Semi-analytic forecasts for *Roman* – the beginning of a new era of deep-wide galaxy surveys

L. Y. Aaron Yung,[1][⋆][†] Rachel S. Somerville,[2] Steven L. Finkelstein,[3]
Peter Behroozi,[4,5] Romeel Davé,[6,7,8] Henry C. Ferguson,[9] Jonathan P. Gardner,[1]
Gergö Popping,[10] Sangeeta Malhotra,[1] Casey Papovich,[11,12] James E. Rhoads,[1]
Micaela B. Bagley,[3] Michaela Hirschmann,[13] and Anton M. Koekemoer[9]

[1]*Astrophysics Science Division, NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA*
[2]*Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA*
[3]*Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA*
[4]*Department of Astronomy, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA*
[5]*Division of Science, National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[6]*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK*
[7]*Department of Physics and Astronomy, University of the Western Cape, Cape Town 7535, South Africa*
[8]*South African Astronomical Observatory, Cape Town 7925, South Africa*
[9]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[10]*European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany*
[11]*Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*
[12]*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University,*
College Station, TX 77843, USA
[13]*Institute of Physics, Laboratory of Galaxy Evolution, Ecole Polytechnique Fédeérale de Lausanne (EPFL),*
Observatoire de Sauverny, 1290 Versoix, Switzerland

ABSTRACT

The *Nancy Grace Roman Space Telescope*, NASA's next flagship observatory, will redefine deep-field galaxy survey with a field of view two orders of magnitude larger than *Hubble* and an angular resolution of matching quality. These future *deep-wide* galaxy surveys necessitate new simulations to *forecast* their scientific output and to optimise survey strategies. In this work, we present five realizations of 2-deg^2 lightcones, containing a total of $\gtrsim 25$ million simulated galaxies with $-16 \gtrsim M_{\rm UV} \gtrsim -25$ spanning $z \sim 0$ to 10. This dataset enables a new set of experiments with the impacts of survey size on the derived galaxy formation and cosmological constraints. The intrinsic and observable galaxy properties are predicted using a well-established, physics-based semi-analytic modelling approach. We provide forecasts for number density, cosmic SFR, field-to-field variance, and angular two-point correlation functions, and demonstrate how the future wide-field surveys will be able to improve these measurements relative to current generation surveys. We also present a comparison between these lightcones and others that have been constructed with empirical models. The mock lightcones are designed to facilitate the exploration of multi-instrument synergies and connecting with current generation instruments and legacy surveys. In addition to *Roman*, we also provide photometry for a number of other instruments on upcoming facilities, including Euclid and Rubin, as well as the instruments that are part of many legacy surveys. Full object catalogues and data tables for the results presented in this work are made available through a web-based, interactive portal.

Key words: galaxies: evolution – galaxies: formation – galaxies: high-redshifts – galaxies: star formation – astronomical data base: surveys

1 INTRODUCTION

Over the past three decades, observations with the *Hubble Space Telescope* have revolutionized our understanding

⋆ E-mail: aaron.yung@nasa.gov
† NPP Fellow

of the assembly histories of galaxies in the context of the Universe's overall evolutionary history. The multi-cycle treasury program Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) has established a handful of relatively well-surveyed legacy fields driven mainly by *Hubble* in conjunction with the *Spitzer Space Telescope* and many ground-based telescopes. These surveys reliably reach a 5σ depth of ~ 26.5 (with some variation across different fields). The five CANDELS legacy fields combined cover a total of ~ 850 arcmin^2, reaching galaxies as far as $z \sim 11$ (e.g. Tacchella et al. 2022; Finkelstein et al. 2022b). Within the coverage of the CANDELS fields, the *Hubble* Ultra Deep Field (HUDF; Beckwith et al. 2006; Ellis et al. 2013; Koekemoer et al. 2013) and the eXtreme Deep Field (XDF; Illingworth et al. 2013; Oesch et al. 2013, 2018) have pioneered imaging the extremely deep universe, reaching a 5σ depth of ~ 29.8, at the expense of long exposure times.

The *Nancy Grace Roman Space Telescope*[1], or *Roman* in short, NASA's next premier space-based observatory, is expected to survey the Universe at unprecedented efficiency with its extreme wide-field survey capability (Spergel et al. 2013, 2015). As of today, the *Roman* mission has been confirmed by NASA in March 2020, passed various critical design reviews in 2021, and is on track towards an anticipated launch date before May 2027. *Roman*'s advanced optical system and its onboard Wide Field Instrument (WFI) together offer a field of view of ~ 0.28 arcmin^2, which is approximately a hundred times bigger than that of *Hubble*, and possesses an infrared sensitivity that is comparable to or exceeding that of *Hubble* (Pasquale et al. 2014, 2018). In other words, the size of a single *Roman* WFI pointing is comparable to the total area of the five legacy CANDELS fields combined. In this new era of deep-field galaxy surveys, *every Roman field is a wide field when compared to current generation observations.*

Roman's wide-field capabilities will enable coverage over larger survey areas with many fewer pointings, which effectively permits longer exposure over the targeted fields and increases the survey depths relative to past *Hubble* galaxy surveys with similar survey size and allocated time. Therefore, the *new generation* of deep-field galaxy surveys enabled by *Roman* is expected to reach (or even surpass) the depth of HUDF over areas that are orders of magnitude larger than current generation CANDELS fields (see the *Roman* Ultra Deep Field concept described in Koekemoer et al. 2019). Furthermore, *Roman* will, for the first time, enable high-redshift (e.g. z > 6) surveys spanning multiple square degrees that reach depths comparable to *Hubble* extragalactic legacy surveys (e.g. CANDELS). Current surveys at this scale are driven largely by ground-based instruments, such as the Spitzer/HETDEX Exploratory Large Area survey (SHELA; Papovich et al. 2016; Stevans et al. 2018, 2021; Wold et al. 2019), the Great Optically Luminous Dropout Research Using Subaru HSC survey (GOLDRUSH; Harikane et al. 2018, 2022a; Ono et al. 2018; Toshikawa et al. 2018), and the Lyman Alpha Galaxies in the Epoch of Reionization survey (LAGER; Zheng et al. 2017; Hu et al. 2019; Wold et al. 2022). These large surveys by ground-based telescopes

have identified millions of sources up to $z \sim 7$, providing robust constraints for the number density and clustering statistics of bright, massive objects. However, ground-based instruments are subject to various disadvantages compared to their space-based counterparts and are less ideal for high-redshift explorations. Future *Roman* surveys will provide higher angular resolution images, and reach depths that are not accessible from the ground, particularly in the near-IR. These surveys are expected to deliver robust statistical constraints on the physical properties and spatial distributions of galaxies across cosmic time and on the clustering of galaxies, which have strong implications for the nature of dark matter and the formation of large-scale structure. Furthermore, these large survey areas will be crucial for robustly constraining the number density of rare, luminous galaxies at extreme redshifts (Harikane et al. 2022b).

While *JWST* (Gardner et al. 2006), NASA's latest space-based observatory, remains the most sensitive infrared telescope in operation and will be the powerhouse for ultra-deep surveys (e.g. $m_{AB} > 29$), its relatively narrow field of view is not expected to significantly increase survey area compared to past *Hubble* surveys. Wide-field *JWST* programs, such as Cycle 1 GO/Treasury program COSMOS-Web (Kartaltepe et al. 2021; Casey et al. 2022, ~ 0.60 deg^2), will still be able to cover areas that are comparable to a handful of WFI pointings and are therefore an important pathfinder to inform future Roman deep-field survey strategies. Thanks to the excellent launch delivered by ESA and Arianespace that resulted in significant conservation of *JWST*'s on-board fuel, its expected mission lifespan has been prolonged to upwards of 15 to 20 years. This implies that *JWST* is expected to overlap significantly with NASA's *Roman*, NSF's *Vera C. Rubin Observatory*, and ESA's *Euclid* mission. Therefore, the synergy across *JWST* and these flagship observatories plays a crucial part in increasing the scientific productivity of all facilities. For instance, *JWST* wide-field surveys are important pathfinders for future *deep-wide* surveys that are expected to be conducted by these next generation wide-field survey instruments. *JWST*'s superb sensitivity, mid-IR coverage, and high-resolution spectroscopic capabilities are also suitable for follow-up investigations complementary to the wide-field surveys.

In anticipation of this new generation of deep-wide surveys, predictions or *forecasts* of galaxy population properties over large volumes are essential for the assessment and development of optimal survey strategies, studying synergies between planned observations with different facilities, and ultimately realizing the full scientific potential of these observations. There are three main approaches for modelling galaxy formation: numerical hydrodynamic simulations, semi-analytic models, and empirical models. The first two approaches are similar in that they are based on a set of *a priori* physical processes, such as gas cooling and inflow, star formation and stellar feedback, chemical enrichment, and black hole growth and feedback (Somerville & Davé 2015; Naab & Ostriker 2017). In empirical approaches, either a mapping is created between dark matter halos and galaxy properties using observational constraints, as in subhalo abundance matching models (SHAMs, Wechsler & Tinker 2018), or models are constructed based purely on existing observations (e.g. JADES, Williams et al. 2018). Numerical hydrodynamic simulations provide very detailed predictions,

but it is currently very challenging or impossible to run simulations with volumes comparable to the anticipated next generation of wide-field surveys. Empirical models can very efficiently populate large volumes, but they have limited predictive or interpretive power, and are not able to produce self-consistent predictions for multiple galaxy components. Semi-analytic models are based on physical processes set within a cosmological framework, and self-consistently predict multiple components of galaxies (such as stars, metals, dust, and different phases of gas) (Somerville & Primack 1999; Somerville et al. 2015; Benson 2010; Cowley et al. 2018; Henriques et al. 2015, 2020). At the same time, they are computationally efficient enough to be able to explore parameter space, and create forecasts for relatively large volumes. In the past, this type of theoretical framework has been shown to be a valuable tool that adds scientific value to high-redshift galaxy surveys, including the interpretation of the CANDELS surveys (Somerville et al. 2021) and can facilitate the planning of future *JWST* survey programs, such as CEERS[2], NGDEEP[3], and PRIMER[4] (Yung et al. 2022). These models are also useful for cross-correlating with a large set of anticipated results from intensity mapping surveys, such as EXCLAIM (e.g. Switzer et al. 2021; Yang et al. 2021; Pullen et al. 2022). Useful predictions for *JWST* have also been made by other groups using semi-analytic models (e.g. Dayal et al. 2014, 2015; Cowley et al. 2018) and hydrodynamic cosmological simulations (e.g. Vogelsberger et al. 2020; Kannan et al. 2022a,b; Wilkins et al. 2022a,b).

Lightcones are an effective tool for bridging the gap between simulations and observations. Conventional numerical techniques are carried out by tracking the positions and velocities of mass particles within a simulated cubical (comoving) volume and provide predictions of galaxy properties in snapshots at discrete output times. However, of course when we observe the Universe, we observe galaxies along a past lightcone. Based on snapshots output by numerical simulations, galaxies and halos within a cubical simulated volume can be sampled along such a past lightcone, and mock catalogs constructed in this way are commonly referred to as *lightcones*. Lightcones can be constructed based on hydrodynamic simulations (e.g. Snyder et al. 2017, 2022) or with halos extracted from N-body simulations and filled in with galaxies from empirical models (e.g. Behroozi et al. 2019; Drakos et al. 2022) or semi-analytic models (Overzier et al. 2013; Bernyk et al. 2016; Smith et al. 2017; Barrera et al. 2022). We note that this is a broad overview of lightcone construction and refer the reader to the above referenced works for detailed descriptions.

While lightcones extracted from hydrodynamic simulations contain spatially resolved galaxies that are tracked self-consistently within the simulated environment, the size and mass resolution of these lightcones are limited due to the relatively high computational expense of the underlying hydro simulations. On the other hand, semi-analytic models

and empirical methods built on top of cheaper dark matter only simulations provide a more cost effective alternative, enabling larger area mock fields to be simulated. In anticipation of upcoming multi-$\deg^2$ *deep-wide* surveys, we present physically-based mock catalogues with tens of millions of galaxies constructed based on the halos from the Bolshoi-Planck simulation (Klypin et al. 2016) and galaxies predicted by the physically-motivated Santa Cruz SAM (Somerville et al. 2015).

This work is built based on the well-established Santa Cruz Semi-analytic models (Somerville & Primack 1999; Somerville et al. 2015), which have been extensively compared with observations (Somerville et al. 2015; Yung et al. 2019a,b, 2021) and with the predictions of numerical hydrodynamic simulations (Pandya et al. 2020; Gabrielpillai et al. 2022). In particular, this work builds on the techniques presented in Somerville et al. (2021) and the *Semi-analytic forecasts for JWST* paper series (Yung et al. 2019a,b, 2020a,b, 2021). In these works, the SC SAM modelling framework was shown to reproduce existing constraints on physical and observable properties of high-redshift galaxies $(4 < z < 10)$ and AGN $(2 < z < 7)$, as well as their subsequent impact on the cosmic hydrogen and helium reionization history. The final paper of the series (Yung et al. 2022) presented a large collection of wide-field $(\sim 1000 \text{ arcmin}^2)$ and ultra-deep (rest-frame $M_{UV} \lesssim -12$) lightcones and associated data products. In this complementary paper, we present a set of 2-$\deg^2$ lightcones, with the aim of providing physically accurate predictions for the large-scale distribution and clustering of galaxies. Using these predictions, we present quantitative predictions for the expected uncertainty due to field-to-field variance in both one-point distributions (object counts) and two-point statistics (two-point correlation functions). In addition to *Roman*, *Euclid*, and *Rubin*, the dataset also includes photometric bands presented in past *CANDELS* and *JWST* mock catalogues, and can be used to explore the synergy across *Roman*, *JWST*, *Hubble*, *Spitzer*, and many ground-based observatories. In this work series, we use two NASA flagships, *Roman* and *JWST* as practical examples to demonstrate how these predictions can be used. The physically motivated predictions made with the Santa Cruz semi-analytic model can be easily adapted to make predictions for other space- and ground-based facilities.

