The Astronomical Journal, 166:134 (20pp), 2023 September https://doi.org/10.3847/1538-3881/ace7ca
© 2023. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

Lensing in the Blue. II. Estimating the Sensitivity of Stratospheric Balloons to Weak Gravitational Lensing

Jacqueline E. McCleary[1], Spencer W. Everett[2], Mohamed M. Shaaban[3,4], Ajay S. Gill[4,5], Georgios N. Vassilakis[1], Eric M. Huff[2], Richard J. Massey[6,7,8], Steven J. Benton[9], Anthony M. Brown[7,8], Paul Clark[8], Bradley Holder[4,10], Aurelien A. Fraisse[9], Mathilde Jauzac[6,7], William C. Jones[9], David Lagattuta[7], Jason S.-Y. Leung[4,5], Lun Li[9], Thuy Vy T. Luu[9], Johanna M. Nagy[11,12], C. Barth Netterfield[3,4,5], Emaad Paracha[3], Susan F. Redmond[9], Jason D. Rhodes[2], Jürgen Schmoll[8], Ellen Sirks[6,13], and Sut Ieng Tam[14]

[1] Department of Physics, Northeastern University, 360 Huntington Avenue, Boston, MA, USA; j.mccleary@northeastern.edu
[2] Jet Propulsion Laboratory (JPL), California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA
[3] Department of Physics, University of Toronto, 60 St. George Street, Toronto, ON, Canada
[4] Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON, Canada
[5] Department of Astronomy, University of Toronto, 50 St. George Street, Toronto, ON, Canada
[6] Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK
[7] Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK
[8] Centre for Advanced Instrumentation (CfAI), Durham University, South Road, Durham DH1 3LE, UK
[9] Department of Physics, Princeton University, Jadwin Hall, Princeton, NJ, USA
[10] University of Toronto Institute for Aerospace Studies (UTIAS), 4925 Dufferin Street, Toronto, ON, Canada
[11] Department of Physics, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO, USA, 63130
[12] McDonnell Center for the Space Sciences, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130, USA
[13] Sydney Consortium for Particle Physics and Cosmology, School of Physics, The University of Sydney, NSW 2006, Australia
[14] Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan
Received 2023 March 29; revised 2023 July 11; accepted 2023 July 15; published 2023 August 31

Abstract

The Superpressure Balloon-borne Imaging Telescope (SUPERBIT) is a diffraction-limited, wide-field, 0.5 m, near-infrared to near-ultraviolet observatory designed to exploit the stratosphere's space-like conditions. SUPERBIT's 2023 science flight will deliver deep, blue imaging of galaxy clusters for gravitational lensing analysis. In preparation, we have developed a weak-lensing measurement pipeline with modern algorithms for PSF characterization, shape measurement, and shear calibration. We validate our pipeline and forecast SUPERBIT survey properties with simulated galaxy cluster observations in SUPERBIT's near-UV and blue bandpasses. We predict imaging depth, galaxy number (source) density, and redshift distribution for observations in SUPERBIT's three bluest filters; the effect of lensing sample selections is also considered. We find that, in three hours of on-sky integration, SUPERBIT can attain a depth of $b = 26$ mag and a total source density exceeding 40 galaxies per square arcminute. Even with the application of lensing-analysis catalog selections, we find b-band source densities between 25 and 30 galaxies per square arcminute with a median redshift of $z = 1.1$. Our analysis confirms SUPERBIT's capability for weak gravitational lensing measurements in the blue.

Unified Astronomy Thesaurus concepts: High altitude balloons (738); Galaxy clusters (584); Weak gravitational lensing (1797); Astronomy data analysis (1858)

1. Introduction

The abundance of galaxy clusters as a function of redshift depends sensitively upon both the geometry of the Universe (Holder et al. 2001) and the ongoing mechanism of structure formation via gravitational collapse (Haiman et al. 2001). Cluster number counts provide a statistically significant constraint on cosmological parameters, and as the largest particle colliders in the Universe, galaxy clusters themselves are proving grounds for alternative models of dark matter (Clowe et al. 2004).

Because most of the mass in a cluster is invisible dark matter, a major challenge confronting cluster cosmology is the difficulty of measuring their masses. The most direct method takes advantage of clusters' weak gravitational lensing signal: the small but coherent magnification of background galaxy fluxes and observed distortion of background galaxy shapes. High-quality weak gravitational lensing studies illuminate the relationship between the true masses of galaxy clusters and their observable gas and stars.

In this context, our collaboration will deploy the Superpressure Balloon-borne Imaging Telescope (SUPERBIT): a stratospheric imaging system that will deliver space-quality imaging from the near-ultraviolet to the near-infrared. SUPERBIT has been optimized for measurement of cluster gravitational lensing: the telescope has a $15' \times 23'$ field of view to enable efficient measurements of the weak-lensing signal of galaxy clusters at $z \geqslant 0.05$, and it provides stable, near-diffraction-limited imaging for well-measured galaxy shapes.

Floating above more than 97% of the Earth's atmosphere, the telescope experiences nearly perfect transmission from 280 to 900 nm. The stratosphere also offers low sky backgrounds: Gill et al. (2020) show that SUPERBIT experiences 23.6–25.5 mag arcsec^{-2} in its b filter (365 nm—575 nm), up to three mag arcsec^{-2} fainter than the darkest ground-based sites with 22.7 mag arcsec^{-2}.

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

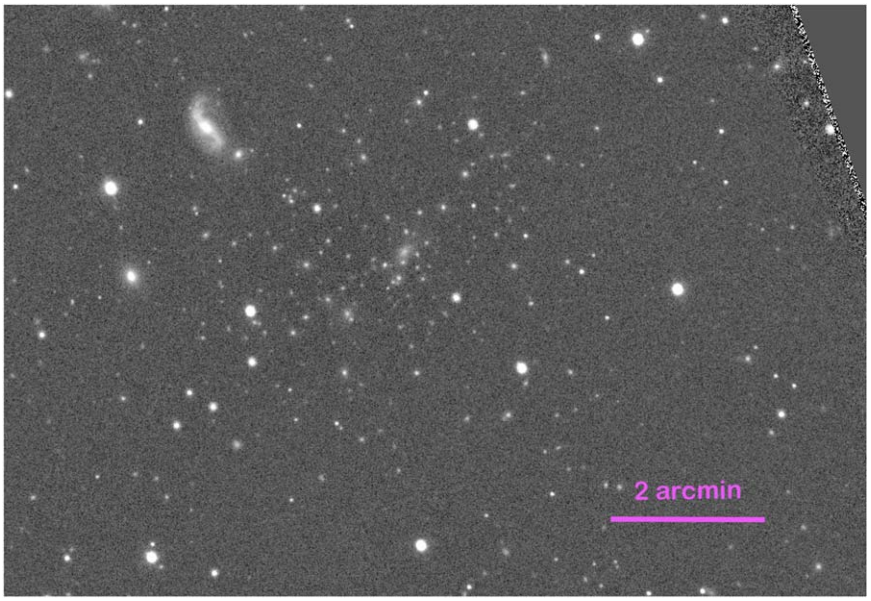

Figure 1. Coadded 15 minute *lum* observation of A2218 made by SUPERBIT during the 2019 engineering test flight.

While most surveys measure weak gravitational lensing at red wavelengths, the dark sky background and diffraction-limited optics in the stratosphere uniquely mean that lensing measurements are more efficient in the blue (Gill et al. 2022; Shaaban et al. 2022).

Beyond weak gravitational lensing measurements, SUPERBIT's deep, blue imaging enables a range of scientific investigations. For example, its near-UV (300–400 nm) photometry spans the Balmer and 4000 Å breaks used to fit galaxy templates for photometric redshift estimation; including NUV photometry can halve uncertainties on the resulting photometric redshifts (Sawicki et al. 2019).

To prepare for SUPERBIT's 2023 science flight, we have created a suite of simulated SUPERBIT galaxy cluster observations with realistic galaxy flux, size, and redshift distributions. PSF models are informed by previous test flights, and background galaxies are gravitationally lensed by foreground cluster halos. We have also developed a weak-lensing analysis pipeline built from modern, publicly available tools like PIFF for PSF characterization, NGMix for galaxy shape measurement, and Metacalibration for galaxy shear calibration.

At a basic level, processing the simulated observations validates our pipeline performance. More interestingly, this procedure enables us to flow down science requirements into an efficient observing strategy. Galaxy clusters have highly localized weak-lensing signal, which makes galaxy number density and average redshift the primary figures of merit for cluster surveys. However, the total number density of galaxies observed is less important than the number that survive cuts on redshift, signal-to-noise, and size for weak-lensing analysis. In this paper, we will forecast imaging depths, source density, and redshift distributions for stratospheric observations in SUPERBIT's near-UV and blue bandpasses.

This paper is organized as follows. We summarize the SUPERBIT platform in Section 2, and lensing theory in Section 3. We describe our galaxy shape measurement pipeline in Section 4, and our mock SUPERBIT observations in Section 5. We present our results in Section 6, provide additional context in Section 7, and conclude with Section 8.

2. The SUPERBIT Observing Platform

2.1. Instrument

SUPERBIT is a 0.5 m mirror telescope that exploits the superpressure balloon capabilities provided by the National Aeronautics and Space Administration (NASA), which offers mid-latitude long-duration balloon flights up to 100 days. SUPERBIT has been developed and iteratively improved through four one-night commissioning flights. Successful recovery after each flight enabled efficient, closed-loop engineering cycles. A complete description of the resulting mechanical, thermal, control systems, and software architecture appears in Redmond et al. (2018), Romualdez et al. (2018), and Li (2016).

The platform consists of a gondola pointing system and an optical assembly that work together to achieve $0\farcs05$ focal plane stability via three successive pointing and stabilization regimes: coarse target acquisition to within $0\farcm5$, fine telescope stabilization at the $0\farcs5$ level, and finally $0\farcs05$ image stabilization at the focal plane. During the most recent test flight in 2019 September, SUPERBIT maintained telescope stability of $0\farcs3$ ($0\farcs5$) over a 5 minute (30 minute) exposure, and image stability of $0\farcs046$ ($0\farcs048$) over a 5 minute (30 minute) exposure. This enabled the first measurements of gravitational lensing from the stratosphere (Tam 2020), using images of A2218 (Figure 1), and defined a fiducial exposure time of 5 minutes for future observations (Romualdez et al. 2020).

Because of the fast development timescales of balloon-borne missions, SUPERBIT has had the ability to upgrade its core technologies between flights. To wit, the 2023 flight camera is a marked improvement over the CCD flown in 2019. The 2023 science camera is a 9600×6422 pixel Sony IMX 455 CMOS detector with $3.76\,\mu$m ($0\farcs141$) square pixels. At operating temperature $-10°$ C, this has low read noise (rms $\sim$1.7 e^-/pixel) and low dark current ($\sim$0.0022 e^-/s/pixel). It is sensitive from 300 to 900 nm; its quantum efficiency (QE) and optical throughput are presented in Gill et al. (2022). Its filter wheel currently includes five broadband filters (*u*, *b*, *g*, *r*, *nir*) plus one very broad filter (*lum*) designed to collect as much light as possible (Table 1 and Figure 2). We also show the *shape* filter, which is

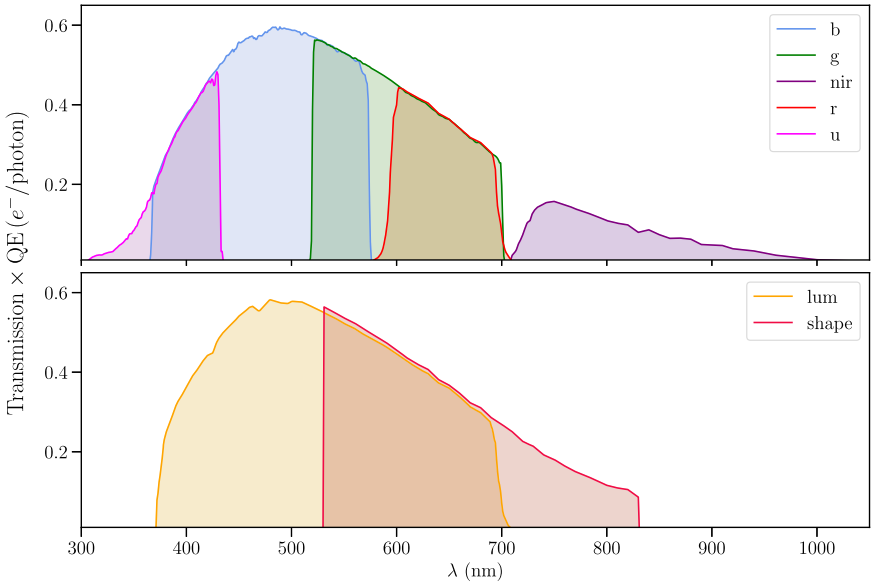

Figure 2. SUPERBIT transmission across the six filters planned for the 2023 flight and the now-deprecated *shape*-band, which will not be flown but is included here for comparative purposes.

very similar in range to the Euclid VIS filter and was at one point designated for galaxy shape measurements (hence the name).

2.2. Survey and Expected Data

During its planned, up to 100 day science flight from NASA's Long Duration Balloon facility in Wanaka, New Zealand (scheduled for 2023 April at the time of writing), SUPERBIT will be able to observe almost anywhere in the Southern Hemisphere and up to 20° North. Target selection will depend on launch date and balloon path, but targets will be automatically drawn from a list of galaxy clusters at redshift $z < 0.5$. These include well-studied clusters from the Hubble Frontier Fields, CLASH, RELICS, LoCuSS, and COSMOS surveys that are required for calibration, plus merging clusters identified principally via bimodality in Chandra X-ray imaging. The clusters have abundant ancillary data: all with X-ray imaging, most with infrared (IR) and radio imaging, and many with substantial investments of ground-based spectroscopy. To these data, SUPERBIT will add deep, wide-field near-UV and optical imaging with angular resolution of 0.″3. With a minimal sample of 45 clusters and assuming a per-cluster scatter in mass of 20% (Becker & Kravtsov 2011), this data can yield weak-lensing masses with an ensemble M_{200c} fractional uncertainty of $0.2/\sqrt{45} = 0.03$. Based on SUPERBIT's 2018 and 2019 test flights, calculations in Shaaban et al. (2022), and results in this work, each cluster target will be observed for 3 hr (36 ×300 s exposures) in b, plus shorter integrations in the u and g bands.

