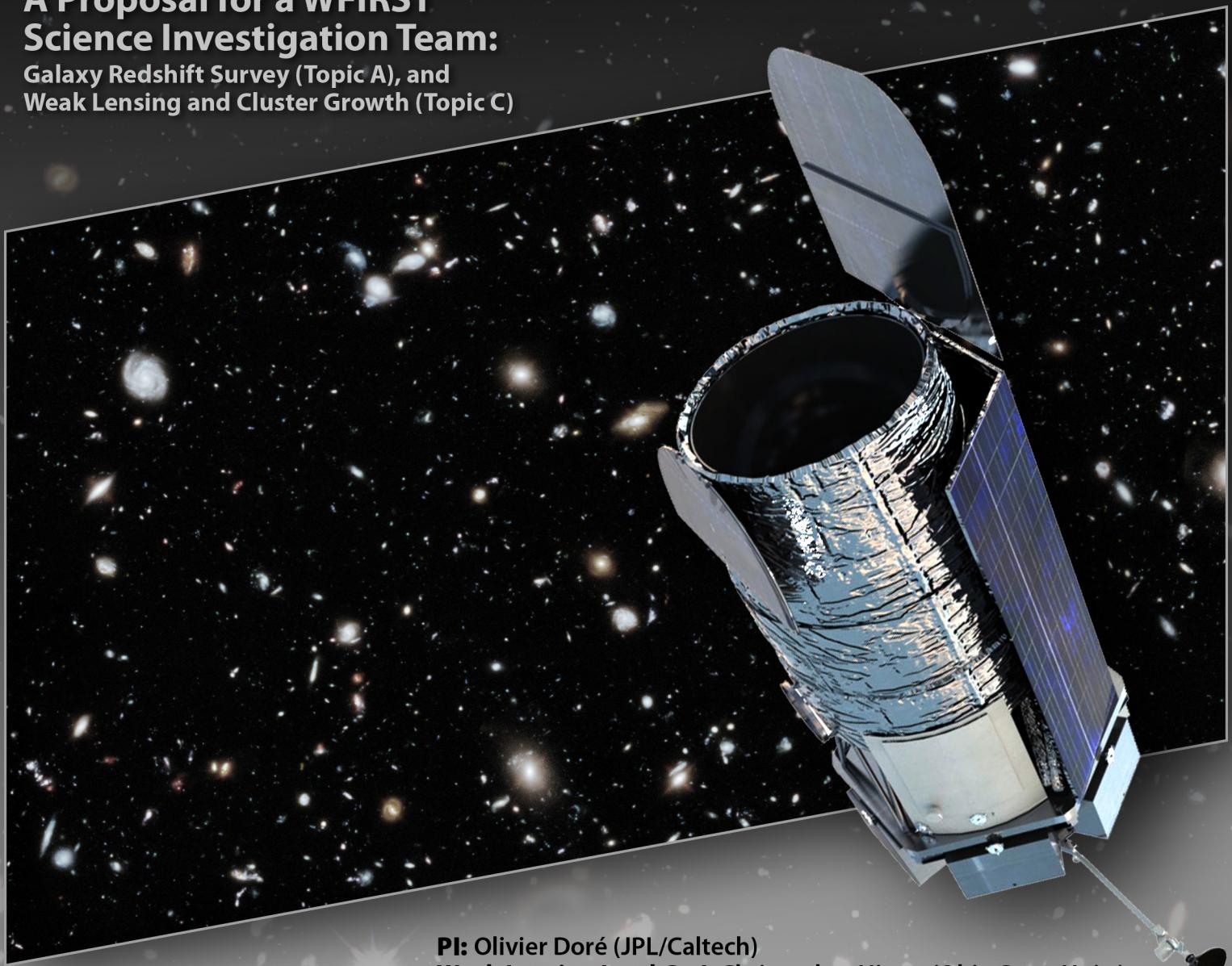


# Cosmology with the WFIRST High Latitude Survey



## A Proposal for a WFIRST Science Investigation Team:

Galaxy Redshift Survey (Topic A), and  
Weak Lensing and Cluster Growth (Topic C)



**PI:** Olivier Doré (JPL/Caltech)  
**Weak Lensing Lead Co-I:** Christopher Hirata (Ohio State Univ.)  
**Galaxy Redshift Survey Lead Co-I:** Yun Wang (IPAC)  
**Cluster Growth Sub-lead Co-I:** David Weinberg (Ohio State Univ.)

**Co-Is:** Rachel Bean (Cornell), Peter Capak (IPAC), Tim Eifler (JPL), Shirley Ho (Carnegie-Mellon Univ.), Bhuvnesh Jain (Univ. of Penn.), Mike Jarvis (Univ. of Penn.), Alina Kiessling (JPL), Robert Lupton (Princeton), Rachel Mandelbaum (Carnegie-Mellon Univ.), Nikhil Padmanabhan (Yale), Lado Samushia (Kansas State Univ.), David Spergel (Princeton), Harry Teplitz (IPAC)

**Collaborators:** Andrew Benson (Carnegie Obs.), Katrin Heitmann (ANL), George Helou (IPAC), Michael Hudson (Univ. of Waterloo), Elisabeth Krause (Stanford), Hironao Miyatake (JPL), Eduardo Rozo (U. of Arizona), Mike Seiffert (JPL), Chaz Shapiro (JPL), Kendrick Smith (Perimeter Insti.), Masahiro Takada (Univ. of Tokyo), Anja von der Linden (Stony Brook Univ.), Naoki Yoshida (Univ. of Tokyo)

## TABLE OF CONTENTS

<b>1 Executive Summary [Oli: Olivier, 2 pages]</b>	<b>1</b>
<b>2 Proposed Deliverables [Oli: Olivier, 4 pages]</b>	<b>2</b>
<b>3 Weak Lensing and Cluster Growth Investigation [Oli: Chris, Rachel M, Mike, David W., 15 pages]</b>	<b>4</b>
3.1 Requirements (D1, D3, D4, D5) . . . . .	4
3.1.1 Polarization effects . . . . .	6
3.1.2 Interpixel capacitance requirements . . . . .	6
3.2 Photometric Calibration . . . . .	6
3.2.1 Dark filter . . . . .	8
3.2.2 Calibration plan . . . . .	12
3.2.3 Detector characterization . . . . .	14
3.3 Photometric Redshifts . . . . .	16
3.4 Cosmological Forecasting, Simulations, and Methodology Development (D2, D8, D9) . . . . .	17
3.5 Systematics Testing and Mitigation (D8) . . . . .	19
<b>4 Galaxy Redshift Survey Investigation [Oli: Yun, Lado, Shirley et al., 15 pages]</b>	<b>20</b>
4.1 Requirements . . . . .	20
4.2 Simulations . . . . .	23
4.3 Cosmological Forecasting, Modeling, and Simulations (D2, D8, D9) . . . . .	23
4.4 Systematics Testing and Mitigation (D8) . . . . .	25
<b>5 Cosmological Forecasts [Oli: Tim, Elisabeth, Olivier, 10 pages]</b>	<b>26</b>
<b>6 Operations Model for the HLS and Evaluation of Trades</b>	<b>26</b>
6.1 Snapshot of the HLS observing plan . . . . .	26
6.2 Further optimizations . . . . .	27
<b>7 Community Engagement and External Data-sets (D11, D12) [Oli: Olivier, 5 pages]</b>	<b>27</b>
<b>8 Future Workplan [Oli: Olivier, all, 5 pages]</b>	<b>32</b>
<b>9 List of Acronyms and Abbreviations and References</b>	<b>34</b>

# *Cosmology with the WFIRST High Latitude Survey*

## **1 Executive Summary [Oli: Olivier, 2 pages]**

[Oli: To be updated]

Cosmic acceleration is the most surprising cosmological discovery in many decades. Even the least exotic explanation of this phenomenon requires an energetically dominant component of the universe with properties never previously seen in nature, pervading otherwise empty space, with an energy density that is many orders of magnitude higher than naive expectations. Testing and distinguishing among possible explanations requires cosmological measurements of extremely high precision that probe the full history of cosmic expansion and structure growth. This program is one of the defining objectives of the Wide-Field Infrared Survey Telescope (WFIRST), as set forth in the *New Worlds, New Horizons* report (NWNH) [19]. The WFIRST-AFTA mission, as described in the Science Definition Team (SDT) reports [105, 104, hereafter SDT13 and SDT15 respectively], has the ability to improve these measurements by 1–2 orders of magnitude compared to the current state of the art, while simultaneously extending their redshift grasp, greatly improving control of systematic effects, and taking a unified approach to multiple probes that provide complementary physical information and cross-checks of cosmological results.

We have assembled a team with the expertise and commitment needed to address the stringent challenges of the WFIRST dark energy (DE) program through the Project’s formulation phase. After careful consideration, we have elected to address investigations A (Galaxy Redshift Survey, GRS) and C (Weak Lensing (WL) and Cluster Growth (CL)) of the WFIRST Science Investigation Team (SIT) NASA Research Announcement (NRA) with a unified team, because the two investigations are tightly linked at both the technical level and the theoretical modeling level. The imaging and spectroscopic elements of the High Latitude Survey (HLS) will be realized as an integrated observing program, and they jointly impose requirements on instrument and telescope performance, operations, and data transfer. The methods for simulating and interpreting weak lensing and galaxy clustering observations largely overlap, and many members of our team have expertise in both areas. The WFIRST supernova cosmology program (investigation B) is more distinct in its methods and requirements, so it is feasible to integrate the supernova and HLS investigations at the level of the Formulation Science Working Group (FSWG).

The team PI, Olivier Doré, is a cosmologist with broad expertise in cosmic microwave background and large scale structure (LSS) studies. He brings extensive experience with complex data analysis (e.g., the Wilkinson Microwave Anisotropy Probe (WMAP), Planck) and mission design (e.g., Joint Dark Energy Mission (JDEM) Destiny and the SMall EXplorer (SME) concept SPHEREx currently under Phase A study, for which he is the Project Scientist). Yun Wang and Chris Hirata will serve as Lead Co-Investigators for topics A and C, respectively and David Weinberg will serve as Lead for sub-topic “Cluster growth” within topic C. Many members of our team have been involved with the design and requirements of a dark energy space mission for a decade or more, including the Co-Chair (Spergel) and four additional members (Hirata, Hudson, Wang, Weinberg) of the 2013-2015 WFIRST-AFTA SDTs. Our team includes authors of the two most comprehensive reviews of observational methods for probing dark energy [116, 120] and the Chair and Vice-Chair (Spergel, Weinberg) of the Astro2010 Science Frontier Panel on Cosmology and Fundamental Physics, whose report played a central role in the NWNH recommendation of WFIRST as the highest priority large space-based program. Our team of Co-Is includes world leading experts on image processing and weak lensing (Eifler, Jain, Jarvis, Kiessling, Lupton, Hirata, Mandelbaum), on design and analysis of galaxy redshift surveys (Ho, Padmanabhan, Samushia,

Wang, Weinberg), on space-based slitless spectroscopy analogous to that planned for WFIRST (Teplitz), on photometric calibration (Padmanabhan), on photometric redshifts (Capak) from large imaging surveys, and on cosmological forecasting and parameter estimation from combinations of cosmic microwave background (CMB), WL, and LSS data (Bean, Doré, Eifler, Hirata, Ho, Jain, Mandelbaum, Samushia, Spergel, Wang, Weinberg).

This team of Co-Is brings close connections to most of the major current or planned cosmological experiments that will provide the context for the WFIRST dark energy program. This includes the WMAP and Planck CMB missions, the Sloan Digital Sky Survey (SDSS), the Baryon Oscillation Spectroscopic Survey (BOSS), the Dark Energy Survey (DES), the Subaru Hyper Suprime-Cam (HSC) and Prime Focus Spectrograph (PFS) projects, the Dark Energy Spectroscopic Instrument (DESI), the Euclid mission and the Large Synoptic Survey Telescope (LSST) Dark Energy Science Collaboration (DESC). Our team of U.S. and international collaborators brings extensive expertise in detector characterization, cosmological simulations, detailed simulations of observational data sets, and the theoretical modeling and cosmological interpretation of weak lensing and galaxy clustering data. Notably, members of our team are responsible for nearly all of the tables and figures in §§ 2.2.3-2.2.5 of the SDT15 report, describing the HLS dark energy program. We therefore have an unparalleled understanding of the current design of WFIRST-AFTA and of the challenges ahead in achieving its science goals.

We have structured our planning and our proposal around the series of deliverables described in §2. Because development of requirements is at the core of our proposed investigation, we present some broad aspects of our strategy in §?? before turning to a more detailed discussion of the WL and GRS program elements in §§3 and 4. We address questions of survey operations and optimization in §6 and our plans for broad community engagement in §7. We conclude with our management plan in §??, re-emphasizing the value of a unified approach to the HLS dark energy science program.

## 2 Proposed Deliverables [Oli: Olivier, 4 pages]

[**Oli:** In this section, we summarize the deliverables we promised and what we actually accomplished]

In §§??,3,4, we describe our work plan and explain how we will iteratively flow down science objectives to the measurements to be conducted, develop observational strategies, simulate synthetic astronomical ‘truth’ data and the observational data output (including calibration), develop a methodology for validating dark energy constraints, and define scientific performance requirements and a complete plan for the science investigation. Our work plan maps the six SIT tasks, identified in parentheses (numbered  $T\ 1\text{--}6$  as in §3.1 of the WFIRST SIT call), to the deliverables below. We also explicitly reference the deliverables (D1-12) in the section titles of our proposal.

- (**D1**) *Full requirements flow-down* from the high-level science goals of the HLS galaxy clustering and weak lensing survey to detailed performance of the telescope, wide field instrument, software, operations, and data transfer. We will evaluate the mission design, as prepared for the Critical Design Review (CDR), against these requirements ( $T_1$ ,  $T_3$ ,  $T_4$ ).

- (**D2**) *Forecasts of the cosmological performance of the HLS Imaging and Spectroscopy data sets*, including expected constraints on dark energy, modified gravity, neutrino masses, and inflation, from analyses that include the measurement of the location of the Baryon Acoustic Oscillations (BAO), Redshift-Space Distortions (RSD), galaxy power spectrum and higher order statistics, cosmic shear, galaxy-galaxy lensing, and cluster demographics. These forecasts will incorporate realistic assessments of observational systematics and theoretical modeling systematics, and they will examine the expected constraints from different probes individually, in concert with each other, and in concert with expected constraints from the WFIRST supernova program,

CMB experiments, and other cosmological surveys such as DESI, LSST, and Euclid. We will use our forecasting tools to investigate trades, e.g., the impact of survey or instrument design choices (area, depth, pixel size, spectral resolution, etc.) on cosmological performance. (*T1, T2*)

- (*D3*) **Simulated imaging and spectroscopic data sets** for testing pipeline performance and evaluating systematic biases — e.g., from confusion, noise, and incompleteness in images and spectra, or errors in Point Spread Function (PSF) determination or shape measurement. These data sets will be created with varying levels of complexity in the source catalogs and instrumental effects, to allow isolation of individual contributions to statistical and systematic uncertainties. Some of these artificial data sets will be made publicly available, and some will take the form of data challenges, where the underlying parameters are initially known only to the creators of the data set, in the spirit of the Shear Testing Program (STEP) and Gravitational Lensing Accuracy Test (GREAT) weak lensing data challenges [33, 67, 13, 52, 62]. (*T3, T5*)

- (*D4*) **Proto-type imaging and spectroscopic pipelines**, including weak lensing shape measurement and galaxy redshift measurement, tested against the above artificial data sets. These proto-type pipelines will provide building blocks for development of full pipelines during the implementation phase, and they will allow us to sharpen definitions of software requirements and to identify challenges to and strategies for meeting these requirements. (*T3, T5*)

- (*D5*) **Calibration strategies** for photometry, shape measurement, spectroscopy, and redshift completeness. Evaluation of the expected performance of these strategies against the science requirements. (*T4*)

- (*D6*) **A strategy for the determination and calibration of photometric redshifts** using WFIRST data and anticipated external data (e.g., LSST optical photometry), and defining ground-based data that are needed to implement this strategy (e.g., spectroscopic training sets, large redshift surveys for calibration via cross-correlation). Evaluation of the impact of remaining photometric redshift uncertainties on statistical and systematic errors in weak lensing and clustering analyses. Definition of requirements for WFIRST photometric redshifts informed by this strategy and evaluation. (*T4*)

- (*D7*) **A detailed operations concept for the HLS Imaging and Spectroscopy program**, extending the work presented in SDT13 and SDT15. (*T2, T6*)

- (*D8*) **Development of methods for modeling and interpreting the cosmological measurements anticipated from WFIRST**. Determination of the effects of non-linear gravitational clustering, realistically complex relations between the galaxy and dark matter distributions, and the influence of the baryon component on matter clustering. The study of techniques to remove systematic biases, e.g., by marginalization over nuisance parameters. Utilization of cosmic shear, galaxy-galaxy lensing, cluster mass functions and cluster weak lensing, BAO, RSD, the galaxy power spectrum, and higher order statistics for galaxy clustering, weak lensing, and various combinations. Identification of areas where further improvements of theoretical modeling would significantly enhance the cosmological return from WFIRST. (*T3*)

- (*D9*) **Simulated light-cone observations** based on cosmological simulations for guiding this methodology development and testing its performance. Most of these data sets will be at the level of galaxy redshift and shape catalogs rather than the pixel-level imaging and spectroscopy simulations described above. They will incorporate varying degrees of complexity regarding galaxy bias, redshift evolution, survey geometry, and observational systematics such as incompleteness, shape measurement errors, and photometric redshift biases. Many of these artificial data sets will be made publicly available, and some will take the form of data challenges, where the underlying parameters are initially known only to the creators of the data set. (*T3, T5*)

- (*D10*) **Pilot survey proposals with associated figures of merits**, to be executed during the first months of WFIRST operations. These would become part of the final dark energy data set but also pin down remaining astrophysical or instrument performance uncertainties at the level needed to optimize the HLS. We will develop the figures of merit required to quickly assess the

data-quality and make operational decisions regarding the cosmological surveys. (*T2, T6*)

- (**D11**) *A prioritized program of observations from other facilities*, ground and space-based, needed to calibrate or finalize strategy decisions on the WFIRST dark energy program. (*T6*)
- (**D12**) *Broad engagement with the cosmological community*, through workshops, talks, publications, and public release of codes and artificial data sets, with the goals of (a) building awareness of and broad support for the WFIRST dark energy program and (b) inspiring the community to develop methods and carry out investigations that will maximize the cosmological return from WFIRST. (*T6*)

### **3 Weak Lensing and Cluster Growth Investigation [Oli: Chris, Rachel M, Mike, David W., 15 pages]**

As discussed in §2.2.3 of SDT15 (written by members of our team), the HLS Imaging survey will (in its current design) measure the shapes of nearly 400 million galaxies in 3 near-infrared (NIR) bands, plus fluxes in a 4th band to improve photometric redshifts (photo- $z$ ). With a data set two orders-of-magnitude larger than the current state of the art [34, 7], the WFIRST weak lensing program will measure the cosmic expansion history and the growth of structure with exquisite statistical precision, demanding corresponding advances in the control of WL systematics. The cosmic shear power spectrum, which is the basic WL observable, depends on both the distance-redshift relation  $D(z)$  and the power spectrum of matter clustering  $(\Omega_m h^2)^2 P_m(k, z)$ . The WL survey will also enable high-precision cosmological constraints from galaxy-galaxy lensing (GGL) and from galaxy clusters, which can be identified in either the HLS or external data sets and characterized with the help of WFIRST WL. The `CosmoLike` forecasting tool can predict the constraints from these methods individually and in combination with complementary probes such as BAO, RSD, supernovae, and the CMB. While all aspects of our investigation are interconnected, we separate the discussion into more tightly connected loops: requirements, image simulations, and algorithm development at level of pixels and shape measurement; issues related to photometric calibration and photometric redshifts; development of methods for cosmological analysis of WFIRST data and testing them on simulated data; and finally methods for systematic error testing and mitigation.

#### **3.1 Requirements (D1, D3, D4, D5)**

[**Authors:** Chris, Rachel M, Mike]

The definition of WFIRST requirements will start in pre-Phase A and continue through CDR. In the early stages, we will focus on identifying the driving requirements for the mission (e.g., those related to total throughput, wavefront error and wavefront stability, and any calibration requirements that would require dedicated hardware or special operational modes). Simple simulation tools will be used for this stage, where fast turnaround times and conservative assumptions will be prioritized over the true end-to-end simulations that will be required for the analysis. As the project matures, we will consider the more complete list of requirements (e.g., specifications on data products) and work with the WFIRST Science Centers (WSC) to build the fully realistic simulations needed to test reduction and analysis pipelines.

Our existing tools (the ETC, operations simulations, and `CosmoLike`) allow us to forecast the statistical power of the WL survey for a specified observing strategy and time allocation and to evaluate the impact of hardware or strategy changes on cosmological constraints (see §6.2). The challenge for the SIT is to define requirements that ensure control of systematic uncertainties at the level of these statistical errors. WL systematics fall into two broad categories, instrumental/algorithmic effects tied directly to the measurements and astrophysical effects tied to the cosmological interpretation of the measurements. The former affect engineering requirements, while

the latter must be predicted theoretically or constrained through observations.

Our basic strategy for defining instrumental/algorithmic requirements is to create simulated HLS images with varying levels of realism, analyze them with proto-type data reduction and WL pipelines, and compare the resulting measurements to the simulation inputs. We will determine the sensitivity of WL shear measurements to each effect, and then flow down requirements from nuisance parameters in `CosmoLike` (e.g., spurious shear) to hardware requirements that can be compared to integrated modeling results (e.g., wavefront stability). Our simulation framework will build on the public code `GalSim` [89], which already includes a WFIRST-AFTA module and to which Co-Is Jarvis and Mandelbaum are lead contributors. Our proto-type pipelines will build on codes for image analysis and shear measurements developed over many years by Co-Is Hirata, Jarvis, Lupton, and Mandelbaum and their collaborators for analysis of SDSS, DES, HSC, and LSST imaging [57, 37, 61, 60, 43, 44]. WL algorithms are an active area of research, and we will integrate promising new approaches as our investigation progresses. Our overall framework is analogous to, but more focused than, the GREAT3 challenge led by Mandelbaum [63], including the use of different simulation branches where effects are turned on and off (separately or together) to probe the biases resulting from individual or combined effects.