All mock catalogues and simulated results presented in this work series are accessible through the project homepage[5] and the Flatiron Institute Data Exploration and Comparison Hub (Flathub[6]).

The key components of this work are summarized as follows: we provide a concise summary of the galaxy formation model and present the simulated lightcones in Sections 2 and 3, respectively. We present the main results in Section 4. We discuss our findings in Section 5, and a summary and conclusions follow in Section 6.

---

2 LIGHTCONE CONSTRUCTION PIPELINE WITH PHYSICAL MODELS

In this section, we provide a concise summary of the semi-analytic model (SAM) for galaxy formation developed over the years by the Santa Cruz group and collaborators (Somerville & Primack 1999; Somerville et al. 2008, 2012, 2015; Popping et al. 2014). We refer the reader to these papers for a full description of the model components. The specific models and configurations for galaxies and AGN are documented in Yung et al. (2019a, 2021) and Somerville et al. (2021). Free parameters in these models are calibrated as described in Yung et al. (2019a) and Somerville et al. (2021). Throughout this work, we adopt cosmological parameters $\Omega_m = 0.308$, $\Omega_\Lambda = 0.692$, $H_0 = 67.8$ km s^{-1}Mpc^{-1}, $\sigma_8 = 0.831$, and $n_s = 0.9665$; which are broadly consistent with the ones reported by the Planck Collaboration in 2015 (Planck Collaboration XIII 2016) and are consistent with the rest of the paper series. All magnitudes presented in this work are expressed in the AB system (Oke & Gunn 1983) and all uses of log are base 10 unless otherwise specified. The calculations in this work are carried out with ASTROPY (Robitaille et al. 2013; Price-Whelan et al. 2018), NUMPY (van der Walt et al. 2011), SCIPY (Virtanen et al. 2020), and pandas (Reback et al. 2022). We provide the code that we used to calculate the co-moving volume of the lightcone slice in Appendix B. This simple calculation is essential to deriving many volume-averaged quantities presented in this work.

2.1 Dark matter cones and merger histories constructions

The set of five realizations of 2-deg^2 lightcones presented in this work are constructed with the same process detailed in Somerville et al. (2021) and Yung et al. (2022). The dark matter halos that form the basis of the 2-deg^2 lightcones are sourced from the Small MultiDark-Planck (SMDPL) simulation from the MultiDark simulation suite (Klypin et al. 2016). This dark matter-only N-body simulation has a volume of $(400 \text{ Mpc } h^{-1})^3$ and dark matter particle mass of $M_{DM} \sim 9.6 \times 10^7 \text{ M}_\odot h^{-1}$. Here h denotes $h_{100} \equiv H_0/100$. Halos in this cosmological simulation are identified using the six-dimensional phase-space halo finder ROCKSTAR and CONSISTENT TREES (Behroozi et al. 2013a,b). We refer the reader to Rodríguez-Puebla et al. (2016) for details on the halo catalogues. As in that work, we adopt the halo virial mass definition from Bryan & Norman (1998). The mass threshold for resolved halos is set to $M_{res} \sim 10^{10} \text{ M}_\odot$, which is equivalent to the mass of ~ 100 dark matter particles.

The dark matter halos in SMDPL are then arranged into mock observed fields spanning 2-deg^2 each between $0 \lesssim z \lesssim 10$, using the lightcone package that is released as part of UNIVERSEMACHINE (Behroozi et al. 2019). For each lightcone realization, the lightcone tool picks a random origin and viewing angle within the base dark matter-only simulation (see Table 1), and includes all halos that fall within the specified survey area. The tool makes use of the periodic boundary conditions when halos lie beyond the boundary of the simulated volume. The distance along the lightcone axis determines the redshift of the simulation snapshot from which halo properties are taken. While the lightcones in this paper were allowed to pass through the same region of the

simulation volume multiple times, since the halos are sampled at a random angle, it is unlikely that a slice of the lightcone will be repeated in the same redshift slice (which happens only if halos are sampled in a slice that is perpendicular to the boundary of the simulation). We refer the reader to Behroozi et al. (2020) for a full description.

For all halos in the lightcone, we use the virial mass of each halo as the 'root mass' and construct a Monte Carlo realization of the merger history using an extended Press-Schechter (EPS)-based method (Lacey & Cole 1993; Somerville & Kolatt 1999; Somerville et al. 2008). These semi-analytically constructed dark matter halo merger trees resolve progenitor halos down to a limiting mass of $\sim 10^{10}$ M$_\odot$ or 1/100th of the root halo mass, whichever is smaller, for all halos. These merger trees are shown to be qualitatively similar to the ones extracted from N-body simulations, and the EPS method enables us to simulate galaxies over a much larger dynamic range than if we had used the merger trees extracted from the N-body simulation. We note that due to the limitations of the EPS algorithm, the halo merger trees do not account for environmental influences, such as assembly bias. We are in the process of developing a merger tree algorithm that account for these effects with machine learning methods (T. Nguyen et al. in preparation) based on N-body cosmological simulations with extremely high mass and temporal resolution (Yung et al. in preparation).

2.2 The semi-analytic galaxy formation model

Using the halos and their merger histories described in the previous section as input, semi-analytic models provide detailed predictions for the star formation histories and a wide variety of other physical properties of galaxies, which can then be forward modelled into observable properties. The Santa Cruz SAM consists of a collection of carefully curated physical processes that are either described analytically or derived from observations and hydrodynamic simulations. These processes include gas cooling and accretion, star formation, stellar feedback, chemical evolution, black hole growth, and AGN feedback. We refer the reader to the schematic flow chart (fig. 1) in Yung et al. (2022) for a comprehensive illustration of the internal workflow of the Santa Cruz SAM.

As in Yung et al. (2022) and the rest of the *Semi-analytic forecasts for JWST* series papers, we adopt the fiducial 'GK–Big2' model, which includes a multi-phase gas partitioning recipe motivated by numerical simulation results from Gnedin & Kravtsov (2011, denoted by GK). The cold gas in the galactic disc is partitioned into a neutral, ionized, and molecular component, and an observationally-motivated H$_2$-based star formation recipe from Bigiel et al. (2008, denoted by Big) is adopted, where the slope of the relation between surface densities of SFR, Σ_{SFR}, and molecular hydrogen, Σ_{H_2}, is unity at lower gas surface densities, and the slope of the relation steepens to $\Sigma_{SFR} \propto (\Sigma_{H_2})^2$ above a critical surface H$_2$ density. Popping et al. (2014) and Somerville et al. (2015) implemented several models for cold gas partitioning and cold gas-based or H$_2$-based SF relations, and showed the impact of these different modelling assumptions on galaxy properties. The predicted cold gas properties are compared with the IllustrisTNG simulations

and ALMA observations in Popping et al. (2019b). Yung et al. (2019a) and Yung et al. (2019b) further experimented with a subset of star-formation relations and found that the fiducial model choices adopted here best reproduce the observed evolution in the galaxy population at $4 \leq z \leq 10$.

Like any physical models that utilize analytic or 'subgrid' prescriptions, the Santa Cruz SAM contains free parameters that must be calibrated to match global galaxy observations (see discussion in Somerville et al. 2015). We refer the reader to Somerville et al. (2008, 2015) for detailed descriptions of the full set of parameters in the model. The Santa Cruz SAMs are typically calibrated 'by hand' to reproduce a set of $z \sim 0$ observational constraints, including stellar-to-halo mass ratio (Rodríguez-Puebla et al. 2017), stellar mass function (Bernardi et al. 2013), $M_{BH}-M_{bulge}$ relation (McConnell & Ma 2013), cold gas metallicity (Andrews & Martini 2013; Zahid et al. 2013; Peeples et al. 2014), stellar metallicity (Gallazzi et al. 2005), and cold gas fraction (Boselli et al. 2014; Peeples et al. 2014; Calette et al. 2018). We do not tune the models to match $z > 0$ observations. In Yung et al. (2019a), we updated the cosmological parameters to be consistent with more recent constraints from Planck, and this necessitated a minor re-calibration of the model parameters to retain agreement with the calibration observations. The set of parameters tuned in this process are supernova (SN) feedback efficiency ϵ_{SN}, SN feedback slope α_{rh}, star formation timescale normalization $\tau_{*,0}$, chemical yield y, and radio mode AGN feedback κ_{AGN}. Yung et al. (2019a) showed that ϵ_{SN} ($\tau_{*,0}$) has a significant effect on the faint (bright) galaxy populations at $z > 4$. The impact of varying α_{rh} and $\tau_{*,0}$ on the predicted galaxy populations at $4 < z < 10$ was explored in Yung et al. (2019a,b). On the other hand, it was shown that AGN feedback has no noticeable effect on the galaxy populations across the full range of mass and luminosity at $z > 4$.

The performance of this model configuration has been tested against observations at low redshifts ($0 \lesssim z \lesssim 6$; Somerville et al. 2015, 2021) and at high redshifts ($4 \lesssim z \lesssim 10$; Yung et al. 2019a,b). It has been shown that these physical models can reproduce the observed distribution functions for stellar mass and star formation rate up to $z \sim 8$, as well as rest-frame UV luminosity functions up to $z \sim 10$. Yung et al. (2019b) has also shown that the predicted stellar-to-halo mass ratio and other scaling relations are in good agreement with other empirical models and hydrodynamic simulations.

We note that the same satellite position re-assignment detailed in Yung et al. (2022) is also applied to the lightcones presented in this work, where the satellite positions are assigned assuming an NFW profile (Navarro, Frenk & White 1997). This improves the agreement of the distributions of radial distances between satellite galaxies and their central galaxies with N-body simulations, and is important to produce the 'one-halo' term in the predicted two-point correlation functions. This is illustrated and further discussed in Section 4.3 and is compared to abundance matching model results from UniverseMachine in Appendix C.

2.3 Observables for *Roman* and other facilities

Based on the predicted star formation and chemical enrichment histories (SFHs, stored mass in bins of stellar age and metallicity), galaxies are assigned spectral energy distributions (SEDs) generated based on the stellar population synthesis (SPS) model of Bruzual & Charlot (2003). The rest-frame SEDs are used to calculate rest-frame luminosities in filter bands as presented in the mock catalogue. In addition, quantities labelled with dust are calculated accounting for the effect of dust in the ISM. We assume the dust attenuation curve of Calzetti et al. (2000). The V-band dust attenuation is calculated based on the surface density of cold gas and metallicity, based on a 'slab' model as described in Somerville et al. (2012), but adopting the latest recalibration of ISM dust optical depth presented in Yung et al. (2021), where the dust optical depth, $\tau_{dust,0}(z)$, decreases slightly at $z \gtrsim 4$ relative to the previous calibration from Yung et al. (2019a) following the implementation of an updated black hole growth model that yields slightly stronger AGN feedback. This update, consistent with Yung et al. (2022), improves the agreement between UV LF predictions and observations at $z > 6$ compared to CANDELS DR1. We have implemented a stellar mass threshold of $\log(M_{*,thres}/M_\odot) = 7$, below which we do not generate an SED, as these galaxies would likely not be observable in the wide-field surveys of interest here. This threshold is set with the mass resolution of the underlying dark matter-only simulation and the detection limit of *Roman* taken into account. We note that in these mock catalogues, we do not include the contribution from nebular line or continuum emission, but plan to do so in future work (Yung, Hirschmann, Somerville et al. in prep).

The rest-frame SEDs are then redshifted according to their redshift in the lightcone, and observed-frame magnitudes are computed, accounting for attenuation effects from the intervening IGM (Madau et al. 1996). In additional to the larger simulated volume of these new lightcones, a new aspect of the lightcone catalogues described in this work is the additional photometry from *Roman* WFI[7], *Euclid*[8] visible imager (VIS, Euclid Collaboration et al. 2022b) and Near Infrared Spectrometer and Photometer (NISP-P, Euclid Collaboration et al. 2022a), and *Rubin Observatory*[9] (Ivezić et al. 2019).