Table 1
SUPERBIT Bandpasses and Sky Backgrounds

Filter Name	Wavelength Range (nm)	Pivot Wavelength (nm)	Sky Brightness (e^- s^{-1} pix^{-1})
u	300–435	395	0.029
b	365–575	476	0.052
g	515–705	597	0.052
r	570–720	640	0.030
nir	706–1100	814	0.064
lum	370–710	522	0.084
shape	530–830	650	0.15

Notes. Summary of the 2023 flight filters and expected sky brightnesses in each. The shape filter is deprecated and included in this analysis for comparison purposes.

3. Gravitational Lensing and Metacalibration

3.1. Weak Gravitational Lensing Formalism

Gravitational lenses like galaxy clusters introduce an isotropic magnification of background galaxies and percent-level distortions in their shapes. The magnification of galaxy images is described by the convergence κ, a scalar quantity equal to the Laplacian of the gravitational potential of the lens projected along the line of sight. The convergence κ can be related to the surface mass density of the galaxy cluster, Σ, as

$$\kappa \equiv \frac{1}{2}\nabla^2\Psi(\theta) = \frac{\Sigma}{\Sigma_{\rm crit}}; \quad \Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l}D_{\rm ls}}, \tag{1}$$

where the critical surface mass density $\Sigma_{\rm crit}$ of the lens depends on the angular diameter distances to the background galaxy $D_{\rm s}$, the lens $D_{\rm l}$, and the lens and source $D_{\rm ls}$, respectively.

The distortion of galaxy images introduced by gravitational lenses is represented as a complex shear γ:

$$\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\phi}. \tag{2}$$

Distortion along the real axes (x/y) is described by the γ_1 component of shear; the γ_2 component describes the galaxy image distortion along axes rotated through $\pi/4$ radians. The shear γ can be related to the cluster gravitational potential Ψ as

$$\gamma(\boldsymbol{\theta}) = \boldsymbol{D}\Psi, \tag{3}$$

$$\boldsymbol{D} = \frac{\partial_1^2 - \partial_2^2}{2} + i\partial_1\partial_2. \tag{4}$$

Observations of gravitationally lensed galaxies actually return the reduced shear $\boldsymbol{g}$

$$\boldsymbol{g} = \frac{\gamma}{1-\kappa} = g_1 + ig_2, \tag{5}$$

where the variables g_1 and g_2 in Equation (5) are the polarization states of background galaxies with reduced shear $\boldsymbol{g}$.

Irrespective of the presence of a gravitational lens, the shapes of galaxies measured on an image can be characterized by an ellipticity e:

$$\boldsymbol{e} = e_1 + ie_2, \tag{6}$$

$$e_1 = e\cos(2\theta), \quad e_2 = e\sin(2\theta), \tag{7}$$

$$e = \frac{a^2 - b^2}{a^2 + b^2}, \tag{8}$$

where a and b are the major and minor axes of the galaxy image ellipse. The shear γ can be extracted from galaxy ellipticities e in the weak-lensing regime, where the distortion introduced by the lens is much smaller than the galaxy images themselves, i.e., where $\kappa, \gamma \ll 1$. In that case, in the absence of intrinsic alignments and for source galaxies at the same redshift,

$$\gamma \simeq g \simeq \frac{\langle e \rangle}{2\mathscr{R}}. \tag{9}$$

The factor $\mathscr{R}$ encodes the shear response factor $1 - \sigma_e^2$. Because the lensing potential induces curl-free distortions in galaxy images, we estimate the reduced shear about a point on the sky with the tangential ellipticity:

$$g_{\tan} = -(g_1\cos(2\phi) + g_2\sin(2\phi)), \tag{10}$$

where ϕ is the azimuthal angle from the fiducial center of mass to the galaxy.

Because it is a curl-free statistic, in analogy with electromagnetism, Equation (10) is sometimes called the E-mode signal. A divergence-free statistic, the B-mode or cross shear, is obtained by rotating Equation (10) through $\pi/4$ radians:

$$g_\times = g_2\cos(2\phi) - g_1\sin(2\phi). \tag{11}$$

Galaxy shapes are also convolved with the point-spread function (PSF) of the telescope and atmosphere. PSFs tend to circularize galaxy shapes, diluting the real weak-lensing signal, while the anisotropic components introduce ellipticities into the galaxy shapes that mimic weak-lensing shear. Accurate shear inference thus requires that the PSF be modeled and deconvolved from galaxy shape measurements. Readers interested in a comprehensive review of galaxy cluster weak gravitational lensing, including considerations of the PSF, may consult Umetsu (2020).[15]

3.2. Metacalibration

In real measurements, the measured galaxy shears g_1, g_2 are biased estimators of the underlying shear distribution and need to be converted into a true estimator for the weak-lensing shear $g_{\tan}$. This is generally accomplished by dividing each galaxy's ellipticity by an appropriate "shear responsivity factor" $\boldsymbol{R}$, which characterizes the response of the galaxy shape estimator $\hat{\boldsymbol{g}}$ to an applied shear γ (Hirata 2005):

$$\langle \hat{\boldsymbol{g}} \rangle = \langle \hat{\boldsymbol{g}} \rangle|_{\gamma=0} + \langle \boldsymbol{R}\gamma \rangle + O(\gamma^2) \tag{12}$$

$$\approx \langle \boldsymbol{R}\gamma \rangle, \tag{13}$$

where in the absence of an external shear field, the average ellipticity should be zero.

Image simulations are often used to obtain the shear calibration (e.g., Fenech Conti et al. 2017; Mandelbaum et al. 2018), but they face the usual difficulties in replicating all the effects that affect real images. Instead, we use the Metacalibration algorithm, which calibrates shear estimators from the galaxy image data itself, without requiring significant prior information about galaxy properties. Metacalibration's data-driven approach is particularly valuable in a new survey like SUPERBIT.

Metacalibration introduces an artificial shear to images and calculates how the shear estimator responds to that applied shear. More specifically, the original galaxy image is deconvolved from the PSF and then sheared by some amount γ along each ellipticity component g_i. The sheared image is reconvolved with a function slightly larger than the original PSF to suppress noise amplified by the deconvolution process, and measurement of $\hat{\boldsymbol{g}}$ is repeated. The shear responsivity $\boldsymbol{R}$ is then obtained through the finite difference derivative:

$$\boldsymbol{R}_{k,l} = \frac{\hat{g}_k^+ - \hat{g}_k^-}{\Delta\gamma_l}, \tag{14}$$

where $\hat{g}^+$ is the measurement made on an image sheared by $+\gamma_l$ and $\hat{g}^-$ is the measurement made on an image sheared by $-\gamma_l$ for all ellipticity components k. The responsivities can be computed for every galaxy in an observation catalog, but they are very noisy because the ellipticity estimators $\hat{\boldsymbol{g}}$ themselves are noisy. Thus, in practice, a shear estimate is obtained by dividing the galaxy ellipticity estimator by the mean responsivity over the entire galaxy sample:

$$\langle \hat{\gamma} \rangle = \langle \boldsymbol{R} \rangle^{-1} \langle \hat{\boldsymbol{g}} \rangle. \tag{15}$$

Estimation of weak-lensing shear commonly requires selection cuts on quantities like galaxy size and signal-to-noise ratio. The probability that a galaxy passes selection cuts changes after the application of an artificial shear. The responsivity then includes both the shear response and the effect of sample selections. We continue to follow the formalism of Sheldon & Huff (2017) and break up the responsivity into two components:

$$\langle \boldsymbol{R} \rangle = \langle \boldsymbol{R}_\gamma \rangle + \langle \boldsymbol{R}_S \rangle, \tag{16}$$

where brackets denote the average over galaxies $k = 1 \ldots N_{\rm gals}$, $\langle \boldsymbol{R}_\gamma \rangle$ captures the ensemble response of galaxy shapes to an applied shear, and $\langle \boldsymbol{R}_S \rangle$ represents the response of the selections to an applied shear.

4. Shape Measurement Pipeline

In anticipation of SUPERBIT's 2023 science flight, we have developed a galaxy shape measurement and weak-lensing analysis pipeline that employs state-of-the-art algorithms, such as NGMix for optimal estimation of galaxy shapes (Sheldon 2015) and Metacalibration to correct for multiplicative shear bias (Huff & Mandelbaum 2017; Sheldon & Huff 2017; also see 3.2). We provide an overview of our pipeline below. Upon acceptance of the paper, we intend to make the pipeline public.

[15] Some authors also use the opposite sign convention, $g_\times = g_1\sin(2\phi) - g_2\cos(2\phi)$.

The pipeline is divided into three modules: creation of the input files for galaxy shape fitting (`medsmaker`); galaxy shape fitting and shear bias correction (`metacal`); and calculation of the galaxy clusters' tangential and cross-shear profiles (`shear_profiles`). For ease of use, the pipeline has code to auto-generate the configuration files needed to run this pipeline from beginning to end, based only on a few user inputs.

4.1. Creation of Shape Measurement Input files

The input for NGMix and Metacalibration is a multi-epoch data structure (MEDS[16]): a kind of FITS binary table with an entry for every object detected in an observation. Each object's MEDS entry contains the following: a postage-stamp cutout of the object, a rendering of the point-spread function (PSF) at the location of the object, a weight, a segmentation map, and a bad pixel mask for every exposure in which the object was detected (Jarvis et al. 2016).

In our pipeline, MEDS files for SUPERBIT observations are created with the `medsmaker` module. Much of the "standard operating procedure" for astronomical imaging is implemented in `medsmaker`; we detail the particulars here for reference in future analyses.

4.1.1. Detection Catalog

Image data supplied to `medsmaker` are assumed to be calibrated CMOS or CCD imaging data for which bias subtraction and flat-fielding have already been performed. Exposure weight maps and bad pixel masks are required as well.

The AstrOmatic tool SWarp is used to combine single-epoch exposures into a deep detection image from which the master observation catalog—the basis of the MEDS file—is obtained with SExtractor (Bertin et al. 2002). To maximize the number of sources detected, we set a relatively low detection threshold of 1.5σ. This necessarily generates spurious detections that do not correspond to galaxies in any exposure. Rather than cut these items out of the MEDS file, which risks introducing an uncontrolled shear selection bias, spurious sources with no cutouts are flagged to be skipped during shape fitting. Segmentation maps and catalogs for single-epoch exposures are also generated with SExtractor; segmentation maps go into the MEDS, and single-epoch catalogs are used to identify stars for PSF model fitting.

4.1.2. PSF Estimation

As discussed in Section 3.1, accurate shear inference hinges upon the successful deconvolution of the observation's PSF and galaxy shape measurements. Before their light passes through the atmosphere and telescope, stars are effectively point sources, so the shape and size of their surface brightness profiles axiomatically define the PSF at that location. Using stars as fixed points, the PSF can be interpolated across the rest of the image.

Star catalogs for PSF modeling are generated with simple selections to the single-exposure detection catalogs based on SExtractor `CLASS_STAR`, a minimum signal-to-noise ratio, and a magnitude range. A sample star catalog is highlighted in Figure 3. Should greater sample purity be required, we have incorporated into `medsmaker` an option to cross-reference candidate stars against a reference catalog, and we have also added the capability to query the Gaia star database (Brown et al. 2018) on the fly. Though the Gaia catalog is relatively shallow, the high purity of the Gaia catalog avoids the problem of star–galaxy confusion. The use of the Gaia catalog for PSF fitting is also considered in Bertin (2011).

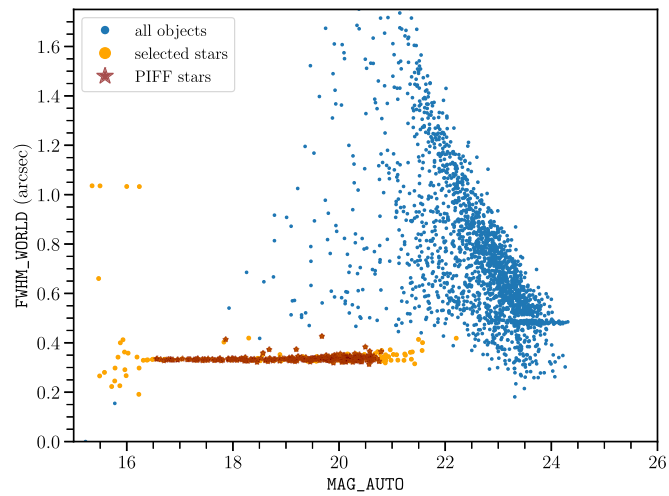

Figure 3. Size–magnitude diagram for objects detected in a single-epoch exposure. Blue points represent all sources; orange points show stars supplied to the PIFF software; dark red star markers show stars selected by the PIFF software for PSF modeling.

We model SUPERBIT PSFs with the recently introduced PIFF software package.[17] Like most PSF fitters, PIFF takes an input catalog of stars, fits their surface brightness profiles with a user-specified model, interpolates the PSF parameters across the FOV following some schema, and saves the resulting description of the observation's PSF to file. A notable feature of PIFF is that PSF models are expressed in sky coordinates, as opposed to the pixel coordinates commonly used in other PSF modeling software. Because high-frequency components of the PSF, e.g., astrometric distortion, vary more smoothly across the detector FOV when considered in sky coordinates, PIFF avoids the "size bias" (mismatch between the real and model PSF size) that can affect other PSF fitting software (Jarvis et al. 2021).

Following the DES Y3 approach, we use the `PixelGrid` model, which treats the PSF profile as a two-dimensional grid of points smoothed by a Lanczos kernel with $n = 3$. The total number of free model parameters is then equal to the number of pixels in the grid. We also follow the DES Y3 approach to interpolate the PSF model across the FOV by using the `BasisPolynomial` scheme, which solves for the `PixelGrid` model parameters (pixel fluxes) in terms of the interpolation coefficients.

