Deliverable 3.7.2: Report on Mock Testing Results

WP 3.7: Low-surface-brightness science using the LSST

Project Acronym	LUSC-B
Project Title	UK Involvement in the Legacy Survey of Space and Time
Document Number	LUSC-B-14

Submission date	04/02/2020
Version	1.0
Status	Final Version
Author(s) inc. institutional affiliation	Aaron Watkins (LJMU Chris Collins (LJMU) Sugata Kaviraj (Hertfordshire)
Reviewer(s)	John Stott (Lancaster) Ivan Baldry (LJMU)

Dissemination level	
Public	*Public*

Version History

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	04/02/21	First draft	Watkins, Collins, Kaviraj
0.2	10/03/21	Second draft	Watkins, Collins, Kaviraj, Stott, Baldry
1.0	12/04/21	Final Version	

Table of Contents

Index of Figures

1 Executive Summary

We have analyzed a new set of model galaxies injected into the LSST pipeline meant to test the pipeline sky-subtraction routines, using the metric we developed in WP3.7.1. This new catalogue includes a wider array of parameter space—magnitudes, sizes, light profile shapes, and model axial ratios. We have found that regardless of profile shape or axial ratio, trends similar to what we identified in WP3.7.1 still appear in this expanded model set: under the current full focal plane sky subtraction routine, faint, low surface brightness objects suffer significantly more than bright, high surface brightness objects, with the faintest models losing on average ~0.5 magnitudes of flux due to the algorithm.

Additionally, we devised a new metric: the average surface brightness at which any given model's radial surface brightness profile is over-subtracted by 0.1 magnitudes/arcsec^2. We found that this value is stable regardless of which sub-sample of models we use to test it: for a typical galaxy, serious flux loss begins below ~26 magnitudes/arcsec^2, with a scatter of ~2.7 magnitudes/arcsec^2. **Through further experimentation, we found that the total amount of flux lost around each model due to the sky-subtraction follows a log-linear relationship with the model magnitudes, which explains much of the behavior described above. All of** this has serious implications for low-surface-brightness (LSB) science, which requires accurately measured fluxes as low as 32 magnitudes/arcsec^2, **orders of magnitude** fainter than is apparently currently achievable with the LSST pipeline. However, knowing this behavior, we can now easily propose a concrete target to aim for when amending the pipeline.

Without amending the pipeline, most of LSST's extra-galactic discovery space, which resides in the LSB regime, will remain inaccessible. In addition, since the pipeline processing will generically affect faint objects, it is likely to compromise the discovery space for many other science areas (e.g. studies of diffuse ISM, which are of interest to Galactic researchers).

2 Introduction

Our statistical understanding of how the Universe evolves is strongly determined by the objects and structures that are brighter than the surface-brightness limits of wide-area surveys. While huge strides have been made in comprehending galaxy evolution over the last few decades, using surveys like the SDSS, our understanding is naturally constrained by aspects of the Universe that are actually observable in such datasets. For example, the completeness of galaxies in surveys like the SDSS decreases rapidly for surface brightnesses fainter than ~24.5 mag arcsec^{-2} (e.g. [4]). However, the low-surface-brightness (LSB) regime, defined as the domain that is invisible in past wide-area surveys, contains a wealth of information that is essential for understanding how the observable Universe evolves over cosmic time.

First, both theory [6] and observation [2] indicate that the bulk of the galaxy population actually resides in the LSB regime. For example, ~50 (85) per cent of galaxies down to 10^8 (10^7) M_* inhabit this regime (see Table 2 in [6]). Second, there are key LSB components around high-surface-brightness galaxies that offer fundamental constraints on the evolution of the observable Universe. Two examples are merger-induced LSB tidal features, galactic stellar haloes, and intra-cluster or intra-group light (ICL, IGL). Tidal features encode the assembly histories of galaxies and constrain our structure-formation model. However, the surface-brightness of tidal features is a strong function of merger mass ratio. Given that low-mass galaxies far outnumber their massive counterparts, most mergers involve low mass ratios (i.e. are 'minor' mergers), which produce faint tidal features that are largely undetectable in past wide-area surveys (e.g. [3]). Nevertheless, both theory and observation suggest that minor mergers are key drivers of galaxy evolution, making the analysis of LSB tidal features an essential component of our galaxy-evolution effort. In a similar vein, ICL is a significant component of galaxy clusters, which are important tests of our cosmological model. Since the ICL contributes anything up to 40% of the baryonic mass budget of clusters at low redshift, a larger fraction than the contribution from the central brightest cluster galaxies in many cases [1], the utility of clusters as cosmological probes is closely linked to our ability to detect and characterize the diffuse ICL over cosmic time.

Under ideal conditions, LSST is capable of reaching depths fainter than **$\mu_\lambda \approx 30$ mag arcsec^{-2} in most photometric bands ([5])** over around 20,000 square degrees. *The LSB Universe thus represents virtually all the extra-galactic discovery space of this transformational survey.* However, LSB structures are acutely sensitive to sky over-subtraction. Preservation of LSB flux in LSST images is, therefore, a key requirement of the data-processing pipeline, without which LSST will not be capable of providing access to this revolutionary regime.

2.1 Purpose

We present here results of a new round of model injections, made using a wider parameter space than our previous round (Deliverable 3.7.1). Models are analyzed in the same manner as before, therefore we only briefly reiterate this methodology here: we perform aperture photometry of the models as injected into the images, with astrophysical sources subtracted out.

2.2 Glossary of Acronyms

SDSS: Sloan Digital Sky Survey

LSST: Legacy Survey of Space and Time

LSB : low surface brightness

HSB: high surface brightness

ICL: intra-cluster light

IGL: intra-group light

DM: data management

HSC: Hyper Suprime-Cam

3 Revised Model Catalogue

3.1 Catalogue creation

Our second model injection greatly expanded the parameter space. We show the new model parameter space in Figure 1. These limits were drawn from the distribution of the same parameters among real local Universe galaxies, using available Galfit decompositions of galaxies in the Spitzer Survey of Stellar Structure in Galaxies (S^4G; [8]; [7]) and the Fornax Deep Survey (FDS; [9] and references therein). To produce the catalogue, we proceeded as follows:

1. Using the limiting values of the parameters as measured in these surveys **(with eyeballed extensions to fainter magnitudes)**, we first created a grid of magnitude, Sérsic index, and effective radius, then culled from the grid any combinations of parameters that fall off of known galaxy scaling relations. This yielded ~40 valid points in the 3D grid, which served as the basis for our models. **The survey data scaling relations we used are shown in Figure 2 (as well as the Kormendy relation in the lower-right panel, for comparison), and an example grid drawn from this data is shown in Figure 3.**
2. From this grid, we created 40 copies of each possible model—20 with an axial ratio of 1.0, and 20 with an axial ratio of 0.5.
3. **We then projected all quantities to a distance of 100 Mpc to produce the final catalogue table.**
4. **Finally, we** randomly assigned each model (total number N ~ 1500) coordinates within Tract 9615 (a field in the GAMA15 Survey) as well as a random position angle. In assigning coordinates, we limited model proximity to some extent by disallowing models from being placed within $2R_{eff}$ of each other, however some significant overlap still occurred. We discuss this further in the following section.

