

# Uncertainties driven by S/N

Here we investigate contribution of the object S/N to the uncertainties in the structural parameters. In large-scale HST imaging surveys, it has been well established that the parameter uncertainties depends to first order on the object S/N (e.g., CANDELS, van der Wel et al., 2012; COSMOS-DASH, Cutler et al., 2022), which is in turn a function of magnitude or surface brightness. This conclusion has been assumed even in cluster fields by various studies, e.g., Coma cluster (Hoyos et al., 2011; Alabi et al., 2020), Fornax cluster (Venhola et al., 2017, 2018), Abell 901/2 supercluster (Gray et al., 2009), and others. The first two approaches, however, are entirely based on the data itself, thus they are not straightforward to infer systematic uncertainties as the real parameters of a given galaxy are unknown. In assessing the fitting accuracy of the structural parameters, simulations have the benefit that one knows exactly a priori knowledge of the input galaxy parameters, which allows us to directly compare with the recovered (output) parameters. The galaxy simulations are therefore a valuable technique to handle both systematic and random uncertainties resulting from profile fitting. In this section, we use simulated galaxies with known light distributions to estimate more reliable error bars for the individual parameters, by comparing the intrinsic values with the values recovered by parametric fitting. Similar simulations have been commonly applied on uncrowded galaxy fields in either HST or ground-based surveys, to determine the uncertainties on fitted single Sérsic model parameters (e.g., Häußler et al., 2007; Gray et al., 2009; Hoyos et al., 2011; van der Wel et al., 2012; Häußler et al., 2013; Delaye et al., 2014; Venhola et al., 2017; Zaritsky et al., 2021; La Marca et al., 2022). Such simulation serves three main purposes. The first is to estimate systematic and random uncertainties of the structural parameters produced by our fitting pipeline. The second purpose is to quantify the limiting magnitude (limiting surface brightness or minimum S/N) below which it is unable to recover reliable structural parameters. The third purpose is to estimate the detection limits and the completeness.

In our approach, the mock galaxies are created using GalfitM as single-component Sérsic profiles. Multi-component galaxies are not included. The galaxy input parameters to be simulated were derived by using the fitting results of real data, in order to be as close as possible to the real galaxy sample. The input parameter ranges are chosen to span the same range in parameter space as real galaxies, except for magnitude whose range has been extended 2 mag fainter than those found in the real distribution to be able to test the fitting quality and derive the survey completeness, i.e.,  $18 \leq \text{mag} \leq 31$ . To generate parameters for given mock galaxy, following similar procedures in Häußler et al. (2013) we divide the  $\text{mag-log}(r_e)$  plane of real galaxy sample into bins of width 0.4 in mag and 0.1 in  $\log(r_e)$ . We randomly select an  $\text{mag-log}(r_e)$  bin uniformly in the plane, and use the bin center as

the starting magnitude. Next, we apply an offset drawn from a Gaussian distribution with standard deviation of 0.1 mag. We then randomly pick a real galaxy within that bin, from which we obtain realistic values of half-light radius, Sérsic index, and axis ratio. These three values are also modified by multiplying factors drawn from a Gaussian distribution with mean 1 and standard deviation of 0.1. The position angle is randomly chosen uniformly from the range  $-90^\circ$  to  $90^\circ$ . Our mock catalog consists of 210,000 objects.

The mock catalog, as well as the existing PSF extracted from the real image, are provided to GalfitM to create mock galaxy images, and appropriate Poisson noise (shot noise) is added into each pixels to match the noise characteristics of the real frame. Then the resulting mock galaxies are embedded into simulated frame that shows the same noise properties as the real imaging. To build such frame, we choose multiple empty areas of same size from the science image (i.e., without objects detected by SExtractor) and paste them together to create a larger empty frame, which could accurately account for real background fluctuations, read noise (Gaussian), and correlated pixel-to-pixel patterns (induced by the drizzling process) in observed real data. Since in this simulation we intend to estimate parameter uncertainties that are limited by the S/N, rather than flux contaminations by the neighbors, the mock galaxies are inserted (into simulated frame) in a more controlled and conservative way to isolate the crowding effects. Instead of inserting a sample of mocks at random positions or evenly spaced in a grid pattern over the simulated empty frame, in each time we only place one mock galaxy in the center of the empty frame with a  $401 \times 401$  pixels area. In other words, we construct an image cutout centered on the centroid of each target galaxy ( $x_0, y_0 = 201, 201$  pixels), with no neighboring galaxies inside the cutout area. This ensures that each target is not contaminated by the stray flux from neighbors. The image cutout size is sufficiently large to bracket the most extended mock profile, but not too large to require a substantial amount of computing time.