PSF model residuals are quantified with the ρ statistics introduced in Rowe (2010) and expanded in Vogelsberger & Zavala (2016). The ρ statistics below summarize the spatial correlations of size and ellipticity residuals between the real (star) and model PSFs; large values imply a systematic error in the model.

$$\rho_1(\theta) \equiv \langle \delta e^*_{\rm PSF}(\boldsymbol{x}) \delta e_{\rm PSF}(\boldsymbol{x} + \boldsymbol{\theta}) \rangle, \tag{17}$$

$$\rho_2(\theta) \equiv \langle e^*_{\rm PSF}(\boldsymbol{x}) \delta e_{\rm PSF}(\boldsymbol{x} + \boldsymbol{\theta}) \rangle, \tag{18}$$

[16] https://github.com/esheldon/meds/wiki/MEDS-Format

[17] https://rmjarvis.github.io/Piff/_build/html/overview.html

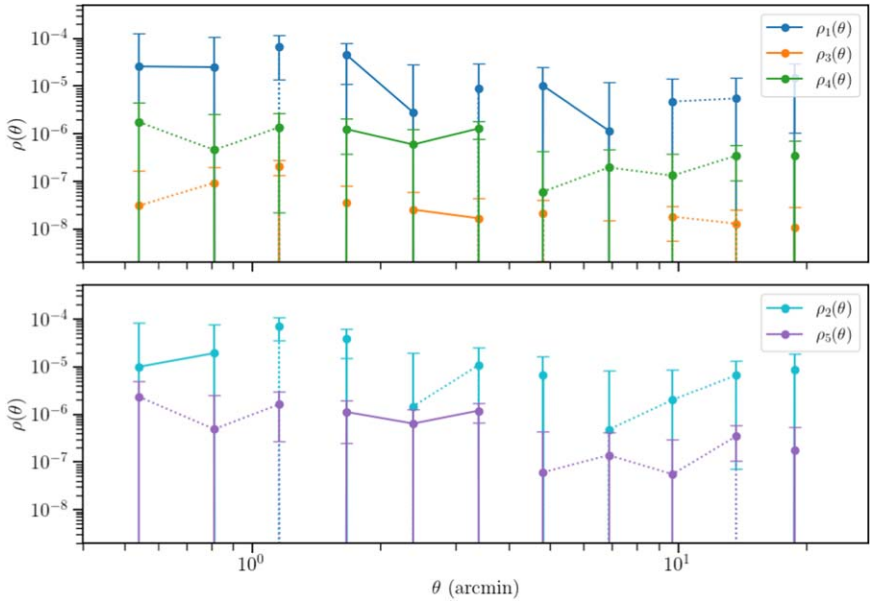

Figure 4. Example of ρ statistics computed from PSF model residuals for a simulated single-epoch exposure. Negative correlations are shown in absolute value and connected to neighboring bins by dotted lines. Isolated points show correlations that changed from positive to negative or vice versa.

$$\rho_3(\theta) \equiv \left\langle \left(e^*_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x})\left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle, \quad (19)$$

$$\rho_4(\theta) \equiv \left\langle \delta e^*_{\rm PSF}(\boldsymbol{x})\left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle, \quad (20)$$

$$\rho_5(\theta) \equiv \left\langle e^*_{\rm PSF}(\boldsymbol{x})\left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle. \quad (21)$$

Here, $e_{\rm PSF}$ is the ellipticity of the real PSF, i.e., the star ellipticity; $T_{\rm PSF}$ is the size of the real PSF; $\delta e_{\rm PSF}$ is the difference between the ellipticity of the real and model PSFs at position $\boldsymbol{x}$; and $\delta T_{\rm PSF}$ is the difference between the sizes of the real and model PSFs at position $\boldsymbol{x}$. Brackets denote averages over all pairs within a separation $\boldsymbol{\theta}$, and asterisks denote complex conjugates. An example of ρ statistics plotted as a function of distance between neighboring stars is shown in Figure 4.

We will also compute the two-point spatial correlations of star and galaxy ellipticities:

$$C_i = \langle e_i(\boldsymbol{x}) \times e_i(\boldsymbol{x}+\boldsymbol{\theta})\rangle, \;\; i = \{1, 2\}, \quad (22)$$

where e_i is the ith ellipticity component of a PSF-corrected star or galaxy at position $\boldsymbol{x}$. The correlations $C_{1/2}$ of galaxy–galaxy pairs should have a relatively high amplitude, reflecting the correlated shear introduced by the galaxy cluster. However, the $C_{1/2}$ functions should vanish when evaluated over star–galaxy pairs, as there should be no correlation between the shapes of circularized stars and PSF-corrected galaxy shapes (McCleary et al. 2015; McCleary 2020).

4.1.3. Multi-epoch Data Structure

With object cataloging and PSF modeling complete, an instance of the MEDS class is created. For each object in the detection catalog, an entry in the MEDS is made to hold a binary table with postage-stamp cutouts from the single-epoch exposures, a PSF model rendering, weights, masks, and segmentation maps. The MEDS is also populated with objects' celestial and image coordinates, original catalog ID number, and WCS information.

4.2. Galaxy Shape Measurement

We measure galaxy ellipticities using the NGMix[18] package, which implements Gaussian mixture models to recover the shear from 2D images with good accuracy for even very low-S/N galaxies. Rather than a single-point estimate, NGMix returns an estimator of the shape from an ensemble of measurements of the galaxy—generally every epoch, in every filter in which the galaxy was observed. We implement NGMix with a Python wrapper script that creates an instance of the NGMIXMEDS class and populates it with observation information for all sources in the supplied MEDS file.

Because of the high source density of SUPERBIT observations, many of the postage stamps in the MEDS contain not one but two sources: the galaxy of interest and an interloping star or galaxy. Left unmasked, the presence of interlopers introduces a large scatter in the final tangential shear measurements, as NGMix treats both sources as a single galaxy. Following the solution used in DES SV and Y1, we mitigate interlopers using so-called überseg masks. These masks are generated using the detection (coadd) image's SExtractor segmentation maps, projected onto the plane of single-epoch exposures. Pixels in the MEDS weight cutouts are set to zero if they are more closely associated with an interloping object than with the galaxy of interest (Jarvis et al. 2016).

4.3. Weak-lensing Shear Profile Calculation

Tangential and cross-shear profiles of galaxy clusters are produced in the `shear_profile` module of our pipeline. At this stage, redshift information is added to the galaxy shape fit catalog, selection cuts (including redshift selection) are applied, and Metacalibration responsivities are calculated and applied to galaxy shapes. The ($g_{\rm tan}$,$g_\times$) shears are computed from (g_1, g_2)

[18] https://github.com/esheldon/ngmix

and then averaged in radial bins from the cluster center. Further details are provided below.

4.3.1. Creation of Galaxy Shear Catalog

A top-level catalog with galaxy shape parameters, responsivity components, detection parameters, and redshifts is generated by joining the SExtractor and NGMix catalogs on-sky coordinates (α, δ) and then matching to a third catalog with redshift information. The galaxy shear catalog is then built from galaxies meeting the following criteria:

1. $10 < \mathrm{S/N} < 1000$, where the signal-to-noise measure is the galaxy `s2n` from NGMix fits;
2. Galaxy size (really, area) $0 < \mathtt{T} < 10$, where `T` is in units of arcsec^2;
3. Ratio of galaxy to PSF size ($\mathtt{T/Tpsf}$) > 1.0;
4. Galaxy redshift z_{gal} greater than the cluster redshift z_{cl}.

When appropriate, i.e., for a nearly round PSF, we base size and signal-to-noise cuts on the "roundified" size and signal-to-noise `T_r` and `s2n_r`. These selections are based on those in the DES analyses (Jarvis et al. 2016; Zuntz et al. 2018; Gatti et al. 2021).

Shear and selection responsivities are calculated from the NGMix `1p/1m/2p/2m` shape fit parameters to produce responsivity-corrected galaxy shear (g_1, g_2). Selection of background galaxies through redshift cuts is included within the calculation of the selection responsivity. Galaxies are weighted by their shape fit covariances $\sigma^2_{g1/2}$, and a shape noise of $\sigma_{\mathrm{SN}} = 0.26$ is based on our own fits to COSMOS galaxies:

$$\frac{1}{\sigma^2_{\mathrm{SN}} + \sigma^2_{g1} + \sigma^2_{g2}}. \tag{23}$$

4.3.2. Shear Profile Calculation

Response-corrected (g_1, g_2) moments are transformed into tangential and cross ellipticities (g_{tan},g_x) using the galaxy image coordinates (x_i, y_i) and the user-specified location of the galaxy cluster center (x_c, y_c):

$$g_{\mathrm{tan}} = -(g_1 \cos(2\phi) + g_2 \sin(2\phi)), \tag{24}$$

$$g_\times = g_1 \sin(2\phi) - g_2 \cos(2\phi), \tag{25}$$

$$\phi = \arctan\left(\frac{y - y_c}{x - x_c}\right). \tag{26}$$

Using the `DescrStatsW` module of the `statsmodels` Python library, the code then computes weighted averages of (g_{tan},$g_\times$) in radial bins about the cluster center. The final outputs are a shear profile catalog with averaged (g_{tan},$g_\times$) and a plot of the cluster's cross and tangential shear profiles. Two examples from simulated observations are shown in Figure 5.

4.4. Shear Bias Estimator

While we do not attempt shear calibration in this analysis, we have developed an estimator for shear bias tailored to cluster shear tangential profiles. It is included in the pipeline to support future efforts.

Borrowing the language of large cosmological surveys, we express the difference between input (simulated) and output (measured) tangential shears as a shear bias α, which we quantify with a maximum likelihood estimator $\hat{\alpha}$. Each galaxy's measured shear g_{tan} is considered a random sample of the true halo shear g_{true} at the galaxy's position. The joint probability distribution (likelihood) $\mathcal{L}$ of the data then follows a multivariate Gaussian in which the mean shear $\langle g_{\mathrm{tan}} \rangle$ converges to $\alpha\, g_{\mathrm{true}}$ in a radial bin, where

$$\alpha = \frac{\langle g_{\mathrm{tan}} \rangle}{\langle g_{\mathrm{true}} \rangle}. \tag{27}$$

If the measurements g_{tan} are unbiased measurements of the true shear g_{true}, then $\alpha \equiv 1$ and $\langle g_{\mathrm{tan}} \rangle = g_{\mathrm{true}}$. Any $\alpha \neq 1$ indicates a biased measurement.

To obtain an optimal estimator $\hat{\alpha}$ for the tangential shear bias α, we express the log-likelihood as

$$\log \mathcal{L} = -[\boldsymbol{g}_{\mathrm{tan}} - \alpha\, \boldsymbol{g}_{\mathrm{true}}]\, \frac{\boldsymbol{C}^{-1}}{2} [\boldsymbol{g}_{\mathrm{tan}} - \alpha\, \boldsymbol{g}_{\mathrm{true}}]^{\mathrm{T}}. \tag{28}$$

The measurement uncertainties of the data $\boldsymbol{g}_{\mathrm{tan}}$ are expressed as the covariance $\boldsymbol{C}$. We note that matrix quantities are written in boldface. Differentiating Equation (28) with respect to α, setting the result to zero, and then solving for α, we obtain the maximum likelihood estimator for shear bias:

$$\hat{\alpha} = \frac{\boldsymbol{g}^{\mathrm{T}}_{\mathrm{true}}\, \boldsymbol{C}^{-1}\, \boldsymbol{g}_{\mathrm{tan}}}{\boldsymbol{g}^{\mathrm{T}}_{\mathrm{true}}\, \boldsymbol{C}^{-1}\, \boldsymbol{g}_{\mathrm{true}}}. \tag{29}$$

The uncertainty on $\hat{\alpha}$ is given by the the Cramér–Rao bound:

$$\sigma^2_{\hat{\alpha}} = \frac{1}{\boldsymbol{g}^{\mathrm{T}}_{\mathrm{true}}\, \boldsymbol{C}^{-1}\, \boldsymbol{g}_{\mathrm{true}}}. \tag{30}$$

An unbiased cluster tangential shear measurement has $\hat{\alpha} = 1 \pm \sigma_{\hat{\alpha}}$. The goal for SUPERBIT shear calibration will be a shear bias consistent with unity within the mass uncertainty of the full cluster sample (2%–3%). As the value of $\hat{\alpha}$ calculated from a large number of simulations is a useful metric for shear calibration, the pipeline contains tools for the calculation of the average $\hat{\alpha}$ as well. An example setup is shown in Figure 6. The SUPERBIT shear calibration analysis will be presented in S. Everett et al. (2023, in preparation).

5. Simulated Galaxy Cluster Observations

To plan observations and calibrate the analysis pipeline for SUPERBIT's science flight, we have used GalSim (Rowe et al. 2015) to produce mock observations of galaxy clusters. These simulate 3 hr observations in each of SUPERBIT's *u*, *b*, *lum*, and *shape* filters, divided into `n_exp` = 36 individual, dithered exposures of `exp_time` = 300 s. The central region of a full 3 hr observation of one simulated cluster is shown in Figure 7. For each cluster, we create 30 mock sets of images with independent distributions of stars, cluster member galaxies, and field galaxies both in front of and behind the galaxy cluster. We store a truth catalog containing the objects' positions, sizes, fluxes, redshifts, and applied lensing distortion (for galaxies behind the cluster). For distortion calculations, we set $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$.

Simulated clusters have mass $M_{200c} = 4.1 \times 10^{14}\, M_\odot\, h^{-1}$ (the mean mass of clusters in the SUPERBIT target list) and three redshifts ($z = 0.059$, 0.3, 0.45). Cluster mass distributions are modeled with Navarro, Frenk, and White (1996; NFW)

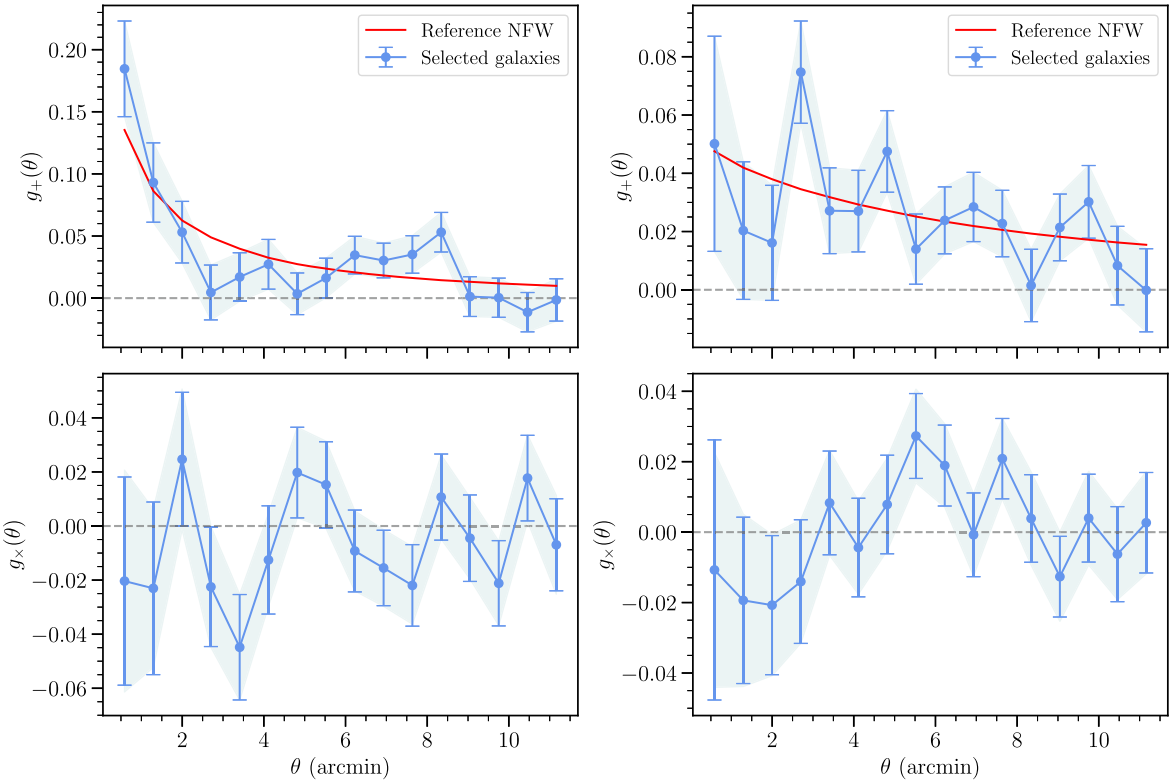

Figure 5. Examples of tangential (top panels) and cross-shear profiles (bottom panels) for single galaxy cluster realizations in *lum*. Left: Realization of a cluster with $M_{200c} = 4.1 \times 10^{14}\ M_\odot\ h^{-1}$ and $z = 0.45$. Right: Realization of a cluster with $M_{200c} = 4.1 \times 10^{14}\ M_\odot\ h^{-1}$ and $z = 0.059$. The selections described in Section 4.3.1 were applied, yielding 31.3 and 31.6 galaxies arcmin^{-2} for the $z = 0.059$ and $z = 0.3$ clusters, respectively.

