Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **421,** 1007–1039 (2012) doi:10.1111/j.1365-2966.2012.20355.x

Galaxy And Mass Assembly (GAMA): Structural Investigation of Galaxies via Model Analysis

Lee S. Kelvin,[1,2]★ Simon P. Driver,[1,2] Aaron S. G. Robotham,[1,2] David T. Hill,[1] Mehmet Alpaslan,[1,2] Ivan K. Baldry,[3] Steven P. Bamford,[4] Joss Bland-Hawthorn,[5] Sarah Brough,[6] Alister W. Graham,[7] Boris Häussler,[4] Andrew M. Hopkins,[6] Jochen Liske,[8] Jon Loveday,[9] Peder Norberg,[10] Steven Phillipps,[11] Cristina C. Popescu,[12] Matthew Prescott,[3] Edward N. Taylor[5,13] and Richard J. Tuffs[14]

[1]*School of Physics and Astronomy, University of St Andrews, North Haugh, St. Andrews, Fife KY16 9SS*
[2]*International Centre for Radio Astronomy Research, 7 Fairway, The University of Western Australia, Crawley, Perth, WA 6009, Australia*
[3]*Astrophysics Research Institute, Liverpool John Moores University, Twelve Quays House, Egerton Wharf, Birkenhead CH41 1LD*
[4]*The Centre for Astronomy and Particle Theory, The School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD*
[5]*Sydney Institute for Astronomy, School of Physics A28, University of Sydney, NSW 2006, Australia*
[6]*Australian Astronomical Observatory, PO Box 296, Epping, NSW 1710, Australia*
[7]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*
[8]*European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany*
[9]*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH*
[10]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE*
[11]*Astrophysics Group, H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL*
[12]*Jeremiah Horrocks Institute, University of Central Lancashire, Preston PR1 2HE*
[13]*School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia*
[14]*Max-Planck-Institut fuer Kernphysik, Saupfercheckweg 1, D-69117 Heidelberg, Germany*

Accepted 2011 December 7. Received 2011 December 7; in original form 2011 September 30

ABSTRACT

We present single-Sérsic two-dimensional (2D) model fits to 167 600 galaxies modelled independently in the *ugrizYJHK* bandpasses using reprocessed Sloan Digital Sky Survey Data Release Seven (SDSS DR7) and UKIRT Infrared Deep Sky Survey Large Area Survey imaging data available from the Galaxy And Mass Assembly (GAMA) data base. In order to facilitate this study we developed Structural Investigation of Galaxies via Model Analysis (SIGMA), an R wrapper around several contemporary astronomy software packages including SOURCE EXTRACTOR, PSF EXTRACTOR and GALFIT 3. SIGMA produces realistic 2D model fits to galaxies, employing automatic adaptive background subtraction and empirical point spread function measurements on the fly for each galaxy in GAMA. Using these results, we define a common coverage area across the three GAMA regions containing 138 269 galaxies. We provide Sérsic magnitudes truncated at $10r_e$ which show good agreement with SDSS Petrosian and GAMA photometry for low Sérsic index systems ($n < 4$), and much improved photometry for high Sérsic index systems ($n > 4$), recovering as much as $\Delta m = 0.5$ mag in the r band. We employ a K-band Sérsic index/$u - r$ colour relation to delineate the massive ($n > \sim 2$) early-type galaxies (ETGs) from the late-type galaxies (LTGs). The mean Sérsic index of these ETGs shows a smooth variation with wavelength, increasing by 30 per cent from g through K. LTGs exhibit a more extreme change in Sérsic index, increasing by 52 per cent across the same range. In addition, ETGs and LTGs exhibit a 38 and 25 per cent decrease, respectively, in half-light radius from g through K. These trends are shown to arise due to the effects of dust attenuation and stellar population/metallicity gradients within galaxy populations.

Key words: astronomical data bases: miscellaneous – catalogues – galaxies: fundamental parameters – galaxies: structure.

★E-mail: lee.kelvin@st-andrews.ac.uk

1 INTRODUCTION

The shapes and sizes of galaxies are not random but are defined by the orbital motions of their constituent stellar populations, arranging themselves into elliptical, bulge, disc and bar-like structures. Exactly why and how these structures come about is somewhat a mystery which no doubt relates to a complex formation history involving collapse, merging, infall, secular evolution and feedback processes as well as the precise nature of the coupling between the dark matter, gas, dust and stars and the influence of the larger halo in which the galaxy might reside (group, cluster, etc.) and the broader environment (filament, void, nexus, etc.). The combination of variations in, for example, galaxy structure, formation history, evolution and relative environment leads to distinct measurable effects on global galaxy parameters such as colour, concentration and size. The ultimate goal of structural analysis is to inform this discussion by robustly isolating and quantifying these parameters and exploring correlations between these properties and those obtained by other means, such as dynamical information.

Once the underlying structure of a galaxy is understood the overarching morphological class may be determined, and from this we can explore correlations with, for example, environment through the well-known morphology–density relation (Dressler 1980), i.e. the apparent preference for red, passive galaxies in the dense cores of galaxy groups and clusters. Several mechanisms have been suggested to explain this feature, most notably the combined effects of strangulation (Larson, Tinsley & Caldwell 1980; Kauffmann, White & Guiderdoni 1993; Diaferio et al. 2001), ram pressure stripping (Gunn & Gott 1972), harassment (Moore et al. 1996) and tidal interactions and merging (Park, Gott & Choi 2008). Recent studies by e.g. van der Wel (2008), Welikala et al. (2008, 2009), Bamford et al. (2009) and others confirm this morphology-environment connection; however, they suggest that the relation between structure and morphology is less apparent. Indeed, it appears that the mass of a galaxy rather than the environment in which it resides is more influential in determining its structure, highlighting the importance of stellar mass estimates.

As an example of the connection between galaxy structure and the physical processes of galaxy formation Dalcanton, Spergel & Summers (1997) and independently Mo, Mao & White (1998), both following on from Fall & Efstathiou (1980), relate the scale length of the disc to the angular momentum of a galaxy's dark matter halo. In addition, numerous properties of the bulge component are now known to relate to the mass of the supermassive black hole (e.g. Häring & Rix 2004; Novak, Faber & Dekel 2006; Graham & Driver 2007). Variations in structural properties as a function of wavelength (e.g. La Barbera et al. 2010) enable the extraction of colour gradients, potentially implying the direction of disc growth (e.g. inside out; Barden et al. 2005; Bakos, Trujillo & Pohlen 2008; Trujillo et al. 2009; Wang et al. 2011), or arguing for the redistribution of populations from the inner to outer regions (Roškar et al. 2008), possibly coupled with bar formation (Debattista et al. 2006).

The physics underpinning galaxy structure is relatively immature, despite the very long history dating back to Knox-Shaw, Reynolds and Hubble and essentially consists of spot-check simulations which focus on a particular phenomena in a non-cosmological context (for recent developments see Roškar et al. 2010; Agertz, Teyssier & Moore 2011). For example, numerical models can readily produce bar, pseudo-bulge, spiral patterns and spheroidal structures through coupled rotation, secular evolution, shock-wave propagation and merging history. Until recently the very thin nature of the spiral discs has presented a particular challenge for numerical models, with numerical simulations in particular forming small thick discs, mainly because of the high level of merging (Navarro & Steinmetz 1997). This is in stark contrast to a number of independent empirical studies (e.g. Driver et al. 2007; Gadotti 2009; Tasca & White 2011) which estimate that approximately 60 per cent of the stellar mass in the Universe today lies within disc systems, suggesting a more quiescent merger history (however, see Hopkins & Quataert 2010 on the stability of gas rich discs). In addition, studies by Menéndez-Delmestre et al. (2007) suggest that up to 67 per cent of spiral galaxies contain a barred structure, further complicating simulation efforts. However, numerical simulations are now starting to produce realistic disc systems (Governato et al. 2007, 2009; Agertz, Teyssier & Moore 2009; Agertz et al. 2011) albeit with heavily controlled initial conditions, more quiescent merger histories and a greater degree of gas infall.

Beyond a distance of $\sim$100 Mpc detailed structural studies have been relatively rare and mostly confined to the deep yet very narrow pencil beam surveys from the *Hubble Space Telescope* (*HST*). It was only following the refurbishment of *HST* that structural analysis once again became a mainstream study (Driver et al. 1995a,b, 1998). *HST* provides kpc resolution across the full path length of the Universe, which is now also becoming possible with adaptive optics (AO) ground-based systems (Huertas-Company et al. 2007). The conjunction of development in numerical models and this new ability to resolve the shapes and sizes of galaxies at any distance has led to a dramatic renewed interest in structural analysis. One interesting claim is the apparent remarkable growth of galaxy sizes since intermediate redshifts (e.g. Barden et al. 2005; McIntosh et al. 2005; Trujillo & Pohlen 2005; Trujillo et al. 2006, 2007; Weinzirl et al. 2011), potentially supporting the notion of recent growth in disc systems following an earlier aggressive merger phase at $z \sim 2$ (see Driver et al. 1996, 2005). An alternative suggestion which does not require galaxy growth through mergers is the transformation of some of these so-called 'red nuggets' into the bulges of disc galaxies via the accretion of a cold gas disc (Graham 2011).

However, structural analysis is not trivial to implement and interpret correctly, and is plagued by a number of key issues. In particular the following.

(i) Wavelength bias. At different wavelengths, light traces varying stellar populations (Block et al. 1999). Typically this is a young stellar population at shorter wavelengths and an older stellar population at longer wavelengths. For this reason, it is vital when comparing structural properties to compare properties measured at the same rest wavelength.

(ii) Dust attenuation. Dust is predicted to modify not only the recovered total flux as a function of wavelength (e.g. Pierini et al. 2004; Tuffs et al. 2004) but also galaxy sizes, shapes and concentrations (see e.g. Möllenhoff, Popescu & Tuffs 2006; Graham & Worley 2008). Dust can vary enormously from system to system with significant environmental dependencies and strong evolution with redshift. Each individual galaxy ultimately requires either a dust correction or analysis at rest-near-infrared (NIR) wavelengths where dust will have a smaller impact (see photon escape fraction curve in Driver et al. 2008). It is obvious that any attempt to model the dust in galaxies raises the larger problem of degeneracies appearing between additive and subtractive flux components.

(iii) Local minima during the minimization process. For a single profile fit there are often seven free parameters, with that number

rising for multicomponent fits. The surface within this parameter space is known to be complex, containing multiple local minima representing potentially non-physical results e.g. a bulge which contributes significantly more flux to the outer regions of a galaxy than the disc (Graham 2001). Other than manual checks of the output, various methods may be employed to reduce the risk of divergence on an incorrect result including constraints applied during the minimization routine and employing an automated logical filter (e.g. Allen et al. 2006).

(iv) What lies below the limiting isophote. Whilst the surface brightness profile of some galaxies behaves as expected out to very faint magnitudes (e.g. NGC 300: Bland-Hawthorn et al. 2005; Vlajić, Bland-Hawthorn & Freeman 2009; NGC 7793: Vlajić, Bland-Hawthorn & Freeman 2011), the potential myriad of phenomena present in the outer wings of many systems may cause deviations away from a typical light profile. These include truncated and antitruncated discs (Erwin, Beckman & Pohlen 2005; Pohlen & Trujillo 2006), ultraviolet (UV) excesses (Bush et al. 2010), tidal debris, haloes (Barker et al. 2009; McConnachie et al. 2009) and minor merger fossil records (Martínez-Delgado et al. 2010). In fact, the outer regions of galaxies may defy any systematic profile fitting into a restricted number of structures. The accuracy of any estimation of the background sky and gradients therein will also no doubt affect analyses of these outer structures.

(v) The number of components required. When considering the structure of very nearby galaxies, the deeper one looks the more one finds. Some galaxies, even in the dust-free 3.8-μm bands, require up to six components (Buta et al. 2010) before a satisfactory fit can be obtained. In many cases there is uncertainty as to how many components are required, how to quantitatively decide this in an automated and repeatable fashion, and which components are fundamental and which perhaps secondary. For example, should bar and pseudo-bulge flux be incorporated into a single disc model or kept distinct.

(vi) Sky estimation. Understanding the background sky level at the position of your primary object of interest is crucial in producing meaningful measurements of that galaxy. Considerations must be made in regards accuracy and speed of estimating the background.

(vii) Systematic selection bias. Sample bias will be introduced due to size, resolution, orientation, profile shape and smoothing scale limitations (Phillipps & Disney 1986; Driver 1999). Samples of galaxies are usually selected based on global criteria, such as magnitude. However, it becomes non-trivial to transcribe these global limits into appropriate limits for galaxy subcomponents e.g. a certain type of disc may only have been detected because it also contains a prominent bulge.

There are several publicly available galaxy modelling codes in common usage including GIM2D (Simard et al. 2002), BUDDA (de Souza, Gadotti & dos Anjos 2004), GASPHOT (Pignatelli, Fasano & Cassata 2006) and GALFIT 3 (Peng et al. 2010). In addition, there are a number of software pipelines, wrappers around contemporary astronomy software that aim to automate the process of galaxy modelling including GALAPAGOS (Barden et al., in preparation) and PYMORPH (Vikram et al. 2010). These packages all have their advantages and disadvantages and have been compared in a number of external studies (e.g. Häussler et al. 2007; Hoyos et al. 2011) in addition to their own internal comparisons, and so we refer the reader to these papers for discussions of the pros and cons between 1D versus 2D fitting and the actual minimization algorithms employed. For this body of work GALFIT was chosen for its ease of use and high-quality realistic model outputs, plus the ability to perform simultaneous modelling of nearby neighbours to the primary galaxy.

In this series of papers we introduce and utilize Structural Investigation of Galaxies via Model Analysis (SIGMA), an automated code designed to produce single-Sérsic and multicomponent profile fits for galaxies in the GAMA data set. Using SIGMA, this paper presents one of the largest catalogues of multiwavelength single-Sérsic model fits; 167 600 galaxies modelled independently across nine bandpasses. This catalogue is currently in use to aid in measurement of the evolution in the size–(stellar mass) distribution of galaxies (Baldry et al., in preparation); explore star formation trends as a function of morphology (Bauer et al., in preparation); to further understand the cosmic spectral energy distribution (SED) from 0.1 μm to 1 mm (Driver et al. 2011); to apply dust corrections to galaxies observed at multiple inclinations (Grootes et al., in preparation); to explore the dust properties and star formation histories of local submillimetre-selected galaxies (Rowlands et al. 2012); better constrain stellar mass measurements by providing total flux corrections (Taylor et al., in preparation); comment on the quenching of star formation in the local Universe (Taylor et al., in preparation); explore the relation between galaxy environments and their star formation rate variations (Wijesinghe et al., in preparation); provide a new method for automatic morphological classification (Kelvin et al., in preparation) and further explore the relation between environment (i.e. halo mass; Robotham et al. 2011), morphology and structure (Kelvin et al., in preparation). Future studies building on SIGMA will incorporate advanced logical filtering and profile management to produce multicomponent fits for a low-redshift sample, allowing full structural decomposition into bulge, disc, bar, etc. (Kelvin et al., in preparation).

This paper is organized as follows. We first outline the GAMA data in Section 2. We describe the SIGMA pipeline developed to process this data and produce robust 2D galaxy models in Section 3 and present SIGMA single-Sérsic results for 167 600 objects modelled independently in *ugrizYJHK* from the GAMA Phase I study in Section 4. From this large catalogue we establish a common coverage sample of 138 269 galaxies. Finally, we further explore the wavelength dependence on recovered structural parameters in Section 5. A standard cosmology of $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{m} = 0.3$, $\Omega_\Lambda = 0.7$ is assumed throughout.

2 DATA

The GAMA survey is a combined spectroscopic and multiwavelength imaging programme designed to study spatial structure in the nearby ($z < 0.25$) Universe on scales of 1 kpc to 1 Mpc (see Driver et al. 2009 for an overview). The survey, after completion of Phase I, consists of three regions of sky each of 4° (Dec.) × 12° (RA), close to the equatorial region, at approximately 9^h (G09), 12^h (G12) and $14^\mathrm{h}.5$ (G15) RA (see Table 1 and Fig. 1). The three regions were selected to enable accurate characterization of the large-scale structure over a range of redshifts and with regard to practical observing considerations and constraints. They lay within areas of sky scheduled for survey by both Sloan Digital Sky Survey (SDSS; Abazajian et al. 2009) as part of its Main Survey, and United Kingdom Infrared Telescope (UKIRT) as part of the UKIRT Infrared Deep Sky Survey Large Area Survey (UKIDSS-LAS; Lawrence et al. 2007). These data provide moderate depth and resolution imaging data in *ugrizYJHK* suitable for analysis of nearby galaxies. The accompanying spectroscopic input catalogue was derived from the SDSS PHOTO parameter (Stoughton et al. 2002) as described in Baldry et al. (2010). The GAMA spectroscopic programme (Robotham

Table 1. GAMA region definitions. The GAMA main survey definitions are based on SDSS extinction corrected *r*-band Petrosian magnitude limits, the depth of which varies between $r = 19.4$ mag in G09/G15 and $r = 19.8$ mag in G12. Comparison magnitude limits are shown in brackets for reference. Number counts and redshift completeness are based on objects which passed star–galaxy separation in the GAMA *TilingCatv11* (see Baldry et al. 2010 for further details).

Region	RA (J2000)	Dec. (J2000)	$r_{\rm lim}$	$N_{\rm obj}$	$z_{\rm comp}$
G09	$129.^{\circ}0 < \alpha < 141.^{\circ}0$	$-1.^{\circ}0 < \delta < 3.^{\circ}0$	19.4 (19.8)	30 289 (48 548)	99.23 per cent (79.36 per cent)
G12	$174.^{\circ}0 < \alpha < 186.^{\circ}0$	$-2.^{\circ}0 < \delta < 2.^{\circ}0$	19.8 (19.4)	50 868 (32 747)	99.12 per cent (99.39 per cent)
G15	$211.^{\circ}5 < \alpha < 223.^{\circ}5$	$-2.^{\circ}0 < \delta < 2.^{\circ}0$	19.4 (19.8)	33 205 (51 217)	98.95 per cent (79.27 per cent)

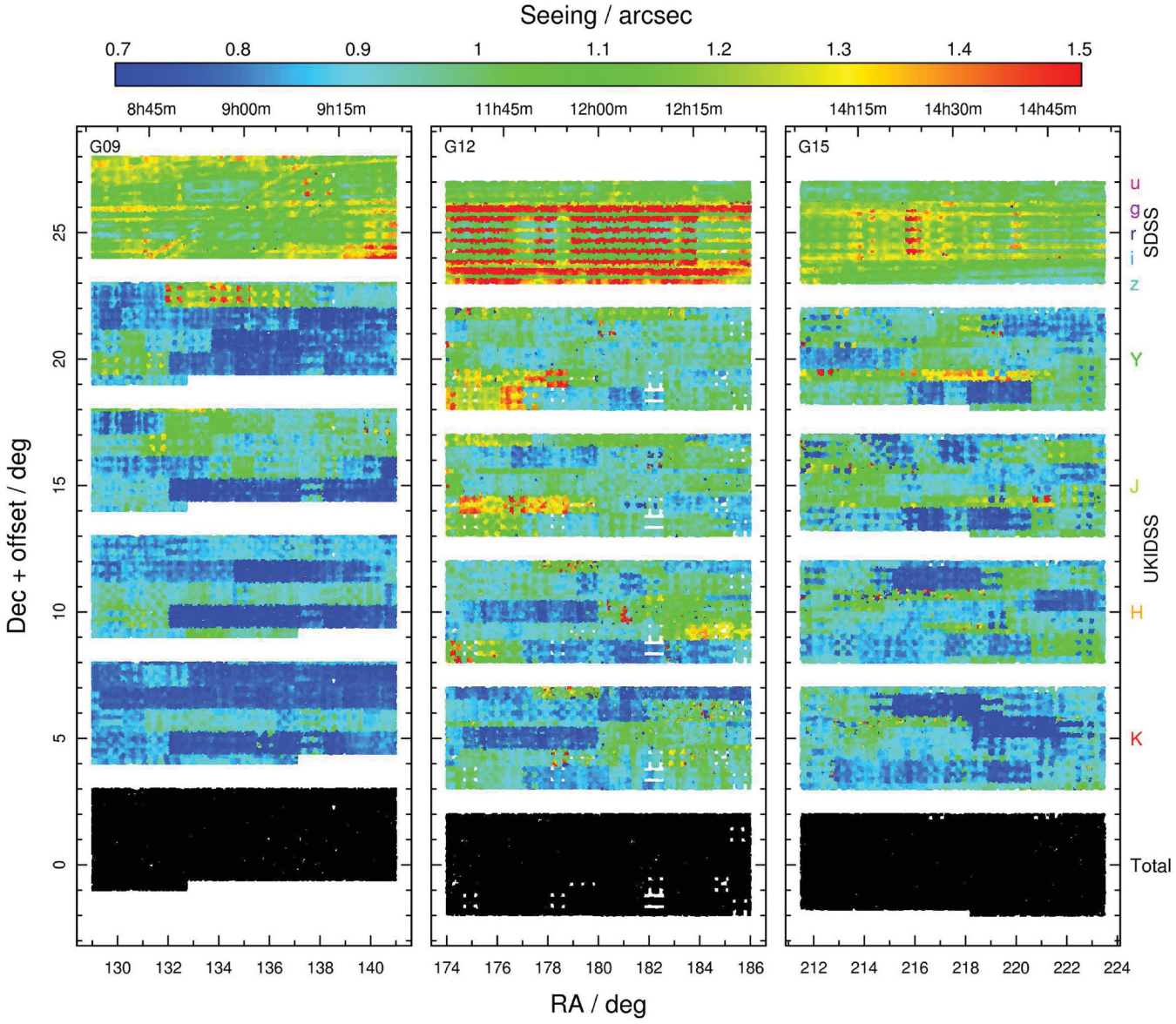

Figure 1. Coverage in each bandpass across the three GAMA regions is shown, where white space indicates a lack of coverage at that position. Different bandpasses are offset in declination for plotting purposes, with the bottom row (black data points) positioned at the correct GAMA coordinates. The five SDSS bands (*ugriz*) were taken simultaneously and therefore are represented together, whilst the four UKIDSS bands (*YJHK*) are shown separately. The data points are coloured according to the local value of the PSF FWHM at that position, giving an indication of the variation in seeing (SDSS data are taken from the *r* band). See Section 3.4 for further details on the derivation of the PSFs. The bottom row represents a common coverage area, containing galaxies that have been observed across all nine bands (92.8 per cent of the total). A common coverage sample is derived from this area, and used for all subsequent figures.

et al. 2010) commenced in 2008 using AAOmega on the Anglo-Australian Telescope to obtain distance information (redshifts) for all galaxies brighter than $r < 19.8$ mag. The survey is ∼99 per cent complete to $r < 19.4$ mag in G09 and G15 and $r < 19.8$ mag in G12 (see Table 1, column 6), with a median redshift of $z \sim 0.2$. Full details of the GAMA Phase I spectroscopic programme, key survey diagnostics and the GAMA public and team data bases are given in Driver et al. (2011).

The data used in this paper are obtained from the GAMA data base (Driver et al. 2011), and include reprocessed imaging from the SDSS (*ugriz*) and UKIDSS-LAS (*YJHK*) archive as described in Hill et al. (2011). The reprocessing involves the creation of large single image mosaics for each region in each filter, commonly referred to as SWARPed images (*swpim*) due to the SWARP software used in their creation (Bertin et al. 2002). Associated weight map mosaics (*swpwt*) are also constructed. The mosaicking process is described in full in Hill et al. (2011). In brief, all native reduced frames are downloaded from the respective archives [SDSS Data Release 7 (DR7) and ROE/WFAU] and scaled to a single uniform zero-point. For SDSS, the input data are the corrected (*fpC*) DR7 files, with the data having already been bias subtracted and flat-fielded as part of the SDSS *frames* pipeline (Stoughton et al. 2002, section 4.4). The UKIDSS-LAS data have been collected from the UKIDSS Early Data Release (EDR; Dye et al. 2006) and data releases 1 and 2 (DR1: Warren et al. 2007a; DR2: Warren et al. 2007b). The UKIDSS project is defined in Lawrence et al. (2007). UKIDSS uses the UKIRT Wide Field Camera (WFCAM; Casali et al. 2007). The photometric system is described in Hewett et al. (2006), and the calibration is described in Hodgkin et al. (2009). The pipeline processing and science archive are described in Irwin et al. (in preparation) and Hambly et al. (2008).

Once these input data have been obtained and calibrated, SWARP is then used to combine them into a single image mosaic at a resolution of 0.339 arcsec pixel^{-1} in the TAN projection system (Calabretta & Greisen 2002) centred within each GAMA region as appropriate. Note that we are using version 2 SWARP mosaics scaled to a slightly higher resolution (0.339 arcsec)[1] than the version 1 mosaics (0.4 arcsec) described in Hill et al. (2011). Version 2 mosaics are a minimum of 193 900 × 79 700 pixels each, with each individual FITS file ∼60 GB in size. The process used to create the version 2 mosaics is identical except the regions have been expanded in preparation for GAMA Phase II operations and at higher resolution in preparation for matching to VISTA data in due course.[2]

As part of the SWARP mosaicking process the background is removed on each individual frame prior to merging using a 256 × 256 pixel median filtered mesh which itself is median filtered within a 3 × 3 mesh. The original SDSS and UKIDSS data are typically held in chunks of 2048 × 1489 and 2072 × 2072 pixels, respectively, at comparable pixel scales (SDSS: 0.396 arcsec pixel^{-1}; UKIDSS: 0.4 arcsec pixel^{-1}). At the native pixel resolution the mesh size therefore equates to 101.4 × 101.4 and 102.4 × 102.4 arcsec^2, respectively, and so structures with half-light radii less than ∼17 arcsec should be unaffected by the background smoothing.[3]

In addition to the science image frames are the associated weight maps. Because of the zero-point normalization across all data, and overlapping edge duplication in the SDSS data, the actual weight map values produced by SWARP are an approximation of their correct value. However, the weight maps remain useful as a record of which stars can be associated with which pre-mosaicked frame for the purposes of detailed point spread function (PSF) modelling (described in Section 3.4).

