard cuts, to define morphology. For the  $R - L$  relation Es, S0s and Sa's are early-types, whereas Sbs and Scds are late. Neither the early- nor the late-type relations are pure power-laws. Both show significant curvature, which is quantified.

This curvature seems to confirm that two mass scales are special for both early- and late-type galaxies:  $M_* \sim 3 \times 10^{10} M_\odot$  and  $2 \times 10^{11} M_\odot$ . Also, although the  $R_{\text{disk}} - L_{\text{disk}}$  and  $R_{\text{disk}} - M_{*\text{disk}}$  relations of disks of disk-dominated galaxies run parallel to the corresponding relations for the total light in late-types (i. e., they are significantly curved),  $R_{\text{bulge}} - L_{\text{bulge}}$  and  $R_{\text{bulge}} - M_{*\text{bulge}}$  for bulge dominated systems show almost no curvature (i. e., unlike for the total light of early-type galaxies). Finally, the intrinsic scatter in the  $R - L$  relation decreases at large  $L$  and/or  $M_*$  and should provide additional constraints on models of how the most massive galaxies formed.

## 11.1 Introduction

The spatial (and color) distribution of star light in a galaxy is thought to encode information about its formation history, so there has been considerable interest in developing accurate descriptions of the projected surface brightness distribution of galaxies. The work presented in this thesis, as well as SDSS, Simard et al. (2011), Lackner and Gunn (2012), etc. , cited throughout this thesis, seek to answer these questions.

Of course, it is not obvious that the light profile should be fit using a single component. The stellar kinematics in many galaxies indicate that the stars define more than one dynamical component. Examples include counter-rotating disks, as well as disk systems with

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

bulges or bars in their centers (e.g. Emsellem et al. 2011). Evidence for more than one component is often seen in the chemical composition as well (e.g. Johnston et al. 2012). In such galaxies, it is interesting to see if the light profile also indicates the presence of more than one component.

This has motivated studies which model the observed profile as the sum of an exponential and a de Vaucouleurs profile; what we will call the **deV-Exp** model. (Of course, since there are now more free parameters to be fit, better, higher resolution data are required. In this context, it is worth noting that Sérsic's initial motivation was to fit a functional form with fewer free parameters which would allow one to interpolate between two-component systems having varying fractions of an  $n = 4$  bulge and an  $n = 1$  disk. It is common to report the result of such two-component fits in terms of the fraction of the total light that is in the bulge (de Vaucouleurs) component: B/T. Correlations of these B/T values with other parameters (e.g. luminosity) are then used to constrain formation history scenarios.

On the other hand, if galaxies really are single component **Ser**, and one attempts to fit them with two component **deV-Exp** profiles, then one will infer an entirely spurious B/T value (the profile was, after all, just a single component). This spurious B/T will correlate with other parameters if  $n$  itself does, complicating the interpretation of such correlations. Indeed, some have argued that the evidence for two-components in the light profile is sometimes just a consequence of trying to fit what is really a single component **Ser** with a linear combination of exponential and de Vaucouleurs profiles (eg Graham et al. 2003), although this leaves unanswered the question of why dynamically or chemically distinct components do not leave a signature in the light.

In recent years, the correlation between size and luminosity for early-type galaxies has received much attention, because high redshift early-types appear to be more compact than their counterparts at low redshift (e.g. Bruce et al. 2012; Cimatti et al. 2008; Trujillo et al. 2006; van Dokkum et al. 2008). However, both the size and the luminosity estimates,  $R$  and  $L$ , are derived parameters, obtained by fitting to the observed surface brightness

## 11.1 Introduction

---

distribution. As a result, they depend on assumptions about the intrinsic shape of the surface brightness profile. For example, if the fit assumes that galaxies are made up of two components or just one, and if two, whether they are modeled as the sum of an exponential and a de Vaucouleurs profile, an exponential and a Sérsic, or two Sérsic profiles.

The main goal of this chapter is to quantify the systematics on the local  $R - L$  relation which are associated with the choice of a particular model. In practice, ‘local’ means the  $5 \times 10^5$  galaxies at  $z \sim 0.1$  in DR7 of the SDSS Main Galaxy sample (Abazajian et al. 2009). Because this sample is apparent magnitude limited ( $m_r < 17.75$ ), in practice, by  $R - L$  relation we always mean  $\log_{10} R$  fitted as a function of absolute magnitude (see Sheth and Bernardi 2012, for a simple description of the bias which would arise from fitting  $L$  as a function of  $R$ ). And  $R$  denotes the radius which encloses half the total light  $L$ . (For exponential disks, this radius is 1.67 times the scale length of the exponential. In the case of two components,  $R$  is a complicated function of the light in each component and the two scale radii.)

Our goal implies that we must fit the observed profiles and determine the associated  $R - L$  relation using a variety of different models. Chapter 11.2 summarizes the relevant properties of the SDSS DR7 sample we study. Chapter 11.3 compares the  $R - L$  relation based on single **Ser**, **Ser-Exp**, **deV-Exp**, and **deV** fits, showing that the relations from **Ser** fits (the standard to date) are offset to larger sizes and those from **deV** fits to smaller sizes, compared to those from the two-component fits.

There is no particular reason why systems supported by rotational motions should define the same  $R - L$  relation as those supported by random motions. Therefore, it makes little sense to speak of a single  $R - L$  relation for the entire galaxy population. Indeed, the  $R - L$  relation has long been known to depend on morphological type (e.g. Shen et al. 2003). Chapter 11.4 quantifies this morphological dependence, and then focuses on the differences between the relations defined by early-type bulge dominated systems, and

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

later type disk-dominated systems. Disk dominated galaxies have small bulges, and bulge dominated galaxies have extended second components.

### 11.2 The SDSS DR7 sample

In this chapter, we study the galaxies in the Seventh Data Release (DR7) of the SDSS Main Galaxy sample (Abazajian et al. 2009). This sample is the same as the sample defined in Bernardi et al. (2010) and used in Chapter 10. It contains about 260,000 SDSS galaxies having  $14.5 \leq m_{r\text{Pet}} \leq 17.7$ . The DR7 database provides crude estimates of the (galactic extinction corrected) apparent brightness and angular size of each galaxy in the catalog. We will use the SDSS-based `CModel` magnitudes and sizes (a weighted combination of separate fits to `Exp` and `deV` profiles) defined in Bernardi et al. (2010). When converting these to physical sizes and luminosities, we assume a flat  $\Lambda$ CDM model with  $\Omega_m = 0.3$  and a Hubble constant whose present value is  $H_0 = 70 \text{ km s}^{-1}\text{Mpc}^{-1}$ . The luminosities are  $k$ -corrected following Bernardi et al. (2003a). For colors we use SDSS DR7 `Model` magnitudes (corrected for galactic extinction). The database also provides estimates of the stellar velocity dispersion of each galaxy. We follow custom and correct these values for aperture effects following Bernardi et al. (2003a).

In the first half of this chapter, we describe a number of other estimates of the total light associated with each  $r$ -band image. However, recent work has focused on stellar masses  $M_*$  rather than luminosity  $L$ . Our estimated stellar masses come from combining our estimates of  $L_r$  with  $(M_*/L_r)$  estimated following Bernardi et al. (2010) assuming a Chabrier IMF.

#### 11.2.1 Estimates of photometric parameters

One of our primary goals is to quantify how the  $R - L$  relation depends on how  $R$  and  $L$  were estimated. This work uses the fits presented in Part I. It is conventional to speak of

the two components as being ‘bulge’ and ‘disk’ components; while this is accurate for disk-dominated systems (typically later-type galaxies), it may be better to think of the ‘disk’ component in bulge-dominated systems (typically early-type galaxies) as simply being a second component that is not necessarily a (thin, inclined) disk (e.g. Gonzalez et al. 2005; Oemler 1976; Schombert 1986).

Before moving on, we note that there is an analytic expression for the light enclosed within a given distance of the center of a single circular  $\text{S}\acute{\text{e}}\text{r}$  profile (e.g. Ciotti and Bertin 1999). From this, the half light radius can be obtained easily. However, if the object has axis ratio  $b/a \neq 1$ , where  $b$  and  $a$  are the half-light radii along the principal axes of the image, then the corresponding expression must be integrated numerically. Since this can be time-consuming, it is usual to approximate this case by using the expression for a circle, but with a suitably chosen effective circular radius. The most common choice is  $\sqrt{ba} = a\sqrt{b/a}$ , but Saglia et al. (2010) have recently shown that  $(b+a)/2$  is more accurate: for bulge dominated systems the difference matters little, but it does matter for disks. Therefore, we use  $(b+a)/2$  except in Chapter 11.3.1 where, to fairly compare with previous work, we use  $\sqrt{ba}$ .

This raises the question of what we should do when we have two components? A natural choice would be to circularize each component using its own  $(b+a)/2$ , and to then determine the half light radius of the sum of the circularized components, where each is weighted by the fraction of the total light that it contains. We have found that this approximation is quite accurate, so we use it throughout.

### 11.2.2 Morphologies

A secondary goal of this work is to quantify the role of galaxy type or morphology on the  $R - L$  relation. In practice this is difficult, because unambiguous determinations of the morphological type are not straightforward, although the task is slightly easier for bulge dominated systems. Previous work has used crude proxies for morphological type: these

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

include isophotal shape and central concentration (Strateva et al. 2001), the Sérsic index  $n$  (Shen et al. 2003), the color, spectral features, and some combination of the above (Baldry et al. 2004; Bernardi et al. 2003a).

In what follows, we use the Bayesian Automated Classifications (hereafter BAC) of Huertas-Company et al. (2011) which are available for our full DR7 sample. The BAC classifications are particularly interesting, because they are expressed as probabilistic weights (determined from an object's  $k$ -corrected  $g - r$  and  $r - i$  colors, and its isophotal shape and light concentration in the  $i$ -band)

We will study an ‘early-type’ sample defined on the basis of hard conservative cuts on two parameters which are available for each galaxy: the value of  $n$  returned by PYMORPH when fitting a `Ser` profile to the image, and the BAC probability  $p(E+S0)$  that the object is an early-type. We require

$$n > 3 \quad \text{and} \quad p(E + S0) > 0.85.$$

These cuts by no means select all early-type galaxies; they are simply designed to select a population which is very unlikely to be contaminated by later-types. Since our goal is to select objects of a single type, we are willing to sacrifice completeness for purity.

To assess how these BAC-based hard-cuts perform, we use the eye-ball classifications of Fukugita et al. (2007, ; hereafter F07) and of Nair and Abraham (2010, ; hereafter N10). These are based on analysis of a much smaller patch of the SDSS sky, and a brighter magnitude limit (e.g. F07 has only  $\sim 7000$  objects restricted to  $m_r < 16$ ; N10 has about twice as many), but for our purposes, the important point is that they are both magnitude limited.

Whereas BAC classifies galaxies into 4 (E,S0,Sab,Scd) morphological types, F07 use  $0 < T < 7$  in steps of 0.5. To convert, we assign E ( $T = 0$  and 0.5), S0 ( $T = 1$ ), Sa ( $T = 1.5$  and 2), Sb ( $T = 2.5$  and 3), and Scd ( $T = 3.5, 4, 4.5, 5$ , and 5.5). Similarly, N10 use the T-Type classification ( $-5 < T < 7$ ) from the modified RC3 classifiers; we assign E ( $T = -5$

## 11.2 The SDSS DR7 sample

---

Selection	E	S0	Sa	Sb	Scd
<b>EARLY-TYPES</b>					
Selected					
P(E+S0)> 0.85 AND $n > 3$	0.70	0.21	0.08	0.01	0
$n > 2.5$	0.44	0.18	0.20	0.13	0.05
Missed					
P(E+S0)< 0.85 OR $n < 3$	0.10	0.43			
$n < 2.5$	0.02	0.12			
<b>LATE-TYPES</b>					
Selected					
P(E+S0)< 0.15 AND $n < 3$	0	0.01	0.08	0.36	0.51
$n < 2.5$	0.01	0.04	0.11	0.33	0.45
Missed					
P(E+S0)> 0.15 OR $n > 3$			0.86	0.47	0.21
$n > 2.5$			0.75	0.41	0.16

**Table 11.1:** Eyeball morphological classifications from Fukugita et al. (2007). We set E ( $T = 0$  and  $0.5$ ), S0 ( $T = 1$ ), Sa ( $T = 1.5$  and  $2$ ), Sb ( $T = 2.5$  and  $3$ ), and Scd ( $T = 3.5, 4, 4.5, 5$ , and  $5.5$ ).

and  $T = -4$ ), S0 ( $T = -3, T = -2$  and  $T = -1$ ), Sa ( $T = 0, T = 1$  and  $T = 2$ ), Sb ( $T = 3$  and  $T = 4$ ), and Scd ( $T = 5, T = 6$  and  $T = 7$ ).

Table 11.1 shows the mix of F07 morphological types in samples which are defined by the hard cuts (on  $n$  and BAC  $p(\text{type})$ ) given above. Table 11.2 reports a similar analysis which is based on the eye-ball classifications of N10 instead of F07. These Tables show that 91% and 86% of the resulting sample are indeed either E and S0. In contrast, requiring only  $n > 2.5$  (as done in the past) yields a sample in which the E+S0 fraction is just 62% and 56% respectively. (A small fraction of the objects are Irregulars, which is why the numbers do not always add up to 100%.) Clearly our selection is much purer.

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

As a measure of its incompleteness, we also indicate the fraction of objects classified as Es and S0s which do not make the cut. These fractions are 10% and 43% for the F07 classifications, and 14% and 43% for N10. This ‘missed’ fraction is much smaller if we only require  $n < 2.5$ , but we believe the price to pay in purity is unacceptable.

The bottom halves of the two tables show a similar analysis of BAC and  $n$  cuts which are designed to produce a pure sample of later types. In this case requiring

$$n < 3 \quad \text{and} \quad p(E + S0) < 0.15$$

yields a sample in which Sa + Sb + Scd account for 95% of the objects. If we only require  $n < 2.5$ , and do not use the BAC-probability at all, then Sa + Sb + Scd account for 89% of the objects so, for later types, the use of the BAC analysis does not make such a dramatic difference.

