

# Fermilab

Dark Energy Survey Year 6 Results: Synthetic-source  
Injection Across the Full Survey Using Balrog

FERMILAB-PUB-24-0940-T

arXiv:2501.05683

This manuscript has been authored by Fermi Research Alliance, LLC  
under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy,  
Office of Science, Office of High Energy Physics.

## Dark Energy Survey Year 6 Results: Synthetic-source Injection Across the Full Survey Using BALROG

D. ANBAJAGANE (安部),<sup>1,2</sup> M. TABBUTT,<sup>3</sup> J. BEAS-GONZALEZ,<sup>3</sup> B. YANNY,<sup>4</sup> S. EVERETT,<sup>5</sup> M. R. BECKER,<sup>6</sup> M. YAMAMOTO,<sup>7,8</sup> E. LEGNANI,<sup>9</sup> J. DE VICENTE,<sup>10</sup> K. BECHTOL,<sup>3</sup> J. ELVIN-POOLE,<sup>11</sup> G. M. BERNSTEIN,<sup>12</sup> A. CHOI,<sup>13</sup> M. GATTI,<sup>12</sup> G. GIANNINI,<sup>9,1</sup> R. A. GRUENDL,<sup>14,15</sup> M. JARVIS,<sup>12</sup> S. LEE,<sup>16</sup> J. MENA-FERNÁNDEZ,<sup>17</sup> A. PORREDON,<sup>10,18</sup> M. RODRIGUEZ-MONROY,<sup>19</sup> E. ROZO,<sup>20</sup> E. S. RYKOFF,<sup>21,22</sup> T. SCHUTT,<sup>22,21,23</sup> E. SHELDON,<sup>24</sup> M. A. TROXEL,<sup>8</sup> N. WEAVERDYCK,<sup>25,26</sup> V. WETZELL,<sup>12</sup> M. AGUENA,<sup>27</sup> A. ALARCON,<sup>28</sup> S. ALLAM,<sup>4</sup> A. AMON,<sup>7</sup> F. ANDRADE-OLIVEIRA,<sup>29</sup> D. BROOKS,<sup>30</sup> A. CARNERO ROSELL,<sup>31,27,32</sup> J. CARRETERO,<sup>9</sup> C. CHANG,<sup>2,1</sup> M. CROCCE,<sup>33,28</sup> L. N. DA COSTA,<sup>27</sup> M. E. S. PEREIRA,<sup>34</sup> T. M. DAVIS,<sup>35</sup> S. DESAI,<sup>36</sup> H. T. DIEHL,<sup>4</sup> S. DODELSON,<sup>2,1,4</sup> P. DOEL,<sup>30</sup> A. DRLICA-WAGNER,<sup>1,4,2</sup> A. FERTÉ,<sup>21</sup> J. FRIEMAN,<sup>1,4</sup> J. GARCÍA-BELLIDO,<sup>19</sup> E. GAZTANAGA,<sup>28,37,33</sup> D. GRUEN,<sup>38</sup> G. GUTIERREZ,<sup>4</sup> W. G. HARTLEY,<sup>39</sup> K. HERNER,<sup>4</sup> S. R. HINTON,<sup>35</sup> D. L. HOLLOWOOD,<sup>40</sup> K. HONSCHIED,<sup>41,42</sup> D. HUTERER,<sup>29</sup> D. J. JAMES,<sup>43</sup> E. KRAUSE,<sup>44</sup> K. KUEHN,<sup>45,46</sup> O. LAHAV,<sup>30</sup> J. L. MARSHALL,<sup>47</sup> R. MIQUEL,<sup>48,9</sup> J. MUIR,<sup>49</sup> J. MYLES,<sup>7</sup> A. PIERES,<sup>27,50</sup> A. A. PLAZAS MALAGÓN,<sup>21,22</sup> J. PRAT,<sup>51,2</sup> M. RAVERI,<sup>52</sup> S. SAMUROFF,<sup>53,9</sup> E. SANCHEZ,<sup>10</sup> D. SANCHEZ CID,<sup>10</sup> I. SEVILLA-NOARBE,<sup>10</sup> M. SMITH,<sup>54</sup> E. SUCHYTA,<sup>55</sup> G. TARLE,<sup>29</sup> D. L. TUCKER,<sup>4</sup> A. R. WALKER,<sup>56</sup> P. WISEMAN,<sup>57</sup> Y. ZHANG,<sup>56</sup>

(DES COLLABORATION)

<sup>1</sup>Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

<sup>2</sup>Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA

<sup>3</sup>Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, WI 53706-1390

<sup>4</sup>Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

<sup>5</sup>California Institute of Technology, 1200 East California Blvd, MC 249-17, Pasadena, CA 91125, USA

<sup>6</sup>Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

<sup>7</sup>Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA

<sup>8</sup>Department of Physics, Duke University Durham, NC 27708, USA

<sup>9</sup>Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain

<sup>10</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

<sup>11</sup>Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

<sup>12</sup>Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>13</sup>NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

<sup>14</sup>Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

<sup>15</sup>Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA

<sup>16</sup>Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA

<sup>17</sup>LPSC Grenoble - 53, Avenue des Martyrs 38026 Grenoble, France

<sup>18</sup>Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute, German Centre for Cosmological Lensing, 44780 Bochum, Germany

<sup>19</sup>Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, 28049 Madrid, Spain

<sup>20</sup>Department of Physics, University of Arizona, Tucson, AZ 85721, USA

<sup>21</sup>SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

<sup>22</sup>Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

<sup>23</sup>Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA

<sup>24</sup>Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA

<sup>25</sup>Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA

<sup>26</sup>Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

<sup>27</sup>Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

<sup>28</sup>Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain

<sup>29</sup>Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

<sup>30</sup>Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

<sup>31</sup>Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain

<sup>32</sup>Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain

<sup>33</sup>Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain

<sup>34</sup>Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany

<sup>35</sup>School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia

<sup>36</sup>Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India

- <sup>37</sup>*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*
- <sup>38</sup>*University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany*
- <sup>39</sup>*Department of Astronomy, University of Geneva, ch. d'Écogia 16, CH-1290 Versoix, Switzerland*
- <sup>40</sup>*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
- <sup>41</sup>*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
- <sup>42</sup>*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
- <sup>43</sup>*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
- <sup>44</sup>*Department of Astronomy/Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA*
- <sup>45</sup>*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
- <sup>46</sup>*Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA*
- <sup>47</sup>*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*
- <sup>48</sup>*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
- <sup>49</sup>*Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada*
- <sup>50</sup>*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
- <sup>51</sup>*Nordita, KTH Royal Institute of Technology and Stockholm University, Hannes Alfvéns väg 12, SE-10691 Stockholm, Sweden*
- <sup>52</sup>*Department of Physics, University of Genova and INFN, Via Dodecaneso 33, 16146, Genova, Italy*
- <sup>53</sup>*Department of Physics, Northeastern University, Boston, MA 02115, USA*
- <sup>54</sup>*Physics Department, Lancaster University, Lancaster, LA1 4YB, UK*
- <sup>55</sup>*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831*
- <sup>56</sup>*Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile*
- <sup>57</sup>*School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK*

## ABSTRACT

Synthetic source injection (SSI), the insertion of sources into pixel-level on-sky images, is a powerful method for characterizing object detection and measurement in wide-field, astronomical imaging surveys. Within the Dark Energy Survey (DES), SSI plays a critical role in characterizing all necessary algorithms used in converting images to catalogs, and in deriving quantities needed for the cosmology analysis, such as object detection rates, galaxy redshift estimation, galaxy magnification, star-galaxy classification, and photometric performance. We present here a source injection catalog of 146 million injections spanning the entire 5000 deg<sup>2</sup> DES footprint, generated using the BALROG SSI pipeline. Through this sample, we demonstrate that the DES Year 6 (Y6) image processing pipeline provides accurate estimates of the object properties, for both galaxies and stars, at the percent-level, and we highlight specific regimes where the accuracy is reduced. We then show the consistency between SSI and data catalogs, for all galaxy samples developed within the weak lensing and galaxy clustering analyses of DES Y6. The consistency between the two catalogs also extends to their correlations with survey observing properties (seeing, airmass, depth, extinction, etc.). Finally, we highlight a number of applications of this catalog to the DES Y6 cosmology analysis. This dataset is the largest SSI catalog produced at this fidelity and will serve as a key testing ground for exploring the utility of SSI catalogs in upcoming surveys such as the Vera C. Rubin Observatory Legacy Survey of Space and Time.

*Keywords:* sky surveys - astronomical simulations - cosmology - dark energy

## 1. INTRODUCTION

Photometric datasets have a storied history spanning more than a century:<sup>1</sup> from the popularization of long-exposure photographic plates by Roberts (1893), to the half-sky survey of Barnard et al. (1919, 1927) revealing the faint structures

(“dark nebulae”) of the Milky Way, to the international Carte du Ciel project (circa 1887) that spanned telescopes from twenty institutions and whose measurements of stars continue to be used today for long-baseline astrometry (Lehtinen et al. 2023). The first hints of the large-scale structure of our Universe were also discovered through photometric surveys: from the work of Hubble (1934) in constructing photometric galaxy (“nebulae”) catalogs from 650 deg<sup>2</sup> and uncovering

<sup>1</sup> Photometry performed *through photography/plates* is a little over a century old, whereas photometry-by-eye has existed for over two millennia and extends back to the catalogs of Hipparchus of Nicea in 120 BC and Ptolemy’s “Almagest” from 137 AD (Miles 2007).

signatures of their spatial clustering, to that of Abell (1958) in discovering and cataloging large, clusters of galaxies.<sup>2</sup>

Wide-field, digital photometric surveys, a more recent endeavor in this long history, have amplified the importance of such datasets (e.g., York et al. 2000; Flaugher 2005; de Jong et al. 2013; Aihara et al. 2018; Dawson et al. 2013, 2016). Thus, the assembly of precision photometric datasets is now a mature field. It requires a multitude of steps (e.g., Stoughton et al. 2002; Magnier 2006; Skrutskie et al. 2006; Valdes et al. 2014; de Jong et al. 2015; Bosch et al. 2018; Morganson et al. 2018; Wright et al. 2024) that have been studied and calibrated in precise detail. However, as the field of cosmology moves further into the precision frontier, we require even higher precision in our understanding of our photometric datasets.<sup>3</sup> Reliable cosmological inference now requires a robust calibration of uncertainties and biases associated with the complexities in the imaging data (such as, for example, the impact of a range of purely observational effects) as well as of the algorithms in the image processing pipelines.

Synthetic source injection (SSI) provides one such avenue for studying these effects. By injecting sources with known properties (such as size, luminosity, color, etc.) into the on-sky images from a survey, and processing those augmented images through the same pipeline as the survey data, we can understand the performance of the image processing pipeline in characterizing different kinds of astronomical sources. Such a calibration is critical for many downstream analyses of photometric datasets (we describe the analyses relevant for our work later in this paper). By injecting synthetic sources into the individual pixel-level images, the SSI catalog reflects the complexities of the on-sky data in a way that is challenging to replicate, at full fidelity, with purely simulated images. Complexities include instrument signatures, spatial varying backgrounds, blending with complicated scenes of real galaxies etc. The method naturally also incorporates the exact distributions of delivered image quality, photometric zeropoints, sky background variances, etc., corresponding to the images used in deriving the DES Y6 data products and the corresponding cosmology results.

For these reasons, SSI has been used in many analyses within the Dark Energy Survey (DES, Flaugher 2005): see

Suchyta et al. (2016, henceforth, S16) and Everett et al. (2022, henceforth, E22) for the DES Year 1 (Y1) and Year 3 (Y3) SSI pipelines, respectively. The DES Y1 SSI campaign emulated some pieces of the full image processing pipeline, while the DES Y3 SSI campaign emulated nearly all pieces of this pipeline. In this work, we describe the DES Y6 SSI campaign, which further enhances the consistency between the image processing used in the SSI pipeline (denoted as BALROG) and that used on the data. Our injection catalog constitutes fits to the light profiles of objects in the DES Y3 deep fields (Hartley & Choi et al. 2021) — where these deep-field measurements have significantly lower noise relative to those from the wide-field — and is therefore completely data-driven.

Other surveys are also undertaking SSI to understand the statistical properties and biases of their datasets; for example, the OBI-WAN pipeline (Kong et al. 2024) used in the Dark Energy Spectroscopic Instrument (DESI Collaboration et al. 2016), the SYNPIPE software used in the Hyper Suprime-Cam (HSC) data (Miyazaki et al. 2018), or the COSMODC datasets (Sánchez et al. 2020) used in the Dark Energy Science Collaboration (DESC, The LSST Dark Energy Science Collaboration et al. 2018).

This work supports the DES Y6 weak lensing and galaxy clustering cosmology analysis by presenting the associated SSI dataset, its validations, and its applications. We detail the source injection pipeline, the image processing pipeline, and all updates made to both in relation to the previous Y3 effort. Of particular note is a new, optimized injection scheme that improves the number of usable injections by up to a factor of 12, and also an extension of SSI to the full,  $\sim 5000 \text{ deg}^2$  footprint of the survey. As a result, the Y6 SSI catalog has nearly two orders of magnitude more sources compared to the Y3 SSI catalog. We also characterize the performance of the image processing pipeline in accurately recovering the input properties of simulated objects, and then showcase the similarities between the synthetic samples and the actual catalogs from the Y6 datasets. The Y3 analysis (E22) also presented a detailed study of their photometric measurements, identifying failure modes (e.g., blending, sky background errors etc.). Given those algorithms studied in E22 are being used in DES Y6 with minor/no changes, we do not repeat all aspects of their study in this work.

The BALROG synthetic dataset plays critical roles in estimating multiple, key quantities for the cosmology analysis: for example, estimating the redshift distributions of the different galaxy samples (Myles & Alarcon et al. 2021; Sánchez et al. 2023; Giannini et al. 2024), the magnification of galaxies due to foreground structure (Garcia-Fernandez et al. 2018; Elvin-Poole & MacCrann et al. 2023), and validating the systematics treatment of the galaxy catalogs (Rodríguez-Monroy et al. 2022). We describe a few of these

<sup>2</sup> While this entire description mentions (a subset of) the history of *photographic* surveys — i.e. those that used photographic plates to image objects — we point out the first precursors to our modern widefield surveys were the visually identified, non-photographic catalogs of Sir William Herschel (Herschel 1786, 1789, 1802), which covered thousands of square degrees of the sky and were eventually integrated into the NGC catalogs (Dreyer 1888) that are still referred to today.

<sup>3</sup> For example, the relative photometric accuracy in DES Y6 is calibrated to better than 1% (Bechtol et al. 2025), and the morphological measurements (primarily the galaxy ellipticity/orientations) are also better than 1% (Yamamoto & Becker et al. 2025; Mau & Becker et al. in prep.).

applications in this work, with a focus on how BALROG is employed in these analyses.

This work is organized as follows: in Sections 2 and 3 we detail the image processing pipeline of Y6 BALROG, including all similarities/differences with the fiducial pipeline applied to the data and with the previous Y3 BALROG pipeline. In Section 4 we validate the synthetic samples, through a number of comparisons with the true, injected samples, and also with the actual data catalogs, and then examine the photometric accuracy of the synthetic sample. In Section 5 we detail some key applications of BALROG in the Y6 cosmology efforts, as detailed above. We conclude in Section 6. In Appendix A we describe the correlations between the galaxy number density of two lens samples and the survey property maps. Appendix B details further photometric validations for the synthetic star sample in BALROG.

## 2. THE BALROG PIPELINE

BALROG—the DES internal nomenclature for our SSI pipeline<sup>4</sup>—injects sources with known photometry and morphology into on-sky images, and then replicates the entire processing of the DES Data Management pipeline (DESDM; Morganson et al. 2018). The processed images are then run through the same postprocessing steps as the data to obtain a variety of object catalogs. These catalogs are then position-matched to the true, injected sources to thereby obtain a distribution of noisy, realistic measurements for each injected source.

We now detail the full pipeline, including both image processing and galaxy catalog generation, in Section 2.1. We describe specific differences between the BALROG and DESDM pipelines in Section 2.2, with explicit justifications for each difference.

### 2.1. Emulating the DES Data Management (DESDM) Pipeline

The image processing in DES follows the methods presented in Morganson et al. (2018), with some updates made during the Y3 (Sevilla-Noarbe et al. 2021) and Y6 analyses (Bechtol et al. 2025). The BALROG pipeline replicates nearly all steps performed by DESDM. Some minor steps are avoided in order to reduce the complexity of the pipeline, to speed up the execution, or due to other limitations. All relevant differences are detailed in Section 2.2.

The primary unit of the DESDM coadd image processing is the “coadd tile” (Morganson et al. 2018), which is a  $0.7 \times 0.7 \text{ deg}^2$  square region of the sky, spanning  $10,000 \times 10,000$

pixels. The DES footprint is partitioned into 10,169 tiles, all of which are re-processed with the BALROG injections; see Figure 1 for the area covered by the BALROG sample. DES accumulated a total of 83,706 exposures in the wide survey, with each location in the footprint being imaged an average of 8 times in each of the *grizY* bands (DES Collaboration et al. 2021). The *Y* band is not used in the cosmology analysis, while the *g* band is not used in object detection. The image processing pipeline performs coaddition on all exposures that overlap with the defined tile region, and generates catalogs for the tile.

The BALROG pipeline for an individual coadd tile undertakes the following steps:

1. **Database query:** Given a coadd tile, we find all CCD images, in the *griz* bands, that are associated with that tile; that is, all images used in generating the coadd image of that tile. These data products are hosted at the National Center for Supercomputing Application (NCSA) and queried locally for processing. Note that this step involves locating not only the raw images, but also models of the image point spread function (PSF) from PIFF (Jarvis et al. 2020)<sup>5</sup>, astrometric solutions from SCAMP (Bertin 2006), sky background, noise, variance, and mask plane images from SOURCEEXTRACTOR (Bertin & Arnouts 1996), etc.
2. **Null-weight images:** The actual images used for generating coadd images (as well as other imaging products therein) are not the raw images downloaded in Step 1 above, but rather *null-weighted* images. These images have had an initial detrending and instrument signature removal applied via the DESDM Final Cut pipeline (Morganson et al. 2018). This includes a masking and interpolation step to remove problematic features (e.g., cosmic rays) in the images, and also corrects for biases and non-linear responses in the CCD pixels (Morganson et al. 2018, see their Section 3.3). The null-weighted CCD images are used as the starting point for the source injection step in BALROG.
3. **Source injection into CCD images:** Synthetic sources are injected into all CCD images from Step 2. The catalog of injections we use is obtained from the DES deep-field data, and is described in Section 3.1. The injection models are ingested and rendered using a combination of the NGMIX (Sheldon 2015)

<sup>4</sup> To paraphrase the footnote from the original introduction of BALROG in Suchyta et al. (2016): “BALROG is not an acronym. The software was born out of the authors digging too greedily and too deep into their data, ergo the name.”

