

1    **Optimal Scheduling of Exoplanet Direct Imaging Single-Visit**  
2    **Observations of a Blind Search Survey**

3    Dean R. Keithly<sup>ac</sup>, Dmitry Savransky<sup>ac</sup>, Daniel Garrett<sup>ac</sup>, Christian Delacroix<sup>b</sup>, and Gabriel  
4    Soto<sup>ac</sup>

5    <sup>a</sup>Cornell University, 124 Hoy Rd, Ithaca, NY, 14853, United States

6    <sup>b</sup>University of Liège, Allée du Six Août 19c, B-4000 Liège, Belgium

7    <sup>c</sup>Carl Sagan Institute, Cornell University, Ithaca, NY, 14853, United States

8    **ABSTRACT**

9    We present an algorithm, effective over a broad range of planet populations and instruments, for optimizing inte-  
10    gration times of an exoplanet direct imaging observation schedule, to maximize the number of unique exoplanet  
11    detections under realistic mission constraints. Our planning process uses “completeness” as a reward metric and  
12    the non-linear combination of optimal integration time per target and constant overhead time per target as a  
13    cost metric constrained by a total mission time. We validate our planned target list and integration times for  
14    a specific telescope by running a Monte Carlo of full mission simulations using EXOSIMS. These simulations  
15    encapsulate dynamic details such as time-varying local zodiacal light for each star, planet keep-out regions, ex-  
16    oplanet positions, and strict enforcement of observatory use over time. We test our methods on the Wide-Field  
17    Infrared Survey Telescope (WFIRST) Coronagraphic Instrument (CGI). We find that planet, sun, and solar  
18    panel keep-out regions limit some target per annum visibility to less than 28%, and that the mean local zodiacal  
19    light flux for optimally scheduled observations is  $22.79 \text{ mag arcsec}^{-2}$ . Both these values are more pessimistic  
20    than previous approximations and impact the simulated mission yield. We find that the WFIRST CGI detects  
21     $5.48 \pm 0.17$  and  $16.26 \pm 0.51$  exoplanets, on average, when observing two different planet populations based on  
22    Kepler Q1-Q6 data, and the full Kepler data release, respectively. Optimizing our planned observations using  
23    completeness derived from the more pessimistic planet population (in terms of overall planet occurrence rates)  
24    results in a more robust yield than optimization based on the more optimistic planet population. We also find  
25    optimization based on the more pessimistic population results in more small planet detections than optimization  
26    with the more optimistic population.

27    **Keywords:** optimization, exoplanet, direct imaging, coronagraph, validation, scheduling

28    **1. INTRODUCTION**

29    The 2010 astronomy and astrophysics decadal survey highly prioritized exoplanet bulk population statistics and  
30    inventorying planets around nearby stars (within 30 pc).<sup>1</sup> The Wide-Field Infrared Survey Telescope (WFIRST),<sup>2</sup>  
31    prioritized by the 2010 decadal survey, will include a coronagraphic instrument (CGI) capable of directly imaging  
32    and detecting new exoplanets unobservable by modern radial velocity or transit techniques. While the expected  
33    performance of CGI in blind search surveys can be estimated through probabilistic methods using comple-  
34    teness,<sup>3,4</sup> we can validate this performance by executing a Monte Carlo of full survey simulations on simulated  
35    universes. This process creates an ensemble of design reference missions (DRMs) containing a list of target stars  
36    observed, the integration time used for each star, when the simulated observations occurred, and the simulated  
37    outcome of each observation. Such a collection of DRMs, produced by our method, effectually certifies the ability  
38    of the instrument to make the expected number of detection observations claimed in a probabilistic evaluation.  
39    It is important to note that both methods are still equally limited in their overall prediction accuracy by the  
40    assumptions made about the true population of exoplanets to be discovered.

41    Detailed DRMs enable requirement definition and design iteration optimization for future telescopes including  
42    the large-scale mission concepts under development by science and technology definition teams for NASA’s

---

Send correspondence to Dean Keithly drk94@cornell.edu

43 2020 decadal survey. Both the Habitable Exoplanet Observatory (HabEx)<sup>5</sup> and the Large UV-Optical-Infrared  
44 Surveyor (LUVOIR)<sup>6</sup> mission concepts contain a significant exoplanet direct imaging component with HabEx  
45 reserving 1.95 years for coronagraph science operations and LUVOIR reserving 50% of the total mission time for  
46 exoplanet science. While target revisits and spectral characterizations could represent a substantial portion of  
47 the executed mission, we purposefully omit optimization with revisits on the grounds of dynamic program non-  
48 linear complexity and need of a multi-objective reward metric. Here, we focus solely on delivering an estimate of  
49 the maximum possible number of uniquely detected exoplanets under realistic mission constraints by evaluating  
50 the number of detections made through single-visit observations of stars, henceforth referred to as yield. We  
51 leave the treatment of revisits, orbit characterizations, and the full optimization with spectral characterizations  
52 for future work.

53 Brown, in Ref. 3, used single-visit completeness to estimate the number of extrasolar planets potentially  
54 discoverable with the Terrestrial Planet Finder (TPF-C), an earlier direct imaging mission concept. Single-  
55 visit completeness, hereafter referred to as completeness, is the probability of detecting a planet, drawn from  
56 a particular population, using a particular instrument, should one exist about a given target star.<sup>3,4,7</sup> When  
57 completeness is evaluated for each star in a full target list, summed, and scaled by the expected number of  
58 planets from that population per star ( $\eta$ ), we arrive at the expected number of planets to be detected from that  
59 population by that mission.<sup>8</sup> While this technique can be used to quickly evaluate a mission’s performance, it  
60 abrogates temporal constraints and uncertainties such as target visibility, variable overhead times, changing local  
61 zodiacal light intensity, and unscheduled characterizations of newly detected planets. Solely using completeness  
62 to evaluate a mission can therefore only provide an upper bound for expected performance.

63 Completeness has previously been used as a reward metric for multiple observation integration time optimiza-  
64 tions. Brown demonstrated a method for finding the group  $\Delta\text{mag}$  (difference in brightness between the planet  
65 and star in magnitudes), and number of target trade-off point by optimizing a target list subset  $\Delta\text{mag}$  against  
66 the number of targets in that subset assuming different fixed mission times.<sup>3</sup> While this method approximates  
67 the reward gradient for achieving a specific group  $\Delta\text{mag}$ , it overlooks the gain made by customizing  $\Delta\text{mag}_i$  for  
68 each individual star,  $i$ . Hunyadi et al., in Ref. 9, advanced Brown’s work by maximizing summed completeness  
69 over all targets assuming a fixed mission time and using integration times as the decision variables. In this new  
70 approach, star integration times are optimized to equivalent slopes beyond the completeness vs log integration  
71 time inflection point and the highest completeness per integration time of these targets is then selected. To prac-  
72 tically achieve this, the authors of that study discretized integration times into 1 hour increments and calculated  
73 completeness values for each integration time. Their final target list contained the set of highest completeness  
74 per integration time targets. Hunyadi et al. specifically investigated Earth-analogue planets in the habitable zone  
75 (as also done in Ref. 3), but also explored Jupiter and Saturn-analogues. While this work advanced the field, the  
76 use of discretized integration times into relatively large  $dt$  blocks and the omission of overhead per observation  
77 make the results sub-optimal and infeasible when observatory overhead times are included for each observation.  
78 Alternatively, a limiting search observation, as defined by Brown,<sup>10</sup> would observe a target for the fixed exposure  
79 time sufficient to achieve the system’s limiting planet-star magnitude,  $\Delta\text{mag}_{\text{lim}}$ , for each target. The creation of  
80 the final observation schedule in Ref. 10 involved selecting the subset of targets with precalculated integration  
81 times and overhead time per observation that fit within the total observing time. While Brown implemented  
82 target revisits, which can improve yield, the use of precalculated  $\Delta\text{mag}_{\text{lim}}$  still makes the resulting target list  
83 sub-optimal. While optimization of integration times and inclusion of revisits mark improvements in yield and  
84 target list planning realism, the omission of overhead times and discretization of integration times limits the  
85 ability to practically implement the desired observations within a finite span of time and under dynamic mission  
86 constraints.

87 A refinement of Brown and Hunyadi’s work—altruistic yield optimization (AYO)<sup>4</sup>—uses completeness vs  
88 integration time as a figure of merit as in Ref. 10 to incrementally sacrifice stars from a target list and re-allocate  
89 the integration time,  $t_i$ , from star  $i$  to the largest  $dc_j(t_j)/dt_j$  star  $j$  in increments of  $dt$ . At its core, this represents  
90 a form of “greedy optimization” which incrementally converges each observation to a constant slope similar to  
91 that described in Ref. 9. This method does not allow for the re-introduction of targets stars to the target list, a  
92 necessity of a dynamically evolving mission schedule. Finally, the original AYO method does not fully account  
93 for overhead times in the calculation of integration times but rather states that the addition of time can occur  
94 after the fact and the use of a finite  $dt$  parameter results in a loss of potential completeness.

95        None of these yield optimization processes use continuous optimization of integration times or test the ability  
96        to schedule observations via full mission simulations accounting mission elasped time (MET) Brown stressed  
97        that Monte Carlo simulations of the mission as a whole should be used to produce confidence in the proposed  
98        mission's integrity.<sup>3</sup> Brown's work seeded the founding pillars of our well book-kept full mission survey simulator  
99        to include:

- 100      1. tracking individual exoplanets around target stars vs MET  
101      2. tracking spacecraft position vs MET  
102      3. accounting for solar system body locations and keep-out regions vs MET  
103      4. accounting for variations in local zodiacal light vs MET  
104      5. potential restriction of telescope observations to prescribed observing blocks and time-sharing with other  
105        observatory instruments vs MET

106      The EXOSIMS<sup>11</sup> code base was specifically developed to book-keep these parameters across MET. In EXOSIMS  
107        and this paper, we account for the locations of individual exoplanets around target stars vs MET, the tracking of  
108        our observatory on a nominal L2 Halo orbit<sup>12</sup> vs MET, solar system body locations vs MET from NAIF kernels,  
109        keep-out occlusion of target stars vs MET, changes in local zodiacal light intensity vs MET, and possesses the  
110        capability to account for cordoned off observing blocks reserved for other instruments at varying MET, and  
111        portion of mission life reserved for observatory science.

112      This paper describes our process for producing a set of planned observations maximizing unique exoplanet  
113        detection yield and subsequently validating this prediction under realistic mission conditions. The observation  
114        planning process incorporates a combination of filters applied to a planet population and star catalog, described  
115        in Section 2.1 and Section 2.2, analytical completeness calculations detailed in Section 2.3, and our optimization  
116        algorithms with Binary-Integer Programming (BIP)<sup>13</sup> and Sequential Least Squares Quadratic Programming  
117        (SLSQP),<sup>14</sup> discussed in Section 2.4. Our validation process is outlined in Section 2.5 where we discuss the  
118        framework of EXOSIMS<sup>11,15–17</sup> survey simulation as well as incorporation of time-varying keep-out regions  
119        (Section 2.5.1), time-varying local zodiacal light noise (Section 2.5.2), and the convergence of our Monte Carlo  
120        simulation (Section 2.5.3). We then show practical application on WFIRST in Section 3 where we discuss  
121        attributes of the observing plan, and a comparison of planet populations input vs instrument capabilities vs  
122        detected planet population. We also include symbols and acronyms in Appendix A, instrument and fit file  
123        parameters<sup>18</sup> in Appendix B, and a full optimized target list in the Appendix C.

## 124                  2. METHODS

### 125                  2.1 Planet Populations

126      To calculate completeness for each target for a specific instrument, we first generate a joint probability density  
127        function of  $\Delta\text{mag}$  and planet-star separation projected onto the image plane,  $s$ , using Monte Carlo as in Ref. 3  
128        for an assumed planet population. We use two planet populations in this paper; one derived from Kepler's  
129        detections from Q1-Q6 data<sup>19</sup> (Kepler Like)<sup>20</sup> and another derived from NASA's Exoplanet Program Analysis  
130        Group (ExoPAG) Study Analysis Group 13 (SAG13).<sup>21,22</sup> We use the same sampling methods and distributions  
131        for calculating completeness and generating planets in the simulated universes on which we run our simulated  
132        missions.

133      For the Kepler Like planet population, we adopt a modified power-law distribution for semi-major axis ( $a$ )  
134        of the form

$$f_a(a) = \frac{a^{-0.62}}{a_{\text{norm}}} \exp\left(-\frac{a^2}{a_{\text{knee}}^2}\right) \quad (1)$$

135      where -0.62 is adopted from Ref. 23 derived from the power law fit from Ref. 24. In this model, we include an  
136        exponential decline in semi-major axis past a "semi-major axis knee" ( $a_{\text{knee}}$ ), which we place at 10 AU, based

<sup>137</sup> on the observed, sharp decline in detected planets with period  $\approx 10^4$  d around an assumed solar mass star.<sup>24</sup>  
<sup>138</sup> The normalization factor is given by the integrating the un-normalized distribution over a specific  $a$  range

$$a_{\text{norm}} = \int_{a_{\min}}^{a_{\max}} a^{-0.62} \exp\left(-\frac{a^2}{a_{\text{knee}}^2}\right) da, \quad (2)$$

<sup>139</sup> where we consider values of  $a$  range in  $a_{\min} = 0.1$  AU to  $a_{\max} = 30$  AU, again based on the paucity of wide-  
<sup>140</sup> separation planets discovered to date. We note, however, that for WFIRST, which has an inner working angle  
<sup>141</sup> (IWA) of 0.15 arcsec, the closest target list star has distance,  $d_i$ , of 2.63 pc and would have the smallest observable  
<sup>142</sup> planet star separation  $s_{\min}$ , given by  $IWA \times d_i \approx s_{\min}$ , at 0.394 AU. Since  $s_{\min} \approx a_{\min}(1 + e_{\max})$ , the smallest  
<sup>143</sup> observable semi-major axis is 0.292 AU for a maximum eccentricity,  $e_{\max}$  of 0.35.

<sup>144</sup> The set of generated Kepler Like planetary radii,  $\mathbf{R}$ , is built off Figure 7 in Ref. 19 with bounding limits of  
<sup>145</sup>  $\mathbf{R} = \{R_k \mid 1 R_{\oplus} < R_k < 22.6 R_{\oplus} \forall k \in 1..N_p\}$ .  $N_p$  is the total number of planets to generate. We see this  
<sup>146</sup> planetary radius range reflected in the y-axis limits of the generated Kepler Like universe in the top left plot  
<sup>147</sup> of Figure 10. Since Figure 7 of Ref. 19 only contains data for planet periods less than 85 d and greater than  
<sup>148</sup> 0.8 d, we normalize the bin values in Fressin ( $\mathbf{R}_{85}$ ) by  $a_{85,\text{norm}}$ , Eq. 2 evaluated at  $a_{\min} = a_{0.8}$  to  $a_{\max} = a_{85}$   
<sup>149</sup> (semi-major axes corresponding to periods of 0.8 d and 85 d around a sun mass star). We get the planetary  
<sup>150</sup> radius bin values  $\mathbf{R}_{\text{vals}}$ , which are normalized and scaled by  $f_{\bar{a}}$  with  $a_{85,\text{norm}}$  of

$$\mathbf{R}_{\text{vals}} = \int_{a_{\min}}^{a_{\max}} a^{-0.62} \exp\left(-\frac{a^2}{a_{\text{knee}}^2}\right) \frac{\mathbf{R}_{85}}{a_{85,\text{norm}}} da. \quad (3)$$

<sup>151</sup> We multiply the last 5 bins in  $\mathbf{R}_{\text{vals}}$  by 2.5 to account for longer orbital baseline data and more closely match  
<sup>152</sup> the larger period orbit distributions available from radial velocity surveys at the time when this distribution was  
<sup>153</sup> first derived.<sup>24,25</sup> Even with this multiplication, the right hand side histogram of the top left plot in Figure 10  
<sup>154</sup> shows the characteristic drop-off in planetary radius clearly observable in Figure 7 of Ref. 19.

<sup>155</sup> To actually generate the number of planets, we calculate the number of samples to take per bin  $q$ ,  $N_q$ , in  
<sup>156</sup>  $\mathbf{R}_{\text{vals}}$  as

$$N_q = N_p \left\lceil \frac{R_{\text{vals},q}}{\sum \mathbf{R}_{\text{vals}}} \right\rceil. \quad (4)$$

<sup>157</sup> We then take  $N_q$  samples from a log-uniform distribution over the  $q^{\text{th}}$  bin range to get  $\mathbf{R}_q$ . Finally, we randomly  
<sup>158</sup> select  $N_p$  radii from the collective 9 sets of planetary radii to get  $\mathbf{R} = \bigcup_{q=1}^9 \mathbf{R}_q$ . At the same time, we use an  
<sup>159</sup> inverse transform sampler to generate  $\mathbf{a} = \{a_k \forall k \in 1..N_p\}$  based on Eq. 1. For our Monte Carlo completeness  
<sup>160</sup> calculations, we use  $N_p = 10^8$ , while the average number of planets generated per star in the Kepler Like universe  
<sup>161</sup> is  $\eta_{KL} = 2.375$ . The per-simulation average distribution of  $\mathbf{a}$  is shown in the histogram above the top left plot  
<sup>162</sup> of Figure 10.

<sup>163</sup> The SAG13 planet population represents a significant update on the Kepler Like population, incorporating  
<sup>164</sup> all available Kepler data circa 2017. Our EXOSIMS SAG13 model implementation is similar to the analysis  
<sup>165</sup> in Ref. 21 and is based on an extrapolation of the broken power law detailed in Ref. 26. There is additional  
<sup>166</sup> discussion of the EXOSIMS SAG13 model implementation in Refs. 22 and 27. In the SAG13 model, we calculate  
<sup>167</sup>  $\eta_{\text{SAG13}}$  using

$$\eta_{\text{SAG13}} = \gamma_0 \frac{R_{\lim,1}^{\alpha_0} - (1 R_{\oplus})^{\alpha_0}}{\alpha_0 C a_0} + \gamma_1 \frac{(22.6 R_{\oplus})^{\alpha_1} - R_{\lim,1}^{\alpha_1}}{\alpha_1 C a_1}, \quad (5)$$

<sup>168</sup> where the input parameters are listed in Table 1 and

$$C a_i(\beta_i) = \int_{a_{\min}}^{a_{\max}} (2\pi\sqrt{a^3/\mu})^{(\beta_i-1)} 3\pi \sqrt{\frac{a}{\mu}} \exp\left(-\left(\frac{a}{a_{\text{knee}}}\right)^3\right) da. \quad (6)$$

<sup>169</sup> Here,  $\mu$  is the gravitational parameter. It is important to note that here, a cubic form of the roll-off on the semi-  
<sup>170</sup> major axis distribution is used, rather than the quadratic form from the Kepler Like case (c.f., Eq. 1), motivated  
<sup>171</sup> by recent limits placed on wide-separation planets by direct imaging and longer-baseline radial velocity data.<sup>28</sup>  
<sup>172</sup> The average number of planets generated per star in the SAG13 universe,  $\eta_{\text{SAG13}}$ , is 5.62.

Table 1: Parameters for the SAG13 planet population as implemented in EXOSIMS

Parameter	$[i = 0, i = 1]$
$R_{\text{lim},i}$	$[0, 3.14]$
$\beta_i$	$[0.26, 0.59]$
$\alpha_i$	$[-0.19, -1.18]$
$\gamma_i$	$[0.38, 0.73]$

We first sample the planetary radius distribution,  $f_{\bar{R}}$ , using an inverse transform sampler on Eq. 7,

$$f_{\bar{R}}(R) = \frac{\gamma_i C a_i}{\eta_{\text{SAG13}}} R^{(\alpha_i - 1)} \quad (7)$$

over the range  $0.666 R_{\oplus} < R < 17.0859375 R_{\oplus}$ .  $f_{\bar{R}}$  is bifurcated by  $R$  based on  $R_{\text{lim},1}$ . In this notation we use  $i = 0$  for all subscripts of Eq. 6 if  $0.666 R_{\oplus} \leq R \leq R_{\text{lim},1}$  and  $i = 1$  if  $R_{\text{lim},1} \leq R \leq 17.0859375 R_{\oplus}$ . The distribution of generated planets in  $\mathbf{R}$  is shown in the right side histogram of the top right plot of Figure 10.