To create our sample of galaxies for modelling, we extracted 167 600 galaxies from the GAMA Tiling Catalogue version 11 (*TilingCatv11*), selecting all galaxy-like objects using the GAMA catalogue flag SURVEY_CLASS > 1.[4] The output from these galaxies is stored in the catalogue *SersicCatv07*, presented in Section 4. These galaxies, the mosaicked images and the weight maps constitute our input data set and are all available from the GAMA data base[5] as *TilingCatv11*, x.mosaic.v2.fits and x.weight.v2.fits, where $x = ugrizYJHK$.

3 SIGMA: AUTOMATED GALAXY MODELLING

SIGMA is an automated front-end wrapper which utilizes a wide-range of image analysis software and a series of logical filters and handlers to perform bulk structural analysis on an input catalogue of galaxies. This is primarily achieved through the use of SOURCE EXTRACTOR (Bertin & Arnouts 1996), PSF EXTRACTOR (PSFEX; Bertin and Delorme, private communication) and GALFIT 3 (Peng et al. 2010), with additional packages also created and utilized to aid in the fitting process. Key to this process is the galaxy modelling software GALFIT. GALFIT is able to create a realistic model of each input galaxy by fitting one or more analytical functions (e.g. Sérsic, exponential, Ferrer, Moffat, Gaussian) in multiple combinations.

Principle in the available GALFIT fitting functions used throughout this paper is the Sérsic profile (Sérsic 1963, 1968; Graham & Driver 2005) which describes how the galaxy light profile varies as a function of radius. The Sérsic equation provides the intensity I at a given radius r as given by

$$I(r) = I_{\rm e} \exp\left[-b_n\left(\left(\frac{r}{r_{\rm e}}\right)^{1/n} - 1\right)\right], \quad (1)$$

where $I_{\rm e}$ is the intensity at the effective radius $r_{\rm e}$, the radius containing half of the total light, and n is the Sérsic index which determines the shape of the light profile (see Fig. 2). The value of b_n is a function of Sérsic index and is such that $\Gamma(2n) = 2\gamma(2n, b_n)$,[6] where Γ and γ represent the complete and incomplete gamma functions, respectively (Ciotti 1991). Varying the Sérsic index parameter n allows one to model a wide range of galaxy profile shapes, with $n = 0.5$ giving a Gaussian profile, $n = 1$ an exponential profile suitable for galactic discs and $n = 4$ a de Vaucouleurs profile commonly associated with massive spheroidal components such as elliptical galaxies.

The only inputs required by SIGMA are the imaging data itself and the locations of the primary galaxies within them which are to be modelled. All additional parameters and starting values for extra neighbouring objects in the field of view (secondary objects),

[1] This increased resolution has been chosen to match that which is expected for future VISTA VIKING data releases, allowing easy cross-wavelength cross-facility comparison of data. Original SDSS and UKIDSS resolutions of 0.396 and 0.4 arcsec, respectively, place a limit on how high one is able to artificially increase the resolution of mosaicked data, requiring increasing amounts of interpolation with increasing artificial resolution. Further details may be found in Liske et al. (in preparation) .

[2] These larger, higher resolution version 2 mosaics will be released shortly via the GAMA website http://www.gama-survey.com

[3] UKIDSS *J*-band data and selected UKIDSS EDR fields in both *H* and *K* bands were microstepped. These data are typically stored in chunks of 4103 × 4103 pixels at a native resolution of 0.2 arcsec pixel^{-1}, giving a mesh size of 51.2 × 51.2 arcsec^2.

[4] SURVEY_CLASS is a flag present in many GAMA catalogues allowing one to quickly select pre-defined subsets of the GAMA data.

[5] The GAMA data base can be found at http://www.gama-survey.org/database

[6] b_n can trivially be calculated within R using the relation b_n = qgamma(0.5, 2n), where qgamma is the quantile function for the Gamma distribution.

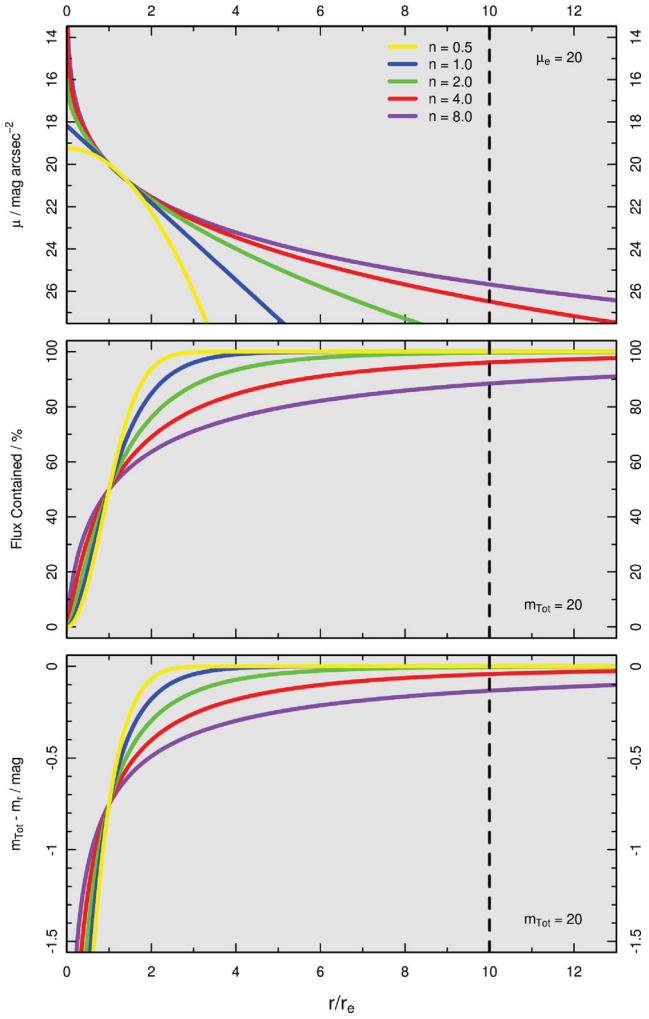

Figure 2. The Sérsic profile (equation 1) describes how a galaxy light profile varies as a function of radius, shown here for five distinct values of Sérsic index *n*. Top: surface brightness at a given radius. Middle: flux contained within a given radius. Bottom: the magnitude offset between the total magnitude of the galaxy and the Sérsic magnitude at a given radius.

including PSF evaluation, are determined by SIGMA on the fly on a per-galaxy basis. All scripting and additional programming is written in the open source and freely available R programming language (R Development Core Team 2010). Further information on the invocation of SIGMA may be found in Appendix A.

SIGMA operates in a semimodular fashion, with an overarching master script calling and linking several key modules within. Each module is specialized in performing and handling a different task. A summary of each module and its purpose is shown in Table 2, and a schematic of the SIGMA data-flow process is shown in Fig. 3. The average run-time to profile a single primary object is 15 s per processor[7] sustained over several hundred thousand objects.

3.1 SIGMA master script

When initializing SIGMA a number of options are specified. In addition to several expected inputs such as band, naming conventions and the combination and types of GALFIT functions to be modelled, SIGMA also allows the workload to be split over several processors. Using a greater number of processors directly decreases the amount of time required to analyse every primary object in the input catalogue, and is limited only by available hardware. A full summary of SIGMA input options is given in Appendix A.

To begin, SIGMA's master script loads into memory the entirety of the GAMA input catalogue and defines a naming convention for each primary object based on its own unique identifier (SIGMA_INDEX). A template master comma-separated variable (CSV) catalogue is created into which all of the output data will accumulate as SIGMA loops across each primary galaxy. Once the set-up is complete, the master script will loop across every primary object in turn. If for any reason a primary galaxy causes a software crash, with attempted fixes as detailed in subsequent sections unsuccessful, SIGMA will report how far it was able to progress and record a NULL result before proceeding on to the next primary galaxy in the input catalogue. We now discuss each module from Table 2 in turn.

[7] Using current computer hardware at the University of St Andrews. This consists of a 16 core Intel Xeon E5520 server with 48 GB RAM.

Table 2. Summary of the modules that comprise SIGMA, a brief description of their purpose and a list of the file outputs produced by each.

Module	Description	File outputs
CUTTERPIPE (Section 3.2)	Creates a science and weight map cutout from the master GAMA mosaics using the CFITSIO routine FITSCOPY (Fig. 4) and performs an additional local background sky subtraction using SOURCE EXTRACTOR (Fig. 5).	*cutim* *cutwt*
STARPIPE (Section 3.3)	Determines which frames contributed flux to the primary galaxy, and creates a catalogue of stars that lie within these frames using SOURCE EXTRACTOR.	*psfws* *psfwt* *psfct*
PSFPIPE (Section 3.4)	Generates an empirical 2D PSF at the primary object position using PSFEX (Fig. 6).	*psfss* *psfim* *psfsr*
OBJECTPIPE (Section 3.5)	Calculates starting parameters for size, brightness, position angle and ellipticity for the primary galaxy and any secondary neighbours (galaxies and stars) using SOURCE EXTRACTOR. A dynamic search algorithm is used to attempt to detect the primary galaxy. Elongated objects (such as satellite trails) are removed from the secondary catalogue and instead added to a bad pixel mask.	*objct* *segim*
GALFITPIPE (Section 3.6)	Fits an analytical function in 2D to the science image using GALFIT. Both primary and secondary objects are modelled, with any detected errors/crashes flagged and a fix attempted on a per-galaxy basis (Figs 7 and 8).	*segfr* *extct* *objim*

3.2 CUTTERPIPE: image cutout and preparation

The CUTTERPIPE module creates and prepares the fitting image to be fed into GALFIT. Version 2 mosaics of the three GAMA regions are used as an input to CUTTERPIPE, with a full description of the construction and manipulation of these files given in Hill et al. (2011) and summarized in Section 2.

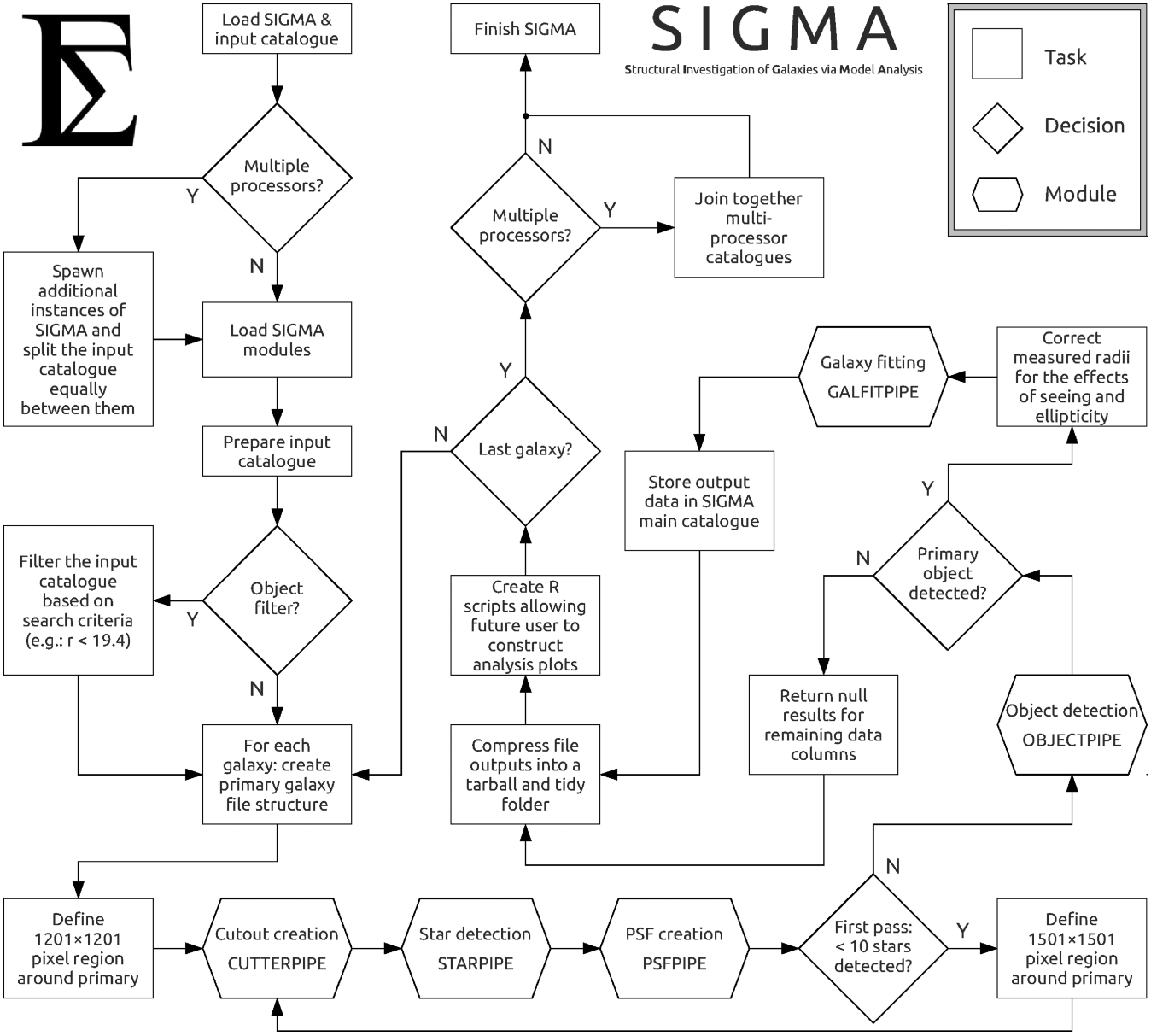

Figure 3. A flowchart describing SIGMA's operation. Required inputs are imaging data in the form of GAMA mosaics (including weight maps), and an input catalogue with a list of galaxy coordinates.

CUTTERPIPE's first task is to create the core cutout of the science image and its associated weight map. Using the WCS information stored in the FITS header of the appropriate mosaicked image, CUTTERPIPE converts the input RA/Dec. into an *x*/*y* pixel coordinate. The upper and lower limits of a 1201 × 1201 pixel (∼400 × 400 arcsec^2) region centred on the primary galaxy are determined. Using the NASA HEASARC package's CFITSIO subroutine library, namely the routine FITSCOPY, cutouts centred on the primary galaxy on both the mosaicked science image, *sw pim*, and mosaicked weight map, *swpwt*, are created. These cutouts are named *cutim* and *cutwt*, respectively. FITSCOPY was found to be the most efficient routine at dealing with the large mosaic files in use, able to quickly analyse the input file and read into memory only the relevant area of interest, thereby reducing file handling time significantly.

The process of creating the GAMA mosaics alters a number of keywords in the FITS header in order to better describe the nature of the mosaicked data. The mosaic headers are copied over to *cutim* and *cutwt* during their creation. Several of these keywords are required later in the fitting process by GALFIT in order to generate a sigma map (an image showing the 1σ confidence interval at every pixel). CUTTERPIPE reverts these to typical pre-mosaic values which are more appropriate for a smaller single image rather than a larger mosaic. GAIN, RDNOISE, NCOMBINE and EXPTIME are set to values of 0.5, 3, 1 and 1, respectively.[8]

An estimate of the local background sky is then made with SOURCE EXTRACTOR (v2.8.6; Bertin & Arnouts 1996) using a variable background grid in a 3 × 3 mesh configuration. Possible grid sizes are 32 × 32, 64 × 64 and 128 × 128 pixels. The size of the chosen background grid is dependent upon the size of the primary galaxy: larger galaxies will lead to a larger background grid being used so

[8] These typical values are averages taken from pre-mosaicked data frames.

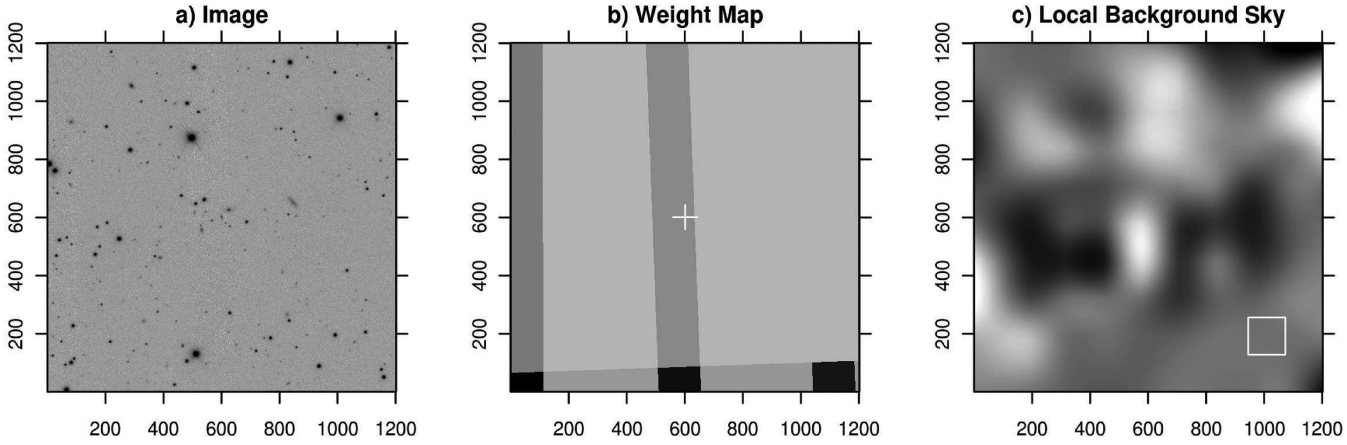

Figure 4. A series of set-up images centred on G00196053. (a) A science image cutout of 1201 × 1201 pixels (∼400 × 400 arcsec^2) in the *r* band, with the primary galaxy centred at 601 × 601. The image is scaled logarithmically from $-1\sigma_{\rm sky}$ to $40\sigma_{\rm sky}$, where $\sigma_{\rm sky}$ is the typical rms of the sky in the *r* band. (b) A weight map cutout of 1201 × 1201 pixels (∼400 × 400 arcsec^2) in the *r* band, with the primary galaxy located in an overlap region at the white cross-hairs. Darker shades indicate larger values, therefore, a greater number of overlapping frames. (c) The additional local background sky removed from this cutout. The *z*-axis range shown is ±0.5 counts. For this galaxy, a background mesh size of 128 × 128 pixels (white square) was chosen in addition to the coarse 256 × 256 background mesh already applied in creating the GAMA mosaics.

as not to contaminate the sky estimate with galaxy flux. An initial basic estimate of the total size of the primary is given by

$$r_{\rm tot} = 2\,r_{99}, \tag{2}$$

where r_{99} is the radius of the primary galaxy which contains 99 per cent of the flux. This is obtained from the SOURCE EXTRACTOR parameter FLUX_RADIUS setting PHOT_FLUXFRAC=0.99.[9] If $r_{\rm tot} < 128$, CUTTERPIPE rounds up $r_{\rm tot}$ to the nearest available grid size and performs a background subtraction on the science image as appropriate. If $r_{\rm tot} \geq 128$, no background subtraction is necessary, as the master GAMA mosaics have a 256 × 256 pixel background grid subtraction already applied. Although the value of actual subtracted sky varies with position on the cutout image, the specific value at the position of the primary galaxy, $\rho_{\rm sky}$ is recorded through the SOURCE EXTRACTOR parameter BACKGROUND. The error on the background sky estimate is then given by

$$\Delta\rho_{\rm sky} = \frac{\sigma_{\rm sky}}{\sqrt{0.9\,n_x\,n_y}}, \tag{3}$$

where $\sigma_{\rm sky}$ is the rms of background sky counts across the cutout, and n_x and n_y are the dimensions of the cutout in the *x* and *y* dimensions, respectively. The background sky typically encompasses ∼90 per cent of any given cutout, and hence a factor of 0.9 is introduced into the above calculation to account for this. After extensive testing, this variable background mesh method for sky estimation was found to be the most robust at removing small-scale sky fluctuations in the data without subtracting real galactic light from objects (see Section 4.2.1 for sky results discussion).

An example image cutout, weight map and background estimation map are shown in Fig. 4.

[9] Elsewhere in this paper, SOURCE EXTRACTOR FLUX_RADIUS will typically refer to $r_{\rm e}$, a radius containing 50 per cent of the flux of the primary galaxy. It is worth noting however that a size estimate produced by SOURCE EXTRACTOR in this manner is known to be smaller than the true galaxy value, scaling as a function of Sérsic index and thus absolute magnitude. This effect has been accounted for, and does not adversely affect any of the analysis or results presented in this paper.

3.3 STARPIPE: star detection

STARPIPE uses SOURCE EXTRACTOR to create a catalogue of star-like objects with which to create a PSF in the subsequent PSFPIPE module (Section 3.4).

The first step is to determine which of the original pre-mosaic frames contain the primary galaxy. This step is non-trivial, as a single cut-out image (*cutim*) may contain data from several pre-mosaicked frames overlapping at random angles to each other, with only some of the frames contributing flux (and therefore seeing information) to the primary galaxy. Calculating frame ownership is crucial in PSF determination, as using stars from non-contributing frames would skew the PSF estimate away from its true shape at the position of the primary galaxy. A method was devised to determine contributing frames using the information within the GAMA weight maps. Each pre-mosaicked frame is assigned a numerical value based on the global variance of the data for that frame. This value, repeated for each pixel, becomes the weight map for that individual frame. Weight-map values are essentially unique to several significant figures, and therefore useful in identifying that particular frame. During the SWARP process, overlapping imaging data are median combined (setting the SWARP argument COMBINE_TYPE to MEDIAN) whereas weight maps are co-added to produce a global weight map representing the change in the variance across the data. When two or more frames overlap, their individual weight-map values are summed. Larger values indicate a greater number of overlapping frames.

The value of the weight map at the primary position is determined, with all pixels of that value clearly contributing flux to the primary galaxy. This defines the initial primary region. However, since this primary region may be an overlap region itself, parent frames must also be determined. The weight-map values of all bordering pixels to the primary region are determined. Higher pixel values indicate a region which contains data from additional frames that did not contribute flux to the primary region, and so these pixels are discarded. Lower values (if any) indicate parent frames of the primary region, and (if they exist), their pixel positions are added to the primary region. This process will continue until all pixels are accounted for across the cutout. As an example of frame determination, contrast Fig. 4(b) with the shaded red regions in Fig. 7.

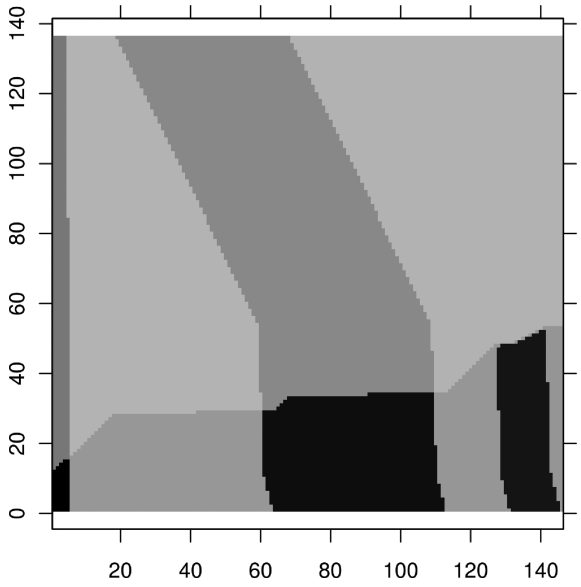

Figure 5. A simplified version of the weight map shown in Fig. 4(b). A single cutout from a GAMA mosaic may contain data from several frames observed on different nights in different seeing conditions. The weight map allows us to determine which frames contributed flux to the primary galaxy and which did not, thereby allowing a PSF to be constructed representing the seeing for that galaxy. Simplification of the weight map removes superfluous information contained within the original cutout weight map, reducing the number of pixels to analyse typically by a factor of ∼150, and significantly speeding up STARPIPE.

Pixel determination via this technique is time intensive for the full 1201 × 1201 cutout region, as it requires analysis of 1.4 million pixels for each galaxy. A more efficient method is to reduce the number of pixels that require analysis by simplifying the weight map to its minimal number of pixels which still describe the nature of the data. Duplicate rows and columns in the weight map are removed, producing a simplified weight map, *psfws*, typically of order ∼100 × 100 pixels. This thereby reduces the number of pixels needing to be analysed by a factor of ∼150, significantly speeding up the primary region determination. Fig. 5 is one such simplification of the cutout weight map shown in Fig. 4(b), in this case reducing the number of pixels from 1.4 million to ∼20 000.

Once a primary region is determined, a local star catalogue must be created. A modified version of *cutwt*, *psfwt*, is created, setting all non-primary pixels to a weight value of zero. This will bar SOURCE EXTRACTOR detecting any objects in those regions. The central 25 × 25 pixel region is also set to a weight value of zero to ensure against the primary galaxy being falsely classified as a star and used in later PSF analysis.

A SOURCE EXTRACTOR parameter file is created to output NUMBER, X_IMAGE, Y_IMAGE, FLUX_RADIUS, FLAGS, FLUX_APER(1), FLUX_MAX, ELONGATION, VIGNET(25,25) and BACKGROUND. The numeric values following VIGNET determine the ultimate size in pixels of the 2D PSF created subsequently in PSFPIPE. A detection threshold of 2σ above the background is specified, along with a minimum object detection area of 10 pixels, a 25 000 count saturation level and a fixed sky pedestal value of zero counts. The image is filtered through a 5 × 5 pixel Gaussian convolution kernel with $\Gamma = 2$ pixels. SOURCE EXTRACTOR defaults are used everywhere else. Using these settings, SOURCE EXTRACTOR is run across the cutout image *cutim* using the weight map image *psfwt*. An output catalogue, *psfct*, is created in the *FITS_LDAC* format, a format which saves the image data as well as the catalogue.