Whereas the extremes of the morphological mix are relatively easy to define, Tables 11.1 and 11.2 indicate that the intermediate regime, the Sa/Sb class, will be difficult to define cleanly. Indeed, we have found that more than a third of the objects with BAC  $p(S0) > 0.6$  are Sa’s, and about a fifth of the objects with  $p(Scd) > 0.6$  are Sb’s. Conversely, of the objects which have  $p(Sab) > 0.6$ , about one third are Scd’s. For this reason, when we provide these BAC weight-derived relations in Tables 11.3 and 11.4, we refer to them as being for E, S0/Sa, Sa/Sb/Scd and Scd samples.

We will use these hard cuts, as well as the BAC-weights themselves. Note that all the results which follow are based on the full SDSS DR7 sample: we use the smaller F07 and N10 samples again only to perform sanity checks (as we did here) in Chapter 11.4.

### 11.3 The $R - L$ relation at $z \sim 0.1$

We now study how the  $R - L$  relation depends on the functional form for the surface brightness profile that was assumed when estimating  $R$  and  $L$ . We would especially like to compare the effects of fitting one versus two-component models to the images, and we

### 11.3 The $R - L$ relation at $z \sim 0.1$

---

Selection	E	S0	Sa	Sb	Scd
<hr/>					
EARLY-TYPES					
<hr/>					
Selected					
$P(E+S0) > 0.85$ AND $n > 3$	0.57	0.29	0.14	0	0
$n > 2.5$	0.32	0.23	0.29	0.12	0.03
<hr/>					
Missed					
$P(E+S0) < 0.85$ OR $n < 3$	0.14	0.43			
$n < 2.5$	0.02	0.07			
<hr/>					
LATE-TYPES					
<hr/>					
Selected					
$P(E+S0) < 0.15$ AND $n < 3$	0	0	0.15	0.41	0.39
$n < 2.5$	0.01	0.03	0.17	0.35	0.36
<hr/>					
Missed					
$P(E+S0) > 0.15$ OR $n > 3$			0.79	0.35	0.17
$n > 2.5$			0.71	0.33	0.10

**Table 11.2:** Eyeball morphological classifications from Nair and Abraham (2010) who used T-Type classification using the modified RC3 classifiers. We set E ( $T = -5$  and  $T = -4$ ), S0 ( $T = -3$ ,  $T = -2$  and  $T = -1$ ), Sa ( $T = 0$ ,  $T = 1$  and  $T = 2$ ), Sb ( $T = 3$  and  $T = 4$ ), and Scd ( $T = 5$ ,  $T = 6$  and  $T = 7$ ).

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

do this separately for the early and late type samples defined by the hard cuts on  $n$  and BAC- $p$ (type) described in Section 2.2.

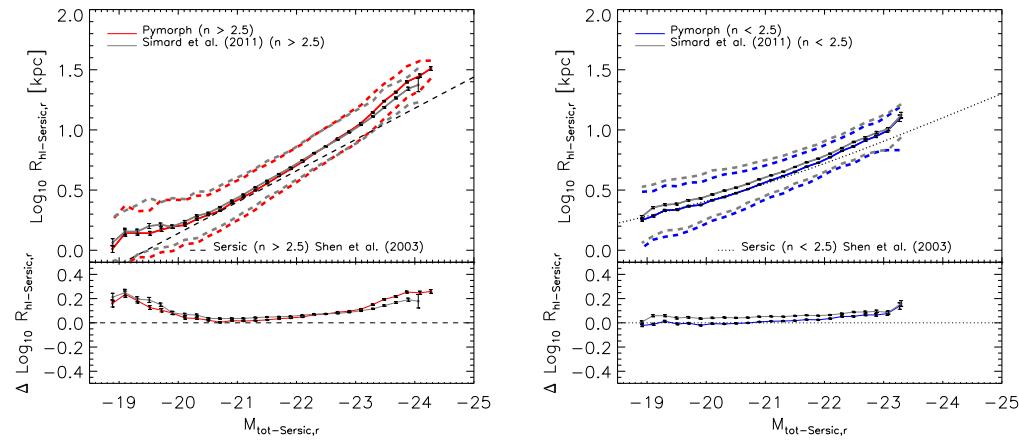
It is conventional to speak of these as ‘bulge’ and ‘disk’ components; while this is accurate for disk-dominated systems (typically later-type galaxies), it may be better to think of the ‘disk’ component in bulge-dominated systems (typically early-type galaxies) as simply being a second component that is not necessarily a (thin, inclined) disk (e.g. Gonzalez et al. 2005; Oemler 1976; Schombert 1986).

### 11.3.1 Single Ser fits and comparison with previous work

To connect with previous work, we begin with a comparison of the PYMORPH derived  $R - L$  relation based on **Ser**-derived parameters, with analogous fits from the literature. Using the objects in an earlier SDSS data release, Shen et al. (2003, ; hereafter S03) reported fits to the  $R - L$  relation for objects which had  $n > 2.5$  and  $n < 2.5$ , where  $R$ ,  $L$  and  $n$  were determined from a **Ser** fit to the light profile. Note that the **Ser** parameters used by S03 were estimated from a 1-dimensional radial surface brightness profile (**profMean**), measured in  $\sim 5 - 10$  azimuthally averaged annuli (Blanton et al. 2003a). Thus, it is expected to be significantly less accurate than a 2-dimensional fit to the whole galaxy image.

The S03 relations for  $n > 2.5$  and  $n < 2.5$  are shown as the dashed and dotted lines in the left and right hand panels of Figure 11.1, respectively. The red and blue symbols with error bars show our determination of the single-**Ser** based relation, where now  $R$ ,  $L$  and  $n$  are from our PYMORPH reductions, and the gray symbols and error bars show the  $R - L$  relation which follows from **Ser** fits performed by S11.

For objects with  $n < 2.5$ , the S11-derived relation runs parallel to that from S03, but is offset to larger sizes by 0.05 dex, whereas the PYMORPH-derived relation transitions from S03 at low luminosities to S11 at high luminosities. For objects with  $n > 2.5$  the PYMORPH-derived relation lies about 0.1 dex above, and is more curved than the fit reported by S03. The PYMORPH and S11 based relations depart significantly from S03 at the low and high



**Figure 11.1:** The r-band single-Ser based half-light radius ( $R_{hl}$ ) versus total absolute magnitude ( $M_{tot}$ ) relation for objects with  $n > 2.5$  (left) and  $n < 2.5$  (right). In the panel on the left, our PYMORPH determination is in good agreement with that based on single-Ser parameters from S11, but lies about 0.1 dex above, and is more curved than the fit reported by S03. Symbols with error bars (joined by a solid curve for clarity) show the median half-light radius in bins of absolute magnitude. Dashed lines show the 16th and 84th percentile. In the panel on the right (objects with  $n < 2.5$ ), except for the brightest objects, the PYMORPH relation lies systematically 0.05 dex below that of S11.

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

luminosity ends, where they curve upwards to larger sizes. For this reason, we are inclined to conclude that, at least at the bright end, S03 is slightly biased. At the low end the curvature could be due to contamination by later-type galaxies.

However, at the highest luminosities, and for objects with  $n > 2.5$ , the PYMORPH and S11 relations are also slightly but significantly different from one another.

### 11.3.2 Dependence on model fitting

Figure 11.1 showed the  $R - L$  relation derived from **Ser** fits to the two-dimensional surface brightness profile. We now compare these to relations based on SDSS fits to a single de Vaucouleurs profile; SDSS-based **CMode1** sizes defined by Bernardi et al. (2010), PYMORPH fits to a two-component **deV-Exp** model; PYMORPH fits to a two-component **Ser-Exp** model; and PYMORPH fits to a **Ser** profile. We expect the **Ser-Exp** reductions to return the least biased estimates of  $R$  and  $L$ , and hence of the  $R - L$  relation. We also show the corresponding  $R - M_*$  relations (with  $M_*$  estimated as described in Section 2).

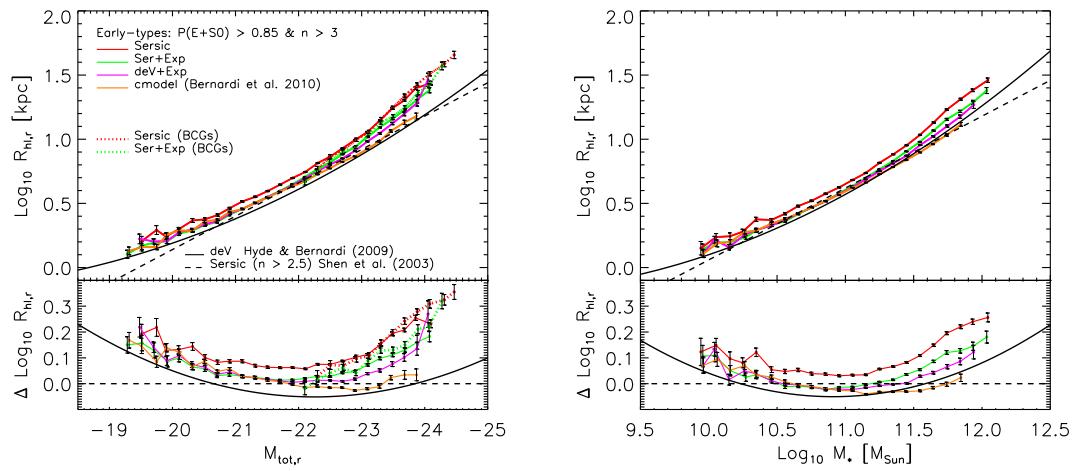
All of these relations show curvature which we quantify by fitting to

$$\left\langle \log_{10} \frac{R}{\text{kpc}} \middle| O \right\rangle = p_0 + p_1 O + p_2 O^2; \quad (11.1)$$

the coefficients of these fits for  $O = M_r$  and  $O = \log_{10}(M_*/M_\odot)$  are reported in Tables 11.3 and 11.4. Although  $p_1$  is the coefficient of the linear part of the relation, the slope on scale  $O$  is  $p_1 + 2p_2 O$ ; this is the value (at some characteristic  $O$ ) which should be compared with the slope of a linear fit to the relation.

#### 11.3.2.1 Early-types

The panel on the left of Figure 11.2 shows the  $R - L$  relation obtained for the early-type sample (i.e.  $n > 3$  and  $p(\text{E+S0}) > 0.85$ ) based on a number of single and two-component fits to the images. There are clear systematic differences between these relations, with the **Ser** and de Vaucouleurs models returning the relations with the largest and smallest sizes,



**Figure 11.2:** Dependence of derived size-luminosity (left panels) and size-stellar mass (right panels) correlations for early-type galaxies on the assumed surface brightness profile. Symbols with error bars (joined by a solid curve for clarity) show the median half-light radius in bins of absolute magnitude (left) and stellar mass (right). The SDSS fits to a single de Vaucouleurs profile return a relation with the smallest sizes; our PYMORPH fits to a **Ser** profile return the largest sizes. Of the relations which lie in between these two extremes, and which are almost indistinguishable at  $M < -21.5$ , the SDSS based **CModel** sizes (defined by Bernardi et al. 2010) are the smallest; those based on our PYMORPH fits to a two-component **deV-Exp** model are slightly larger; and those based on PYMORPH fits to a **Ser-Exp** model are largest. The curvature at the bright end appears to be due to an increasing incidence of BCGs, which define steeper relations (dotted lines) than the bulk of the early-type population. To highlight this curvature, bottom panel shows the ratio of all quantities in the top panel to the dashed line.

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

respectively. The various two-component based relations are in good agreement except at the highest luminosities ( $M_r < -22$ ), where the sample becomes increasingly contaminated by BCGs which are known to define steeper relations than the bulk of the population (e.g. Bernardi 2009; Bernardi et al. 2007).

The panel on the right shows a similar analysis of the  $R - M_*$  relation. Note that both  $R - L$  and  $R - M_*$  are significantly curved, consistent with previous work (e.g. Binggeli et al. 1984; Graham 2013). Of course, if the stellar population models used to estimate  $M_*/L$  are incorrect, or if the IMF is mass-dependent, then this will modify the curvature in  $R - M_*$ .

While not the main focus of the present work, we note that Figure 11.2 is consistent with recent work showing that a variety of other early-type galaxy scaling relations change slope at  $\sim 3 \times 10^{10} M_\odot$  and again at  $\sim 2 \times 10^{11} M_\odot$  (Bernardi et al. 2011b). Crude estimates of these two mass scales are given by the values of  $M_*$  at which a linear fit intersects the parabola given by the quadratic fit. For the Ser-Exp fits to the Early-types shown above the best linear fit has  $\langle R|M \rangle = -5.9155 + 0.598 \log_{10}(M_*/M_\odot)$ , so it crosses the associated parabola (parameters from Table 11.4) at  $\log_{10}(M_*/M_\odot) = 10.4$  and 11.6. We could assign uncertainties to these scales by propagating the uncertainties on the fitted coefficients  $p_i$ , but these are smaller than the systematics associated with this particular choice of defining the mass scales. Another estimate comes from adjusting the amplitude of the straight line fit so that it is tangent to the parabola, and then identifying the scales on which the parabola lies sufficiently far from it. Figure 11.3 shows such an analysis for the Ser-Exp  $R - M_*$  relation; the vertical lines show  $3 \times 10^{10} M_\odot$  and  $2 \times 10^{11} M_\odot$ . This demonstrates our main point – that the existence of these two scales is not an artifact of the model used to estimate  $R$  and  $L$ .