In the SAG13 model, the semi-major axis distribution is dependent on  $R_k$ ,  $f_{\bar{a}|R}$ . For  $R_k < R_{\text{lim},1}$  we use  $i = 0$  in

$$f_{\bar{a}|R}(a) = \left(2\pi\sqrt{\frac{a^3}{\mu}}\right)^{(\beta_i - 1)} \frac{3\pi}{C a_i} \sqrt{\frac{a}{\mu}} \exp\left(-\frac{a^3}{a_{\text{knee}}^3}\right), \quad (8)$$

over the range  $a_{\min} < a < a_{\max}$ , using the same range of values as the Kepler Like model. We also use the same  $a_{\text{knee}}$  as the Kepler Like planet population. For  $R_k \geq R_{\text{lim},1}$  we use Eq. 8 with  $i = 1$ . SAG13's  $a$  frequency distribution is included above the top right plot in Figure 10. The 2D contour plot and associated grid values of the top right plot in Figure 10 shows  $\mathbf{R}$  and  $\mathbf{a}$  interdependence.

We assume a Rayleigh distribution for orbital eccentricities in both the Kepler Like and SAG13 planet populations as done in Ref. 23, such that the eccentricity ( $e_k$ ) of the  $k^{\text{th}}$  simulated planet is

$$e_k = \sigma_e \sqrt{-2 \ln \left[ \exp\left(-\frac{e_{\min}^2}{2\sigma_e^2}\right) - n \left( \exp\left(-\frac{e_{\min}^2}{2\sigma_e^2}\right) - \exp\left(-\frac{e_{\max}^2}{2\sigma_e^2}\right) \right) \right]}, \quad (9)$$

where  $n$  is a uniform random variable between 0 and 1,  $\sigma_e$  is the Rayleigh parameter for eccentricity,  $e_{\min}$  is the minimum allowed eccentricity, defined as 0, and  $e_{\max}$  is the upper 95<sup>th</sup> percentile for  $e$ . The mean eccentricity  $\bar{e} = \sigma_e \sqrt{\pi/2}$ . In Ref. 23, the authors fit the Rayleigh distribution to radial velocity-detected planets and found  $\bar{e} \approx 0.225$  to be a best fit with  $p$ -value of 0.5. They additionally found  $0.125 < \bar{e} < 0.25$  has a  $p$ -value above 0.05 and placed strict limits of  $\bar{e} < 0.35$  and  $\bar{e} > 0$ .<sup>23</sup> For this work, we use  $\bar{e} = 0.175$  (splitting the difference between 0.225 and 0.125) as done in Ref. 19 and adopt their independence of star spectral type assumption.

In both the Kepler Like and SAG13 planet population, we calculate albedo,  $p$ , using a cubic interpolant over metallicity and  $a$  using the V-band column from Table 4 of Ref. 29. For each  $p_k$  interpolated, we use the randomly generated  $a_k$  truncated to be between 0.8 and 10 AU. We assume a uniform random planet metallicity multiplication factor between 1 $\times$  and 30 $\times$ , representative of the planet's atmospheric metallic composition relative to solar abundances. The  $p_{\max} = 0.766$  at a  $p(2 \text{ AU}, 3\text{x})$  and  $p_{\min} = 0.142$  at a  $p(0.8 \text{ AU}, 30\text{x})$  for Table 4 in Ref. 29.

The final remaining parameters needed to describe planets at the time of an observation are the physical location of the planets relative to their host stars,  $\underline{r}_{k/i}$ , which is defined by Eq. (1) in Ref. 30. We assume that the direction of the orbit eccentricity vector is uniformly distributed in space, so that the orbit inclination is sinusoidally distributed between 0 and  $\pi$ , while the remaining Keplerian Elements (longitude of the ascending node, argument of periapsis, and mean anomaly) are all uniformly distributed from 0 to  $2\pi$ .<sup>30</sup> The planet orbital radius vector,  $\underline{r}_{k/i}$ , can be projected onto the plane of the sky to get the planet-star projected separation,  $s_k$ , as

$$s_k = \left| \underline{r}_{k/i} - \underline{r}_{k/i} \cdot \frac{\underline{r}_{i/SC}}{|\underline{r}_{i/SC}|} \right|. \quad (10)$$

Sampling all of the parameters described above allows us to calculate the joint probability density of projected separation and star-planet magnitude difference,  $f_{\bar{s}, \Delta\text{mag}}(s, \Delta\text{mag})$ . As in Eq. (7) in Ref. 7 we assume independence between all parameters, except for semi-major axis and planet radius in the SAG13 population. The resulting  $f_{\bar{s}, \Delta\text{mag}}(s, \Delta\text{mag})$  densities for the two populations are shown in Figure 1.

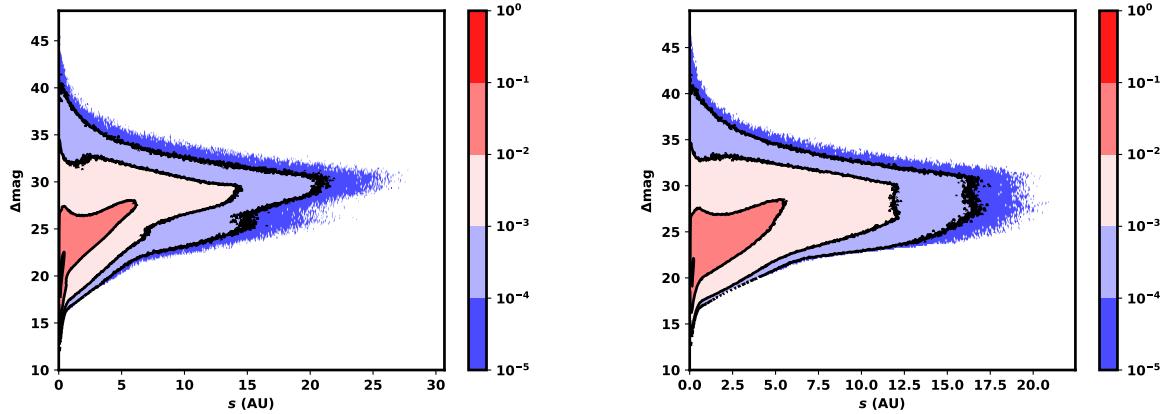


Figure 1: Joint probability density functions of projected separation and  $\Delta\text{mag}$  based on Kepler Like (left) and SAG13 (right) Planet Populations. The Kepler Like distribution produces larger orbital radii (and therefore projected separations) than SAG13 for the same  $a_{\text{knee}}$  values due to the use of the quadratic and cubic semi-major axis rollover functions (see Eq. 1 and 6).

## 2.2 Star Catalog

We need a list of target stars, along with their positions on the sky, distance ( $d_i$ ), and apparent brightness in B and V bands for our calculations of completeness and integration time. We derive the list of targets stars from the EXOCAT-1 star catalog discussed in Ref. 31, which contains a variety of targets out to a distance of approximately 30 pc. From the  $N = 2396$  targets we have an initial set of target stars **I**. We pare down **I** using a missing data filter, binary star filter, and integration time cut-off filter which reduces **I** down to 651 targets (the filters are not mutually exclusive). The

**Missing Data Filter** removes 429 targets with missing star parameters necessary for calculating completeness and integration times,

**Binary Star Filter** removes 164 targets using the Washington Double Star catalog (filters targets with companion stars within less than 10 arcseconds),<sup>32</sup>

**Integration Time Cut-off Filter** removes 1436 targets with integration times  $t_i > 30\text{d}$ , where  $t_i$  is calculated assuming local zodiacal light to be  $Z = 23.0 \text{ mag arcsec}^{-2}$ , exo-zodiacal light with magnitude  $EZ = 22.0 \text{ mag arcsec}^{-2}$ ,  $\Delta\text{mag}_0 = 22.5$  (used on the  $t_i(\Delta\text{mag})$  equations in Ref. 33 and Ref. 4), and a working angle  $WA_0 = 0.28 \text{ arcsec}$  (see Appendix B).

## 2.3 Calculating Completeness

Completeness is calculated for the  $i^{\text{th}}$  target by integrating over the joint probability density function of  $s$  and  $\Delta\text{mag}$ ,<sup>15</sup>

$$c_i = \int_0^{\Delta\text{mag}_i} \int_{s_{\min,i}}^{s_{\max,i}} f_{\bar{s}, \Delta\text{mag}}(s, \Delta\text{mag}) \, ds \, d\Delta\text{mag}. \quad (11)$$

The limits on the inner integrand are strictly obscurational. For star  $i$ , at a distance  $d_i$  from the sun, the minimum planet-star separation observable is  $s_{\min,i} = \text{IWA } d_i$  and the maximum planet-star separation is

227  $s_{\max,i} = \text{OWA } d_i$  where IWA and OWA are the instrument's inner and outer working angles, respectively. For  
228 the outer integrand, We use a lower limit on  $\Delta\text{mag}_{\min,i} = 0$  as opposed to the analytical lower bound in Eq. (18)  
229 of Ref. 7. The upper limit on  $\Delta\text{mag}$  relates the completeness to the integration time. Typically, integration  
230 times ( $t_i$ ) are defined as a function of a limiting  $\Delta\text{mag}$  and background flux levels, which are functions of the  
231 assumed instrument operating characteristics (i.e., throughput, contrast, etc.)<sup>4,33</sup> We invert the integration time  
232 model based on Ref. 33 to find  $\Delta\text{mag}$  as a function of integration time to find

$$\Delta\text{mag}_i(t_i) = -2.5 \log_{10} \frac{SNR \sqrt{\frac{C_{b,i}}{t_i} + C_{sp,i}^2}}{C_{\mathcal{F}_0} 10^{-0.4\nu_i(\lambda)} T(\lambda, WA) \epsilon_{PC}}. \quad (12)$$

233 Here,  $\nu_i(\lambda)$  is the target star's B-V color, implemented as an empirical fit to data from Ref. 34 (see Appendix  
234 B), which is accurate to about 7% in the wavelength range  $400 \text{ nm} < \lambda < 1000 \text{ nm}$  as discussed in Ref. 35.  
235  $SNR$  is the signal to noise ratio threshold chosen for determining planet detections.<sup>3</sup>  $\epsilon_{PC}$  is the photon counting  
236 efficiency of the system.  $T(\lambda, WA)$  is the instrument's core throughput (see Appendix B).  $C_{b,i}$  is the net  
237 background count rate, and  $C_{sp,i}$  is the net speckle residual count rate, including all post-processing assumptions.  
238 We use the calculations of  $C_{p,i}$ ,  $C_{b,i}$ , and  $C_{sp,i}$  from Ref. 33 and include them in Appendix B. The spectral flux  
239 density,  $C_{\mathcal{F}_0}$ , is given by

$$C_{\mathcal{F}_0}(\lambda) = \mathcal{F}_0(\lambda) A \Delta\lambda \epsilon_q(\lambda) \epsilon_{inst} \epsilon_{syst}. \quad (13)$$

240 Here,  $\mathcal{F}_0$  is the zero-magnitude flux calculated as in Eq. B.1 of Appendix B and presented in Ref. 35,  $A$  is the  
241 pupil area,  $\Delta\lambda$  is the spectral bandwidth,  $\epsilon_q(\lambda)$  is the detector quantum efficiency from Figure 12 in Appendix  
242 B,  $\epsilon_{inst}$  is the optical attenuation due to the science instrument, and  $\epsilon_{syst}$  is the optical attenuation due to  
243 the coronagraph. By integrating Eq. 11 with the limiting  $\Delta\text{mag}_i$  from Eq. 12, we arrive at a formulation for  
244 completeness as a direct function of integration time,  $c_i(t_i)$ .

245 The theoretical maximum completeness for the  $i^{\text{th}}$  target ( $c_{\infty,i}$ ) is found by integrating Eq. 11 to the upper  
246 limit of  $\Delta\text{mag}_i$  at  $t_{\infty}$ . Using the WFIRST parameters and local zodi-minimum for each target, this produces  
247  $\Delta\text{mag}_i$  varying from 23.137 to 23.138. It is important to note that the inclusion of speckle residuals means that  
248 integrating past a certain point will not produce any improvements in the achieved SNR, meaning there is a  
249 specific time for every target past which it makes no sense to integrate further.

250 While the equations derived thus far are sufficient to perform continuous integration time optimization,  
251 gradient-based solvers, such as the ones we employ below, all perform significantly better with analytical gradient  
252 functions. To expedite the optimization process in section 2.4, we calculate the derivative of completeness with  
253 respect to integration time as a function of integration time. As in Ref. 15, the derivative of Eq. 12 with respect  
254 to time is

$$\frac{d\Delta\text{mag}_i}{dt_i}(t_i) = \frac{5C_{b,i}}{4\ln(10)} \frac{1}{C_{b,i}t_i + (C_{sp,i}t_i)^2}, \quad (14)$$

255 and the derivative of completeness with respect to integration time is therefore

$$\begin{aligned} \left. \frac{dc_i}{dt_i} \right|_{t_i} &= \frac{d}{dt_i} \left[ \int_0^{\Delta\text{mag}_i(t_i)} \int_{s_{\min,i}}^{s_{\max,i}} f_{\bar{s}, \Delta\text{mag}}(s, \Delta\text{mag}) ds d\Delta\text{mag} \right] \Big|_{t_i} \\ &= \frac{d}{d\Delta\text{mag}_i} \left[ \int_0^{\Delta\text{mag}_i(t_i)} \int_{s_{\min,i}}^{s_{\max,i}} f_{\bar{s}, \Delta\text{mag}}(s, \Delta\text{mag}) ds d\Delta\text{mag} \right] \left. \frac{d\Delta\text{mag}_i}{dt_i} \right|_{t_i} \\ &= \left[ \int_{s_{\min,i}}^{s_{\max,i}} f_{\bar{s}, \Delta\text{mag}}(s, \Delta\text{mag}(t_i)) ds \right] \left. \frac{d\Delta\text{mag}_i}{dt_i} \right|_{t_i}. \end{aligned} \quad (15)$$

256 We now have all the analytical expressions needed to optimize our planned observing schedule and have  
257 filtered down the original >2,000 degrees of freedom represented by the required integration times for each star  
258 in the input catalog to 651 (see section 2.2). Input decision variables  $\mathbf{t} = t_i \forall i \in \mathbf{I}$  of 651 degrees of freedom is  
259 still quite large and could take a long time to compute, especially given the non-convexity of the sigmoid-shaped

*c*<sub>*i*</sub>(*t*<sub>*i*</sub>) curves. In order to ensure fast convergence of the nonlinear optimization, we need to provide a feasible starting guess preferably close to the final solution. We know from experimentation with AYO<sup>4</sup> that an optimal observation schedule will converge to a fixed  $\varepsilon = dc_i/dt_i \forall i \in \mathbf{I}$ . By combining Eq. 15 with Eq. 12 and inverting, we can find integration time as a function of  $dc_i/dt_i \equiv \varepsilon$ ,

$$t_i = \frac{1}{2\varepsilon C_{sp,i}^2 \sqrt{\ln(10)}} \left( -C_{b,i}\varepsilon \sqrt{\ln(10)} + \sqrt{5C_{b,i}\varepsilon C_{sp,i}^2 + C_{b,i}^2 \ln(10)\varepsilon^2} \right), \quad (16)$$

which allows us to analytically solve for integration time of all targets in a specific sub-group of  $\mathbf{I}$  at  $\varepsilon$ .<sup>15</sup> This provides us with everything we need to both formulate the optimization problem and find a good initial solution, as described in the next section.

## 2.4 Optimization Process

We formulate the summed completeness maximization problem to ensure any non-zero fractional observation incurs the observatory overhead and instrument settling time costs of making an observation. Settling time,  $T_{settling}$ , is time required by the observatory to start a new observation. This includes waiting out transient vibrations from the slew, time needed to reach thermal equilibrium, and time for the initial generation of the high-contrast region, either by “digging the dark hole” for a coronagraph, or by completing the precision alignment required by an external starshade. Overhead time,  $T_{OH}$ , on the other hand, is any additional time required by the observatory during the science integration. This includes time reserved for momentum dumping and orbit maintenance (if these operations will interrupt science data collection), dark hole maintenance for coronagraphs, and stationkeeping burns for starshades. The inclusion of  $T_{settling}$  and  $T_{OH}$  makes it difficult to find an initial feasible solution in most cases, as the overhead time required for observing every target in the target list is typically greater than the total amount of mission time available. In the specific case of the WFIRST CGI explored here, we have over 650 days of overhead time associated with observing the full target list, and less than 100 days of allotted exoplanet observing time. This means that we cannot simply evenly distribute our available time between targets to get an initial state for the optimization, as this would generate a constraint-breaking total required time. In general, initializing gradient-based optimizations on nonlinear and non-convex search spaces leads to poor optimizer performance and frequently results in no feasible solution being found. To find an initial, feasible solution and solve this problem of accounting for overheads, we use a Binary Integer Program (BIP) solver and a Sequential Least Squares Quadratic Programming (SLSQP) solver in succession.

Our optimization process is broken down into 3 major steps. The first is the calculation of an initial feasible solution via a Binary Integer Program (BIP). The second is a scalar minimization problem solving the target subset and collective derivative using Brent’s Method wrapped around a BIP. The third step uses the output from the first or second step depending on which produces a higher summed completeness as an initial solution to optimize the solution by adding, removing, and finely tuning integration times via SLSQP.

Step 1 uses an initial calculation of background count rate ( $\mathbf{C}_{b0}$ ), and residual speckle count rate ( $\mathbf{C}_{sp0}$ ) using  $f_{Z0}$ ,  $f_{EZ0}$ ,  $WA_{int}$ , and a Δmag of 25.0 mag.  $f_{Z0}$  is the zodiacal light surface brightness, in arcsec<sup>-2</sup>, calculated using

$$f_{Z0} = 10^{-0.4Z}, \quad (17)$$

where the default Z we use in Step 1 is a static 23.0 mag arcsec<sup>-2</sup>.<sup>3</sup>  $f_{EZ0}$  is the exo-zodiacal light surface brightness in arcsec<sup>-2</sup> calculated using

$$f_{EZ0} = 10^{-0.4EZ}, \quad (18)$$

where the default EZ we use in Step 1 is 22.0 mag arcsec<sup>-2</sup>.<sup>4</sup>  $WA_{int}$  is the working angle used for calculating integration times (this sets the specific values of instrument contrast, throughput and other angular-separation dependent terms). We use 0.3 arcsec for all targets in Step 1. The aforementioned parameters are used to calculate an initial  $\mathbf{c}_0$  and  $\mathbf{t}_0$ . We additionally impose a constraint on the total time spent using  $T_{max}$  as the maximum amount of time to spend observing and  $T_{OH} + T_{settling}$  as a fixed overhead time for making any observation. We use the the Coin-OR Branch and Cut solver,<sup>13</sup> as provided by Google OR-Tools to solve our BIPs, as described in Algorithm 1. This gives us an initial feasible solution of targets, denoted by  $\mathbf{x}_1^*$ , with integration times,  $\mathbf{t}_0$ , and summed completeness,  $\mathbf{x}_1^* \mathbf{c}_0$ .

---

**Algorithm 1:** Binary Integer Program -  $\mathbf{x}_1^* = \text{BIP}(\mathbf{c}_0, \mathbf{t}_0)$ 


---

**Input:**  $\mathbf{I}$ ,  $\mathbf{c}_0$ ,  $\mathbf{t}_0$ ,  $T_{OH}$ ,  $T_{settling}$ ,  $T_{\max}$ , and an optimization time limit maximum of 5 minutes  
**Output:**  $\mathbf{x}_1^*$ , the list of binary values signaling to keep (1) or remove (0) each target

$$\begin{aligned}\mathbf{x}_1^* = \min_{\mathbf{x}} \quad & - \sum_{i=0}^{N-1} x_i c_{0,i} \\ \text{s.t.} \quad & \sum_{i \in \mathbf{I}} x_i (t_{0,i} + T_{OH} + T_{settling}) \leq T_{\max}, \\ & x_i \in \{0, 1\}, \quad \forall i \in \mathbf{I}\end{aligned}$$


---

304 In Step 2, we reuse  $\mathbf{C}_{p0}$ ,  $\mathbf{C}_{b0}$ , and  $\mathbf{C}_{sp0}$  specifying a solution tolerance of  $10^{-2}$  on a bounded scalar mini-  
305 mization problem with bounds on  $\varepsilon$  of  $[0, 7]$ . We know  $\varepsilon$  at  $t = \infty$  and  $t = 0$  is 0, but we have determined from  
306 software experiments that 7 works well as an upper bound in this case. This minimization in Step 2 uses the  
307 Python implemented scipy ‘‘minimize scalar’’ function, as described in Algorithm 2. This procedure produces a  
308 separate, feasible solution, distinct from the solution arrived at in Step 1. By varying the  $\varepsilon$  and solving the BIP  
309 sub-problem, we achieve a different set of targets  $\mathbf{x}_2^*$  with integration times  $\mathbf{t}_2$ .

