



# Simulations of the *WFIRST* Supernova Survey and Forecasts of Cosmological Constraints

R. Hounsell<sup>1,2</sup>, D. Scolnic<sup>3</sup>, R. J. Foley<sup>1</sup>, R. Kessler<sup>3</sup><sup>ID</sup>, V. Miranda<sup>4</sup>, A. Avelino<sup>5</sup><sup>ID</sup>, R. C. Bohlin<sup>6</sup><sup>ID</sup>, A. V. Filippenko<sup>7,8,17</sup><sup>ID</sup>, J. Frieman<sup>3,9</sup>, S. W. Jha<sup>10</sup><sup>ID</sup>, P. L. Kelly<sup>11</sup><sup>ID</sup>, R. P. Kirshner<sup>12</sup>, K. Mandel<sup>5,13,14</sup><sup>ID</sup>, A. Rest<sup>6,15</sup>, A. G. Riess<sup>6,15</sup>, S. A. Rodney<sup>16</sup><sup>ID</sup>, and L. Strolger<sup>6</sup>

<sup>1</sup> Department of Astronomy and Astrophysics, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA; [rebekahhounsell@gmail.com](mailto:rebekahhounsell@gmail.com)

<sup>2</sup> Department of Astronomy, University of Illinois Urbana Champaign, 1002 W Green Street, Urbana, IL 61801, USA

<sup>3</sup> Kavli Institute for Cosmological Physics at the University of Chicago, 5620 S Ellis Avenue, Chicago, IL 60637, USA

<sup>4</sup> University of Pennsylvania, Department of Physics & Astronomy, 209 South 33rd Street, Philadelphia, PA 19104-6396, USA

<sup>5</sup> Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

<sup>6</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>7</sup> Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA

<sup>8</sup> Miller Institute for Basic Research in Science, University of California, Berkeley, CA 94720, USA

<sup>9</sup> Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, USA

<sup>10</sup> Department of Physics and Astronomy, Rutgers, the State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA

<sup>11</sup> School of Physics and Astronomy, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA

<sup>12</sup> Gordon and Betty Moore Foundation, 1661 Page Mill Road, Palo Alto, CA 94304, USA

<sup>13</sup> Institute of Astronomy and Kavli Institute for Cosmology, Madingley Road, Cambridge, CB3 0HA, UK

<sup>14</sup> Statistical Laboratory, DPMMS, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WB, UK

<sup>15</sup> Department of Physics and Astronomy, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

<sup>16</sup> Department of Physics and Astronomy, University of South Carolina, 712 Main Street, Columbia, SC 29208, USA

Received 2017 February 7; revised 2018 March 14; accepted 2018 March 23; published 2018 October 25

## Abstract

The *Wide Field InfraRed Survey Telescope (WFIRST)* was the highest-ranked large space-based mission of the 2010 *New Worlds, New Horizons* decadal survey. It is now a NASA mission in formulation with a planned launch in the mid 2020s. A primary mission objective is to precisely constrain the nature of dark energy through multiple probes, including Type Ia supernovae (SN Ia). Here, we present the first realistic simulations of the *WFIRST* SN survey based on current hardware specifications and using open-source tools. We simulate SN light curves and spectra as viewed by the *WFIRST* wide-field channel (WFC) imager and integral field channel (IFC) spectrometer, respectively. We examine 11 survey strategies with different time allocations between the WFC and IFC, two of which are based upon the strategy described by the *WFIRST* Science Definition Team, which measures SN distances exclusively from IFC data. We propagate statistical and, crucially, systematic uncertainties to predict the Dark Energy Task Force figure of merit (FoM) for each strategy. Of the strategies investigated, we find the most successful to be WFC focused. However, further work in constraining systematics is required to fully optimize the use of the IFC. Even without improvements to other cosmological probes, the *WFIRST* SN survey has the potential to increase the FoM by more than an order of magnitude from the current values. Although the survey strategies presented here have not been fully optimized, these initial investigations are an important step in the development of the final hardware design and implementation of the *WFIRST* mission.

**Key words:** dark energy – space vehicles: instruments – supernovae: general – surveys – techniques: imaging spectroscopy

**Supporting material:** machine-readable table

## 1. Introduction

The *Wide-Field InfraRed Space Telescope (WFIRST)* is a NASA mission that will constrain the nature of dark energy through multiple probes. It was the top large space-based mission from *New Worlds, New Horizons*, the most recent U.S. astronomy and astrophysics decadal survey (National Research Council 2010). As its name suggests, *WFIRST* is optimized for near-infrared (NIR) observations, and it possesses a large field of view (FoV). The mission is in formulation at NASA, and several concepts have been suggested so far (Spergel et al. 2015). The current design utilizes a telescope that was donated in 2012 by the National Reconnaissance Office. The aperture of the telescope is the same as that of the *Hubble Space Telescope (HST)*, both having 2.37 m primary mirrors. Two main

instruments are proposed for *WFIRST*: a coronagraph, which will be used for exoplanet and planetary disk studies, and a wide-field instrument, which will be used to probe dark-energy models. The wide-field instrument is itself composed of a wide-field channel (WFC) imager and integral field channel (IFC) spectrometer.

Two major *WFIRST* goals are to measure the cosmological growth of the universe and to probe its geometry on large scales. To achieve these milestones, *WFIRST* will conduct multiple observational programs, one of which is a supernova (SN) survey. Type Ia supernovae (SNe Ia) have played a critical role in the discovery of acceleration in the expansion of the universe (Riess et al. 1998; Perlmutter et al. 1999). Recent analyses using multiple cosmological probes (e.g., Betoule et al. 2014; Planck Collaboration et al. 2016; Alam et al. 2017) are all consistent with a universe that is geometrically flat, filled with dark energy that behaves like a cosmological

<sup>17</sup> Miller Senior Fellow.

constant, and cold dark matter (the  $\Lambda$ CDM model; e.g., Peebles 1984; Efstathiou et al. 1990; Frieman et al. 2008b). There remain, however, theoretical arguments for alternatives to the cosmological constant (e.g., Weinberg 1989; Frieman et al. 2008b), which can serve as additional motivation for a new generation of experiments.

The dark-energy equation of state can be used to distinguish between many alternative explanations for the accelerated expansion of the universe (e.g., see Joyce et al. 2016 for a review of dark energy and modified gravity), and it is parameterized as

$$P = w\rho c^2, \quad (1)$$

where  $P$  and  $\rho$  are the dark-energy pressure and energy density, respectively, and  $w$  is its equation-of-state parameter. In some models, the dark-energy equation of state evolves with time, and one common parameterization (proposed by Chevallier & Polarski 2001; Linder 2003), which we adopt in this work, is

$$w = w_0 + (1 - a)w_a, \quad (2)$$

where  $a = (1 + z)^{-1}$  is the scale factor of the universe,  $w_0$  is the current value of the equation-of-state parameter, and  $w_a$  parameterizes its evolution. For a cosmological constant,  $w_0 \equiv -1$  and  $w_a \equiv 0$ .

Given the importance of measuring  $w$ , the Dark Energy Task Force (DETF; Albrecht et al. 2006) suggested the use of a figure of merit (FoM), defined as the inverse of the area enclosed within the 95% confidence contour in the  $w_0$ – $w_a$  plane, to compare the capabilities of different surveys in constraining the dark-energy equation of state. Current constraints on  $(w_0; w_a)$  are

$$(w_0; w_a) = (-0.91 \pm 0.10; -0.39 \pm 0.34), \quad (3)$$

which correspond to an FoM of 32.6 in Alam et al. (2017; see also Betoule et al. 2014, where the FoM = 31.3). This FoM value includes the use of SNe; without SNe, Alam et al. (2017) obtain FoM = 22.9.

An alternative parameterization of the same linearly evolving  $w_a$  model is expressed as  $w_p$ , and its relation to  $(w_0; w_a)$  and to the FoM described above are defined as

$$w_p = w_0 + (1 - a_p)w_a, \quad (4)$$

$$\text{FoM} \propto [\sigma(w_a)\sigma(w_p)]^{-1}, \quad (5)$$

where  $a_p$  is the pivot value of  $a$  and represents the point at which the uncertainty in  $w_a$ , for a given data model, is minimized (Albrecht et al. 2006). Details of  $w_p$  and its application within SN surveys can be found within Astier et al. (2014). For the purposes of our paper, however, we have chosen to investigate the  $w_0$ – $w_a$  plane only.

Understanding the nature of the largest component of the universe is an important goal, one in which the community has invested significant resources. The DETF identified different “stages” of dark-energy experiments starting with initial studies (Stage 1) and progressing toward Stage 4 surveys in the mid 2020s. Stage 3 experiments are currently underway (e.g., the Dark Energy Survey DES Collaboration 2005)<sup>18</sup> and are expected to increase the FoM by a factor of 3–5 over Stage 2 experiments. WFIRST is a Stage 4 experiment, and it is designed to reach a factor of 10 gain over Stage 2 experiments (i.e.,  $\text{FoM} \geq 320$ ) via

a combination of larger statistical samples and a reduction of systematic uncertainties.

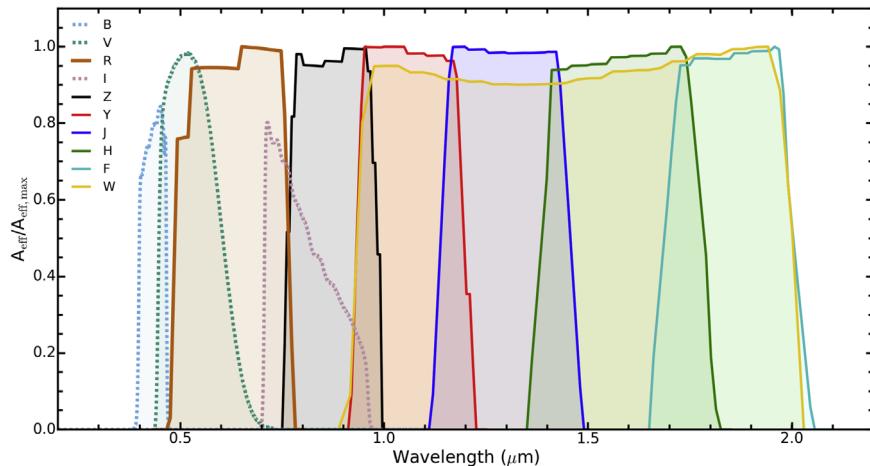
In order for the combined probes from Stage 3 and Stage 4 experiments to reach their projected constraints, SNe Ia are critical. Several surveys have been working to gather data on SNe Ia over a broad range of redshifts. Low-redshift ( $0.01 < z < 0.1$ ) SN Ia data have been obtained by groups/surveys such as the Center for Astrophysics 1–4 (CfA; Riess et al. 1999; Jha et al. 2006; Hicken et al. 2009a, 2009b, 2012), the Carnegie Supernova Project (CSP; Contreras et al. 2010; Folatelli et al. 2010; Stritzinger et al. 2011), the Lick Observatory Supernova Search (LOSS; Ganeshalingam et al. 2010, 2013), and the Foundation SN survey (Foley et al. 2018). SNe Ia at higher redshifts ( $1.0 < z < 1.1$ ) have been examined by surveys including ESSENCE (Miknaitis et al. 2007; Wood-Vasey et al. 2007; Narayan et al. 2016), the SuperNova Legacy Survey (SNLS; Conley et al. 2011; Sullivan et al. 2011), the Sloan Digital Sky Survey (SDSS; Frieman et al. 2008a; Kessler et al. 2009a; Sako et al. 2018), and Pan-STARRS1 (PS1; Rest et al. 2014; Scolnic et al. 2014a). To date, some of the highest redshift ( $z > 1.0$ ) SNe Ia have been observed by the Supernova Cosmology Project (SCP; Suzuki et al. 2012), GOODS (Riess et al. 2007), the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Rodney et al. 2014), and the Dark Energy Survey’s SN program (DES-SN; Bernstein et al. 2012). These surveys form our current state-of-the-art cosmology sample, consisting of over 1000 spectroscopically confirmed SNe Ia, and extending the Hubble diagram out to  $z \approx 2$ .

Each SN Ia light curve must be corrected for both color and shape (“stretch”) in order to standardize the SN brightness and to reduce the Hubble residual dispersion. In addition, redshift-dependent bias corrections are needed, particularly at higher redshifts where fainter SNe are excluded from the sample. Finally, systematic uncertainties must be evaluated and propagated for the inference of cosmological parameters.

Using a simple model for statistical and systematic uncertainties, the WFIRST Science Definition Team (SDT) outlined a baseline six-year mission, including a two-year SN survey, corresponding to six months of “on-sky” time (Spergel et al. 2015). A key SDT assumption is that systematic uncertainties can be characterized in discrete, independent redshift bins ( $\Delta z = 0.1$ ). For many systematics, however, this is an oversimplification, and correlations are found across much broader redshift ranges (e.g., calibration and SN color). The focus of our paper is to expand the discussion of survey strategies, progress toward a more optimized WFIRST SN strategy, and include systematic uncertainties with state-of-the-art analysis tools. Using a covariance matrix approach to systematics, we account for correlations among all redshifts.

A greater understanding of systematic uncertainties and their impact on cosmological constraints is obtained by accurately simulating the survey with sophisticated analysis software called the SuperNova ANAlysis (SNANA; Kessler et al. 2009b) package. SNANA is designed to generate highly realistic simulations of SN surveys and to model the impact of systematic uncertainties. The best SN Ia cosmology constraints (Betoule et al. 2014; Rest et al. 2014; Rodney et al. 2014; Scolnic et al. 2018) are from analyses where SNANA has been used to perform light-curve fitting and predict bias corrections for a variety of surveys including those at low redshift, SDSS, PS1, SNLS, and HST. It is routinely updated with the most current techniques for

<sup>18</sup> See <http://www.darkenergysurvey.org>.



**Figure 1.** *WFIRST* WFC imaging filter bandpass effective areas,  $A_{\text{eff}}$ , divided by the maximum effective area (solid lines) as described by the *WFIRST* Cycle 7 instrument parameter release. Also shown are the *HST* WFC3 filters used for this work (dotted lines). The WFC3 throughputs presented here have been scaled for comparison.

simulations and analysis. Using SNANA in addition to several other open-source tools, we have designed and evaluated various *WFIRST* SN survey strategies, created detailed simulations, and conducted a thorough investigation of uncertainties. Our simulations are the first of their kind for the *WFIRST* mission, and they allow us to predict and compare the potential scientific impact of each strategy. Furthermore, our work acts as a reference for future simulations and provides a guide for the ongoing planning of the *WFIRST* mission.

This paper is structured as follows. We describe *WFIRST* and its instruments in Section 2. Section 3 presents an outline of the SDT SN survey strategy, while Section 4 provides a comprehensive description of how we applied all tools to create the various SN simulations. Additional survey strategies as well as analyses of those strategies examined are presented in Section 5. We explore different assumptions for various systematic uncertainties and outline their impact on the FoM measured by simulated *WFIRST* SN surveys in Section 6. Section 7 compares the simulated survey strategies described in this work, with Section 8 providing a discussion of future considerations for the optimization of the *WFIRST* SN survey. Section 9 presents our conclusions.

## 2. *WFIRST* Hardware

Planned for launch in the mid 2020s, *WFIRST* is expected to be placed into an L2 orbit (1.5 million km away from Earth at the second Lagrange point), where it will reside for the duration of its mission. Analogous to *HST*, *WFIRST* consists of a primary mirror that is 2.37 m in diameter. Light from the primary is reflected to the on-axis secondary mirror, which then feeds into the paths of its various instruments. The design of the telescope is not yet finalized; however, current plans call for both a wide-field instrument (WFI) and a coronagraph.<sup>19</sup> For the purposes of this paper, we focus only on the WFI. When preparing our simulations we used the best-available WFI hardware specifications; these were taken from the 2017 July 30 (Cycle 7) spacecraft and instrument parameter release,<sup>20</sup> and an operational temperature of 260 K is assumed.

<sup>19</sup> For more information on the coronagraph, see <http://wfirstr.gsfc.nasa.gov/observatory.html>.

<sup>20</sup> [https://wfirstr.ipac.caltech.edu/sims/Param\\_db.html](https://wfirstr.ipac.caltech.edu/sims/Param_db.html).

### 2.1. The Wide Field Instrument

The WFI has two optical channels: the first is the Wide Field Channel (WFC), the second the Integral Field Channel (IFC). The WFC possesses an imager and has the ability to perform slitless grism spectroscopy, while the IFC has two small-field integral field units (IFUs). Combined, these instruments will be used to perform the dark-energy survey as well as the microlensing and high-latitude surveys.

*The Wide Field Channel:* In its most simplified form, the optical layout of the WFC consists of three mirrors, two fold mirrors, and a nine-slot filter wheel. Currently, seven of these slots are dedicated to imaging filters, one is for a grism that will provide low-resolution spectra of the full WFC FoV, and the last is a blank position dedicated to dark- and flat-field calibration.

Eighteen  $4\text{k} \times 4\text{k}$  HgCdTe detectors (H4RG-10) will be used by the WFC and will be arranged into a  $6 \times 3$  array to generate an effective FoV<sup>21</sup> of  $0.281 \text{ deg}^2$ .

The seven imaging filters of the WFC are named  $F062$ ,  $Z087$ ,  $Y106$ ,  $J129$ ,  $H158$ ,  $F184$ , and  $W149$  (a very wide filter); hereafter, these filters will be referred to in the text as  $R$ ,  $Z$ ,  $Y$ ,  $J$ ,  $H$ ,  $F$ , and  $W$ . The central wavelengths of these filters are 0.62, 0.87, 1.09, 1.30, 1.60, 1.88, and  $1.40 \mu\text{m}$  (respectively), and combined they cover the  $0.44\text{--}2.0 \mu\text{m}$  range, as illustrated in Figure 1.<sup>22</sup>

The spatial resolution of the imaging component of the WFC is  $\sim 0.11'' \text{ pixel}^{-1}$  with an interpixel capacitance (IPC) of 0.02 in each of the four neighboring pixels. IPC is a form of crosstalk in NIR detectors, in which some of the charge from one pixel will transfer to a neighboring pixel during readout. The effect of IPC is to redistribute charge, which can alter the full width at half-maximum (FWHM) intensity and change the impact of cosmic rays and hot pixels. IPC must therefore be taken into account when calculating the point-spread function (PSF) FWHM of a source for each *WFIRST* filter.

<sup>21</sup> See [https://wfirstr.ipac.caltech.edu/sims/Param\\_db.html?csvfile=WFIRSTParameters\\_v5.0.csv](https://wfirstr.ipac.caltech.edu/sims/Param_db.html?csvfile=WFIRSTParameters_v5.0.csv) for a list of more detailed WFI parameters.

<sup>22</sup> More filter information is provided within [https://wfirstr.gsfc.nasa.gov/science/sdt\\_public/wps/references/instrument/WFIRST-WFI-Transmission\\_160720.xls](https://wfirstr.gsfc.nasa.gov/science/sdt_public/wps/references/instrument/WFIRST-WFI-Transmission_160720.xls), pages 5 through 10.

**Table 1**  
The WFC Imaging Filters

Filter	Central Wavelength ( $\mu\text{m}$ )	Filter FWHM ( $\mu\text{m}$ )	AB Zero-point <sup>a</sup>	PSF FWHM (pixel)
<i>F</i> 062	0.62	0.28	26.99	1.68
Z087	0.87	0.22	26.39	1.69
<i>Y</i> 106	1.09	0.27	26.41	1.86
<i>J</i> 129	1.30	0.32	26.35	2.12
<i>H</i> 158	1.60	0.40	26.41	2.44
<i>F</i> 184	1.88	0.31	25.96	2.71
<i>W</i> 149	1.40	1.1	27.50	2.19

**Note.**

<sup>a</sup> Here, the zero-point is calculated using each filter's effective area; it is equivalent to the magnitude that results in one count per second for an infinite detection aperture.

The gain for the WFC is assumed to be unity. A more detailed description of the WFC filters, including their zero-points and FWHM,<sup>23</sup> can be found in Table 1.

The WFC grism is designed such that it provides spectroscopic coverage within the 1.00–1.89  $\mu\text{m}$  range. It possesses a dispersion of 1.04–1.14 nm pixel<sup>-1</sup>, with a spectral resolving power of  $\lambda/\Delta\lambda \approx 435$ –865 (2 pixels). However, we do not focus on the use of the grism in this paper.

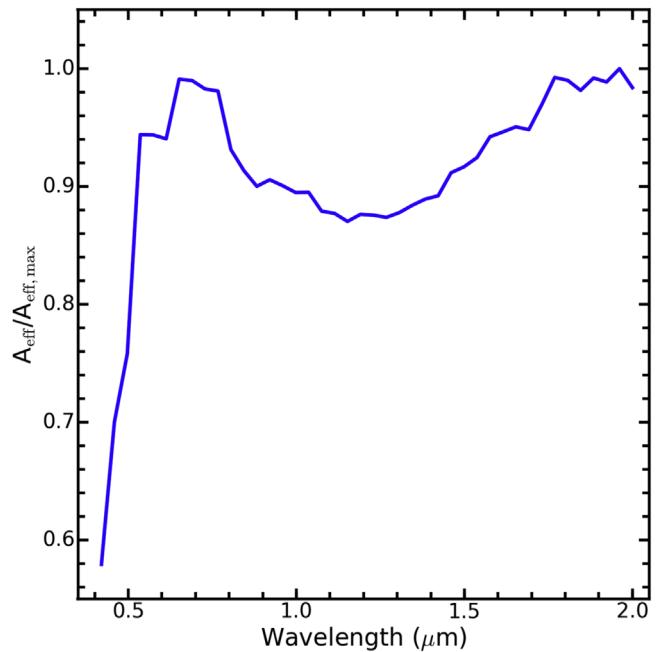
*The Integral Field Channel:* The IFC contains two image slicers that feed a spectrograph. Each image slicer corresponds to a different FoV: the smaller FoV is for the higher spatial-resolution IFC-S, which is designed for SN observations; the larger FoV is for the lower spatial resolution IFC-G, which is designed for galaxy observations (unrelated to the SN survey). The IFC-S has a  $3''00 \times 3''15$  FoV that is composed of  $0''.15$  wide slices, a  $0''.05$  pixel<sup>-1</sup> scale, and a wavelength range of 0.42–2.0  $\mu\text{m}$ . The instrument has a spectral resolution of  $\lambda/\Delta\lambda \approx 70$ –225 (per two-pixel resolution element) and like the WFC, contains H4RG detectors. The IFC-S consists of 352 spectral bins, and the properties of each bin (wavelength range, PSF, noise) are given in Table 14 of Appendix A. The resolution of the IFC-S is based on the design described within Content et al. (2013), but with two recent modifications: the extension of the IFC-S blueward of 6000 to 4200 Å and the use of H4RG detectors, which affect the pixel scale. The PSF FWHM values presented in Table 14 were calculated using an Airy disk approximation for each bin. The wavelength coverage of the IFC-S is illustrated in Figure 2.<sup>24</sup>

### 3. An Outline of the SDT SN Survey Strategy

The WFIRST SDT final report (Spergel et al. 2015) presents a supernova survey strategy in which the imaging component of the WFC is used for SN discovery and the IFC-S for classification and obtaining distances. An outline of this strategy is described below.

<sup>23</sup> As the WFIRST point-spread function (PSF) is non-Gaussian, the PSF FWHM values presented and used for this analysis are derived from the noise-equivalent areas.

<sup>24</sup> See page 13 of [https://wfir.st.gsfc.nasa.gov/science/sdt\\_public/wps/references/instrument/WFIRST-WFI-Transmission\\_160720.xlsx](https://wfir.st.gsfc.nasa.gov/science/sdt_public/wps/references/instrument/WFIRST-WFI-Transmission_160720.xlsx) for more IFC-S information.



**Figure 2.** Relative throughput efficiency ( $A_{\text{eff}}/A_{\text{eff,max}}$ ) vs. wavelength for the WFIRST IFC-S. Wavelengths beyond those displayed have not had their throughputs calculated.

1. The SN survey spans two years in a selected SN field, with a five-day cadence. There are therefore 146 visits to the SN field.
2. Each visit, or epoch of observation, is 30 hr long, including overhead, resulting in a total survey time of 4380 hr (six months).
3. Within each visit, 8 hr of imaging is used exclusively for SN discovery. These data are obtained every five observer-frame days.
4. The imaging is split into three subsurveys (hereafter referred to as tiers) of differing area/depth and using different discovery filters (see Table 2).
5. The remaining 22 hr in each visit are for IFC-S observations, used to classify SN candidates and to synthesize broadband photometry.
6. IFC-S observations are designed to be taken at a cadence of roughly five rest-frame days, with the goal of obtaining spectrophotometry to measure distances.
7. There are three different IFC-S exposure times: typical short exposures, medium classification exposures, and long “deep” exposures. These three exposure times are the first three IFC-S spectra taken for each SN candidate.
8. The short- and medium-exposure spectra are used for initial classification, and if these spectra meet certain criteria (outlined in more detail below), the IFC-S obtains a long exposure through which a final classification is obtained. If classified as an SN Ia, further follow-up observations are initiated.
9. The follow-up observations consist of six short-exposure spectra plus one medium exposure of the host galaxy, taken after the SN has faded, to use as a template.
10. The exposure times for the long- and medium-exposure spectra are on average approximately 3.9 and 1.9 times longer than the short-exposure spectra, respectively.

**Table 2**  
Description of the Three-tier WFC Survey as Outlined in the SDT Report

Survey Tier	Redshift Range	Area (deg <sup>2</sup> )	Discovery Filters	Depth per Exposure <sup>a</sup> (mag)	Total Depth <sup>a,b</sup> (mag)
Shallow	$0.1 \leq z < 0.4$	27.44	<i>Y, J</i>	22.0, 22.0	24.7, 24.7
Medium	$0.4 \leq z < 0.8$	8.96	<i>J, H</i>	24.8, 24.8	27.5, 27.5
Deep	$0.8 \leq z \leq 1.7$	5.04	<i>J, H</i>	26.2, 26.2	28.9, 28.9

**Notes.**<sup>a</sup> Accounts for overhead from slew-and-settle time.<sup>b</sup> Total depth is for co-add over all 146 visits.

