

A synthetic WFIRST survey: Simulation suite and the impact of wavefront errors on weak lensing

and others

(WFIRST HLS Cosmology SIT)

(Dated: August 22, 2019)

I. INTRODUCTION

no dust extinction on galaxies

- DE/cosmology context
- motivation for WFIRST in context of other stage iii and stage iv surveys
- cosmology sit
- goals of paper

The nature of dark energy, which drives cosmic acceleration in the Universe, remains one of the most fundamental mysteries in physics twenty years after its discovery [? ? ? ? ?]. A number of new experiments have been undertaken to probe dark energy using a variety of physical phenomena, including baryon acoustic oscillations, numbers and masses of galaxy clusters, galaxy clustering, redshift space distortions, Type Ia supernovae, and weak gravitational lensing. Current generation experiments are limited to some subset of these probes, but have already begun to expose interesting questions about the soundness of our standard cosmological model, Lambda-Cold Dark Matter (Λ CDM) that will require more and better data to resolve. The Wide-Field InfraRed Space Telescope (WFIRST) has been designed to take advantage of all of these probes to study dark energy with unprecedented systematic control [? ?].

In the past few years, the current generation of ground-based weak lensing experiments like the Dark Energy Survey (DES), Hyper-Suprime Cam (HSC), and Kilo-Degree Survey (KiDS) have reached levels of precision that rival the previously best possible cosmological constraints including dark energy []. These surveys have spurred the development of revolutionary algorithms and methods for galaxy shape measurement and weak lensing analysis [], highlighting the immense power of weak lensing to unravel the fundamental mysteries we face in cosmology today.

By the planned launch of WFIRST in 2025, we will have final results from the ongoing generation of weak lensing experiments (DES, HSC, and KiDS) and preliminary results from the Dark Energy Spectroscopic Instrument (DESI), the Large Synoptic Survey Telescope (LSST), and the Euclid mission. Faced with the unknown discovery potential of these experiments in the early 2020s, it is vital to maintain the agility and flexibility of the WFIRST mission to respond with the best possible science, particularly in what is likely to be a systematics dominated weak lensing field. The process of quantifying empirically the robustness of the design requirements of the WFIRST mission for weak lensing now in Phase B of the mission development is a critical task that this paper will attempt to address. Precise control of these systematics at the

statistical precision offered by current WFIRST mission forecasts will enable WFIRST to play the likely role of arbiter in the study of new discoveries made in the early years of LSST and Euclid.

We present in this paper a set of synthetic WFIRST imaging surveys covering approx. 6 sq. deg. in one filter: a fiducial set of images and 11 variations in ways the PSF could be misestimated. It incorporates realistic distributions of photometric properties for galaxies and stars; complex analytic galaxy models; a simulated fiducial five year, 2000 sq. deg. observation strategy for the survey; and realistic detector effects, PSF models, and WCS solutions that match current (Cycle 7) design specifications. We use a blending-free version of this simulation to test the impact on weak lensing science of a variety of PSF or wave-front errors, including static, low-, and high-frequency biases.

(insert standard toc para)

II. WFIRST WEAK LENSING, THE REQUIREMENTS PROCESS, AND THE NEED FOR SIMULATIONS

Contributing: Hirata

Chris: need to write this; will make extensive reference to the appendices for technical details that aren't really necessary to understand the simulations

We now proceed to describe requirements and the role of this suite of image simulations in verifying that the requirements flowdown is correct. We begin with a description of weak lensing on WFIRST that emphasizes the issues most closely tied to the image simulations (§II A), and a high-level review of the requirements process in a cosmology project (§II B). There we describe where in this process we need the mapping between the wavefront error and galaxy ellipticities, $\partial e_i / \partial \psi_j$ (where ψ_j denotes a Zernike mode of the wavefront error). This mapping was obtained using a simplified analytic model, calibrated by toy simulations, in the Phase A requirements flowdown; this approach is described at a high level in §II C, with technical details placed in the appendices. In the rest of this paper, we will use much more advanced image simulations, based on the GALSIM package, to estimate $\partial e_i / \partial \psi_j$.

A. WFIRST weak lensing

B. The requirements process

Every precision cosmology project has a requirements process to control both its statistical and systematic errors and

ensure that the overall mission can achieve its science objectives. In the case of WFIRST, requirements on the Project (e.g., flight hardware and software or ground system support) were baselined early in the mission (the Science Requirements Document was placed under configuration control in 2018), but requirements on science analyses are more flexible and will be fixed at a later date. The statistical error requirements are usually formulated in terms of survey area, depth and image quality in each filter, cadence (for time-domain programs), etc.; their relation to the science reach of the mission is handled by forecasting tools [ref. forecasting papers](#). Systematic error control is much more difficult, and the approach may differ depending on whether a source of systematic error is *observational* or *astrophysical*. Usually, observational systematics (e.g., PSF calibration for weak lensing) can be budgeted within the systems requirements framework of a large project, whereas astrophysical systematics (e.g., baryonic feedback) are addressed through a combination of nuisance parameters, additional observations, and theory/simulation. These astrophysical systematics are important science team responsibilities but are not part of Project requirements and engineering reviews.

We focus now on the approach to observational systematic errors; our focus is on the WFIRST process, but note that something similar has been done for other large weak lensing programs such as LSST and Euclid [cites](#).

First, one identifies a *data vector* that will contain the cosmological information. For setting WFIRST weak lensing requirements, the data vector is the concatenated list of shear power spectra and cross-power spectra \mathbf{C} across tomographic bins. (Other choices, such as including higher-order statistics, using all 3×2 -point information, or working in correlation function space are possible, but given the tools available at the time of Project start these would have required additional tool development that did not fit in the schedule.)

Second, one identifies an *error metric* that summarizes the impact of a systematic error on the data vector. We have chosen the error metric $Z^2 = \Delta\mathbf{C} \cdot \Sigma^{-1} \Delta\mathbf{C}$, where $\Delta\mathbf{C}$ is the bias on the data vector and Σ is the statistics-only covariance matrix. The metric Z is essentially a metric for the ratio of the systematic to the statistical error, and this depends on the solid angle Ω covered by the survey ($Z \propto \sqrt{\Omega}$). One also sets a limit on the maximum allowed error Z ; in our case, we set $Z = 0.5$ at 2500 deg^2 (or $Z = 1$ at $10,000 \text{ deg}^2$), which means that the observational systematic errors are required to be below 50% of the statistical errors in a 2500 deg^2 survey and below 100% of the statistical errors if the survey were to be extended to $10,000 \text{ deg}^2$. **there are other choices; do we want to talk about the advantages of this choice here or put it in an appendix?**

Third, we note that each category of observational systematic error contributes to Z . In cases where the errors are presumed independent, the Z^2 values can be summed (i.e., Z obeys root-sum-square or RSS addition), and the “top-level” budget for Z can be broken down into contributions from different sources. If a source of observational systematic is parameterized by a parameter p (e.g., overall shear calibration), then a requirement on knowledge of p can be obtained by

computing the sensitivity $d\mathbf{C}/dp$ and setting

$$\sqrt{\frac{d\mathbf{C}}{dp} \Delta p \cdot \Sigma^{-1} \frac{d\mathbf{C}}{dp} \Delta p} \quad (1)$$

equal to the allocation for Z from that contribution. An important aspect of this budgeting is that, like a requirements flow-down, it is *hierarchical* – a top-level requirement on observational systematics may contain an allocation for shear calibration (one of several contributions), which itself may contain a branch for PSF size (one of several contributions), which itself may contain a branch for detector non-linearity, etc. In the life cycle of a cosmology project, more detail will be filled in first on the branches that have hardware impacts, and then branches related to algorithms or simulations later on.

Chris: Lots of references to the appendix; need to clean that up so it contains only the needed information

Chris: need to write something describing how we get from shear power spectrum down to caring about galaxy shapes and then Zernikes; also describe in very simple terms what the S-factor is

C. Mapping from waveform error to galaxy ellipticities – analytic approach

D. Limitations of the analytical approach

III. SIMULATION SUITE

- We have developed a realistic synthetic survey
- into to a synthetic survey and necessary components
- galsim framework and cycle 7 ([6?](#)) baseline
- Detailed description of simulation suite components

To empirically test weak lensing requirements, methods, and algorithms in WFIRST, we have designed a sufficiently complex synthetic survey that, while not entirely realistic in all object properties, contains sufficiently complex and representative objects as to enable informative tests and preliminary algorithm development. This synthetic survey utilizes several external simulation and data sources, and generates WFIRST-like imaging using the GALSIM framework and its WFIRST module. The simulation framework is generally capable of producing a full WFIRST HLS imaging survey in all filters matching Cycle 7 specifications. The code is publicly available at https://github.com/matrixel/wfirst_imsim, where configuration files are provided for the fiducial simulation used in this work. Figure 1 shows an example pointing produced in the fiducial simulation, and Fig. 2 shows a larger view of one of the SCAs. Performance details are provided in Appendix [[?](#)].

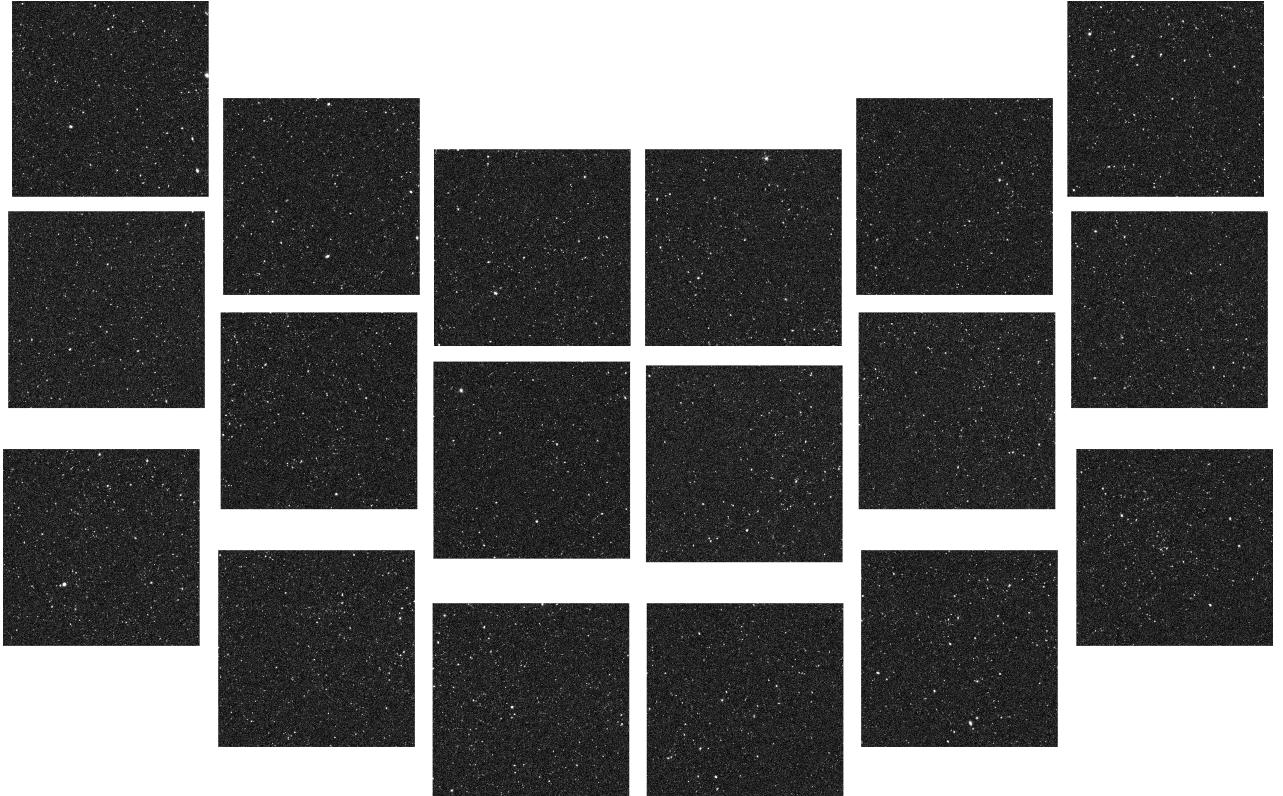


FIG. 1. A simulated field-of-view (FOV) from the fiducial simulation matching Cycle 7 specifications. The WFIRST FOV is 0.282 deg^2 , composed of 18 near-IR Sensor Chip Arrays (H4RGs). This FOV is 200 times the area of the Hubble Space Telescope Wide Field Camera 3 FOV and 90 times that of its Advanced Camera for Surveys. (check chip orientation)

A. Simulation stages

The simulation is broken into several stages:

Truth catalog generation – A truth catalog is generated from the input galaxy distribution, photometric galaxy catalog, and Milky Way simulation. The following true object properties are assigned to each galaxy in the input galaxy distribution (from which the unique galaxy id is generated): 1) The position in RA and Dec from the galaxy distribution; 2) Photometric properties drawn from a random object in the photometric galaxy catalog (consistent $YJHF$ magnitudes, size, and redshift); 3) Intrinsic ellipticity components drawn from a Gaussian distribution of width 0.27 (truncated at ± 0.7); 4) A random rotation angle; 5) The ratio of fluxes in each of the three galaxy components: a) de Vaucouleurs bulge, b) exponential disk, and c) random walk star-forming knots (a maximum of 25% of flux can exist in the knots); 6) The gravitational shear applied to the object, drawn from a discrete list of $(e_1, e_2) \in \{\pm 0.1, \pm 0.1\}$, though including a coherent shear field instead is a simple modification. Further details on the provenance of the galaxy catalogs and Milky Way simulation can be found in Secs. III C and III D, respectively. These truth properties are saved in a light-weight FITS

format that is accessed by the following stages.

