

Predicting the Yields of $z > 6.5$ Quasar Surveys in the Era of Roman and Rubin

WEI LEONG TEE,¹ XIAOHUI FAN,¹ FEIGE WANG,¹ JINYI YANG,¹ SANGEETA MALHOTRA,² AND JAMES E. RHOADS²

¹*Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85719, USA*

²*Astrophysics Science Division, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, Maryland, 20771, USA*

ABSTRACT

Around 70 $z > 6.5$ luminous quasars have been discovered, strongly biased toward the bright end, thus not providing a comprehensive view on quasar abundance beyond cosmic dawn. We present the predicted results of Roman/Rubin high-redshift quasar survey, yielding 3 times more, 2–4 magnitudes deeper quasar samples, probing high-redshift quasars across broad range of luminosities, especially faint quasars at $L_{\text{bol}} \sim 10^{10} L_{\odot}$ or $M_{1450} \sim -22$ that are currently poorly explored. We include high- z quasars, galactic dwarfs and low- z compact galaxies with similar colors as quasar candidates. We create mock catalogs based on population models to evaluate selection completeness and efficiency. We utilize classical color dropout method in z and Y bands to select primary quasar candidates, followed up with Bayesian selection method to identify quasars. We show that overall selection completeness $> 80\%$ and efficiency $\sim 10\%$ at $6.5 < z < 9$, with 180 quasars at $z > 6.5$, 20 at $z > 7.5$ and 2 at $z > 8.5$. The quasar yields depend sensitively on the assumed quasar luminosity shape and redshift evolution. Brown dwarf rejection through proper motion up to 50% can be made for stars brighter than 25 mag, low- z galaxies dominate at fainter magnitude. Our results show that Roman/Rubin are able to discover a statistical sample of the earliest and faintest quasars in the Universe. The new valuable datasets worth follow up studies with James Webb Space Telescope and Extremely Large Telescopes, to determine quasar luminosity function faint end slope and constraint the supermassive black holes growth in the early Universe.

Keywords: Quasars (1319), Supermassive black holes (1663), Wide-field telescopes (1800)

1. INTRODUCTION

High-redshift quasars are key probes of the early Universe. Spectra of $z \gtrsim 6$ quasars provide crucial information regarding the properties of the intergalactic medium (IGM) during the epoch of reionization (EoR). Gunn-Peterson (GP) absorption in high resolution quasar spectra can be used to measure the neutral hydrogen content in the IGM and map the history of reionization (Fan et al. 2006; Becker et al. 2015). Recent quasar studies suggest that the IGM changes rapidly from being highly neutral to ionized in a relatively short period of time at $5.5 < z < 7.5$ (e.g. Bañados et al. 2018; Davies et al. 2018; Wang et al. 2020; Yang et al. 2020a). However, more sightlines are needed to sample the IGM evolution, especially at $z > 7$, to probe whether this transition is homogeneous or highly patchy.

The highest redshift quasars also directly measure the evolution of early supermassive black holes (SMBHs) and constrain the formation and growth history of the first generation black hole seeds. Recent observations

of quasars at $z > 6$ confirm the existence of massive SMBHs with $\sim 10^{8-10} M_{\odot}$ (e.g. Mortlock et al. 2011; Wu et al. 2015; Bañados et al. 2018; Matsuoka et al. 2019a,b; Shen et al. 2019; Wang et al. 2021; Yang et al. 2020a, 2021) in the young Universe, with the highest redshift quasars at $z \sim 7.5$ hosting $10^9 M_{\odot}$ SMBH. Standard SMBH formation models, in which SMBHs grow via Eddington-limited accretion from stellar mass black hole (e.g. Volonteri & Begelman 2010; Volonteri 2012), fail to explain the existence of these observed SMBHs, given the extreme short growth time. Several theoretical approaches have been investigated to explain the formation of observed SMBHs, through the direct collapse of a primordial cloud (e.g. Bromm & Loeb 2003; Begelman et al. 2006; Ferrara et al. 2014; Habouzit et al. 2016; Schauer et al. 2017; Dayal et al. 2019) into massive black hole seeds ($10^{4-6} M_{\odot}$), or by rapid growth of low mass seeds ($10^{2-3} M_{\odot}$) with periods of Eddington and super-Eddington accretion (e.g. Madau & Rees 2001; Ohsuga et al. 2005; Tanaka & Haiman 2009; Inayoshi et al. 2016), or growth in radiatively inefficient accretion modes (e.g.

Trakhtenbrot et al. 2017; Davies et al. 2019). Quasars at higher redshift and lower luminosity than those in the current sample will allow stronger tests to these models.

The discovery of high- z quasars starts at the beginning of this century, mostly using wide-field imaging data from the Sloan Digital Sky Survey (SDSS; e.g. Fan et al. 2001a, 2006), the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS; e.g. Willott et al. 2009), the Panoramic Survey Telescope and Rapid Response System 1 (Pan-STARRS 1; e.g. Bañados et al. 2016), the United Kingdom Infrared Telescope Infrared Deep Sky Survey (UKIDSS; e.g. Bañados et al. 2018), the VISTA Kilo-degree Infrared Galaxy survey (VIKING; e.g. Venemans et al. 2013); the VLT Survey Telescope ATLAS (VST-ATLAS; e.g. Reed et al. 2017); the Dark Energy Survey (DES; e.g. Yang et al. 2019; Wang et al. 2019); and the Hyper Suprime-Cam (HSC; e.g. Matsuoka et al. 2018) on the Subaru telescope. More than 200 quasars have now been discovered at $z \gtrsim 6$ from different wide-field surveys (see review Fan et al. 2022), and the sample size grows rapidly in the last decade due to the combined effort from deep optical and near-infrared (NIR) observations.

The search for quasars at $z \gtrsim 6.5$ requires photometric information from both optical and near-infrared (NIR) bands. Quasar flux passing through the high- z IGM leaves the trademark of a distinct Ly α break with flux blueward being completely absorbed, and therefore distinguishable from other types of objects with blue dropout in colors. Utilizing multi-band all sky survey information over past decade, ~ 70 quasars at $z > 6.5$ have been discovered, including 3 at $z > 7.5$ with the highest redshift at 7.6 (e.g. Mortlock et al. 2011; Bañados et al. 2018; Wang et al. 2018, 2021; Yang et al. 2019, 2020a,b; Matsuoka et al. 2019a,b). However, the current generation of quasar searches is limited to $z < 8$, due to a combination of shallow survey depth, insufficient area coverage, and enormous number of contaminants or junk introduced in the flux-limited sample. The contaminants are pickup up with similar NIR colors, majority of them are early type galaxies at $z \sim 1-2$ (ETGs), and Galactic cool stars or brown dwarfs, in particular M stars, L and T dwarfs (MLTs). The large number of contaminants in quasar candidates has elevated observational challenges in spectroscopic follow up observations. The success rate or efficiency, defined as the ratio of real quasars classified over possible candidates through spectroscopy, are in the order of few percents in present survey depths of $J < 22.5$ for $z \sim 6.5$ quasar searches (Wang et al. 2019). The efficiency at higher redshift and fainter magnitude drops drastically.

The next generation ground based and space based telescopes, such as the Vera C. Rubin Observatory, the Euclid Telescope, the James Webb Space Telescope (JWST), and the Roman Space Telescope, should transform quasar research at the highest redshift. Combining deep optical photometry from Rubin Observatory Legacy Survey of Space and Time (LSST) and Euclid Near Infrared Spectrometer and Photometer (NISP) wide-field imaging, Euclid Collaboration et al. (2019) investigates the quasar detection to Euclid depth. They report prediction of ~ 250 quasars at $7 < z < 9$, ~ 25 quasars at $8 < z < 9$ within Euclid 15000 deg² area, with moderate contamination at $J \sim 23$. Euclid selection however is flux-limited to high redshift sources at $J < 24$. To maximize the quasar yield and probe full understanding on quasar properties, deeper near-infrared surveys are required to push toward fainter magnitudes at which quasars are more abundant. Roman and Rubin combination is therefore a natural solution.

We aim to use the multi-band photometry of Roman and Rubin to develop sophisticated search methodology for $z > 6.5$ quasars. The current state of the art methods in discovering the Type 1 high- z quasars include color dropout and color-color cuts (e.g. Wang et al. 2019), Bayesian model comparison (BMC; e.g. Mortlock et al. 2012), SED model fitting (e.g. Reed et al. 2017), random forest classification (e.g. Schindler et al. 2017), and extreme de-convolution (XD; Nanni et al. 2022). In this paper we study joint selection methods: color dropout and color-color cut, followed with the Bayesian model comparison technique, to optimize the quasar yield and minimize the contamination. The paper is structured as follow: we describe the survey properties, especially the data sets that we are using, and the photometric models in Sec 2. We explain the population models and mock catalog generation in Sec 3. In Sec 4 we present the detail of each selection method. In Sec 5 we show the results from the simulation in terms of selection completeness. We apply selections to mock catalogs to justify our selection functions, discuss the limitations and summarize the expected yield. In Sec 6 we discuss the effect of proper motion and the benefit of this study. Finally we conclude in Sec 7. Through the paper we adopt a flat Λ CDM cosmology model with $H_0 = 70 \text{ kms}^{-1}\text{Mpc}^{-1}$, $\Omega_m = 0.272$, and $\Omega_\Lambda = 0.728$ (Komatsu et al. 2009), following McGreer et al. (2013) and Yang et al. (2016) where the quasar simulation has been inherited. All magnitudes and colors used, unless specifically mentioned, are on AB system, while uncertainties quoted are at 1σ confidence level.

2. SURVEY PROPERTIES

In this section, we consider photometric properties of Roman and Rubin wide-field surveys. The surveys' wavelength coverage and depth are shown in Table 1 and Fig 1. We introduce additional astrometry information to aid proper motion selection discussed in Sec 6.1. The construction of photometric catalog data is presented in Sec 2.5.

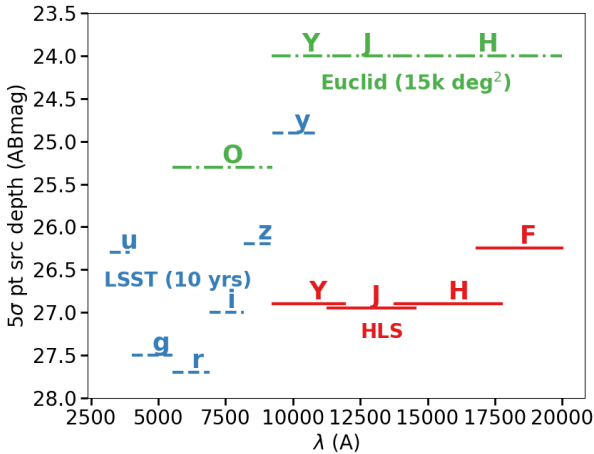


Figure 1. Wavelength coverage and depth of each survey up for the full mission time span. Both Roman HLS and Rubin LSST will reach 5σ depth of 26 in general. We do not use LSST near-IR Y because it is shallower than Y-band with Roman WFI.