Photometry from the large collection of filters from *Hubble*, *Spitzer*, and other ground- and space-based instruments as presented in the CANDELS lightcones (Somerville et al. 2021) and the NIRCam broad- and medium-band photometry presented in the *JWST* lightcones (Yung et al. 2022) are also included in these new lightcones. Studies utilising the predicted photometry in these bands should also reference these works. In addition, in this work, we also added photometry for instruments utilized in the SHELA survey, including DECam, NEWFIRM K-band, and VISTA. Having a large collection of predictions for photometric bands from existing instruments has been shown to be useful for characterizing foreground contaminants in wide-field surveys

---

[7] https://roman.gsfc.nasa.gov/science/Roman_Reference_Information.html (filter transmission dated Jun 14th, 2021, accessed on Nov 1st, 2021)
[8] https://euclid.esac.esa.int/msp/refdata/nisp/NISP-PHOTO-PASSBANDS-V1, access to filters granted via private communication
[9] https://github.com/lsst/throughputs/tree/master/baseline, v1.7, accessed on May 28th, 2021
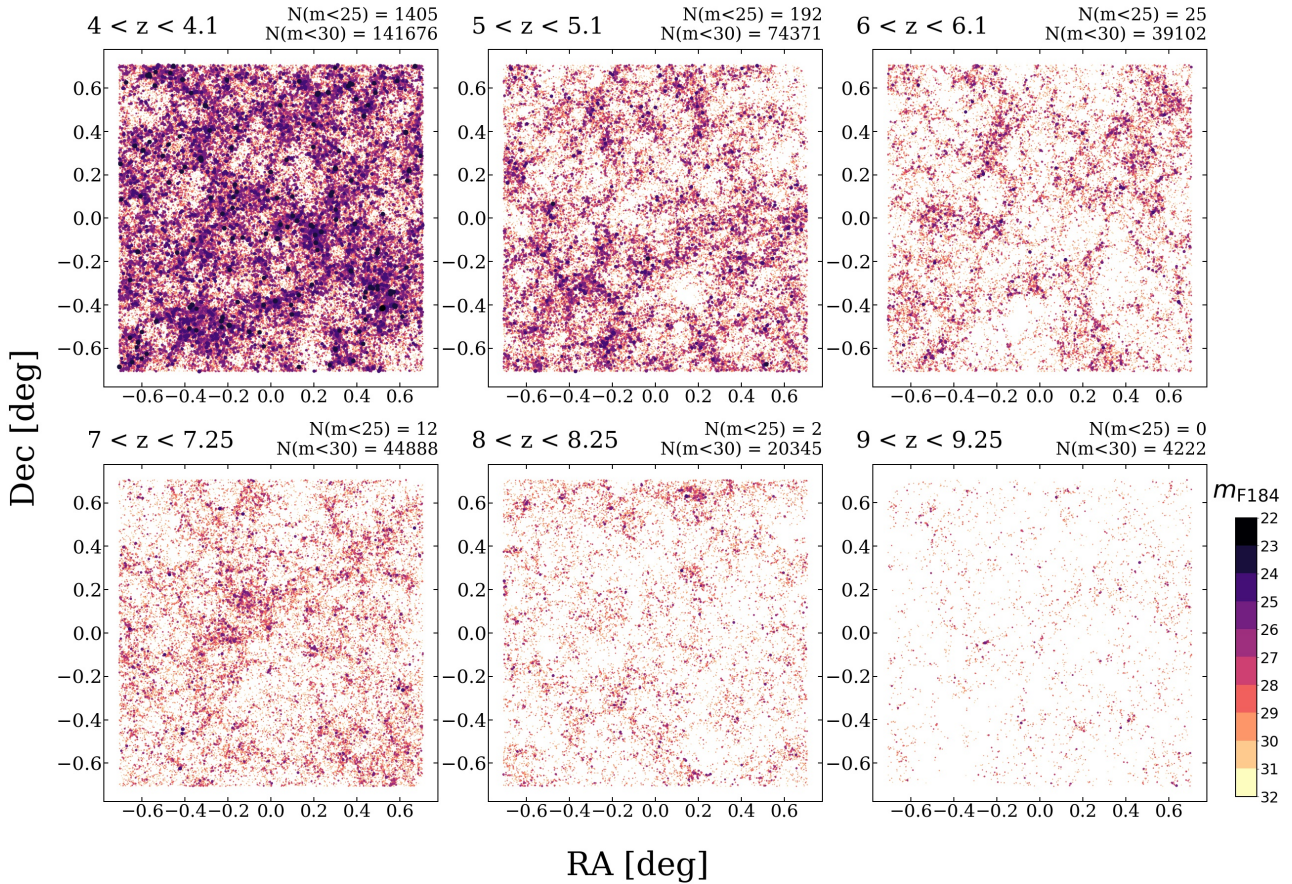
Figure 1. A summary of the footprint and galaxy populations in the first realization of a 2-deg^2 lightcone at various redshift slices between $z \sim 4$ to 10. The data points are colour-coded by the observed-frame IR magnitude in the *Roman* WFI F184 band. The sizes of the data points are also scaled to emphasize brighter objects and do not reflect their predicted angular sizes. In addition, the number of bright and faint objects ($m_{F184} < 25$ and $m_{F184} < 30$, respectively) within each slice is indicated at the top right corner of each panel. The specifications of this lightcone are summarized in Table 1.

Table 1. This table summarizes the dimension, area, and key configurations for the lightcones release with this work.

Specification	2-deg^2 lightcones
Dimension (arcmin)	84.85×84.85
Area (arcmin^2)	7200
Base simulation	MultiDark - SMDPL
$\log M_{h,res}/ M_\odot$	10.00
$\log M_{*,lim}/ M_\odot$	7.00
M_{UV} range	-16 to -25
redshift range	$0 < z \lesssim 10$

and in the search for extreme redshift galaxies (Leung et al. 2022, Bagley et al. in preparation).

Rest-frame luminosities in the mock catalogues are indicated with '`_rest`'. Luminosities without such labels are in the observed frame. Similarly, the luminosities of the bulge component alone are labelled with '`_bulge`'. See Table A1 in Appendix A for a complete list of all physical properties and photometric bands available in the mock catalogues.

3 SIMULATED DATA PRODUCTS

Based on the physical models that have been extensively tested and, in previous works, shown to reproduce existing observations up to $z \sim 10$, we present predictions for five independently sampled 2-deg^2 fields that spans $0 < z \lesssim 10$, providing a comprehensive compilation of photometric and physical properties of galaxies (see Table A1 for full list of available quantities). In addition, full high-resolution spectra and star formation histories are also available. These data products are useful for a wide variety of post-processing applications, such as implementing an alternative SPS model or computing photometry for additional filters.

Each of the 2-deg^2 lightcones contains ~ 12 million galaxies with $\log(M_*/M_\odot) \geq 7$, among which ~ 5 million are in the rest-frame luminosity range $-16 \lesssim M_{UV} \lesssim -25$ or observed-frame magnitude $31 \lesssim m_{F184} \lesssim 22$ (both at $z \sim 4$). We note that all predicted observed- and rest-frame magnitudes presented in this work include dust attenuation, unless specified otherwise. The key specifications of these lightcones are summarized in Table 1. These 2-deg^2 lightcone have the same mass resolution as the wide-field lightcones presented in Yung et al. (2022). However, the approximately seven times larger footprint increases the chance of includ-
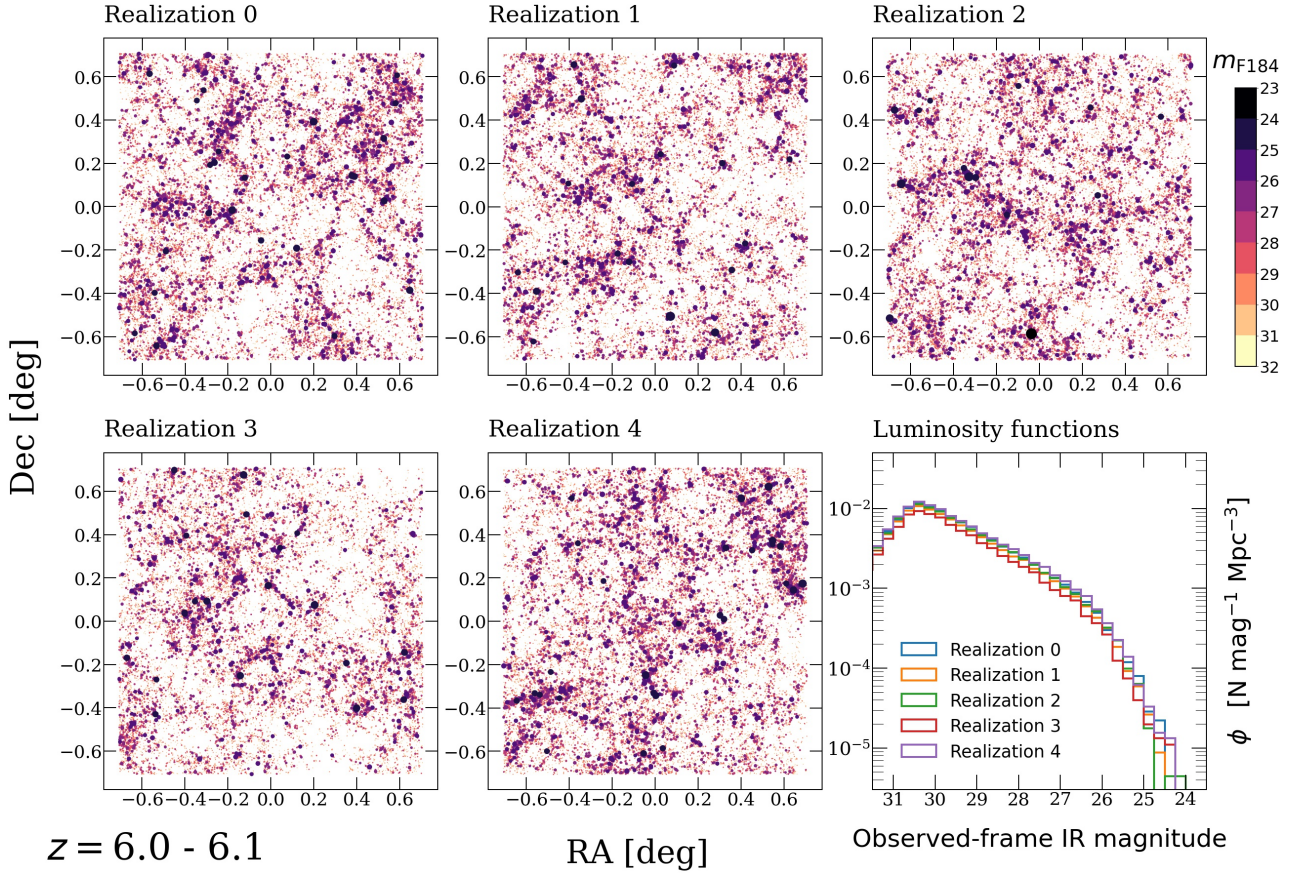
Figure 2. A side-by-side comparison of thin slices between $6.0 \leq z \leq 6.1$ from each of the five realizations of the 2-deg^2 lightcones. The predicted galaxies are both size- and colour-coded by their rest-frame UV luminosity, where darker coloured, larger data points represent brighter galaxies. The size of the data point do not reflect their predicted angular sizes. The lower-right panel shows the predicted observed-frame IR luminosity functions for $6.0 \leq z \leq 6.1$ across the five realizations.

ing galaxies forming in more massive halos, and therefore results in better sampling at the bright end.

Fig. 1 shows the predicted star-forming galaxies from the first realization of the simulated lightcones in several redshift slices: $4.0 < z < 4.1$, $5.0 < z < 5.1$, $6.0 < z < 6.1$, $7 < z < 7.25$, $8 < z < 8.25$, and $9 < z < 9.25$. The data points are colour-coded by their rest-frame UV luminosities. While the sizes of the data points in this figure are scaled with their observed-frame IR luminosities in the WFI F184 filter, m_{F184}, to emphasize the bright objects, we note that the sizes of these data points do not reflect the galaxies' predicted angular sizes. In addition, we also show the counts of galaxies in the slice within each of these panels. This figure gives an intuitive, qualitative view of the possible evolution of object number density and large-scale structure along the light of sight in a 2-deg^2 field.

Fig. 2 shows a side-by-side, qualitative comparison of thin slices of the lightcone between $6.0 \leq z \leq 6.1$ from all five realizations. Similar to the previous figure, the data points are colour- and size-coded by the observed-frame m_{F184} of the simulated galaxies. We note that the colour- and size-code for this figure is chosen to be slightly different from the previous figure to better highlight the galaxy populations and large-scale structure within this redshift range. In the last panel of Fig. 2, we compare the distribution functions of m_{F184} across the five realizations. We show that the dif-

ference in the object counts across these five realizations is very small. The flattening of the luminosity functions shows where the galaxy population is becoming unresolved due to the limited mass resolution of the underlying halo populations. We deliberately show a very thin slice of the lightcone to highlight the differences across the multiple lightcone realizations. We note that the redshift slice shown here is much narrower than the typical redshift range used in studies with photometric redshifts, typically $\Delta z \sim 1$ (Bouwens et al. 2015; Larson et al. 2022), due to the relatively large uncertainties in the photometric redshift estimates. We also note that the variance across the field realizations is larger on the bright end than in the faint end as expected (luminous galaxies are more strongly clustered at all redshifts). We will further explore the impacts of survey size and limiting magnitude on field-to-field variance in Section 4.2.

These wide-field lightcones have been shown to provide more robust number statistics and simulated volume that are necessary for preparatory studies for large galaxy surveys (e.g. Finkelstein et al. 2022b; Kakos et al. 2022; Chworowsky et al., in preparation; Hellinger et al., in preparation) and for intensity mapping (e.g. Yang et al. 2021).

Given the large number of galaxies included, these lightcones are delivered in slices by redshift. The redshift slices have $\Delta z = 0.1$ between $0 \leq z \leq 1$; $\Delta z = 0.25$ between $1 \leq z \leq 4$; $\Delta z = 0.50$ between $4 \leq z \leq 6$
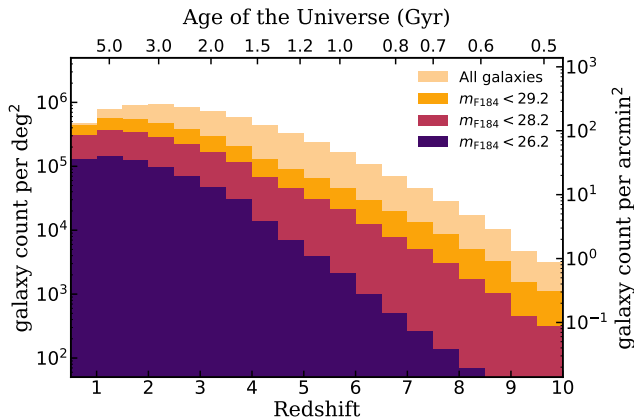
Figure 3. This histogram shows the (non-cumulative) number of galaxies expected normalized to unit deg^2 *(left axis)* and unit arcmin^2 *(right axis)* between $0.5 < z < 10$ in bins of $\Delta z = 0.5$ for galaxies in all five realizations of 2-deg^2 lightcones, spanning a total mock survey area of 10 deg^2. On the top axis, we show the age of the Universe corresponding to the redshift indicated on the bottom axis. We show the number of objects above several survey limits expected for the high-latitude survey, moderate depth galaxy survey, and ultra-deep survey. See Table 2 and text for details about the survey configurations.