In the domain of instrument- and pipeline-related systematic errors, each error must be evaluated according to the following criteria: (i) What is the raw magnitude of the systematic error compared to statistical errors? (ii) What is the approach to modeling and removing that systematic error? What cross-checks will be necessary to validate that this has been done correctly? (iii) Are there implications to the optimal observing strategy (e.g., dithering, repeat observations, dedicated calibration observations)? (iv) Are there implications for the hardware requirements (e.g., stability, pre-flight characterization, or dedicated flight calibration hardware)?

As an example: optical aberrations induce PSF ellipticity that is 2 orders of magnitude greater than WFIRST weak lensing requirements if uncorrected. It is not practical to eliminate the effect in hardware by requiring the wavefront to be perfect at the few nm level in a wide-field instrument, so a model of the PSF will have to be built. Some contributions to the PSF (e.g., jitter and guiding errors) will need to be measured separately for each exposure. Others vary on longer time scales (e.g., due to thermal changes in the optics) and imply a trade-off between stability requirements and the noise on a measurement of the relevant parameter from a set of  $N$  exposures. We will use these considerations to turn high-level requirements on knowledge of the PSF into lower-level requirements on thermal stability and on the operations plan.

Detector systematics are special because their absolute amplitudes may be difficult for the vendor to control or even test. We thus anticipate absolute requirements only on the largest systematic effects such as inter-pixel capacitance (IPC) and persistence, which are being addressed as part of the technology development program. There will be many subtler effects arising in the detectors (examples might include NIR-detector analogues of the brighter-fatter effect, color-dependent charge diffusion, etc.); we will set requirements on the knowledge of these. We will pay special attention to methods of measuring these effects on-orbit using a combination of ground characterization, dedicated calibration modes, and the survey data itself, and the interaction with survey operations and stability requirements (on e.g., focal plane temperature). We will coordinate the best approach to these effects with the Project and with collaborators Seiffert and Shapiro. Shapiro leads the Precision Projector Laboratory (PPL), a JPL facility designed to emulate WFIRST weak lensing data using scenes focused onto WFIRST NIR-detectors. Systematics identified thusly will be studied and mitigated in collaboration with the SIT. Co-I Hirata advises the PPL on image analysis software and interpretation of results [88, 99].

Cluster cosmology is generally considered to be less demanding in terms of hardware requirements than cosmic shear, since the large galaxy over-densities and shear signals are not as easily masked by subtle optical aberrations or detector behaviors. Nevertheless, it may place new requirements on survey footprint/operations (to ensure overlap with other data sets); pipeline behavior in

crowded fields (e.g., [102]); and ancillary data products and simulations to describe, e.g., changes in selection effects and source redshift distributions in the presence of blending and magnification. WL requirements lead Hirata will work closely with cluster lead Weinberg to ensure that these requirements are captured in the flow-down.

**Below is some new text from Rachel M. It needs to be put in the right place. If more information or detail is needed, please let me know.**

### 3.1.1 Polarization effects

During early 2017, work was carried out to assess the approximate level of an effect that could cause weak lensing systematics, but that had never been previously considered by the weak lensing community. This effect is polarization-dependent quantum efficiency (due to e.g. different reflectivity of various coatings for different polarizations of light). Since the light from edge-on disk galaxies typically has some low level polarization perpendicular to the disk, any polarization-dependence of the QE could result in a preferential selection of such galaxies based on their orientation in the focal plane. This would violate the baseline assumption in a weak lensing analysis, which is that all coherent galaxy alignments are due to gravitational lensing.

A student at CMU, Brent Tan, worked with Rachel Mandelbaum and Chris Hirata on a simple toy model for this effect. The toy model had two parameters: the fraction of the disk galaxy light that is polarized, and the relative attenuation of that perpendicular polarization component (both numbers in the range  $[0, 1]$ ). For each point in that parameter space, the coherent shear due to selection bias was calculated; see results in Figure 1. Finally, the results were modified to account for the fact that not all disk galaxies are viewed edge-on and that not all galaxies are disks, giving a net coherent shear due to this selection bias of  $\sim 3 \times 10^{-4}$ . The results are still quite uncertain because our fiducial values for the disk polarization fraction were based on observations of nearby galaxies, not  $z \sim 1$  disks. However, this is large enough to be relevant for WFIRST, so this systematic needs to be evaluated more carefully and requirements placed in future. A publication on this topic will be prepared during summer 2017.

Another possible polarization-related systematic is a polarization-dependent PSF. That will be the subject of future work.

### 3.1.2 Interpixel capacitance requirements

The WFIRST detectors will suffer from electrical crosstalk between the pixels, unlike the optical detectors that are based on CCDs. This effect, known as the *interpixel capacitance* (IPC), appears as a systematic effect in the weak lensing shear measurements and causes a bias in the measurements if not properly taken into account. The effect of IPC on the point-spread function (PSF) was already studied by members of our SIT in [47], and requirements were placed on the level of uncertainty in the IPC based on how that uncertainty affects the PSF.

More recently, in late 2016, members of our SIT (Mandelbaum and student Kannawadi) carried out and analyzed simulations to determine whether additional requirements on IPC are needed to ensure that weak lensing shear estimation is not biased beyond our tolerances. To calibrate the shear multiplicative bias to an accuracy of  $2 \times 10^{-3}$ , we find that the requirements on the IPC placed by the PSF requirements are sufficient, so no new requirement is needed. A paper on this result is in preparation.

## 3.2 Photometric Calibration

**[Authors: Nikhil, Chris]**

Our SIT has been instrumental in the Calibration Working Group, since precision cosmology measurements depend sensitively on calibration; subtle effects that might not be noticeable in other

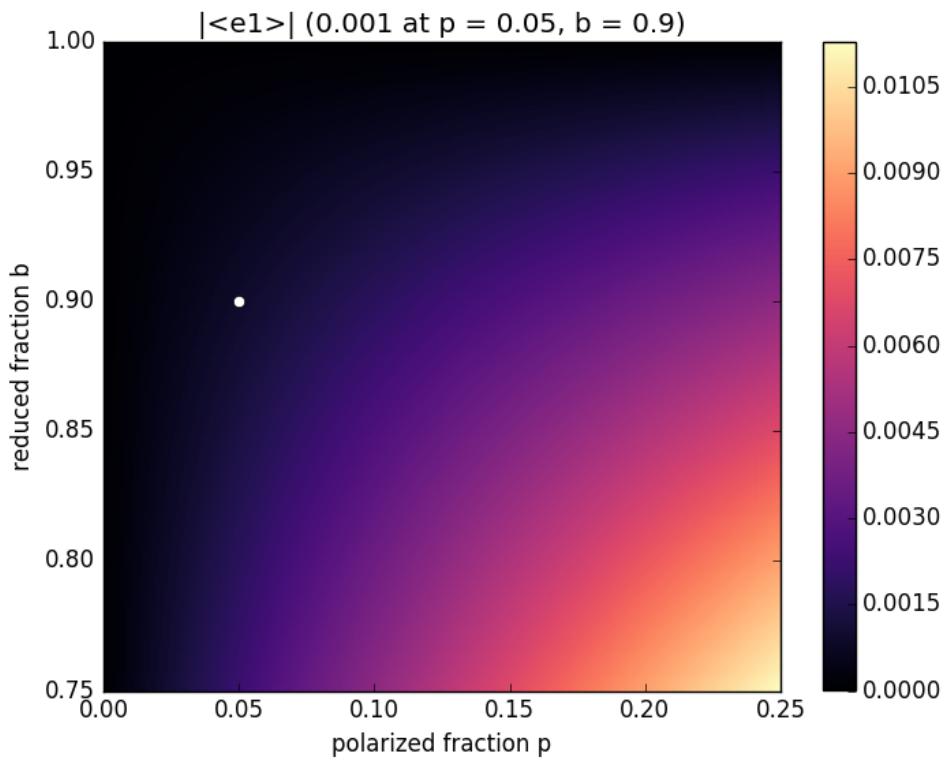


Figure 1: The average shear due to weak lensing selection biases due to polarization-dependent quantum efficiency, as a function of the fraction of polarized light from edge-on disks (horizontal axis) and the fractional attenuation of the perpendicular polarization (vertical axis). The dot at  $(0.05, 0.9)$  is our fiducial point in parameter space, and has  $\langle e \rangle \approx 0.001$ .

areas of astrophysics can become important when trying to measure galaxy shapes to  $< 0.1\%$ . Activities over the past year have included:

- *Dark filter*: Co-I's Wang, Capak, and Hirata participated extensively in the analyses and discussions that led the FSWG to recommend a dark position in the element wheel on WFIRST.
- *Calibration plan*: Our SIT has contributed extensively to the WFIRST WFI Calibration Plan, including detailed quantitative assessments of calibration approaches and their ability to meet requirements. In some areas, such as dark current and the point spread function, our contributions to the calibration plan are now traceable all the way from science measurements (WL shear) down to the specific calibration approaches and the hardware stability requirements needed for them to work. A major area of work leading up to SRR/MDR is to complete this flow-down for the other areas of calibration.
- *Detector characterization*: We have made use of the H4RG data provided by the Detector Characterization Laboratory to measure some of the non-linear effects relevant to weak lensing in real H4RG detectors. This is an important practice step toward building calibration pipelines that will support WL science.

In what follows, we provide some highlights from our calibration activities. The list is not exhaustive.

### 3.2.1 Dark filter

In the summer and fall of 2016, the FSWG was tasked with determining whether a dark filter was needed for WFIRST calibration. This required the FSWG to enumerate the list of calibration tests that might use the dark filter, and establish whether alternative options were possible. We led the effort to assemble this list of tests based on input from the SITs (both ours and others), the SOC, and Project personnel. The list<sup>1</sup> included 14 items: (i) the dark current (including internal instrument backgrounds); (ii) unstable pixels; (iii) post-reset transients; (iv) read noise correlations; (v) inter-pixel capacitance; (vi) gain measurement; (vii) the high spatial frequency flat; (viii) the low spatial frequency flat; (ix) persistence from previous observations; (x) persistence from slews; (xi) classical linearity; (xii) count rate dependent non-linearity; (xiii) the brighter-fatter effect; and (xiv) persistence re-activation.

The problem of persistence from slews (i.e. streaks across the detector following a slew from one observation to another) is of particular importance to weak lensing, because it leads to a coherent, highly directional pattern on the detector that has the correct symmetry to induce a coherent systematic error in the galaxy ellipticities. This is a concern without a dark capability, or even with a dark capability if it is not (or cannot be) used during every slew. Our group identified two budgets in WL that flow down into slew persistence requirements. First is the total systematic shear error budget of  $2.7 \times 10^{-4}$ . Second is the masked pixel budget.

The details of the slew persistence study are provided in the Calibration Plan. It consisted of several stages: first, assessing the magnitude distribution of the stars that would be encountered in the High Latitude Survey; then assessing the probability of stimulus levels in a slew, given the distribution of slews from our operations model (§6); and then folding this through a persistence model (based on DCL data for the development H4RG detectors) to predict the probability distribution of persistent pixels in the HLS imaging survey. The stimulus distribution ( $x$  in e: the well depth to which a pixel is filled during a slew) from the Calibration Plan is shown in Figure 2, and the persistence signal distribution ( $y$  in e: the persistence signal in a pixel over the course of an exposure) is shown in Figure 3.

---

<sup>1</sup>DarkAlternativesMatrix\_161030.docx

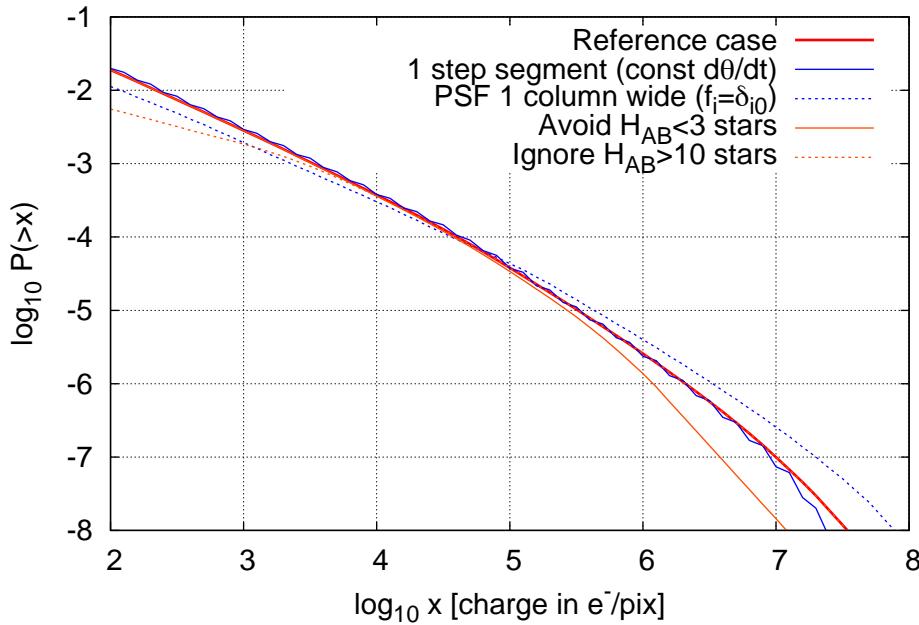


Figure 2: Comparison of stimulus levels predicted under different assumptions and approximations. The vertical axis shows the log probability to exceed a given stimulus level during a slew of 0.4 degrees (a step along the short axis of the field, executed frequently during the HLS). The thick red line indicates reference assumptions. The solid blue line treats the slews as being at constant  $\dot{\theta}$ . The dashed blue line approximates the PSF as 1 column wide (all the flux from the star is concentrated in the central column). The orange lines show what happens if bright ( $H_{AB} < 3$ ) or faint ( $H_{AB} > 10$ ) stars are excluded from the model.

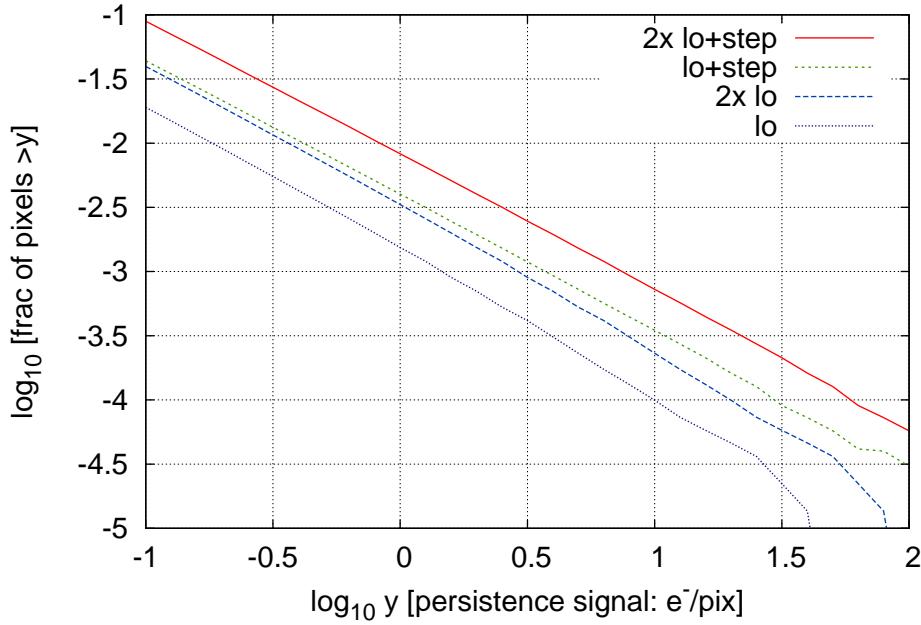


Figure 3: The cumulative distribution function of slew-induced persistence in the HLS imaging survey,  $P(>y)$ . The persistence signal  $y$  is estimated in electrons per pixel; the decaying persistence curve is integrated over a 160 s. Several persistence models are shown, including the “lo” case (typical of the development detectors), and a “lo+step” case (including an order of magnitude step at saturation, as seen in portions of some detectors). This figure has not been updated yet to go from the 6-year to the 5-year observing plan, although we expect only minor differences.

After negotiating with the Project, we settled on a mitigation strategy for slew persistence that involved saving the spacecraft orientation information from the Attitude Control System (ACS), using this to predict the locations of persistence from bright star streaks, and masking  $\pm 2\sigma$  on either side of these streaks. Unmasked streaks are simply accepted as part of the systematic error budget. Their impact on shape measurement is based on an analytic result derived by our SIT and tested against Monte Carlo simulations:

$$\Delta\gamma_1 + i\Delta\gamma_2 = \frac{M\Omega_{\max}\sigma_n^2 R^4}{2F^2 N_{\text{ind}} \text{Res}} f_{\text{scale}} f_{\text{aniso}}, \quad (1)$$

where  $\Delta\gamma_{1,2}$  are the two components of spurious shear;  $M$  is a margin factor;  $\Omega_{\max} = 421.3$ ;  $\sigma_n^2$  is the variance of the persistence image;  $R$  is the radius of the galaxy in pixels;  $F$  is the signal from the galaxy in electrons per exposure;  $N_{\text{ind}}$  is the number of *independent* exposures of the galaxy<sup>2</sup>; Res is the galaxy resolution factor [?];  $f_{\text{scale}}$  and  $f_{\text{aniso}}$  are factors  $\leq 1$  describing the scale dependence and anisotropy of the persistence power spectrum (defined to be 1 in the worst case).

The results of this study – shown in Figure 4 – are promising, given the top-level systematic shear budget of  $2.7 \times 10^{-4}$  and that the modern detectors typically show “lo” or (in some regions) “lo+step”-like behavior, rather than the much larger persistence characteristic of the WFC3-IR model (third column). The masking algorithm will continue to be revisited as part of the mission optimization. However, the small number of masked pixels led the FSWG to conclude that a dark shutter that operated during every slew was not required for the WFIRST HLS.

We carried out a related study, also using Eq. (1) and related machinery, to assess how well we need to know the dark current for WFIRST. Dark current measurements without a dark filter are possible, e.g. via median algorithms that combine many exposures from a survey, but are subject to: (i) a degeneracy in which the “true” sky brightness is unknown and hence the zero level of the dark current cannot be established, and (ii) possible correlated errors from imprinted celestial sources. The requirements, as derived in the appendix to the calibration plan, are:

- The error in the dark current + bias determination in a 140 s HLS imaging exposure shall be no more than  $0.0096 f_{\text{corr}}^{-1/2}$  e/p/s (uncorrelated part) or  $0.0017 f_{\text{corr}}^{-1/2}$  e/p/s (imprinted celestial sources).
- The error in the dark current + bias determination in a 297 s HLS spectroscopy exposure shall be no more than  $0.0059 f_{\text{corr}}^{-1/2}$  e/p/s (uncorrelated part) or  $0.00072 f_{\text{corr}}^{-1/2}$  e/p/s (imprinted celestial sources).

Here “ $f_{\text{corr}}$ ” denotes the factor by which we plan to correct biases induced by errors in the dark current map (we normally choose  $f_{\text{corr}} = 1$  to be conservative). The requirements are traceable to additive shear biases from non-circular imprinted celestial sources; multiplicative shear biases as the noise in the dark current map results in e.g. galaxy centroids getting “pulled” toward pixels whose measured dark current fluctuates below the true dark current of that pixel; and Eddington-like biases for sources detected in the GRS. While the semi-analytic estimates in the calibration plan based on source counts suggest that the HLS imaging requirement can be met without a dark filter, our SIT and the Calibration Working Group had concerns about possible degeneracies in the self-calibration procedure that can only be addressed by a detailed simulation. Moreover, the approach requires empty space in the images, which we will not have in the case of grism spectroscopy. As the imaging exposures are shorter than the spectroscopy exposures, this would require dedicated long imaging exposures (of HLS spectroscopy exposure length) just for the purpose of self-calibrating

---

<sup>2</sup>This may be less than the total number of exposures of the galaxy, since slew persistence from successive exposures will be correlated.

# Masked pixels & systematic shear results for ACS-based flagging

Example parameters:

- Mask trails of  $H_{AB} < 9$  stars,  $\pm 2\sigma$  on either side of track, if expected persistence is  $> 8$  e.
- For fainter stars, assume masking of science data for  $> 64$  e outliers ( $4.3\sigma$ ).

	H4RG-lo	H4RG-lo + step		WFC3-IR		
ACS error (arcsec rms per axis)	Masked pixel fraction	Systematic shear per component	Masked pixel fraction	Systematic shear per component	Masked pixel fraction	Systematic shear per component
1.0	0.16%	3.4E-5	0.68%	8.6E-5	2.85%	6.6E-4
2.0	0.32%	3.4E-5	1.31%	8.6E-5	5.50%	6.6E-4
4.0	0.62%	3.4E-5	2.58%	8.6E-5	10.78%	6.6E-4

Note that further tuning of the parameters may yield some improvement, or enable different trades between masked pixels and systematics in the unmasked data.

Figure 4: The outcome of the October 2016 slew persistence study. This shows the masked pixel fraction and the predicted systematic shear due to unmasked streaks as a function of both the persistence model and the accuracy of pointing information.

the dark. Due to sky Poisson noise, we would need many of these images – our February 2017 estimate was for  $N = 73$  exposures, which, if done every week, would consume 4% of the wall clock time. In light of these and other issues, the Calibration Working Group recommended that WFIRST maintain the dark filter.

### 3.2.2 Calibration plan

Our SIT has contributed extensively to the WFIRST WFI Calibration Plan. This includes extensive quantitative analysis of proposed calibration techniques, as detailed in the appendix to the plan. Some highlights follow.

The requirement on knowledge of the dark current and the calibration approaches are fully defined, based on analysis done during the dark filter trade (October 2016 – February 2017).