While we also supplied the DM team with model ICL profiles, a memory error during the injection prevented their inclusion in the final catalogue, so they will not be discussed here. Similarly, we supplied a handful of stamps from the New Horizon simulation, but a bug in the software resulted in the stamps' orientation not being preserved properly, resulting in improperly coadded images of the stamps. Hence, we must exclude these as well from this analysis. These issues should be resolved by the next pipeline run, but given the consistency of the trends we uncover here with what we have uncovered previously, we do not believe the inclusion of the missing models will alter the results much.

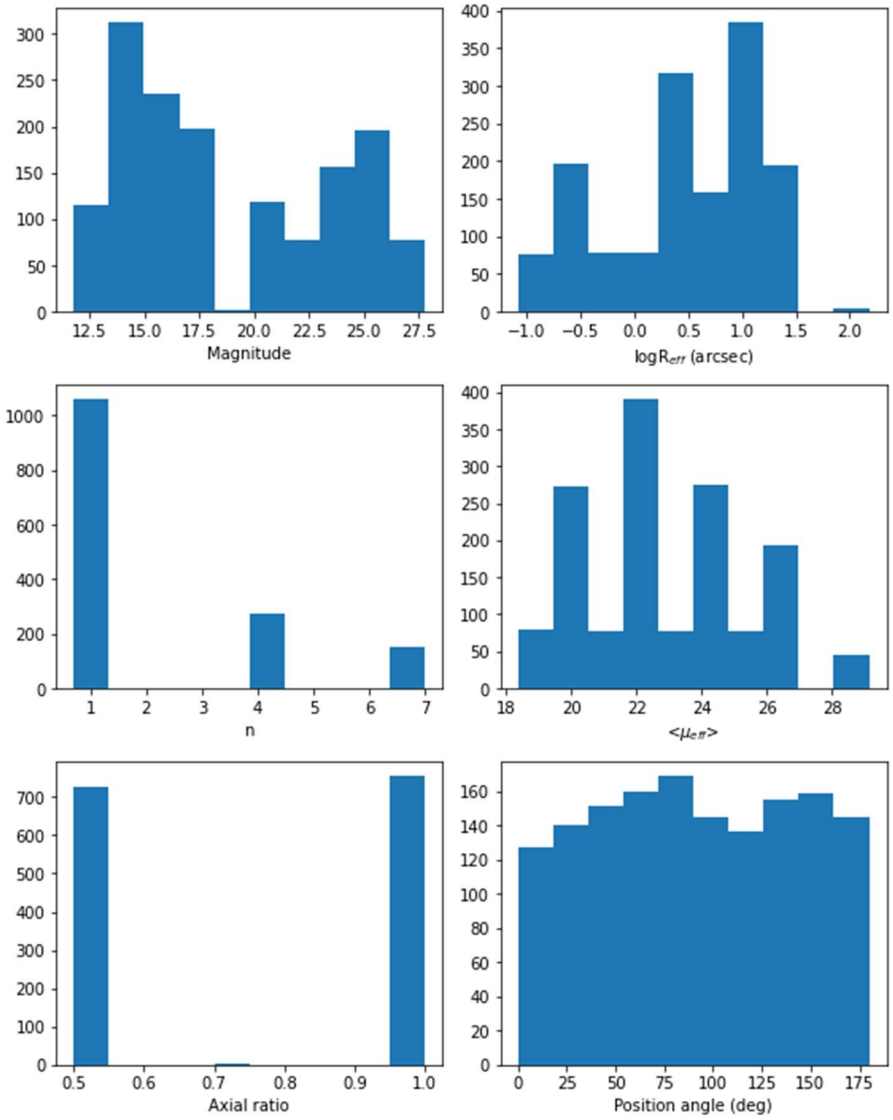

Figure 1 : Distributions of model parameters

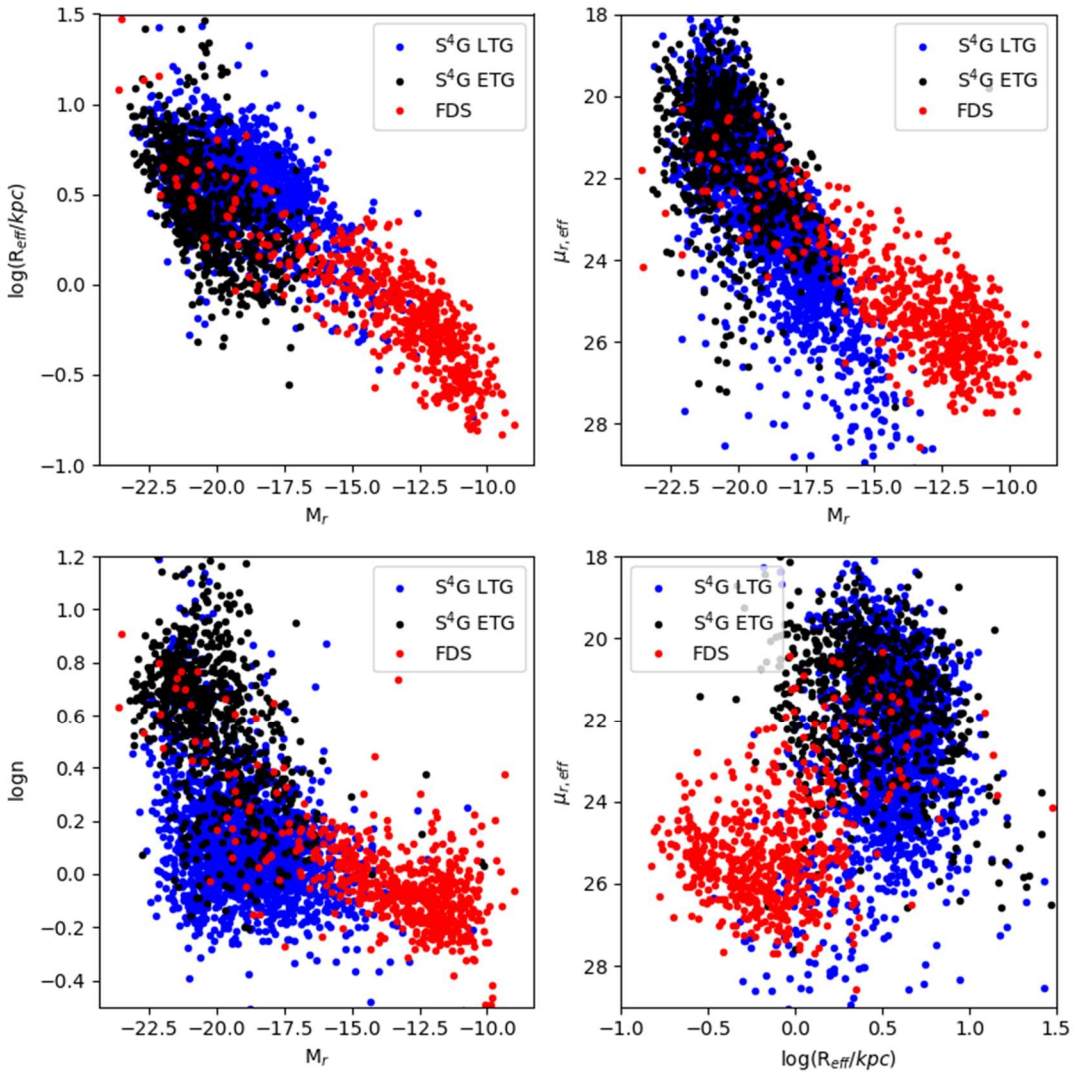

Figure 2: Scaling relations used to create grid

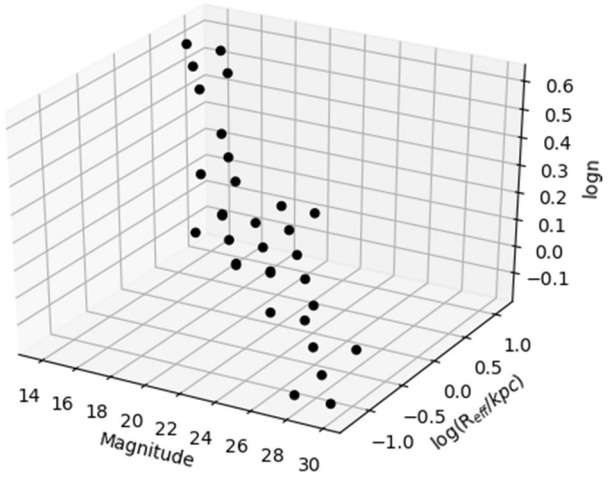

Figure 3: Final model grid points

3.2 Catalogue analysis

In this report, we assess only the full focal plane sky subtraction routine, and not the weak-lensing optimised routine that occurs as the final pipeline step. The latter, we feel, we have already shown to be incompatible with LSB science, and so our efforts are best spent improving the former.

To compare the models pre- and post-sky-subtraction, we **did aperture photometry on** the models at the injection level and at the coadd level, respectively**, with all astrophysical objects subtracted out, as described in the WP3.7.1 report. To briefly summarize, subtracting an image without models from an image with models leaves behind only the models; if one subtracts a pre-sky-subtraction image without models from a sky-subtracted image with models, this yields only the models and the sky subtraction. Our aperture photometry thus measures only the model fluxes pre- and post-sky-subtraction, without any noise, residual background, or contamination from real astrophysical objects.**

Models were injected in individual visits, a visit being the label for a single CCD image in a given band taken at a given date and time. **For the full-focal-plane sky subtraction, the only way to isolate injected models prior to any sky subtraction is at this step, hence we did photometry of the unmodified models at the visit level**. Models that landed close to the edge of a CCD were truncated, however—the pipeline had no capability of injecting them across CCDs. Also, to save processing time (the bulk of which was consumed by image retrieval), we used only the first such injection of all available, even if the model may have appeared less truncated in other visits due to the observational dithering pattern **(the model's fixed RA and Dec means that