Table 2. This table summarizes the 5σ detection limits, exposure time, and survey areas for anticipated *Roman* surveys. The *Roman* detection limits assume the use of the WFI F184 filter. See text for full description.

survey type	$m_{\mathrm{lim}, 5\sigma}$	exposure	survey area
high-latitude survey	26.2	146 sec	1700 deg^2
moderate depth	28.2	5 hr	2.5 deg^2
ultra-deep survey	29.2	20 hr	0.5 deg^2

and $\Delta z = 1$ between $6 \leq z \leq 10$. The data products presented in this work are accessible through the interactive portal Flathub (https://flathub.flatironinstitute.org/group/sam-forecasts), which allows more flexible access and download capability for galaxy catalogues across multiple redshift slices. Users may inspect and filter the data (e.g. by magnitude and/or by physical properties) and selectively download a subset of the data as needed. Alternatively, the mock catalogues in ASCII format can be download in full at https://users.flatironinstitute.org/~rsomerville/Data_Release/SAM_lightcones/.

4 RESULTS FROM WIDE-FIELD LIGHTCONES

In this section, we present a set of quantitative, key predictions at high redshift that are derived from the set of 2-deg^2 lightcones. We show the evolution of object counts (per survey area) as a function of redshift and field-to-field variance estimated for a range of survey areas. These results are selected specifically to demonstrate the advantages of the large area coverage of these simulated lightcones.

4.1 Evolution of galaxy demographics across redshift

In this subsection, we show predictions for the redshift evolution of volume-averaged galaxy counts and the cosmic SFR density. Using galaxies in all five realizations of the 2-deg^2 lightcones and the observed-frame IR magnitude in the WFI F184 filter, m_{F184}, as an example, in Fig. 3 we give an overview of the number of galaxies expected per deg^2 and per arcmin^2 surveyed as a function of redshift. This histogram is made with galaxies in a combined simulated area spanning 10 deg^2, which contains a total of ~ 67 million galaxies (centrals and satellites), among which ~ 34 million, ~ 21 million, and ~ 7 million galaxies have $m_{\mathrm{F184}} < 29.2$, 28.2, and 26.2, respectively. In addition, we show the corresponding age of the universe on the top horizontal axis.

This figure can be used as a look-up table to quickly estimate the number of objects expected in a large survey. Note that this figure is based on a simulated area that is ~ 35 times bigger than one of the wide-field *JWST* lightcones considered in Yung et al. (2022), and therefore is statistically more robust and less susceptible to field-to-field variation. In addition, we consider three discrete magnitude limits that are representative of anticipated future *Roman* surveys. These include an extremely wide but relatively shallow high-latitude survey[10] that is expected to cover ~ 1700 deg^2. We also include scenarios of a moderate depth survey comparable to approximately ten WFI fields, and an ultra-deep survey that spans approximately two WFI fields. The magnitude limits for both scenarios are estimated assuming a total ~ 500 hours of imaging for each type of hypothetical survey, distributed equally across all available WFI filters. The estimated survey depths for given exposure times are calculated based on the latest available Anticipated Performance Tables[11] and are summarized in Table 2. In addition to showing counts per deg^2, we also show the count per arcmin^2 for quick reference. For instance, we show that a moderate-depth survey reaching $m_{\mathrm{F184}} \sim 28$ would yield significantly more sources at $z > 8$ than the shallower high-latitude survey per unit survey area.

The number counts of objects shown in this figure take into account several effects that impact the galaxy populations at low redshift in a degenerate manner, including the increasing dust attenuation, which preferentially affects massive galaxies. The WFI F184 filter is ideal for detecting rest-frame UV radiation for galaxies at $4 < z < 10$, but the wavelength range of the same filter begins to probe rest-frame optical and longer wavelengths for galaxies at $z \lesssim 4$. Furthermore, the number of objects *per unit area* decreases as expected as the physical volume enclosed decreases. Together, these factors result in the decrease and flatting of the total number of objects expected in the F184 band at $z \lesssim 2.5$. We note that the number of objects per co-moving volume does not decrease as shown in the top panel of Fig. 4.

We note that this plot provides an idealized estimate of counts of galaxies down to the stated limit. We refer the reader to Bagley et al. (in preparation) for a more detailed

---

[10] https://roman.gsfc.nasa.gov/high_latitude_wide_area_survey.html
[11] https://roman.gsfc.nasa.gov/science/apttables2021/table-exposuretimes.html
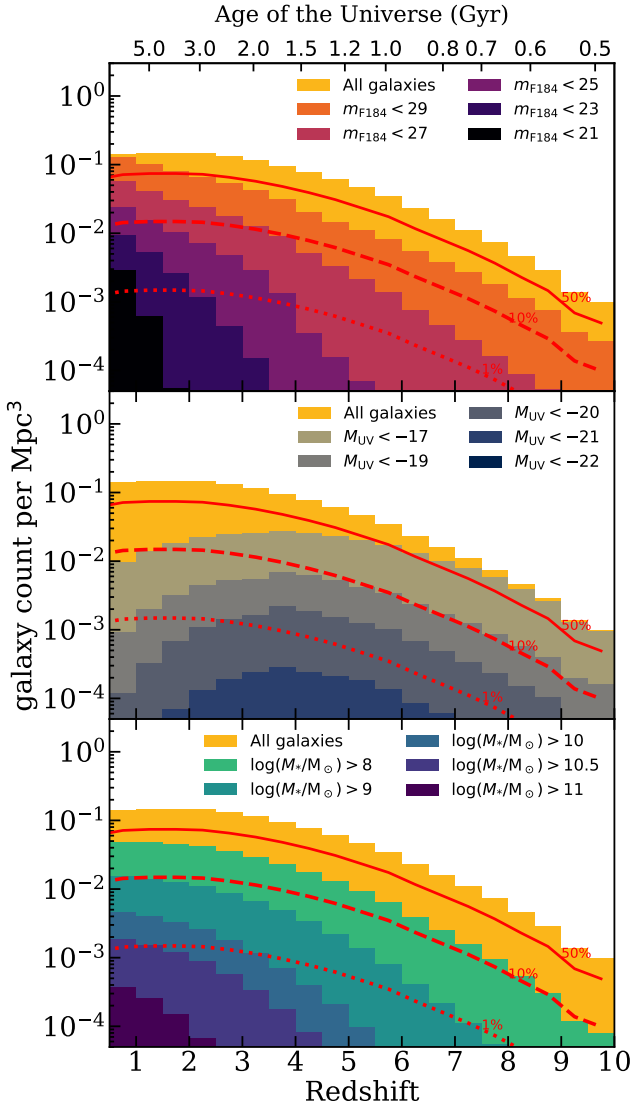
Figure 4. These histograms show the predicted number of galaxies normalized to unit cubic co-moving-Mpc between $0.5 < z < 10$ in bins of $\Delta z = 0.5$. The number density of all predicted galaxies is coloured in yellow. The galaxy populations are then broken down by observed-frame IR magnitude in the *Roman* WFI F184 band (*top*), rest-frame UV magnitude (*middle*), and stellar mass (*bottom*). In all three panels, we mark the levels of 50%, 10%, and 1% of the total galaxy number counts with red solid, dashed, and dotted lines, respectively.

Figure 5. The cosmic SFR density as a function of redshift (bottom axis) or age of the Universe (top axis) between $0.5 < z < 10$ in bins of $\Delta z = 0.5$. The SFRD of all predicted galaxies is coloured in yellow. The galaxy populations are then broken down by observed-frame IR magnitude in the *Roman* WFI F184 band (*top*), rest-frame UV magnitude (*middle*), and stellar mass (*bottom*). In all three panels, we show lines corresponding to 50%, 10%, and 1% of the total cosmic SFRD with red solid, dashed, and dotted line types, respectively.

look into galaxy colour selection based on the mock *Roman* photometry, including Lyman-break selection.

    In a similar spirit, we also provide predictions for the number density of galaxies as a function of redshift. Fig. 4 shows histograms of the predicted counts of galaxies per co-moving Mpc³ as a function of redshift. In the panels of this figure, the galaxy population is broken down by observed-frame IR magnitude in F184, rest-frame dust-attenuated UV luminosity, and stellar mass. This provides an overview of the evolution in the galaxy population across cosmic time broken down by these key properties. However, we note that the vertical axis is shown in log scale and therefore the area
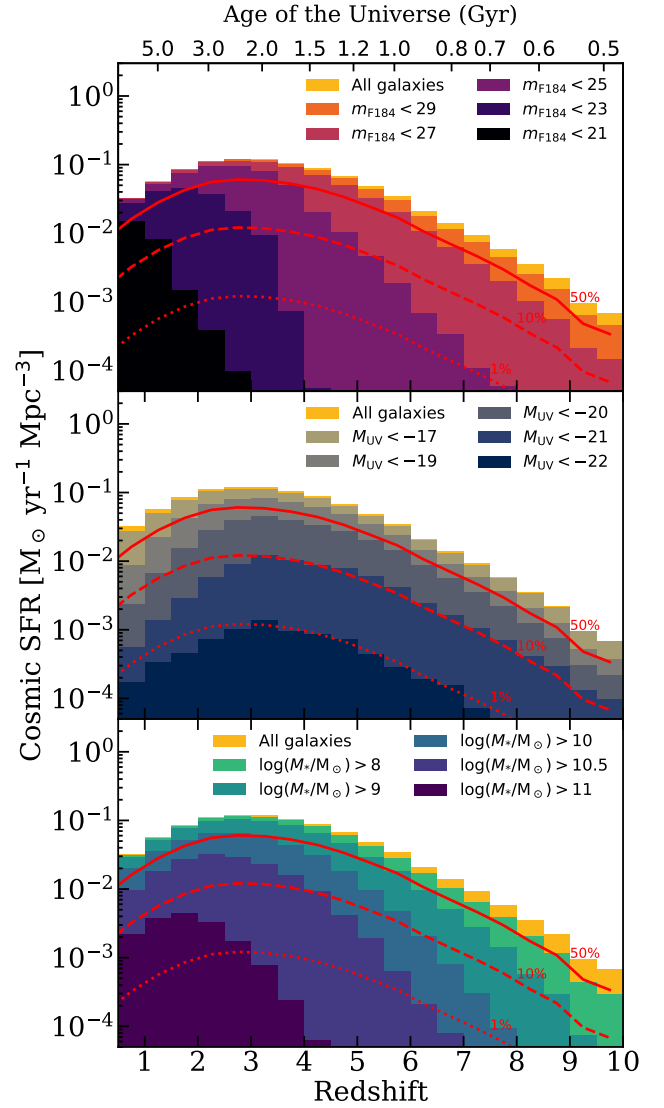
carved out by the histogram does not directly correspond to the proportion of their contribution. We therefore mark the 50%, 10%, and 1% of the total number of galaxies available in the lightcone in all three panels.

    We note that the full set of galaxy samples shown here include all galaxies available in the predicted lightcones, which contains halo populations that are complete down to $\log M_h/M_\odot = 10$. However, the corresponding $m_{\rm F184}$, $M_{\rm UV}$, and M_* cut-off due to the halo mass resolution limit evolves with redshift. We refer the reader to figs. 6 and 7 in Yung et al. (2022) for where the flattening occurs in the one-point distribution functions for the observed-frame $m_{\rm F184}$

and rest-frame M_{UV}. Given that these lightcones are configured with identical halo mass and stellar mass limit as the wide-field lightcones presented in that work, where the flattening occurs is expected to be the same as those presented in Yung et al. (2022). We also refer the reader to fig. 17 in Yung et al. (2019b) for the stellar-to-halo mass relation from $z = 4$ to 10 predicted with the same model used in this work. Those results can be used to estimate the stellar mass for a given halo mass limit at a given redshift.

The top panel of this figure effectively shows the evolution of cumulative number densities of galaxies as a function of redshift at various m_{F184} detection limits. This is equivalent to showing the redshift-evolution of the cumulative counts (number density) of galaxies with $n(< m_{lim})$ (also see fig. 14 in Yung et al. (2019a)). The middle and bottom panels also work the same way, which illustrate the redshift-evolution of the cumulative number density of galaxies by rest-frame UV luminosity function and stellar mass function. This set of figure panels offers a new perspective on the predicted evolution of apparent magnitude functions, rest-frame UV luminosity functions, and stellar mass functions that are difficult to achieve with conventional one-point distribution functions. For example, a quick comparison between the M_{UV} *(middle)* and M_* *(bottom)* panel shows that dust plays an important role in suppressing the rest-frame luminosity for massive galaxies, as the number of galaxies across all mass ranges increases steadily at $z \lesssim 4$ towards lower redshift, but the number of UV-bright galaxies declines.

Similarly, Fig. 5 show the redshift evolution of cosmic SFR density (SFRD) as a function of redshift, which is a commonly used diagnostic quantity that measures the collective evolution of the integrated SFR for a galaxy population across time (see Madau & Dickinson 2014 for review and discussion therein). The cosmic SFRD predicted by the Santa Cruz SAM has been shown to be in good agreement with a variety of observations integrated down to $M_{UV} \sim -17$ (Bouwens et al. 2015; Finkelstein et al. 2015a; McLeod et al. 2016; Oesch et al. 2018) and the empirical model of Behroozi et al. (2013c) (see fig. 9 in Yung et al. (2019a)). We also show a comparison of the predicted cosmic SFR with other empirical models and a subset of observational constraints in Section 4.4. In this figure, we break down the contribution to the cosmic SFR budget from galaxies by the m_{F184}, M_{UV}, and M_* of predicted galaxies with thresholds matching the ones in Fig. 4. We also mark the 50%, 10%, and 1% of the total cosmic SFR budget to guide the eye.

A side-by-side comparison of Figs. 4 and 5 also helps visually correlate the redshift evolution of galaxy number densities and their contributions to the cosmic SFR budget. For instance, the bottom panels from both figures show that galaxies with $\log(M_*/M_\odot) < 8$ made up over 50% of the galaxy populations but contribute to very little to the overall cosmic SFR. Similarly, the middle panels show that galaxies with $M_{UV} < -20$ made up approximately 3% of the population by number between $4 \lesssim z \lesssim 10$, and are responsible for $\sim 50\%$ for the cosmic SFR budget.