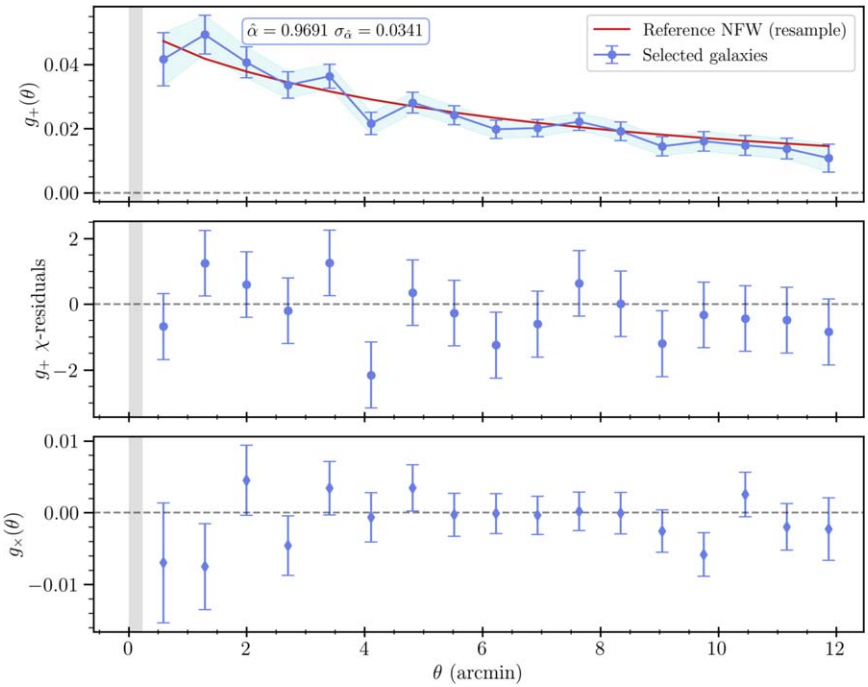

Figure 6. Example of the setup for shear bias (α) estimation. Top panel: mean tangential shear. Blue points are the mean galaxy shears in a radial bin for 30 realizations of $z = 0.059$ clusters; error bars are standard errors of the mean. The input NFW tangential shear is plotted as a red line. Middle panel: difference between measured and input tangential shears. Bottom panel: cross shear. The gray-shaded regions indicate the regime where the linearized implementation of Metacalibration is invalid. The SUPERBIT shear calibration analysis will be presented in a forthcoming paper by S. Everett et al.

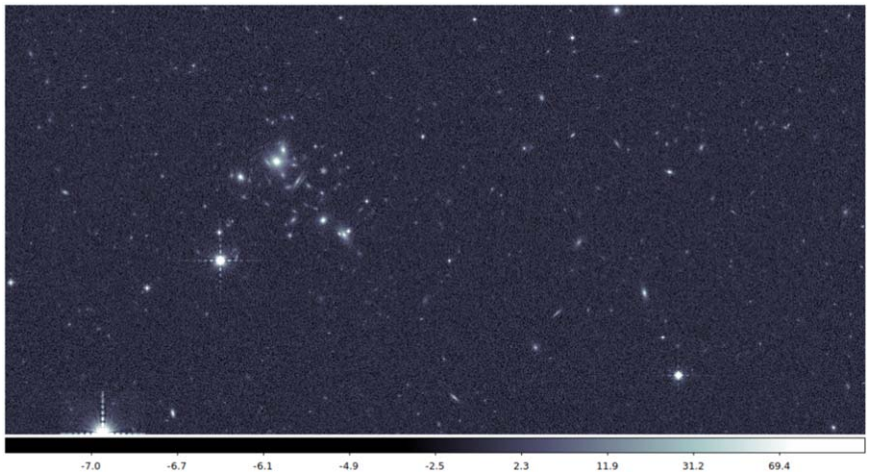

Figure 7. Simulated three-hour SUPERBIT observation of a galaxy cluster in the b band using the jitter+optics ("optics-on") PSF. The central $6' \times 3'$ region of the simulation is shown. Pixel values represent background-subtracted ADU.

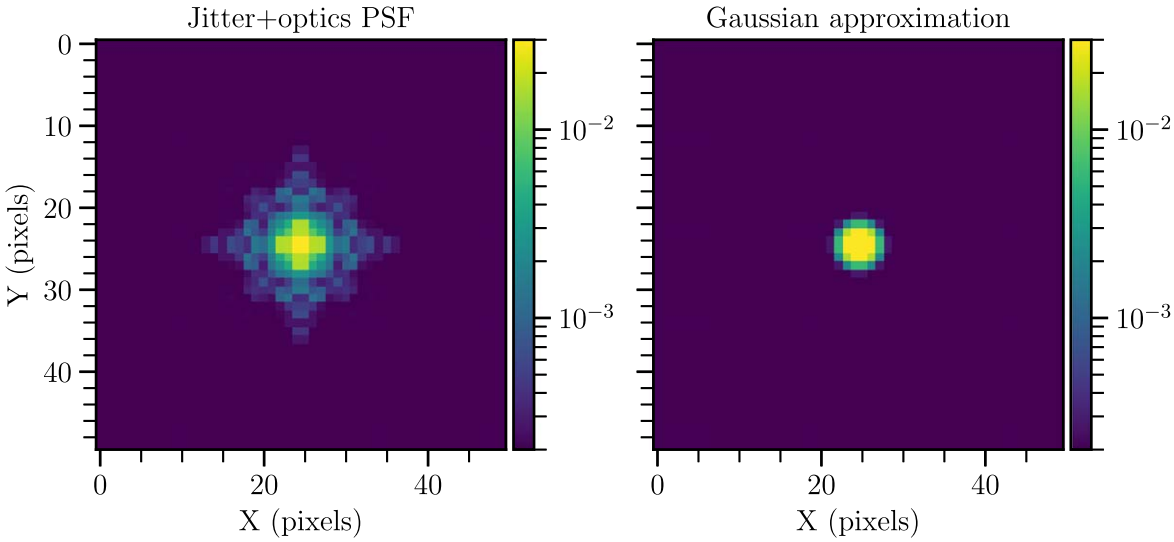

Figure 8. Left: rendering of SUPERBIT effective (pixel-convolved) PSF in the b filter. Right: Gaussian approximation to b PSF with FWHM $= 0\farcs315$. In both panels, intensity values have been normalized so that the total flux in each image sums to unity; the color scale represents relative flux or intensity.

density profiles:

$$\rho(r) = \frac{\rho_0}{\frac{r}{R_S}\left(1 + \frac{r}{R_S}\right)^2}. \tag{31}$$

5.1. Point-spread Function and Stars

The SUPERBIT PSF is well modeled with two components (jitter+optics or "optics-on") that combine the residual telescope jitter measured during test flights with spherical aberrations for the SUPERBIT optical train derived with ray-tracing software. However, we base survey forecasts on an "optics-off" Gaussian approximation to the PSF, because of a temporary limitation in the NGMix method that we use for shape measurement. NGMix does not currently include templates for diffraction-limited PSFs, and it tends to overestimate the PSF size $T_{\rm PSF}$ by about $\sim$50%. Given that almost all weak-lensing analyses select galaxies based on their size relative to PSF size ($T/T_{\rm PSF}$), this artificially decreases the source density. An extension to the NGmix template set will be presented in the shear calibration paper by S. Everett et al. Meanwhile, we implement Gaussian approximations to the jitter+optics PSF, with FWHM of $0\farcs278$ in u, $0\farcs315$ in b, $0\farcs333$ in *lum*, and $0\farcs37$ in *shape*. These values are the combination of the jitter FWHM of $0\farcs05$ and the FWHM obtained with ray-tracing models of the SUPERBIT optical train in each bandpass. The "optics-on" and "optics-off" versions of the b-band PSF are compared in Figure 8.

We simulate the spatial clustering and magnitude distribution of foreground stars by sampling Gaia DR2 catalogs (Gaia Collaboration 2018) at the RA and Dec coordinates of 52 of SUPERBIT's target clusters. We convert Gaia $G/G_{\rm BP}$ fluxes to SUPERBIT AB fluxes, then choose one of these star fields at random for each realization of mock images. Because the star fields span a range of galactic latitudes, this effectively marginalizes over stellar number density when predicting shear biases.

5.2. Galaxies

Our simulation input catalog is a hybrid of two different COSMOS catalogs. A full description for generating the mock

SUPERBIT source galaxy catalog will appear in a paper by A. Gill et al. (2023, in preparation); a high-level overview is presented here.

The baseline is the UltraVISTA-DR2 region of the COSMOS 2015 catalog (Laigle et al. 2016), which contains 518,404 galaxies with high-quality redshifts spread out over 1.5 deg^2. The number density, redshift, and magnitude distributions of our simulated background source galaxies are drawn directly from COSMOS 2015. To convert COSMOS 2015 fluxes to their equivalent in SUPERBIT bandpasses, we access the spectral energy distribution fits from the EL-COSMOS project (Saito et al. 2020), convolve these with the wavelength-dependent OTA throughput, detector QE, and filter transmission curves, and finally integrate counts over the collecting area of the SUPERBIT mirror (see Section 2.1).

We add morphological information to COSMOS 2015 with a heuristic match in luminosity ($m_{\rm C15}$) and redshift ($z_{\rm C15}$) to galaxies in the GalSim-COSMOS F814W<25.2 catalog. In our simulations, galaxies are drawn as single-component Sersíc profiles with half-light radius $R_{1/2}$ and index n, position angle ϕ, and major-to-minor axis ratio q.

Parameter values are chosen with the following algorithm:

1. In the best-case scenario of $z_{\rm C15} < 5$ and $18 < m_{\rm C15} < 25.2$, a source is selected from the GalSim-COSMOS that best matches the COSMOS 2015 galaxy and its shape parameters are assigned to the COSMOS 2015 galaxy.
2. If $z_{\rm C15} < 5$ and $25.2 < m_{\rm C15} < 30$, a source is selected from GalSim-COSMOS that best matches the COSMOS 2015 galaxy redshift and used to set the half-light radius. The Seríc index n is selected from a uniform distribution $U[0, 4]$. The position angle ϕ is also chosen from a uniform distribution $U[-2, 2]$ radians. The axis ratio q is selected from a uniform distribution $U[0.1, 1]$.
3. If $z_{\rm C15} > 5$ but $18 < m_{\rm C15} < 25.2$, n, q, and ϕ are chosen based on the closest match in $m_{\rm F814W}$ between GalSim-COSMOS and COSMOS 2015. The half-light radius $R_{1/2}$ is randomly chosen from a uniform distribution $U = [5, 20]$ pixels (plate scale $= 0\farcs03$ pixel^{-1}).
4. All other $z_{\rm C15}$ and $m_{\rm C15}$ cases correspond to outliers with no equivalents in the GalSim-COSMOS catalog. In this instance, all galaxy shape parameters are chosen from uniform distributions.

While GalSim does have ready-made galaxy catalogs available, their maximum depth of F814W = 25.2 would limit our ability to simulate deep SUPERBIT observations. Moreover, the number of galaxies with photometric redshifts has increased since 2007 (the year of the original GalSim-COSMOS catalog's release). These limitations motivated us to create our own galaxy catalog for simulations.

5.3. Simulation Procedure

First, we initialize the random number generators for stars, source galaxies, cluster galaxies, noise, and dither offsets, passing any seeds set in the GalSim configuration file.

The blank exposure is represented with an instance of the GalSim object (GSObject) `ImageF` set to match the SUPERBIT instrument properties of Section 2.1, and it includes a model world coordinate system (WCS). The image is filled with the raw sky background derived in Gill et al. (2020); approximately 45 ADU for a 300 s exposure in the b filter.

The cluster lensing potential is represented with an instance of the `NFWHalo` class. The halo concentration is set to 4 in all simulations.

For each source galaxy to be injected into the image, the following process is repeated. A galaxy entry is randomly drawn from the SUPERBIT mock galaxy catalog and assigned some R.A. and decl. on the observation. The galaxy's photometric redshift, shape parameters, and flux in the SUPERBIT filter of choice are accessed from our mock galaxy catalog. The galaxy image is created as an instance of `Sérsic` with shape parameters set to the catalog values. To convert the catalog flux from units of photoelectrons s^{-1} to equivalent observed analog-to-digital units (ADU), we multiply the flux by the exposure time and the gain.

The source galaxy object is sheared and magnified according to its redshift with the `NFWHalo` object, or if the source galaxy redshift is below the cluster redshift, the galaxy's magnification and distortion are set to 1 and 0, respectively. The galaxy image is then convolved with the PSF model. For later reference, the galaxy position, lensing magnification, reduced shear moments, redshift, and stamp flux are passed to a truth catalog. Finally, the galaxy image is converted to a "stamp" GSObject and drawn onto the observation at the appropriate coordinates. Based on the COSMOS 2015 $i+$ magnitude limit of 26.2 ($3''$ aperture), we inject a fiducial number of 99 galaxies per square arcminute.