3.4 PSFPIPE: PSF estimation and creation

The PSF describes the blurring effect of both the atmosphere and the telescope optics on our imaging data. Observed galaxy images have had their flux redistributed according to this PSF. The galaxy flux most affected by the PSF blurring is that which emanates from the core regions, where the gradient of the light profile is at its steepest. It is therefore crucial to have a good understanding of the PSF when considering fitting smooth analytical galaxy models to imaging data. Furthermore, most current galaxy modelling software weights model fitting towards higher signal-to-noise ratio regions (typically the same core regions), increasing the importance of accurate, reliable PSF estimation.

The PSFPIPE module is a wrapper around the PSF extraction software PSFEX v3.3.4 (Bertin, private communication),[10] and produces a 2D PSF model to be taken into account at the later galaxy modelling stage. PSFEX extracts precise models of the PSF from images pre-processed by SOURCE EXTRACTOR, allowing for a wide range of PSF's to be quickly and accurately constructed, including arbitrary non-parametric features present in the PSF.

In brief, the sample of objects from the *psfct* catalogue created by STARPIPE is initially used for analysis. PSFEX reduces this object list to a star sample based on a set of pre-defined criteria. A signal-to-noise ratio limit of at least 10 is required, and objects with an eccentricity of $((a - b)/(a + b)) > 0.05$ are removed, where a and b refer to the semimajor and semiminor axes, respectively.[11] Each star's full width at half-maximum (FWHM), Γ, is estimated, with only stars in the pixel range $2 < \Gamma < 10$ accepted. Furthermore, variability in the star sample is limited to the central 50 per cent quantile. After extensive testing on the variation in PSF quality with star sample size, and communication with the authors of PSFEX, we found that a star sample size of at least 10 stars is necessary to ensure that the resultant PSF is not adversely affected by small-number biases. Therefore, if fewer than 10 stars remain in the star sample after selection criteria have been applied, SIGMA will loop back to CUTTERPIPE and expand the cutout region to 1501 × 1501 pixels (∼500 × 500 arcsec^2). The mean number of stars used for PSF estimation in the r band is 24.4, with 10.2 per cent of cutouts containing fewer than 10 stars after the cutout region has been expanded.

Cutout images of each star are pre-stored in the *FITS_LDAC* format of *psfct*, the size of the cutout having been specified at the Source Extraction stage. PSFEX uses the positional information from SOURCE EXTRACTOR to mask nearby neighbours to the final star sample, and presents this sample in the output *psfss* FITS image (Fig. 6a).

The variation in the shape of the stars in the star sample is then modelled in both x and y as a function of position in the field by a 2D nth order polynomial function. Higher order terms in the fit (i.e. x, x^2, etc.) describe the variation in the PSF at positions away from

[10] More information on the PSFEX software may be found at http://www.astromatic.net/software/psfex

[11] PSFEX refers to this quantity as ellipticity rather than eccentricity, however its definition is more akin to that of the latter. We adopt the terminology eccentricity here to avoid confusion with the standard definition of ellipticity used throughout the remainder of this paper, namely $e = 1 - (b/a)$. An eccentricity of 0.05 therefore corresponds to an ellipticity of $e \sim 0.095$.

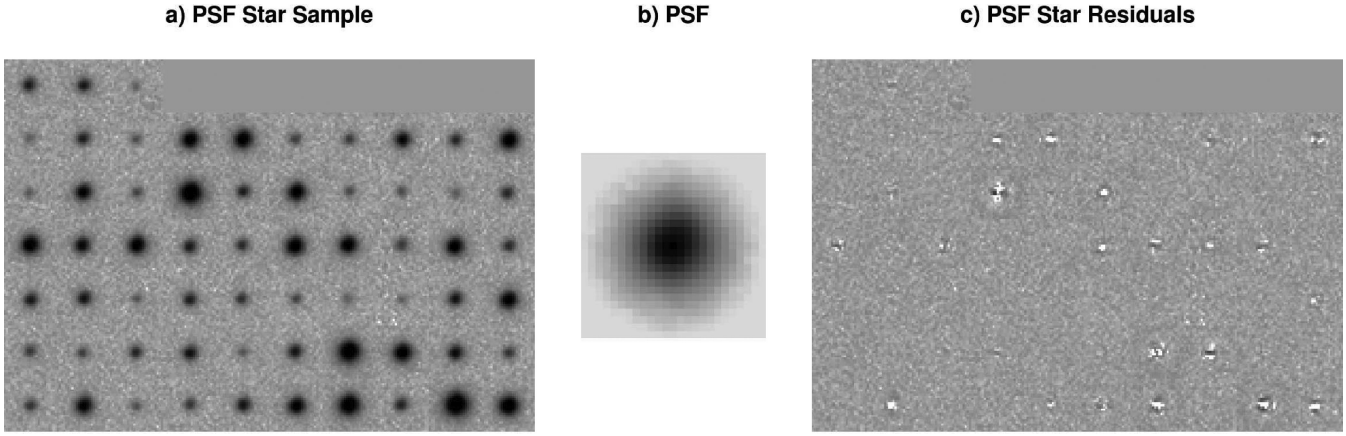

Figure 6. PSFEX generates an empirical PSF (b) from a host sample of representative stars (a). This figure shows 63 sample star cutouts of 25 × 25 pixels each chosen from around GAMA object G00196053, whose real sky positions may be noted in Fig. 7 (orange circles). Panel (c) represents the residual of each star sample with a scaled form of the PSF subtracted from each. (a) and (c) are scaled logarithmically from $-1\sigma_{\rm sky}$ to $40\sigma_{\rm sky}$, where $\sigma_{\rm sky}$ is the typical rms of the sky in the *r* band.

the centre of the frame. Within SIGMA, the primary galaxy is always centred in the cutout image, and so a zeroth-order polynomial was found to adequately describe the PSF. The best-fitting polynomial is sampled at a 1:1 ratio relative to the input data, and an output PSF image *psfim* is produced of the same size as the input cutout stars, 25 × 25 pixels (Fig. 6b). As a consistency check, scaled models of *psfim* are fit to each of the input stars in *psfss*, and a residual map *psfsr* produced (Fig. 6c). Note that some of the PSF residuals still show noticeable structure once the PSF model has been subtracted from the star sample. This is as expected when subtracting a zeroth-order PSF model (only accurate at the location of the primary galaxy in the centre) from a star sample taken over a large area on the sky. Those stars with noticeable residuals therefore are typically either significantly spatially separated in the field of view from the primary galaxy or approaching saturation (or both). Both of these factors are accounted for by PSFEX when constructing the model PSF. The stars chosen as part of the star sample are shown in orange circles in Fig. 7, with each circle numbered according to their position in Fig. 6(a), starting at 1 in the bottom left and increasing horizontally left-to-right and then bottom-to-top.

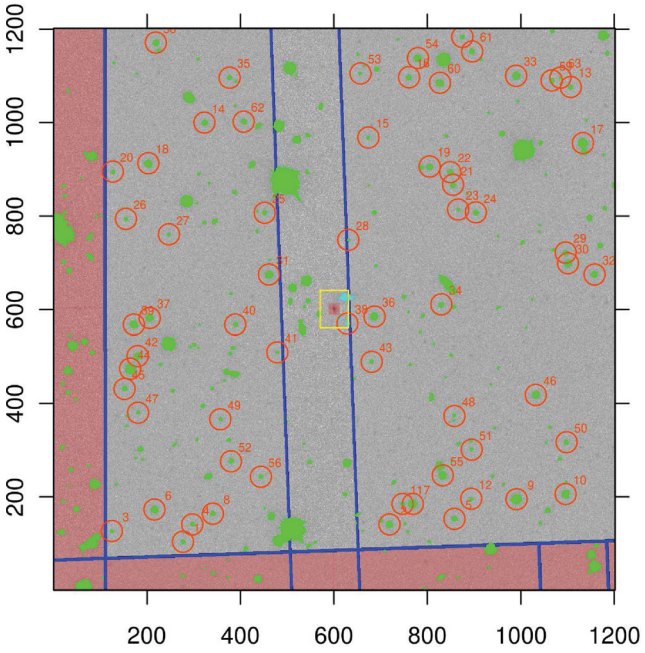

Figure 7. Final detail analysis plot for G00196053. Green shaded areas represent all detected objects, with stars chosen as part of the PSF star sample circled in orange. The GALFIT fitting area is outlined in yellow, and the weight map frame edges in blue. Secondary objects which will be modelled as nearby neighbours in the fitting process are represented in turquoise. Note how no sample stars are taken from frames which did not contribute directly to the flux of the primary galaxy (areas represented with red shading). Consequently, no stars are taken for PSF analysis from the red shaded areas. The image is scaled logarithmically from $-1\sigma_{\rm sky}$ to $40\sigma_{\rm sky}$, where $\sigma_{\rm sky}$ is the typical rms of the sky in the *r* band.

3.5 OBJECTPIPE: object detection

A second catalogue of objects optimized for galaxy detection, *objct*, is created in OBJECTPIPE, to be later fed into GALFITPIPE. This catalogue provides the basic starting parameters for the primary galaxy and any secondary galaxies and stars in the frame. OBJECTPIPE also creates a segmentation map of the frame, modified to mask any erroneous regions of flux in the image which may cause fitting problems (e.g. satellite trails). A SOURCE EXTRACTOR parameter file is created containing X_IMAGE, Y_IMAGE, MAG_AUTO, FLUX_RADIUS, KRON_RADIUS, A_IMAGE, B_IMAGE, THETA_IMAGE, ELLIPTICITY and CLASS_STAR. These give position (*x*/*y*), luminosity, size, position angle and ellipticity for the primary and all secondaries in the field. SOURCE EXTRACTOR settings are similar to those used in STARPIPE, excepting a lower detection threshold of 1.8σ above the background (where σ is the rms as estimated by SOURCE EXTRACTOR), and a lower deblending contrast parameter of 0.0001. The SOURCE EXTRACTOR Neural-Network Weights V1.3 file is used in the creation of the CLASS_STAR parameter, as well as a standard 5 × 5 pixel Gaussian filter, $\Gamma = 2$ pixels, used during object detection.

OBJECTPIPE calls SOURCE EXTRACTOR, and records the results. If initially the primary galaxy is unable to be located within a 5 pixel radius of the input coordinates, OBJECTPIPE will in the first instance decrease the detection threshold in steps of 0.4σ down to 1σ above the background until an object is found, re-running SOURCE EXTRACTOR as appropriate. This usually occurs with faint objects in the

field, or in crowded regions. If the primary object is still unable to be located, the threshold is reset to 1.8σ, and a larger search radius of up to 15 pixels from the input coordinates in 5 pixel steps is tried. This stage accounts for large nearby galaxies whose centroids are not matched to better than 5 pixels, hence requiring a larger detection area. If multiple matches are found, the largest object will be taken to be the primary galaxy. If at this stage the primary galaxy is still not found, OBJECTPIPE will report a null detection, and move on to the next primary in the input catalogue.

Output parameters from SOURCE EXTRACTOR are modified by OBJECTPIPE before being fed into GALFITPIPE with the exception of magnitude which is used unaltered. Position angle is modified to the GALFIT standard (by adding 90°), increasing anticlockwise from the positive x-axis. Ellipticity e is converted to an axis ratio using the relation

$$e = 1 - \frac{b}{a}, \tag{4}$$

with semiminor axis b and semimajor axis a. Half-light radius r_e is estimated using the relation

$$r_e = \sqrt{\left(r_{50}^2 \, \frac{a}{b}\right) - \left(0.32\,\Gamma^2\right)}, \tag{5}$$

where r_{50} is the (unmodified) SOURCE EXTRACTOR half-light radius as given by FLUX_RADIUS (setting PHOT_FLUXFRAC = 0.5) and Γ is the FWHM of the PSF of the primary galaxy. A minimum bound on r_e of 1 pixel is enforced. This conversion corrects for the fact that SOURCE EXTRACTOR's output half-light radii are circularized and based on PSF convolved data, whereas GALFIT radii are along the semimajor axis and intrinsic (non-PSF convolved). The value of 0.32 was derived from simulated test data, see appendix A of Driver et al. (2005) for further details. Fig. B1 shows a comparison of corrected and uncorrected radii against modelled GALFIT radii for all GAMA objects in the r band. This suggests that the revised starting value for r_e is appropriate and an accurate first estimate of the true half-light radius of the primary galaxy. Because of the downhill minimization employed by GALFIT, it is important to provide input parameters as close as possible to the desired solution in order to avoid local minima.

Once physical parameters for the primary galaxy have been determined, a segmentation map of the frame is created to be used as a potential mask for secondary features should modelling them fail. Secondary objects whose ellipticity is greater than 0.95 are excluded from modelling and will instead be masked, as these are determined to be satellite trails or bad data, and consequently difficult to model. Similarly, secondary objects with a stellaricity index of CLASS_STAR>0.8 (see Bertin & Arnouts 1996) are modelled by a PSF within GALFITPIPE, with all others being modelled using a single Sérsic function. A relatively low CLASS_STAR boundary is chosen as tests have shown that a more reasonable fit is produced when fitting a PSF rather than a Sérsic function to ambiguous objects. A graphical representation of detected galaxies, stars, weight-map areas and secondary neighbour determination for G00196053 is shown in Fig. 7.

3.6 GALFITPIPE: galaxy fitting

The actual modelling is handled by GALFITPIPE, which is a wrapper around the GALFIT image analysis software (v3.0.2; Peng et al. 2010) along with several event handlers and logical filters written in R. GALFIT is a 2D parametric galaxy-fitting algorithm written in the C language. It allows for multiple parametric functions (such as Sérsic, exponential, Ferrer, Moffat, Gaussian, etc.) to be modelled simultaneously as either multiple components of a single object, multiple objects in a single frame or combinations thereof.

GALFIT uses a Levenberg–Marquardt algorithm to fit a 2D function to 2D data, in doing so minimizing the global χ^2 until the gradient $\Delta\chi^2$ has become negligible and convergence is reached. When a global minimum is thought to be found, GALFIT introduces a 10 iteration cool-down period, sampling the parameter space around the best-fitting parameters in an attempt to overcome the problem of converging on a local rather than global minimum.

In this paper we fit each primary galaxy with a single Sérsic function containing seven free parameters: object centres x_0 and y_0; total integrated magnitude m_{tot}; effective radius along the semimajor axis r_e; Sérsic index n; ellipticity e and position angle θ. Secondaries (galaxies and stars) will also be modelled by either a Sérsic function or a scaled PSF as appropriate. The PSF contains three free parameters: x_0, y_0 and m_{tot}. For additional information on the operation of GALFIT, refer to Peng et al. (2010).

All primary inputs to GALFIT are taken from SOURCE EXTRACTOR and modified as described in Section 3.5, with the exception of Sérsic index which starts at $n_{initial} = 2.5$. After extensive testing on the r-band data, it was found that the chosen starting Sérsic index has little to no effect on the end result, and so choosing a value in the centre of the expected parameter range was deemed appropriate. No explicit constraints were put on the range of acceptable Sérsic indices upon which GALFIT may converge, however, GALFIT has internal limits of $0.05 < n < 20$, where the lower limit is a 'soft' limit (indices scatter around this value) and the upper is a 'hard' limit (indices may not converge above this value). More conservative limits were not enforced on Sérsic index so as not to lead the final results and make presumptions about Sérsic index distributions. More detail on chosen initial conditions may be found in Appendix B.

In order for GALFITPIPE to function correctly, it needs the cutout science image from CUTTERPIPE, *cutim*; the associated segmentation map and object catalogue from OBJECTPIPE, *segim* and *objct*, respectively, and a 2D FITS image PSF representing the PSF at the primary galaxy location from STARPIPE and PSFPIPE, *psfim*. Note that the weight map (*cutwt*) is no longer required at this modelling stage.

Once the aforementioned files are in place, an initial fitting region radius on the cutout is defined by

$$r_x = 2r_{Kron}\left(\left|\cos\left(\theta\right)\right| + \left(1 - e\right)\left|\sin\left(\theta\right)\right|\right), \tag{6}$$

$$r_y = 2r_{Kron}\left(\left|\sin\left(\theta\right)\right| + \left(1 - e\right)\left|\cos\left(\theta\right)\right|\right) \tag{7}$$

in order to account for the ellipticity e of the object and its position angle θ. Objects within the central $2r_x \times 2r_y$ of the fitting region will be convolved with the supplied PSF at the modelling stage. The segmentation map is modified to unmask all secondary objects in the fitting region, with the resultant map saved into a new *segfr* file.

A GALFIT *feedme* file is created containing the starting values for every object being modelled (primary and secondary) as described in Section 3.5 and above. A constraints file is used to constrain secondary objects. These objects are constrained in order to reduce the fitting time, and reduce the size of the allowed parameter space. x_0 and y_0 are constrained to ± 3 pixels of their input parameters, ellipticity is constrained to $0 < e < 0.95$ and half-light radius is constrained to $r_{e,initial}/4 < r_{e,final} < 4r_{e,initial}$. A final parameter for *sky* is added to the bottom of the GALFIT *feedme* file, fixing the value of the sky to zero counts.

GALFIT is then initialized, fixing the sky rms to that measured in OBJECTPIPE.the time taken to converge on a fit scales with the size of

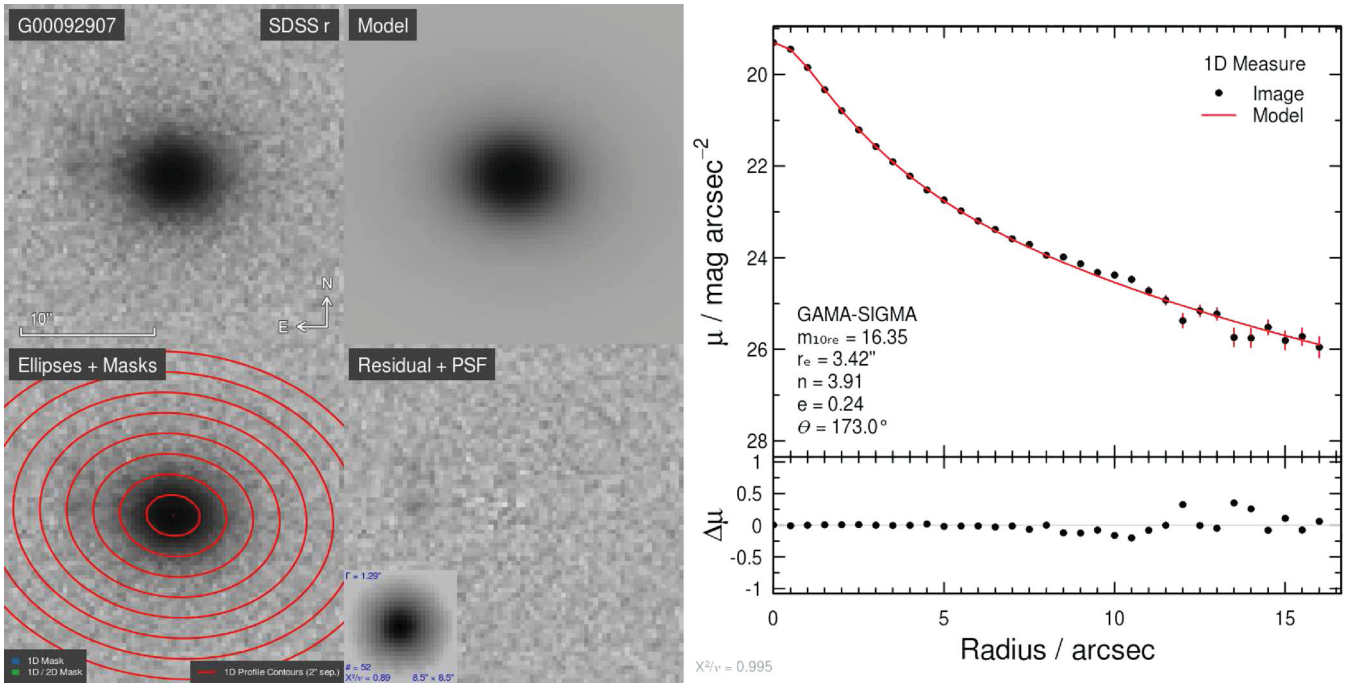

Figure 8. An example model output for GAMA galaxy G00092907. The original SDSS *r*-band image is shown in the top left. The 2D model of this galaxy and its residual (image − model) are also shown as indicated. Inset into the residual is a postage stamp of the PSF constructed for this galaxy. Blue captions within the PSF postage stamp indicate (anticlockwise from top left) the FWHM of the PSF; the number of stars used in creating the PSF; the χ^2/ν for the PSF model fit and the size of the PSF postage stamp (8.5×8.5 arcsec^2). The 1D profile for this galaxy, calculated by taking the average counts along ellipses centred on the primary galaxy and displayed against the semimajor axis, was created using the IRAF package ELLIPSE. Example ellipses from ELLIPSE have been added to the image in the bottom left to guide the eye, spaced evenly at intervals of 2 arcsec along the semimajor axis. Note that inset into the 1D profile is relevant output Sérsic modelling parameters for this galaxy, and the overall χ^2/ν for the Sérsic model is shown in grey at the bottom of the figure. The 1D profile is a 1D measure of 2D data. Any flux from secondary objects (neighbouring galaxies and stars) lying outside the 1D mask and overlapping with the primary galaxy will be counted as belonging to the primary galaxy by ELLIPSE. The 1D profile therefore should chiefly be used as a guide as to the true light distributions of both the 2D image and 2D model. For plotting, the image data are divided by some scaling constant (150 counts in the *r* band) and scaled using the arctan function with cut levels at $-(\pi/4)$ and $(\pi/2)$.

the fitting region, and the number of secondaries being fit. Once the GALFIT process has finished, its output (if any) is read and processed. GALFITPIPE scans the primary galaxy for a number of problems in this order:

(i) crash or a segmentation fault;
(ii) galaxy centre migration of $\sqrt{x^2 + y^2} > r_{\rm e,initial}$;
(iii) an exceptionally large radius of $\log_{10}(\frac{r_{\rm e,final}}{r_{\rm e,initial}}) > 3$;
(iv) an exceptionally small radius of $\log_{10}(\frac{r_{\rm e,final}}{r_{\rm e,initial}}) < 3$;
(v) a high ellipticity of $e > 0.95$.

If any of these are detected, a fix will be attempted and GALFIT re-run as appropriate. Fixes attempted vary depending on the problem encountered. If a crash or segmentation fault is detected, GALFIT will be re-run modelling only the primary galaxy, with all secondaries masked. This usually occurs for large nearby objects with a high number of secondary neighbours and foreground stars, providing the Levenberg–Marquart minimization routine in GALFIT with many local minima. If the centre migrates away to fit a secondary feature, GALFIT will be re-run with the primary centroids fixed to their starting values. Large or small radii are initially handled by suggesting a lower starting shape parameter ($n = 0.5$). This usually assists GALFIT in finding a way out of any local minima. If this attempt still provides a wildly different size to the input parameter, the size is fixed to the input and GALFIT re-run. Finally, a high ellipticity usually indicates the model has migrated away to fit flocculent secondary features. Re-running GALFIT with a starting ellipticity of $e = 0.1$, i.e. highly circular, in most cases mitigates this problem. If all fixes have been attempted and problems persist, GALFITPIPE will record GALFIT's best-guess model parameters and move on to the next object in its catalogue, with a flag updated to reflect the fitting history.

If the fit has been successful, the output multi-HDU FITS file from GALFIT is saved as *objim*, with a catalogue of final modelled secondary objects saved as *extct*. An example model output for GAMA galaxy G00092907 is shown in Fig. 8. A series of value added measurements are calculated and added to the structural measurements already taken for the primary galaxy. These include μ_0 (central surface brightness), $\mu_{\rm e}$ (surface brightness at the half-light radius), $\langle\mu_{\rm e}\rangle$ (average surface brightness within the half-light radius) and r_{90} (radius along the semimajor axis that contains 90 per cent of the total[12] Sérsic flux) amongst others. These values along with the output parameters from GALFIT and previous SIGMA modules are added to a CSV catalogue, allowing SIGMA to move to the next primary galaxy in the input catalogue.

4 SIGMA OUTPUT

The SIGMA master catalogue, *SersicCatv07*, provides measurements of Sérsic index, half-light radius, position angle, ellipticity and magnitude in addition to extra pre-modelling sky estimation, Source Extraction and PSF Extraction measurements and post-modelling

[12] Integrated to infinity.

value added measurements as detailed in Section 3. Magnitudes contained within this catalogue are defined according to the AB magnitude system, and have not been corrected for the effects of foreground Milky Way dust extinction. The catalogue is an output of the GAMA *SersicPhotometry* data management unit, and contains 527 columns of data, 58 columns per passband, and five additional common descriptive columns. Here we discuss the results of the modelling pipeline.