Weak lensing was found to place demanding requirements on measurement of the count rate-dependent non-linearity (CRNL). The weak lensing program is sensitive to CRNL because it enhances the bright center of a PSF star relative to its wings, thereby making the star appear slightly smaller, but does not have a similar effect on the faint galaxies used for shape measurement. The PSF second moment is biased by a factor of  $1 - \alpha$  (where  $\alpha$  is the CRNL exponent), and has a top-line systematic error budget of  $7.2 \times 10^{-4}$ . This means that if  $\alpha$  is measured to  $\pm 3 \times 10^{-4}$  (the requirement from the supernova SITs), then CRNL consumes 17% of the PSF size error budget, in an RSS sense. Given that CRNL is a pernicious bias for two of the dark energy probes, we recommended a multi-faceted approach to CRNL calibration, including a lamp-on/lamp-off capability for WFIRST (this was not available on WFC3-IR).

Our team has revisited the wavefront stability requirements for weak lensing, using a set of

codes and scripts on the team’s GitHub site. This begins with a Fisher matrix analysis of the uncertainties in the shear power spectrum, and our top-line requirement that the systematic errors be equivalent to the statistical errors even if the survey is extended to 10,000 deg<sup>2</sup> (i.e. in an RSS sense, the systematic errors should be 20% of the statistical errors in the nominal 2,000 deg<sup>2</sup> survey). Requirements are assessed using the significance, defined by

$$Z = \sqrt{\Delta\mathbf{C} \cdot \boldsymbol{\Sigma}^{-1} \Delta\mathbf{C}}, \quad (2)$$

which is the number of sigmas at which one could distinguish the correct power spectrum from the power spectrum containing a systematic error. We built sub-allocations for multiplicative (shear calibration) errors, and for additive (spurious shear) errors in each angular bin. An early discovery was that this process depends on the redshift dependence of the shear error: some redshift dependences are “worse” than others by the  $Z$ -metric. The worst possibility is *not* for the error to be redshift-independent, but rather for it to change sign, as this can mimic a change in redshift evolution of the growth of structure.

In our current formalism, for each angular template, we introduce a limiting amplitude  $A_0^{\text{flat}}(\alpha)$ , defined to be the RMS spurious shear per component  $A_0$  at which we would saturate the requirement on  $Z(\alpha)$  for angular bin  $\alpha$  in the case of a redshift-independent systematic  $w_i = 1 \forall i$  (here  $\alpha$  denotes an angular bin and  $i$  a redshift bin). That is, if the additive systematics did not depend on redshift, we could tolerate a total additive systematic shear of  $A_0^{\text{flat}}$  (RMS per component) in band  $\alpha$ . We also introduce a scaling factor  $S[\mathbf{w}, \alpha]$  for a systematic error

$$S[\mathbf{w}, \alpha] = \frac{Z(\alpha) \text{ for this } w_i}{Z(\alpha) \text{ for all } w_i = 1} \quad (3)$$

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

$$\sum_{\text{systematics}} [A(\alpha)]^2 \times S[\mathbf{w}, \alpha] \leq [A_0^{\text{flat}}(\alpha)]^2, \quad (4)$$

where  $A(\alpha)$  is the RMS additive shear per component due to that systematic. We 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. In most cases, however, something is known about the redshift dependence of the systematic error (e.g. for PSF errors the error scales with the size of the galaxy, and hence has a redshift dependence tied to the measured redshift evolution of galaxy sizes). In these cases, we use the correct redshift weighting factor  $S$ . This approach has been critical in order to set stability requirements that are consistent with the Project’s integrated modeling results.

We have begun incorporating the HLS observing strategy (§6) in studies of self-calibration of time-dependent drifts in the response of the system (i.e. time dependence of the conversion from  $\mu\text{Jy}$  on the sky to DN/s in the digitized detector system outputs). This model is in a state of flux as we add parameters to it, but here we show a current snapshot allowing for time-dependent drifts of the response of each of the 18 SCAs making up the focal plane, with time dependence parameterized in calibration periods of  $\Delta t$  (assessed down to a period of 3 hours) each. Both individual-SCA drifts and common-mode drifts are allowed, with an assumed intrinsic variation (calibration prior) of 1%

RMS drift in each  $e$ -fold of timescales. A network of randomly distributed stars with a density of 500 stars/deg<sup>2</sup> and  $S/N = 50$  was assumed; in self-calibration, the magnitudes of these stars are *not* known a priori, but are assumed to be stable across multiple repeated observations of the same field. These are preliminary parameters being used to test our tools and are not currently held as requirements. The stellar density model is very conservative since the Trilegal model predicts star counts of 572, 803, 990, and 1137 stars/deg<sup>2</sup> at  $H_{AB} = 18 - 19$ ,  $19 - 20$ ,  $20 - 21$ , and  $21 - 22$  at the SGP, and even an  $H_{AB} = 22$  star will have  $S/N > 50$ . The temporal stability of the system needs further study and will be varied as an input parameter in future versions of this model. The current model uses the April 19, 2017 update to the HLS observing strategy. The number of calibration parameters varies depending on the filter, since there are no parameters for periods of time when the instrument is not observing in that filter; the current version has 17262 parameters for the H band.

Despite the intrinsic stability assumed, in which each SCA can have its response fluctuate by 1.67% RMS from one time interval to the next, the repeated observations do an excellent job of tracking these changes and reducing the posterior uncertainty. Even for  $\Delta t = 0.125$  days = 3 hours, the posterior calibration errors are at the level of 0.14–0.17% RMS (here “RMS” is weighted by number of observations), depending on the filter. An example of the model output (predicted uncertainties in the calibration parameters for each SCA at each epoch) is shown in Figure 5. It must be remembered that this analysis is overly simplistic in some ways – particularly that we have not yet allowed for shorter-timescale variations (i.e. on timescales  $< \Delta t$ ), nor have we allowed for separate gain drifts among the different readout channels. These will have to be included in a future version of the model. On the other hand, the stellar density and  $S/N$  assumptions were extremely conservative (e.g. the full range of stellar magnitudes 18–22 should have 7 times more stars than were assumed, even at the Galactic pole), so there is margin to absorb these additional degrees of freedom. The next iteration of the model for time-dependent calibration drifts will include additional parameters, as well as updated priors reflecting expected detector system stability rather than the place-holder requirements shown here.

### 3.2.3 Detector characterization

The WFIRST dark energy analyses will place enormous demands on our understanding of the detectors. Some aspects of this problem can be anticipated in advance – for example, we know that effects such as inter-pixel capacitance, count-rate-dependent non-linearity, etc. will need to be carefully characterized, and we are working as part of the Calibration Working Group to build these measurements into the mission. However, with systematic error budgets at the level of a few  $\times 10^{-4}$ , it is likely that WFIRST analyses will turn up new effects that were not apparent in past missions. Therefore a key task for our SIT is to analyze the data from development detectors and identify these new effects early enough to inform the calibration plan.

In ground-based weak lensing projects using thick CCDs (e.g. DES), one of the key detector issues has been the *brighter-fatter effect* (BFE). This is an electrostatic effect in which as a pixel fills up with collected charge, it changes the electric field geometry and new charges generated are more likely to be deflected into neighboring pixels. This has the effect of making bright stars appear larger than faint stars, as the repulsion effect is non-linear and increases with signal level. The field geometry is very different in a NIR detector, but a brighter-fatter effect is still possible.

We have searched for the brighter-fatter effect in the H4RG detector arrays using the flat fields for two devices H4RG-17940 and H4RG-18237, provided to us by the DCL. The BFE imprints a signature in the auto-correlation function of a flat field; using the correlations in multiple non-destructive reads in a flat field, one can separate linear IPC from the BFE. Preliminary brighter-fatter effect results for H4RG-17940 are shown in Figure 6. The BFE coefficients are  $a'_{\Delta i, \Delta j}$ , which is the fractional change in effective area of pixel  $(i, j)$  when an electron is placed in pixel

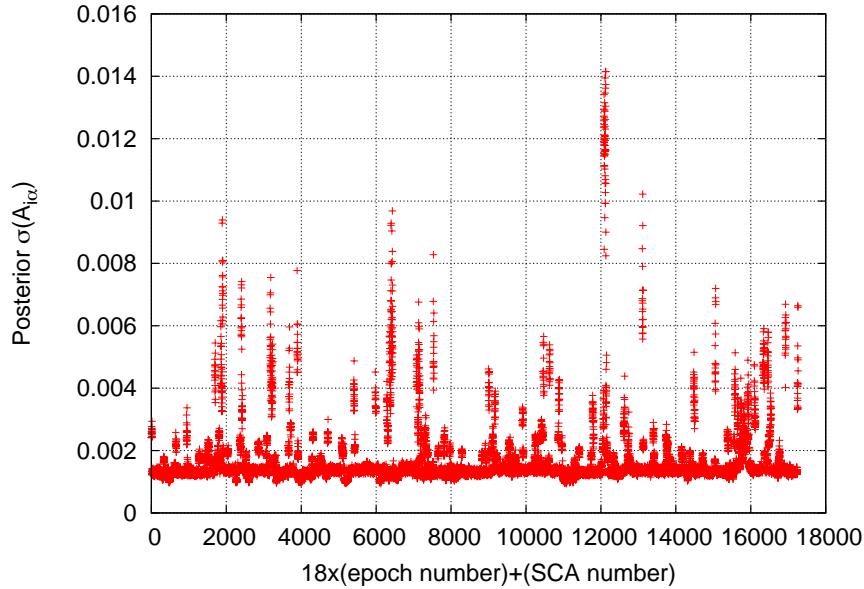


Figure 5: An example of the posterior calibration error from an HLS self-calibration calculation. The horizontal scale displays the time intervals (left-to-right in time order), with 18 points per time interval indicating the various SCAs. The vertical axis shows the standard deviation of the calibration solution for that SCA at that time,  $\sigma(A_{i\alpha})$ , relative to the survey mean. The figure shows the case of H band with  $\Delta t = 0.125$  days. This example had 17262 calibration parameters. A few epochs, mostly containing only a few observations, are poorly constrained due to minimal overlap observations. It is subject to refinement and the input parameters of the model will be varied as we work toward requirements on the stability of the detector system.

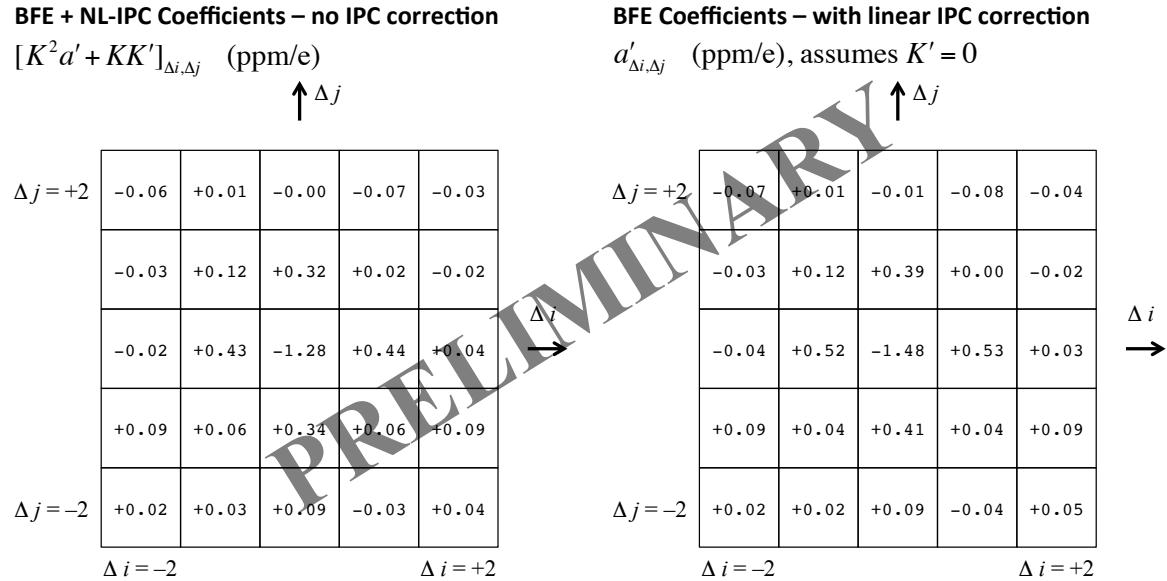


Figure 6: The BFE + NL-IPC coefficients  $[K^2 a' + KK']_{\Delta i, \Delta j}$  (left panel) and IPC-corrected coefficients  $a'_{\Delta i, \Delta j}$  (right panel), for H4RG-17940. Note that the IPC-corrected coefficients assume that the IPC is linear, i.e. the non-overlapping correlations are ascribable entirely to the BFE and not NL-IPC. The  $1\sigma$  uncertainty in each pixel is 0.07 ppm/e.

$(i+\Delta i, j+\Delta j)$ ; they have units of parts per million per electron (ppm/e). The flat auto-correlations are sensitive to both the brighter-fatter effect and non-linear inter-pixel capacitance (NL-IPC); we are currently working on distinguishing the two effects.

### 3.3 Photometric Redshifts

[Authors: Peter, Shooby, Dan]

Accurate photo-zs are crucial to all WFIRST probes of dark energy. We plan to combine calibrations of the color-redshift relation [68] and cross-correlation techniques to achieve the required photo-z performance. These will be supplemented by consistency checks such as the expected level of WL cross-correlations for sources in different photo-z bins. Our team is actively developing the color-redshift mapping for WFIRST (supported by WFIRST Preparatory Support (WPS); PI Capak), with initial tests demonstrating a path to meeting the WFIRST requirements on photo-z bias. Our team has worked with NASA HQ to allow the NASA Strategic Keck Proposals to obtain the required spectra. A key piece of the proposed work will be carrying out these spectroscopic surveys. Cross-correlation methods [42, 75, 70, 76] require spectroscopic overlap and rely on assumptions about the bias of galaxies. We will determine the applicability of these methods and their implications for survey design. We also plan to investigate promising new methods of photo-z inference [12, 45, 8] as additional consistency checks.

Our team will refine the performance estimates and calibration data requirements for each of these photo-z methods. This will include studying the impact on WFIRST operations (e.g., incorporating observations of spectroscopic calibration fields in the observing plan). Co-I Capak will lead the photo-z effort, with support from other team members (Mandelbaum, Hirata, Hudson, Jain, Eifler, Krause, Miyatake) who have extensive experience using photo-z calibration techniques for WL applications with SDSS (e.g., [74, 64]), Canada-France-Hawaii Telescope (CFHT), DES and HSC survey data.

### 3.4 Cosmological Forecasting, Simulations, and Methodology Development (D2, D8, D9)

**Forecasting.** Cosmological forecasting plays an essential role in connecting strategies and requirements defined at the instrument and observation level to WFIRST’s top-level science goals. Co-Is Bean, Hirata, Padmanabhan, Wang and Weinberg have been involved in the forecasting for the WFIRST reports and for cosmological surveys and space missions in general. One of our early activities will be to develop a forecasting framework for WFIRST building on Eifler & Krause’s `CosmoLike`, incorporating all of the dark energy probes from the HLS Imaging, Spectroscopy, and Supernova surveys. We will incorporate the impact of the leading observational and theoretical systematics through “nuisance parameters” that can be marginalized over in cosmological parameter estimates. This framework will be much more thorough than the one used for the SDT reports, and it will enable complete flow-down and flow-up analyses between the instrument, strategy, and software requirements and the expected cosmological performance of WFIRST. We will initially use analytic approximations for error covariance matrices and the dependence of observables on model parameters. We will steadily update these approximations using the simulations described below. For cosmic shear, we will extend existing work on the WL bispectrum (e.g., [48, 29]), which should produce complementary constraints to the power spectrum and may have similar statistical power given the high source density expected in HLS Imaging. Figure 1 illustrates the application of `CosmoLike` to WFIRST multi-probe forecasting.

**Cosmological simulations.** As emphasized in SDT15, WFIRST will improve the statistical precision of cosmic expansion and structure growth measurements by factors of 5 – 50 compared to current data, and will require commensurate improvements in modeling the lensing and clustering signals and the astrophysical contaminants thereof. We will devote significant effort to modeling the nonlinear regime via a combination of numerical simulations, perturbation theory, and empirical measurements, to address issues such as baryonic effects on matter clustering [125, 112]; the validity of the Born approximation [95]; and galaxy intrinsic alignments [39, 55, 53]. We will develop analysis strategies that take advantage of our methodological improvements and marginalize over remaining uncertainties. The computational requirements for cosmological simulations will increase significantly in the years just prior to launch. Where these resources will be sourced from is currently an unsolved issue that is being investigated at the project level. Co-I Kiessling has been part of this effort at the request of the project and she will be responsible for leading the team efforts to quantify computing requirements beyond FY20.

Most of our work will use large-volume  $N$ -body simulations (gravity only) in concert with statistical recipes or semi-analytic models for assigning galaxies to dark matter (sub-)halos. A large set of such simulations ranging in size from tens of billions to a trillion particles already exists. These will be made available to the collaboration by collaborator Heitmann and will be augmented over time with more simulations as needed. We will use hydrodynamic simulations of smaller volumes, provided by collaborator Yoshida, for targeted investigations such as the impact of baryonic physics on the matter power spectrum or as a realistically complicated test of non-linear galaxy bias models. From the  $N$ -body simulations we will produce mock galaxy catalogs both in fixed redshift cubes and in light cones with realistic survey geometry and redshift evolution. We will use these mocks to understand the effect of observational complications (survey geometry, variable depth and completeness, etc.) on WL and clustering measurements and to provide realistically clustered inputs for the pixel-level simulations described earlier.

Our team includes Co-I Kiessling and collaborators (Heitmann, Takada, Yoshida) who are leading similar efforts for Euclid, Subaru HSC and PFS, DES, DESI and LSST. In §?? we describe the computing facilities that will enable these computations. We will publish papers documenting our methodology development and will simultaneously release the associated simulations. This growing library of publicly available simulations will encourage the broader community to develop

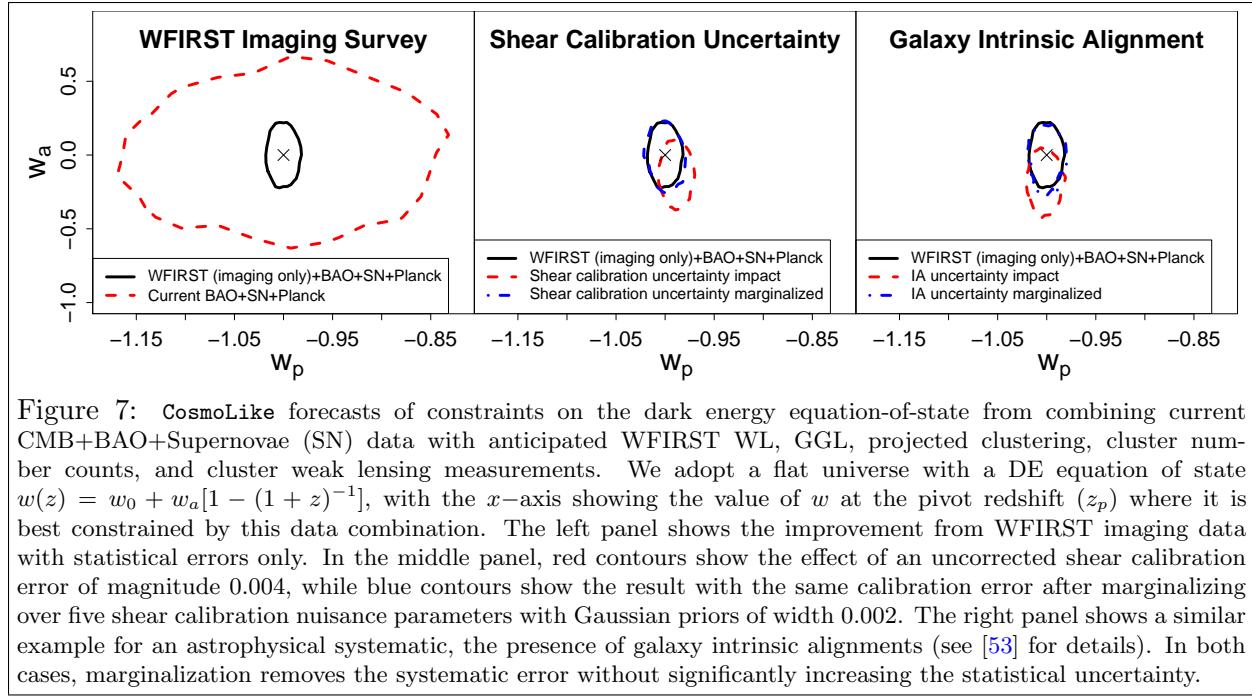


Figure 7: *CosmoLike* forecasts of constraints on the dark energy equation-of-state from combining current CMB+BAO+Supernovae (SN) data with anticipated WFIRST WL, GGL, projected clustering, cluster number counts, and cluster weak lensing measurements. We adopt a flat universe with a DE equation of state  $w(z) = w_0 + w_a[1 - (1 + z)^{-1}]$ , with the  $x$ -axis showing the value of  $w$  at the pivot redshift ( $z_p$ ) where it is best constrained by this data combination. The left panel shows the improvement from WFIRST imaging data with statistical errors only. In the middle panel, red contours show the effect of an uncorrected shear calibration error of magnitude 0.004, while blue contours show the result with the same calibration error after marginalizing over five shear calibration nuisance parameters with Gaussian priors of width 0.002. The right panel shows a similar example for an astrophysical systematic, the presence of galaxy intrinsic alignments (see [53] for details). In both cases, marginalization removes the systematic error without significantly increasing the statistical uncertainty.

analysis strategies that will ultimately be applied to WFIRST (and other surveys).

**Galaxy-galaxy lensing.** The combination of GGL with galaxy clustering is an alternative route to extracting cosmological constraints from an imaging survey. Systematics are significantly different from those affecting cosmic shear analysis, and theoretical studies suggest that the statistical power is comparable [123]. The need to mitigate systematics favors a joint modeling approach to cosmic shear, GGL, and galaxy clustering, which requires devising and testing models of non-linear galaxy clustering and its dependence on redshift, building on studies such as [124, 6, 14, 65, 20, 71, 127] and including the possible impact of “assembly bias” connected to halo formation histories. We will include GGL in our cosmological forecasting framework and identify any GGL-specific requirements distinct from those tied to cosmic shear.