Image generation – In this stage, an empty SCA image is initialized (4088×4088 pixels), and a model is built for each galaxy and star is turn, then drawn into the image. The galaxy models are built chromatically from the truth parameters for the object, with each component being assigned a different representative SED of types: S0 (bulge), SBa (disk), and Im (knots), respectively. The assigned SED is the same for all objects, since after redshifting the spectrum and applying the appropriate flux in each component and size, the model is converted to be achromatic in each passband to speed up the drawing.¹ The intrinsic ellipticity, random rotation, and gravitational shear is then applied. For stars, we apply the SED of Alpha Lyra to a chromatic delta function model. Stars are also converted to be achromatic before drawing. Both stars and galaxies are then convolved with the appropriate PSF for the SCA (constant across the SCA in the fiducial simulation). An example of the PSF model for an object is shown in Fig. 10. We save images of the true PSF model both at native pixel

¹ This simplification can be removed, however, at the cost of multiple factors of increased runtime.



FIG. 2. A simulated WFIRST Sensor Chip Array (SCA). Each SCA (HgCdTe H4RG) has a useable pixel grid of 4088×4088 , with a pixel scale of $0.11''$. One interesting feature can be seen in the relatively bright galaxy in the lower middle part of the SCA, for which six faint diffraction spikes due to the WFIRST secondary mirror support struts are visible.

scale and oversampled by a factor of 8, in stamps of native pixel size 8×8 at the position of each galaxy.

The models are drawn in dynamically-sized squares stamps, the sizes of which are chosen automatically by GAL-SIM to include at least 99.5% of the flux. These stamps are then added to the SCA image and saved separately (if drawing a galaxy) to provide an isolated image of each simulated galaxy to allow for tests of the impact of blending. Objects that would have a postage stamp that overlaps the SCA image are drawn, such that light from objects in chip gaps are appropriately drawn onto the SCA, but we only save postage stamps for objects that have a centroid that falls on the SCA. We do not save isolated postage stamps of objects that have a stamp size of greater than 288×288 pixels, but they are drawn into the images. Finally, each isolated postage stamp is processed through the steps described in Sec. III B 3 to simulate the WFIRST observatory and detectors and written to disk. When all objects are added to the full SCA image, it is also processed through these steps and written to a FITS image file. The simulation is parallelized at the SCA level, such that each SCA in each pointing can be run independently. (send sentence to appendix) The truth file is lightweight and completion time semi-random, such that even remote disk I/O has not been a limiting factor at the level of thousands of parallel jobs.

MEDS creation – We then compile the output across pointings of the isolated object stamps into MEDS (Multi-Epoch

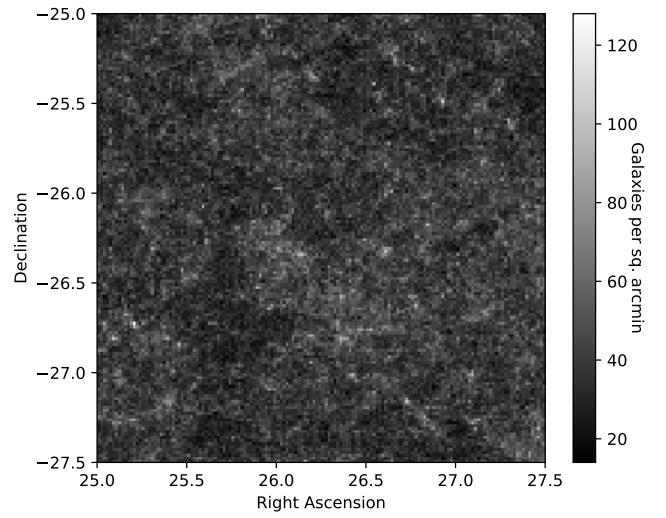


FIG. 3. The distribution of simulated galaxies. The mean galaxy density is 40 arcmin^{-2} .

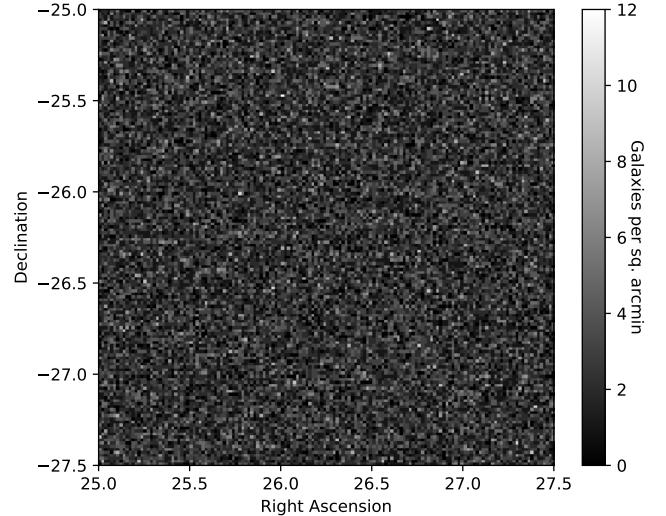


FIG. 4. The distribution of simulated stars. The mean stellar density is 2.5 arcmin^{-2} .

Data Structure) files². These files concatenate all exposures of unique objects to allow for fast access for object-by-object data processing (like shape measurement). Each MEDS file also stores for each object (and stamp) its original SCA, the object position and the stamp position within the SCA, the WCS for each stamp, the PSF model for each object, and other ancillary information. Each MEDS file contains all objects within a $n_{\text{side}} = 512$ Healpixel.

Shape measurement – We utilize the MEDS files to measure the shapes of each galaxy and the PSF. The galaxy shape

² <https://github.com/esheldon/meds>

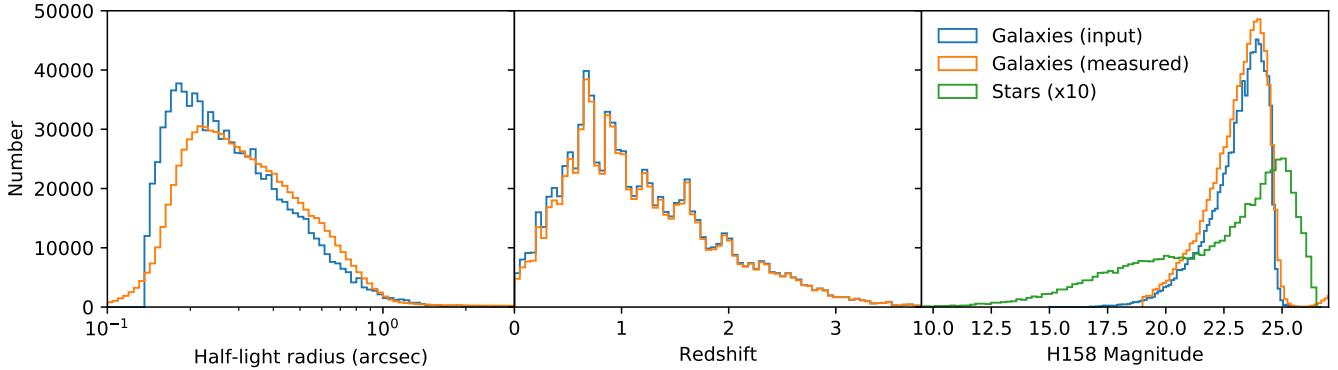


FIG. 5. The input distributions of half-light radius, redshift, and H158 magnitude for galaxies (blue) and stars (orange).

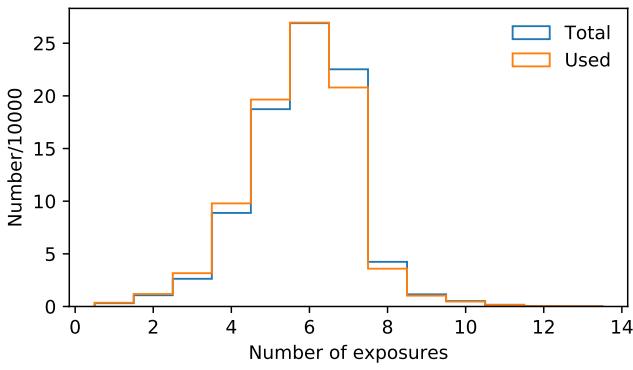


FIG. 6. The number of exposures per object simulated (total) compared to the number used in the shape measurement.

is measured by jointly fitting a two-component model, de Vaucouleurs bulge and exponential disk, across all suitable exposures. Exposures where more than 20% of the pixels are masked (i.e., the centroid falls too close to the edge of the SCA) are rejected. The model fit has x parameters: $e_{1,2}$, $p_{x,y}$, half-light radius, flux, and bulge flux fraction, where $e_{1,2}$ is the component of the ellipticity and $p_{x,y}$ is the pixel centroid offset. The fits are done using the NGMIX³ and MOF⁴ packages [?]. We also measure the PSF size and shape using an adaptive moments method [?]. This stage writes a set of FITS files containing the galaxy and PSF measurement results and relevant truth catalog information.

B. GALSIM

As mentioned above, the image rendering uses the GALSIM software package [?]. This package has been extensively

tested and has been shown to yield very accurate rendered images of galaxies and stars. Notably, the image rendering process has been shown to impart biases in the shapes of galaxies at a level much less than 10^{-3} for the kinds of objects we are simulating here.

The GALSIM package is mostly generic with respect to the telescope and observational strategy, allowing for a wide variety of options in performing the simulation. However, it does have a sub-module (`galsim.wfirst`) that has a number of WFIRST-specific implementation details. Some of the code in this module pre-dates this work, but some of it was developed specifically for this project, especially updating some of the details to match Cycle 7 information, and to reflect new information from laboratory tests of persistence in WFIRST sensors. The values used for this project correspond to the `galsim.wfirst` module in GALSIM release version 2.2.0.

MJ: Note: We should reference a real version number for all this. The code used was technically master branch when Troxel ran it, but we should tag it as an official release soon, which will match all the important aspects of what Troxel used.

1. World coordinate system

The `galsim.wfirst` module has code to provide an estimate of the WFIRST WCS (world coordinate system) for each SCA given a rotation angle, date, and pointing direction. The WCS gives the two-dimensional mapping from (x, y) coordinates on the image to right ascension (RA) and declination (Dec) on the sky. The specific orientations and gaps between the sensors was updated to match the Cycle 7 data as part of the development work for this project.

MJ: ref?

We create our scene of objects in sky coordinates (RA, Dec). All surface brightness profiles in GALSIM are defined in sky coordinates as well, so units for things like the half-light radius are typically arcseconds. We then use a series of pointings and rotation angles designed to match a plausible WFIRST observational schedule.

³ <https://github.com/esheldon/ngmix>

⁴ <https://github.com/esheldon/mof>

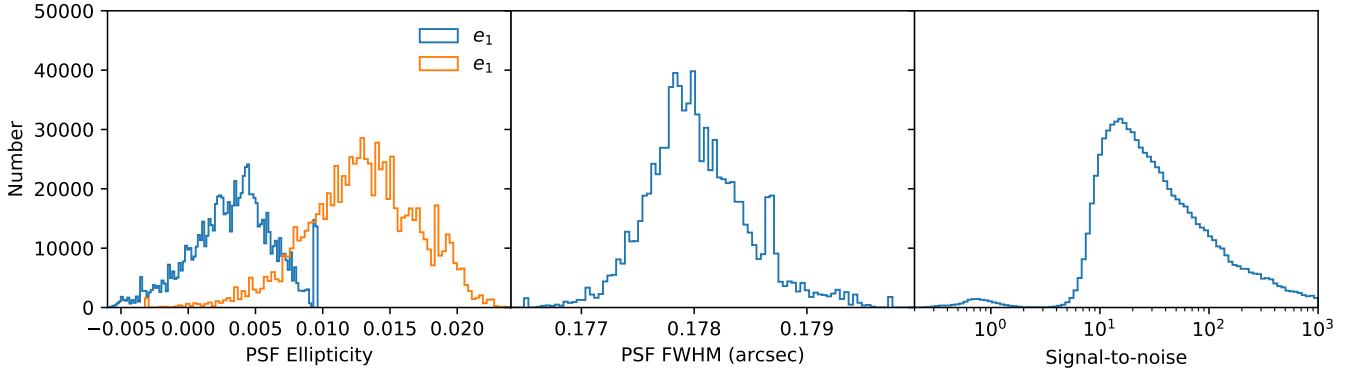


FIG. 7. The measured PSF ellipticity and size, and the signal-to-noise of the galaxy measurements.

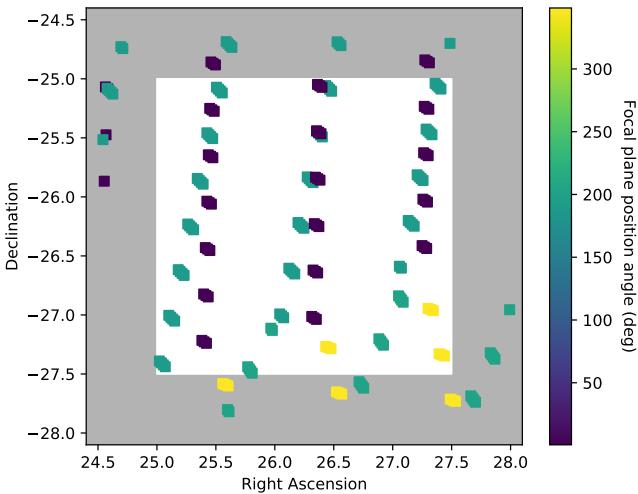


FIG. 8. Pointings that overlap the simulated region (non-shaded). The color of each marker shows the position angle of the focal plane.

MJ: Probably say more about this? Chri's observing sequence of pointings and such. Or maybe you already did elsewhere.

GALSIM is then able to turn these observational parameters into full WCS functions for all 18 SCAs and determine which objects in the scene would be observed by each sensor.

GALSIM automatically accounts for the Jacobian of the WCS transformation when rendering the surface brightness profiles on each sensor's pixels. Details such as the telescope distortion and variable pixel area are correctly accounted for in this process.