Cluster galaxies are generated in much the same way as source galaxies, except that they are concentrated in the center of the observation and no lensing distortion is applied. The number of cluster galaxies (30) is set to approximately match the source density of bright cluster galaxies in the 2019 A2218 observation. They are uniformly distributed in a circle of radius 200 pixels ($28''$). A random offset of ± 50 pixels is added, i.e., about $7''$ per galaxy. Because the cluster galaxies are generally large and bright, the default GalSim-COSMOS F814W < 23.5 sample catalog is sufficient for modeling cluster galaxy sizes and brightnesses. For recording in the truth catalog, they are assigned a redshift equal to the redshift in the `NFWHalo` class.

We have an ensemble of catalogs containing star positions and brightness. These catalogs are made using the Gaia satellite observations of the galaxy clusters in SUPERBIT's planned target list. For each simulation, we select a catalog and draw the same number of stars as observed by Gaia, using their fluxes to accurately represent the stars' brightness, while the spatial density is also preserved. Star positions, however, are not specifically replicated.

Pre-seeing stars are modeled as `DeltaFunction` objects, with a flux randomly drawn from the selected cluster's Gaia catalog of real stars. The star model is convolved with the same PSF model as above, before itself being drawn into the observation. Unless otherwise specified in the configuration file, the total number of stars injected over the entire field of view matches the number of entries in the selected Gaia catalog.

Once injection of all stars and background and cluster galaxies is complete, we add dark current to the image. The final step is application of the `CCDNoise` method, which adds Poisson noise to the image based on the pixel values (including read noise). At this point, the simulated observation may be saved to file, and the process is repeated up to the total desired exposure time, in each desired filter.

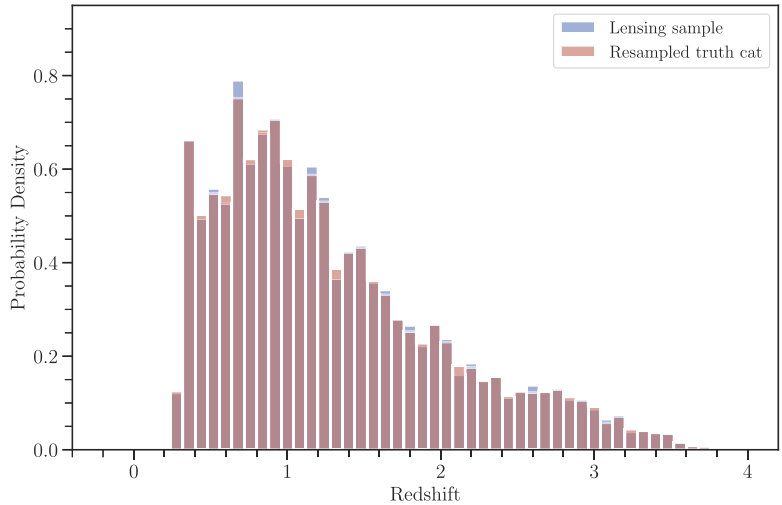

Figure 9. Redshift distribution for weak-lensing source galaxies and bootstrapped reference NFW catalog in one simulated observation of a cluster at $z = 0.3$.

To provide a reference for shear bias calculations, NFW tangential shear catalogs are generated in every (M, z) bin with a modified version of the simulations code. The redshift distributions of the reference NFW catalogs are identical to the input COSMOS catalog; however, they will differ significantly from the redshift distributions of the final mock observation catalogs. We circumvent this problem by resampling the NFW references catalog with a Monte Carlo rejection sampling algorithm until the redshift distributions match the mock observation catalogs. Figure 9 shows an example of the resulting, nearly indistinguishable redshift distributions.

6. Results

Having developed this data analysis infrastructure, we now consider its application to our simulated galaxy cluster observations. Table 2 summarizes the mean source density, imaging depth, and galaxy redshift distributions for mock SUPERBIT observations of clusters in three redshift bins: $z = 0.059$, $z = 0.3$, and $z = 0.45$. These results are computed from 30 independent realizations in each redshift bin for a total of 90 unique cluster fields.

We estimate survey properties for the total number of galaxies observed ("all galaxies" in Table 2) and lensing-analysis galaxies that pass selection cuts in Section 4.3.1 ("lensing"). To separate the effect of redshift cuts from the rest of the lensing selections in 4.3.1, we also compute survey properties for the background galaxies for each cluster ("$z_{\rm gal} > z_{\rm clust}$") without any size selections. All quantities are computed on the coadded images and obey the following color convention in plots: *u* is shown in pink, *b* in blue, *lum* in orange, and *shape* in red. All magnitudes are expressed in the AB magnitude system.

6.1. Source Density

We compute the galaxy number (source) density as a function of exposure time as follows. Upon completion of a cluster realization, a script generates a list of exposures that are a subset of the total number. Next, a pared-down version of `medsmaker` combines the exposures into a coadd and produces a source catalog, which is then matched to the galaxy and lensing analysis catalogs of the full observation. The process is repeated for 1–6 exposures and then intervals of 3 exposures. Once this process is complete for all 30 realizations in that (M, z) and bandpass bin, the script computes summary statistics such as mean and standard deviation of galaxy catalog lengths for the *n*-exposure coadds.

Figures 10 and 11 show mean number of galaxies per square arcminute as a function of on-sky integration time. Results are shown for the *u*, *b*, *lum*, and *shape* bands. Error bars are standard error of the mean across the 30 cluster realizations of each redshift bin. Integration time is expressed in number of coadded five-minute exposures (the fiducial exposure time) to reach a total of 3 hr (36×5 minutes).

Total galaxy number densities are shown in Figure 10; these samples have no selections on galaxy shape fits or redshifts beyond a SExtractor `SNR_WIN` >5 cut. The source densities for three hours of integration time are 45.5 galaxies per square arcminute in *lum*, 43.1 in *b*, 36.5 in *shape*, and 15.4 in *u*.

The growth of source density is well fit by a logarithmic function. In the planned shear measurement band *b*,

$$N_{\rm gals} = 11.01 \log_2(2.99 + N_{\rm exp}) - 15.34. \quad (32)$$

Extrapolating outward, increasing the *b* source density from 43 galaxies arcmin^{-2} to 50 galaxies arcmin^{-2} would take an additional 1.3 hr of observation.

The background galaxy number densities in Figure 11 include the lensing sample selections of Section 4.3.1. Lensing-analysis samples for clusters at $z = 0.059$ have mean source densities of 33.5 galaxies arcmin^{-2} in *lum*, 31.4 galaxies arcmin^{-2} in *b*; 26.0 in *shape*, and 9.1 in *u* (though we would not attempt weak-lensing measurements in *u*). For clusters at $z = 0.3$, the corresponding source densities are 30.2 galaxies arcmin^{-2} in *lum*, 28.3 galaxies arcmin^{-2} in *b*, 23.2 in *shape*, and 7.6 in *u*. Source densities clusters at $z = 0.45$ (the highest-redshift bin considered) have mean source densities of 27.2 galaxies arcmin^{-2} in *lum*, 25.3 galaxies arcmin^{-2} in *b*, 20.5 in *shape*, and 6.4 in *u*.

To separate the effect of redshift cuts from the rest of the lensing selections in 4.3.1, we also calculate source densities of background galaxies with no additional selections. Table 2 shows that redshift cuts alone produce more modest drops in source density than the lensing selections. The change in source density for a cluster at $z = 0.059$ is insignificant within error bars, but lensing selections reduce the source density in *b* by 27%, from 43.1 to 31.4 galaxies arcmin^{-2}. The source density behind $z = 0.3$

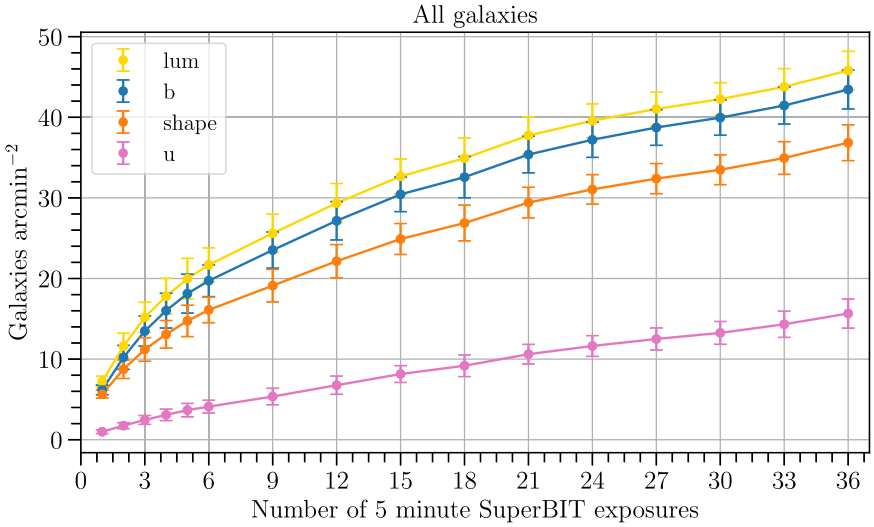

Figure 10. Galaxy number density as a function of integration time (measured in five-minute increments) in four SUPERBIT bandpasses. Points are the mean values across 30 simulated galaxy clusters with $M = 4.1 \times 10^{14}\, M_{\odot}\, h^{-1}$, $z = 0.059$; error bars are standard errors of the mean. Beyond a requirement that galaxy S/N >5, no selections on galaxy redshifts or fit parameters are made.

Table 2
Forecast Observation Depths and Redshifts

Cluster z	Galaxy Sample	Filter	Source Density ($N_{\rm gals}$ arcmin^{-2})	S/N = 10 Depth (AB mag)	Median z	Mean z
0.059	All galaxies	*u*	15.4	25.5	0.9	1.0
0.059	All galaxies	*b*	43.1	26.3	1.1	1.3
0.059	All galaxies	*lum*	45.5	26.3	1.1	1.3
0.059	All galaxies	*shape*	36.5	25.2	0.9	1.2
0.059	$z > z_{\rm clust}$	*u*	15.2	25.5	0.9	1.1
0.059	$z > z_{\rm clust}$	*b*	42.9	26.3	1.1	1.3
0.059	$z > z_{\rm clust}$	*lum*	45.2	26.3	1.1	1.3
0.059	$z > z_{\rm clust}$	*shape*	36.3	25.2	0.9	1.2
0.059	Lensing	*u*	9.1	25.4	0.9	1.0
0.059	Lensing	*b*	31.4	26.3	1.0	1.2
0.059	Lensing	*lum*	33.5	26.2	1.0	1.2
0.059	Lensing	*shape*	26.0	25.1	1.0	1.2
0.3	$z > z_{\rm clust}$	*u*	12.9	25.4	1.0	1.1
0.3	$z > z_{\rm clust}$	*b*	38.5	26.3	1.2	1.3
0.3	$z > z_{\rm clust}$	*lum*	40.6	26.3	1.1	1.3
0.3	$z > z_{\rm clust}$	*shape*	32.1	25.2	1.0	1.3
0.3	Lensing	*u*	7.6	25.4	0.9	1.1
0.3	Lensing	*b*	28.3	26.2	1.1	1.3
0.3	Lensing	*lum*	30.2	26.2	1.1	1.3
0.3	Lensing	*shape*	23.2	25.1	1.0	1.2
0.45	$z > z_{\rm clust}$	*u*	11.1	25.5	1.1	1.2
0.45	$z > z_{\rm clust}$	*b*	34.5	26.3	1.2	1.4
0.45	$z > z_{\rm clust}$	*lum*	36.4	26.3	1.2	1.4
0.45	$z > z_{\rm clust}$	*shape*	28.2	25.2	1.1	1.3
0.45	Lensing	*u*	6.4	25.4	1.0	1.2
0.45	Lensing	*b*	25.3	26.2	1.2	1.4
0.45	Lensing	*lum*	27.2	26.2	1.2	1.4
0.45	Lensing	*shape*	20.5	25.1	1.1	1.3

Notes. Results are based on three hours of integration time per band per cluster. The $z > z_{\rm clust}$ and "all galaxies" samples have an S/N >5 selection. "Lensing" galaxies pass the selection criteria listed in Section 4.3.1.

is 38.5 galaxies arcmin^{-2} (about 10% drop from 43.1), but the rest of the lensing selections leaves 28 galaxies arcmin^{-2} (two-thirds of the original source density). Similarly, the source density behind $z = 0.45$ is 34.5 galaxies arcmin^{-2} in *b* (a 20% drop), while the lensing sample has a source density of 25.3 galaxies arcmin^{-2} (40% lower than the full galaxy sample). We find that

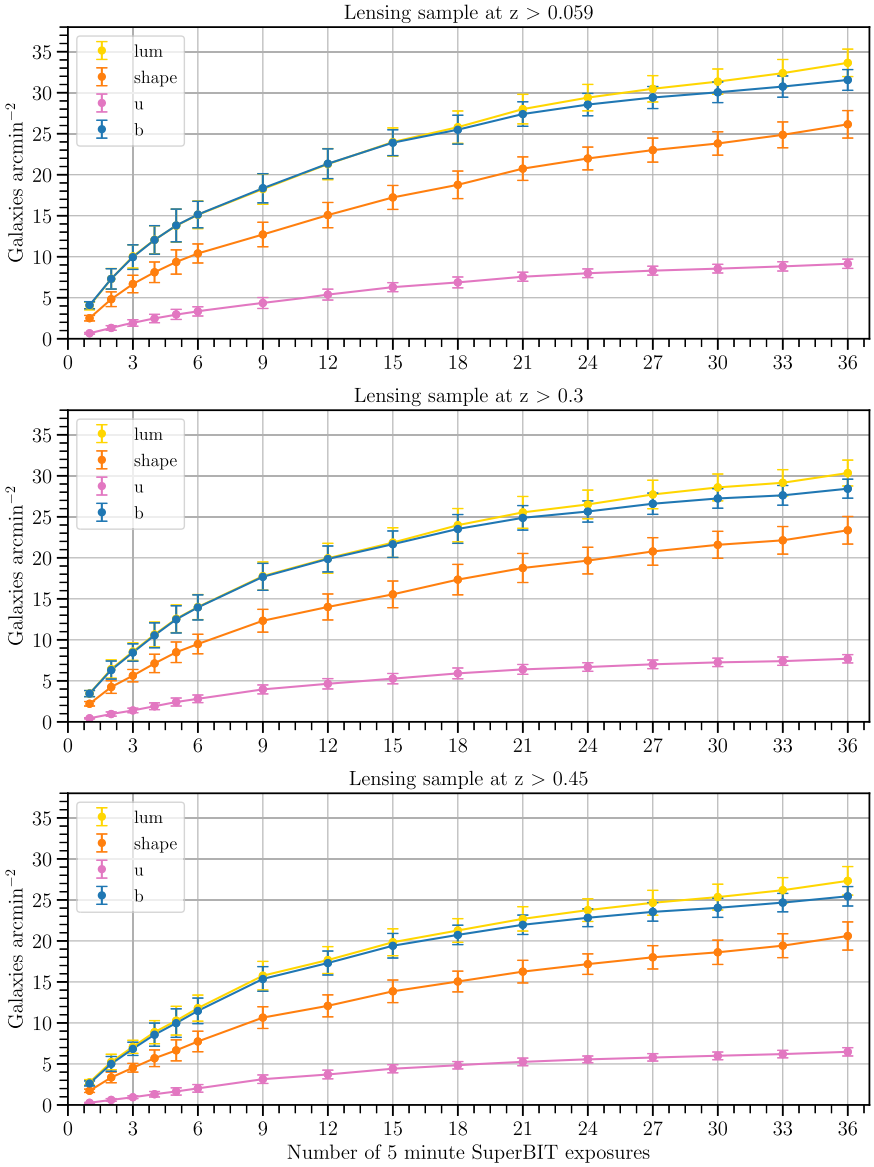

Figure 11. Background galaxy number density as a function of total integration time on clusters at $z = 0.059$ (top), $z = 0.3$ (middle), and $z = 0.45$ (bottom). Plotted values are the mean of 30 simulated observations of clusters with $M = 4.1 \times 10^{14}$ and the indicated redshift; error bars are the standard errors of the mean. The source densities reflect the selection criteria for clusters' respective "lensing" samples.

lensing-analysis selections tend to decrease the source densities more strongly than redshift cuts alone.