### 11.3 The $R - L$ relation at $z \sim 0.1$

---

Sample/fit	$p_0$	$p_1$	$p_2$
Ser (Early-types)	12.8145	1.3788	0.0377
Ser (Late-types)	8.4847	0.9092	0.0254
Ser ( $n > 2.5$ )	8.1624	0.9821	0.0292
Ser ( $n < 2.5$ )	4.7207	0.5601	0.0173
Ser (E)	7.0946	0.8650	0.0262
Ser (S0/Sa)	10.9232	1.2218	0.0344
Ser (Sa/Sb/Scd)	13.9656	1.4694	0.0395
Ser (Scd)	12.6494	1.3128	0.0352
Ser-Exp (Early-types)	8.6032	0.9979	0.0290
Ser-Exp (Late-types)	7.3204	0.7929	0.0226
Ser-Exp ( $n > 2.5$ )	6.0716	0.7770	0.0242
Ser-Exp ( $n < 2.5$ )	4.2848	0.5151	0.01615
Ser-Exp (E)	7.4437	0.8922	0.0266
Ser-Exp (S0/Sa)	9.6010	1.0903	0.0311
Ser-Exp (Sa/Sb/Scd)	9.3135	1.0182	0.0286
Ser-Exp (Scd)	7.8056	0.8396	0.0237
Ser-Exp (Early-type-Bulges)	-2.0733	0.0956	0.0098
Ser-Exp (Late-type-Disks)	6.4982	0.6934	0.0199

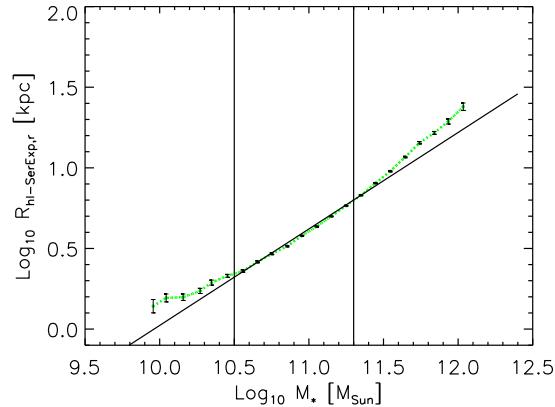
**Table 11.3:** Luminosity-size relation. Early-types:  $p(E+S0) > 0.85$  and  $n > 3$ . Late-types:  $p(E+S0) < 0.15$  and  $n < 3$ . Early-type-bulges: the bulge half-light radius versus the bulge luminosity for galaxies with  $p(E+S0) > 0.85$  and  $n > 3$ . Late-type-disks: the disk half-light radius versus the disk luminosity for galaxies with  $p(E+S0) < 0.15$  and  $n < 3$ .

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

Sample/fit	$p_0$	$p_1$	$p_2$
Ser (Early-types)	19.0933	-3.9536	0.2070
Ser (Late-types)	13.0054	-2.6438	0.1393
Ser ( $n > 2.5$ )	14.4995	-3.1767	0.1742
Ser ( $n < 2.5$ )	8.6098	-1.8301	0.1014
Ser (E)	13.6593	-2.9799	0.1635
Ser (S0/Sa)	20.1092	-4.1549	0.2166
Ser (Sa/Sb/Scd)	22.3082	-4.4655	0.2275
Ser (Scd)	17.9815	-3.6102	0.1862
Ser-Exp (Early-types)	13.4131	-2.9324	0.1607
Ser-Exp (Late-types)	11.2699	-2.3026	0.1227
Ser-Exp ( $n > 2.5$ )	12.5026	-2.7875	0.1551
Ser-Exp ( $n < 2.5$ )	9.5210	-1.9963	0.1090
Ser-Exp (E)	12.8394	-2.8246	0.1557
Ser-Exp (S0/Sa)	19.2830	-3.9866	0.2079
Ser-Exp (Sa/Sb/Scd)	18.6150	-3.7425	0.1922
Ser-Exp (Scd)	11.7537	-2.3957	0.1271
Ser-Exp (Early-type-bulges)	4.0853	-1.4159	0.0992
Ser-Exp (Late-type-disks)	17.9763	-3.5683	0.1831

**Table 11.4:** Stellar mass-size relation. Early-types:  $p(E+S0) > 0.85$  and  $n > 3$ . Late-types:  $p(E+S0) < 0.15$  and  $n < 3$ . Early-type-bulges: the bulge half-light radius versus the bulge stellar mass for galaxies with  $p(E+S0) > 0.85$  and  $n > 3$ . Late-type-disks: the disk half-light radius versus the disk stellar mass for galaxies with  $p(E+S0) < 0.15$  and  $n < 3$ .



**Figure 11.3:** The  $R - M_*$  relation for **Ser-Exp** fits, compared to a power-law showing  $R \propto M_*^{0.6}$ . Vertical lines show the two mass scales identified by Bernardi et al. (2011b) on the basis of a variety of other early-type galaxy scaling relations; the  $R - M_*$  relation clearly curves away from the power law at these scales.

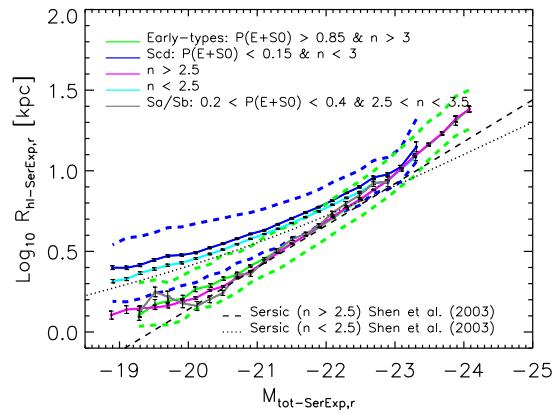
### 11.3.2.2 Comparison of early- and late-types

We have repeated this analysis for the late-type sample (defined by requiring  $n < 3$  and  $p(E+S0) < 0.15$ ). Although we do not show the corresponding plots here, we again see curvature (coefficients of fits to equation 11.1 are given in Tables 11.3 and 11.4). Rather, we illustrate this in Figure 11.4, which compares the **Ser-Exp**-based  $R - L$  relation for our way of selecting early- and late-type samples, with the more traditional cuts on  $n$  (larger or smaller than 2.5). The two ways of selecting the samples lead to very similar results, with the low luminosity early-types having smaller sizes, but defining a steeper relation, so they would cross the  $R - L$  relation of late-types at about  $M_r < -23$  (beyond which there are few late-types anyway).

We have also selected an intermediate-type population (i.e. Sa/Sb) by requiring  $2.5 < n < 3.5$  and  $0.2 < p(E + S0) < 0.4$ . Notice that this sample defines the same  $R - L$  relation

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---



**Figure 11.4:** Similar to previous figure, but now objects are selected using different hard cuts which define early-, late- or intermediate-type (i.e. Sa/Sb) samples. Symbols with error bars (joined by a solid curve for clarity) show the median half-light radius in bins of absolute magnitude. Dashed lines show the 16th and 84th percentile. Note that this definition of intermediate's (i.e. Sa/Sb) yields an  $R - L$  relation which is essentially the same as for the population with  $n > 2.5$ .

---

## 11.4 Dependence on Bayesian Automated Classifier morphologies

---

as when we require our early-type selection (i.e.,  $n > 3$  and  $p(E + S0) > 0.85$ ), as well as that when we only require  $n > 2.5$ ; we return to this in Chapter 11.4.

## 11.4 Dependence on Bayesian Automated Classifier morphologies

Prior to now, we used a hard cut on the BAC probability to determine morphology. Since this is not quite in the spirit of why such probabilities were generated in the first place, this section shows the result of using the BAC probabilities as weights when determining the  $R - L$  scaling relation. As this is one of the first such studies of weight-based scaling relations, and Tables 11.1 and 11.2 indicate that the relation between these weights and eye-ball classifications can be complex in detail, we include an explicit comparison of these BAC-based scaling relations with those based on eye-ball classifications.

Figure 11.5 shows the size-luminosity (left) and size- $M_*$  (right) relations obtained by weighting objects by  $p(\text{type})$  as determined by BAC. The results of fitting equation (11.1) to these curved relations are reported in Tables 11.3 and 11.4.

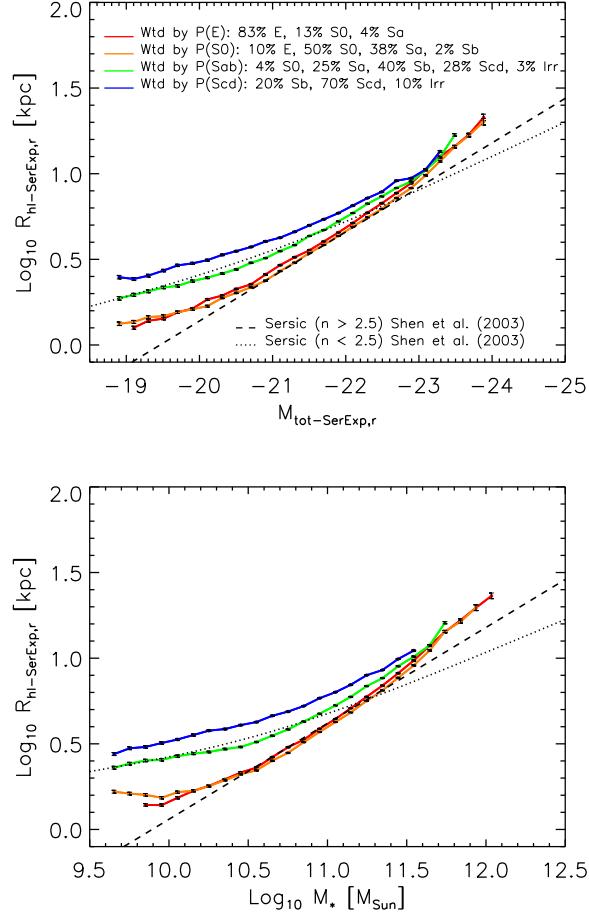
Notice that galaxies weighted by  $p(\text{Sab})$  define a relation which lies between that defined by  $p(\text{Scd})$  on the one hand and  $p(\text{E})$  and  $p(\text{S0})$  on the other. However, the Sa/Sb class is difficult to define (c.f. discussion of Tables 11.1 and 11.2).

### 11.4.1 Small but statistically significant difference between Ellipticals and S0s

Above, we noted that there is essentially one  $R - L$  relation for E, S0 and Sa galaxies. However, our sample is large enough to detect small but significant differences within the early-type (E and S0) sample. A closer look at Figure 11.5 indicates that S0s are slightly smaller than Es of the same luminosity. Figure 11.6 shows that this offset is about 0.06 dex,

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

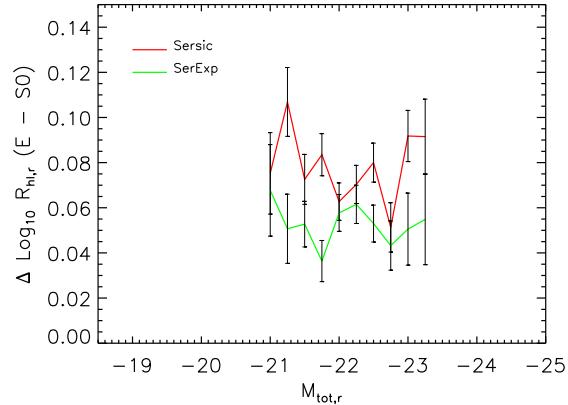
---



**Figure 11.5:** Ser-Exp based size-luminosity (top) and size- $M_*$  (bottom) relations, obtained by weighting objects by the BAC  $p(\text{type})$ . The low  $L$  or  $M_*$  part of the relation for Scds has the same slope as that reported by S03 for their  $n < 2.5$ ; and the intermediate  $L$  or  $M_*$  part has the same slope they report for  $n > 2.5$ . The relations for S0s are very similar to those for Es, and the Sab relations always lie between the E and Scd relations. Numbers in legend show the percentage of E, S0, Sa, Sb, Sbc and Irr galaxies classified by F07 with BAC  $p > 0.6$ . Using this selection we miss about 18% of Es, 60% of S0s, 64% of Sab (37% Sa and 27% Sb) and 56% of Scd, respectively.

## 11.4 Dependence on Bayesian Automated Classifier morphologies

---



**Figure 11.6:** At fixed luminosity, Es tend to be about 0.06 dex larger than S0s, although this offset depends slightly on how  $R$  and  $L$  were determined.

although it depends slightly on how  $R$  and  $L$  were determined. This is particularly interesting in view of recent work at  $z \sim 1$ , based on the S11 reductions, which shows a similar offset of about 15% for the SDSS sample growing to  $\sim 40\%$  at  $z \sim 1$  (Huertas-Company et al. 2013b). Both the sign of the trend and its evolution deserve further study, because, the sign of the trend is not what one might naively have expected.

### 11.4.2 Scatter in $\log(\text{size})$ around the mean relation for early-types

In addition to the mean  $R - L$  relation, the scatter around the mean relation is expected to constrain galaxy formation models (Shankar and Bernardi 2009; Shankar et al. 2012; Shen et al. 2003). As we show below, our analysis of the mean  $R - L$  relation allows us to make two interesting statements about the scatter around it for early-type galaxies: (i) there is intrinsic scatter and (ii) it is smaller at the high mass end.