We then run GalfitM on the whole pool of mock image created, using exactly the same procedures and setup as used for real data. Note that each image cutout contains only one galaxy. This means that the fits are simply carried out one by one. We remove the objects with any fitted parameters running into a fitting constraint from further analysis, because such fit is unlikely to produce a valid fitting result, and also has the danger of biasing the results of the uncertainty estimations. After cleaning the catalog, there are  $\sim 160,000$  objects left, and then we compare the fitted values with input values of the same objects, allowing us to derive systematic bias and random uncertainties. Figure 1 gives the differences between the output and input parameters (output minus input) for the four key parameters of Sérsic model as a function of MEASURED (OUTPUT) mean effective surface brightness  $\langle \mu_{\text{eff}} \rangle$ , given by

$$\langle \mu_{\text{eff}} \rangle = \text{mag} + 2.5 \log[2(b/a)\pi r_e^2] \quad (1)$$

where mag is the (output) total magnitude,  $b/a$  is the axis ratio, and  $r_e$  is the half-light radius of the object in arcseconds. Hence,  $\langle \mu_{\text{eff}} \rangle$  is defined as the mean surface brightness within effective (half-light) radius. We adopt this quantity as a proxy for object S/N in this study. From top to bottom in Figure 1, the panels present the differences of mag,  $r_e$ ,  $n$ , and  $b/a$  between the input and measured values for the whole pool of the cleaned mock catalog, which are shown as purple points. The distributions are color-coded according to the number density of the underlying points. The red solid lines indicate the running medians

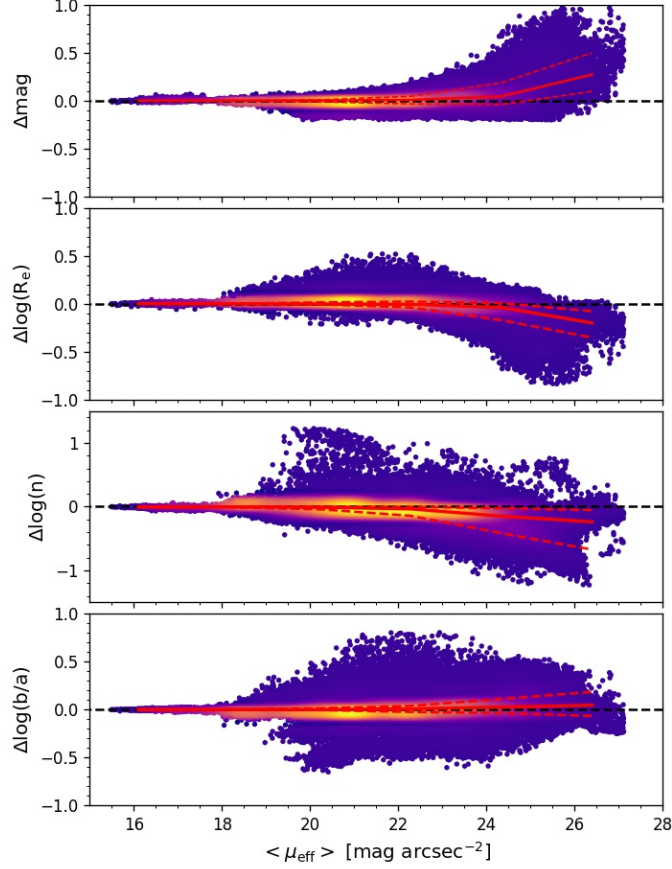
of the differences between the input and output values, which should be identical to zero for a perfect fit. The two dashed lines in each panel delimit the 16th and 84th percentiles ( $1\sigma$ ) of the value differences (i.e., the widths of the distributions). These median and associated  $1\sigma$  values of the differences are calculated by dividing  $\langle\mu_{\text{eff}}\rangle$  into equally spaced bins.