2. Point-spread function

For the point-spread function (PSF) we use a model of the WFIRST PSF from the `galsim.wfirst` module. While this module includes a high-resolution Cycle 7 estimate of the

WFIRST spider pattern (i.e. the obscuration of the struts and camera in the pupil plane), we use a faster, low-resolution approximation, which gets the qualitative features correct, but has a slightly different detailed diffraction pattern. For the purposes of this study, we are insensitive to the differences between the two spider patterns, so we did not enable the slower, more accurate option.

The PSF uses position-dependent (Zernike) aberrations, based on an investigation of the field-dependent wavefront errors for Cycle 5.0.6⁵. We are not aware of a more recent update to these numbers. Aberrations between the tabulated positions are estimated using bilinear interpolation of the tabulated values.

The wavelength-dependent features of the PSF, such as the width of the Airy diffraction pattern, and the wavelength-dependence of the aberrations, are taken at the effective wavelength of the observation bandpass. This is an approximation, which leads to an enormous speed up in the rendering time. However, it does omit some interesting and subtle chromatic effects as different parts of a galaxy, with different effective SEDs, would be convolved by slightly different effective PSFs. There are plans to improve the implementation of this aspect of GALSIM, but it cannot currently simulate such effects efficiently enough for our needs.

There are also plans to enable the use of WebbPSF⁶ in `galsim.wfirst` to leverage the work being done on that project to simulate the WFIRST PSF. That PSF is qualitatively similar to what we are using from `galsim.wfirst`, but there are slight differences. We expect that the WebbPSF model is probably more accurate, so we look forward to being able to use that in future simulations.

MJ: Both of the above planned improvements will be done by us. (Probably me in fact.) Not sure if we should allude to that fact or not...

⁵ https://wfirst.gsfc.nasa.gov/science/sdt_public/wps/references/instrument/README_AFTA_C5_WFC_Zernike_and_Field_Data.pdf

⁶ <https://webbpsf.readthedocs.io/en/stable/>

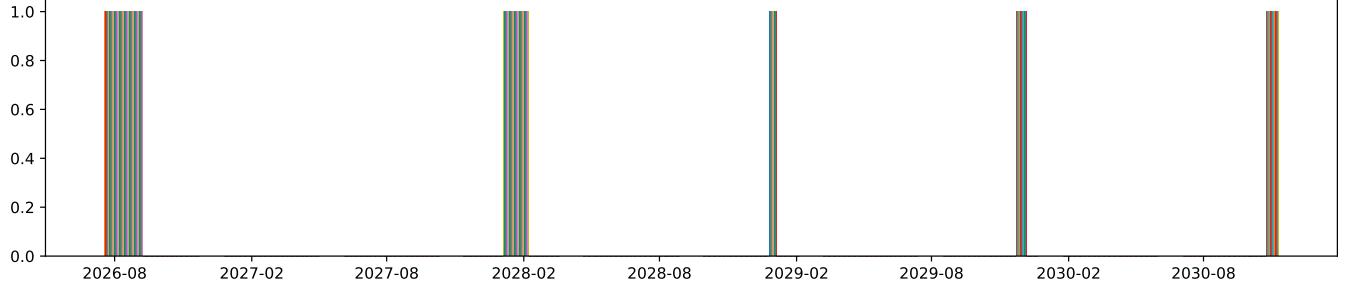


FIG. 9. Dates of observations (is this useful? hard to make it readable)

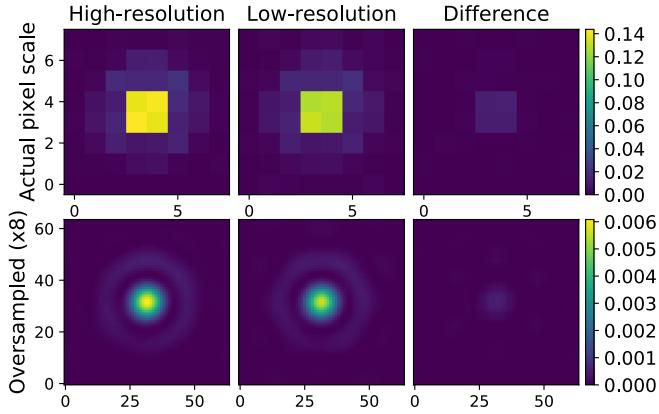


FIG. 10. PSF model for SCA 1. The top row shows the model in native pixel scale, while the bottom row is oversampled by a factor of 8. From left to right: a comparison of the high-resolution ('true') model, the low-resolution model used in the simulation, and the difference of the two models. The color bars are defined by the range of the high-resolution model.

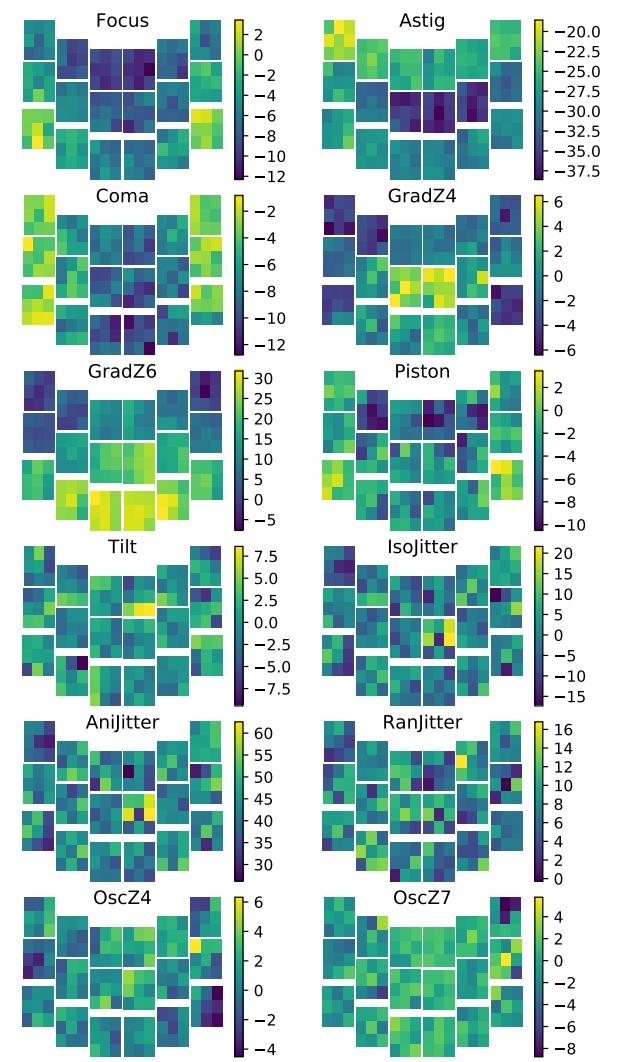
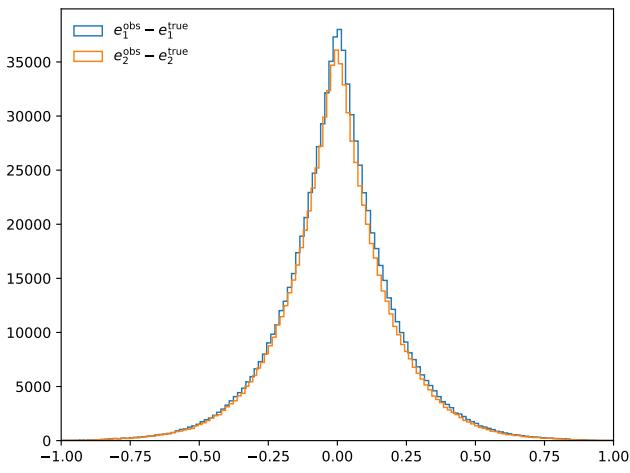


FIG. 12. The binned mean difference in measured e_1 compared to the Fiducial run. Each color bar is in units of 1×10^{-4} .

FIG. 11. A histogram of the difference in the measured shape from the true shape in the Fiducial simulation.



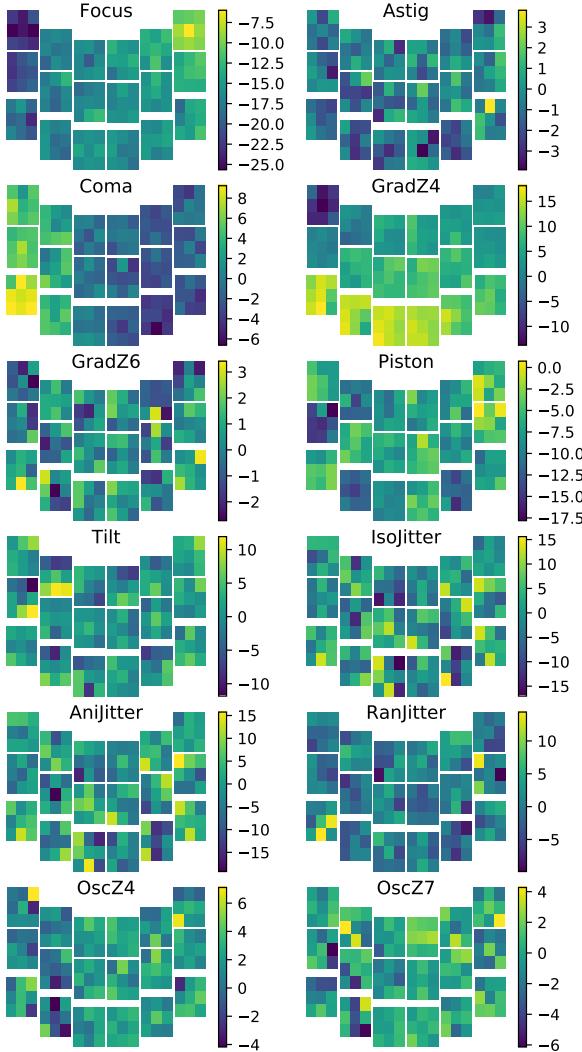


FIG. 13. The binned mean difference in measured e_2 compared to the Fiducial run. Each color bar is in units of 1×10^{-4} .

3. Implemented detector effects

Most of the development of GALSIM has been to render simulations of CCD images. The HgCdTe detectors used by WFIRST are qualitatively similar, but there are significant differences in the physics, which lead to differences in some of the simulation steps.

Reciprocity failure is a non-linear relationship between the voltage response in the detector to the incident flux of photons at low light levels. The exact mechanism of this effect is unknown and hence we lack a good theoretical model. GALSIM uses a power law

$$\frac{p}{p_{\text{nominal}}} = \left(\frac{p_{\text{nominal}}}{f_0 t_{\text{exp}}} \right)^{\frac{\alpha}{\log(10)}} \quad (2)$$

where p_{nominal} is the pixel response (in electrons) that would

have occurred in the absence of reciprocity failure, p is the actual observed response due to reciprocity failure, f_0 is the base flux rate (in electrons/sec) at which the nominal gain was calibrated, t_{exp} is the exposure time, and α is taken to be 6.5×10^{-3} for the WFIRST sensors.

A particularly pernicious effect present in the HgCdTe detectors is known as “persistence”. In a series of images taken sequentially, some small fraction of the charge accumulated in earlier exposures apparently remains in the sensor and appears in later exposures. The effect lasts for many minutes across multiple reset cycles. Therefore, for simulating the effect, we need to keep track of the precise order of the observations, the time of each, and the electron-level (i.e. pre-read-out) images of multiple prior exposures.

The exact functional form of this effect is not very well understood, although some progress is being made in laboratory tests. The functional form for this effect was updated during the Cycle 7 updates to account for recently improved understanding from laboratory measurements. GALSIM now uses a fermi profile when the deposited flux is above the half-well level, and linear below. Above the half-well, the functional form is

$$n_{\text{persist}} = \frac{A (n/n_0)^a \left(\frac{t}{1000 \text{sec}} \right)^{-r}}{\exp(-\frac{n-n_0}{dn}) + 1} \quad (3)$$

where A , n_0 , a , r , and dn are constants estimated from laboratory measurements (and stored in the `galsim.wfirst` module).

In addition to the non-linear pixel response, known as reciprocity failure, there is also a non-linearity in the conversion of accumulated charge to the measured voltage. This is a different effect, which occurs at a different point in the simulation – namely, after the application of dark current and persistence. GALSIM treats this as a modification in the effective number of electrons:

$$n'_e = n_e - 6 \times 10^{-7} n_e^2 \quad (4)$$

where n_e is the actual number of electrons accumulated and n'_e is the effective number to account for the voltage response nonlinearity.

Inter-pixel capacitance (IPC) essentially amounts to a convolution of the image by a 3×3 kernel in pixel coordinates. However, the timing of the convolution is during the readout process, which means that some (but not all) of the noise has already occurred. Thus it cannot be treated as part of the PSF for the purpose of the simulation. It needs to be applied separately after the dark current and Poisson shot noise have been applied, but before the read noise. The IPC coefficients have been measured in the lab for WFIRST detectors; the values used in the `galsim.wfirst` module come from the Cycle 5 estimates.

MJ: It looks like there were Cycle 7 updates to these coefficients. But we don't seem to be using them. Slightly awkward....

C. Galaxy catalogs

The input galaxy catalog is created using a simulated galaxy distribution on the sky taken from one realization of the Buzzard simulation [? ?], to introduce realistic galaxy clustering. Each galaxy is then assigned a random set of photometric properties matching a galaxy from a sample based on the Candels survey that simulates the fiducial WFIRST weak lensing sample selection [?]. We show the galaxy distribution in Fig. 3. We use a galaxy density that is approx. 40 arcmin^{-2} . In Fig. 5, we show the distributions of size, redshift, and H158 magnitude in the Candels sample. We discard less than 1% of the largest objects in the shape measurement stage, however, due to a maximum postage stamp size restriction. In general, the input distribution and properties of galaxies can be easily modified by configuration (i.e., specifying a different input galaxy catalog).

Need lensing selection cuts and maybe short description of candels sample.