11. The total set of observations for any given SN Ia (i.e., excluding the host template) is equivalent to an average of  $\sim 12.8$  short exposures. The exposure times are set by the redshift of the SN.

The SN survey strategy proposed by the SDT report is designed to achieve a relatively flat redshift distribution using a three-tier survey, where each successive tier is deeper and covers less area than the previous tier. The first tier consists of a shallow wide field for SNe with  $z < 0.4$ , over an area of 27.44 deg<sup>2</sup>, using the *Y+J* filters for discovery. The second is a medium tier for SNe with  $0.4 \leq z < 0.8$ , over a moderate 8.96 deg<sup>2</sup> area, using the *J+H* filters. Finally, the last is a deep tier for SNe with  $z \leq 1.7$ , over a small 5.04 deg<sup>2</sup> area, again using the *J+H* filters. Table 3 lists the exposure times for each of the three tiers and the number of spacecraft pointings required to make up their designated areas. The different filter combinations for each tier were chosen in order to probe similar rest-frame wavelengths. However, for the shallow tier, the *Z*-band filter is the only band that covers a rest-frame wavelength range, which is sufficiently modeled for cosmological analysis. One might assume, therefore, that redder wavelengths (i.e.,  $> 7000$  Å in the rest frame) will be accurately trained either with data from *WFIRST* or from precursor data.

Of the 146 planned visits, the discovery search will be implemented in only 132. The remaining survey time will be used for host-galaxy follow-up observations, acquiring a template. The host-galaxy template spectrum is to be taken a year after the peak brightness of the SN, when the relative amount of light from the SN compared to the galaxy is negligible. Thus, in the first year, only 27 of the total 30 hr in each five-day visit will be used, with the remainder deferred to year 2. SNe discovered during the second year will have their galaxy reference spectrum taken in year 3, after the discovery component of the two-year SN survey has concluded.

The spectroscopic observations planned in the SDT report are designed to observe one SN at a time, using the IFC-S. The exposure times were tailored to achieve a signal-to-noise ratio (S/N) high enough to clearly identify key spectral features. The longest exposure times are therefore required for the highest-redshift SNe, i.e.,  $z \approx 1.7$  events. For each SN classified as an SN Ia, a series of 10 spectra is obtained. The first three spectra vary in exposure time and are used not only for obtaining time-critical data on the SN, but also for selection and identification purposes. It is expected that by the third spectrum, core-collapse (CC) SNe are eliminated from the sample (see Section 3.1). A list of exposure times for each SN (excluding the host-galaxy template) per redshift tier is given in Table 4. The exposure times listed here are based on the initial estimates provided by contributing SDT report authors, but adjusted for the number of spectra per SN (10) as specified within the SDT report.

**Table 3**  
WFC Exposure Times ( $t_{\text{exp}}$ ) and Number of Pointings ( $N_p$ ) for Each Filter and Each Tier<sup>a</sup>

Survey Tier	<i>Y</i> band $t_{\text{exp}}$ (s)	<i>J</i> band $t_{\text{exp}}$ (s)	<i>H</i> band $t_{\text{exp}}$ (s)	$N_p$	$t_{\text{tot}}$ (hr)
Shallow	13	13	0	98	3.0
Medium	0	67	67	32	2.0
Deep	0	265	265	18	3.0

**Note.**<sup>a</sup> The exposure times listed here do not include the 42 s slew. However, slew-and-settle times are incorporated into our overall calculations, so that the total observatory time allocated to the SN survey is still six months.

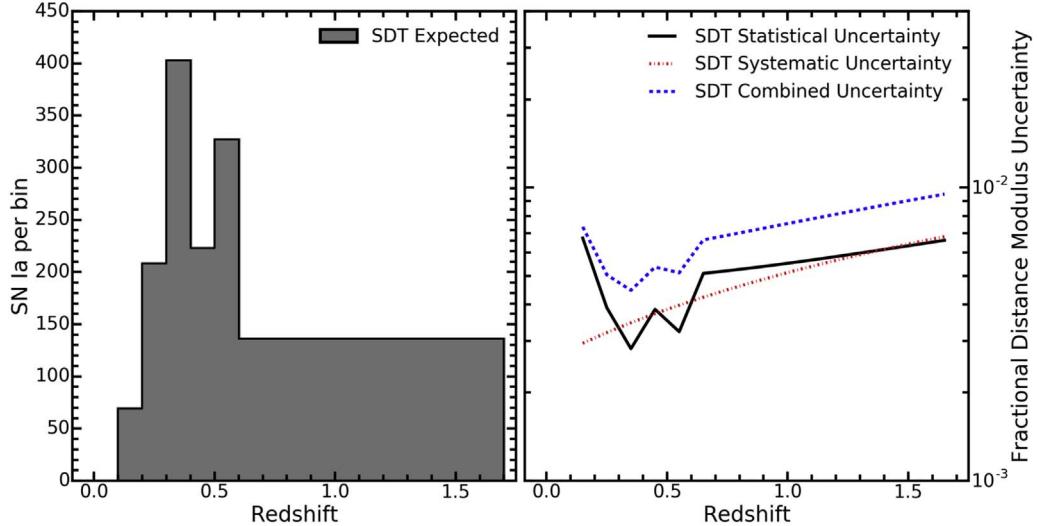
**Table 4**  
Exposure Times per 0.1 Redshift Bin for the *WFIRST* IFC-S Component

Mean Redshift	Short $t_{\text{exp}}$ (s)	Medium $t_{\text{exp}}$ (s)	Long $t_{\text{exp}}$ (s)	Total <sup>a</sup> $t_{\text{tot}}$ (s)
0.15	30.4	47.0	76.0	335.8
0.25	52.1	83.8	143.7	592.2
0.35	80.5	134.1	241.6	939.2
0.45	118.5	205.2	387.5	1422.2
0.55	162.7	291.5	571.8	2002.2
0.65	184.5	337.2	675.7	2304.4
0.75	208.6	386.3	785.1	2631.6
0.85	229.5	428.7	879.0	2914.2
0.95	267.6	508.5	1060.8	3442.5
1.05	319.9	621.8	1325.7	4186.8
1.15	368.9	729.0	1578.0	4889.3
1.25	427.9	862.2	1899.8	5757.3
1.35	493.3	1012.3	2268.4	6733.8
1.45	550.4	1146.9	2604.2	7603.9
1.55	603.5	1274.1	2926.5	8425.1
1.65	629.9	1336.1	3081.0	8826.4

**Note.**<sup>a</sup> Total time observing one SN within a given 0.1 redshift bin (not including the template host-galaxy spectrum).

Both the SDT report and Spergel et al. (2013; an earlier SDT publication) assumed a combined instrumental slew-and-settle time of 42 s. Although these overheads were mentioned by Spergel et al. (2013), they were not incorporated within the SDT's SN strategy, and as such the exposure times and search depths presented for each imaging tier were overestimated. In our simulations, the slew-and-settle overheads are incorporated within each strategy, and we present<sup>25</sup> updated depths in Table 2. Note that this 42 s overhead is a severe underestimate

<sup>25</sup> Derived from ETC calculations, see <https://wfir.st.ipac.caltech.edu/sims/ETC.html>.



**Figure 3.** Left: redshift distribution of the *WFIRST* SN survey presented (and assumed) by the SDT report. Right: fractional statistical (black curve), systematic (red dotted-dashed line), and total (blue dashed line) distance uncertainty per  $\Delta z = 0.1$  bin as assumed in the SDT report.

of the actual value (the most recent estimates of slew-and-settle time from mission HQ are almost double that quoted here), and therefore these overheads remain an uncertain aspect of mission performance.

The total time ( $t_{\text{tot}}$ ) listed per imaging tier in Table 3, including overheads, is therefore calculated as

$$t_{\text{tot}} (\text{s}) = (t_{\text{exp}} + t_{\text{oh}}) \times N_f \times N_p, \quad (6)$$

where  $t_{\text{exp}}$  is the exposure time on the sky in seconds,  $t_{\text{oh}}$  is the 42 s overhead,  $N_f$  is the number of filters used (which for discovery is always 2), and  $N_p$  is the number of pointings.

### 3.1. SDT Detection and Classification

The detection and selection of SNe Ia for follow-up observations as outlined in the SDT report is a complex process, influenced by the cost of single-object follow-up observations with the IFC-S. The process starts with all possible SN candidates, including both SNe Ia and CC SNe, and then progressively removes SNe that do not satisfy certain conditions. The first part of this selection procedure involves a S/N requirement. It is not clear within the SDT report if their S/N is based on image-subtracted data, but in this paper, the S/N includes noise from both the search and template images. Note also that pre-existing spectroscopic redshifts for all host galaxies are assumed by the SDT report, thus enabling the classification procedure outlined. At each stage of the selection process, SN candidates are removed and cannot re-enter. Therefore, each step in the selection process is considered a selection cut, which we list below.

1. *Cut 0:* An object is “detected” if  $S/N \geq 4$  in both imaging discovery bands ( $Y+J$  or  $J+H$ ) within a single epoch. Of these objects, those that have a discovery-epoch color inconsistent with an SN Ia at their host-galaxy redshift are removed. All remaining objects are scheduled for a short-exposure IFC-S spectrum during the next visit to the SN field.
2. *Cut 1:* An object is removed if its flux does not increase between the first and second epochs (in both filters), or if the color at the second epoch is inconsistent with an SN

Ia at its host-galaxy redshift. All remaining objects are scheduled for a medium-exposure IFC-S spectrum.

3. *Cut 2:* An object is removed if its flux does not increase between the second and third epochs (in both filters), or if the color at the third epoch is inconsistent with an SN Ia at its host-galaxy redshift.
4. *Cut 3:* If the medium-exposure spectrum does not resemble that of an SN Ia (see Section 4 for how we determine this), the object is removed. All remaining objects are scheduled for a long-exposure IFC-S spectrum.
5. *Cut 4:* An object that is not identified (again, see Section 4 for how we determine this) as an SN Ia with the long-exposure IFC-S spectrum is removed. The remaining objects are scheduled for additional IFC-S observations and are included in the final cosmology sample.

### 3.2. SDT Statistical and Systematic Uncertainties

The SDT survey strategy is designed such that statistical uncertainties match an assumed systematic uncertainty budget. This means that the assumptions about systematic uncertainties motivate the whole SN survey strategy, as these assumptions dictate the desired sample size, which in turn sets the required discovery rate and redshift distribution. The final distribution of SNe Ia per 0.1 redshift bin, as expected by the SDT report, is shown in Figure 3 (left panel).

The systematic uncertainties presented in the SDT report for the *WFIRST* SN survey follow the description of distance modulus uncertainties used for the SNAP design outlined by Kim et al. (2004; see also Frieman et al. 2003; Perlmutter & Schmidt 2003). In the SDT report, the magnitude of the uncertainties was reduced roughly by a factor of two compared to the SNAP design. The formulation assumes that the systematic uncertainties are uncorrelated on scales larger than  $\Delta z = 0.1$  and can be treated equivalently to statistical uncertainties. The total systematic uncertainty is assumed to increase with redshift, following

$$\sigma_{\text{sys}} = \frac{0.01(1+z)}{1.8} \text{ (mag)}. \quad (7)$$

However, there are known systematics that contradict this assumption. Specifically, uncertainties related to calibration and SN color are correlated across a wide redshift range. The SDT systematic model (Equation (7)) is overly simplistic and not used for our analysis. The SDT functional form for the systematic uncertainty model also drives the broad, flat redshift distribution seen in Figure 3 (right panel).

The SDT report assumes that the distance precision per SN is  $\sigma_{\text{meas}} = 0.08$  mag, and this includes both statistical measurement uncertainties and statistical model uncertainties. This uncertainty is a constant since the SDT strategy adjusts the exposure time for each SN observation based on redshift so that all SNe have approximately the same distance uncertainty. The intrinsic scatter in the corrected SN Ia distances is set to be  $\sigma_{\text{int}} = 0.09$  mag. This value is more optimistic than what is currently measured for optical data, where  $\sigma_{\text{int}} \simeq 0.13$  mag (see Section 7.1 of Kessler & Scolnic 2017). The lensing uncertainty is modeled as  $\sigma_{\text{lens}} = 0.07 \times z$  mag, which is an average of the values derived by Holz & Hughes (2005), Gunnarsson et al. (2006), and Jönsson et al. (2010). The total statistical uncertainty for a given redshift bin is therefore given in the SDT report as

$$\sigma_{\text{stat}} = \frac{(\sigma_{\text{meas}}^2 + \sigma_{\text{int}}^2 + \sigma_{\text{lens}}^2)^{1/2}}{N_{\text{SN}}^{1/2}} \text{ (mag)}, \quad (8)$$

where  $N_{\text{SN}}$  is the number of SNe Ia in a given redshift bin.

The statistical, systematic, and combined uncertainty budgets of the SDT report SN survey are illustrated in Figure 3 (right panel). To be clear, the SDT analysis is not based on SN simulations or light-curve analysis, but instead is based on assumptions about statistical and systematic uncertainties that would arise from such an analysis.

#### 4. Simulation and Analysis Tools

Within this paper, we test the various assumptions made in the SDT report for the SN survey, evaluate its statistical and systematic uncertainty budget, and develop a framework to explore other strategies and optimize parameters for the future *WFIRST* mission. To accomplish this, we simulate and analyze a realistic survey and include the most significant uncertainties. Here we describe software tools that we have used to implement the simulation, apply selection criteria, and determine cosmological constraints used to compute the FoM.

To examine a variety of possible *WFIRST* survey strategies, we used the SNANA simulation package (Kessler et al. 2009b).<sup>26</sup> SNANA has been extensively used for the simulation of SN surveys and analysis of SN samples (see, e.g., Betoule et al. 2014; Scolnic et al. 2014a). The goal of the *WFIRST* simulation is to provide the same fidelity as an ideal image-level simulation by using image properties (zero-points, sky noise, PSFs) rather than the images themselves. As this is a “catalog-level” simulation rather than a pixel-level simulation, we assume that Poisson noise correctly describes the uncertainties from the image subtraction.

To characterize a *WFIRST* SN strategy, we provide SNANA with information about the observatory (e.g., filter/spectrograph properties and noise sources), the survey (e.g., cadence, exposure time, selection requirements), and the physical universe (e.g., SN spectral models, SN rates,

cosmological parameters, lensing assumptions). Each of these components is described below, along with other analysis tools needed to determine the FoM. Our analysis has resulted in several publicly available upgrades to SNANA.

*Imaging filters and spectroscopic bins:* Tables 1 and 14 (in Appendix A) describe the WFC imaging filters and IFC-S wavelength bins used within our simulations. SNANA was originally designed only to simulate broadband SN light curves. In order to simulate the IFC-S, we added a new SNANA module for simulating spectra and “synthetic” broadband filters.

While it may be possible to directly infer distances from SN spectral time series, an examination of that approach is beyond the scope of this paper. Instead, we implement the SDT report’s IFC-S strategy in SNANA by integrating each simulated spectrum into a set of 52 synthetic filters. These synthetic filters were determined by binning together the 352 spectral elements of the IFC-S by a factor of  $\sim 7$  and taking the upper and lower wavelength limits. The SNANA software allows up to 62 broadband filters, 10 of which are used for broadband imaging filters, leaving 52 for the IFC-S synthetic filters (note that there is no limit on spectral binning within SNANA). Once binned, SNANA treats the resulting “synthetic photometry” in a similar manner to any broadband photometry for estimating distances. As the SDT analysis only uses spectral data from the rest-frame optical (3000–8000 Å), we have limited our IFC-S simulations/data accordingly. This choice likely limits the full capability of the IFC-S; however, the various published analyses of IFU data have only probed SNe Ia in the rest-frame optical (e.g., Fakhouri et al. 2015; Saunders et al. 2015). Note, however, that the SDT discovery imaging still makes use of the NIR filters to enable follow-up spectroscopy with the IFC-S.

*Cadence and exposure time:* The cadence of both the WFC and IFC-S components of the SN survey is described in Section 3. The exposure time per imaging tier of the survey is given in Table 3, with IFC-S redshift-dependent times presented in Table 4. The exposure time of the IFC-S within a given 0.1 redshift bin is identical between imaging tiers. Our simulations do not make adjustments to account for the mean SN brightness shifting slightly within a redshift bin (i.e., changes in brightness at  $z = 0.45$  to  $z = 0.46$ , etc.) as it is unlikely that any actual SN survey executed would have specific exposure times for individual objects of interest.

*Sources of noise:* For all simulated SN observations, we include four sources of noise: zodiacal light, thermal background, dark current, and read noise. The contributions from each of these sources are presented in Tables 5, 6, and 14, within Appendix A. Host-galaxy Poisson noise is also included in both the SN-search and template observations, where possible.

The zodiacal light is calculated using a broken power law as described by Aldering (2001). Thermal noise contributions are calculated using code developed by D. Rubin (2018, private communication) under the assumption of a 260 K operating temperature and are comparable to values produced when using the *WFIRST* ETC.<sup>27</sup> The zodiacal and thermal noise for the IFC-S, as a function of wavelength, are presented in Table 14 in Appendix A. The higher resolution of the IFC-S leads

<sup>26</sup> <http://snana.uchicago.edu>

<sup>27</sup> See <https://wfir.ipac.caltech.edu/sims/ETC.html>.

**Table 5**  
WFC Imaging Filters: Sources of Noise

Filter	Zodiacal Noise ( $e^- s^{-1} \text{ pixel}^{-1}$ )	Thermal Noise ( $e^- s^{-1} \text{ pixel}^{-1}$ )
F062	0.44	0
Z087	0.34	0
Y106	0.38	0
J129	0.36	0
H158	0.35	0.005
F184	0.20	0.125
W149	0.97	0.099

**Table 6**  
Read Noise, WFC Imaging Survey

Survey Tier	Read Noise ( $e^- \text{ pixel}^{-1}$ )
Shallow	26.38
Medium	14.53
Deep	8.67

Note. Calculated via Equation (9).

to smaller zodiacal and thermal noise contributions when compared to the WFC.

From Hirata & Penny (2014), we assume a dark current for the WFC to be  $0.015 e^- s^{-1} \text{ pixel}^{-1}$ , and for the IFC-S it is  $0.003 e^- s^{-1} \text{ pixel}^{-1}$  (a conservative estimate based on current measurements of  $0.001 e^- s^{-1} \text{ pixel}^{-1}$ ). The read noise is a function of exposure and readout time and is calculated using a modified version of the expression described by Rauscher et al. (2007). For any given WFC exposure time, the read noise,  $\sigma_{\text{read}}$ , is

$$\sigma_{\text{read}} (e^- s^{-1}) = \sqrt{25 + 4800 \times \frac{(t_{\text{exp}}/t_{\text{read}} - 1)}{(t_{\text{exp}}/t_{\text{read}})} \times \frac{1}{(t_{\text{exp}}/t_{\text{read}} + 1)}}, \quad (9)$$

where  $t_{\text{exp}}$  is the exposure time of the observation in seconds and  $t_{\text{read}}$  is the read time in seconds, which is taken to be 2.825 s.

For each SN, the underlying sky and host-galaxy flux is constant in time, meaning that the associated “template” noise for a supernova is coherent across exposures. The inclusion of this noise source is particularly important to the analysis of IFC-S observations. In the SDT report, each template is planned to be a single medium exposure. As this exposure is not particularly long (and shorter than the long exposures), it adds significant noise to the template-subtracted SN spectrophotometry. On the other hand, this source is negligible for WFC photometry, as imaging templates can be generated from several images, significantly reducing the template noise.

For each WFC simulated SN, we draw an underlying host-galaxy flux from a distribution determined from the high- $z$  *HST* SN survey portion of the CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) program. From the CANDELS SN sample, we determine the host-galaxy surface brightness at the SN position for a  $0''2$  radius aperture in the *F606W*, *F775W*, *F850L*, *F105W*, *F125W*, *F140W*, and *F160W* *HST* filters. We then fit spectral models to the host-galaxy measurements. From

this sample, we determine the expected flux through each *WFIRST* filter as a function of redshift. SNANA has the ability to include host-galaxy noise in simulations for a variety of galaxy profiles and brightnesses. Since we have measured the host-galaxy flux at the SN position, we force the SN position to be at the center of an appropriate-brightness galaxy with a Sérsic profile of index 0.5; see Appendix B for more details.

SNANA includes the impact of host-galaxy Poisson noise when calculating the SN photometry, but it cannot currently simulate the same effect on IFC-S spectrophotometry. Investigations of this noise in WFC simulations show that this is a negligible (<5%) source of uncertainty for these observations.

*Volumetric SN rates:* To accurately determine the number of SNe Ia (and CC SNe) that can be discovered by *WFIRST*, we parameterize the rate as a function of redshift and fit it to rate measurements that extend to  $z = 2.5$  from Rodney et al. (2014) and Graur et al. (2014 and references therein). For SNe Ia, the volumetric rates used are

$$R_{\text{Ia}}(z) = \begin{cases} 2.5 \times (1+z)^{1.5} (10^{-5} \text{ yr}^{-1} \text{ Mpc}^{-3}), & \text{for } z < 1. \\ 9.7 \times (1+z)^{0.5} (10^{-5} \text{ yr}^{-1} \text{ Mpc}^{-3}), & \text{for } 1 < z < 3. \end{cases} \quad (10)$$

For CC SNe, we use the volumetric rate from Strolger et al. (2015; green line, Figure 6, Equation (9)). As the expected detection rate for  $z > 3$  SNe is low, we do not attempt to simulate SNe at those redshifts.

*Spectral model for SNe Ia:* We base all of our SN Ia simulations on the SALT2 spectral model (Guy et al. 2007, 2010). Accurate spectrophotometry can be produced from this model, covering a range of phases and light-curve shapes. The SALT2 model flux ( $F$ ) in the rest frame is parameterized by Guy et al. (2007) as

$$F(t, \lambda) = X_0 [M_0(t, \lambda) + x_1 M_1(t, \lambda)] \times e^{c \cdot \text{CL}(\lambda)}, \quad (11)$$

$$X_0 = 10^{0.4(\alpha x_1 - \beta c - M)}. \quad (12)$$

The two SN-dependent parameters are the light-curve shape parameter,  $x_1$ , which characterizes the brightness as a function of phase, and the color,  $c$ , which describes the slope of an empirically determined color law (CL) that changes the spectral shape but does not depend on phase (typical values of  $x_1$  and  $c$  are between  $\pm 3$  and  $\pm 0.3$ , respectively, as seen in Table 6 of Betoule et al. 2014). The parameter  $M$ , in Equation (12), represents the magnitude for which  $x_1 = c = 0$ . The overall scale,  $X_0$ , depends on the global standardization parameters taken from the JLA analysis (Betoule et al. 2014):  $\alpha = 0.14$  and  $\beta = 3.1$ . The fixed parameters from light-curve training are the spectral surfaces ( $M_0$  and  $M_1$ ), which describe the spectral-energy distribution (SED) versus phase, and the CL, which describes the wavelength dependence. For each epoch, the flux in Equation (11) is redshifted, integrated over the *WFIRST* passband throughput, and dimmed according to the distance modulus based on the cosmology parameters given below.

For our simulations and analysis, the SN Ia spectral model is an extension of the SALT2 model in Betoule et al. (2014). While *WFIRST* will observe SNe in the rest-frame NIR, the fiducial SALT2 model is limited to optical wavelengths below 9200 Å. To extend the model into the NIR up to 25,000 Å, we follow the procedure used to simulate SNe for the CANDELS and the Cluster Lensing And Supernova survey with Hubble (Graur et al. 2014; Rodney et al. 2014; Strolger et al. 2015),

which is implemented with the *SNSEDextend* software package (Pierel et al. 2018).

This SALT2 extension uses a compilation of 118 well-sampled, low- $z$  SNe Ia with both optical and NIR light curves (A. Avelino et al. 2018, in preparation; A. Friedman et al. 2018, in preparation). NIR light-curve data are obtained from nearby SN surveys, principally from CfA IR1-2 (Wood-Vasey et al. 2008; Friedman et al. 2015) and CSP (Contreras et al. 2010; Stritzinger et al. 2011), as well as other sources (see Table 3 of Friedman et al. 2015 and references therein). Corresponding optical photometry comes largely from CfA1-4 (Riess et al. 1999; Jha et al. 2006; Hicken et al. 2009a, 2012), CSP (Contreras et al. 2010; Stritzinger et al. 2011), and LOSS (Ganeshalingam et al. 2010). Each SN light curve in this sample is used to generate a spectrophotometric model by warping the SN Ia spectral template from Hsiao et al. (2007) to match the observed photometric colors at each epoch.

From the resulting set of 118 warped spectral time series models, a median SED is derived for each phase and smoothly joined with the zeroth-order component of the SALT2 model (the  $M_0$  component in Guy et al. 2007). The higher-order SALT2 model components, including variance and covariance terms, are extrapolated using flat-line extensions.<sup>28</sup> This model has not yet been calibrated to produce accurate distance estimates from real data. However, this SALT2 extrapolation is sufficient for producing realistic simulations for the purposes of investigating the *WFIRST* SN survey optimization.

Finally, the SALT2 CL was extended to infrared wavelengths using a modification of the polynomial function from Guy et al. (2010). The polynomial coefficients were set so that the effective CL approximately matches the extinction curve of Cardelli et al. (1989), with  $R_V = 3.1$ .

To model intrinsic scatter, we use the spectral variation model in Kessler et al. (2013) that is derived from the uncertainty model of Guy et al. (2010). This model results in 0.13 mag of scatter to the Hubble residuals:  $\sim 70\%$  of this scatter is from achromatic luminosity variation at all wavelengths and phases, and  $\sim 30\%$  of the scatter is from color variation.