---

**Algorithm 2:** Bounded Scalar Minimization Wrapping Binary Integer Program

---

**Input:**  $\mathbf{I}$ ,  $\mathbf{C}_{p0}$ ,  $\mathbf{C}_{b0}$ ,  $\mathbf{C}_{sp0}$ ,  $T_{OH}$ ,  $T_{settling}$ ,  $T_{\max}$ , and an optimization time limit maximum of 5 minutes  
**Output:**  $\varepsilon^*$ , the value of  $dc/dt$  evaluated for each target which maximizes yield  
**Output:**  $\mathbf{t}^*$ , integration times for each target evaluated at  $\varepsilon^*$   
**Output:**  $\mathbf{x}_2^*$ , the list of binary values signaling to keep or remove each target

$$\begin{aligned}\varepsilon^* = \min_{\varepsilon} \quad & - \sum_{i \in \mathbf{I}} \text{BIP}(c_i(t_i^*(\varepsilon)), t_i^*(\varepsilon), T_{OH}, T_{settling}, T_{\max})_i c_i(t_i^*(\varepsilon)) \\ \text{s.t.} \quad & \varepsilon \leq 7, \\ & -\varepsilon \leq 0\end{aligned}$$

$$\begin{aligned}\mathbf{t}_2^* &\Leftarrow [t_i^*(\varepsilon^*), \forall i \in \mathbf{I}] \\ \mathbf{x}_2^* &\Leftarrow [\text{BIP}(c_i(t_i^*(\varepsilon^*)), t_i^*(\varepsilon^*), T_{OH}, T_{settling}, T_{\max}), \forall i \in \mathbf{I}]\end{aligned}$$


---

310 In Step 3, we formulate the SLSQP optimization process with an initial solution seeded with  $\mathbf{x}_1^* \mathbf{t}_0$  or  $\mathbf{x}_2^* \mathbf{t}_2^*$ ,  
311 whichever produces a larger summed completeness. This seeded solution should prove sufficiently close to an  
312 optimal solution such that the  $\mathbf{c}(\mathbf{t})$  sigmoid-like inflection point is exceeded. We now replace our previous assumed  
313  $f_{Z0}$  and  $f_{EZ0}$  with a more optimistic surface brightness. We calculate zodiacal light surface brightness every  
314  $\approx 1/3$  of a day for a year, interpolating the lookup tables from Ref. 36 as shown in Figure 4. For our optimization,  
315 we specifically use the per annum minimum,  $\mathbf{f}_{Z,\min}$  for all targets excluding  $f_{Z,i}$  in keep-out regions shown in  
316 Figure 3. This is crucial as the  $f_{Z,i}$  intensity of targets with  $0^\circ$  heliocentric ecliptic latitude, (b) have local  
317 zodiacal light minimum within  $56^\circ$  of the observatory’s anti-solar point (see Figure 4), a region not visible to  
318 WFIRST due to solar panel pointing requirements in Table 2.

319 The output  $\mathbf{t}_3^*$  is an optimal allocation of integration times to each target accounting for the non-linear  
320 overhead time assignment. There are technically infinite ‘‘near-optimal’’ solutions. Equivalent maximum summed  
321 completeness target lists can be achieved for a range of different numbers of target stars. This is directly caused

---

**Algorithm 3:** SLSQP Optimization

---

**Input:**  $\mathbf{I}$ ,  $\mathbf{f}_Z$ ,  $\mathbf{t}$ ,  $T_{OH}$ ,  $T_{settling}$  and  $T_{\max}$

**Output:**  $\mathbf{t}_3^*$ , the integration times to spend on each star

$$\begin{aligned} \mathbf{t}^* = \min_{\mathbf{t}} \quad & - \sum_{i=0}^{N-1} c_i(t_i) \\ \text{s.t.} \quad & t_i < T_{\max}, \quad \forall i \in \mathbf{I}, \\ & -t_i < 0, \quad \forall i \in \mathbf{I}, \\ & \sum_{i \in \mathbf{I}} x_i(T_{OH} + T_{settling}) + t_{0,i} < T_{\max} \end{aligned}$$

---

322 by the increasing number of approximately equivalent target stars at further stellar distances. These integration  
323 times are used as inputs to our validation process discussed in section 2.5 and implemented in EXOSIMS which  
324 runs a Monte Carlo of full mission survey simulations, observing each target for the integration times prescribed  
325 in  $\mathbf{t}_3^*$  and bookkeeping all dynamic aspects of the mission.

## 326 2.5 Validation

327 We validate the ability to schedule our optimized integration times from section 2.4, and listed in Table 9 of  
328 Appendix C, via a Monte Carlo of full survey simulations using the EXOSIMS codebase, detailed in Ref. 16.  
329 This software allows us to bookkeep dynamic aspects of the mission while scheduling observations on randomly  
330 generated planetary systems about a real set of target stars. By using EXOSIMS, we account for exoplanet,  
331 solar system planet, and observatory orbit propagation as well as recalculating look vectors, keep-out regions,  
332 and local zodiacal light noise.

333 At the start of each survey simulation, we randomly generate planets around stars; drawn from the Kepler  
334 Like or SAG13 planet populations discussed in section 2.1 using sampling methods described in Ref. 16. To  
335 define the observable times of individual stars, we use solar system planet locations from NAIF Kernels<sup>37,38</sup>  
336 combined with instrument specific keep-out angles and observatory look vectors for each target throughout the  
337 mission. The EXOSIMS framework places our observatory on a halo orbit with period of  $\approx 180$  days; propagated  
338 along the orbit throughout the mission. In each simulation, we incrementally filter available targets, simulate  
339 observations and their outcomes, and propagate orbits as shown in Figure 2.

340 At the start of each main loop depicted in Figure 2 (Steps 1 - 9), filters remove targets with too long of  
341 integration times ( $t_i > 30$  d); targets currently in keep-out regions of planetary bodies; previously observed stars  
342 not currently scheduled for revisits; and filters targets not observable within the nearest time constraint. In this  
343 paper, we do not revisit targets, so we filter out any targets that have been observed during the mission. We then  
344 use an intelligent method of choosing the next target at the current mission time,  $t_c$ . Our method of selecting  
345 targets for observation advances time by the smallest amount  $\min(t_c - \mathbf{t}_{f_Z, \min})$ , where  $\mathbf{t}_{f_Z, \min}$  are the mission  
346 times when targets in  $\mathbf{I}$  have their next local zodiacal light minimum,  $\mathbf{f}_Z, \min$ . This identifies target star  $i$  which  
347 we proceed to observe for the pre-calculated time  $t_i$  from the method described in section 2.4. We divide the  
348 observations into discrete time intervals and calculate the signal and noise at each interval, calculate the total  
349 achieved SNR, and propagate planets around the star as well as the observatory and solar system planets. The  
350 splitting of intervals allows for a better approximation of achieved SNR (via Riemann sum) than achieved by  
351 keeping all values constant in a single calculation. We then advance the mission time by  $t_i + T_{OH} + T_{settling}$  and  
352 check if the mission is over.

353 The EXOSIMS framework relies upon probabilistic planet generation and random draws, however, our Python  
354 implementation is capable of not only replicating results, but also fully reproducing each survey simulation run  
355 by resetting the simulation from the simulation's random seed. EXOSIMS also keeps track of all inputs required  
356 to replicate the simulation.

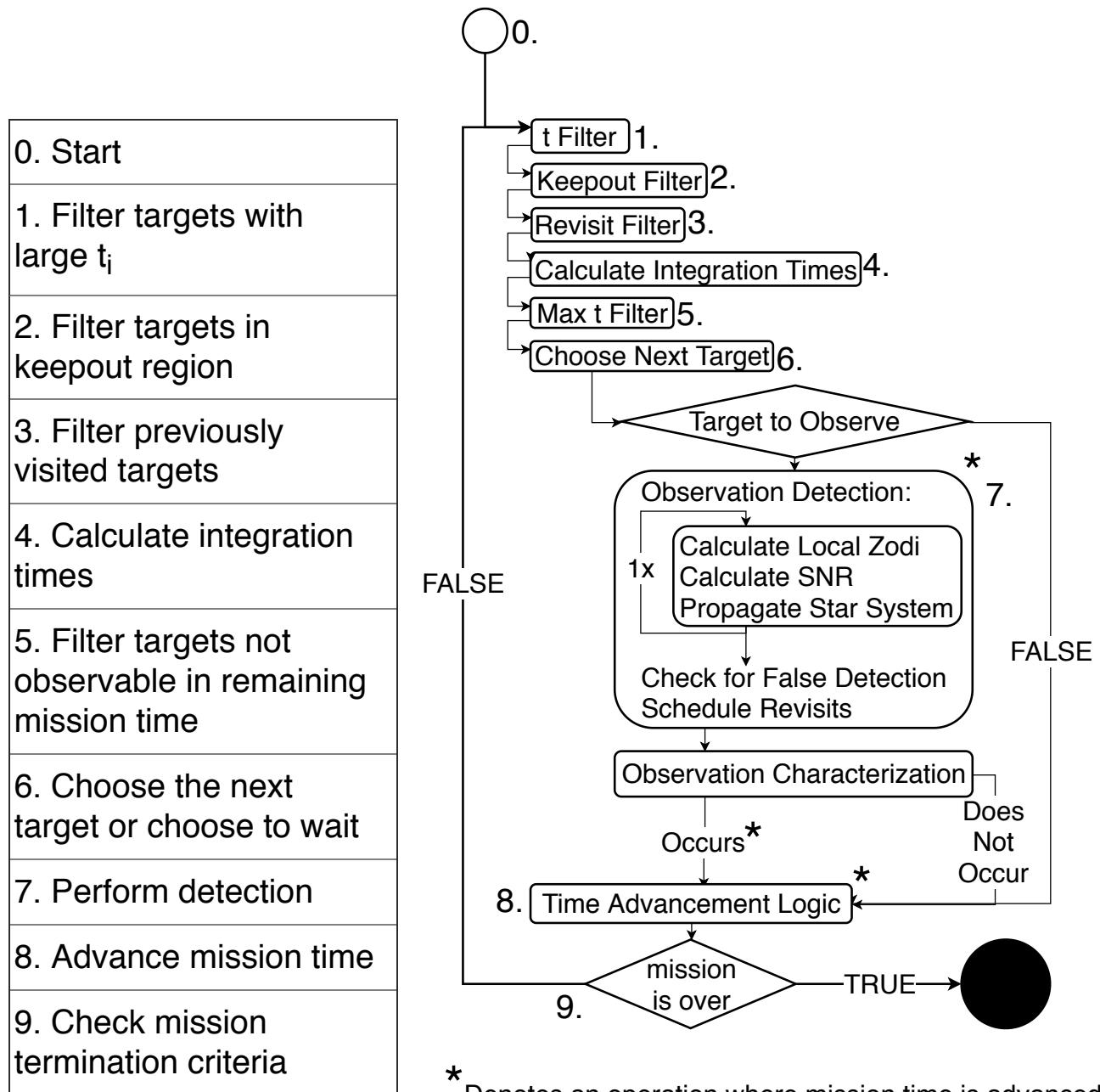


Figure 2: EXOSIMS survey simulation simplified flowchart depicting major filtering steps discussed in section 2.2, calculating integration time, selecting the next target, making the detection observation, making the characterization observation under conditions outlined in Appendix B, and advancing time. EXOSIMS is additionally capable of strictly adhering to pre-defined observing blocks, but this functionality was not used for the results presented here.

357 **2.5.1 Keep-out Regions**

358 WFIRST has keep-out regions specified in Table 2. We define these keep-out regions as a subset of the sky which  
 359 cannot be entered by the telescope pointing vector throughout an observation. Our strict accounting for time  
 360 and geometry in simulations enables us to ensure any observation’s look vector ( $\underline{r}_{i/SC}$ ) does not start, stop, or  
 361 pass through a keep-out region. Nominally, each body in Table 2 has a minimum keep-out region, however, the  
 362 sun has a maximum keep-out region requirement set by the minimum incidence angle of light on the spacecraft’s  
 363 solar panels to power the observatory and instrument. The bore-axis vector of WFIRST cannot point farther  
 364 than  $124^\circ$  away from the spacecraft sun vector ( $\underline{r}_{\odot/SC}$ ) in order to meet spacecraft power requirements. Since  
 365 our implementation in EXOSIMS in this specific case enforces observations at local zodiacal light minimum  
 366 indicated by the black squares shown in Figure 4, these observations either occur near the spacecraft anti-solar  
 367 point or edge of the solar panel keep-out region.

Table 2: Keep-out regions for WFIRST defined as the (minimum angle between  $\underline{r}_{Body/SC}$  and  $\underline{r}_{i/SC}$ , maximum angle between  $\underline{r}_{Body/SC}$  and  $\underline{r}_{i/SC}$ )

Body	Keep-out Angle (deg)
Earth	$45^\circ$
Moon	$45^\circ$
Sun	$45^\circ$
Small Bodies	$1^\circ$
Solar Panels	$124^\circ$ <sup>2</sup>

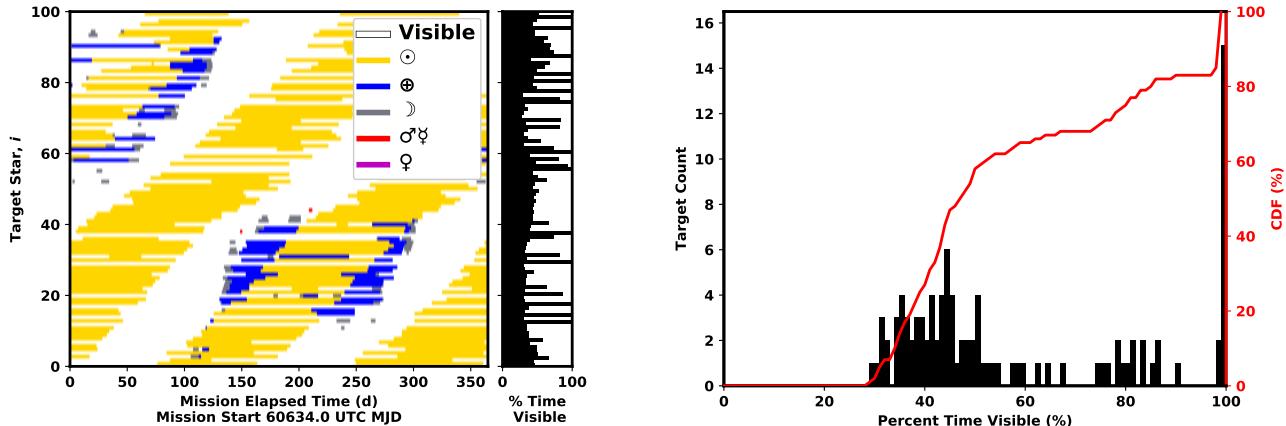


Figure 3: Keep-out map of WFIRST/targets over the first mission year showing visible times of targets (white) and different sources of keep-out occlusion including the Sun (yellow), Earth (blue), Moon (grey), Mercury/Venus (red). The horizontal histogram shows percentage of time each target is visible (left). The keep-out map filter is implemented in EXOSIMS at the Step 2 shown in Figure 2. The histogram and cumulative distribution of visibility of all targets is shown (right). Minimum target visibility is 28%.<sup>39,40</sup>

368    **2.5.2 Local Zodiacal Light**

369    The local zodiacal light intensity ( $f_{Z,i}$ ) is the largest, known, time varying noise source we account for and  
 370    manifests itself in the background noise count rate introduced in Eq. 12 and explicitly written in Appendix B.  
 371    Variations in local zodiacal light can vary the summed completeness by up to 10%.<sup>17,41,42</sup>

372    In section 2.4, for Step 1 for our integration time optimization, we use a static  $Z_0$  of 23.0 mag arcsec<sup>-2</sup> as  
 373    done in Ref. 4. However, in Step 3, we optimize our final integration times using  $f_{Z,\min}$  (or  $Z_{\max}$ ) which implies  
 374    specific times of the year when these observations can be made, as shown by the black lines in Figure 4.

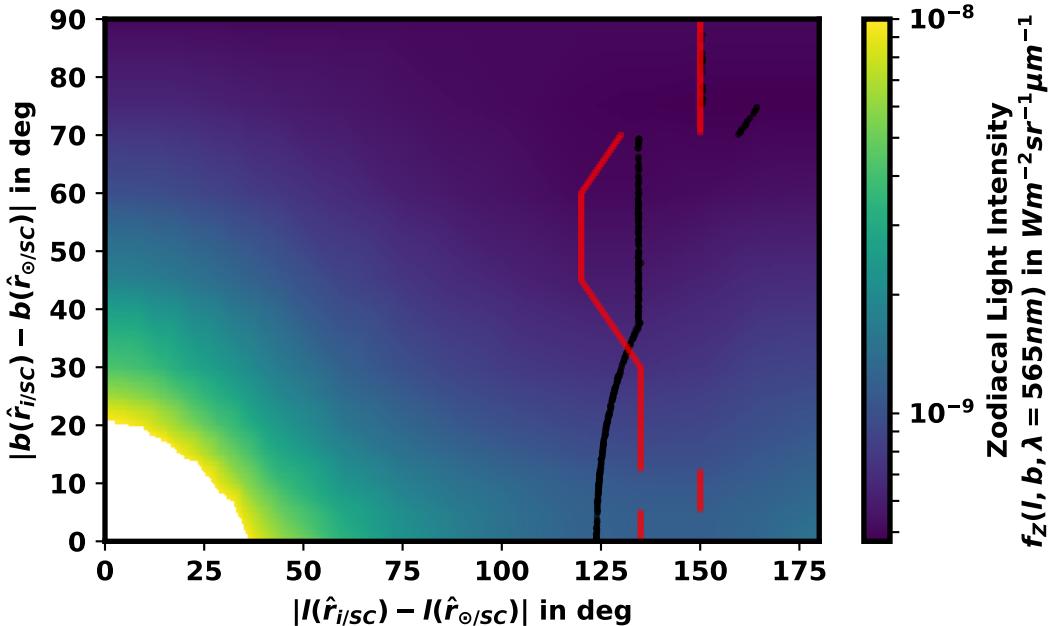


Figure 4: Local zodiacal light intensity interpolant ( $f_Z(l, b)$ ) of Eq. 19 at target star heliocentric ecliptic longitude ( $|l(\hat{r}_{i/SC}) - l(\hat{r}_{\odot/SC})|$ ) relative to the sun-spacecraft heliocentric ecliptic longitude ( $|l(\hat{r}_{\odot/SC})|$ ), target star heliocentric ecliptic latitude ( $|b(\hat{r}_{i/SC}) - b(\hat{r}_{\odot/SC})|$ ) relative to spacecraft heliocentric ecliptic latitude ( $|b(\hat{r}_{\odot/SC})|$ ), and  $\lambda = 565$  nm. The minimum zodiacal light intensity ( $f_{Z,\min}(b)$ ) for each  $b$  is indicated by densely packed red squares which appear to form lines caused by using a linear interpolant and the coarseness of the underlying datapoints. Our implementation makes observations only when target stars are coincident with black squares, where minimum  $f_Z$  restricted by keep-out ( $b < 37^\circ$ ). Note Ref. 36 is significant to 3 decimal places but the interpolant introduces machine precision numbers, the red and black dots have functionally equivalent values for  $b > 37^\circ$ . Only one quadrant is shown as the data is reflection symmetric.