D. Star catalog

We simulate the positions and magnitudes in WFIRST bandpasses of input stars using the galaxy simulation Galaxia⁷ [?]. Galaxia uses an analytic model [?] to simulate stars in the galaxy that includes a thin and thick disk with warp and flaring, bulge, and halo components. Stars are simulated to 27th magnitude in V band, extinction is added, and they are uniformly translated to WFIRST bandpasses using the stellar SED of Alpha Lyra derived from HST CALSPEC and packaged with GALSIM.

E. Survey strategy

Contributing: Hirata

F. Simulation implementation for this study

In this work, we are interested in the impact of how a variety of biases in the PSF model propagate to shape measurement and the weak lensing signal. To study this, we produce a set of 13 image simulations that are identical, including noise, modulo a single PSF model change relative to the fiducial simulation in each case. The details of these changes and their impacts are described in more detail in Sec. ???. Shape measurement is then performed on the images with some PSF model bias, but using the fiducial, unbiased PSF model for deconvolution, to simulate an unknown wavefront error.

TABLE I. Table
A summary of the 13 simulation runs. (+ means done)

Run name	PSF change	Mode	Notes
Fiducial	—	—	—
Focus	Z_4	Static	—
Astig	Z_5	Static	—
Coma	Z_7	Static	—
GradZ4	Z_4	Static	Gradient in focal plane
GradZ6	Z_6	Static	Gradient in focal plane
Piston	Z_4	Static	Random per SCA
Tilt	Z_4	Static	Random gradient per SCA
IsoJitter	Gaussian	High-Freq.	Isotropic
AniJitter	Gaussian	High-Freq.	Anisotropic
RanJitter	Gaussian	High-Freq.	15% of pointings
OscZ4	Z_4	Low-Freq.	Time-dependent
OscZ7	Z_7	Low-Freq.	Time-dependent

Several simplifications are employed relative to the generic synthetic survey generation described in Sec. III A to accommodate the computational load of the many realizations of the survey we are producing.

- We simulate objects in a $2.5 \times 2.5 \text{ deg}^2$ patch of the sky.
- We only simulate pointings targeted for the H158 filter. Since we are not simulating chromatic effects, the specific filter choice does not make a large difference in our results. **This will have an impact on the 'averaging' of PSF errors, however, ...need justification or explanation**
- We use a lower-resolution version of the PSF, which significantly speeds up the convolution. The impact of this approximation on the PSF model, in both native and oversampled pixels, can be seen in Fig. 10.
- To better isolate the effects of PSF changes, we only utilize the isolated object postage stamps in shape measurement.
- We do not simulate objects with photometry that would fall below the fiducial weak lensing selection criteria.
- We do not implement a shear calibration scheme like metacalibration, since we only care about changes to the recovered shape between simulation runs.

We simulate a total of 907,170 unique galaxies and 56,128 unique stars across 189 pointings in each of the runs. The distribution of PSF properties and exposures per galaxy are shown in Fig. ??.

⁷ <http://galaxia.sourceforge.net>

IV. WAVEFRONT MODEL ERRORS

- background on psf and wavefront errors
- details of wfirst psf in cycle 7 (6?) baseline
- summary of approach

In this paper, we focus on empirical tests of weak lensing requirements for wave front model control (i.e., the PSF) in WFIRST. **Chris: insert text connecting to early motivation section when written.**

We simulate 13 identical $2.5 \times 2.5 \text{ deg}^2$ HLS survey cutouts: a fiducial survey that represents perfect knowledge of the PSF and 12 iterations to simulate failure modes in the PSF reconstruction. These are split into three types of failure modes: 1) static biases in the model, which are constant as a function of time, 2) high-frequency biases in the model, which correspond to rapidly changing conditions compared to the timescale of a single exposure, and 3) low-frequency biases in the model, which change over lifetime of the mission, but can be considered static over the timescale of a single exposure. In each mode, the (rms) amplitude of the bias corresponds to 0.005 wavelengths (a fiducial wavelength is taken to be 1293 nm), which is equivalent to approx. 6.5 nm. **Chris: connect value back to requirement discussion when written.** These runs are summarized in Table ??.

A. Static biases

We simulate seven static sources of bias in the PSF model. Three of these simulations include a coherent change in the PSF model Zernike coefficients, where the fiducial value is changed by 0.005 wavelengths in each of defocus (Z_4) – *Focus*, oblique astigmatism (Z_5) – *Astig*, and vertical coma (Z_7) – *Coma*. Two simulations include a coherent gradient in the defocus (Z_4) – *GradZ4* – and vertical astigmatism (Z_6) – *GradZ6* – across the focal plane with equivalent rms of 0.005 wavelengths. For speed, these are simulated such that the PSF is constant within a single SCA. Finally, two simulations approximate errors in the mounting of the SCAs: 1) a random vertical mounting offset of up to 0.005 wavelengths is assigned to each SCA – *Piston*, and 2) a random tilt in the x or y direction is assigned to each SCA – *Tilt*, with equivalent rms of up to 0.005 wavelengths. These are modeled as changes in the Z_4 coefficient, with the PSF being evaluated based on the object x - y position within the SCA (i.e., each object assigned a different PSF consistent with this random tilt of the SCA). Potential correlated biases in the WCS model due to these changes are ignored in this work, but should be considered in future studies of the WCS model recovery.

B. High-frequency biases

Three high-frequency resonant modes are simulated to represent residual vibrations of the telescope after orienting to a

new pointing. These are represented by an additional convolution of the image with a Gaussian PSF of rms 0.005 wavelengths. We simulate three cases: 1) an isotropic (about the pointing axis) vibration – *IsoJitter*, 2) an anisotropic vibration modeled with a shear of $e_2 = 0.3$ of the Gaussian – *AniJitter*, and 3) only applying this asymmetric vibration to a random 15% of pointings – *RanJitter*.

C. Low-frequency biases

Heyang: Do we need to describe the details of the gaussian random field generation process in appendix?

Two low-frequency biases are simulated to represent the Zernike coefficients changes caused by thermal drift through exposures. Thermal perturbation propagates to the 4th and 7th Zernike terms, which we simulate respectively as *OscZ4* and *OscZ7*. Here we generate a random time-dependent function $f(t)$ following a power spectrum to quantify the perturbation of Zernike terms. The form of power spectrum is $P(\nu) = \frac{A}{1+(\nu/\nu_0)^2}$, with $A = \frac{\sigma^2}{\pi\nu_0}$, $\sigma = 6\text{nm}$, $\nu_0 = \frac{1}{2\pi\tau} = 3.14 \times 10^{-4}\text{Hz}$ as the time constant $\tau = 1\text{hr}$.

V. RESULTS

Each simulation is analyzed in an identical way, except that shape measurement for each simulation assumes the Fiducial PSF model is the true model, which introduces varying levels of bias. All estimates of the multiplicative and additive bias will be explored relative to the Fiducial simulation run, since we have not employed a calibration scheme. This is justified to first order, since we are only interested in the relative impacts of the PSF model biases. We find that 3% of objects are not included in the shape measurement stage in the Fiducial simulation, due to being too large/bright or because too large a fraction of all cutouts are masked (fall off the edge of an SCA) – see Sec. III A for more details on these selections.

Since the simulated objects have already been pre-selected as objects that should pass the fiducial WFIRST weak lensing selection, we are able to successfully recover a shape fit for more than 99% of the remaining objects – a total of 871,841 galaxies. We do not make an additional selection on objects that would pass the fiducial WFIRST shape selection based on measured properties, since we expect all objects to be within this selection if we were to simulate all remaining pointings in other bandpasses. The recovered multiplicative shear bias is only approx. 2% smaller and the mean shear is unchanged if we make this selection, which removes an additional 35% of objects, almost exclusively due to the signal-to-noise cut.

We present results for the non-Fiducial simulations only for objects that lie in the intersection of successful shape measurement between each simulation and the Fiducial simulation, to allow for 1-1 comparison of the shapes and cancellation of shape noise and sources of photon noise, which are identical in each simulation. We neglect the impact of selection biases here, since implementing a calibration scheme to correct them for objects with an undersampled PSF is beyond the scope of

Run name	$\partial e_1 / \partial \psi$	$\partial e_2 / \partial \psi$
Fiducial	—	—
Focus	$-8.7 \times 10^{-5} \pm 3.2 \times 10^{-6}$	$-2.5 \times 10^{-4} \pm 3.1 \times 10^{-6}$
Astig	$-4.5 \times 10^{-4} \pm 3.1 \times 10^{-6}$	$-1.0 \times 10^{-4} \pm 3.0 \times 10^{-6}$
Coma	$-1.0 \times 10^{-4} \pm 3.2 \times 10^{-6}$	$3.0 \times 10^{-6} \pm 3.1 \times 10^{-6}$
GradZ4	$-8.9 \times 10^{-6} \pm 3.0 \times 10^{-6}$	$1.0 \times 10^{-4} \pm 2.9 \times 10^{-6}$
GradZ6	$2.2 \times 10^{-4} \pm 2.9 \times 10^{-6}$	$4.4 \times 10^{-6} \pm 2.8 \times 10^{-6}$
Piston	$-5.6 \times 10^{-5} \pm 4.8 \times 10^{-6}$	$-1.2 \times 10^{-4} \pm 4.8 \times 10^{-6}$
Tilt	$-3.6 \times 10^{-6} \pm 6.3 \times 10^{-6}$	$1.0 \times 10^{-5} \pm 6.3 \times 10^{-6}$
IsoJitter	$1.2 \times 10^{-3} \pm 7.2 \times 10^{-3}$	$1.8 \times 10^{-3} \pm 7.0 \times 10^{-3}$
AniJitter	$0.29 \pm 6.9 \times 10^{-3}$	$-1.9 \times 10^{-3} \pm 6.8 \times 10^{-3}$
RanJitter	$0.049 \pm 3.2 \times 10^{-3}$	$1.7 \times 10^{-3} \pm 3.1 \times 10^{-3}$
OscZ4	$1.3 \times 10^{-5} \pm 3.9 \times 10^{-6}$	$3.0 \times 10^{-5} \pm 3.9 \times 10^{-6}$
OscZ7	$-1.9 \times 10^{-6} \pm 4.5 \times 10^{-6}$	$-4.4 \times 10^{-6} \pm 4.4 \times 10^{-6}$

this investigation. We anticipate this additional bias will be small, however, since the intersection criteria excludes on average only 0.3% of objects.

A. Summary statistics

The bias in an ensemble shear measurement is typically characterized in the weak limit by

$$e_i^{\text{obs}} = (1 + m_i)e_i^{\text{true}} + b_i. \quad (5)$$

We find the following multiplicative and additive biases in the Fiducial simulation:

$$m_1 = -0.0756 \pm 0.0019$$

$$m_2 = -0.0940 \pm 0.0019$$

$$b_1 = 0.00120 \pm 0.00017$$

$$b_2 = -0.00157 \pm 0.00016.$$

The difference in measured shape vs true input shape (intrinsic shape and shear) is shown in Fig. 11.

For each simulation, we compare the recovered shear to the Fiducial simulation in several ways. First, we compare the change to the recovered values of m and b , along with the changes in the mean galaxy and PSF sizes and ellipticities, where we use the true model PSF images for each simulation and not the biased PSF model used in the shape measurement.

We will also represent the changes in m and b as a response to the wavefront error in units of nm.

B. Spatially-Correlated biases

We are also interested in the spatial correlation of any biases, since this will propagate into the measured shear correlation functions. ...

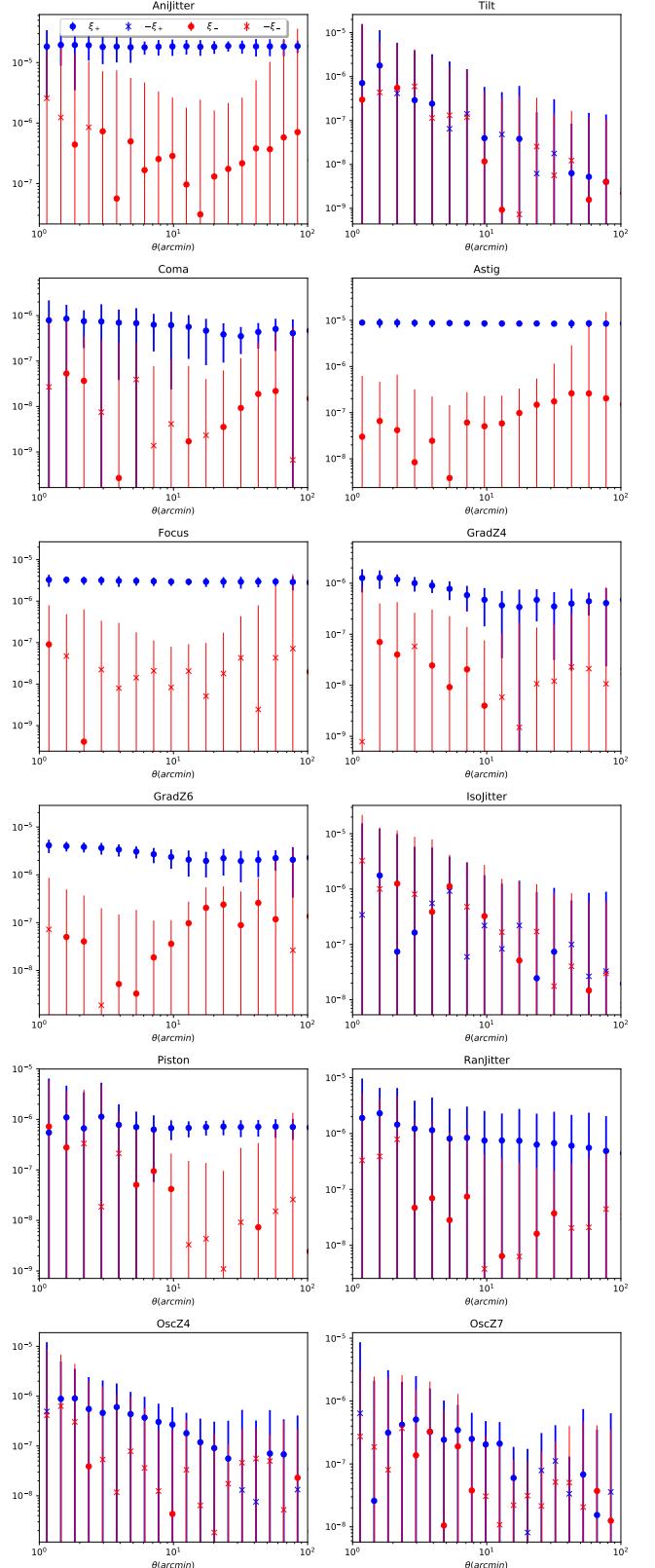


FIG. 14.