While the SDT report assumes the intrinsic scatter is entirely achromatic, the scatter model used here does not. The population parameters for the color and stretch distributions of our simulations are those derived by Scolnic & Kessler (2016) for the high- $z$  SN sample.

*Spectral model for CC SNe:* The CC SED models used in our simulations are described by Kessler & Scolnic (2017) and Kessler et al. (2010), and were generated from a combination of SDSS (Sako et al. 2018) and CSP (Hamuy et al. 2006) light-curve data using  $u$ ,  $g$ ,  $r$ ,  $i$ ,  $z$  filters. These optical CC templates have been extended into the NIR by warping a CC SED model<sup>29</sup> to match the  $V - H$  and  $V - K$  colors in Bianco et al. (2014). The synthetic colors are  $V - H = V - K = 1$  for Types Ib and II, and  $V - H = V - K = 1.5$  for Type Ibc (see Filippenko 1997 for a review of SN spectral classification).

*Selection requirements (cuts):* Within the SDT report, a supernova (both Ia and CC) is detected if it has an observation with  $S/N \geq 4$  in both of the discovery filters ( $Y+J$  or  $J+H$ ) within the same epoch. To reduce CPU time without loss of accuracy, we simulate a trigger that requires  $S/N \geq 3$  in both

discovery bands on the same epoch (a trigger of  $S/N \geq 3$  rather than 4 was chosen in order to prevent the loss of SNe close to the detection limit).

After the simulation has generated light curves satisfying the two-band trigger, we apply the photometric selection criteria defined in Section 3.1. Spectra of the objects that successfully pass these criteria are analyzed via a modified, NIR-enabled version of the Supernova Identification (SNID; Blondin & Tonry 2007) package. SNID compares each input SN spectrum to a library of template spectra and determines how closely template spectra match the input. In our SN classification analysis, we define a “good SN Ia” if 80% of the matches and the top match are an SN Ia at the correct redshift, and if the SN is discovered roughly 7–12 days before peak. This SNID spectral analysis is used to implement the spectroscopic cuts described in Section 3.1.

For imaging-only strategies, the selection criteria for the final sample occur only in the final analysis; no choices are made during the survey itself. First, we require that each SN have at least one epoch with  $S/N \geq 10$  and at least two epochs with  $S/N \geq 5$ . Next, we require that the light-curve parameters of each SN fall within a “well-trained” range of color and stretch values such as those defined by Betoule et al. (2014 and references therein), i.e.,  $-3 < x_1 < 3$  and  $-0.3 < c < 0.3$ . Note that no cut on distance uncertainty is applied.

*Lensing:* We simulate distance dispersion from line-of-sight gravitational lensing. For SNe at  $z > 1.4$ , we use the log-normal approximation in Marra et al. (2013). For  $0.4 < z < 1.4$ , lensing is computed from a 900 deg<sup>2</sup> patch of the MICECATv1<sup>30</sup> simulation. For  $z < 0.4$ , the distance dispersion at  $z = 0.4$  is scaled by  $z/0.4$ . The resulting root-mean square (rms) dispersion is  $0.04 \times z$ , which is on the low side of predictions (e.g., Jönsson et al. 2010). We therefore scale the dispersion by  $\sim 1.4$  to achieve the average predicted dispersion of  $0.055 \times z$ .

*Galactic extinction:* Since the SN fields have not been chosen, we assume that each field will have a low value of  $E(B - V) = 0.015$  mag.

*Low-redshift sample:* To provide an anchor for the SN Ia Hubble diagram, we include 800 simulated SNe Ia with  $z < 0.1$  from a source other than *WFIRST*, which we model as having the characteristics of the Foundation SN survey (Foley et al. 2018). The Foundation survey uses the PS1 telescope and observes low- $z$  ( $0.01 < z < 0.1$ ) SNe in *griz* every five days with typical distance modulus uncertainties due to measurement error  $< 0.1$  mag. A similar external low- $z$  SN Ia sample is a requirement specified in the SDT report.

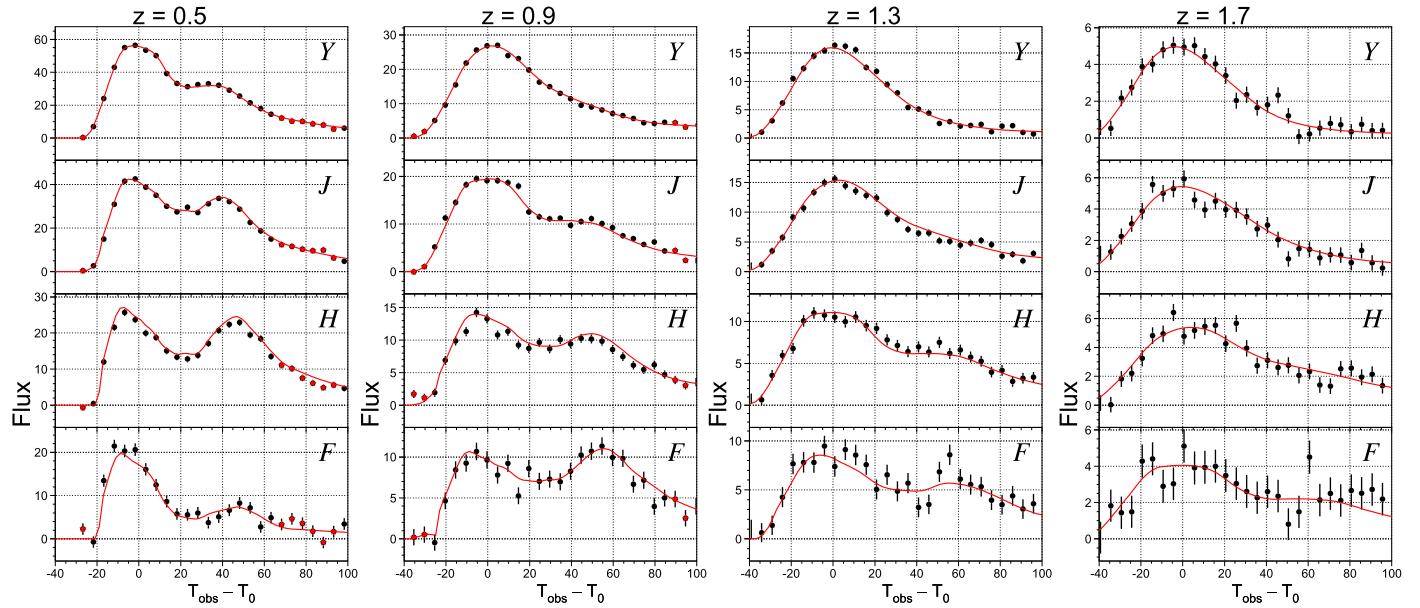
*Cosmology parameters:* Distance moduli are generated with the following  $w$ CDM model parameters:  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $w_0 = -1$ , and  $w_a = 0$ .

#### 4.1. Cosmology Analysis

Here we describe the general analysis strategy for measuring cosmological parameters and FoM. After the simulation, selection cuts are applied (Section 5), and each light curve is fit with the SALT2 model to determine three parameters: stretch ( $x_1$ ), color ( $c$ ), and amplitude ( $x_0$ ). In the limit of no intrinsic scatter and no measurement noise,  $x_0 = X_0 10^{-0.4\mu}$ . The distance modulus for each event is obtained from the Tripp

<sup>28</sup> For more details, see <http://github.com/srodney/wfirst>.

<sup>29</sup> See [https://c3.lbl.gov/nugent/nugent\\_templates.html](https://c3.lbl.gov/nugent/nugent_templates.html).



**Figure 4.** Example WFIRST broadband (YJHF) simulated light curves (black circles) and best-fit light curves (smooth curve) for SNe at redshifts 0.5, 0.9, 1.3, and 1.7. Magnitudes are  $(27.5 - 2.5) \times \log_{10}(\text{Flux})$ ; e.g.,  $m = 25.0$  mag for Flux = 10. These data are generated using the medium-imaging exposure time of 67 s. The red stars seen in some of the plots indicate a point that is excluded from the SNANA fit.

(1998) formulation,

$$\mu = m_B - M + \alpha \cdot x_1 - \beta \cdot c, \quad (13)$$

where  $m_B \equiv -2.5 \log_{10}(x_0)$ .  $M$ ,  $\alpha$ , and  $\beta$  are global nuisance parameters determined from a fit to minimize the Hubble residuals for an entire sample. Since  $m_B$  and  $M$  are degenerate,  $M$  is often quoted to be around  $-19.35$  mag with a corresponding  $m_B$  offset such that  $m_B$  appears to be a Bessel  $B$ -band magnitude. Here we leave out the  $m_B$  offset to make it clear that  $m_B$  is related to the amplitude ( $x_0$ ) and is not related to the  $B$  band.

For the WFC, the light-curve fit includes the imaging filters. For the IFC-S, the light-curve fit includes the 52 synthetic filters. Typical WFC light-curve fits are shown in Figure 4 for a range of redshifts.

We use the “BEAMS with Bias Correction” (BBC) method (Kessler & Scolnic 2017) to determine distances from the Tripp equation (Equation (13)). BBC determines the bias on each fitted parameter ( $c$ ,  $x_1$ ,  $m_B$ ) for each event and minimizes the Hubble residuals in a global fit to determine  $\alpha$ ,  $\beta$ , a scale parameter ( $S_{CC}$ ) for the CC probabilities, and an average distance modulus in approximately 40 log-spaced redshift bins. The output of BBC is a redshift-binned Hubble diagram that is corrected for biases from selection effects and from CC contamination. The BBC distance modulus uncertainties include statistical uncertainties on the fitted parameters ( $c, x_1, m_B$ ), Gaussian lensing scatter ( $\sigma_\mu = 0.055 \times z$ ), peculiar velocities ( $\sigma_v = 150 \text{ km s}^{-1}$ ), and intrinsic scatter. In this analysis we have not trained a photometric classifier to determine CC probabilities, and therefore we do not include the CC SN likelihood but fix  $S_{CC} = 0$ . This makes our estimate of the systematic uncertainty from CC contamination conservative because the analysis does not take advantage of the BEAMS approach to account for contaminants.

Following the SDT report, we often plot the “fractional distance uncertainty (FDU),” defined as

$$\text{FDU} \equiv 10^{0.2\sigma_\mu} - 1, \quad (14)$$

where  $\sigma_\mu$  is the BBC-fitted uncertainty in the distance modulus.

For each systematic uncertainty (Section 6), the SALT2 light-curve fits and BBC fit are repeated. The resultant set of BBC Hubble diagrams are used to compute a total covariance matrix ( $\mathcal{C}_{\text{tot}}$ ) that includes both statistical and systematic uncertainties. Compared to using individual SN distances, using redshift-binned distances greatly reduces the required computing time for constructing  $\mathcal{C}_{\text{tot}}$ .

The redshift-binned distance moduli, their respective uncertainties, and  $\mathcal{C}_{\text{tot}}$  are passed to CosmoMC (Lewis 2013) to determine cosmological constraints. FoMs corresponding to the inverse area of the 95% confidence contours in the  $w_0-w_a$  space (Albrecht et al. 2006) are calculated. For each FoM determination, we assume a flat universe and marginalize over  $H_0$  and  $\Omega_M$ . Furthermore, we include prior constraints from both baryon acoustic oscillation (BAO; Anderson et al. 2014) and cosmic microwave background (CMB; Planck Collaboration et al. 2016) data sets.

When assessing the effect of each individual systematic uncertainty, a modified version of CosmoMC is used to reduce the computational complexity of determining the FoM. This version of CosmoMC, which we call “CosmoMC\*,” encodes the CMB information using the compressed Gaussian likelihood presented by the Planck Collaboration et al. (2016, see their Table 4; the version that does not marginalize over  $A_L$ ) and only accounts for the geometric effects of dark energy. Therefore, CosmoMC\* fixes  $\tau$ , the reionization optical depth, and  $\log(A)$  (equivalent to  $\ln(10^{10}A_s)$ , where  $A_s$  is the inflation power spectrum amplitude). These changes significantly reduce the time to compute the FoM. We have done extensive checks to ensure that the CosmoMC\* model produces accurate results relative to the original version that uses the full set of Planck

**Table 7**  
Simulated Strategies Investigated for the *WFIRST* SN Survey, Including the Strategy Suggested within the SDT Report

Name	Redshift Range			Filter Set Used			Area (deg <sup>2</sup> )			Number of SNe Ia Selected		
	Shallow	Medium	Deep	Shallow	Medium	Deep	Shallow	Medium	Deep	Shallow	Medium	Deep
SDT	0.10–0.39	0.40–0.79	0.80–1.70	IFC-S, YJ	IFC-S, JH	IFC-S, JH	27.44	8.96	5.04	27	300	1181
SDT*	0.10–0.39	0.40–0.79	0.80–1.70	IFC-S, YJ	IFC-S, JH	IFC-S, JH	27.44	8.96	5.04	149	598	1221
SDT* Highz	...	0.10–0.79	0.80–1.70	...	IFC-S, JH	IFC-S, JH	...	22.80	5.04	...	1214	1217
SDT Imaging	0.01–2.99	0.01–2.99	0.01–2.99	YJ	JH	JH	27.44	8.96	5.04	1	546	3046
Imaging:Allz	0.01–2.99	0.01–2.99	0.01–2.99	RZYJ	RZYJ	YJHF	48.82	19.75	8.87	1225	5723	6640
Imaging:Lowz	0.01–2.99	0.01–2.99	...	YJ	JH	...	142.30	66.91	...	6	4560	...
Imaging:Lowz*	0.01–2.99	0.01–2.99	...	RZYJ	RZYJ	...	73.57	32.24	...	1828	9396	...
Imaging:Lowz+	0.01–2.99	0.01–2.99	...	RZYJHF	RZYJHF	...	50.66	20.68	...	1237	5990	...
Imaging:Lowz-Blue	0.01–2.99	0.01–2.99	...	BVRJYJ	BVRJYJ	...	50.66	20.68	...	1169	5644	...
Imaging:Highz*	...	0.01–2.99	0.01–2.99	...	RZYJ	YJHF	...	32.06	13.24	...	9354	9640
Imaging:Highz+	...	0.01–2.99	0.01–2.99	...	RZYJHF	RZYJHF	...	20.50	9.14	...	5965	6759

likelihoods. Fluctuations of a few percent in the FoM value are expected to arise from the sampling variance of the finite Markov Chain Monte Carlo (*MCMC*).

## 5. Simulated Strategies

Here we describe simulations of several different strategies for the *WFIRST* SN survey. The survey variations examined here are summarized in Table 7. This table lists the strategy names, filters, imaging tiers, areas, and the resultant number of simulated SNe Ia.

We include strategies that use both the WFC imager and the IFC-S spectrograph (the SDT, SDT\*, and SDT\* Highz strategies) as well as strategies that employ imaging exclusively (the Imaging, Imaging:Lowz, Imaging:Highz strategies).

For each strategy, we account for the 42 s slew-and-settle time per exposure, and we satisfy the constraint of six-month total observing time.

For the imaging component of the survey, the exposure time per tier, filter zero-points, and sources of noise for each filter are specified in Tables 1, 3, 5, and 6. When the IFC-S is used, its wavelength range, redshift-dependent exposure times, and sources of noise remain set to the values given in Tables 14 and 4.

If an instrument (i.e., the IFC-S), tier (shallow or deep), or filter within a survey simulation is removed or added, the areas (listed in Table 2) of the remaining tiers are scaled evenly (except in the SDT\* Highz and SDT Imaging cases; see Section 5.2) to account for the loss or gain of time.

Note that we have not changed the cadence, depth of a given tier, IFC-S strategy (epochs and number of SNe), or *WFIRST* filter bandpass for any strategy outlined within this paper. Such investigations/optimizations will be the focus of future papers.

For strategies that only have an imaging component, we consider the impact of adding bluer filters that are not in the current *WFIRST* design. We do not consider filters redder than the F band. For any given strategy, we choose no more than six filters in total.

For simplicity, we assume that the additional filters are similar to those from *HST*'s WFC3, and as such, we have used their throughputs and taken the average AB magnitudes of the two WFC3 chips to be our zero-points (see Table 8). The FWHM values for these filters are calculated in part via the use of the WebbPSF for *WFIRST*<sup>31</sup> tool. This tool allows the user

to input appropriate SN spectra and account for wavefront aberrations, in order to calculate binned and unbinned PSF data. WebbPSF, however, is not designed for filters bluer than the Z band. We therefore modified this tool to calculate bluer wavefront aberrations via the extrapolation and application of higher-order Zernike coefficients. Pixelation is then applied to these results along with an IPC effect of the order of ∼2%. As PSF FWHM values change slightly between each tier, the resultant average values are presented in Table 8.

In the current SDT strategy, a set number of SNe in the  $0.1 < z \leq 1.7$  range are followed-up with the IFC-S (2726 SNe). For imaging-only strategies, there is no need to fix the number of SNe or the redshift range. We therefore allow the redshift range to extend from 0.01 to 2.99. However, additional selection criteria as described in Section 4 are implemented, and when combined with typical cuts on the color and stretch of the SN light curves, the photometric classification purity is >99%.

This small CC SN contamination of the SN Ia sample is included as a systematic uncertainty within our work. Note that host-galaxy redshifts in an imaging-only survey could be collected after the *WFIRST* survey is completed since they are not needed to define the imaging sequence (as is the case for the SDT survey).

The design of each survey strategy is discussed below.

### 5.1. The SDT and SDT\* Strategies

Here we present the simulated SDT survey strategy (see Section 3) along with a slightly modified version of this strategy (called SDT\*), in which the efficiency of SN Ia detection and selection has been significantly improved. These strategies use both WFI channels: the WFC imager and IFC-S.

The number of generated SNe is set by the volumetric rates, survey area, depth, and duration; the numbers are reported in Table 9 and do not include selection requirements. Within the appropriate redshift ranges, a total of 21,094 SNe are generated: 3608 are SNe Ia, and the remaining 17,486 are CC SNe. The initial SDT S/N requirement (Section 3.1) reduces the total to 6640 “detectable” events (3514 of which are SNe Ia; see Table 10). For these detectable events, 2621 pass the photometric cuts specified within the SDT report (see the list given in Section 3.1). A breakdown of the number of SNe passing each cut is given in Table 10.

<sup>31</sup> <http://webbpsf.readthedocs.io/en/stable/wfirst.html>

**Table 8**  
HST WFC3 Filters Included within WFIRST SN Simulations

Filter	Zero-point (AB Mag)	PSF FWHM (pixel)
<i>F425W (B)</i>	24.75	1.62
<i>F555W (V)</i>	25.72	1.62
<i>F814W (I)</i>	25.03	1.67

**Table 9**  
Number of SNe Generated per SDT Survey Tier

Survey Tier	Redshift Range	Number of SNe Ia	Number of CC SNe	Total SNe per Tier
Shallow	$0.1 \leq z < 0.4$	520	2437	2957
Medium	$0.4 \leq z < 0.8$	932	5080	6012
Deep	$0.8 \leq z \leq 1.7$	2156	9969	12125
SN Total:		3608	17,486	21,094

Our analysis implements the photometric cuts by defining a range of acceptable color and rise values. Details on how these ranges were defined are provided in Appendix C.

A short- and medium-exposure spectrum is obtained for each SN that satisfies photometric Cuts 0 and 1, respectively. A long-exposure spectrum is obtained once the SN passes both the photometric Cut 2 and spectroscopic Cut 3. Within SNID (Blondin & Tonry 2007, see references therein for a list of spectra used), each spectrum is compared to a library of real SN spectra. The number of SNe passing these additional spectroscopic selection criteria (see Section 3.1) are reported in Table 11. Analysis within SNANA reduces the sample to 1957 SNe Ia, i.e., 56% of the total number of SNe Ia detected (3514 SNe Ia). As is done with all current cosmological analyses, we then apply additional color and light-curve shape requirements, which further reduces the number of SNe in the cosmological sample.

While the number of SNe observable with IFC-S-dominated surveys is limited by the observing time required to reach the desired S/N in each spectrum, our work has shown that the original SDT-defined selection effects are inefficient in obtaining the 2726 SNe Ia desired by the report. In addition, the distribution of SNe Ia obtained in our simulations does not match that of the SDT, which results in larger fractional distance uncertainties at low  $z$  ( $z \leq 0.6$ ). The dependence of the SDT strategy on a long-exposure spectrum for final classification is also inefficient in that for tens of SNe (see difference in SNe between Cuts 3 and 4 of Table 11), this spectrum indicates that they are not SNe Ia, resulting in several hours of exposure time spent on contamination.

When considering the purity of the SN Ia sample, the photometric selection criteria (Cuts 0, 1, and 2) result in a purity of  $\sim 79\%$  (2080 of the 2621 SNe within Cut 2 are SNe Ia). Of the  $\sim 21\%$  CC SN contaminants,  $\sim 67\%$  are SNe Ib/c and  $\sim 33\%$  are SNe II. The SNe Ib/c that make up the majority of the contaminants are also the objects that are most spectroscopically similar to SNe Ia, and therefore the most difficult to remove with low-S/N spectra. Example spectra of a Type Ia SN that passes all cuts, a Type Ia SN that is excluded based on its long-exposure spectrum, and a CC SN

**Table 10**  
Number of SNe that Satisfy the Photometric Cuts Defined in the SDT Report<sup>a</sup>

Cut	Shallow		Medium		Deep		Total
	Ia	CC	Ia	CC	Ia	CC	
0	426	175	932	672	2156	2279	6640
1	159	67	718	202	2073	782	4001
2	27	28	305	110	1748	403	2621

**Note.**

<sup>a</sup> Photometric cuts are defined in Section 3.1.

**Table 11**

Number of SNe that Pass the Photometric and Spectroscopic Cuts Defined by the SDT Report, and the Number of SNe Ia within this Sample which then Undergo Further Cuts based on Color and Stretch Parameters

Cut	Number of SNe	Number of SN Ia Analyzed
3	2109	1958
4	2013	1957

**Note.** SNe that pass these final cuts are then used within our analysis.

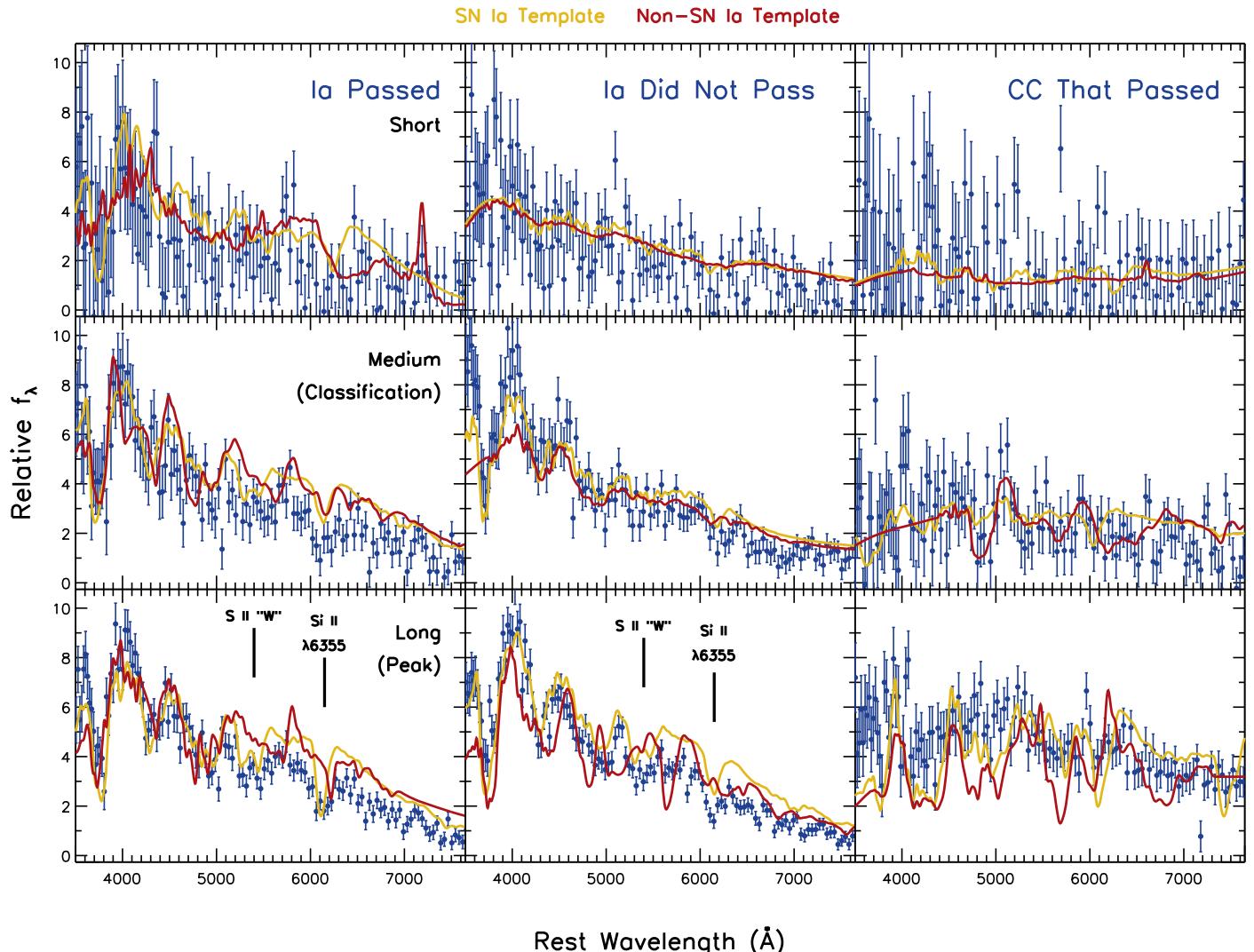
(SN Ic) that passes all cuts and is included in the cosmology sample are illustrated in Figure 5. This figure demonstrates the difficulty of classification using the SDT report strategy. The middle panel of Figure 5 illustrates an SN Ia with spectral features that cannot be identified given its S/N and resolution within each of the different IFC-S exposures. In particular, the sulfur “W,” which the SDT report uses as a clear example of an SN Ia feature, is not detected in the long-exposure SN Ia spectrum (bottom row, center column of of Figure 5), and thus this object is rejected.