The top two jagged solid curves in the left panel of Figure 11.8 show the measured scatter around the mean  $R - L$  relation for SDSS early-types, when  $R$  and  $L$  are determined by

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

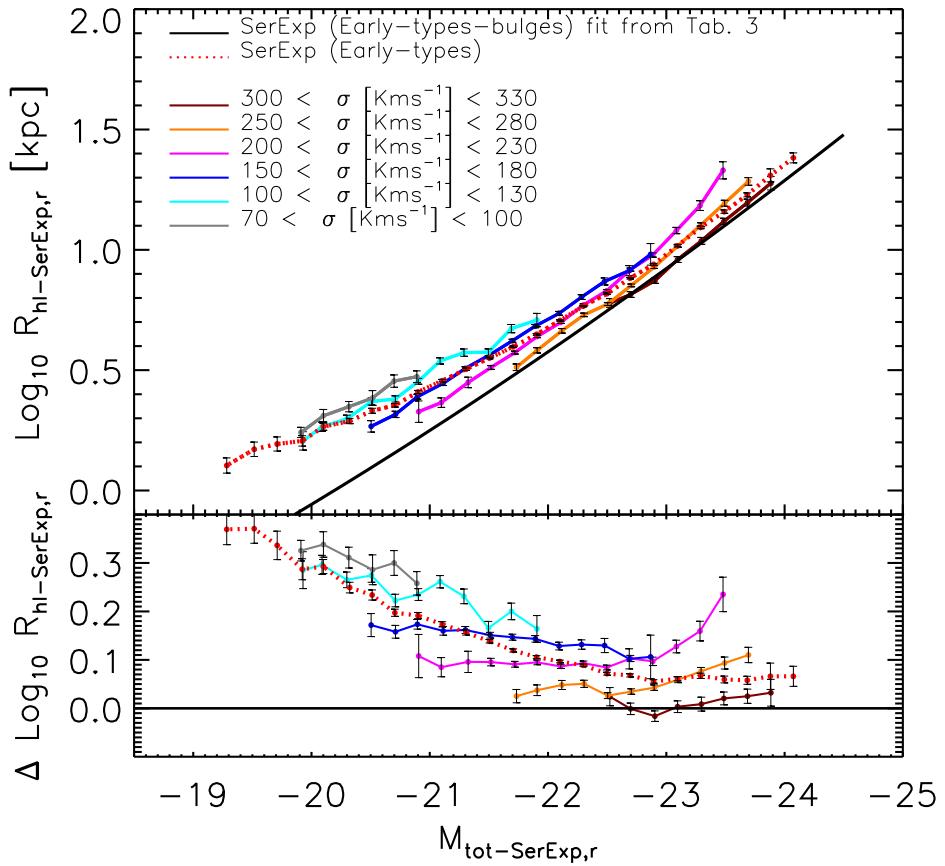
fits to a single **Ser** (larger scatter) and to a **Ser-Exp** model (lower scatter). This scatter is broader than the intrinsic one, because it includes a contribution from the measurement errors. For comparison, the smooth black curve shows the corresponding (error broadened) scatter reported by S03. It is in reasonably good agreement with ours, except at the faint end, where we believe the enhanced scatter is due to increased contamination by spirals, for which the scatter is larger (as we show later).

To estimate the intrinsic scatter, we must account for the broadening due to measurement errors. We estimate the errors on the sizes from fitting to the objects in the mock catalogs used in Chapter 4, where we know the input values. The dotted and dashed lines show these simulation-based estimates of the measurement error on the sizes for an early-type sample. The lowest dotted line shows the RMS scatter in  $\log_{10} R$  around the input value if the input profile is a **Ser**, and we fit it with a **Ser**. In this, and all the cases which follow, we show this scatter as a function of the fitted (as opposed to the input) absolute magnitude. The other dotted line, which lies only slightly above the previous one, shows what happens if we fit a **Ser-Exp** with a **Ser-Exp**. These curves certainly underestimate the full measurement error, since they are based on fits to smooth images, whereas real images may be lumpy, have spiral arms, etc.

To get an idea of the magnitude of such effects, the dot-dashed curve in the left panel of Figure 11.8 shows the RMS difference between PYMORPH and S11 sizes returned by two-component **Ser-Exp** fits to SDSS images, plotted as a function of the PYMORPH **Ser-Exp** absolute magnitude for the early-type sample. This is almost certainly an overestimate of the measurement error on the sizes, since it includes systematic effects which arise from the differences between PYMORPH and S11; we have included it just to get a sense of the overall magnitude with which systematic rather than random errors might affect the scatter in the  $R - L$  relation.

The two dashed curves show results from using PYMORPH to fit a **Ser** with a **Ser-Exp** (lower) and a **Ser-Exp** with a **Ser** (upper). The differences between these and the dotted

## 11.4 Dependence on Bayesian Automated Classifier morphologies



**Figure 11.7:** At fixed velocity dispersion  $\sigma$ , the  $R - L$  relation is almost a pure power law whose slope is almost the same for all  $\sigma$  (the black solid line shown is very close to  $R \propto L^{0.83}$ ), but whose zero-point increases as  $\sigma$  decreases. In contrast, the relation upon averaging over all  $\sigma$  (red dotted curve) is much more curved. The bottom panel shows the result of dividing the measurements in the top panel by the black solid line.

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---

curves give an idea of the effect on the scatter of fitting an incorrect model to the data. The upper dashed curve is particularly interesting, in view of the fact that the **Ser-Exp** model is more realistic (see Chapters 6-8) whereas the **Ser** model is most often fit. Clearly, subtracting it in quadrature from the upper solid curve will lead to negative values at large luminosities. This is shown by the lower of the two curves in the bottom panel: at  $M_r < -23$  or so, the intrinsic scatter is consistent with zero. This, of course, does *not* mean that the  $R - L$  relation is intrinsically a line with negligible scatter. Rather, it is entirely a consequence of fitting an incorrect model.

Recently, Nair et al. (2011) have used just such an argument to claim that the  $R - L$  relation has no scatter. However, their argument is based on Petrosian sizes and luminosities; these are known to be inaccurate at large  $L$ , so the analysis above illustrates why their claim should be treated with skepticism. Indeed, the upper curve shows the result of subtracting (in quadrature) the upper dotted curve from the lower solid one, since both these are based on fitting to what we argued were more realistic models of the light profile (i.e. **Ser-Exp**). In this case, the intrinsic scatter is well-behaved: although it decreases steadily with  $M_r$ , it does not go negative.

Of course, since our estimate of the measurement error is really an underestimate, it is still possible that the intrinsic scatter is smaller than we show. Therefore, we turn to what we believe is a much more effective way of showing that there is some intrinsic scatter. This method studies if the residuals from the relation correlate with other parameters, once correlations between the measurement errors have been accounted for. If they do, then there must be some intrinsic scatter.

Figure 11.7 shows the  $R - L$  relation for a number of narrow bins in velocity dispersion  $\sigma$ . At fixed  $\sigma$ , the  $R - L$  relation is a power law whose slope is 0.85 for all  $\sigma$  but whose zero-point increases as  $\sigma$  decreases. The scaling is quite well described by

$$\langle R | L, \sigma \rangle \propto L^{0.85} \sigma^{-0.73}. \quad (11.2)$$

The dependence on  $\sigma$  shows clearly that the scatter around the mean  $R - L$  relation correlates with  $\sigma$ ; it is not all due to measurement errors. We are not the first to have made this point (see Bernardi et al. 2003b, for an explicit discussion of this, although the long-studied Fundamental Plane is the result of this correlation), so it is surprising that Nair et al. (2011) ignored it. The slope of 0.85 (at fixed  $\sigma$ ) is consistent with previous work (Bernardi 2009; Bernardi et al. 2003a). While steeper than the slope of 0.64 associated with averaging over all  $\sigma$ , it is less than unity.

Finally, one final observation is that the intrinsic scatter appears to be smallest for the most luminous objects. Since it is commonly believed that mergers will affect the scatter of scaling relations such as this one, our overestimate of the intrinsic scatter in the  $R - L$  relation provides a new constraint on models of how the most massive galaxies must have formed. For example, Shen et al. (2003) argue that many minor mergers may be more consistent with the shape and scatter of the  $R - L$  relation than are few major mergers. Other work has also explored constraints which come from the scatter (Shankar et al. 2012); it will be interesting to revisit this question in light of the mass-dependence we believe we see.

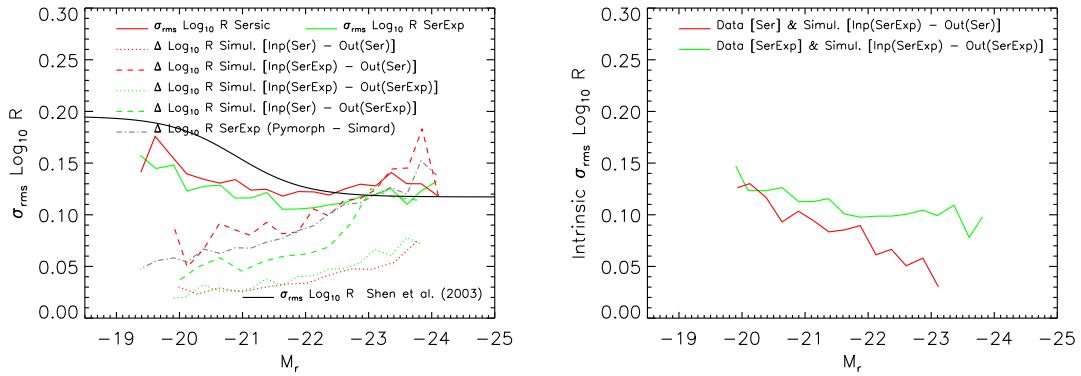
## 11.5 Summary

We used our automated image decomposition algorithm PYMORPH to study the effects of systematics in the size-luminosity relation of galaxies in the SDSS main sample (i.e. at  $z \sim 0.1$ ) which arise from fitting different models to the images.

We studied how the  $R - L$  relation depends on galaxy morphology. Our Figure 11.5 shows one of the first uses of Bayesian classifier-based weights in the estimation of the  $R - L$  scaling relation for different morphologies (e.g. Aguerri et al. 2012). We found that, even if we allow for finer bins in morphology, there seem to be only two fundamental  $R - L$  relations, both of which are slightly but statistically significantly curved (Figures 11.4 and 11.5 and Tables 11.3 and 11.4).

## 11. SYSTEMATIC EFFECTS ON THE SIZE-LUMINOSITY RELATIONS OF EARLY AND LATE TYPE GALAXIES: DEPENDENCE ON MODEL FITTING AND MORPHOLOGY

---



**Figure 11.8:** Left: Observed scatter around the mean  $\langle R|L \rangle$  relations for early-types based on fitting **Ser** (solid red) and **Ser-Exp** models (solid green) to the images. Black solid curve shows the corresponding (error-broadened) measurement from Shen et al. (2003). Dashed and dotted curves show a number of estimates from simulations of the measurement errors (see text for details). Grey dot-dashed line in top panel shows the RMS difference between PYMORPH and S11 sizes (both based on fitting a two-component **Ser-Exp**). Right: Estimate of the intrinsic scatter around the **Ser** (lower, red curve) and **Ser-Exp** (upper, green) derived relations for early-types, obtained by subtracting in quadrature the red-dashed and green dotted curves from the corresponding red and green solid curves shown in the left panel.

## 11.5 Summary

---

Of course, a closer inspection does reveal subtle dependences on morphology. Amongst early-types, S0s tend to be about 0.06 dex smaller than Es of the same luminosity (Figure 11.6). This difference is smaller than the  $\sim 40\%$  reported by Huertas-Company et al. (2013a) at  $z \sim 1$ . We argued that this subtle difference argues against interpreting the second component in **Ser-Exp** fits as a disk. It is particularly interesting in view of the fact that the two types show very different trends as a function of age Bernardi et al. (2010), so we expect that it, and its evolution, should yield interesting new constraints on models of how early-type galaxies assembled their stellar mass. Similarly, amongst late-types, faint Sbs tend to be  $\sim 0.1$  dex smaller than Scds of the same luminosity, but these differences decrease as luminosity increases.

Whether or not these scales are associated with the onset of different physics is the subject of ongoing debate. For example, Graham and Worley (2008) have pointed out that curvature in the  $R - L$  relation may arise as a consequence of linear  $n - L$  and central surface brightness- $L$  relations; the linearity of these would not be suggestive of changing physics along the relation. Our **Ser**-based fits do indeed yield pure power laws for these relations, at least for Early-types above  $M_* = 3 \times 10^{10} M_\odot$ . But they are not pure power laws for our **Ser-Exp** reductions, which we believe we have demonstrated are more appropriate.

Bernardi et al. (2011b) attribute the change at  $M_* \sim 2 \times 10^{11} M_\odot$  to a change in the assembly histories – to ones in which major dry mergers become important. So it is interesting that it is at this mass scale that the bulge and total  $R - M_*$  relations become the same, despite being very different at smaller masses. This is particularly remarkable in light of recent work showing that early-types below this mass scale tend to be fast rotators Cappellari et al. (2013). It may be that our **Ser-Exp** bulge-disk decompositions of the images are reflecting this change in the kinematics. Using the PYMORPH results, colors and color gradients of the individual components are topics for future study.

## **Part III**

# **Future Work and Concluding Remarks**

# Chapter 12

## Future Work

### 12.1 Further exploration of *g*- and *i*-bands

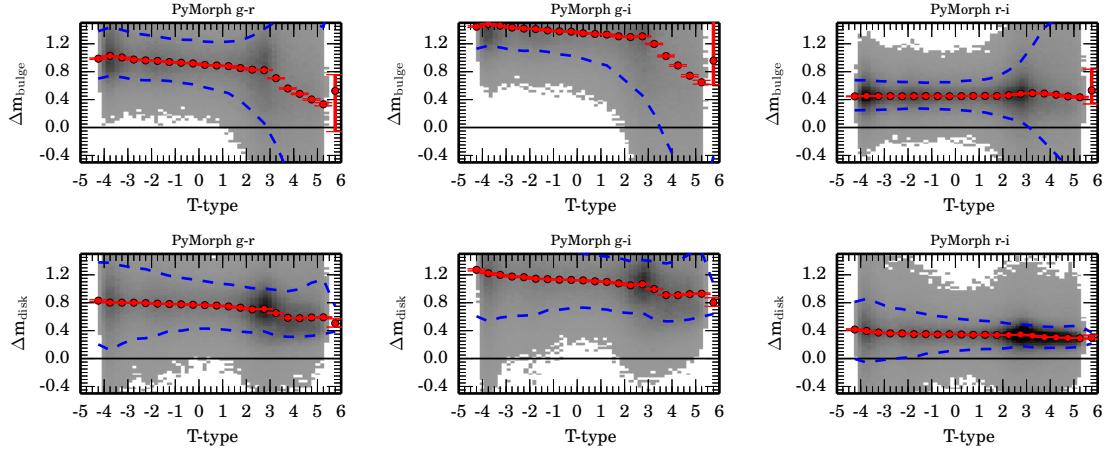
A full exploration of the *g*- and *i*-band data produced as part of this work is the next step with this project. These results as well as the total colors, bulge and disk colors, and the interband relationships will be explored further.