4.1 Sample definitions

Table 3 summarizes the various sample definitions in use throughout this paper. Our initial input is the GAMA tiling catalogue, *TilingCatv11*, which contains 169 850 sources. Of these sources, 167 600 are classed as galaxy like (as defined in Baldry et al. 2010). SIGMA was run across this galaxy-like subset independently in all nine bands, the output of which defines the SIGMA master catalogue, *SersicCatv07*, which is available via the GAMA data base. However, we define additional subsamples in order to facilitate further analysis of the data throughout the remainder of this paper. This is to ensure that selection bias does not adversely affect our conclusions. *SersicCatv07* contains sources fainter than the deepest nominal GAMA limit of $r_{\rm petro} = 19.8$, and so a cut was made limiting sources to $r_{\rm petro} < 19.8$. A common coverage sample was constructed so as not to compare galaxies between bands whose observations are incomplete or have missing data. This was defined using the SOURCE EXTRACTOR Auto magnitude SEX_MAG_X from *SersicCatv7*, a product of the OBJECTPIPE module, where $X =$ *UGRIZY JHK*. A common region is defined as having a detected SOURCE EXTRACTOR magnitude in any of the SDSS bands (*ugriz*) as well as in each of the UKIDSS bands (*YJHK*). Incompleteness mainly affects the NIR bands, with noticeable UKIDSS footprint gaps visible in the final common coverage area shown in Fig. 1. The number of detected sources in individual SDSS bands is typically very high, >97 per cent, with the exception of the *u* band. The *u*-band data have a detection percentage via this method of 50.8 per cent, indicating the poorer quality of the data in that band. For this reason, *u*-band data are excluded from further fits to the data, with relations instead extrapolating into the *u*-band wavelength for reference. The common coverage area reduces the sample to 138 269 galaxies and is used throughout Section 4 and the beginning of Section 5 (with the exception of Fig. 17). A full listing of detected and modelled sources used in each band may be found in Table 4. In any analysis that makes use of data from the *StellarMassesv03* catalogue (rest-frame *K*-corrected $u - r$ colours or stellar masses), a reduced matched coverage sample of 116 951 objects is defined.

Table 3. Table defining various sample definitions in use throughout this paper. Cuts are sequential, and include the definitions from previous rows.

Name	Number	Definition
TilingCatv11	169 850	Complete GAMA tiling catalogue
SersicCatv07	167 600	Removes star-like objects
Survey	150 633	Removes $r_{\rm petro} < 19.8$
Common	138 269	Requires SIGMA SOURCE EXTRACTOR coverage in $(ugriz) + Y + J + H + K$
Matched	116 951	Requires a match in the *StellarMassesv03* catalogue

Table 4. The number of detected and modelled galaxies in *SersicCatv07* for each band. *SersicCatv07* contains 167 600 galaxies in total.

	Detected	Modelled
u	85 138	81 120
g	165 367	165 196
r	166 506	166 384
i	166 675	166 377
z	163 902	160 684
Y	156 702	156 280
J	152 316	151 612
H	159 464	158 797
K	157 537	156 662

4.2 Analysis

4.2.1 Additional sky subtraction

As part of the cutout creation process, SIGMA uses a variable background mesh to estimate and subtract the background sky for each galaxy in each band before any other image analysis takes place. Sky correction distributions are mostly Gaussian in shape, with a small bias to recovering positive sky values most likely owing to background source contamination at the sky estimation stage. The additional correction on top of that already applied at the GAMA mosaicking stage is usually small. In the *r* band, for example, the sky correction distribution has a 3σ-clipped mean of 0.56 ADUs (∼0.01 per cent of the sky pedestal) and a standard deviation of $\sigma = 2.06$ ADUs. Longer wavelengths produce larger corrections as expected. The accuracy with which we were able to estimate the sky using this method was found to produce good quality sky estimates in an efficient and relatively fast manner.

An additional spike feature at zero counts relates to objects whose determined preferred background mesh size was larger than that already used in the creation of the GAMA mosaics. If this has occurred, SIGMA performs no sky subtraction, and returns zero counts. This feature affects 0.39 per cent of galaxies in the *r* band, and so whilst a larger mosaicking background mesh may be preferred for future surveys, it is not believed to be a major issue affecting these data.

4.2.2 Astrometry

Initial checks were made on the output astrometric accuracy of the SIGMA models. Fig. 9 shows the computed astrometric offsets between the input SDSS positions and their GALFIT modelled SIGMA positions for all bands, with each subplot representing 2×2 pixels. Generally speaking, the astrometry is in good agreement with SDSS, with the *r*-band offset $c_r = 0.010$ arcsec (0.029 pixels) and a 1σ spread of $1\sigma_r = 0.044$ arcsec (0.130 pixels). The one exception to this is the *u*-band data, showing a much larger spread in the recovered centroids owing to the poor quality and depth of the data in this band. There are however two interesting features worthy of note in this figure. First, the apparent asymmetry in the SDSS astrometry, particularly noticeable in the higher quality *r* and *i* bands. Secondly, the global systematic offsets in the NIR UKIDSS bands (*YJHK*) of approximately 0.07 arcsec.

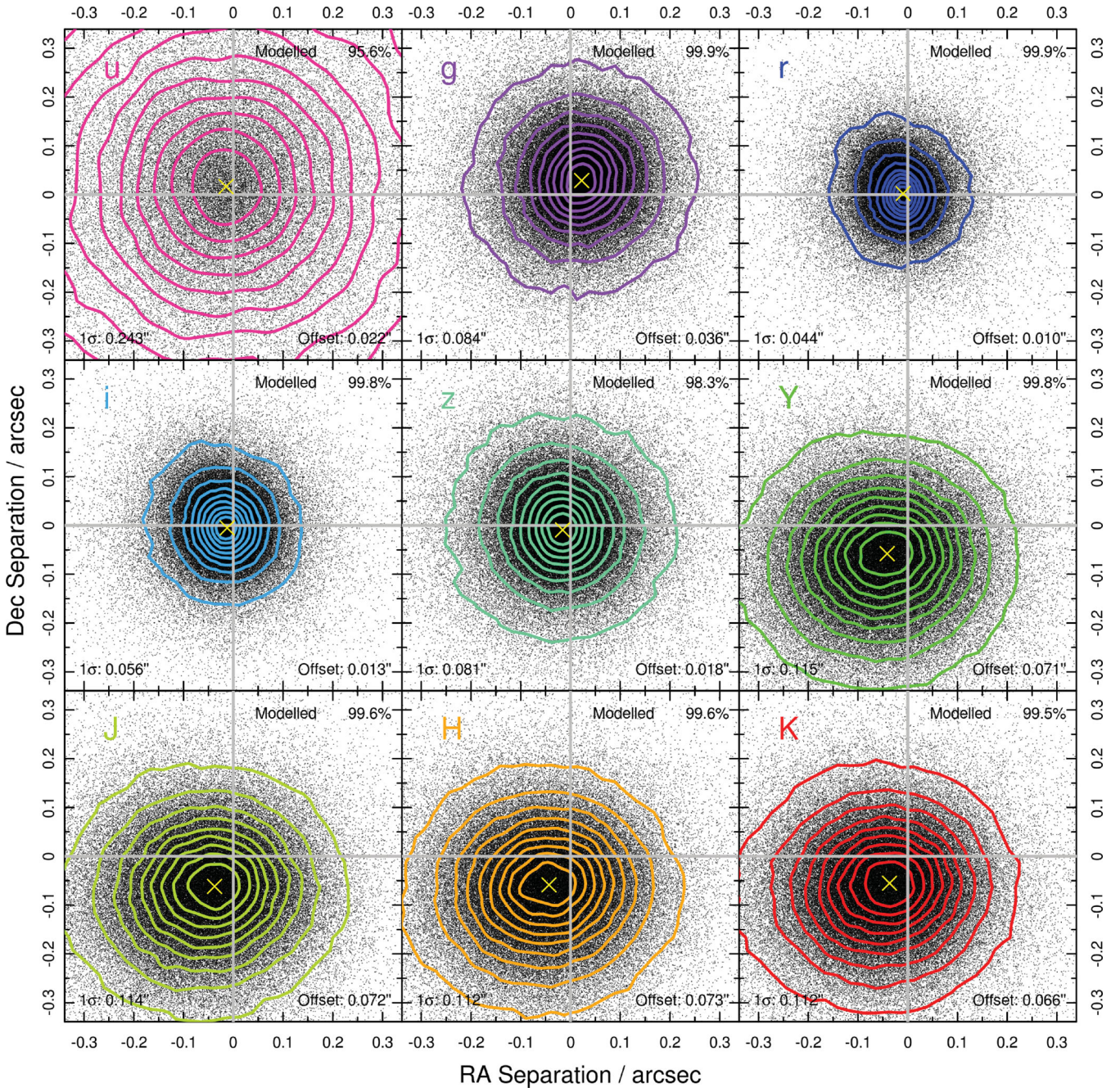

Figure 9. Astrometric offsets in RA and Dec. between the input SDSS positions and the modelled SIGMA positions for all nine bands. Contours range from the 10th to the 90th percentile in steps of 10 per cent with the peak density of each distribution represented by a yellow cross. Each subplot is exactly 2×2 pixels in dimension. The global systematic offsets in the NIR UKIDSS data (*YJHK*), typically ∼0.07 arcsec (∼0.2 pixels), are caused by minor variations in the WCS definitions between SDSS and UKIDSS. SIGMA accounts for this during the modelling phase.

The asymmetry present in the SDSS astrometric data is found to be associated with an individual SDSS strip[13] that crosses the G09 field at a large angle of incidence with respect to the equatorial plane. Galaxies whose input imaging data lie in this strip appear to have centroids scattered around $\Delta\mathrm{RA} \sim -0.05$ arcsec, $\Delta\mathrm{Dec.} \sim 0.10$ arcsec rather than the origin. This feature is less prominent in the lower quality SDSS bands as it becomes lost in the random scatter, and consequently the effect is most noticeable in the *r* and *i* SDSS bands. Since this asymmetry affects only one strip of an SDSS stripe, the error must have been introduced at the splicing stage within the SDSS pipeline. These offsets remain small however, and are not believed to seriously affect this study as they are accounted for during the SIGMA modelling pipeline.

Global systematic offsets in the NIR bands represent minor differences in the WCS calibration between the SDSS and UKIDSS data. Any discrepancy between the imaging data would be

[13] In SDSS imaging data, a single run covers a strip. Two strips constitute a stripe, with the second strip offset from the first in order to cover a continuous area.

carried through to the larger GAMA mosaics. This feature also varies according to GAMA region, with measured offsets of approximately 0.05, 0.11 and 0.09 arcsec in G09, G12 and G15, respectively. As with the previous feature, whilst consistent, offsets remain small subpixel variations (∼0.2 pixels) and therefore are not believed to be a major factor affecting cross-band matching between sources within GAMA. These features do not arise at the GALFIT modelling stage, as similar plots comparing input SDSS positions against pre-modelling SOURCE EXTRACTOR centroids from SIGMA show similar results albeit with larger spreads. On the contrary, SIGMA should do a better job of recovering true centroids due to GALFIT's model extrapolation method in estimating centroids. This makes SIGMA robust against astrometric errors such as this by recentring every galaxy at the modelling stage, emphasizing the strength of full modelling against basic source extraction.

4.2.3 *Seeing*

An independent measure of the seeing and the form of the PSF for each galaxy in each band is a necessary requirement when considering galaxy modelling. Through PSFEX in PSFPIPE, SIGMA is able to provide robust measurements of the PSF for each galaxy as described in Section 3.4 prior to the GALFIT modelling stage. Fig. 10 shows the recovered PSF FWHM Γ for every galaxy within the SIGMA common sample for all nine bands. Each density curve has a main peak in the range $0.7 < \Gamma < 1.4$ arcsec, and an additional peak at $\Gamma = 0.4$ arcsec, which shall be discussed later.

We note that on average the NIR data are of better seeing than the optical, with the former in the range $0.6 < \Gamma < 1.3$ arcsec and peaking at around $\Gamma = 0.9$ arcsec, and the latter in the range $0.8 < \Gamma < 1.7$ arcsec with variable peaks. These ranges are in good agreement with UKIDSS (K band) and SDSS (r band) seeing targets of $\Gamma_{\mathrm{UKIDSS},K} < 1.2$ arcsec and $\Gamma_{\mathrm{SDSS},r} < 1.5$ arcsec, respectively. Some of the NIR data display a secondary peak, particularly in the K band, possibly due to the use of microstepping in the taking of some of the UKIDSS data. The worst quality seeing data is in the u band, exhibiting the largest width distribution, and the highest seeing data on average. This distribution of its mean across the GAMA regions is represented in Fig. 1, with the data points coloured according to the measured seeing at that location. This figure shows significant striping in the SDSS r band due to the drift scan mode of collection, and a measure of consistency coupled with lower average values across each of the UKIDSS bands. This could cause significant problems for image analysis routines, with average seeing doubling on the scale of a few pixels. Modelling the PSF and using that model at the galaxy modelling stage, as in SIGMA, goes some way towards mitigating this issue.

An additional peak at $\Gamma = 0.4$ arcsec represents those frames where no stars were detected in order to compute the PSF in that region, and so a generic value of $\Gamma = 0.4$ arcsec is returned. Note that for the majority of bands this problem is minimal, becoming most noticeable in the lower quality u- and z-band data.

4.2.4 *Surface brightness limits*

Consideration of the surface brightness limit beyond which data becomes unreliable at the 1σ level is also important. An estimate of the surface brightness limit at any given position may be given by

$$\mu_{\mathrm{lim}} = \mathrm{ZP} - 2.5 \log I_{\mathrm{rms}}, \quad (8)$$

where ZP is the zero-point of the imaging data, and I_{rms} is the root-mean-square of the background sky per square arcsecond. Note however that this provides a worst-case scenario value to the surface brightness limit, with the real limit likely to be deeper on a per-galaxy basis dependent upon the number of pixels (n) used in constructing the 2D model at large radii from the core region, and scaling as $\sqrt{n}$. Fig. 11 shows the global distributions in μ_{lim} for the SIGMA common coverage sample across

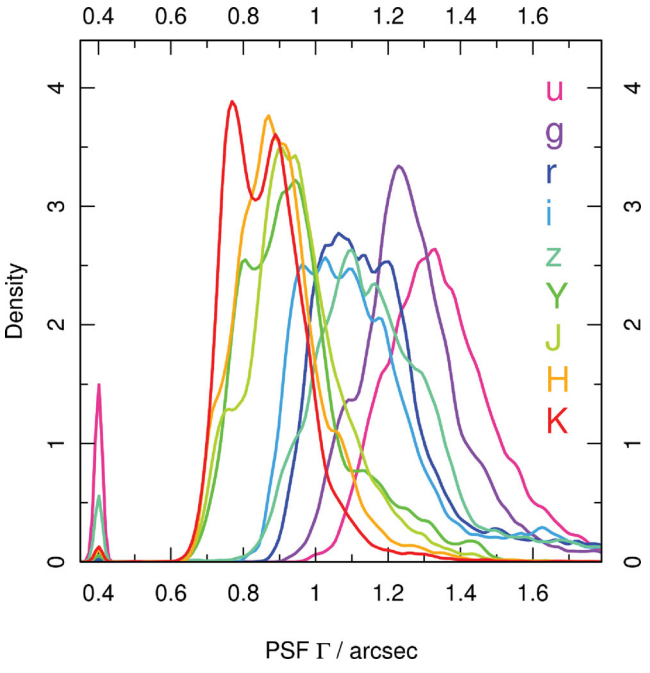

Figure 10. Recovered FWHM PSF values from the SIGMA common coverage sample.

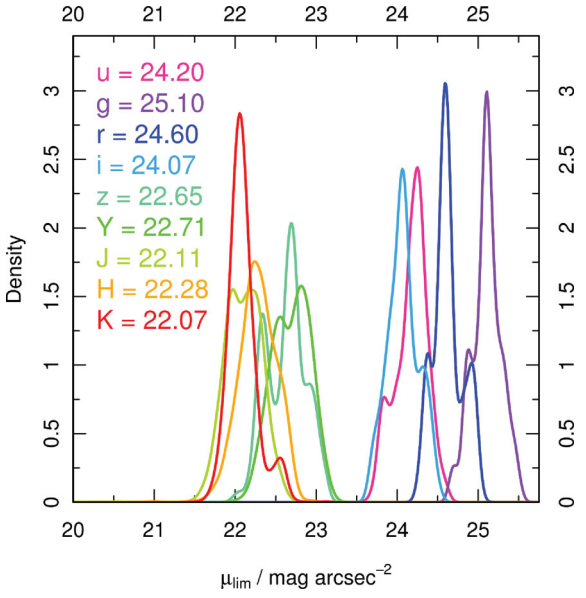

Figure 11. Apparent surface brightness limits for all galaxies within the SIGMA common coverage sample, with median surface brightness values for each band inset.

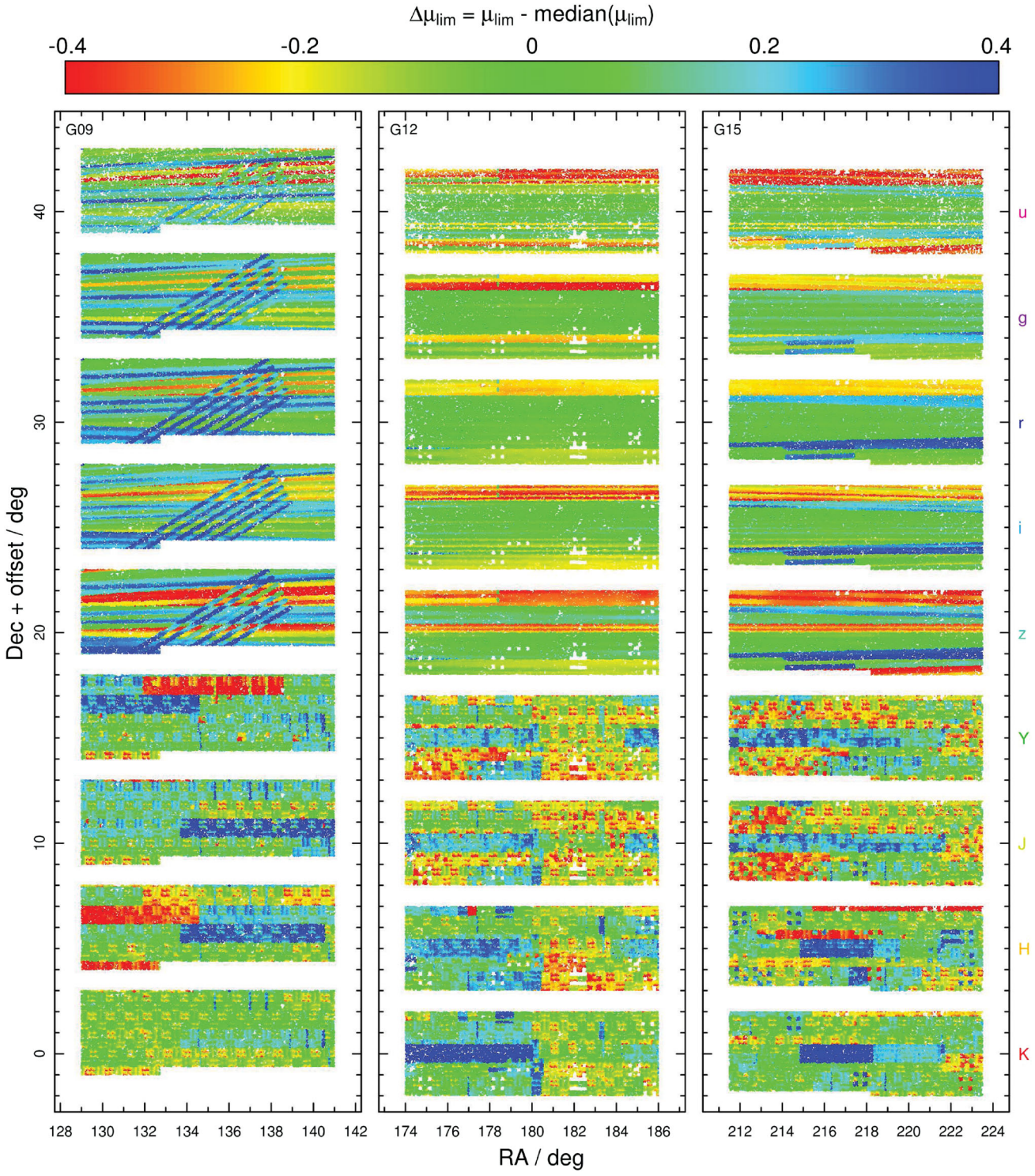

Figure 12. Apparent surface brightness limits for all galaxies within the SIGMA common coverage sample as a function of their position on the sky in RA and Dec. The three GAMA regions are displayed, as indicated, with each band labelled along the right-hand side of the figure. Bands are offset in declination in order to differentiate them from one another. The *K* band is situated at the correct GAMA coordinates. Surface brightnesses are shown as offsets from the median surface brightness for that band, the values of which are found in Fig. 11. Blue data points represent the deepest limits and red the shallowest.

each bandpass, with the median surface brightness limits in each band given inset into the figure. We note that the shorter wavelengths typically exhibit deeper limits, as expected, with a transition occurring to shallower limits beyond the *i*–*z* interface. Fig. 12 shows the spatial variation of $\mu_{\rm lim}$ across the GAMA fields. The deepest $\mu_{\rm lim}$ data are represented by blue data points, the shallowest by red. The centroid weighting mechanism employed by GALFIT should minimize the impact of a spatially varying $\mu_{\rm lim}$, and therefore should not heavily affect the output results from SIGMA.

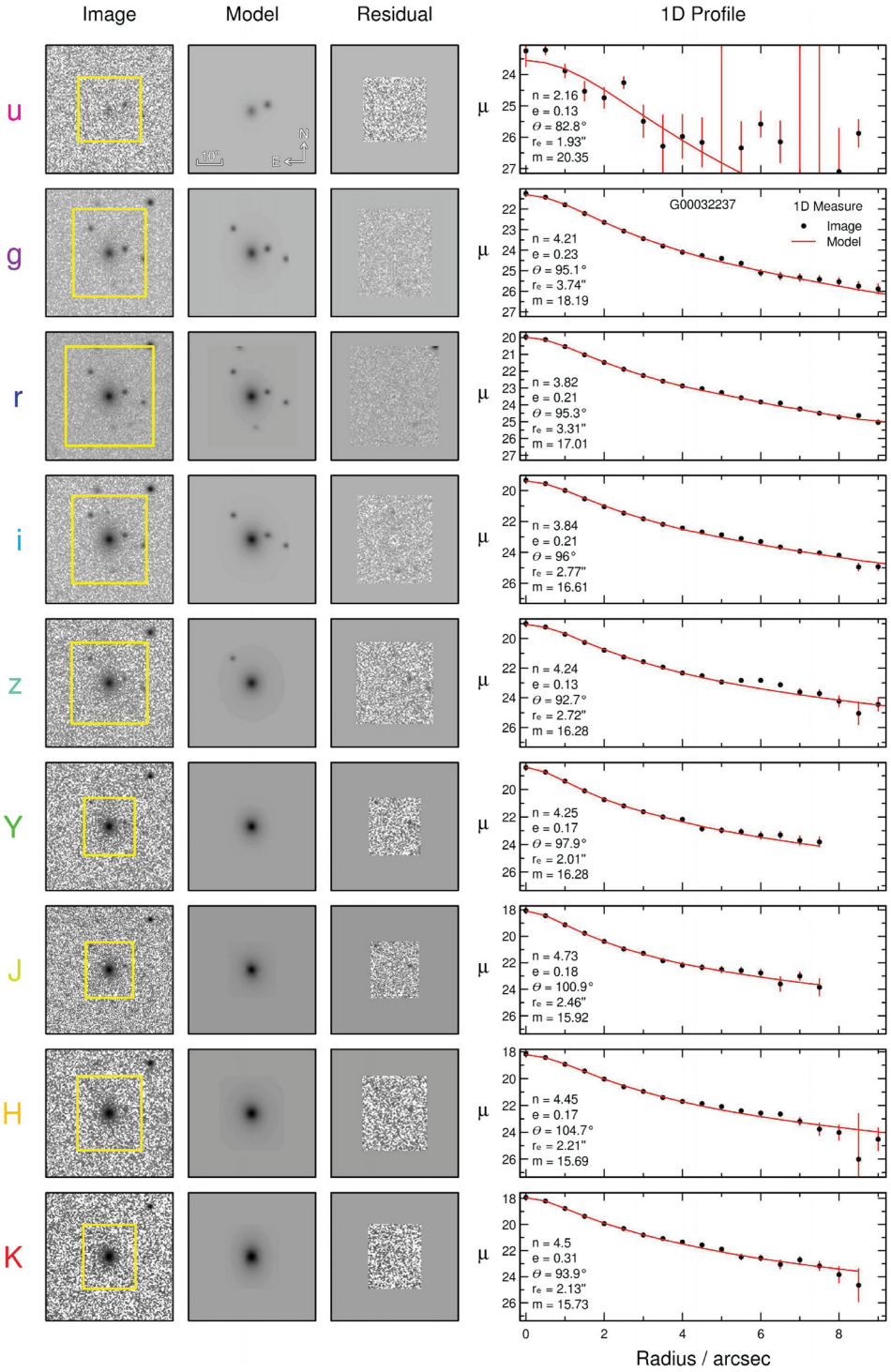

Figure 13. Model fits for G00032237 across all nine bands. Each column represents (from left to right) the original input image, the model fit to the input image, the residual image (input − model) and the 1D surface brightness profile along the semimajor axis (averaged along the annulus). The fitting region (the region within which 2D modelling takes place) is represented by a yellow box. Recovered 2D intrinsic (i.e. prior to PSF convolution) Sérsic parameters are listed inset into the 1D profile plot. The images are scaled logarithmically from $-1\sigma_{\rm sky}$ to $n\sigma_{\rm sky}$, where $\sigma_{\rm sky}$ is the typical rms of the sky in that band, and n is some scaling constant (generally, $n \sim 40$).