375    There is a distinction between the  $f_{Z,i}$  used in the integration time optimization and the  $f_{Z,i}$  used in the  
 376    Monte Carlo validation. When we evaluate whether a detection has been made at the top of Step 7 in Figure  
 377    2, we calculate the planet *SNR* using  $f_{Z,i}$  based on  $r_{i/SC}(t_c)$  where  $t_c$  is the current time in the simulation.  
 378    We calculate the interpolant in Figure 4 using Eq. 19, a 2D linear interpolation of intensity from Table 17 of  
 379    Ref. 36 ( $f_\beta(l, b)$ ), a quadratic interpolation of wavelength dependence from Table 19 in Ref. 36 ( $f_\lambda(\log_{10}(\lambda))$ ),  
 380    and applying a sun color correction of ( $\mathcal{F}_0(\lambda)$ ) to get

$$f_Z(l, b, \lambda) = \frac{f_\lambda(\log_{10}(\lambda)) f_\beta(l, b)}{hc' \mathcal{F}_0(\lambda)}. \quad (19)$$

381 Here,  $h$  and  $c'$  are the planck constant and speed of light in a vacuum, and  $f_\lambda(\log_{10}(\lambda))$  is a quadratic 1D  
 382 interpolant of Ref. 36 data in Figure 13 of Appendix B. Since the Table 17 from Ref. 36 is in the geocentric  
 383 ecliptic coordinates, but the spacecraft will physically be located on an Earth-Sun L2 orbit, we assume the table's  
 384  $b = 0^\circ$  and  $l = 0^\circ$  is coincident with  $r_{\odot/SC}$  and the additional distance of the spacecraft from the Sun has no  
 385 effect. Further discussion of these two assumptions is included in Appendix B.

386 Using the 2D interpolation of Figure 4 combined with instrument specific filters in section 3, and optimization  
 387 process in section 2.4, we can plot the histogram of minimum (black dots in Figure 4) and maximum (at edge  
 388 of solar keep-out) local zodiacal light intensity in Figure 5. Depending on when observations are made, the  
 389 variation in zodiacal light intensity can change throughout the year (excluding when targets are in keep-out)  
 390 by upwards of 2 magnitudes. We also see using an estimated  $Z_0$  of 23.0 mag arcsec $^{-2}$  is an overly optimistic  
 391 approximation of the local zodiacal light noise. Realistically, for our filtered set of target stars, the appropriate  
 392  $\mu_{Z_{\min}} \approx 22.79$  mag arcsec $^{-2}$  and  $\mu_{Z_{\max}} \approx 21.59$  mag arcsec $^{-2}$ .

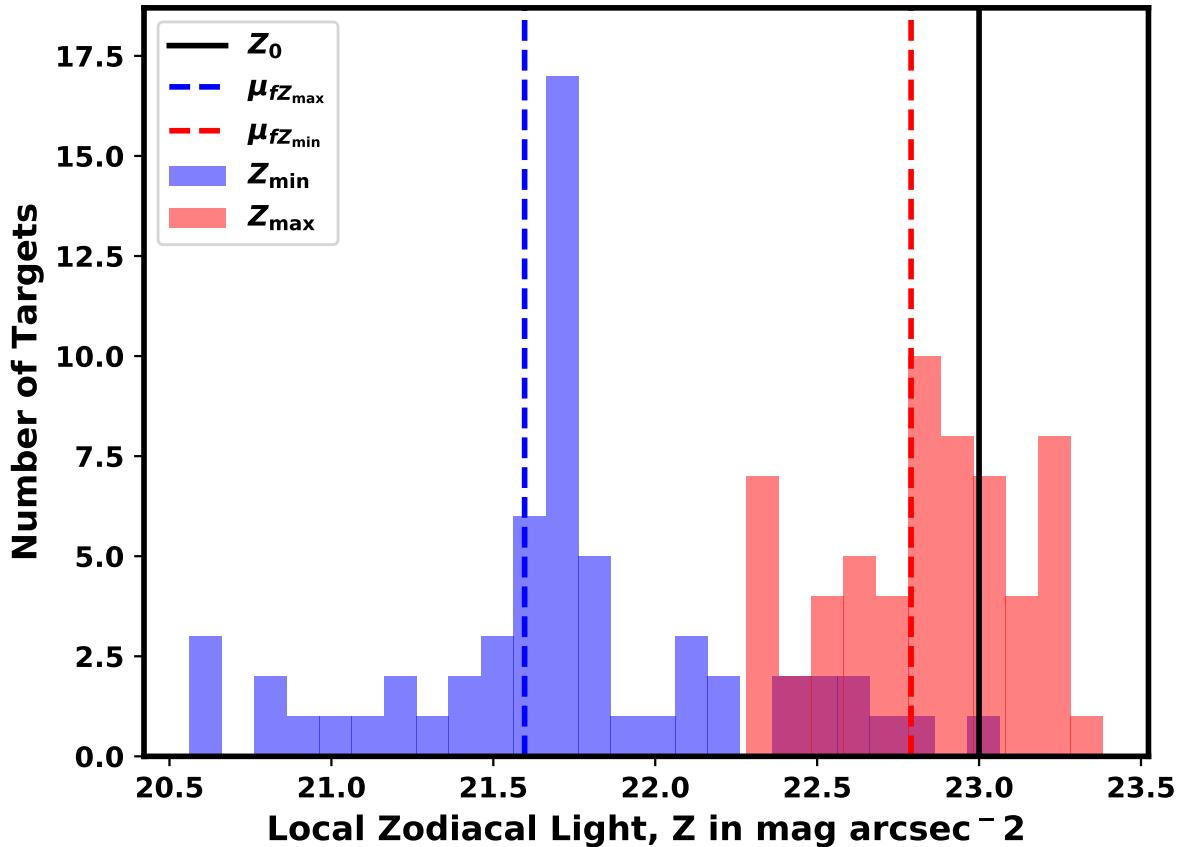


Figure 5: A histogram of Zodiacal Light intensity in mag arcsec $^{-2}$  ( $Z$ ), for stars at the minimum observed  
 intensity ( $Z_{\min}$ ), and stars at maximum observed intensity ( $Z_{\max}$ ), taking into account keep-out regions.  $Z_0$   
 corresponds to the optimistic zodiacal light intensity in mag arcsec $^{-2}$  used by Stark in Ref. 4 and Brown in  
 Ref. 3. Dashed lines represent target list  $\mu_{Z_{\min}} = 22.79$  mag arcsec $^{-2}$  and  $\mu_{Z_{\max}} = 21.59$  mag arcsec $^{-2}$ .

### 393 2.5.3 Convergence

394 We determine the required number of simulations to ensure the accuracy of our results by executing 10,000  
 395 simulations of a generic input specification similar to that used in section 3. By incrementally calculating the

396 mean of the ensemble of runs using Eqn. 38 from Ref. 43, we demonstrate convergence of our simulation yield  
 397  $\hat{X} = \mu_{det,10000}$  shown in Figure 6. If we assume the ensemble mean is the true mean ( $\hat{X} = \bar{X}$ ), we can calculate the  
 398 convergence of incremental error shown in Figure 6, which demonstrates convergence of our expected simulation  
 399 yield for the ensemble.

400 By assuming the mean number of detections per run are normally distributed, we can show the  $1\sigma$ ,  $2\sigma$ , and  
 401  $3\sigma$  confidence intervals of the mean shown in Figure 6. In reality, the ensembles are some form of gamma distri-  
 402 bution because the numbers are all positive and priors are exponentially distributed. This normal distribution  
 403 assumption fits better for high yield telescopes. Excluding Table 3 and Figure 6, all results are derived from  
 404 single simulations or ensembles of 1000 simulations. For the results in this paper, we take the  $3\sigma$  results as being  
 405 significant which is shown to be 3.19% in Table 3.

Table 3: Absolute percent error confidence intervals for 100 and 1000 simulations. This table references data created using runs from Dean22May18RS09CXXfZ01OB01PP01SU01 and file convergence-  
 DATA\_Dean22May18RS09CXXfZ01OB01PP01SU01\_2019\_04\_09\_01\_23\_.txt

# Sims	Confidence Interval	Absolute Percent Error (%)
1000	$1\sigma$	1.16
1000	$2\sigma$	2.33
1000	$3\sigma$	3.19
100	$1\sigma$	3.45
100	$2\sigma$	6.95
100	$3\sigma$	9.58

### 406 3. WFIRST RESULTS

407 The results in this section are generated using all previously discussed assumptions and parameters as well as  
 408 the parameters in Table 7 of Appendix B based on the cycle 6<sup>18</sup> description of the CGI.

#### 409 3.1 Completeness and Planned Observations

410 By applying the optimization process from section 2.4 to the target list filtered in section 2.2 and assuming a  
 411 Kepler Like planet population with per observation  $T_{OH} = T_{settling} = 0.5$  d and maximum observing time of  
 412  $T_{max} = 91.3125$  d (3 months), we get the optimal integration times in Table 9 in Appendix C. The planned  
 413 summed completeness of this target list is  $\sum c_i = 2.31$ . Since completeness is the probability of detecting  
 414 planets from a population around a star, multiplying  $\sum c_i$  by the population occurrence rate ( $\eta_{KL}$  from section  
 415 2.1) gives the expectation value of planets detected, in this case equal to 5.48 detections. We calculate the  
 416 ultimate completeness<sup>10</sup> for all 651 potential targets by evaluating  $\sum c_i$  at an infinite integration time ( $t_\infty$ ) to  
 417 get  $\sum c_\infty = 9.01$  and an expectation value of 21.40 exoplanets detected. The ratio  $2.31/9.01 \times 100 = 25.6\%$  is  
 418 a measure of planned target list yield to the maximum theoretical yield observing all 651 targets with  $t_\infty$ . The  
 419 ratio  $2.31/2.99 \times 100 = 77.2\%$  is a measure of planned target list yield to the maximum theoretical summed  
 420 completeness for all targets in Table 9. These ratios represent the fraction of planet phase space about all  
 421 scheduled targets that could be probed in the assumed, finite, total integration time.

422 There is a non-negligible difference between the planned completeness,  $c_{3,i}$ , and the completeness actually  
 423 observed  $c_{t_{obs},i}$ . The planned observations and Monte Carlo simulations entirely based on the Kepler Like  
 424 population are shown in Figure 7. Note the loss of 1 observation between the planned and actual observations,  
 425 which we attribute to accumulation of machine precision errors and our strict adherence to the  $T_{max}$  upper bound  
 426 on observing time as well as the allowance of characterizations. The loss of an observation is evident by the single  
 427 red square without an associated blue circle in Figure 7. The  $\sum c_i$  actually observed in this particular simulation  
 428 of the Monte Carlo was 2.28. For each observation made in a survey simulation, observed completeness (blue  
 429 circles) coincident with planned completeness (red squares) indicate each simulated observation occurs under  
 430 optimal conditions. Because we are observing targets solely at the local zodiacal light minimum and do not  
 431 modify integration times, all observations are optimal. If targets were observed at sub-optimal zodiacal light  
 432 levels, the  $c_{t_{obs},i}$  would be below the  $c_{0,i}$  as shown in Ref. 17.

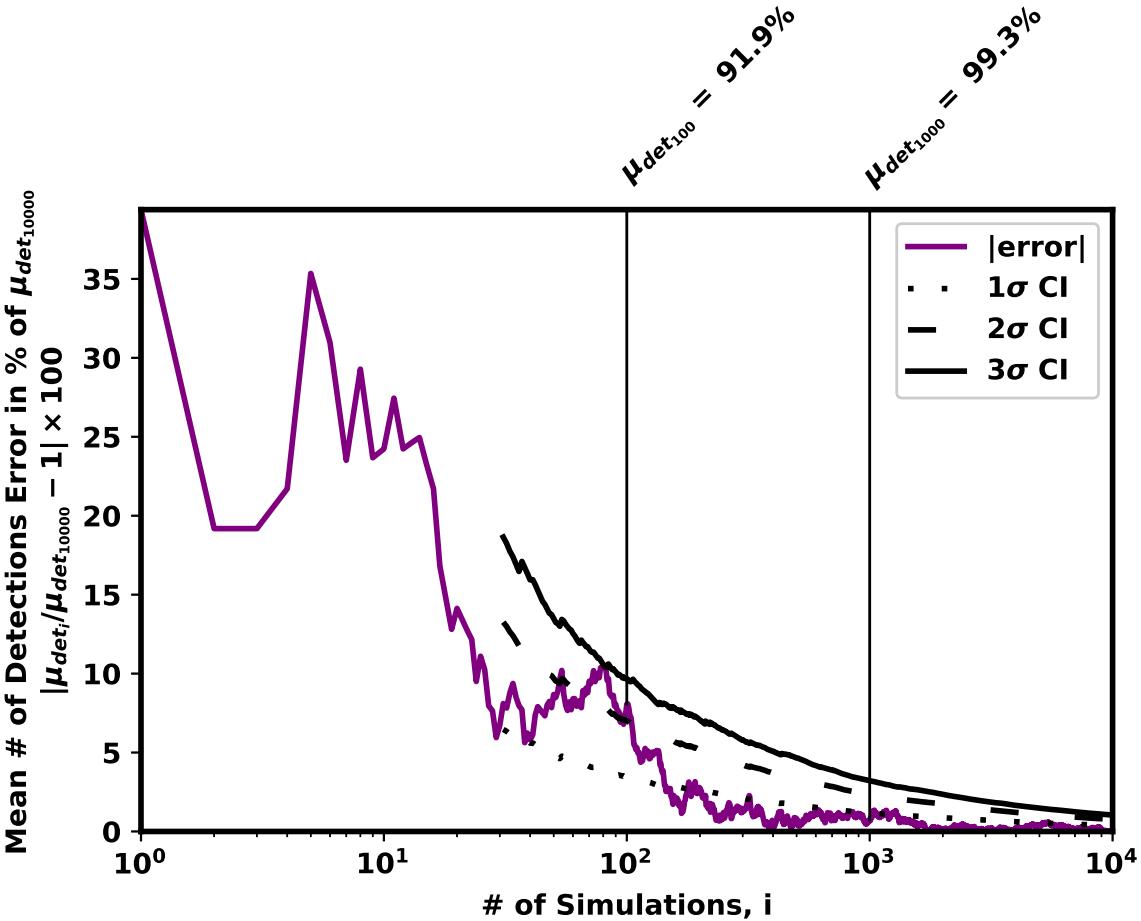


Figure 6: Absolute percent error from  $\mu_{det_{10000}}$  convergence combined with  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  confidence intervals.

433 From the combined time varying limit in Eq. 12 applied to Eq. 11 we get the black sigmoid-like lines in  
 434 Figure 7, specifically plotting  $c_i(t_i)$  for the top 10, median, and lowest completeness optimized targets. The  
 435 median and lowest  $c_i(t_i)$  lines are characteristic for the majority of similar completeness targets and the addition  
 436 of targets will typically be below the lowest completeness target in this list. The  $c_i(t_i)$  and completeness side  
 437 histogram shows a clustering of targets at lower completeness values, which can be attributed to the increased  
 438 number of targets at larger  $d_i$ . The upper limit of completeness lines are consistent with the theoretical maximum  
 439 completeness values.

440 Demonstrating the importance of including overhead and settling times in observations are the max  $c_i/t_i$   
 441 diamonds in Figure 7. These are universally located at some small integration time ( $t_i < 10^{-3}$  d,  $\forall i \in \mathbf{I}$ ) and  
 442 small completeness meaning any optimized target list without overhead constraints will have strictly non-zero  $t_i$ .  
 443 This means optimizing summed completeness without  $T_{OH}$  and  $T_{settling}$  results in an observation target list of  
 444 length N (651 in the case of WFIRST) which cannot be executed under realistic conditions. Similarly, observing  
 445 at  $\Delta\text{mag}_{\text{lim}}$  additionally results in sub-optimal individual target completeness and also leaves a substantial  
 446 amount of unobserved phase space around each target.

447 Calculating completeness using the SAG13 planet population shows the maximum theoretical summed com-  
 448 pleteness of all targets is 13.538. The max theoretical completeness of the observed targets is 3.641. The  
 449 minimum completeness of these targets has increased due to the increased likelihood of larger  $\Delta\text{mag}$  planets at  
 450 larger  $s$  in the SAG13 distribution in Figure 1. As in the Kepler Like optimized target list, we see a single target  
 451 is not being observed due to strict enforcement of observing time constraints. The major difference between

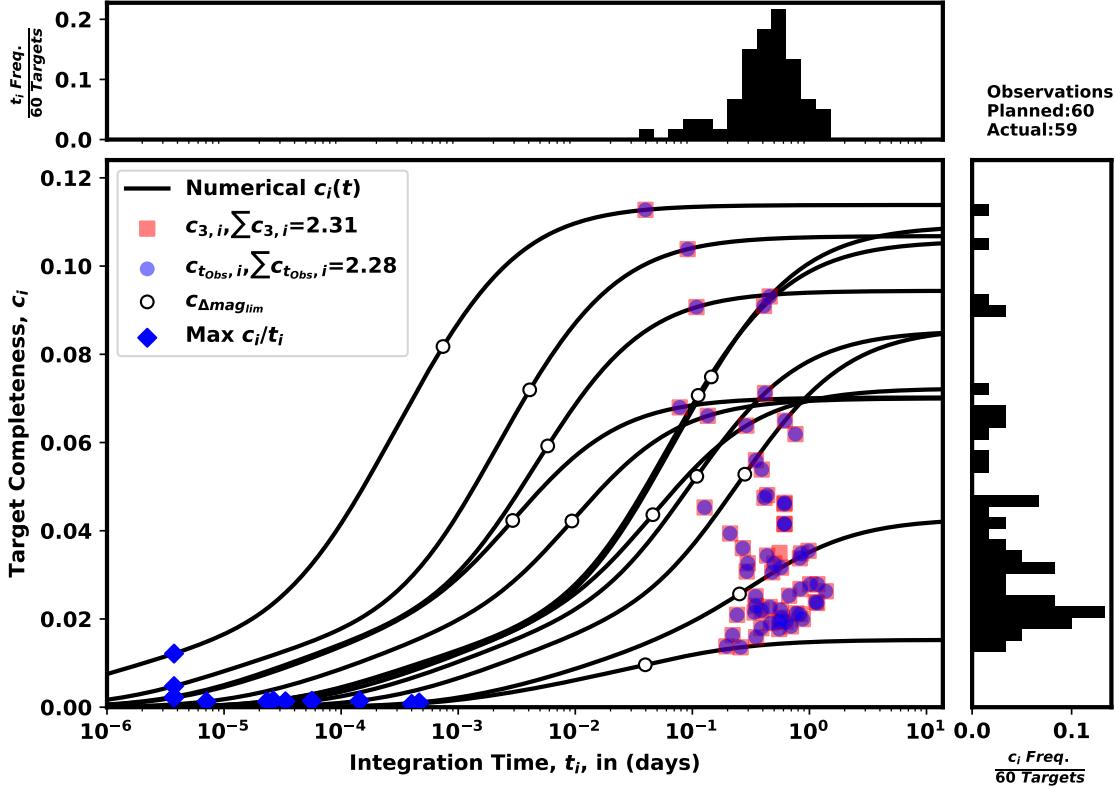


Figure 7: Completeness as a function of integration times calculated using the Kepler Like distribution, EXOCAT-1 star catalog, Nemati SNR model,<sup>33</sup> and Leinert Table Zodiacial Light.<sup>36</sup> Black lines show completeness vs integration time for the 10 highest planned completeness targets, the median completeness target, and the lowest completeness target. Red squares indicate planned integration time and planned completeness based on  $\varepsilon$  maximizing summed completeness. Blue dots indicate observation integration time and observation completeness of the simulated observation. White dots represent completeness at  $\Delta\text{mag}_{lim}$  and is plotted for the 10 highest, median, and lowest completeness planned targets. Blue diamond shows the completeness and integration time of the maximum  $c_i/t_i$  point for the 10 highest, median, and lowest completeness targets. This plot references data generated using C0vsT0andCvsTDATA\_WFIRSTcycle6core\_CKL2\_PPKL2\_2019\_04\_05\_19\_18\_.txt with specific target stars, integration times, and completeness included in Table 9.

452 Figure 7 and Figure 8 is the change in shape of the  $c_i(t_i)$  lines, which mostly have a higher theoretical maximum  
 453 completeness and general shift towards higher completeness at lower integration times. There is additionally  
 454 a larger separation of lower completeness targets in Figure 8 compared to Figure 7. In general, completeness  
 455 calculated using the SAG13 population is larger than completeness calculated using the Kepler Like population.