the models' positions with respect to the CCD edges move from visit to visit)**. All of this is reflected in the photometry, such that model magnitudes measured from individual CCDs are sometimes fainter than their input catalogue values **due to the model images being truncated at the CCD edge at the point of injection. This** problem does not persist at the same level to the coadds, **as there *all* dithered injections are stacked, including visits in which the model landed farther from the CCD edge (hence was not as truncated).**

We show the comparison between measured pre-sky-subtraction magnitudes and input catalogue magnitudes for all models in Figure 4, where this problem can be seen as points lying to the right of the 1:1 line **(here the colorbar represents the density of points)**. This resulted in many outlier points in the trends of over-subtraction shown in the following sections. We simply exclude these from consideration, as their outlier values of Δm result from an error in measurement, not from the sky-subtraction procedure.

Additionally, despite our efforts at avoiding overlap, many models still overlapped enough to result in measured magnitudes significantly brighter than their catalogue values (points left of the 1:1 line). **Because we do photometry of** the models at the **pixel** level, **when models overlap, our photometry apertures include flux from any overlapping models in both the pre- and post-sky-subtraction cases in the same manner. We thus find it reasonable to only compare magnitudes from our pre- and post-sky-subtraction aperture photometry, rather than considering the input catalogue magnitudes, which reflect the models' magnitudes as they would appear in isolation.**

Finally, during the model injection procedure, many of the models (mostly bright, large angular size models) were left out due to the same memory error that excluded the model ICL during the stamp creation process. Our photometry of these models at the CCD level thus resulted in extremely low catalogue values, as we inevitably drew

apertures in regions of the image with no flux save where they might overlap with neighbors. Because these models do not exist in the images, we rejected all models with total magnitudes measured to be more than 1 magnitude fainter than their catalogue values, as well as models in which we measured their post-sky-subtraction magnitudes to be > 40 or < 0 (suggesting a serious error in the photometry). We show these as red x's in Figure 4**; some of the latter case examples (post-sky-subtraction magnitudes >40 or <0) appear as x's nearby or to the left of the 1:1 line**. All of the following plots show values derived from only the black points in this figure.