All spectral classifications are conducted using SNID, and a supernova is classified as an SN Ia if 80% of the matches and the top match are SNe Ia at the correct redshift (see Section 4 for further details). The SDT report does not indicate any use of the photometric light-curve data beyond that of the first three epochs after detection. If all epochs for a given SN light curve were used in the classification, and S/N cuts on the light curves are applied (as discussed below), we would likely reduce CC SN contamination to a negligible level (see Section 6.2).

The current number of correct spectral classifications for the SDT strategy is likely optimistic. While correlated template sky noise is included in the simulations, we have optimistically ignored flux and noise from unsubtracted host-galaxy light that will be there as a consequence of not having a galaxy template at the time of classification. Even if a spectrum of the host galaxy does exist (e.g., from a ground-based spectrograph), the exact galaxy SED at the position of the SN will not be accurately measured.

The SN Ia efficiency for a given selection cut (see Section 3.1) is shown in Figure 6. Efficiency is defined as the number of SNe correctly identified as Type Ia and passing each cut, divided by the number of SNe Ia that pass Cut 0. This calculation is per 0.1 redshift bin. SNe that pass Cut 4 have short-, medium-, and deep-exposure IFC-S spectra, in addition to the six short follow-up spectra.

The efficiency is low at particular redshifts,  $0.3 < z \leq 0.4$  and  $0.7 < z \leq 0.8$ . This efficiency gap is partially the result of



**Figure 5.** Simulated rest-frame WFIRST IFC-S spectra of  $z = 1$  SNe. The left panels correspond to a Type Ia SN that passes all cuts and for which full follow-up observations would be obtained. The middle panels correspond to an SN Ia that is not identified as a Type Ia SN based on its long-exposure spectrum and is thus rejected. The right panels correspond to a CC SN that passes all requirements, including photometric cuts, and would undergo the full set of follow-up observations. The top, middle, and bottom rows correspond to short-, medium-, and long-exposure spectra, respectively, for each SN. The WFIRST spectra are plotted as blue points with error bars. Note the changing resolution with wavelength. The best-matching SN Ia and non-SN Ia spectrum are plotted as gold and red, respectively.

the survey design producing insufficient SN discoveries at the high- $z$  end of each tier. However, photometric selection criteria that require the color to be consistent with an SN Ia at their host-galaxy redshift, and that the flux rise between epochs, are the main contributors to the low efficiency. For the shallow imaging tier, which covers  $0.1 \leq z < 0.4$ , these criteria are problematic due to the tier’s short 13 s exposure, which results in noisy light curves and an undetectable rise value.

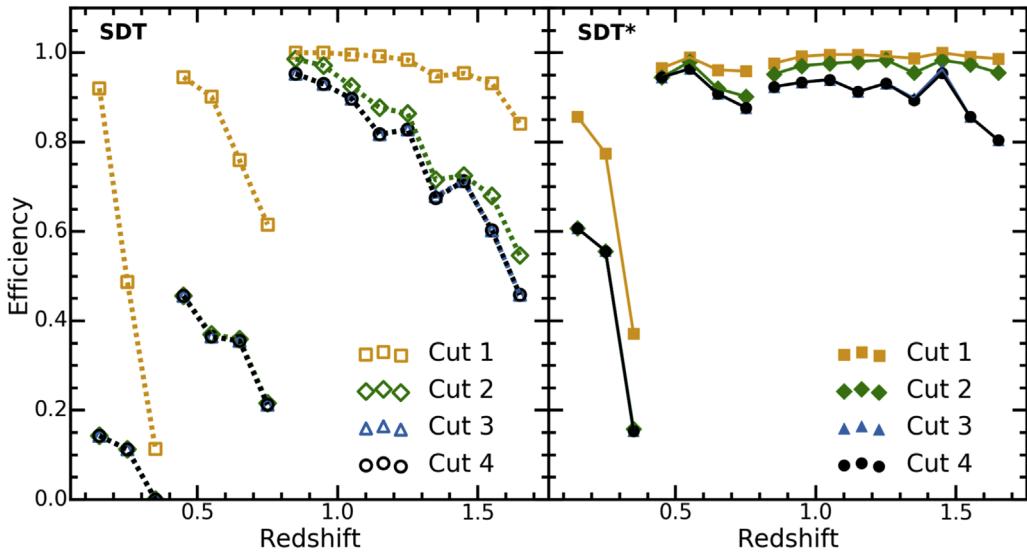
The large reduction in the number of SNe between Cuts 1 and 2 is also caused by the required increase in flux between epochs. In many cases, statistical noise causes a candidate to appear to fade between two successive epochs. To reduce this selection artifact, this criterion is loosened via the iterative examination of a range of measured rise values (including negative values) for the simulated SNe Ia as a function of redshift for each tier, and a looser cut is applied.

The requirements on the discovery colors ( $Y - J$  for shallow,  $J - H$  for medium and deep) are also tightened, excluding some

of the most extreme SNe Ia and significantly reducing the number of CC SNe at each step. The effect of these improved selection criteria is to increase the SN Ia acceptance from 56% to  $\sim 81\%$  (2909 SNe Ia make it to the final sample) and to decrease the number of misclassified CC SNe, all with minimal SN Ia losses. Hereafter, we refer to this sample as SDT\*, a simulated SDT strategy where the selection criteria have been modified.

Using the results of both the SDT and SDT\* selection procedure, the efficiency of each strategy is shown in Figure 6. The SDT methodology results in a rapidly falling efficiency at high  $z$ , while the SDT\* strategy has a relatively flat efficiency versus redshift.

Spectroscopic classification for the SDT\* strategy, however, suffers from the same issues as the SDT strategy, reducing the efficiency to  $\sim 82\%$  (average of efficiency measurement for each 0.1 redshift bin). While we have not yet examined potential biases related to the spectroscopic selection, previous



**Figure 6.** SDT and SDT\* (left and right panels, respectively) SN Ia selection efficiency as a function of redshift. Efficiency is defined as the number of SNe correctly identified as Type Ia within each cut, divided by the number of SNe Ia that pass Cut 0 (see Section 3.1). The efficiency is calculated per 0.1 redshift bin. The gold squares, green diamonds, blue triangles, and black circles represent the efficiency of SNe Ia that pass Cuts 1, 2, 3, and 4 respectively. Lines connect data from the same tier of the survey. The large drop in efficiency from the first to the second cut at  $z < 0.8$  for the SDT strategy is caused primarily by the SDT requirement that the SN flux rises between each epoch. Comparison of the two strategies suggests that the looser selection criteria of the SDT\* strategy is significantly more efficient than that of the SDT.

experience with spectroscopically confirmed SN samples shows that this selection will introduce a distance bias that must be corrected. The resultant combined photometric and spectroscopic selection procedure has a  $\sim 99\%$  purity, which will further increase when considering full light curves and all spectral data.

To accurately match the SDT description of their survey strategy, we select SN redshifts only from their corresponding tiers. This is particularly important for  $z < 0.4$ , where the shallow tier is conducted in the  $Y$  and  $J$  filters, while deeper tiers use the  $J$  and  $H$  filters. To be specific, SNe with  $z < 0.4$ ,  $0.4 \leq z < 0.8$ , and  $0.8 \leq z \leq 1.7$  are selected exclusively from the shallow, medium, and deep tiers, respectively. The overall efficiency is not strongly affected by this decision since the IFC-S exposure time is only a function of redshift and not, for instance, brightness.

The choice to only select SNe within their corresponding redshift tiers, however, does reduce the number of SNe available for follow-up observations. For the SDT\* simulation, the final redshift distribution (Figure 7(a)) has only 475 SNe, compared to the 1230  $z < 0.6$  SNe Ia expected in the SDT report (i.e., only 39%; see Figure 7(a)). This deficit is caused primarily by the low S/N for objects in the shallow survey tier and the choice not to obtain follow-up observations of low-z SNe from the medium and deep tiers.

If SNe were to be selected from any of the three imaging tiers regardless of their redshift, then the number of SNe per redshift range would likely increase (see results of SDT\* Highz, Section 5.2). The total number of SNe discovered, however, will be similar to that in the SDT report, and additional selection criteria would likely reduce the final number below that desired.

For the simulated SDT and SDT\* samples, the statistical uncertainties on the fractional distance (Equation (14)) as a function of redshift are presented in Figure 7(f).

The significant disagreement between statistical uncertainties forecast in the SDT report (henceforth SDT-required) and

SDT\* surveys at  $z < 0.6$  is due to the lack of low- $z$  SNe Ia in the SDT\* sample.

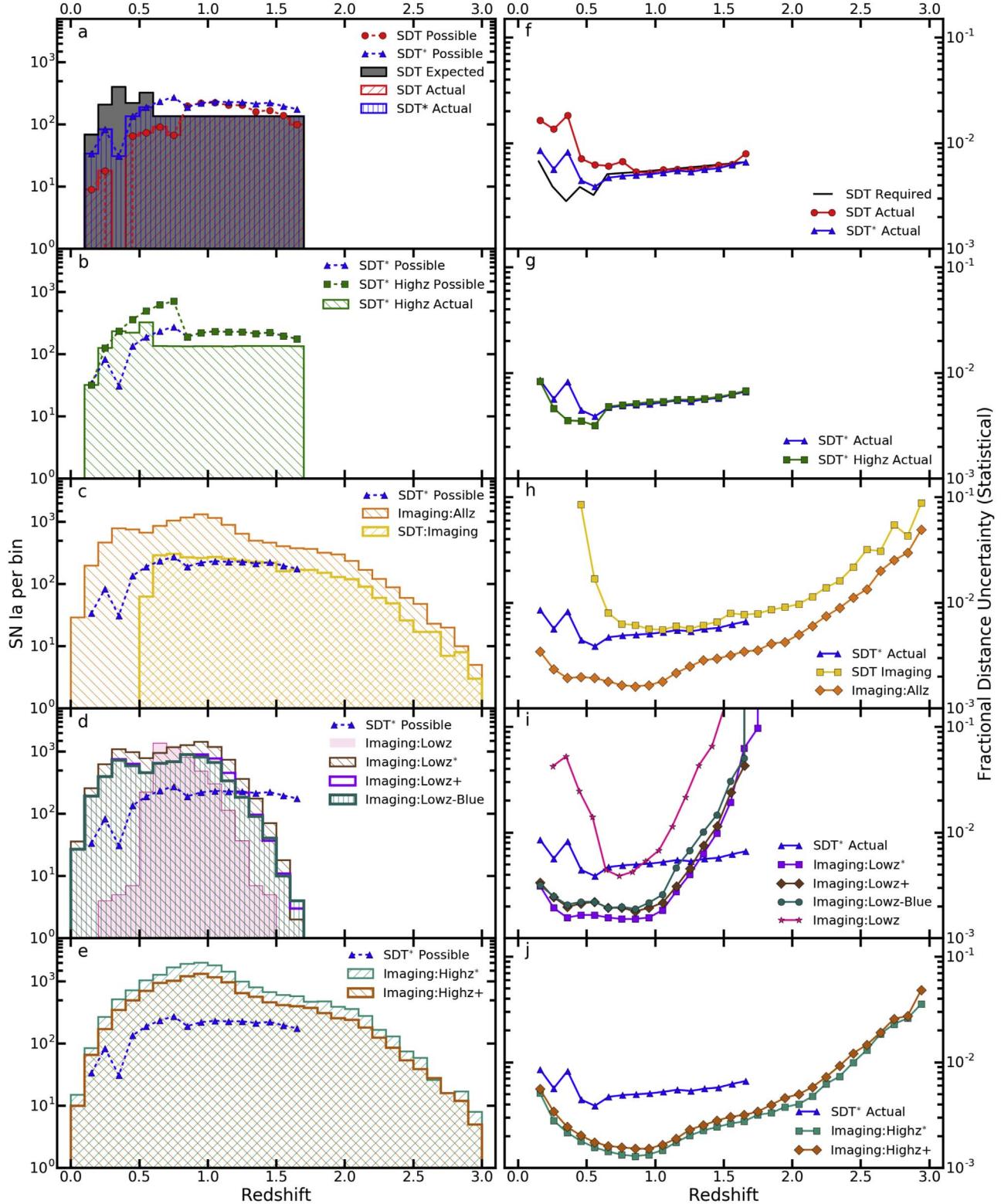
### 5.2. SDT\* Highz

As the shallow tier did not yield many SNe for the simulated SDT\* survey (Section 5.1), we examined the effects of removing that component and reallocating the time to the medium tier. This IFC-S+imaging-based strategy therefore consists of only two tiers: medium and deep. This simulation allows the medium tier to sample SNe within a greater redshift range,  $0.1 \leq z < 0.8$  (instead of the previously defined  $0.4 \leq z < 0.8$  range). The deep tier is unchanged from its description in Section 5.1. The area of the medium WFC imaging component is therefore increased by a factor of  $\sim 2.3$  by using the survey time from the shallow tier. The numbers presented in Table 7 are limited (where applicable) to the maximum number of SNe per 0.1 redshift bin as outlined in the SDT report. In addition, the modified selection criteria (Section 5.1) are implemented.

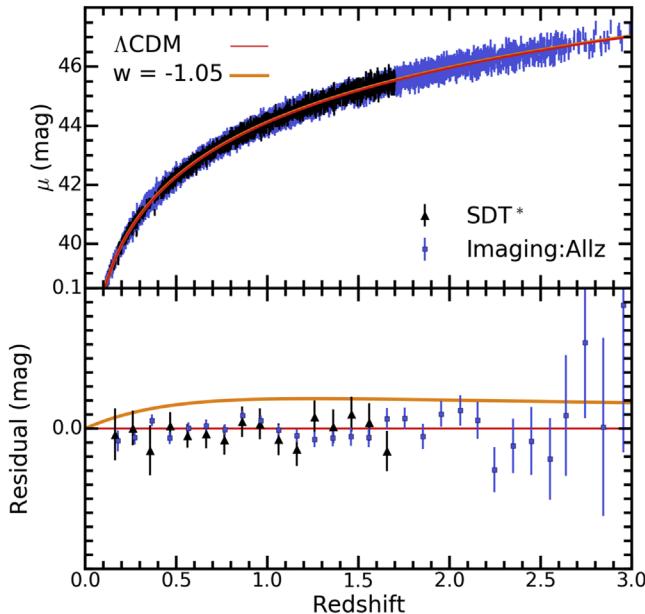
The SDT versions of the selection criteria were also applied to this simulation, but as in the previous SDT\* scenario, our modified selection yields a larger statistical sample. The results of this strategy can be seen in Figures 7(b) and (g). The redistribution of time to the medium tier results in 24% more SNe Ia within the final classified sample in comparison to the actual SDT\* results.

### 5.3. SDT Imaging

This simulation is based on a worst-case scenario where the SDT strategy is executed, but after the fact it is determined that the IFC-S data are unusable, resulting in exclusive use of the existing WFC imaging data. Presumably, this analysis can happen even if the IFC-S works perfectly. There is no increase in the areas of this strategy as it is exploring the idea of data obtained when an instrument is “faulty.” There are also no SN selection criteria as outlined in Section 3.1 as there are



**Figure 7.** Left panels: redshift distribution for each simulated *WFIRST* SN survey examined. For comparison, the SDT-required distribution is presented as a gray histogram in panel (a) (equivalent to Figure 3). Panel (a) also shows the “possible” SDT distribution for discovered SNe Ia that pass all SDT cuts (red circles). The red histogram represents the “actual” SDT distribution as described in the text. A similar curve and histogram for the SDT\* strategy are shown as blue triangles and a blue histogram. The remaining panels on the left present the redshift distributions of the other strategies examined. Right panels: fractional statistical distance uncertainties for each simulated *WFIRST* SN survey as a function of redshift. The assumed SDT uncertainties are plotted as the thick black line in the top-right panel (see Figure 3), with the measured uncertainties for the SDT (SDT\*) strategies represented by red circles (blue triangles). The remaining right panels present the fractional statistical distance uncertainties of the other strategies, with the left and right panels of a given row corresponding to the same strategy. For comparison, the “actual” SDT\* fractional distance uncertainties are presented as blue triangles in each panel.



**Figure 8.** BBC-fitted Hubble diagram of the simulated *WFIRST* SDT\* sample (black points) and Imaging:Allz sample (blue points). The red lines represent the  $\Lambda$ CDM model used to generate the simulation. Gold lines show a  $w$ CDM model with  $w = -1.05$ . The bottom panel displays binned Hubble residuals, relative to the  $\Lambda$ CDM model, for SDT\* (black triangles) and Imaging:Allz (blue squares).

no spectra. The purity of the resulting SN Ia sample is implemented via the aforementioned S/N requirements made on fitting (see Section 4). The results of this simulated survey are presented in Figures 7(c) and (h).

While the number of SNe Ia obtained within the simulation is a factor of  $\sim 1.24$  more than the possible SDT\* strategy sample, only  $\sim 76\%$  of these have  $0.1 \leq z \leq 1.7$ , and the rest are spread over higher  $z$ . In addition, only two SNe Ia are detected at  $z < 0.5$ . This issue can again be attributed to the insufficient S/N of the low- $z$  SNe in the shallow tier of the survey.

As stated above, this SDT Imaging strategy is a worst-case scenario and unlikely to happen. The strategy does, however, indicate the usefulness of limited imaging-only data. We do not consider the case where discovery filters in the WFC were to fail, as the mission would no longer be self-reliant for SN discovery.

#### 5.4. Imaging:Allz

This simulated WFC imaging-only survey uses all three SDT tiers, but four broadband filters instead of two. The four filters used are *RZYJ* for the shallow and medium tiers and *YJHF* for the deep tier. To account for the removal of the IFC-S component and the increase in the number of filters, the areas have been adjusted by factors of  $\sim 1.8$ ,  $2.2$ , and  $1.8$  for the shallow, medium, and deep tiers, respectively. These filters are chosen to span the rest-frame optical and NIR wavelength range, where our spectral models are defined.

The redistribution of IFC-S time, lack of IFC-S selection criteria, and additional filter selection, results in a factor of  $\sim 4.7$  increase in the final SN sample over the possible SDT\*. See Figures 7(c) and (h) for the results of this strategy. This is the first scenario for which the number of SNe Ia in each 0.1 redshift bin has exceeded the requirements set by the SDT.

After selection requirements and the BBC fit, we present a Hubble diagram for the Imaging:Allz and SDT\* SN surveys in Figure 8.

The SDT\* survey covers  $0.1 \leq z \leq 1.7$ , while the Imaging: Allz data set covers  $0.0 \leq z \leq 3.0$ . Within the  $0.1 \leq z \leq 1.7$  range, the SDT\* distance uncertainties are on average a factor of  $\sim 2.4$  worse than those from Imaging:Allz. By contrast, the uncertainties within the high- $z$  end ( $z > 2.0$ ) of the Imaging: Allz simulation are on average  $\sim 3.2$  times worse than the average uncertainty within the  $0.1 \leq z \leq 1.7$  range for the SDT\* survey. To illustrate the sensitivity of *WFIRST*, the gold lines show a  $w$ CDM model with  $w = -1.05$ .

#### 5.5. Imaging:Lowz

We simulated a WFC imaging-only survey that consists of the shallow and medium tiers only, and the use of the two SDT discovery filters (*Y+J* and *J+H*). The SDT time from the deep imaging and IFC-S are used to increase the area of the shallow tier by a factor of  $\sim 5.2$  and that of the medium tier by  $\sim 7.5$ .

This scenario results in a factor of  $\sim 1.6$  increase in the number of SNe Ia compared to the possible sample in the SDT\* scenario (see Figures 7(d) and (i)). There are no SNe Ia at  $z < 0.2$ , and only  $\sim 5\%$  of the total sample have  $z < 0.6$ . The fraction of SNe Ia with  $z \geq 1.2$  is only  $\sim 1\%$ .

#### 5.6. Imaging:Lowz\*

This is another imaging-only simulation that consists of the shallow and medium tiers, where the area of each respective tier has been increased by a factor of  $\sim 2.7$  and  $\sim 3.6$  to account for not using the IFC-S and deep-tier components. *RZYJ* filters are used in both tiers, maximizing our coverage of the rest-frame optical and extending to the rest-frame NIR. Comparing to the Imaging:Lowz strategy, we have included two additional filters, *R+Z*, and thus decreased the observed areas. This Imaging:Lowz\* strategy results in a factor of  $\sim 3.9$  increase in the number of SNe Ia in the final sample. Of the total sample, only  $\sim 6\%$  have  $z \geq 1.2$  (see Figures 7(d) and (i)).

#### 5.7. Imaging:Lowz+

Six filters, *RZYJHF*, are used in this imaging-only strategy with two tiers (each tier uses six filters). The areas of the shallow and medium tiers are increased by factors of  $\sim 1.8$  and  $2.3$ , respectively, with respect to SDT\*. This six-filter strategy leads to a factor of  $\sim 2.5$  increase in the number of SNe Ia compared to SDT\*. Only  $\sim 5\%$  of the sample has  $z \geq 1.2$  (see Figures 7(d) and (i)).

#### 5.8. Imaging:Lowz-blue

The number of tiers and their area on the sky within this strategy are the same as for Imaging:Lowz+; however, bluer filters have been selected, which include the WFC3 *F425W* (*B*), *F555W* (*V*), and *F814W* (*I*) filters in combination with the *WFIRST* SDT discovery filters, *Y+J*. The number of SNe Ia found by this strategy is  $\sim 2.3$  times greater than in the SDT\* survey. Only  $\sim 5\%$  of the SN sample has  $z \geq 1.2$  (see Figures 7(d) and (i)).

#### 5.9. Imaging:Highz\*

This strategy is similar to the Imaging:Lowz\* strategy, but here the medium and deep tiers are used rather than the shallow

and medium tiers. Time from the IFC-S and shallow components is used to increase tier areas by factors of  $\sim 3.6$  and 2.6, respectively. Filters selected are *RZYJ* for the medium tier and *YJHF* for the deep tier. The number of SNe Ia found by this strategy is  $\sim 6.5$  times greater than in the SDT\*. This two-tier survey results in a more complete sample of SNe Ia across the required redshift range. This is the second scenario for which the number of SNe Ia per 0.1 redshift bin exceeds the requirements set by the SDT report (see Figures 7(e) and (j)).

### 5.10. Imaging:Highz+

This strategy is similar to the Imaging:Lowz+ strategy, but here the medium and deep tiers are used rather than the shallow and medium tiers, and six bands are used: *RZYJHF* for the medium and *RZYJHF* for the deep. The areas of the two tiers have been increased by factors of  $\sim 2.3$  and 1.8, respectively, accounting for not using the IFC-S and shallow tier. The number of SNe Ia found by this strategy is a factor of  $\sim 4.4$  times greater than that found in the possible SDT\* survey sample. There is a reduction in the number of SNe Ia found here in comparison to the Imaging:Highz\* strategy due to the addition of two filters (*HF* in medium and *RZ* in deep), and thus a comparative reduction in area size. See Figures 7(e) and (j) for the results of this strategy.

### 5.11. Summary of Simulated Surveys

The redshift distributions of SNe Ia and their associated fractional distance uncertainties (Equation (14)) are shown in Figure 7. For the Imaging:Highz and Imaging:Allz strategies, the number of SN Ia detected per 0.1 redshift bin is significantly greater compared to the SDT\* results. However, for each low-z strategy the number of SNe Ia at  $z \geq 1.3$  drops dramatically owing to the loss of the long-exposure deep-tier component with redder bands.

The SDT Imaging and Imaging:Lowz strategies indicate how ineffective the shallow tier is. The dearth of SNe Ia at  $z < 0.5$  is driven by a combination of the shallow exposure time, resulting in objects with a low S/N, and poor filter selection, which provides minimal rest-frame coverage. The slew-and-settle time has a drastic impact on the shallow imaging tier, explaining the much lower utility we find for this tier relative to the SDT report.

In many of the imaging-only scenarios, the redshift range extends beyond  $z > 2$ , but note that the SN Ia rate has a larger uncertainty for  $z > 2$ . The increased number of SNe Ia for particular redshift ranges leads to an increase in the statistical precision per redshift bin, as much as a factor of  $\sim 2.4$  better than SDT\* for the Imaging:Highz\* strategy.

## 6. Systematic Uncertainties

In addition to the statistical uncertainties included in our analysis, several sources of systematic uncertainty have been investigated. These investigations are the first attempt to quantify the systematic uncertainties of the *WFIRST* SN survey without the use of ad hoc functions such as Equation (7).

We compute a covariance matrix to describe the distance modulus uncertainties such that  $C_{\text{tot}} = D_{\text{stat}} + C_{\text{sys}}$  (Conley et al. 2011).

$D_{\text{stat}}$  is a purely diagonal matrix, where the diagonal elements correspond to the individual SN distance modulus uncertainties given by Equation (8). The systematic

component,  $C_{\text{sys}}$ , can be described as the summation over each systematic uncertainty such that

$$C_{\text{sys},ij} = \sum_k \left( \frac{\delta\mu_i}{\delta S_k} \right) \left( \frac{\delta\mu_j}{\delta S_k} \right) (\Delta S_k)^2, \quad (15)$$

where  $\delta\mu_i/\delta S_k$  expresses the dependence of the distance modulus on the  $k^{\text{th}}$  systematic uncertainty for the  $i^{\text{th}}$  SN, and the magnitude of the  $k^{\text{th}}$  systematic uncertainty is given as  $\Delta S_k$ . To calculate  $C_{\text{sys}}$ , we change each systematic effect by  $1\sigma$  and determine the distance modulus difference from the BBC fit.

During this process, we fix  $\alpha$  and  $\beta$  from Equation (13) to the values found from the BBC fit using only the statistical uncertainties.