Run name	m_1	b_1	m_2	b_2
Fiducial	-0.0756 ± 0.0019	0.00120 ± 0.00017	-0.0940 ± 0.0019	-0.00157 ± 0.00016
Focus	-0.0710 ± 0.0019	0.00064 ± 0.00017	-0.0896 ± 0.0019	-0.00315 ± 0.00016
Astig	-0.076 ± 0.0019	-0.00171 ± 0.00017	-0.0931 ± 0.0019	-0.00162 ± 0.00016
Coma	-0.0769 ± 0.0019	0.00054 ± 0.00017	-0.0954 ± 0.0019	-0.00155 ± 0.00016
GradZ4	-0.0783 ± 0.0019	0.00115 ± 0.00017	-0.0966 ± 0.0019	-0.00090 ± 0.00016
GradZ6	-0.0754 ± 0.0019	0.00268 ± 0.00017	-0.0943 ± 0.0019	-0.00153 ± 0.00016
Piston	-0.0896 ± 0.0019	0.00083 ± 0.00016	-0.1089 ± 0.0019	-0.00232 ± 0.00016
Tilt	-0.1133 ± 0.0019	0.00116 ± 0.00016	-0.1303 ± 0.0019	-0.00149 ± 0.00016
IsoJitter	-0.1164 ± 0.0019	0.00123 ± 0.00016	-0.1333 ± 0.0018	-0.00153 ± 0.00016
AniJitter	-0.0870 ± 0.0019	0.00553 ± 0.00016	-0.1054 ± 0.0019	-0.00158 ± 0.00016
RanJitter	-0.0884 ± 0.0019	0.00195 ± 0.00016	-0.1066 ± 0.0019	-0.00154 ± 0.00016
OscZ4	-0.08278 ± 0.0019	0.0013 ± 0.00016	-0.10050 ± 0.0018	-0.00135 ± 0.00016
OscZ7	-0.08761 ± 0.0019	0.00119 ± 0.00016	-0.10529 ± 0.0019	-0.00158 ± 0.00016

Run name	$\Delta m_1 \times 10^3$	$\Delta b_1 \times 10^4$	$\Delta m_2 \times 10^3$	$\Delta b_2 \times 10^4$
Fiducial	–	–	–	–
Focus	4.649 ± 0.237	-5.57 ± 0.238	4.358 ± 0.228	-15.87 ± 0.304
Astig	-0.226 ± 0.241	-29.14 ± 0.301	0.901 ± 0.209	-0.549 ± 0.189
Coma	-1.328 ± 0.259	-6.66 ± 0.262	-1.469 ± 0.211	0.224 ± 0.341
GradZ4	-2.632 ± 0.267	-0.507 ± 0.210	-2.515 ± 0.228	6.71 ± 0.423
GradZ6	0.216 ± 0.226	14.72 ± 0.624	-0.216 ± 0.226	0.441 ± 0.227
Piston	-13.596 ± 4.986	-3.70 ± 0.313	-14.972 ± 5.256	-7.51 ± 0.34
Tilt	-37.978 ± 8.044	-0.403 ± 0.382	-36.629 ± 7.996	0.731 ± 0.381
IsoJitter	-40.797 ± 8.612	0.244 ± 0.835	-39.489 ± 8.549	0.425 ± 0.905
AniJitter	-11.399 ± 1.057	43.27 ± 0.845	-11.473 ± 0.958	-0.132 ± 0.857
RanJitter	-12.820 ± 4.413	7.50 ± 0.637	-12.603 ± 4.099	0.326 ± 0.431
OscZ4	-7.172 ± 3.070	0.891 ± 0.307	-6.513 ± 3.126	2.148 ± 0.491
OscZ7	-12.067 ± 6.832	-0.083 ± 0.308	-11.351 ± 6.315	-0.173 ± 0.437

C. Fiducial simulation results

VI. CONCLUSION

wrap-up
future work and timeline

ACKNOWLEDGEMENTS

This work used resources on the CCAPP condo of the Ruby Cluster at the Ohio Supercomputing Center [?] and ...OSG, Duke.... Plots in this manuscript were produced partly with MATPLOTLIB [?], and it has been prepared using NASA's

Run name	$\Delta m_1/m_{1fid} \times 10^3$	$\Delta m_2/m_{2fid} \times 10^3$	$\Delta b_1/b_{1fid}$	$\Delta b_2/b_{2fid}$
Fiducial	–	–	–	–
Focus	5.029 ± 0.257	-4.809 ± 0.254	-0.462862 ± 0.055216	1.01220 ± 0.111853
Astig	-2.44 ± 0.260	0.994 ± 0.231	-2.420371 ± 0.291650	0.034980 ± 0.012969
Coma	-1.436 ± 0.281	-1.622 ± 0.233	-0.552890 ± 0.068065	-0.014333 ± 0.021670
GradZ4	-2.847 ± 0.289	-2.775 ± 0.251	-0.042084 ± 0.018697	-0.417965 ± 0.051765
GradZ6	0.233 ± 0.244	-0.337 ± 0.236	1.222809 ± 0.160794	-0.026453 ± 0.014106
Piston	-15.100 ± 5.402	-16.524 ± 5.805	-0.307647 ± 0.041169	0.479172 ± 0.056785
Tilt	-41.084 ± 8.696	-40.428 ± 8.822	-0.033508 ± 0.031404	-0.046589 ± 0.002347
IsoJitter	-44.134 ± 9.314	-43.584 ± 9.439	0.020227 ± 0.069474	-0.027127 ± 0.056551
AniJitter	-12.332 ± 1.129	-12.663 ± 1.050	3.594399 ± 0.441964	0.00838 ± 0.054719
RanJitter	-13.869 ± 4.774	-13.911 ± 4.525	0.622797 ± 0.095830	-0.020775 ± 0.027366
OscZ4	–	–	–	–
OscZ7	–	–	–	–

Astrophysics Data System Bibliographic Services.

Chris: assorted material on analytic flowdown here. this is material from the appendices the WFI Calibration Plan (the relevant parts were written by me, with some commentary from people I should add to the acknowledgements), with data from the wlrequirements github. this was a Phase A snapshot. We can talk about whether numbers should be updated, but my inclination is *not* to do so and just identify the version, since requirements don't update when we revise, e.g., the galaxy population or instrument model (if they do need updating, we should tell the project). I will also clean up some of the references and broken links in here. many can be made shorter in this document than the calibration plan as they refer to standard WL stuff we can just cite. I plan to do some cleanup of this in the next few days, if you want to wait until that is done to read go ahead.

Appendix A: Overview of weak lensing systematics budgeting

This appendix describes the requirements flowdown and error budgeting for the weak lensing program on the WFIRST mission, and documents the detailed rationale behind the summary requirements listed in the WFIRST SRD. This kind of error budgeting has been performed elsewhere in the literature [? ?], but this document focuses on the error terms relevant to WFIRST. For example, the PSFs are based on an obstructed pupil with low-order aberrations rather than using generic formulae involving second moments (some such formulae, including those used in the JDEM and WFIRST-IDRM studies, were for Gaussians).

We set most systematics requirements for this mission on the basis of having systematic errors sub-dominant to statistical errors in the WL shear power spectra or cross-power spectra (or any linear combinations thereof). Exceptions to this policy should be considered in cases where meeting the original systematic budget becomes a cost or complexity driver, or is not possible. Most measurement biases fall into the “additive” or “multiplicative” forms (see §A 1) and will be treated according to the formalism therein. Other sources of systematic error will be treated in their own dedicated sections.

1. Additive and multiplicative biases

The cosmic shear measurement is sensitive to two major types of measurement errors. *Additive bias* or “spurious shear” is a shear signal that is detected even when none is present. *Multiplicative bias* or “calibration bias” is an incorrect response to a real shear, e.g. a shear γ is present in the sky but the measurement yields 1.01γ . Normally, we think of additive biases as resulting from mis-estimation of the PSF ellipticity (or its variation across the sky), whereas multiplicative biases result from mis-estimation of the size of the PSF. However, detector nonlinearities, approximations used in the data processing/analysis pipelines, and uncertainties about the distribution of galaxy morphologies in the sky can also contribute to both types of biases. These biases produce an effect on the observed shear:

$$\gamma(\boldsymbol{\theta}, z; \text{obs}) = [1 + m(z)]\gamma(\boldsymbol{\theta}, z; \text{true}) + c(\boldsymbol{\theta}, z), \quad (\text{A1})$$

where m is the multiplicative bias parameter (possibly redshift-dependent) and c is the additive bias field. The E -mode shear cross-power spectrum between two redshift bins z_i and z_j is modified in the presence of these biases:

$$C_\ell^{z_i, z_j}(\text{obs}) = (1 + m_i)(1 + m_j)C_\ell^{z_i, z_j}(\text{true}) + C_\ell^{c_i, c_j}, \quad (\text{A2})$$

where we write $m_i \equiv m(z_i)$ as a shorthand. To linear order in the biases, the correction to the power spectrum can be written as

$$\Delta C_\ell^{z_i, z_j} = C_\ell^{z_i, z_j}(\text{obs}) - C_\ell^{z_i, z_j}(\text{true}) = (m_i + m_j)C_\ell^{z_i, z_j} + C_\ell^{c_i, c_j}. \quad (\text{A3})$$

2. Setting requirements

We arrange the power spectra are arranged into a vector \mathbf{C} with a covariance matrix Σ . For the WL power spectrum, with N_z redshift bins and N_ℓ angular scale bins, there are $N_\ell N_z(N_z + 1)/2$ power spectra $C_\ell^{z_i, z_j}$; hence \mathbf{C} is a vector of length $N_\ell N_z(N_z + 1)/2$, and Σ is a matrix of size $N_\ell N_z(N_z + 1)/2 \times N_\ell N_z(N_z + 1)/2$. A contaminant that changes the power spectrum by $\Delta \mathbf{C}$ can have its significance assessed by

$$Z = \sqrt{\Delta \mathbf{C} \cdot \Sigma^{-1} \Delta \mathbf{C}}, \quad (\text{A4})$$

which is the number of sigmas at which one could distinguish the correct power spectrum from the contaminated power spectrum. Note that as the survey area Ω is increased, Z will increase as $\propto \Omega^{1/2}$, and hence contaminants $\Delta \mathbf{C}$ must be reduced to keep them below statistical errors. If $Z = 1$, then the power spectrum is biased at the same level as the statistical errors. We use Z as a metric for contaminants, rather than e.g. biases in (w_0, w_a) -space, for generality: if $Z < 1$ then the bias due to $\Delta \mathbf{C}$ in *any* cosmological parameter from the combination of the WFIRST WL power spectrum with *any* other data set(s) from WFIRST or other experiments is $< 1\sigma$; whereas if one based the analysis on biases in (w_0, w_a) then we would need a separate requirement derived from every cosmological analysis planned on WFIRST WL data. Using Z as a metric also enables us to write requirements that do not depend on other cosmological probes (e.g. the WFIRST WL systematic error budget does not change if we discover a new way to reduce the scatter in the SN Ia Hubble diagram), which will help to ensure the stability of our requirements going forward.

Technically the above discussion applies only to the E -mode of spurious shear; we have not set a specific requirement on the B -mode, which contains no cosmological information to linear order and is used as a null test. For the latter reason, we set a requirement on the B -mode that is equal to the requirement on the E -mode, so that the B -mode null test will pass if requirements are met. We also note that the WL analysis includes a range of angular scales, $\ell_{\min, \text{tot}} \leq \ell \leq \ell_{\max, \text{tot}}$; requirements apply to sources of systematic error that affect these scales, i.e. are “in-band” for the WL measurement. The “in-band” qualifier is critical: as an example, pixelization errors can cause shape measurement errors in galaxies that depend on whether the galaxy lands on a pixel center, corner, vertical edge, or horizontal edge. For some shape measurement methods, this error may dramatically exceed the additive systematic error budget, but it is concentrated at very small angular scales (multiples of 2π divided by the pixel scale P , or $2\pi/P = 1.2 \times 10^7$). Our requirements are set on the portion of this power that is within (or mixes into) the band limit, $\ell \leq \ell_{\max, \text{tot}}$ due to e.g. edge effects, selection effects, etc.

Equation (A4) still does not completely define a requirement, since we have not described the redshift or scale dependence of the spurious shear in question. Neither dependence is expected to be trivial: errors in PSF models have a greater impact on shape measurements for higher redshift galaxies, since they tend to be smaller; and the angular power spectrum of PSF model errors should be non-white in a survey strategy that “marches” across the sky, even if heavily cross-linked (there may also be a characteristic scale at the size of the field; for example, a repeating error at the $\sim 0.8 \times 0.4^\circ$ size of the WFIRST field has reciprocal lattice frequencies at $\ell = 450$ and 900 , so a large scale error in the instrument PSF model that is “tessellated” as we tile the sky will appear at these frequencies or multiples thereof). At first, we considered assuming a particular scale and redshift dependence for the errors, but in order to be conservative we would have to assume the worst combination of angular and

redshift dependences. Many of our large sources of systematic error, such as PSF ellipticity due to astigmatism, have predictable dependences (e.g. the systematic error induced in galaxy shears is of the same sign in all redshift bins) that are far from the worst case, and this could lead to over-conservatism in the requirements. Therefore we need a more nuanced approach to the requirements, where the allowed amplitude of each term in the error budget is informed by the structure of the correlations it produces.