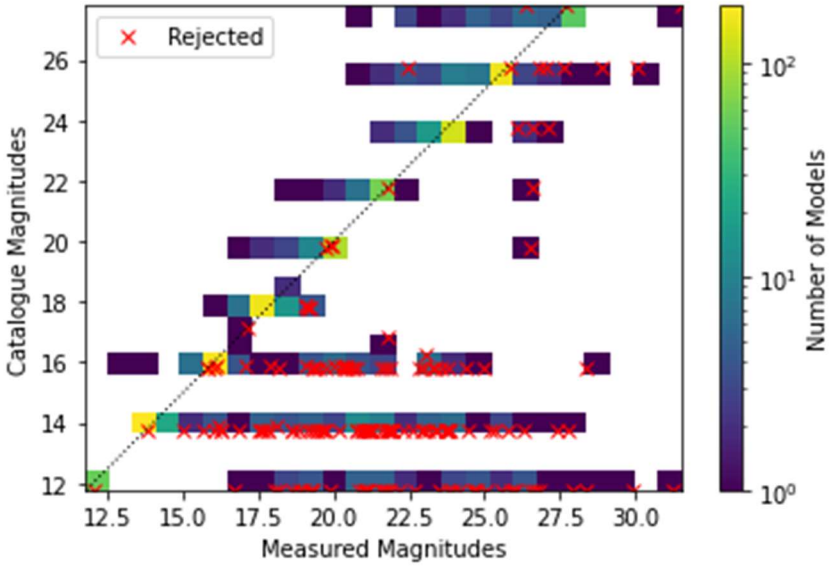

Figure 4 : Catalogue vs. pre-sky-subtraction aperture magnitude comparison

4 Results

We have expanded both the parameter space and the kinds of metrics we use to evaluate the pipeline sky subtraction algorithm. While we continue to measure the change in magnitude, Δm, we also now measure the change in surface brightness of individual radial profiles, a metric we describe in Section 4.2. **The latter, we found, was required to fully appreciate the behavior of the former on a model-by-model basis through the analysis of individual radial profiles, and in our analysis we found behavior that was noteworthy in its consistency. Finally, we show trends in flux loss directly rather than magnitude, which appears to complete the picture regarding how the sky subtraction is affecting the various models.**

4.1 First Metric: Δm

As described in our previous deliverable report, our primary metric is Δm, the change in total magnitude for each model resulting from the pipeline's sky-subtraction algorithm. We show these values for all models in Figure 5 and Figure 6. The left panels of Figure 5 show Δm plotted against model magnitudes (these were derived from the g-band; other bands appear similar and so are not shown), while the right panels show the same but for the mean effective surface brightnesses of each model within R_{eff}. Figure 6 shows similar plots for model effective radius and Sérsic index. Red squares show the median values in bins designated by the model grid spacing. In each panel, model points are color-coded by the values of the three other parameters not plotted on the x-axes, as well as by axial ratio.

We see some similar behavior to what we saw using our previous injection of flat models. Over-subtraction (positive Δm) occurs mostly for faint, low surface brightness models, with upward of half a magnitude over-subtraction for the faintest models injected. Trends with other parameters seem driven mostly by magnitude and surface brightness; for example, at an identical surface brightness, fainter (in magnitude) models are over-subtracted more than brighter models. Among the models we injected, faint models are also small in angular size, which appears as a systematic over-subtraction of models with low values of R_{eff}. Up to the largest models that made it into the images (~10"), there is no other trend of over-subtraction with size. Among the other parameters—axial ratio and Sérsic index—there are no obvious trends of over-subtraction other than those already followed by the models (e.g., faint models also have low Sérsic indexes).

This shows that the over-subtraction we demonstrated with the flat models persists among a normal population of galaxies in much the same way. With repeated models placed randomly around the tract, we are also able to measure the model-to-model scatter around these over-subtraction trends. We show this explicitly in Figure 7, where it becomes apparent that even among model categories that show no *systematic* over-subtraction, individual models are still frequently greatly over- or under-subtracted by the pipeline. Improving the pipeline will thus require not only removing the systematic over-subtraction trends we have uncovered here, but also will require reducing this galaxy-to-galaxy scatter.