**Clusters.** Our efforts in cluster analysis methodology will parallel those for cosmic shear and GGL, facing many of the same issues but in the (somewhat simpler) high mass halo regime. Co-I Weinberg will lead the cluster effort with support from collaborators Rozo and von der Linden. While the WFIRST data set presents some unique issues, we will draw extensively on the machinery being developed for DES by Rozo, which follows the broad strategy laid out in chapter 6 of [120], and for LSST, described in [56].

The first branch of the cluster effort will focus on the identification and characterization of clusters in WFIRST+LSST (or WFIRST+HSC) imaging. This optical+NIR combination will lead to the best statistical precision, since optical/IR selection can select clusters (at typical redshifts  $z \sim 0.5 - 1.5$ ) down to mass thresholds significantly lower than X-ray or Sunyaev-Zeldovich detection; we expect  $\sim 40,000$  clusters in the HLS with masses  $> 10^{14} M_\odot$ . Building on DES methodology [92], we will design and test (in simulations) prototype cluster finders for WFIRST+LSST data. Among the issues we will address are: completeness and contamination of the cluster catalogs as a function of redshift and richness; biases in estimated richness or cluster redshift; expected scatter between richness and halo mass; accuracy of galaxy photometry in cluster regions; accuracy of shear measurement in cluster regions; reliable separation of foreground, member, and background galaxy populations for weak lensing analysis; and ellipticity- or orientation-dependent cluster selection.

A second branch recognizes WFIRST’s unique potential for calibrating cluster mass-observable

relations at  $z \gtrsim 1$  via weak lensing. Cluster-galaxy lensing (CGL, e.g., [100, 114]) is the method of choice for all cluster surveys overlapping weak-lensing surveys, but requires sufficient number densities of galaxies behind the clusters with accurate photo-zs. For cluster mass calibration at  $z \gtrsim 1$ , space-based shear measurements and photo-zs from the combination of LSST and deep NIR photometry, such as delivered by WFIRST, are therefore necessary. Calibrating cluster mass-observable relations furthermore benefits tremendously from the availability of multi-wavelength mass proxies (e.g. [121]), hence we will consider the synergies with X-ray (eROSITA) and SZ (Planck, AdvACT, SPT-3G, CMB-S4) measurements, with special attention to the impact on survey footprint placement. We will also investigate the potential for HLS-CMB cross-correlations to extract the kinetic Sunyaev-Zeldovich signature to constrain dark energy, gravity and neutrino mass sum [73, 72]. Magnification of special galaxy subsamples can provide a cross-check on the shear-based lensing effort (e.g., [36]).

For modeling methods, we will develop a comprehensive approach that combines CGL and cluster-galaxy cross-correlations to extract cosmological information from fully non-linear, trans-linear, and linear scales, extending and unifying methods based on the cluster mass function (e.g., [90], [66]), cluster mass-to-light or mass-to-number ratios [110, 109], large scale cluster-mass correlations [128]. We will test the robustness of these methods to all of the observational effects listed above. For our cosmological forecasting, we will integrate clusters with our `CosmoLike` treatment of WL and the GRS to account for correlated statistical errors (from large scale structure in the HLS survey volume) and common systematics (shear calibration, photo-z errors).

### 3.5 Systematics Testing and Mitigation (D8)

Achieving WFIRST’s precision cosmology goals requires eliminating systematic biases while minimizing statistical losses, and *demonstrating* that biases have been removed. We will develop a 3-pronged approach to this challenge.

**Marginalization.** As illustrated in Figure 1, a general strategy for both instrumental and astrophysical systematics is to describe their possible impact by nuisance parameters and marginalize over these in cosmological analysis. Important elements in making this strategy effective are: (i) devising concise templates that describe the systematics with minimal numbers of parameters; (ii) setting realistic priors on nuisance parameters; (iii) combining multiple observables that can break degeneracies. We will develop this approach for treating observational effects, particularly shear measurement systematics and photo-z biases, and astrophysical effects, particularly the impact of baryons on the matter power spectrum [46, 91, 111, 97] and galaxy intrinsic alignments (IA). For the observational effects, we will use our simulations (§4.1) and photo-z investigations (§4.2) to design templates and determine appropriate priors. For the astrophysical effects, we will draw on our team’s extensive experience with analytic and numerical modeling of IAs [39, 59, 38, 58, 32, 49, 50, 103, 107] and on the work of Eifler and Krause in incorporating baryonic and IA effects into cosmological WL analysis [55, 51, 27, 53]. This will include the principal components approach, which allows one to marginalize out the directions in observable space that are most sensitive to the choice of prescriptions for baryonic physics in simulations [27]. Cosmic shear, GGL, CGL, and galaxy clustering each respond differently to these systematics, so we anticipate that joint analyses will allow much tighter constraints on both systematics and cosmology than any probe in isolation.

**Systematics maps.** A second approach to identifying and removing observational systematics is to cross-correlate the signal being measured with maps of possible systematic effects, such as stellar density, PSF size, or Galactic extinction (e.g., [87]). These methods both measure the impact of systematics and provide a template for removing them. We will devise a system of such methods for WFIRST WL and galaxy clustering measurements and test their efficacy on our simulated data sets. This effort will be led by Co-Is Ho and Padmanabhan, who developed such approaches for their analyses of large scale galaxy and quasar clustering in the SDSS [40, 79].

**Null tests, internal consistency, and external data sets.** WL analyses must be validated using internal tests that the measurements should pass for any set of cosmological parameters but may fail in the presence of systematics (e.g., [44]). For example, the cross-correlation between PSF-corrected galaxy shapes and star shapes should be consistent with zero, many statistics associated with  $B$ -mode shear (e.g., [26]) should vanish, and consistent cosmological results should be obtained when using the largest or smallest 50% of the source galaxies. Drawing on our team’s experience with other surveys, we will ensure that there is a coherent pipeline for carrying out these standard tests on WFIRST data, including the many consistency tests enabled by having shape measurements in multiple bands. We will pay close attention to tests that make use of unique properties of WFIRST data; for example, comparison of shapes measured on subsets of an exposure with multiple non-destructive reads would test for the impact of detector non-linearities.

Cross-correlations with external imaging and spectroscopic surveys (Kilo Degree Survey (KiDS), HSC, DES, PFS, DESI, LSST, Euclid) offer multiple opportunities for improving the HLS analysis, including photo- $z$  calibration and tests for shear systematics. For example, the cross-correlation of WFIRST and LSST shapes will evade additive systematics that impact only one or the other survey, e.g., those coming from the atmosphere for LSST or from detector effects specific to WFIRST. Other validations can be performed using surveys that measure similar quantities as the HLS but using different techniques, such as cluster masses estimated via CMB lensing. We will investigate a variety of possible tests, evaluate the hardware and operations implications (e.g., footprint overlap, joint data management), and ensure that they are represented in the FSWG process.

## 4 Galaxy Redshift Survey Investigation [Oli: Yun, Lado, Shirley et al., 15 pages]

As discussed extensively in §2.2.4 of SDT15 (written by members of our team), the defining goal of HLS spectroscopy is to derive constraints on dark energy from a slitless spectroscopic (grism) redshift survey of approximately 20 million emission line galaxies (ELG) in the redshift range  $z = 1 - 3$ . The galaxy redshift survey will enable high-precision measurements of the cosmic expansion history via BAO and structure growth via RSD. Acoustic oscillations in the pre-recombination universe imprint a characteristic scale on matter clustering, which can be measured in the transverse and line-of-sight directions to determine the angular-diameter distance  $D_A(z)$  and Hubble parameter  $H(z)$ , respectively [11, 98, 17]. Anisotropy of clustering caused by galaxy peculiar velocities constrains (in linear perturbation theory) the combination  $\sigma_m(z)f_g(z)$ , where  $\sigma_m$  describes the rms amplitude of matter fluctuations and  $f_g(z) \equiv d \ln \sigma_m(z) / d \ln a$  is the fluctuation growth rate. Thus the GRS on its own can address the key questions identified by NWNH: whether cosmic acceleration is caused by modified gravity or by dark energy, and whether (in the latter case) the dark energy density evolves in time [30, 115]. These tests become more powerful in combination with weak lensing and cluster measurements from HLS Imaging and high-precision relative distance measurements from the Supernova Survey [24, 23]. The broadband shape of the galaxy power spectrum and higher order measures of galaxy clustering provide additional diagnostics of dark energy, neutrino masses, and inflation, and insights on the physics of galaxy formation. While all aspects of our GRS investigation are interconnected, we organize it in a structure similar to that of §3 for clarity: requirements, simulations, and prototype pipelines in §4.1, cosmological forecasting, modeling, and cosmological simulations in §4.3, and systematics testing and mitigation in §4.4.

### 4.1 Requirements

[Authors: Yun, Lado]

The most important task of the SIT in guiding development of the WFIRST HLS spectroscopy

is to set and validate the requirements of the instrument, the data reduction software, and the survey. The GRS Lead Co-I Wang will work closely with PI Doré and Co-Is Hirata and Teplitz on setting requirements for the GRS. To make fully informed decisions, the team requires high fidelity simulations of both instrument performance and the observable sky that the instrument will measure. The team must also ensure that the analysis of these data by the reduction pipeline will be of sufficient quality to enable measurement with the high precision needed for cosmology. These simulation and pipeline activities will require the team to coordinate with the WSC. Several members of our team (Wang, Teplitz, Capak, Helou) are located at the Infrared Processing and Analysis Center (IPAC), and work closely with the WSC. To the extent practical, we will draw on tools created by the WSC and design our own software and simulations to be useful to them. We note that our work primarily demands the ability to quickly and flexibly simulate different configurations and analyze the results with different algorithms, while the WSC has the task of developing tools for the community and production ready pipelines that integrate with the full WFIRST data system.

**Deriving requirements.** As with WL (§3.1), we will focus first on GRS requirements that may drive hardware choices, i.e., those that may be demanding in terms of grism design, detector properties, stability and repeatability of pointing, or dedicated calibration hardware. We will include a prioritized list of effects to incorporate in grism simulations. Over time, we will use our increasingly realistic network of simulations to evaluate the impact of requirements and possible trades from the pixel level through to cosmological inferences. The starting points for this process are the WFIRST ETC and survey planning software and the `CosmoLike` forecasting tool.

The maximum achievable statistical power of the GRS is determined mainly by the telescope aperture, throughput, detector area and pixel size, and allotted observing time. However, the statistical power and uniformity of the GRS are further affected by numerous aspects of instrument performance and survey design, e.g., spectral resolution, detector read noise and persistence, dither and roll angle pattern, image quality, complexity of non-1<sup>st</sup> order features, spatially varying thermal background due to the warm telescope, scattered light from bright stars, repeatability of the grism positioning, and accuracy of calibration of the wavelength-dependent PSF and distortion map.

Non-uniformity of the survey, which is inevitable to some degree, can be corrected in clustering measurements by weighting galaxies to account for incompleteness. However, large corrections typically come at a cost in statistical power, and imperfect knowledge of the non-uniformity leads to systematic errors in the inferred clustering. The other important source of observational systematics is contamination of the redshift catalog by artifacts or objects with incorrectly determined redshifts, and loss of objects from the catalog because of catastrophic redshift errors or uncertainties in the flux calibration. We will define requirements such that (a) the statistical power of the GRS is close to the maximum allowed by the telescope aperture and detector area and (b) the impact of uncorrected observational systematics is small compared to the statistical errors. The expected precision of the galaxy power spectrum provides a useful guide to the statistical power of the GRS, but the full question of cosmological constraining power depends on the astrophysical modeling techniques used to interpret the measured clustering, as discussed in §4.3 below. By the end of our investigation we will have a complete set of tools to evaluate the impact of hardware or strategy trades, changes in requirements, or changes in astrophysical inputs on the expected cosmological return from the GRS.

**Simulations.** Our development of requirements and a prototype spectroscopic pipeline will rely critically on realistic simulations of the pixel-level grism images. We will work closely with the WSC on developing these simulations for a variety of cases, ranging from simple widely separated sources to realistically clustered galaxy populations drawn from the cosmological simulations described in §4.3. Our team has extensive experience in producing such simulations for the Hubble Space Telescope (HST) and Euclid. Co-I Tepliz is one of the leaders of the HST Wide-Field Camera 3 (WFC3) IR Spectroscopic Parallel survey (WISPs), for which pixel simulations are vital in as-

sessing completeness and other parameters [18]; Co-I Wang (with Teplitz and Capak) is developing simulation techniques for the Euclid grism survey. Co-I Teplitz will lead our grism simulations for WFIRST. A critical astrophysical input for these simulations is the redshift-dependent luminosity function of H $\alpha$  and [OIII] emitters, which is currently uncertain at levels that have an important impact on WFIRST strategy and performance forecasts. Co-Is Teplitz and Wang are part of an HST archival study to reprocess existing data from multiple HST projects to mitigate systematic uncertainties of the H $\alpha$  luminosity function (LF) measurement. Through WISPs, Teplitz is also working to obtain significantly more HST data to improve the H $\alpha$  LF measurement. In addition, realistic galaxy templates are vital to the forecasting of grism measurements, and we are working to improve both line diagnostics and prediction of line vs. continuum properties.

**Prototype pipeline.** We will build a prototype spectroscopic pipeline for the analysis of slitless spectroscopic data to produce a redshift catalog. This is a complex, multi-step process. Co-I Teplitz has extensive experience with HST slitless spectroscopy using the WFC3, HST Near Infrared Camera and Multi-Object Spectrometer (NICMOS), and HST Space Telescope Imaging Spectrograph (STIS) instruments [5, 101, 108]; he will lead our work on prototype pipelines for the GRS. We will take the basic steps implemented for WFC3 processing as the starting point for a prototype WFIRST pipeline. This pipeline must clean the grism images of contamination from detector artifacts and cosmic rays, register and combine images from separate dithers and roll angles, match objects in the dispersed and direct imaging exposures, extract wavelength- and flux-calibrated 2D spectra, infer redshifts based on detected emission lines, and measure emission-line fluxes and other spectroscopic characteristics. The resulting catalogs are the input for the clustering analyses discussed further in §4.3.

**Analysis challenges.** Slitless spectroscopic analysis presents several important challenges. First, the pipeline must mitigate the confusion caused by overlapping spectra. The standard solution (used by the HST data pipeline) is to subtract a model of neighboring objects from each source as it is extracted. We will investigate the use of HST-like algorithms for WFIRST, as well as more sophisticated solutions that could produce better results, such as fitting the full pixel set for regions of the frame. Model dependent solutions, with iterative fitting, are potentially promising, but biases would have to be carefully understood. While the HLS obtains exposures at multiple roll angles, these may be greatly separated in time. This could introduce new problems for variable sources, or in fields with foreground moving objects. We will also develop methods to automate quality assessment and flagging of extracted spectra, as the sheer volume of WFIRST GRS will make human review of spectra (standard practice in current grism surveys) impossible.

A second major challenge is the need to mitigate catastrophic redshift errors. Such failures arise from the misidentification of redshifts (e.g., confusing [OIII] for H $\alpha$  in low signal to noise (S/N) spectra) or false-positive line detections caused by noise peaks or unflagged cosmic rays. Redshift fidelity can be greatly improved by using the photometric redshift estimates derived from the multi-band photometry as a prior in the redshift determination. Co-I Capak is spearheading multidimensional analysis of galaxy color information for WFIRST photometric redshifts, and that work will be folded into the spectroscopic pipeline.

**Completeness maps.** Nearly as important as the redshift catalogs themselves is the production of completeness maps that characterize the spatially varying depth of the survey, the level of contamination, and regions that should be masked because the catalog is unreliable. These completeness maps are used to weight galaxies in clustering analysis and/or to create random catalogs such that the local number density of points is proportional to the likelihood of successfully measuring a redshift of a galaxy if it were at that point. Co-I Samushia is leading a similar effort in DESI and has previously worked on quantifying and removing systematic effects associated with inaccuracies in random catalogues [93]. Co-I Ho has also led the effort in creating the SDSS-BOSS LSS catalog and randoms [83] and led the effort in removing observational effects in BOSS LSS catalog [86]. Co-I Samushia will lead our work to develop tools for creating these completeness maps

by a full forward-modeling method, where artificial sources are assigned random angular positions and redshifts, added to grism images, and pushed through the data pipeline. Compared to existing large redshift surveys (from ground-based fiber spectroscopy), the WFIRST completeness map will have much more complex small scale structure because of the varying numbers of exposures at individual points on the sky and sensitivity variations across the focal plane. Because of source confusion and sky background effects, the completeness and contamination will be a function of the local galaxy surface density. We will develop strategies and tools for recording these large and complex completeness maps in formats that can be efficiently used to create random catalogs and weight galaxies for clustering analyses. By the end of the investigation period, we will be able to create full pixel-level simulations from an input cosmological simulation (see §4.3), run them through our proto-type pipeline to create a redshift catalog and completeness map, and analyze the resulting artificial data set with our clustering analysis tools to compare to the idealized case that has the complete galaxy catalog of the cosmological simulation.

**Calibration strategies.** We will define the absolute and relative calibration requirements for the GRS, such as the angular scale and temporal stability. We will develop methods for calibrating the relative and absolute flux measurements along with wavelength calibration and redshift accuracy and completeness. For flux and wavelength calibration we will set requirements on the ground testing, in flight calibration sources, and calibration observations based on experience with other missions including Spitzer, HST, and Euclid. Furthermore, we will investigate self calibration strategies based on optimizing dither patterns and exposure times for the science observations and the use of touch-stone fields that both calibrate and provide long-term trending of the data. Both the primary and self calibration procedures will be tested with simulations specified by this SIT and conducted by the WSC. Finally, we will use the large spectroscopic surveys necessary for the weak lensing photo- $z$  calibration to verify the calibration by directly testing the redshift accuracy and completeness estimates from the simulations. Co-Is Capak and Padmanabhan, both with extensive experience from similar work for Euclid and BOSS, will lead our calibration work.

## 4.2 Simulations

[Authors: Shirley, Elena, Andrew, Alina, Alex, Yun]

## 4.3 Cosmological Forecasting, Modeling, and Simulations (D2, D8, D9)

**Forecasting.** Our initial forecasts for the cosmological constraints from the WFIRST GRS will adopt the model-independent approach (incorporating both BAO and RSD) that Co-I Wang has developed [118] and used for the WFIRST SDT reports and similar forecasts for Euclid [119]. This approach offers a fast way to forecast how uncertainties in  $H(z)$ ,  $D_A(z)$ , growth rates, and other cosmological parameters change in response to changes of the survey strategy, instrument performance, or astrophysical inputs such as the  $\mathrm{H}\alpha$  luminosity function. Early in this investigation, we will incorporate a full description of redshift-space galaxy clustering into `CosmoLike`, using a halo occupation density (HOD) framework similar to that already implemented for angular galaxy clustering [54]. In the medium term, we will also incorporate effects of clustering measurement and theoretical modeling systematics via nuisance parameters, analogous to our existing treatments of observational and theoretical systematics in weak lensing analysis. The expected level of these systematics will be informed by the studies described in §4.1 and below. As with the imaging survey, this comprehensive forecasting framework will enable us to connect low-level technical requirements to our top-level science goals.

**GRS modeling.** The development of the methodology for the interpretation of GRS data is centered on the mitigation of the astrophysical systematic effects for galaxy clustering measurements: nonlinear effects, RSD (growth rate signal on large scales and contamination on small and intermediate scales), and galaxy bias (the difference between galaxy and matter distributions).

Co-I Padmanabhan is a leading expert in BAO/RSD data analysis [79, 78, 80, 81, 122]; he will lead our work in GRS modeling/interpretation methods, with participation from the PI and Co-Is Bean, Ho, Samushia, Spergel, Wang, and Weinberg [17, 41, 4, 117, 77, 3, 2, 21].

BAO measurement is now a mature field, but WFIRST probes new regimes of precision and redshift using different instrumental choices and different classes of galaxy tracers from previous surveys. Effects of non-linear clustering and galaxy bias are expected to influence BAO measurements at the  $\sim 0.5\%$  level [80], which is significant compared to WFIRST statistical errors. Reconstruction methods [28, 81, 113], which attempt to reverse the nonlinear evolution of the BAO feature, appear to remove most of this effect while simultaneously improving the precision of BAO measurements. Current observational studies use very simple reconstruction algorithms. We will explore more sophisticated reconstruction methods, building on low redshift work [15], and test their performance on simulations of WFIRST galaxy redshift catalogs, including realistic treatments of survey geometry, redshift evolution, source space density, and variable completeness. Building on current work by Co-Is [81, 113, 77, 126], we will also investigate improved clustering estimators that can sharpen the precision and improve the robustness of BAO measurements.

In sharp contrast to BAO measurements, cosmological inference from RSD measurements is already limited mainly by uncertainties in theoretical modeling, with application of different models to the same underlying data yielding differences at the  $\sim 10\%$  level. Furthermore, the statistical signal-to-noise ratio of RSD measurements increases rapidly with decreasing scale, so there are potentially large gains from modeling that extends into the fully non-linear regime. We will pursue a variety of approaches to improving and testing RSD models, including the efficient method of computing predictions numerically by populating the halos of N-body simulations with galaxies. Co-I Spergel has expertise in combining imaging and spectroscopic data to predict the relationship between galaxies and halos [35]. One can think of this method as producing “emulators” [31] that predict galaxy clustering statistics as a function of cosmological parameters and parameters that describe the relation between galaxies and dark matter halos. The approach shows promise (e.g., [85]), but it relies on parameterized models for populating halos, and the accuracy of these needs to be tested against galaxy catalogs constructed in ways that do not share the same assumptions (e.g., by semi-analytic models or abundance-age matching). We will use similar techniques to investigate the impact of non-linear evolution and bias on the broadband galaxy power spectrum, and thus improve our ability to extract cosmological information from this measurement.