2.1. Roman High Latitude Wide Area Survey (HLS)

The Roman HLS consists of both imaging (HLImagingS, HLIS) and low resolution (grism) spectroscopic (HLSpectroscopyS, HLSS) observations over a common survey region using the Wide Field Instrument (WFI) on the Roman Space Telescope, which has 100 times larger field of view than Wide Field Camera 3 (WFC3) on the Hubble Space Telescope (HST) (Green et al. 2012; Spergel et al. 2015). The survey will map twice covering approximately 2000 deg² in 5 years after the expected launched date of 2026/2027. Survey footprint has been selected to avoid Ecliptic and Galactic plane, and overlapped with optical ground telescopes. Current survey region has been placed near to $ra = [0, 47.5]$ deg, $dec = [-47.5, 0]$ deg, or $\sin b \sim 1$. HLIS will reach $\sim 26 - 27$ (5σ point source) in four NIR bands: Y, J, H, F.

At the time of this writing, the survey design is not yet finalized and the observing strategy may continue to evolve. For example, the possibility of adding a redder K_s filter ($2.0 - 2.4 \mu m$) has been discussed in Stauffer et al. (2018), with which quasars and galaxies can be better separated from stars with using $Y - K_s$ color than

using $Y - H$ alone. We do not include the additional K_s filter in this work. The final observing program details and cadence will be defined by future community process. In this work we only use information based on HLIS imaging for quasar search, adopting the $YJHF$ filters for our photometric measurements, and we assign random sky coordinates for sources drawn from HLIS survey region.

2.2. Rubin Legacy Survey of Space and Time (LSST)

Quasars at high- z are expected to have negligible flux blueward of redshifted Ly α . Deep optical data are needed to characterize the strong Lyman break. Possible crossover surveys that meet the criteria are the 6 years space-based Euclid Wide Survey and 10 years ground based LSST. Euclid Wide Survey (Euclid Collaboration et al. 2022) has 10σ depth at 24.5 in O band, which is an optical wide filter spanning 4900-9300 Å. LSST has i, z, y -bands reaching 5σ depth at 26.9, 26.1 and 24.9 at the end of 10 years, assuming uniform depth across the survey region. LSST broad bands provide sufficient sensitivity and contrast for color dropout selection. We consider LSST as the main source for optical data.

2.3. Astrometry

Cool dwarfs with high proper motions can be distinguished from stationary quasars using imaging surveys with multiple epochs data. Various studies have been carried out to characterize the kinematics of cool brown dwarfs in Galactic disk and halo (Faherty et al. 2009; Best et al. 2018). We can cross-validate target positional information using forced photometry to identify moving cool dwarfs. Roman WFI has pixel size and full width half maximum (FWHM) for point-spread function (PSF) of $0''.1$ and $0''.27$ in H filter (Troxel et al. 2023). The expected 5σ detection in proper motion is $4 - 16 \text{ mas year}^{-1}$, depending on the timing of Roman second pass imaging (WFIRST Astrometry Working Group et al. 2019). Rubin LSST maps its survey region with 20 observations each year. Estimated astrometry accuracy is $\sim 50 \text{ mas}$ for sufficiently bright objects at the end of the ten-year survey (Ivezić et al. 2019). To estimate the performance of Rubin in measuring the cool dwarf proper motion, we query `rubin_sim`¹, a simulation tool of Rubin Observatory, look for all detected LT cool dwarfs (primary spectral type of dwarf contaminants for high-redshift quasar search) over a 10 years baseline. We obtain the distances of the population, and with an assumed tangential velocity 30 km s^{-1}

¹ https://github.com/lsst/rubin_sim

(see Sec 3.2.3), we derive their respective proper motions. The furthest cool dwarf marks the lower bound of proper motion measured by Rubin, and it is a function of spectral type. The median of minimum proper motion detectable by Rubin across spectral types is found to be 30 mas year^{-1} .

2.4. Advantages of Roman and Rubin

LSST is scheduled to start science operation no later than 2024. By 2028 Roman will have collected 1 year imaging data while LSST will have at least 4 years of imaging data within HLIS survey area. Euclid and Roman are best-suited for different types of high- z quasar science: Euclid eventually will map much larger sky area for the discoveries of luminous quasars, which is crucial for studying massive SMBHs growth in the early Universe and the history of reionization; Roman will discover more low luminosity quasars, with bolometric luminosity close to that of typical AGN ($L_{\text{bol}} \sim 10^{10} L_{\odot}$), which enable the studies of evolution and abundance of SMBHs in the early Universe.

2.5. Photometric Model

Simulated photometric data with realistic noise that matches the expected observed datasets is necessary for investigating selection efficiency and completeness. Troxel et al. (2023) reports that the dominant noises are thermal background, Poisson noise and read-out noise in Roman detectors performance with Rubin/Roman imaging simulation. Our primary interests are quasars near the IR detecting limit, which are usually background-limited sources. After the noiseless photometry has been generated by integrating the simulated spectrum in Sec 3 over the respective filter throughput, the observed photometric errors are added as Gaussian noise with a lower bound of 0.01 for systematic error. We also include an additional 10% scatter for the construction of observed photometry to match with observational magnitude error trend. We list the photometry information used in this work in Table 1.

3. SIMULATION & CATALOG CONSTRUCTION

There are two different kind of simulated catalogs used in this study. We first generate a complete set of quasars distributed uniformly in a (J, z) grid. Simulation details are presented in Sec 3.1. This quasar simulation is used to determine the optimum color dropout and color-color cut selection. We then describe the population modelling, mainly on the number densities and color templates of population as a function of magnitude or redshift in Sec 3.2. Sec 3.3 describes the process of

Table 1. Photometric information used in this work.

Surveys	Band	λ (Å)	Depth (5σ)
(1)	(2)	(3)	(4)
Roman HLS (5 yrs)	Y	9270 - 11920	26.9
	J	11310 - 14540	26.95
	H	13800 - 17740	26.9
Rubin LSST (10 yrs)	F	16830 - 20000	26.25
	z	8030 - 9380	26.2

mock catalogs construction to enable realistic estimation of the completeness and efficiency of quasar selection.

3.1. Simulation of Quasars for Selection Function

Construction of color selection function requires a complete set of simulated quasars matched with observation. Fan (1999) proposes an empirical method to construct quasar spectrum using observed continuum slope and emission line strengths, assuming shape of quasar spectral energy distributions (SEDs) does not evolve with redshift, as well as models of intergalactic absorption blueward of Ly α emission. This approach has been extended to include higher redshift quasars in a number of recent works (e.g. McGreer et al. 2013, 2018; Yang et al. 2016; Jiang et al. 2016; Wang et al. 2019). In this section we briefly describe the spectral model and extend the simulation to include $z = 6-10$ quasars. The simulated quasar continuum spectrum is assumed to be a broken power law continuum $f_{\nu} \propto \nu^{-\mu}$ with breaks at 1100, 5700, 9730, 23820, and 30000 Å. The slopes in between these break points are chosen from Gaussian distributions with $\mu = -0.44, -0.48, -1.74, -1.17$ and $\sigma = 0.3$. Various emission lines, including Ly α and CIV, and FeII lines (Boroson & Green 1992; Vestergaard & Wilkes 2001; Glikman et al. 2006; Tsuzuki et al. 2006) are added to the continuum spectrum as Gaussian profiles. The distribution of emission line features, such as the continuum spectral slope, line equivalent width (EW), and line FWHM are designed to match the colors of SDSS BOSS quasars in the redshift range of $2.2 < z < 3.5$ (Ross et al. 2012). Quasars are generated via uniform grid of (J, z) , where $18 < J < 28$ with $\Delta J = 0.1$ and $6 < z < 10$ with $\Delta z = 0.1$; there are 100 quasars at each (J, z) grid position, with a total of 404,000 quasars generated with absolute magnitude at 1450 Å, M_{1450} , ranges from -30 to -18. The simulated spectra correctly reproduce quasar spectral characteristics seen in observation, e.g. Baldwin effect (Baldwin 1977) and blueshifted lines (Gaskell 1982; Richards et al. 2011).

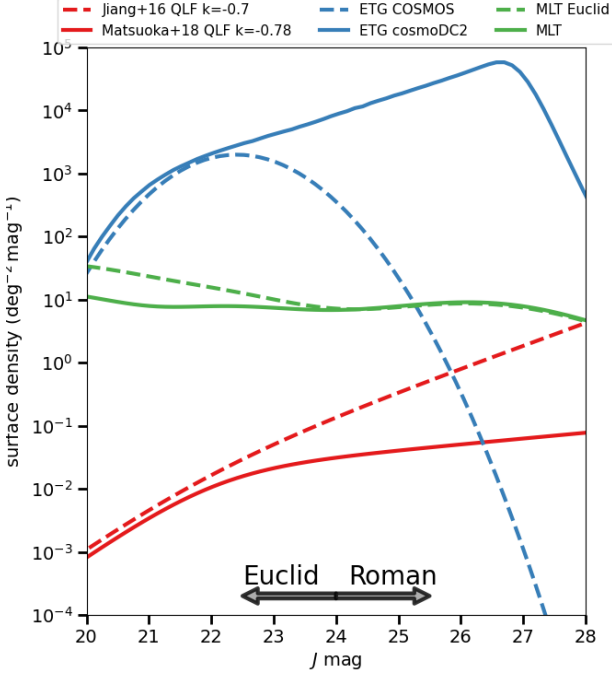


Figure 2. Candidate population surface densities as a function of J band magnitude. Detailed models can be found in Sec 3.2. The vast difference in quasar and galaxy densities in faint magnitudes points out the distinct target strategy between Barnett et al. (2021) (see Fig 3 in their work) and this work.

Red dashed: Quasars, by extrapolating the luminosity function and spatial density decline rate in Euclid Collaboration et al. (2019), and integrating over redshift $z = 6.5 - 9.0$. *Red solid:* Quasars using parameters from Wang+19 and Matsuoka+18, integrate over redshift $z = 6.5 - 9.0$. *Blue dashed:* COSMOS quiescent-flagged galaxies at redshift $z = 1-2$. *Blue solid:* cosmoDC2 red sequence galaxies at redshift $z = 1-2$. *Green dashed:* MLTs summed over all spectral types (M0-T8) modeled in Euclid Collaboration et al. (2019). *Green solid:* MLTs summed over all spectral types in this work (M0-T9) with slight difference in early M population densities.