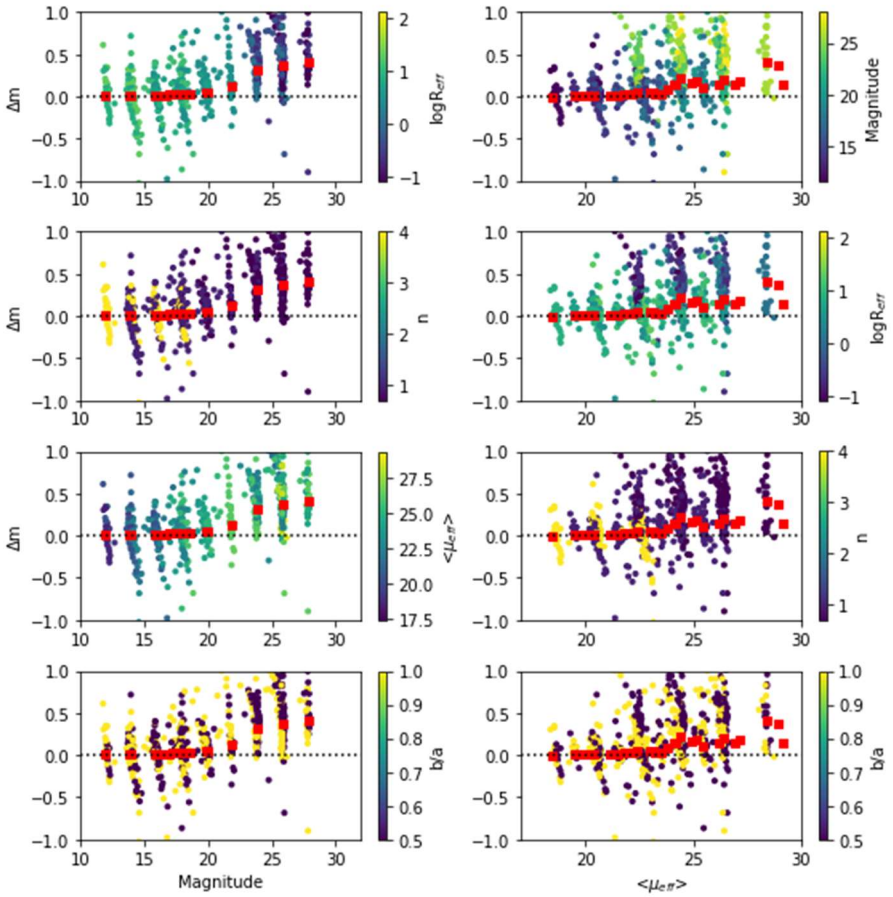

Figure 5 : Change in magnitude vs. magnitude and surface brightness

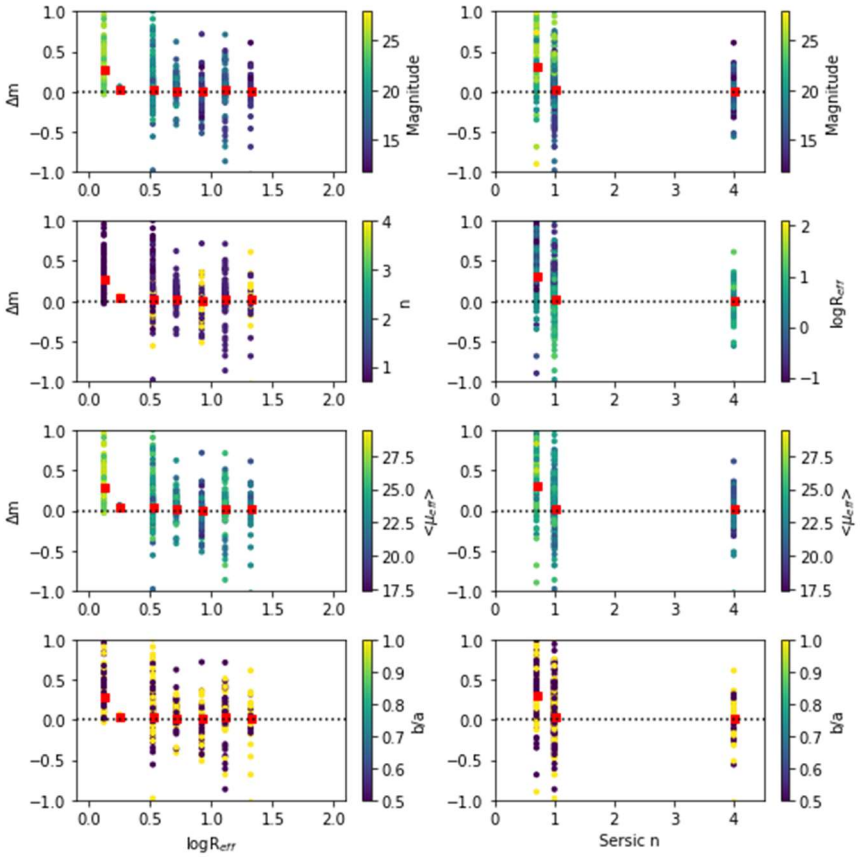

Figure 6 : Change in magnitude vs. size and Sérsic index

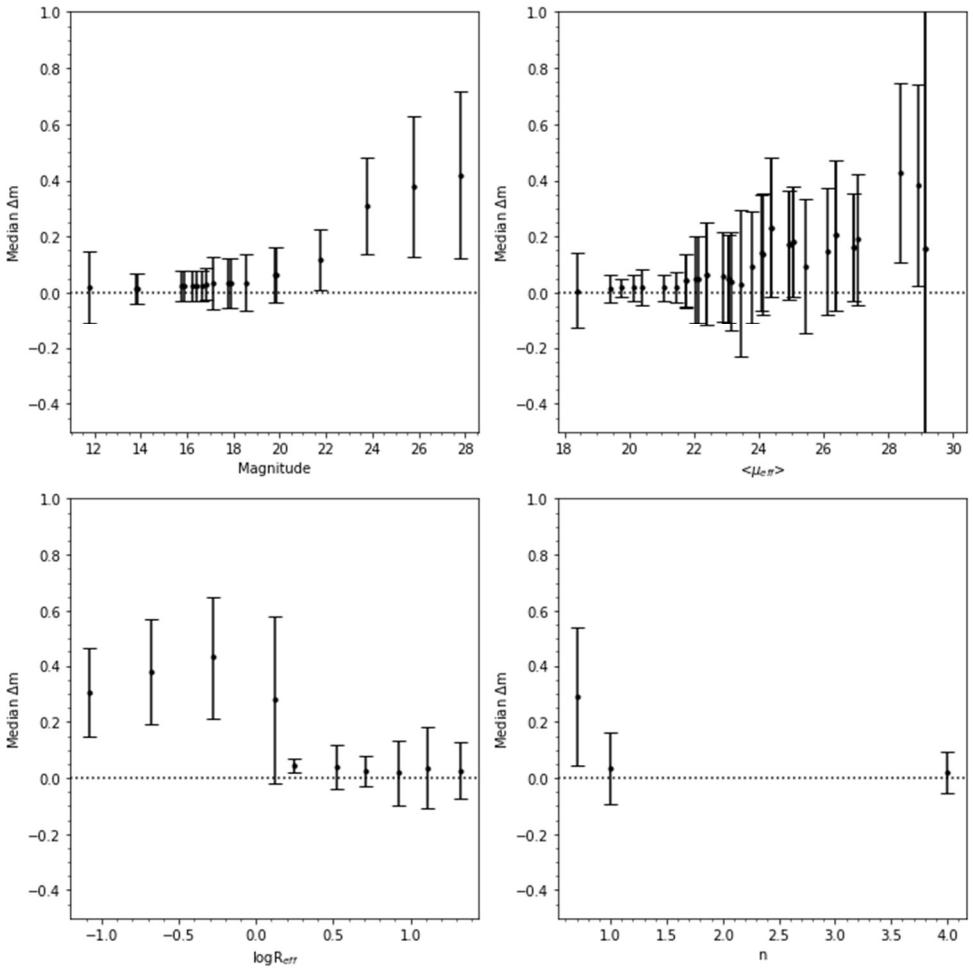

Figure 7 : Median and scatter in change of magnitude vs. all parameters

4.2 Second Metric: $\langle \mu_{\Delta\mu=0.1} \rangle$

In order to more fully understand the behavior of Δm by model type, we decided as well to explore the behavior shown by the models' individual radial surface brightness profiles. In so doing, we have **now** devised a second metric for measuring the over-subtraction by the pipeline. In this case, we measure the radial surface brightness profile of each model**—that being the mean surface brightness within consecutive 1-pixel-wide annuli from the model center out—**before and after sky-subtraction, then take the difference between the two profiles. We then derive, for each profile, the surface brightness at which the profile post-sky-subtraction differs from that pre-sky-subtraction by +0.1 magnitudes/arcsec^2 (roughly a 10% **loss of flux in a given annulus**). We then use the mean surface brightness at which this occurs—as well as the scatter about that mean—as our second metric: $\langle \mu_{\Delta\mu=0.1} \rangle$. We demonstrate this in Figure 8, which shows the distributions of this value for different model selections, both normalized and cumulative.

The three panels in this figure show, from left to right: only models following the selection criteria shown in Figure 4; only models with measured magnitudes within ±1 magnitude of their catalogue values; and no exclusions (thus including photometry of models that were not injected but still showed >0 flux due to neighbor contamination).