### 456 3.2 Sky Distribution of Completeness, Integration Times, and Targets

457 We are able to take the optimal target list included in Table 9 and bin the heliocentric ecliptic coordinates of each  
 458 target,  $l$  and  $b$ , into triangular regions of approximately equivalent size and approximately isotropic distribution  
 459 on the sky. When we sum integration time for all targets in each bin and normalize by bin area we get a skymap  
 460 distribution shown in Figure 9. Since Figure 7 shows all integration times are between 0.1 and 2 days, we can  
 461 conclude the  $l = -140^\circ, b = 0^\circ$  and  $l = 20^\circ, b = 50^\circ$  bins have the highest concentrations of observing time. By  
 462 summing the bins over heliocentric ecliptic latitudes ( $b$ ), we see a large disparity in  $\sum t_i$  vs  $l$ , target count vs  $l$ ,

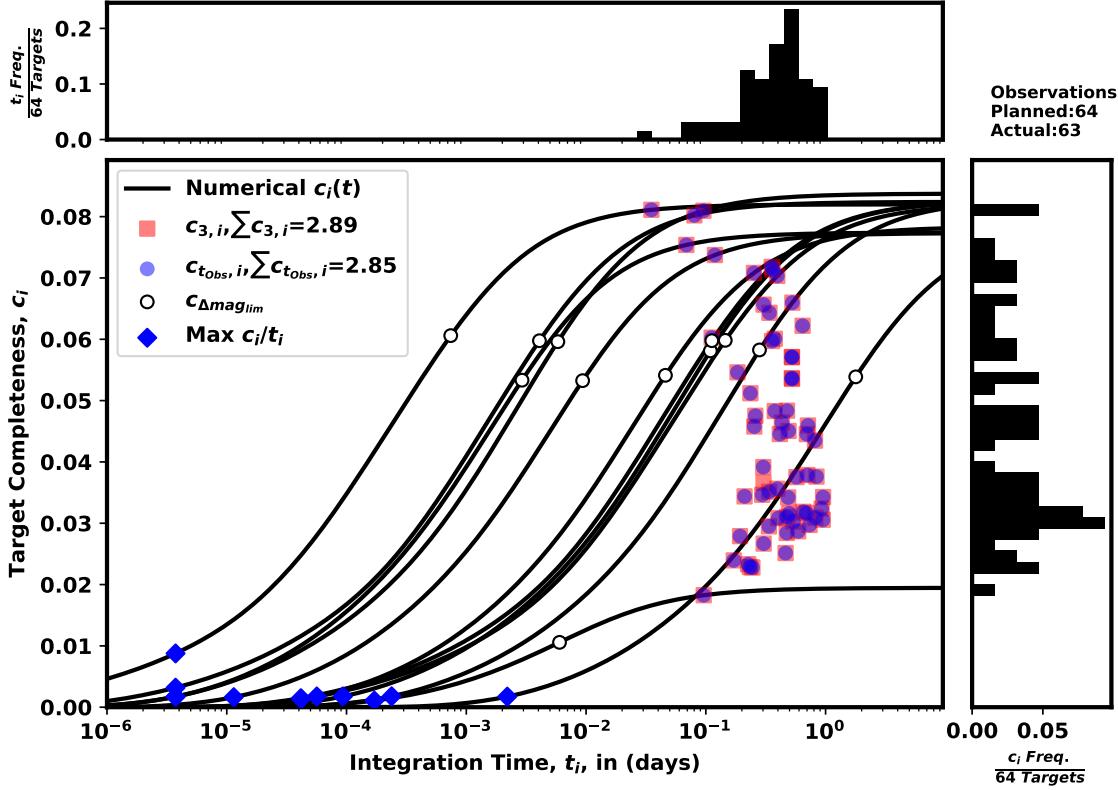


Figure 8: Completeness as a function of integration times calculated using the SAG13 distribution, EXOCAT-1 star catalog, Nemati SNR model,<sup>33</sup> and Leinert Table Zodiacal Light.<sup>36</sup> Black lines show completeness vs integration time for the 10 highest planned completeness targets, the median completeness target, and the lowest completeness target. Red squares indicate planned integration time and planned completeness based on  $\varepsilon$  maximizing summed completeness. Blue dots indicate observation integration time and observation completeness of the simulated observations. White dots represent completeness at  $\Delta\text{mag}_{\text{lim}}$  and is plotted for the 10 highest, median, and lowest completeness planned targets. Blue diamond shows the completeness and integration time of the maximum  $c_i/t_i$  point for the 10 highest, median, and lowest completeness targets. This plot references data generated using C0vsT0andCvsTDATA\_WFIRSTcycle6core\_CSAG13\_PPSAG13\_2019\_04\_05\_19\_18\_.txt.

463 and  $\sum c_i$  vs  $l$ . Since WFIRST is planned to be on an L2 halo orbit and has a sun-orbital period of approximately  
 464 365.25 days, the Leinert local zodiacal light<sup>36</sup> is fixed in this rotating frame, and the time-distribution of stars is  
 465 uneven, the optimally scheduled mission will have preferential observing times consistent with the distribution  
 466 in Figure 9. This result is important for optimally distributing limited CGI time under the constraints of a 5 yr  
 467 mission shared with multiple other instruments. Using this distribution, we can create preferentially distributed  
 468 observing blocks for the CGI.

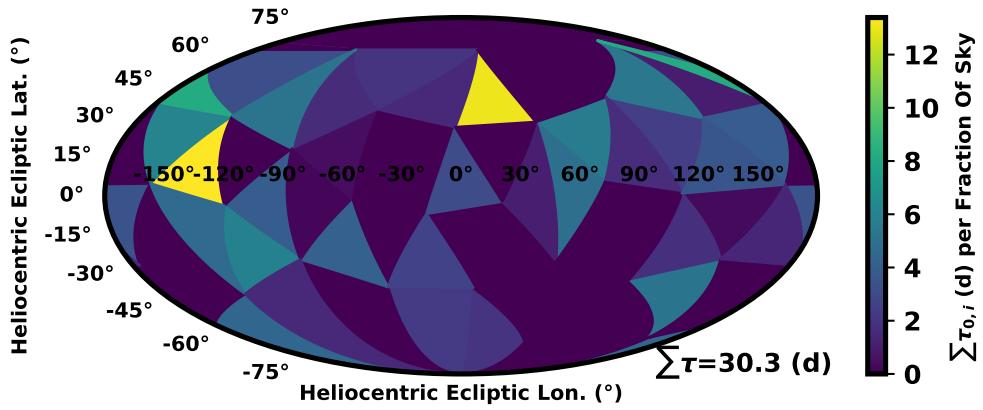
### 469 3.3 Detected Planet Properties

470 The top row of 10 are the average distribution of  $R$  and  $a$  for all generated planets in a universe. In each sub-  
 471 plot of Figure 10, the top left and bottom right show the number of simulations used to generate the resulting  
 472 distribution. Each 2D contour plot is normalized such that the integral over the area is 1 for the individual  
 473 ensemble so the color scale can be shared. The white number in each gridded region shows the average number  
 474 of planets generated or detected per simulation in that bin over the ensemble of simulations. The number in

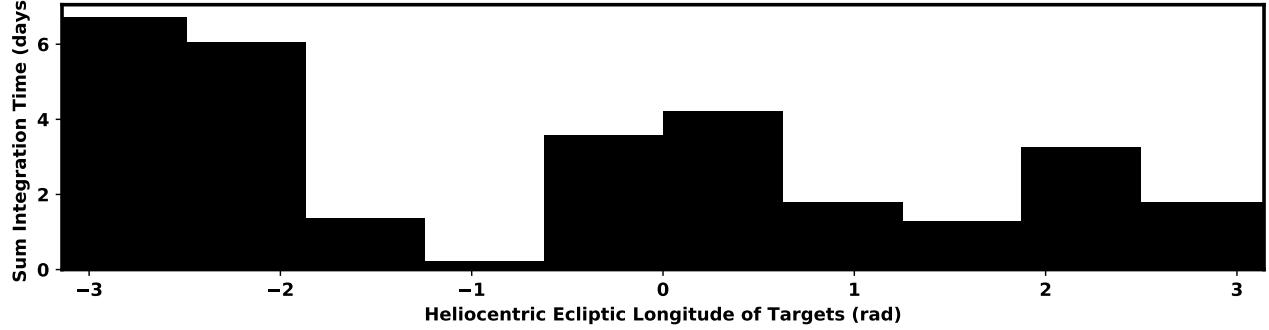
475 the top right of (a) and (b) of Figure 10 are the sum total of all planets generated in the ensemble of universes.  
476 Sub-plots (c)-(f) of Figure 10 show the average distribution of detected planets over an ensemble. The number in  
477 the top right of (c)-(f) of Figure 10 are the sum total of all planets detected in the ensemble. These summations  
478 have been tabulated as average yields in Table 4. Sub-plots (c) and (d) of Figure 10 use a target list optimized for  
479 the Kepler Like planet population to observe universes of Kepler Like and SAG13 simulated planets. Similarly,  
480 sub-plots (e) and (f) of Figure 10 use a target list optimized for the SAG13 planet population to observe universes  
481 of Kepler Like and SAG13 simulated planets.

482 We can do an analysis on the kinds of planets WFIRST is expected to detect in a blind search survey.  
483 Since our universe is randomly generated, we show the distribution of generated  $R$  vs  $a$  planets for all stars for  
484 the Kepler Like and SAG13 planet populations in Figure 10. The implemented planet generation rate for the  
485 universe of Kepler Like planets is  $\eta_{KL} \approx 2.377$ , consistent to 2 decimal places with the planet population model  
486 in section 2.1. Here,  $\eta$  is simply calculated by dividing the sum total of planets in the ensemble of universes by  
487 the number of simulations in the ensemble (1000 from section 2.5.3) and number of target stars (651 from section  
488 2.2). The  $\eta_{SAG13} \approx 5.618$  for SAG13 is also consistent to within 2 decimal places with the model in section  
489 2.1. The generated planet populations are consistent with the limits presented in section 2.1. In section 2.1 we  
490 showed the smallest observable planet-star separation observable to be 0.292 AU and each of the four detected  
491 planets plots show no detections for planets with  $a < 0.292$  AU. A distinctive feature of the planet  $a$  generation  
492 is the “knee” applied at 10 AU which can be seen by the sharp drop-off in both populations.

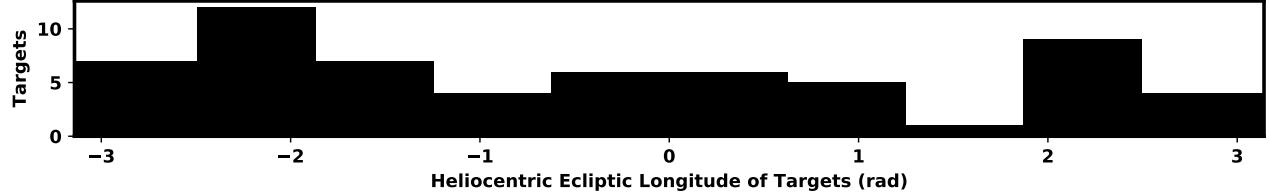
493 There is a marginal difference in the y-axis of the contour plots between Kepler Like and SAG13 populations,  
494 as SAG13 generates smaller planet radii than Kepler like, so direct comparisons cannot be made between individ-  
495 ual gridspace averages across universes. We also note the SAG13 universe generates an order of magnitude more  
496 large  $R$  and large  $a$  planets as indicated by the 24.29 and 221.19 grids from Kepler Like and SAG13 respectively  
497 (Figure 10 (a) and (b) gridspace (4,4)). Comparisons between planet populations used to calculate completeness  
498 observing the same planet population universe indicates that optimizing with the Kepler Like population results  
499 in marginally more smaller planet  $R$  detections. Specifically for the Kepler Like universe’s largest average detec-  
500 tion bin, we see that optimizing with the Kepler Like planet population results in 1.35 detected exoplanets on  
501 average and optimizing with the SAG13 planet population results in 1.11 detected exoplanets on average, which  
502 is significant, given our convergence results above. For the SAG13 universe’s largest averaged detection bin, we  
503 see optimizing with the Kepler Like planet population results in the 3.51 detections and optimizing with SAG13  
504 yields 3.55 detections but scaling by the number of detected planets gives  $3.54 = 3.51 \times 16.266 / 16.101$  which is  
505 nearly identical. Observing a universe of SAG13 planets with integration times optimized using completeness  
506 from a Kepler Like planet population results in more detections of small  $R$  planets. Each gridspace in the bottom  
507 two rows of the center right plot in Figure 10 are greater than or equal to each gridspace in the bottom two  
508 rows of the bottom right plot in Figure 10. We can specifically point to the 0.81 and 0.69 grid spaces that most  
509 evidently confirms this observation.



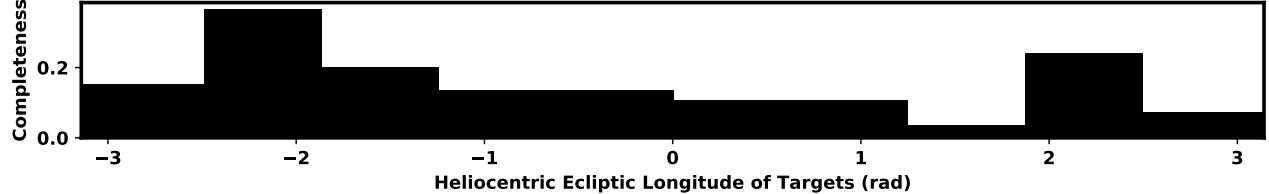
(a)



(b)



(c)



(d)

Figure 9: The distributions of a Kepler Like optimized target list including (a) a skymap divided into approximately evenly sized triangular bins with isotropic sky distribution showing the time/area density of observations, (b) a histogram of total sky time vs  $l$ , (c) a histogram of target counts vs  $l$ , (d) a histogram of summed completeness vs  $l$

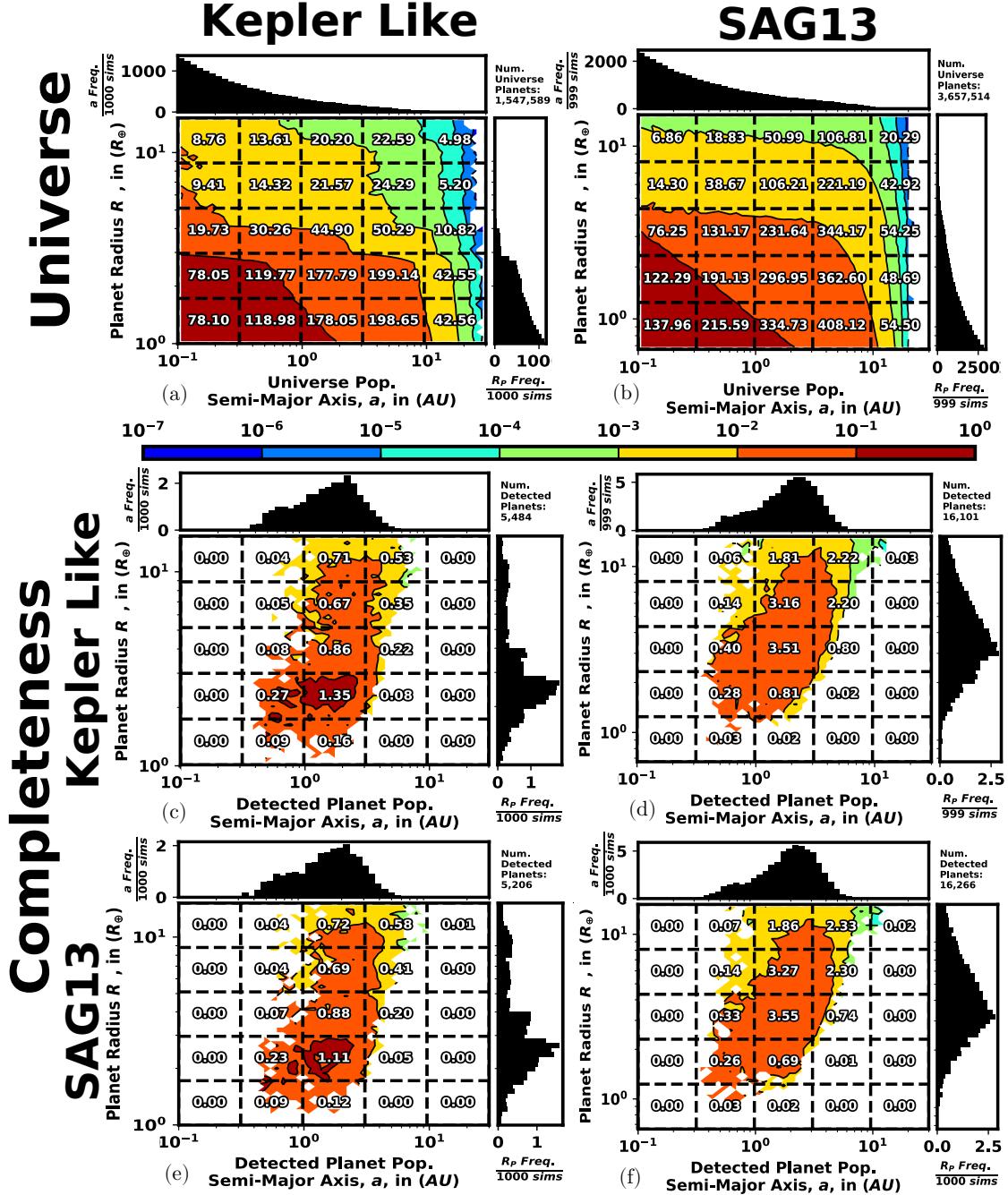


Figure 10: Universe of planets generated over all simulations for Kepler Like (top left) and SAG13 (top right) planet populations. Center row shows the population of detected planets for simulations run with completeness calculated using the Kepler Like planet population (center row) observing a universe of Kepler Like planets (center left) and SAG13 universe of planets (center right). Bottom row shows the population of detected planets for simulations run with completeness calculated using the SAG13 planet population (bottom row) observing a universe of Kepler Like planets (bottom left) and SAG13 universe of planets (bottom right). Overlay text shows average occurrences (top row) or detections (bottom rows) per grid-space per simulation. These plots reference RpvSMA detections DATA\_WFIRSTcycle6core\_CKL2\_PPKL2\_2019\_04\_05\_19\_34\_.txt.

510 **3.4 Overfitting**

511 We have chosen to use the Kepler Like and SAG13 planet populations to optimize target lists in this paper,  
 512 both of which are created based on known exoplanets. Since a motivation for the WFIRST CGI is to observe  
 513 new exoplanets in an unexplored region of space, we must investigate how yield for a target list of integration  
 514 times optimized for one planet population changes when observing a universe full of planets based off another  
 515 planet population. In Table 4, optimizing a target list using completeness based off a planet population and  
 516 observing that same planet population results in the highest yield. However, observing a universe of Kepler  
 517 Like planets with a target list optimized for SAG13 planets results in a 5.06% decrease in exoplanets detected, a  
 518 greater decrease than observing SAG13 planets with a target list optimized using Kepler Like planets (a 1.01%  
 519 decrease). A possible explanation is our inclusion of the rare characterization but the characterization part of  
 520 Table 3.4 does not indicate this is the case. A characterization observation is triggered whenever a planet is  
 521 detected with sufficiently small  $\Delta\text{mag}$  and separation such that the immediate re-observation could achieve a  
 522  $SNR > 10$  with a newly calculated  $t_i < 30$  d and all other detection observation filters in Section 2.5. Since the  
 523 observation would be carried out with a new instrument, the changed parameters are included in Appendix B.