As one might expect, in low S/N regimes both the systematic errors (or bias) and random errors (scatter around the median) on the recovered structural parameters increase. At fainter  $\langle\mu_{\text{eff}}\rangle$  ends, the recovered total luminosities,  $r_e$ , and  $n$  are systematically underestimated, primarily because the outer profiles of low surface brightness objects are below the sky brightness. Slight systematic trend for  $b/a$  also exists, but is almost negligible compared to the random uncertainties. Note that the difference of  $r_e$ ,  $n$ , and  $b/a$ , are displayed in logarithmic scale to facilitate the visualization. We fit exponential functions to these systematic trends, and linear functions to the trends of associated  $1\sigma$  deviations (see below). These relations allows us to assess the uncertainties of the output parameters of real galaxies. For example, the best-fitting bias relation enables us to correct the structural parameter measurement for the real individual galaxies, while the  $1\sigma$  relation can be used to assign the random uncertainties.

In order to assign more realistic magnitude uncertainties to the GalfitM measurements, we further divide the mock sample into 4 bins (subsamples) based on the range of output Sérsic index  $n$  that are labelled above the top row of panels, as indicated in Figure 2 ( $n \leq 1.5$ ,  $1.5 < n < 2.5$ ,  $2.5 \leq n \leq 3.5$ , and  $3.5 \leq n \leq 9.0$ , respectively). For each bin (i.e., column of rows) we show the parameter differences as a function of output  $\langle\mu_{\text{eff}}\rangle$ . Similar to Figure 1, in each panel we also overplot the systematic and  $1\sigma$  scatter trends in red solid and dashed lines, respectively. These  $1\sigma$  lines will be used to estimate the random uncertainties of model parameters. Dividing the mock sample into different bins of  $n$  is motivated by the assumption that the galaxies in similar parameter space should, in principle, have comparable measured uncertainties in structural parameters (van der Wel et al., 2012; Dimauro et al., 2018; Cutler et al., 2022), and Sérsic index reflects the galaxy morphology. Therefore, by utilizing a sample of mock galaxies with similar parameters with the real targets, we could reasonably estimate the true uncertainties for the latter. To characterize the random uncertainties of the parameters, in line with previous analysis (e.g., Hoyos et al., 2011; Alabi et al., 2020) we fit these  $1\sigma$  deviations (with respect to the running medians) in bins of  $\langle\mu_{\text{eff}}\rangle$  with a simple linear function of the form:

$$\log(\sigma) = \alpha \times \langle\mu_{\text{eff}}\rangle + \beta \quad (2)$$

where  $\alpha$  and  $\beta$  are free coefficients, and  $\langle\mu_{\text{eff}}\rangle$  is the measured mean effective surface brightness. The red symbols in Figure 3 show the running random errors as a function of the output  $\langle\mu_{\text{eff}}\rangle$  for bins of increasing  $n$  (the panel layout is identical to that of Figure 2). The blue lines in Figure 3 present the fits to the running errors using above formula, with best-fitting coefficients  $\alpha$  and  $\beta$  shown in the middle left of each panel and also summarized in Table 1. As is expected, the amplitudes of the random uncertainties increase towards fainter surface brightness regimes. To assign typical errors to the fitted parameters of real galaxies, we evaluate the concrete functional form as given in Table 1, based on the measured  $n$ . The first column of Table 1 tabulates the adopted parameter ranges of  $n$ , which correspond to these ranges labeled in the panels of Figure 2 and Figure 3. Then the measured  $\langle\mu_{\text{eff}}\rangle$  are provided to functions (i.e., Formula 7) with given  $\alpha$  and  $\beta$ . The final adopted errors are the



**Figure 1:** Comparisons between the intrinsic values (input parameters) and the parametric measurements (output parameters) when using Galfitm to fit one Sérsic models to 160,765 simulated Sérsic galaxies, after excluding those in which the fit ran into any of the fitting constraints. From top to bottom panels, we plot the discrepancies (output - input) between the measured and true values for parameters magnitude,  $r_e$ ,  $n$ , and  $b/a$ , as a function of mean effective surface brightness  $\langle \mu_{\text{eff}} \rangle$ , which can be used as a proxy for S/N. Note that  $\langle \mu_{\text{eff}} \rangle$  is a measured/output quantity. In each panel, the purple points show individual simulated galaxies. The red solid lines represent the running medians of the differences, which are calculated in equally spaced bins of  $\langle \mu_{\text{eff}} \rangle$ , while the red dashed lines are the corresponding 16th and the 84th percentiles. The black dashed lines indicate the differences of zero. The color coding shows the number density of underlying data points.