Regardless of the selection criteria, we find a consistent value for this metric: **~25.6** magnitudes/arcsec^2. The scatter is also similar: from left to right panels, σ = 2.74, 2.79, and 3.13. Therefore, we can conclude that for **~50% of the** models run through the

pipeline, noticeable **(>10%)** over-subtraction occurs already at **or above** this surface brightness, regardless of the model parameters. This explains to some extent the behavior demonstrated by Δm: faint, LSB models contain a larger fraction of their total flux below **25.6** magnitudes/arcsec^2 than large, HSB models, hence show larger magnitude differences post-sky subtraction. **We investigate this further in the following section.**

This behavior, combined with that shown via Δm, suggests that the way in which the sky-subtraction routine is over-subtracting models is quite consistent regardless of photometric profile shape or brightness: over-subtraction becomes problematic beginning at around 26 magnitudes/arcsec^2 across the entire field, with a large model-to-model scatter. This provides a solid target for pipeline improvement—for example, an acceptable pipeline might be able to maintain flux to within 10% down to at least 32 magnitudes/arcsec^2, with a low object-to-object scatter (preferably tenths of a magnitude).

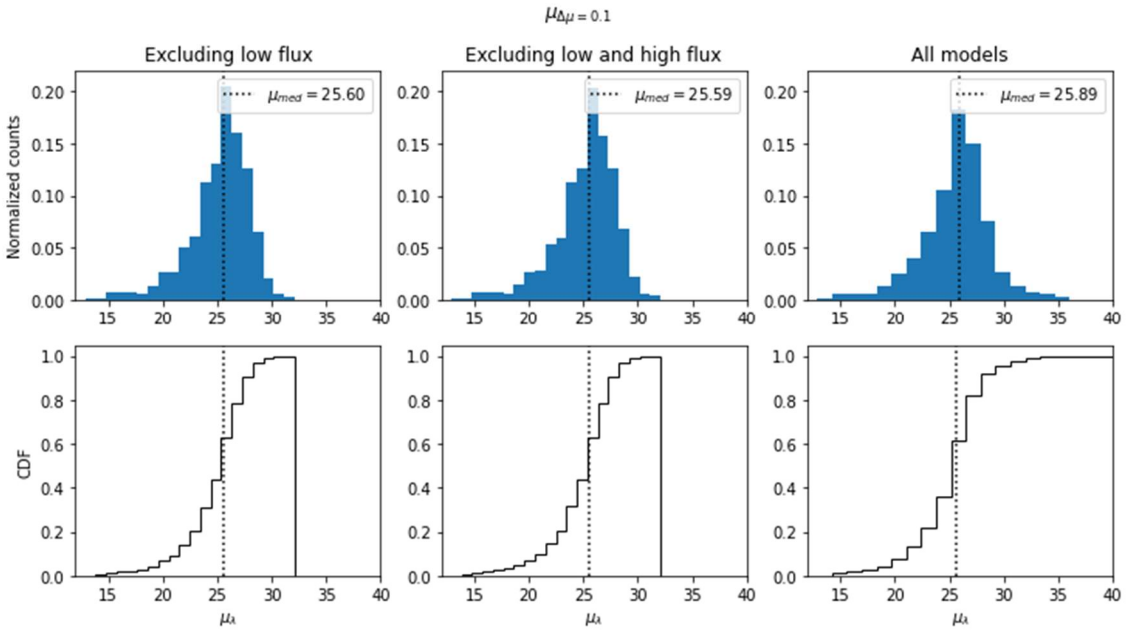

Figure 8 : Distribution of surface brightness at which Δμ = 0.1 magnitudes/arcsec^2

4.3 Testing linear flux offsets

Given the consistency of the behavior in $\langle\mu_{\Delta\mu=0.1}\rangle$, we now investigate the behavior of the models' flux loss in linear space rather than magnitude space. We therefore recreate the Figures from Section 4.1 but in linear space (Figure 9, Figure 10, and Figure 11). Here we define ΔF_ν as the difference between the models' total flux pre-sky-subtraction and the models' total flux post-sky-subtraction. Flux in HSC imaging is calibrated to the AB magnitude system, hence here 'cgs' refers to standard CGS units: ergs s^{-1} cm^{-2} Hz^{-1}. We show these flux losses as log(ΔF_ν), as the scales involved in the over-subtraction appear to require this.

This presentation clarifies the situation. Brighter, larger models clearly lose more flux than smaller, fainter models, in a surprisingly regular manner, with the correlations between magnitude and the other parameters again driving the behavior of most of the trends shown in these figures. Model axial ratio again appears to have no particular influence on the local sky subtraction. Though bright models appear to induce a greater local over-subtraction of sky, the extent of this flux loss is not large enough for the brighter models to appear as a

significant change in total magnitude. Likewise, fainter models lose less flux through sky-subtraction by value than brighter models, but the fraction lost is more substantial, resulting in greater changes in magnitude.

The trend of over-subtraction vs. model magnitudes is the cleanest and shows the least scatter, suggesting that it may be the primary driver behind the local over-subtraction of sky. Indeed, the colorbars show that models clearly separate in magnitude space when trends with other parameters are being tested. Figure 12, however, demonstrates that the exact nature of the over-subtraction is complex and asymmetric. This shows the models injected into Patch 4,4 as they appear post-sky-subtraction (with all astrophysical objects and background noise removed). Though we are not privy to the details of the sky subtraction algorithm, it appears some manner of grid-based sky modelling is still being imprinted on the models at low surface brightnesses.