Table 4: Summary of overfitting average unique detection yield and average characterizations from four Monte Carlo ensembles with optimized target list integration times calculated for different planet populations observing universes of different planets WFIRSTCompSpecPriors\_WFIRSTcycle6core\_3mo\_405\_19

	Planet Population		
	Completeness	Kepler Like	SAG 13
Average Yield	Kepler Like	5.484	16.117
	SAG13	5.206	16.266
Average Characterizations	Kepler Like	0.214	1.003
	SAG13	0.217	0.718

524

## 4. CONCLUSIONS

525 In this work, we presented our sequential least squares quadratic programming technique (SLSQP) for optimizing  
 526 an observation target list and integration times under non-linear overhead time constraints for blind search single-  
 527 visit observations of a generalized coronagraph on a generalized population of planets. We discuss EXOSIMS,  
 528 our framework for bookkeeping all time varying aspects of the mission including keep-out regions of major  
 529 planets from NAIF Kernels, spacecraft orbital position, local zodiacal light, positions of target stars, positions of  
 530 simulated exoplanets around target stars, and strict enforcement of total observing time. We show the optimal  
 531 heliocentric ecliptic longitudes for target observation based on local zodiacal light intensity as well as more realistic  
 532 local zodiacal light average minimum and average maximum magnitudes of 22.79 mag arcsec $^{-2}$  and 21.59 mag  
 533 arcsec $^{-2}$ , respectively. The strictness of WFIRST's keep-out regions causes the visibility of some targets to  
 534 drop below < 28% which severely constrains both single visits and would constrain revisits. By making only  
 535 optimal observations in our survey simulations, we see the optimized target list's preferential observing times are  
 536 unevenly distributed on the sky and across heliocentric ecliptic longitudes but these longitudes do not necessarily  
 537 line up with the completeness longitude distribution. This mismatch between required time across the sky and  
 538 the potential benefit of observing targets at different longitudes has a specific impact on the optimal scheduling  
 539 of observations over the finite time of a mission. Most importantly, it means that in the case where a mission is  
 540 designed with pre-allocated observing blocks, blocks reserved for exoplanet imaging must take into account both  
 541 the integration time and completeness distributions over ecliptic longitudes.

542 We validated our WFIRST optimized target list's summed completeness of 2.31 by performing a Monte Carlo  
 543 of survey simulations on the Kepler Like planet population, achieving an observed summed completeness of 2.28  
 544 using our model of WFIRST. We have also demonstrated the convergence of the mean number of detected  
 545 exoplanets from Monte Carlo of 1000 simulations to have 3.19% error at  $3\sigma$ , giving additional confidence in  
 546 the Monte Carlo results. Our results also indicate that the WFIRST CGI detects 5.48 exoplanets, on average,  
 547 when observing the Kepler Like planet population and 16.26 exoplanets on average when observing the SAG13  
 548 distribution (driven entirely by the different occurrence rates of these two simulated populations). Observing a

549 population of Kepler Like planets with a target list optimized for a SAG13 population of planets results in a  
550 decrease in detections of  $> 5\%$  whereas the inverse indicates a yield decrease  $< 1\%$  indicating the Kepler Like  
551 planet population optimized target list is less sensitive and should be used to optimize CGI target lists. More  
552 generally, based on multiple such results with mismatched planning and execution planet populations, lead us  
553 to conclude that we should strive to use more pessimistic (in terms of planet occurrence) populations for survey  
554 planning and optimization when maximizing unique detections.

## 555 5. FUTURE WORK

556 In this work, we included spectral characterizations of newly detected planets with sufficiently large achievable  
557 SNR. Characterization time was counted against total observing time and therefore reduced total time spent  
558 making detections. Since characterizations only occur immediately following a detection observation, some char-  
559 acterization observations may be systematically omitted. This is due to the pre-planned observations occurring  
560 at  $f_{Z,min}$  with adjacency of heliocentric ecliptic longitude of  $r_{i/SC}$  to keep-out regions where characterization  
561 observations would be filtered if a detection was made on a leading edge of a keep-out region. Additional work  
562 on intelligently scheduling future characterization observations is required. EXOSIMS makes such intelligent  
563 tracking of planet positions possible.

564 Since characterization observations are substantially longer than detection observations, a dynamic program  
565 should be implemented to test strategies for making revisit and characterization observations that minimize  
566 observing time when the detected planet is not visible. An option is to constrain the planet's orbit and observe  
567 at optimal planet phase but even modest uncertainty in orbital parameters results in large planet position and  
568 brightness uncertainty. Another option is to make detection observations immediately prior to characterizing,  
569 thereby saving the time that would otherwise be spent characterizing. We should also plan for the expected  
570 decreased observing time caused by characterizations. WFIRST is not particularly sensitive to this issue because  
571 of a relatively short CGI mission time with a lower throughput resulting in a target starved environment and  
572 less characterization candidates. Running SLSQP on larger, more capable observatory designs like HabEx and  
573 LUVOIR results in unplanned characterizations at nearly every star which consumes a large portion of mission  
574 time and many planned detections are not made.

575 Science reward for various images and discoveries needs to be enumerated, categorized, and weighted to  
576 accommodate more than single-visit unique detections.<sup>44</sup> Such metrics would include planet characterizations  
577 and detections by planet sub-type (i.e. hot/cold Earths/Jovians), planetary system packing, planet re-detection,  
578 marginal planet orbit constraint improvement ( $a$ ,  $e$ , inclination, and mean anomaly), and planet property ( $R$   
579 and albedo) constraint improvement. These metrics could have non-linear weighting to simulating diminishing  
580 returns of similar data products. Crafting a dynamic program out of these weights with probabilistic modeling  
581 can produce optimal observing programs for a board spectrum of science objectives.

582 The solar panel power keep-out region is substantial and, while not significantly impacting single-visit direct  
583 detection yield, will greatly impact revisit yield and orbit characterization of planets with orbital period similar  
584 to Earth. An analysis on the sub-space of Earth-Like planets unobservable caused by the selection of an orbit  
585 with an Earth-like period should be conducted.

586 The planet-star systems generated in these simulations are statistically consistent with known planet occur-  
587 rence rates. However, dynamically unstable planet-star systems may be created. If we used solely dynamically  
588 packed star systems, we could use detections of "easy to detect" planets to constrain and calculate the probability  
589 space of a harder to detect planet (i.e.  $P(\text{Planet2}(\Delta\text{mag}, s) \mid \text{Planet 1})$ ).

## 590 ACKNOWLEDGMENTS

591 This research has made use of NASA's NAIF planetary data system kernels. This research has made use of  
592 the Washington Double Star Catalog maintained at the U.S. Naval Observatory. This work was supported by  
593 the NASA Space Grant Graduate Fellowship from the New York Space Grant Consortium, NASA Grant Nos.  
594 NNX14AD99G (GSFC), NNX15AJ67G (WFIRST Preparatory Science), and NNG16PJ24C (WFIRST Science  
595 Investigation Teams). This research made use of Astropy, a community-developed core Python package for  
596 Astronomy (Astropy Collaboration, 2018) and OR-Tools, an optimization utility package made by Google Inc.

597 with community support. This research has made use of the Imaging Mission Database, which is operated by  
598 the Space Imaging and Optical Systems Lab at Cornell University. The database includes content from the  
599 NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with  
600 the National Aeronautics and Space Administration under the Exoplanet Exploration Program, and from the  
601 SIMBAD database, operated at CDS, Strasbourg, France.

## 602 REFERENCES

- 603 [1] National Research Council, *Panel Reports—New Worlds, New Horizons in Astronomy and Astrophysics*,  
604 National Academies Press, Washington, D.C., may 2011.
- 605 [2] D. Spergel, N. Gehrels, C. Baltay, D. Bennett, J. Breckinridge, M. Donahue, A. Dressler, B. S. Gaudi,  
606 T. Greene, O. Guyon, C. Hirata, J. Kalirai, N. J. Kasdin, B. Macintosh, W. Moos, S. Perlmutter, M. Post-  
607 man, B. Rauscher, J. Rhodes, Y. Wang, D. Weinberg, D. Benford, M. Hudson, W.-S. Jeong, Y. Mellier,  
608 W. Traub, T. Yamada, P. Capak, J. Colbert, D. Masters, M. Penny, D. Savransky, D. Stern, N. Zimmerman,  
609 R. Barry, L. Bartusek, K. Carpenter, E. Cheng, D. Content, F. Dekens, R. Demers, K. Grady, C. Jackson,  
610 G. Kuan, J. Kruk, M. Melton, B. Nemati, B. Parvin, I. Poberezhskiy, C. Peddie, J. Ruffa, J. K. Wallace,  
611 A. Whipple, E. Wollack, and F. Zhao, “wide-field infrared survey telescope-astrophysics focused telescope  
612 assets wfIRST-afta 2015 report,” *arXiv preprint arXiv:1503.03757*.
- 613 [3] R. A. Brown, “SingleVisit Photometric and Obscurational Completeness,” *The Astrophysical Journal* **624**,  
614 pp. 1010–1024, may 2005.
- 615 [4] C. C. Stark, A. Roberge, A. Mandell, and T. D. Robinson, “Maximizing the exoearth candidate yield from  
616 a future direct imaging mission,” *The Astrophysical Journal* **795**(2), p. 122, 2014.
- 617 [5] Jet Propulsion Laboratory, “HabEx Interim Report,” tech. rep., 2018.
- 618 [6] Goddard Space Flight Center, “LUVOIR Interim Report,” tech. rep., 2018.
- 619 [7] D. Garrett and D. Savransky, “ANALYTICAL FORMULATION OF THE SINGLE-VISIT COMPLETE-  
620 NESS JOINT PROBABILITY DENSITY FUNCTION,” *The Astrophysical Journal* **828**, p. 20, aug 2016.
- 621 [8] D. Garrett, D. Savransky, and B. Macintosh, “A Simple Depth-of-Search Metric for Exoplanet Imaging  
622 Surveys,” *The Astronomical Journal* , 2017.
- 623 [9] S. L. Hunyadi, S. B. Shaklan, and R. A. Brown, “The lighter side of TPF-C: evaluating the scientific gain  
624 from a smaller mission concept,” International Society for Optics and Photonics, sep 2007.
- 625 [10] R. A. Brown and R. Soummer, “NEW COMPLETENESS METHODS FOR ESTIMATING EXOPLANET  
626 DISCOVERIES BY DIRECT DETECTION,” *The Astrophysical Journal* **715**, pp. 122–131, may 2010.
- 627 [11] D. Savransky, C. Delacroix, and D. Garrett, “EXOSIMS,” 2017.
- 628 [12] R. W. Farquhar, “The utilization of halo orbits in advanced lunar operations,” *NASA Technical Note* ,  
629 pp. 1–99, Jul 1971.
- 630 [13] R. Lougee-Heimer, “The Common Optimization INterface for Operations Research: Promoting open-source  
631 software in the operations research community,” *IBM Journal of Research and Development* **47**, pp. 57–66,  
632 jan 2003.
- 633 [14] P. T. Boggs and J. W. Tolle, “Sequential Quadratic Programming,” *Acta Numerica* **4**, p. 1, 1995.
- 634 [15] D. Savransky, C. Delacroix, and D. Garrett, “Multi-mission modeling for space-based exoplanet imagers,”  
635 in *Techniques and Instrumentation for Detection of Exoplanets VIII*, S. Shaklan, ed., p. 54, SPIE, sep 2017.
- 636 [16] D. Savransky and D. Garrett, “WFIRST-AFTA coronagraph science yield modeling with EXOSIMS,”  
637 *Journal of Astronomical Telescopes, Instruments, and Systems* **2**, p. 011006, dec 2015.
- 638 [17] D. Keithly, D. Savransky, D. Garrett, and C. Delacroix, “Scheduling and target selection optimization  
639 for exoplanet imaging spacecraft,” in *Space Telescopes and Instrumentation 2018: Optical, Infrared, and  
640 Millimeter Wave*, H. A. MacEwen, M. Lystrup, G. G. Fazio, N. Batalha, E. C. Tong, and N. Siegler, eds.,  
641 **10698**, p. 191, SPIE, jul 2018.
- 642 [18] J. Krist, B. Nemati, and B. Mennesson, “Numerical modeling of the proposed WFIRST-AFTA coronagraphs  
643 and their predicted performances,” *Journal of Astronomical Telescopes, Instruments, and Systems* **2**(1),  
644 p. 011003, 2015.

- [19] F. Fressin, G. Torres, D. Charbonneau, S. T. Bryson, J. Christiansen, C. D. Dressing, J. M. Jenkins, L. M. Walkowicz, and N. M. Batalha, "THE FALSE POSITIVE RATE OF KEPLER AND THE OCCURRENCE OF PLANETS," *The Astrophysical Journal* **766**, p. 81, mar 2013.
- [20] D. Savransky, "Space mission design for exoplanet imaging," in *Proc. SPIE*, **8864**, 2013.
- [21] R. K. Kopparapu, E. Hébrard, R. Belikov, N. M. Batalha, G. D. Mulders, C. Stark, D. Teal, S. Domagal-Goldman, and A. Mandell, "Exoplanet Classification and Yield Estimates for Direct Imaging Missions," *The Astrophysical Journal* **856**, p. 122, mar 2018.
- [22] D. Garrett and D. Savransky, "Building Better Planet Populations for EXOSIMS," *American Astronomical Society, AAS Meeting #231, id.#246.04* **231**, 2018.
- [23] A. V. Moorhead, E. B. Ford, R. C. Morehead, J. Rowe, W. J. Borucki, N. M. Batalha, S. T. Bryson, D. A. Caldwell, D. C. Fabrycky, T. N. Gautier, D. G. Koch, M. J. Holman, J. M. Jenkins, J. Li, J. J. Lissauer, P. Lucas, G. W. Marcy, S. N. Quinn, E. Quintana, D. Ragozzine, A. Shporer, M. Still, and G. Torres, "THE DISTRIBUTION OF TRANSIT DURATIONS FOR KEPLER PLANET CANDIDATES AND IMPLICATIONS FOR THEIR ORBITAL ECCENTRICITIES," *The Astrophysical Journal Supplement Series* **197**, p. 1, nov 2011.
- [24] A. Cumming, R. P. Butler, G. W. Marcy, S. S. Vogt, J. T. Wright, and D. A. Fischer, "The Keck Planet Search: Detectability and the Minimum Mass and Orbital Period Distribution of Extrasolar Planets," *Publications of the Astronomical Society of the Pacific* **120**, pp. 531–554, may 2008.
- [25] A. Howard, G. Marcy, J. Johnson, D. Fischer, J. Wright, H. Isaacson, J. Valenti, J. Anderson, D. Lin, and S. Ida, "The occurrence and mass distribution of close-in super-earths, neptunes, and jupiters," *Science* **330**(6004), pp. 653–655, 2010.
- [26] R. Belikov, C. Stark, D. Angerhausen, D. Apai, E. Bendek, D. Bennett, G. Blackwood, A. Boss, R. Brown, G. Bryden, S. Bryson, K. Cahoy, D. Caldwell, J. L. Callas, J. Catanzarite, D. Ciardi, C. Clanton, N. Cowan, and W. Danchi, "Exoplanet Occurrence Rates and Distributions," tech. rep., 2017.
- [27] D. Garrett, D. Savransky, and R. Belikov, "Planet occurrence rate density models including stellar effective temperature," *Publications of the Astronomical Society of the Pacific* **130**(993), p. 114403, 2018.
- [28] R. B. Fernandes, G. D. Mulders, I. Pascucci, C. Mordasini, and A. Emsenhuber, "Hints for a turnover at the snow line in the giant planet occurrence rate," *The Astrophysical Journal* **874**, p. 81, mar 2019.
- [29] K. L. Cahoy, M. S. Marley, and J. J. Fortney, "EXOPLANET ALBEDO SPECTRA AND COLORS AS A FUNCTION OF PLANET PHASE, SEPARATION, AND METALLICITY," *The Astrophysical Journal* **724**, pp. 189–214, 2010.
- [30] D. Savransky, E. Cady, and N. J. Kasdin, "PARAMETER DISTRIBUTIONS OF KEPLERIAN ORBITS," *The Astrophysical Journal* **728**(7pp), p. 66, 2011.
- [31] M. C. Turnbull, "ExoCat-1: The Nearby Stellar Systems Catalog for Exoplanet Imaging Missions," *arXiv preprint arXiv:1510.01731* , 2015.
- [32] G. L. Wycoff, B. D. Mason, and S. E. Urban, "DATA MINING FOR DOUBLE STARS IN ASTROMETRIC CATALOGS," *The Astronomical Journal* , pp. 50–60, 2006.
- [33] B. Nemati, "Detector selection for the WFIRST-AFTA coronagraph integral field spectrograph," International Society for Optics and Photonics, aug 2014.
- [34] M. J. Pecaut and E. E. Mamajek, "INTRINSIC COLORS, TEMPERATURES, AND BOLOMETRIC CORRECTIONS OF PRE-MAIN-SEQUENCE STARS," *The Astrophysical Journal Supplement Series* **208**(22pp), p. 9, 2013.
- [35] W. A. Traub, J. Breckinridge, T. P. Greene, O. Guyon, N. Jeremy Kasdin, and B. Macintosh, "Science yield estimate with the Wide-Field Infrared Survey Telescope coronagraph," *Journal of Astronomical Telescopes, Instruments, and Systems* **2**, p. 011020, mar 2016.
- [36] C. Leinert, S. Bowyer, L. K. Haikala, M. S. Hanner, M. G. Hauser, A.-C. Levasseur-Regourd, I. Mann, K. Mattila, W. T. Reach, W. Schlosser, H. J. Staude, and G. N. Toller, "The 1997 reference of diffuse night sky brightness," *PAGE 1 SUPPLEMENT SERIES Astron. Astrophys. Suppl. Ser* **127**, pp. 1–99, 1998.
- [37] C. H. Acton, "Ancillary data services of NASA's Navigation and Ancillary Information Facility," *Planetary and Space Science* **44**, pp. 65–70, jan 1996.

- 695 [38] C. Acton, N. Bachman, B. Semenov, and E. Wright, “A look towards the future in the handling of space  
 696 science mission geometry,” *Planetary and Space Science* **150**, pp. 9–12, jan 2018.
- 697 [39] G. Soto, D. Savransky, D. Garrett, and C. Delacroix, “Parameterizing the Search Space of Starshade Fuel  
 698 Costs for Optimal Observation Schedules,” *Journal of Guidance, Control, and Dynamics* (**submitted**, 2018).
- 699 [40] G. Soto, D. Keithly, D. Savransky, C. Delacroix, and D. Garrett, “Optimal starshade observation scheduling,”  
 700 2018.
- 701 [41] D. Keithly, D. Garrett, C. Delacroix, and D. Savransky, “WFIRST: Exoplanet Target Selection and Schedul-  
 702 ing with Greedy Optimization,” in *American Astronomical Society, AAS Meeting #231, id. 246.06*, **231**,  
 703 2018.
- 704 [42] D. Keithly, D. Savransky, D. Garrett, and R. Morgan, “Blind Search Single-Visit Exoplanet Direct Imaging  
 705 Yield for Space Based Telescopes,” in *American Astronomical Society, AAS Meeting #233, id.140.40*, **233**,  
 706 2019.
- 707 [43] D. Savransky, “SEQUENTIAL COVARIANCE CALCULATION FOR EXOPLANET IMAGE PROCESS-  
 708 ING,” *The Astrophysical Journal* **800**, pp. 100–9, 2015.
- 709 [44] N. S. Budden and P. D. Spudis, “Evaluating Science Return in Space Exploration Initiative Architectures,”  
 710 tech. rep., Houston, 1993.
- 711 [45] B. Nemati, J. E. Krist, and B. Mennesson, “Sensitivity of the WFIRST coronagraph performance to key  
 712 instrument parameters,” in *Techniques and Instrumentation for Detection of Exoplanets VIII*, S. Shaklan,  
 713 ed., p. 7, SPIE, sep 2017.
- 714 [46] J. E. Krist, “End-to-end numerical modeling of AFTA coronagraphs,” *Space Telescopes and Instrumentation  
 715 2014: Optical, Infrared, and Millimeter Wave* **9143**, p. 91430V, 2014.
- 716 [47] G. Soto, J. Lloyd, D. Savransky, K. Grogan, and A. Sinha, “Optimization of high-inclination orbits using  
 717 planetary flybys for a zodiacal light-imaging mission,” in *Techniques and Instrumentation for Detection of  
 718 Exoplanets VIII*, S. Shaklan, ed., **10400**, p. 68, SPIE, sep 2017.