Figure 12.1 shows the bulge (top row) and disk (bottom row) colors as a function of T-type for the three combinations of bands g-r, g-i, and r-i in the left-hand, centre, and right-hand columns, respectively. Colors are stable across the T-types for the bulge and disk components using the r-i color. The g-r and g-i colors show trends toward bluer colors. The trend is small between  $T = -5$  and  $T = 3$  changing roughly 0.2 mags across the range. There is a break in the trend at  $T = 3$  and later types show bluer bulges by 0.5 mags compared to those at  $T = 3$ . This change corresponds to the transition from majority two-component to majority “pure disk” galaxies seen in Figure 8.3. The transition is also present in the disk colors, although it is somewhat smaller.

Interband relationships are particularly interesting for future analysis. Choosing an appropriate functional relationship between the fitted parameters in the different bands is useful for constraining bands with low S/N. For instance, the MEGAMORPH project (Häußler

## 12. FUTURE WORK

---



**Figure 12.1:** Bulge (top row) and disk (bottom row) colors for the three combinations of bands g-r, g-i, and r-i in the left-hand, centre, and right-hand columns, respectively. Median values are plotted in as red symbols with 95 per cent CI as error bars. The 68 per cent contours are shown as blue dashed lines.

et al. 2013) uses the same GALFIT fitting routine used in this work and seeks to automate multiband fitting, using information from neighboring bands to improve fitting stability and accuracy in bands with low S/N. Neighboring bands are used to constrain fitting by requiring variation in the fitting parameters (e.g. bulge radius, Sérsic index, etc.) to vary across the bands according to low-order polynomial functions. This is likely appropriately smooth, as the figures in Chapters 6.2 and 8 show, there is a systematic change in the parameters across the bands. However, more optimal methods may exist. For instance, also informing the direction of the change across the bands may be useful (e.g. the B/T is expected to decrease toward the rest-frame blue bands and increase toward redder bands in the rest-frame). Further exploration of this is necessary as the next generation of surveys (DES and LSST) begin to map out more dim galaxies.

## 12.2 Pushing to higher resolution and redshift

The simulations in Chapter 4 showed that improving the S/N as well as the resolution offered substantial improvement in the accuracy of the fitted parameters. This was particularly true in the case of two-component models. SDSS reached limits in fitting accuracy for two-component fits at magnitudes dimmer than the apparent magnitude cut of the spectroscopic sample. However, Chapter 6.3 showed that many of the bulges fit in this catalog have radii below the PSF radius and may suffer from fitting bias as a result. PSF effects reduce the accuracy of the fitted radius as well as the sensitivity of the fit to profile Sérsic index. Newer surveys offer better resolution and S/N relative to the SDSS data (e.g. DES has almost a factor of 2 better seeing in the  $r$ -band). Extending this study to newer surveys will improve the fitted models further lowering the uncertainty on the LF, other scaling laws, and B/T values; all of which will improve constraints on formation models.

In addition, newer surveys offer the opportunity to extend the fits to galaxies at higher redshift. For example, DES will find many galaxies near  $z \sim 0.5$  at similar physical resolution(i. e., kpc per image pixel rather than arcsec per image pixel) and S/N as the SDSS galaxies studied here. This offers an important opportunity to extend the analysis to higher redshifts where galaxy evolution is more active. Constructing an accurate LF at these different redshifts would constrain merger rates and relative fractions of major vs minor mergers through evolution of LF, MF, and other scaling relations.

## 12.3 Exploring the Green Valley

The green valley (e.g. Baldry et al. 2004; Strateva et al. 2001) was only briefly explored in Chapter 8.5. This region of parameter space offers interesting opportunities to explore the mechanisms for galaxy evolution from actively star-forming (blue galaxies) to passive galaxies (i. e., “red and dead”) galaxies. Recently, Schawinski et al. (2014) challenged the idea of the green valley being a common transitional state for galaxies by combining Galaxy

## **12. FUTURE WORK**

---

Zoo data (Lintott et al. 2011) with SDSS and GALEX UV data. Galaxy Zoo uses citizen scientists to classify galaxies using somewhat subjective scales in place of numerical values (e.g. How prominent is the bulge? Extremely, moderately, barely evident, ...). Numerical measurements of B/T produced by this catalog would offer a more quantitative approach that either supports or questions these results.

# Chapter 13

## Conclusion

A catalog of `deV`, `Ser`, `deV-Exp`, and `Ser-Exp` galaxies was constructed for the SDSS DR7 spectroscopic sample using the PYMORPH pipeline, including the SExtractor and GALFIT programs, to perform 2D decompositions (see Chapters 2 and 3). A physically motivated flagging system that removes poor fits and accurately identifies pure bulge, pure disc, and two-component systems was presented (Chapters 5.3 and 8.1). After applying the flagging system to the `Ser-Exp` fit, about 94 per cent of the fitted sample as having reliable total magnitude and half-light measurements in the *r*-band. About 39 per cent of the sample are two-component galaxies with well-behaved components. An additional  $\approx 11$  per cent may be two-component fits, but with difficult-to-interpret components. The remaining 44 per cent are pure bulge and disc galaxies.

The fits were compared to various results from the literature in Chapter 7. These comparisons showed that some measurable systematic differences exist in sky brightness as well as percent-level systematic differences in size and magnitude. 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. Chap-

## 13. CONCLUSION

---

ter 4, Davari et al. 2014; Mosleh et al. 2013) 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 catalog using the flags (i. e., bulge, disc, or two component) also correlate well with estimated  $T$ -types (Section 8.2). The fits of this catalog also behave appropriately in B/T, bulge size, and bulge Sérsic index with respect to  $T$ -type (Section 8.3.1).

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 catalog.

- Bias due to resolution effects for the **Ser** bulge components of the **Ser-Exp** fits is likely present in the catalog. 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 6.3).
- 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 8.4).
- 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 8.3.2 and Figures 8.6 and 8.7) 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.

After presenting and analyzing the catalog, two applications of the catalog were presented in Part II These applications applied the catalog to test and update important

---

predictions of formation models, namely the luminosity function (Chapter 10) and the size-luminosity relation (Chapter 11).

As Astronomy becomes more data-rich and newer surveys become main sources of astronomical information (e.g. LSST), large-scale numerical analysis similar to this catalog will become more important and also more feasible, providing further refinement of models of galaxy formation and evolution.

# Glossary

## Astronomical Terms

**$L_*$**  The characteristic luminosity above which the number of galaxies decline exponentially in the LF, page 7

**Cepheid Variables** Unstable stars that have a periodic pulsation. The period of the pulsation is related to the brightness of the star in a known way., page 1

**K-correction** The correction to broadband magnitudes to account for cosmological redshift shifting the spectrum of the observed object relative to the observer's telescope filter band, page 7

**LF** The function describing the number density of galaxies at a given magnitude. This can vary with environment or redshift. See Johnston (2011) for a review of the LF., page 5

**$M/L$**  The ratio of mass-to-light in a galaxy or other astronomical object., page 8

**MF** Similar to the LF, but using additional assumptions to convert the galaxy luminosities to masses. This is easier to

compare to theoretical models and numerical simulations, although more difficult to estimate., page 8

## Symbols

$b/a$  The ratio of semi-minor to semi-major axis for ellipses, page 35

$B/T$  The B/T ratio is the ratio of light contributed by the bulge of a galaxy to the total light within the galaxy. This parameterizes the relative strength of the bulge and disk components in the galaxy and correlates with morphology., page 4

$\chi^2_\nu$  The chi-squared per degree of freedom, a common measure of goodness-of-fit, page 40

$\mu_{50}$  The Petrosian half-light surface brightness, page 23

## Surveys and Catalogs

**DES** The Dark Energy Survey, page 13

**GAMA** Galaxy Mass and Assembly survey presented in (Barden et al. 2012), page 94

**H2011** H2011 is an automated morphological classification that used a Bayesian SVM algorithm to classify all galaxies in the spectroscopic sample based on data available as part of the SDSS DR7. Objects are matched to our catalog based on the SDSS DR7 `objID`., page 28

**LSST** The Large Synoptic Survey Telescope, page 13

**Men14** A reanalysis of S11, see Mendel et al. (2014), page 187

## GLOSSARY

---

<b>S03</b>	The study of size-luminosity relation of SDSS galaxies published in Shen et al. (2003), page 282	many distinct cells on it called pixels. The pixels produce free electrons at a rate proportional to the amount of light falling on the chip. The device is constructed such that these free electrons can be moved around and counted by the astronomer taking data., page 16
<b>S11</b>	The work of Simard et al. (2011) was a previous attempt at decomposing the sample used here using different software., page 23	
<b>SDSS</b>	The Sloan Digital Sky Survey, page 13	
<b>GZ2</b>	Galaxy Zoo 1 and 2 are an ongoing project using citizen scientists to visually classify galaxies. See Lintott et al. (2008) and Lintott et al. (2011), page 223	
<b>Acronyms</b>		
<b>CCD</b>	A charge-coupled device. This is the standard for imaging in optical wavelengths and consists of a chip with	
<b>DN</b>	‘data numbers’ or, equivalently, counts are the actual measured quantity and are related to the number of photo-electrons produced by the detector scaled by the gain of the detector amplifier. Photo-electrons, not DN, obey Poisson statistics, making this distinction important, page 25	
<b>MCC</b>	Matthews Correlation Coefficient, see Baldi et al. (2000), page 122	

# References

- ABAZAJIAN, K. N., ADELMAN-MCCARTHY, J. K., AGÜEROS, M. A., ALLAM, S. S., ALLENDE PRIETO, C., AN, D., ANDERSON, K. S. J., ANDERSON, S. F., ET AL. The Seventh Data Release of the Sloan Digital Sky Survey. *ApJS*, **182**:543–558 (2009). doi:10.1088/0067-0049/182/2/543. 20, 59, 251, 275, 276
- ABRAHAM, R. G. Perspectives in Physical Morphology. ArXiv Astrophysics e-prints (1998). 1, 2
- ADELMAN-MCCARTHY, J. K., AGÜEROS, M. A., ALLAM, S. S., ANDERSON, K. S. J., ANDERSON, S. F., ANNIS, J., BAHCALL, N. A., BAILER-JONES, C. A. L., ET AL. The Fifth Data Release of the Sloan Digital Sky Survey. *ApJS*, **172**:634–644 (2007). doi:10.1086/518864. 17
- AGUERRI, J. A. L., HUERTAS-COMPANY, M., SÁNCHEZ ALMEIDA, J., AND MUÑOZ-TUÑÓN, C. Dissecting the morphological and spectroscopic properties of galaxies in the local Universe. I. Elliptical galaxies. *A&A*, **203**:A136 (2012). doi:10.1051/0004-6361/201117632. 297
- AHN, C. P., ALEXANDROFF, R., ALLENDE PRIETO, C., ANDERS, F., ANDERSON, S. F., ANDERTON, T., ANDREWS, B. H., AUBOURG, É., ET AL. The Tenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment. *ApJS*, **211**:17 (2014). doi:10.1088/0067-0049/211/2/17. 21
- AIHARA, H., ALLENDE PRIETO, C., AN, D., ANDERSON, S. F., AUBOURG, É., BALBINOT, E., BEERS, T. C., BERLIND, A. A., ET AL. The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III. *ApJS*, **193**:29 (2011). doi:10.1088/0067-0049/193/2/29. 188
- ALLEN, P. D., DRIVER, S. P., GRAHAM, A. W., CAMERON, E., LISKE, J., AND DE PROPRIS, R. The Millennium Galaxy Catalogue: bulge-disc decomposition of 10095 nearby galaxies. *MNRAS*, **371**:2–18 (2006). doi:10.1111/j.1365-2966.2006.10586.x. 114
- ANDRAE, R., SCHULZE-HARTUNG, T., AND MELCHIOR, P. Dos and don'ts of reduced chi-squared. arXiv:astro-ph/1012.3754 (2010). 111, 113
- ANDREON, S. AND CUILLANDRE, J.-C. Dissecting the Luminosity Function of the Coma Cluster of Galaxies Using Canada-France-Hawaii Telescope Wide-Field Images. *ApJ*, **569**:144–156 (2002). doi:10.1086/339261. 256, 268
- BALDI, P., BRUNAK, S., CHAUVIN, Y., ANDERSEN, C. A. F., AND NIELSEN, H. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, **16**(5):412–424 (2000). doi:10.1093/bioinformatics/16.5.412. 119, 309
- BALDRY, I. K., DRIVER, S. P., LOVEDAY, J., TAYLOR, E. N., KELVIN, L. S., LISKE, J., NORBERG, P., ROBOTHAM, A. S. G., ET AL. Galaxy And Mass Assembly (GAMA): the galaxy stellar mass function at  $z < 0.06$ . *MNRAS*, **421**:621–634 (2012). doi:10.1111/j.1365-2966.2012.20340.x. 252, 255, 259, 261, 263, 265, 266, 271
- BALDRY, I. K., GLAZEBROOK, K., BRINKMANN, J., IVEZIĆ, Ž., LUPTON, R. H., NICHOL, R. C., AND SZALAY, A. S. Quantifying the Bimodal Color-Magnitude Distribution of Galaxies. *ApJ*, **600**:681–694 (2004). doi:10.1086/380092. 278, 303
- BARDEN, M., HÄUSSLER, B., PENG, C. Y., MCINTOSH, D. H., AND GUO, Y. GALAPAGOS: from pixels to parameters. *MNRAS*, **422**:449–468 (2012). doi:10.1111/j.1365-2966.2012.20619.x. 94, 308
- BELL, E. F., MCINTOSH, D. H., KATZ, N., AND WEINBERG, M. D. The Optical and Near-Infrared Properties of Galaxies. I. Luminosity and Stellar Mass Functions. *ApJS*, **149**:289–312 (2003). doi:10.1086/378847. 252, 260, 262, 263, 265, 266
- BERLIND, A. A. AND WEINBERG, D. H. The Halo Occupation Distribution: Toward an Empirical Determination of the Relation between Galaxies and Mass. *ApJ*, **575**:587–616 (2002). doi:10.1086/341469. 270
- BERNARDI, M. Evolution in the structural properties of early-type brightest cluster galaxies at small lookback time and dependence on the environment. *MNRAS*, **395**:1491–1506 (2009). doi:10.1111/j.1365-2966.2009.14601.x. 5, 13, 87, 286, 297
- BERNARDI, M., HYDE, J. B., SHETH, R. K., MILLER, C. J., AND NICHOL, R. C. The Luminosities, Sizes, and Velocity Dispersions of Brightest Cluster Galaxies: Implications for Formation History. *AJ*, **133**:1741–1755 (2007). doi:10.1086/511783. 87, 221, 254, 286
- BERNARDI, M., MEERT, A., SHETH, R. K., VIKRAM, V., HUERTAS-COMPANY, M., MEI, S., AND SHANKAR, F. The massive end of the luminosity and stellar mass functions: dependence on the fit to the light profile. *MNRAS*, **436**:697–704 (2013). doi:10.1093/mnras/stt1607. 14, 61, 63, 87, 167, 182, 217, 221, 222, 250