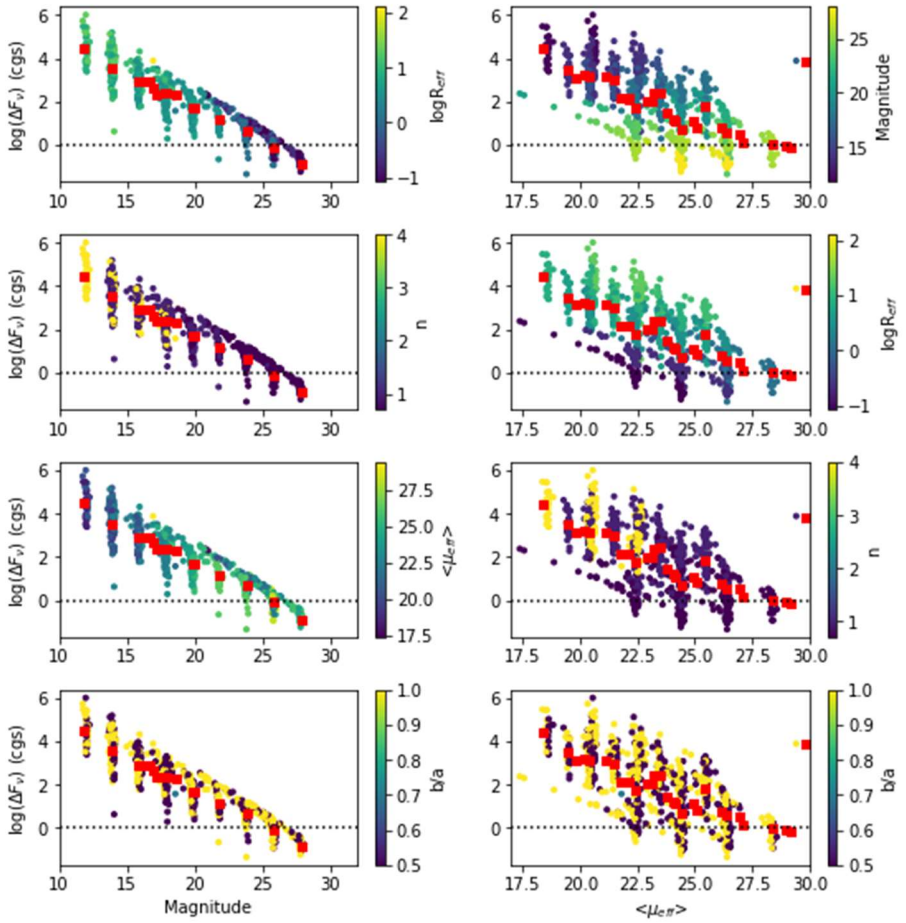

Figure 9: Linear flux loss vs. magnitude and surface brightness

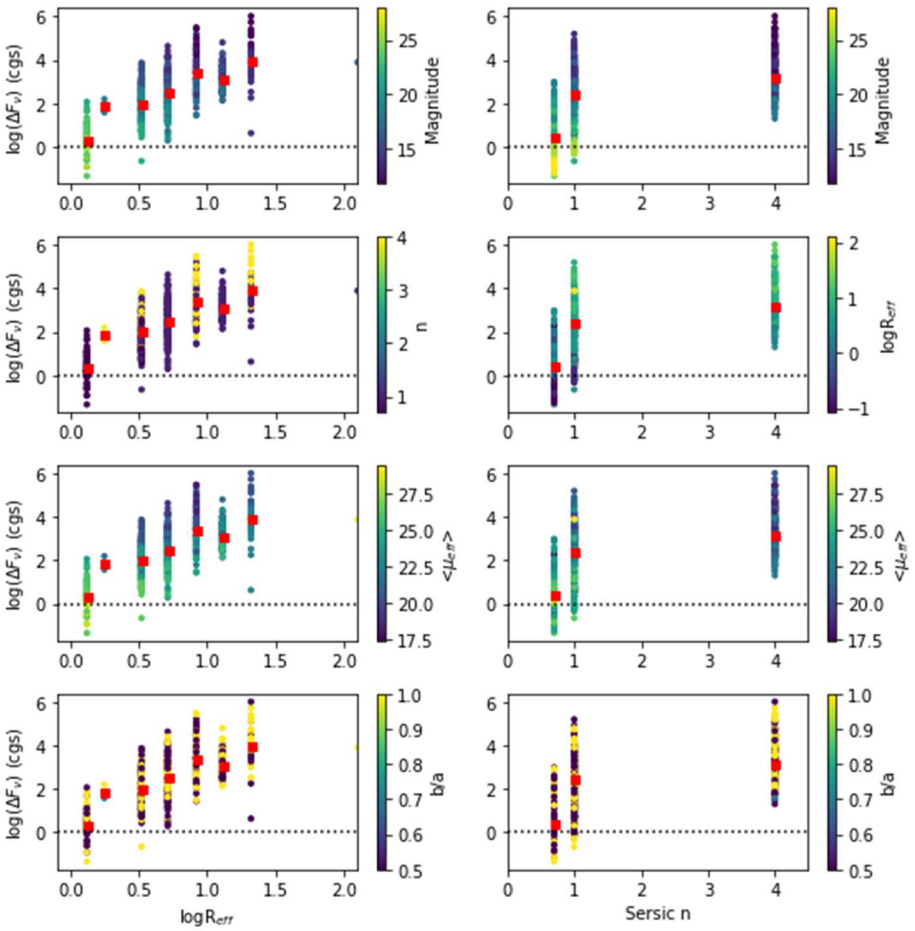

Figure 10: Linear flux loss vs. size and Sérsic index

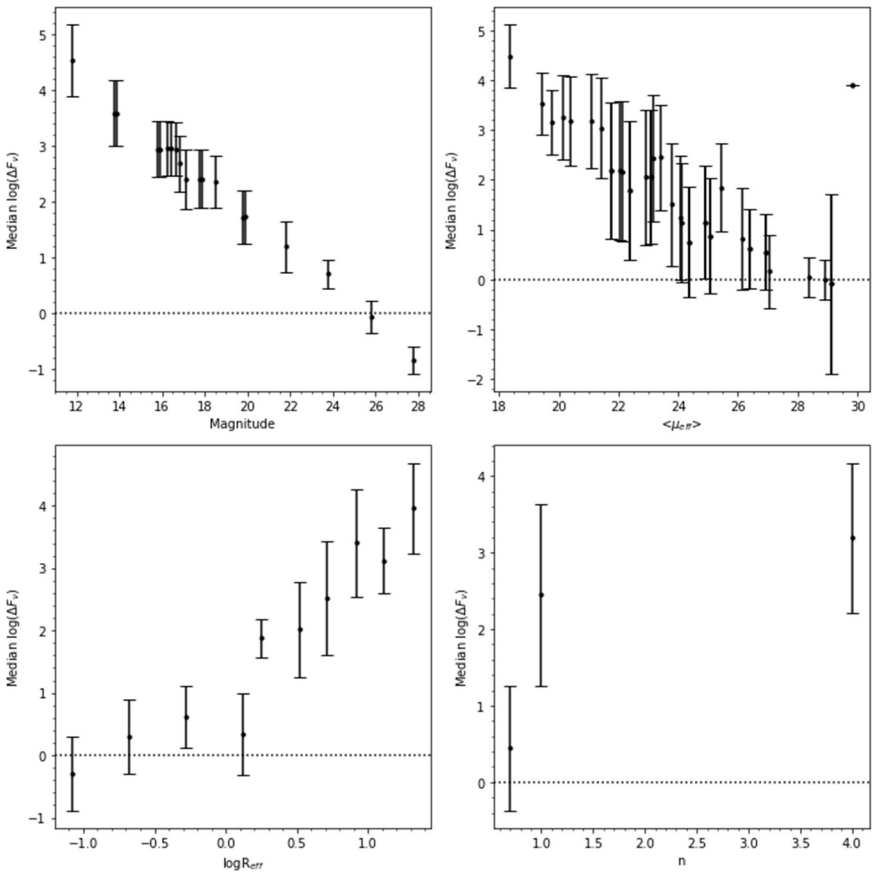

Figure 11: Median and scatter in flux loss vs. all parameters

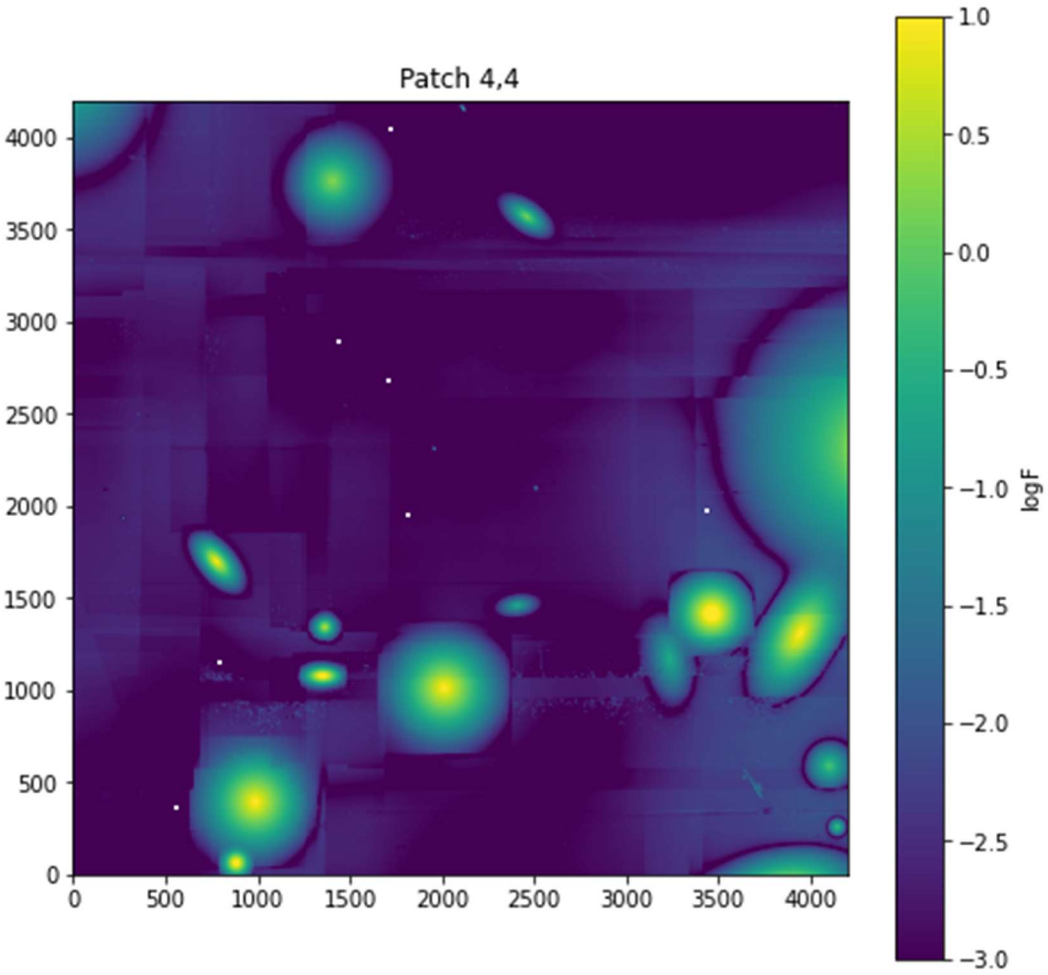

Figure 12: Example of over-subtraction on coadd

5 Summary

- We have expanded our testing of the LSST sky-subtraction algorithm using a new set of models covering a much wider parameter space, one which follows realistic galaxy scaling relations.
- We uncover very similar trends to those we uncovered using flat models:
  - Over-subtraction of models gets worse for fainter, lower surface brightness models, up to ~0.5 magnitudes on average for the worst cases. Scatter is high among all models regardless of parameter values.
  - No trend of over-subtraction appears with size for models up to ~10" effective radius; though we submitted larger models to the DM team, a problem with the injection software prevented their inclusion in this analysis.
  - Other parameters, such as Sérsic index and axial ratio, show no obvious trends with over-subtraction.
- We devised a secondary metric as well, that being the average surface brightness at which individual surface brightness profiles are over-subtracted by ~10%, or $\langle \mu_{\Delta\mu=0.1} \rangle$.
  - We measured $\langle \mu_{\Delta\mu=0.1} \rangle$ ~ 25.6; therefore, 50% of model galaxies were over-subtracted by 10% at ~26 magnitudes/arcsec^2.
  - The scatter among models about this value was also fairly stable: ~2.7 magnitudes/arcsec^2.