## APPENDIX A.

## 720 A.1 Symbols

Table 5: Table of Symbols

$\eta$	expected number of planets per star
$\eta_{KL}$	$\eta$ for the Kepler Like planet population
$\eta_{SAG13}$	$\eta$ for the SAG13 planet population
$\Delta\text{mag}$	difference in brightness between the planet and star in magnitudes
$\Delta\text{mag}_{\text{lim}}$	observatory limiting planet-star magnitude
$\Delta\text{mag}_0$	planet-star magnitude used to calculate integration time cut-off for filter
$\Delta\text{mag}_{\min,i}$	lower limit of integration for $\Delta\text{mag}$ for the $i^{\text{th}}$ star
$t_i$	integration time for observation of the $i^{\text{th}}$ target star
$dc_i(t_i)/dt_i$	the derivative of $c$ with respect to integration time for the $i^{\text{th}}$ target star
$dt$	infinitesimally small amount of time
$e$	planet eccentricity
$a$	planet semi-major axis
$IWA$	inner working angle
$d_i$	sun to target star distance
$s_{\min}$	minimum planet-star separation observable by the instrument
$\mathbf{I}$	list of target stars to observe
$N$	number of target stars in $\mathbf{I}$
$N_p$	number of planets generated in a universe
$N_q$	number of planets in planetary radius bin $q$
$WA_0$	working angle used to calculate integration time cut-off for filter
$s_k$	planet
$s_{\max,i}$	maximum planet-star separation observable by the instrument for the $i^{\text{th}}$ target star
$s_{\min,i}$	minimum planet-star separation observable by the instrument for the $i^{\text{th}}$ target star
$OWA$	outer working angle of the instrument
$SNR$	signal to noise ratio of the planet signal
$\mathcal{F}_0$	zero-magnitude flux at $\lambda$ see Eq. B.1
$\nu_i(\lambda)$	$i^{\text{th}}$ star visual magnitude in B-V color
$BV_i$	color of the star as a difference between B and V bands in magnitudes
$A$	pupil area
$\Delta\lambda$	observing wavelength bandwidth
$\epsilon_q(\lambda)$	detector quantum efficiency see Figure 12
$\epsilon_{inst}$	attenuation due to optics specific to the science instrument
$\epsilon_{syst}$	attenuation due to optics specific to the coronagraph e.g. polarizer, Lyot stop, extra flat mirror
$\epsilon_{PC}$	photon counting efficiency
$\epsilon_{pp}$	post processing efficiency
$\lambda$	wavelength
$T(\lambda, WA)$	instrument core throughput at $\lambda$
$\Psi(\lambda, WA)$	core mean intensity see Figure 11
$\gamma(\lambda, WA)$	intensity transmission of extended background sources see Figure 11
$\Gamma(\lambda, WA)$	area of the Full Width Half Maximum region of the planet Point Spread Function
$\varepsilon$	$\equiv dc_i/dt_i$ the derivative of completeness w.r.t $t$ for a subset of stars
$T_{OH}$	time per observation needed for the CGI to dig the dark hole and reach thermal equilibrium at the new spacecraft attitude
$T_{\text{settling}}$	time per observation for necessary observatory upkeep
$\mathbf{C}_{p0}$	a vector of $C_{p,i}$ used for calculation of $\mathbf{c}_0$

Continued on next page

---

$C_{p,i}$	planet signal electron count rate of the $i^{\text{th}}$ planet in $s^{-1}$
$\mathbf{C}_{b0}$	a vector of $C_{b,i}$ used for calculation of $\mathbf{c}_0$
$C_{b,i}$	background noise electron count rate of the $i^{\text{th}}$ planet in $s^{-1}$
$\mathbf{C}_{sp0}$	a vector of $C_{sp,i}$ used for calculation of $\mathbf{c}_0$
$C_{sp,i}$	speckle spatial structure of the system in $s^{-1}$
$Z$	Zodiacal light intensity in magnitudes
$EZ$	Exo-Zodiacal light intensity in magnitudes
$f_{Z0}$	local zodiacal light surface brightness in intensity
$f_{EZ0}$	exo-zodiacal light surface brightness in intensity
$WA_{int}$	WA used for integration time calculations
$\mathbf{c}_0$	a vector of $c_{0,i}$
$\mathbf{t}_0$	a vector of $t_{0,i}$
$T_{\max}$	maximum amount of time to spend observing
$\mathbf{x}_1^*$	optimal binary vector output solution to Algorithm 1
$x_i$	binary decision variable to include the $i^{\text{th}}$ target in target list in Algorithm 1
$c_{0,i}$	nominal completeness for the $i^{\text{th}}$ target used in Algorithm 1
$t_{0,i}$	nominal integration time for the $i^{\text{th}}$ target used in Algorithm 1
$\mathbf{x}_2^*$	optimal binary vector output solution to Algorithm 2
$\varepsilon^*$	the optimal $\varepsilon$
$\mathbf{t}_2$	optimal integration time solutions to Algorithm 2
$\mathbf{x}_3^*$	a binary vector output solution to Algorithm 3
$c_i(t_i)$	completeness for the $i^{\text{th}}$ target star at integration time $t_i$
$\mathbf{c}(\mathbf{t})$	$c_i(t_i) \forall i \in \mathbf{I}$
$c_{\infty,i}$	theoretical maximum completeness for the $i^{\text{th}}$ target star for this instrument
$t_\infty$	an infinite integration time
$c_{3,i}$	planned completeness of target $i$ to be observed
$c_{t_{obs},i}$	completeness of target $i$ observed in a survey simulation
$\mathbf{t}_3$	a binary vector output solution to Algorithm 3
$f_{Z,i}$	local zodiacal light of the $i^{\text{th}}$ target star
$\mathbf{f}_{Z,\min}$	$f_{Z,i} \forall i \in \mathbf{I}$
$t_{f_{Z,\min}}$	times past $t_c$ when $\mathbf{f}_{Z,\min}$ occur
$t_c$	the current mission time
$\mu_{Z_{\min}}$	mean magnitude of minimum zodiacal light intensity of stars in $\mathbf{I}$
$\mu_{Z_{\max}}$	mean magnitude of maximum zodiacal light intensity of stars in $\mathbf{I}$
$N_{pix}$	number of pixels
$PPL$	is the number of pixels per lenslet and is simply the lenslSamp <sup>2</sup>
$PS$	is the detector pixel scale in arcsec per pixel
$NCTE$	is the net charge transfer efficiency
$C_{\mathcal{F}_0}(\lambda)$	spectral flux density as a function of $\lambda$
$ENF$	excess noise factor
$C_{sr,i}$	starlight residual count rate
$C_{z,i}$	zodiacal light count rate
$C_{ez}$	exozodiacal light count rate
$C_{dc}$	dark current count rate
$C_{cc}$	clock induced charge count rate
$C_{rn}$	readout noise count rate
$f_\lambda(\log_{10}(\lambda))$	zodiacal light wavelength correction factor
$h$	planck constant
$c'$	speed of light in a vacuum

---

Continued on next page

$a_{\text{knee}}$	planet semi-major axis occurrence drop-off knee
$a_{\text{norm}}$	normalization factor for $f_{\bar{a}}$
$f_{\bar{a}}$	planet semi-major axis occurrence distribution for Kepler Like planets
$f_{\bar{a} R}$	planet semi-major axis occurrence distribution given $R$ for SAG13 planets
$a_{\text{min}}$	minimum planet semi-major axis
$a_{\text{max}}$	maximum planet semi-major axis
$\mathbf{a}$	vector of planet semi-major axis $a_k$
$a_k$	semi-major axis of the $k^{\text{th}}$ planet
$f_{\bar{a}} \text{ integrated over } a_{0.8} \text{ to } a_{85}$	$f_{\bar{a}}$ integrated over $a_{0.8}$ to $a_{85}$
$a_{0.8}$	semi-major axis for a 0.8 day orbital period planet around a sun mass star
$a_{85}$	semi-major axis for an 85 day orbital period planet around a sun mass star
$R$	general planetary radius
$\mathbf{R}$	vector of planet radii with length $N_p$
$R_k$	planetary radius of the $k^{\text{th}}$ planet
$R_{\oplus}$	radius of the Earth
$\mathbf{R}_{85}$	vector of planet radius occurrence frequency values <sup>19</sup>
$\mathbf{R}_{\text{vals}}$	$\mathbf{R}_{85}$ normalized and scaled by $a_{85,\text{norm}}$
$R_{\text{vals},q}$	planet occurrence frequency for $q^{\text{th}}$ bin of $\mathbf{R}_{\text{vals}}$
$q$	index for each bin in Figure 7 of Ref. 19
$\mathbf{R}_q$	set of $N_q$ randomly generated $R$ from $R_{\text{vals},q}$
$\mu$	without subscripts, this is gravitational parameter for the sun
$\mu_{\text{det}}$	mean number of detections
$\gamma_i$	SAG13 planet population parameter in Table 1
$\alpha_i$	SAG13 planet population parameter in Table 1
$R_{\lim,i}$	SAG13 planetary radius parametric equation cut-off from $i = 0$ to $i = 1$
$C a_i$	intermediate constant of the SAG13 planet population
$\beta_i$	SAG13 planet population constant
$f_{\bar{R}}$	SAG13 planetary radius probability density function
$\sigma_e$	Rayleigh parameter for eccentricity
$e_{\text{min}}$	minimum allowed eccentricity
$e_k$	eccentricity of the $k^{\text{th}}$ planet
$e_{\text{max}}$	upper 95 <sup>th</sup> percentile for $e$
$n$	uniform random variable between 0 and 1
$\bar{e}$	mean eccentricity
$p$	planet albedo
$p_k$	albedo of the $k^{\text{th}}$ planet
$p_{\text{max}}$	maximum allowed albedo
$p_{\text{min}}$	minimum allowed albedo
$r_{k/i}$	position vector from star $i$ to planet $k$
$\underline{r}_{i/SC}$	vector from the spacecraft position to the $i^{\text{th}}$ target star
$\hat{r}_{i/SC}$	unit vector from the spacecraft to the target star $i$
$\underline{r}_{\odot/SC}$	vector from the spacecraft to the sun
$\hat{r}_{\odot/SC}$	unit vector from the spacecraft to the sun
$\underline{r}_{\text{Body}/SC}$	vector from the spacecraft to the planetary body

722 **A.2 Acronyms**

Table 6: Table of Acronyms

WFIRST	Wide-Field Infrared Survey Telescope
CGI	Coronagraphic Instrument
DRMs	Design Reference Missions
HabEx	Habitable Exoplanet Observatory
LUVOIR	Large Ultra-Violet-Optical-Infrared Surveyor
TPF-C	Terrestrial Planet Finder Coronagraph
AYO	Altruistic Yield Optimization
SLSQP	Sequential Least Squares Quadratic Programming
NASA	National Aeronautics and Space Administration
ExoPAG	Exoplanet Program Analysis Group
SAG13	ExoPAG Study Analysis Group
EXOSIMS	Exoplanet Open-Source Imaging Mission Simulator
IWA	Inner Working Angle
OWA	Outer Working Angle
MET	Mission Elapsed Time

723 **APPENDIX B.**

724 This Appendix contains functions and interpolants used in the calculation of  $\Delta\text{mag}$  in Eq. 12, including expres-  
 725 sions for the zero-magnitude flux, star apparent V magnitude, throughput, intensity transmission of extended  
 726 background sources, core mean intensity, core area, and quantum efficiency.

727 **B.1 Zero-Magnitude Flux**

728 The zero-magnitude flux,  $\mathcal{F}_0$ , used in the calculation of spectral flux density,  $C_{\mathcal{F}_0}$ , in Eqs. 13, 19 and 12, is given  
 729 by

$$\mathcal{F}_0(\lambda) = 10^4 \times 10^{(4.01 - \frac{\lambda - 550\text{nm}}{770\text{nm}})} \text{ph/s/m}^2/\text{nm} \quad (20)$$

730 from Ref. 35.

731 **B.2 Star Apparent V Magnitude**

732 The calculation of  $\Delta\text{mag}_i$  relies upon the star's apparent V magnitude,  $\nu_i(\lambda)$ , given by the parametric equation  
 733 from Ref. 35

$$\nu_i = \begin{cases} V_{\text{mag},i} + 2.20BV_i(1/\lambda - 1.818) & \lambda < 550\text{nm} \\ V_{\text{mag},i} + 1.54BV_i(1/\lambda - 1.818) & \lambda \geq 550\text{nm}. \end{cases} \quad (21)$$

734  $V_{\text{mag},i}$  is the V-band apparent magnitude of the  $i^{\text{th}}$  target star (in magnitudes) and  $BV_i$  is the color of the star  
 735 as measured by the difference between B and V bands (in magnitudes). The parameters for the target stars used  
 736 in the final target list from Table 9 in Appendix C.

737 **B.3 WFIRST Cycle 6 Parameters**

738 This section presents the WFIRST cycle 6 parameters<sup>18</sup> in Table 7 used to calculate completeness and integration  
 739 times.

## 740 B.4 Characterization Parameters

741 In this work, we include characterization observations, which slightly detract from executing the planned ob-  
 742 servation schedule in Table 9 of C. These characterization observations immediately follow the detection of a  
 743 planet so long as the star  $r_{i/SC}$  is not obstructed by a keep-out region, is not filtered by integration time, and  
 744 there is sufficient time to make the observation before the target enters a keep-out region. A detected planet  
 745 is considered for characterization if  $SNR > 10$  of a detected planet. These characterization observations use  
 746 slightly modified set of parameters.

747 Our characterization observation integration time is calculated using  $SNR = 10$ , at  $\lambda = 660$  nm,  $PPL = 4.0$ ,  
 748  $\epsilon_{syst} = 0.46584$ ,  $N_{pix} = 76$ ,  $PS = 0.02631$ , and  $Rs = 50$ .  $Rs$ , the spectral resolving power specifically changes  
 749 the  $\Delta\lambda$  to  $\lambda/Rs$ .

## 750 B.5 WFIRST Cycle 6 Derivative 2D Interpolants

751 The fits files used for 2D interpolants of core throughput  $T(\lambda, WA)$ , intensity transmission of extended back-  
 752 ground sources  $\gamma(\lambda, WA)$ , core mean intensity  $\Psi(\lambda, WA)$ , core area  $\Gamma(\lambda, WA)$ , and quantum efficiency  $\epsilon_q$  are  
 753 derivative of Ref. 45, Ref. 46, and Ref. 18.

## 754 B.6 Electron Count Rates

755 This section lays out Nemati's  $SNR$  equation and all sub-components from Ref. 33. The  $SNR$  equation used in  
 756 this paper is

$$757 t_i = \frac{SNR^2 \times C_{b,i}}{C_{p,i}^2 - (SNR \times C_{sp,i})^2}. \quad (22)$$

758 This uses  $C_{p,i}$ ,  $C_{b,i}$ , and  $C_{sp,i}$  which are respectively the planet signal, background signal, and residual speckle  
 759 spatial structure in electron count rates in  $s^{-1}$ .

We calcualte  $C_{p,i}$  using

$$759 C_{p,i} = C_{F_0} \times 10^{-0.4(\nu_i + \Delta mag_i)} \times T(\lambda, WA) \times \epsilon_{PC} \times NCTE. \quad (23)$$

760  $NCTE$  is the net charge transfer efficiency.

761 We calculate  $C_{b,i}$  using

$$762 C_{b,i} = ENF^2 \times (C_{sr,i} + C_{z,i} + C_{ez}) + (ENF^2 \times (C_{dc} + C_{cc}) + C_{rn}). \quad (24)$$

763  $C_{sr,i}$ ,  $C_{z,i}$ ,  $C_{ez}$ ,  $C_{dc}$ ,  $C_{cc}$ , and  $C_{rn}$  are electron count rates with units of  $s^{-1}$  for starlight residual, zodiacal  
 764 light, exozodiacal light, dark current, clock-induced-charge, and readout noise respectively.  $ENF$  is an excess  
 noise factor.

765 We calculate the starlight residual count rate,  $C_{sr,i}$ , for each target star  $i$  by

$$766 C_{sr,i} = C_{F_0} \times 10^{-0.4 \times \nu_i} \times \Psi(\lambda, WA) \times N_{pix}. \quad (25)$$

767 Where  $N_{pix}$  is the number of pixels in the photometric aperture ( $\Gamma(\lambda, WA)/\theta^2$ ) calculated by

$$768 N_{pix} = PPL \times \frac{\Gamma(\lambda, WA)}{PS^2}. \quad (26)$$

769 Where PPL is the number of pixels per lenslet and is simply the lenslSamp<sup>2</sup>, a parameter specified in Table  
 770 7. Where PS is the detector pixel scale in arcsec per pixel.  $\Psi(\lambda, WA)$  is the core mean intensity from Figure 11  
 in Appendix B.  $\nu_i$  is given by the Eq. B.2 in Appendix B.

770 We calculate the zodiacal light count rate,  $C_{z,i}$ , using

$$771 C_{z,i} = C_{F_0} \times f_{Z,i} \times \Gamma(\lambda, WA) \times \gamma(\lambda, WA). \quad (27)$$

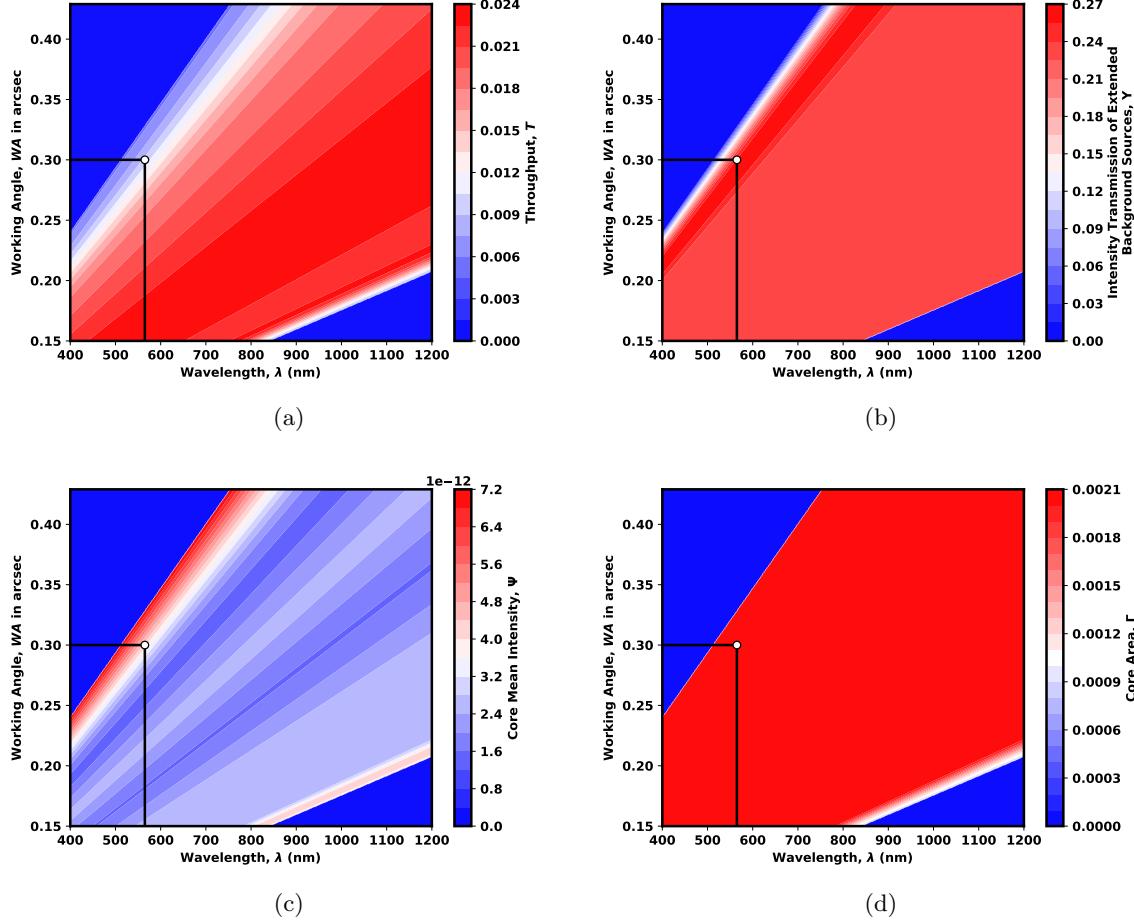


Figure 11: System throughput in the FWHM region of the planet PSF core (top left) from /WFIRST\_cycle6/G22\_FIT\_565/G22\_FIT\_565.thruput.fits. Intensity transmission of extended background sources such as zodiacal light including the pupil mask, Lyot stop and polarizer (top right) from /WFIRST\_cycle6/G22\_FIT\_565/G22\_FIT\_565\_occ\_trans.fits. Core mean intensity as a function of wavelength and working angle. Black lines and dot represent detection mode wavelength (565 nm) and inner working angle (0.3 arcsec) for integration time calculation (bottom left) from /WFIRST\_cycle6/G22\_FIT\_565/G22\_FIT\_565\_mean\_intensity.fits. Area of the FWHM region of the planet PSF in arcsec<sup>2</sup> (bottom right) from /WFIRST\_cycle6/G22\_FIT\_565/G22\_FIT\_565\_area.fits.