To characterize the dependence of the FoM on the value of each systematic uncertainty, we introduce a bias in our measurements that mimics the effect of each systematic uncertainty. We vary each systematic effect with multiplicative scaling from 0 (no effect) to 12 times the value of our current constraints for that uncertainty. For each case, we compare the BBC-fitted distance moduli determined with the included uncertainty to that determined without the effect. We display the absolute median distance modulus bias as a function of redshift for the nominal case (multiplicative factor of 1) in Figures 9 and 10. The  $\mu$  differences are used to compute the derivative term in Equation (15). Note that although absolute values are presented in Figures 9 and 10, the signs of the differences are used in the computation of the derivative. In addition, the values presented in these figures are defined to have zero residual at  $z = 0$ .

To determine the FoM, we input the derived distances and the associated covariance matrix to `CosmoMC*`. Additional constraints from CMB and BAO measurements are included in the fitting (as discussed in Section 4). For a subset of systematics, Figures 9–10 show the  $\mu$  shift versus redshift (left panels) and the relative FoM change with respect to the statistical-only FoM,  $\text{FoM}_{\text{stat}}$  (right panels).

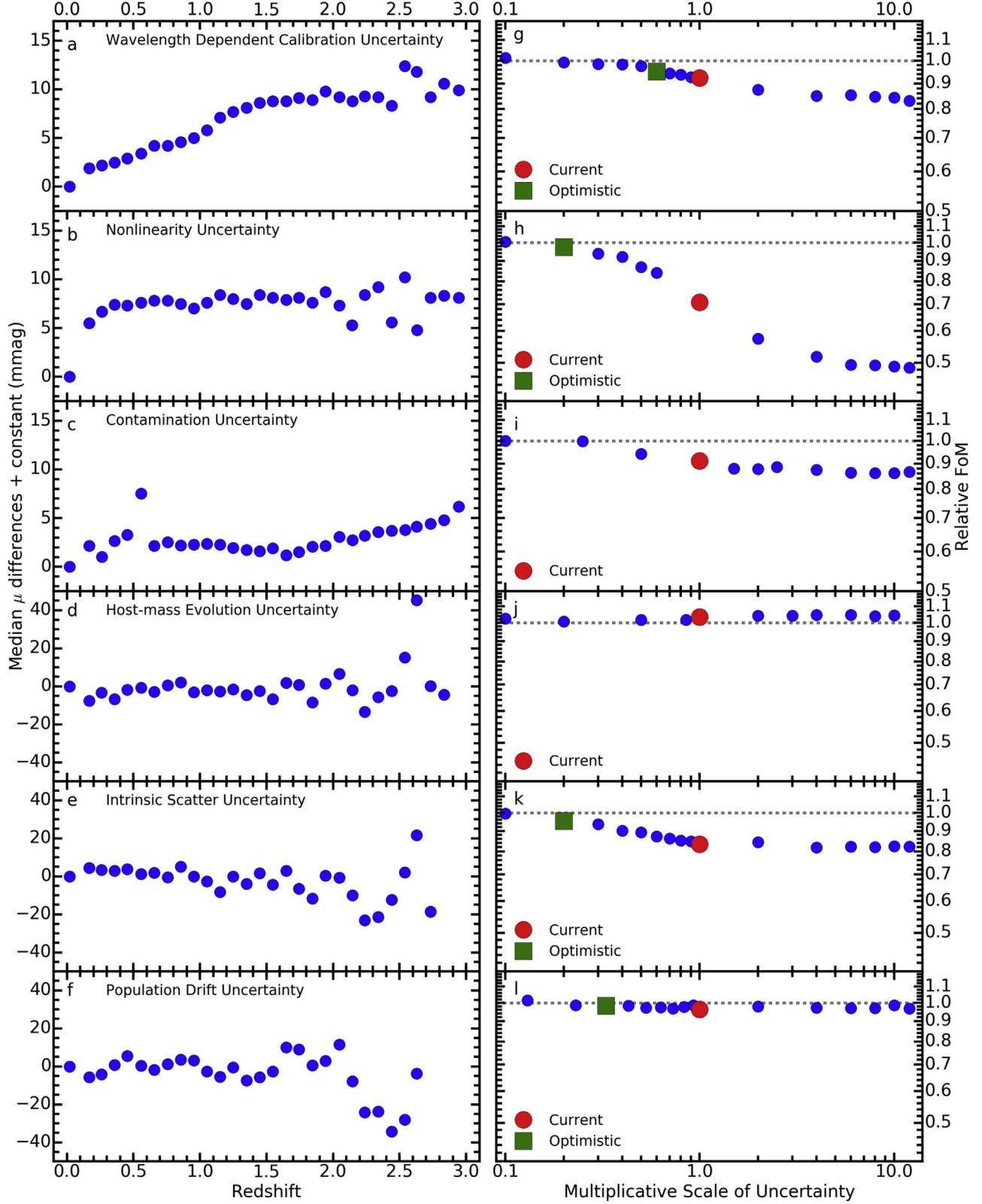
The points marked as “current” represent the FoM calculated with our present understanding of the systematic uncertainty (i.e., a multiplicative scaling of 1),  $\text{FoM}_{\text{tot,curr}}$ . Points marked “optimistic” represent the FoM values calculated with assumptions for improved systematic uncertainties,  $\text{FoM}_{\text{tot,opt}}$ . These optimistic systematic uncertainties are values that we hope will be achievable by launch, and are based on reasonable assumptions.

The limited precision of `CosmoMC*` (Section 4) and the artifacts of light-curve fitting add small numerical noise to FoM measurements, making their values deviate, on order of a few percent, from a smooth interpolation. However, all of our main findings are robust against these small variations.

### 6.1. Calibration

Calibration uncertainty is currently the largest systematic uncertainty of all recent ground-based SN cosmology analyses (e.g., Scolnic et al. 2014a). The primary sources of calibration uncertainty can be split into three separate components, which are listed and discussed below. The nominal size of each component is set to match the current values determined for the *HST* system. This is likely a conservative assumption and is varied within the present analysis.

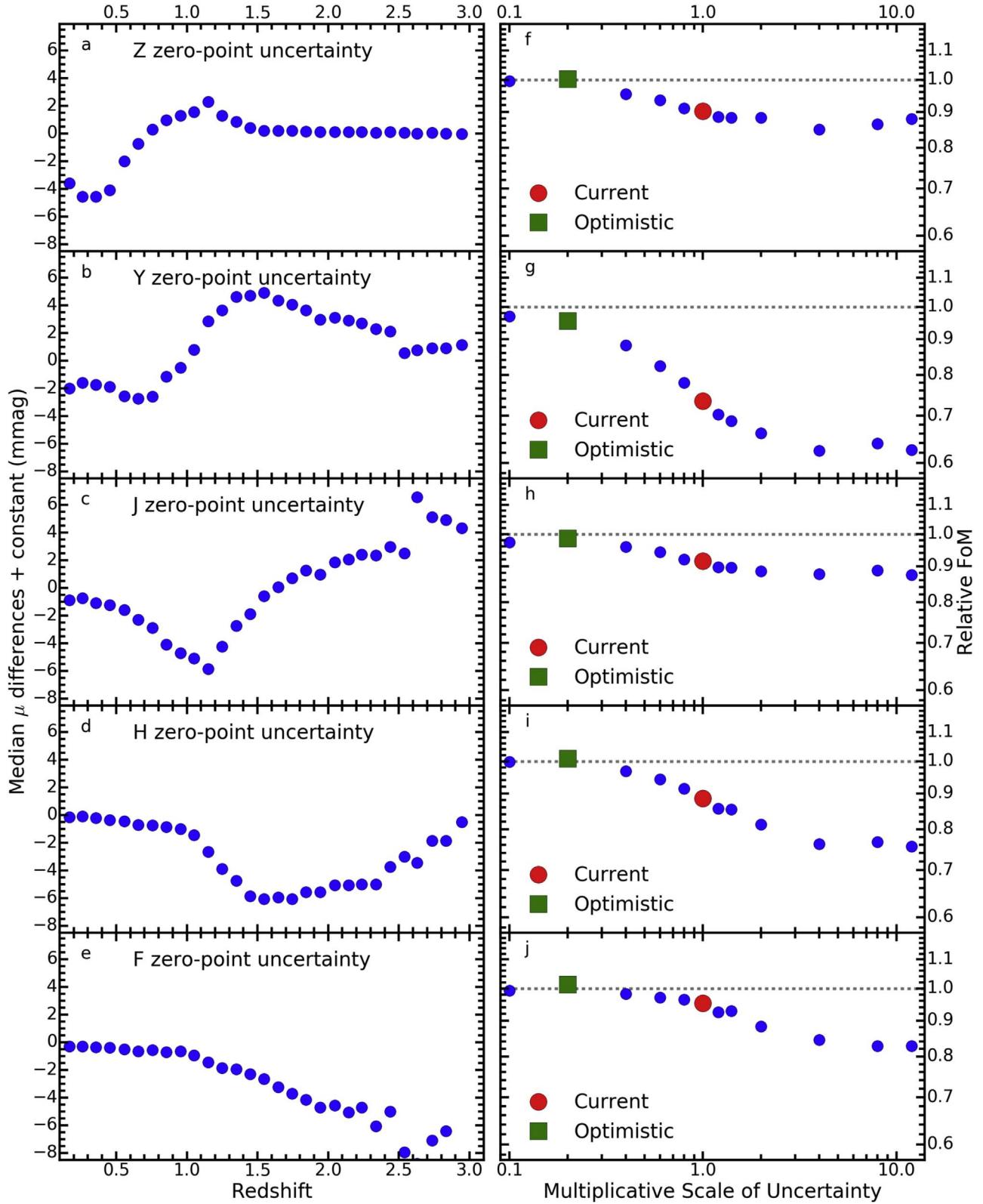
1. The wavelength-dependent flux uncertainty: The accuracy of the *HST* Calspec system (Bohlin et al. 2014) is



**Figure 9.** Left panels: for the Imaging:Allz strategy, median distance modulus difference vs. redshift for the current systematic effect labeled on each panel. Note that the values presented in these figures are defined to have zero residual at  $z = 0$ . Right panels: FoM/FoM<sub>stat</sub> for different values of each systematic uncertainty (with the scaling relative to the current value). The dashed line represents the statistical FoM. Red circles and green squares represent the current and optimistic values of each uncertainty, respectively.

described as a linear function with a slope of roughly 5 mmag per 7000 Å (Bohlin 2007). Assuming that the functional form of the calibration of WFIRST is similar to

that of *HST*, we use the magnitude of the *HST* systematic uncertainty as the nominal uncertainty for the WFC (see Figures 9(a) and (g)). For the IFC-S, we assume this



**Figure 10.** Same as Figure 9, but for additional systematic uncertainties.

uncertainty to be 50 mmag per 7000 Å, as there are many unknown calibration issues for this instrument. This higher uncertainty for the IFC-S is based on work conducted in Section 4.6 of Bacon et al. (2015), which compares synthesized broadband magnitudes from the Multi Unit

Spectroscopic Explorer (MUSE—a panoramic integral field spectrograph) to those from the *HST* and finds a mean bias of 50 mmag, with a statistical uncertainty of 40 mmag. In addition, similar values were also found by Childress et al. (2016) in which data from a supernova survey using the Wide

Field Spectrograph on the Australian National University Telescope enabled the determination of a color variation ranging from 40 mmag in the red to 90 mmag in the blue.

For both the IFC-S and WFC, we assume an optimistic wavelength-dependent flux uncertainty of 3 mmag per 7000 Å, a factor of  $\sim 1.7$  better than *HST*. This is the main calibration systematic uncertainty for the IFC-S and also a significant uncertainty for the WFC filters.

2. Nonlinearity of the detector: Detector response nonlinearity can impact photometric precision in astronomical observations. Recent work (e.g., Riess 2010) suggests that a count-rate dependent nonlinearity is common in the HgCdTe detectors that will be used by *WFIRST*'s WFI. Riess (2010) measured a WFC3-IR nonlinearity of  $\sim 1\%$  per dex over a range of 10 mag (4 dex), which was independent of wavelength. We take this value as our baseline systematic<sup>32</sup> (see Figures 9(b) and (h)).

Our optimistic nonlinearity systematic uncertainty is assumed to be a factor of 5 better than the values obtained from *HST* studies. This is reasonable given current and future improvements of detectors.

3. Zero-point uncertainty: Recent ground-based imaging surveys have achieved 5 mmag uncertainties in their filter zero-points (Betoule et al. 2013; Scolnic et al. 2015). Because of the color term in the SN distance equation, the  $\mu$  bias can be  $\beta \simeq 3$  times larger than the zero-point bias.

A space-based observatory, being above the atmosphere, should have zero-point uncertainties no larger than those from state-of-the-art ground-based surveys. We have therefore included this uncertainty with a 5 mmag shift to each filter zero-point (see Figure 10). Optimistic WFC imaging zero-point uncertainties are assumed to be 1 mmag.

As stated in Section 4, we include a sample of 800 low- $z$  SNe simulated based on predictions of the Foundation survey from Foley et al. (2018). These low- $z$  SNe provide an anchor to the SN Ia Hubble diagram.

Here we consider two main sources of calibration uncertainties that impact the low- $z$  sample. The first is how well the photometric system of the low- $z$  sample is calibrated and tied to *HST* Calspec standards. Since the Foundation survey uses the Pan-STARRS photometric system (Chambers & Pan-STARRS Team 2017), we adopt  $g'$ -,  $r'$ -, and  $i'$ -band zero-point uncertainties of 5 mmag (Scolnic et al. 2015) and assume that these uncertainties will improve with a dedicated standard-star observing program to more optimistic values of 3 mmag. The second source of calibration uncertainty is that of the *HST* Calspec standards themselves (as discussed earlier), which affects both the low- $z$  and *WFIRST* samples.

Given that we aim to obtain a relative calibration accuracy of 1–2 mmag between ground-based SN surveys (e.g., Large Synoptic Survey Telescope (LSST)/PS1/DES) and *WFIRST*, it might be appropriate to drop the low- $z$  (shallow) tier of the *WFIRST* SN survey. This shallow tier has proved to be inefficient in obtaining SNe with the S/N required for detection and accurate classification within the *WFIRST* cosmology sample. The redistribution of the shallow-tier time into that of the medium or deep tiers serves to benefit the overall efficiency

<sup>32</sup> Note added in proof. For the IFC-S, we assumed an optimistic nonlinearity for both the current and optimistic case; we realized this only after the analysis, and thus all quoted IFC-S FoM values presented within the paper use an optimistic nonlinearity value. If a nominal nonlinearity is applied, the effect is to reduce the current FoM listed in Table 13 by a maximum of 13%.

of the survey (see, for example, the SDT\* Highz or Imaging: Highz\* surveys).

If, however, the relative calibration between ground and space-based data is suboptimal, then a low- $z$  *WFIRST* SN survey may be required. Such a survey would require a different strategy than those proposed within this paper and will be the focus of future investigations.

Within this section, we have addressed some of the larger systematic uncertainties affecting the calibration of the *WFIRST* instruments. We have not, however, considered calibration uncertainties at the pixel level. For these systematics, the implicit assumption is that they average out due to the distribution of observations of SNe over the detector, resulting in a random measurement uncertainty, and thus their effect is considered minimal compared to the more dominant uncertainties listed within this section. In addition, we ignore subpixel sensitivity variations and persistence uncertainties, which may introduce redshift-dependent effects owing to the reduced resolution of host galaxies at high redshift.

## 6.2. Core-collapse SN Contamination

For surveys that use the IFC-S, we assume that there are no CC SNe in the cosmological sample, i.e., we set the current and optimistic contamination to be 0%. Although some CC SN contamination was present within the SDT/SDT\* results (see Section 5.1), we expect that by using all spectra and light-curve data (not just the first five imaging data points and three IFC-S spectra), contamination will drop significantly. In addition, we expect that improved classification techniques in the future will aid in CC SN contamination reduction.

For each imaging-only scenario (Sections 5.3 through 5.10), however, CC SN contamination must be considered. Our analysis results in a photometric classification purity of  $>99\%$  (for the Imaging:Allz survey). To account for the systematic uncertainty introduced by any remaining contamination, BBC-fitted Hubble residuals are obtained with and without CC SNe, and differences in distance versus redshift are used as a systematic uncertainty. We further reduce the contamination by a factor of 5 based on the assumed improvements from photometric classifiers using machine learning (see Figures 9(c) and (i)). This assumption is based on a factor of 3.6 contamination reduction for DES simulations (Kessler & Scolnic 2017) using a simple nearest-neighbor (NN) method on only three variables (redshift, color, stretch). Using a wider range of light-curve features, better machine-learning methods have been demonstrated, such as by Lochner et al. (2016). The addition of NIR data in *WFIRST* can also be used to reduce contamination. Training a photometric classifier is beyond the scope of this work but will be considered in future investigations.

Our optimistic core-collapse SN contamination uncertainty for imaging-only simulations is assumed to be negligible, as we expect classification methods to have improved substantially by launch and to be able to take advantage of the rest-frame NIR data.

## 6.3. SN Physics

Our analysis includes five systematic uncertainties related to SN physics.

1. The host-galaxy–SN luminosity relation: After correcting for SN light-curve shape and color, SN Ia Hubble residuals still correlate with host-galaxy properties (e.g., Kelly et al. 2010;

Lampeitl et al. 2010; Sullivan et al. 2010). Although the cause of this effect is still unknown, it is possible that it is related to different progenitor properties, such as metallicity or age, that correlate with environment.

Currently, cosmology analyses (e.g., Betoule et al. 2014) correct SN luminosities based on the mass of the SN host galaxy relative to a central split value. The exact functional form of this correction is still poorly constrained, but most assume a binary population split at  $10^{10} M_{\odot}$  (Sullivan et al. 2010). It is possible that the magnitude of this correction and the form could change with redshift (e.g., Rigault et al. 2013, 2015; Childress et al. 2014). However, the size of the systematic uncertainty due to the mass-dependent evolution can be mitigated by measuring the relation between distance residuals and mass at different redshifts. This method is similar to ideas presented by Shafer & Huterer (2014). Therefore, the size of the systematic uncertainty is actually dependent on how well the evolution of the relation is measured.

To study this effect, we label half of our SN Ia sample as “high mass” and introduce a redshift-dependent offset in the fitted  $m_B$ ,

$$m_{B,\text{shift}} = m_B + 0.06 - [0.06 \times (1 - z)]. \quad (16)$$

We then determine the redshift dependence of the difference in Hubble residuals between our altered “high-mass” sample and the unaltered “low-mass” sample. The difference in our recovered dependence and our input dependence, given in Equation (16), is used as the size of our nominal systematic uncertainty (see Figures 9(d) and (j)). For this analysis, we assume that the uncertainty in the difference of the Hubble residuals for the two host-mass bins is dominated by the distance uncertainty rather than by uncertainties in the mass estimates of the host galaxies. There may be a larger systematic uncertainty related to a population drift of the host galaxies; however, we choose this particular kind of systematic for the host-galaxy–SN luminosity relation because it can be reduced with increasing statistics. As such, there is no optimistic value for this bias because it is based purely upon the statistics of the survey.

We note that within Figure 9(j), the FoM appears to increase with the multiplicative scale factor. Further analysis of the host-mass systematic indicates that within the 0.1–10 range of the fiducial value,  $w_a$  appears to have shifted by approximately one half sigma. Because  $w_a$  correlates with other cosmological parameters that are fixed by the CMB, this bias spuriously increases the FoM. For a more detailed discussion of similar issues, see Scolnic et al. (2014a).

2. Intrinsic scatter uncertainty: There is still uncertainty in the relative proportion of color variation and luminosity variation in the intrinsic scatter model for SNe Ia (see Scolnic & Kessler 2016, for a review). The distance bias corrections applied depend on the assumption of the intrinsic scatter model. The differences between the bias corrections are typically largest where selection effects are strongest because the intrinsic scatter model will determine whether predominantly bluer objects are selected or predominantly brighter objects are selected. To determine the impact on our cosmological measurements from this uncertainty, we first simulated our samples with two different intrinsic scatter models: G10, a model from Guy et al. (2010), which has 70% luminosity variation and 30% color variation, and a model from Chotard et al. (2011, hereafter C11), which has 25% luminosity variation and 75% color variation. Following Kessler et al. (2013),

we converted the G10 and C11 models into spectral-variation models for SNANA. The difference between the recovered distances from these two models is the systematic uncertainty, shown in Figure 9(e). The structure of the distance differences with redshift shown in Figure 9(e) is due to the impact of various selection effects (from the tiered surveys) on the different scatter models.

The optimistic intrinsic scatter uncertainty is assumed to be a factor of 5 better than current estimates due to improved models in the IR (e.g., Mandel et al. 2011) and improved spectra to empirically model intrinsic SED variations.

3. Population drift: Related to uncertainty in the intrinsic scatter model, there is uncertainty in whether this form of the scatter could evolve with redshift. This issue is conflated in past analyses with the possibility that the color of the SN population could evolve with redshift (Scolnic & Kessler 2016; Mandel et al. 2017), and this evolution is not accounted for in the analysis. To determine the impact on our cosmological measurements from this uncertainty, we introduced a supernova color population drift of  $0.01 \times z$  mag, keeping the defined color range and bifurcated Gaussian  $\sigma$  identical to previous simulations. This means that the center of the color ( $c$ ) population for a sample of SNe Ia, at fixed absolute magnitude,  $M$ , increases as a function of redshift by  $0.01 \times z$ .

While there is evidence for an  $x_1$  population drift (see Scolnic & Kessler 2016; Scolnic et al. 2018), it will have less impact on possible distance biases than a  $c$  population drift because of the different correlations between  $c$  and  $x_1$  with luminosity. Therefore, in this analysis, we do not include an uncertainty from  $x_1$  population drift. The difference between the recovered distances from the color shift and the nominal simulation are shown in Figure 9(f), with relative FoM values given in Figure 9(l).

It is possible that with the IFC-S, evolution of the intrinsic color can be constrained by measuring the SN ejecta velocities (Foley & Kasen 2011; Foley et al. 2011; Foley 2012; Mandel et al. 2014). This claim is analyzed further in Appendix D, though for our nominal systematic uncertainty, we do not assume any improvement in the constraint on intrinsic color evolution or population drift from the IFC-S.

The optimistic color-versus-redshift drift is taken to be  $0.0033 \times z$ .

This is an estimate, as it is unclear from recent studies (e.g., Rubin et al. 2015; Rubin & Hayden 2016; Scolnic & Kessler 2016) what an optimistic constraint should be.

The following systematic uncertainties are not included in our FoM<sub>tot</sub> predictions, but their effects have been considered.

4. Beta evolution: The properties of interstellar dust may change with redshift, affecting the ratio of total to selective extinction. This evolution would manifest itself in a change in the recovered value of  $\beta$  (Scolnic et al. 2014b) with redshift (Conley et al. 2011). Furthermore, as shown in Mandel et al. (2017), the color law (CL in Equation (11)) may be composed of a reddening law as well as a separate relation between SN intrinsic color with luminosity, and the relative components of the two may change with redshift. Similar to the correlation of Hubble residuals with host mass,  $\beta$  evolution can be included with additional BBC fit parameters, and their uncertainties will decrease with increasing sample size.

Therefore, the  $\beta$ -evolution uncertainty is expected to be small compared to the systematic uncertainty from the intrinsic uncertainty or population drift, so it is not included here.

**Table 12**  
Current and Optimistic Systematic Uncertainties Investigated for Both the WFC and IFC-S

Systematic Uncertainty	Current		Optimistic	
	WFC	IFC-S	WFC	IFC-S
Wavelength-dependent calibration	5 mmag per 7000 Å	50 mmag per 7000 Å	3 mmag per 7000 Å	3 mmag per 7000 Å
Nonlinearity	1% per dex over 10 mag	0.2% per dex over 10 mag	0.2% per dex over 10 mag	0.2% per dex over 10 mag
Zero-point offsets	5 mmag	...	1 mmag	...
CC contamination	1/5 of derived systematic <sup>a</sup>	0%	0%	0%
Population drift	$10 \text{ mmag} \times z$	$10 \text{ mmag} \times z$	$3.3 \text{ mmag} \times z$	$3.3 \text{ mmag} \times z$
Intrinsic scatter	The difference between the G10 and C11 models		1/5 of current	1/5 that of current
Host-mass evolution	Calculated for each strategy <sup>b</sup>			
Beta evolution	Considered Negligible <sup>c</sup>			
K-corrections	Considered Negligible <sup>c</sup>			
MW extinction	Considered Negligible <sup>c</sup>			

**Notes.**<sup>a</sup> See Section 6.2 for details on this systematic uncertainty.<sup>b</sup> For each simulated survey strategy, the host-mass systematic uncertainty was calculated as described within Section 6.3.<sup>c</sup> As the effect of this systematic uncertainty is considered negligible (see Section 6.3), we have not included it within our final analysis.

5. *K*-corrections: The SDT report lists *K*-corrections as a top systematic uncertainty and a large motivation for the use of the IFC-S over broadband imaging. However, since modern distance-fitting algorithms employ spectral models to fit SN data in the observer frame, no true *K*-correction is ever applied. Instead, a *K*-correction uncertainty should be described as an imperfect knowledge of a supernova SED. Since most of our SN Ia training set is at  $z \approx 0$ , certain regions of the spectral model (near the effective wavelength of certain filters) are better constrained than others. If the deredshifted observer-frame and rest-frame filters are not well aligned, the diversity of spectral features could cause an additional statistical uncertainty of up to 0.04 mag (e.g., Saunders et al. 2015). With IFC-S measurements, one can synthesize photometry over any wavelength range, largely eliminating this uncertainty.

It has been argued (e.g., Aldering et al. 2002; Perlmutter & Schmidt 2003) that this uncertainty will be dominant for a space-based SN mission. However, even with the most pessimistic scenario (Saunders et al. 2015), the additional 0.04 mag of scatter is still negligible compared to the 0.13 mag intrinsic scatter of SNe Ia (Betoule et al. 2014; Scolnic et al. 2014a). Saunders et al. (2015) also predict systematic biases on the 0.01 mag level, which follow oscillatory behavior with a period of  $\Delta z \approx 0.1$ . For a redshift range of an SN sample with  $\Delta z \geq 1.0$ , the impact of this 0.01 mag bias should be reduced to the sub-mmag level.