- **Inspired by the regularity of this second metric's behavior, we also have now examined the linear flux loss for all models as functions of the various parameters.**
  - **There is a regular power-law correlation between model magnitude and the total flux lost due to sky-subtraction, which seems to be imprinted on the correlations between flux loss and all other parameters (save axial ratio, which again shows no correlation).**
  - **This seems to provide a concrete explanation for the behavior seen using the previous two metrics.**
- This expanded catalogue, as well as our choice of metrics, provide a concrete goal for the pipeline that the DM team can try to meet using future iterations of this catalogue, or future catalogues.
- **In total, we find that the current sky subtraction routine will greatly hamper LSB science efforts with LSST, thus greatly reducing the survey's possible extragalactic discovery space.**

6 References

[1] Burke C., Collins C. A., Stott J. P., Hilton M., 2012, MNRAS, 425, 2058

[2] Dalcanton J. J., Spergel D. N., Gunn J. E., Schmidt M., & Schneider D. P., 1997, AJ, 114, 635

[3] Kaviraj S., 2014, MNRAS, 440, 2944

[4] Kniazev A. Y., Grebel E. K., Pustilnik S. A., Pramskij A. G., Kniazeva T. F., Prada F., & Harbeck D., 2004, AJ, 127, 704

[5] Laine, S., et al., 2018, eprint arXiv:1812.04897

[6] Martin G., et al., 2019, MNRAS, 485, 796

[7] Salo, H. et al. 2015, ApJS, 219, 4

[8] Sheth, K. et al. 2010, PASP, 122, 1397

[9] Su, A. H., et al. 2020, arXiv preprint: 2101.05699

Annex A. Acknowledgements

We are grateful to the DM pipeline team without whom this project will not be possible. Johan Knapen is thanked for many interesting discussions.