- BERNARDI, M., MEERT, A., VIKRAM, V., HUERTAS-COMPANY, M., MEI, S., SHANKAR, F., AND SHETH, R. K. Systematic effects on the size-luminosity relations of early- and late-type galaxies: dependence on model fitting and morphology. *MNRAS*, **443**:874–897 (2014). doi:10.1093/mnras/stu1106. 14, 181, 202, 203, 251, 254, 257, 258, 268, 272, 305
- BERNARDI, M., ROCHE, N., SHANKAR, F., AND SHETH, R. K. Curvature in the colour-magnitude relation but not in colour- $\sigma$ : major dry mergers at  $M_* > 2 \times 10^{11} M_\odot$ ? *MNRAS*, **412**:684–704 (2011a). doi:10.1111/j.1365-2966.2010.17984.x. 269
- BERNARDI, M., ROCHE, N., SHANKAR, F., AND SHETH, R. K. Evidence of major dry mergers at  $M_* > 2 \times 10^{11} M_\odot$  from curvature in early-type galaxy scaling relations? *MNRAS*, **412**:L6–L10 (2011b). doi:10.1111/j.1745-3933.2010.00982.x. 269, 286, 289, 299
- BERNARDI, M., SHANKAR, F., HYDE, J. B., MEI, S., MARULLI, F., AND SHETH, R. K. Galaxy luminosities, stellar masses, sizes, velocity dispersions as a function of morphological type. *MNRAS*, **404**:2087–2122 (2010). doi:10.1111/j.1365-2966.2010.16425.x. 221, 250, 251, 252, 253, 254, 257, 258, 259, 260, 263, 265, 268, 269, 270, 276, 284, 285, 299
- BERNARDI, M., SHETH, R. K., ANNIS, J., BURLES, S., EISENSTEIN, D. J., FINKBEINER, D. P., HOGG, D. W., LUPTON, R. H., ET AL. Early-Type Galaxies in the Sloan Digital Sky Survey. I. The Sample. *AJ*, **125**:1817–1848 (2003a). doi:10.1086/367776. 276, 278, 297
- BERNARDI, M., SHETH, R. K., ANNIS, J., BURLES, S., EISENSTEIN, D. J., FINKBEINER, D. P., HOGG, D. W., LUPTON, R. H., ET AL. Early-type Galaxies in the Sloan Digital Sky Survey. II. Correlations between Observables. *AJ*, **125**:1849–1865 (2003b). doi:10.1086/374256. 5, 13, 297
- BERNARDI, M., SHETH, R. K., ANNIS, J., BURLES, S., EISENSTEIN, D. J., FINKBEINER, D. P., HOGG, D. W., LUPTON, R. H., ET AL. Early-type Galaxies in the Sloan Digital Sky Survey. III. The Fundamental Plane. *AJ*, **125**:1866–1881 (2003c). doi:10.1086/367794. 5, 13
- BERNARDI, M., SHETH, R. K., ANNIS, J., BURLES, S., FINKBEINER, D. P., LUPTON, R. H., SCHLEGEL, D. J., SUBBARAO, M., ET AL. Early-Type Galaxies in the Sloan Digital Sky Survey. IV. Colors and Chemical Evolution. *AJ*, **125**:1882–1896 (2003d). doi:10.1086/367795. 5, 13
- BERTIN, E. AND ARNOUDS, S. SExtractor: Software for source extraction. *A&AS*, **117**:393–404 (1996). 37, 38, 44
- BINGGELI, B. AND CAMERON, L. M. Dwarf galaxies in the Virgo cluster. I - The systematic photometric properties of early-type dwarfs. *A&A*, **252**:27–52 (1991). 256
- BINGGELI, B., SANDAGE, A., AND TARENghi, M. Studies of the Virgo Cluster. I - Photometry of 109 galaxies near the cluster center to serve as standards. *AJ*, **89**:64–82 (1984). doi:10.1086/113484. 286
- BLANTON, M. R., DALCANTON, J., EISENSTEIN, D., LOVEDAY, J., STRAUSS, M. A., SUBBARAO, M., WEINBERG, D. H., ANDERSON, J. E., JR., ET AL. The Luminosity Function of Galaxies in SDSS Commissioning Data. *AJ*, **121**:2358–2380 (2001). doi:10.1086/320405. 167, 256, 268
- BLANTON, M. R., EISENSTEIN, D., HOGG, D. W., SCHLEGEL, D. J., AND BRINKMANN, J. Relationship between Environment and the Broadband Optical Properties of Galaxies in the Sloan Digital Sky Survey. *ApJ*, **629**:149–157 (2005a). doi:{10.1086/422897}. 5, 13
- BLANTON, M. R., HOGG, D. W., BAHCALL, N. A., BALDRY, I. K., BRINKMANN, J., CSABAI, I., EISENSTEIN, D., FUKUGITA, M., ET AL. The Broadband Optical Properties of Galaxies with Redshifts  $0.02 < z < 0.22$ . *ApJ*, **594**:186–207 (2003a). doi:10.1086/375528. 282
- BLANTON, M. R., HOGG, D. W., BAHCALL, N. A., BRINKMANN, J., BRITTON, M., CONNOLLY, A. J., CSABAI, I., FUKUGITA, M., ET AL. The Galaxy Luminosity Function and Luminosity Density at Redshift  $z = 0.1$ . *ApJ*, **592**:819–838 (2003b). doi:10.1086/375776. 268
- BLANTON, M. R., KAZIN, E., MUNA, D., WEAVER, B. A., AND PRICE-WHELAN, A. Improved Background Subtraction for the Sloan Digital Sky Survey Images. *AJ*, **142**:31 (2011). doi:10.1088/0004-6256/142/1/31. 23, 67, 95
- BLANTON, M. R. AND ROWEIS, S. K-Corrections and Filter Transformations in the Ultraviolet, Optical, and Near-Infrared. *AJ*, **133**:734–754 (2007). doi:10.1086/510127. 29
- BLANTON, M. R., SCHLEGEL, D. J., STRAUSS, M. A., BRINKMANN, J., FINKBEINER, D., FUKUGITA, M., GUNN, J. E., HOGG, D. W., ET AL. New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys. *AJ*, **129**:2562–2578 (2005b). doi:10.1086/429803. 61, 75
- BROWN, R. J. N., FORBES, D. A., SILVA, D., HELSDON, S. F., PONMAN, T. J., HAU, G. K. T., BRODIE, J. P., GOUDFROOIJ, P., ET AL. Near-infrared imaging of ellipticals: surface brightness profiles and photometry. *MNRAS*, **341**:747–759 (2003). doi:10.1046/j.1365-8711.2003.05947.x. 256
- BRUCE, V. A., DUNLOP, J. S., CIRASUOLO, M., McLURE, R. J., TARGETT, T. A., BELL, E. F., CROTON, D. J., DEKEL, A., ET AL. The morphologies of massive galaxies at  $1 < z < 3$  in the CANDELS-UDS field: compact bulges, and the rise and fall of massive discs. *MNRAS*, **203**:1666–1701 (2012). doi:10.1111/j.1365-2966.2012.22087.x. 274
- CAPACCIOLI, M. Photometry of early-type galaxies and the  $R \exp{1/4}$  law. In H. G. Corwin, Jr. and L. Bottinelli, editors, *World of Galaxies (Le Monde des Galaxies)*, page 208. Springer-Verlag, Berlin (1989). 3, 35

## 13. REFERENCES

---

- CAPPELLARI, M., McDERMID, R. M., ALATALO, K., BLITZ, L., BOIS, M., BOURNAUD, F., BUREAU, M., CROCKER, A. F., ET AL. The ATLAS<sup>3D</sup> project - XX. Mass-size and mass- $\sigma$  distributions of early-type galaxies: bulge fraction drives kinematics, mass-to-light ratio, molecular gas fraction and stellar initial mass function. *MNRAS*, **432**:1862–1893 (2013). doi: 10.1093/mnras/stt644. 299
- CAROLLO, C. M., BSCHORR, T. J., RENZINI, A., LILLY, S. J., CAPAK, P., CIBINEL, A., ILBERT, O., ONODERA, M., ET AL. Newly Quenched Galaxies as the Cause for the Apparent Evolution in Average Size of the Population. *ApJ*, **773**:112 (2013). doi: 10.1088/0004-637X/773/2/112. 269
- CHABRIER, G. The Galactic Disk Mass Function: Reconciliation of the Hubble Space Telescope and Nearby Determinations. *ApJ*, **586**:L133–L136 (2003). doi:10.1086/374879. 252, 259
- CIMATTI, A., CASSATA, P., POZZETTI, L., KURK, J., MIGNOLI, M., RENZINI, A., DADDI, E., BOLZONELLA, M., ET AL. GMASST ultradeep spectroscopy of galaxies at  $z \sim 2$ . II. Superdense passive galaxies: how did they form and evolve? *A&A*, **482**:21–42 (2008). doi:10.1051/0004-6361:20078739. 274
- CIOTTI, L. AND BERTIN, G. Analytical properties of the  $R^{1/m}$  law. *A&A*, **352**:447–451 (1999). 277
- COLE, S., LACEY, C. G., BAUGH, C. M., AND FRENK, C. S. Hierarchical galaxy formation. *MNRAS*, **319**:168–204 (2000). doi: 10.1046/j.1365-8711.2000.03879.x. 9
- CONSELICE, C. J. The Relationship between Stellar Light Distributions of Galaxies and Their Formation Histories. *ApJS*, **147**:1–28 (2003). doi:10.1086/375001. 40
- COORAY, A. AND SHETH, R. Halo models of large scale structure. *Phys. Rep.*, **372**:1–129 (2002). doi:10.1016/S0370-1573(02)00276-4. 270
- CROTON, D. J., SPRINGEL, V., WHITE, S. D. M., DE LUCIA, G., FRENK, C. S., GAO, L., JENKINS, A., KAUFFMANN, G., ET AL. The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *MNRAS*, **365**:11–28 (2006). doi:10.1111/j.1365-2966.2005.09675.x. 270
- DAVARI, R., HO, L. C., PENG, C. Y., AND HUANG, S. How Robust are the Size Measurements of High-redshift Compact Galaxies? *ApJ*, **787**:69 (2014). doi:10.1088/0004-637X/787/1/69. 169, 182, 306
- DE VAUCOULEURS, G. Recherches sur les Nébuleuses Extragalactiques. *Ann. Astrophys.*, **11**:247 (1948). 2, 3, 35
- DE VAUCOULEURS, G. Structure, dynamics and statistical properties of galaxies. In J. R. Shakeshaft, editor, *The Formation and Dynamics of Galaxies*, volume 58 of IAU Symposium, pages 1–52 (1974). 4
- DJORGOVSKI, S. AND DAVIS, M. Fundamental properties of elliptical galaxies. *ApJ*, **313**:59–68 (1987). doi:10.1086/164948. 8, 13
- DONZELLI, C. J., MURIEL, H., AND MADRID, J. P. The Luminosity Profiles of Brightest Cluster Galaxies. *ApJS*, **195**:15 (2011). doi:10.1088/0067-0049/195/2/15. 36
- DRESSLER, A. Galaxy morphology in rich clusters - Implications for the formation and evolution of galaxies. *ApJ*, **236**:351–365 (1980). doi:10.1086/157753. 8, 13
- DRIVER, S. P., HILL, D. T., KELVIN, L. S., ROBOTHAM, A. S. G., LISKE, J., NORBERG, P., BALDRY, I. K., BAMFORD, S. P., ET AL. Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, **413**:971–995 (2011). doi:10.1111/j.1365-2966.2010.18188.x. 257
- EMSELLEM, E., CAPPELLARI, M., KRAJNOVIĆ, D., ALATALO, K., BLITZ, L., BOIS, M., BOURNAUD, F., BUREAU, M., ET AL. The ATLAS<sup>3D</sup> project - III. A census of the stellar angular momentum within the effective radius of early-type galaxies: unveiling the distribution of fast and slow rotators. *MNRAS*, **414**:888–912 (2011). doi:10.1111/j.1365-2966.2011.18496.x. 274
- FABER, S. M. AND JACKSON, R. E. Velocity dispersions and mass-to-light ratios for elliptical galaxies. *ApJ*, **204**:668–683 (1976). doi:10.1086/154215. 8, 13
- FUKUGITA, M., NAKAMURA, O., OKAMURA, S., YASUDA, N., BARENTINE, J. C., BRINKMANN, J., GUNN, J. E., HARVANEK, M., ET AL. A Catalog of Morphologically Classified Galaxies from the Sloan Digital Sky Survey: North Equatorial Region. *AJ*, **134**:579–593 (2007). doi:10.1086/518962. xiv, 228, 278, 279
- GADOTTI, D. A. Image decomposition of barred galaxies and AGN hosts. *MNRAS*, **384**:420–439 (2008). doi:10.1111/j.1365-2966.2007.12723.x. 178, 181, 238
- GADOTTI, D. A. Structural properties of pseudo-bulges, classical bulges and elliptical galaxies: a Sloan Digital Sky Survey perspective. *MNRAS*, **393**:1531–1552 (2009). doi: 10.1111/j.1365-2966.2008.14257.x. 13, 178, 181, 306
- GONZALEZ, A. H., ZABLUDOFF, A. I., AND ZARITSKY, D. Intracluster Light in Nearby Galaxy Clusters: Relationship to the Halos of Brightest Cluster Galaxies. *ApJ*, **618**:195–213 (2005). doi: 10.1086/425896. 36, 277, 282
- GRAHAM, A. W. Elliptical and Disk Galaxy Structure and Modern Scaling Laws, page 91 (2013). doi:10.1007/978-94-007-5609-0\_2. 286
- GRAHAM, A. W. AND DRIVER, S. P. A Concise Reference to (Projected) Sérsic  $R^{1/n}$  Quantities, Including Concentration, Profile Slopes, Petrosian Indices, and Kron Magnitudes. *PASA*, **22**:118–127 (2005). doi:10.1071/AS05001. 167