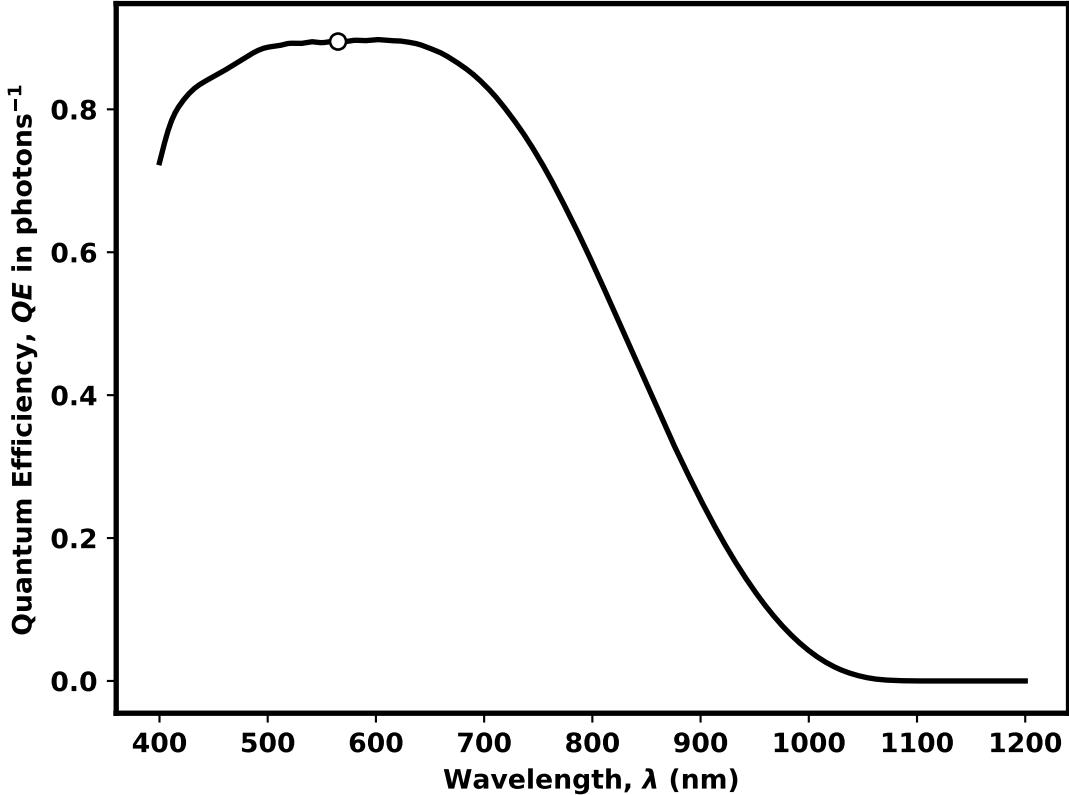


Figure 12: Quantum Efficiency of the detector as a function of wavelength. White dot represents detection mode wavelength (565 nm). From /WFIRST\_cycle6/QE/Basic\_Multi2\_std\_Si.fits

771 Where  $\gamma(\lambda, WA)$  comes from Figure 11 in Appendix B.

772 We calculate the exozodiacal light attributed count rate,  $C_{ez}$ , as

$$C_{ez} = C_{\mathcal{F}_0} \times f_{EZ} \times \Gamma(\lambda, WA) \times T(\lambda, WA). \quad (28)$$

773 We calculate the dark current count rate,  $C_{dc}$ , as

$$C_{dc} = N_{pix} \times i_{dark}. \quad (29)$$

774 Where  $i_{dark}$  is the dark current per pixel.

775 We calculate clock induced charge count rate,  $C_{cc}$ , as

$$C_{cc} = N_{pix} \times CIC/t_{exp}. \quad (30)$$

776 Where CIC is the clock induced charge per pixel and  $t_{exp}$  is the exposure time.

777 We calculate the readout noise count rate,  $C_{rn}$ , as

$$C_{rn} = N_{pix} \times s_{read}/t_{exp}. \quad (31)$$

778 Where  $s_{read}$  is the readout noise per pixel.

779 We calculate  $C_{sp,i}$  using

$$C_{sp,i} = C_{sr,i} \times \epsilon_{pp}. \quad (32)$$

780 Where  $\epsilon_{pp}$  is the post processing efficiency.

781 **B.7 Local Zodiacal Light**

782 Calculation of local zodiacal light is broken down into two major components: an intensity wavelength dependence  
 783 correction factor,  $f_\lambda(\lambda)$ , and intensity at the spacecraft centered look vector,  $f_\beta(\underline{r}_{i/SC})$ .

784 For  $f_\beta$ , we know the zodiacal dust cloud has structure, but the degree to which structure and phase/scattering  
 785 properties contribute to the zodiacal light intensity from a general observer location in space is currently uncertain  
 786 (although missions have been proposed to model such a dust cloud<sup>47</sup>). Knowing the degree of contribution  
 787 determines whether the anti-solar point of Table 17 of Ref. 36 should be modeled as fixed relative to the Earth  
 788 or fixed relative to the observer. To simplify our work, we assume the latter so  $\lambda - \lambda_0 = 0^\circ$  when  $\underline{r}_{i/SC} = \underline{r}_{\odot/SC}$   
 789 and the corresponding anti-solar point is when  $\underline{r}_{i/SC} = \underline{r}_{SC/\odot}$  and  $\lambda - \lambda_0 = 180^\circ$ .

790 To calculate local zodiacal light, we first find the position of the observatory in the heliocentric ecliptic frame  
 791  $\underline{r}_{SC/\odot}(t_c)$ . We then calculate

$$l_{SC/\odot}(t_c) = \text{sgn}(\underline{r}_{SC/\odot}(t_c) \cdot \hat{\underline{y}}) \cos^{-1} \left( \frac{\underline{r}_{SC/\odot}(t_c) \cdot \hat{\underline{x}}}{|\underline{r}_{SC/\odot}(t_c)|} \right). \quad (33)$$

792 We get the longitude of the sun relative to the spacecraft in the heliocentric frame  $l_{\odot/SC} = (l_{SC/\odot} + 180) \% 360$ .  
 793 We find the position vector describing the star position in the heliocentric true ecliptic frame  $\underline{r}_{i/\odot}$  and calculate  
 794 the star position with respect to the observatory  $\underline{r}_{i/SC} = \underline{r}_{i/\odot} - \underline{r}_{SC/\odot}(t_c)$ . We then transform  $\underline{r}_{i/SC}$  into spherical  
 795 coordinates using Astropy’s SkyCoord and extract the target star’s latitude ( $b_{i/SC}$ ) and longitude ( $l_{i/SC}$ ) relative  
 796 to the spacecraft. We then convert to absolute values for interpolation in the latitude and longitude range of  
 797 Figure 4 ( $0^\circ < b_i < 90^\circ$  and  $0^\circ < l < 180^\circ$ ) by  $b_i = |b_{i/SC}|$  and  $l_i = |(l_{i/SC} + 180^\circ) \bmod 360^\circ - 180^\circ|$  respectively.  
 798 This  $l$  and  $b$  are used in  $f_\beta(l, b)$ , a linear gridded interpolation of Table 17 in Ref. 36 and by extension  $f_Z(l, b, \lambda)$   
 799 in Eq. 19.

800 To assess the validity of our spacecraft centered vs geocentric ecliptic frame, we need to assess how much  
 801 the angular position of zodiacal light intensity interpolant inputs would differ. Since WFIRST is on a L2 Halo  
 802 orbit, its out of ecliptic motion is  $< 0.004$  AU and orbital distance from the Sun is  $\approx 1.010$  AU, resulting in a  
 803 geocentric ecliptic frame interpolation input deviation of  $\Delta b < 0.22^\circ$ . When  $\underline{r}_{i/SC}$  is  $180^\circ$  or  $0^\circ$  from  $\underline{r}_{\odot/SC}$ ,  
 804 the  $l$  and  $b$  used for interpolation is correct, however, interpolating for a target at say  $l = 90^\circ$  has the value  
 805 somewhere between  $89^\circ < l < 90^\circ$  due to the actual position of the spacecraft at the L2 Halo orbit and not  
 806 Earth. We expect  $\Delta l < 1^\circ$  for a Sun-Earth L2 orbit. We now make note that the smallest griddspacing of the  
 807 input data is  $5^\circ$  meaning  $\Delta l$  and  $\Delta b$  are within these bounds. It is also important to note the accuracy of this  
 808 Zodiacal light model is, at best, 10%.<sup>36</sup> The final component necessary to complete Eq. 19 is  $f_\lambda(\lambda)$ , a wavelength  
 809 correction factor. Which has a detailed explanation in Ref. 36.

810

## APPENDIX C.

Table 9: Planned observation target list optimized using the Kepler Like planet population. sInd refers to the index of the planet in the filtered, initial target list of 651 targets. The “Known Planet” column in Table 9 was generated by taking all target star names in the optimized target list derived from the EXOCAT-1 star catalog, and cross-referencing them using a list of aliases from SIMBAD and the NASA Exoplanet Archive. In total, 9 of the 60 planned targets already have known exoplanets.

sInd	Name	V mag	$d_i$ (pc)	B-V	$t_{obs,i}$ (d)	$c_i(t_{obs,i})$	Known Planet
1	HIP 746	2.26	16.78	0.36	0.221	0.016	0
9	HIP 1599	4.23	8.59	0.58	0.611	0.046	0
12	HIP 2021	2.82	7.46	0.62	0.289	0.064	0
25	HIP 3765	5.74	7.45	0.89	1.171	0.028	0
46	HIP 7513	4.09	13.49	0.54	0.568	0.022	1
51	HIP 7918	4.96	12.74	0.62	0.878	0.020	0
52	HIP 7981	5.24	7.53	0.84	0.983	0.035	0

Continued on next page

53	HIP 8102	3.49	3.65	0.73	0.409	0.091	1
54	HIP 8362	5.63	10.07	0.8	1.134	0.024	0
70	HIP 10644	4.86	10.78	0.61	0.838	0.027	0
79	HIP 12777	4.1	11.13	0.49	0.572	0.032	0
80	HIP 12843	4.47	14.22	0.48	0.695	0.018	0
90	HIP 14632	4.05	10.54	0.6	0.555	0.035	0
97	HIP 15457	4.84	9.14	0.68	0.828	0.034	0
98	HIP 15510	4.26	6.04	0.71	0.618	0.065	1
100	HIP 16537	3.71	3.21	0.88	0.457	0.093	1
101	HIP 16852	4.29	13.96	0.58	0.631	0.019	0
104	HIP 17378	3.52	9.04	0.93	0.412	0.047	0
136	HIP 22449	3.17	8.07	0.46	0.349	0.056	0
153	HIP 24813	4.69	12.63	0.61	0.773	0.021	0
165	HIP 27072	3.59	8.93	0.48	0.436	0.048	0
176	HIP 28103	3.71	14.88	0.34	0.467	0.019	0
199	HIP 32349	-1.44	2.63	0.01	0.040	0.113	0
229	HIP 37279	0.4	3.51	0.43	0.091	0.104	0
233	HIP 37826	1.16	10.36	0.99	0.127	0.045	1
270	HIP 44127	3.1	14.51	0.21	0.340	0.021	0
280	HIP 46853	3.16	13.48	0.47	0.347	0.025	0
288	HIP 49908	6.6	4.87	1.34	1.379	0.026	0
295	HIP 51459	4.82	12.78	0.54	0.825	0.021	0
311	HIP 54872	2.56	17.91	0.13	0.258	0.014	0
318	HIP 56997	5.31	9.61	0.72	1.015	0.028	0
319	HIP 57443	4.89	9.22	0.66	0.849	0.035	1
321	HIP 57632	2.14	11	0.09	0.210	0.039	0
322	HIP 57757	3.59	10.93	0.52	0.435	0.034	0
328	HIP 59199	4.02	14.94	0.33	0.552	0.018	0
341	HIP 61317	4.24	8.44	0.59	0.614	0.046	0
353	HIP 64394	4.24	9.13	0.57	0.615	0.042	0
372	HIP 67927	2.68	11.4	0.58	0.270	0.036	0
376	HIP 68933	2.06	18.03	1.01	0.195	0.014	0
385	HIP 70497	4.04	14.53	0.5	0.554	0.019	0
398	HIP 71908	3.16	16.57	0.26	0.350	0.016	0
429	HIP 77257	4.41	12.12	0.6	0.671	0.025	0
433	HIP 77952	2.81	12.38	0.32	0.291	0.031	0
434	HIP 78072	3.85	11.25	0.48	0.501	0.033	0
488	HIP 86974	3.41	8.31	0.75	0.392	0.054	0
508	HIP 91262	0.03	7.68	0	0.078	0.068	0
527	HIP 95501	3.36	15.53	0.32	0.388	0.018	0
529	HIP 96100	4.67	5.75	0.79	0.760	0.062	0
540	HIP 97649	0.76	5.12	0.22	0.108	0.091	0
543	HIP 98036	3.71	13.7	0.86	0.458	0.023	0
557	HIP 99240	3.53	6.11	0.76	0.417	0.071	0
561	HIP 99825	5.72	8.91	0.91	1.163	0.024	1
571	HIP 102422	3.41	14.27	0.91	0.389	0.022	0
581	HIP 105199	2.43	15.04	0.24	0.241	0.021	0
585	HIP 105858	4.22	9.26	0.47	0.610	0.041	0
594	HIP 107556	2.85	11.87	0.3	0.298	0.033	0
602	HIP 109176	3.77	11.73	0.44	0.481	0.031	0
623	HIP 113368	1.23	7.7	0.14	0.135	0.066	1

Continued on next page

642	HIP 116727	3.21	14.1	1.03	0.349	0.023	1
645	HIP 116771	4.13	13.71	0.51	0.581	0.020	0

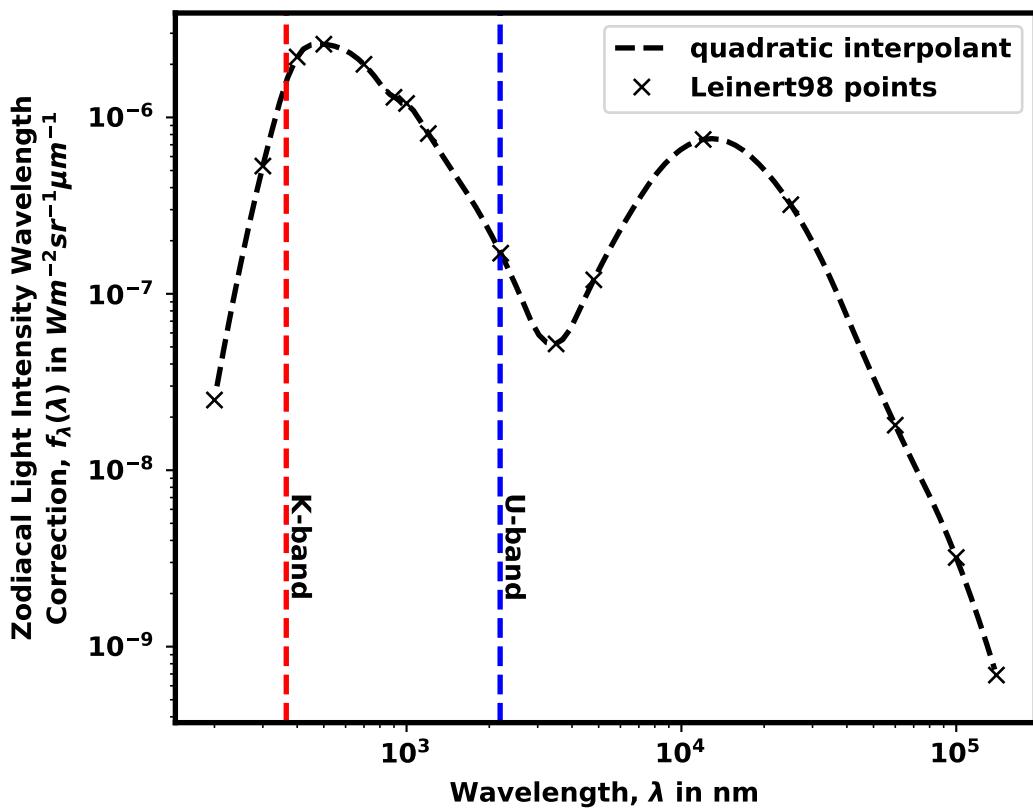


Figure 13: Local zodiacal light correction factor from Table 19 of Ref. 36 with region of high accuracy between 200 nm (red dashed line) and 2.0  $\mu m$  (blue dashed line) and decreasing accuracy at larger wavelengths due to infrared and scattering parity in contribution.

Table 7: WFIRST Cycle 6 input parameters<sup>18</sup> as defined by the EXOSIMS<sup>11</sup> JSON input script

Subgroup	Description	Value	EXOSIMS JSON Name
INST	Clock Induced Charge	0.01	CIC
INST	Excess Noise Factor	1.0	ENF
INST	Field Of View	9.5 arcsec	FoV
INST	Photon counting efficiency	0.8	PCeff
INST	Quantum Efficiency	Figure 12	QE
INST	Spectral Resolving Power (specific to spectrometers)	1	Rs
INST	Detector f-number	60.977	fnumber
INST	focal length	144.515 m	focal
INST	Dark current per pixel	0.000114 s <sup>-1</sup>	idark
INST	Lenslet sampling, number of pixel per lenslet rows or cols (specific to spectrometers)	1.0	lenslSamp
INST	Attenuation due to optics specific to the science instrument	0.518018	optics
INST	Detector array format, # of pixels per detector	1024	pixelNumber
INST	Pixel Scale in arcsec per pixel	0.01855469 arcsec	pixelScale
INST	Pixel pitch	1.3e-5 m	pixelSize
INST	Detector effective read noise per frame per pixel	0	sread
INST	Exposure Time	100 s	texp
SYST	Bandwidth Fraction	0.1	BW
SYST	Inner Working Angle	0.15 arcsec	IWA
SYST	Outer Working Angle	0.428996 arcsec	OWA
SYST	Area of FWHM region of planet PSF, in arcsec <sup>2</sup>	Figure 11	core_area
SYST	Mean starlight residual normalized intensity per pixel, required to calculate total core intensity as $\Psi \times N_{pix}$	Figure 11	core_mean_intensity
SYST	Core platescale	0.3	core_platescale
SYST	System throughput in the FWHM region of the planet PSF core	Figure 11	core_thrput
SYST	Bandwidth	56.5 nm	deltaLam
SYST	Central wavelength	565 nm	lam
SYST	Coronagraph name	HLC-565	name
SYST	Intensity transmission of extended background sources such as $f_Z$ . Includes pupil mask, occultor, Lyot stop and polarizer.	Figure 11	occ_trans
SYST	Overhead time	0.5 d	ohTime
SYST	Attenuation due to optics specific to the CGI e.g. polarizer, Lyot stop, extra flat mirror	0.983647	optics
SYST	Signal to noise ratio threshold	5	SNR
Observatory	Anti-solar keep-out region	124°	koAngleMax
Observatory	Maximum time past mission start an observation can be made	6 yrs	missionLife
Observatory	Fraction of mission life spent observing	0.04166	missionPortion
Observatory	Settling time after repoint	0.5 d	settlingTime
-	Post processing efficiency	0.1	ppFact
Technical	Δmag used to calculate minimum integration times for inclusion in target list	22.5	dMag0
Technical	Maximum allowed integration time in units of day	30d	intCutoff