<sup>5</sup> PIFF is a more recent addition to the DES processing pipeline: it was not in the original pipeline of Morganson et al. (2018) but was later developed and integrated into the pipeline to improve the PSF modelling (which in turn improves the galaxy shape estimation needed for weak-lensing cosmology).

and GALSIM packages (Rowe et al. 2015). We convolve the galaxy models using the PIFF PSF (Jarvis et al. 2020), evaluated at the location of the injected source. See Schutt et al. (2025) for details on the Y6 PSF model. All galaxies are rotated randomly prior to injection. The source photometry is reddened using the average  $E(B - V)$  value of that given CCD image — obtained from the map of (Schlegel et al. 1998) — and the band-dependent  $R_b$  value for the *griz* filters.<sup>6</sup> This follows the same procedure done in BALROG from DES Y3. After this, the object’s light profile is added to the CCD image at the relevant location (see Section 3.2 for more details). In Y6, we also utilize multiple modes of source injection that aid different science goals within the Y6 cosmology analysis: see Section 3.2 and Section 3.3.

4. **Generate coadd image with SWARP:** Step 3 provides us a set of new “null-weighted” images that include real sources *and* the synthetic sources. These are passed into the SWARP (Bertin et al. 2002) routine to produce the coadd image. A coadd is generated for each band, and the coadds from the *riz* bands are further coadded together to obtain a “detection” coadd. This latter image is used for source detection. While the *g* band is simulated in BALROG and used in the measurements, it is not used for detection following the procedure done to the data.
5. **Source catalogs and MEDS files:** The detection coadd image from Step 4 is analyzed with SOURCEEXTRACTOR (Bertin & Arnouts 1996) to obtain the detection catalog for the given coadd tile. We run SOURCEEXTRACTOR in “dual mode” to generate photometric measurements in each of the *griz* bands for all detections. We then make Multi-Epoch Data Structure (MEDS; Jarvis et al. 2016) files, which consistently collate all images of an object (in small, square cutouts of at least  $32 \times 32$  pixels<sup>7</sup> or more) from all exposures/CCDs and all bands, for all detected objects. These MEDS files are used in generating almost all of the measurements for galaxy samples that we discuss; the notable exception is the METADETECTION sample, which requires a slightly different detection and MEDS-making procedure.
6. **FITVD catalogs:** The first catalog we generate — that serves as a general purpose photometric measurement

in DES, and also as a base sample from which we derive lens samples used in the galaxy clustering analysis — are the FITVD fits, named after the codebase used to derive them.<sup>8</sup> During fitting of a given object *A*, all pixels in its MEDS data are checked and we mask any pixel whose closest detection is not object *A*. This is done using the segmentation map provided by SOURCEEXTRACTOR and collected in the MEDS data. In DES Y3, we also had a multi-object fitting routine (MOF) that jointly fit groups of galaxies (rather than one galaxy at a time) to handle blending effects. This measurement was foregone in Y6 given it is computationally expensive, and was ultimately not used for cosmology analyses in Y3.

7. **Cell-based coadds:** In DES Y6, we utilize the METADETECTION algorithm (Sheldon et al. 2020, 2023) for estimating the shear of galaxies. This algorithm does not work on the coadds built in Step 4 and instead generates its own coadds, denoted as “cell-based” coadds. The latter are smaller than the coadd tiles we discuss here, and are constructed to have a continuous, well-defined PSF model. These cell-based coadds are generated using the PIZZACUTTER codebase.<sup>9</sup> See Yamamoto & Becker et al. (2025) for more details on METADETECTION and the cell-based coadds.
8. **METADETECTION source sample:** The METADETECTION estimator generates five versions of the cell-based coadds in step 7. The first is the original data, and the remaining four have an artificial shear applied to the full cell-based coadd in different directions, to empirically calibrate the response of object properties to external shear (e.g., Sheldon et al. 2020, 2023; Yamamoto & Becker et al. 2025). The SOURCEEXTRACTOR detection algorithm is run on all five images,<sup>10</sup> thereby providing five detection catalogs. We estimate shapes for all objects in each of the five catalogs, once again using the NGMIX package. This serves as one of our “source” samples, which is used for measuring weak lensing.
9. **Bayesian Fourier Domain (BFD) source sample:** BFD (Bernstein & Armstrong 2014) is a second estimator of the weak-lensing shear of a galaxy. For this method, we only need the MEDS files built in Step 5.

<sup>6</sup> We show below, in Section 4.1.3, that BALROG matches the data in capturing the spatial correlations between the galaxy catalogs and the dust extinction map.

<sup>7</sup> The pixel scale of a coadd image is, to good approximation, 0.263 arcseconds.

<sup>8</sup> <https://github.com/esheldon/fitvd>

<sup>9</sup> <https://github.com/beckermr/pizza-cutter>

<sup>10</sup> For the METADETECTION step alone, we match the choices of Yamamoto & Becker et al. (2025) and use the SEP (Barbary 2016) package, which is a python-based wrapper of SOURCEEXTRACTOR.

The BFD pipeline includes two steps: measuring the moments of the galaxy, and performing a bayesian integration of the moments to assign a shear estimate to the galaxy. We only perform the first step within the BALROG pipeline. See [Gatti & Wetzell et al. \(in prep.\)](#) for more details.

10. **Catalog matching:** Finally, for each type of catalog we generate — FITVD<sup>11</sup> (step 6) and METADETECTION (step 8) — we match the resulting output catalogs to the injection truth catalogs. For each synthetic source, injected at a given sky position  $x = \{\alpha_{2000}, \delta_{2000}\}$ , we find the nearest object in the output catalog. If the distance to the nearest object is within  $< 0.5''$  it is considered a successful match. More details on the matching process are also given in Section 3.4. This matching is done separately for the FITVD and METADETECTION catalogs. For METADETECTION this procedure is repeated for all five detection catalogs.
11. **MAGLIM++ lens sample:** The MAGLIM++ catalog ([Weaverdyck et al. in prep.](#)) is one of the lens samples<sup>12</sup> in the cosmology analysis and its selection requires a preliminary redshift estimate of the galaxies. Thus, we take all synthetic objects (after the postprocessing mentioned above) and we assign a redshift to each using the Directional Neighborhood Fitting (DNF) algorithm ([De Vicente et al. 2016](#)). The inputs to this are the magnitudes from the FITVD catalog. The MAGLIM++ sample is selected using the DNF redshifts as well as color selections performed using photometry from both FITVD and from measurements in WISE infrared survey ([Wright et al. 2010](#)). BALROG does not simulate the WISE photometry, and is instead limited to the DES photometric bands. However, the deep-field galaxies we inject have associated WISE photometry, so all wide-field realizations of a given deep-field galaxy are assigned the WISE photometry estimated for that deep-field galaxy. While this is an approximation, we find that the synthetic MAGLIM++ catalog still accurately reproduces the data. For example, it exhibits spatial correlations with survey properties (including the WISE magnitude limit) that are consistent with the data catalog (see Figure 11).
12. **REDMAGIC lens sample:** The REDMAGIC sample ([Rozo et al. 2016](#)) is a second lens sample for the

cosmology analysis, but is also used in constructing the redshift distributions of other galaxy samples (e.g., [Gatti & Giannini et al. 2022](#); [Giannini et al. 2024](#); [d’Assignies & Bernstein et al. in prep.](#)). In DES Y6, we use it for informing the redshifts distributions of the METADETECTION, BFD, and MAGLIM++ samples. The sample is constructed from color-based selections, in order to identify red galaxies whose redshifts can be estimated with higher accuracy. The inputs to the algorithm are the FITVD magnitudes (and magnitude errors) in addition to survey observing conditions, such as the sky background estimates mentioned in Section 5.

All pipelines for image processing and catalog generation are explicitly version-matched to follow those used in the processing of the actual data. Therefore BALROG uses the exact software as the data.

## 2.2. Differences from DESDM

Not all steps from the DESDM pipeline are reproduced in BALROG. We highlight here four relevant steps that are simplified in BALROG, and discuss the motivation for altering them.

1. **PSF estimation:** For all data processing of a given image, we use the PIFF model associated with a given CCD image (see [Schutt et al. \(2025\)](#) for details on the Y6 PSF model), and do not remeasure them after injecting our sources. This makes the approximation that the PSF model of the image is unchanged with/without the synthetic sources. Our choice follows the same approach used in all BALROG and image simulation pipelines within DES Y1 and Y3 ([S16](#); [E22](#); [MacCrann & Becker et al. 2022](#)), and we have no indication (from BALROG catalogs of previous years or from the catalog in Y6) that this approximation causes artifacts in the catalog. Thus, the same, fiducial PSF model is used for the SSI step and for all detection/measurement steps that follow.
2. **Astrometric solutions:** In the DESDM processing pipeline, all single-epoch exposures are fit with an astrometric solution. However, during the coadd processing, all CCD images belonging to a coadd tile (which necessarily includes images from different exposures) are jointly fit using SCAMP ([Bertin 2006](#)) to determine a new astrometric solution. In BALROG we take the existing SCAMP solutions of all images contributing to the coadd tile and do not run SCAMP again during our coaddition of the source-injected images. Similar to point (1), this approximation is used in all BALROG and image simulation pipelines, with no in-

<sup>11</sup> All other catalogs used in this work — BFD, MAGLIM++, REDMAGIC — use a subset of the detections present in the FITVD catalog and therefore have the same matching as the latter catalog.

<sup>12</sup> Following existing DES works, we use the term “lens sample” to denote a sample used for making galaxy clustering measurements.

dication of deficiencies in the data quality due to skipping this step.

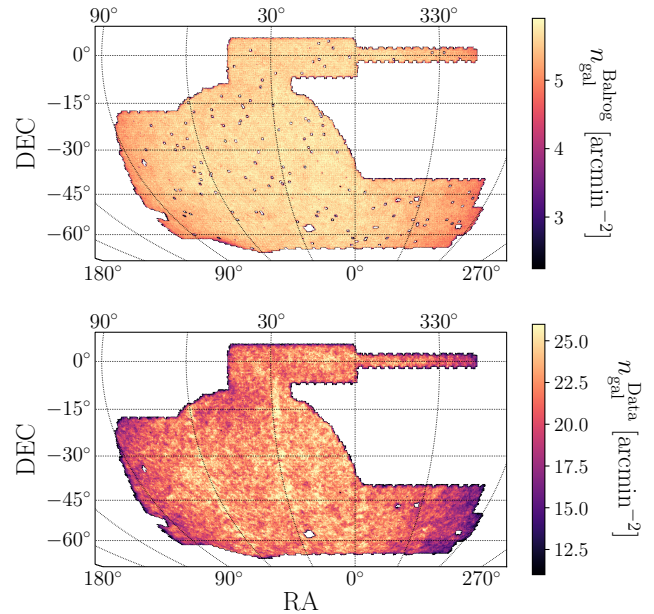
3. **Photometric calibration:** the photometric calibration (i.e., zeropoint of each CCD image) is not recomputed in the SSI pipeline. The value of the zeropoint is primarily sensitive to instrument response and atmospheric conditions (Burke et al. 2017), and not on the number density of sources. We use the zeropoints estimated for the original CCD images to compute the signal amplitude of injected sources, and therefore, to avoid circularity, we do not repeat the zeropoint calculation on the source-injected images.
4. **Reduced SOURCEEXTRACTOR parameter list:** Finally, we do not generate the exact same list of SOURCEEXTRACTOR properties for objects as is done in the data. For example, properties associated with the PSF are skipped. This choice is consistent with what is done in BALROG for DES Y3, and is motivated by selecting only those properties that are needed for the key DES cosmology science goals. The reduction in number of output properties greatly reduces the computational expense of the SOURCEEXTRACTOR step, and allows us to generate synthetic samples for a larger set of tiles (i.e., the full DES footprint), which in turn improves all DES science cases.

In general, BALROG does not redo any of the “single-epoch” processing and only performs the “multi-epoch” processing; see Figure 3 (and text therein) from Morganson et al. (2018) for more details on these two classes of pipelines.

### 3. BALROG IN DES YEAR 6

A number of updates have occurred in the BALROG pipeline between the DES Y3 analysis and the current Y6 analysis. These include changes to the source injection catalog (Section 3.1), the injection schemes (Section 3.2), and the specific runs (and variants) performed in BALROG (Section 3.3). We also highlight other smaller, but salient, differences in Section 3.4.

The runs were performed primarily at the FERMIGRID computing center at Fermilab, and also at the National Energy Research Science Computing (NERSC) system. The total computing cost of all BALROG runs was approximately eight million CPU hours, split between two million CPU hours for 7000 tiles run at Fermilab, and six million CPU hours for 5000 tiles run at NERSC. The differences in the required CPU hours per tile reflect the differences in the configurations of the two distributed computing systems. Some parts of the pipeline are not designed to use all available processors (for example, the processing pipeline uses a single-threaded version of SOURCEEXTRACTOR) and thus have CPU loads that can change depending on the system.



**Figure 1. Top:** the number density of all detections in the Y6 Balrog sample, using only high purity galaxies of the Gold catalog (defined with the star-galaxy classifier `EXT_MASH = 4`; see Bechtol et al. (2025, Section 4.2) for more details). The isolated, lower-density locations correspond to a (random) handful of failed tiles in the BALROG run, contributing  $< 1\%$  of area. **Bottom:** the same but for the actual DES Y6 data. The BALROG sample covers the entire DES imaging area. The density in Balrog is lower than data by construction, and is done to prevent blending of injected sources.

#### 3.1. Injection catalog

In Y3 BALROG, the injection catalog was obtained from Hartley & Choi et al. (2021) and contained FITVD models of galaxies detected in the DES Y3 deep fields. A similar, but slightly different, approach is done in Y6 BALROG: we continue using the same deep-field detections, but the galaxy models were re-fit using the latest version of the FITVD algorithm, to be consistent with the Y6 processing. We use the de-reddened fluxes of these objects during injection, where the de-reddening is performed using the same image-by-image correction as done in Y3 Balrog (E22, see their section 3.1.1). We show in Figure 5 that our synthetic galaxies show correlations with the extinction map that are consistent with that found in the data catalogs.

Our injection catalog is obtained by taking the catalog of Hartley & Choi et al. (2021) and then applying the following selections, which is the same as those motivated in E22 (see their Section 3.4):



```

in.VHS_footprint == 1
AND flags == 0
AND mask_flags == 0
AND SNR_griz > -3
AND avg_bdf_mag < 25.4
AND bdf_T < 100 arcsec2

```

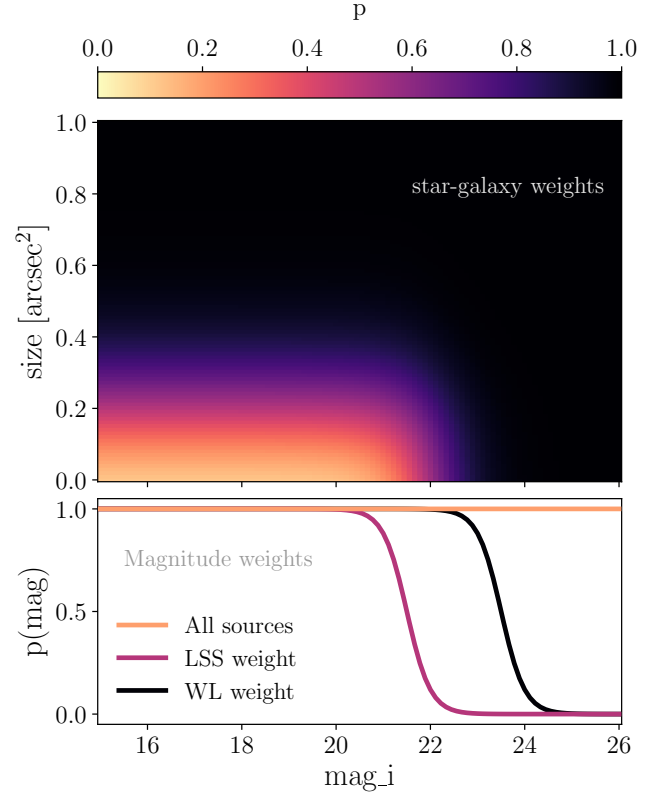
(1)

All selections above are made using quantities found in (or derived from) the DES Y3 deep-field catalog. The first three selections remove all objects with problematic fits or objects in regions of the deep fields that do not have the multi-band coverage required by some applications of BALROG (see Section 5.1). The SNR selection removes objects that were determined to be fit failures during an inspection done in the Y3 BALROG analysis. The `avg_bdf_mag` selection removes objects too faint to be detected in the DES Y6 sample. The quantity is computed by averaging the fluxes in the *riz* bands, and then converting this average flux to a magnitude. The selection of `avg_bdf_mag < 25.4` follows the Y3 BALROG analysis, where the chosen threshold is the magnitude of injected sources that are detected with 1% efficiency (E22). We have used the same choice in Y6 BALROG, even though the Y6 data has a fainter magnitude limit than the Y3 data, to further optimize the number of usable injections from BALROG Y6 relative to Y3. The selection criteria for the source and lens galaxy samples are at least one magnitude brighter than this limit (see histograms in Figure 9). Finally, the `bdf_T` selection removes large objects in order to reduce the blending of injected, synthetic sources with other synthetic sources. These selections follow exactly those used in E22 for Y3 BALROG.

For a subset of injections in Y6 BALROG, we apply a weighting to preferentially inject galaxies from the deep field rather than probable stars.<sup>13</sup> While BALROG will still contain injections of stars found in the deep field (as only a subset, and *not all*, of the injections are weighted in this manner), we reduce the ratio of stellar injections to galaxy injections by performing this selection, as the latter are the sole focus of all the cosmology analyses. In practice this weighting down-weights objects in an approximately rectangular region in the size-magnitude plane. These star weights are defined as,

$$\begin{aligned}
 w(m, T) = 1 - & \left[ \left( 0.9995 - \frac{0.9995}{1 + \exp\left[-\frac{T-0.2}{0.1}\right]} \right) \right. \\
 & \left. \times \left( 0.9995 - \frac{0.9995}{1 + \exp\left[-\frac{m-22}{0.5}\right]} \right) \right] \quad (2)
 \end{aligned}$$

<sup>13</sup> We use the terminology of “probable stars” to refer to the objects in the DES Y3 deep fields (Hartley & Choi et al. 2021) that were classified as stars using photometric colors (and *not* morphology).



**Figure 2.** **Top:** the star/galaxy weights of the deep-field objects as a function of their size and magnitude. The weighting reduces the number of small and bright objects that are injected, which are predominantly stars. **Bottom:** The magnitude-based weights for the LSS and WL samples we inject into BALROG. They are both sigmoid functions, with the transition centered at 21.5 and 23.5 in *i* band for the LSS and WL sample, respectively.

The weighting above removes objects with  $\text{size} \lesssim (0.2'')^2$  and  $\text{mag}_i \lesssim 22$ . Figure 2 shows the region of magnitude-size space that is down-weighted by this procedure. We apply this down-weighting for only a certain fraction of the injections in a tile. See Section 3.2 below for more details on the subsamples.