6.2. Depths

Depth, or the limiting magnitude for some threshold, is a commonly used figure of merit in astronomical surveys. We adopt the magnitude limit corresponding to a fixed $\sim$S/N = 10 threshold (9.8–10.2) based on $\delta F/F \sim 0.1$, where $F =$ FLUX_AUTO and $\delta F =$ FLUXERR_AUTO (Abbott et al. 2018).

Three hours of observation in *b* yields an $\sim$S/N = 10 depth of 26.3 before any lensing selections are made; lensing selections do not significantly change the depth. S/N = 10 depths in *lum* are similar to *b*, while *u* and *shape* depths are about a magnitude shallower. Values for all filters are listed in Table 2.

Figure 13 and the top panel of Figure 12 show galaxy brightness distributions, displayed as histograms of detected galaxy counts. Magnitudes are obtained from Source Extractor FLUX_AUTO values, using the IMX455 detector gain and quantum efficiency to convert to the AB system. Distributions are shown for each of *u*, *b*, *lum*, and *shape*, and the histograms are normalized such that the product of bin width and probability is equal to unity.

The top panel of Figure 12 shows the distribution of the "all galaxies" sample with only an S/N >5 selection criterion. Figure 13 presents number counts as a function of brightness for galaxies behind clusters at $z = 0.059$ (top row), $z = 0.3$ (middle row), and $z = 0.45$ (bottom row). The left panels show the $z_{\rm gal} > z_{\rm clust}$ sample for each cluster, and the right panels show the lensing-analysis galaxy samples.

6.3. Redshift Distributions

The strength of weak-lensing signal depends on the relative distances of the cluster and background galaxies. Accordingly,

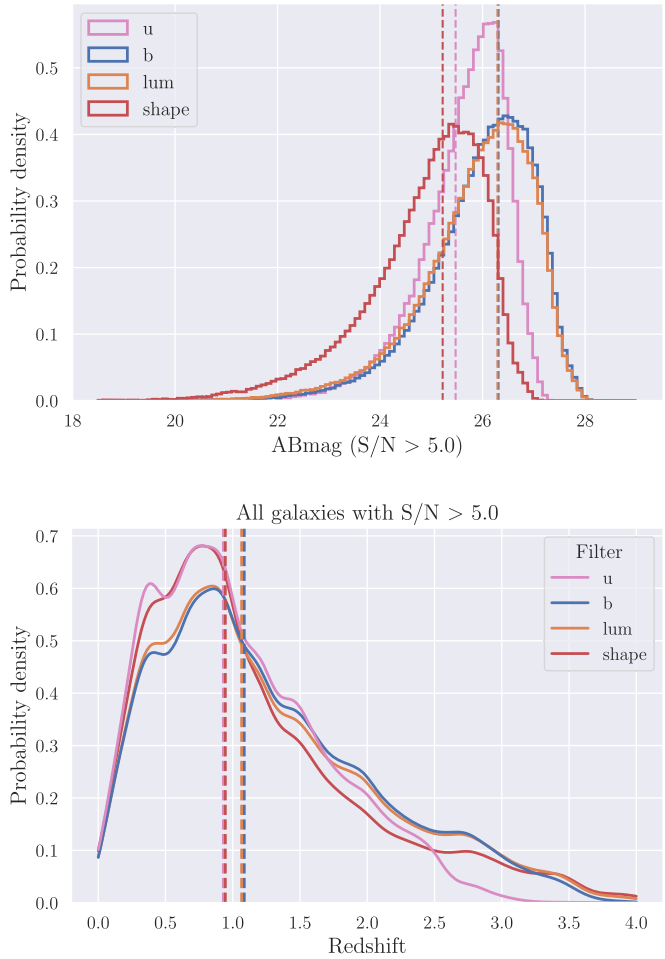

Figure 12. Galaxy brightness histograms (top) and redshift distributions (bottom) for the "all galaxies" sample with only an S/N >5 selection criterion. Dotted lines mark the S/N = 10 limiting magnitudes (top) and the median redshift (bottom), respectively, in each filter.

we calculate galaxy redshift distributions (Figure 12, bottom panel; Figure 14). Distributions are obtained with kernel density estimation and are normalized to show relative probability density within a bandpass (the product of bin width and probability density equals unity). Dotted lines mark the median redshift in a given filter.

The bottom panel of Figure 12 shows redshift distributions for all galaxies detected in *u* (pink), *b* (blue), *lum* (orange), and *shape* (red) coadds. The probability density in all bands peaks around z=0.83, with long tails past redshift z=1.5. Consistent with the depths in Section 6.2, *u* and *shape* observations have mean and median redshifts about 0.15 units lower than the deeper bandpasses.

As in Section 6.2, redshift distributions are shown both for a $z_{\rm gal} > z_{\rm clust}$ selection and for galaxies that pass all lensing analysis selections. The mean redshift of the background galaxies increases slightly with increasing cluster redshift, from a mean *b* redshift of $\bar{z} = 1.3$ for $z_{\rm clust} = 0.059$ to $\bar{z} = 1.4$ for $z_{\rm clust} = 0.45$. However, the changes are small, and lensing selections do not appear to change the mean or median background galaxy redshifts. The mean and median values of redshift in all bandpasses and galaxy samples are summarized in Table 2.

It is reasonable to ask whether the redshift distributions of the galaxy lensing samples are actually distinct from the input COSMOS 2015 catalog. We compare these in Figure 15, which compares the redshift distributions of the COSMOS 2015 simulation input catalog and the lensing-analysis catalogs of *b* observations of clusters with $(M,\ z) = (4.1 \times 10^{14}\, M_\odot\, h^{-1},\ 0.059)$ and $(4.1 \times 10^{14}\, M_\odot\, h^{-1},\ 0.3)$. The probability densities of both the COSMOS 2015 catalog and SUPERBIT lensing samples are maximized at $z \sim 0.9$. However, the COSMOS 2015 catalog has a higher probability density at $z > 1$. The mean redshift of the COSMOS 2015 catalog is $\bar{z} = 1.5$, compared with $\bar{z} = 1.2$ for the $z > 0.059$ lensing-analysis sample and $\bar{z} = 1.3$ for the $z > 0.3$ lensing-analysis sample.

The calculations above assume perfect knowledge of the redshift. In real SUPERBIT observations, we will separate background (lensed) galaxies from foreground (unlensed) galaxies with galaxy color cuts. To optimize the exposure time per bandpass for an effective foreground/background separation, we investigated the evolution of $(u - b)$, $(b - g)$, and $(g - r)$ colors with redshift for a range of galaxy types.

We sampled galaxy redshifts in the range $0 < z_{\rm gal} < 1.5$ at $\delta z = 0.02$ intervals for spectral templates from elliptical to starburst (Kinney et al. 1996). For each $\delta z = 0.02$ point, we transformed the galaxy's spectral energy distribution (SED) to the desired redshift and scaled the SED flux to achieve an integrated S/N of 10 in the *b*-band filter, representing the minimum S/N for inclusion in lensing analysis. Based on the scaled SED flux, we calculated the *b*-, *g*-, and *u*-band magnitudes along with their respective magnitude errors. By calculating uncertainties for a galaxy with S/N = 10, we obtained conservative error bars that allowed us to define realistic color-cut boundaries for our galaxy selection.

For a fiducial cluster redshift of $z = 0.5$, we determined that 3 hr of integration time in *b*, 1.5 hr in *g*, and 3 hr in *u* provided optimal separation for galaxies of most spectral types. Figure 16 illustrates the color evolution for two spectral types (elliptical and disk-dominated spiral), depicted by solid lines. The small points represent the galaxy colors calculated at each $\delta z = 0.02$ interval. To aid the reader, we highlight specific foreground ($z_{\rm gal} < 0.5$) and background ($z_{\rm gal} \geqslant 0.6$) locations as blue stars and red squares, respectively, at intervals of $\delta z = 0.1$.

Error bars in Figure 16 represent predicted 1σ color uncertainties for the aforementioned exposure times and a galaxy *b*-band S/N = 10. Cyan error bars correspond to galaxies in the foreground of a $z = 0.5$ cluster, while magenta error bars indicate galaxies behind the cluster ($z_{\rm gal} > 0.6$). The dashed black lines in Figure 16 demarcate a "clean" color–color space for a $z = 0.5$ cluster. The galaxy sample below the black lines is dominated by background galaxies at $z_{\rm gal} > 0.6$, with minimal contamination from foreground galaxies ($z_{\rm gal} \leqslant 0.5$).

6.4. Mean Shear Profiles

As part of the pipeline validation effort, we also produce weak gravitational lensing shear profiles for all cluster realizations. Two examples of single-realization cluster shear profiles were shown in Figure 5.

To examine the claim that SUPERBIT is capable of weak-lensing measurements in blue bandpasses, we compare the mean tangential shear profiles of cluster observations in *b* (SUPERBIT's intended filter for galaxy shape measurement), *lum*, and the Euclid VIS–like *shape* filter in Figure 17. The

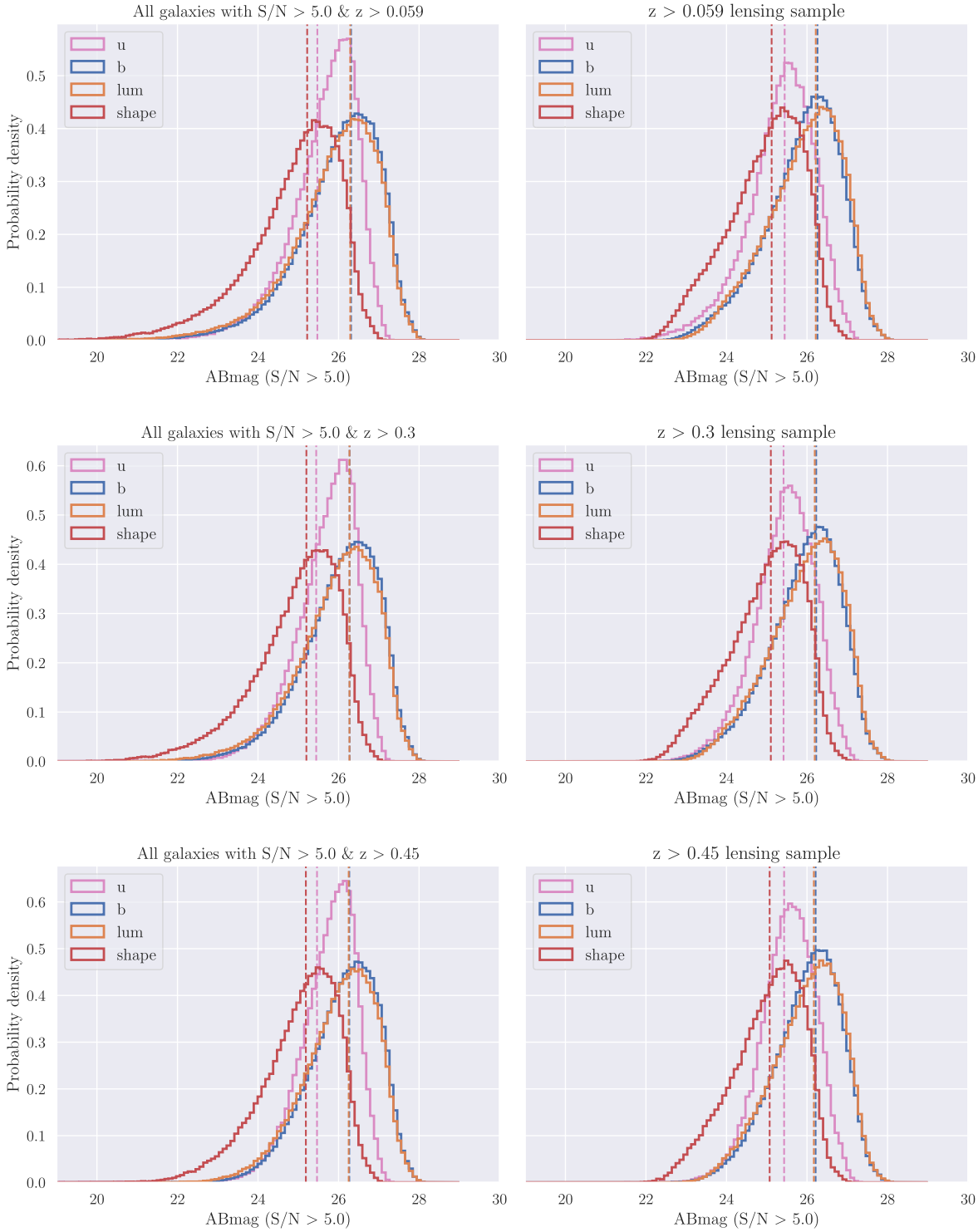

Figure 13. Galaxy brightness distributions shown as normalized histograms of galaxy counts, binned by AB magnitude. Dotted lines mark the S/N = 10 depth. Left panels: all galaxies with S/N > 5 and $z > z_{\rm clust}$. Right: galaxies that pass lensing analysis selections including $z > z_{\rm clust}$.

mean tangential shear profiles of 30 realizations of $z = 0.059$ clusters are shown in the top panel and $z = 0.45$ clusters in the bottom panel. Each point represents the mean value of the cluster tangential shear profiles, while error bars show the standard deviation of the mean in each radial bin.