3.2. Population Models

3.2.1. Quasars

The quasar luminosity function can be generally parametrized as a double power-law (e.g. Boyle et al. 1988):

$$\Phi(M_{1450}, z) = \frac{\Phi^*(z)}{10^{0.4(\alpha+1)(M_{1450}-M_{1450}^*)} + 10^{0.4(\beta+1)(M_{1450}-M_{1450}^*)}} \quad (1)$$

and the spatial density decline rate k is embedded within $\Phi^*(z)$,

$$\Phi^*(z) = \Phi^*(z=6) \times 10^{k(z-6)} \quad (2)$$

where M_{1450} is the absolute magnitude at 1450 Å, α and β are the faint end and bright end slopes, respec-

tively, M_{1450}^* is the characteristic absolute magnitude or break magnitude measured at 1450 Å (assume no redshift evolution), and Φ^* is the normalization of the LF. With high completeness quasar samples, QLF has been measured at $z \sim 5$ and extended to $z = 6$ across wide luminosity ranges ($-30 \lesssim M_{1450} \lesssim -23$), with $(\alpha, \beta, M_{1450}^*) \sim (-2.0, -3.0, -27)$ (e.g. Willott et al. 2010; McGreer et al. 2013, 2018; Jiang et al. 2016). Previous studies of SDSS quasars at $z > 3.5$ show that $k = -0.47$ and this evolution has been extrapolated to predict higher redshift quasars yield (e.g. Fan et al. 2001b,a; Jiang et al. 2008; Willott et al. 2010). Using deeper data from SDSS Stripe 82 region quasars, McGreer et al. (2013) work shows k redshift evolution between $4 < z < 6$, and proposes $k = -0.7$ at $z = 5 - 6$.

However, the measurement of QLF at $z \geq 5$ remains highly uncertain because of the small sample sizes, especially at fainter magnitudes (e.g. Willott et al. 2010; McGreer et al. 2013, 2018; Venemans et al. 2013, 2015; Kashikawa et al. 2015; Yang et al. 2016; Jiang et al. 2016; Wang et al. 2019). Jiang et al. (2016, hereafter Jiang+16) measures the QLF parameters to be $(\alpha, \beta, M_{1450}^*, k) = (-1.90, -2.8, -25.2, -0.72)$ using a complete sample of 54 quasars at $5.7 < z < 6.4$. Their result of $k = -0.72 \pm 0.11$ suggests that the quasar spatial density drops rapidly after $z = 5$. Recently Matsuoka et al. (2018, hereafter Matsuoka+18) measures the QLF at $5.7 < z < 6.5$ with 110 quasars, including the largest low luminosity quasars in the Subaru High- z Exploration of Low-Luminosity Quasars project (SHELLQs; Matsuoka et al. 2016). They find a clear break at $M_{1450}^* = -24.90$ and a much flatter faint end slope $\alpha = -1.23^{+0.44}_{-0.34}$, compared with Jiang+16 $\alpha = -1.90^{+0.58}_{-0.44}$, although still within 2σ confidence interval. Wang et al. (2019, hereafter Wang+19) use a complete bright quasar samples at $6.4 < z < 6.9$ and find $k = -0.78 \pm 0.18$ at $z = 6.0 - 6.7$, further strengthen the claim of a rapid declination in quasar spatial density from $z = 3.0 - 5.0$ to $z > 6.0$.

We adopt the latest Matsuoka+18 QLF parameters and Wang+19 decline rate in our QLF-sampled mock catalog construction. We fix the parameters as $\alpha = -1.23$, $\beta = -2.73$, $M_{1450}^* = -24.90$, $\Phi^* = 10.9 \times 10^{-9} \text{ Gpc}^{-3} \text{ mag}^{-1}$, $k = -0.78$.

3.2.2. Early Type Galaxies

cosmoDC2 is a synthetic galaxy catalog (Korytov et al. 2019) produced for the LSST second data challenge (DC2) by LSST Dark Energy Science Collaboration (LSST DESC) et al. (2021). The catalog is based on cosmological N-body simulation matching up with empirical and semi-analytic galaxy models. It covers 440 deg^2 of sky area within redshift $0 < z < 3$. The cat-

alog contains detailed galaxy properties including redshift, luminosity, stellar mass, size and shape, as well as host halo information and lensing-related parameters.

Quasar selection based on color-color cuts depends primarily on color dropout in the bluer bands. Red, compact, low- z galaxies (early type or quiescent galaxies; ETGs) detected in optical-infrared bands could be misclassified as quasar candidates. Galaxies show tendency in increasing compactness at higher redshift (e.g. van der Wel et al. (2014)), and will appear unresolved especially at low signal-to-noise (S/N) level. Euclid Collaboration et al. (2019) investigates the degree of contamination by unresolved $z = 1.0 - 2.0$ ETGs for Euclid quasar search. They use the quiescent galaxies from COSMOS2015 catalog in Laigle et al. (2016), which are selected by the rest frame NUV/optical-optical/NIR colour cut, or similarly by the rest-frame UVJ color cut. They simulate the Euclid Near Infrared Spectrometer and Photometer (NISIP) instrument detection and observed photometry within $1''$ aperture on COSMOS ETGs, and conclude that galaxies with $J > 22.0$ may be mistaken as unresolved point sources (see Fig 6 in Euclid Collaboration et al. 2019) due to their small sizes. Both Euclid and WFI have similar pixel resolution of $0''.1$ in NIR bands, we expect the galaxy size information would not contribute significantly in discriminating low S/N quasars and galaxies in our work. We inspect the cosmoDC2 galaxies half-light radii and find that faint galaxies being misidentified as quasar with our selection criteria ($J \sim 26$ with $S/N \sim 15$) indeed have radii $< 0''.1$.

cosmoDC2 catalog simulates both star-forming and red sequence galaxies populations, with rest frame colours matching observational datasets. We extract the $z = 1 - 2$ red sequence galaxies and treat them as primary galaxy contaminants ETG in this work. The cosmoDC2 ETG population density is determined by counting the ETG galaxy number in (J, z) bins within redshift space $z = 1 - 2$ and is shown in Fig 2. Densities of cosmoDC2 ETG and COSMOS ETG match well in the brighter magnitudes, while the lack of COSMOS ETG at $J > 23$ leads to underestimation of the density of ETG (note that $J = 24$ is the Euclid 5σ point source magnitude) at fainter magnitude.

3.2.3. Dwarfs (MLTs)

Low mass stars and brown dwarfs with low surface temperature, including late types M, L and T spectral types, have red infrared colors, and are known as one of the major contaminant populations for high- z quasar selection in infrared bands from past wide surveys (e.g. Mortlock et al. 2012; Wang et al. 2016). The majority

of MLTs lie in the Galactic thin disk with lower scale height, which are considered to be a young (~ 3 Gyrs) and slow moving ($v_{\text{tan}} \sim 30 \text{ km s}^{-1}$) population (Faherty et al. 2009; Ferguson et al. 2017).

The procedure to obtain number density is similar to that described in Euclid Collaboration et al. (2019); we briefly address the main process and refer readers to their paper for more detail. The number density of the thin disk MLT population can be modeled as $\rho = \rho_0 e^{-Z/Z_s}$, where ρ_0 is the number density of MLT spectral type at the Galactic central plane, Z is the vertical distance from the plane, and Z_s is the scale height, assumed to be 300 pc (Gilmore & Reid 1983; Kilic et al. 2017). Given the Roman HLS location in the high galactic latitude region ($\sin b \sim 1$) with a comparatively small survey area, the dependence of number density on Galactic latitude should be negligible for the purpose of quasar selection. We derive the density value from luminosity function and photometry information in Bochanski et al. (2010), Dupuy & Liu (2012) and Skrzypek et al. (2016). We construct the color template based on a dwarfs catalog (Best et al. 2020) compiled from several whole sky surveys observations, the population model is consistent with Euclid Collaboration et al. (2019).

Astrometric information of MLTs are obtained from the UltracoolSheet catalog (Best et al. 2020). They provide ra, dec, parallax and distance measurements for spectroscopic confirmed MLTs. Best et al. (2021) additionally identify a complete MLT sample within 20 pc. We derive the medians of the tangential velocity v_{tan} , total velocity v_{tot} , and velocity dispersion σ_v for each spectral type [Spt-0.5, Spt+0.5] in km s^{-1} , for the complete sample in 20 pc and the entire dataset, respectively. We find that the result are not significantly different. $(v_{\text{tan}}, \sigma_v) = (30, 23)$ and $(v_{\text{tot}}, \sigma_v) = (49, 35)$ are broadly in good agreement with Faherty et al. (2009) and Best et al. (2018). By assuming no evolution of velocity distribution between bright and faint dwarfs, we sample the MLTs total velocity as a Gaussian distribution on top of the Milky Way rotation curve generated by python package galpy² (Bovy 2015). We compute distance modulus for each target and measure their distances and proper motions.

3.3. Mock Catalog Construction

To simulate the realistic survey data and the impact of contaminants on quasar selection, we produce photometric mock catalogs for each population, both

² <https://github.com/jobovy/galpy>

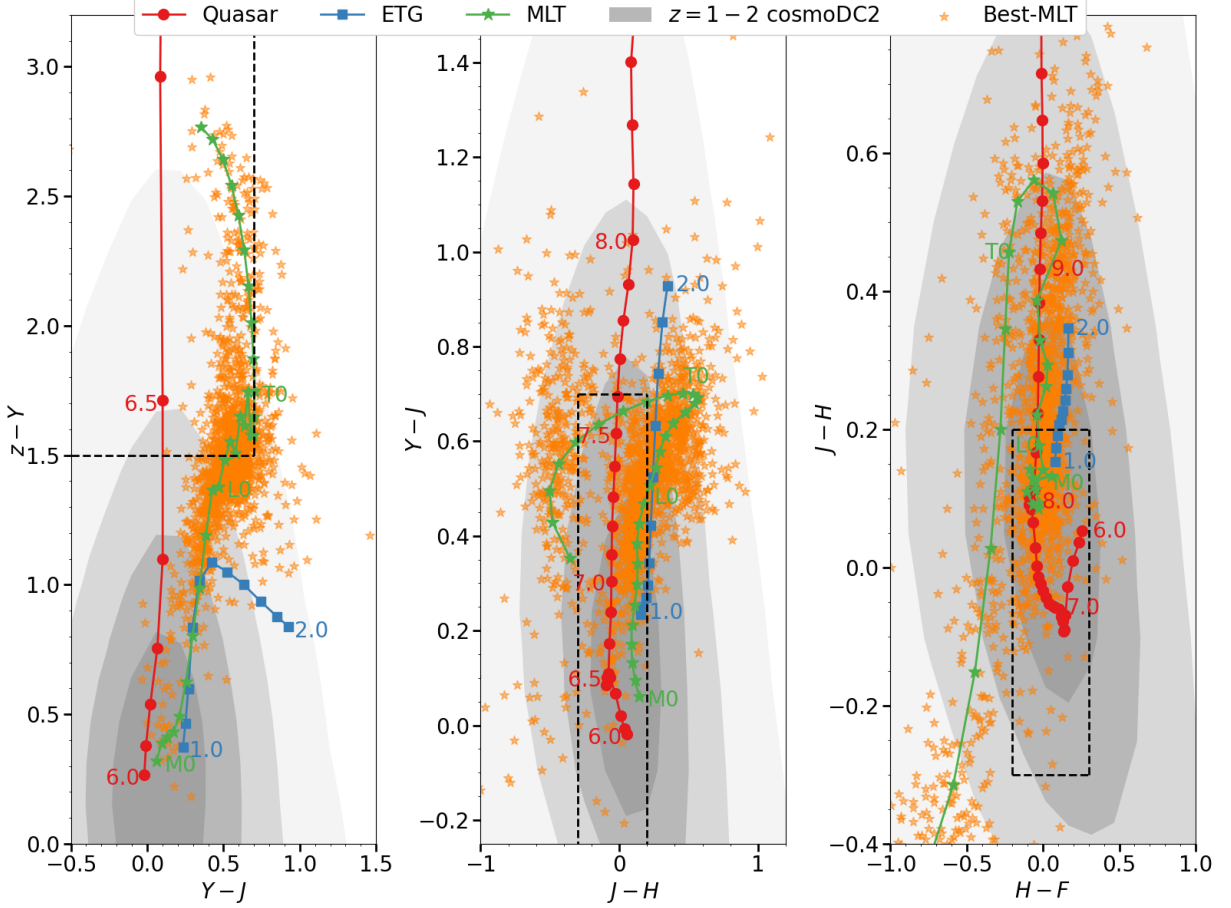


Figure 3. Color-color diagram for the $z = 6.5 - 7.5$ quasar color selection. Detail modelling and full color selections can be found in Sec 3.2 and Sec 4.1. The red tracks are quasar model colors based on the simulation described in Sec 4.1, red circles for a redshift grid with $\Delta z = 0.1$ spacing. The blue squares are median cosmoDC2 ETG colors with $z = 1.0, 2.0$ labelled, grey contours are the entire galaxy population with 4 iso-proportions of densities from 0.2-0.8 (limited to $J = 25.5$ since fainter galaxies are excluded as discussed in Sec 5.1). The green stars are the MLT model colors with M0, L0 and T0 labelled. We also show orange points of spectroscopic confirmed MLTs in Best et al. (2020).