### 3.2. Injection and weighting scheme

BALROG runs are performed for a single coadd tile at a time. For a given tile/run, the injections are assigned to a regular grid spanning the  $0.7 \times 0.7 \text{ deg}^2$  region of the sky covered by the coadd tile (with a padding of  $10''$  from the edge). Once the sky location is defined, the source is injected into the relevant subsection of each CCD image that covers that region of the sky. Similar to DES Y3, the sky locations are a hexagonal grid (in pixel coordinates) with a separation of  $20''$ . This separation is chosen as we only inject sources with half light radii below  $10''$  (see Equation 1). Once the locations are defined, we randomly draw objects from the sample defined above in Section 3.1. A new choice of note

in Y6 BALROG is separating our source injections into five categories, in order to boost statistics for the objects that are most important for downstream DES Y6 cosmology analyses. In each coadd tile, we inject sources belonging to each of the five categories:

1. **All sources (Y3-scheme):** A third of all injections in a given tile are of objects defined in Section 3.1 above, through Equation (1). Note that this includes both galaxies *and stars*, i.e. we have not applied any star-galaxy selections in this sample.
2. **LSS sample:** A sixth of the injections are a weighted sampling of the objects. These down-weight all galaxies with magnitudes fainter than  $m_{i,\text{ref}} = 21.5$ , and also include the star-galaxy selection (see Figure 2). The magnitude weights are defined as,

$$w(m, m_{\text{ref}}) = 1 - \frac{1}{1 + \exp[-4(m - m_{\text{ref}})]} \quad (3)$$

and we apply them using the de-reddened  $i$ -band magnitudes of the deep-field objects.

3. **LSS high-quality redshifts sample:** Another sixth of the injections are the same as the LSS sample described above, but now with the additional requirement that the deep-field object being injected also has a redshift estimate from COSMOS (Laigle et al. 2016; Weaver et al. 2022), PAUS (Alarcon et al. 2021), or C3R2 (Masters et al. 2017, 2019; Stanford et al. 2021). These objects are more vital in the redshift estimation method used in DES Y3 (Myles & Alarcon et al. 2021; Giannini et al. 2024) and Y6 (Yin et al. in prep.; Anbagane & Cross et al. in prep.).
4. **WL sample:** Another sixth is once again the same as the LSS sample (including the star-galaxy selections), but with the magnitude-based weights moving to higher magnitudes as the weak lensing sample is fainter; here, we down-weight all objects fainter than  $m_{i,\text{ref}} = 23.5$ .
5. **WL high-quality redshifts sample:** The final sixth is similar to the LSS high-quality redshifts sample, but now tuned to the weak lensing sample by moving the magnitude-based weights to higher magnitude as mentioned above.

Figure 2 presents the different weights we use. The top panel shows the star-galaxy weights used to remove the stellar locus, and the bottom panel shows the different magnitude-based weights. The magnitude weights are sigmoid functions centered at values detailed above. This weighting scheme greatly increases the number of usable

injections — defined as injections that are actually detected in a given sample/method — in Y6 BALROG and thereby improves the statistical power of all analyses using BALROG. The number of selected sources in Y6 BALROG is increased, relative to what would be obtained using the “Y3-scheme” alone, by factors of 12, 2.7 and 1.6 for the LSS sample (MAGLIM++, REDMAGIC), WL sample (METADETECTION, BFD), and FITVD sample (with no selections on magnitude and color) respectively. All estimates come from comparing the number counts of selected/detected objects from the “Y3-scheme” to those from the other four weighting schemes (while accounting for differences in the number of injections assigned under each scheme).

We note that the magnitude cutoffs defined above for the LSS and WL samples are slightly brighter relative to the actual magnitude distributions of the real samples (Legnani et al. in prep.; Yamamoto & Becker et al. 2025; Gatti & Wetzel et al. in prep.), as our cutoff choice was defined prior to the availability of the Y6 sample definitions. However, we stress that the BALROG synthetic catalog — given the unweighted “Y3-scheme” subset of the catalog — still contains objects across the full magnitude range relevant for all the Y6 cosmology samples. Any difference in our chosen magnitude cutoff and the true magnitude distribution of the Y6 samples will only result in a potential sub-optimality (i.e. increased noise) of measurements derived from BALROG and *not* in biases in these measurements.

### 3.3. Fiducial run and magnification run

In Y3 BALROG, a number of different variations were run, each serving a slightly different role in the main DES analysis (E22, see their Table 1). In this work, we perform two variations: a fiducial run, which injects the sources described in Section 3.1, and a magnification run, where we inject the same sources but with their fluxes and area increased by 2%. Note that all random seeds (during both the injection step and the subsequent postprocessing of images to catalogs) are fixed between the two runs. The magnification run is necessary for the Y6 cosmology analysis, as it enables estimates of the galaxy magnification coefficients for the specific samples used in DES Y6; for example, see Elvin-Poole & MacCrann et al. (2023) for how BALROG was used to do this in DES Y3, and also Section 5.2 below. These magnification coefficients are folded into both the cosmological inference of galaxy clustering (Pandey et al. 2022; Prat et al. 2022; DES Collaboration 2022) and the redshift calibration analysis (Gatti & Giannini et al. 2022; Giannini et al. 2024; d’Assignies & Bernstein et al. in prep.).

### 3.4. Improvements relative to BALROG in DES Year 3

**Adopting the EASTLAKE pipeline:** A key (logistical) difference from Y3 is the unification of our BALROG pipeline

with that used for the image simulations (Mau & Becker et al. in prep.). Both BALROG and the image simulations require source injection, though with some significant differences in final analysis choices: for example, the image simulations generate fully simulated images, by adding together the synthetic galaxies with simulated noise, star sources, data-based background and mask etc. Whereas in BALROG we inject galaxies (and probable stars) directly into the real images and therefore do not need to simulate any additional image properties. However, the vast majority of the processing steps are common between the two pipelines, and thus in DES Y6 both pipelines derive from the EASTLAKE codebase which performs most of the image processing steps discussed in Section 2.1. The specific BALROG injection (Step 3) is performed outside EASTLAKE, as are the BFD measurements (Step 9) and matching (Step 10). All remaining steps associated with image processing (Step 1 to 8) are performed using EASTLAKE and the packages within, and are therefore *identical* between this work and the image simulations used for shear calibration in DES Y6 (Mau & Becker et al. in prep.).

**Moving to PIFF PSF models:** In Y6 BALROG, our multi-epoch multi-band object fits use the PIFF models (Jarvis et al. 2020) rather than following Y3 BALROG in using the PSFEX models (Bertin 2013). This improves the consistency of our synthetic catalog for the two source samples, METADETECTION and BFD, and leaves a minor mismatch with the FITVD catalog in Y6 (Bechtol et al. 2025) as the latter is constructed using PSFEX models.

**Identification of blended objects:** In Y3 BALROG, the matching step included an additional component — beyond that described in step 10 of Section 2.1 — to identify BALROG injections that were blended with a real source in the image. In this step, an injected source was considered problematic only if it was blended *and* it was the fainter object in the blended set. In DES Y6, we identify all synthetic detections that have a real object within  $1.5''$  of their location. All such synthetic objects are flagged in our catalogs, regardless of whether or not they are the brighter source in the blended group. This is a more conservative selection relative to DES Y3. Such objects constitute only  $< 0.2\%$  of the entire catalog, and are removed from all analyses that use BALROG, including all those presented below.

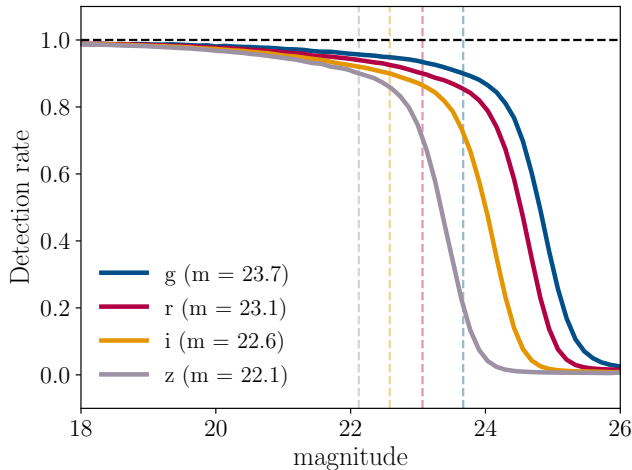
**Weight scheme and star catalog:** As described above, the weighting scheme used in Y6 was not part of the DES Y3 analysis. Under this new scheme, a subset of the injections (1/3rd of all injections) consists of any real, detected object (i.e., objects that are not artifacts) in the deep fields. This includes stars. Thus, in Y6 we have used real stars as injection sources rather than a simulated star sample as was done in Y3. At brighter magnitudes this enhances the realism of the

stellar data as we are directly injecting stars from the deep field. However, at fainter magnitudes this means we rely on the accuracy of the data-driven, color-based star-galaxy classifier of Hartley & Choi et al. (2021) to determine whether a given deep-field object is a faint star or a faint galaxy. Hartley & Choi et al. (2021, see their Figure 15) show that the star sample is  $> 95\%$  pure down to  $\text{mag}_i < 24$ . Additionally the color/magnitude distribution of the injected stars will, by construction, follow that found in the Y3 deepfields whereas in reality this distribution varies across the DES footprint.

**BALROG for the full DES footprint:** In DES Y3, the BALROG sample was restricted to around 2500 tiles due to computational limitations. In DES Y6, this is increased 5 fold, with over 12000 tiles being run: 10,169 tiles for the fiducial run, and 2,000 for the magnification run. The full coverage of the survey enables new science cases, particularly any area-dependent studies of all BALROG-related science, such as magnification, transfer functions, imaging systematics, etc. Figure 1 shows the area covered by BALROG in comparison to the DES Y6 data. There is already visual indication that BALROG traces the large scale density fluctuations across the survey (due to variations in depth and other observing conditions);<sup>14</sup> for example, the galaxy number density in both BALROG and data is lower near the east and west (left and right) edges of the footprint, which are regions of high stellar density at lower Galactic latitude. These regions have a lower effective area due to masking around bright stars, and are regions where objects are more frequently observed as blended sources. Other correlations with observing conditions are noted below in Section 4.1.3. Note that these large-scale fluctuations in the number density of BALROG objects are not generated by hand, and manifest solely from the interaction of image properties (such as noise, sky background, source crowding, etc.) and the object detection/measurement algorithms.

Finally, we make a brief note on the connection between BALROG and the image simulations (Mau & Becker et al. in prep.) used in DES Y6. BALROG injects sources somewhat sparsely (i.e., with  $20''$  separations) into real on-sky images, and therefore accurately captures the relationship between complexities in the image data and the measured photometric/morphological properties of detected objects. This relationship is a necessary input to many application in DES Y6, which are detailed in Section 5. The image simulations, on the other hand, inject at least an order of magnitude more

<sup>14</sup> The amplitude of fluctuations is suppressed because the weighted injection scheme in Y6 Balrog (see Section 3.2) results in the majority of injected objects being above the object detection threshold. The fluctuation amplitude in Figure 1 will increase if we limit the sample to the unweighted injections alone.



**Figure 3.** The detection rate of objects as a function of their magnitude. Objects are detected if they have  $\text{SNR} = \text{flux}/\text{flux\_err} > 5$  in order to avoid any spurious detections. Vertical lines denote the magnitudes at which we achieve 90% completeness, and the numerical values are given in the legend. This rate is measured using BALROG objects across the full Y6 footprint.

sources than BALROG, and do so into fully simulated images rather than the real images. These simulated images are then made sufficiently realistic by adding simulated noise, star injections, image masks, sky backgrounds, etc. These simulations are used for characterizing any biases in the galaxy shape measurements from BFD or METADETECTION. Given the biases are at the 1% level or below, a large number of injections (relative to BALROG) are needed to precisely measure this bias.

#### 4. VALIDATION OF THE SYNTHETIC SAMPLE

As mentioned previously (and detailed in Section 5), the BALROG data products are used extensively to meet many different DES science requirements. In this section, we highlight both the quality of the BALROG catalog — checking its consistency with data — to justify its use in the DES Y6 cosmology pipeline, and also the science performance of algorithms involved in going from raw images to robust object catalogs. We first highlight in Section 4.1 the consistency of our catalog with DES data, then in Section 4.2 and 4.3 the performance of the photometric pipeline for the fiducial sample and star sample, respectively. Finally in Section 4.4, we detail the performance of the different star-galaxy classifications used in defining the various samples used in the DES Y6 cosmology analyses.

##### 4.1. Consistency with DES Data

###### 4.1.1. Completeness

Figure 3 shows the detection probability of a galaxy, as a function of its *griz* magnitudes. As expected, the rate nears 1 for bright objects and drops to 0 for faint objects. The vertical

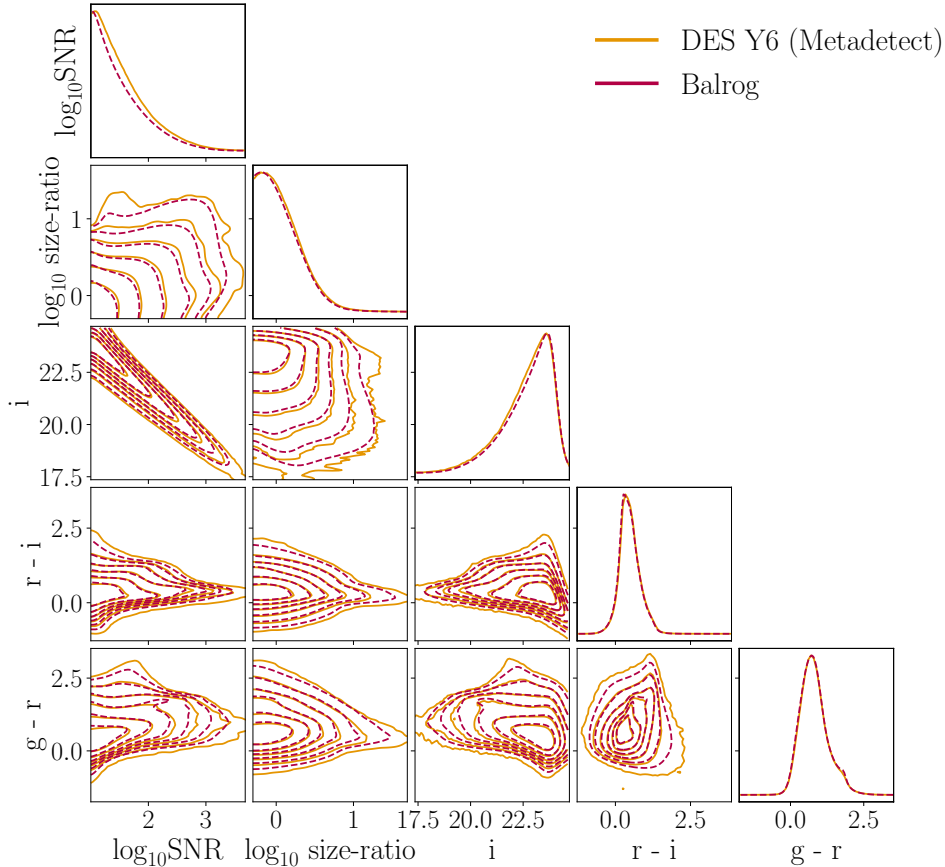
dashed lines show the magnitude where we have 90% completeness of objects with  $\text{SNR} = \text{flux}/\text{flux\_err} > 5$ . We impose a signal-to-noise selection in each band to remove any spurious false-positive BALROG detections due to matching injected sources with noise fluctuations; this is relevant only near the faintest end ( $m \geq 25.5$ ) and particularly for the *g*-band as it is not used in the detection step. When considering all objects with magnitudes above the threshold set by the 90% completeness rate, more than 99.9% of the sample have  $\text{flux}/\text{flux\_err} > 5$ . For objects at the 90% completeness rate we find  $\langle \text{flux}/\text{flux\_err} \rangle \approx 20$ .

While our test of completeness does not include direct comparison to the Y6 data — unlike in the Y3 analysis, where a data-based estimate of completeness was done by matching the deep-field detections to their wide-field counterparts — this measurement shows us that BALROG produces physically reasonable completeness rates for the extreme ends of bright and faint objects, which in itself is a test of consistency. Estimates of the improvement in depth from DES Y3 to DES Y6 are presented alongside the Y6 Gold Catalog (Bechtol et al. 2025) — and the DES DR2 (DES Collaboration et al. 2021) — and we find consistent improvements in BALROG. Table 1 below details one such comparison.

###### 4.1.2. Photometry and morphology

A key consistency check is ensuring the distribution of measured properties in a given BALROG catalog matches that of the corresponding data catalog. Figure 4 shows a range of properties of the METADETECTION sample from BALROG and data; we have shown only properties for the unsheared sample here, but the remaining four samples show similar consistency as the unsheared sample. In general, the data and BALROG distributions match across  $\approx 99\%$  of the data volume. Note that we expect some variation, as an implicit assumption in BALROG is that the deep-field galaxies are a representative sample of the wide-field galaxies, whereas in reality, simple factors like sample variance (arising from the small footprint of the deep fields) will break this assumption slightly. Figure 4 shows such factors are likely a minor contributor given that the data and BALROG catalogs match for 99% of the data volume. The distributions are similarly consistent when repeating this check for other galaxy samples (BFD, MAGLIM++, and REDMAGIC). We do not show those comparisons here for brevity.

Note that the above measurements are performed using only the unweighted subsample in BALROG (the “Y3-scheme” defined in Section 3.2). Adding morphology/photometry-based weighting to the deep-field sample during source injection will generate large differences between the injected sample and the actual data. For example, under the LSS weights any galaxy with  $m_i \gtrsim 23$  is never injected, even though the weak lensing sample con-



**Figure 4.** A corner plot of the distributions of galaxy properties from the METADETECTION catalog. The contours show the  $0.5\sigma$  to  $3\sigma$  bounds, in increments of 0.5. The Balrog density contours agree well with the real data except for the rarest ( $3\sigma$ ) sources’ distribution. Thus, nearly 99% of the catalog is consistent between BALROG and data. The properties shown are the signal-to-noise, the ratio of object size and PSF size, the  $i$ -band magnitude, and the  $r - i$  and  $g - r$  colors.

tains objects down to  $m_i \sim 24.5$ . Thus, comparing the distribution of properties in data while using the weighted subsample in BALROG will naturally lead to discrepancies in the comparisons. While this could be handled by “inverse weighting” the different, weighted distributions, we opt for the simpler approach of just using the unweighted “Y3-scheme” sub-sample.