We find that the tangential shear profiles are easily detected in all three SUPERBIT bandpasses. No differences in the mean values for *b*, *lum*, and *shape* are readily apparent for either cluster in Figure 17. Qualitatively, the *shape*-band error bars appear slightly larger than the *lum* and *b* error bars, which is consistent with the lower *shape* source densities in Table 2.

We emphasize that the shear profiles of Figure 17 are averages of averages, and would not be used for shear calibration or mass fitting. Instead, the figure highlights the variability and reliability of the measured tangential shear across the sample of clusters.

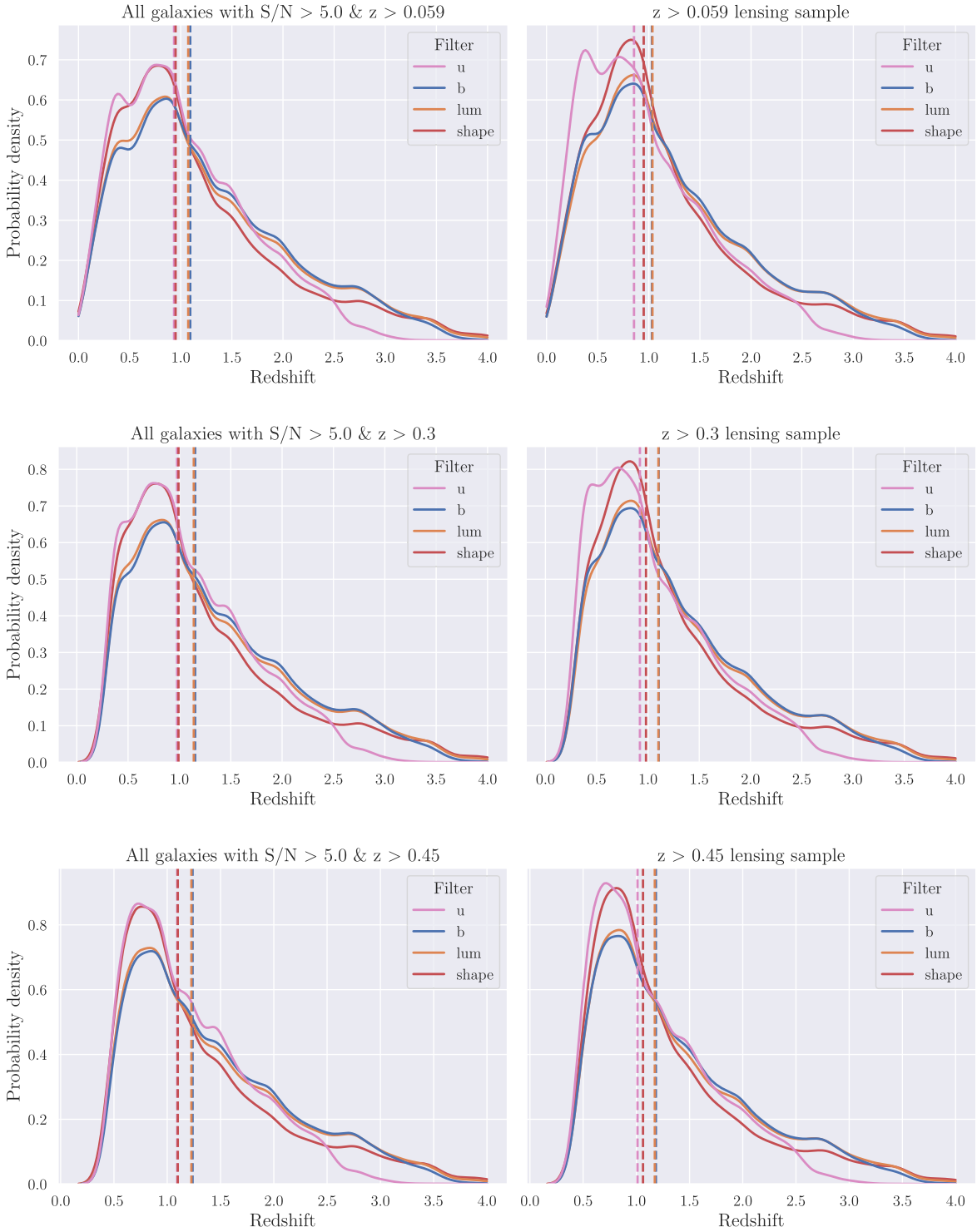

Figure 14. Normalized galaxy redshift distributions behind a cluster at $z = 0.059$ (top), $z = 0.3$ (middle), and $z = 0.45$. Left panels: all galaxies with S/N > 5 and $z_{\rm gal} > z_{\rm clust}$. Right: galaxies that pass lensing analysis selections, including $z_{\rm gal} > z_{\rm clust}$.

7. Discussion

We provide some additional commentary on our analyses and results here.

7.1. Simulation Inputs and Effect on Forecasts

The simulations that are presented in Section 5 and form the basis for Section 6 have many realistic features: an NFW cluster weak-lensing profile, star flux and densities from Gaia coverage of SUPERBIT targets, measured stratospheric sky brightnesses from Gill et al. (2020), and real galaxy redshifts and luminosities from COSMOS catalogs transformed to SUPERBIT bandpasses.

Although the simulated observations incorporate considerable complexity, there are a few limitations. First, they use a Gaussian approximation of the SUPERBIT PSF. In reality, the space-like SUPERBIT PSF features Airy rings and diffraction spikes (see Figure 8).

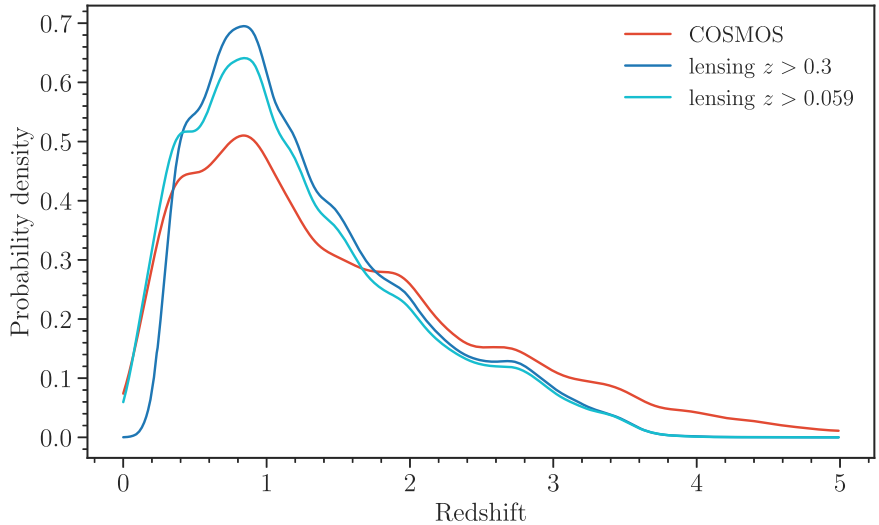

Figure 15. Normalized redshift distributions of the COSMOS 2015 simulation input catalog (red line), the $z > 0.3$ lensing sample (blue line), and the $z > 0.059$ lensing sample (cyan line). The two lensing samples have lower mean redshifts ($\bar{z} = 1.2$ and $\bar{z} = 1.3$) than the COSMOS 2015 catalog ($\bar{z} = 1.5$).

We also do not model uncertainties of galaxy redshifts. This was a deliberate choice, as the systematic errors that redshift uncertainties introduce to weak-lensing analysis are orthogonal to the pipeline validation aspect of this work and the shear calibration in forthcoming efforts. Even if we did attempt to incorporate redshift uncertainties, SUPERBIT's strategy for determining redshifts may evolve as the campaign progresses, rendering such forecasting estimates moot.

In addition, the validity of our forecast is limited by the simulation input catalog. There are few deep, high-resolution observations in the blue and near-UV. A workaround is presented in Section 5, but it assumes that galaxy morphology parameters in the ($\bar{z} = 0.9$) GalSim-COSMOS catalog can be extrapolated to the ($\bar{z} = 1.5$) galaxies in the COSMOS 2015 photometric catalog. A more theory-driven approach could involve hydrodynamical simulations. However, the morphology of intermediate- to high-z galaxies is itself a very active area of research. On balance, the high accuracy of the galaxy fluxes and realistic redshift distributions in our input COSMOS 2015 catalog outweigh any uncertainty in galaxy shapes. A full treatment of the galaxy catalog will be presented in a forthcoming paper by A. Gill et al. (2023, in preparation).

Finally, the input Gaia star catalogs are incomplete, as illustrated by the gap in the stellar locus of Figure 3. We do not believe that the dearth of faint stars affects our conclusions, as very faint stars near the "zone of confusion" would be excluded from PSF fits anyway. Future simulations will incorporate theoretical `TRILEGAL` star distributions.

7.2. Estimated Observation Depths, Source Densities, and Redshifts

A major goal of this analysis was to quantify the effect of weak-lensing selections on galaxy number density. Table 2 shows that weak-lensing analysis selections cause a more significant decrease in source density (30–40%) than redshift selections alone. The addition of lensing selections does not appear to significantly change the mean and median redshifts of the samples any more than a redshift cut alone.

A surprising result of Section 6.2 is the high depth and source density in u. The deep NUV CLAUDS survey (Sawicki et al. 2019) provides one of the few points of comparison for our own u findings. At a similar depth to ours (25.5 mag), they report a 5 σ source density of $\log_{10} N = 4.58/\mathrm{deg}^2/0.5$ mag, or 10.7 galaxies per arcmin^2. This is 50% lower than our maximum reported value of 15.4 galaxies per arcmin^2. The change in source density with redshift is also noteworthy: we report a 27% decline in u source density from $z = 0.059$ to $z = 0.45$, while over approximately the same redshift range, the CLAUDS survey reports a decline of $\sim$12% (Moutard et al. 2020).

One possibility for the divergence is SUPERBIT's smaller PSF: the CLAUDS survey experienced an average PSF FWHM of $0\farcs92$, but the SUPERBIT u PSF FWHM is about $0\farcs278$. A smaller PSF translates to a higher source density, as objects that might otherwise be blended or smeared out over noisy pixels become resolvable. A more likely explanation is that our UV luminosities do not account for foreground extinction by Milky Way dust, which is significant in the UV and will certainly depress u source counts in real SUPERBIT observations. If the GalSim-COSMOS shape parameters cannot be extrapolated to bluer bands and fainter galaxies, it is also possible that the galaxy morphologies in our catalog are inaccurate for u observations. The ultimate calibration for our simulations will be provided by the analysis of real SUPERBIT observations in u and b.

7.3. Impact on Observation Strategy

Figures 10, 11, and 17 show that the source density achieved in three hours of observation in b or *lum* is completely adequate for shear profile measurements. Observations in b or *lum* longer than three hours would confer limited advantages at a high cost in integration time (see Equation (32)). In fact, future analysis may reveal that shorter integration times would suffice, saving time during flight and allowing a greater number of targets to be observed.

The final observation strategy will depend on the results of ongoing optics-on (jitter+optics) simulations in all SUPERBIT bandpasses as well as a redshift analysis that is currently underway. However, Figures 10 and 11 strongly support the conclusions of Shaaban et al. (2022) that *lum* and b observations are both faster and deeper than the Euclid VIS–like *shape* when observing from the stratosphere. Our

estimated source densities in these bandpasses also agree with Shaaban et al. (2022) within uncertainties.

Finally, Figure 17 shows the feasibility of measuring galaxy cluster weak-lensing signal in *b* and *lum*, and that a broadband red filter like *shape* offers no noticeable advantage over the bluer filters. This result supports our planned observing strategy of deep *b* observations for galaxy shape measurements.

8. Conclusions and Outlook for 2023

In this work, we have presented a first iteration of the galaxy shape measurement pipeline for SUPERBIT's weak-lensing analysis. The software and algorithms we employ—GalSim, SExtractor, PIFF, Metacalibration, and NGMix—have been rigorously tested and were intended for widespread adoption by the community. Processing simulated observations has allowed us to test their implementation in this pipeline. Several years after the release of these tools, there is now a growing number of pipelines similar to ours, e.g., `ShapePipe` (Guinot et al. 2022) and `run_steps` (Fu et al. 2022), with more likely to come.

Beyond pipeline validation, our simulated observations and catalogs provide estimates for the expected number density, depth, and redshift distribution of galaxies in deep, stratospheric imaging. We predict that SUPERBIT can attain a depth of 26.3 mag in the *b* filter and 25.5 mag in the *u* filter—competitive with even the deepest ground-based surveys. We also find a total source density greater than 40 galaxies arcmin^{-2} in three hours of integration time in both the *b* and *lum* bands. The source density remains high even after the application of lensing catalog selections: 25–30 galaxies arcmin^{-2} in the *b* bandpass. We expect that instrumental effects (including the optical PSF) will depress the source density. However, the relative performance of *b*, *lum*, and *shape* is unlikely to be affected and supports SUPERBIT's observation strategy.

This work also offers a look at the weak-lensing tangential shear profiles expected for SUPERBIT cluster observations, further confirming SUPERBIT's capacity for weak gravitational lensing measurements in the blue. As with the other forecast survey properties, these weak-lensing profiles are based on Gaussian approximations to the SUPERBIT PSF and do not include redshift uncertainties. The vagaries of real observations will add some scatter to the final weak-lensing measurements. Even with these caveats, the relative performance of different filters also supports SUPERBIT's observation strategy.

The SUPERBIT pipeline and simulations remain in active development. Forthcoming improvements include source detection on a multi-bandpass composite image; galaxy shape measurement with the full SUPERBIT PSF; inclusion of faint stars in simulated observations using stellar population synthesis models; and the addition of redshift uncertainty to the input galaxy catalog. While the pipeline includes tools for shear calibration, we do not validate them here. Instead, a complete shear calibration analysis will be presented in a forthcoming paper by S. Everett et al. (2023, in preparation).

Though our pipeline has been developed specifically for SUPERBIT weak-lensing measurements, it is generic and can be refactored for weak-lensing observations with other instruments. An obvious example is SUPERBIT's successor mission, GIGABIT: a planned 1.3 m gigapixel class balloon-borne observatory (Li et al. 2022). Future pipeline developments will facilitate forecasting and survey planning for both SUPERBIT and GIGABIT.