Galaxy bias is likely scale-dependent; its testing will require realistic ELG mocks, and its mitigation will require the successful measurement of the higher-order statistics of galaxy clustering, which in turn requires a sufficiently high galaxy number density for the GRS. An essential difference between the WFIRST and Euclid spectroscopic surveys is that due to the much smaller pixel scale ( $0.11''$  for WFIRST versus  $0.3''$  for Euclid) and larger telescope aperture, WFIRST is capable of carrying out a significantly deeper GRS, which can result in a much higher space density of the WFIRST sample over most of its redshift range. This high sampling density represents a significant science opportunity for WFIRST; in particular it will boost the significance of higher-order correlations. Building on previous work by members of our team [106, 94, 25, 16], we will investigate techniques that use the galaxy bispectrum and other higher-order statistics to sharpen cosmological constraints, by directly probing the matter density and velocity fields and by improving knowledge of “nuisance parameters” that describe galaxy bias. We will examine ways that RSD measurement precision can be improved by cross-correlating multiple tracer populations with different clustering bias to suppress cosmic variance [69, 10, 22], weighting galaxies by mass to suppress shot noise [96], and building group catalogs to collapse fingers-of-God [84]. We will investigate potential gains from cross-correlating the WFIRST galaxy redshift catalogs with H I 21 cm “intensity mapping” measurements, or with CMB measurements, or (at  $z > 2$ ) the Ly $\alpha$  forest. In all of these studies, we will pay particular attention to the influence of the sampling density, as this directly informs the trade between depth and area in the GRS (see §??).

**Cosmological simulations.** The cosmological simulations described in §3.4 will also be useful for the methodology development outlined above. However, the optimal simulations for BAO and RSD studies will typically be larger volume and lower resolution than those for weak lensing: large volumes are needed for good statistics and to eliminate finite box effects, but we do not need to model the small scale matter distribution or baryonic effects (which are encoded in the models used to populate halos with galaxies). Furthermore, simulations tuned to the WFIRST GRS need only be evolved to  $z = 1$ . Co-I Ho will lead our cosmological simulations for the GRS, with participation from team members Benson, Heitmann, Kiessling, Wang, and postdocs.

We will leverage the participation of several of our team members in Euclid and DESI to produce large simulated ELG catalogs, building on work we have already begun for these projects. To model emission-line selection, we will use both semi-analytic galaxy formation models (SAM) and HOD models that are tuned to produce observed number densities and clustering. We expect to be able to provide ELG mocks for WFIRST similar to those used by Euclid on a short time scale. We will incorporate these into the early pixel-level simulations described in §4.1, which will in turn be used to assess impacts of incompleteness, contamination, and redshift errors on the galaxy distribution.

We will base our first cosmological volume ELG catalogs on two very large simulations that are already available to us through collaborator Heitmann: the Outer Rim simulation, covering a volume of  $4.225 \text{ Gpc}^3$  with a particle mass of  $\sim 2 \times 10^9 M_\odot$ , and the Q Continuum simulation, covering a volume of  $1.3 \text{ Gpc}^3$  with particle mass of  $\sim 10^8 M_\odot$ . The Outer Rim simulation was used to create the simulated ELG catalog for DESI. We will combine the halo populations from these simulations with the Galacticus SAM code [9] to create clustered ELG catalogs, some in fixed-redshift cubes for methodology tests and some in light cones with realistic survey geometry and redshift evolution. Our most ambitious simulation efforts, later in the investigation period, will take full light-cone ELG catalogs, create pixel-level simulations that span the entire HLS area, and analyze these simulations with the proto-type pipeline to produce “observed” galaxy catalogs. We can then apply the full clustering measurement machinery (including corrections for varying completeness) to these catalogs and apply our cosmological inference tools to understand the impact of observational systematics on cosmology from the WFIRST GRS. As with the weak lensing investigation, we will investigate computational requirements beyond FY20 and techniques to reduce them. We will also make our simulations publicly available so that others can develop and test their own methods, with some of them released in the form of blind data challenges.

#### 4.4 Systematics Testing and Mitigation (D8)

The galaxy clustering measurements are susceptible to observational and astrophysical systematic effects. We discussed the astrophysical systematic effects and their mitigation in §4.3. We now focus on the observational systematic effects.

As we discussed in the §4.3, we will take full light cone ELG catalogs, create pixel-level simulations that span the entire HLS area, and produce the observed galaxy and corresponding random catalogs. We envision that multiple LSS catalogs will be produced by extracting the galaxies (and their corresponding randoms) according to their specific continuum levels and/or line-fluxes (or other selection criteria). For each of these catalogs, we will test for systematics such as effects of stellar density, dependencies on line luminosity, continuum luminosity and varying exposure number. Co-I Ho led investigations in effects of observational systematics on galaxy over-density in BOSS [41]; we will apply the same techniques in detecting galaxy over-density variations due to the various potential systematic sources. Co-I Ho will lead our work to design and perform internal empirical tests that involve dividing the galaxies (and corresponding randoms) into subsets of different continuum luminosity, line luminosity, galaxy environment, exposure number and other relevant parameters. We will pay particular attention to potential systematic effects caused by the complex structure of the completeness function, as a function of redshift. This is particularly important as the number of exposures can vary significantly, from 0 to 10 with a median of 7 in

the SDT15 observing strategy.

Team Co-Is have experience in using template projection method and cross-correlation method in photometric clustering to mitigate observational systematics such as stellar density, PSF variations, magnitude error fluctuations [82, 1]. We will adapt these methodologies to remove systematics in 3D clustering. We will test our 3D systematics detection and mitigation methodology on our simulated catalogs and check whether we achieve unbiased results in BAO distances and the growth rate of large scale structure. These potential systematics will also affect the photometric clustering and the full shape of the galaxy power spectrum, which can be powerful in constraining the sum of neutrino masses.

## 5 Cosmological Forecasts [Oli: Tim, Elisabeth, Olivier, 10 pages]

## 6 Operations Model for the HLS and Evaluation of Trades

Co-I Hirata is leading the development of the HLS observing plan, extending his previous tools used for the SDT. These tools incorporate observing constraints in the chosen orbit, an exposure-by-exposure observing sequence optimized with detailed model of overheads, and tiling/coverage maps including field distortions and curved sky effects. These tools treat both imaging and spectroscopy with unified functions and scripts, and are well suited to joint survey optimization when both hardware parameters (e.g., reaction wheel orientations) and the observing program (e.g., depth vs. area) are considered. This effort is coordinated with the scheduling and Design Reference Mission Working Groups. We are both providing an example detailed plan for the HLS to the DRM working group, and cross-checking the Project’s spreadsheet-level survey calculators against our simulations. The HLS observing plan is also being transferred to the Calibration Working Group, since the HLS observing strategy feeds directly into the issue of self-calibration.

### 6.1 Snapshot of the HLS observing plan

Our team provided a “snapshot” of the HLS observing sequence to the full FSWG on April 19, 2017. This is by no means a final or even optimized version of the HLS, but is a work in progress as a result of trades in Phase A, as well as the recent decision to reduce the primary mission to 5 years.

Major updates relative to the SDT plan have included:

- A Lissajous orbit around L2. This is presently a place-holder, as the exact orbit has not yet been selected (and would depend on the launch date), but it gives a possible sampling of Sun, Earth, and Moon constraints.
- A rotated WFI (by 90° relative to the Cycle 6 design).
- Recommended slew and settle times provided by the Project.
- Faster detector readout (200 kHz instead of 100 kHz).
- Changes to the exposure time and dithering strategy to accommodate a 5-year baseline mission (as is to be presented to the WIETR). Specifically, we reduced exposure time to 140.2 s (imaging) and 297.0 s (spectroscopy); and changed the dither pattern in J band. (We are working on checking this strategy with image simulations, if it causes a problem we may have to revert.)
- Implemented bright star avoidance (observations are skipped if there is an  $H_{AB} \leq 3$  star within 6 arcmin of any SCA).

Known current issues with the snapshot plan include:

- The SN and coronagraph programs in the code haven't been updated since the SDT (except to cut the mission time by a factor of 5/6), even though they will likely change significantly. As in the SDT report, the coronagraph has blocks of time reserved. This will evolve in order to align the HLS plan with the other groups, as well as any changes to the scheduling architecture that we are directed to implement (e.g. block scheduling).
- We have begun putting the deep fields into this document, but right now they are (i) not fully specified, (ii) the tiling is not optimized, and (iii) some roll angles don't align with WFIRST constraints (hence didn't schedule). These issues will be solved in the next snapshot.

**No policy decisions should be inferred from this sequence**, as these will come from a higher level.

The survey bounding box is  $2097 \text{ deg}^2$ . The area covered with  $\geq 3$  exposures in every filter and the grism, including edge effects and holes around the bright stars, is  $1947 \text{ deg}^2$ . The time required for this version of the HLS is 394 days (imaging) + 215 days (spectroscopy).

The scheduling tools output a set of charts, included in this package:

- Fig. 8: Graphical display of the 5-year observing sequence.
- Fig. 9: HLS distribution of number of exposures in each filter.
- Fig. 10: HLS distribution of dust column [ $E(B - V)$  in magnitudes]. Cosmological forecasts are based on a dust column of  $E(B - V) = 0.035 \text{ mag}$ .
- Fig. 11: HLS distribution of zodiacal light (normalized to 1 at the ecliptic poles averaged over the year). Cosmological forecasts are based on a zodiacal brightness of 1.60 (except for 1.75 in the F184 filter, which is the least sensitive to zodiacal light and therefore was scheduled at inferior times of year).
- Fig. 12: The footprint of the HLS on the sky. This is an area of ongoing optimization, as we consider the needs of the deep fields, overlap with LSST, and the fraction of the survey footprint accessible from Northern observatories such as Subaru.

## 6.2 Further optimizations

Our team plans to study further optimizations to the HLS – including more drastic changes such as multi-tiered surveys, or a significant re-balancing of area vs. depth – in Phase B. However, in preparation for SRR/MDR, our main focus has been on demonstrating at least one survey configuration that meets requirements, and the construction of tools that link the observing strategy to calibration studies (§3.3) and image simulations (§??).

## 7 Community Engagement and External Data-sets (D11, D12) [Oli: Olivier, 5 pages]

The coming decade will be an exciting time for cosmology. Before WFIRST launch, major cosmological imaging surveys (KiDS, HSC, DES) and the DESI and PFS spectroscopic surveys will significantly advance our current understanding. WFIRST, Euclid and LSST will then go further and survey the sky at optical and infrared wavelengths, the James Webb Space Telescope (JWST) and the Extremely Large Telescopes (ELTs) will make very deep maps of the sky; eROSITA will

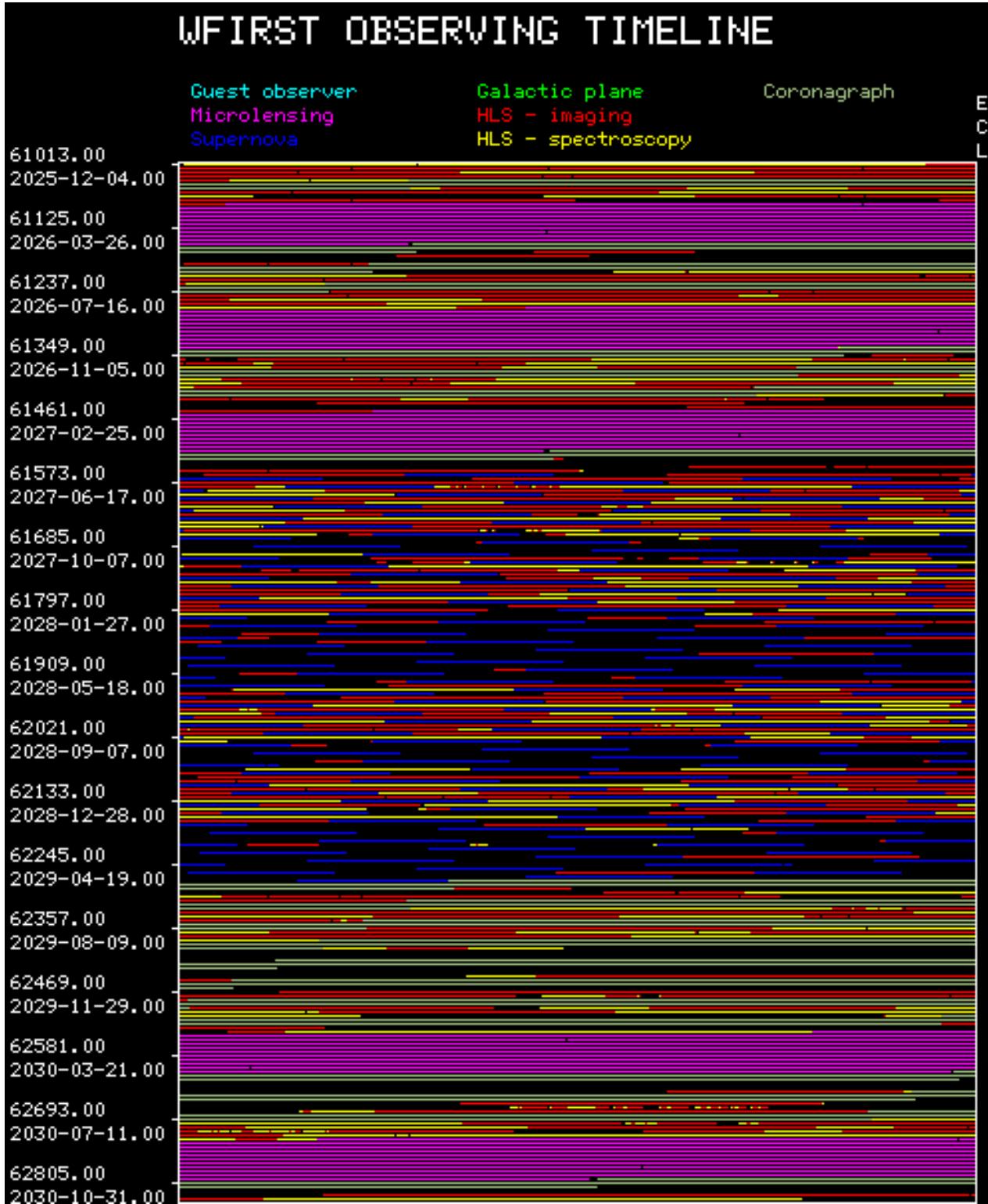


Figure 8: Observing timeline. Each row represents 7 days of observations, and is color-coded according to the observing program. Note the microlensing seasons (magenta), supernova survey (blue:  $\sim$ 5-day cadence), and HLS (red+yellow). Blank areas are not allocated. Labels on the left-hand side are shown every 16 weeks.

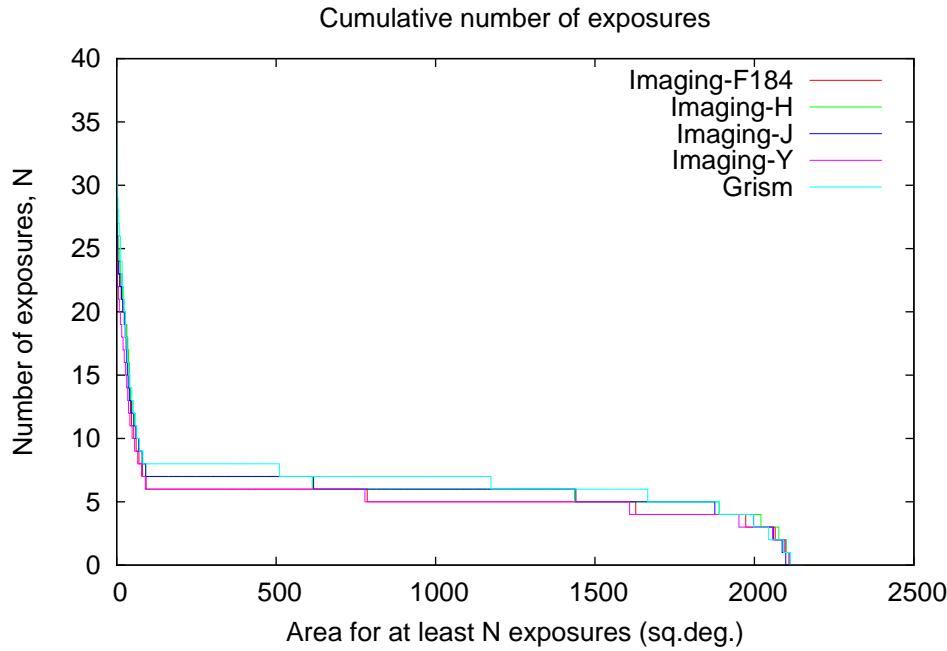


Figure 9: The cumulative distribution of HLS exposure depths above a certain area. The pile-up with many exposures at small area is the result of the deep fields.

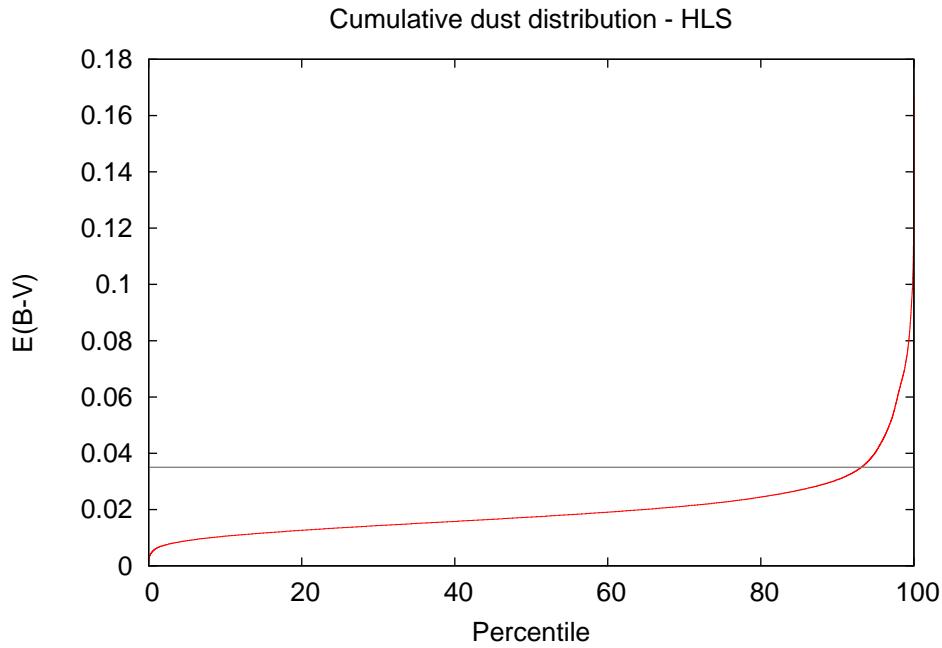


Figure 10: The cumulative distribution of Galactic dust in the HLS.

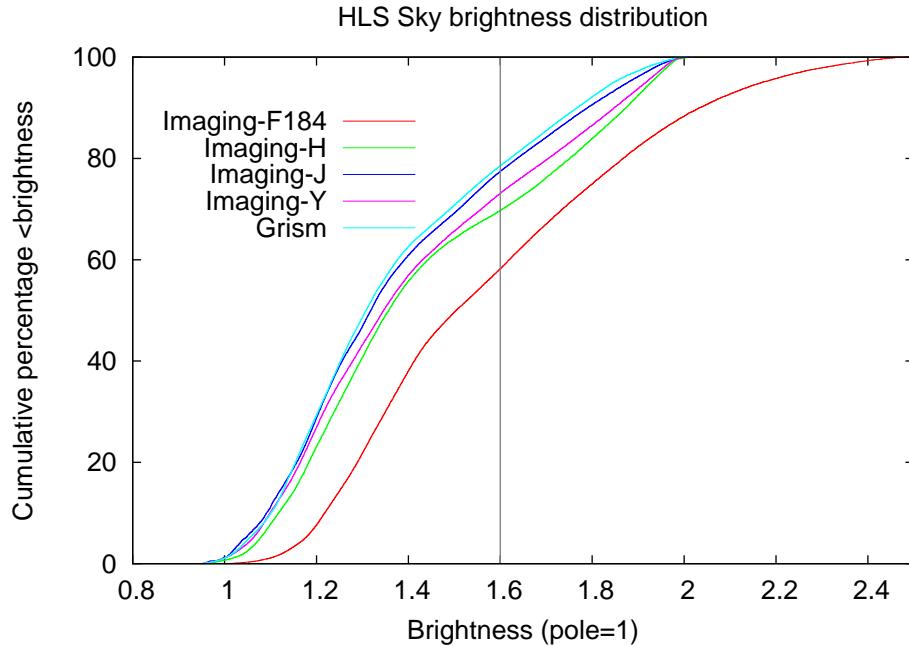


Figure 11: The cumulative distribution of zodiacal light in the HLS.

survey the X-ray sky; CMB-S4 will make a deep map of the millimeter sky; and the Canadian Hydrogen Intensity Mapping Experiment (CHIME) and other radio surveys will map the large-scale distribution of H I. A goal of the pre-CDR activities will be to determine the analysis infrastructure and observations needed to achieve the full potential of WFIRST in combination with these surveys. This is best done through a broad community effort that brings together scientists from these complementary projects. Our SIT team is extremely well-placed to do this, through team members' significant roles in a number of these projects and ties to the other major surveys. We will engage the community to identify and pursue the key areas where WFIRST and the concurrent projects will provide new opportunities to mitigate systematics and enhance the combined cosmological science return.

Through organizing a number of open workshops (with some modest travel support budgeted for the team and community members) our SIT will incorporate the interplay between major planned surveys and WFIRST into the WFIRST strategy, to identify: (i) pre-launch observations, (ii) how these external data sets affect the WFIRST observing strategy (e.g., deep fields) and the instrument, and (iii) the software needed (to be built post-CDR) for combining these data sets.

Themes we have identified for these multi-experiment workshops include: (i) Statistical methods for next generation cosmological analyses (ii) Strong lensing: synergies between ground and space (iii) Photometric redshift accuracy for next generation surveys (iv) LSST and WFIRST: data simulations and joint analyses to tackle systematics and reveal the dark sector.