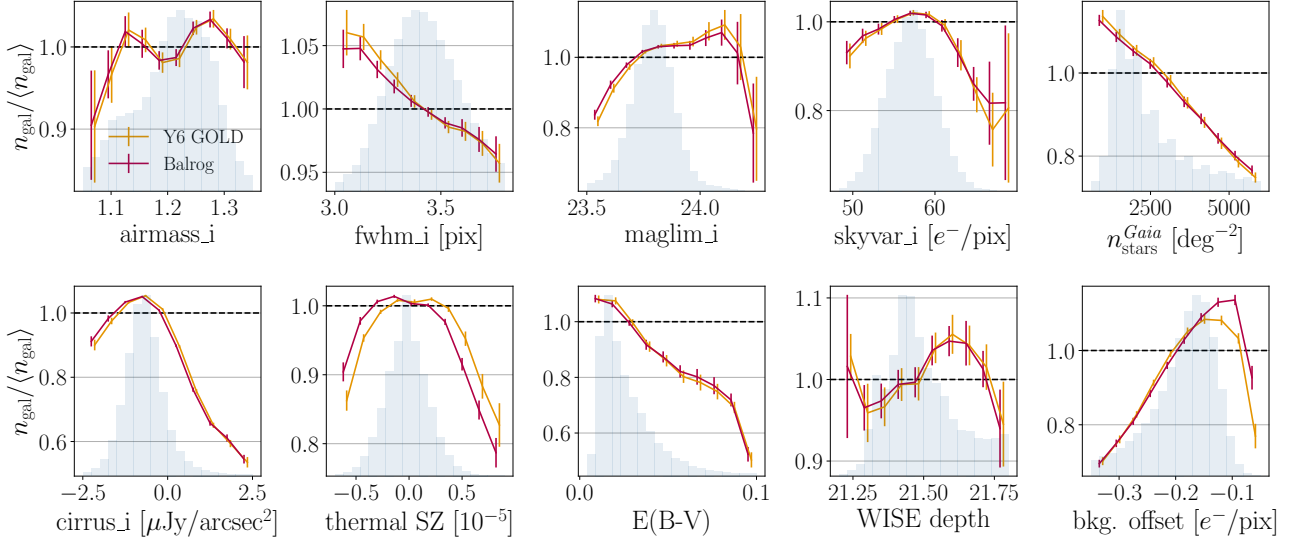
#### 4.1.3. Spatial Variation and Property Maps

The detection and characterization of an object on the sky will depend on the observing conditions of the processed images, and on other non-cosmological fields observed in the images, such as Galactic dust or stellar density (Ho et al. 2012; Ross et al. 2012; Leistedt et al. 2013; Rykoff et al. 2015). We represent variations of these effects across the DES footprint with “survey property” (SP) maps (Bechtol et al. 2025; Rodríguez-Monroy & Weaverdyck et al. in prep.). The BALROG catalogs will be a more accurate representation of the data if they also capture the dependence of object properties on such survey property maps. This correlation of galaxy properties with observing conditions is a key source of systematic uncertainty to be addressed for measure-

ments of the cosmological clustering of galaxies (Rodríguez-Monroy et al. 2022; Kong et al. 2024), and thus reproducing it in BALROG is of importance.

We compute the correlation between galaxy number counts and survey properties in BALROG and in the Y6 Gold sample, and compare the two. At zeroth order, survey property maps alter the measured properties of galaxies in our images, and this will lead to variations in the number of detected/selected galaxies in different regions of the sky. For example, regions of high airmass have larger PSFs and can lead to fewer faint/small galaxies being observed, which in turn will reduce the overall galaxy number density in such regions. We measure these galaxy–SP map correlations by making maps of the galaxy number counts,  $N_{\text{gal}}$ , and measuring the average  $N_{\text{gal}}$  as a function of SP map value. We use HEALPIX (Górski et al. 2005) maps of  $N_{\text{SIDE}} = 512$ , but note that other choices result in similar agreement between BALROG and data.

For a given SP map, we assign healpixels into 20 equal-size bins, spanning the 1-99 percentile range of the map values in order to avoid outliers. Within each of these 20 bins,



**Figure 5.** The trend in number density fluctuations  $n_{\text{gal}}/\langle n_{\text{gal}} \rangle$  as a function of various survey observing properties, for the Balrog (red) and Y6 Gold (yellow) samples for overlapping HEALPix pixels of  $N_{\text{SIDE}}=512$ . We apply basic selections, such as star-galaxy classifications, to both samples. The distribution of survey property values for the Y6 Gold map is shown in the background (blue) of each plot to highlight typical values. The errors have been estimated using a jackknife resampling of the pixels with 100 patches. We select only a few properties to show in this figure and, for those maps with different values on the *griz* photometric bands, we only show the *i*-band. Briefly defined, these survey properties are: the airmass, the PSF size, the magnitude limit at  $S/N = 10$ , the sky brightness variation, the *Gaia* stellar density, the Galactic cirrus nebulosity, the Sunyaev Zel’dovich effect, the interstellar extinction, the WISE depth, and the background offset. The  $\chi^2$  between the data and the BALROG measurements follows  $\chi^2/N_{\text{dof}} \lesssim 1$  for all cases except the correlations with the thermal SZ and the background offset, where  $\chi^2/N_{\text{dof}} \gg 1$ . This is expected, due to the impact of large-scale structure on these maps and we detail this in Section 4.1.3.

we compute the mean galaxy number density. This process is repeated for both the BALROG and Y6 Gold galaxy samples. We use the FITVD catalogs, with `EXT_MASH = 4`, so that our sample only includes galaxies that could be part of the lens sample used in the cosmology analysis. In Appendix A, we also show results from repeating this test but now using the full selection function of the MAGLIM++ and REDMAGIC samples. The uncertainty on the measurements is estimated by splitting the DES footprint into 100 patches and performing a leave-one-out jackknife resampling.

We use SP maps spanning survey observing conditions, maps of Galactic properties, and maps of other large-scale structure probes. In particular, our observing conditions include the airmass, full-width half-max of the PSF, the magnitude limit (or depth), the variation in the sky background, and the background offset. These are measured using DECASU<sup>15</sup>. We only present the correlations with the *i*-band maps but our results and conclusions are the same for the other bands as well. We then use the depth map in the WISE survey (Wright et al. 2010), as the MAGLIM++ sample is defined using photometry from the UNWISE program (Meisner et al. 2017), which includes data from the WISE survey. We then use the Galactic stellar density, estimated through *Gaia* (Gaia Collaboration et al. 2023), the Galactic cirrus (Bechtol et al.

2025), and the extinction  $E(B - V)$  map from Schlegel et al. (1998). Finally, we also include the Sunyaev-Zeldovich (SZ) map from the Atacama Cosmology Telescope (ACT, Fowler et al. 2007; Thornton et al. 2016), which is a tracer of the large-scale structure, and contains emissions from galaxies over a wide range of redshifts (see Carlstrom et al. 2002, for a review of the SZ effect).<sup>16</sup> We note that the background offset map mentioned prior is also a potential tracer of large-scale structure as the excess presence of (unresolved) galaxies can increase the mean sky background and lead to larger background offsets.

Together, these maps span a wide variety of quantities that can impact measurements of galaxy properties (or correlate with them) in different ways. Figure 5 shows the mean galaxy number density as a function of SP map value. The background of each panel (light gray histogram) shows the distribution of survey property values across the DES Y6 footprint. The galaxy-SP correlations in BALROG are statistically consistent ( $\chi^2/N_{\text{dof}} \lesssim 1$ ) with those present in the data. The two exceptions are the correlations with the ACT SZ map and the background offset map. However, these two exceptions are in fact expected. Both maps are tracers of

<sup>15</sup> <https://github.com/erykoff/decasu>

<sup>16</sup> This map is not used as a systematics map but rather as a tracer of large-scale structure (not measured in DES) for validating the systematics correction of the lens samples; see Weaverdyck et al. (in prep.) for more details.

large-scale structure as mentioned above. The correlations between the Y6 data and these maps will therefore contain a *cosmological* correlation. Such a cosmological term is not present in BALROG by construction as the synthetic galaxies are distributed uniformly across the coadd tile. Thus, there will be an expected difference in the galaxy–SP correlations between BALROG and the Y6 data for these maps.

There is also a correlation between the galaxy counts and the WISE depth (for both BALROG and data). The galaxy sample here does not use WISE data for its selection, and therefore any correlation observed is a consequence of the WISE depth being correlated with other observing conditions that *do* impact the galaxy sample used in this analysis.

We note that Kong et al. (2024, see their Figure 12 and 13), who also perform SSI in DECam images — but to study the targets selected for the DESI survey — find significant differences in the correlations of SSI galaxies with the extinction map compared to that of real galaxies with this map. However, these differences are prominent in the northern Galactic hemisphere and not present in the southern hemisphere, where the DES footprint is located. They suggest this qualitative difference across areas is also related to how their imaging data in the northern half was observed using a depth-based strategy that could lead to residual impact from extinction corrections (see their Section 7.2). The same is not present in the southern half, as that region is dominated by DES data which does not follow this observing strategy. They also suggest this difference between the north/south regions highlights systematics in the northern regions of the dust extinction map of Schlegel et al. (1998). As shown above, our work does not find any significant discrepancies in the correlations of BALROG galaxies with the extinction map. This is also true for the individual lens samples, which we detail in Appendix A.

#### 4.2. Photometric Performance of Galaxy sample

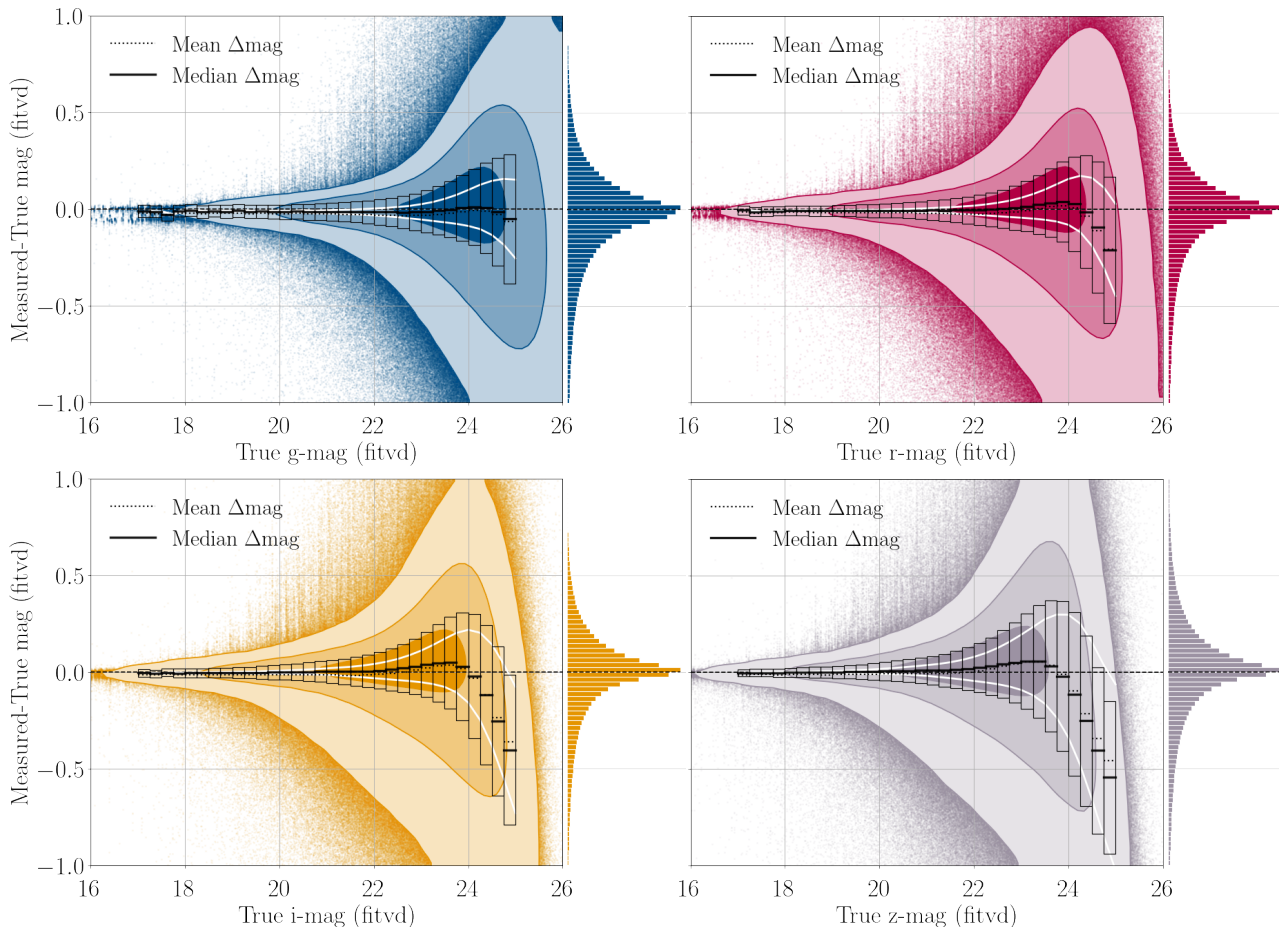
The DESDM processing pipeline is multi-faceted, and includes many different steps (each validated and tested) that are integrated to perform the entire image processing procedure; see Morganson et al. (2018) for details of the pipeline used in DES, and DES Collaboration et al. (2021); Bechtol et al. (2025) for modifications made specifically in DES Y6. BALROG provides a precise test of the pipeline’s performance by comparing the properties of injected objects to their measured properties in noisy images. We perform this test with the Y6 dataset, using the FITVD measurements of the galaxies. The deep-field galaxies also have FITVD measurements made using the same algorithm, and therefore can be used to precisely quantify any biases in the wide-field measurements.

Figure 6 compares the differences between the measured and true deep-field magnitudes  $\Delta\text{mag}_{\text{DF}}$  as a function of

true magnitude in the *griz* bands. The density contours include 39.3%, 86.5%, and 98.9% of the population, which correspond to the 1, 2, and  $3\sigma$  regions of a 2D Gaussian. The mean bias, median bias, and standard deviation of the differences are over-plotted as a dotted line, solid line, and black bars, respectively. We compute these in bins of 0.25 magnitudes. The standard deviation is computed using the 16% and 84% percentiles to avoid biases from outliers. The median and standard deviation of the magnitude differences are tabulated in Table 2. The white lines for each also show the  $1\sigma$  magnitude region — as computed by the FITVD algorithm — around the median difference. Similar to the results of BALROG in DES Y3 (Everett et al. 2022, see their Figure 16), the estimated errors are smaller relative to the scatter in the measured magnitudes of the BALROG synthetic sources (the black, vertical bars in Figure 6). The BALROG-based estimate of this error is generally expected to be larger as it includes systematic effects like the variations across different observing conditions, which is not accounted for in the internal, FITVD-based error estimates. The differences between the BALROG-based and FITVD-based error estimates are larger for brighter objects, where such systematic effects are the dominant contributor to the magnitude differences, and are smaller for fainter objects where noise fluctuations are the dominant contributor.

The tails to negative magnitude offsets in Figure 6 was also found in E22 (see their Section 4.3.3), and can arise from a number of different causes, which include excessive pixel masking, residual light from neighboring sources etc. Their Figures 8 and 19 show the same asymmetric tails, at faint magnitudes, that we find in Figure 6. We also note that some negative bias (i.e. measured magnitudes are brighter than the true magnitudes, on average) is expected for faint objects due to the Malmquist bias produced by applying a signal-to-noise selection during object detection. If we consider multiple wide-field realizations of a faint object, we preferentially detect realizations where the noise fluctuations boost the measured brightness. This is a natural consequence from using noisy data, and does not indicate problems with the image processing algorithms.

The estimates in Figure 6 are made using galaxies that pass all the Y6 Gold selection criteria as well as the star-galaxy selection,  $\text{EXT\_MASH} \geq 4$  (Bechtol et al. 2025). This is similar to, but still strictly distinct from, the galaxy samples used for the various cosmology analyses in DES Y6. Our estimates (see also Table 2) show we recover the true magnitudes to within 3-6 millimag (below  $m < 20$ ) for the *riz* bands, and within 12 millimag for *g*-band. This is generally larger than, but the same order as, the relative photometric uniformity which is at  $\lesssim 2$  millimag (DES Collaboration et al. 2021) and also smaller than the 11 mmag uncertainty on the absolute photometric calibration (DES Collaboration



**Figure 6.** The distribution of differences in recovered FITVD  $griz$  magnitude vs the injected magnitude ( $\Delta\text{mag}_{\text{DF}}$ ) as a function of input magnitude for the galaxy sample. The contour lines correspond to areas containing 39.2%, 86.5%, and 98.9% of the data volume, which are the 1, 2, 3 $\sigma$  bounds of a 2D Gaussian, respectively. The mean (dotted), median (solid), and standard deviation (bars) of the magnitude responses, in bins of size 0.25 magnitude, are also shown. The white lines are the average magnitude error of objects in a given bin, where the error is estimated during FITVD fitting. There is some discretization in true magnitude (particularly in the brighter end) given we have only a small number of bright objects in the deep-field sample.

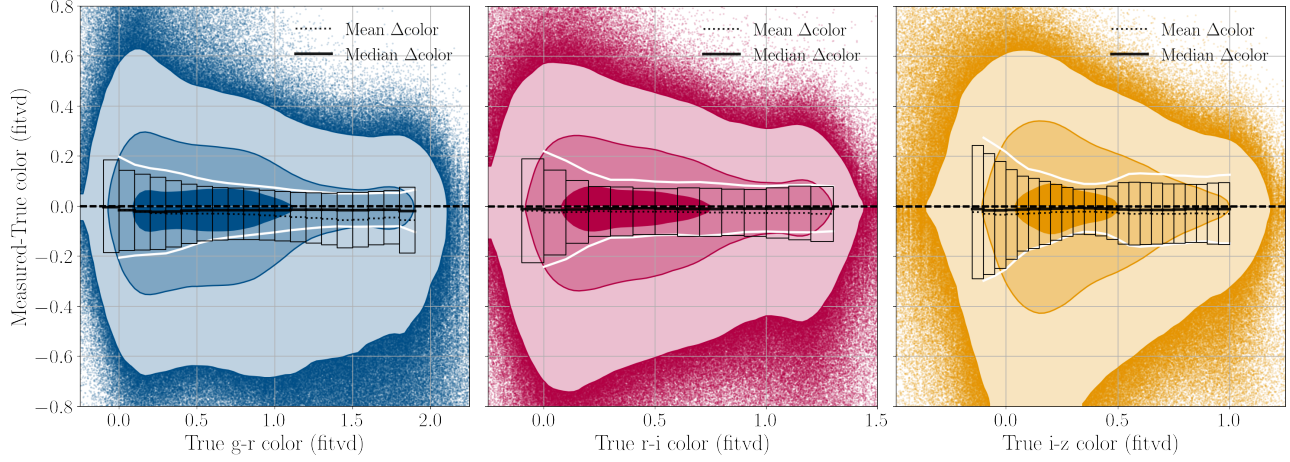
et al. 2021). For fainter objects ( $m > 23$ ), we find larger biases of up to 50 millimag that increase to  $\approx 500$  millimag at  $m \sim 25$ . We note that for  $m < 24$ , the median difference is always well below the standard-deviation. The differences we discuss here were also found in E22 (see their Figure 16), where they show that the bias grows increasingly negative for fainter galaxies. However, the Y6 dataset extends to fainter objects due to the increased depth, and also reduced the signal-to-noise threshold for detection from  $S/N = 10$  to  $S/N = 3$  (Bechtol et al. 2025), relative to Y3 and we therefore observe stronger biases than was found in their work.