Left: zYJ colors for LSST and WFI. Quasars at $z > 6.5$ have their $\text{Ly}\alpha$ line gradually shifted into longer wavelength infrared bands, leaving them almost undetectable flux in the z band, increasing their $z - Y$ colors with redshifts. The majority of non-quasar population have bluer colors on vertical axis. *Middle:* YJH colors for LSST and WFI. At $z > 8.5$ $\text{Ly}\alpha$ line shifts from Y to J , resulting in redder $Y - J$ colors. From the color tracks it is obvious that $z > 8.0$ quasars suffer less contamination from other populations. *Right:* JHF colors for LSST and WFI. Note that the late T population (T6 and later) generally have bluer JHF colors and are easily excluded using this color plot.

luminosity- and spatial-sampled, reaching 1 mag fainter than the survey limit, to allow possible scatter above the S/N limit.

We generate a QLF-sampled quasar catalog (detail in Sec 3.2.1) with the parameters in Eq (1)-(2). The number of quasars generated are 10 times larger than the number in 2000 deg^2 to minimize the impact of small number statistics. The cosmoDC2 catalog contains rest-frame SED (100 nm to 2000 nm) for individual galaxy. We measure the corresponding synthetic WFI magnitudes by integrating against the filter throughput. Note that in constructing the galaxy contaminants we do not explicitly requires the object to be an early type galaxy.

The entire galaxy catalog is used to include all possible sources. The MLT catalog is generated with their number densities matching the distributions in Fig 2. SpeX Prism Library (SPL) and SpeX Prism Library Analysis Toolkit (SPLAT) (Burgasser & Splat Development Team 2017) provide over 3000 low-resolution, near-infrared MLT spectra with spectral class classification. We construct the composite MLT spectrum by inspecting the suitable MLT spectra, and derive the synthetic magnitudes based on the median colors in each band with respect to J . Observed photometric catalog with multiple realizations are created to minimize possible scatter across experiments. In all, our mock catalog

contains 4×10^3 WFI- J 5σ detected quasars, 9×10^5 MLTs and 4.2×10^9 galaxies (20% ETG).

4. HIGH-REDSHIFT QUASAR SELECTION

4.1. Color Dropout Selection Method

$$\begin{aligned}
 & S/N(u, g, r) < 5.0 \\
 & S/N(J, H, F) > 5.0 \\
 \text{For } 6.5 < z < 7.5: & \begin{cases} S/N(Y) > 5.0 \\ S/N(i) < 3.0 \text{ or } i - Y > 3.5 \\ S/N(z) < 3.0 \text{ or } z - Y > 1.5 \\ Y - J < 0.7 \\ -0.3 < J - H < 0.2 \\ -0.2 < H - F < 0.3 \end{cases} \\
 \text{For } 7.5 < z < 8.5: & \begin{cases} S/N(i) < 3.0 \\ S/N(z) < 3.0 \text{ or } z - Y > 3.5 \\ S/N(Y) < 3.0 \text{ or } 0.6 < Y - J < 2.0 \\ -0.2 < J - H < 0.3 \\ -0.3 < H - F < 0.1 \end{cases} \\
 \text{For } 8.5 < z < 9.0: & \begin{cases} S/N(i, z) < 3.0 \\ S/N(Y) < 3.0 \text{ or } Y - J > 1.5 \\ 0.1 < J - H < 0.6 \\ -0.2 < H - F < 0.2 \end{cases} \quad (3)
 \end{aligned}$$

To select quasar candidates, we first require 5σ S/N non-detection in optical u, g, r and 5σ S/N detection in J, H, F . In literature works (e.g. McGreer et al. 2018) S/N of blue dropouts are generally set to $< 3\sigma$, we initially use 5σ to avoid missing quasar candidates, follow up investigation on real quasar candidates reveals that they also satisfy $S/N(u, g, r) < 3\sigma$ criteria. We do not modify the dropout S/N to conserve the high completeness in selecting quasar candidates. Also note that F band 5σ limit is shallower than those in other bands. As a result, we are actually requiring $S/N \sim 10$ in J band for any source to be selected. We split the color-color cut selection into three redshift ranges: $6.5 < z < 7.5$, $7.5 < z < 8.5$, $8.5 < z < 9.0$. The regions are characterized by optical-NIR drop out in $i - Y$, $z - Y$ and infrared drop out $Y - J$. The regions of $J - H$ and $H - F$ are chosen to optimize the quasar selection and minimize the contaminants cross-over. We summarize the color dropout and color-color cut criteria in Eq 3 and Fig 3.

4.2. Bayesian Model Comparison

Mortlock et al. (2012) propose a Bayesian-based method to discover high-redshift quasars. The main idea is to calculate a posterior quasar probability, P_q , for a target to be quasar rather than others. In this work we consider three closely related population in color space,

namely quasars (q); early type galaxies (g); and MLTs (s). P_q can be expressed as,

$$P_q \equiv P(q|\mathbf{d}) = \frac{W_q(\mathbf{d})}{W_q(\mathbf{d}) + W_g(\mathbf{d}) + W_s(\mathbf{d})} \quad (4)$$

where $\mathbf{d}$ is the photometric data.

To calculate the weights of each population for a given source, we make use of the available photometric fluxes and uncertainties, combined with the surface density model of the population as a prior. For every single source, assume it belongs to a particular population, we compute a Gaussian likelihood function based on model fluxes,

$$p(\mathbf{d}|\theta_i, i) = \prod_{b=1}^{N_b} \frac{1}{\sqrt{2\pi}\hat{\sigma}_b} \exp\left\{-\frac{1}{2} \left[\frac{\hat{f}_b - f_b(\theta_i)}{\hat{\sigma}_b}\right]^2\right\} \quad (5)$$

where b is the photometric band, $\hat{f}_b$ and $\hat{\sigma}_b$ are the observed flux and error, $f_b(\theta_i)$ is the model flux, θ_i is the model parameters govern the population color, $i = q, g, s$ respectively. Note that current formalism only support scatter in brown dwarf colors, as seen in some studies (e.g. Skrzypek et al. 2016). For simplicity we assume no scatter in population model colors, the generalization of including scattering of population model color will be considered in future works. We combine the Gaussian likelihood function with the surface density function Σ_t , integrate over parameter spaces to obtain corresponding weights,

$$W_i(\mathbf{d}) = \int \Sigma_t(\theta_i) p(\mathbf{d}|\theta_i, i) d\theta_i \quad (6)$$

in order to calculate the final posterior quasar probability. In Mortlock et al. (2012) $P_q = 0.1$ was used as the threshold to select quasar candidates. The value is chosen to balance the quasar selection completeness and contamination rate, and proved to be an effective cut in UKIDSS LAS and other high-redshift quasar surveys (Mortlock et al. 2012; Barnett et al. 2021), where changing $P_q = 0.07 - 0.2$ does not greatly impact the overall selection (e.g. Barnett et al. 2021; Nanni et al. 2022). In this work we adopt similar value of $P_q = 0.1$ for quasar candidate selection.

5. RESULTS

Our goal is to investigate the potential of joint color selection and BMC. We first apply color selection, then follow up with Bayesian probability selection to the mock catalogs generated to identify high- z quasars. We discuss the expected yield, completeness and efficiency of color selection method in Sec 5.1. Improvement over Bayesian method is addressed in Sec 5.2.

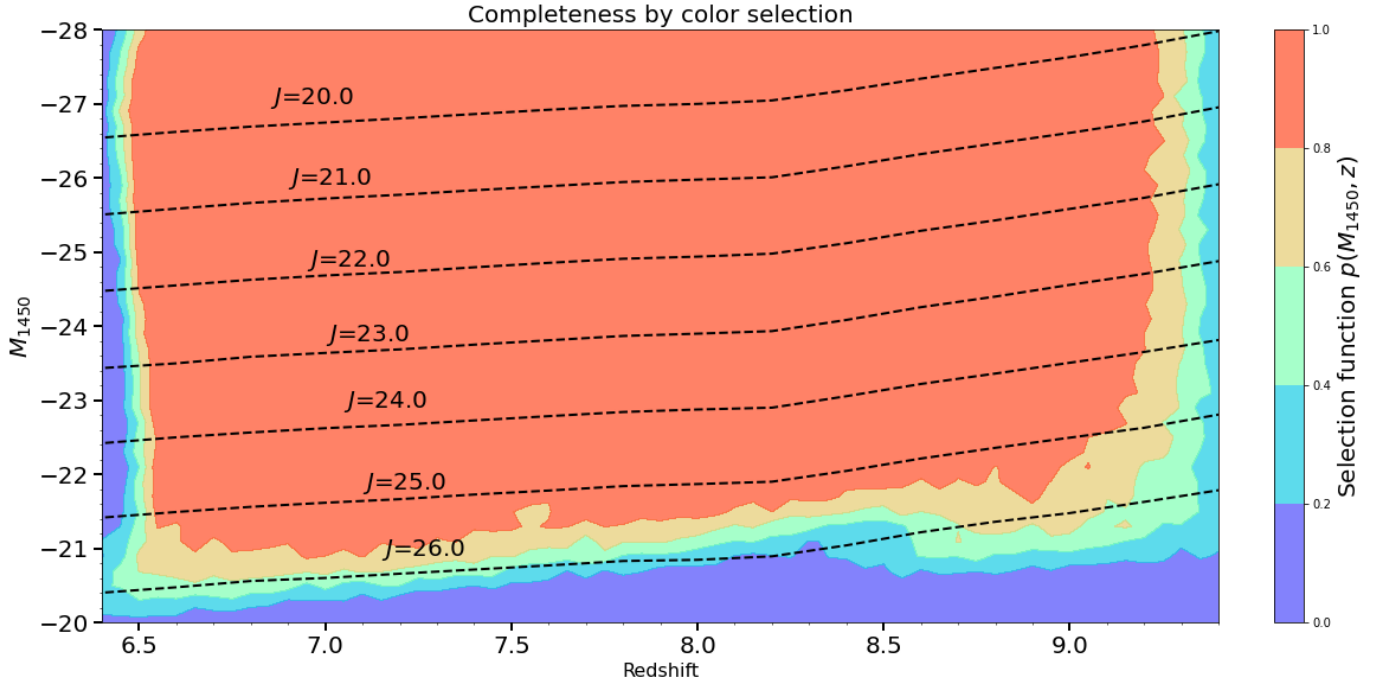


Figure 4. Completeness of color selection function of $6.5 < z < 9.0$ quasars. The probability is the fraction of simulated quasars selected by color selection method among all simulated quasars in the (M_{1450}, z) bin. Black dashed line is the J band magnitude across the redshift range. Color selection is capable to select quasar down to $J \sim 25.5$ at all redshifts. The selection function is mostly complete at $J < 25.5$ ($S/N > 20$).