Our approach to this problem is to write a script that accepts a specific angular and redshift dependence (“template”) for a systematic error, and returns the amplitude A_0 of the systematic error at which we would have $Z = 1$ (i.e. a 1σ bias on the most-contaminated direction in power spectrum space). For cases where the template is not known (or where we have not done the analysis), the script is capable of searching the space of templates and finding the most conservative choice, i.e. the choice that leads to the smallest value of A_0 . The combined results enable us to build an error tree, where the overall top-level systematics requirement (a limit on Z) can be flowed down to upper limits on each source of systematic error. Finally, some portions of the systematic error budget sum in quadrature (“root-sum-square” or RSS addition) and others linearly; in this document, we carefully account for which is which.

a. Data vector and covariance model

Chris: I think the tools used here are basically obsolete, there are several that are better. but they were what was available in ~2014/5 when much of this work was done. I will shorten the section accordingly.

We use for our data vector the $N_\ell N_z(N_z + 1)/2$ power spectra and cross-spectra. Each ℓ is treated separately, so there are $N_\ell = \ell_{\max,\text{tot}} - \ell_{\min,\text{tot}}$ angular bins; we use $\ell_{\min,\text{tot}} = 10$ and $\ell_{\max,\text{tot}} = 3161$, thereby covering 2.5 orders of magnitude in scale. WFIRST provides little cosmological constraining power at the larger scales due both to the finite size of its survey and due to the large cosmic variance of the lowest multipoles. The smallest scales are generally not used in cosmic shear analyses because the baryonic effects are severe (e.g. [? ?]). We use $N_z = 15$ redshift slices, as shown in Table II. In order to ensure that WFIRST would not become systematics-limited in an extended mission, we set the top-level requirement on systematics to $Z = 1$ for a survey of area $\Omega = 10^4 \text{ deg}^2$ (3.05 sr).

The power spectra were obtained from CLASS [?] using the fiducial cosmology from the *Planck* 2015 “TT,TE,EE+lowP+lensing+ext” results [?]. The shape noise contribution was added to construct \mathbf{C}^{tot} according to

$$C_\ell^{\text{tot},z_i,z_j} = C_\ell^{z_i,z_j} + \frac{\gamma_{\text{rms}}^2}{\bar{n}_i} \delta_{ij}, \quad (\text{A5})$$

where \bar{n}_i is the mean effective number density in galaxies per steradian in redshift slice i , and γ_{rms} is the shape noise expressed as an equivalent RMS shear per component; we take $\gamma_{\text{rms}} = 0.22$.

We approximate Σ using the usual Gaussian covariance matrix formula,

$$\Sigma[C_\ell^{z_i,z_j}, C_{\ell'}^{z_k,z_m}] = \frac{1}{(2\ell + 1)f_{\text{sky}}} \delta_{\ell\ell'} [C_\ell^{\text{tot},z_i,z_k} C_\ell^{\text{tot},z_j,z_m} + C_\ell^{\text{tot},z_i,z_m} C_\ell^{\text{tot},z_j,z_k}], \quad (\text{A6})$$

where $f_{\text{sky}} = \Omega/(4\pi)$. The non-Gaussian contributions to the error covariance matrix are turned off, because since the FoMSWG [?] there has been an ongoing program of using nonlinear transformations on the data to remove them (e.g. [? ?]) and we do not want applications of these novel statistics to WFIRST data to run into systematic error limits. We also turn off astrophysical systematic errors (e.g. intrinsic alignments and baryonic corrections): we do not want to define a weak requirement on WFIRST based on the expectation that astrophysical systematics will be large enough to swamp the observational contribution.

The formal length of the data vector \mathbf{C} is 378240. However, the inversion of Eq. (A6) is accelerated by two realizations: first, that each ℓ is independent, so that Σ is block-diagonal and the contributions of each ℓ to Z^2 can be linearly summed. Secondly, the particular form of Eq. (A6) allows us to write⁸

$$\Delta \mathbf{C} \cdot \Sigma^{-1} \Delta \mathbf{C} = \sum_\ell \frac{(2\ell + 1)f_{\text{sky}}}{2} \sum_{ijkm} \Delta C_\ell^{ij} [C_\ell^{\text{tot}-1}]_{jk} \Delta C_\ell^{km} [C_\ell^{\text{tot}-1}]_{mi}, \quad (\text{A7})$$

where the matrix inverses are $N_z \times N_z$.

⁸ This equation is trivially verified by forming, for each ℓ , linear combina-

tions of the redshift slices that diagonalize $\mathbf{C}_\ell^{\text{tot}}$.

TABLE II. The effective number density in each redshift bin, in units of galaxies/arcmin², used for setting requirements. These are *per bin*, i.e. are $dn_{\text{eff}}/dz \times \Delta z$.

z	n_{eff}	z	n_{eff}	z	n_{eff}
0.10 ± 0.10	3.73	1.10 ± 0.10	3.95	2.10 ± 0.10	1.41
0.30 ± 0.10	2.26	1.30 ± 0.10	3.37	2.30 ± 0.10	1.04
0.50 ± 0.10	3.23	1.50 ± 0.10	2.75	2.50 ± 0.10	0.90
0.70 ± 0.10	6.26	1.70 ± 0.10	1.55	2.70 ± 0.10	0.76
0.90 ± 0.10	2.99	1.90 ± 0.10	1.83	2.90 ± 0.10	0.24

TABLE III. The bands used for the additive systematic errors. There are $N_{\text{band}} = 5$ bands ranging over a total signal band from $\ell_{\text{min,tot}} = 10$ to $\ell_{\text{max,tot}} = 3161$. The fraction of the error budget allocated to each band is also indicated, as are the maximum allowed redshift-independent spurious shear ($A_0^{\text{flat}}(\alpha)$, RMS per component), and the maximum scaling factors for redshift dependence, $S_{\text{max},\pm}(\alpha)$ and $S_{\text{max},+}(\alpha)$.

Band α	$\ell_{\text{min}}(\alpha)$	$\ell_{\text{max}}(\alpha)$	Allocation $Z(\alpha)$	$A_0^{\text{flat}}(\alpha)$	$S_{\text{max},\pm}(\alpha)$	$S_{\text{max},+}(\alpha)$
0	31	99	0.2500	6.791×10^{-5}	8.876	2.878
1	100	315	0.2500	9.697×10^{-5}	5.862	2.100
2	316	999	0.2500	1.360×10^{-4}	3.712	1.544
3	1000	3161	0.2500	1.855×10^{-4}	2.193	1.166

b. Implementation: additive systematics

Each additive systematic error is taken to have an angular dependence given by some template T_ℓ , and a redshift dependence given by a set of weights $w_i = w(z_i)$. That is, there is a reference additive shear c_{ref} , with the additive shear in redshift bin i given by $c(z_i) = w_i c_{\text{ref}}$. For example, a systematic error independent of redshift bin would be specified with $w_i = 1$ for all i . The reference signal is taken to have a power spectrum proportional to the template: $C_\ell^{c_{\text{ref}}} = A_0^2 T_\ell$, and the template is normalized so that c_{ref} has variance 1 per component (from in-band fluctuations):

$$\sum_{\ell=\ell_{\text{min,tot}}}^{\ell_{\text{max,tot}}} \frac{2\ell+1}{4\pi} T_\ell = 1. \quad (\text{A8})$$

The additive cross-power spectrum is then

$$C_\ell^{c_i, c_j} = A_0^2 w_i w_j T_\ell, \quad (\text{A9})$$

and the total RMS per component of the spurious shear in bin i is $A_0 |w_i|$.

The additive systematic errors can have various scale dependences. We therefore consider a suite of N_{band} disjoint angular templates that cover the shape measurement band. Each template satisfies the normalization rule, Eq. (A8), and has $\ell(\ell+1)T_\ell/(2\pi) = \text{constant}$:

$$T_\ell^{(\alpha)} = \left[\sum_{\ell'=\ell_{\text{min}}(\alpha)}^{\ell_{\text{max}}(\alpha)} \frac{2\ell'+1}{\ell'(\ell'+1)} \right]^{-1} \frac{4\pi}{\ell(\ell+1)} \times \begin{cases} 1 & \ell_{\text{min}}(\alpha) \leq \ell \leq \ell_{\text{max}}(\alpha) \\ 0 & \text{otherwise} \end{cases}, \quad \alpha = 0, 1, \dots, N_{\text{band}} - 1. \quad (\text{A10})$$

The current bands are displayed in Table III. Each band α is allowed a contribution to the total error $Z(\alpha)$. Since there are no statistical correlations between different ℓ s in the covariance matrix Σ , the $Z(\alpha)$ can be quadrature-summed (see Eq. A4). However, additive systematic error is positive in the sense that it adds rather than subtracts power; thus the power spectrum error vectors ΔC from two sources of additive systematic error contributing to the same angular bin are not orthogonal and the Z 's should be added linearly. Another way to think of this is that since Z is proportional to the square of the RMS shear, $Z \propto A_0^2$, quadrature-summation of the additive shear is equivalent to linear summation of the Z -values.

The allocations for each bin $Z(\alpha)$ have been set to $\sqrt{0.25/N_{\text{band}}}$, so that in an RSS sense 25% of the systematic error budget is allocated to additive shear.

The construction of Z -values for each angular band and each additive systematic is mathematically sufficient to build the error budget. However, they can be difficult to conceptualize. Therefore, we introduce some equivalent notation to describe the WL error budget. For each angular template, we introduce a limiting amplitude $A_0^{\text{flat}}(\alpha)$, defined to be the amplitude A_0 at which we would saturate the requirement on $Z(\alpha)$ for bin α in the case of a redshift-independent systematic $w_i = 1 \forall i$. That is, if the additive systematics did not depend on redshift, we could tolerate a total additive systematic shear of A_0^{flat} (RMS per component) in band α . We also introduce a scaling factor $S[\mathbf{w}, \alpha]$ for a systematic error

$$S[\mathbf{w}, \alpha] = \frac{Z(\alpha) \text{ for this } w_i}{Z(\alpha) \text{ for all } w_i = 1} \quad (\text{A11})$$

that depends on the redshift dependence w_i . An additive systematic error that is independent of redshift will have $S = 1$. A systematic that is “made worse” by its redshift dependence will have $S > 1$, and a systematic that is “made less serious” by its redshift dependence will have $S < 1$. The requirement that the (linear) sum of Z s not exceed $Z(\alpha)$ thus translates into

$$\sum_{\text{systematics}} [A(\alpha)]^2 \times S[\mathbf{w}, \alpha] \leq [A_0^{\text{flat}}(\alpha)]^2, \quad (\text{A12})$$

where $A(\alpha)$ is the RMS additive shear per component due to that systematic.

In most cases, we will take the “reference” additive shear to be the additive shear in the most contaminated redshift slice; in this case, $w_i = 1$ for that slice, and $|w_i| \leq 1$ for the others. Under such circumstances, we can determine a *worst-case scaling factor* $S_{\max, \pm}(\alpha)$, which is the largest value of $S[\mathbf{w}, \alpha]$ for any weights satisfying the above inequality. We may also determine a worst-case scaling factor $S_{\max, +}(\alpha)$ conditioned on $0 \leq w_i \leq 1$, i.e. for sources of additive shear that have the same sign in all redshift bins. The search within these spaces is simplified by the fact that – according to Eq. (A7) – the contribution to $S[\mathbf{w}, \alpha]$ considering only a single value of ℓ reduces to a semi-positive-definite quadratic function of w . Therefore the worst-case weights $\{w_i\}_{i=1}^{N_z}$ always occur at the corners of the allowed cube in N_z -dimensional \mathbf{w} -space, and we can simply search the 2^{N_z} corners by brute force.

c. Implementation: multiplicative systematics

Chris: write this

3. Flow-down to PSF requirements

We need a model to associate high-level requirements for each source of systematic error (e.g. a limit on Z , A_0 , etc.) to low-level requirements on its source (e.g. wavefront drift rate in nm/hr). This section collects the models used for this flow-down for PSF-related systematics. Note that none of the models described here anywhere near a complete description of the WFIRST telescope, instrument, and pipeline. Rather the intention is to understand the sensitivity of WFIRST weak lensing to specific types of errors.

a. Shape model

In order to translate a requirement on spurious shear c into requirements on lower-level quantities, we need to know how a given effect – e.g. an error in the PSF model – affects the shear measurement. In principle this depends on the shape measurement algorithm, which presents a difficulty because the final shape measurement algorithm used for WFIRST has not yet been selected (and the true “final” version may not be set until the analysis stage). However, most methods of measuring shear have some basic properties in common – if e.g. the true PSF has greater e_1 than the model (i.e. is elongated in the x -direction), then the inferred shear in that region of the sky will also have greater c ; the coefficient depends on the exact distribution of galaxy morphologies and sizes, the weighting of galaxy isophotes, the way in which the PSF ellipticity varies with radius, etc. Therefore we will consider a range of methods when we determine the quantity $\partial \gamma_{\text{obs}}(z_i)/\partial X$, where γ_{obs} is the measured shear in a region (and in redshift slice i) and X is any quantity on which we want to set a knowledge requirement. The spurious shear in bin i is taken to be

$$c_i = c(z_i) = \frac{\partial \gamma_{\text{obs}}(z_i)}{\partial X} \Delta X, \quad (\text{A13})$$

where $\Delta X = X_{\text{true}} - X_{\text{model}}$ is the error in knowledge of X . In the context of the additive systematic errors, the ratios of the partial derivatives $\partial \gamma_{\text{obs}}(z_i)/\partial X$ set the redshift slice dependence: if $i(\max)$ is the redshift bin with the largest derivative (in absolute value) then

$$w_i = \frac{\partial \gamma_{\text{obs}}(z_i)/\partial X}{\partial \gamma_{\text{obs}}(z_{i(\max)})/\partial X} \quad (\text{A14})$$

and the reference additive shear is $c_{\text{ref}} = c_{i(\max)}$.