4.3 Results

4.3.1 Case study examples

We present example model fits for an individual galaxy across all nine bands and various galaxies separated in magnitude space in Figs 13 and 14, respectively. These figures represent the input 2D science image, model and residual along with a 1D profile radiating outwards from the core region of the galaxies along the semimajor axis. The displayed input image is a postage-stamp subregion of the background corrected cutout from the original GAMA mosaicked

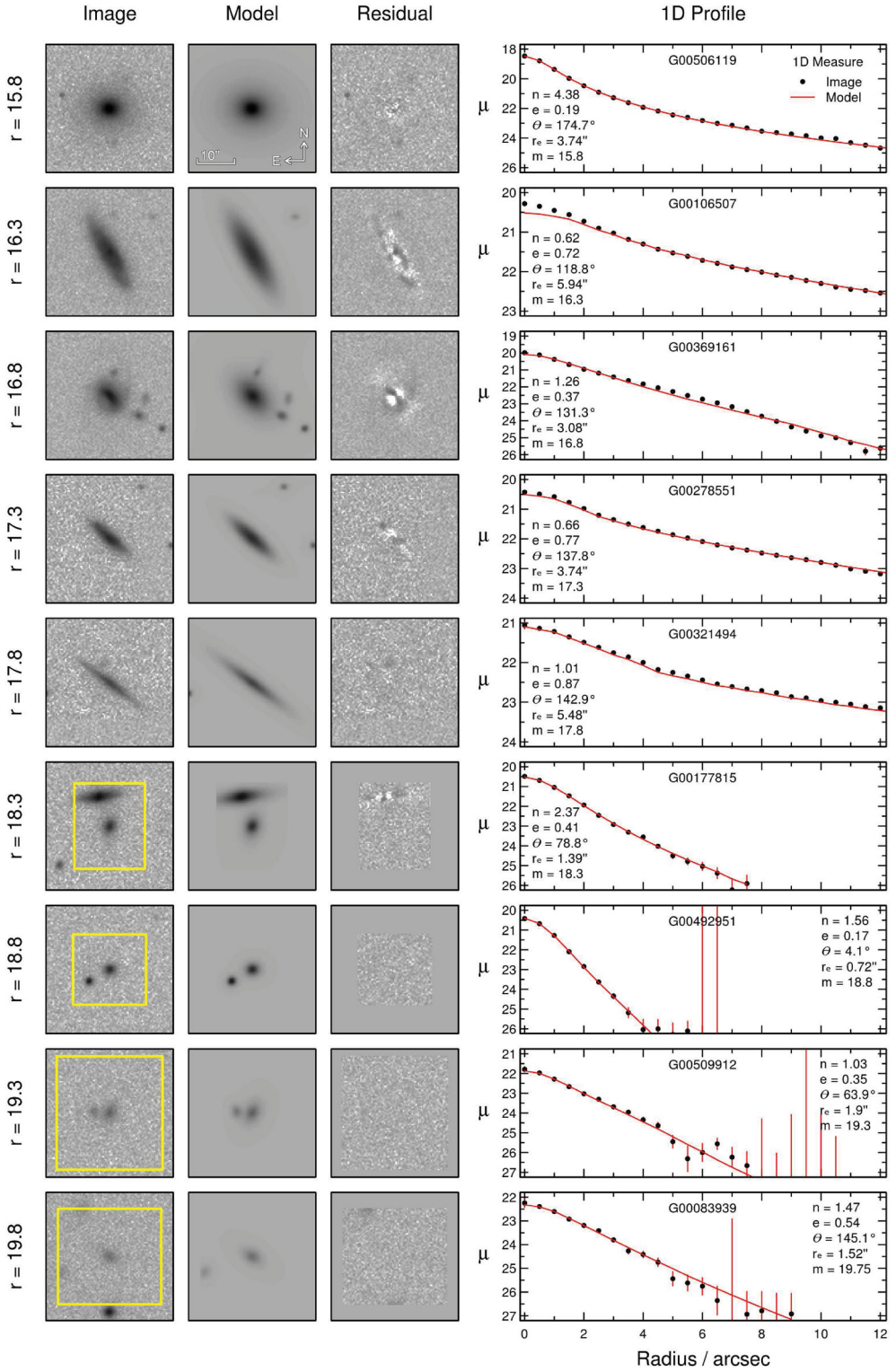

Figure 14. Model fits for nine galaxies in the *r* band separated in magnitude space by approximately $\Delta m_r = 0.5$ in the range $15.8 < m_r < 19.8$. Each column represents (from left to right) the original input image, the model fit to the input image, the residual image (input − model) and the 1D surface brightness profile along the semimajor axis (averaged along the annulus). If the fitting region (the region within which 2D modelling takes place) lies within the image thumbnail above, it is represented by a yellow box. If no yellow box is visible, the fitting region is larger than the thumbnail. Recovered 2D intrinsic (i.e. prior to PSF convolution) Sérsic parameters are listed inset into the 1D profile plot. The images are scaled logarithmically from $-1\sigma_{\rm sky}$ to $40\sigma_{\rm sky}$, where $\sigma_{\rm sky}$ is the typical rms of the sky in the *r* band.

image. The yellow box in each of the input image postage stamps represents the size of the fitting region as determined in GALFITPIPE, the dimensions of which specify the size of the output model FITS image. Recovered Sérsic parameters are listed inset into the 1D profile plot.

Fig. 13 shows the output SIGMA results for the elliptical galaxy G00032237 across all nine bands. Each image is modelled independently in each band, leading to a variable fitting region size dependent upon local conditions including object density and the physical size of the primary galaxy. The residuals in each band show the high quality of the fit for this particular galaxy, bar some minor core disturbance in the higher quality *r* and *i* bands. These bands cover wavelengths that are more sensitive to dust attenuation, in this case potentially highlighting small quantities of dust in the centre of the galaxy possibly related to a recent minor merger or some form of morphological disturbance. Dust has the effect of perturbing the light profile slightly away from that of a purely single-Sérsic object. Interestingly however, no evidence for dust lanes are evident in the lowest wavelength *u* band residual. This should not be surprising considering the lower quality data of the *u* band, hence these small perturbations would be lost in the noise of the image. Barring the *u*-band data, and despite dust attenuation, the recovered Sérsic indices remain relatively stable, ranging from $n = 4.21$ to 4.50 in *g* to *K*. Sérsic index peaks in the *J* band ($n_{\rm max} = 4.73$) and reaches a minimum in the *r* band ($n_{\rm min} = 3.82$, excluding *u*-band data). Modelled ellipticity *e* and position angle θ are also in good agreement, with recovered Sérsic magnitude evolving as expected across this wavelength range. Interestingly, the recovered half-light radii show a size–wavelength dependence, ranging from $r_{\rm e} = 3.74$ to 2.13 arcsec in *g* to *K*.

Secondary objects whose object centres lie outside the fitting region but whose flux leaks into it are masked so as not to affect the model fit. One such object can be seen in the upper right-hand corner of the fitting region in the *r* band postage stamp in Fig. 13. The GALFITPIPE module creates a bad-pixel mask using the segmentation map provided by OBJECTPIPE. Should GALFIT reach an error whilst trying to converge on a model, a potential additional fix is to mask all secondary objects in the field of view, and re-run GALFIT. This dynamic masking, whilst not the first choice for producing a model, typically allows high-quality and consistent model data to be extracted from difficult regions where otherwise none would exist.

Fig. 14 shows nine example galaxies in the *r* band from the SIGMA common coverage sample, separated in magnitude space in approximately equal SDSS *r*-band magnitude steps of $\Delta m_r = 0.5$ from the faintest GAMA limit of $m_r = 19.8$. These galaxies span a wide range of morphologies and environments, exhibiting the large variance in the input data processed by SIGMA. In each case, the residual images show the quality of the fits are relatively high, more so for obvious single-component objects such as the huge elliptical galaxy G00506119 ($m_r = 15.8$) than multicomponent objects such as the barred-spiral galaxy G00369161 ($m_r = 16.8$). In the case of the latter, despite a global single-Sérsic fit to a multicomponent object, the resultant model does a good job at describing global parameters such as magnitude. Despite a relatively disturbed fit to the secondary neighbour of G00177815 ($m_r = 18.3$), the quality of the primary galaxy model remains high. This outlines the exclusive use of secondary objects in accounting for additional flux in the wings of primary galaxies. Note that whilst the quality of model fits reflected through the residuals appears to become better at fainter magnitudes, this effect is more likely an example of non-resolved components of a galaxy.

4.3.2 *Global results*

Complete distributions for the SIGMA common coverage sample of 138 269 galaxies are shown in Fig. 15. Here we plot 1D density distributions for recovered model Sérsic indices, half-light radii and ellipticities in each band. Alongside these distributions are displayed the average model galaxies based on median values from the aforementioned parameters.

Recovered Sérsic parameters peak primarily in the range $0.2 < n < 10$, with additional peaks at $n \sim 0.05$ and $n = 20$ arising due to failed fits, discussed in more detail below. The primary range appears bimodal in nature, consisting of two approximately Gaussian-like distributions whose means are $n \sim 1$ and $\sim$3.5. These two peaks, for the most massive/brightest systems in GAMA, correspond to the two main galaxy morphologies as originally identified by Hubble, namely, late-type disc-dominated galaxies and early-type spheroid-dominated galaxies for $n = 1$ and 3.5, respectively. Interestingly, the second of these two peaks does not appear at $n = 4$, which is typically expected for a classic de Vaucouleurs profile. The relative strength of these two peaks shifts with increasing wavelength, with the stronger disc-dominated peak at $n = 1$ giving way to the spheroid-dominated $n = 3.5$ peak at wavelengths longer than the *i* band. This is believed to be an indicator of the shifting in observed stellar population with wavelength, however, see Section 5.3 for further discussion. In addition, the centroid of the $n = 1$ peak at wavelengths longer than the *i* band appears to move towards higher Sérsic index values, merging into an elongated shoulder of the relatively stable $n = 3.5$ peak. In the *K* band, the first peak appears to have a mean centred on $n \sim 1.5$. This should not be surprising, as optical bands are more likely to probe the young stellar populations in the discs of galaxies whereas the longer wavelengths pick out the older stellar populations within the core regions of a spiral galaxy or in elliptical galaxies. Dust may also be an issue at shorter wavelengths, blocking light from the core regions of galaxies and therefore biasing recovered Sérsic indices towards lower values (e.g. see Pastrav et al. 2012). It is important to note that these data are derived from the same *r*-band selected sample of galaxies observed in different wavelengths, and so these relative shifts in peak positions represent real variances in observed stellar populations, highlighting a wavelength dependence on structural measurements.

The additional peaks at $n \sim 0.05$ and $n = 20$ represent failed fits. For these galaxies, the fitting procedure drifted into an unrealistic parameter space during the downhill minimization routine employed by GALFIT. Despite attempts to force a better fit from the data within GALFITPIPE, the fits to the images of these objects remain corrupted, and are not appropriate for further analysis. Bad fits occur for many reasons. Typical reasons are overdense regions introducing too many free parameters into the minimization routine, or bad sky subtraction for that region. The corrupt peak values of $n \sim 0.05$ and $n = 20$ arise due to constraints placed by the fitting code GALFIT, and unchanged for the purposes of this study. The upper peak is a hard limit, with galaxies unable to obtain a Sérsic index beyond this value. The lower peak is a result of a consistency check within the GALFIT code that attempts to force a fit at $n > 0.05$, hence leading to a small distribution around this value. These errors that caused these additional peaks are also the cause of those found in the ellipticity distribution, discussed further below. The density of objects within these failed regions scales with wavelength, with the higher quality bands exhibiting fewer cases than poorer quality bands such as the *u* band. A conservative estimate ($n < 0.07$ or $n > 19$) places

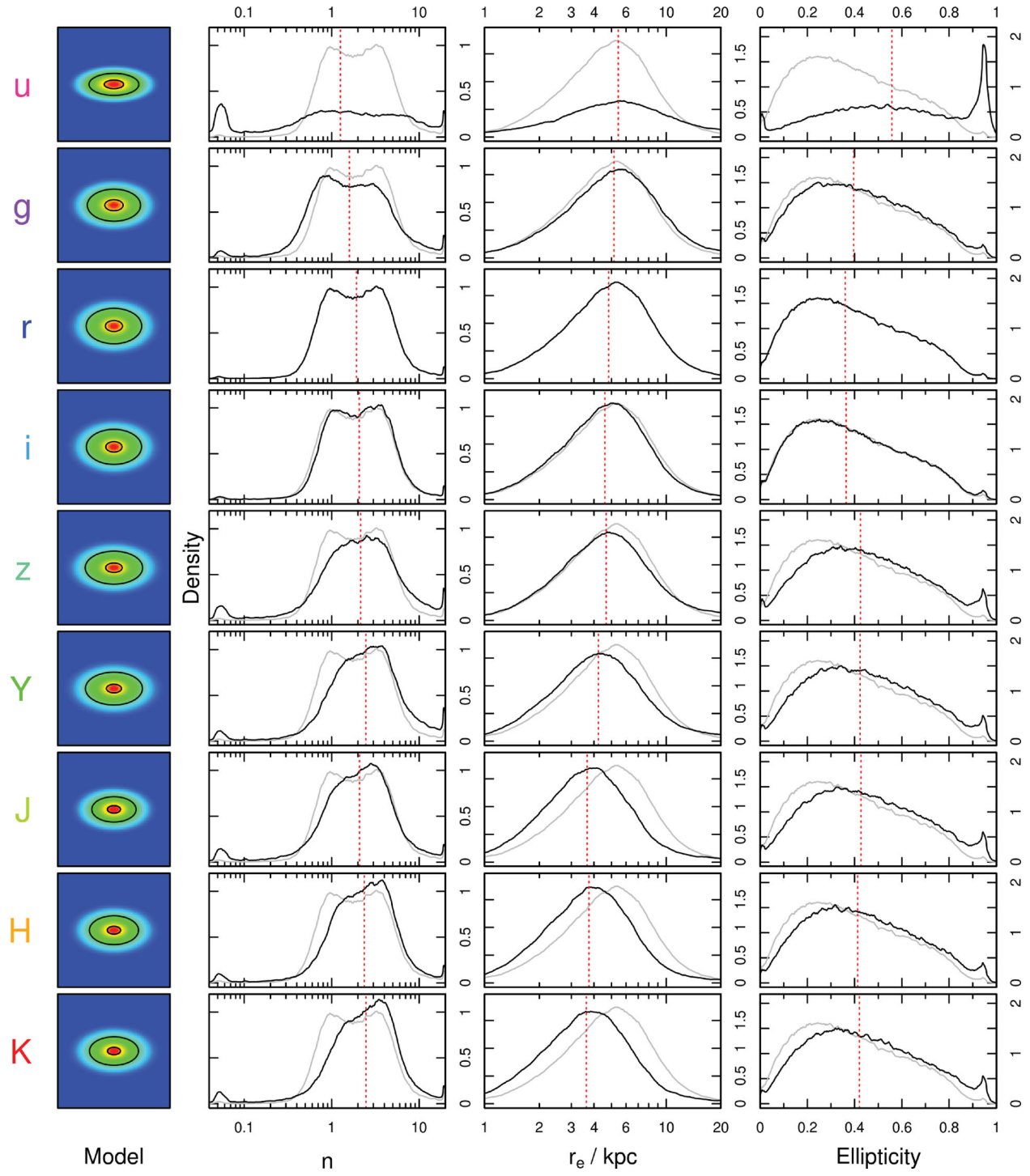

Figure 15. Global results from the SIGMA common coverage sample for all nine bands. Each column represents (from left to right) the average model galaxy based on median values for Sérsic index, half-light radius and ellipticity and the distributions for all recovered Sérsic indices, half-light radii (converted into kpc) and ellipticities. The *y*-axis for each distribution shows the probability density function convolved with a rectangular top-hat kernel with standard deviations of 0.05, 0.05 and 0.02 for index, size and ellipticity, respectively. Median values for each distribution are represented by a red dashed line and are used in creating the average model galaxy in the left-hand column. The *r*-band distributions are shown in grey for reference.

1.1 per cent (1456) of *r*-band galaxies within these extremely non-physical regions, rising to 9.1 per cent (12 630) in the worst affected *u* band.

Distributions of recovered effective radii (along the semimajor axis), converted to kpc, are also shown. Density profiles at all wavelengths appear relatively smooth, approximating a skewed Gaussian distribution. Note that these distributions exhibit no additional peaks as observed in the Sérsic index and ellipticity plots. When regarding the median values of these distributions, represented in Fig. 15 by red dashed lines, we note that the median effective radius of a galaxy ranges from 5.5 kpc in the *u* band to 3.5 kpc in the *K* band. This marked decrease in physical size with

observed wavelength is again as expected if one expects the longer wavelengths to probe core stellar populations in the bulges of galaxies, whilst shorter wavelengths probe recently formed populations in the discs of galaxies (i.e. inside-out growth; see Trujillo & Pohlen 2005). The transition wavelength appears to be the *Y* band, with a median size of ∼4.5 kpc in the *z* band, and ∼3.5 kpc in the *J* band, highlighting once more the wavelength dependence on structural measurements and the importance of the right choice of wavelength when comparing galaxy samples. Indeed, there appears to be little size variation at wavelengths longer than the *Y* band. Clearly, care must be taken when comparing the sizes of galaxies observed at different rest wavelengths.

Ellipticity $(1 - (b/a))$ measurements remain relatively consistent across all bands, peaking in the range $0.25 < e < 0.35$, and displaying additional peaks at $e = 0$ and 0.95. Bands *g*–*K*, excepting the *r* and *i* bands, show very similar distributions, with a consistent median value of $e \sim 0.4$ and a modal value of $e \sim 0.35$. The higher quality *r* and *i* bands appear to have, on average, more circular recovered ellipticities, with median and modal values of $e \sim 0.35$ and ∼0.25, respectively, however the shift is minimal. The lower quality *u*-band data lead the ellipticity measurements in that band to be biased towards higher values, caused by the fitting routine being more susceptible to background noise fluctuations and random noise in the frame. Ellipticity measurements presented here are global ellipticities, and there will be variability in ellipticity with increasing radius from the core on a per-galaxy basis caused by additional factors, for example, the effect of seeing or the presence of a bar. The additional peaks directly correspond to those already discussed previously, and represent failed fits. The values of $e = 0$ and 0.95 correspond to internal GALFIT boundaries.

4.3.3 *Sérsic magnitudes*

It is not known exactly how the light profile of a galaxy behaves at large radii away from the core regions. The exact nature of any profile will undoubtedly be influenced by many factors including, but not limited to, recent merger history, star formation rate, gas and dust content and local environment. Magnitudes within isophotal radii, not surprisingly, systematically underestimate the total galaxy light, in particular, relative to the Sérsic magnitude (e.g. Caon, Capaccioli & Rampazzo 1990; Caon, Capaccioli & D'Onofrio 1993). Graham & Driver (2005) show for example that Kron magnitudes may underestimate the total galaxy flux by as much as ∼55 per cent dependent upon choice of the multiple of Kron radii chosen to integrate out to and the profile shape of the galaxy. The comparative value for Petrosian magnitudes is considerably worse, underestimating flux by as much as ∼95 per cent in the extreme case of a high-Sérsic index object integrating out to thrice the Petrosian radius. In addition to these considerations, below $\mu_B = 27\,\mathrm{mag\,arcsec^{-2}}$ environmental effects begin to play a large role in profile shape determination.

In contrast to traditional aperture methods, studies have repeatedly shown the strength of Sérsic profiling for the majority of elliptical galaxies (e.g. Caon et al. 1993; Graham & Guzmán 2003; Trujillo et al. 2004; Ferrarese et al. 2006). Tal & van Dokkum (2011) support this viewpoint, showing the light profiles of massive ellipticals are well described by a single Sérsic component out to $\sim 8r_e$, with evidence for additional flux beyond these radii possibly related to unresolved intragroup light. With regards to disc systems, Bland-Hawthorn et al. (2005) use one of the deepest imaging studies of spiral galaxy NGC 300 to show that an exponential profile ($n = 1$) is a good descriptor of its light profile out to $\sim 17r_e$. From a sample of 90 face-on late-type galaxies (LTGs), Pohlen & Trujillo (2006) confirm the accuracy of Sérsic profiling down to $\mu = 27\,\mathrm{mag\,arcsec^{-2}}$, and suggest up to 10 per cent of their sample show evidence for a deviation from a standard $n = 1$ Sérsic fit (Type I), instead showing a broken exponential profile. These breaks appear in the form of either a *downbending* (Type II; steeper flux drop-off) or *upbending* (Type III; shallower flux drop-off) with increasing radii. Importantly, this study also suggests this observed feature is independent of local environment.

It is clear that opinion is divided amongst the community as to how a galaxy behaves below the typical limiting isophote, particularly so in the case of a disc galaxy. Each of these studies does however suggest a more complex structure at large radii than a Sérsic profile extrapolated out to infinity would imply. In order to account for the lack of reliable profile information at large radii, Sérsic magnitudes require some form of profile truncation so as not to extrapolate flux into regimes of which we know little. Two schools of thought exist in terms of appropriate truncation methods, extrapolating flux down to a fixed surface brightness limit or integrating under the profile out to a fixed multiple of the half-light radii. A constant surface brightness limit is more closely related to galaxy gas content, and so has physical meaning. However, this method introduces a redshift dependence on truncated flux, causing different fractions of light to be missed at different redshifts. Truncating at a given multiple of the effective radius assumes that the effective radius is well understood prior to truncation, which owing to the interdependency between output Sérsic parameters, is not always evident. It does not display any redshift dependence however, and is trivial to subsequently recorrect if desired. Corrections are typically minor for most galaxies, becoming most acute in high-index systems (see Fig. 2).

A sufficiently large truncation radius must be adopted to provide a close estimate of total flux without extrapolating too deep into the region of uncertainty. SDSS model magnitudes employ a smooth truncation at $3r_e$ down to zero flux at $4r_e$ for exponential ($n = 1$) profiles and $7r_e$ down to zero flux at $8r_e$ for de Vaucouleurs ($n = 4$) profiles. We adopt a sharp truncation radius of $10r_e$ for all Sérsic indices, which corresponds to an isophotal detection limit of $\mu_r \sim 30\,\mathrm{mag\,arcsec^{-2}}$, the limit to which galaxy profiles have been studied. Fig. 16 shows the magnitude offset between the Sérsic

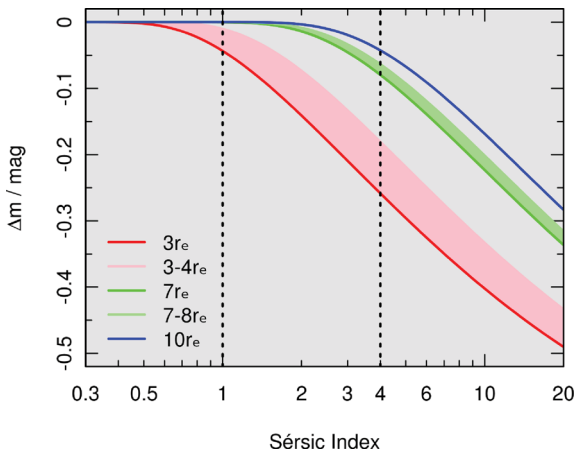

Figure 16. Magnitude offset between the Sérsic profile integrated out to infinity and that truncated at a given multiple of the effective radius. SDSS model magnitudes [forcing either exponential ($n = 1$) or de Vaucouleurs ($n = 4$) profile fits; dotted vertical lines] employ smooth tapering truncation radii, represented by the shaded red and green areas. SIGMA Sérsic magnitudes within GAMA adopt a sharp $10r_e$ truncation radius; blue line.

profile integrated to infinity and that truncated at 3, 7 and $10r_e$ as red, green and blue lines, respectively, with shaded areas representing the SDSS tapered limits. A $10r_e$ truncation gives a negligible magnitude offset for $n = 1$, and an offset of $\Delta m \sim -0.04$ for $n = 4$, with larger corrections for higher Sérsic indices. Fig. 2, middle panel, shows the flux contained within $10r_e$ (dashed vertical line) for various values of n. Given a $10r_e$ truncation, ∼100 per cent of the pre-truncation flux is retained for $n = 1$, reducing to 96.1 per cent for $n = 4$. Sérsic magnitudes truncated at $10r_e$ for each galaxy processed by SIGMA are adopted as the standard Sérsic magnitude system throughout the remainder of this paper, however, full (i.e. integrated out to infinity) Sérsic magnitudes are provided alongside truncated Sérsic magnitudes in *SersicCatv07* for reference.

Fig. 17 shows the offsets between various magnitude systems discussed previously as a function of Sérsic index. When comparing Sérsic magnitudes integrated to infinity against SDSS Petrosian magnitudes we see the two systems are in good agreement until $n_r \sim 2$, beyond which the magnitude offset relation begins to turn-off from the $\Delta m = 0$ relation. This trend argues that Sérsic magnitudes are recovering an additional ∼0.4 mag for an $n_r = 8$ object which would otherwise have been missed by traditional photometric methods. However, for the reasons previously discussed, this value should be taken as a rough estimate of the true amount of flux missed for an object of a given Sérsic index. Truncating the Sérsic index at $10r_e$ reduces the scale of this turn-off, as expected, keeping the two magnitude systems in agreement out to $n_r \sim 3$, however, still providing some measure of turn-off beyond this point. We would expect SDSS Petrosian magnitudes, or indeed any aperture-based photometry, to underestimate total flux for objects with large wings, and so this result is not surprising and a good indication that truncated Sérsic magnitudes are performing as expected. The final panel in Fig. 17 compares truncated Sérsic magnitudes against SDSS model magnitudes. The SDSS force fit either an exponential or de Vaucouleurs profile fit (marked on the figure by dashed grey lines) depending upon which profile an individual galaxy most approximates. We see clearly the inadequacy of model magnitudes when a more comprehensive Sérsic magnitude is available, with the population of galaxies split into two distinct subpopulations based upon their SDSS forced fit. For a galaxy at $n = 2$ for example, the model magnitude for the galaxy may be offset from its correct magnitude by as much as $\Delta m = \pm 0.3$ mag, with larger offsets observed for high-index galaxies. If one constructs a line of best fit for each of these two artificial subpopulations the lines pass through $\Delta m = 0$ and $n = 1$ or 4 as appropriate, confirming that SDSS and SIGMA agree for exponential and de Vaucouleurs type galaxies. As highlighted previously, the peak Sérsic index for the second subpopulation does not lie at $n = 4$ but rather at $n \sim 3.5$.