4.2 Field-to-field variance and survey area

In this section, we leverage the combined 10 deg^2 simulated fields to conduct controlled experiments quantifying the dependence of field-to-field variance on survey area. A similar
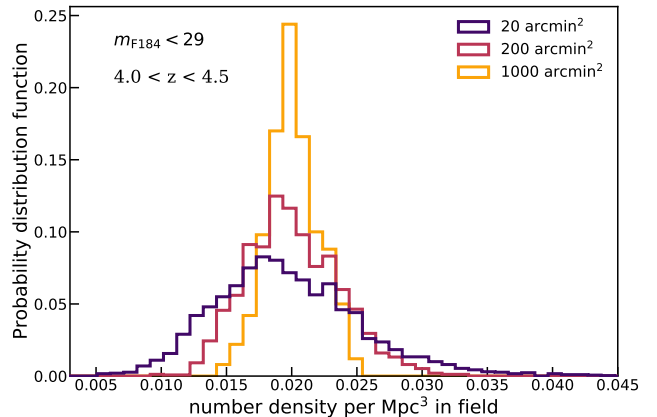

Figure 6. Probability distribution function for the expected number density per Mpc^3 for galaxies with $m_{F184} < 29$ between $4.0 < z < 4.5$ in subsampled fields of three different sizes: 20, 200, and 1000 $arcmin^2$.

experiment was carried out in (Yung et al. 2022) utilizing 40 simulated lightcones, each ~ 1000 $arcmin^2$. The larger contiguous areas covered by the 2-deg^2 lightcones presented in this work enable more flexible sampling of sub-fields with different field sizes and geometry.

The current generation of high-redshift surveys (with *HST* and *JWST*) span areas of the order of hundreds of $arcmin^2$; field-to-field variation has therefore been a major source of uncertainty in the estimates of object counts. To illustrate how future generation wide-field instruments will be able to improve this situation, in this exercise, we consider three field sizes across multiple orders of magnitude, chosen to represent iconic deep surveys of the past and give a flavor for future surveys. We consider a case of 20 $arcmin^2$, which is comparable to the legacy HUDF or two *JWST* NIRCam pointings, 200 $arcmin^2$, which is comparable to the size of a legacy CANDELS field, and 1000 $arcmin^2$, which is comparable to a single pointing of *Roman* WFI. Fig. 6 shows the probability distribution for the number density per Mpc^3 for galaxies with $m_{F184} < 29$ between $4.0 < z < 4.5$ when we repeatedly subsample the 2-deg^2 lightcones for sub-fields of the aforementioned sizes. This example demonstrates how the field-to-field variance is affected by the survey area. We note that the distribution is skewed to the left, with a long tail towards higher number densities. This is expected in the standard theory of structure growth via gravitational collapse. Thus it is important to keep in mind that although the field-to-field *variance* is often quoted, the probability distribution for over-densities can be significantly non-Gaussian.

Finkelstein et al. (2022b) have utilized a similar subsampling method to investigate a possible over-density in the observed EGS field compared to the cosmic average at $z \gtrsim 9$. A companion work will explore a similar effect in rest-frame luminosity functions (Bagley et al., in preparation).

Somerville et al. (2004) and Moster et al. (2011) provided an analytic prescription to calculate the field-to-field variance for CANDELS-sized deep pencil beam surveys for $z \sim 0.5$ to 4, based on an empirically established sub-halo abundance matching (SHAM) model for stellar mass selected galaxies. In this work, we make use of the very wide
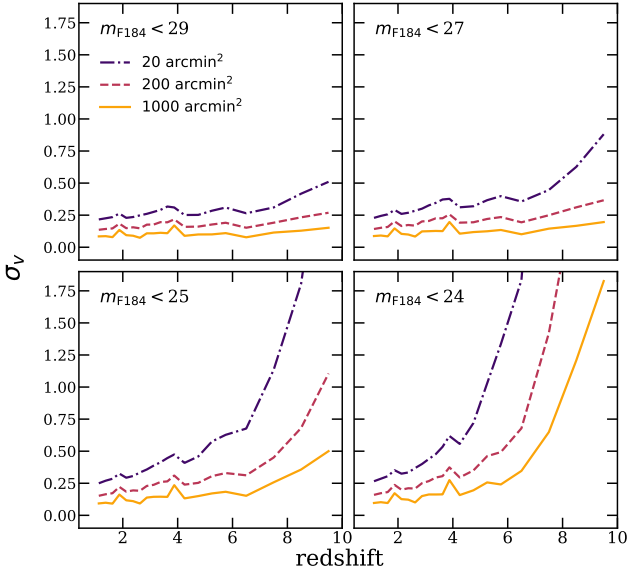
Figure 7. Root field-to-field variance as a function of redshift between $1 < z < 10$ for galaxies with $m_{\mathrm{F184}} < 29$, 27, 25, and 24 in survey sizes that represent past and future deep galaxy surveys. We consider survey areas of 20 arcmin^2, 100 arcmin^2, and 1000 arcmin^2 represented by dashed, dot-dashed, and dotted lines, respectively. See text for details regarding sampling.

Figure 8. Root field-to-field variance as a function of redshift between $1 < z < 10$ for galaxies with $\log(M_*/\mathrm{M}_\odot) > 7$, 8, 9, and 10 in survey sizes that represent past and future deep galaxy surveys. We consider survey areas of 20 arcmin^2, 100 arcmin^2, and 1000 arcmin^2 represented by dashed, dot-dashed, and dotted lines, respectively. See text for details regarding sampling.

lightcones to empirically calculate the field-to-field variance for magnitude- and mass-selected galaxies at high redshift in survey fields with different sizes. Following the steps detailed in Somerville et al. (2004) and the modification in Yung et al. (2022), the relative cosmic variance (with shot noise removed) can be expressed as

$$\sigma_v^2 = \frac{\langle n^2 \rangle - \langle n \rangle^2}{\langle n \rangle^2} - \frac{1}{V\langle n \rangle}, \tag{1}$$

where $\langle n \rangle$ and $\langle n^2 \rangle$ denote the mean and variance of object number density n, respectively, and are the first and second moments of the probability distribution function $P_N(V)$ with the volume, V, factored out and cancelled. In this calculation, V is fixed at an assigned value such that $V\langle n \rangle > 1$ even in subregions with extremely low number density (e.g. $n < 1$ per unit volume when the number count in the field $N > 0$).

In this exercise, we calculate the field-to-field variance empirically with predicted galaxies over the range of $1 \lesssim z \lesssim 10$ by repeatedly subsampling the lightcones with square regions and calculating the cosmic variance based on the number of galaxies (both central and satellite) found within the subregions. We consider three distinct survey sizes: 20 arcmin^2 (approximately the size of HUDF), 200 arcmin^2 (on the order of current generation surveys), and 1000 arcmin^2 (approximately the size of wide-field lightcones from Yung et al. (2022)). For 20 arcmin^2, we sample a total of a total of 3000 regions, 1000 boxes each from the first three realizations; for 200 arcmin^2, we sampled a total of 2500 regions, 500 boxes each from five realizations; and for 1000 arcmin^2, we sampled a total of 500 regions, 100 boxes each from five realizations. These calculations share the same redshift binning as the lightcone dataset (see Sec-
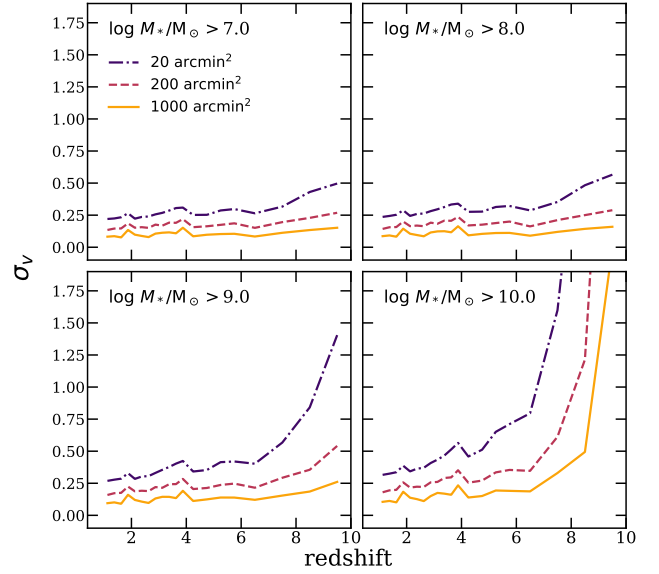
tion 3). The total volume covered by the one thousand 20 arcmin^2 fields is about 2.8 times larger than the total volume of the three 2-deg^2 lightcones from which they are drawn, and for the 200 arcmin^2 and 1000 arcmin^2 fields, the ratio is 13.9. As a result, our calculation will not fully sample the density probability distribution of the larger equivalent volume. However, we have confirmed that this has a negligible effect on the estimates of the variance of the distribution, and mainly affects the tails.

We show the calculated field-to-field variance as a function of redshift for observed magnitude-selected and stellar-mass-selected galaxies in Figs. 7 and 8, respectively. In Fig. 7, we show the evolution of the expected cosmic variance as a function of redshift for galaxies above certain observed-frame m_{F184} thresholds at $m_{\mathrm{F184,lim}} = 29$, 27, 25, and 24. It is well known that variance in HUDF-like and even CANDELS-like fields can be significant both at high redshift and for rare luminous objects. We show that the variance decreases significantly, down to only a few per cent, with a single WFI pointing. Similarly, in Fig. 8, we show the evolution of cosmic variance for galaxies above several stellar mass thresholds at $\log(M_{*,\mathrm{lim}}/\mathrm{M}_\odot) = 7.0$, 8.0, 9.0, and 10.0. We note that the two instances where σ_v slightly drops at $z \sim 4$ and ~ 6 are caused by the increase of redshift bin size (e.g. $\Delta z = 0.25$ to 0.50 at $z \sim 4$ and $\Delta z = 0.50$ to 1.00 at $z \sim 6$). While we are considering the number density of objects found throughout the volume, doubling the redshift bins has effectively increased the volume sampled and decreases the variance among the object counts. While this can be handled with finer redshift bins, we keep the current binning as it is, closer to the redshift range from colour selection at high redshift.

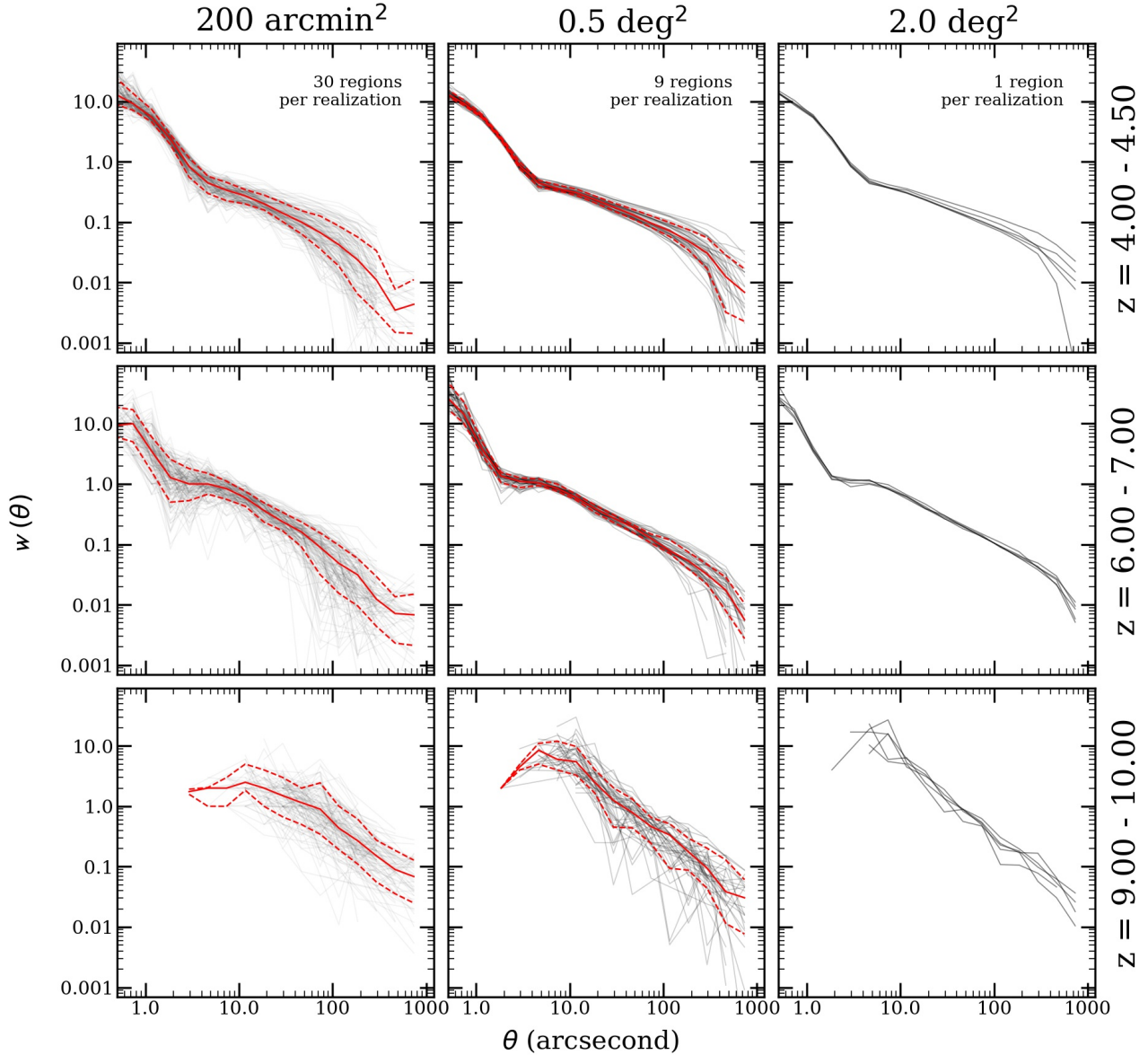These calculations indicate that the variance of past

Figure 9. Predicted angular 2PCF, $w(\theta)$, calculated for galaxies with $m_{F184} < 28$, which is chosen to represent the depth expected to be reached by *Roman* moderate depth surveys. Galaxies are subsampled from 30, 9, and 1 subregions per lightcone realization in three distinct survey areas of 200 arcmin^2 (*left column*), 0.5 deg^2 (*middle column*), and 2.0 deg^2 (*right column*), respectively. See text for details regarding the subsampling. The full set of 2PCF are shown individually in grey. This calculation is repeated for three redshift bins: $4.0 < z < 4.5$ (*top row*), $6.0 < z < 7.0$ (*middle row*), $9.0 < z < 10.0$ (*bottom row*). For the 200 arcmin^2 and the 0.5 deg^2 cases, the red solid line shows the median of the sampled correlation functions, and the dashed lines mark the 16th and 84th percentiles to characterize the spread among the total of 150 and 45 individually calculated 2PCFs, respectively.