- GRAHAM, A. W. AND DRIVER, S. P. A Log-Quadratic Relation for Predicting Supermassive Black Hole Masses from the Host Bulge Sérsic Index. *ApJ*, **655**:77–87 (2007). doi:10.1086/509758. 3, 35
- GRAHAM, A. W., DRIVER, S. P., PETROSIAN, V., CONSELICE, C. J., BERSHADY, M. A., CRAWFORD, S. M., AND GOTO, T. Total Galaxy Magnitudes and Effective Radii from Petrosian Magnitudes and Radii. *AJ*, **130**:1535–1544 (2005). doi:10.1086/444475. 256, 268
- GRAHAM, A. W., ERWIN, P., TRUJILLO, I., AND ASENSIO RAMOS, A. A New Empirical Model for the Structural Analysis of Early-Type Galaxies, and A Critical Review of the Nuker Model. *AJ*, **125**:2951–2963 (2003). doi:10.1086/375320. 274
- GRAHAM, A. W. AND WORLEY, C. C. Inclination- and dust-corrected galaxy parameters: bulge-to-disc ratios and size-luminosity relations. *MNRAS*, **388**:1708–1728 (2008). doi:10.1111/j.1365-2966.2008.13506.x. 299
- GRANATO, G. L., DE ZOTTI, G., SILVA, L., BRESSAN, A., AND DANESE, L. A Physical Model for the Coevolution of QSOs and Their Spheroidal Hosts. *ApJ*, **600**:580–594 (2004). doi:10.1086/379875. 270
- GUNN, J. E., CARR, M., ROCKOS, C., SEKIGUCHI, M., BERRY, K., ELMS, B., DE HAAS, E., IVEZIĆ, Ž., ET AL. The Sloan Digital Sky Survey Photometric Camera. *AJ*, **116**:3040–3081 (1998). doi:10.1086/300645. 16
- GUNN, J. E., SIEGMUND, W. A., MANNERY, E. J., OWEN, R. E., HULL, C. L., LEGER, R. F., CAREY, L. N., KNAPP, G. R., ET AL. The 2.5 m Telescope of the Sloan Digital Sky Survey. *AJ*, **131**:2332–2359 (2006). doi:10.1086/500975. 16
- GUO, Y., MCINTOSH, D. H., MO, H. J., KATZ, N., VAN DEN BOSCH, F. C., WEINBERG, M., WEINMANN, S. M., PASQUALI, A., ET AL. Structural properties of central galaxies in groups and clusters. *MNRAS*, **398**:1129–1149 (2009). doi:10.1111/j.1365-2966.2009.15223.x. 75, 97, 99
- HALL, M., COURTEAU, S., DUTTON, A. A., McDONALD, M., AND ZHU, Y. An investigation of Sloan Digital Sky Survey imaging data and multiband scaling relations of spiral galaxies. *MNRAS*, **425**:2741–2765 (2012). doi:10.1111/j.1365-2966.2012.21290.x. 268
- HAO, J., MCKAY, T., KOESTER, B., RYKOFF, E., ROZO, E., ANNIS, J., WECHSLER, R. H., ÉVRARD, A., ET AL. A GMBCG Galaxy Cluster Catalog of 55,437 Rich Clusters from SDSS DR7. In American Astronomical Society Meeting Abstracts #217, volume 43 of Bulletin of the American Astronomical Society, page 213.06 (2011). 86
- HÄUSSLER, B., BAMFORD, S. P., VIKA, M., ROJAS, A. L., BARDEN, M., KELVIN, L. S., ALPASLAN, M., ROBOTHAM, A. S. G., ET AL. Megamorph - multiwavelength measurement of galaxy structure: complete Sérsic profile information from modern surveys. *MNRAS*, **430**:330–369 (2013). doi:10.1093/mnras/sts633. 13, 27, 301
- HÄUSSLER, B., MCINTOSH, D. H., BARDEN, M., BELL, E. F., RIX, H.-W., BORCH, A., BECKWITH, S. V. W., CALDWELL, J. A. R., ET AL. GEMS: Galaxy Fitting Catalogs and Testing Parametric Galaxy Fitting Codes: GALFIT and GIM2D. *ApJS*, **172**:615–633 (2007). doi:10.1086/518836. 58, 61, 64, 76, 188, 202
- HE, Y. Q., XIA, X. Y., HAO, C. N., JING, Y. P., MAO, S., AND LI, C. Photometric Properties and Luminosity Function of Nearby Massive Early-type Galaxies. *ApJ*, **773**:37 (2013). doi:10.1088/0004-637X/773/1/37. 256, 268
- HILL, D. T., KELVIN, L. S., DRIVER, S. P., ROBOTHAM, A. S. G., CAMERON, E., CROSS, N., ANDRAE, E., BALDRY, I. K., ET AL. Galaxy and Mass Assembly: FUV, NUV, ugrizYJHK Petrosian, Kron and Sérsic photometry. *MNRAS*, **412**:765–799 (2011). doi:10.1111/j.1365-2966.2010.17950.x. 94, 254, 258, 259, 260, 265, 268
- HOGG, D. W. Distance measures in cosmology. *arXiv:astro-ph/9905116* (1999). 29
- HUANG, S., HO, L. C., PENG, C. Y., LI, Z.-Y., AND BARTH, A. J. The Carnegie-Irvine Galaxy Survey. III. The Three-component Structure of Nearby Elliptical Galaxies. *ApJ*, **766**:47 (2013). doi:10.1088/0004-637X/766/1/47. 36, 169
- HUBBLE, E. P. A general study of diffuse galactic nebulae. *ApJ*, **56**:162–199 (1922). doi:10.1086/142698. 1
- HUBBLE, E. P. Extragalactic nebulae. *ApJ*, **64**:321–369 (1926). doi:10.1086/143018. 1
- HUBBLE, E. P. A spiral nebula as a stellar system, Messier 31. *ApJ*, **69**:103–158 (1929). doi:10.1086/143167. 1
- HUERTAS-COMPANY, M., AGUERRI, J. A. L., BERNARDI, M., MEI, S., AND SÁNCHEZ ALMEIDA, J. Revisiting the Hubble sequence in the SDSS DR7 spectroscopic sample: a publicly available Bayesian automated classification. *A&A*, **525**:A157+ (2011). doi:10.1051/0004-6361/201015735. 28, 252, 264, 266, 278
- HUERTAS-COMPANY, M., MEI, S., SHANKAR, F., DELAYE, L., RAICHOOR, A., COVONE, G., FINOGUENOV, A., KNEIB, J. P., ET AL. The evolution of the mass-size relation for early-type galaxies from  $z \sim 1$  to the present: dependence on environment, mass range and detailed morphology. *MNRAS*, **428**:1715–1742 (2013a). doi:10.1093/mnras/sts150. 299
- HUERTAS-COMPANY, M., SHANKAR, F., MEI, S., BERNARDI, M., AGUERRI, J. A. L., MEEART, A., AND VIKRAM, V. No Evidence for a Dependence of the Mass-Size Relation of Early-type Galaxies on Environment in the Local Universe. *ApJ*, **779**:29 (2013b). doi:10.1088/0004-637X/779/1/29. 14, 293

## 13. REFERENCES

---

- HYDE, J. B. AND BERNARDI, M. Curvature in the scaling relations of early-type galaxies. *MNRAS*, **394**:1978–1990 (2009). doi: 10.1111/j.1365-2966.2009.14445.x. 268
- JOHNSTON, E. J., ARAGÓN-SALAMANCA, A., MERRIFIELD, M. R., AND BREDREGAL, A. G. Spectroscopic bulge-disc decomposition: a new method to study the evolution of lenticular galaxies. *MNRAS*, **422**:2590–2599 (2012). doi:10.1111/j.1365-2966.2012.20813.x. 274
- JOHNSTON, R. Shedding light on the galaxy luminosity function. *A&A Rev.*, **19**:41 (2011). doi:10.1007/s00159-011-0041-9. 308
- KAUFFMANN, G., HECKMAN, T. M., WHITE, S. D. M., CHARLOT, S., TREMONTI, C., BRINCHMANN, J., BRUZUAL, G., PENG, E. W., ET AL. Stellar masses and star formation histories for  $10^5$  galaxies from the Sloan Digital Sky Survey. *MNRAS*, **341**:33–53 (2003). doi:10.1046/j.1365-8711.2003.06291.x. 13
- KELVIN, L. S., DRIVER, S. P., ROBOTHAM, A. S. G., HILL, D. T., AL-PASLAN, M., BALDRY, I. K., BAMFORD, S. P., BLAND-HAWTHORN, J., ET AL. Galaxy And Mass Assembly (GAMA): Structural Investigation of Galaxies via Model Analysis. *MNRAS*, **421**:1007–1039 (2012). doi:10.1111/j.1365-2966.2012.20355.x. 13, 255, 257, 263, 268
- KORMENDY, J. Brightness distributions in compact and normal galaxies. II - Structure parameters of the spheroidal component. *ApJ*, **218**:333–346 (1977). doi:10.1086/155687. 8, 13
- KORMENDY, J., FISHER, D. B., CORNELL, M. E., AND BENDER, R. Structure and Formation of Elliptical and Spheroidal Galaxies. *ApJS*, **182**:216–309 (2009). doi:10.1088/0067-0049/182/1/216. 3, 35
- KRAVTSOV, A., VIKHILIN, A., AND MESHSCHERYAKOV, A. Stellar mass – halo mass relation and star formation efficiency in high-mass halos. *arXiv:astro-ph/1401.7329* (2014). 13, 14, 185, 203, 305
- LACKNER, C. N. AND GUNN, J. E. Astrophysically motivated bulge-disc decompositions of Sloan Digital Sky Survey galaxies. *MNRAS*, **421**:2277–2302 (2012). doi:10.1111/j.1365-2966.2012.20450.x. 5, 13, 27, 58, 88, 89, 102, 113, 273
- LEAUTAUD, A., GEORGE, M. R., BEHROOZI, P. S., BUNDY, K., TINKER, J., WECHSLER, R. H., CONROY, C., FINOGUENOV, A., ET AL. The Integrated Stellar Content of Dark Matter Halos. *ApJ*, **746**:95 (2012). doi:10.1088/0004-637X/746/1/95. 270
- LINTOTT, C., SCHAWINSKI, K., BAMFORD, S., SLOSAR, A., LAND, K., THOMAS, D., EDMONDSON, E., MASTERS, K., ET AL. Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies. *MNRAS*, **410**:166–178 (2011). doi:10.1111/j.1365-2966.2010.17432.x. 223, 304, 309
- LINTOTT, C. J., SCHAWINSKI, K., SLOSAR, A., LAND, K., BAMFORD, S., THOMAS, D., RADICK, M. J., NICHOL, R. C., ET AL. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *MNRAS*, **389**:1179–1189 (2008). doi:10.1111/j.1365-2966.2008.13689.x. 115, 223, 309
- LOTZ, J. M., PRIMACK, J., AND MADAU, P. A New Nonparametric Approach to Galaxy Morphological Classification. *AJ*, **128**:163–182 (2004). doi:10.1086/421849. 40
- LUPTON, R. Statistics in theory and practice (1993). 102
- MEERT, A., VIKRAM, V., AND BERNARDI, M. Simulations of single- and two-component galaxy decompositions for spectroscopically selected galaxies from the Sloan Digital Sky Survey. *MNRAS*, **433**:1344–1361 (2013). doi:10.1093/mnras/stt822. 14, 252, 258
- MEERT, A., VIKRAM, V., AND BERNARDI, M. A catalogue of 2D photometric decompositions in the SDSS-DR7 spectroscopic main galaxy sample: preferred models and systematics. *MNRAS*, **446**:3943–3974 (2015). doi:10.1093/mnras/stu2333. 14, 252, 268
- MENDEL, J. T., SIMARD, L., PALMER, M., ELLISON, S. L., AND PATTON, D. R. A Catalog of Bulge, Disk, and Total Stellar Mass Estimates for the Sloan Digital Sky Survey. *ApJS*, **210**:3 (2014). doi:10.1088/0067-0049/210/1/3. 114, 187, 308
- METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., AND TELLER, E. Equation of State Calculations by Fast Computing Machines. *JCP*, **21**:1087–1092 (1953). doi: 10.1063/1.1699114. 188
- MITCHELL, P. D., LACEY, C. G., BAUGH, C. M., AND COLE, S. How well can we really estimate the stellar masses of galaxies from broad-band photometry? *MNRAS*, **435**:87–114 (2013). doi: 10.1093/mnras/stt1280. 271
- MO, H., VAN DEN BOSCH, F. C., AND WHITE, S. Galaxy Formation and Evolution (2010). 9, 10
- MOSLEH, M., WILLIAMS, R. J., AND FRANX, M. On the Robustness of  $z = 0\text{--}1$  Galaxy Size Measurements through Model and Non-parametric Fits. *ApJ*, **777**(2):117 (2013). 169, 182, 254, 306
- MOSTER, B. P., NAAB, T., AND WHITE, S. D. M. Galactic star formation and accretion histories from matching galaxies to dark matter haloes. *MNRAS*, **428**:3121–3138 (2013). doi: 10.1093/mnras/sts261. 270
- MOUSTAKAS, J., COIL, A. L., AIRD, J., BLANTON, M. R., COOL, R. J., EISENSTEIN, D. J., MENDEZ, A. J., WONG, K. C., ET AL. PRIMUS: Constraints on Star Formation Quenching and Galaxy Merging, and the Evolution of the Stellar Mass Function from  $z = 0\text{--}1$ . *ApJ*, **767**:50 (2013). doi:10.1088/0004-637X/767/1/50. 252, 261, 263, 265, 266, 271