**Table 1:** Summary of the best-fitting coefficients of Formula 7 for calculating random uncertainties

Parameter ranges of n	Mag		$r_e$		n		b/a		pa	
	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$
$0.2 < n < 1.5$	0.210	-6.194	0.220	-6.598	0.225	-6.279	0.184	-5.646	0.186	-4.002
$1.5 < n < 2.5$	0.155	-4.830	0.132	-4.388	0.177	-5.101	0.107	-3.839	0.151	-3.296
$2.5 < n < 3.5$	0.283	-7.689	0.248	-6.930	0.138	-4.143	0.091	-3.488	0.173	-3.893
$3.5 < n < 9.0$	0.178	-5.339	0.168	-5.046	0.095	-3.264	0.096	-3.917	0.147	-3.313

**Notes.** This table serves as the template to assign the random uncertainties caused by the object S/N, which are unaccounted for during the Galfitm fitting process. It lists the best-fitting coefficients  $\alpha$  and  $\beta$  that are used to derive the  $1\sigma$  deviations (i.e., random errors) on the total magnitude,  $r_e$ , n, b/a, and pa, as a function of measured  $\langle\mu_{\text{eff}}\rangle$ . The first column shows the various ranges of output n from the GalfitM fit, which correspond to these ranges labeled in the panels of Figure 2, Figure 3, and Figure 4, while the rest columns show the best-fitting coefficients in Formula 7 for the given Sérsic parameter.

**Table 2:** Summary of the best-fitting coefficients of Formula 8 for calculating systematics

Parameter ranges of n	Mag			$r_e$			n		
	$\alpha$	$\beta$	$\gamma$	$\alpha$	$\beta$	$\gamma$	$\alpha$	$\beta$	$\gamma$
$0.2 < n < 1.5$	0.645	-2.586	1.148e-15	0.548	-1.337	-1.665e-14	0.331	-2.736	-6.726e-0
$1.5 < n < 2.5$	0.333	-2.262	1.170e-07	0.309	-1.230	-2.932e-08	0.189	-0.557	-1.283e-0
$2.5 < n < 3.5$	0.328	-0.417	3.409e-09	0.463	-2.858	-5.489e-10	0.291	-1.217	-2.632e-0
$3.5 < n < 9.0$	0.214	-0.528	3.676e-06	0.212	-0.435	-2.734e-06	0.164	-0.259	-2.588e-0