CANDELS surveys ($m_{F160W} < 26.5$, a few hundred arcmin^2) and deep *JWST* surveys ($m_{F200W} < 29$, a few hundred arcmin^2) have a potential variance of $\sigma_v \sim 1$ and ~ 0.25 toward $z \sim 10$, which is likely the onset of cosmic reionization (see Yung et al. 2020b and discussion therein). The variance depends strongly on the luminosity or mass of the objects selected, and therefore it decreases with increasing survey depth as fainter, less clustered objects are probed. In addition, the variance decreases with increasing survey area as the clustering of all objects on these large scales is smaller. Additionally, the variance tends to increase towards higher redshift because objects of a given mass are more strongly clustered (more biased) at high redshift. While with the current set of lightcones, we are only able to explore such calculations up to ~ 1000 arcmin^2, it is clear that future *Roman* surveys, with moderate depth surveys reaching $m_{lim} \sim 28.2$ to span multiple square degree and the high-latitude survey reaching $m_{lim} \sim 26.2$ spanning thousands of deg^2, will further suppress the variance to negligible values, as shown by the solid orange line in Figs. 7 and 8.
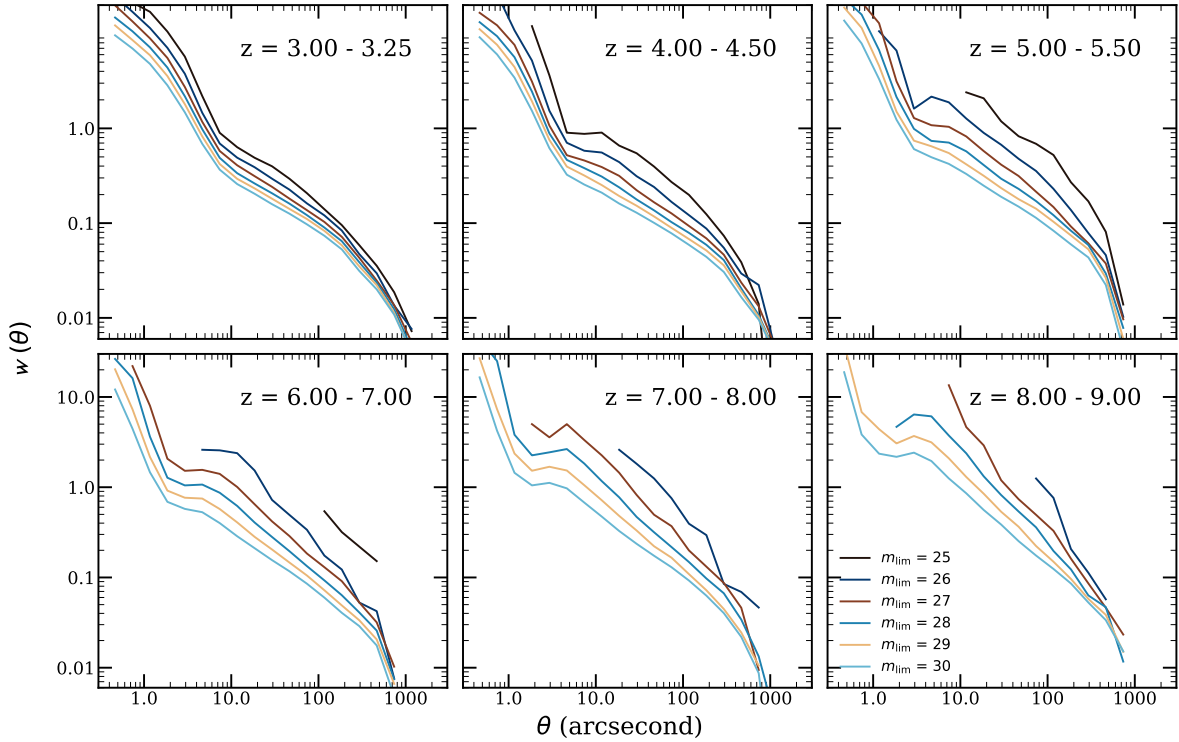
Figure 10. Angular correlation functions, $w(\theta)$, for bins across a wide range of redshifts computed for galaxy samples above several observed-IR magnitude limits. The ACFs are computed for all galaxies in each of the 2-deg^2 lightcones, and in this figure we show the median of the five realizations.

4.3 Galaxy clustering

The two-point correlation function (2PCFs) measures the excess of galaxy pair counts relative to a randomly distributed sample (Peebles 1980). The angular auto-correlation function, $w(\theta)$, is a specific type of 2PCF that accounts for the projected angular separation on the sky θ between pairs of objects. Yung et al. (2022) showed that $w(\theta)$ predicted from ~ 1000 arcmin^2 lightcones constructed with the same physical model used here are in excellent agreement with observations across $1.25 \lesssim z \lesssim 7.5$ derived from the CANDELS legacy fields, *Hubble* legacy deep imaging and analysis of Subaru/Hyper Suprime Cam (HSC) data presented by Harikane et al. (2016). In addition, it was also shown that the projected 2PCF, $w_p(r_p)$, from the same lightcones are in good agreement with the PRIMUS and DEEP2 observations at $0.2 < z < 1.2$ presented by Skibba et al. (2015).

Similar to the previous subsection, here we quantify the field-to-field variance on $w(\theta)$ for different survey areas. The angular correlation functions are calculated in the same manner described in Yung et al. (2022), utilizing the publicly available, CPU-optimised code CORRFUNC[12] (Sinha & Garrison 2020). This package adopts the Landy-Szalay estimator (Landy & Szalay 1993). We note that the positions of satellite galaxies within the host halos are assigned assuming an NFW halo mass profile (Navarro, Frenk & White 1997). We refer the reader to section 2.1 in Yung et al. (2022) for more details.

[12] https://github.com/manodeep/Corrfunc/, v2.3.4

In this section, we make use of lightcones with much larger simulated area than the ones presented in our previous work to quantify the benefits of scaling up survey areas on mitigating field-to-field variance. Similar to the process described in Section 4.2, we iteratively subsample regions within the lightcones. In this exercise, we calculate $w(\theta)$ for galaxies in a number of sub-fields and characterize the variance among the resulting auto-correlation functions. We consider three distinct field sizes: 200 arcmin^2, which is comparable to the size of legacy CANDELS fields and current generation wide-field surveys (e.g. with *JWST*); 0.5 deg^2, which is approximately the size of future *Roman* ultra-deep surveys (approximately 2 WFI pointings); and 2.0 deg^2, which is approximately the size of future *Roman* moderate-depth surveys (see Table 2 and associated text for justification). The galaxy samples are uniformly subject to a $m_{F184} = 28$ detection limit, which roughly correspond to the expected depth of the moderate-depth surveys. For the 200 arcmin^2 field, we randomly sampled 30 sub-regions per simulated lightcones, totalling 150 fields. For 0.5 deg^2, we sample 9 sub-regions for each simulated lightcones, where the centres of these sub-regions are at $(x, y) \in [1/4, 2/4, 3/4]$ of the length of the axes, totalling 45 fields. This arrangement is adopted to minimize overlap among these sub-regions. For 2 deg^2, the full span of the lightcone is utilized, which gives a total of 5 fields. In Fig. 9, we show ACFs computed for galaxies with $m_{F184} < 28$ at $4.0 < z < 4.5$, $6.0 < z < 7$, and $9.0 < z < 10$. For field sizes of 200 arcmin^2 and 0.5 deg^2, we also mark the median and 16th and 84th percentiles with solid and dashed red lines, respectively.

This figure demonstrates the impact of variance in smaller fields, which suffer much larger uncertainties across all scale lengths. Furthermore, the smaller fields are unable to meaningfully capture $w(\theta)$ at both large separation (e.g. $\theta \gtrsim 500$ arcsecond), as the maximum pair separation is limited by the field size, and small separation (e.g. $\theta \lesssim 20$ arcsecond), since objects at these small separations are relatively rare and only a few pairs are captured by a small field. Similar to the cosmic variance exercise done in Section 4.2, the set of scenarios explored in this figure demonstrates that field to field variance in deep clustering studies at high redshift can be reduced to very low levels already with a single *Roman* WFI pointing.

In Fig. 10, using the case of the full 2-deg^2 as an example, we explore the effects of survey depth. We calculate the ACF individually for each realisation and show only the median among the five realizations. For instance, the $m_{F184,lim} = 28$ line (shown in blue) in Fig. 10 corresponds to the right column in Fig. 9 in matching redshift bins. This figure provides a general overview for how the clustering statistics are affected by the survey depth. One can also estimate the increase in scatter in the relation due to field-to-field variance by referencing results from Section 4.2 (e.g. Fig. 7). As noted above, the survey depth strongly impacts the clustering of the selected galaxy population, as deeper surveys will be dominated by fainter galaxies, which are more weakly clustered. In addition, this effect does not scale uniformly across redshift, as the strength of galaxy clustering at a given mass/luminosity also depends on redshift. This is because halos of a given mass represent rarer, more clustered 'peaks' at earlier times in cosmic history. For instance, $w(\theta)$ seems to be uniformly sensitive across all scales at $z \sim 3$. However, the small-scale end becomes increasingly sensitive to the survey depth towards higher redshifts. We also note that the small separation pairs disappear for the shallow survey depths at high redshift as expected, as such bright objects at such small separations are extremely rare.

In Fig. 11, we show the ACF calculated for galaxies in four redshift bins spanning across $4.0 \lesssim z \lesssim 9.0$ for galaxies $M_{UV} < -18$. In addition to the $w(\theta)$ calculated for the full galaxy populations, we show the one- and two-halo terms, where the 'one-halo' term refers to pairs that are within the same halo and the 'two-halo' term refers to pairs where each galaxy is in a distinct dark matter halo. The two-halo term is calculated only for central galaxies. The one-halo term is calculated by subtracting the two-halo term from the $w(\theta)$ for the full galaxy population. This figure illustrates the angular scales where we expect the one vs. two-halo term to dominate at high redshift. The angular scale where the two terms cross ranges from about 3 arcsec at $z \sim 4$ (equivalent to ~ 25 kpc) to ~ 1 arcsec at $z \sim 8$–9 (equivalent to ~ 5 kpc).

4.4   Comparison with other lightcones

We compare our predictions with those from two other studies in which similar lightcones have been constructed, with galaxy properties assigned via empirical models. Here we show a direct comparison of stellar-to-halo mass ratio (SHMR), galaxy counts per Mpc^3, and the ACFs across these models. Both simulations adopted cosmological parameters that are similar to the ones adopted in this work.
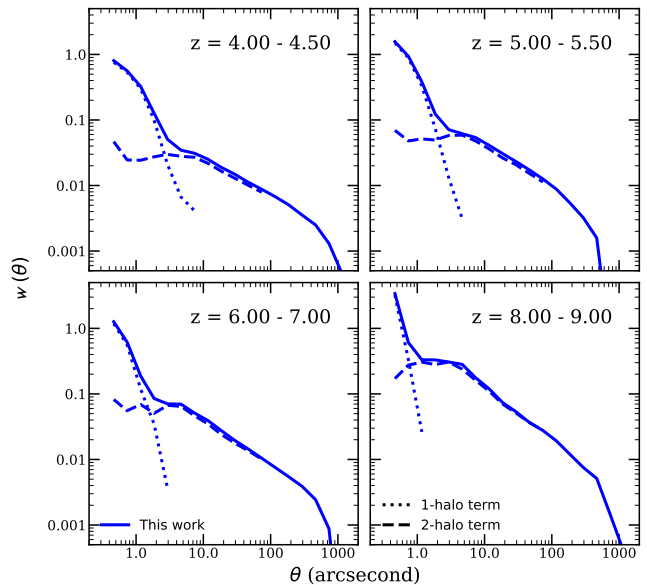

Figure 11. Angular correlation functions, $w(\theta)$, calculated for galaxies with $M_{UV} < -18$ in redshift bins across $z \sim 4$ to ~ 9 from the first realization of the 2-deg^2 lightcones presented in this work. In addition to the $w(\theta)$ calculated for the full galaxy populations (solid line), we show the one- and two-halo terms with dotted and dashed lines, respectively (see text for details).

In addition, halo catalogues underlying all three simulations are generated using Rockstar and Consistent Trees (Behroozi et al. 2013a,b). These simulations also all adopt a Bryan & Norman (1998) virial mass definition, a Chabrier (2003) IMF, and a Calzetti et al. (2000) attenuation curve.

UniverseMachine is an empirical model that is optimized to reproduce a wide variety of observational constraints, including stellar mass functions, UV luminosity functions, cosmic star formation rate, specific star formation rate, etc (Behroozi et al. 2019, 2020). In this empirical model, the star formation is modelled as a function of $v_{M_{peak}}$, z, and Δv_{max}, where $v_{M_{peak}}$ is the circular velocity of the halo at the redshift where it reached its peak mass and Δv_{max} corresponds to the relative change in v_{max} over the past dynamical time. The SFR as a function of v_{max} and quenched fraction $f_{quenched}(v_{max})$ are iteratively tuned using Markov Chain Monte Carlo until the modelled galaxy populations matches a range of observational constraints, including stellar mass functions between $z = 0$ to 4, cosmic star formation rates between $z = 0$ to 10, etc. (see table 1 in Behroozi et al. (2019) associate text for detail). This model has previously been used to interpret CANDELS observations and has been compared to the Santa Cruz SAM in Somerville et al. (2021). In this work, we compared to a custom version of UniverseMachine lightcone, which is constructed with the same underlying halo populations as the ones used in this work. Therefore, the UniverseMachine lightcone compared here has exactly the same footprint and halo mass resolution as the 2-deg^2 lightcones presented in this work. We note that the UniverseMachine lightcone used in this comparison is constructed based on the the exact same underlying set of halos.