To further examine the impact on the measurement uncertainties using the IFC-S, the IFC-S spectra were rebinned, maintaining the overall S/N, mimicking progressively lower-resolution spectra (or wider filters). Note that intrinsic spectral variations are included in the simulated spectra, but the analysis does not account for these variations except for a global intrinsic scatter term in the distance modulus uncertainty. No systematic bias is found and distance modulus uncertainties do not increase, confirming that this uncertainty will be subdominant. Therefore, *K*-corrections are not included as an additional systematic uncertainty.

6. Milky Way extinction: Systematic uncertainties in the amount of Milky Way (MW) extinction along the line of sight to the SNe will propagate to systematic uncertainties in the recovery of the cosmological parameters. The *WFIRST* SN fields have not yet been chosen, but it is likely they will be picked to minimize the amount of MW extinction: MW  $E(B - V) < 0.02$  mag. As discussed by Scolnic et al. (2014b),

systematic uncertainties in the MW extinction take the form of a multiplicative component and additive component. Assuming a 10% multiplicative uncertainty and a separate 3 mmag additive uncertainty, we find that the impact on the FOM is small (<10%), relative to the other systematic uncertainties discussed above. Therefore, it is not included in our analysis.

The list of systematic uncertainties presented within this section, although not complete, covers some of the most dominant uncertainties found today in ground-based (and space-based) SN surveys. It is important to note, however, that some of the uncertainty descriptions may be oversimplifications of a more complex problem. Future studies into these systematic uncertainties as well as additional biases could include the consideration of population drift outside of the SALT2 framework; the redshift-dependent efficiency and bias of imaging-only surveys; selection biases in the IFC-S-focused surveys, including the effect of SN classifications with photo-redshifts instead of spectroscopically determined redshifts; and host-subtraction uncertainty, i.e., random noise or redshift dependency. See Section 8 for a more detailed discussion of these biases.

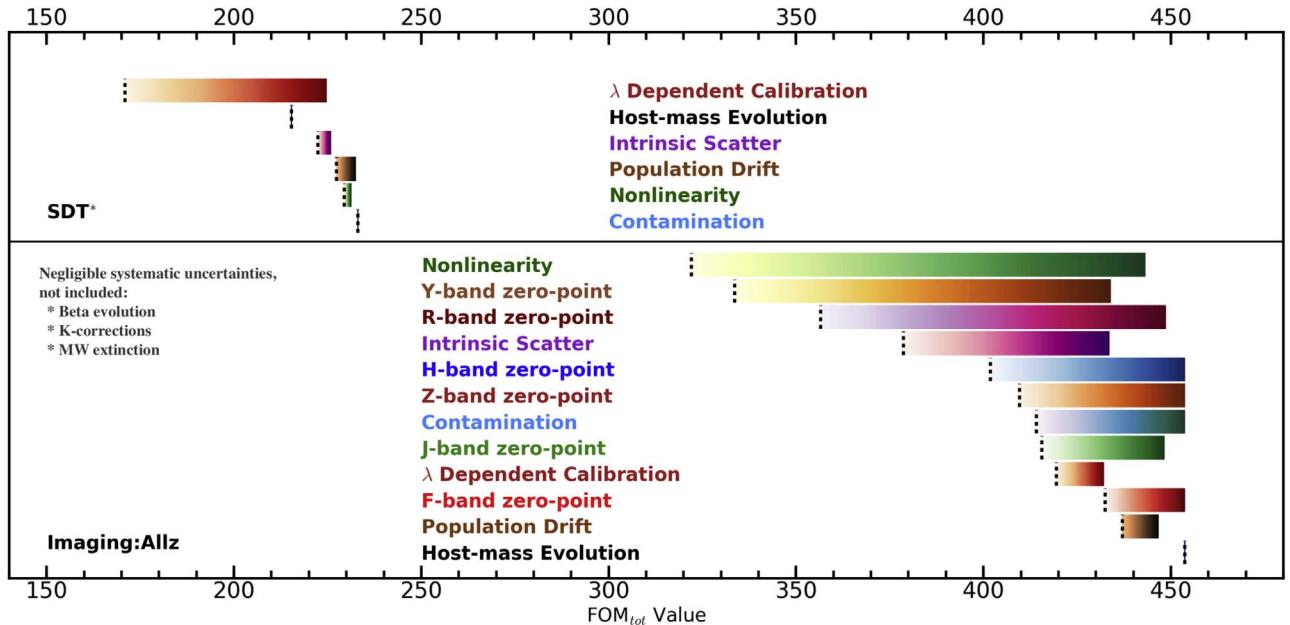
#### 6.4. Statistical and Systematic Uncertainty Summary

A summary of the current and optimistic systematic uncertainties investigated by our various simulations is presented in Table 12.

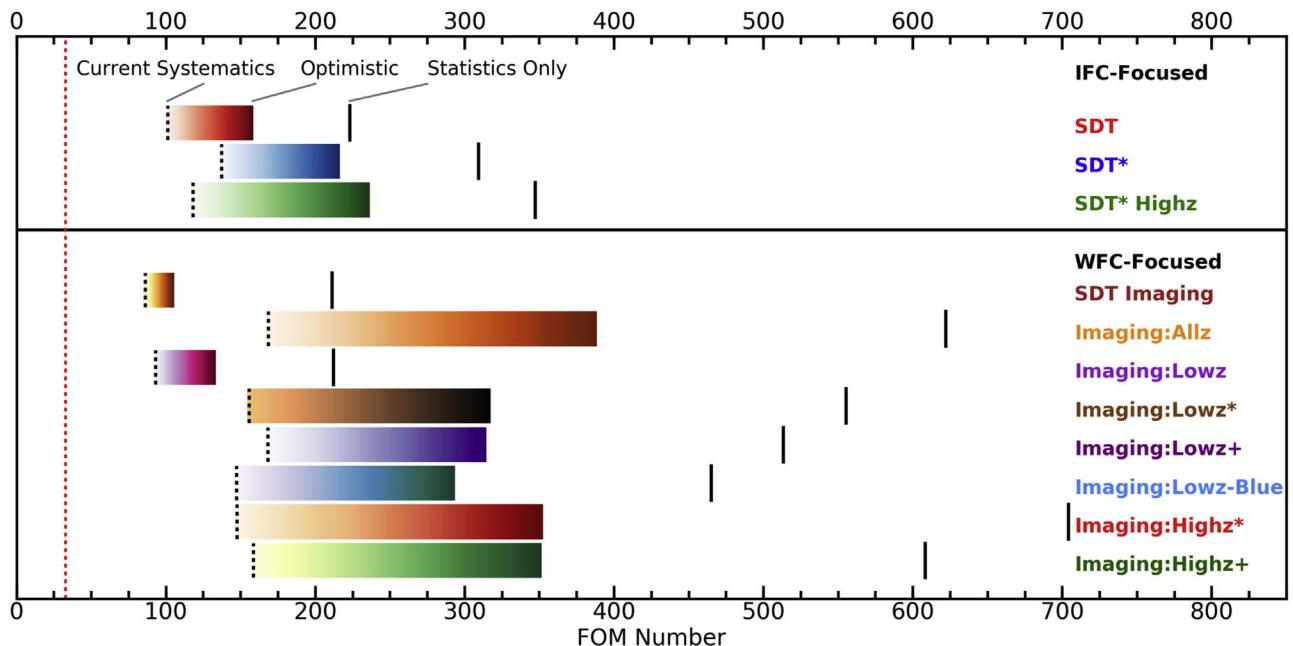
The effect of each individual systematic uncertainty, both current and optimistic, is presented within Figure 11. The values plotted here are produced using *CosmoMC\**. For the SDT\* survey simulation, the largest uncertainty is the wavelength-dependent calibration uncertainty. Our current estimate for this uncertainty is 50 mmag per 7000 Å. We hope that this value will decrease by over a factor of 10 by launch, leading to the much larger relative  $FOM_{tot,opt}$ . For the imaging-only scenario, the largest systematic uncertainties are the nonlinearity, intrinsic scatter, and zero-point offsets, specifically for the *Y* and *R* bands. Further evaluation of these uncertainties is required in order to fully understand their effects and enable optimization of survey strategies.

#### 7. Comparison of Simulated Survey Strategies

Within this paper, we have simulated a total of 11 different SN survey strategies for the *WFIRST* mission. Here we



**Figure 11.**  $\text{FoM}_{\text{tot}}$  when only a single systematic uncertainty is included, as indicated in the figure. The range for each  $\text{FoM}_{\text{tot}}$  spans current (dotted line) to optimistic. The top and bottom panels include systematic uncertainties for the SDT\* (see Section 5.1) and Imaging:Allz strategies (see Section 5.4), respectively. In cases where noise fluctuations make the  $\text{FoM}_{\text{tot}}$  slightly greater than  $\text{FoM}_{\text{stat}}$ , the FoM is set as  $\text{FoM}_{\text{stat}}$ . Note that the SDT\* results do not include the effect of zero-point uncertainties as this strategy does not use any imaging to measure distances. Negligible uncertainties (such as beta evolution, K-corrections, MW extinction) are not included.



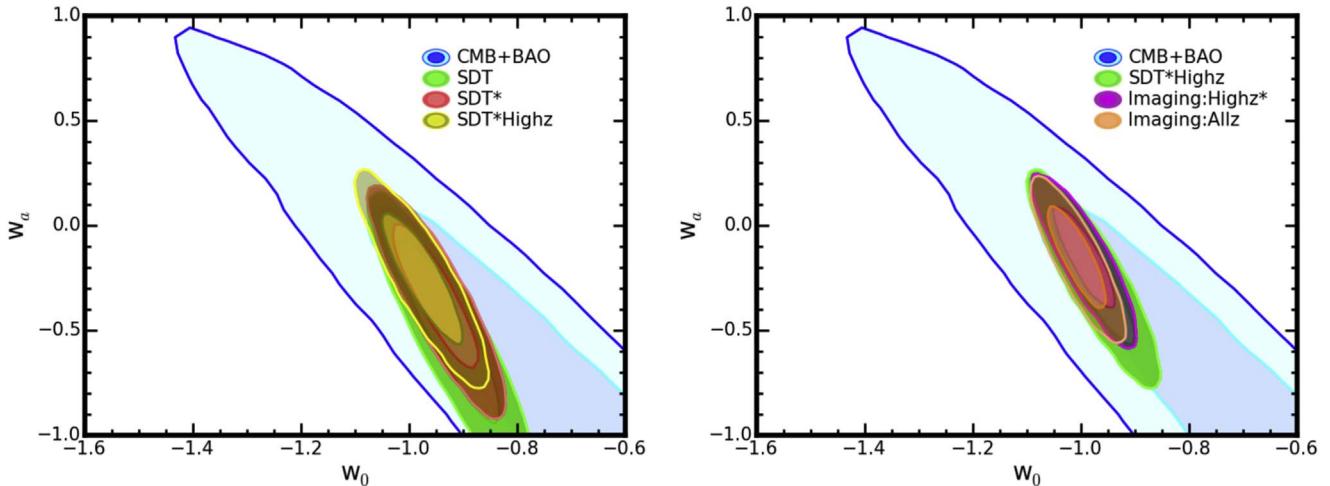
**Figure 12.** Predicted dark-energy FoMs for the simulated WFIRST SN survey strategies outlined in Section 5. IFC-S-focused and WFC-focused strategies are presented in the top and bottom panels, respectively. The gradients for each strategy represent the range of FoMs from  $\text{FoM}_{\text{tot},\text{curr}}$  (dotted lines) to  $\text{FoM}_{\text{tot},\text{opt}}$ . The thick black lines represent  $\text{FoM}_{\text{stat}}$ . The red dashed vertical line indicates the current FoM value of 32.6 (Alam et al. 2017).

compare each strategy, assessing how successful they are at constraining dark-energy models, via their FoM values. We also examine the details of these strategies, such as redshift distribution of SNe Ia, and suggest how they may be improved.

Using the “optimistic” systematic uncertainties described above, we have evaluated the impact of each uncertainty on each strategy, the results of which are presented in Figure 12, with Table 13 listing the  $\text{FoM}_{\text{stat}}$ ,  $\text{FoM}_{\text{tot},\text{curr}}$ , and  $\text{FoM}_{\text{tot},\text{opt}}$  values determined for each case. For completeness, the FoM values presented here are calculated using the original version of CosmoMC, where the full set of Planck likelihoods is used.

Figure 12 shows that these strategies result in a wide range of  $\text{FoM}_{\text{stat}}$  (211–704) values compared to a much narrower range of  $\text{FoM}_{\text{tot},\text{curr}}$  (86–169) values.

Examination of strategies that use both the IFC-S and WFC (e.g., SDT, SDT\*, and SDT\* Highz) lead to several important conclusions. The SDT strategy as outlined by Spergel et al. (2015) results in a lower than expected number of SNe Ia at  $z < 0.6$ . This decrease in low- $z$  SNe is a result of the short exposure time within the shallow tier of the imaging survey and the strict SDT spectrophotometric selection criteria.



**Figure 13.** The  $w_0-w_a$  68% and 95% confidence contours for the simulated SDT, SDT\*, and SDT\* Highz *WFIRST* SN surveys (left panel) and the SDT, Imaging:Allz, and Imaging:Highz\* *WFIRST* SN surveys (right panel). Each contour represents the total (statistical plus optimistic systematic uncertainties) SN Ia constraints combined with CMB and BAO constraints. For comparison, we have included the confidence contours created using CMB+BAO data only.

**Table 13**  
FoM Values and Cosmological Parameter Uncertainties for Each Strategy

* Strategy <sup>a</sup>	Statistical			Current			Optimistic		
	FoM	$\sigma(w_0)$	$\sigma(w_a)$	FoM	$\sigma(w_0)$	$\sigma(w_a)$	FoM	$\sigma(w_0)$	$\sigma(w_a)$
SDT	223	0.053	0.26	101	0.067	0.29	158	0.061	0.28
SDT*	309	0.042	0.20	137	0.058	0.26	216	0.050	0.22
SDT* Highz	347	0.037	0.18	118	0.060	0.28	236	0.048	0.20
SDT Imaging	211	0.059	0.29	86	0.071	0.27	105	0.068	0.25
<b>Imaging:Allz</b>	<b>622</b>	0.024	0.12	169	0.061	0.25	388	0.035	0.16
Imaging:Lowz	212	0.057	0.29	93	0.078	0.33	133	0.067	0.31
Imaging:Lowz*	555	0.027	0.13	156	0.066	0.28	317	0.042	0.19
Imaging:Lowz+	513	0.029	0.14	168	0.063	0.27	314	0.043	0.19
Imaging:Lowz-Blue	465	0.031	0.15	147	0.067	0.27	293	0.044	0.19
<b>Imaging:Highz*</b>	<b>704</b>	0.022	0.11	148	0.069	0.29	352	0.039	0.17
<b>Imaging:Highz+</b>	<b>608</b>	0.024	0.12	159	0.063	0.27	351	0.035	0.16

Note.

<sup>a</sup> The ordering of this table follows the ordering in Section 5. Strategy names in bold have the highest FoM<sub>tot,opt</sub> values.

A slight modification of these selection criteria, as implemented in SDT\* (see Figure 7(a)), increases the total number of low- $z$  ( $z < 0.6$ ) SNe Ia by  $\sim 300\%$ , with an increase of  $\sim 31\%$  for  $0.1 \leq z \leq 1.7$ . The short exposure time of the shallow imaging tier leads to low-S/N SNe, and thus even with modified selection criteria this shallow survey still hinders the number of  $z < 0.6$  SNe Ia obtained, reaching only  $\sim 39\%$  of the fraction stated within the SDT report.

Based on these results, we conclude that any selection criteria implemented must be very carefully chosen to maximize both efficiency and purity, and that the shallow tier of the imaging survey (with current overhead estimates) is a suboptimal use of survey time. Shifting exposure time from the shallow imaging tier to the medium tier, and applying the modified selection criteria, significantly increases the SN Ia sample size (460 and 923 more SNe Ia compared to the SDT\*, and SDT surveys, respectively) as indicated in our SDT\* Highz survey simulation (see Figure 7(b)). This SDT\* Highz strategy results in a much higher FoM<sub>tot,opt</sub> value of 236 in comparison to the FoM<sub>tot,opt</sub> = 158 value of the SDT.

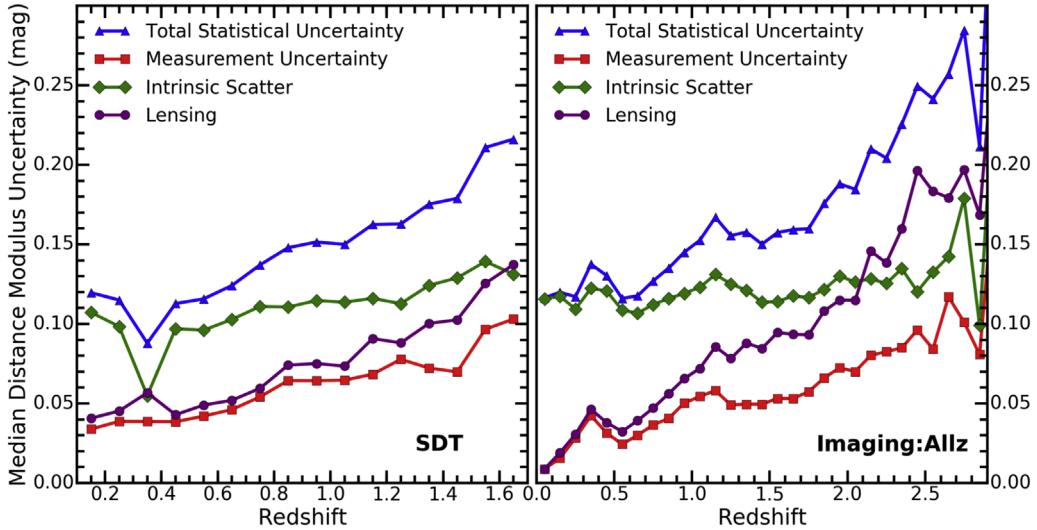
Figure 13 (left) presents the  $w_0-w_a$  68% and 95% confidence contours for the simulated SDT, SDT\*, and SDT\* Highz

surveys. These contours illustrate how a slight modification of the classification criteria presented by the SDT report (see Section 3.1) can lead to an increase in FoM<sub>tot,opt</sub>, whereas moving time to focus on the medium tier of the survey results in a more significant impact.

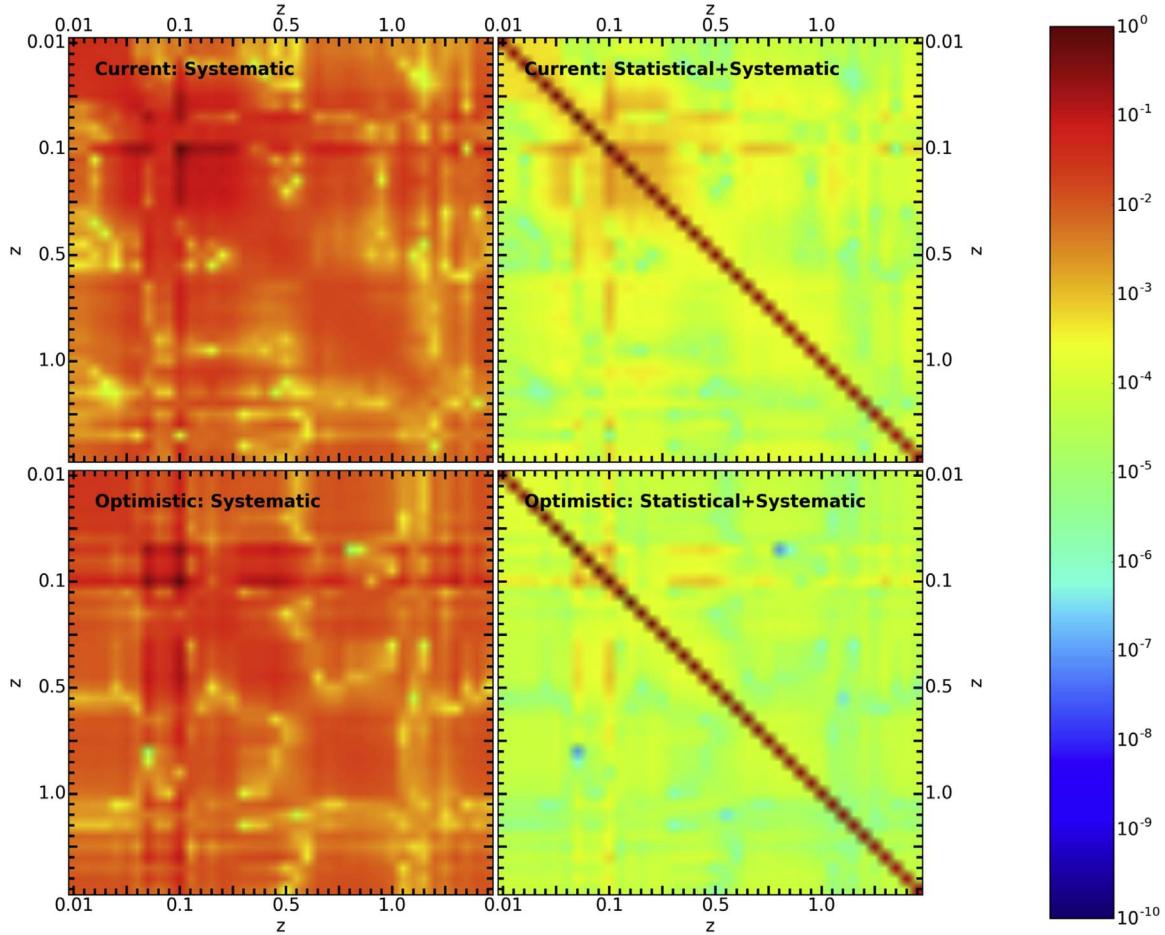
As an informative worst-case scenario, the SDT Imaging simulation replicates a situation where all IFC-S data are determined to be unusable, but only after completion of the *WFIRST* mission. As a result, only the imaging data as part of the SDT survey would be used for cosmological analyses. Unsurprisingly, this SN survey produces too few SNe Ia at  $z < 0.6$  and delivers a low FoM<sub>tot,opt</sub> = 105.

The Imaging:Allz simulation has FoM<sub>tot,opt</sub> = 388 and is our most successful imaging-only strategy. It is a three-tier imaging strategy that uses four broadband filters (*RZYJ* or *YJHF*). This strategy results in  $>5$  times as many SNe Ia as any IFC-S strategy and has FoM<sub>stat</sub> = 622. Zero-point uncertainties are some of the largest systematic uncertainties for this strategy (see Figure 11), as with all imaging strategies.

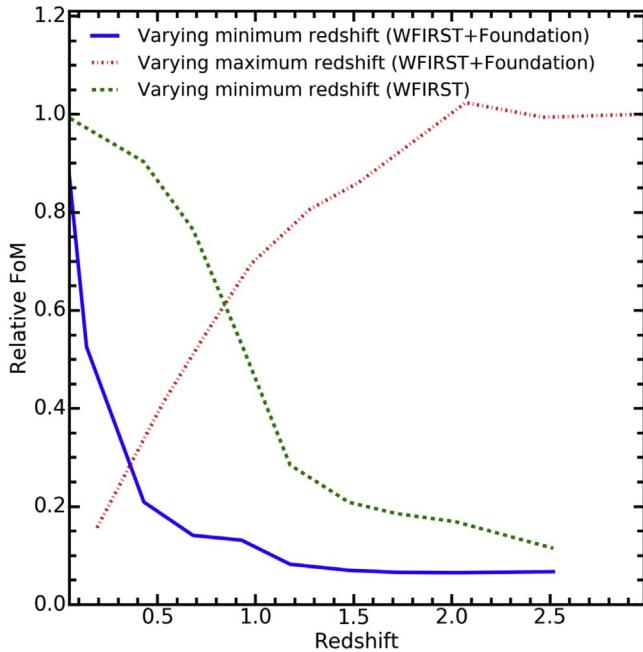
For both the SDT and Imaging:Allz simulations, Figure 14 shows the contribution of measurement (red line), lensing (purple line), and intrinsic scatter (green line) uncertainties on



**Figure 14.** Contribution to the median distance modulus uncertainty per SN within a 0.1 redshift bin, from the measurement uncertainty (red squares), intrinsic scatter (green diamonds), and lensing uncertainty (purple circles) components for both the IFC-S-focused SDT (left) and Imaging:Allz (right) strategies. The blue triangles show the combined statistical uncertainties added in quadrature.



**Figure 15.** Redshift correlation matrices for the Imaging:Allz strategy. Within a given matrix, the data have been normalized to the maximum array value (absolute values are displayed). A log color bar is used to show the relative distribution. Left panels show the correlation matrix for current (top) and optimistic (bottom) systematic uncertainties. Right panels show the full correlation matrix, which includes statistical uncertainties. Note that although the off-diagonal elements from the systematic uncertainties appear to have a smaller magnitude than the diagonal elements from the statistical uncertainties, the cumulative impact of the off-diagonal elements can be greater than the impact of the diagonal elements on the FoM.



**Figure 16.** Effect on the  $\text{FoM}_{\text{stat}}$  if the redshift range of the Imaging:Allz survey is modified. The red dotted-dashed line shows the ratio with a high- $z$  cut and minimum redshift  $z \approx 0.02$ . The blue line shows the ratio with a low- $z$  cut and maximum redshift  $z \approx 3$ . The green dashed line is similar to the blue line, except that the low- $z$  cut applies only to the WFIRST component of the survey such that the “Foundation-like” ( $0.01 < z < 0.1$ ) data always remain.

the total statistical uncertainty. Within the SDT simulation, the dominant uncertainty is that of intrinsic scatter, with the lensing uncertainty becoming comparable at  $z \approx 1.7$ . At higher redshift (i.e.,  $z \geq 2$ ), the Imaging:Allz simulation shows that lensing is the dominant source of uncertainty.