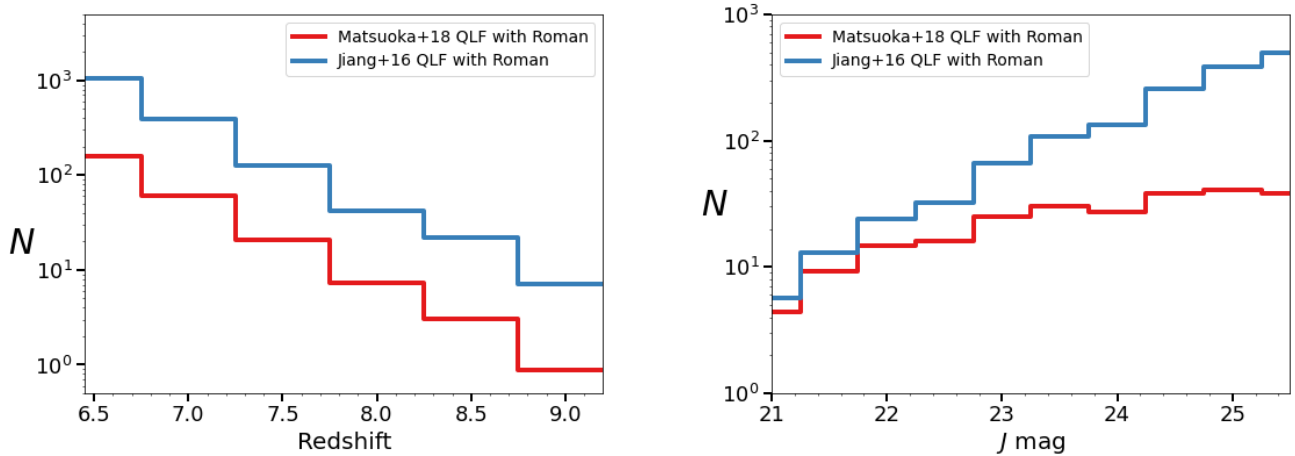


Figure 5. *Left:* Predicted quasar yield in 0.5 redshift bin, *Right:* Predicted quasar yield in 0.5 magnitude bin. Results are derived from color cut selection in a 2000 deg^2 Roman HLS survey. ~ 6 times difference in number yield from two QLF models. Using QLF from Matsuoka et al. (2018) and Wang et al. (2019), we predict of ~ 180 quasars in total at $z > 6.5$, with ~ 8 at $z = 8.0 - 9.0$.

Table 2. Summary of predicted quasar yields of Roman HLS in redshift bins, determined by integrating the QLFs over the color dropout selection functions. Results from Euclid are based on using both Euclid and Rubin optical data, and assuming Jiang+16 QLF. We present the results from both Jiang+16 and Matsuoka+18 QLF, representing the optimistic and conservative estimates.

Redshift range	Roman	Roman	Euclid
	Matsuoka+18	Jiang+16	Jiang+16
$6.5 < z < 7.0$	158	616	-
$7.0 < z < 7.5$	60	240	204
$7.5 < z < 8.0$	20	84	45
$8.0 < z < 8.5$	7	30	16
$8.5 < z < 9.0$	3	11	7

5.1. High-redshift Quasar Survey Predictions For Color Selection Method

We apply our proposed color selection in Sec 4.1 to the uniform grid of simulated quasars, and show the color selection function completeness in (M_{1450}, z) space in Fig 4. Enforce common 5σ detection in WFI infrared bands with different 5σ limiting magnitudes (J : 26.95, F : 26.2) results in lifting the faintest quasar magnitude limit to $J \sim 26$. Color selection successfully recovered $J < 25$ ($M_{1450} > -22$) quasars at redshift range $6.5 < z < 9$, with extended coverage through $J \sim 25.5$ ($S/N \sim 20$).

With color dropout technique, bright quasars at $J < 22$ are completely recovered, while selection function completeness drops rapidly after $J = 25$. The main reason of the drop is primarily the common 5σ S/N ratio requirement in the shallower F band, eliminating lower S/N (< 20 in J) quasars which scatter above the 26.95 detection limit. Besides that, faint quasar selection is also sensitive to $J - H$ and $H - F$ colors. Faint quasars exhibit larger color scatter approaching fainter magnitudes, and subsequently are being missed at > 25.5 . To allow more feasible follow up studies, we limit our discussion to $J \leq 25.5$, or $J_{\text{lim}} = 25.5$, unless specifically noticed.

The predicted quasar yields per redshift bin and magnitude bin are shown in Fig 5, and the quasar yield is summarized in Table 2. In the Roman HLS field we show the predictions for both Jiang+16 ($k = -0.72$) and Matsuoka+18 ($k = -0.78$) QLF in 0.5 redshift bins. Jiang+16 QLF is derived from brighter quasars ($M_{1450} < -24$) sample at $z \approx 6$, and their faint end slope $\alpha = -1.90^{+0.58}_{-0.44}$ is not well constrained; Mat-

suoka+18 QLF includes more faint quasars $M_{1450} \approx -22$ and gives $\alpha = -1.23^{+0.44}_{-0.34}$, however the sample becomes less complete at $M_{1450} > -22.5$, which may underestimate the the real number of quasars at the faintest end. The two QLF functions represent the optimistic and conservative cases for high-redshift quasar yields, and the ~ 6 times differences between these two predictions underlies the need for a deep, wide-field high-redshift quasar survey with high completeness, beyond what the current generation surveys could offer.

We select the candidates based on color selection from three mock catalogs to estimate the performance of color selection on recovering quasar targets and discriminating against contaminants. The results are scaled to meet the Roman survey area yields. The selected candidate list includes 263 quasars, 1.6×10^4 MLTs and 2.9×10^6 galaxies (43% ETG). We define J band 5σ detected quasars as the underlying real quasars with $J_{\text{lim}} = 25.5$. The number of candidates reduce to 215 quasars out of total 230 quasars, 1.2×10^5 galaxies (38% ETG) and 1.0×10^4 MLTs. Our results show that color selections give exceptionally good completeness, on average $\gtrsim 0.85$ at all redshift ranges. The primary color drop out selections are $z - Y$ and $Y - J$, effectively drop sustainable amount of contaminants. At $z = 7.5 - 8.0$, Ly α break lies between Y and J , the completeness significantly depends on the photometric scatter in these two filters, results in overall lower completeness. We plot the cumulative completeness and efficiency as a function of J_{lim} , shown in Fig 6 and summarize the result in Table 3. The overall completeness and efficiency are high at bright magnitudes. Efficiency falls rapidly after $J_{\text{lim}} = 23.5$, with only ~ 0.2 efficiency at $J_{\text{lim}} = 24$, or finding 1 quasar out of 5 possible candidates.

5.2. High-redshift Quasar Survey Predictions For Bayesian Modelling Comparison

BMC is capable of clean selection of bright quasar to $J \sim 23$ at $z = 7 - 9$ (Euclid Collaboration et al. 2019). Color selection gives similar cleanliness results at similar depths. We do not repeat the calculation for bright quasars, and intend to use BMC in improving selection efficiency in the expense of dropping completeness to include fainter quasars. After the color selection, we continue with the Bayesian probability selection. The results are shown in Fig 6 and Table 3.

We adopt a similar Bayesian probability value of $P_q = 0.1$ for quasar as in Mortlock et al. (2012) and Euclid Collaboration et al. (2019). Quasar candidates composes of 211 out of 230 quasars, 2×10^4 galaxies (40% ETG), 2.4×10^3 MLTs at $J_{\text{lim}} = 25.5$. We also show cumulative efficiency and completeness down

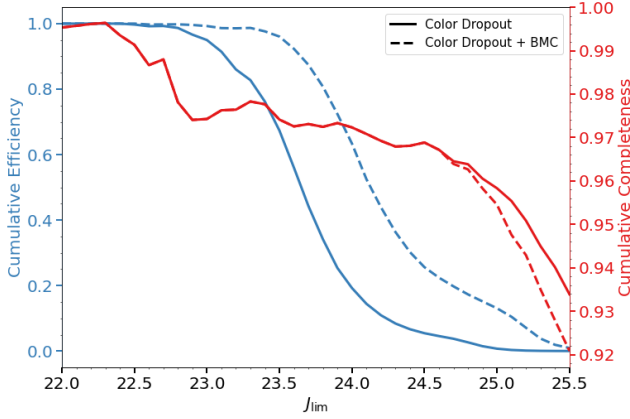


Figure 6. Cumulative efficiency and completeness as a function of J_{lim} . Overall completeness and efficiency are high at bright magnitude. Optimized selections results in efficiency ~ 0.1 at $J_{\text{lim}} \sim 25$, around 180 quasars discoverable. Pushing toward $J_{\text{lim}} > 25.5$, efficiency drop below 0.01, meaning we are expecting one quasar out of several tens to hundreds contaminants.

to the faintest J_{lim} , where selections are still effective. Cumulative completeness is quite high even at the faint end, but the cumulative efficiency escalates rapidly after $J_{\text{lim}} = 24$, reaching 0.25/0.12/0.001 at $J_{\text{lim}} = 24.5/25.0/25.5$, with BMC pushes the identification limit 0.5-1 mag deeper. Cumulative completeness remain high at the faint end, BMC greatly improve cumulative efficiency at $J_{\text{lim}} < 24$, where similar efficiency level is now 0.5-1 mag deeper than in color dropout selection alone. We choose a practical 0.1 selection efficiency for sustainable spectroscopic follow up, and conclude that we are able to identify faint quasars at $J \sim 25$ ($S/N \sim 15$) or $M_{1450} \sim -21$. Quantitatively, we are able to recover ~ 180 quasars at $z > 6.5$, 23 quasars at $z > 7.5$ and 2 quasars at $z > 8.5$.