The Deep Realistic Extragalactic Model (DREaM)

Galaxy Catalogue[13] is a lightcone created specifically for future deep galaxy surveys with *Roman* (Drakos et al. 2022). This 1-deg^2 lightcone is constructed based on a dark matter-only *N*-body simulation and galaxies from an empirical model (Springel 2005; Williams et al. 2018). The underlying dark matter-only *N*-body simulation has a box size of 115 h^{-1} Mpc and dark matter particle mass of 1.5×10^7 M$_\odot$ h^{-1}. The stellar masses of galaxies are assigned to dark matter halos using the subhalo abundance matching (SHAM) technique and are constructed to match the observed stellar mass functions at $z \leq 4$ and the available observed UV luminosity functions. Galaxies are proportionately but randomly assigned to be star-forming or quiescent, and depending on the galaxy type, various free parameters, such as *e*-folding time (τ) and the age of the Universe at the time star formation started ($t_{\rm start}$), are sampled from a 'parent catalogue' constructed based on known scaling relations. The star formation histories of these galaxies are then modelled using the 'delayed tau model', given by SFR(t) $\propto t\,e^{-t/\tau}$ with t being the time elapsed since $t_{\rm start}$ (see Williams et al. 2018 and references therein). Based on the combination of fixed parameters (M_* and z), randomly assigned galaxy type, and inferred parameters, synthetic SEDs are generated using the Flexible Stellar Population Synthesis (FSPS) package (Conroy et al. 2009). The SFR for each galaxy is then inferred based on the synthetic stellar SED, which also gives $M_{\rm UV}$ and other colour-related predictions, as well as the *Roman* and *JWST* photometry. We note that the cosmic SFR presented in Drakos et al. (2022) is averaged over the last 100 Myr based on the assumed SFH.

It is important to note that while these three simulations share a number of basic components (e.g. cosmological parameters, dark matter-only simulations, halo finder, dust attenuation model), which should not be a major source of differences, these three models took very different approaches to modelling SFR, M_*, and other galaxy properties and observables. Fig. 12 shows the stellar-to-halo mass ratio (SHMR) for galaxies in the first realization of the 2-deg^2 lightcones compared to galaxies from a 2-deg^2 UniverseMachine lightcone and the DREaM lightcone. This figure highlights the difference in the predicted halo occupation between the SAM and both empirical models. In particular, the SAM predicts a milder redshift evolution in SHMR compared to the empirical models. We discuss this further in Section 5.3. We add that this is a mass range where abundance matching, empirical models, and hydrodynamic simulations have very little consensus on even at $z \sim 0$ (e.g. Behroozi et al. 2019; Munshi et al. 2021).

Fig. 13 shows the number density of galaxies per Mpc^3 above specific cut-off values in $m_{\rm F184}$, $M_{\rm UV}$, and M_*. Each of the panels in this plot breaks down the predicted galaxy populations roughly into two groups. We note that these limits are chosen to compare the number of galaxies at both a brighter (or more massive) and a fainter (or less massive) limit. These limits are not meant to be correlated across panels, as the mass-to-luminosity ratio and the observed- and rest-frame magnitude evolve as a function of redshift. For instance, while the SC SAM predicts a similar number of galaxies with $m_{\rm F184} < 27$, this model predicts noticeably
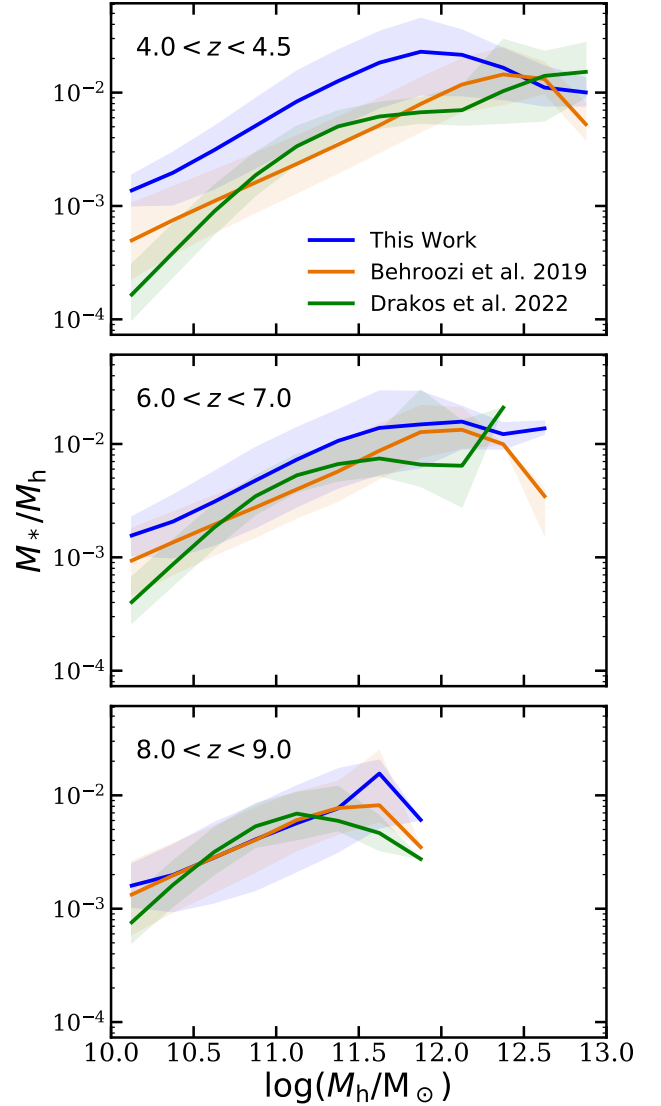
[13] https://www.nicoledrakos.com/dream

Figure 12. Stellar-to-halo mass ratio (SHMR) for predicted galaxies in one of the 2-deg^2 lightcones (blue) between $4.0 < z < 4.5$ (*top*), $6.0 < z < 7.0$ (*middle*), and $8.0 < z < 9.0$ (*bottom*). These results are compared to UniverseMachine (orange) and DREaM (green). For all models, the solid line marks the median of the relation, and the shaded region represents the 17th and 84th percentiles.

more galaxies with $m_{\rm F184} < 29$ at $z > 4$. It is not surprising to see that the number of objects as a function of rest-frame and observed-frame magnitude from these models are in good agreement, as the number density of observed IR-bright galaxies at intermediate redshift are relatively well-constrained and all of the models agree well with these observables. However, the discrepancy in stellar mass reveals that a very different mass-to-magnitude relation is predicted by these different model approaches. Overall, we see that the SAM predicts higher stellar mass content across $2 \lesssim z \lesssim 6$ in halos across all mass ranges at lower redshift. Yung et al. (2019b) has shown a compilation of model predictions, including the Santa Cruz SAM, UniverseMachine, and Williams et al. (2018), and show that these models pre-
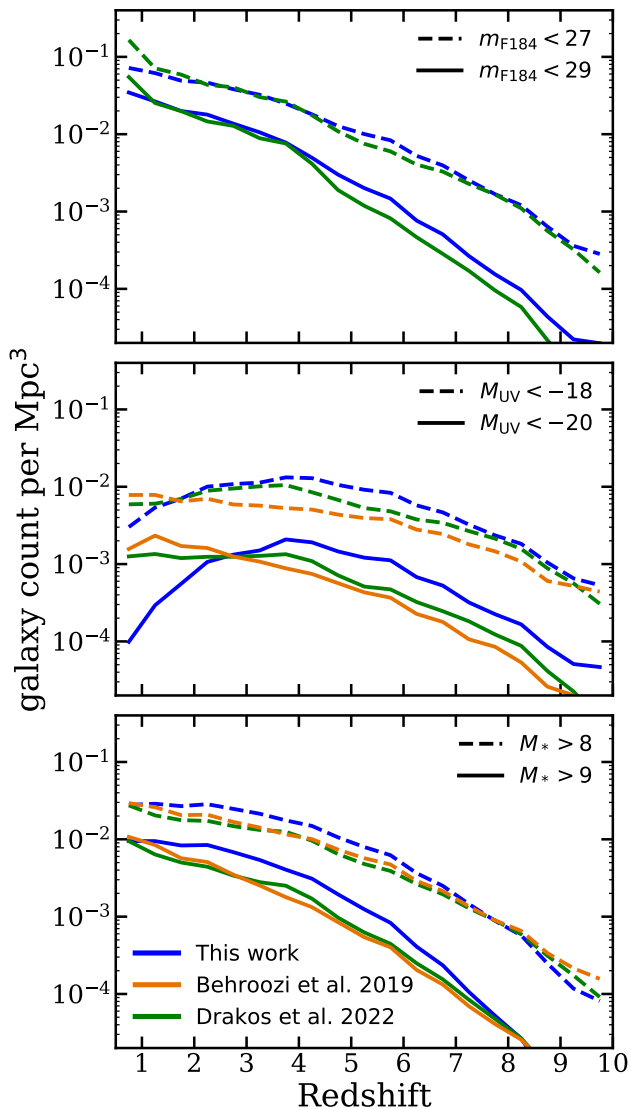
Figure 13. Number density of galaxies per cubic comoving-Mpc between $0.5 < z < 10$ in bins of $\Delta z = 0.5$ from one realization of the 2-deg^2 lightcones presented in this work (blue), UniverseMachine (orange), and DREaM (green). We apply selection criteria by observed-frame IR magnitude in the *Roman* WFI F184 band (*top*), rest-frame UV magnitude (*middle*), and stellar mass (*bottom*).

dicted very different evolution of the underlying stellar mass and star formation rate. Among the models included in the comparison, the Santa Cruz SAM predicts more galaxies across a wide range of M_* and SFR than the other two models at $z \lesssim 6$. We note that the SMFs predicted by these models are well within the uncertainties of the observed constraints currently available (e.g. Duncan et al. 2014; Song et al. 2016; Katsianis et al. 2017a,b).

In Fig. 14, we show a comparison of the predicted ACFs from this work to the two empirically modelled lightcones. This is calculated for all galaxies with $M_{\rm UV} < -18$. As shown in (Yung et al. 2022), the Santa Cruz SAM reproduces the projected 2PCF measured by PRIMUS and DEEP2 in at $0.2 \lesssim z \lesssim 1.2$ reported by Skibba et al. (2015), at inter-
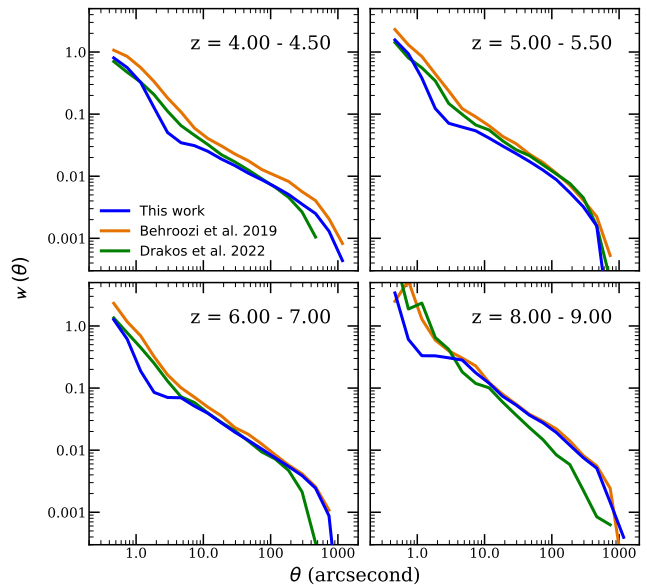

Figure 14. Angular correlation functions calculated for galaxies with $M_{\rm UV} < -18$ in redshift bins across $z \sim 4$ to ~ 9 from one realization of the 2-deg^2 lightcones presented in this work (blue), a 2-deg^2 lightcone from UniverseMachine (orange), and a 1-deg^2 lightcone from DREaM (green).

mediate redshift $1.25 < z < 4.5$ in legacy CANDELS surveys (computed via theory catalogue presented in Somerville et al. 2021), and high-redshift measurement at $3.5 < z < 7.5$ reported by Harikane et al. (2016). The UniverseMachine model has been shown to reproduce the measured clustering of massive galaxies in the nearby universe up to $z \sim 0.7$ from the PRIMUS and DEEP2 survey (Coil et al. 2017). The DREaM lightcone has been shown to reproduce the $0.1 < z < 0.2$ measurements from SDSS (Yang et al. 2012). It is interesting that, in spite of the very different modelling approaches, all three models arrive at very similar predictions for the clustering of rest-UV selected galaxies at $z \gtrsim 4$.

Differences in the ACF on the small-scale end likely arise mainly from the modelling of satellite galaxies, which is very uncertain. The satellite galaxies in UniverseMachine are modelled using $\Delta v_{\rm max}$ over the past dynamical time, which captures the rapid drop of $v_{\rm max}$ following a major merger (Behroozi et al. 2014). $\Delta v_{\rm max}$ has been shown to be a more robust measurable quantity for satellites and yield more clearly orbit- and profile-dependent satellite SFRs (Onions et al. 2012). On the other hand, DREaM adopts the peak of the maximum circular velocity, $V_{\rm peak}$, over the entire merger history as a halo mass proxy. SHAM is first performed between the SMF and halo populations, and then a random number is assigned to distinguish whether a galaxy is quiescent or star-forming. This assignment does not distinguish whether a halo is a central or sub-halo, and therefore the environmental dependence of galaxy quenching is not accounted for. In the SC SAMs, galaxies are starved of new gas cooling once they become satellites, and are therefore probably over-quenched. For that reason, our satellite galaxies are in general fainter compared to the other two models and are more susceptible to the magnitude cut applied. Additional comparison for the satellite populations in

these models is available in Appendix C. It is not surprising that the agreement at small scales is closer between the two empirical models, as although the modelling of satellites differs in detail, it is quite similar in spirit. On the other hand, the Santa Cruz SAM seems to predict a stronger 'shoulder' between the one- and two-halo terms in the ACF. Future observations should thus be able to constrain the physical processes that shape satellite galaxy properties. The differences on the very large-scale end of the ACFs reflect the difference in lightcone sizes between DREaM (1 deg^2) and UNIVERSEMACHINE and the SC SAM (2 deg^2).