Although measurement uncertainties for high-redshift SNe are significant, Figure 14 clearly illustrates that these uncertainties are not dominant.

Figure 15 shows the redshift correlation matrices for the Imaging:Allz strategy. The statistical uncertainty dominates across all redshifts with some systematic contribution around  $0.1 < z < 0.3$ . While the statistical uncertainty per redshift bin is larger than the systematic uncertainty, the impact of systematic uncertainties depends on the covariance between redshift bins. The effects of these systematic uncertainties are significantly reduced when considering their optimistic values.

To demonstrate the dependence of  $\text{FoM}_{\text{stat}}$  on the redshift range of the Imaging:Allz sample, Figure 16 shows how the FoM changes when we vary the minimum and maximum redshifts of the sample. We find that at a maximum redshift of  $z = 1.5$  (red line),  $\text{FoM}_{\text{stat}}$  is  $\sim 86\%$  of the full-sample  $\text{FoM}_{\text{stat}}$ , and that  $\text{FoM}_{\text{stat}}$  no longer increases with additional SNe past  $z \approx 2.0$ . The dependence on the minimum redshift (blue line) shows the importance of a low- $z$  sample. For instance, a cut at  $z = 0.4$  reduces  $\text{FoM}_{\text{stat}}$  by  $\sim 79\%$ . In addition, we note the importance of the ground-based “Foundation-like” data (i.e.,  $z \leq 0.1$ ); their removal from the survey reduces  $\text{FoM}_{\text{stat}}$  by  $47\%$ .

When comparing the Imaging:Allz and SDT\* Highz  $\text{FoM}_{\text{stat}}$  values, it is interesting to note that even though the final cosmology sample of the Imaging:Allz survey contains a factor of 6 more SNe Ia at  $z \geq 1.2$ , the addition of these SNe increases the relative  $\text{FoM}_{\text{stat}}$  value by only  $\sim 79\%$ . The reason for this relatively small improvement in  $\text{FoM}_{\text{stat}}$  is that the majority of additional SNe in the Imaging:Allz strategy are at  $z \geq 1.5$ , and as

shown in Figure 16, these high- $z$  SNe contribute little to  $\text{FoM}_{\text{stat}}$ . This effect is likely due to the fact that at high  $z$ , dark energy becomes dynamically unimportant in conventional models, and within our chosen  $w_0-w_a$  parameterization, there is little room to vary  $w(z)$  in a way that substantially changes high-redshift distances relative to those at lower redshift. Statistical and systematic uncertainties are also greater at higher  $z$  (see Figure 14), and the inclusion of zero-point uncertainties within each imaging scenario also means that such strategies are hitting more of a systematic floor than those that are IFC-S-focused.

We have examined four low- $z$  imaging-only strategies with two tiers, in which time from the IFC-S and deep tier have been redistributed among the shallow and medium tiers of the discovery survey. This method has allowed for the addition of several filters and an increase in each tier’s observational area. Each of these strategies failed to meet the required number of SNe Ia (as outlined in the SDT report) at  $z > 1.2$ . The Imaging:Lowz survey lacks the desired number at  $z < 0.6$  and  $z > 1.1$ , resulting in the small  $\text{FoM}_{\text{tot,opt}} = 133$ .

For the two high- $z$  imaging-only strategies, time from the IFC-S and the shallow-tier observations was redistributed to the medium and deep tiers and observations were made with additional filters. The Imaging:Highz\* survey is one of the most successful imaging-only strategies with  $\text{FoM}_{\text{tot,opt}} = 352$ . It also has very small statistical uncertainties with  $\text{FoM}_{\text{stat}} = 704$ , the largest statistical-only FoM for any strategy examined.

Figure 12 shows the FoM estimate for each strategy. The SDT Imaging and Imaging:Lowz strategies are less precise than other strategies since their  $\text{FoM}_{\text{tot,opt}}$  values are close to or below the  $\text{FoM}_{\text{tot,curr}}$  values of most other surveys. These strategies are clearly inferior to other options.

Excluding the SDT Imaging and Imaging:Lowz strategies, IFC-S-focused strategies are the least successful. Since the SDT and SDT\* strategies are equivalent except in classification, the final systematic uncertainty for either strategy would be essentially equivalent. Therefore, the SDT\* strategy is superior to that of the SDT report. Of the IFC-S-focused strategies, however, the SDT\* Highz survey is the most successful.

For the remaining strategies, it is difficult to assign a clear ranking. The effectiveness of each strategy has different dependencies on specific improvements in systematic uncertainties. The  $\text{FoM}_{\text{tot,curr}}$  value for many WFC-focused strategies are comparable, an effect that can likely be attributed to the fact that at this point, we are becoming systematics limited.

Our simulations provide important information about where to focus efforts. Considering the  $\text{FoM}_{\text{tot,opt}}$  values, the top three strategies are Imaging:Allz, Imaging:Highz\*, and Imaging: Highz+, which all have similar  $\text{FoM}_{\text{tot,opt}}$  values. There is no obvious optimal strategy among those investigated here and with current knowledge. Importantly, imaging-only strategies have been shown to constrain dark energy as well as, and even better than, our current IFC-S strategies. Figure 13 (right) presents the  $w_0-w_a$  68% and 95% confidence contours for the simulated SDT, Imaging:Allz, and Imaging:Highz\* surveys. These contours illustrate how competitive imaging-only strategies are with respect to an IFC-S-focused strategy.

The wavelength-dependent calibration uncertainty for the IFC-S system is currently large enough to significantly hamper the effectiveness of any IFC-S-focused strategy. We have optimistically assumed that by launch it will improve by a factor of 17 (see Figure 11). However, since no clear path has

been presented for this improvement, we have also investigated how factors of 2 and 10 improvement (i.e., 25 and 5 mmag per 7000 Å) affect the final FoM values of the SDT\* strategy. For improvement factors of 1 (no improvement; current value), 5, 10, and 17 (optimistic value), we find  $\text{FoM}_{\text{tot}} = 166, 171, 209$ , and 216, respectively. For these calculations, the values of the other systematic uncertainties (i.e., nonlinearity, host-mass evolution, population drift, and intrinsic scatter) are set to their optimistic values. It is clear that a precision of at least 5 mmag per 7000 Å is required for optimal implementation of an IFC-S-focused strategy.

In addition, imaging-only strategies like Imaging:Allz and Highz\* may have an advantage because their data can be divided into subsamples for further systematic studies, including high- and low- $z$  host-mass and high- and low- $z$  Galactic extinction studies. If new effects are found, such as  $\beta(z)$  or a better host-mass function, then imaging-only strategies with superior statistics will prove better for measuring these additional parameters.

## 8. Discussion and Future Work

The strategies outlined in this paper illustrate how the *WFIRST* SN survey can be modified to increase the number of SNe Ia and to increase the redshift range over which they are found. These strategies are intended as reference options that can be updated and expanded upon to perform more rigorous optimizations.

Future optimization of the survey may include trading depth or area and adjusting the cadence of the light curves. In addition, the current redshift distribution proposed by the SDT report could be further optimized with relatively small modifications to the survey. Below we discuss in more detail some of the ways in which the survey could be optimized.

Our analysis currently assumes that the redshift of each SN or its host galaxy is perfectly known. In reality, we will know the redshift of the SN with varying levels of accuracy based on how well it is determined. The accuracy of the redshift affects observation choices such as exposure times, the precision of classification routines, and potential biases that propagate to the Hubble diagram. Meanwhile, the uncertainty in the redshift propagates directly to constraining cosmological parameters.

To obtain redshifts, the *WFIRST* strategy will likely use a combination of high-resolution spectroscopic host-galaxy redshifts, lower-resolution *WFIRST* grism host-galaxy redshifts, SN+galaxy photometric redshifts, and spectroscopic redshifts from the SNe themselves. Further complicating the issue, the redshifts (and their uncertainties) will be updated and improved during the course of the survey.

If redshifts were to be obtained primarily via photometry of the SN Ia and its host galaxy (i.e., photo- $z$ ), then it is important to understand how well these photo- $z$  values can be determined.

The photometric redshift noise,  $\Delta z/(1+z)$ , scales roughly as  $S/N^{-1}$ , and at some point, a “degeneracy floor” is reached because of the degeneracy between intrinsic color and redshift. Ideally, we would obtain uncertainties no greater than  $0.02 \times (1+z)$ , which may be possible for a subset ( $\sim 20\%$ ) of SNe found within luminous red galaxies where the S/N is high and calibration biases are well controlled. The majority of SN galaxies, however, will possess photo- $z$  uncertainties within the range of  $(0.03-0.04) \times (1+z)$  (see the weak-lensing requirements of the SDT report); the extra scatter in the Hubble diagram is high, and there are huge pathologies between redshift and color, leading to significant issues. Modern photo- $z$  codes contain many

assumptions and are constantly evolving, and thus further work is required within this field to improve the results found.

A full analysis of the various effects associated with each method of redshift determination (e.g., photo- $z$ , high-resolution host-galaxy spectroscopy, host-galaxy grism data, SN spectra) requires an accurate assessment of the redshift catalogs present at the beginning of the SN survey, the ground-based resources available during the survey, the exact *WFIRST* survey strategy, and resources available upon completion of the survey. With estimates of the available resources, we can assign redshifts with appropriate accuracy to each simulated SN and determine how each survey strategy is affected.

Our simulations have followed the current NASA mandate that all SN discovery and follow-up observations be performed exclusively by *WFIRST*. However, we will likely perform follow-up observations of *WFIRST*-discovered SNe using ground-based facilities; furthermore, there will be *WFIRST* follow-up observations of SNe discovered in ground-based surveys. There could be significant efficiency gains, for example, by using SNe Ia discovered by LSST (Ivezic et al. 2008) and by allowing an overlap in the LSST and *WFIRST* SN fields. In IFC-S-focused strategies, the inclusion of LSST data would essentially remove the need for a shallow imaging tier (potentially up to  $z \approx 0.8$ ), meaning that at low  $z$ , *WFIRST* would be used for spectrophotometric follow-up observations of the LSST SN sample only. The discovery search time from the shallow tier could then be redistributed to the higher- $z$  imaging tiers, increasing the likelihood of detection for SNe at these redshift. Conversely, time gained from the shallow imaging tier could be used to obtain additional low- $z$  *WFIRST* spectra and thus improve the spectrophotometric calibration between the two surveys. In addition, if SN classifications can be obtained using ground-based 8–30 m telescopes (e.g., Gemini, TMT), then low- $z$  IFC-S spectra are no longer required (assuming that the relative calibration between the facilities is well known) and additional time can be spent observing SNe with  $z \geq 0.8$ .

For imaging-driven strategies, the inclusion of LSST data would remove the shallow imaging tier, allowing more time to be dedicated to SNe at high  $z$ . Alternatively, a series of short *WFIRST* WFC exposures similar to that of LSST might enable the reduction of calibration systematics (including variable seeing, extended periods of bad weather affecting cadence, etc., for ground-based data) between the two surveys, providing a necessary template between ground-based and space-based data sets. Future simulations will examine the possibility of multiple scenarios for using ground-based observatories to enhance the *WFIRST* SN survey.

The use of the grism has not yet been fully explored or simulated. We performed preliminary simulations, finding that grism spectroscopy would be effective for classification, but only with longer exposure times than those for the IFC-S-focused strategies (see Jones et al. 2013 for a detailed examination of the *HST* WFC3 grism for this purpose). More extensive simulations are necessary to determine if the grism is useful for the SN survey.

Usage of both the IFC-S and WFC imaging components with their current five day cadence would require considerable resources for scheduling.

Within five days, data would have to be downloaded, processed, searched for transients, objects fit and selected, IFC-S follow-up observation schedules built and sent, and finally the instrument set to observing again.

A longer cadence of seven days, or even a flexible cadence, may simplify some of the scheduling issues and have little to

no scientific impact on the mission. Modified cadence investigations are needed for strategies using the IFC-S.

The idea of using parallel observing for the IFC-S and WFC must also be considered. Parallel observing would allow *WFIRST* to operate both the WFC and IFC-S at the same time. Preliminary calculations suggest that given the huge number of IFC-S observations, the WFC imaging fields will be almost completely covered. However, within this basic calculation, we have ignored the possibility of selecting particular roll angles or the likely correlations between angles for a given SN. This means that there will be different cadences for each SN, and thus patchy and incomplete imaging of that SN.

A WFC imaging-focused survey that utilizes parallelization to obtain IFC-S spectra during deep  $H$ - or  $F$ -band observations (required to obtain SNe at high  $z$  with  $S/N \geq 10$ ) could, however, be an important hybrid strategy. Obtaining even a small fraction (10%–15%) of SN Ia spectra may provide data vital to the analysis, such as improving the underlying SALT2 SED model, obtaining detailed information about the host-galaxy environment, allowing characterization of systematic effects like population drift, and helping us to potentially explore the effects of unknown systematics. Future work will include the use of parallel fields within our simulations and the creation of such hybrid strategies.

An important limitation in our analysis is the training sample used to determine the underlying SED model. As described by Astier et al. (2014), the SN model uncertainty can be reduced by using the same rest-frame wavelength range at all redshifts. For a rest-frame wavelength range of 2000–25000 Å, corresponding to the current extended SALT2 spectral model (the nominal model is 2800–7000 Å), the mean effective wavelength for *RZYJHF* filters will fall in redshift ranges of  $z < 2.10$ ,  $z < 3.35$ ,  $z < 4.45$ ,  $z < 5.5$ ,  $z < 7.00$ , and  $z < 8.4$ , respectively (for the nominal case, the redshift ranges are  $z < 1.21$ ,  $0.24 < z < 2.11$ ,  $0.56 < z < 2.89$ ,  $0.86 < z < 3.64$ ,  $1.29 < z < 4.71$ , and  $1.69 < z < 5.71$  for the *RZYJHF* filters, respectively).

There have been several efforts to obtain NIR SN Ia data (e.g., Krisciunas et al. 2004, 2017; Wood-Vasey et al. 2008; Contreras et al. 2010; Stritzinger et al. 2011; Friedman et al. 2015). In total, there are  $\approx 200$  SNe Ia with published NIR light curves, with some objects observed independently by multiple groups including the Center for Astrophysics (Friedman et al. 2015) and the CSP (Krisciunas et al. 2017). While most of these data are for low- $z$  SNe, the RAISIN (anagram of SN IA in the IR) project (Kirshner 2012) has collected rest-frame NIR data on  $\sim 45$  moderate-redshift ( $0.2 < z < 0.6$ ) SNe Ia with *HST*/WFC3. In addition to contributing to the NIR model, these data will be useful for investigating systematic uncertainties related to intrinsic scatter, dust, and color.

There are additional calibration issues for the WFC and IFC-S that need to be taken into consideration when examining systematic uncertainties. Required instrumental characterizations include persistence, flat fields, astrometric mapping of detector to the sky, out-of-band stray light, etc. Initial assessments show that these calibration uncertainties are all second-order systematics that are significantly below the ones included in our current analysis, but should be reviewed in future work.

When assessing multiple survey strategies, it is best to have a single, predefined metric by which one can compare. With multiple metrics, one can generally choose the metric that is optimal for a particular strategy. That said, there can be critical aspects of a problem that do not affect a metric. For instance, the DETF FoM

that we use to compare strategies does not contain any information related to mission cost/risk or enabling ancillary science.

Furthermore, the DETF FoM is not the only metric by which we can optimize our understanding of dark energy. For instance, eigenvectors (Huterer & Starkman 2003) have been a popular approach (although Linder & Huterer 2005 argue that something like the DETF FoM is sufficient for most needs). It will be straightforward to implement modified dark-energy characterizations into our simulations, but additional cosmology-fitting tools may be needed to implement a more comprehensive analysis.

The SN survey defined within the SDT report limits the number of SNe Ia at high  $z$  and focuses on achieving a larger sample within  $0.2 \leq z \leq 0.6$ . Our imaging-only strategies, on the other hand, place no limit on the number of SNe Ia within a given redshift bin and explore out to  $z \leq 3.0$ . As our surveys have not been optimized, we have not specifically considered the effects of focusing observations within any given redshift range. However, our preliminary studies have indicated (see Section 7) that an increase in the fraction of SNe Ia with higher redshifts (i.e.,  $z > 1.2$ ) does not necessarily provide the same fractional increase in a survey’s FoM (see Figure 16). This is likely due to the nature of dark energy and the  $w_0-w_a$  parameterization. Variations on the redshift distribution should be considered in future optimization studies.

Our work has used constraints on the cosmological parameters from both the BAO (Anderson et al. 2014) and CMB (Planck Collaboration et al. 2016) data sets. However, there is ongoing work to include external constraints from projections of CMB S4 (Abazajian et al. 2016) and future BAO missions such as DESI (DESI Collaboration et al. 2016). Weinberg et al. (2013) showed that the impact of Stage 4 SN constraints on the FoM has a strong dependence on the relative constraints from the Stage 4 BAO and future weak-lensing probes. Future work is needed to replicate this analysis in the context of various SN strategies.

Although the key objective of the *WFIRST* SN survey is to obtain data on SNe Ia, it will image a significant fraction of the sky with a cadence of five days. Within these data, a plethora of other transient events may be found, which could lead to significant scientific gains (e.g., Scolnic et al. 2017). When designing the survey, such synergies with other areas of astrophysics should be considered.

## 9. Conclusion

Using open-source tools, including newly created ones,<sup>33</sup> we have produced the first realistic simulations of the *WFIRST* SN survey. We examined 11 strategies in detail, including the strategy presented in the SDT report. For each simulated SN survey strategy, several statistical and systematic uncertainties have been examined and included in order to calculate the  $FoM_{tot,opt}$  value, which we have used as our final measure of success.

One key aspect of this paper is that the proper incorporation of slew-and-settle overheads (which were omitted in the forecasts in the SDT report) drastically reduces the efficiency of the low-redshift, wide-area tier of the SDT’s proposed SN survey. These overheads remain an uncertain aspect of mission performance, and optimal design of the *WFIRST* SN survey will have a strong dependence on their values.

For our forecasts of the SDT survey and all additional strategies examined, we have retained the SDT’s proposed exposure times minus the 42 s overhead (see Table 3). For each

<sup>33</sup> See <https://jet.uchicago.edu/blogs/WFIRST/>.

strategy presented, the total observing time is always six months. Recovered survey time from not using the IFC-S, or the removal of a survey tier, goes into increasing the survey area of the remaining tiers and/or the addition of filters.

Examination of the results produced by our SDT simulation (see Section 5.1) shows that this strategy results in fewer SNe Ia than outlined in the SDT report. The selection efficiency of the SDT strategy is low, and the noise is significantly underestimated, resulting in many SNe Ia being cut or misclassified in the final sample by the strict SDT classification routine outlined in Section 3.1. With  $\text{FoM}_{\text{tot,opt}} = 158$ , this is one of the least successful survey strategies investigated (see Figure 12 and Table 13). Modification of the selection criteria (SDT\*) to account for statistical fluctuations in the rise in flux between epochs, and the restriction of colors to reduce CC SN contamination (and exotic events), increased  $\text{FoM}_{\text{tot,opt}}$  to 216. Even within the SDT\* simulation, however, there are still too few SNe Ia selected at  $z < 0.6$ , due in part to the short 13 s exposure of the shallow imaging tier. Time devoted to the shallow tier of the SDT\* survey is placed into the medium tier to produce our SDT\* Highz scenario. For this strategy,  $\text{FoM}_{\text{tot,opt}}$  improves to 236.

Our imaging-only strategies include the shallow+medium tiers (suffix Lowz), all three imaging tiers (SDT Imaging and Imaging:Allz), and the medium+deep tiers (suffix Highz). Tier areas were increased and additional filters added to compensate for removing the IFC-S and/or discovery tiers. Additional filters resulted in a broader coverage of the rest-frame optical wavelengths (via the addition of R and Z bands), a region where the SALT2 SED model is well defined, and extension into the rest-frame NIR (via the addition of the F band). The Imaging:Allz and Imaging:Highz\* surveys have  $\text{FoM}_{\text{tot,opt}}$  values of 388 and 352 (respectively), making them some of the most successful strategies. The Imaging:Highz\* survey also has the highest  $\text{FoM}_{\text{stat}}$  value at 704. For many of the imaging-only strategies, the number of SNe Ia within the final sample is significantly higher than that obtained by the SDT strategy.

Using  $\text{FoM}_{\text{tot,opt}}$  to measure the success of each strategy, there is no clear winner. The Imaging:Allz, Imaging:Highz\*, and Imaging:Highz+ simulated surveys all have similar current and optimistic FoM values. Although our work suggests that WFC-focused strategies are more successful than IFC-S-focused ones, further constraints on systematic uncertainties are needed, especially for the IFC-S.

There are several additional concerns related to an IFC-S-focused strategy. Specifically, an IFC-S-focused strategy requires active target selection (likely with human decisions included), which increases operational resources and locks in selection bias at the time of target selection. The ability to produce high-precision spectrophotometry with an IFC has yet to be demonstrated, resulting in a higher risk of reaching systematic uncertainty goals than with an imaging-only strategy. A further limitation of IFC-S-focused strategies is the limited ability to divide their relatively small sample sizes into subsamples for systematic studies.

A parallel imaging survey conducted during IFC-S observations could mitigate these concerns, but the power of such a survey relative to an optimized imaging-only survey has not yet been assessed.

One virtue of an IFC-S-focused strategy that has not yet been considered is the stacking of multiple low-S/N spectra. Each individual spectrum may not be useful for classification, but stacked spectra may be useful for measuring the evolution of

spectral diversity with redshift. In general, IFC-S-focused strategies may afford advantages in calibration, in that they will have less sensitivity than filter photometry to SN SED evolution, and spectra can provide more information in determining population drift systematics.

While the strategies we have presented are not fully optimized, they provide a broader understanding of the possibilities for the survey. Moreover, at this stage in the mission, such an investigation is critical for mitigating risk and ensuring the ultimate success of *WFIRST*. Our initial investigations have determined that there is no single correct survey scenario for the mission, yet all our top-performing strategies provide a significant improvement in comparison to current surveys that utilize SNe Ia as cosmological probes, and they progress toward that which is expected by a Stage 4 experiment.

Our work has focused on establishing a reliable and reproducible set of baseline strategies, using well-defined methods and software tools. In future work, we plan to continue optimizing the *WFIRST* mission to achieve a more successful strategy.

We would like to dedicate this paper to the memory of Neil Gehrels. His hard work and enthusiasm as the *WFIRST* Project Scientist helped make the mission possible.

This manuscript is based upon work supported by the National Aeronautics and Space Administration (NASA) under contract No. NNG16PJ34C issued through the *WFIRST* Science Investigation Teams Program. It was also supported in part by the U.S. Department of Energy under contract DE-AC02-76CH03000. Analysis was done using the Midway-RCC computing cluster at the University of Chicago.

R.H., D.S., and R.J.F. were supported in part by NASA grant 14-WPS14-0048. The UCSC group is supported in part by fellowships to R.J.F. from the Alfred P. Sloan Foundation and the David and Lucile Packard Foundation. D.S. and R.K. acknowledge support from the Kavli Institute for Cosmological Physics at the University of Chicago through National Science Foundation (NSF) grant PHY-1125897 and an endowment from the Kavli Foundation and its founder Fred Kavli. Hubble/KICP Fellow D.S. is also supported by NASA through Hubble Fellowship grant HST-HF2-51383.001 awarded by the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. V.M. was supported in part by the Charles E. Kaufman Foundation, a supporting organization of the Pittsburgh Foundation. Supernova cosmology at the Harvard College Observatory is supported in part by the NSF through grants AST-1516854 and AST-1211196, and NASA grant NNX15AJ55G. R.P.K. and A.A. were supported in part by the RAISIN1 GO-13046 and RAISIN2 GO-14216 *HST* grants, which were administered by STScI. A.V.F. and P.L.K. received support from the Christopher R. Redlich Fund, the TABASGO Foundation, the Miller Institute for Basic Research in Science (U.C. Berkeley), and NASA/*HST* grants GO-14041 and GO-14199.

We would like to thank A. Friedman for providing his compilation of public NIR photometry and C. Heinrich for their contribution to our cosmological analysis.

In addition, we would also like to thank our referee, Prof. D. Weinberg. Professor Weinberg's thoughtful comments and insight served to significantly improve the paper.

**Table 14**  
Each IFC-S Bin between 0.42 and  $2.1 \mu\text{m}$

Maximum Wavelength (Å)	Minimum Wavelength (Å)	PSF FWHM (pixels)	Zodiacal Noise ( $e^- \text{s}^{-1} \text{pixel}^{-1}$ )	Thermal Noise ( $e^- \text{s}^{-1} \text{pixel}^{-1}$ )
4200.00	4209.35	1.540	0.000	0.000
4209.35	4218.76	1.541	0.000	0.000
4218.76	4228.23	1.542	0.000	0.000
4228.23	4237.76	1.543	0.000	0.000
4237.76	4247.35	1.543	0.000	0.000
4247.35	4257.00	1.544	0.000	0.000
4257.00	4266.71	1.545	0.000	0.000
4266.71	4276.49	1.546	0.000	0.000
4276.49	4286.34	1.546	0.000	0.000

**Note.** The minimum and maximum wavelength ranges for each bin are given along with the FWHM in pixels and associated sources of noise.

(This table is available in its entirety in machine-readable form.)