E22 (see their Section 4.3.1) discuss how this bias originates from biases in the sky-background estimates of the SOURCEEXTRACTOR algorithm (which are then propagated to the rest of the galaxy catalog processing). Such a bias grows in the presence of more crowded fields, and crowding increases towards the redder bands, resulting in the chromatic effect that we observe. We stress that the source and

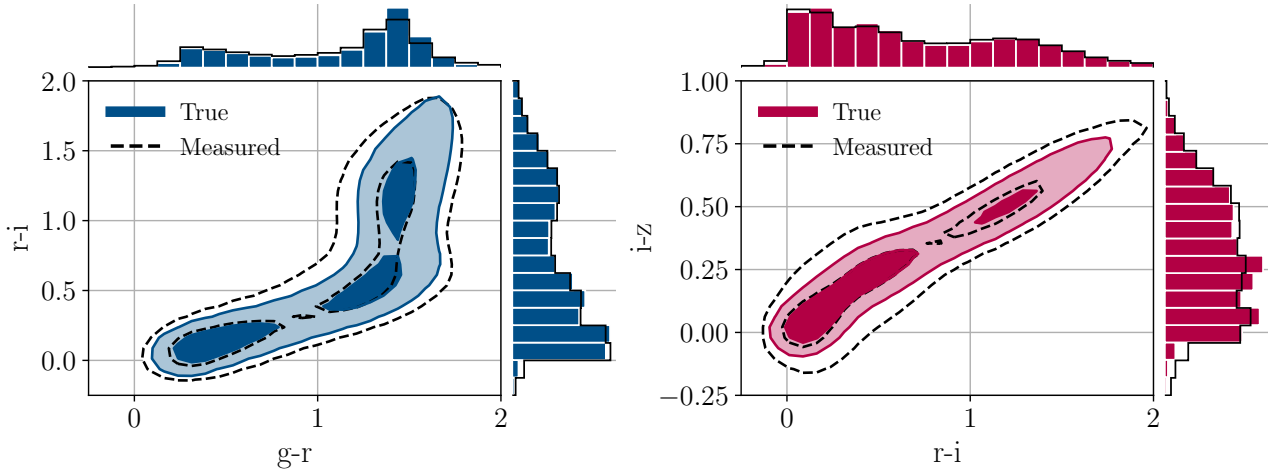
lens galaxy samples used in DES Y6 occupy a magnitude range of  $\text{mag}_{\text{g,i}} \lesssim 24$  (see Figure 9 below) — in addition, once weights are included, the sample is dominated by even brighter objects — and these effects are not catastrophic in this magnitude range. Malmquist bias can also explain the relatively low bias in the  $g$ -band related to the  $riz$ -bands, as the former is not used in object detection and is therefore akin to a forced photometry measurement, which will naturally avoid Malmquist bias. Similarly the Malmquist bias is expected to be stronger in Y6, relative to Y3, due to the lower signal-to-noise threshold.

We also investigate the photometric accuracy of the object colors. Such colors are used extensively in applications such as photometric calibration, star-galaxy classification, redshift estimation, and studies of the Milky Way structure/dynamics (e.g., Burke et al. 2017; Salvato et al. 2018; Myles & Alarcon et al. 2021; Giannini et al. 2024; De Vicente et al. 2016; Bechtol et al. 2025; De Vicente et al. 2016; Shipp et al. 2018;





**Figure 7.** Similar to Figure 6 but for the differences in measured FITVD  $g-r$ ,  $r-i$ , and  $i-z$  color. The statistics are computed in bins of 100 mmag magnitude for  $g-r$  and  $r-i$  and 50 mmag for  $i-z$ . The median bias in the recovered colors is below 0.02. The FITVD-based uncertainty on the colors, which is computed by adding the individual magnitude uncertainties in quadrature, is in good agreement with the BALROG-based estimate.



**Figure 8.** The  $g-r$  vs  $r-i$  and  $r-i$  vs  $i-z$  color-color distributions of the stellar sample for the injected colors in blue and the measured colors in black. The density contour lines correspond to the percentiles of the first two sigmas of a 2D Gaussian, containing 39.2%, 86.5% of the data volume respectively. The marginal distributions are included for comparison.

Pieres et al. 2020). Figure 7 shows the difference in measured and true colors for  $g-r$ ,  $r-i$ , and  $i-z$ . The density contours and over-plotted summary statistics are defined in the same way as in Figure 6. One additional selection is done to this sample: a signal-to-noise of greater than 5, where this signal-to-noise is directly estimated during the FITVD fluxes and flux errors, to remove color measurements from particularly noisy detections. Our results show that the difference in colors is consistently negative, suggesting a small bias for the recovered colors to be bluer than the true colors. However, this difference is always within 10-20 millimag, and also similar across all three colors we present, indicating there are no band-dependent effects. This behavior is also found in E22 (see their Figure 18).

In conclusion, we find that the photometric accuracy of the DES Y6 dataset is similar to that of the Y3 dataset. The accuracy in recovering the true magnitudes degrades for fainter objects (particularly those objects that are too faint to be found in DES Y3 but are now detected in the deeper, DES Y6 data), as is expected. While we also discuss the faint regime, where the photometric recovery shows growing biases, all the main source and lens galaxy samples used in the DES Y6 cosmology analysis occupy the brighter end of the magnitude range in Figure 6, which show fluxes are recovered at the 2% level.

#### 4.3. Photometric Performance of star sample

The image processing pipeline — in particular, the fitting of models to objects — can show different accuracies when

processing galaxies or stars, given the differences in morphology of the categories. We repeat the analyses above using only the stars injected into BALROG; these are determined through the color-based classifier built and presented in Hartley & Choi et al. (2021). Their work (see their Figure 15) shows that the star sample is  $> 95\%$  pure down to  $\text{mag}_i < 24$ . For brevity we present the results in Appendix B but summarize here that photometric measurements of stars are accurate to a similar level as seen in the galaxy sample (Section 4.2). Table 2 lists the photometric accuracy for the galaxy and star samples, and shows that the median magnitude differences are similar for the two samples but the standard deviations of these differences is roughly two times smaller for the star sample.

In addition, Figure 8 shows the color-color diagrams for  $g-r$  vs  $r-i$ , and  $r-i$  vs  $i-z$  colors, as measured in FITVD, for both the true and measured star samples. As expected, the measured colors have a broader distribution due to being noisy measures, but the non-linear shape of the 2D distribution is still preserved by the image processing pipeline.

#### 4.4. Contamination & efficiency of the star-galaxy classification

The DES Y6 analyses contain many different galaxy samples: the METADETECTION and BFD source samples, and the MAGLIM++ and REDMAGIC lens samples. Through BALROG we can quantify the contamination and efficiency of the star-galaxy classification as defined in the different sample selections. To do so, we estimate the “galaxy efficiency” defined as the fraction of true galaxies in the selected sample relative to *all detected* true galaxies, and we estimate the “stellar contamination”, defined as the fraction of all selected objects that are true stars. We classify BALROG injections into galaxies and stars based on the color-based classifier of Hartley & Choi et al. (2021). The stellar contamination rate depends on the ratio of stars and galaxies in the truth catalogs. Therefore all our results below are valid for ratios of stars to galaxies similar to those found in the DES deep fields (Hartley & Choi et al. 2021). Estimates of stellar contamination from different subsets of a catalog can vary slightly from each other; for example, our estimates using BALROG are slightly different from the estimates from the Y6 Gold catalog (Bechtol et al. 2025, see their Figure 3) which matches DES objects to a combination of data from the Vista Hemisphere Survey (VHS DR5, McMahan et al. 2013) and HSC (Miyazaki et al. 2018).

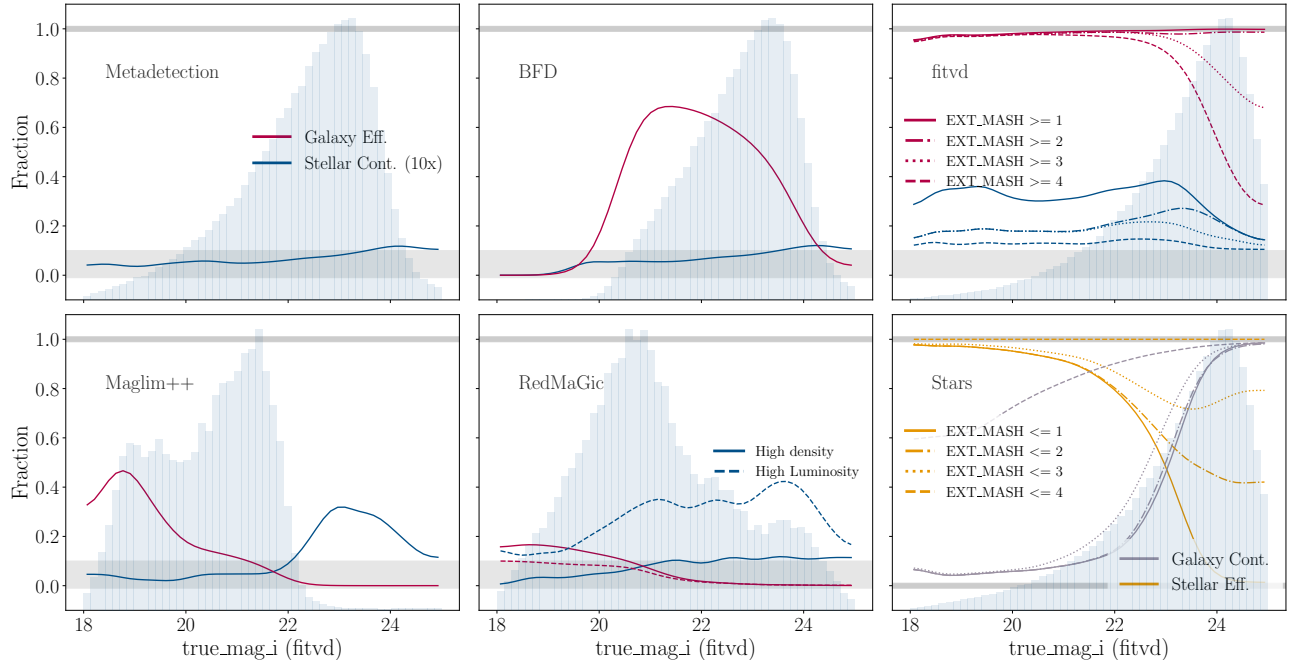
Figure 9 shows our results. The four leftmost panels present the galaxy efficiency and stellar contamination for the different galaxy samples. The contamination in the two source samples is below  $\lesssim 1\%$  (shown as the horizontal gray band) across the full magnitude range  $18 < m_i < 25$ . For the lens samples, the contamination is below  $1\%$  for the brighter

magnitudes of the MAGLIM++ sample, and similarly for the REDMAGIC high-density sample. The contamination rises towards fainter magnitudes since the number of objects selected from those magnitudes decreases significantly, given the lens galaxy sample contains predominantly brighter objects. The background of each panel presents a histogram of the number of objects (showing the counts on a linear rather than log scale) at a given  $i$ -band magnitude. As expected, any increases in stellar contamination fractions are in magnitude ranges that are at the faint tail of the distribution for each of the four main galaxy samples.

The source samples are the most efficient in their galaxy selection, as they discard the fewest true galaxies in their selection. The lens samples’ selections are far more restrictive, with REDMAGIC being the most restrictive. This is expected as the lens samples select specific types of galaxies that are best suited for obtaining precise redshift estimates. We do not present an efficiency estimate for the METADETECTION sample as we do not store the necessary data (i.e. the full catalog, without any selections applied) due to storage limitations; recall that METADETECTION provides five different detection catalogs, which increases our storage footprint by five times compared to other samples.

Finally, the rightmost panels of Figure 9 show the same estimates but for the full Y6 Gold sample, from varying the FITVD-based star-galaxy classifier (Bechtol et al. 2025). Note that this classifier is used in defining the lens samples, but not the source samples. The panel shows that  $\text{EXT\_MASH} \geq 4$  results in a percent-level contamination rate. The bottom right panel shows an inverted estimate, where we compute the *galaxy* contamination and the *stellar* efficiency. Such estimates are valuable for any science derived from the DES star catalogs. Our results affirm the findings in previous DES works (e.g., Everett et al. 2022; Sevilla-Noarbe et al. 2021), that the morphological classifier still results in significant ( $> 10\%$ ) contamination from galaxies when selecting a star sample. In DES Y6, we also develop an improved estimator — based on the gradient-boosted decision tree algorithm, XGBOOST (Chen & Guestrin 2016) — for selecting purer star/galaxy samples from the Y6 data (Bechtol et al. 2025, see their Section 4.2). We have not tested its efficiency as it has not been implemented in the BALROG catalog.

We make a brief note about two subtleties of the analysis in this section. First, the accuracy of star-galaxy classification can suffer if there are non-negligible errors in the PSF model. For example, if the true PSF is larger/smaller than the PSF model, then classifiers that use the size of the object and/or the size of the PSF would incur errors. This error will not be observed in BALROG, which uses the same PSF for convolving a source during injection and for deconvolving observations of the source during model fitting. Note that Schutt et al. (2025) (see their tests in Section 4) show that the



**Figure 9.** We show the efficiency (the number of selected good objects over the number of all good objects) and contamination (the number of selected bad objects over the total number of selected objects) for different samples used in DES Y6. We multiply the stellar contamination by 10x for easier visibility. The top row shows the two source samples — METADETECTION and BFD — and the base galaxy catalog from FITVD. The bottom row shows the lens samples, from MAGLIM++ and REDMAGIC, while the right-most panel shows the efficiency and contamination for the *star* sample based on the FITVD catalog. The histograms in the background show the number counts of objects as a function of true, injected magnitude. For METADETECTION, we cannot compute a galaxy efficiency curve as the necessary data was not kept due to storage limitations. For REDMAGIC, we show the high density sample in the histogram, and for FITVD and stars, we show the full sample without selections in the histogram. The gray bands show the 1% region around fractions of 1 and 0, with the latter region being increased by 10x when corresponding to stellar (not galaxy) contamination curves.

PSF model in Y6 reproduces the size of the true PSF (estimated using observed stars) within 1%, and the ellipticities within  $\delta e \sim 5 \times 10^{-4}$ . Thus, this subtlety is not a notable concern.

Second, stars in partially resolved binary systems will be measured as slightly extended objects, and can therefore be incorrectly classified as galaxies during the star-galaxy classification. BALROG will include such binaries as such sources will be present in the deep-field catalog from which we draw our injections. However, their impact of the stellar contamination crucially relies on accurate classification of such binary systems in the deep-field catalog (given this classification is used to determine the stellar contamination). We do not estimate the accuracy of the color-based estimator from Hartley & Choi et al. (2021) in classifying partially resolved binaries.

## 5. APPLICATIONS TO DES Y6 PROJECTS

The BALROG pipeline and its derived data products are necessary for key pieces of the DES cosmology analysis (e.g., redshift calibration and magnification estimates), and for informing the quality of our photometric/pipeline datasets. We have thus far detailed the latter in the section

above, and now discuss a few examples of the former, with specific use-cases in DES Y6.

### 5.1. Photometric redshift calibration

A key application of BALROG is in the photometric redshift calibration of the source sample and the lens sample (Myles & Alarcon et al. 2021; Sánchez et al. 2023; Giannini et al. 2024). The main DES  $3 \times 2$  point analysis — which uses galaxy clustering, cosmic shear, and galaxy-galaxy lensing to infer cosmology — relies on a robust characterization the source and lens samples’ redshift distributions in order to infer cosmology (DES Collaboration 2018, 2022). The characterizations for Y6 will be presented in Yin et al. (in prep.); Anbajagane & Cross et al. (in prep.); Giannini et al. (in prep.) for the METADETECTION, BFD, and MAGLIM++ sample, respectively. A significant challenge in estimating redshifts in photometric surveys is the presence of degeneracies in the color-redshift relation. These degeneracies can be broken by including additional photometric measurements per galaxy, but this is not always available for all galaxies in the full DES footprint. Following Buchs & Davis et al. (2019), DES has used a fully Bayesian technique for learning the color-redshift relation in regions of the sky with low-noise, multi-

band photometry measurements, and then transferring that knowledge to all galaxies across the full DES footprint.

In our case, the low-noise data is obtained from the DES Y3 deep fields (Hartley & Choi et al. 2021), which are four patches of the sky totaling  $\approx 5.9 \text{ deg}^2$  (Hartley & Choi et al. 2021) that are observed to significantly greater depth (i.e., the photometric measurements have lower noise). A majority of galaxies in these patches also have measurements in eight bands ( $u, g, r, i, z, J, H, K_s$ ), as opposed to the four bands of the wide field ( $g, r, i, z$ ), due to supplemental data from other surveys (e.g., Jarvis et al. 2012; Le Fèvre, O. et al. 2013). A large fraction of these available galaxies also have redshift estimates, provided by other surveys. The color-redshift relation is learned in these deep fields, and then extrapolated to the fiducial galaxy samples, which only have measurements in four bands. BALROG provides a robust, probabilistic way of performing this conversion, while accounting for the complexities in the images.

In brief, the probability distribution for the redshift of a given set of galaxies can be written as

$$p(z|\hat{c}, \hat{s}) = \sum_c p(z|c, \hat{c}, \hat{s})p(c|\hat{c}, \hat{s}), \quad (4)$$

where  $z$  is redshift,  $c$  is a galaxy phenotype, defined using the deep-field measurements (the 8-band, low-noise photometry);  $\hat{c}$  is the phenotype defined using the wide-field measurements (the 4-band, noisy photometry), and  $\hat{s}$  is the sample selection function. The second factor,  $p(c|\hat{c}, \hat{s})$ , is the transfer function, which is measured with BALROG. This quantity represents the probability of a galaxy being in deep-field phenotype  $c$ , given it is already in wide-field phenotype  $\hat{c}$  and has passed the selection  $\hat{s}$ . We use “phenotype” to refer to the classification of a galaxy into a low-dimension space using data from a higher-dimensional space. As a simple example, galaxies can be classified into red and blue “phenotypes” based on their photometric measurements. The classification step in DES Y6 is performed using self-organizing maps (SOMs), following Buchs & Davis et al. (2019); Myles & Alarcon et al. (2021), which is an unsupervised, data-driven algorithm.

In practice, we estimate redshift distributions for galaxies placed in different tomographic bins. In this case, the redshift distribution in each tomographic bin  $\hat{b}$  is estimated as

$$p(z|\hat{b}, \hat{s}) = \sum_{\hat{c} \in \hat{b}} \sum_c p(z|c, \hat{b}, \hat{s})p(c|\hat{c}, \hat{b}, \hat{s})p(\hat{c}|\hat{b}, \hat{s}), \quad (5)$$

The BALROG data is used for assigning galaxies of phenotype  $\hat{c}$  to tomographic bins, by computing the mean redshift of all BALROG galaxies in a given phenotype (e.g., Myles & Alarcon et al. 2021; Sánchez et al. 2023; Giannini et al. 2024). Then, BALROG provides a transfer function — which is the probability of a galaxy of wide-field phenotype

$\hat{c}$  also being in deep-field phenotype  $c$  — corresponding to all galaxies in a given tomographic bin.

We have already presented a rough version of the transfer function in Section 4.2. The phenotypes  $c$  and  $\hat{c}$  are discretizations of the deep-field and wide-field photometry, respectively. Figures 6 and 7 show the distributions of measured fluxes/colors as a function of the true, noiseless quantities. This can be thought of as one slice of the distribution  $p(\hat{c}|c)$ . Those figures show that for fainter objects (which dominate the sample, given their larger number density relative to that of bright objects) there can be large differences between the true, deep-field photometry and the observed, wide-field photometry. Such behaviors are accurately captured by computing the transfer function through BALROG.