Furthermore, through our team's international scientists in Canada and Japan, we plan to also engage these communities in WFIRST science. If these countries become partners, we hope that our international co-Investigators will receive support from their agencies and that they will serve as points of contact with their scientific communities and with the CHIME project (Smith) and the Subaru HSC/PFS project (Takada/Yoshida).

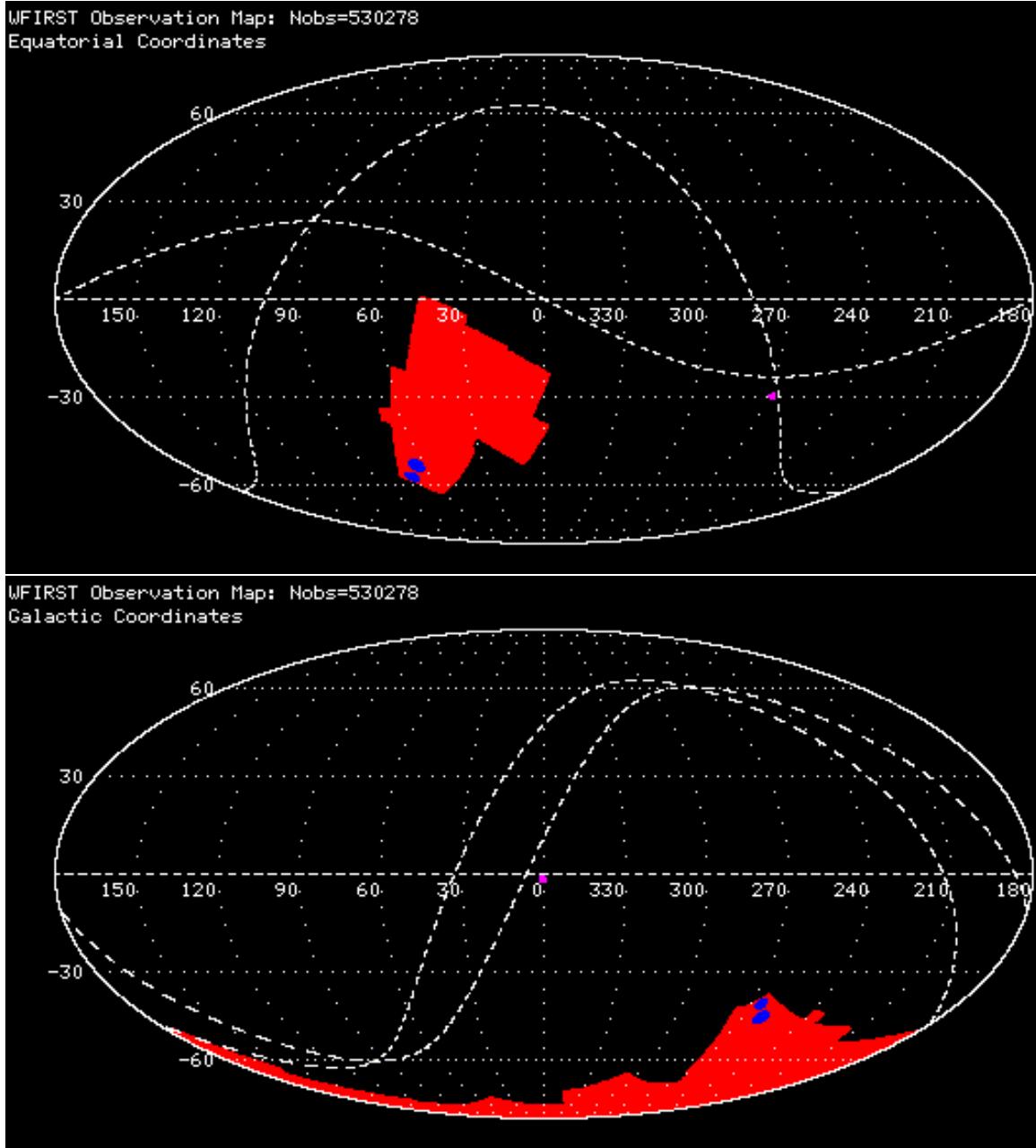


Figure 12: The footprint of the HLS (red) in Equatorial (top) and Galactic (bottom) coordinates.

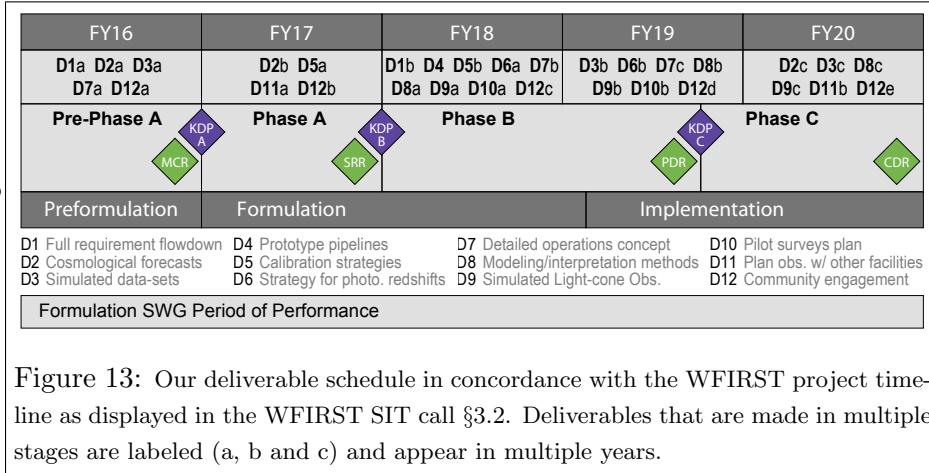
## 8 Future Workplan [Oli: Olivier, all, 5 pages]

**Team Management.** We have structured our team to cover all the areas of expertise required for the proposed work, and to maximize the synergies between WFIRST and the cosmology community. Scientific decisions including the allocation of funds will be made by the PI in consultation with a Steering Committee consisting of the PI, the two topic leads (Hirata and Wang), the topic sub-lead Weinberg plus Spergel, who have extensive organizational experience through WMAP, ACT, SDSS, National Research Council (NRC) and NASA committees, and other activities. While the PI will have final authority on these decisions, this Committee has the mission experience and breadth of knowledge needed to advise the PI. Monthly Steering Committee telecons will review the budget and priorities so that our effort remains focused on the scientific success of WFIRST. Each topic lead will report to the PI on a monthly basis, to facilitate monitoring of all team work, enable advice on budget priorities, and ensure a regular review of work effort. The PI will ensure proper communication with the WFIRST program office and will maximize collaborations with other WFIRST SITs. For example, we anticipate jointly developing and sharing low-level image simulation tools with SN-focused SITs. Topic leads will closely monitor the progress of the technical tasks (listed in §§ 2,3,4) in association with the deliverables and will adjust the scope of the work according to developments in the project and in the field.

**Risk Management.** Inefficient collaboration/coordination presents the most significant risk for accomplishing our program. Within our SIT, this is mitigated partly by most Co-Is having experience working together in existing (or past) projects. The PI will coordinate the team's effort, and ensure it is well-integrated into and aligned with the broader WFIRST effort. The PI will lead the following: (i) Monthly telecons with the Steering Committee, to optimize the team effort; (ii) Regular telecons with all team members, with status updates from the deliverables leads (Fig. ??) so that progress, issues, and solutions are broadly visible across the team. (iii) Yearly face-to-face meetings and more focused working meetings will be added as needed (the associated travel costs have been budgeted). Monthly, the Steering Committee will evaluate progress reports and redistribute responsibilities within the team if necessary. Finally, team members will present recent progress, results, and goals for the upcoming year to an external review panel at this annual meeting. This will also serve as a NASA management review with representatives of NASA HQ and the JPL WFIRST Project Office invited to attend. This follows the successful model used by the US Planck and US Euclid teams with which the PI and many Co-Is have experience. The PI will also set-up wikis and repositories for efficient sharing and discussion of software, documents, plans, progress and issues. We plan to share with the community the software developed for this program. Key results will be published in peer review papers.

**Program Milestones.** In Fig. 13 we outline milestones for our effort in conjunction with the project timeline.

**Postdocs and Students.** The largest component of our budget is support for post-doctoral researchers who will work under the supervision of the PI and Co-Is to carry out the many calculations, simulations, and tests needed to accomplish our tasks. Our budget incorporates an initial



plan for which postdocs will be located at which institutions in which years to work on which tasks.

However, the optimal division of labor may shift over time, and the Steering Committee will consider reallocations of work annually based on internal proposals from team members. Where appropriate, we will consider raising collaborators to funded Co-Is, with NASA approval, if they are best positioned to carry out particular tasks. Many of the methodology development efforts, and some of the technical tasks, are well suited to graduate students, supported by other sources or by this Investigation if their work clearly falls within its scope. In addition to accomplishing the work of this proposal, the involvement of many postdocs and students will build a cadre of young researchers who are ideally positioned to exploit the scientific opportunities of WFIRST.

**A Unified Team.** As is evident from our previous discussion, jointly addressing the weak lensing and redshift-space clustering elements of the WFIRST dark energy program leads to an ambitious task list that offers critical advantages relative to studying them separately. First is the effective use of expertise; many members of our team have expertise in both of these program elements and will contribute to both during the course of this investigation. Second is the commonality of the hardware; HLS Imaging and Spectroscopy use the same telescope, detectors, and data systems, and it makes sense to develop associated requirements through joint consideration of the two surveys. Third is the need to develop a unified operations concept for these two interleaved surveys; in overall footprint and in details of dither patterns and roll angles, choices made for one survey affect the performance of the other.

Finally, a joint investigation allows us to take a broad perspective on the WFIRST dark energy program, including a common forecasting framework that can account for complementary information and the ability of one survey to mitigate systematics in the other, common sets of cosmological simulations and models of galaxy bias that are relevant to both topics, and investigation of trades between HLS Imaging and Spectroscopy. We have carefully constructed a team capable of addressing all of the essential tasks for elements A and C of the WFIRST SIT NRA, and we have requested resources that will enable us to accomplish this work. If NASA decides to select additional teams in these areas we will be glad to collaborate with them. As emphasized in §7, WFIRST is a national and international resource, and it is valuable to engage as much of the astronomical community as possible in its formulation.

## 9 List of Acronyms and Abbreviations and References

$\sigma_m$	– rms amplitude of matter fluctuations	LSS	– large scale structure
$\Omega_m$	– dimensionless density of the Universe	LSST	– Large Synoptic Survey Telescope
$a$	– scale-factor of the Universe	NICMOS	– HST Near Infrared Camera and Multi-Object Spectrometer
ACT	– Atacama Cosmology Telescope	NIR	– near-infrared
AdvACT	– Advanced ACT	NRA	– NASA Research Announcement
AFTA	– Astrophysics Focused Telescope Asset	NRC	– National Research Council
BAO	– baryon acoustic oscillations	NWNH	– New Worlds, New Horizons
BOSS	– Baryon Oscillation Spectroscopic Survey	$P_m$	– matter power spectrum
CDR	– critical design review	PFS	– Subaru Prime Focus Spectrograph
CFHT	– Canada-France-Hawaii Telescope	photo- $z$	– photometric redshift
CGL	– cluster-galaxy lensing	PSF	– point spread function
CHIME	– Canadian Hydrogen Intensity Mapping Experiment	RSD	– redshift-space distortions
CL	– galaxy clusters / cluster growth	SAM	– semi-analytic galaxy formation models
CMB	– cosmic microwave background	SDSS	– Sloan Digital Sky Survey
CMB-S4	– CMB stage 4 experiment	SDT	– Science Definition Team
$D(z)$	– distance-redshift relation	SDT13	– 2013 WFIRST SDT report
$D_A(z)$	– angular-diameter distance	SDT15	– 2015 WFIRST SDT report
DE	– dark energy	SIT	– Science Investigation Team
DES	– Dark Energy Survey	SMEX	– NASA Small Explorer
DESC	– Dark Energy Science Collaboration	SN	– supernovae
DESI	– Dark Energy Spectroscopic Instrument	S/N	– signal-to-noise
ELG	– emission line galaxies	SPHEREx	– Spectrophotometer for the History of the Universe, Epoch of Reionization, and Ices Explorer
ELTs	– Extremely Large Telescopes	SPT-3G	– South Pole Telescope Third-Generation Camera Survey
eROSITA	– extended Roentgen Survey with an Imaging Telescope Array	STEP	– Shear Testing Program
ETC	– Exposure Time Calculator	STIS	– HST Space Telescope Imaging Spectrograph
$f_g$	– fluctuation growth rate	SZ	– Sunyaev-Zeldovich
FSWG	– Formulation Science Working Group	$w(z)$	– dark energy equation-of-state
GEO	– geostationary earth orbit	WFC3	– HST Wide-Field Camera 3
GGL	– galaxy-galaxy lensing	WFIRST	– Wide-Field Infrared Survey Telescope
GREAT	– Gravitational Lensing Accuracy Test	WISPs	– HST WFC3 IR Spectroscopic Parallel survey
GREAT3	– The third GREAT challenge	WL	– weak lensing
GRS	– Galaxy Redshift Survey	WMAP	– Wilkinson Microwave Anisotropy Probe
$H(z)$	– Hubble parameter	WPS	– WFIRST Preparatory Science
HLS	– High Latitude Survey	WSC	– WFIRST Science Centers
HOD	– halo occupation distribution	$z$	– redshift
HSC	– Subaru Hyper Suprime-Cam	$z_p$	– pivot redshift
HST	– Hubble Space Telescope		
IA	– intrinsic galaxy alignments		
IPAC	– Infrared Processing and Analysis Center		
IPC	– inter-pixel capacitance		
IR	– infrared		
JDEM	– Joint Dark Energy Mission		
JWST	– James Webb Space Telescope		
KiDS	– Kilo Degree Survey		
L2	– Lagrange point 2 orbit		
LF	– luminosity function		

We highlighted in bold the team members in the following bibliography.

## References

- [1] N. Agarwal, S. Ho, and S. Shandera. Constraining the initial conditions of the Universe using large scale structure. *JCAP*, 2:38, Feb. 2014. [26](#)
- [2] S. Alam, S. Ho, and A. Silvestri. Testing deviations from  $\Lambda$ CDM with growth rate measurements from 6 Large Scale Structure Surveys at  $z=0.06$  to 1. *ArXiv e-prints*, Sept. 2015. [24](#)
- [3] S. Alam, S. Ho, M. Vargas-Magaña, and D. P. Schneider. Testing general relativity with growth rate measurement from Sloan Digital Sky Survey - III. Baryon Oscillations Spectroscopic Survey galaxies. *MNRAS*, 453:1754–1767, Oct. 2015. [24](#)
- [4] L. Anderson, É. Aubourg, S. Bailey, F. Beutler, V. Bhardwaj, M. Blanton, A. S. Bolton, J. Brinkmann, J. R. Brownstein, A. Burden, C.-H. Chuang, A. J. Cuesta, K. S. Dawson, D. J. Eisenstein, S. Escoffier, J. E. Gunn, H. Guo, S. Ho, K. Honscheid, C. Howlett, D. Kirkby, R. H. Lupton, M. Manera, C. Maraston, C. K. McBride, O. Mena, F. Montesano, R. C. Nichol, S. E. Nuza, M. D. Olmstead, N. Padmanabhan, N. Palanque-Delabrouille, J. Parejko, W. J. Percival, P. Petitjean, F. Prada, A. M. Price-Whelan, B. Reid, N. A. Roe, A. J. Ross, N. P. Ross, C. G. Sabiu, S. Saito, L. Samushia, A. G. Sánchez, D. J. Schlegel, D. P. Schneider, C. G. Scoccola, H.-J. Seo, R. A. Skibba, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. V. Magaña, L. Verde, D. A. Wake, B. A. Weaver, D. H. Weinberg, M. White, X. Xu, C. Yèche, I. Zehavi, and G.-B. Zhao. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples. *MNRAS*, 441:24–62, June 2014. [24](#)
- [5] H. Atek, M. Malkan, P. McCarthy, H. I. Teplitz, C. Scarlata, B. Siana, A. Henry, J. W. Colbert, N. R. Ross, C. Bridge, A. J. Bunker, A. Dressler, R. A. E. Fosbury, C. Martin, and H. Shim. The WFC3 Infrared Spectroscopic Parallel (WISP) Survey. *ApJ*, 723:104–115, Nov. 2010. [22](#)
- [6] T. Baldauf, R. E. Smith, U. Seljak, and R. Mandelbaum. Algorithm for the direct reconstruction of the dark matter correlation function from weak lensing and galaxy clustering. *Phys. Rev. D*, 81(6):063531, Mar. 2010. [18](#)
- [7] M. R. Becker, M. A. Troxel, N. MacCrann, E. Krause, T. F. Eifler, O. Friedrich, A. Nicola, A. Refregier, A. Amara, D. Bacon, G. M. Bernstein, C. Bonnett, S. L. Bridle, M. T. Busha, C. Chang, S. Dodelson, B. Erickson, A. E. Evrard, J. Frieman, E. Gaztanaga, D. Gruen, W. Hartley, B. Jain, M. Jarvis, T. Kacprzak, D. Kirk, A. Kravtsov, B. Leistedt, E. S. Rykoff, C. Sabiu, C. Sanchez, H. Seo, E. Sheldon, R. H. Wechsler, J. Zuntz, T. Abbott, F. B. Abdalla, S. Allam, R. Armstrong, M. Banerji, A. H. Bauer, A. Benoit-Levy, E. Bertin, D. Brooks, E. Buckley-Geer, D. L. Burke, D. Capozzi, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, M. Crocce, C. E. Cunha, C. B. D’Andrea, L. N. da Costa, D. L. DePoy, S. Desai, H. T. Diehl, J. P. Dietrich, P. Doel, A. Fausti Neto, E. Fernandez, D. A. Finley, B. Flaugher, P. Fosalba, D. W. Gerdes, R. A. Gruendl, G. Gutierrez, K. Honscheid, D. J. James, K. Kuehn, N. Kuropatkin, O. Lahav, T. S. Li, M. Lima, M. A. G. Maia, M. March, P. Martini, P. Melchior, C. J. Miller, R. Miquel, J. J. Mohr, R. C. Nichol, B. Nord, R. Ogando, A. A. Plazas, K. Reil, A. K. Romer, A. Roodman, M. Sako, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, J. Thaler, D. Thomas, V. Vikram, A. R. Walker,

- and The DES Collaboration. Cosmic Shear Measurements with DES Science Verification Data. *ArXiv e-prints*, July 2015. [4](#)
- [8] J. Benjamin, L. Van Waerbeke, C. Heymans, M. Kilbinger, T. Erben, H. Hildebrandt, H. Hoekstra, T. D. Kitching, Y. Mellier, L. Miller, B. Rowe, T. Schrabback, F. Simpson, J. Coupon, L. Fu, J. Harnois-Déraps, M. J. Hudson, K. Kuijken, E. Sembolini, S. Vafaei, and M. Velander. CFHTLenS tomographic weak lensing: quantifying accurate redshift distributions. *MNRAS*, 431:1547–1564, May 2013. [16](#)
  - [9] A. J. Benson and R. Bower. Galaxy formation spanning cosmic history. *MNRAS*, 405:1573–1623, July 2010. [25](#)
  - [10] G. M. Bernstein and Y.-C. Cai. Cosmology without cosmic variance. *MNRAS*, 416:3009–3016, Oct. 2011. [24](#)
  - [11] C. Blake and K. Glazebrook. Probing Dark Energy Using Baryonic Oscillations in the Galaxy Power Spectrum as a Cosmological Ruler. *ApJ*, 594:665–673, Sept. 2003. [20](#)
  - [12] R. Bordoloi, S. J. Lilly, and A. Amara. Photo-z performance for precision cosmology. *MNRAS*, 406:881–895, Aug. 2010. [16](#)
  - [13] S. Bridle, S. T. Balan, M. Bethge, M. Gentile, S. Harmeling, C. Heymans, M. Hirsch, R. Hosseini, M. Jarvis, D. Kirk, T. Kitching, K. Kuijken, A. Lewis, S. Paulin-Henriksson, B. Schölkopf, M. Velander, L. Voigt, D. Witherick, A. Amara, G. Bernstein, F. Courbin, M. Gill, A. Heavens, R. Mandelbaum, R. Massey, B. Moghaddam, A. Rassat, A. Réfrégier, J. Rhodes, T. Schrabback, J. Shawe-Taylor, M. Shmakova, L. van Waerbeke, and D. Wittman. Results of the GREAT08 Challenge: an image analysis competition for cosmological lensing. *MNRAS*, 405:2044–2061, July 2010. [3](#)
  - [14] M. Cacciato, F. C. van den Bosch, S. More, H. Mo, and X. Yang. Cosmological constraints from a combination of galaxy clustering and lensing - III. Application to SDSS data. *MNRAS*, 430:767–786, Apr. 2013. [18](#)
  - [15] J. Carrick, S. J. Turnbull, G. Lavaux, and M. J. Hudson. Cosmological parameters from the comparison of peculiar velocities with predictions from the 2M++ density field. *MNRAS*, 450:317–332, June 2015. [24](#)
  - [16] Y.-C. Chen, S. Ho, P. E. Freeman, C. R. Genovese, and L. Wasserman. Cosmic Web Reconstruction through Density Ridges: Method and Algorithm. *MNRAS*, 454:1140–1156, Nov. 2015. [24](#)
  - [17] C.-H. Chuang and Y. Wang. Measurements of  $H(z)$  and  $D_A(z)$  from the two-dimensional two-point correlation function of Sloan Digital Sky Survey luminous red galaxies. *MNRAS*, 426:226–236, Oct. 2012. [20](#), [24](#)
  - [18] J. W. Colbert, H. Teplitz, H. Atek, A. Bunker, M. Rafelski, N. Ross, C. Scarlata, A. G. Bedregal, A. Dominguez, A. Dressler, A. Henry, M. Malkan, C. L. Martin, D. Masters, P. McCarthy, and B. Siana. Predicting Future Space Near-IR Grism Surveys Using the WFC3 Infrared Spectroscopic Parallels Survey. *ApJ*, 779:34, Dec. 2013. [22](#)
  - [19] N. R. Council. *New Worlds, New Horizons in Astronomy and Astrophysics*. The National Academies Press, Washington, DC, 2010. [1](#)