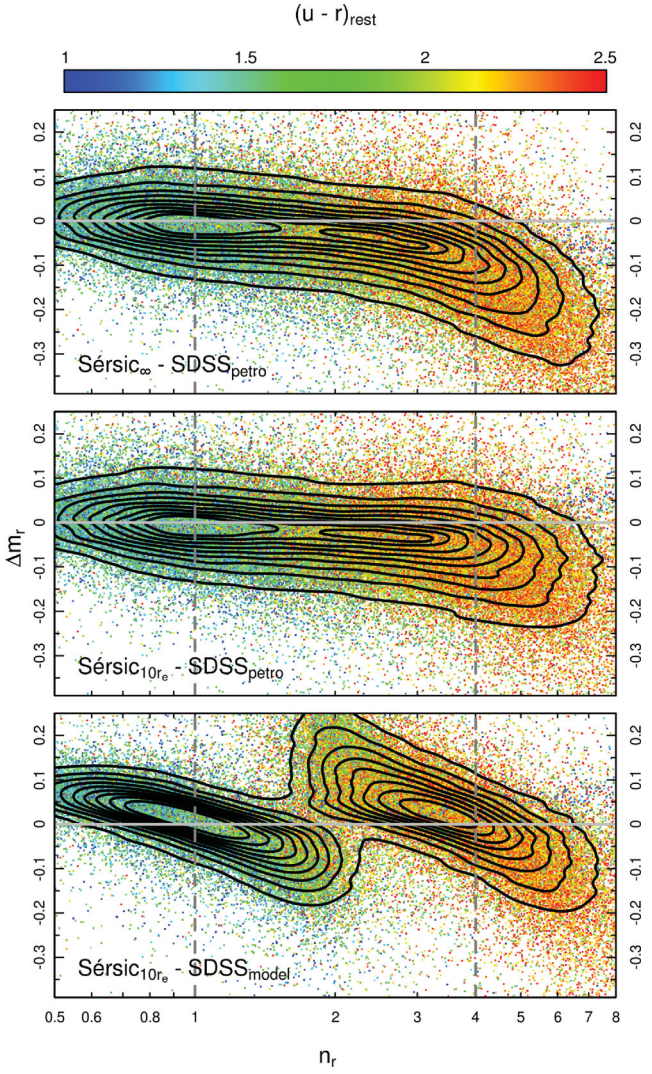

Figure 17. A series of plots displaying offsets between various magnitude systems as a function of r-band Sérsic index, with the data points coloured according to their $u - r$ rest colour, as shown. Contours range from the 10th percentile to the 90th percentile in 10 per cent steps. Top: the Sérsic profile integrated out to infinity minus the SDSS Petrosian magnitude. Middle: the Sérsic profile truncated at $10r_e$ minus the SDSS Petrosian magnitude. These figures show how Sérsic profiling is able to recover flux in the wings of galaxies that would otherwise be missed by traditional aperture-based methods, such as Petrosian apertures. Bottom: the Sérsic profile truncated at $10r_e$ minus the SDSS model magnitude. SDSS force fit either an exponential ($n = 1$) or de Vaucouleurs ($n = 4$) profile fit to attain their model magnitudes. This figure shows how model magnitudes provide an inaccurate measure of flux for a galaxy whose Sérsic index differs significantly from either $n = 1$ or $n = 4$. Vertical dashed grey lines at exponential ($n = 1$) and de Vaucouleurs ($n = 4$) Sérsic indices are added for reference.

5 VARIATIONS IN STRUCTURAL PARAMETERS WITH WAVELENGTH

5.1 Magnitude comparisons

The observed nature and form of a galaxy varies dependent upon the wavelength at which the observation is taken. These variations reflect physical mechanisms occurring within the galaxy, including but not limited to; dust attenuation and; intrinsic gradients in stellar population, age and/or metallicity (e.g. Block et al. 1999). Below, we show the variance in recovered Sérsic parameters with wavelength, and discuss how this behaviour is characterized.

Fig. 18 compares SDSS *ugriz* Petrosian photometry against truncated Sérsic magnitudes as recommended in Section 4.3.3 as a function of SDSS magnitude. Each row represents a different band, with the mode and standard deviation for varying magnitude subsets inset into the left-hand column subplots. Across each band we see a good global agreement between SDSS and recovered Sérsic photometry at all magnitudes, with the variance between the two photometric methods increasing towards fainter magnitudes as expected. The global *total* spread is larger in the lower quality u band than in the higher quality r band, ranging from $\sigma_u = 0.72$ mag in the former and $\sigma_r = 0.21$ mag in the latter. This trend should not be surprising, as lower quality data presents a unique challenge

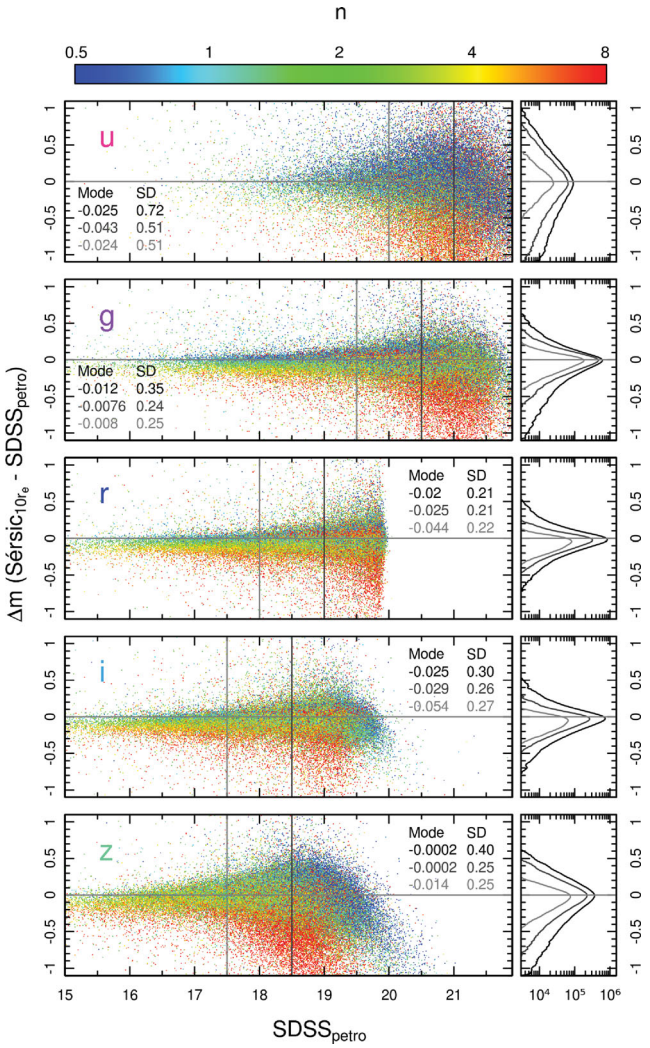

Figure 18. Comparison between Sérsic magnitudes truncated at $10r_e$ and SDSS Petrosian magnitudes for the five SDSS bandpasses as a function of SDSS Petrosian magnitude, with the data points coloured according to their Sérsic index in that band (left-hand column). Vertical lines define subsets at magnitudes brighter than those values, with corresponding statistics for mode and standard deviation inset into the figure. Density plots (right-hand column) show the relative density of objects in Δm space for each of the aforementioned subsets.

in recovering 'correct' structural parameters, with larger errors expected between different photometric systems for fainter galaxies. In all cases, the peak modal values are typically less than $\Delta m = 0.03$ mag, re-enforcing the notion of good photometric agreement between these two different methods. This is also in good agreement with the offsets previously laid out in Hill et al. (2011). The data points are coloured according to the recovered Sérsic index, and highlights the importance of Sérsic modelling in recovering accurate structural parameters. At all wavelengths, the largest offset between SDSS Petrosian and Sérsic magnitudes is observed in those well-resolved galaxies whose Sérsic indices are large ($n > 4$; red data points). These high-index systems are typified by being highly centrally concentrated with large extended wings, with the flux in the wings of these galaxies most likely to be missed by traditional photometric methods such as Petrosian aperture photometry. By extrapolating the fitted light profile, Sérsic photometry is able to recover this missing flux and provide a more accurate measure of 'total' magnitude. Unresolved compact high-index galaxies agree well with the Petrosian aperture method. As an example of flux recovery, Sérsic photometry for a relatively bright ($r = 18$) high-index[14] galaxy observed in the r band may recover as much as $\Delta r = 0.5$ mag. This missing flux recovery may be seen more clearly in the turn-off shown in Fig. 17 (middle panel), with the scale of the Petrosian correction a direct function of Sérsic index. In contrast, intermediate and low Sérsic index galaxies ($n < 4$; blue and green data points) agree much more closely with SDSS Petrosian photometry. It is interesting to note that high-index galaxies are also those whose Sérsic magnitudes have been truncated by the largest amount, and so one must take into consideration the arguments laid out in Section 4.3.3 when analysing these systems.

Fig. 19 shows a comparison between GAMA AUTO and Sérsic magnitudes as a function of GAMA AUTO magnitude. It should be noted that the GAMA AUTO photometry presented here is version 2 data and different to the version 1 data presented in Hill et al. (2011). Briefly, version 2 photometry employs an updated source detection pipeline over a larger area, with a small fraction of version 1 input frames discarded due to erroneous data in that region (e.g. badly focused frames). The process used in deriving these renewed data is similar in approach to that previously employed, a full description of which may be found in Liske et al. (in preparation). As in Fig. 18, there is good agreement between the two photometric systems, with the larger magnitude offsets observed in the resolved high-index systems.

5.2 Two distinct populations

Figs 18 and 19 show that Sérsic index plays an important role when considering magnitudes, with higher index galaxies typically recovering more missing flux than their lower index counterparts when compared to traditional photometric methods. In addition to this, there appears to be a wavelength dependence on the flux difference between high- and low-index galaxies, which has implications for variations in other structural parameters with wavelength. In order to further analyse this wavelength relationship with structural parameters, we define two galaxy subpopulations based on Sérsic index and $u - r$ rest-frame colour (AUTO aperture defined, as found in the GAMA catalogue *StellarMassesv03* as described in Taylor et al., in preparation). Fig. 20 shows this relation for the K-band Sérsic index, with the data points coloured according to stellar mass. The bulk of the galaxies appear to lie in two distinct populations, the nature of which have most recently been explored in Baldry et al. (2006), Driver et al. (2006), Allen et al. (2006), Cameron & Driver (2009), Cameron et al. (2009) and Mendez et al. (2011) amongst others, and are typically well described by two overlapping Gaussians. Blue low-index systems correspond to late-type disc-dominated galaxies and red high-index systems to early-type spheroid-dominated galaxies. This is well supported by galaxy stellar mass, with the least massive galaxies appearing disc dominated, and the most massive appearing spheroid dominated, as expected. The faintest types of galaxy, namely dwarf systems (dE, dS0), are not represented in our common coverage sample, and so these two peaks do not relate to those morphological classes. We used the positions of peak

[14] We note here that by high index we are referring to $n \sim 8$ objects, however, Caon et al. (1993) show that for a deep data set and employing good sky-subtraction methods it is possible to find galaxies whose central concentrations are of the order $n \sim 15$.

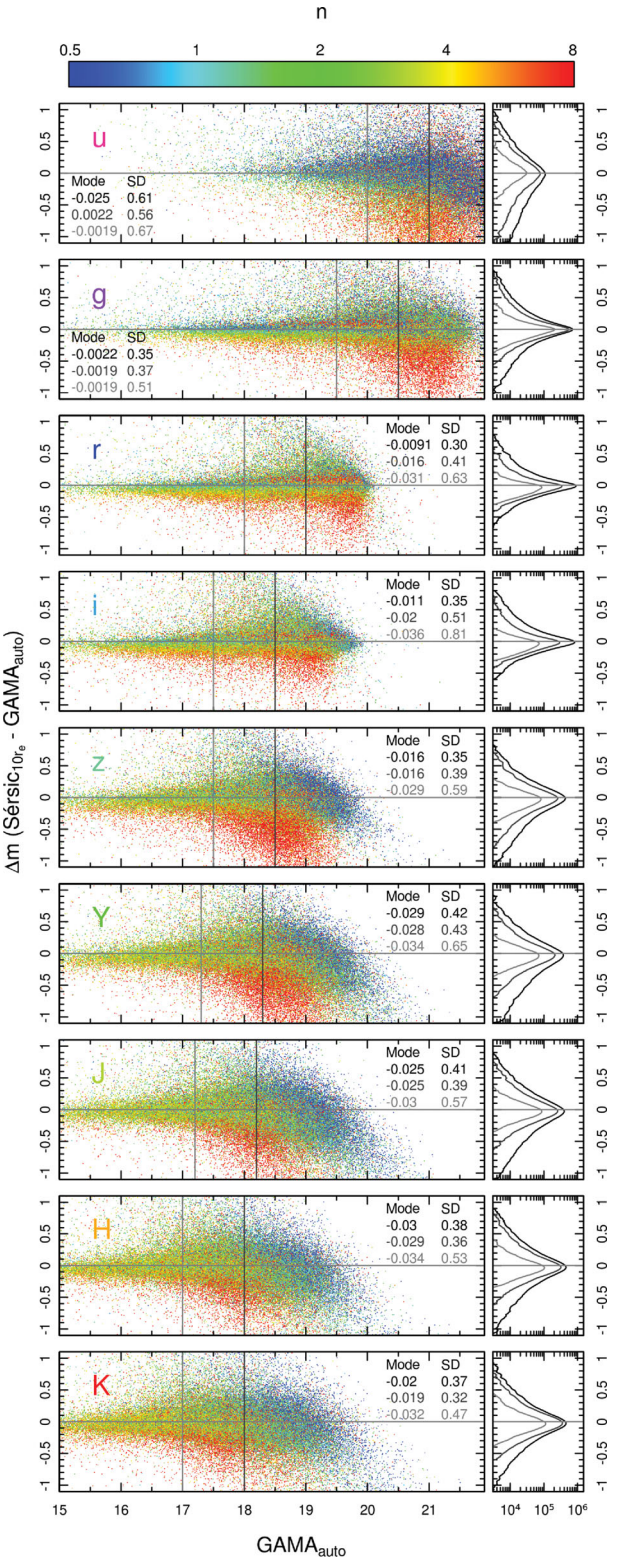

Figure 19. Comparison between Sérsic magnitudes truncated at $10r_e$ and GAMA AUTO magnitudes for all wavelengths as a function of GAMA AUTO magnitude, with the data points coloured according to their Sérsic index in that band (left-hand column). Vertical lines define subsets at magnitudes brighter than those values, with corresponding statistics for mode and standard deviation inset into the figure. Density plots (right-hand column) show the relative density of objects in Δm space for each of the aforementioned subsets.

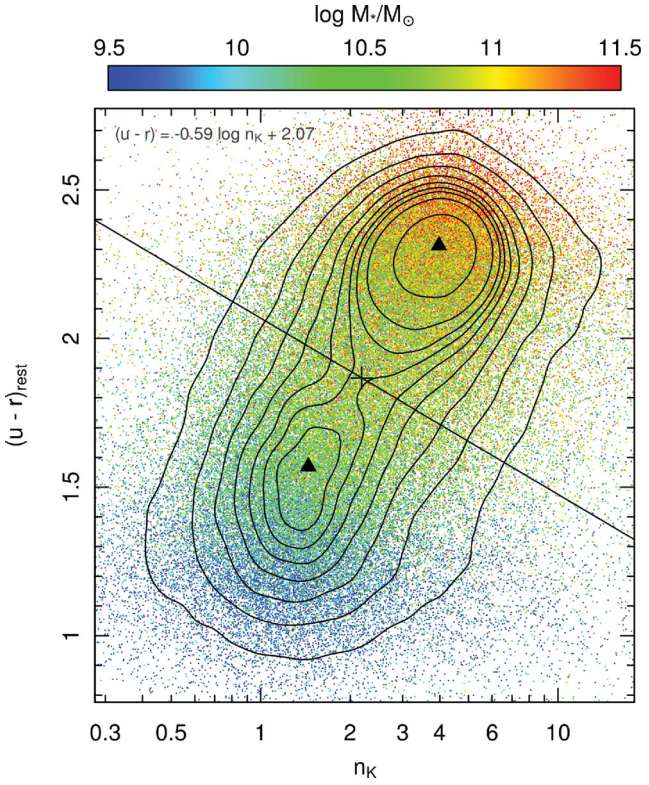

Figure 20. *K*-band Sérsic index versus $u - r$ rest-frame colour, with the data points coloured according to their galaxy stellar mass estimates, as shown. Contours range from the 10th to the 90th percentile in steps of 10 per cent. The two highest peaks in object density, corresponding to two distinct galaxy populations (late- and early type for low- and high index, respectively), are represented by filled black triangles. The diagonal line lies perpendicular to the line connecting these two peaks, and bisects it at the point of lowest object density along the connecting line, marked on the figure with a plus sign. This dividing line defines two subsamples, which for the most massive galaxies, relate to disc-dominated systems below the line (LTGs) and spheroid-dominated systems above (ETGs), the equation of which is inset into the figure.

object density for each sub-population to define a dividing line between them, specifically, the line which lies perpendicular to one connecting the two peaks in object density, bisecting it at the point of lowest object density along the connecting line. The equation of the dividing line is given by

$$(u - r)_{\rm rest} = -0.59 \log n_K + 2.07. \quad (9)$$

In order to avoid any potential misclassifications due to the effects of dust attenuation, our longest wavelength *K*-band data were chosen as a measure of central concentration. Sérsic indices recovered at shorter wavelengths return a steeper dividing line, with the gradient only becoming stable at wavelengths longer than the *z* band. This effect is characteristic of the effects of dust, and shall be explored in more depth in Section 5.3. Interestingly, Mendez et al. (2011) show that the choice of bands used to quantify colour is less important, and so a standard $u - r$ colour definition is employed in equation (9) for comparison with much of the current literature. Objects bluer than this dividing line will be referred to as disc-dominated LTGs, whereas objects redder than this line will be referred to as spheroid-dominated early-type galaxies (ETGs) throughout the remainder of this paper. It is well known that two 2D Gaussians are able to aptly describe these two populations. It

follows therefore that a harsh cut of this nature will no doubt introduce a small fraction of cross-contaminants for galaxies occupying a parameter space in close proximity to this dividing boundary, namely, those galaxies that lie in the wings of the opposing Gaussian function. The amount of contamination will be small however, with the overall trends entirely sufficient for analysing global trends within each subpopulation. Improvements to the nature of automatic morphological classification based on global structural measurements exhibited in this paper will be the focus of future studies presented in Kelvin et al. (in preparation).

5.3 Variations with wavelength

The key galaxy measurements produced by SIGMA, in addition to improved object centring accuracy are position angle, ellipticity, Sérsic magnitude, Sérsic index and half-light radius. Understanding how each of these parameters varies with wavelength is crucial to remove biases when comparing measurements made in different bandpasses. Wavelength bias may also represent real physical bias caused by dust attenuation and stellar population gradients.

Recovered position angle should show little variance with wavelength, instead varying mainly as a function of input data quality. In line with the cosmological principle, recovered position angle should merely be a random quantity assuming no detector bias, although for small area samples it may be coupled with filamentary structure. On a per-galaxy basis, one might expect minor variations with wavelength to occur in the presence of stellar population gradients in transient local features such as star-forming regions and bars, with different bands being more sensitive to different stellar populations that trace distinct structural components. However, for the SIGMA common sample of 138 269 galaxies one does indeed find no noticeable trend with wavelength for recovered position angle.

Recovered ellipticity remains relatively stable at all wavelengths, instead varying primarily as a function of the quality of the input data, as shown in Fig. 15. The highest quality *r* and *i* bands typically return the most circular galaxy models, whereas the lowest quality *u* and *z* bands return more elongated profiles across the same galaxy sample. This is as expected as one reduces the signal-to-noise ratio of the data, with the fitting routine becoming increasingly sensitive to nearby background noise, however, further studies and deeper data are required in order to comment further on this effect.

Finally, recovered Sérsic magnitude is expected to vary as a function of wavelength as per each galaxy's individual SED, the theory of which is well understood and will not be discussed further. This leaves the apparent central concentration (Sérsic index; n) and size (half-light radius; r_e) of each galaxy as a function of wavelength to be discussed.

5.3.1 Sérsic index with wavelength

Fig. 21 shows the recovered Sérsic indices for each galaxy in the SIGMA matched coverage sample at their rest-frame wavelength, coloured according to their population classification as described in Section 5.2. Considering the population definitions are based on *K*-band Sérsic indices it is reassuring to note that the spheroidal population primarily retain their high Sérsic index values across all wavelengths, and similarly for the disc population. This indicates a significant level of consistency in recovered parameters with wavelength, i.e. a galaxy that appears disc like in the *g* band is likely to appear disc like in the *H* band, for example. 3σ clipped mean Sérsic indices are shown for each population in each band, represented by large filled circles coloured as appropriate. Polynomial fits to these mean data points, excluding *u*-band values due to their lower quality imaging data, reveal general trends in Sérsic index with wavelength. The best-fitting Sérsic index for the disc-dominated population is given by

$$\log n_{\rm disc} = -0.715 \log^2 \lambda_{\rm rest} + 4.462 \log \lambda_{\rm rest} - 6.801, \quad (10)$$

and similarly for the spheroid-dominated population

$$\log n_{\rm sph} = -0.210 \log^2 \lambda_{\rm rest} + 1.394 \log \lambda_{\rm rest} - 1.753, \quad (11)$$

where $\lambda_{\rm rest}$ is the rest-frame wavelength of the observation of the galaxy. It is important to remind the reader to be mindful of our sample selection when considering these relations. Note that we have adopted log-quadratic relations for equations (10) and (11). Whilst the spheroid-dominated population may not appear to require a quadratic fit, the disc-dominated population most likely does. For this reason, the functional form of both equations has been kept the same. The linear relation for the disc-dominated population is given by

$$\log n_{\rm disc} = 0.267 \log \lambda_{\rm rest} - 0.676, \quad (12)$$

and the spheroid-dominated population is given by

$$\log n_{\rm sph} = 0.170 \log \lambda_{\rm rest} + 0.024. \quad (13)$$

These linear relations are provided for reference only and are not used in any subsequent calculations, with the log-quadratic forms instead being the preferred descriptors of the two populations.

We find that the spheroid population Sérsic indices remain relatively stable at all wavelengths, exhibiting slightly lower Sérsic indices at shorter wavelengths and becoming essentially stable at wavelengths longer than the *z*/*Y* interface. Mean ETG Sérsic indices range from $n_g = 2.79$ to $n_K = 3.63$ from *g* through *K*, an increase of 0.11 dex, equivalent to 30 per cent. This increase is consistent with the 23 per cent increase reported in La Barbera et al. (2010) over a similar wavelength range. However, whilst the fractional increase is comparable, the absolute values are not; La Barbera et al. find on average Sérsic indices $n \sim 2$–3 larger than those reported here. Whilst it is unclear what causes this difference, a potential difference in sample definitions may be important. That study defines ETGs based on a number of SDSS parameters including fracDeV$_r$; a parameter that describes how well the global light profile of the galaxy is fit by a de Vaucouleurs profile. A cut of this nature is somewhat analogous to making a Sérsic index cut alone which, as can be seen in Fig. 17, and again in Fig. 20, would introduce an element of contamination from the LTG population. If a relatively large number of the ETG sample in La Barbera et al. are in fact bulge-dominated systems with a weak underlying disc then one might expect the Sérsic indices of their bulges to differ somewhat from a traditional de Vaucouleurs profile. A Sérsic index of $n \sim 6$ is the value found in Simard et al. (2011) for bulge+disc systems with a well-defined bulge, in good agreement with the offset found here.