Myles & Alarcon et al. (2021, see their Section 5.4) have also already shown that the DES Y3 BALROG dataset contributes an infinitesimal uncertainty to the final redshift distribution estimates. We expect this contribution to be even smaller in Y6, as the previous estimate accounted for the fact that the Y3 BALROG tiles only sampled  $\approx 20\%$  of the DES Y3 footprint. In Y6, we sample the entire footprint and therefore remove this uncertainty factor.

## 5.2. Lensing magnification

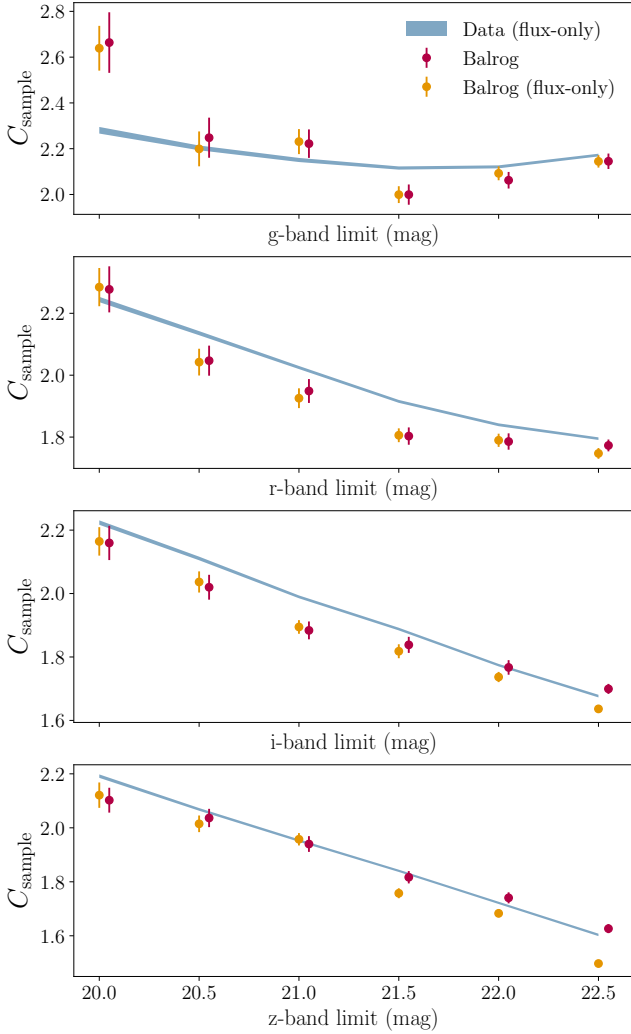
Another key application of BALROG is in the estimation of magnification coefficients. Gravitational lensing from the foreground matter distribution induces a magnification on the background structure (Bartelmann & Schneider 2001). This alters the observed surface area from a given redshift, and also impacts the measured properties of galaxies at that redshift, such as flux and area, which in turn changes the number of galaxies that pass a given selection function. As a result, magnification impacts any analyses that relies of the counts of galaxies, such as the galaxy clustering and galaxy-galaxy lensing analyses, as shown in DES Y3 (Elvin-Poole & MacCrann et al. 2023). It is also a key input in one of the redshift calibration methods, using the clustering of galaxies (Gatti & Giannini et al. 2022; d’Assignies & Bernstein et al. in prep.).

Magnification can be expressed in terms of the convergence  $\kappa$  and the shear  $\gamma$  (Bartelmann & Schneider 2001). In the weak lensing regime,  $\kappa \ll 1$  and  $\gamma \ll 1$  — where  $\kappa$  is the lensing convergence, and  $\gamma$  is the two-component shear — so the local magnification can be written as

$$\mu = \frac{1}{(1 - \kappa)^2 - |\gamma|^2} \approx \frac{1}{1 - 2\kappa} \approx 1 + 2\kappa. \quad (6)$$

Given such a magnification, we can rewrite the observed galaxy number density field  $\delta_g^{\text{obs}}$  into two additive components: an intrinsic density fluctuation, and a contribution due to magnification,

$$\delta_g^{\text{obs}} = \delta_g^{\text{int}} + \delta_g^{\text{mag}}. \quad (7)$$



**Figure 10.** Estimates of the magnification bias  $C_{\text{sample}}$  from boosted detection efficiency in samples defined with a flux magnitude limit (shown in the x-axis). The BALROG estimates (red) use the magnified runs, in which a 2% magnification is applied to every input object and the entire image processing is rerun. The data estimate (blue) applies the same magnification to galaxy magnitudes from the Y6 Gold sample and re-applies the same selection function. Finally, we show a BALROG-based flux-only estimate for comparison. In all cases, the error for each magnitude-limited sample is obtained via a jackknife estimate over 100 patches of the footprint.

As mentioned before, the contribution from magnification is the result of two separate and competing effects: an increase in the observed area element that decreases the observed number density of galaxies, and an enhancement in the detection efficiency of faint galaxies that increases the observed number density. With these effects, the observed local density fluctuation is given by

$$\delta_g^{\text{mag}} = (C_{\text{sample}} + C_{\text{area}}) \delta\kappa, \quad (8)$$

where  $C_{\text{sample}}$  is the factor associated to the boost in detection efficiency and  $C_{\text{area}}$  is the factor related to the decrease in number density. We can set  $C_{\text{area}} \equiv -2$  given the terms encapsulates a known geometric effect (Elvin-Poole & MacCrann et al. 2023) whereas  $C_{\text{sample}}$  depends on the exact definition of the sample selection function. To linear order, it can be estimated as the numerical derivative

$$C_{\text{sample}} \delta\kappa = \frac{N_{\delta\kappa} - N_0}{N_0}, \quad (9)$$

where  $N_{\delta\kappa}$  is the number of objects, whose properties have been impacted by magnification  $\delta\kappa$ , that pass the sample selection function. Then  $N_0$  is the number of objects, without any magnification, that pass the selection function.

To estimate  $C_{\text{sample}}$ , we run a magnified variant of BALROG for 2000 tiles ( $1,000 \text{ deg}^2$ ), or 20% of the full DES Y6 footprint, where each variant tile has the exact same injections (including the same positions on the sky) as in the fiducial version but now with a 2% increase in the flux and area of all injected objects, corresponding to magnification due to a lensing convergence of  $\delta\kappa = 0.01$ . Given a galaxy sample with a specific selection function, we can calculate  $C_{\text{sample}}$  from Equation (9), with  $N_{\delta\kappa}$  measured on the magnified runs and  $N_0$  measured on the fiducial runs. An approximate estimate of this method can also be made using a “flux-only” method where one takes the fiducial catalog, then increases the measured flux by 2% and applies the selection function of interest. This method will only trace a subset of the selection effects captured by the BALROG-based estimate.

In Figure 10 we show estimates of  $C_{\text{sample}}$ , where the selection function is just a simple magnitude threshold in a given photometric band. We make three different estimates, using two samples and two techniques. In all cases, we only use galaxies that pass the Y6 Gold selection and the star-galaxy selection; our selection function therefore includes a magnitude limit *and* a star-galaxy selection. We estimate  $C_{\text{sample}}$  using the fiducial and magnified BALROG samples, which contain the full impact of magnification on all measured properties of a galaxy. We then also perform a flux-only of magnification using the fiducial BALROG catalog and then also using the data catalog. In all cases, the error is obtained by performing a jackknife estimate over 100 patches of the footprint.

We present all three estimates in Figure 10. Overall, the BALROG magnification estimates are consistent with the BALROG flux-only estimates, with some discrepancies at the faint end for redder bands. This can occur because the flux-only estimate will not account for all impacts of magnification of the selection function. Next, the data flux-only estimate is generally similar to, but statistically distinct from, the BALROG flux-only estimate. While the BALROG data is largely representative of the real data, differences in the distribution of object properties between the two samples will

band	BALROG	Y6 GOLD
<i>g</i>	24.14 <sup>+0.21</sup> <sub>-0.19</sub>	24.16 <sup>+0.11</sup> <sub>-0.09</sub>
<i>r</i>	23.82 <sup>+0.20</sup> <sub>-0.17</sub>	23.84 <sup>+0.11</sup> <sub>-0.09</sub>
<i>i</i>	23.26 <sup>+0.20</sup> <sub>-0.18</sub>	23.28 <sup>+0.11</sup> <sub>-0.09</sub>
<i>z</i>	22.61 <sup>+0.21</sup> <sub>-0.20</sub>	22.62 <sup>+0.12</sup> <sub>-0.10</sub>

**Table 1.** The magnitude limit for extended objects with  $9.5 < \text{flux}/\text{flux\_err} < 10.5$  in a given band, computed using FITVD quantities from the BALROG synthetic data and the Y6 Gold data. The limits estimates from the two samples are consistent at the 2-3% level. The BALROG estimate is noisier as the sample contains fewer objects relative to Y6 Gold. See text for details on the calculation.

cause differences in the estimates of  $C_{\text{sample}}$ . Figure 10 shows that such differences are at the few percent level for most of the datapoints. The main magnification estimates in DES Y6 find that such BALROG–data differences are subdominant to the uncertainties of the  $C_{\text{sample}}$  estimates (Legnani et al. in prep.).

### 5.3. Depth estimates

In Figure 3 (and Section 4.1.1) we show the detection efficiency for galaxies of different brightness, in different bands. We have motivated above how this efficiency, especially as a function of the object’s true properties, is relevant for many science cases in DES Y6; with the redshift estimation (Section 5.1) and magnification estimates (Section 5.2) being two pertinent examples. It is also useful to quantify the raw completeness of the general purpose catalog. In DES Y3 this was done through a variety of techniques, one of which was to cross-match objects from the DES deep fields with the same object in the wide-field data (which is shallower) and compute a detection efficiency accordingly. This was accompanied by a BALROG-derived estimate as well. The latter is a purer test as we have exact control on the objects being injected into the data, but retain the realism of the DES dataset by injecting into real images. Similarly, in DES Y6, the BALROG dataset presented here is used to characterize the exact detection efficiency of the survey data; particularly the magnitude limit, per band, at which a galaxy is detected 90% of the time.

In Table 1 we present an estimate of survey depth, using the Y6 BALROG and Y6 Gold samples, for sources with  $S/N = 10$ . This is computed through the following process: (1) selecting all objects with  $9.5 < \text{flux}/\text{flux\_err} < 10.5$  in a given band, and with `EXT_MASH = 4` so that we select only extended objects (2) computing the mean magnitude in pixels of map with `NSIDE = 256`, and (3) computing the median and 16/84% percentile values of this magnitude over the footprint. The magnitude limits for the two samples are consistent at the 2-3% level, which is significantly smaller than the variation of this limit across the sky.

### 5.4. Large-scale survey systematics

Observations of the galaxy number density field, and in particular of the spatial correlations in this field, are sensitive to many cosmological processes (Peebles 1980) and have been used to extract constraints on parameters (e.g., Gil-Marín et al. 2015; Alam et al. 2017, 2021; DES Collaboration 2022). As a result, they have been extensively used in the DES Y1 and Y3 analyses (DES Collaboration 2018, 2022). A critical part of such analyses is the characterization, and subsequent correction, of spatial correlations in the field that arise from non-cosmological sources, such as the observing conditions of the survey. For example, we have shown in Figure 1 that the number of detected galaxies (either in BALROG or the Y6 Gold data) is lower near the east and west edges of the DES footprint, where the stellar density is higher. This causes a spatial correlation in the number density that does not originate from any cosmological process.

Such non-cosmological correlations can be corrected by assigning weights to the galaxies, where the weights are determined using the survey property maps (see Section 4.1.3) in conjunction with a variety of data-driven methods (e.g., Rodríguez-Monroy et al. 2022). BALROG provides a powerful test dataset for the robustness of these methods. The synthetic sources in BALROG are injected on a uniform grid (see Section 3.2). Under ideal conditions with no systematic effects, the number density of synthetic sources will have no correlation with that of the real sources in the data. However, observational effects will necessarily generate correlations between the two samples, as both samples are impacted by the presence of such observational effects in the real images. Vice-versa, if these non-cosmological correlations are completely corrected for, then there will be no spatial correlations between the two samples. Thus, BALROG can provide a vital estimate of the residual systematics in the galaxy correlation measurements. Such a test was performed in DES Y3 and was one aspect of the lens sample validations (Rodríguez-Monroy et al. 2022, see their Section 8.5 and Figure 12). The same validation in Y6 will be significantly more precise given the factor of 5 increase in area covered by BALROG and a factor of 12 increase in number of usable synthetic lens galaxies.

## 6. SUMMARY AND CONCLUSION

Photometric surveys are powerful tools for studying astrophysics and cosmology across many epochs and scales. In this work, we present the synthetic source catalogs used to characterize the DES Y6 image processing pipelines and the resulting galaxy samples derived from the catalogs. The synthetic sources are injected into real CCD images from DES Y6, and the entire processing pipeline is rerun on these source-injected images. This work is a supporting analysis for the main DES Y6 cosmology efforts. The BALROG cata-

log from DES Y6 is also currently the largest synthetic catalog at this fidelity level, and enables new science cases. We list our main results/findings below:

- BALROG in DES Y6 now provides synthetic sources for the entire  $\sim 5000 \text{ deg}^2$  footprint of the survey, which both increases the number of injections over Y3 (from  $\approx 25$  million to 146 million) and also enables more precise analyses of area-dependent effects, as we point out below.
- A “tiered” injection scheme for the synthetic sources greatly increases the size of the BALROG sample relevant for the cosmology analysis, while still injecting a large/representative sample of sources. For source and lens samples, we find increases of 300% and 1200%, respectively (Figure 2).
- The synthetic sample is representative of the data (Figure 4) and accurately captures the correlation between galaxy properties and observing conditions of the survey (Figure 5 and 11). The latter is not constructed by hand, and instead manifests directly from the complexities of the real images that BALROG injects synthetic sources into.
- Through BALROG, we show the DES Y6 image processing pipeline can recover the mean photometry (averaged over many detections) at the percent level (Figure 6 and 7), with deviations for the faintest objects and reddest bands that is consistent with Malmquist bias and potential background subtraction biases from SOURCEEXTRACTOR.
- The photometric accuracy is similar for stars, which have a much simpler morphology than galaxies, with a slightly better accuracy in general (Figure 12 and 13).
- The star-galaxy classification of the general purpose FITVD catalog, and of the two source samples and two lens samples, is found to limit stellar contamination to the percent level or lower over the magnitude range of interest (Figure 9).

This work has focused on the use-cases of BALROG that are particularly relevant for the DES Y6 cosmology analysis. There are also other use-cases beyond those discussed above: for example, BALROG in DES has been used to study the impact of unresolved sources and background subtraction (Eckert et al. 2020), to learn the distribution of observed properties as a function of survey conditions, to improve precision measurements of Milky Way substructure etc.

SSI has become a critical tool to robustly characterize and use data from surveys. In the upcoming decade, the Vera C. Rubin Observatory’s Legacy Survey of Space and

Time (LSST) will observe over 20,000  $\text{deg}^2$  of the sky, and will achieve higher precision in photometric and morphological measurements. SSI pipelines, similar to the BALROG pipeline used in this work, will therefore be necessary for most survey tasks. The dataset presented in this work is the largest synthetic source catalog built for a photometric survey and therefore will prove a viable testing ground for how SSI datasets can benefit/complement analyses with LSST data.

#### AUTHOR CONTRIBUTIONS

DA, MT and BY worked on the pipeline for BALROG DES Y6 — which was based on the original pipeline of DES Y3 from SE — with assistance from MB, MY, GB, RG, MJ and ES for various technical details/improvements of the processing pipeline done in Y6. DA and MT generated the entire BALROG dataset and produced all the resulting catalogs. DA and JB performed all analyses, and wrote all the text, as presented in this work. KB and JE served as our internal review committee and contributed to the writing and presentation of this paper. The MAGLIM++ sample for BALROG was generated by EL and GG, and the DNF redshift catalog by JD. The Y6 METADETECTION pipeline run on BALROG (as well as data) was developed by MY, MB, ES, MJ, RG, ERY and TS. The same for BFD was developed by MG, GB, VW, and MJ. The MAGLIM++ sample definitions was constructed by NW, MR, SL, JE, AC, JM and AP. The survey property maps we use were developed by ERY, while the REDMAGIC sample selection was defined by ERY and ERO. The remaining authors have made contributions to this paper that include, but are not limited to, the construction of DECam and other aspects of collecting the data; data processing and calibration; developing broadly used methods, codes, and simulations; running the pipelines and validation tests; and promoting the science analysis.

#### ACKNOWLEDGEMENTS

DA is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1746045. DA also thanks Richard Kron for discussions on the history of photometric surveys and observational techniques.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos

Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l’Espai (IEEC/CSIC), the Institut de Física d’Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University

of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at NSF Cerro Tololo Inter-American Observatory at NSF NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants PID2021-123012, PID2021-128989 PID2022-141079, SEV-2016-0588, CEX2020-001058-M and CEX2020-001007-S, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya.