- [20] J. Coupon, S. Arnouts, L. van Waerbeke, T. Moutard, O. Ilbert, E. van Uitert, T. Erben, B. Garilli, L. Guzzo, C. Heymans, H. Hildebrandt, H. Hoekstra, M. Kilbinger, T. Kitching, Y. Mellier, L. Miller, M. Scodéglio, C. Bonnett, E. Branchini, I. Davidzon, G. De Lucia, A. Fritz, L. Fu, P. Hudelot, M. J. Hudson, K. Kuijken, A. Leauthaud, O. Le Fèvre, H. J. McCracken, L. Moscardini, B. T. P. Rowe, T. Schrabback, E. Semboloni, and M. Velander. The galaxy-halo connection from a joint lensing, clustering and abundance analysis in the CFHTLenS/VIPERS field. *MNRAS*, 449:1352–1379, May 2015. [18](#)
- [21] A. J. Cuesta, M. Vargas-Magaña, F. Beutler, A. S. Bolton, J. R. Brownstein, D. J. Eisenstein, H. Gil-Marín, S. Ho, C. K. McBride, C. Maraston, N. Padmanabhan, W. J. Percival, B. A. Reid, A. J. Ross, N. P. Ross, A. G. Sánchez, D. J. Schlegel, D. P. Schneider, D. Thomas, J. Tinker, R. Tojeiro, L. Verde, and M. White. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in the correlation function of LOWZ and CMASS galaxies in Data Release 12. *ArXiv e-prints*, Sept. 2015. [24](#)
- [22] R. de Putter and O. Doré. Designing an Inflation Galaxy Survey: how to measure  $\sigma(f_{\text{NL}}) \sim 1$  using scale-dependent galaxy bias. 2014. [24](#)
- [23] R. de Putter, O. Doré, and S. Das. Using Cross-Correlations to Calibrate Lensing Source Redshift Distributions: Improving Cosmological Constraints from Upcoming Weak Lensing Surveys. *Astrophys. J.*, 780:185, 2014. [20](#)
- [24] R. de Putter, O. Doré, and M. Takada. The Synergy between Weak Lensing and Galaxy Redshift Surveys. 2013. [20](#)
- [25] O. Doré, J. Bock, M. Ashby, P. Capak, A. Cooray, R. de Putter, T. Eifler, N. Flagey, Y. Gong, S. Habib, K. Heitmann, C. Hirata, W.-S. Jeong, R. Katti, P. Korngut, E. Krause, D.-H. Lee, D. Masters, P. Mauskopf, G. Melnick, B. Mennesson, H. Nguyen, K. Öberg, A. Pullen, A. Raccanelli, R. Smith, Y.-S. Song, V. Tolls, S. Unwin, T. Venumadhav, M. Viero, M. Werner, and M. Zemcov. Cosmology with the SPHEREX All-Sky Spectral Survey. *ArXiv e-prints*, Dec. 2014. [24](#)
- [26] T. Eifler, M. Kilbinger, and P. Schneider. Comparing cosmic shear measures. Optimizing the information content of cosmic shear data vectors. *A&A*, 482:9–19, Apr. 2008. [20](#)
- [27] T. Eifler, E. Krause, S. Dodelson, A. Zentner, A. Hearin, and N. Gnedin. Accounting for baryonic effects in cosmic shear tomography: Determining a minimal set of nuisance parameters using PCA. *ArXiv e-prints*, May 2014. [19](#)
- [28] D. J. Eisenstein, H.-J. Seo, E. Sirk, and D. N. Spergel. Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak. *ApJ*, 664:675–679, Aug. 2007. [24](#)
- [29] L. Fu, M. Kilbinger, T. Erben, C. Heymans, H. Hildebrandt, H. Hoekstra, T. D. Kitching, Y. Mellier, L. Miller, E. Semboloni, P. Simon, L. Van Waerbeke, J. Coupon, J. Harnois-Déraps, M. J. Hudson, K. Kuijken, B. Rowe, T. Schrabback, S. Vafaei, and M. Velander. CFHTLenS: cosmological constraints from a combination of cosmic shear two-point and three-point correlations. *MNRAS*, 441:2725–2743, July 2014. [17](#)
- [30] L. Guzzo, M. Pierleoni, B. Meneux, E. Branchini, O. Le Fèvre, C. Marinoni, B. Garilli, J. Blaizot, G. De Lucia, A. Pollo, H. J. McCracken, D. Bottini, V. Le Brun, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodéglio, L. Tresse, G. Vettolani, A. Zanichelli, C. Adami, S. Arnouts, S. Bardelli, M. Bolzonella, A. Bongiorno, A. Cappi, S. Charlot, P. Ciliegi, T. Contini, O. Cucciati, S. de la Torre, K. Dolag, S. Foucaud, P. Franzetti, I. Gavignaud, O. Ilbert,

- A. Iovino, F. Lamareille, B. Marano, A. Mazure, P. Memeo, R. Merighi, L. Moscardini, S. Paltani, R. Pellò, E. Perez-Montero, L. Pozzetti, M. Radovich, D. Vergani, G. Zamorani, and E. Zucca. A test of the nature of cosmic acceleration using galaxy redshift distortions. *Nature*, 451:541–544, Jan. 2008. [20](#)
- [31] K. Heitmann, E. Lawrence, J. Kwan, S. Habib, and D. Higdon. The Coyote Universe Extended: Precision Emulation of the Matter Power Spectrum. *ApJ*, 780:111, Jan. 2014. [24](#)
- [32] C. Heymans, E. Grocott, A. Heavens, M. Kilbinger, T. D. Kitching, F. Simpson, J. Benjamin, T. Erben, H. Hildebrandt, H. Hoekstra, Y. Mellier, L. Miller, L. Van Waerbeke, M. L. Brown, J. Coupon, L. Fu, J. Harnois-Déraps, M. J. Hudson, K. Kuijken, B. Rowe, T. Schrabback, E. Semboloni, S. Vafaei, and M. Velander. CFHTLenS tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments. *MNRAS*, 432:2433–2453, July 2013. [19](#)
- [33] C. Heymans, L. Van Waerbeke, D. Bacon, J. Berge, G. Bernstein, E. Bertin, S. Bridle, M. L. Brown, D. Clowe, H. Dahle, T. Erben, M. Gray, M. Hetterscheidt, H. Hoekstra, P. Hudelot, M. Jarvis, K. Kuijken, V. Margoniner, R. Massey, Y. Mellier, R. Nakajima, A. Refregier, J. Rhodes, T. Schrabback, and D. Wittman. The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations. *MNRAS*, 368:1323–1339, May 2006. [3](#)
- [34] C. Heymans, L. Van Waerbeke, L. Miller, T. Erben, H. Hildebrandt, H. Hoekstra, T. D. Kitching, Y. Mellier, P. Simon, C. Bonnett, J. Coupon, L. Fu, J. Harnois Déraps, M. J. Hudson, M. Kilbinger, K. Kuijken, B. Rowe, T. Schrabback, E. Semboloni, E. van Uitert, S. Vafaei, and M. Velander. CFHTLenS: the Canada-France-Hawaii Telescope Lensing Survey. *MNRAS*, 427:146–166, Nov. 2012. [4](#)
- [35] C. Hikage, M. Takada, and D. N. Spergel. Using galaxy-galaxy weak lensing measurements to correct the finger of God. *MNRAS*, 419:3457–3481, Feb. 2012. [24](#)
- [36] H. Hildebrandt, L. van Waerbeke, D. Scott, M. Béthermin, J. Bock, D. Clements, A. Conley, A. Cooray, J. S. Dunlop, S. Eales, T. Erben, D. Farrah, A. Franceschini, J. Glenn, M. Halpern, S. Heinis, R. J. Ivison, G. Marsden, S. J. Oliver, M. J. Page, I. Pérez-Fournon, A. J. Smith, M. Rowan-Robinson, I. Valtchanov, R. F. J. van der Burg, J. D. Vieira, M. Viero, and L. Wang. Inferring the mass of submillimetre galaxies by exploiting their gravitational magnification of background galaxies. *MNRAS*, 429:3230–3237, Mar. 2013. [19](#)
- [37] C. Hirata and U. Seljak. Shear calibration biases in weak-lensing surveys. *MNRAS*, 343:459–480, Aug. 2003. [5](#)
- [38] C. M. Hirata, R. Mandelbaum, M. Ishak, U. Seljak, R. Nichol, K. A. Pimbblet, N. P. Ross, and D. Wake. Intrinsic galaxy alignments from the 2SLAQ and SDSS surveys: luminosity and redshift scalings and implications for weak lensing surveys. *MNRAS*, 381:1197–1218, Nov. 2007. [19](#)
- [39] C. M. Hirata and U. Seljak. Intrinsic alignment-lensing interference as a contaminant of cosmic shear. *Phys. Rev. D*, 70(6):063526, Sept. 2004. [17](#), [19](#)
- [40] S. Ho, N. Agarwal, A. D. Myers, R. Lyons, A. Disbrow, H.-J. Seo, A. Ross, C. Hirata, N. Padmanabhan, R. O’Connell, E. Huff, D. Schlegel, A. Slosar, D. Weinberg, M. Strauss, N. P. Ross, D. P. Schneider, N. Bahcall, J. Brinkmann, N. Palanque-Delabrouille, and C. Yèche. Sloan Digital Sky Survey III photometric quasar clustering: probing the initial conditions of the Universe. *JCAP*, 5:40, May 2015. [19](#)

- [41] S. Ho, A. Cuesta, H.-J. Seo, R. de Putter, A. J. Ross, M. White, N. Padmanabhan, S. Saito, D. J. Schlegel, E. Schlafly, U. Seljak, C. Hernández-Monteagudo, A. G. Sánchez, W. J. Percival, M. Blanton, R. Skibba, D. Schneider, B. Reid, O. Mena, M. Viel, D. J. Eisenstein, F. Prada, B. A. Weaver, N. Bahcall, D. Bizyaev, H. Brewinton, J. Brinkman, L. Nicolaci da Costa, J. R. Gott, E. Malanushenko, V. Malanushenko, B. Nichol, D. Oravetz, K. Pan, N. Palanque-Delabrouille, N. P. Ross, A. Simmons, F. de Simoni, S. Snedden, and C. Yéche. Clustering of Sloan Digital Sky Survey III Photometric Luminous Galaxies: The Measurement, Systematics, and Cosmological Implications. *ApJ*, 761:14, Dec. 2012. [24](#), [25](#)
- [42] S. Ho, C. Hirata, N. Padmanabhan, U. Seljak, and N. Bahcall. Correlation of CMB with large-scale structure. I. Integrated Sachs-Wolfe tomography and cosmological implications. *Phys. Rev. D*, 78(4):043519, Aug. 2008. [16](#)
- [43] E. M. Huff, C. M. Hirata, R. Mandelbaum, D. Schlegel, U. Seljak, and R. H. Lupton. Seeing in the dark – I. Multi-epoch alchemy. *MNRAS*, 440:1296–1321, Mar. 2014. [5](#)
- [44] M. Jarvis, E. Sheldon, J. Zuntz, T. Kacprzak, S. L. Bridle, A. Amara, R. Armstrong, M. R. Becker, G. M. Bernstein, C. Bonnett, C. Chang, R. Das, J. P. Dietrich, A. Drlica-Wagner, T. F. Eifler, C. Gangkofner, D. Gruen, M. Hirsch, E. M. Huff, B. Jain, S. Kent, D. Kirk, N. MacCrann, P. Melchior, A. A. Plazas, A. Refregier, B. Rowe, E. S. Rykoff, S. Samuroff, C. Sánchez, E. Suchyta, M. A. Troxel, V. Vikram, T. Abbott, F. B. Abdalla, S. Allam, J. Annis, A. Benoit-Lévy, E. Bertin, D. Brooks, E. Buckley-Geer, D. L. Burke, D. Capozzi, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, M. Crocce, C. E. Cunha, C. B. D’Andrea, L. N. da Costa, D. L. DePoy, S. Desai, H. T. Diehl, P. Doel, A. Fausti Neto, B. Flaugher, P. Fosalba, J. Frieman, E. Gaztanaga, D. W. Gerdes, R. A. Gruendl, G. Gutierrez, K. Honscheid, D. J. James, K. Kuehn, N. Kuropatkin, O. Lahav, T. S. Li, M. Lima, M. March, P. Martini, R. Miquel, J. J. Mohr, E. Neilsen, B. Nord, R. Ogando, K. Reil, A. K. Romer, A. Roodman, M. Sako, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, M. E. C. Swanson, G. Tarle, J. Thaler, D. Thomas, A. R. Walker, and R. H. Wechsler. The DES Science Verification Weak Lensing Shear Catalogs. *ArXiv e-prints*, July 2015. [5](#), [20](#)
- [45] J. Jasche and B. D. Wandelt. Bayesian inference from photometric redshift surveys. *MNRAS*, 425:1042–1056, Sept. 2012. [16](#)
- [46] Y. P. Jing, P. Zhang, W. P. Lin, L. Gao, and V. Springel. The Influence of Baryons on the Clustering of Matter and Weak-Lensing Surveys. *ApJ*, 640:L119–L122, Apr. 2006. [19](#)
- [47] A. Kannawadi, C. A. Shapiro, R. Mandelbaum, C. M. Hirata, J. W. Kruk, and J. D. Rhodes. The Impact of Interpixel Capacitance in CMOS Detectors on PSF Shapes and Implications for WFIRST. *PASP*, 128(9):095001, Sept. 2016. [6](#)
- [48] I. Kayo, M. Takada, and B. Jain. Information content of weak lensing power spectrum and bispectrum: including the non-Gaussian error covariance matrix. *MNRAS*, 429:344–371, Feb. 2013. [17](#)
- [49] A. Kiessling, M. Cacciato, B. Joachimi, D. Kirk, T. D. Kitching, A. Leonard, R. Mandelbaum, B. M. Schäfer, C. Sifón, M. L. Brown, and A. Rassat. Galaxy alignments: Theory, modelling and simulations. *ArXiv e-prints*, Apr. 2015. [19](#)
- [50] D. Kirk, M. L. Brown, H. Hoekstra, B. Joachimi, T. D. Kitching, R. Mandelbaum, C. Sifón, M. Cacciato, A. Choi, A. Kiessling, A. Leonard, A. Rassat, and B. M. Schäfer. Galaxy alignments: Observations and impact on cosmology. *ArXiv e-prints*, Apr. 2015. [19](#)

- [51] D. Kirk, I. Laszlo, S. Bridle, and R. Bean. Optimising cosmic shear surveys to measure modifications to gravity on cosmic scales. *ArXiv e-prints*, Sept. 2011. [19](#)
- [52] T. D. Kitching, S. T. Balan, S. Bridle, N. Cantale, F. Courbin, T. Eifler, M. Gentile, M. S. S. Gill, S. Harmeling, C. Heymans, M. Hirsch, K. Honscheid, T. Kacprzak, D. Kirkby, D. Margala, R. J. Massey, P. Melchior, G. Nurbaea, K. Patton, J. Rhodes, B. T. P. Rowe, A. N. Taylor, M. Tewes, M. Viola, D. Witherick, L. Voigt, J. Young, and J. Zuntz. Image analysis for cosmology: results from the GREAT10 Galaxy Challenge. *MNRAS*, 423:3163–3208, July 2012. [3](#)
- [53] E. Krause, T. Eifler, and J. Blazek. The impact of intrinsic alignment on current and future cosmic shear surveys. *ArXiv e-prints*, June 2015. [17](#), [18](#), [19](#)
- [54] E. Krause, C. M. Hirata, C. Martin, J. D. Neill, and T. K. Wyder. Halo occupation distribution modelling of green valley galaxies. *MNRAS*, 428:2548–2564, Jan. 2013. [23](#)
- [55] I. Laszlo, R. Bean, D. Kirk, and S. Bridle. Disentangling dark energy and cosmic tests of gravity from weak lensing systematics. *MNRAS*, 423:1750–1765, June 2012. [17](#), [19](#)
- [56] LSST Dark Energy Science Collaboration. Large Synoptic Survey Telescope: Dark Energy Science Collaboration. *arXiv:1211.0310*, Nov. 2012. [18](#)
- [57] R. Lupton, J. E. Gunn, Z. Ivezić, G. R. Knapp, and S. Kent. The SDSS Imaging Pipelines. In F. R. Harnden, Jr., F. A. Primini, and H. E. Payne, editors, *Astronomical Data Analysis Software and Systems X*, volume 238 of *Astronomical Society of the Pacific Conference Series*, page 269, 2001. [5](#)
- [58] R. Mandelbaum, C. Blake, S. Bridle, F. B. Abdalla, S. Brough, M. Colless, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, K. Glazebrook, B. Jelliffe, R. J. Jurek, I.-H. Li, B. Madore, C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, and T. Wyder. The WiggleZ Dark Energy Survey: direct constraints on blue galaxy intrinsic alignments at intermediate redshifts. *MNRAS*, 410:844–859, Jan. 2011. [19](#)
- [59] R. Mandelbaum, C. M. Hirata, M. Ishak, U. Seljak, and J. Brinkmann. Detection of large-scale intrinsic ellipticity-density correlation from the Sloan Digital Sky Survey and implications for weak lensing surveys. *MNRAS*, 367:611–626, Apr. 2006. [19](#)
- [60] R. Mandelbaum, C. M. Hirata, A. Leauthaud, R. J. Massey, and J. Rhodes. Precision simulation of ground-based lensing data using observations from space. *MNRAS*, 420:1518–1540, Feb. 2012. [5](#)
- [61] R. Mandelbaum, C. M. Hirata, U. Seljak, J. Guzik, N. Padmanabhan, C. Blake, M. R. Blanton, R. Lupton, and J. Brinkmann. Systematic errors in weak lensing: application to SDSS galaxy-galaxy weak lensing. *MNRAS*, 361:1287–1322, Aug. 2005. [5](#)
- [62] R. Mandelbaum, B. Rowe, R. Armstrong, D. Bard, E. Bertin, J. Bosch, D. Boutigny, F. Courbin, W. A. Dawson, A. Donnarumma, I. Fenech Conti, R. Gavazzi, M. Gentile, M. S. S. Gill, D. W. Hogg, E. M. Huff, M. J. Jee, T. Kacprzak, M. Kilbinger, T. Kuntzer, D. Lang, W. Luo, M. C. March, P. J. Marshall, J. E. Meyers, L. Miller, H. Miyatake, R. Nakajima, F. M. Ngolé Mboula, G. Nurbaea, Y. Okura, S. Paulin-Henriksson, J. Rhodes, M. D. Schneider, H. Shan, E. S. Sheldon, M. Simet, J.-L. Starck, F. Sureau, M. Tewes, K. Zarb Adami, J. Zhang, and J. Zuntz. GREAT3 results - I. Systematic errors in shear estimation and the impact of real galaxy morphology. *MNRAS*, 450:2963–3007, July 2015. [3](#)

- [63] R. Mandelbaum, B. Rowe, J. Bosch, C. Chang, F. Courbin, M. Gill, M. Jarvis, A. Kannawadi, T. Kacprzak, C. Lackner, A. Leauthaud, H. Miyatake, R. Nakajima, J. Rhodes, M. Simet, J. Zuntz, B. Armstrong, S. Bridle, J. Coupon, J. P. Dietrich, M. Gentile, C. Heymans, A. S. Jurling, S. M. Kent, D. Kirkby, D. Margala, R. Massey, P. Melchior, J. Peterson, A. Roodman, and T. Schrabback. The Third Gravitational Lensing Accuracy Testing (GREAT3) Challenge Handbook. *ApJS*, 212:5, May 2014. [5](#)
- [64] R. Mandelbaum, U. Seljak, C. M. Hirata, S. Bardelli, M. Bolzonella, A. Bongiorno, M. Carollo, T. Contini, C. E. Cunha, B. Garilli, A. Iovino, P. Kampczyk, J.-P. Kneib, C. Knobel, D. C. Koo, F. Lamareille, O. Le Fèvre, J.-F. Le Borgne, S. J. Lilly, C. Maier, V. Mainieri, M. Mignoli, J. A. Newman, P. A. Oesch, E. Perez-Montero, E. Ricciardelli, M. Scodéggi, J. Silverman, and L. Tasca. Precision photometric redshift calibration for galaxy-galaxy weak lensing. *MNRAS*, 386:781–806, May 2008. [16](#)
- [65] R. Mandelbaum, A. Slosar, T. Baldauf, U. Seljak, C. M. Hirata, R. Nakajima, R. Reyes, and R. E. Smith. Cosmological parameter constraints from galaxy-galaxy lensing and galaxy clustering with the SDSS DR7. *MNRAS*, 432:1544–1575, June 2013. [18](#)
- [66] A. B. Mantz, A. von der Linden, S. W. Allen, D. E. Applegate, P. L. Kelly, R. G. Morris, D. A. Rapetti, R. W. Schmidt, S. Adhikari, M. T. Allen, P. R. Burchat, D. L. Burke, M. Cataneo, D. Donovan, H. Ebeling, S. Shandera, and A. Wright. Weighing the giants - IV. Cosmology and neutrino mass. *MNRAS*, 446:2205–2225, Jan. 2015. [19](#)
- [67] R. Massey, C. Heymans, J. Bergé, G. Bernstein, S. Bridle, D. Clowe, H. Dahle, R. Ellis, T. Erben, M. Hetterscheidt, F. W. High, C. Hirata, H. Hoekstra, P. Hudelot, M. Jarvis, D. Johnston, K. Kuijken, V. Margoniner, R. Mandelbaum, Y. Mellier, R. Nakajima, S. Paulin-Henriksson, M. Peeples, C. Roat, A. Refregier, J. Rhodes, T. Schrabback, M. Schirmer, U. Seljak, E. Semboloni, and L. van Waerbeke. The Shear Testing Programme 2: Factors affecting high-precision weak-lensing analyses. *MNRAS*, 376:13–38, Mar. 2007. [3](#)
- [68] D. Masters, P. Capak, D. Stern, O. Ilbert, M. Salvato, S. Schmidt, G. Longo, J. Rhodes, S. Paltani, B. Mobasher, H. Hoekstra, H. Hildebrandt, J. Coupon, C. Steinhardt, J. Speagle, A. Faisst, A. Kalinich, M. Brodwin, M. Brescia, and S. Cavauti. Mapping the Galaxy Color-Redshift Relation: Optimal Photometric Redshift Calibration Strategies for Cosmology Surveys. *ArXiv e-prints*, Sept. 2015. [16](#)
- [69] P. McDonald and U. Seljak. How to evade the sample variance limit on measurements of redshift-space distortions. *JCAP*, 10:7, Oct. 2009. [24](#)
- [70] B. Ménard, R. Scranton, S. Schmidt, C. Morrison, D. Jeong, T. Budavari, and M. Rahman. Clustering-based redshift estimation: method and application to data. *ArXiv e-prints*, Mar. 2013. [16](#)
- [71] S. More, H. Miyatake, R. Mandelbaum, M. Takada, D. N. Spergel, J. R. Brownstein, and D. P. Schneider. The Weak Lensing Signal and the Clustering of BOSS Galaxies. II. Astrophysical and Cosmological Constraints. *ApJ*, 806:2, June 2015. [18](#)
- [72] E.-M. Mueller, F. de Bernardis, R. Bean, and M. D. Niemack. Constraints on gravity and dark energy from the pairwise kinematic Sunyaev-Zeldovich effect. *Astrophys. J.*, 808(1):47, 2015. [19](#)
- [73] E.-M. Mueller, F. de Bernardis, R. Bean, and M. D. Niemack. Constraints on massive neutrinos from the pairwise kinematic Sunyaev-Zel'dovich effect. *Phys. Rev.*, D92(6):063501, 2015. [19](#)