## Appendix A IFC-S Extended Table

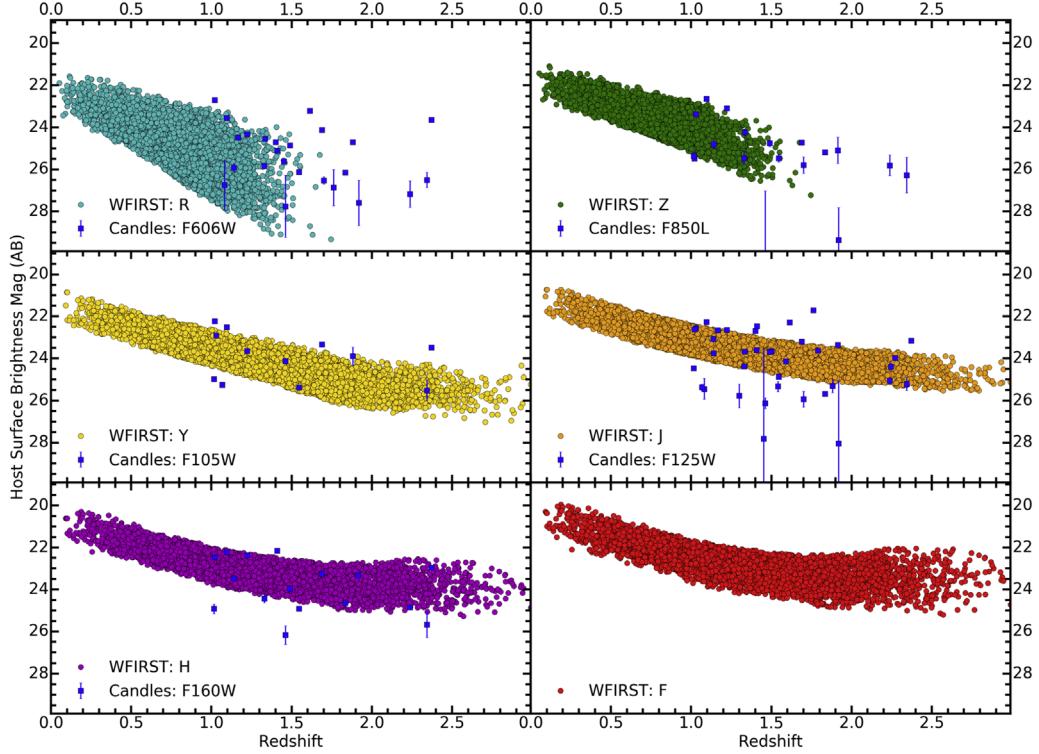
Table 14, presented within this appendix, provides a complete listing of each IFC-S bin used within our simulations. A machine-readable version of this table is available.

## Appendix B Host-galaxy Surface Brightness Library

Using CANDELS data, we obtained host-galaxy surface brightnesses at the site of  $\sim 40$  SN events. Local surface brightness fluxes were obtained on images without SN light, using a  $0.^{\prime\prime}2$  radius aperture with the following *HST* filters:

*F606W*, *F775W*, *F850L*, *F105W*, *F125W*, *F140W*, and *F160W*. Rest-frame SED galaxy models were redshifted and warped so that their synthetic photometry agreed with the observed *HST* photometry. The resulting best-fit SEDs were used to compute synthetic photometry in the *WFIRST* filters as a function of redshift.

SNANA can generate galaxies with arbitrary profiles and magnitudes; however, since our *HST* data were measured only at the site of the SN, we forced our simulated SN location to be at the center of its host galaxy with a Sérsic profile index of 0.5 and a slight offset applied to the magnitude. To test how well this approximation works, Figure 17 compares the simulated *WFIRST* host-galaxy surface brightnesses to that of



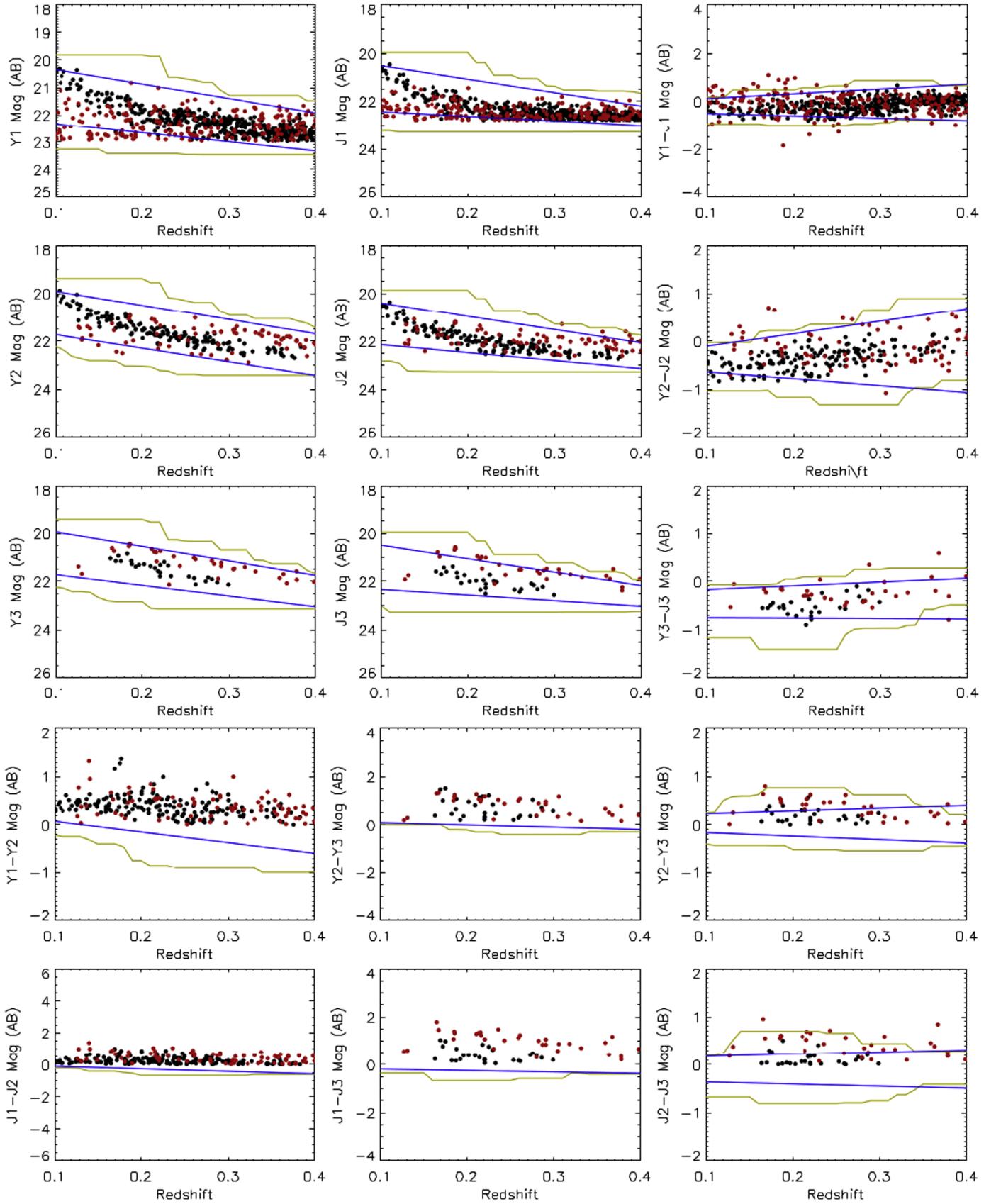
**Figure 17.** Simulated host-galaxy surface brightness vs. redshift for each *WFIRST* filter, and comparison with data from CANDELS.

corresponding data in a CANDELS filter. Although not perfect, this simple approximation was implemented throughout our study.

### Appendix C Photometric Cuts in the IFC-S Data

The SDT report describes a number of photometric cuts to be implemented on imaging data used in the discovery of SNe. These cuts are listed within Section 3.1 and require that both the color and flux of a supernova be examined from one epoch to the next.

Within our analysis of IFC-S-focused strategies, the photometric cuts were applied to imaging data by defining acceptable color and rise value (i.e., increase in flux) ranges from one epoch to the next via an iterative process. The ranges were defined using the simulated SN sample as a whole and encompassed the impact of both the photometric and intrinsic scatter of the SNe. Figure 18 depicts the photometric cuts applied to SNe within the shallow tier of the SDT survey. A supernova that occupies the resultant ranges has a color that is consistent (with respect to the simulated sample) with an SN Ia at its host-galaxy redshift and has an acceptable flux increase from one epoch to the next.



**Figure 18.** Photometric cuts applied to SNe Ia (black circles) and CC SNe (red circles) discovered within the shallow imaging tier of the SDT survey strategy. The top row shows the  $Y$  and  $J$  magnitudes within the first imaging epoch ( $Y1/J1$ ) and their corresponding colors vs. redshift. The second and third rows represent the same information but for the second and third epochs, respectively ( $Y2/J2$  and  $Y3/J3$ ). The fourth row shows the rise data for the  $Y$  band, with the fifth showing rise values for the  $J$ . The blue and gold lines represent rise or color-cut regions, defined by the scatter within the SN sample.

The photometric uncertainties associated with each individual SN, however, were not included when defining these ranges. We note that the exclusion of individual photometric uncertainties leads to rather restrictive color and rise values, but at the time this was our best interpretation of the SDT report. The method of selection adopted within our SDT\* strategy is likely more representative of what was originally planned, although not stated.

## Appendix D Measuring the Population Drift

For a fixed light-curve shape, the evolutionary change of intrinsic color with redshift has been shown to occur with marginal significance (Foley et al. 2012; Maguire et al. 2012; Milne et al. 2015). The color variation has been demonstrated to be empirically correlated with SN ejecta velocity (Foley & Kasen 2011; Foley et al. 2011; Foley 2012; Mandel et al. 2014) and therefore measuring the ejecta velocity can remove this potential redshift-dependent bias and improve the distance precision.

Since the color change correlation with ejecta velocity is restricted to  $\lambda < 4500 \text{ \AA}$  in the rest frame, SNe across the *WFIRST* redshift range will be affected by different amounts.

Although we could exclude all data blueward of  $4500 \text{ \AA}$ , this would result in using only 35% of the pixels for a  $z = 1.5$  SN Ia, greatly diminishing the distance precision of these SNe. Alternatively, we should be able to use all data if we can measure a precise velocity. We have already shown that this measurement is possible for spectral resolution  $R > 75$ , if the data are of high quality (Foley 2013). In fact, an  $R = 130$  spectrum has already been used to measure a Si II velocity for a  $z = 1.55$  SN Ia (Rodney et al. 2012).

The most important feature for measuring the ejecta velocity is Si II  $\lambda 6355$ , the hallmark feature of SNe Ia. This feature is blueshifted to  $\sim 6100 \text{ \AA}$  in the rest frame, making it accessible for all IFC-S spectra and all grism spectra at  $z > 1.2$ . To determine if we can measure the ejecta velocity with realistic IFC-S data, we measured the Si II velocity for all long-exposure spectra in the final SDT sample.

We found that the typical velocity uncertainty is  $1000 \text{ km s}^{-1}$ , with a  $\sim 5\%$  failure rate. The ejecta velocities are biased low by  $\sim 500 \text{ km s}^{-1}$ , although this bias can be corrected with measurements from higher-resolution spectra (perhaps from the ground) or from simulations.

This large velocity uncertainty propagates into a 0.10 mag distance modulus uncertainty (Foley et al. 2011). This relatively large uncertainty (comparable to the total distance uncertainty) is caused by a combination of low resolution and low S/N of the IFC-S spectra. For instance, at infinite S/N, we find a scatter of  $340 \text{ km s}^{-1}$  (close to the limit from galactic rotation) and a bias of  $180 \text{ km s}^{-1}$ . For the grism, the uncertainty decreases to  $800 \text{ km s}^{-1}$  for the same (binned) S/N as the long-exposure IFC-S spectrum, indicating that most of the uncertainty is caused by the low S/N.

This spectroscopic study shows that *WFIRST* has the potential to measure the SN ejecta velocity, but for this velocity to be helpful for improving distance estimates, we have found that we require  $S/N > 20$ , which is beyond the current SDT design. However, a slight modification to the survey design and/or strategy (higher resolution and/or a higher S/N spectrum) would alleviate this problem while simultaneously improving spectral classification. Although this

modification requires additional exposure time per SN, it would reduce the statistical uncertainty of each SN.

Other studies have indicated that additional spectral features, including flux ratios from narrow wavelength bands, can improve distance measurements slightly (Bailey et al. 2009; Blondin et al. 2011). Using the simulated long-exposure spectra, we measured the Bailey et al. (2009) flux ratios and find that the uncertainties are generally 20%, which propagates into a  $\sim 0.4$  mag distance modulus uncertainty. This large uncertainty rules out using such flux ratios from the IFC-S at the current S/N.

## ORCID iDs

- R. Kessler [DOI](https://orcid.org/0000-0003-3221-0419) <https://orcid.org/0000-0003-3221-0419>
- A. Avelino [DOI](https://orcid.org/0000-0002-2938-7822) <https://orcid.org/0000-0002-2938-7822>
- R. C. Bohlin [DOI](https://orcid.org/0000-0001-9806-0551) <https://orcid.org/0000-0001-9806-0551>
- A. V. Filippenko [DOI](https://orcid.org/0000-0003-3460-0103) <https://orcid.org/0000-0003-3460-0103>
- S. W. Jha [DOI](https://orcid.org/0000-0001-8738-6011) <https://orcid.org/0000-0001-8738-6011>
- P. L. Kelly [DOI](https://orcid.org/0000-0003-3142-997X) <https://orcid.org/0000-0003-3142-997X>
- K. Mandel [DOI](https://orcid.org/0000-0001-9846-4417) <https://orcid.org/0000-0001-9846-4417>
- S. A. Rodney [DOI](https://orcid.org/0000-0003-1947-687X) <https://orcid.org/0000-0003-1947-687X>

## References

- Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, arXiv:1610.02743
- Alam, S., Ata, M., Bailey, S., et al. 2017, *MNRAS*, **470**, 2617
- Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv:astro-ph/0609591
- Aldering, G. 2001, LBNL Report LBNL-51157, <http://www-supernova.lbl.gov/aldering/zodi.pdf>
- Aldering, G., Adam, G., Antilogus, P., et al. 2002, *Proc. SPIE*, **4836**, 61
- Anderson, L., Aubourg, É., Bailey, S., et al. 2014, *MNRAS*, **441**, 24
- Astier, P., Balland, C., Brescia, M., et al. 2014, *A&A*, **572**, A80
- Bacon, R., Brinchmann, J., Richard, J., et al. 2015, *A&A*, **575**, A75
- Bailey, S., Aldering, G., Antilogus, P., et al. 2009, *A&A*, **500**, L17
- Bernstein, J. P., Kessler, R., Kuhlmann, S., et al. 2012, *ApJ*, **753**, 152
- Betoule, M., Kessler, R., Guy, J., et al. 2014, *A&A*, **568**, A22
- Betoule, M., Marriner, J., Regnault, N., et al. 2013, *A&A*, **552**, A124
- Bianco, F. B., Modjaz, M., Hicken, M., et al. 2014, *ApJS*, **213**, 19
- Blondin, S., Mandel, K. S., & Kirshner, R. P. 2011, *A&A*, **526**, A81
- Blondin, S., & Tonry, J. L. 2007, *ApJ*, **666**, 1024
- Bohlin, R. C. 2007, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric and Polarimetric Standardization, ed. C. Sterken (San Francisco, CA: ASP), 315
- Bohlin, R. C., Gordon, K. D., & Tremblay, P.-E. 2014, *PASP*, **126**, 711
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
- Chambers, K. C. & Pan-STARRS Team 2017, AAS Meeting, **229**, 223.03
- Chevallier, M., & Polarski, D. 2001, *IJMPD*, **10**, 213
- Childress, M. J., Tucker, B. E., Yuan, F., et al. 2016, *PASA*, **33**, e055
- Childress, M. J., Wolf, C., & Zahid, H. J. 2014, *MNRAS*, **445**, 1898
- Chotard, N., Gangler, E., Aldering, G., et al. 2011, *A&A*, **529**, L4
- Conley, A., Guy, J., Sullivan, M., et al. 2011, *ApJS*, **192**, 1
- Content, D., Aaron, K., Abplanalp, L., et al. 2013, *Proc. SPIE*, **8860**, 88600E
- Contreras, C., Hamuy, M., Phillips, M. M., et al. 2010, *AJ*, **139**, 519
- DES Collaboration 2005, arXiv:astro-ph/0510346
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00036
- Efstathiou, G., Sutherland, W. J., & Maddox, S. J. 1990, *Natur*, **348**, 705
- Fakhouri, H. K., Boone, K., Aldering, G., et al. 2015, *ApJ*, **815**, 58
- Filippenko, A. V. 1997, *ARA&A*, **35**, 309
- Folatelli, G., Phillips, M. M., Burns, C. R., et al. 2010, *AJ*, **139**, 120
- Foley, R. J. 2012, *ApJ*, **748**, 127
- Foley, R. J. 2013, *MNRAS*, **435**, 273
- Foley, R. J., Filippenko, A. V., Kessler, R., et al. 2012, *AJ*, **143**, 113
- Foley, R. J., & Kasen, D. 2011, *ApJ*, **729**, 55
- Foley, R. J., Sanders, N. E., & Kirshner, R. P. 2011, *ApJ*, **742**, 89
- Foley, R. J., Scolnic, D., Rest, A., et al. 2018, *MNRAS*, **475**, 193
- Friedman, A. S., Wood-Vasey, W. M., Marion, G. H., et al. 2015, *ApJS*, **220**, 9
- Frieman, J. A., Bassett, B., Becker, A., et al. 2008a, *AJ*, **135**, 338
- Frieman, J. A., Huterer, D., Linder, E. V., & Turner, M. S. 2003, *PhRvD*, **67**, 083505
- Frieman, J. A., Turner, M. S., & Huterer, D. 2008b, *ARA&A*, **46**, 385
- Ganeshalingam, M., Li, W., Filippenko, A. V., et al. 2010, *ApJS*, **190**, 418

- Ganeshalingam, M., Li, W., & Filippenko, A. V. 2013, *MNRAS*, **433**, 2240  
 Graur, O., Rodney, S. A., Maoz, D., et al. 2014, *ApJ*, **783**, 28  
 Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, **197**, 35  
 Gunnarsson, C., Dahlén, T., Goobar, A., Jónsson, J., & Mörtzell, E. 2006, *ApJ*, **640**, 417  
 Guy, J., Astier, P., Baumont, S., et al. 2007, *A&A*, **466**, 11  
 Guy, J., Sullivan, M., Conley, A., et al. 2010, *A&A*, **523**, A7  
 Hamuy, M., Folatelli, G., Morrell, N. I., et al. 2006, *PASP*, **118**, 2  
 Hicken, M., Challis, P., Jha, S., et al. 2009a, *ApJ*, **700**, 331  
 Hicken, M., Challis, P., Kirshner, R. P., et al. 2012, *ApJS*, **200**, 12  
 Hicken, M., Wood-Vasey, W. M., Blondin, S., et al. 2009b, *ApJ*, **700**, 1097  
 Hirata, C., & Penny, M. 2014, White Paper, 3, 1  
 Holz, D. E., & Hughes, S. A. 2005, *ApJ*, **629**, 15  
 Hsiao, E. Y., Conley, A., Howell, D. A., et al. 2007, *ApJ*, **663**, 1187  
 Huterer, D., & Starkman, G. 2003, *PhRvL*, **90**, 031301  
 Ivezić, Z., Tyson, J. A., Abel, B., et al. 2008, arXiv:0805.2366  
 Jha, S., Kirshner, R. P., Challis, P., et al. 2006, *AJ*, **131**, 527  
 Jones, D. O., Rodney, S. A., Riess, A. G., et al. 2013, *ApJ*, **768**, 166  
 Jónsson, J., Sullivan, M., Hook, I., et al. 2010, *MNRAS*, **405**, 535  
 Joyce, A., Lombriser, L., & Schmidt, F. 2016, *ARNPS*, **66**, 95  
 Kelly, P. L., Hicken, M., Burke, D. L., Mandel, K. S., & Kirshner, R. P. 2010, *ApJ*, **715**, 743  
 Kessler, R., Bassett, B., Belov, P., et al. 2010, *PASP*, **122**, 1415  
 Kessler, R., Becker, A. C., Cinabro, D., et al. 2009a, *ApJS*, **185**, 32  
 Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2009b, *PASP*, **121**, 1028  
 Kessler, R., Guy, J., Marriner, J., et al. 2013, *ApJ*, **764**, 48  
 Kessler, R., & Scolnic, D. 2017, *ApJ*, **836**, 56  
 Kim, A. G., Linder, E. V., Miquel, R., & Mostek, N. 2004, *MNRAS*, **347**, 909  
 Kirshner, R. 2012, HST Proposal ID 13046  
 Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, **197**, 36  
 Krisciunas, K., Contreras, C., Burns, C. R., et al. 2017, *AJ*, **154**, 211  
 Krisciunas, K., Suntzeff, N. B., Phillips, M. M., et al. 2004, *AJ*, **128**, 3034  
 Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, *ApJ*, **722**, 566  
 Lewis, A. 2013, *PhRvD*, **87**, 103529  
 Linder, E. V. 2003, *PhRvL*, **90**, 091301  
 Linder, E. V., & Huterer, D. 2005, *PhRvD*, **72**, 043509  
 Lochner, M., McEwen, J. D., Peiris, H. V., Lahav, O., & Winter, M. K. 2016, *ApJS*, **225**, 31  
 Maguire, K., Sullivan, M., Ellis, R. S., et al. 2012, *MNRAS*, **426**, 2359  
 Mandel, K. S., Foley, R. J., & Kirshner, R. P. 2014, *ApJ*, **797**, 75  
 Mandel, K. S., Narayan, G., & Kirshner, R. P. 2011, *ApJ*, **731**, 120  
 Mandel, K. S., Scolnic, D. M., Shariff, H., Foley, R. J., & Kirshner, R. P. 2017, *ApJ*, **842**, 93  
 Marra, V., Quartin, M., & Amendola, L. 2013, *PhRvD*, **88**, 063004  
 Miknaitis, G., Pignata, G., Rest, A., et al. 2007, *ApJ*, **666**, 674  
 Milne, P. A., Foley, R. J., Brown, P. J., & Narayan, G. 2015, *ApJ*, **803**, 20  
 Narayan, G., Rest, A., Tucker, B. E., et al. 2016, *ApJS*, **224**, 3  
 National Research Council 2010, *New Worlds, New Horizons in Astronomy and Astrophysics* (Washington, DC: The National Academies Press)  
 Peebles, P. J. E. 1984, *ApJ*, **284**, 439  
 Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, **517**, 565  
 Perlmutter, S., & Schmidt, B. P. 2003, *LNP*, **598**, 195  
 Pierel, J. D. R., Rodney, S. A., Avelino, A., et al. 2018, *SNSEDextend, Astrophysics Source Code Library*, ascl:1805.017  
 Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, **594**, A14  
 Rauscher, B. J., Fox, O., Ferruit, P., et al. 2007, *PASP*, **119**, 768  
 Rest, A., Scolnic, D., Foley, R. J., et al. 2014, *ApJ*, **795**, 44  
 Riess, A. 2010, in *Proc. Space Telescope Science Institute Calibration Workshop, Hubble after SM4*, ed. S. Deustua & C. Oliveira (Baltimore, MD: STSci), 28  
 Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, **116**, 1009  
 Riess, A. G., Kirshner, R. P., Schmidt, B. P., et al. 1999, *AJ*, **117**, 707  
 Riess, A. G., Strolger, L.-G., Casertano, S., et al. 2007, *ApJ*, **659**, 98  
 Rigault, M., Aldering, G., Kowalski, M., et al. 2015, *ApJ*, **802**, 20  
 Rigault, M., Copin, Y., Aldering, G., et al. 2013, *A&A*, **560**, A66  
 Rodney, S. A., Riess, A. G., Dahlen, T., et al. 2012, *ApJ*, **746**, 5  
 Rodney, S. A., Riess, A. G., Strolger, L.-G., et al. 2014, *AJ*, **148**, 13  
 Rubin, D., Aldering, G., Barbary, K., et al. 2015, *ApJ*, **813**, 137  
 Rubin, D., & Hayden, B. 2016, *ApJL*, **833**, L30  
 Sako, M., Bassett, B., Becker, A. C., et al. 2018, *PASP*, **130**, 064002  
 Saunders, C., Aldering, G., Antilogus, P., et al. 2015, *ApJ*, **800**, 57  
 Scolnic, D., Casertano, S., Riess, A., et al. 2015, *ApJ*, **815**, 117  
 Scolnic, D., & Kessler, R. 2016, *ApJL*, **822**, L35  
 Scolnic, D., Kessler, R., Brout, D., et al. 2017, *ApJL*, **852**, L3  
 Scolnic, D., Rest, A., Riess, A., et al. 2014a, *ApJ*, **795**, 45  
 Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, *ApJ*, **859**, 101  
 Scolnic, D. M., Riess, A. G., Foley, R. J., et al. 2014b, *ApJ*, **780**, 37  
 Shafer, D. L., & Huterer, D. 2014, *PhRvD*, **89**, 063510  
 Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv:1503.03757  
 Spergel, D., Gehrels, N., Breckinridge, J., et al. 2013, arXiv:1305.5425  
 Stritzinger, M. D., Phillips, M. M., Boldt, L. N., et al. 2011, *AJ*, **142**, 156  
 Strolger, L.-G., Dahlen, T., Rodney, S. A., et al. 2015, *ApJ*, **813**, 93  
 Sullivan, M., Conley, A., Howell, D. A., et al. 2010, *MNRAS*, **406**, 782  
 Sullivan, M., Guy, J., Conley, A., et al. 2011, *ApJ*, **737**, 102  
 Suzuki, N., Rubin, D., Lidman, C., et al. 2012, *ApJ*, **746**, 85  
 Tripp, R. 1998, *A&A*, **331**, 815  
 Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, *PhR*, **530**, 87  
 Weinberg, S. 1989, *RvMP*, **61**, 1  
 Wood-Vasey, W. M., Friedman, A. S., Bloom, J. S., et al. 2008, *ApJ*, **689**, 377  
 Wood-Vasey, W. M., Miknaitis, G., Stubbs, C. W., et al. 2007, *ApJ*, **666**, 694