We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

## REFERENCES

- Abell, G. O. 1958, *ApJS*, 3, 211, doi: [10.1086/190036](https://doi.org/10.1086/190036)
- Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, *Publications of the Astronomical Society of Japan*, 70, S8, doi: [10.1093/pasj/psx066](https://doi.org/10.1093/pasj/psx066)
- Alam, S., Ata, M., Bailey, S., et al. 2017, *MNRAS*, 470, 2617, doi: [10.1093/mnras/stx721](https://doi.org/10.1093/mnras/stx721)
- Alam, S., Aubert, M., Avila, S., et al. 2021, *PhRvD*, 103, 083533, doi: [10.1103/PhysRevD.103.083533](https://doi.org/10.1103/PhysRevD.103.083533)
- Alarcon, A., Gaztanaga, E., Eriksen, M., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 501, 6103, doi: [10.1093/mnras/staa3659](https://doi.org/10.1093/mnras/staa3659)
- Anbajagane & Cross, D., Cross, D., Sanchez, C., et al. in prep.
- Barbary, K. 2016, *Journal of Open Source Software*, 1, 58, doi: [10.21105/joss.00058](https://doi.org/10.21105/joss.00058)
- Barnard, E. E., E, E., & E. 1919, *ApJ*, 49, 1, doi: [10.1086/142439](https://doi.org/10.1086/142439)
- Barnard, E. E., Frost, E. B., & Calvert, M. R. 1927, *A Photographic Atlas of Selected Regions of the Milky Way*
- Bartelmann, M., & Schneider, P. 2001, *PhR*, 340, 291, doi: [10.1016/S0370-1573\(00\)00082-X](https://doi.org/10.1016/S0370-1573(00)00082-X)
- Bechtol, K., Drlica-Wagner, A., Balbinot, E., et al. 2015, *ApJ*, 807, 50, doi: [10.1088/0004-637X/807/1/50](https://doi.org/10.1088/0004-637X/807/1/50)
- Bechtol, K., Sevilla-Noarbe, I., Drlica-Wagner, A., et al. 2025
- Bernstein, G. M., & Armstrong, R. 2014, *MNRAS*, 438, 1880, doi: [10.1093/mnras/stt2326](https://doi.org/10.1093/mnras/stt2326)
- Bertin, E. 2006, in *Astronomical Society of the Pacific Conference Series*, Vol. 351, *Astronomical Data Analysis Software and Systems XV*, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 112
- Bertin, E. 2013, *PSFEX: Point Spread Function Extractor*, *Astrophysics Source Code Library*, record ascl:1301.001, <http://ascl.net/1301.001>
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393, doi: [10.1051/aas:1996164](https://doi.org/10.1051/aas:1996164)
- Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in *Astronomical Society of the Pacific Conference Series*, Vol. 281, *Astronomical Data Analysis Software and Systems XI*, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228
- Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, *PASJ*, 70, S5, doi: [10.1093/pasj/psx080](https://doi.org/10.1093/pasj/psx080)
- Buchs & Davis, R., Davis, C., Gruen, D., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 489, 820–841, doi: [10.1093/mnras/stz2162](https://doi.org/10.1093/mnras/stz2162)
- Burke, D. L., Rykoff, E. S., Allam, S., et al. 2017, *The Astronomical Journal*, 155, 41, doi: [10.3847/1538-3881/aa9f22](https://doi.org/10.3847/1538-3881/aa9f22)
- Carlstrom, J. E., Holder, G. P., & Reese, E. D. 2002, *ARA&A*, 40, 643, doi: [10.1146/annurev.astro.40.060401.093803](https://doi.org/10.1146/annurev.astro.40.060401.093803)
- Chen, T., & Guestrin, C. 2016, *arXiv e-prints*, arXiv:1603.02754, doi: [10.48550/arXiv.1603.02754](https://doi.org/10.48550/arXiv.1603.02754)



- d'Assignies & Bernstein, W., Bernstein, G., Sanchez, C., Alarcon, A., & Giannini, G. in prep.
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *The Astronomical Journal*, 145, 10, doi: [10.1088/0004-6256/145/1/10](https://doi.org/10.1088/0004-6256/145/1/10)
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, *The Astronomical Journal*, 151, 44, doi: [10.3847/0004-6256/151/2/44](https://doi.org/10.3847/0004-6256/151/2/44)
- de Jong, J. T. A., Kuijken, K., Applegate, D., et al. 2013, *The Messenger*, 154, 44
- de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, *A&A*, 582, A62, doi: [10.1051/0004-6361/201526601](https://doi.org/10.1051/0004-6361/201526601)
- De Vicente, J., Sánchez, E., & Sevilla-Noarbe, I. 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 3078, doi: [10.1093/mnras/stw857](https://doi.org/10.1093/mnras/stw857)
- DES Collaboration. 2018, *Physical Review D*, 98, doi: [10.1103/physrevd.98.043526](https://doi.org/10.1103/physrevd.98.043526)
- . 2022, *PhRvD*, 105, 023520, doi: [10.1103/PhysRevD.105.023520](https://doi.org/10.1103/PhysRevD.105.023520)
- DES Collaboration, Abbott, T. M. C., Adamó w, M., et al. 2021, *The Astrophysical Journal Supplement Series*, 255, 20, doi: [10.3847/1538-4365/ac00b3](https://doi.org/10.3847/1538-4365/ac00b3)
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints, arXiv:1611.00036, doi: [10.48550/arXiv.1611.00036](https://doi.org/10.48550/arXiv.1611.00036)
- Dreyer, J. L. E. 1888, *MmRAS*, 49, 1
- Eckert, K., Bernstein, G. M., Amara, A., et al. 2020, *MNRAS*, 497, 2529, doi: [10.1093/mnras/staa2133](https://doi.org/10.1093/mnras/staa2133)
- Elvin-Poole & MacCrann, J., MacCrann, N., Everett, S., et al. 2023, *Monthly Notices of the Royal Astronomical Society*, 523, 3649, doi: [10.1093/mnras/stad1594](https://doi.org/10.1093/mnras/stad1594)
- Everett, S., Yanny, B., Kuropatkin, N., et al. 2022, *The Astrophysical Journal Supplement Series*, 258, 15, doi: [10.3847/1538-4365/ac26c1](https://doi.org/10.3847/1538-4365/ac26c1)
- Flaugher, B. 2005, *International Journal of Modern Physics A*, 20, 3121, doi: [10.1142/S0217751X05025917](https://doi.org/10.1142/S0217751X05025917)
- Fowler, J. W., Niemack, M. D., Dicker, S. R., et al. 2007, *ApOpt*, 46, 3444, doi: [10.1364/AO.46.003444](https://doi.org/10.1364/AO.46.003444)
- Gaia Collaboration, Vallenari, A., Brown, A., et al. 2023, *Astronomy & Astrophysics*, 674, A1, doi: [10.1051/0004-6361/202243940](https://doi.org/10.1051/0004-6361/202243940)
- García-Fernández, M., Sánchez, E., Sevilla-Noarbe, I., et al. 2018, *MNRAS*, 476, 1071, doi: [10.1093/mnras/sty282](https://doi.org/10.1093/mnras/sty282)
- Gatti & Giannini, M., Giannini, G., Bernstein, G. M., et al. 2022, *MNRAS*, 510, 1223, doi: [10.1093/mnras/stab3311](https://doi.org/10.1093/mnras/stab3311)
- Gatti & Wetzell, M., Wetzell, V., Bernstein, G., et al. in prep.
- Giannini, G., Alarcon, A., Bernstein, G., Troxel, M., & Anbajagane, D. in prep.
- Giannini, G., Alarcon, A., Gatti, M., et al. 2024, *MNRAS*, 527, 2010, doi: [10.1093/mnras/stad2945](https://doi.org/10.1093/mnras/stad2945)
- Gil-Marín, H., Verde, L., Noreña, J., et al. 2015, *MNRAS*, 452, 1914, doi: [10.1093/mnras/stv1359](https://doi.org/10.1093/mnras/stv1359)
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *The Astrophysical Journal*, 622, 759, doi: [10.1086/427976](https://doi.org/10.1086/427976)
- Hartley & Choi, W. G., Choi, A., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 509, 3547–3579, doi: [10.1093/mnras/stab3055](https://doi.org/10.1093/mnras/stab3055)
- Herschel, W. 1786, *Philosophical Transactions of the Royal Society of London Series I*, 76, 457
- . 1789, *Philosophical Transactions of the Royal Society of London Series I*, 79, 212
- . 1802, *Philosophical Transactions of the Royal Society of London Series I*, 92, 477
- Ho, S., Cuesta, A., Seo, H.-J., et al. 2012, *ApJ*, 761, 14, doi: [10.1088/0004-637X/761/1/14](https://doi.org/10.1088/0004-637X/761/1/14)
- Hubble, E. 1934, *ApJ*, 79, 8, doi: [10.1086/143517](https://doi.org/10.1086/143517)
- Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 460, 2245, doi: [10.1093/mnras/stw990](https://doi.org/10.1093/mnras/stw990)
- Jarvis, M., Bernstein, G. M., Amon, A., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 501, 1282, doi: [10.1093/mnras/staa3679](https://doi.org/10.1093/mnras/staa3679)
- Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 428, 1281, doi: [10.1093/mnras/sts118](https://doi.org/10.1093/mnras/sts118)
- Kong, H., Ross, A. J., Honscheid, K., et al. 2024, arXiv e-prints, arXiv:2405.16299, doi: [10.48550/arXiv.2405.16299](https://doi.org/10.48550/arXiv.2405.16299)
- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *The Astrophysical Journal Supplement Series*, 224, 24, doi: [10.3847/0067-0049/224/2/24](https://doi.org/10.3847/0067-0049/224/2/24)
- Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, *A&A*, 559, A14, doi: [10.1051/0004-6361/201322179](https://doi.org/10.1051/0004-6361/201322179)
- Legnani, E., Elvin-Poole, J., Anbajagane, D., & Tabbutt, M. in prep.
- Lehtinen, K., Prusti, T., de Bruijne, J., et al. 2023, *A&A*, 671, A16, doi: [10.1051/0004-6361/202142929](https://doi.org/10.1051/0004-6361/202142929)
- Leistedt, B., Peiris, H. V., Mortlock, D. J., Benoit-Lévy, A., & Pontzen, A. 2013, *MNRAS*, 435, 1857, doi: [10.1093/mnras/stt1359](https://doi.org/10.1093/mnras/stt1359)
- MacCrann & Becker, N., Becker, M. R., McCullough, J., et al. 2022, *MNRAS*, 509, 3371, doi: [10.1093/mnras/stab2870](https://doi.org/10.1093/mnras/stab2870)
- Magnier, E. 2006, in *The Advanced Maui Optical and Space Surveillance Technologies Conference*, ed. S. Ryan, E50
- Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2017, *ApJ*, 841, 111, doi: [10.3847/1538-4357/aa6f08](https://doi.org/10.3847/1538-4357/aa6f08)
- . 2019, *ApJ*, 877, 81, doi: [10.3847/1538-4357/ab184d](https://doi.org/10.3847/1538-4357/ab184d)
- Mau & Becker, S., Becker, M. R., Jarvis, M., & Anbajagane, D. in prep.
- McMahon, R. G., Banerji, M., Gonzalez, E., et al. 2013, *The Messenger*, 154, 35

- Meisner, A. M., Lang, D., & Schlegel, D. J. 2017, *AJ*, 154, 161, doi: [10.3847/1538-3881/aa894e](https://doi.org/10.3847/1538-3881/aa894e)
- Miles, R. 2007, *Journal of the British Astronomical Association*, 117, 172
- Miyazaki, S., Komiyama, Y., Kawanomoto, S., et al. 2018, *PASJ*, 70, S1, doi: [10.1093/pasj/psx063](https://doi.org/10.1093/pasj/psx063)
- Morganson, E., Gruendl, R. A., Menanteau, F., et al. 2018, *Publications of the Astronomical Society of the Pacific*, 130, 074501, doi: [10.1088/1538-3873/aab4ef](https://doi.org/10.1088/1538-3873/aab4ef)
- Myles & Alarcon, J., Alarcon, A., Amon, A., et al. 2021, *MNRAS*, 505, 4249, doi: [10.1093/mnras/stab1515](https://doi.org/10.1093/mnras/stab1515)
- Pandey, S., Krause, E., DeRose, J., et al. 2022, *PhRvD*, 106, 043520, doi: [10.1103/PhysRevD.106.043520](https://doi.org/10.1103/PhysRevD.106.043520)
- Peebles, P. J. E. 1980, *The large-scale structure of the universe*
- Pieres, A., Girardi, L., Balbinot, E., et al. 2020, *MNRAS*, 497, 1547, doi: [10.1093/mnras/staa1980](https://doi.org/10.1093/mnras/staa1980)
- Prat, J., Blazek, J., Sánchez, C., et al. 2022, *PhRvD*, 105, 083528, doi: [10.1103/PhysRevD.105.083528](https://doi.org/10.1103/PhysRevD.105.083528)
- Roberts, I. 1893, *A Selection of Photographs of Stars, Star-Clusters and Nebulae, together with Information concerning the Instruments and the Methods employed in the pursuit of Celestial Photography*
- Rodríguez-Monroy, M., Weaverdyck, N., Elvin-Poole, J., et al. 2022, *Monthly Notices of the Royal Astronomical Society*, 511, 2665, doi: [10.1093/mnras/stac104](https://doi.org/10.1093/mnras/stac104)
- Rodríguez-Monroy, M., Weaverdyck, N., Elvin-Poole, J., et al. 2022, *MNRAS*, 511, 2665, doi: [10.1093/mnras/stac104](https://doi.org/10.1093/mnras/stac104)
- Rodríguez-Monroy & Weaverdyck, M., Weaverdyck, Noahand Lee, S., Elvin-Poole, J., et al. in prep.
- Ross, A. J., Percival, W. J., Sánchez, A. G., et al. 2012, *MNRAS*, 424, 564, doi: [10.1111/j.1365-2966.2012.21235.x](https://doi.org/10.1111/j.1365-2966.2012.21235.x)
- Rowe, B., Jarvis, M., Mandelbaum, R., et al. 2015, *GalSim: The modular galaxy image simulation toolkit*. <https://arxiv.org/abs/1407.7676>
- Rozo, E., Rykoff, E. S., Abate, A., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 1431, doi: [10.1093/mnras/stw1281](https://doi.org/10.1093/mnras/stw1281)
- Rykoff, E. S., Rozo, E., & Keisler, R. 2015, *arXiv e-prints*, arXiv:1509.00870, doi: [10.48550/arXiv.1509.00870](https://doi.org/10.48550/arXiv.1509.00870)
- Salvato, M., Ilbert, O., & Hoyle, B. 2018, doi: [10.48550/ARXIV.1805.12574](https://doi.org/10.48550/ARXIV.1805.12574)
- Sánchez, C., Alarcon, A., Bernstein, G. M., et al. 2023, *MNRAS*, 525, 3896, doi: [10.1093/mnras/stad2402](https://doi.org/10.1093/mnras/stad2402)
- Sánchez, J., Walter, C. W., Awan, H., et al. 2020, *MNRAS*, 497, 210, doi: [10.1093/mnras/staa1957](https://doi.org/10.1093/mnras/staa1957)
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *The Astrophysical Journal*, 500, 525, doi: [10.1086/305772](https://doi.org/10.1086/305772)
- Schutt, T., Jarvis, M., Roodman, A., et al. 2025
- Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. 2021, *ApJS*, 254, 24, doi: [10.3847/1538-4365/abeb66](https://doi.org/10.3847/1538-4365/abeb66)
- Sevilla-Noarbe, I., Bechtol, K., Kind, M. C., et al. 2021, *The Astrophysical Journal Supplement Series*, 254, 24, doi: [10.3847/1538-4365/abeb66](https://doi.org/10.3847/1538-4365/abeb66)
- Sheldon, E. 2015, *NGMIX: Gaussian mixture models for 2D images*, *Astrophysics Source Code Library*, record ascl:1508.008
- Sheldon, E. S., Becker, M. R., Jarvis, M., Armstrong, R., & LSST Dark Energy Science Collaboration. 2023, *The Open Journal of Astrophysics*, 6, 17, doi: [10.21105/astro.2303.03947](https://doi.org/10.21105/astro.2303.03947)
- Sheldon, E. S., Becker, M. R., MacCrann, N., & Jarvis, M. 2020, *ApJ*, 902, 138, doi: [10.3847/1538-4357/abb595](https://doi.org/10.3847/1538-4357/abb595)
- Shipp, N., Drlica-Wagner, A., Balbinot, E., et al. 2018, *ApJ*, 862, 114, doi: [10.3847/1538-4357/aacdad](https://doi.org/10.3847/1538-4357/aacdad)
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, 131, 1163, doi: [10.1086/498708](https://doi.org/10.1086/498708)
- Stanford, S. A., Masters, D., Darvish, B., et al. 2021, *ApJS*, 256, 9, doi: [10.3847/1538-4365/ac0833](https://doi.org/10.3847/1538-4365/ac0833)
- Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *The Astronomical Journal*, 123, 485, doi: [10.1086/324741](https://doi.org/10.1086/324741)
- Suchyta, E., Huff, E. M., Aleksić, J., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 457, 786, doi: [10.1093/mnras/stv2953](https://doi.org/10.1093/mnras/stv2953)
- The LSST Dark Energy Science Collaboration, Mandelbaum, R., Eifler, T., et al. 2018, *arXiv e-prints*, arXiv:1809.01669, doi: [10.48550/arXiv.1809.01669](https://doi.org/10.48550/arXiv.1809.01669)
- Thornton, R. J., Ade, P. A. R., Aiola, S., et al. 2016, *ApJS*, 227, 21, doi: [10.3847/1538-4365/227/2/21](https://doi.org/10.3847/1538-4365/227/2/21)
- Valdes, F., Gruendl, R., & Project, D. 2014, in *Astronomical Society of the Pacific Conference Series*, Vol. 485, *Astronomical Data Analysis Software and Systems XXIII*, 379
- Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, *ApJS*, 258, 11, doi: [10.3847/1538-4365/ac3078](https://doi.org/10.3847/1538-4365/ac3078)
- Weaverdyck, N., Rodríguez-Monroy, M., Lee, S., et al. in prep.
- Wright, A. H., Kuijken, K., Hildebrandt, H., et al. 2024, *A&A*, 686, A170, doi: [10.1051/0004-6361/202346730](https://doi.org/10.1051/0004-6361/202346730)
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868, doi: [10.1088/0004-6256/140/6/1868](https://doi.org/10.1088/0004-6256/140/6/1868)
- Yamamoto & Becker, M., Becker, M. R., Sheldon, E. S., et al. 2025
- Yin, B., Amon, A., Campos, A., & Alarcon, A. in prep.
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *The Astronomical Journal*, 120, 1579, doi: [10.1086/301513](https://doi.org/10.1086/301513)

## APPENDIX

## A. CORRELATIONS OF LENS GALAXIES WITH SURVEY PROPERTIES: MAGLIM++ &amp; REDMAGIC

In Section 4.1.3 we detail the correlations between the number of detected/selected galaxies and the survey property maps; see that Section for more details. An important finding was that the correlations seen in BALROG were consistent with those measured in the data (except in two cases where we *a priori* expect disagreement due to cosmological correlations). We now repeat this measurement for the two lens sample used in DES Y6: MAGLIM++ and REDMAGIC.

Figure 11 repeats the test of Figure 5 but now using these two lens samples. Note that these samples are a significantly smaller subset of the main Y6 Gold catalog and therefore more sparsely cover the survey footprint. For this reason, we use  $N_{\text{SIDE}} = 128$  when generating maps for the correlations. In general, we find that the galaxy–survey property correlations are consistent between BALROG and the data, for both MAGLIM++ and REDMAGIC. The main outliers are the thermal SZ map and the background offset map, where the difference between the correlations — quantified as a  $\chi^2$  metric — exhibits  $\chi^2/N_{\text{dof}} \gg 1$ . This difference is expected, as we discussed in Section 4.1.3. We have also repeated this test for the galaxies in each tomographic bin of the MAGLIM++ sample (Figure not shown) and find the same qualitative conclusions within each bin. The MAGLIM++ sample used in Figure 11 has a redshift range of  $0.2 < z < 1.05$ .

This test is more analogous to the work of Kong et al. (2024) — when compared to our analysis of Y6 Gold in Section 5 — given their work tested the selection of luminous red galaxies (LRGs), which is a color-based selection. The MAGLIM++ sample is primarily a magnitude-based selection, with minor color selections to reduce stellar contamination, while the REDMAGIC sample is a color-based selection of LRGs. Even after using samples with such selections, we find that the galaxy–survey property correlations are consistent between the real data and the BALROG synthetic sample. This further corroborates the findings of Kong et al. (2024), where the main differences of the galaxy–extinction correlations between the real and synthetic samples were found in the northern Galactic cap (i.e. *not* in the sky region encompassing the DES data).