Chris: Finish! Also describe implementation for multiplicative systematics.

In both the additive and multiplicative cases, to compute the derivatives relating shear systematics quantities (with requirements set in §A 1) to lower-level quantities (knowledge of X) we need a simulation package that generates galaxies, measures their photometric properties, generates a shear estimate, and repeats this varying the quantity X . For the purpose of Phase A requirements budgeting, we used a very simple model that enables rapid exploration of the parameter space.

The “galaxy” used is based on a galaxy with an exponential profile, $f_{\text{circ}}(\mathbf{x}) \propto e^{-1.67834|\mathbf{x}|/r_{\text{eff}}}$, where r_{eff} is the half-light radius. It can optionally be sheared by applying a *finite* shear γ to arrive at the galaxy $f(\mathbf{x})$. This is convolved with the PSF (see the description of models: §A 3 b) to arrive at a simulated image $I(\mathbf{x})$. The ellipticity of this image is then measured as in Appendix ??.

The observed 2-component ellipticity e_I of the galaxy is related to the shear by a 2×2 responsivity matrix

$$R_{ij} = \frac{\partial e_{I,i}}{\partial \gamma_j} = \mathcal{R} \delta_{ij} + R_{ij}^{\text{aniso}}, \quad (\text{A15})$$

which we have decomposed into an isotropic part \mathcal{R} and a traceless matrix R^{aniso} characterizing the anisotropic part of the responsivity. The inverse of the responsivity matrix relates a bias in the galaxy ellipticities to a bias in the shear:

$$c_i = \sum_{j=1}^2 [\mathbf{R}^{-1}]_{ij} \frac{\partial e_{I,j}}{\partial X} \Delta X. \quad (\text{A16})$$

Since the isotropic part of the responsivity dominates except for extreme PSF ellipticity, anisotropic noise correlations, etc., we take the isotropic part and write

$$\frac{\partial \gamma_{\text{obs},i}(z_k)}{\partial X} = \left\langle \mathcal{R}^{-1} \frac{\partial e_{I,i}}{\partial X} \right\rangle, \quad (\text{A17})$$

where the average is taken over the source galaxies in that redshift bin. The various partial derivatives are easily computed as finite differences of the galaxy simulation and ellipticity measurement process.

As this model is intended to be simple, the average is taken only over the distribution of source sizes r_{eff} – we do not include the intrinsic source ellipticity or a distribution of Sersic indices.

b. PSF models

The PSF models used to assess most requirements are highly simplified. In most cases, the intention of these models is *not* to be accurate to the level of a few parts in 10^4 (which would require a great deal more sophistication, and has not been achieved by pure first principles modeling in any telescope used for weak lensing – some empirical tuning of the final PSF model, e.g. by principal components analysis, has always been required). Rather the intention is to address sensitivities: e.g. how far the focus would need to move during an exposure before we care; whether wavefront error jitter is important. As the mission develops, PSF modeling will include many more effects, and we aim in this document to develop requirements that can be flowed to individual sub-groups in the mission (e.g. a STOP analysis should tell us how much of the total PSF error budget is consumed, without having to put in a new model for interpixel capacitance).

Just as for shape measurement (§A 3 a), we take a simplified model here. This model used here takes the Fourier transform of an annular pupil with aberrations appearing as contributions to the phase. The resulting “optical” PSF is then convolved with a detector response that includes a tophat and charge diffusion. For HgCdTe detectors, we take the charge diffusion length to be $2.94 \mu\text{m}$ rms per axis [?].⁹

⁹ This was measured on an H2RG. At the time we had to fix this for Phase

A requirements flow-down, we did not have a measurement on the H4RG.

Chris: The data on the H4RGs show that this was too large, but we are still working on the measurement, will need to figure out what to say here.

Other effects that are either known to be significant at the level needed for WFIRST weak lensing or are likely to be significant, such as polarized transfer through the telescope, angle of incidence variations in mirror reflectivity, interpixel capacitance in the detectors, etc. are not included. Cases where they may influence stability requirements in a way not captured by the simple pupil model are considered separately.

The default model had the spider turned off. The spider has a major effect on the morphology of the wings of the PSF, producing 12 spikes; studies of the AFTA pupil in the past have found effects of the spider on the coefficients herein at the few percent level. The spider further leads to an asymmetric pupil, i.e. with odd-order modes in the decomposition of the amplitude, but this has no appreciable effects on the relation of ellipticity to low-order Zernike modes.¹⁰ Therefore the spider is *not* expected to have a significant effect on the PSF determination procedures or the stability requirements.

Chris: maybe not necessary to go into detail here, we will probably never update the analytic model to treat the spider

Studies were carried out with both monochromatic and polychromatic PSFs. From the perspective of stability requirements, the most important issue is that while the ellipticity of a *monochromatic* PSF is *quadratic* in the wavefront error [?], the ellipticity of a *polychromatic* PSF can contain linear terms in the wavefront error if there is lateral color, as occurs in WFIRST since the beam intersects the filter at a nonzero angle of incidence.¹¹ Thus for small wavefront errors it is possible that a wavefront stability budget based on the quadratic dependences of PSF ellipticity (e.g. defocus \times astigmatism) and monochromatic PSFs may be too optimistic. The lateral color mixes with odd aberrations to produce PSF ellipticity, so a special “coma \times lateral color” contribution is included.

Chris: Need to re-visit the lateral color issue. We didn't specify a separate requirement on this. but not for this paper, maybe

Appendix B: Requirements on wavefront stability for the PSF calibration

The determination of the PSF in imaging mode (§??) will be based on an empirical (principal components or more advanced version thereof) approach. Central to this approach is that we must limit the number of possible principal components in the data by limiting the number of properties of the PSF that vary from one image to another. The WFIRST approach is to keep the PSF stable during an exposure so that no parameters are needed to describe time dependence of the PSF during an exposure. We make one exception to this policy for image motion, since at the WFIRST WL level of precision this is unavoidable. Thus the requirement is for the optics + image motion PSF to be the convolution of the optics PSF with a kernel coming from the image motion, with small residuals. Here “small” means that the residual error must fit within the overall error budget for PSF (or shear) errors.

Requirements are derived for the two major sources of wavefront change: thermally induced drifts (§B 1) and jitter induced by vibrations from e.g. the reaction wheels (§B 2).

1. Wavefront drift

a. Flowdown methodology

In general, we suppose that there is a vector of parameters \mathbf{p} that determines the PSF in each exposure (including its field dependence). Some of these are associated with the equilibrium wavefront – this is the subject of this section – whereas others are associated with image motion, jitter, detector properties, etc. The amplitudes $\psi_i(\theta)$ of each Zernike component of the wavefront error – which depend on field position θ – are functions of these parameters, and will each have their own time dependence $\psi_i(\theta; t)$. This induces a time dependence in the PSF $G(\mathbf{x}; \theta; t)$, and hence in the observed shear γ_{obs} for an object.

We may write the amplitudes ψ_i at a given position as a vector ψ of length N_{Zern} , where N_{Zern} is the number of Zernike coefficients kept. We normalize the Zernike modes to unit RMS, so that $|\psi(\theta)|$ is the RMS wavefront error at position θ . That

¹⁰ It is known that an odd-order mode in the phase can mix with other asymmetric phase modes to produce PSF ellipticity, e.g. if one introduces a large trefoil t then the ellipticity develops a linear term in coma, proportional to tc^* [?]. However, an amplitude feature with 3-fold or other odd symmetry, such as the spider, does not lead to such an effect.

¹¹ In CCD systems lateral color can be produced in the detector due to the large mean free path of red photons in silicon. For WFIRST, using thin direct-bandgap detectors, this effect is much smaller than the lateral color induced by the optics.

is, we write the wavefront error at pupil position $\boldsymbol{\eta}$ and field position $\boldsymbol{\theta}$ as

$$\psi(\boldsymbol{\eta}; \boldsymbol{\theta}) = \sum_{n=2}^{\infty} \sum_m \sqrt{n+1} \psi_{nm}(\boldsymbol{\theta}) R_n^m(\rho) \times \begin{cases} 1 & m = 0 \\ \sqrt{2} \cos m\varphi & m > 0 \\ \sqrt{2} \sin m\varphi & m < 0 \end{cases}, \quad (\text{B1})$$

where ρ is the radius of the pupil position normalized to 1 at the edge, and φ is the polar angle in the pupil plane, m is summed over integers with the same parity as n (both odd or both even) and $|m| \leq n$ (so that there are $n+1$ terms in the m -sum), and R_n^m is the Zernike polynomial with normalization $R_n^m(1) = 1$. The factor of $\sqrt{n+1}$ and (sometimes) $\sqrt{2}$ guarantee the unit normalization of the RMS over the unit disc.

If the wavefront is drifting over time, then to first order in the drift rate we may write

$$\psi_i(\boldsymbol{\theta}; t) = \psi_i(\boldsymbol{\theta}; t_0) + \dot{\psi}_i(\boldsymbol{\theta})(t - t_0), \quad (\text{B2})$$

where t_0 is the central epoch chosen and $-\frac{1}{2}\Delta t < t - t_0 < \frac{1}{2}\Delta t$. Again to linear order in $t - t_0$, the PSF that is determined by a least-squares fit with uniform weighting in time will have an expectation value that is $G(\mathbf{x}; \boldsymbol{\theta}; t_0)$. There is then a corresponding error in the shear in a given redshift bin z_k :

$$c_{k,i}(t) = \sum_j \frac{\partial \gamma_{\text{obs},i}(z_k)}{\partial \psi_j} \dot{\psi}_j(\boldsymbol{\theta})(t - t_0), \quad (\text{B3})$$

where in this equation k denotes a redshift bin and i denotes a component. Taking just the most strongly affected (in the sense of $|c|$) redshift bin to start as the reference, we see that

$$|c_{\text{ref}}(t)| \leq \left\| \frac{\partial \gamma_{\text{obs},\text{ref},i}}{\partial \psi_j} \right\| |\dot{\psi}(\boldsymbol{\theta})| |t - t_0|, \quad (\text{B4})$$

where $\|\cdot\|$ denotes an operator norm (i.e. the maximum singular value of the $2 \times N_{\text{Zern}}$ matrix). The variance of c per component (i.e. divided by 2) is

$$A^2 \equiv \frac{1}{2} \langle |c_{\text{ref}}|^2 \rangle \leq \frac{1}{2} \left[\left\| \frac{\partial \gamma_{\text{obs},\text{ref},i}}{\partial \psi_j} \right\| |\dot{\psi}(\boldsymbol{\theta})| \right]^2 \langle (t - t_0)^2 \rangle; \quad (\text{B5})$$

the last expectation value is $\frac{1}{12}\Delta t^2$ with the average taken over a uniform interval, leading to

$$A \leq \frac{1}{\sqrt{24}} \left\| \frac{\partial \gamma_{\text{obs},\text{ref},i}}{\partial \psi_j} \right\| |\dot{\psi}(\boldsymbol{\theta})| \Delta t. \quad (\text{B6})$$

Thus from a requirement on A , a determination of the matrix $\partial \gamma_{\text{obs},\text{ref},i} / \partial \psi_j$, and an interval of time Δt , we can set a requirement on the wavefront drift rate $|\dot{\psi}|$. The matrix $\partial \gamma_{\text{obs},\text{ref},i} / \partial \psi_j$ depends on the static aberration pattern and its determination is described below. The interval Δt for PSF fitting is a free parameter, and the wavefront drift rate requirement is tighter if Δt is increased. This must be traded against the *statistical* error in the PSF solution, where the target precision is easier to achieve if the time baseline Δt used in fitting the model is increased.

b. Sensitivity matrix

From Eq. (B6), we see that a key step is to compute the sensitivity matrix $\partial \gamma_{\text{obs},\text{ref},i} / \partial \psi_j$. Unfortunately, this matrix depends on the specific combination of static wavefront errors, because $\gamma_{\text{obs},\text{ref}}$ is not a linear function of ψ . Indeed, due to symmetries the possible form of $\gamma_{\text{obs},\text{ref}}$ is restricted, with the result that $\partial \gamma_{\text{obs},\text{ref},i} / \partial \psi_j$ may be suppressed at zero wavefront error ($\psi = 0$) and be much larger in the realistic case where $\psi \neq 0$ [e.g. ?]. Therefore we must search the entire space of possible wavefront errors ψ – bounded by the top-level requirement that $|\psi| < 90 \text{ nm}$ – to find the place where the operator norm is maximized.