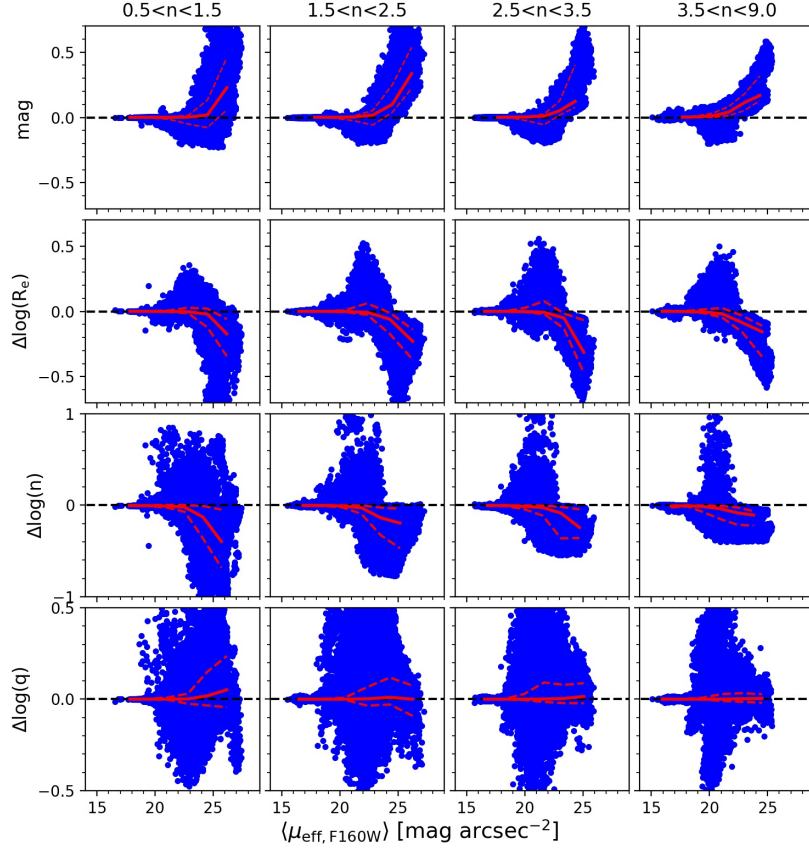
**Notes.** This table serves as the template to assign the systematic uncertainties caused by the object S/N for during the Galfitm fitting process. It lists the best-fitting coefficients  $\alpha$ ,  $\beta$ , and  $\gamma$  that are used to derive the  $1\sigma$  deviations (i.e., bias) on the total magnitude,  $r_e$ , n, b/a, and pa, as a function of measured  $\langle\mu_{\text{eff}}\rangle$ . The first column shows the various ranges of output n from the GalfitM fit, which are labeled in the panels of Figure 2, Figure 3, and Figure 4, while the rest columns show the best-fitting coefficients in Formula 8 for each Sérsic parameter. Since the systematic uncertainties on b/a and pa are substantially smaller than the random uncertainties, we do not correct the systematic effects for these parameters.

anti-logarithm of the results. Therefore, Table 1 serves as the template to infer the fitting uncertainties of the galaxy parameters returned by the pipeline. It is important to point out that these empirically estimated uncertainties (driven by object S/N) are significantly larger than the formal uncertainties output by Galfitm.

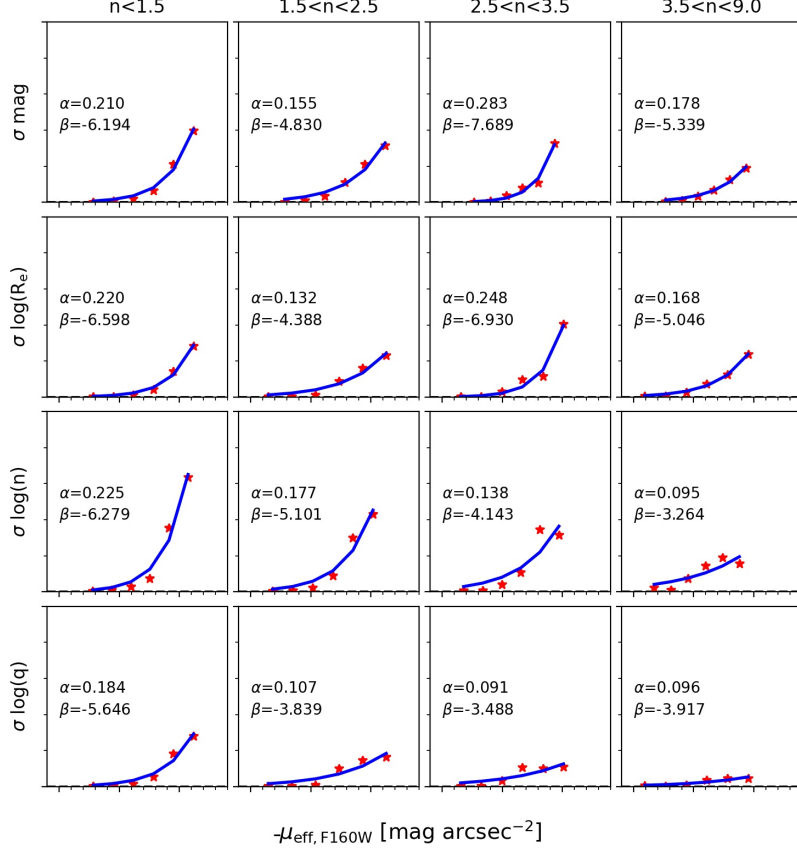
Similar to the determination of random uncertainties, we fit these systematic trends with exponential functions, which is given by

$$b = \gamma \times 10^{\alpha \times \langle\mu_{\text{eff}}\rangle + \beta} \quad (3)$$