## B. PHOTOMETRIC PERFORMANCE OF THE STAR SAMPLE

The main body of our work focuses on validating the synthetic galaxy sample, and testing the accuracy of the image processing pipeline in measuring the properties of galaxies. However, a vast variety of science performed with the DES data use star catalogs (e.g., Bechtol et al. 2015; Shipp et al.

2018; Pieres et al. 2020). To aid such science targets, we now explicitly validate the performance of the Y6 processing pipeline in recovering the properties of stars. We stress one notable difference between the validation in our work in comparison to that done in E22; the latter performed this test using simple, simulated point-sources as the injected star sample whereas we injected actual measurements of stars from the DES Y3 deep-field catalog (Hartley & Choi et al. 2021). This deep-field sample is determined using the photometric classifier built in Y3, which provides a  $> 95\%$  pure star sample down to the 24th magnitude in the *i*-band (see their Figure 15).

Figure 12 shows the difference in recovered magnitudes compared to the injected magnitude  $\Delta\text{mag}_\delta$  as a function of injected magnitude for the *griz* bands. We show the summary statistics of the bias as done in Figure 6. The results are generally similar to that discussed in Figure 6: the pipeline is accurate at the few millimag level for brighter objects, and shows larger biases, at the tens of millimag, for fainter objects. The effects of Malmquist bias and potentially incorrect background subtraction in SOURCEEXTRACTOR, as discussed in Section 4.2 for the galaxy sample, can explain the slightly negative biases found in this analysis as well. Note that we show the FITVD PSF magnitudes, which are photometric measurements that use the PSF for the functional form of the light profile and are therefore optimized for studying point sources. We still show the standard FITVD measurements on the x-axis to facilitate easy comparison with Figure 6.

In comparison with the analysis done on the galaxy sample (section 4.2), we find that the deviations in recovered magnitude of the stellar sample has a scatter that is on average half of that measured in the galaxy sample, across the full magnitude range (see Table 2). This is expected because the galaxy sample contains injected objects with a wide variety of morphologies, compared to the purer stellar injections we study here, and also because the galaxy samples could be more easily affected by residual light from nearby sources, given they are more extended objects than their stellar counterparts.

In Figure 13, we plot the difference in measured and input  $g-r$ ,  $r-i$ , and  $i-z$  color as a function of input color. The contours and summary statistics are calculated the same way as with the magnitudes mentioned above (and first discussed in Figure 6), except now we use narrowed magnitude bins, of size 100 millimag for  $g-r$  and  $r-i$ , and 50 millimag for  $i-z$ . We continue to use only objects with signal-to-noise greater than 5. For the three color definitions we study, the accuracy of the recovered colors is consistently below 10 millimag for most of the color range. For the redder colors (right-end of

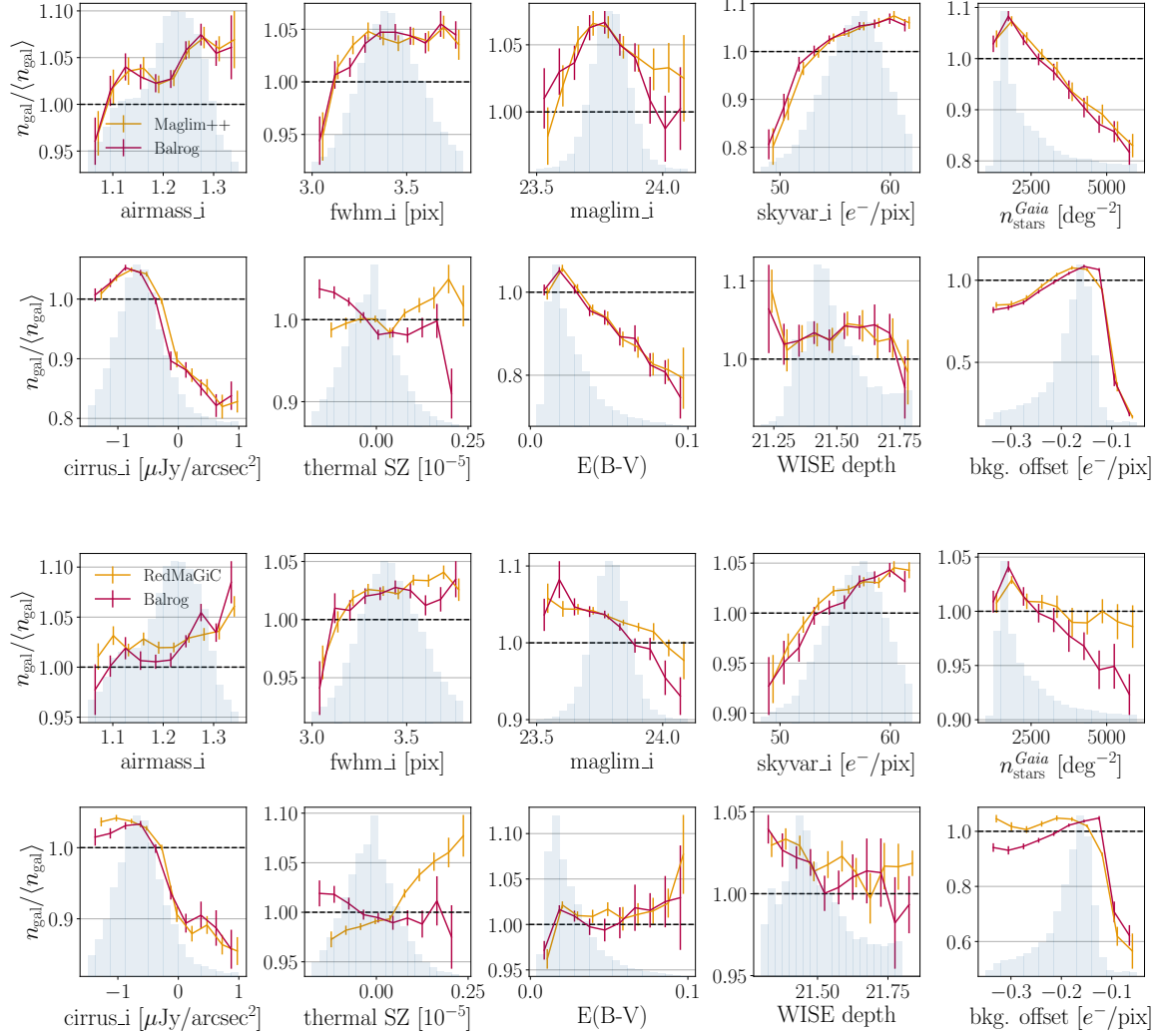
Magnitude	Galaxies (mmag)				Stars (mmag)			
	<i>g</i>	<i>r</i>	<i>i</i>	<i>z</i>	<i>g</i>	<i>r</i>	<i>i</i>	<i>z</i>
17.0 - 17.25	-8 (31)	-4 (22)	-4 (18)	-3 (18)	—	—	—	—
17.25 - 17.5	-12 (31)	-12 (23)	-7 (19)	-2 (18)	—	—	—	—
17.5 - 17.75	-27 (31)	-9 (23)	-2 (19)	-2 (20)	—	—	—	—
17.75 - 18.0	-8 (30)	-9 (23)	-5 (21)	-1 (22)	—	—	—	—
18.0 - 18.25	-6 (30)	-6 (24)	-4 (22)	-1 (23)	-23 (21)	-9 (19)	-3 (14)	0.5 (12)
18.25 - 18.5	-12 (32)	-6 (25)	-3 (23)	-0.6 (24)	-10 (27)	-8 (19)	-2 (15)	1 (13)
18.5 - 18.75	-9 (30)	-7 (26)	-3 (25)	0.7 (27)	-23 (28)	-0.1 (19)	-2 (15)	3 (13)
18.75 - 19.0	-12 (32)	-8 (27)	-2 (26)	0.1 (28)	-0.9 (28)	-1 (18)	1 (15)	1 (14)
19.0 - 19.25	-7 (32)	-6 (28)	-2 (28)	0.9 (31)	-7 (28)	-7 (21)	-0.2 (16)	3 (14)
19.25 - 19.5	-12 (33)	-7 (29)	-3 (30)	2 (34)	-3 (27)	-4 (19)	-0.1 (16)	3 (15)
19.5 - 19.75	-11 (34)	-6 (31)	-1 (32)	3 (37)	-20 (29)	-6 (21)	-0.7 (17)	3 (17)
19.75 - 20.0	-12 (35)	-7 (33)	-0.7 (36)	4 (41)	-14 (28)	-4 (20)	-0.2 (17)	3 (18)
20.0 - 20.25	-13 (37)	-6 (34)	0.1 (39)	6 (47)	-18 (30)	-4 (22)	-0.1 (18)	4 (19)
20.25 - 20.5	-11 (38)	-6 (37)	0.6 (43)	7 (52)	-17 (30)	-3 (21)	0.2 (19)	5 (21)
20.5 - 20.75	-14 (40)	-5 (41)	1 (48)	8 (59)	-17 (30)	-4 (23)	1 (21)	6 (24)
20.75 - 21.0	-13 (43)	-5 (45)	2 (53)	9 (66)	-20 (31)	-5 (24)	0.2 (22)	7 (27)
21.0 - 21.25	-14 (45)	-4 (50)	3 (59)	9 (73)	-18 (33)	-5 (25)	1 (24)	9 (31)
21.25 - 21.5	-14 (49)	-4 (55)	2 (66)	11 (84)	-16 (33)	-5 (26)	4 (26)	10 (36)
21.5 - 21.75	-14 (53)	-4 (61)	3 (74)	13 (95)	-13 (34)	-4 (28)	5 (30)	13 (44)
21.75 - 22.0	-14 (59)	-4 (69)	5 (83)	17 (110)	-16 (36)	-4 (31)	6 (35)	16 (52)
22.0 - 22.25	-15 (66)	-2 (77)	8 (94)	23 (127)	-16 (37)	-0.8 (34)	11 (39)	21 (63)
22.25 - 22.5	-15 (73)	-0.8 (87)	13 (107)	32 (148)	-15 (40)	-0.6 (39)	14 (47)	27 (76)
22.5 - 22.75	-15 (83)	2 (98)	20 (123)	42 (175)	-12 (45)	3 (44)	19 (56)	31 (95)
22.75 - 23.0	-14 (94)	8 (113)	29 (143)	51 (207)	-13 (50)	6 (51)	25 (67)	38 (114)
23.0 - 23.25	-12 (108)	15 (130)	39 (168)	58 (246)	-10 (57)	10 (62)	30 (83)	47 (140)
23.25 - 23.5	-7 (124)	24 (150)	47 (199)	56 (291)	-8 (68)	15 (73)	40 (102)	55 (177)
23.5 - 23.75	-1 (144)	33 (176)	48 (237)	33 (340)	-6 (82)	22 (90)	50 (130)	59 (223)
23.75 - 24.0	6 (168)	38 (207)	30 (278)	-19 (386)	-4 (98)	29 (110)	63 (158)	58 (271)
24.0 - 24.25	11 (198)	28 (245)	-21 (322)	-113 (423)	-1 (119)	42 (135)	74 (197)	31 (316)
24.25 - 24.5	6 (233)	-12 (291)	-117 (360)	-250 (440)	6 (145)	51 (169)	77 (241)	-36 (346)
24.5 - 24.75	-13 (279)	-93 (339)	-252 (384)	-404 (429)	13 (180)	67 (205)	67 (285)	-176 (367)
24.75 - 25.0	-51 (333)	-210 (380)	-402 (387)	-543 (395)	25 (221)	82 (253)	3 (329)	-331 (371)

**Table 2.** The median difference between the true and measured magnitudes as shown in Figure 6 and 12, in units of millimag. The numbers in the brackets show the standard deviation of the differences, shown as black bars in the aforementioned figures. Our synthetic star sample contains too few objects below  $m < 18$  so we do not tabulate the bias there.

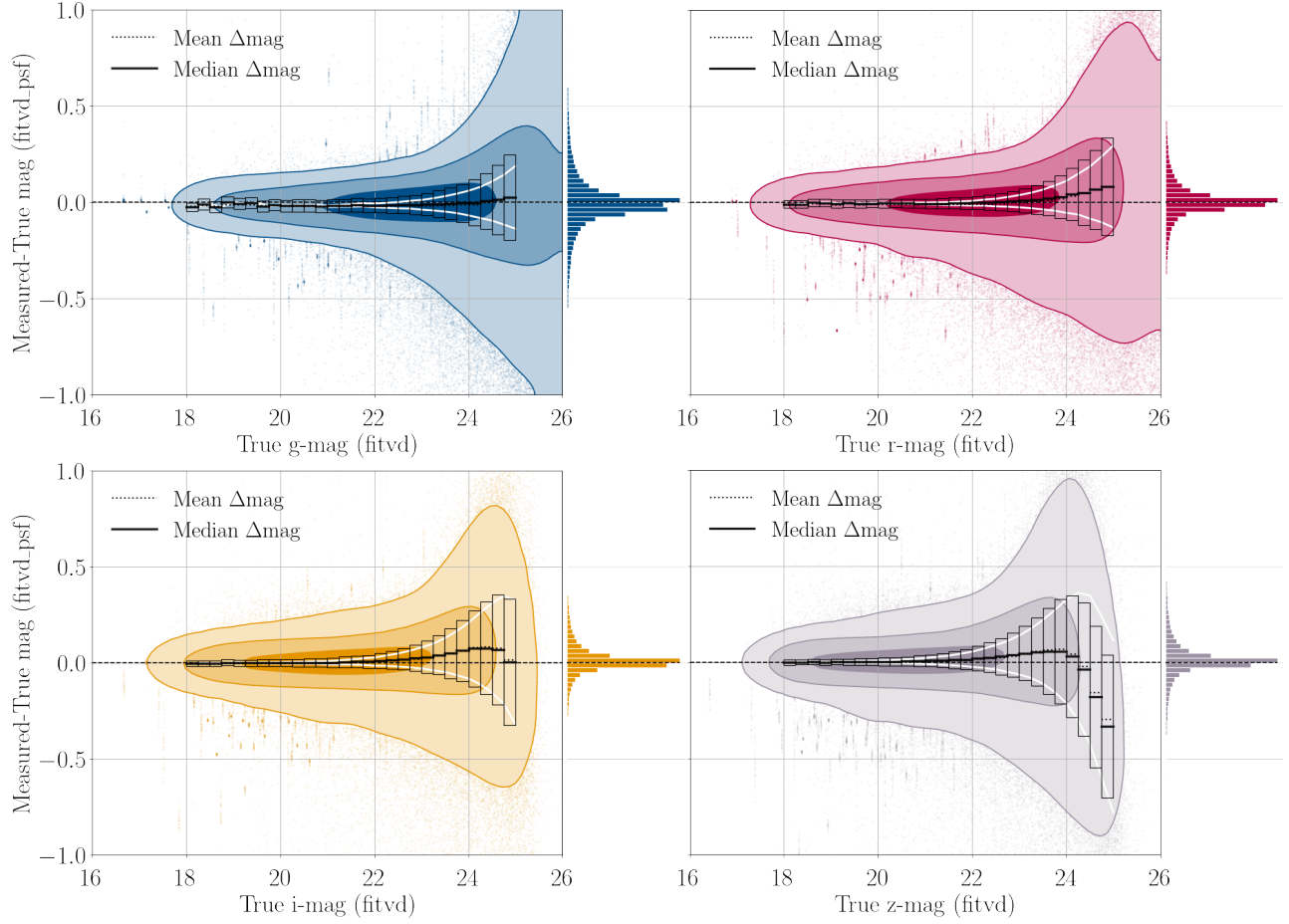
all panels), the median bias grows to the 5% level for  $g - r$  alone. We find no evidence for chromatic effects in the bias as the behavior is consistent across all three color definitions.

### C. TABULATION OF PHOTOMETRIC ACCURACY

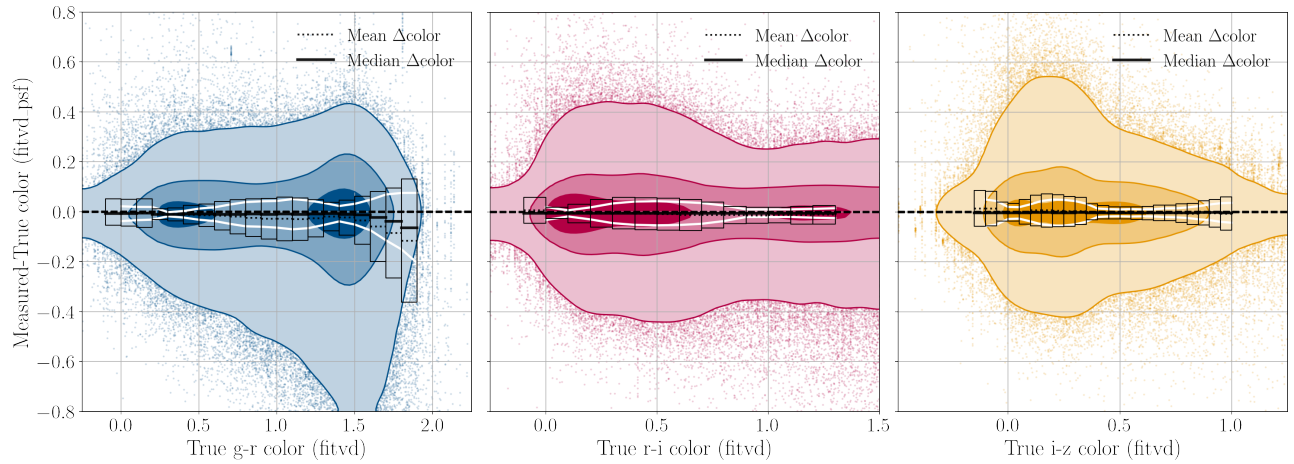
Table 2 presents the photometric accuracy, in the *griz* bands, as presented in Figure 6 and 12. We show the median magnitude difference (in millimag) and scatter on this difference (also in millimag) in the brackets. The latter is computed by measuring the 68% range and taking the scatter to be half the width of this interval.



**Figure 11.** The trend in number density fluctuations  $N/\langle N \rangle$  as a function of various survey observing properties for the Balrog (red) and data (yellow) samples. Analogous to Figure 5 but now for the MAGLIM++ (top) and REDMAGIC (bottom) sample. We use  $N_{\text{SIDE}} = 128$  as this is a smaller sample than the Gold catalog used in Figure 5 and so covers the footprint in a more sparse manner. The  $\chi^2$  between the measurements on data and BALROG (for either lens sample) exhibits  $\chi^2/N_{\text{dof}} \gg 1$  only for the thermal SZ and the background offset, as seen previously. For the correlation of the REDMAGIC samples with stellar density appears more discrepant but is still consistent within  $2\sigma$ . We show results for the Redmagic high-luminosity sample, but note that the high-density samples show consistent behavior as well.



**Figure 12.** Similar to Figure 6 but for the star sample. We compute the differences for the FITVD *psf* magnitudes (instead of the standard model fits) as these are more relevant for stellar studies. The features are all broadly similar to those found in Figure 6, and a detailed numerical comparison is shown in Table 2.



**Figure 13.** Similar to Figure 12 but for the FITVD  $g - r$ ,  $r - i$ , and  $i - z$  colors.