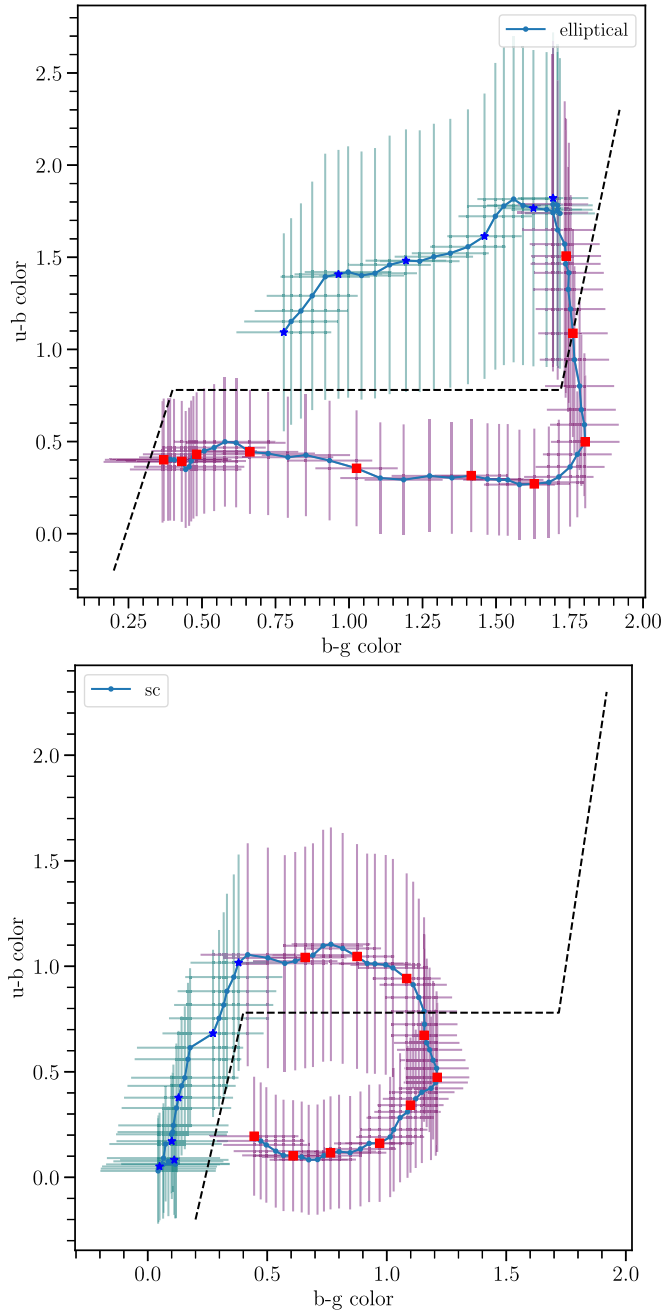

Figure 16. Redshift evolution of elliptical and spiral galaxies in $(b-g)$, $(u-b)$ color–color space. Blue solid lines show the observed galaxy color evolution between $0 < z_{\rm gal} < 1.5$; small points mark $\delta z = 0.02$ intervals. Large markers are included as a visual aid: blue stars and cyan error bars mark $z_{\rm gal} < 0.5$, while red squares and magenta error bars mark $z_{\rm gal} \geqslant 0.6$. Error bars are 1σ color uncertainties for galaxies with *b*-band S/N = 10. Color–color tracks assume exposure times of 3 hr, 1.5 hr, and 3 hr in the *b*, *g*, and *u* bands, respectively. A "clean" color–color space, with low foreground galaxy contamination (no blue stars), suitable for a cluster at $z = 0.5$, lies below the dashed black line.

Since the initial submission of this paper, we are excited to announce the successful launch and completion of the SUPERBIT mission, which spent 40 days afloat. The data calibration process is currently underway, and we will subsequently conduct an analysis along the lines described in this paper.

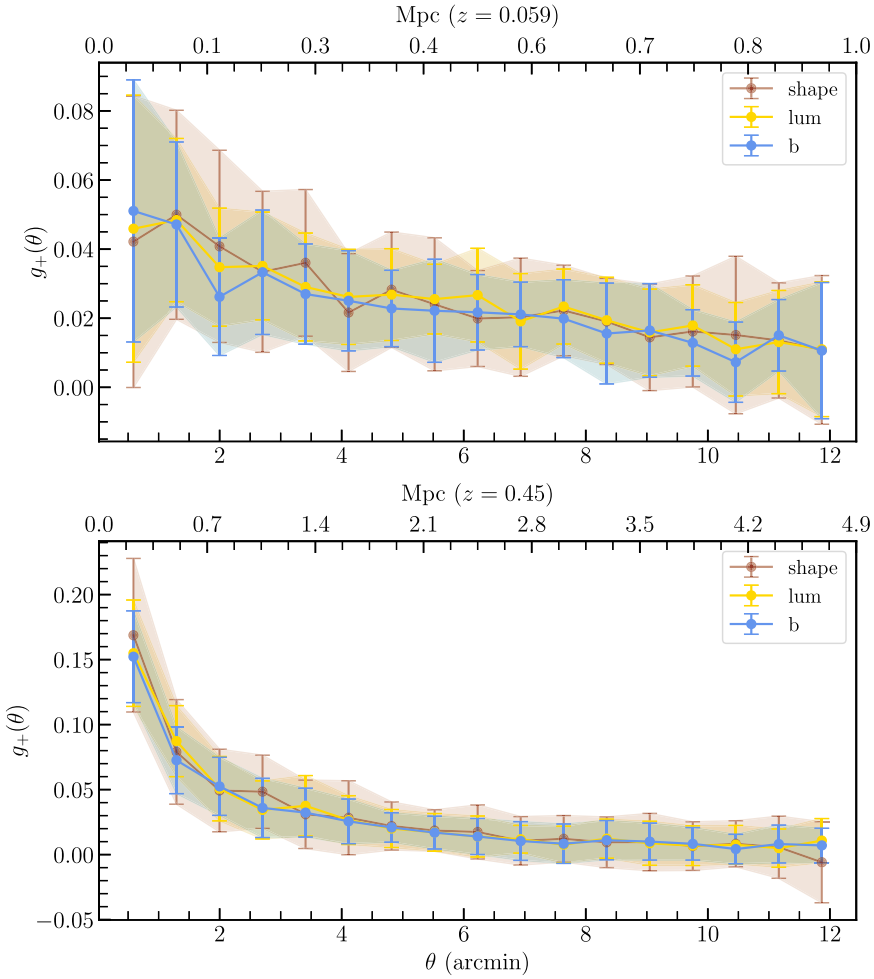

Figure 17. Average tangential shear profile for 30 simulated galaxy clusters with identical masses and redshifts. Each point represents the mean value of the tangential shear profiles across the cluster sample; error bars represent the standard deviation of the mean in each radial bin and illustrate the fluctuation of individual cluster measurements. Profiles are shown for observations in *b*, *lum*, and *shape* (3 hr of exposure time each). Top: clusters with $M = 4.1 \times 10^{14}\, M_\odot\, h^{-1}$, $z = 0.059$. Bottom: clusters with $M = 4.1 \times 10^{14}\, M_\odot\, h^{-1}$, $z = 0.45$. The bottom *x*-axes show distance from the cluster center in arcminutes, while top axes are scaled to show the projected distance from the cluster center in megaparsecs. The tangential shear profiles are easily detected in SUPERBIT *b* and *lum* observations, shown in blue and gold, with no obvious advantage for *shape*-band measurements (shown in red).

SUPERBIT offers a new data product: wide-field, diffraction-limited $\lambda < 600$ nm imaging deep enough to enable galaxy cluster weak-lensing analysis. Our forecast galaxy number density and redshift distribution confirm SUPERBIT's capability for weak-lensing mass measurement in blue wavelengths. This demonstrates that, even in the era of multibillion-dollar space telescopes like JWST, Roman, and Euclid, nimble and low-cost missions like SUPERBIT offer immense scientific potential and a complementary paradigm for space-based scientific observations.

Acknowledgments

Support for the development of SUPERBIT is provided by NASA through APRA grant NNX16AF65G. Launch and operational support for the sequence of test flights from Palestine, Texas are provided by the Columbia Scientific Balloon Facility (CSBF) under contract from NASA's Balloon Program Office (BPO). Launch and operational support for test flights from Timmins, Ontario are provided by the Centre National d'Études Spatiales (CNES) and the Canadian Space Agency (CSA).

J.R., E.H., and S.E. are supported by JPL, which is run under a contract by Caltech for NASA. Canadian coauthors acknowledge support from the Canadian Institute for Advanced Research (CIFAR) as well as the Natural Science and Engineering Research Council (NSERC). The Dunlap Institute is funded through an endowment established by the David Dunlap family and the University of Toronto. UK coauthors acknowledge funding from the Durham University Astronomy Projects Award, STFC [grant ST/P000541/1], and the Royal Society [grants UF150687 and RGF/EA/180026].

The simulation input catalog is based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO program ID 179.A-2005 and on data products produced by TERAPIX and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium.

Software: Astropy (Astropy Collaboration et al. 2013, 2018), GalSim (Rowe et al. 2015), Source Extractor (Bertin & Arnouts 1996), NGMIX (Sheldon 2015), Seaborn (Waskom 2021), Matplotlib (Hunter 2007).

ORCID iDs

Jacqueline E. McCleary https://orcid.org/0000-0002-9883-7460
Spencer W. Everett https://orcid.org/0000-0002-3745-2882

Mohamed M. Shaaban https://orcid.org/0000-0002-7600-3190
Ajay S. Gill https://orcid.org/0000-0002-3937-4662
Georgios N. Vassilakis https://orcid.org/0009-0006-2684-2961
Eric M. Huff https://orcid.org/0000-0002-9378-3424
Richard J. Massey https://orcid.org/0000-0002-6085-3780
Steven J. Benton https://orcid.org/0000-0002-4214-9298
Mathilde Jauzac https://orcid.org/0000-0003-1974-8732
William C. Jones https://orcid.org/0000-0002-3636-1241
David Lagattuta https://orcid.org/0000-0002-7633-2883
Jason S.-Y. Leung https://orcid.org/0000-0001-7116-3710
Lun Li https://orcid.org/0000-0002-8896-911X
Johanna M. Nagy https://orcid.org/0000-0002-2036-7008
Emaad Paracha https://orcid.org/0000-0001-5101-7302
Susan F. Redmond https://orcid.org/0000-0002-9618-4371
Jason D. Rhodes https://orcid.org/0000-0002-4485-8549
Jürgen Schmoll https://orcid.org/0000-0001-5612-7535
Ellen Sirks https://orcid.org/0000-0002-7542-0355

References

Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ApJS, 239, 18
Astropy Collaboration, Price-Whelan, A. M., & Sipőcz, B. M. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Becker, M. R., & Kravtsov, A. V. 2011, ApJ, 740, 25
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in ASP Conf. Ser. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley (San Francisco, CA: ASP), 228
Bertin, E. 2011, in ASP Conf. Ser. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans et al. (San Francisco, CA: ASP), 435
Brown, A., Vallenari, A., Prusti, T., et al. 2018, A&A, 616, A1
Clowe, D., Gonzalez, A., & Markevitch, M. 2004, ApJ, 604, 596
Fenech Conti, I., Herbonnet, R., Hoekstra, H., et al. 2017, MNRAS, 467, 1627
Fu, S., Dell'Antonio, I., Chary, R.-R., et al. 2022, ApJ, 933, 84
Gaia Collaboration, Brown,, A. G. A,, Vallenari, A. , et al. 2018, A&A, 616, A1
Gatti, M., Sheldon, E., Amon, A., et al. 2021, MNRAS, 504, 4312
Gill, A., Benton, S. J., Brown, A. M., et al. 2020, AJ, 160, 266
Gill, A. S., Shaaban, M. M., Tohuvavohu, A., et al. 2022, Proc. SPIE, 12191, 1219114
Guinot, A., Kilbinger, M., Farrens, S., et al. 2022, A&A, 666, A162
Haiman, Z., Mohr, J. J., & Holder, G. P. 2001, ApJ, 553, 545
Hirata, C. M. 2005, PhD thesis, Princeton University
Holder, G., Haiman, Z., & Mohr, J. J. 2001, ApJL, 560, L111
Huff, E., & Mandelbaum, R. 2017, arXiv:1702.02600
Hunter, J. D. 2007, CSE, 9, 90
Jarvis, M., Meyers, J., Leget, P.-F., & Davis, C. 2021, Piff: PSFs In the Full FOV, Astrophysics Source Code Library, ascl:2102.024
Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, MNRAS, 460, 2245
Kinney, A. L., Calzetti, D., Bohlin, R. C., et al. 1996, ApJ, 467, 38
Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
Li, L. 2015, Master's Thesis, University of Toronto
Li, L., Kasdin, N. J., Jones, W. C., & Benton, S. J. 2022, Proc. SPIE, 12182, 1218204
Mandelbaum, R., Lanusse, F., Leauthaud, A., et al. 2018, MNRAS, 481, 3170
McCleary, J. 2020, ApJ, 893, 8
McCleary, J., Dell'Antonio, I., & Huwe, P. 2015, ApJ, 805, 40
Moutard, T., Sawicki, M., Arnouts, S., et al. 2020, MNRAS, 494, 1894
Redmond, S., Benton, S., Brown, A. M., et al. 2018, Proc. SPIE, 10700, 1687
Romualdez, L. J., Benton, S. J., Brown, A. M., et al. 2018, Proc. SPIE, 10702, 107020R
Romualdez, L. J., Benton, S. J., Brown, A. M., et al. 2020, RScI, 91, 034501
Rowe, B. 2010, MNRAS, 404, 350
Rowe, B., Jarvis, M., Mandelbaum, R., et al. 2015, A&C, 10, 121
Rowe, B. T. P., Jarvis, M., Mandelbaum, R., et al. 2015, A&C, 10, 121
Saito, S., de la Torre, S., Ilbert, O., et al. 2020, MNRAS, 494, 199
Sawicki, M., Arnouts, S., Huang, J., et al. 2019, MNRAS, 489, 5202
Shaaban, M. M., Gill, A. S., McCleary, J., et al. 2022, AJ, 164, 245
Sheldon, E. 2015, NGMIX: Gaussian mixture models for 2D images, Astrophysics Source Code Library, ascl:1508.008
Sheldon, E. S., & Huff, E. M. 2017, ApJ, 841, 24
Tam, S.-I. 2020, PhD thesis, Durham University, UK
Umetsu, K. 2020, A&ARv, 28, 7
Vogelsberger, M., & Zavala, J. 2016, MNRAS, 460, 1399
Waskom, M. L. 2021, JOSS, 6, 3021
Zuntz, J., Sheldon, E., Samuroff, S., et al. 2018, MNRAS, 481, 1149