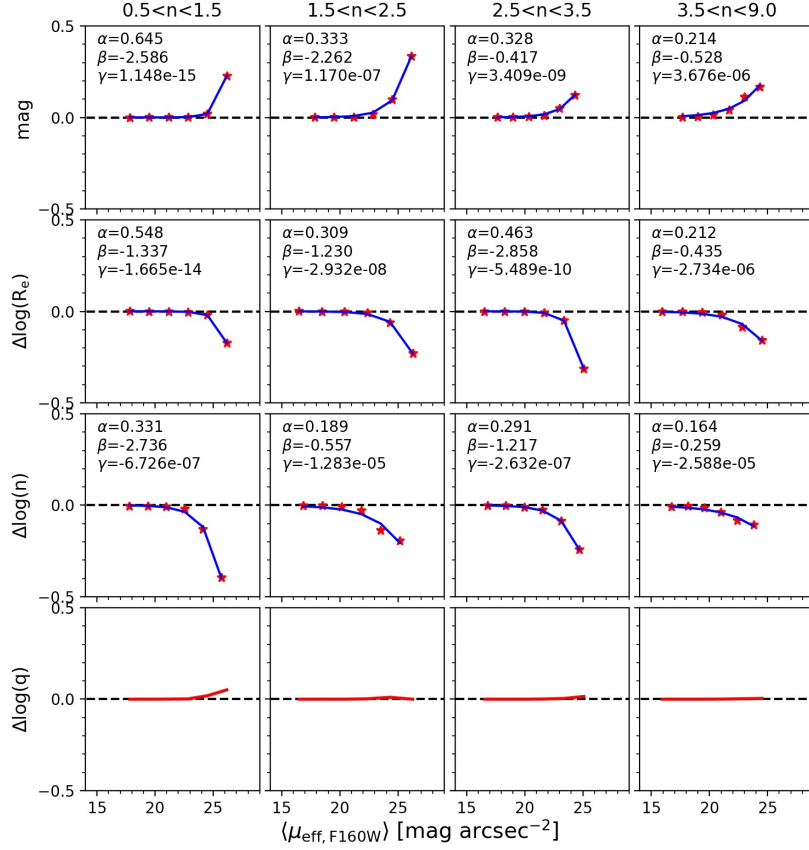
where  $\alpha$ ,  $\beta$ , and  $\gamma$  are free coefficients, and b (bias) indicates systematics (also see Formula 5). The fitting results are listed in Table 2, which displays the same format as in Table 1. They are then applied to extract systematic errors for real galaxy parameters. Since the systematic trends in b/a and pa are almost negligible compared to random uncertainties, we do not perform the fit and therefore set them to zero.



**Figure 2:** Parameter differences (output minus input) as a function of output  $\langle\mu_{\text{eff}}\rangle$  in our simulations. The sample is divided into 4 bins of increasing Sérsic index (i.e.,  $n \leq 1.5$ ,  $1.5 < n < 2.5$ ,  $2.5 \leq n \leq 3.5$ , and  $3.5 \leq n \leq 9.0$ , respectively), which are labelled as the titles on the top row of panels. In each panel, the overplotted red solid and dashed lines represent the running median and  $1\sigma$  deviation around the median, respectively.



**Figure 3:** Dependence of parameter random uncertainties on the  $\langle \mu_{\text{eff}} \rangle$ , as a function of  $n$  (labelled on the top row of panels). The panel layout is the same as that of Figure 2. The red symbols in each panel are the running  $1\sigma$  deviations around medians in bins of  $\langle \mu_{\text{eff}} \rangle$  from Figure 2. The blue solid lines represent the fits using Formula 7. The resulting best-fitting coefficients  $\alpha$  and  $\beta$  are explicitly shown in the middle left of the panel, and also given in Table 1.