5.3. Summary on Joint Selection Functions

At brighter magnitude $J < 23$, color drop out and color cut effectively remove majority of the contaminants. Similar completeness and efficiency at bright end has been reported in Euclid Collaboration et al. (2019), as a result we focus on the performance of BMC at faint end. Toward fainter magnitude, Bayesian probability method improves the efficiency by a factor of ten, while retains similar completeness. We examine the performance of BMC by adjusting $P_q = 0.01 - 0.9$, with trade off of worse completeness for exchange of higher efficiency. The results are insensitive to the choice ex-

cept for extreme threshold values. We decide to keep $P_q = 0.1$ as a fair comparison to literature studies.

To investigate the effect of selection functions optimized in both identifying quasars and rejecting contaminants, we display the number of recovered quasars as a function of J_{lim} magnitude, shown in Table 3. At $J \sim 25$, we are expecting to find at least 1 $z > 8.5$ quasar simply by color dropout given HLS 2000 deg² survey area. At $J < 24.5$, galaxies are less of concern since they can be generally eliminated by purely color dropout, MLTs are the major contaminants because of color similarity. Moving toward fainter magnitudes, unresolved galaxies of high surface densities and low S/N photometry begin to dominate contamination. We note that galaxies interfere more at $z = 6.5 - 7.5$ than $z = 7.5 - 8.5$ color dropout selection. Inclusion of BMC effectively remove galaxies from candidates at $J_{\text{lim}} = 25$. MLTs are the main contaminant in this study at brighter magnitude. They are extremely difficult to remove based on our current models, especially for spectral types L2-L4 and T2-T4. These cool dwarfs have low-temperature atmospheres where their red infrared colors are dominated by strong H₂O, CH₄ absorption features that interfere with the $J - H$ and $H - F$, and are indistinguishable from quasar colors. The inclusion of the surface density prior does remove some MLTs from the contamination, but additional parameters are needed to distinguish faint MLTs from quasars, i.e. the proper motion, which will be discussed in the following Sec 6.1.

5.4. Highest Redshift Quasar in the Universe

Quasar spatial density at $z > 6$ has shown to decline more rapidly than at lower redshifts (Richards et al. 2006; McGreer et al. 2013; Jiang et al. 2016; Wang et al. 2019). Wang et al. (2019) predicts ~ 1 luminous quasar available over all sky region at $z > 9$ with $J \sim 21$. We investigate how to maximize the probability of finding the highest redshift quasars, $\sim 5 - 7$ quasars with $J_{\text{lim}} = 25.5 - 26.95$, discoverable by Roman. At $z > 8.4$, the quasar Ly α break redshifts into J . Candidates can be selected through dropping out all bluer bands. We require candidates exhibits non-detection in u, g, r, i, z . We use $Y - J > 1.5$ if detected, or $S/N(Y) < 3$ if undetected, to select potential $z > 8.4$ quasar candidates. We apply the above color selection on the mock populations at $J_{\text{lim}} = 25/25.5/26$. The resulting candidates include 3/4/5 out of 4/5/6 quasars, $25/1.7 \times 10^4/6.7 \times 10^5$ galaxies (64/63/61% ETG), 1/20/680 MLTs, cumulative completeness ≥ 0.75 and cumulative efficiency 0.12/ $2.5 \times 10^{-4}/7.5 \times 10^{-6}$. The efficiency declines rapidly after $J_{\text{lim}} = 25$ or $S/N=30$. We further inves-

Table 3. Summary result of high- z quasar selection functions and quasars recovered from QLF-sampled mock catalogs as described in Sec 3.3 and Sec 4.

J_{lim}	Redshift	Color Dropout		Color Dropout+BMC($P_q \geq 0.1$)		N_q Recovered
		C(%)	E(%)	C(%)	E(%)	
24.0	$6.5 < z < 7.5$	97.2	20.4	96.9	74.8	45
	$7.5 < z < 8.5$	99.3	12.5	99.3	29.6	7
	$8.5 < z < 9.0$	100.0	100.0	100.0	100.0	1
	Overall	97.5	19.0	97.2	63.4	53
24.5	$6.5 < z < 7.5$	96.8	8.7	96.5	48.2	119
	$7.5 < z < 8.5$	99.0	1.5	99.0	6.0	19
	$8.5 < z < 9.0$	100.0	100.0	100.0	100.0	1
	Overall	97.1	5.2	96.9	25.6	139
25.0	$6.5 < z < 7.5$	95.6	1.2	95.2	31.7	151
	$7.5 < z < 8.5$	98.3	0.2	97.1	2.4	23
	$8.5 < z < 9.0$	100.0	95.0	100.0	95.0	2
	Overall	96.0	0.7	95.5	13.1	176
25.5	$6.5 < z < 7.5$	93.4	0.03	92.5	1.7	182
	$7.5 < z < 8.5$	93.8	0.007	88.9	0.2	27
	$8.5 < z < 9.0$	95.7	0.2	95.7	1.2	2
	Overall	93.4	0.02	92.1	0.9	211

tigate if detection and color cut in H, F would improve our selection. We find that using 1. $S/N(H, F) > 5\sigma$ 2. $J - H > 0.1$ and 3. $-0.2 < H - F < 0.1$ result in slightly better efficiency at $J_{\text{lim}} = 25.5$, where there are 4 quasars, 1.6×10^3 galaxies (61% ETG) and 2 MLTs. We summarized the color selection as below,

$$\begin{aligned}
& S/N(u, g, r, i, z) < 5.0 \\
& S/N(J, H, F) > 5.0 \\
\text{For } z > 8.4: & \begin{cases} S/N(Y) < 3.0 \text{ or } Y - J > 1.5 \\ J - H > 0.1 \\ -0.2 < H - F < 0.2 \end{cases} \quad (7)
\end{aligned}$$

Color selection alone enables high completeness selection of highest redshift quasars down to $J_{\text{lim}} = 25$ with 0.1 efficiency. We further proceed with $P_q = 0.1$ selection. At $J_{\text{lim}} = 25.5$, there are 4 out of 5 quasars, 47 galaxies (54% ETG) and 1 MLT, selection efficiency is around 0.1 but 0.4 mag deeper than in Sec 5.2. Number of $z > 8.4$ quasars have doubled, with the highest possible redshift at $z \sim 9.6$ with $J \sim 25$. Finding these earliest and faintest quasars is important because they may host the first generation of SMBHs and are critical for understanding SMBH growth and SMBH-galaxy coevolution in their earliest stages.

6. DISCUSSIONS

6.1. Proper Motion

To further reduce the amount of faint MLTs contaminants, we investigate the possibility to improve the selection by considering proper motion based on the study in [WFIRST Astrometry Working Group et al. \(2019\)](#). We find that Rubin LSST ability to identify the cool brown dwarfs through proper motion is limited for our application, because the majority MLTs that we are concerned are at $J \gtrsim 23.5$ (500 pc). These targets are largely undetected or marginally detected in the LSST bands, with poor astrometry. When LSST completes 5 years survey, Roman will have completed 1 year mapping on the HLS region, by using forced photometry, it may be able to identify moving objects and place constraint on the position offsets. Dwarf thin disk population with $v_{\text{tan}} = 30 \text{ km s}^{-1}$ move with $\mu \sim 10 \text{ mas year}^{-1}$, which is significantly smaller than Rubin LSST limit $\mu_{\text{LSST}} = 35 \text{ mas year}^{-1}$, and $< 5\%$ of original MLT population can be rejected.

The exact cadence of Roman/HLS is not yet finalized. To investigate the possibility of using proper motion obtained from Roman/HLS, we assume a second scan of the same survey area after a time lag of 1 year and 4 years, with 5σ requirement in localization, correspond to $\mu_{\text{WFI}} = 16 \text{ mas year}^{-1}$ and 4 mas year^{-1} , resulting

in 6% and 46% removal of targets. The result is shown in Fig 7. In the most optimistic scenario, we expect 50% MLT population can be rejected from quasar candidates. Follow results from joint selection functions, assuming optimistic $\mu_{\text{WFI}} = 4 \text{ mas year}^{-1}$ threshold, improvement on selection efficiency are 70% at $J_{\text{lim}} = 25$ and 6% at $J_{\text{lim}} = 25.5$. Proper motion measurements improve the selection efficiency at brighter magnitude regime ($J \sim 24$), the improvement is only marginal at the faintest magnitudes, where contamination is dominated by unresolved galaxies rather than MLT dwarfs.

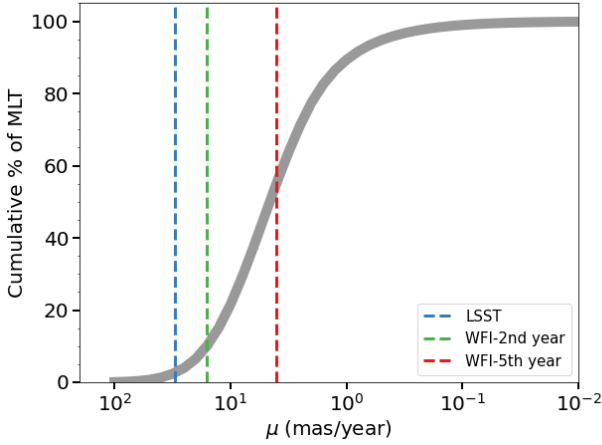


Figure 7. Cumulative percentage of the MLTs population as a function of proper motion. LSST can observe cool brown dwarf with $\mu_{\text{LSST}} = 30 \text{ mas year}^{-1}$. WFI has the ability to measure $\mu_{\text{WFI}} = 4 - 16 \text{ mas year}^{-1}$. The curve represents the percentage of MLTs with proper motion greater than μ , dashed lines represent the proper motion limits of each survey.

6.2. Constraint on QLF Faint End Slope

Bridging the high-redshift faintest quasars/AGNs and brightest galaxies can provide understanding on the co-evolution between AGNs and their host galaxies. Recent discussions have been focusing on transition at $M_{\text{UV}} \sim -22$ (e.g. Kulkarni et al. 2019; Harikane et al. 2022; Adams et al. 2022) where AGNs LF appears to flatten. Matsuoka+18 use 18 $M_{1450} > -23.25$ low luminosity quasars for faint end slope determination. Shen et al. (2020) and Finkelstein & Bagley (2022) compile all known quasars and galaxies with rest-frame UV detection at $z \sim 6$, their models favor a shallow faint end slope for AGNs $M_{\text{UV}} > -23$, although they do not completely rule out the steeper slope due to limited sample sizes. With Roman/Rubin capabilities in finding $J_{\text{lim}} = 25$ quasars at $z > 6.5$, we are going to discover > 80 at $M_{1450} > -23$, 30 at $M_{1450} > -22$. With 4

times increase in low luminosity quasar sample sizes, the faint end slope can be better constrained with 2 times smaller in uncertainties, we are able to distinguish between Jiang+16 and Matsuoka+18, finally determining the underlying faint population density.