The apparent stability found in Sérsic index with wavelength is as expected for relatively dust-free single-component early-type systems, and is interesting to re-confirm empirically. Since the spheroid-dominated population is likely to include a small fraction of misclassified galaxies, as previously discussed, due to the nature of the harsh cut presented in Section 5.2, a small gradient with wavelength should not be surprising. Recent work by Rowlands et al. (2012) suggest that as many as 5.5 per cent of ETGs contain significant fractions of previously unaccounted-for dust, introducing an additional secondary deviation in recovered Sérsic indices with wavelength. Dust in a galaxy is typically centrally concentrated, and so has the effect of masking stellar light

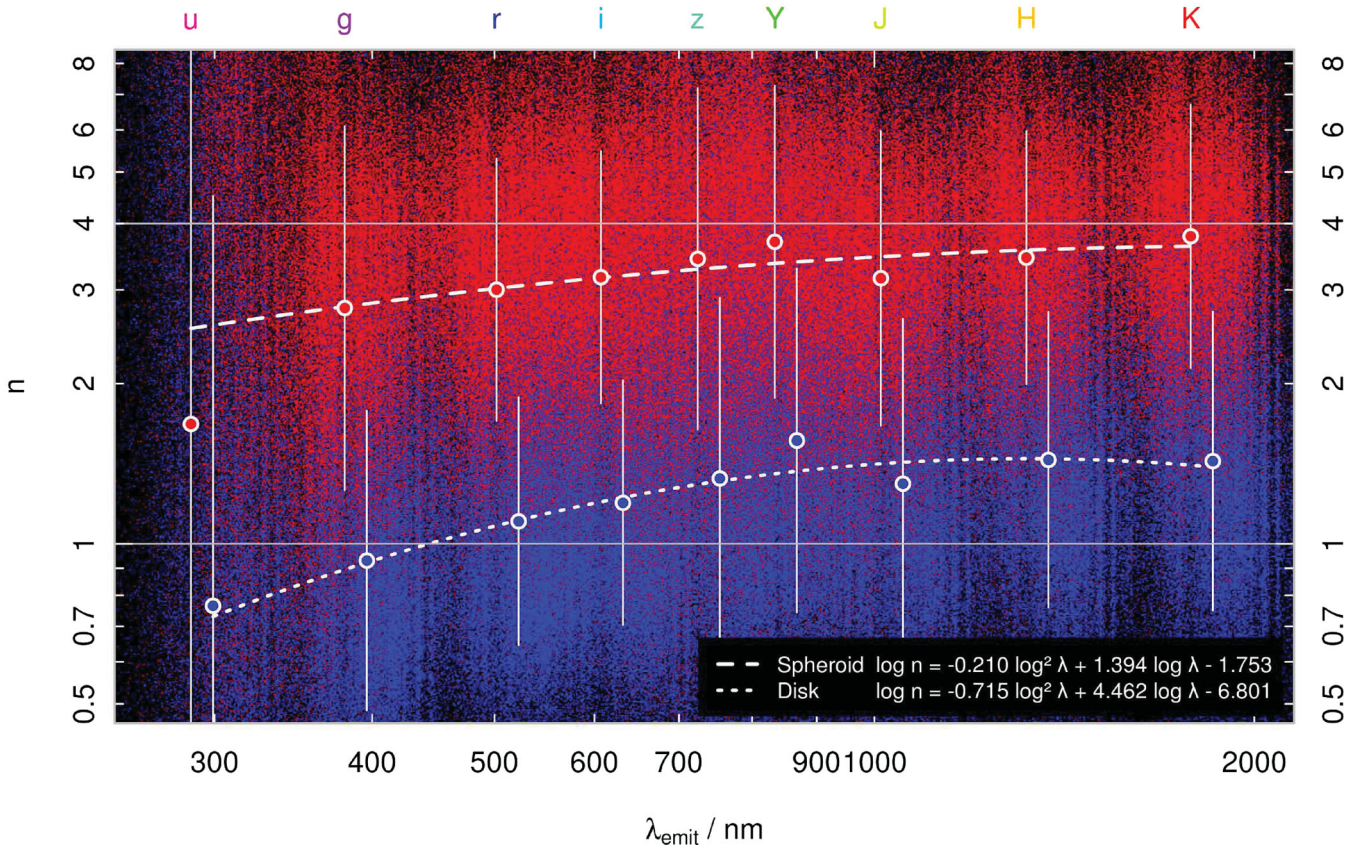

Figure 21. Recovered Sérsic index shown as a function of rest-frame wavelength in log–log space, coloured according to the population definitions described in Section 5.2. Blue data points correspond to disc-dominated galaxies whereas red data points correspond to spheroid-dominated galaxies. Large red and blue circles show the 3σ-clipped mean Sérsic indices for each respective population in each band, positioned at the median-redshift rest-frame wavelength for that population. Polynomial fits to these mean Sérsic indices are shown for both populations, the equations of which are inset into the figure. Owing to its lower quality imaging data, we exclude the *u*-band data in the calculation of these lines. Vertical lines show the 1σ spread in the data. For reference, grey horizontal lines at $n = 1$ and 4, equating to exponential and de Vaucouleurs profiles, respectively, are added.

emanating from the core regions of a galaxy. Since galaxy fitting algorithms such as GALFIT apply larger weighting to higher signal-to-noise ratio regions, minor deviations at small radii have the potential to drastically affect the recovered structural parameters, including the Sérsic index. Therefore, the addition of dust to the core region of a galaxy would subdue the *cuspiness* of the galaxy and bias the model towards a lower Sérsic index.

The disc population exhibits a larger change in Sérsic index variation with wavelength than that observed for the spheroid population. The recovered mean disc Sérsic indices range from $n_g = 0.92$ to $n_K = 1.40$ from *g* through *K*, an increase of 0.18 dex, equivalent to 52 per cent. As with the spheroid population, disc Sérsic indices also become increasingly stable at wavelengths longer than the *z*/*Y* interface. Since we typically expect disc-dominated systems to be dustier than their early-type counterparts, owing to the prevalence of ongoing star formation in many of these galaxies, then a significant variation in Sérsic index with wavelength should be expected as a consequence of the arguments previously laid out. Since the disc Sérsic index appears stable beyond the *z*/*Y* interface, we can conclude that the effect of dust in these regimes is minimal, and therefore if 'intrinsic' disc Sérsic indices are required, one should look to the longest wavelength data available, typically longwards of the *z* band. In addition to the effects of dust attenuation, we may also consider stellar population gradients. Since this sample is not a pure-disc sample, and instead contains a host of disc-dominated systems, a fraction of galaxies in the disc-dominated population will no doubt contain additional structures such as a bulge and/or a bar. Bulges tend to contain older, redder stars of a higher metallicity than the younger, bluer stars found in the discs of galaxies. Shorter wavelengths are more sensitive to the blue population found in the disc whereas longer wavelengths become increasingly sensitive to the red population. Therefore, any real colour gradients that exist in a galaxy, which are indicative of metallicity and age gradients in the underlying stellar population distribution, would also lead to a change in the measured Sérsic index, dependent upon the wavelength at which that galaxy was observed. A short wavelength is therefore more likely to probe the disc stellar population than a longer wavelength. It is unclear whether the effect of dust attenuation or stellar population gradients are the dominant factor in determining the variation in Sérsic index with wavelength, with a combination of both likely to contribute globally.

We note that the disc-dominated and spheroid-dominated populations, once stabilized, tend towards $n_{\rm disc} \rightarrow 1.4$ and $n_{\rm sph} \rightarrow 3.6$, respectively. These values differ from the Sérsic indices typically used to describe late- and early-type systems (excluding dwarf galaxies, for which there is a magnitude–Sérsic index relation), namely $n_{\rm late} = 1$ and $n_{\rm early} = 4$, respectively (represented in Fig. 21 by horizontal grey lines). This may indicate morphological contamination between populations as previously discussed, with some galaxies

exhibiting bulge-to-disc ratios away from values of either zero or unity.

5.3.2 *Half-light radius with wavelength*

Fig. 22 displays the recovered half-light radii as a function of their rest-frame wavelength. 3σ clipped means are represented by solid red and blue circles for the disc-dominated and spheroid-dominated systems, respectively, with linear fits to these data shown. The best-fitting linear relation describing the half-light radii in physical units (kpc) for the disc-dominated population is given by

$$\log r_{\rm e,disc} = -0.189 \log \lambda_{\rm rest} + 1.176, \quad (14)$$

and for spheroid-dominated systems

$$\log r_{\rm e,sph} = -0.304 \log \lambda_{\rm rest} + 1.506, \quad (15)$$

where $\lambda_{\rm rest}$ is the observed rest-frame wavelength for the galaxy. Again, it is important to remind the reader to be mindful of our sample selection when considering these relations.

Using these relations, we observe significant variation in the recovered sizes of galaxies as a function of wavelength. The disc population mean half-light radii range from $r_{{\rm e},g} = 4.84$ kpc to $r_{{\rm e},K} = 3.62$ kpc from g through K, a decrease in size of 0.13 dex, equivalent to a drop of 25 per cent. The spheroid population exhibits a larger spread from $r_{{\rm e},g} = 5.27$ kpc to $r_{{\rm e},K} = 3.29$ kpc over the same wavelength range, a decrease in size of 0.20 dex, equivalent to a drop of 38 per cent. This variation in the spheroid population size is in good agreement with studies by La Barbera et al. (2010) and Ko & Im (2005), reporting decreases of 29 and 39 per cent,[15] respectively, over a similar wavelength range. Explanations for the variation in recovered size with wavelength include dust attenuation at shorter wavelengths or metallicity gradients within the galaxy. The effects of dust on a galaxy light profile have previously been discussed. Obscuring the central region of a galaxy would shift the balance of total flux towards larger radii, artificially increasing the half-light radius. This effect is well understood for late-type systems, and indeed has been predicted in the literature notably by Evans (1994) and Cunow (2001), and more recently by Möllenhoff et al. (2006) and Graham & Worley (2008). Data from these studies are added into Fig. 22 for reference (as indicated), normalized to the disc half-light radius predicted at $\lambda_{\rm obs} = 900$ nm. Where possible, we adopt or infer a dust face-on optical depth of $\tau_B^{\rm f} = 3.8$ (central face-on B-band opacity) as recommended in Driver et al. (2007). This value is close to that found in detailed modelling of nearby galaxies by Popescu et al. (2000), Misiriotis et al. (2001), Popescu et al. (2011) and Hermelo et al. (in preparation). We also adopt an average inclination galaxy of $\cos i = 0.5$. Evans analyses face-on galaxies alone, whereas Cunow only produces detailed models for $\tau_B^{\rm f} = 3.0$ systems, and so care should be taken when comparing these data. In all cases we see an excellent agreement between the observed size–wavelength variation and that predicted by these dust model simulations, notably so when compared with the work of Möllenhoff et al., as they employ dust models that can account for both dust attenuation and emission. As a caveat, we note that the population shown in Figs 21 and 22 does not constitute a volume-limited sample, and so conclusions must remain tentative until further studies can confirm this relation. However, our initial interpretation is that dust models more than adequately account for the apparent size–wavelength relation in late-type disc-dominated galaxies.

It is interesting to note that whilst the spheroid population shows less variation in central concentration, i.e. Sérsic index, than the disc population, it exhibits a larger size variation with wavelength. Dust is not expected to be a dominant factor in the attenuation of light within these systems (although see the earlier discussion regards early-type dust fractions). However, higher optical depth values than that recommended in Driver et al. (2007) would have the effect of skewing the gradient of the dust attenuated size–wavelength relation to match the observed distribution. In addition to the possibility of age/metallicity gradients within spheroids, an alternative explanation for the apparent size variation with wavelength in early-type systems relies upon the interdependency between recovered Sérsic index and half-light radius. A change in the Sérsic index arising due to e.g. small core dust components, additional unresolved or disturbed structure in the core arising due to recent environmental interactions, the influence of an active galactic nucleus (AGN) or uncertainty in the PSF may lead to an equivalent corrective change in half-light radius. MacArthur, Courteau & Holtzman (2003) show that uncertainty on the PSF FWHM by as much as $\Gamma = 1.5 \pm 0.5$ arcsec, a range encompassing most of the optical data as shown in Fig. 10, would yield an equivalent measured size variation of 25 per cent for the worst affected compact systems. One concern might be that small numerical uncertainty in Sérsic index would yield artificial changes in recovered size. Using the Sérsic index ranges for both disc and spheroid populations shown in Fig. 21, and fixing the effective surface brightness and total magnitude in the total luminosity variant of equation (2) in Graham & Driver (2007), we would expect to see equivalent changes in Sérsic half-light radii of $\Delta r_{\rm e,disc} = 17$ per cent and $\Delta r_{\rm e,sph} = 8$ per cent for the disc and spheroid populations, respectively, i.e. far less than that seen here.

5.3.3 *Covariation of Sérsic index and half-light radius*

We have shown how the Sérsic index and half-light radius for the spheroid and disc populations vary with wavelength, however, one must not consider these variations in isolation. All of the output model parameters have a combined effect on the final light profile of a galaxy. Several of these parameters including Sérsic index, half-light radius, total magnitude and the background sky display certain levels of interdependence (e.g. Caon et al. 1993; Graham et al. 1996). For example, an overestimation of the sky level would lead to an underestimation in the total magnitude of that galaxy, and consequently Sérsic index and half-light radius also. In the case of sky however, the signal-to-noise ratio weighting employed by GALFIT somewhat ensures against sky offsets of a few counts from the true value noticeably adversely affecting the fit. Since we would not expect the systematic error in the background sky to show significant trends with wavelength, and the variation in total magnitude with wavelength is well understood, we exclude them from our investigation into the covariation of structural parameters with wavelength. We now consider the combined effect of varying the Sérsic index and half-light radius in unison, and how this impacts on the overall light profile of a test galaxy from $u \rightarrow K$.

Using equations (10), (11), (14) and (15) we generate estimates of Sérsic indices and half-light radii at equal steps in log-wavelength space for both the spheroid and disc populations. Using the Sérsic relation, and assuming a constant total magnitude for both spheroids and discs of $m_{\rm tot} = 15$, we create a series of surface brightness light

[15] A linear extrapolation of the trends reported in Ko & Im (2005) was applied in order to convert V–K offsets into the g–K wavelength range used here.

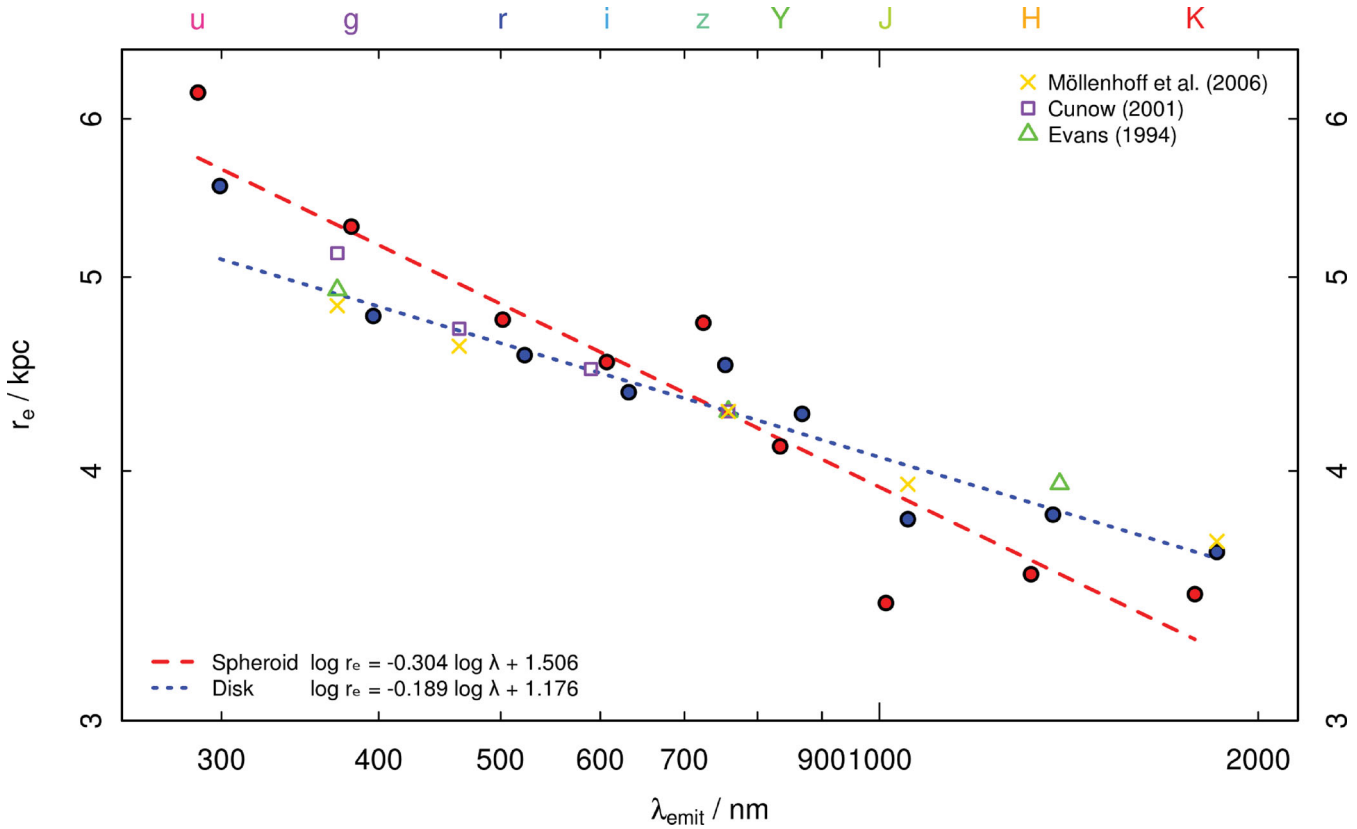

Figure 22. Recovered half-light radii in kpc as a function of rest-frame wavelength. Red and blue circles show the 3σ-clipped mean half-light radii for spheroid-dominated and disc-dominated galaxies, respectively, in each band, positioned at the median-redshift rest-frame wavelength for that population. Linear fits to these mean half-light radii are shown for both populations, the equations of which are inset into the figure. Owing to its lower quality imaging data, we exclude the *u*-band data in the calculation of these lines. Overlaid are data from several authors who predict an increase in the measured half-light radii in late-type systems due to the effects of dust. Further details are available in the text.

profiles from *u* through *K*. Fig. 23 shows the change in the recovered surface brightness light profiles over the $u \rightarrow K$ wavelength range, with the shaded areas representing the maximal area swept out by these light profiles as they vary in wavelength. This gives us an indication of how changes in recovered structural parameters affect the underlying surface brightness profile. The hatched region represents the worst-case limit at which these light profiles may be accepted as containing significant signal above the background sky level, as given in Section 4.2.4. The vertical dashed line represents a 1 pixel distance from the centre.

Despite the relatively large size variation observed in the spheroid population (a decrease of 38 per cent in $g \rightarrow K$), when considered in conjunction with the Sérsic index variation (an increase of 30 per cent in $g \rightarrow K$) the combined effect amounts to a relatively modest impact on the majority of the recovered light profile. It appears that as the spheroidal size decreases the Sérsic index increases at a comparative rate. The most noticeable surface brightness variation is found in the central core region, fluctuating by 0.49 mag at the 1 pixel boundary. Since a significant fraction of total flux lies in the core regions of high-index systems, it should not be surprising that a small variation in Sérsic index would produce a relatively large variation in half-light radius. Despite this effect, the majority of surface brightness profile out to large radii remains relatively stable with wavelength, vastly reducing the need for more complex mechanisms as previously discussed.

The variation in size for the disc-dominated population (a decrease of 25 per cent in $g \rightarrow K$) coupled with a relatively large increase in Sérsic index (an increase of 52 per cent in $g \rightarrow K$) yields a similar effect on the surface brightness profile variation as previously described for the spheroid population. Surface brightness fluctuates by ∼0.86 mag at the 1 pixel boundary, an increase of 75 per cent on the variation in the spheroid population. Here it appears that the impact of dust attenuation has a particularly distinct effect on the light profile in disc-dominated galaxies, agreeing well with the theoretical predictions for size variation with dust presented in Section 5.3.2.

Whilst no single mechanism can be shown to be entirely responsible for the relations between Sérsic index, half-light radius and wavelength observed across the two populations, it is clear that the large apparent size fluctuations in the spheroid population appear to be initially misleading. Only when considering Sérsic index in conjunction with half-light radius does the true nature of these effects come to the fore. The spheroid population, despite exhibiting large changes in half-light radius with wavelength, maintains a relatively stable surface-brightness profile from *u* through *K*. The variation in the disc population with wavelength appears well described by current dust models, however, it is most likely a combination of dust attenuation, stellar population/metallicity gradients, unresolved secondary features in the core region affecting profile fits, and uncertainty on additional parameters such as the PSF that affect the underlying physics in these systems. Future studies aim to further inform this discussion for a limited subsample to be presented in Kelvin et al. (in preparation).

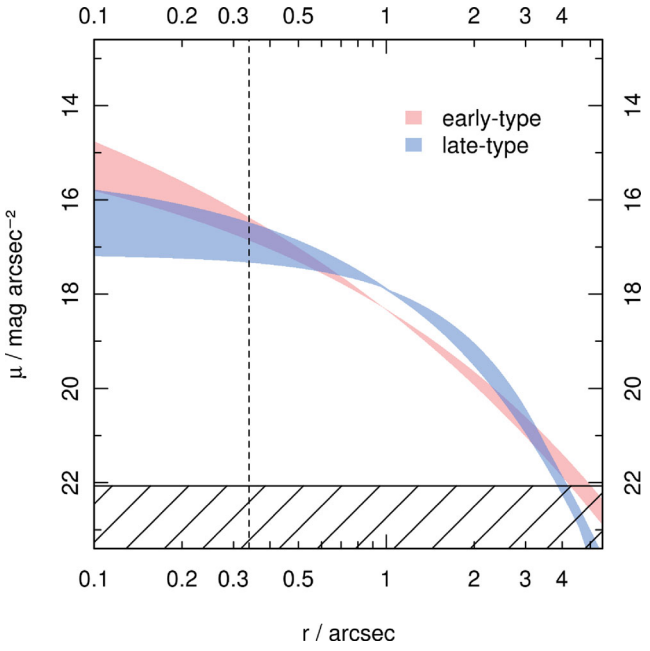

Figure 23. Surface brightness variation from u through to K for the early-type spheroid-dominated and late-type disc-dominated populations. We generate Sérsic indices and half-light radii in wavelength steps from $u \rightarrow K$ for each population using the trends as described in equations (10), (11), (14) and (15). Using these values, surface brightness profiles (without PSF convolution) may be constructed for each wavelength bin. The shaded regions shown here represent the maximal area swept out by these light profiles along the transition from u through K, and represent how much of an effect the reported changes in Sérsic index and half-light radius have on the overall light profiles. The hatched region indicates the brightest limit at which light profiles may be trusted ($\mu_{\mathrm{lim},K} = 22.07\,\mathrm{mag\,arcsec^{-2}}$), and the vertical dashed line represents 1 pixel in distance from the centre. Profiles are produced assuming a constant total magnitude of $m_{\mathrm{tot}} = 15$ for both the spheroid and disc populations.

6 CONCLUSION

We have produced high-fidelity automated two-dimensional single-Sérsic model fits to 167 600 galaxies selected from the GAMA input catalogue. These have been modelled independently across *ugrizYJHK* using reprocessed SDSS and UKIDSS-LAS imaging data. These data have subsequently been delivered to the GAMA data base in the form of the catalogue *SersicCatv07*. In order to facilitate the construction of this data set, SIGMA, an extensive multi-processor enabled galaxy modelling pipeline, was developed. SIGMA is a wrapper and handler of several contemporary astronomy software packages, employing adaptive background subtraction routines and empirical PSF generation on a per-galaxy per-band basis to tailor input data into the galaxy modelling software GALFIT 3. Output results from GALFIT are analysed for pre-determined modelling errors such as positional migration, extreme model shape and/or size parameters and adverse nearby neighbour flux. Nearby object masking is employed as a last resort, with secondary neighbours being preferentially modelled simultaneously with the primary galaxy in the first instance.

Using this data set, we have defined a common coverage area across the three GAMA regions that encompasses 138 269 galaxies, 82.5 per cent of the full sample. This common area contains only those galaxies which have been observed in all nine bands, providing a useful basis upon which to further explore wavelength trends. We define a Sérsic magnitude system that truncates Sérsic magnitudes at $10r_{\mathrm{e}}$. This ensures that flux is not extrapolated below the typical limiting isophote into regions where data quality and quantity is not sufficient to constrain the form of the galaxy light profile. Truncated Sérsic magnitudes appear to be a good descriptor of global galaxy colours and total galaxy flux. For well-resolved disc-like galaxies ($n < 2$), traditional aperture-based methods are in good agreement with truncated Sérsic magnitudes. For high centrally concentrated systems however ($n > 4$), it appears that traditional aperture based, such as Petrosian magnitudes, may miss as much as $\Delta m_r = 0.5$ mag from the total flux budget which is only recovered through Sérsic modelling.

When considering the data set in n–colour space we find galaxies appear to exist in two distinct groups. For the most massive systems, we associate these two groups with the spheroid-dominated ETG and disc-dominated LTG populations. Owing to the nature of our input sample selection, these definitions do not extend down to the fainter dwarf population, and so subsequent trends will not represent those systems. We use the longest wavelength K-band Sérsic index measurements in conjunction with rest-frame $u - r$ colour to define these two populations. Using these definitions, we are able to further probe the variations in recovered structural parameters with wavelength for each population.

We find that the Sérsic indices of ETGs remain reasonably stable at all wavelengths, increasing by 0.11 dex (+30 per cent) from g to K and becoming very stable beyond the z/Y interface. In contrast to this, we find that LTGs exhibit larger variations in Sérsic index with wavelength, increasing by 0.18 dex (+52 per cent) across the same wavelength range. Recovered sizes for both the spheroid and disc systems show a significant variation with wavelength, showing a reduction in half-light radii of 0.20 dex (−38 per cent) in ETGs and 0.13 dex (−25 per cent) in LTGs from g to K. Size variation of this scale for disc systems has been well predicted by dust models, highlighting the important role dust attenuation plays when considering structural variations across a broad wavelength range.

We note that spheroidal systems exhibit a larger size variation with wavelength than that found in disc systems. Possible physical explanations for this behaviour include low levels of unresolved dust or the effects of AGN feedback in the core of the galaxy, both of which would affect Sérsic profiling. Significant amounts of dust, such as an increased dust attenuation optical depth parameter τ_B^{f}, may allow current dust models to accurately describe the variation in half-light radii we find. It is unlikely however that a significant fraction of our spheroid-dominated population contains sufficient amounts of dust for this to be the case. Large stellar population/metallicity gradients present within individual structures of the galaxy would cause galaxies to look markedly different in different wavelengths, contributing to any concentration–wavelength/size–wavelength variation. In addition to these factors, uncertainties on the measured PSF and background sky must be considered.