The requirement that the PSF inverts (i.e. preserves ellipticity and hence spurious shear) under $\psi \rightarrow -\psi$ implies that $\gamma_{\text{obs},\text{ref}}$ is an even function of ψ (this statement remains true even for an asymmetric pupil, due e.g. to the spider). For a circularly symmetric pupil, we find the further restrictions that

$$\begin{aligned} \gamma_{\text{obs},\text{ref}1} &= C_{fa} \psi_{20} \psi_{22} + C_{sa} \psi_{40} \psi_{22} + C_{cc} (\psi_{31}^2 - \psi_{3-1}^2) + C_{ct} (\psi_{31} \psi_{33} + \psi_{3-1} \psi_{3-3}) + \dots \text{ and} \\ \gamma_{\text{obs},\text{ref}2} &= C_{fa} \psi_{20} \psi_{2-2} + C_{sa} \psi_{40} \psi_{2-2} + 2C_{cc} \psi_{31} \psi_{3-1} + C_{ct} (\psi_{31} \psi_{3-3} - \psi_{3-1} \psi_{33}) + \dots, \end{aligned} \quad (\text{B7})$$

where we have taken the lowest-order aberrations (focus, astigmatism, coma, trefoil, and spherical) as these dominate the wavefront stability budget. With the wavefront error vector written in this order, $\psi = (\psi_{20}; \psi_{22}, \psi_{2-2}; \psi_{31}, \psi_{3-1}; \psi_{33}, \psi_{3-3}; \psi_{40})$, we find a sensitivity matrix

$$\mathbf{M}^T = \left[\frac{\partial \gamma_{\text{obs,ref},i}}{\partial \psi_j} \right]^T = \begin{pmatrix} C_{fa}\psi_{22} & C_{fa}\psi_{2-2} \\ C_{fa}\psi_{20} + C_{sa}\psi_{40} & 0 \\ 0 & C_{fa}\psi_{20} + C_{sa}\psi_{40} \\ 2C_{cc}\psi_{31} + C_{ct}\psi_{33} & 2C_{cc}\psi_{3-1} + C_{ct}\psi_{3-3} \\ -2C_{cc}\psi_{3-1} + C_{ct}\psi_{3-3} & 2C_{cc}\psi_{31} - C_{ct}\psi_{33} \\ C_{ct}\psi_{31} & -C_{ct}\psi_{3-1} \\ C_{ct}\psi_{3-1} & C_{ct}\psi_{31} \\ C_{sa}\psi_{22} & C_{sa}\psi_{2-2} \end{pmatrix} \quad (\text{B8})$$

(we show the transpose here for ease of display; the operator norm is the same).

We want a limit on the maximum singular value of Eq. (B8), subject to a limit on $|\psi|$. To do so, let us first consider writing the singular value decomposition $\mathbf{M} = \mathbf{UDV}^T$, where \mathbf{U} is a 2×2 orthogonal matrix, \mathbf{D} has 2 diagonal non-negative entries in non-increasing order ($D_{11} \geq D_{22}$) and is otherwise zeroes (and has dimension $2 \times N_{\text{Zern}}$), and \mathbf{V} is $N_{\text{Zern}} \times N_{\text{Zern}}$. Here \mathbf{U} is simply a rotation of the shear derivative, and due to circular symmetry can be set to the identity by rotating the entire aberration pattern. Thus without loss of generality we can consider cases where \mathbf{U} is the identity, and then

$$\|\mathbf{M}\| = \sqrt{\sum_j \left(\frac{\partial \gamma_{\text{obs,ref},1}}{\partial \psi_j} \right)^2} = \sqrt{\psi^T \Lambda \psi} \leq \|\Lambda\| |\psi|, \quad (\text{B9})$$

where we used the fact that \mathbf{M} is a linear function of ψ and defined the matrix Λ to be the matrix of derivatives of the first row of \mathbf{M} :

$$\Lambda = \begin{pmatrix} 0 & C_{fa} & 0 & 0 & 0 & 0 & 0 & 0 \\ C_{fa} & 0 & 0 & 0 & 0 & 0 & 0 & C_{sa} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2C_{cc} & 0 & C_{ct} & 0 & 0 \\ 0 & 0 & 0 & 0 & -2C_{cc} & 0 & C_{ct} & 0 \\ 0 & 0 & 0 & C_{ct} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{ct} & 0 & 0 & 0 \\ 0 & C_{sa} & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{B10})$$

which has norm

$$\|\Lambda\| = \max \left\{ \sqrt{C_{fa}^2 + C_{sa}^2}, |C_{cc}| + \sqrt{C_{cc}^2 + C_{ct}^2} \right\}. \quad (\text{B11})$$

(There are both even-aberration and odd-aberration sectors of this matrix; the operator norm is determined by whichever has greater leverage on the spurious shear. In most cases, the even sector – the first term – is dominant.)

We are not quite done because we have not specified the redshift or scale dependence of this systematic. Since C_{fa} is usually dominant, we adopt its redshift dependence to determine the weights $w(z_i)$, with the last bin as the reference bin because it is the most heavily contaminated – the galaxies are smallest in that bin and the C -coefficients are largest. However, the weights $w(z_i)$ obtained from C_{ct} (the next largest coefficient) are only slightly different.

Using model P1 for the PSF and G1 for the shear determination, the greatest contamination is in the J -band, where $\|\Lambda\| = 1.25 \times 10^{-5} \text{ nm}^{-2}$. With a limit on the total wavefront error of $|\psi(\theta)| < 92 \text{ nm}$, we find

$$A \leq \frac{1}{\sqrt{24}} \|\Lambda\| |\psi| |\dot{\psi}(\theta)| \Delta t = 2.35 \times 10^{-4} \text{ nm}^{-1} \times |\dot{\psi}(\theta)| \Delta t. \quad (\text{B12})$$

The current wavefront drift allocation is that $\Delta t = 1$ exposure (140 s) and $|\dot{\psi}(\theta)| \Delta t < 0.37 \text{ nm}$, which produces a spurious shear of 8.7×10^{-5} , RMS per component. However the S -factor for the focus×astigmatism mode is 0.49 in the worst angular bin, so the implied spurious shear is $AS^{1/2} = 6.1 \times 10^{-5}$. The requirements in Table III give a top-level error of 2.6×10^{-4} ; 5.5% of the additive shear systematic error, in an RSS sense, is currently being taken up by wavefront drift.

2. Wavefront jitter

The wavefront jitter is handled by a similar calculation to the wavefront drift. The principal difference is that we are now interested in the spurious shear from a PSF that is the superposition of many instantaneous PSFs with different wavefronts. Moreover, the PSFs can have different line-of-sight positions, so instead of simply considering the covariance matrix of the Zernike amplitudes, we must also consider the line-of-sight motion (parameterized by θ_x and θ_y). The spurious shear thus depends on the full covariance matrix of the Zernike amplitudes ψ and the line-of-sight motion θ . Of this covariance matrix, the “line-of-sight block” $\text{Cov}(\theta, \theta)$ corresponds to simple image motion, and is not related to wavefront jitter. On the other hand, the blocks $\text{Cov}(\theta, \psi)$ and $\text{Cov}(\psi, \psi)$ involve the wavefront jitter, and their effects on $\gamma_{\text{obs,ref}}$ must be treated here.

We can then write the matrix of second derivatives:

$$\gamma_{\text{obs},i}(z_k) = \gamma_{\text{obs},i}(z_k)|_{\text{no wf jitter}} + \sum_{aj} K_{iaj}^{\text{LOS,WFE}}(z_k) \text{Cov}(\theta_a, \psi_j) + \frac{1}{2} \sum_{jj'} K_{ija}^{\text{WFE,WFE}}(z_k) \text{Cov}(\psi_j, \psi_{j'}). \quad (\text{B13})$$

The matrix $\mathbf{K}^{\text{WFE,WFE}}$, describing how much small high-frequency vibrations of the wavefront impact the shear, has a dependence on redshift bin z_k , shear component i , and the Zernike modes j and j' . The matrix $\mathbf{K}^{\text{LOS,WFE}}$ describes the effects of correlations between LOS motion and wavefront jitter.

The matrix \mathbf{K} in principle varies with the wavefront error, but since it is a second derivative it is nonzero even for zero aberrations. One option is to take this leading term (i.e. \mathbf{K} evaluated at $\psi = 0$) to set requirements. Another would be to also include the linear dependences on ψ ; this would be necessary if we were to separately write requirements on the individual Zernike modes, since due to symmetries some entries in \mathbf{K} are exactly zero in the unaberrated case.

Following the methodology of §B 1 b, and again exploiting the symmetries of the problem and suppressing the z_k index, we find that at $\psi = 0$, the terms involving the covariance of line-of-sight motion and wavefront jitter are

$$K_{1aj}^{\text{LOS,WFE}} = \begin{pmatrix} 0 & 0 & 0 & K_{\theta c} & 0 & K_{\theta t} & 0 & 0 \\ 0 & 0 & 0 & 0 & -K_{\theta c} & 0 & K_{\theta t} & 0 \end{pmatrix} \quad (\text{B14})$$

and

$$K_{2aj}^{\text{LOS,WFE}} = \begin{pmatrix} 0 & 0 & 0 & 0 & K_{\theta c} & 0 & K_{\theta t} & 0 \\ 0 & 0 & 0 & K_{\theta c} & 0 & -K_{\theta t} & 0 & 0 \end{pmatrix}, \quad (\text{B15})$$

where the two rows are $a = 1, 2$ and the eight columns are the low-order Zernikes. Similarly, for the wavefront jitter variance, we have

$$K_{1jj'}^{\text{WFE,WFE}} = \begin{pmatrix} 0 & K_{fa} & 0 & 0 & 0 & 0 & 0 & 0 \\ K_{fa} & 0 & 0 & 0 & 0 & 0 & 0 & K_{sa} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2K_{cc} & 0 & K_{ct} & 0 & 0 \\ 0 & 0 & 0 & -2K_{cc} & 0 & K_{ct} & 0 & 0 \\ 0 & 0 & 0 & K_{ct} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & K_{ct} & 0 & 0 & 0 \\ 0 & K_{sa} & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{B16})$$

and

$$K_{2jj'}^{\text{WFE,WFE}} = \begin{pmatrix} 0 & 0 & K_{fa} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ K_{fa} & 0 & 0 & 0 & 0 & 0 & 0 & K_{sa} \\ 0 & 0 & 0 & 0 & 2K_{cc} & 0 & K_{ct} & 0 \\ 0 & 0 & 0 & 2K_{cc} & 0 & -K_{ct} & 0 & 0 \\ 0 & 0 & 0 & 0 & -K_{ct} & 0 & 0 & 0 \\ 0 & 0 & 0 & K_{ct} & 0 & 0 & 0 & 0 \\ 0 & 0 & K_{sa} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (\text{B17})$$

TABLE IV. The coefficients of spurious shear at $\psi = 0$, appearing in Eqs. (B14–B20). Coefficients are shown for the worst (most contaminated) redshift bin. This redshift bin and the S -factor are shown in the two right-most columns for the J129 band (which is always the most sensitive to wavefront jitter because it has the shortest wavelength). The S -factor is shown for the worst angular bin, which is always the smallest scales ($3.0 < \log_{10} \ell < 3.5$).

Band:	J129	H158	F184	Units	Worst z -bin	Worst S -factor
$K_{\theta c}$	8.11	6.43	4.91	$10^{-6} \text{ mas}^{-1} \text{ nm}^{-1}$	2.8–3.0	0.536
$K_{\theta t}$	3.59	4.14	4.46	$10^{-6} \text{ mas}^{-1} \text{ nm}^{-1}$	2.8–3.0	0.368
K_{fa}	11.30	9.78	8.52	10^{-6} nm^{-2}	2.8–3.0	0.487
K_{cc}	3.68	2.87	2.31	10^{-6} nm^{-2}	2.8–3.0	0.598
K_{ct}	5.86	5.02	4.31	10^{-6} nm^{-2}	2.8–3.0	0.507
K_{sa}	4.31	3.32	2.67	10^{-6} nm^{-2}	2.8–3.0	0.624

Image simulations are required to determine the specific values of $K_{\theta c}$, $K_{\theta t}$, K_{fa} , K_{sa} , K_{cc} , and K_{ct} . These depend on the galaxy sizes, and hence indirectly on redshift slice z_k . The coefficients in the “worst” redshift slice are shown in Table IV.

Once again, the maximum value of the apparent shear induced by wavefront error can be determined from the eigenvalues of the \mathbf{K} matrices, the RMS wavefront jitter, and the line of sight motion per axis. We note that the RMS wavefront jitter is

$$\sigma_{\text{wfe-jitter}} = \sqrt{\sum_j \text{Var}\psi_j}, \quad (\text{B18})$$

and that covariances between WFE jitter and LOS jitter are limited by the rule that the covariance matrix be positive-definite (in particular, the correlation coefficients cannot exceed 1). This implies the limits

$$\left| \sum_{aj} K_{1aj}^{\text{LOS,WFE}}(z_k) \text{Cov}(\theta_a, \psi_j) \right| \leq \sqrt{K_{\theta c}^2 + K_{\theta t}^2} \sigma_{\text{los-jitter}} \sigma_{\text{wfe-jitter}} \quad (\text{B19})$$

and

$$\left| \frac{1}{2} \sum_{jj'} K_{ija}^{\text{WFE,WFE}}(z_k) \text{Cov}(\psi_j, \psi_{j'}) \right| \leq \frac{1}{2} \max \left(\sqrt{K_{fa}^2 + K_{sa}^2}, |K_{cc}| + \sqrt{K_{cc}^2 + K_{ct}^2} \right) \sigma_{\text{wfe-jitter}}^2, \quad (\text{B20})$$

where $\sigma_{\text{los-jitter}}$ is the RMS line-of-sight jitter per axis, i.e. we set $\text{Cov}(\theta_a, \theta_b) = \sigma_{\text{los-jitter}}^2 \delta_{ab}$. (Note that only the jitter contributes here: the controlled motion of the line of sight does not correlate with the wavefront jitter since it is not in the same frequency band.) The sum of Eqs. (B19) and (B20) represents a bound on the RMS spurious shear in the γ_1 component (a similar bound applies to γ_2).

The RMS spurious shear per component in J-band, weighted by the S -factor (which accounts for redshift dependence), is then

$$\gamma_{\text{rms}} S^{1/2} \leq \frac{7.01 \times 10^{-6} \text{ nm}^{-1} \text{ mas}^{-1} \sigma_{\text{los-jitter}} \sigma_{\text{wfe-jitter}} + 4.78 \times 10^{-6} \text{ nm}^{-2} \sigma_{\text{wfe-jitter}}^2}{N_{\text{ind}}^{1/2}}. \quad (\text{B21})$$

This should be compared to the requirement of 2.60×10^{-4} . The line-of-sight jitter is required to be $\sigma_{\text{los-jitter}} = 12 \text{ mas rms}$ per axis, and the observing strategy has two passes at epochs separated by many months so we take $N_{\text{ind}} = 2$. From Eq. (B21), we find that the entire error budget is taken up for $\sigma_{\text{wfe-jitter}} = 3.63 \text{ nm rms}$.