7. SUMMARY

In this study we predict quasar yields at $z > 6.5$ in the era of Roman and Rubin. We incorporate the public information from Roman and Rubin, including optical/IR filter wavelengths and limited to Roman 2000 deg^2 HLS survey area. The results are summarized below.

1. With Roman and Rubin imaging capabilities, using joint selection of color and Bayesian Model Comparison, $\sim 180/200$ quasars with $J_{\text{lim}} = 25.0/25.5$ will be discovered at $6.5 < z < 9.0$ within 2000 deg^2 assuming Matsuoka+18 and Wang+19 prescription, expanding at least 3 times the current high- z quasar samples. The estimated cumulative completeness is ≥ 0.8 , while the cumulative efficiency drops below 0.1 after $J_{\text{lim}} = 25$.
2. We expect to find around 4 quasars at $z > 8.4$, with the possible highest redshift quasar at $z > 9.5$ with $J \sim 25$.
3. QLF plays important role in determining the predicted high- z quasar yields. Using the conservative (Matsuoka+18; Wang+19) and optimistic (Jiang+16) QLF parameters result in ~ 6 times difference in the quasar yields. With Roman/Rubin capability in finding a statistical faint quasar sample, we can put stronger constrain on the QLF faint end slope.
4. Color selection alone is sufficient to select $J_{\text{lim}} < 24$ quasars with significant Ly α break. BMC with surface density prior can push the selection threshold 1 mag deeper.
5. Contamination from low- z sources, mainly low S/N unresolved galaxies and cool LT dwarfs, is prominent at $J > 24$. Selection efficiency at brighter magnitudes can be improved through proper motion detection of galactic dwarfs. The improvement is marginal at faintest magnitudes where low- z galaxies dominates contamination.

We emphasize that current deployment time of Roman and Euclid is still uncertain. Predictions in Euclid Collaboration et al. (2019) and this work provide an estimated range for the number of the earliest high- z quasars across wide range of luminosities discoverable in the next decade. This work focuses on exploring the

earliest and faintest quasar population, which are essential part in constraining the high- z supermassive black hole abundance and understanding their growth in the early Universe.

8. ACKNOWLEDGEMENT

WLT appreciate the comments and suggestions from R. Green for useful and informative discussion. We thank the community feedback from WFIRST Science Investigation Team (SIT) supported via NASA contract NNG16PJ33C ‘Studying Cosmic Dawn with WFIRST’, Early Universe/Reionization Era Conversations at Arizona and Roman Science Team Community. WLT and

XF acknowledge supports by NSF grants AST 19-08284. FW thanks the support provided by NASA through the NASA Hubble Fellowship grant #HST-HF2-51448.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS526555. WLT would like to thank the support by Lia YCC throughout the work.

Software: IPython (Pérez & Granger 2007), matplotlib (Hunter 2007), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Astropy (Astropy Collaboration et al. 2013, 2018, 2022), seaborn (Waskom 2021), GCRCatalogs (Mao et al. 2018), SIMQSO (McGreer et al. 2013), Pq_server (Barnett et al. 2021).

REFERENCES

- Adams, N. J., Bowler, R. A. A., Jarvis, M. J., Varadaraj, R. G., & Häußler, B. 2022, arXiv e-prints, arXiv:2207.09342. <https://arxiv.org/abs/2207.09342>
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, apj, 935, 167, doi: [10.3847/1538-4357/ac7c74](https://doi.org/10.3847/1538-4357/ac7c74)
- Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, ApJS, 227, 11, doi: [10.3847/0067-0049/227/1/11](https://doi.org/10.3847/0067-0049/227/1/11)
- Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, Nature, 553, 473, doi: [10.1038/nature25180](https://doi.org/10.1038/nature25180)
- Baldwin, J. A. 1977, ApJ, 214, 679, doi: [10.1086/155294](https://doi.org/10.1086/155294)
- Barnett, R., Warren, S. J., Cross, N. J. G., et al. 2021, MNRAS, 501, 1663, doi: [10.1093/mnras/staa3808](https://doi.org/10.1093/mnras/staa3808)
- Becker, G. D., Bolton, J. S., & Lidz, A. 2015, PASA, 32, e045, doi: [10.1017/pasa.2015.45](https://doi.org/10.1017/pasa.2015.45)
- Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, MNRAS, 370, 289, doi: [10.1111/j.1365-2966.2006.10467.x](https://doi.org/10.1111/j.1365-2966.2006.10467.x)
- Best, W. M. J., Dupuy, T. J., Liu, M. C., Siverd, R. J., & Zhang, Z. 2020, The UltracoolSheet: Photometry, Astrometry, Spectroscopy, and Multiplicity for 3000+ Ultracool Dwarfs and Imaged Exoplanets, 1.0.1, Zenodo, doi: [10.5281/zenodo.4570814](https://doi.org/10.5281/zenodo.4570814)
- Best, W. M. J., Liu, M. C., Magnier, E. A., & Dupuy, T. J. 2021, AJ, 161, 42, doi: [10.3847/1538-3881/abc893](https://doi.org/10.3847/1538-3881/abc893)
- Best, W. M. J., Magnier, E. A., Liu, M. C., et al. 2018, ApJS, 234, 1, doi: [10.3847/1538-4365/aa9982](https://doi.org/10.3847/1538-4365/aa9982)
- Bochanski, J. J., Hawley, S. L., Covey, K. R., et al. 2010, AJ, 139, 2679, doi: [10.1088/0004-6256/139/6/2679](https://doi.org/10.1088/0004-6256/139/6/2679)
- Borison, T. A., & Green, R. F. 1992, ApJS, 80, 109, doi: [10.1086/191661](https://doi.org/10.1086/191661)
- Bovy, J. 2015, ApJS, 216, 29, doi: [10.1088/0067-0049/216/2/29](https://doi.org/10.1088/0067-0049/216/2/29)
- Boyle, B. J., Shanks, T., & Peterson, B. A. 1988, MNRAS, 235, 935, doi: [10.1093/mnras/235.3.935](https://doi.org/10.1093/mnras/235.3.935)
- Bromm, V., & Loeb, A. 2003, ApJ, 596, 34, doi: [10.1086/377529](https://doi.org/10.1086/377529)
- Burgasser, A. J., & Splat Development Team. 2017, in Astronomical Society of India Conference Series, Vol. 14, Astronomical Society of India Conference Series, 7–12. <https://arxiv.org/abs/1707.00062>
- Davies, F. B., Hennawi, J. F., & Eilers, A.-C. 2019, ApJL, 884, L19, doi: [10.3847/2041-8213/ab42e3](https://doi.org/10.3847/2041-8213/ab42e3)
- Davies, F. B., Hennawi, J. F., Bañados, E., et al. 2018, ApJ, 864, 142, doi: [10.3847/1538-4357/aad6dc](https://doi.org/10.3847/1538-4357/aad6dc)
- Dayal, P., Rossi, E. M., Shiralilou, B., et al. 2019, MNRAS, 486, 2336, doi: [10.1093/mnras/stz897](https://doi.org/10.1093/mnras/stz897)
- Dupuy, T. J., & Liu, M. C. 2012, ApJS, 201, 19, doi: [10.1088/0067-0049/201/2/19](https://doi.org/10.1088/0067-0049/201/2/19)
- Euclid Collaboration, Barnett, R., Warren, S. J., et al. 2019, A&A, 631, A85, doi: [10.1051/0004-6361/201936427](https://doi.org/10.1051/0004-6361/201936427)
- Euclid Collaboration, Scaramella, R., Amiaux, J., et al. 2022, A&A, 662, A112, doi: [10.1051/0004-6361/202141938](https://doi.org/10.1051/0004-6361/202141938)
- Faherty, J. K., Burgasser, A. J., Cruz, K. L., et al. 2009, AJ, 137, 1, doi: [10.1088/0004-6256/137/1/1](https://doi.org/10.1088/0004-6256/137/1/1)
- Fan, X. 1999, AJ, 117, 2528, doi: [10.1086/300848](https://doi.org/10.1086/300848)
- Fan, X., Banados, E., & Simcoe, R. A. 2022, arXiv e-prints, arXiv:2212.06907, doi: [10.48550/arXiv.2212.06907](https://doi.org/10.48550/arXiv.2212.06907)
- Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001a, AJ, 122, 2833, doi: [10.1086/324111](https://doi.org/10.1086/324111)
- Fan, X., Strauss, M. A., Schneider, D. P., et al. 2001b, AJ, 121, 54, doi: [10.1086/318033](https://doi.org/10.1086/318033)
- Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117, doi: [10.1086/504836](https://doi.org/10.1086/504836)