- [74] R. Nakajima, R. Mandelbaum, U. Seljak, J. D. Cohn, R. Reyes, and R. Cool. Photometric redshift requirements for lens galaxies in galaxy-galaxy lensing analyses. *MNRAS*, 420:3240–3263, Mar. 2012. [16](#)
- [75] J. A. Newman. Calibrating Redshift Distributions beyond Spectroscopic Limits with Cross-Correlations. *ApJ*, 684:88–101, Sept. 2008. [16](#)
- [76] J. A. Newman, A. Abate, F. B. Abdalla, S. Allam, S. W. Allen, R. Ansari, S. Bailey, W. A. Barkhouse, T. C. Beers, M. R. Blanton, M. Brodwin, J. R. Brownstein, R. J. Brunner, M. Carrasco Kind, J. L. Cervantes-Cota, E. Cheu, N. E. Chisari, M. Colless, J. Comparat, J. Coupon, C. E. Cunha, A. de la Macorra, I. P. Dell’Antonio, B. L. Frye, E. J. Gawiser, N. Gehrels, K. Grady, A. Hagen, P. B. Hall, A. P. Hearin, H. Hildebrandt, C. M. Hirata, S. Ho, K. Honscheid, D. Huterer, Ž. Ivezić, J.-P. Kneib, J. W. Kruk, O. Lahav, R. Mandelbaum, J. L. Marshall, D. J. Matthews, B. Ménard, R. Miquel, M. Moniez, H. W. Moos, J. Moustakas, A. D. Myers, C. Papovich, J. A. Peacock, C. Park, M. Rahman, J. Rhodes, J.-S. Ricol, I. Sadeh, A. Slozar, S. J. Schmidt, D. K. Stern, J. Anthony Tyson, A. von der Linden, R. H. Wechsler, W. M. Wood-Vasey, and A. R. Zentner. Spectroscopic needs for imaging dark energy experiments. *Astroparticle Physics*, 63:81–100, Mar. 2015. [16](#)
- [77] K. Osumi, S. Ho, D. J. Eisenstein, and M. Vargas-Magaña. Robust New Statistic for fitting the Baryon Acoustic Feature. *ArXiv e-prints*, May 2015. [24](#)
- [78] N. Padmanabhan, D. J. Schlegel, D. P. Finkbeiner, J. C. Barentine, M. R. Blanton, H. J. Brewington, J. E. Gunn, M. Harvanek, D. W. Hogg, Ž. Ivezić, D. Johnston, S. M. Kent, S. J. Kleinman, G. R. Knapp, J. Krzesinski, D. Long, E. H. Neilsen, Jr., A. Nitta, C. Loomis, R. H. Lupton, S. Rowels, S. A. Snedden, M. A. Strauss, and D. L. Tucker. An Improved Photometric Calibration of the Sloan Digital Sky Survey Imaging Data. *ApJ*, 674:1217–1233, Feb. 2008. [24](#)
- [79] N. Padmanabhan, D. J. Schlegel, U. Seljak, A. Makarov, N. A. Bahcall, M. R. Blanton, J. Brinkmann, D. J. Eisenstein, D. P. Finkbeiner, J. E. Gunn, D. W. Hogg, Ž. Ivezić, G. R. Knapp, J. Loveday, R. H. Lupton, R. C. Nichol, D. P. Schneider, M. A. Strauss, M. Tegmark, and D. G. York. The clustering of luminous red galaxies in the Sloan Digital Sky Survey imaging data. *MNRAS*, 378:852–872, July 2007. [19](#), [24](#)
- [80] N. Padmanabhan and M. White. Calibrating the baryon oscillation ruler for matter and halos. *Phys. Rev. D*, 80(6):063508, Sept. 2009. [24](#)
- [81] N. Padmanabhan, X. Xu, D. J. Eisenstein, R. Scalzo, A. J. Cuesta, K. T. Mehta, and E. Kazin. A 2 per cent distance to  $z = 0.35$  by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey. *MNRAS*, 427:2132–2145, Dec. 2012. [24](#)
- [82] A. R. Pullen and C. M. Hirata. Systematic Effects in Large-Scale Angular Power Spectra of Photometric Quasars and Implications for Constraining Primordial Non-Gaussianity. *PASP*, 125:705–718, June 2013. [26](#)
- [83] B. Reid, S. Ho, N. Padmanabhan, W. J. Percival, J. Tinker, R. Tojeiro, M. White, D. J. Eisenstein, C. Maraston, A. J. Ross, A. G. Sanchez, D. Schlegel, E. Sheldon, M. A. Strauss, D. Thomas, D. Wake, F. Beutler, D. Bizyaev, A. S. Bolton, J. R. Brownstein, C.-H. Chuang, K. Dawson, P. Harding, F.-S. Kitaura, A. Leauthaud, K. Masters, C. K. McBride, S. More, M. D. Olmstead, D. Oravetz, S. E. Nuza, K. Pan, J. Parejko, J. Pforr, F. Prada, S. Rodriguez-Torres, S. Salazar-Albornoz, L. Samushia, D. P. Schneider, C. G. Scoccola, A. Simmons, and

- M. Vargas-Magana. SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large scale structure catalogues. *ArXiv e-prints*, Sept. 2015. [22](#)
- [84] B. A. Reid, W. J. Percival, D. J. Eisenstein, L. Verde, D. N. Spergel, R. A. Skibba, N. A. Bahcall, T. Budavari, J. A. Frieman, M. Fukugita, J. R. Gott, J. E. Gunn, Ž. Ivezić, G. R. Knapp, R. G. Kron, R. H. Lupton, T. A. McKay, A. Meiksin, R. C. Nichol, A. C. Pope, D. J. Schlegel, D. P. Schneider, C. Stoughton, M. A. Strauss, A. S. Szalay, M. Tegmark, M. S. Vogeley, D. H. Weinberg, D. G. York, and I. Zehavi. Cosmological constraints from the clustering of the Sloan Digital Sky Survey DR7 luminous red galaxies. *MNRAS*, 404:60–85, May 2010. [24](#)
- [85] B. A. Reid, H.-J. Seo, A. Leauthaud, J. L. Tinker, and M. White. A 2.5 per cent measurement of the growth rate from small-scale redshift space clustering of SDSS-III CMASS galaxies. *MNRAS*, 444:476–502, Oct. 2014. [24](#)
- [86] A. J. Ross, S. Ho, A. J. Cuesta, R. Tojeiro, W. J. Percival, D. Wake, K. L. Masters, R. C. Nichol, A. D. Myers, F. de Simoni, H. J. Seo, C. Hernández-Monteagudo, R. Crittenden, M. Blanton, J. Brinkmann, L. A. N. da Costa, H. Guo, E. Kazin, M. A. G. Maia, C. Maraston, N. Padmanabhan, F. Prada, B. Ramos, A. Sanchez, E. F. Schlaflly, D. J. Schlegel, D. P. Schneider, R. Skibba, D. Thomas, B. A. Weaver, M. White, and I. Zehavi. Ameliorating systematic uncertainties in the angular clustering of galaxies: a study using the SDSS-III. *MNRAS*, 417:1350–1373, Oct. 2011. [22](#)
- [87] A. J. Ross, W. J. Percival, A. G. Sánchez, L. Samushia, S. Ho, E. Kazin, M. Manera, B. Reid, M. White, R. Tojeiro, C. K. McBride, X. Xu, D. A. Wake, M. A. Strauss, F. Montesano, M. E. C. Swanson, S. Bailey, A. S. Bolton, A. M. Dorta, D. J. Eisenstein, H. Guo, J.-C. Hamilton, R. C. Nichol, N. Padmanabhan, F. Prada, D. J. Schlegel, M. V. Magaña, I. Zehavi, M. Blanton, D. Bizyaev, H. Brewington, A. J. Cuesta, E. Malanushenko, V. Malanushenko, D. Oravetz, J. Parejko, K. Pan, D. P. Schneider, A. Shelden, A. Simmons, S. Snedden, and G.-b. Zhao. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: analysis of potential systematics. *MNRAS*, 424:564–590, July 2012. [19](#)
- [88] B. Rowe, C. Hirata, and J. Rhodes. Optimal Linear Image Combination. *Astrophys. J.*, 741:46, 2011. [5](#)
- [89] B. T. P. Rowe, M. Jarvis, R. Mandelbaum, G. M. Bernstein, J. Bosch, M. Simet, J. E. Meyers, T. Kacprzak, R. Nakajima, J. Zuntz, H. Miyatake, J. P. Dietrich, R. Armstrong, P. Melchior, and M. S. S. Gill. GALSIM: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150, Apr. 2015. [5](#)
- [90] E. Rozo, R. H. Wechsler, E. S. Rykoff, J. T. Annis, M. R. Becker, A. E. Evrard, J. A. Frieman, S. M. Hansen, J. Hao, D. E. Johnston, B. P. Koester, T. A. McKay, E. S. Sheldon, and D. H. Weinberg. Cosmological Constraints from the Sloan Digital Sky Survey maxBCG Cluster Catalog. *ApJ*, 708:645–660, Jan. 2010. [19](#)
- [91] D. H. Rudd, A. R. Zentner, and A. V. Kravtsov. Effects of Baryons and Dissipation on the Matter Power Spectrum. *ApJ*, 672:19–32, Jan. 2008. [19](#)
- [92] E. S. Rykoff, E. Rozo, M. T. Busha, C. E. Cunha, A. Finoguenov, A. Evrard, J. Hao, B. P. Koester, A. Leauthaud, B. Nord, M. Pierre, R. Reddick, T. Sadibekova, E. S. Sheldon, and R. H. Wechsler. redMaPPer. I. Algorithm and SDSS DR8 Catalog. *ApJ*, 785:104, Apr. 2014. [18](#)

- [93] L. Samushia, W. J. Percival, and A. Raccanelli. Interpreting large-scale redshift-space distortion measurements. *MNRAS*, 420:2102–2119, Mar. 2012. [22](#)
- [94] E. Schaan, M. Takada, and D. N. Spergel. Joint likelihood function of cluster counts and n-point correlation functions: Improving their power through including halo sample variance. *Phys. Rev. D*, 90(12):123523, Dec. 2014. [24](#)
- [95] B. M. Schäfer, L. Heisenberg, A. F. Kalovidouris, and D. J. Bacon. On the validity of the Born approximation for weak cosmic flexions. *MNRAS*, 420:455–467, Feb. 2012. [17](#)
- [96] U. Seljak, N. Hamaus, and V. Desjacques. How to Suppress the Shot Noise in Galaxy Surveys. *Physical Review Letters*, 103(9):091303, Aug. 2009. [24](#)
- [97] E. Semboloni, H. Hoekstra, J. Schaye, M. P. van Daalen, and I. G. McCarthy. Quantifying the effect of baryon physics on weak lensing tomography. *MNRAS*, 417:2020–2035, Nov. 2011. [19](#)
- [98] H.-J. Seo and D. J. Eisenstein. Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys. *ApJ*, 598:720–740, Dec. 2003. [20](#)
- [99] S. Seshadri, C. Shapiro, T. Goodsall, J. Fucik, C. M. Hirata, J. Rhodes, B. Rowe, and R. Smith. Initial results from a laboratory emulation of weak gravitational lensing measurements. *Publ. Astron. Soc. Pac.*, 125:1065, 2013. [5](#)
- [100] E. S. Sheldon, D. E. Johnston, R. Scranton, B. P. Koester, T. A. McKay, H. Oyaizu, C. Cunha, M. Lima, H. Lin, J. A. Frieman, R. H. Wechsler, J. Annis, R. Mandelbaum, N. A. Bahcall, and M. Fukugita. Cross-correlation Weak Lensing of SDSS Galaxy Clusters. I. Measurements. *ApJ*, 703:2217–2231, Oct. 2009. [19](#)
- [101] H. Shim, J. Colbert, H. Teplitz, A. Henry, M. Malkan, P. McCarthy, and L. Yan. Global Star Formation Rate Density over  $0.7 \leq z \leq 1.9$ . *ApJ*, 696:785–796, May 2009. [22](#)
- [102] M. Simet and R. Mandelbaum. Background sky obscuration by cluster galaxies as a source of systematic error for weak lensing. *MNRAS*, 449:1259–1269, May 2015. [6](#)
- [103] S. Singh, R. Mandelbaum, and S. More. Intrinsic alignments of SDSS-III BOSS LOWZ sample galaxies. *MNRAS*, 450:2195–2216, June 2015. [19](#)
- [104] D. Spergel, N. Gehrels, C. Baltay, D. Bennett, J. Breckinridge, M. Donahue, A. Dressler, B. S. Gaudi, T. Greene, O. Guyon, C. Hirata, J. Kalirai, N. J. Kasdin, B. Macintosh, W. Moos, S. Perlmutter, M. Postman, B. Rauscher, J. Rhodes, Y. Wang, D. Weinberg, D. Benford, M. Hudson, W.-S. Jeong, Y. Mellier, W. Traub, T. Yamada, P. Capak, J. Colbert, D. Masters, M. Penny, D. Savransky, D. Stern, N. Zimmerman, R. Barry, L. Bartusek, K. Carpenter, E. Cheng, D. Content, F. Dekens, R. Demers, K. Grady, C. Jackson, G. Kuan, J. Kruk, M. Melton, B. Nemati, B. Parvin, I. Poberezhskiy, C. Peddie, J. Ruffa, J. K. Wallace, A. Whipple, E. Wollack, and F. Zhao. Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report. *ArXiv e-prints*, Mar. 2015. [1](#)
- [105] D. Spergel, N. Gehrels, J. Breckinridge, M. Donahue, A. Dressler, B. S. Gaudi, T. Greene, O. Guyon, C. Hirata, J. Kalirai, N. J. Kasdin, W. Moos, S. Perlmutter, M. Postman, B. Rauscher, J. Rhodes, Y. Wang, D. Weinberg, J. Centrella, W. Traub, C. Baltay, J. Colbert, D. Bennett, A. Kiessling, B. Macintosh, J. Merten, M. Mortonson, M. Penny, E. Rozo, D. Savransky, K. Stapelfeldt, Y. Zu, C. Baker, E. Cheng, D. Content, J. Dooley, M. Foote, R. Goullioud, K. Grady, C. Jackson, J. Kruk, M. Levine, M. Melton, C. Peddie, J. Ruffa, and

- S. Shaklan. Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report. *ArXiv e-prints*, May 2013. [1](#)
- [106] M. Takada and B. Jain. Cosmological parameters from lensing power spectrum and bispectrum tomography. *MNRAS*, 348:897–915, Mar. 2004. [24](#)
- [107] A. Tenneti, R. Mandelbaum, T. Di Matteo, A. Kiessling, and N. Khandai. Galaxy shapes and alignments in the MassiveBlack-II hydrodynamic and dark matter-only simulations. *MNRAS*, 453:469–482, Oct. 2015. [19](#)
- [108] H. I. Teplitz, N. R. Collins, J. P. Gardner, R. S. Hill, S. R. Heap, D. J. Lindler, J. Rhodes, and B. E. Woodgate. Emission-Line Galaxies in the Space Telescope Imaging Spectrograph Parallel Survey. I. Observations and Data Analysis. *ApJS*, 146:209–248, June 2003. [22](#)
- [109] J. L. Tinker, E. S. Sheldon, R. H. Wechsler, M. R. Becker, E. Rozo, Y. Zu, D. H. Weinberg, I. Zehavi, M. R. Blanton, M. T. Busha, and B. P. Koester. Cosmological Constraints from Galaxy Clustering and the Mass-to-number Ratio of Galaxy Clusters. *ApJ*, 745:16, Jan. 2012. [19](#)
- [110] J. L. Tinker, D. H. Weinberg, Z. Zheng, and I. Zehavi. On the Mass-to-Light Ratio of Large-Scale Structure. *ApJ*, 631:41–58, Sept. 2005. [19](#)
- [111] M. P. van Daalen, J. Schaye, C. M. Booth, and C. Dalla Vecchia. The effects of galaxy formation on the matter power spectrum: a challenge for precision cosmology. *MNRAS*, 415:3649–3665, Aug. 2011. [19](#)
- [112] M. P. van Daalen, J. Schaye, I. G. McCarthy, C. M. Booth, and C. Dalla Vecchia. The impact of baryonic processes on the two-point correlation functions of galaxies, subhaloes and matter. *MNRAS*, 440:2997–3010, June 2014. [17](#)
- [113] M. Vargas-Magaña, S. Ho, S. Fromenteau, and A. J. Cuesta. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Effect of smoothing of density field on reconstruction and anisotropic BAO analysis. *ArXiv e-prints*, Sept. 2015. [24](#)
- [114] A. von der Linden, M. T. Allen, D. E. Applegate, P. L. Kelly, S. W. Allen, H. Ebeling, P. R. Burchat, D. L. Burke, D. Donovan, R. G. Morris, R. Blandford, T. Erben, and A. Mantz. Weighing the Giants - I. Weak-lensing masses for 51 massive galaxy clusters: project overview, data analysis methods and cluster images. *MNRAS*, 439:2–27, Mar. 2014. [19](#)
- [115] Y. Wang. Differentiating dark energy and modified gravity with galaxy redshift surveys. *JCAP*, 5:21, May 2008. [20](#)
- [116] Y. Wang. *Dark Energy*. Wiley-VCH, 2010. [1](#)
- [117] Y. Wang. Model-independent measurements of cosmic expansion and growth at  $z = 0.57$  using the anisotropic clustering of CMASS galaxies from the Sloan Digital Sky Survey Data Release 9. *MNRAS*, 443:2950–2956, Oct. 2014. [24](#)
- [118] Y. Wang, C.-H. Chuang, and C. M. Hirata. Towards more realistic forecasting of dark energy constraints from galaxy redshift surveys. *MNRAS*, 430:2446–2453, Apr. 2013. [23](#)
- [119] Y. Wang, W. Percival, A. Cimatti, P. Mukherjee, L. Guzzo, C. M. Baugh, C. Carbone, P. Franzetti, B. Garilli, J. E. Geach, C. G. Lacey, E. Majerotto, A. Orsi, P. Rosati, L. Samushia, and G. Zamorani. Designing a space-based galaxy redshift survey to probe dark energy. *MNRAS*, 409:737–749, Dec. 2010. [23](#)

- [120] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo. Observational probes of cosmic acceleration. *Phys. Rep.*, 530:87–255, Sept. 2013. [1](#), [18](#)
- [121] H.-Y. Wu, E. Rozo, and R. H. Wechsler. Annealing a Follow-up Program: Improvement of the Dark Energy Figure of Merit for Optical Galaxy Cluster Surveys. *ApJ*, 713:1207–1218, Apr. 2010. [19](#)
- [122] X. Xu, N. Padmanabhan, D. J. Eisenstein, K. T. Mehta, and A. J. Cuesta. A 2 per cent distance to  $z = 0.35$  by reconstructing baryon acoustic oscillations - II. Fitting techniques. *MNRAS*, 427:2146–2167, Dec. 2012. [24](#)
- [123] J. Yoo and U. Seljak. Joint analysis of gravitational lensing, clustering, and abundance: Toward the unification of large-scale structure analysis. *Phys. Rev. D*, 86(8):083504, Oct. 2012. [18](#)
- [124] J. Yoo, J. L. Tinker, D. H. Weinberg, Z. Zheng, N. Katz, and R. Davé. From Galaxy-Galaxy Lensing to Cosmological Parameters. *ApJ*, 652:26–42, Nov. 2006. [18](#)
- [125] A. R. Zentner, E. Semboloni, S. Dodelson, T. Eifler, E. Krause, and A. P. Hearin. Accounting for baryons in cosmological constraints from cosmic shear. *Phys. Rev. D*, 87(4):043509, Feb. 2013. [17](#)
- [126] F. Zhu, N. Padmanabhan, and M. White. Optimal redshift weighting for baryon acoustic oscillations. *MNRAS*, 451:236–243, July 2015. [24](#)
- [127] Y. Zu and R. Mandelbaum. Mapping stellar content to dark matter halos using galaxy clustering and galaxy-galaxy lensing in the SDSS DR7. *MNRAS*, 454:1161–1191, Dec. 2015. [18](#)
- [128] Y. Zu, D. H. Weinberg, E. Rozo, E. S. Sheldon, J. L. Tinker, and M. R. Becker. Cosmological constraints from the large-scale weak lensing of SDSS MaxBCG clusters. *MNRAS*, 439:1628–1647, Apr. 2014. [19](#)