- NAIR, P., VAN DEN BERGH, S., AND ABRAHAM, R. G. A Fundamental Line for Elliptical Galaxies. *ApJ*, **734**:L31 (2011). doi: 10.1088/2041-8205/734/2/L31. 296, 297
- NAIR, P. B. AND ABRAHAM, R. G. A Catalog of Detailed Visual Morphological Classifications for 14,034 Galaxies in the Sloan Digital Sky Survey. *ApJS*, **186**:427–456 (2010). doi: 10.1088/0067-0049/186/2/427. xiv, 223, 231, 267, 278, 281
- OEMLER, A., JR. The structure of elliptical and cD galaxies. *ApJ*, **209**:693–709 (1976). doi:10.1086/154769. 277, 282
- OOHAMA, N., OKAMURA, S., FUKUGITA, M., YASUDA, N., AND NAKAMURA, O. Properties of Disks and Bulges of Spiral and Lenticular Galaxies in the Sloan Digital Sky Survey. *ApJ*, **705**:245–254 (2009). doi:10.1088/0004-637X/705/1/245. xv, 5, 6
- PADMANABHAN, N., SCHLEGEL, D. J., FINKBEINER, D. P., BARENTINE, J. C., BLANTON, M. R., BREWINGTON, H. J., GUNN, J. E., HARVANEK, M., ET AL. An Improved Photometric Calibration of the Sloan Digital Sky Survey Imaging Data. *ApJ*, **674**:1217–1233 (2008). doi:10.1086/524677. 189
- PARANAJE, A. AND SHETH, R. K. The luminosities of the brightest cluster galaxies and brightest satellites in SDSS groups. *MNRAS*, **423**:1845–1855 (2012). doi:10.1111/j.1365-2966.2012.21008.x. 251
- PENG, C. Y., HO, L. C., IMPEY, C. D., AND RIX, H.-W. Detailed Structural Decomposition of Galaxy Images. *AJ*, **124**:266–293 (2002). doi:10.1086/340952. 37, 39, 64
- PLANCK COLLABORATION, ADE, P. A. R., AGHANIM, N., ARNAUD, M., ASHDOWN, M., ATRIO-BARANDELA, F., AUMONT, J., BACCIGALUPI, C., ET AL. Planck intermediate results. XI. The gas content of dark matter halos: the Sunyaev-Zeldovich-stellar mass relation for locally brightest galaxies. *A&A*, **557**:A52 (2013). doi: 10.1051/0004-6361/20120941. 270
- PRESS, W. H. AND SCHECHTER, P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, **187**:425–438 (1974). doi:10.1086/152650. 5
- PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., AND FLANNERY, B. P. Numerical recipes in C. The art of scientific computing (1992). 39, 76
- SAGLIA, R. P., SÁNCHEZ-BLÁZQUEZ, P., BENDER, R., SIMARD, L., DESAI, V., ARAGÓN-SALAMANCA, A., MILVANG-JENSEN, B., HALLIDAY, C., ET AL. The fundamental plane of EDisCS galaxies. The effect of size evolution. *A&A*, **524**:A6 (2010). doi:10.1051/0004-6361/201014703. 277
- SCHAWINSKI, K., URRY, C. M., SIMMONS, B. D., FORTSON, L., KAVIRAJ, S., KEEL, W. C., LINTOTT, C. J., MASTERS, K. L., ET AL. The green valley is a red herring: Galaxy Zoo reveals two evolutionary pathways towards quenching of star formation in early- and late-type galaxies. *MNRAS*, **440**:889–907 (2014). doi:10.1093/mnras/stu327. 241, 303
- SCHECHTER, P. An analytic expression for the luminosity function for galaxies. *ApJ*, **203**:297–306 (1976). doi:10.1086/154079. 7
- SCHMIDT, M. Space Distribution and Luminosity Functions of Quasi-Stellar Radio Sources. *ApJ*, **151**:393 (1968). doi: 10.1086/149446. 7, 258
- SCHOMBERT, J. M. The structure of brightest cluster members. I - Surface photometry. *ApJS*, **60**:603–693 (1986). doi: 10.1086/191100. 277, 282
- SÉRSIC, J. L. Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy. *Bol. Asociacion Argentina Astron. La Plata Argentina*, **6**:41 (1963a). 3, 35
- SÉRSIC, J. L. Photometry of southern galaxies. IX:NGC 1313. *Bol. Asociacion Argentina Astron. La Plata Argentina*, **6**:99 (1963b). 2
- SHANKAR, F. AND BERNARDI, M. The age dependence of the size-stellar mass relation and some implications. *MNRAS*, **396**:L76–L80 (2009). doi:10.1111/j.1745-3933.2009.00665.x. 293
- SHANKAR, F., LAPI, A., SALUCCI, P., DE ZOTTI, G., AND DANESE, L. New Relationships between Galaxy Properties and Host Halo Mass, and the Role of Feedbacks in Galaxy Formation. *ApJ*, **643**:14–25 (2006). doi:10.1086/502794. 270
- SHANKAR, F., MARULLI, F., BERNARDI, M., BOYLAN-KOLCHIN, M., DAI, X., AND KHOCHFAR, S. Further constraining galaxy evolution models through the size function of SDSS early-type galaxies. *MNRAS*, **405**:948–960 (2010a). doi:10.1111/j.1365-2966.2010.16540.x. xv, 5, 9, 12, 13
- SHANKAR, F., MARULLI, F., BERNARDI, M., DAI, X., HYDE, J. B., AND SHETH, R. K. Sizes and ages of SDSS ellipticals: comparison with hierarchical galaxy formation models. *MNRAS*, **403**:117–128 (2010b). doi:10.1111/j.1365-2966.2009.16137.x. 5, 13
- SHANKAR, F., MARULLI, F., BERNARDI, M., MEI, S., MEERT, A., AND VIKRAM, V. Size evolution of spheroids in a hierarchical Universe. *MNRAS*, **428**:109–128 (2013). doi:10.1093/mnras/sts001. 14, 269
- SHANKAR, F., MARULLI, F., MATHUR, S., BERNARDI, M., AND BOURNAUD, F. Black holes in pseudobulges: demography and models. *A&A*, **540**:A23 (2012). doi:10.1051/0004-6361/201118387. 293, 297
- SHEN, S., MO, H. J., WHITE, S. D. M., BLANTON, M. R., KAUFFMANN, G., VOGES, W., BRINKMANN, J., AND CSABAI, I. The size distribution of galaxies in the Sloan Digital Sky Survey. *MNRAS*, **343**:978–994 (2003). doi:10.1046/j.1365-8711.2003.06740.x. 7, 24, 29, 30, 61, 170, 275, 278, 282, 293, 297, 298, 309

## 13. REFERENCES

---

- SHETH, R. K. AND BERNARDI, M. Plain fundamentals of Fundamental Planes: analytics and algorithms. *MNRAS*, **422**:1825–1834 (2012). doi:10.1111/j.1365-2966.2011.19757.x. 275
- SILK, J. AND MAMON, G. A. The current status of galaxy formation. *Research in Astronomy and Astrophysics*, **12**:917–946 (2012). doi:10.1088/1674-4527/12/8/004. 270
- SIMARD, L., MENDEL, J. T., PATTON, D. R., ELLISON, S. L., AND McCONNACHE, A. W. A Catalog of Bulge+disk Decompositions and Updated Photometry for 1.12 Million Galaxies in the Sloan Digital Sky Survey. *ApJS*, **196**:11 (2011). doi:10.1088/0067-0049/196/1/11. 5, 13, 23, 27, 61, 84, 102, 104, 253, 255, 256, 258, 268, 273, 309
- SIMARD, L., WILLMER, C. N. A., VOGT, N. P., SARAJEDINI, V. L., PHILLIPS, A. C., WEINER, B. J., KOO, D. C., IM, M., ET AL. The DEEP Groth Strip Survey. II. Hubble Space Telescope Structural Parameters of Galaxies in the Groth Strip. *ApJS*, **142**:1–33 (2002). doi:10.1086/341399. 58, 104, 187
- SIMIEN, F. AND DE VAUCOULEURS, G. Systematics of bulge-to-disk ratios. *ApJ*, **302**:564–578 (1986). doi:10.1086/164015. 5
- STOUGHTON, C., LUPTON, R. H., BERNARDI, M., BLANTON, M. R., BURLES, S., CASTANDER, F. J., CONNOLLY, A. J., EISENSTEIN, D. J., ET AL. Sloan Digital Sky Survey: Early Data Release. *AJ*, **123**:485–548 (2002). doi:10.1086/324741. 17, 190, 241
- STRATEVA, I., IVEZIĆ, Ž., KNAPP, G. R., NARAYANAN, V. K., STRAUSS, M. A., GUNN, J. E., LUPTON, R. H., SCHLEGEL, D., ET AL. Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data. *AJ*, **122**:1861–1874 (2001). doi:10.1086/323301. 278, 303
- STRAUSS, M. A., WEINBERG, D. H., LUPTON, R. H., NARAYANAN, V. K., ANNIS, J., BERNARDI, M., BLANTON, M., BURLES, S., ET AL. Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Main Galaxy Sample. *AJ*, **124**:1810–1824 (2002). doi:10.1086/342343. 17, 23, 24, 47
- TREMONTI, C. A., HECKMAN, T. M., KAUFFMANN, G., BRINCHMANN, J., CHARLOT, S., WHITE, S. D. M., SEIBERT, M., PENG, E. W., ET AL. The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey. *ApJ*, **613**:898–913 (2004). doi:10.1086/423264. 13
- TRUJILLO, I., AGUERRI, J. A. L., CEPA, J., AND GUTIÉRREZ, C. M. The effects of seeing on Sérsic profiles - II. The Moffat PSF. *MNRAS*, **328**:977–985 (2001). doi:10.1046/j.1365-8711.2001.04937.x. 256, 268
- TRUJILLO, I., FEULNER, G., GORANOVA, Y., HOPP, U., LONGHETTI, M., SARACCO, P., BENDER, R., BRAITO, V., ET AL. Extremely compact massive galaxies at  $z \sim 1.4$ . *MNRAS*, **373**:L36–L40 (2006). doi:10.1111/j.1745-3933.2006.00238.x. 274
- TULLY, R. B. AND FISHER, J. R. A new method of determining distances to galaxies. *A&A*, **54**:661–673 (1977). 8, 13
- VALE, A. AND OSTRIKER, J. P. Linking halo mass to galaxy luminosity. *MNRAS*, **353**:189–200 (2004). doi:10.1111/j.1365-2966.2004.08059.x. 270
- VAN DOKKUM, P. G., FRANX, M., KRIEK, M., HOLDEN, B., ILLINGWORTH, G. D., MAGEE, D., BOUWENS, R., MARCHESE, D., ET AL. Confirmation of the Remarkable Compactness of Massive Quiescent Galaxies at  $z \sim 2.3$ : Early-Type Galaxies Did Not Form in a Simple Monolithic Collapse. *ApJ*, **677**:L5–L8 (2008). doi:10.1086/587874. 274
- VIKRAM, V., WADADEKAR, Y., KEMBAVI, A. K., AND VIJAYAGOVINDAN, G. V. PYMORPH: automated galaxy structural parameter estimation using PYTHON. *MNRAS*, **409**:1379–1392 (2010). doi:10.1111/j.1365-2966.2010.17426.x. 37
- WILLETT, K. W., LINTOTT, C. J., BAMFORD, S. P., MASTERS, K. L., SIMMONS, B. D., CASTEELS, K. R. V., EDMONDSON, E. M., FORTSON, L. F., ET AL. Galaxy Zoo 2: detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey. *MNRAS*, **435**:2835–2860 (2013). doi:10.1093/mnras/stt1458. 231, 238
- WILMAN, D. J., FONTANOT, F., DE LUCIA, G., ERWIN, P., AND MONACO, P. The hierarchical origins of observed galaxy morphology. *MNRAS*, **433**:2986–3004 (2013). doi:10.1093/mnras/stt941. xv, 9, 11
- ZEHAVI, I., ZHENG, Z., WEINBERG, D. H., BLANTON, M. R., BAHCALL, N. A., BERLIND, A. A., BRINKMANN, J., FRIEMAN, J. A., ET AL. Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity. *ApJ*, **736**:59 (2011). doi:10.1088/0004-637X/736/1/59. 270

# Index

- $L_*$ , 7  
 $M/L$ , 8  
 $b/a$ , 35  
 $K$ -correction, 7, 29, 189  
b/a, 235  
B/T, 4, 9, 60, 125, 235  
BCG, 14, 36, 86, 285  
CCD, 16  
DES, 13  
DN, 25, 33, 65, 66  
Elliptical Galaxy, 1, 36  
F-test, 102, 114, 188  
Fundamental Plane, 9  
GZ2, 223, 238  
H2011, 28, 228  
ICL, 87, 271  
LF, 5, 185, 217, 251, 258, 303  
LG12, 188  
LSST, 13  
M/L, 252  
MCC, 122  
Men14, 188  
MF, 8, 259  
PSF, 28, 33, 41, 67, 83, 113, 169, 303  
S/N, 21, 24, 61, 88, 105, 301  
S03, 282  
S11, 23, 27, 187, 217  
Sérsic index, 235  
Sérsic profile, 3, 35  
SDSS, 13, 16, 187, 217  
Spiral Galaxy, 1  
Sérsic profile, 60, 64  
T-type, 4, 223, 228