- Ferguson, D., Gardner, S., & Yanny, B. 2017, *ApJ*, 843, 141, doi: [10.3847/1538-4357/aa77fd](https://doi.org/10.3847/1538-4357/aa77fd)
- Ferrara, A., Salvadori, S., Yue, B., & Schleicher, D. 2014, *MNRAS*, 443, 2410, doi: [10.1093/mnras/stu1280](https://doi.org/10.1093/mnras/stu1280)
- Finkelstein, S. L., & Bagley, M. B. 2022, arXiv e-prints, arXiv:2207.02233. <https://arxiv.org/abs/2207.02233>
- Gaskell, C. M. 1982, *ApJ*, 263, 79, doi: [10.1086/160481](https://doi.org/10.1086/160481)
- Gilmore, G., & Reid, N. 1983, *MNRAS*, 202, 1025, doi: [10.1093/mnras/202.4.1025](https://doi.org/10.1093/mnras/202.4.1025)
- Glikman, E., Helfand, D. J., & White, R. L. 2006, *ApJ*, 640, 579, doi: [10.1086/500098](https://doi.org/10.1086/500098)
- Green, J., Schechter, P., Baltay, C., et al. 2012, arXiv e-prints, arXiv:1208.4012. <https://arxiv.org/abs/1208.4012>
- Habouzit, M., Volonteri, M., Latif, M., Dubois, Y., & Peirani, S. 2016, *MNRAS*, 463, 529, doi: [10.1093/mnras/stw1924](https://doi.org/10.1093/mnras/stw1924)
- Harikane, Y., Ono, Y., Ouchi, M., et al. 2022, *ApJS*, 259, 20, doi: [10.3847/1538-4365/ac3dfc](https://doi.org/10.3847/1538-4365/ac3dfc)
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357, doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)
- Hunter, J. D. 2007, *Computing in Science & Engineering*, 9, 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- Inayoshi, K., Haiman, Z., & Ostriker, J. P. 2016, *MNRAS*, 459, 3738, doi: [10.1093/mnras/stw836](https://doi.org/10.1093/mnras/stw836)
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, 873, 111, doi: [10.3847/1538-4357/ab042c](https://doi.org/10.3847/1538-4357/ab042c)
- Jiang, L., Fan, X., Annis, J., et al. 2008, *AJ*, 135, 1057, doi: [10.1088/0004-6256/135/3/1057](https://doi.org/10.1088/0004-6256/135/3/1057)
- Jiang, L., McGreer, I. D., Fan, X., et al. 2016, *ApJ*, 833, 222, doi: [10.3847/1538-4357/833/2/222](https://doi.org/10.3847/1538-4357/833/2/222)
- Kashikawa, N., Ishizaki, Y., Willott, C. J., et al. 2015, *ApJ*, 798, 28, doi: [10.1088/0004-637X/798/1/28](https://doi.org/10.1088/0004-637X/798/1/28)
- Kilic, M., Munn, J. A., Harris, H. C., et al. 2017, *ApJ*, 837, 162, doi: [10.3847/1538-4357/aa62a5](https://doi.org/10.3847/1538-4357/aa62a5)
- Komatsu, E., Dunkley, J., Nolte, M. R., et al. 2009, *ApJS*, 180, 330, doi: [10.1088/0067-0049/180/2/330](https://doi.org/10.1088/0067-0049/180/2/330)
- Korytov, D., Hearin, A., Kovacs, E., et al. 2019, *ApJS*, 245, 26, doi: [10.3847/1538-4365/ab510c](https://doi.org/10.3847/1538-4365/ab510c)
- Kulkarni, G., Worsack, G., & Hennawi, J. F. 2019, *MNRAS*, 488, 1035, doi: [10.1093/mnras/stz1493](https://doi.org/10.1093/mnras/stz1493)
- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, 224, 24, doi: [10.3847/0067-0049/224/2/24](https://doi.org/10.3847/0067-0049/224/2/24)
- LSST Dark Energy Science Collaboration (LSST DESC), Abolfathi, B., Alonso, D., et al. 2021, *ApJS*, 253, 31, doi: [10.3847/1538-4365/abd62c](https://doi.org/10.3847/1538-4365/abd62c)
- Madau, P., & Rees, M. J. 2001, *ApJL*, 551, L27, doi: [10.1086/319848](https://doi.org/10.1086/319848)
- Mao, Y.-Y., Kovacs, E., Heitmann, K., et al. 2018, *The Astrophysical Journal Supplement Series*, 234, 36, doi: [10.3847/1538-4365/aaa6c3](https://doi.org/10.3847/1538-4365/aaa6c3)
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2016, *ApJ*, 828, 26, doi: [10.3847/0004-637X/828/1/26](https://doi.org/10.3847/0004-637X/828/1/26)
- Matsuoka, Y., Strauss, M. A., Kashikawa, N., et al. 2018, *ApJ*, 869, 150, doi: [10.3847/1538-4357/aaee7a](https://doi.org/10.3847/1538-4357/aaee7a)
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2019a, *ApJL*, 872, L2, doi: [10.3847/2041-8213/ab0216](https://doi.org/10.3847/2041-8213/ab0216)
- Matsuoka, Y., Iwasawa, K., Onoue, M., et al. 2019b, *ApJ*, 883, 183, doi: [10.3847/1538-4357/ab3c60](https://doi.org/10.3847/1538-4357/ab3c60)
- McGreer, I. D., Fan, X., Jiang, L., & Cai, Z. 2018, *AJ*, 155, 131, doi: [10.3847/1538-3881/aaaab4](https://doi.org/10.3847/1538-3881/aaaab4)
- McGreer, I. D., Jiang, L., Fan, X., et al. 2013, *ApJ*, 768, 105, doi: [10.1088/0004-637X/768/2/105](https://doi.org/10.1088/0004-637X/768/2/105)
- Mortlock, D. J., Patel, M., Warren, S. J., et al. 2012, *MNRAS*, 419, 390, doi: [10.1111/j.1365-2966.2011.19710.x](https://doi.org/10.1111/j.1365-2966.2011.19710.x)
- Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, *Nature*, 474, 616, doi: [10.1038/nature10159](https://doi.org/10.1038/nature10159)
- Nanni, R., Hennawi, J. F., Wang, F., et al. 2022, *MNRAS*, 515, 3224, doi: [10.1093/mnras/stac1944](https://doi.org/10.1093/mnras/stac1944)
- Ohsuga, K., Mori, M., Nakamoto, T., & Mineshige, S. 2005, *ApJ*, 628, 368, doi: [10.1086/430728](https://doi.org/10.1086/430728)
- Pérez, F., & Granger, B. E. 2007, *Computing in Science and Engineering*, 9, 21, doi: [10.1109/MCSE.2007.53](https://doi.org/10.1109/MCSE.2007.53)
- Reed, S. L., McMahon, R. G., Martini, P., et al. 2017, *MNRAS*, 468, 4702, doi: [10.1093/mnras/stx728](https://doi.org/10.1093/mnras/stx728)
- Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, *AJ*, 131, 2766, doi: [10.1086/503559](https://doi.org/10.1086/503559)
- Richards, G. T., Kruczek, N. E., Gallagher, S. C., et al. 2011, *AJ*, 141, 167, doi: [10.1088/0004-6256/141/5/167](https://doi.org/10.1088/0004-6256/141/5/167)
- Ross, N. P., Myers, A. D., Sheldon, E. S., et al. 2012, *ApJS*, 199, 3, doi: [10.1088/0067-0049/199/1/3](https://doi.org/10.1088/0067-0049/199/1/3)
- Schauer, A. T. P., Regan, J., Glover, S. C. O., & Klessen, R. S. 2017, *MNRAS*, 471, 4878, doi: [10.1093/mnras/stx1915](https://doi.org/10.1093/mnras/stx1915)
- Schindler, J.-T., Fan, X., McGreer, I. D., et al. 2017, *ApJ*, 851, 13, doi: [10.3847/1538-4357/aa9929](https://doi.org/10.3847/1538-4357/aa9929)
- Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2020, *MNRAS*, 495, 3252, doi: [10.1093/mnras/staa1381](https://doi.org/10.1093/mnras/staa1381)
- Shen, Y., Wu, J., Jiang, L., et al. 2019, *ApJ*, 873, 35, doi: [10.3847/1538-4357/ab03d9](https://doi.org/10.3847/1538-4357/ab03d9)
- Skrzypek, N., Warren, S. J., & Faherty, J. K. 2016, *A&A*, 589, A49, doi: [10.1051/0004-6361/201527359](https://doi.org/10.1051/0004-6361/201527359)
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757. <https://arxiv.org/abs/1503.03757>
- Stauffer, J., Helou, G., Benjamin, R. A., et al. 2018, arXiv e-prints, arXiv:1806.00554. <https://arxiv.org/abs/1806.00554>

- Tanaka, T., & Haiman, Z. 2009, *ApJ*, 696, 1798, doi: [10.1088/0004-637X/696/2/1798](https://doi.org/10.1088/0004-637X/696/2/1798)
- Trakhtenbrot, B., Volonteri, M., & Natarajan, P. 2017, *ApJL*, 836, L1, doi: [10.3847/2041-8213/836/1/L1](https://doi.org/10.3847/2041-8213/836/1/L1)
- Troxel, M. A., Lin, C., Park, A., et al. 2023, *MNRAS*, doi: [10.1093/mnras/stad664](https://doi.org/10.1093/mnras/stad664)
- Tsuzuki, Y., Kawara, K., Yoshii, Y., et al. 2006, *ApJ*, 650, 57, doi: [10.1086/506376](https://doi.org/10.1086/506376)
- van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, 788, 28, doi: [10.1088/0004-637X/788/1/28](https://doi.org/10.1088/0004-637X/788/1/28)
- Venemans, B. P., Findlay, J. R., Sutherland, W. J., et al. 2013, *ApJ*, 779, 24, doi: [10.1088/0004-637X/779/1/24](https://doi.org/10.1088/0004-637X/779/1/24)
- Venemans, B. P., Bañados, E., Decarli, R., et al. 2015, *ApJL*, 801, L11, doi: [10.1088/2041-8205/801/1/L11](https://doi.org/10.1088/2041-8205/801/1/L11)
- Vestergaard, M., & Wilkes, B. J. 2001, *ApJS*, 134, 1, doi: [10.1086/320357](https://doi.org/10.1086/320357)
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
- Volonteri, M. 2012, *Science*, 337, 544, doi: [10.1126/science.1220843](https://doi.org/10.1126/science.1220843)
- Volonteri, M., & Begelman, M. C. 2010, *MNRAS*, 409, 1022, doi: [10.1111/j.1365-2966.2010.17359.x](https://doi.org/10.1111/j.1365-2966.2010.17359.x)
- Wang, F., Wu, X.-B., Fan, X., et al. 2016, *ApJ*, 819, 24, doi: [10.3847/0004-637X/819/1/24](https://doi.org/10.3847/0004-637X/819/1/24)
- Wang, F., Yang, J., Fan, X., et al. 2018, *ApJL*, 869, L9, doi: [10.3847/2041-8213/aaf1d2](https://doi.org/10.3847/2041-8213/aaf1d2)
- . 2019, *ApJ*, 884, 30, doi: [10.3847/1538-4357/ab2be5](https://doi.org/10.3847/1538-4357/ab2be5)
- Wang, F., Davies, F. B., Yang, J., et al. 2020, *ApJ*, 896, 23, doi: [10.3847/1538-4357/ab8c45](https://doi.org/10.3847/1538-4357/ab8c45)
- Wang, F., Yang, J., Fan, X., et al. 2021, *ApJL*, 907, L1, doi: [10.3847/2041-8213/abd8c6](https://doi.org/10.3847/2041-8213/abd8c6)
- Waskom, M. L. 2021, *Journal of Open Source Software*, 6, 3021, doi: [10.21105/joss.03021](https://doi.org/10.21105/joss.03021)
- WFIRST Astrometry Working Group, Sanderson, R. E., Bellini, A., et al. 2019, *Journal of Astronomical Telescopes, Instruments, and Systems*, 5, 044005, doi: [10.1117/1.JATIS.5.4.044005](https://doi.org/10.1117/1.JATIS.5.4.044005)
- Willott, C. J., Delorme, P., Reyl  , C., et al. 2009, *AJ*, 137, 3541, doi: [10.1088/0004-6256/137/3/3541](https://doi.org/10.1088/0004-6256/137/3/3541)
- . 2010, *AJ*, 139, 906, doi: [10.1088/0004-6256/139/3/906](https://doi.org/10.1088/0004-6256/139/3/906)
- Wu, X.-B., Wang, F., Fan, X., et al. 2015, in *IAU General Assembly*, Vol. 29, 2251223
- Yang, J., Wang, F., Wu, X.-B., et al. 2016, *ApJ*, 829, 33, doi: [10.3847/0004-637X/829/1/33](https://doi.org/10.3847/0004-637X/829/1/33)
- Yang, J., Wang, F., Fan, X., et al. 2019, *AJ*, 157, 236, doi: [10.3847/1538-3881/ab1be1](https://doi.org/10.3847/1538-3881/ab1be1)
- . 2020a, *ApJL*, 897, L14, doi: [10.3847/2041-8213/ab9c26](https://doi.org/10.3847/2041-8213/ab9c26)
- . 2020b, *ApJ*, 904, 26, doi: [10.3847/1538-4357/abbc1b](https://doi.org/10.3847/1538-4357/abbc1b)
- . 2021, *ApJ*, 923, 262, doi: [10.3847/1538-4357/ac2b32](https://doi.org/10.3847/1538-4357/ac2b32)