However, when considering variations in half-light radius and Sérsic index together with wavelength we find that the large fluctuations in spheroidal parameters amount to a relatively modest impact on the recovered light profile. A comparatively larger effect is noted for the disc systems, particularly in the core region, supporting the presence and effect of dust attenuation in addition to stellar population/metallicity gradients. At a distance of 1 pixel from the central region, spheroid systems display a variation in surface brightness of 0.49 mag from u through K. In disc systems, the comparative figure is 0.86 mag, an increase of 75 per cent. This highlights the importance of not considering recovered parameters in isolation, as the interplay between them has the possibility of masking underlying trends.

The effects of dust attenuation appear to be the dominant factor constraining the variations in structural parameters with wavelength, notably so for the disc-dominated population. In contrast with this, apparent large structural variations in the spheroid-dominated population appear to have a relatively minor effect on the underlying surface-brightness profile than might have been expected. Future studies in Kelvin et al. (in preparation), focusing on a limited subsample of this data set, will provide a deeper understanding of these structural variations with wavelength, enabling us to comment further on the key mechanisms involved in varying structural parameters with wavelength for a host of different morphologies.

ACKNOWLEDGMENTS

We thank Emmanuel Bertin and Philippe Delorme for providing an early release version of the PSFEX software, and their many useful conversations on its correct use. We also thank the referee for his suggestions and comments which helped improve the paper. LSK thanks the Science and Technology Facilities Council and The University of Western Australia for their financial support during the writing of this article. JL acknowledges support from the Science and Technology Facilities Council (grant numbers ST/F002858/1 and ST/I000976/1). AWG acknowledges support from the Australian Research Council (grant DP110103509). BH is supported by a NPRP grant from the Qatar National Research Fund (a member of the Qatar Foundation). GAMA is a joint European–Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the SDSS and the UKIDSS. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programs including GALEX MIS, VST KIDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO and the participating institutions. The GAMA website is http://www.gama-survey.org/.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Agertz O., Teyssier R., Moore B., 2009, MNRAS, 397, L64
Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391
Allen P. D., Driver S. P., Graham A. W., Cameron E., Liske J., de Propris R., 2006, MNRAS, 371, 2
Bakos J., Trujillo I., Pohlen M., 2008, ApJ, 683, L103
Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469
Baldry I. K. et al., 2010, MNRAS, 404, 86
Bamford S. P. et al., 2009, MNRAS, 393, 1324
Barden M. et al., 2005, ApJ, 635, 959
Barker M. K., Ferguson A. M. N., Irwin M., Arimoto N., Jablonka P., 2009, AJ, 138, 1469
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, ASP Conf. Ser. Vol. 281, Astronomical Data Analysis Software and Systems XI. Astron. Soc. Pac., San Francisco, p. 228
Bland-Hawthorn J., Vlajić M., Freeman K. C., Draine B. T., 2005, ApJ, 629, 239
Block D. L., Puerari I., Frogel J. A., Eskridge P. B., Stockton A., Fuchs B., 1999, Ap&SS, 269, 5
Bush S. J., Cox T. J., Hayward C. C., Thilker D., Hernquist L., Besla G., 2010, ApJ, 713, 780
Buta R. J. et al., 2010, ApJS, 190, 147
Calabretta M. R., Greisen E. W., 2002, A&A, 395, 1077
Cameron E., Driver S. P., 2009, A&A, 493, 489
Cameron E., Driver S. P., Graham A. W., Liske J., 2009, ApJ, 699, 105
Caon N., Capaccioli M., Rampazzo R., 1990, A&AS, 86, 429
Caon N., Capaccioli M., D'Onofrio M., 1993, MNRAS, 265, 1013
Casali M. et al., 2007, A&A, 467, 777
Ciotti L., 1991, A&A, 249, 99
Cunow B., 2001, MNRAS, 323, 130
Dalcanton J. J., Spergel D. N., Summers F. J., 1997, ApJ, 482, 659
Debattista V. P., Mayer L., Carollo C. M., Moore B., Wadsley J., Quinn T., 2006, ApJ, 645, 209
de Souza R. E., Gadotti D. A., dos Anjos S., 2004, ApJS, 153, 411
Diaferio A., Kauffmann G., Balogh M. L., White S. D. M., Schade D., Ellingson E., 2001, MNRAS, 323, 999
Dressler A., 1980, ApJ, 236, 351
Driver S. P., 1999, ApJ, 526, L69
Driver S. P., Windhorst R. A., Ostrander E. J., Keel W. C., Griffiths R. E., Ratnatunga K. U., 1995a, ApJ, 449, L23
Driver S. P., Windhorst R. A., Griffiths R. E., 1995b, ApJ, 453, 48
Driver S. P., Windhorst R. A., Phillipps S., Bristow P. D., 1996, ApJ, 461, 525
Driver S. P., Couch W. J., Phillipps S., Smith R., 1998, MNRAS, 301, 357
Driver S. P., Liske J., Cross N. J. G., De Propris R., Allen P. D., 2005, MNRAS, 360, 81
Driver S. P. et al., 2006, MNRAS, 368, 414
Driver S. P., Popescu C. C., Tuffs R. J., Liske J., Graham A. W., Allen P. D., de Propris R., 2007, MNRAS, 379, 1022
Driver S. P., Popescu C. C., Tuffs R. J., Graham A. W., Liske J., Baldry I., 2008, ApJ, 678, L101
Driver S. P. et al., 2009, Astron. Geophys., 50, 050000
Driver S. P. et al., 2011, MNRAS, 413, 971
Dye S. et al., 2006, MNRAS, 372, 1227
Erwin P., Beckman J. E., Pohlen M., 2005, ApJ, 626, L81
Evans R., 1994, MNRAS, 266, 511
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Ferrarese L. et al., 2006, ApJS, 164, 334
Gadotti D. A., 2009, MNRAS, 393, 1531
Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, MNRAS, 374, 1479
Governato F. et al., 2009, MNRAS, 398, 312
Graham A. W., 2001, AJ, 121, 820
Graham A. W., 2011, ArXiv e-prints (1108.0997)
Graham A. W., Driver S. P., 2005, Publ. Astron. Soc. Australia, 22, 118
Graham A. W., Driver S. P., 2007, ApJ, 655, 77
Graham A. W., Guzmán R., 2003, AJ, 125, 2936
Graham A. W., Worley C. C., 2008, MNRAS, 388, 1708
Graham A., Lauer T. R., Colless M., Postman M., 1996, ApJ, 465, 534
Gunn J. E., Gott J. R., III, 1972, ApJ, 176, 1
Hambly N. C. et al., 2008, MNRAS, 384, 637
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Häussler B. et al., 2007, ApJS, 172, 615
Hewett P. C., Warren S. J., Leggett S. K., Hodgkin S. T., 2006, MNRAS, 367, 454
Hill D. T. et al., 2011, MNRAS, 412, 765
Hodgkin S. T., Irwin M. J., Hewett P. C., Warren S. J., 2009, MNRAS, 394, 675
Hopkins P. F., Quataert E., 2010, MNRAS, 407, 1529
Hoyos C. et al., 2011, MNRAS, 411, 2439
Huertas-Company M., Rouan D., Soucail G., Le Fèvre O., Tasca L., Contini T., 2007, A&A, 468, 937
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Ko J., Im M., 2005, J. Korean Astron. Soc., 38, 149
La Barbera F., de Carvalho R. R., de La Rosa I. G., Lopes P. A. A., Kohl-Moreira J. L., Capelato H. V., 2010, MNRAS, 408, 1313
Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692
Lawrence A. et al., 2007, MNRAS, 379, 1599
MacArthur L. A., Courteau S., Holtzman J. A., 2003, ApJ, 582, 689
McConnachie A. W. et al., 2009, Nature, 461, 66
McIntosh D. H. et al., 2005, ApJ, 632, 191

Martínez-Delgado D. et al., 2010, AJ, 140, 962
Mendez A. J., Coil A. L., Lotz J., Salim S., Moustakas J., Simard L., 2011, ApJ, 736, 110
Menéndez-Delmestre K., Sheth K., Schinnerer E., Jarrett T. H., Scoville N. Z., 2007, ApJ, 657, 790
Misiriotis A., Popescu C. C., Tuffs R., Kylafis N. D., 2001, A&A, 372, 775
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Möllenhoff C., Popescu C. C., Tuffs R. J., 2006, A&A, 456, 941
Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613
Navarro J. F., Steinmetz M., 1997, ApJ, 478, 13
Novak G. S., Faber S. M., Dekel A., 2006, ApJ, 637, 96
Park C., Gott J. R., III, Choi Y.-Y., 2008, ApJ, 674, 784
Pastrav B. A., Popescu C. C., Tuffs R. J., Sansom A., 2012, in Tuffs R. J., Popescu C. C., eds, Proc. IAU Symp. 284, The Spectral Energy Distribution of Galaxies. Cambridge Univ. Press, Cambridge, p. 3
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097
Phillipps S., Disney M., 1986, MNRAS, 221, 1039
Pierini D., Maraston C., Bender R., Witt A. N., 2004, MNRAS, 347, 1
Pignatelli E., Fasano G., Cassata P., 2006, A&A, 446, 373
Pohlen M., Trujillo I., 2006, A&A, 454, 759
Popescu C. C., Misiriotis A., Kylafis N. D., Tuffs R. J., Fischera J., 2000, A&A, 362, 138
Popescu C. C., Tuffs R. J., Dopita M. A., Fischera J., Kylafis N. D., Madore B. F., 2011, A&A, 527, A109
R Development Core Team, 2010, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
Robotham A. et al., 2010, Publ. Astron. Soc. Australia, 27, 76
Robotham A. S. G. et al., 2011, MNRAS, 416, 2640
Roškar R., Debattista V. P., Quinn T. R., Stinson G. S., Wadsley J., 2008, ApJ, 684, L79
Roškar R., Debattista V. P., Brooks A. M., Quinn T. R., Brook C. B., Governato F., Dalcanton J. J., Wadsley J., 2010, MNRAS, 408, 783
Rowlands K. et al., 2012, MNRAS, 419, 2545
Sérsic J. L., 1963, Bol. Asociacion Argentina Astron. La Plata Argentina, 6, 41
Sérsic J. L., 1968, Atlas de Galaxias Australes. Univ. Cordoba, Cordoba
Simard L. et al., 2002, ApJS, 142, 1
Simard L., Mendel J. T., Patton D. R., Ellison S. L., McConnachie A. W., 2011, ApJS, 196, 11
Stoughton C. et al., 2002, AJ, 123, 485
Tal T., van Dokkum P. G., 2011, ApJ, 731, 89
Tasca L. A. M., White S. D. M., 2011, A&A, 530, A106
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
Trujillo I., Pohlen M., 2005, ApJ, 630, L17
Trujillo I., Erwin P., Asensio Ramos A., Graham A. W., 2004, AJ, 127, 1917
Trujillo I. et al., 2006, ApJ, 650, 18
Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, MNRAS, 382, 109
Trujillo I., Martinez-Valpuesta I., Martínez-Delgado D., Peñarrubia J., Gabany R. J., Pohlen M., 2009, ApJ, 704, 618
Tuffs R. J., Popescu C. C., Völk H. J., Kylafis N. D., Dopita M. A., 2004, A&A, 419, 821
van der Wel A., 2008, ApJ, 675, L13
Vikram V., Wadadekar Y., Kembhavi A. K., Vijayagovindan G. V., 2010, MNRAS, 409, 1379
Vlajić M., Bland-Hawthorn J., Freeman K. C., 2009, ApJ, 697, 361
Vlajić M., Bland-Hawthorn J., Freeman K. C., 2011, ApJ, 732, 7
Wang J. et al., 2011, MNRAS, 412, 1081
Warren S. J. et al., 2007a, MNRAS, 375, 213
Warren S. J. et al., 2007b, ArXiv e-prints (astro-p/0703037)
Weinzirl T. et al., 2011, ApJ, 743, 87
Welikala N., Connolly A. J., Hopkins A. M., Scranton R., Conti A., 2008, ApJ, 677, 970
Welikala N., Connolly A. J., Hopkins A. M., Scranton R., 2009, ApJ, 701, 994

APPENDIX A: SIGMA INPUT OPTIONS

On starting SIGMA, a number of input options can be specified. Some of these are essential in its use, whereas others are designed for testing purposes only. The available input options can be found in the help document, reproduced below.

Listing 1: SIGMA help lists the available input options that may be specified when starting SIGMA.

```
$ sigma −h
  ----- SIGMA Version 0.9-0 -- 23 Jul 2010
DESCRIPTION
  SIGMA (Structural Investigation of Galaxies via
  Model Analysis) is a 2 dimensional fitting code
  taking inputs from the GAMA SWarped regions and
  producing models using the GALFIT software.
OPTIONS
      −a A      −append A to output log files
      −b A      −A-band (default: r)
      −c A      −input catalogue [img/csv] (needs
                 at least RA & DEC)
      −d        −show program defaults
      −e #      −error generation method
                 (1=GALFIT, 2=BOOTSTRAP)
      −h        −help (this screen)
      −i        −interactive mode
      −m        −make a plot of output .fits
                 files (png format)
      −n A      −output catalogue name
      −o        −no headers in output catalogue,
                 only data
      −p #      −number of sub-processes to spawn
      −r #      −number of bootstrap runs to
                 generate errors in GALFIT
      −s # #    −subsample, from lower to upper
                 quantile
      −t #,#    −principle allowed multi-
                 component types (eg: 1,2,5,10)
      −v        −version number
      −x #      −GAMA ID
      −y #      −SIGMA ID
      −z #      −SDSS OBJID
CONTACT
      Lee Kelvin
      University of St Andrews
      lsk9@st-andrews.ac.uk
```

In this study, we initialized SIGMA using the following command:
sigma −b x −n sigmacat_x.csv −p 16 −t 1,
where *x* represents the band for modelling (*ugrizYJHK*). This command initializes SIGMA on 16 processors, restricting the fits to single component (single-Sérsic) only. Running SIGMA across each band individually, we produced nine individual catalogues for later matching in TOPCAT (Taylor 2005).

APPENDIX B: INITIAL CONDITIONS

The galaxy modelling phase has been discussed extensively in Section 3. Most of the input parameters fed into GALFIT come directly or trivially from the pipeline OBJECTPIPE (Section 3.5), a module wrapper around SOURCE EXTRACTOR. The two main exceptions to this are half-light radius and Sérsic index.

Half-light radii from SOURCE EXTRACTOR are modified before being fed into GALFIT. This is to account for the difference in

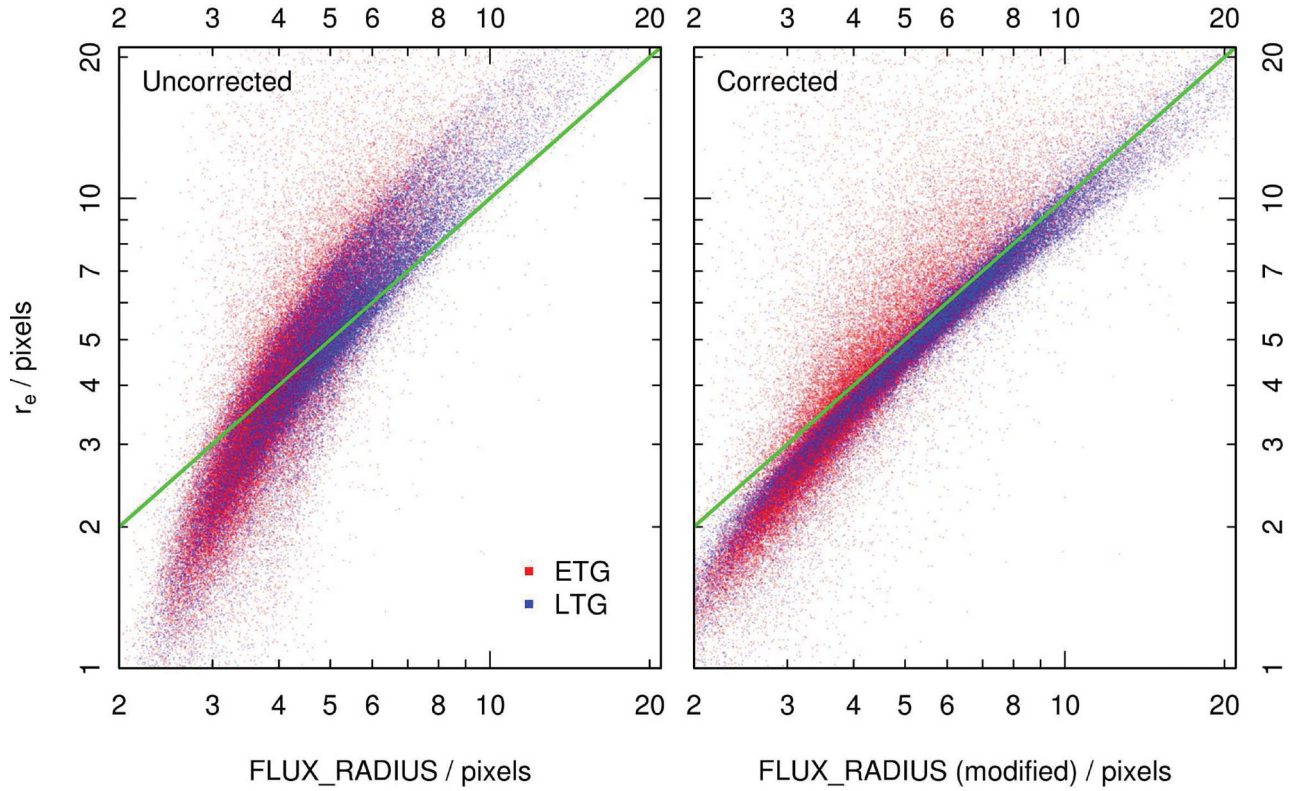

Figure B1. A comparison between SOURCE EXTRACTOR half-light radii (FLUX_RADIUS) and modelled GALFIT half-light radii (r_e) in the r band, with data points coloured according to their predicted morphological type as described in Section 5.2. Left: uncorrected half-light radii from SOURCE EXTRACTOR are a poor initial condition for modelling an object in GALFIT, as SOURCE EXTRACTOR radii are circularized and make no attempt to correct for the effect of seeing. Right: using equation (5) to correct for ellipticity and the PSF produces a much better starting value for GALFIT, hence reducing the chance of finding local minima during the minimization phase. The green line represents a 1:1 ratio, for reference.

radii definitions between the two programs. SOURCE EXTRACTOR's FLUX_RADIUS parameter outputs a circularized radius which is based on PSF convolved imaging data. The format of GALFIT's initial estimate of the half-light radius is that along the semimajor axis which is intrinsic to the object (i.e. deconvolved from the PSF). Equation (5) converts SOURCE EXTRACTOR circularized radii into semimajor intrinsic radii appropriate for GALFIT. Fig. B1 displays the before (uncorrected) and after (corrected) SOURCE EXTRACTOR half-light radii against their output modelled half-light radii from the SIGMA full sample, coloured according to their predicted morphological type as detailed in Section 5.2. Unmodified SOURCE EXTRACTOR half-light radii are a poor predictor of final modelled GALFIT half-light radii, as expected. Large galaxies ($r_e > 4$ pixels) tend to have their sizes underestimated by SOURCE EXTRACTOR by as much as 50 per cent. Following a turn off at $r_e \sim 4$ pixels, small galaxies tend to have their sizes overestimated by SOURCE EXTRACTOR. Once these data have been corrected, we find a marked increase in the agreement between the two measures, notably so for LTGs (blue data points). Data above $r_e > 4$ pixels has been reduced to minimal scatter about a 1:1 correlation. The effect of the turn-off has been significantly mitigated, yet not entirely diminished. This indicates the difficulty in accurate size estimation of galaxies that only occupy a matter of a few pixels. Correcting radii in this manner significantly reduces the chance of GALFIT finding a local minima during the minimization phase, and consequently reduces the risk of convergence on a non-physical solution.

There is no obvious proxy for Sérsic index in the default SOURCE EXTRACTOR parameters file. An approximation was created based on the trend between the output Sérsic index and the ratio between the corrected half-light radius to the Kron radius for a small test sample of trusted galaxies. From this we were able to derive a relation for a predicted variable Sérsic index:

$$n_{\rm var} = 10^{-8.6\left(\frac{r_e}{r_{\rm Kron}}\right)+2.8}, \quad \text{(B1)}$$

where r_e is the corrected SOURCE EXTRACTOR half-light radius (equation 5) and $r_{\rm Kron}$ is the SOURCE EXTRACTOR Kron radius. Fig. B2 shows the density distributions between the variable Sérsic index, $n_{\rm var}$, and several other initial conditions for a sample of 49 395 galaxies in the r band. The other input parameters (size, position angle, ellipticity, magnitude, position) are not modified. These results show that the final recovered Sérsic index is largely independent of its initial condition, with the notable exception of a bump in the distribution at $n = 0.1$ for $n_{\rm initial} = 0.1$ and a variable height spike of failed objects at $n \sim 20$. The $n = 0.1$ bump represents galaxies whose initial Sérsic index guess is placed too far away from its true value, and so fails to successfully migrate away from the initial parameter space using the Levenberg–Marquart method employed by GALFIT. It appears that the minor fluctuations found in the main body of the distributions directly correspond to the varying height of the $n \sim 20$ spike. Despite these features, it is clear that the initial Sérsic index is afforded a great deal of variability in order to achieve a successful and consistent fit. The majority of distributions presented in Fig. B2 show little variation, with similar levels of success and failure. It was therefore felt that a simple $n_{\rm initial} = 2.5$ would be an appropriate initial condition as it lies in the middle of the expected parameter space, yet not at the value of either of the bimodal peaks.

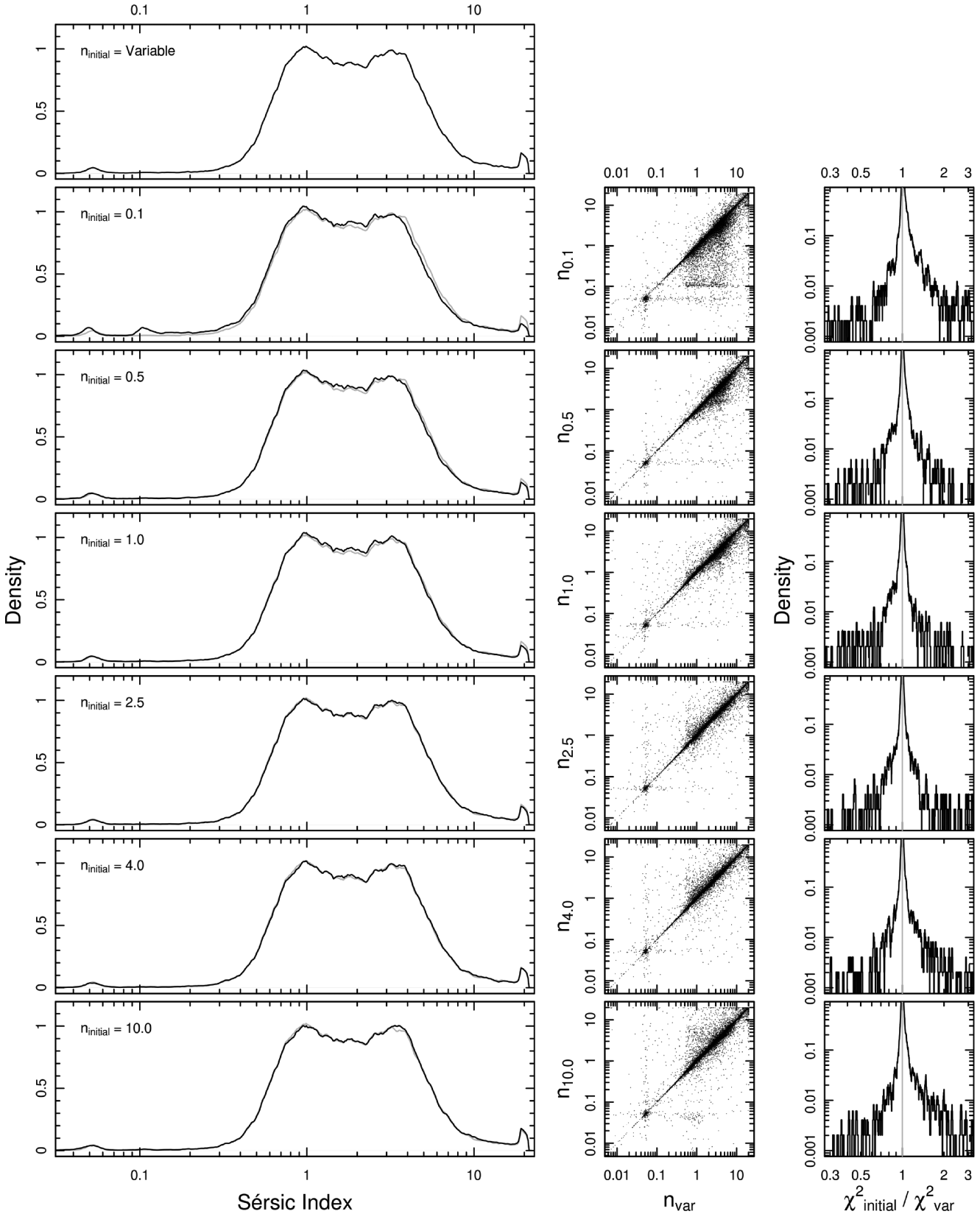

Figure B2. A plot comparing final modelled Sérsic indices for a sample of 49 395 galaxies in the *r* band given different initial Sérsic indices, as shown. From top to bottom, the initial Sérsic indices fed into GALFIT are variable (see equation B1), 0.1, 0.5, 1.0, 2.5, 4.0 and 10.0. Underlying the fixed initial Sérsic index distributions is the distribution for the variable Sérsic index coloured in grey, for reference.

This paper has been typeset from a TEX/LATEX file prepared by the author.