**Figure 4:** Dependence of systematic uncertainties on the  $\langle \mu_{\text{eff}} \rangle$ , as a function of  $n$  (labelled on the top row of panels). The panel layout is the same as that of Figure 2 and 3. The red symbols in each panel are the running  $1\sigma$  deviations around medians in bins of  $\langle \mu_{\text{eff}} \rangle$  from Figure 2. The blue solid lines represent the fits using Formula 7. The resulting best-fitting coefficients are explicitly shown in the middle left of the panel, and also given in Table 1. The systematic effects for  $b/a$  are negligible compared to random fitting uncertainties. We thus do not correct the measurements for these systematic effects



**Table 3:** Systematic and random uncertainties of the Sérsic parameters for the real sources, driven by source S/N.

id	mag_sys (mag)	mag_err (mag)	re_sys (pixels)	re_err (pixels)	n_sys	n_err	q_sys	q_err	pa_sys (degree)	pa_err (degree)
1	0.0001	0.0132	-0.0048	0.2872	-0.0063	0.0173	0.0	0.0066	0.0	0.6614
2	0.0002	0.0217	-0.0038	0.1025	-0.0199	0.0412	0.0	0.0033	0.0	1.0243
3	0.0002	0.0005	-0.0	0.0006	-0.0044	0.0804	0.0	0.001	0.0	0.0644
4	0.0095	0.0714	-0.0902	0.362	-0.1532	0.1393	0.0	0.0503	0.0	2.9439
5	0.0602	0.0838	-0.4975	0.6313	-0.7472	0.695	0.0	0.0451	0.0	1.4102
6	0.0013	0.0372	-0.0151	0.1738	-0.0726	0.1076	0.0	0.0159	0.0	1.6536
7	0.0005	0.027	-0.0024	0.047	-0.069	0.1254	0.0	0.0312	0.0	1.243
8	0.0089	0.0699	-0.1831	0.7618	-0.1044	0.0967	0.0	0.0317	0.0	2.8868
9	0.0027	0.0058	-0.0081	0.0891	-0.0509	0.2466	0.0	0.0204	0.0	0.2763
10	0.0089	0.0082	-0.0117	0.0161	-0.1227	0.2439	0.0	0.0122	0.0	0.237
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

**Notes.** This table lists systematic and random uncertainties of single Sérsic parameters. The first column shows the object ID number in the catalog, while the rest columns show two type of uncertainties of magnitude (column 2, 3), effective radius (column 4, 5), Sérsic index (column 6, 7), axis ratio (column 8, 9), and position angle (column 10, 11).

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

In case of the galaxies with extremely faint  $\langle\mu_{\text{eff}}\rangle$ , where our simulated data may fail to cover, we will linearly extrapolate relevant best-fitting relations to fainter surface brightness regimes to obtain the uncertainties of interest. We caution the readers that our simulations are performed under the assumption that galaxies light profiles are perfectly described by Sérsic functions. However, this is not true for some galaxies with complex substructures, e.g., irregular galaxies, mergers, and nearby galaxies observed with sufficient details. Our simulations therefore could only provide a lower limit on the error bars for real galaxies. One should be also aware that this simulation is not designed to test the performance of the fitting code in crowded situations, where the impacts from nearby neighbors become more pronounced. We will extensively discuss the crowding errors in the next section.

## References

- Alabi, A. B., Romanowsky, A. J., Forbes, D. A., Brodie, J. P., & Okabe, N. 2020, MNRAS, 496, 3182
- Cutler, S. E. et al. 2022, ApJ, 925, 34
- Delaye L. et al., 2014, MNRAS, 441, 203
- Dimauro P., et al., 2018, MNRA, 478, 5410
- Gray, M. E., Wolf, C., Barden, M., et al. 2009, MNRAS, 393, 1275
- Häußler B., et al., 2013, MNRAS, 430, 330
- Häußler B. et al., 2007, ApJS, 172, 615
- Hoyos, C., den Brok, M., Verdoes Kleijn, G., et al. 2011, MNRAS, 411, 2439
- La Marca, A., Peletier, R., Iodise, E., et al. 2022, A&A, 659, 92
- van der Wel, A., Bell, E. F., Häußler B., et al. 2012, ApJS, 203, 24
- Venhola, A., Peletier, R., Laurikainen, E., et al. 2017, A&A, 608, A142
- Venhola A. et al., 2018, A&A, 620, A165
- Zaritsky D., Donnerstein R., Karunakaran A., Barbosa C. E., Dey A., Kadowaki J., Spekkens K., Zhang H., 2021, ApJS, 257, 60