5 DISCUSSION

The highly anticipated *Roman Space Telescope* marks the beginning of a new era of *deep-wide* surveys. *Roman*'s wide field of view, sensitivity, and spatial resolution promise to deliver a new generation of *deep-wide* surveys, which will reach depths comparable to the legacy ultra-deep *Hubble* surveys (e.g. HUDF) or wide-field *JWST* surveys while covering areas several times those of current generation wide-field galaxy surveys by space-based telescopes (e.g. CANDELS). These future generation deep-wide surveys are expected to deliver more robust statistical constraints through detecting large populations of high-redshift galaxies. They are also expected to detect some very rare objects, including massive galaxies, luminous quasars, or perhaps even the first stars. Semi-analytic models for galaxy formation have been shown to be a promising tool for interpreting legacy CANDELS observations (Somerville et al. 2021) and forecasting for upcoming *JWST* surveys (Yung et al. 2022). The physically-motivated Santa Cruz SAM has been calibrated to observed constraints from the local universe and its performance in terms of reproducing the observed evolution of the high-redshift galaxy population has been rigorously tested in the series of *Semi-analytic forecasts for JWST* paper series (e.g. Yung et al. 2019a) and other works. In this work, we present a set of multi-deg^2 lightcones that have been created to prepare for future deep surveys with *Roman* and other next generation wide-field survey telescopes, such as Euclid and Rubin. These lightcones also include photometry for a large set of current generation of space- and ground-based instruments, as well as instruments used in legacy surveys. See Table A1 for the full list of available photometry provided in the mock catalogues.

5.1 The promise of deep-wide-field surveys

The new generation of highly efficient wide-field survey instruments will redefine *deep surveys*, as surveys of the size of a handful of telescope pointings will surpass the coverage of the widest 'deep' surveys conducted with current generation instruments, and the largest 'deep' surveys achievable with many pointings will reach at least tens of square degrees. The many benefits of wide-field, high resolution space-based imaging with *Hubble* that were discussed in Somerville (2005) still hold today in a new era of high-redshift surveys with the next generation of wide-field survey instruments. These wide-field surveys will be able to measure the stellar mass assembly history across cosmic time and disaggregate it by morphological type, large scale environment, and other

galaxy properties. The large galaxy sample expected from these surveys will also strengthen our understanding of the connection between the processes that regulate star formation on sub-galactic scales and the overall global trends in the star formation and mass assembly. Furthermore, detecting galaxies with embedded accreting supermassive black holes forming in the early part of cosmic history will help reveal the relationship between star-forming galaxies and AGN, and the processes that regulate black hole feeding and growth.

CANDELS and and other large legacy surveys have addressed many of these questions and produced observational constraints that tremendously improved our understanding of galaxy formation in the context of cosmic evolution. However, as observations push towards the very early Universe, some of these questions remain incompletely or imprecisely answered, limited by the capability of current generation instruments, which in turn limit feasible survey sizes and depths. Currently, only a few dozen objects at extreme redshifts (e.g. $z > 8$) have been detected to date. While *JWST* is expected to find many more faint objects within the survey areas of familiar legacy surveys, it is not expected to find exotic objects such as massive, bright galaxies, as that would require a larger survey area. *Roman*, on the other hand, will deliver next generation wide-field surveys that are expected to efficiently detect large numbers of massive galaxies in the early Universe and will provide more robust statistics on the number counts of bright galaxies.

As demonstrated in Sections 4.2 and 4.3, the current generation surveys are susceptible to significant uncertainties arising from field to field variance, and these increase at higher redshift for a given galaxy mass or rest-frame luminosity. Somerville et al. (2004) and Moster et al. (2011) have presented predictions for the cosmic variance expected in the legacy CANDELS fields at $0.5 < z < 4$. As shown in Finkelstein et al. (2022b), based on a 2-deg^2 lightcone presented in this work, the EGS field as observed by *Hubble* and *Spitzer* could be over-dense at $z \sim 9$ relative to the cosmic mean. Similarly, Yung et al. (2022) presented a controlled experiment with 40 realizations of of the order of ~ 1000 deg^2 lightcones and showed that the field-to-field variance can be up to $\sigma_v \sim 0.40$ for survey reaching observed-frame IR magnitude in *JWST*/NIRCam F200W $m_{F200W} \sim 27$ at $z \sim 9$.

Taking advantage of the new suite of 2-deg^2 lightcones, we presented detailed predictions of field-to-field variance between $1 < z < 10$ as a function of survey area and depth. We explored survey fields of different sizes ranging from approximately the size of UDF or a handful of *JWST* NIRCam pointings to a single *Roman* WFI pointing, and for depths that were reachable by legacy CANDELS surveys to those expected for ultra deep *Roman* surveys. This experiment presents quantitative predictions that enable informed estimates of the area needed to reduce field-to-field variance to a desired level for a population with specified intrinsic or observable properties. For instance, Fig 7 shows that an ultra-deep *Roman* survey (with area ~ 1000 arcmin^2 reaching depth $m_{F184} < 29$) will reduce $\sigma_v \sim 0.75$ at $z \sim 9$ from a CANDELS-like survey (with ~ 200 arcmin^2 reaching depth $m_{IR} < 25$) to $\sigma_v \sim 0.10$. Obviously, our publicly available lightcones can be utilized to create similar estimates for any desired survey configuration that fits within their constraints on area and depth.

5.2 Advantages of physics-based models

Based on a well-established, versatile model for galaxy formation, we are able to provide a wide range of self-consistent, physically-backed predictions for high-resolution synthetic SEDs, multi-wavelength photometry across observatories, as well as the underlying physical properties and star formation history. The semi-analytic modelling approach incorporates a wide range of physical processes and simultaneously and self-consistently explores their effects in a computationally efficient way. SAMs can also be used to explore the effects of varying the uncertain parameters that characterize these processes in current models, as shown in Yung et al. (2019a,b). Achieving the dynamic range of the mock lightcones presented here (in galaxy mass/luminosity and area/volume) is well out of reach for current conventional numerical hydrodynamics simulations. Our mock catalogues provide tens of millions of galaxies, which provide the statistically robust sample size required for forecasting future wide-field surveys.

On the other hand, although empirical models can very inexpensively fill large volumes with galaxies, they provide limited insights into the physical processes that shape galaxy properties. Moreover, they probably do not accurately represent the diversity of star formation histories in the galaxy population, which is important for producing a realistic suite of galaxy SEDs. These are important for designing and testing multi-instrument synergies and strategies. For example, narrow- and intermediate-band flux on HSC can efficiently refine the redshift estimates for Lyman-break galaxy candidates detected with broad-band dropout in VISTA/VIRCam (Endsley et al. 2022). Combining flux measurements from filters on different instruments with a similar but slightly offset wavelength range can be used as a pseudo narrow band to refine redshift estimates (e.g. *JWST*/NIRCam F150W and *HST*/WFC3 F160W, see Finkelstein et al. 2022c), or two instruments with complementary wavelength coverage can be used (e.g. *HST*/WFC3 Near-IR capability and *Spitzer*/IRAC mid-IR capability, see Finkelstein et al. 2022b). In addition, some of the most exciting applications of next generation wide-field surveys will be cross-correlation studies using galaxies detected via different multi-wavelength observational tracers. Current empirical models are not able to self-consistently predict, for example, the stellar mass, SFR, H I and H_2 content, and dust content of galaxies. These different components of galaxies, as probed by different tracers such as UV, optical, and IR continuum emission, 21-cm emission, and sub-mm lines such as CO and [C II], will ultimately provide very stringent constraints on and help to break degeneracies between different physical processes in galaxies. The same semi-analytic models presented here have also been used very successfully to predict some of these tracers (Popping et al. 2019a,b; Yang et al. 2021, 2022).

5.3 Our results in the context of other model predictions

In this work, we performed a comparison between our predicted lightcones and two other mock lightcone suites constructed with the empirical models UNIVERSEMACHINE (Behroozi et al. 2019, 2020) and DREaM (Drakos et al. 2022). We compared the stellar-to-halo mass ratio, bright

and faint galaxy populations, and the angular correlation functions. Given that these lightcones are comparable in size and depth with the ones presented in this work, we are able to apply selection criteria for m_{F184}, M_{UV}, and M_* similar to the ones adopted in earlier sections in this work. While all three lightcones presented in this comparison share some modelling assumptions, including a ΛCDM cosmology (with comparable cosmological parameters), virial mass definition, and underlying IMF, it is worth noting that these models took very different approaches to modelling star formation histories, which is the main reason for the discrepancy seen in this comparison. We note that although we do not make a direct comparison with them, related predictions for field-to-field variance have also been presented in the literature by Somerville et al. (2004); Moster et al. (2011); Trenti & Stiavelli (2008); Bhowmick et al. (2020); Endsley et al. (2020). We also note that there are also lightcones that have been generated based on hydrodynamic simulations (e.g. Snyder et al. 2017, 2022; Kaviraj et al. 2017), which are excellent for mock images and for studies that require realistic structural information for galaxies (e.g. Costantin et al. 2022; Garcia-Argumanez et al. 2022; Rose et al. 2022). However, in general these lightcones span much smaller areas than the ones that we focus on in this work, and are therefore not included in the comparison.

Given the nearly identical assumptions that went into simulating the halo populations, the contribution to the difference among the models from the halo properties is expected to be minuscule. Therefore, the stellar-to-halo mass ratios presented in Fig. 12 serve as direct diagnostics for the relationship between the stellar mass of galaxies and the mass of their host dark matter halos. It is not surprising that the two empirical models, driven by a similar set of observational constraints, are in broad agreement. Our model directly predicts the stellar-to-halo mass relation and the dispersion in it, and it is a result of the physical processes included in the model, including merger-induced episodic star bursts and reduced star-formation activity in low-mass halos due to the ejection of gas by stellar feedback. While the model is calibrated to reproduce the observed stellar mass function (e.g. Bernardi et al. 2013) and SHMR (e.g. Rodríguez-Puebla et al. 2017) from abundance matching at $z \sim 0$, it is not explicitly *'tuned'* to match higher redshift constraints. On the other hand, the empirical models are optimized to reproduce observed constraints over a wide range of redshifts, where the connection between stellar mass and halo mass is obtained by varying specific scaling relation(s) (e.g. SFR–v_{max} and $f_{quenched}$–v_{max} in UNIVERSEMACHINE). The SAM predicts very little redshift-evolution in the SHMR (for low-mass halos), while predictions from both empirical models increase monotonically towards high redshift. This is a direct consequence of the parametrization of the mass-loading of stellar driven winds with halo maximum circular velocity in the SAM.

Similarly, the comparison in Fig. 13 helps demonstrate the differences between these models by breaking down the predicted galaxy populations, and sheds light on the differences seen in the predicted SHMR. All three models produce similar predictions for the *observable* quantity, the rest-UV luminosity function, as reflected in the top two panels of Fig. 13. However, the SAM accomplishes this with a different ratio of rest-UV light to stellar mass than the empiri-

cal models, which is due to the galaxies containing stellar populations with a different age and/or metallicity distribution, and/or a different assumed dust attenuation. The SAM self-consistently *predicts* the joint age-metallicity distribution in each galaxy based on the star formation and assembly history computed from physical prescriptions, and the dust attenuation law is adjusted empirically to match the observations. The empirical models effectively construct star formation histories by matching to observational estimates of SFR in high redshift galaxies, and assume a relationship between stellar mass and metallicity that is obtained by extrapolating a lower redshift observational estimate. The SAM generally predicts that there is more mass in stars at $4 \lesssim z \lesssim 7$ than the empirical models (bottom panel of Fig. 13). As previously discussed in Yung et al. (2019b), these predictions from our model are in tension with the evolution of $M_*/M_{\rm h}$ over the interval $z \sim 4$–7 as derived from observations as presented in Finkelstein et al. (2015b). See also fig. 4 in Yung et al. 2019b for the predicted stellar mass function from the Santa Cruz SAM and how it compares to the uncertainties in existing observations. We note that the differences in the predicted number density of both massive and low-mass galaxies between the SAM and the two empirical models gradually decreases towards lower redshift at $z < 4$, which indicate the discrepancy in the predicted SHMR would narrow and eventually converge at $z \lesssim 1$. It is important to note that there are very large uncertainties in the estimates of physical parameters such as stellar mass, SFR, and metallicity from observations, and these flow through to the empirical models. These constraints will be improved significantly by *JWST* and other upcoming facilities including *Roman*.

The predicted angular correlation functions from the different lightcones are in remarkably good agreement when galaxies are selected via rest-UV luminosity. The large-scale separation ACF calculated for galaxies in the DREaM lightcone is slightly affected by the size of the lightcone, which is a quarter of the size of the 2-deg^2 lightcones made with the Santa Cruz SAM and UniverseMachine. Furthermore, differences across the ACF predicted by these different models are most noticeable in the small-scale separation, which is dominated by galaxies that reside within the same host halo (also referred to as the 'one-halo term'). As already noted, the Santa Cruz SAM uses a rather different approach for modeling the positions of satellite galaxies within their host halos.

5.4 Caveats, limitations, and uncertainties

Given that the 2-deg^2 lightcones presented in this work are constructed with the same physical model and tools as the wide-field and ultra-deep lightcones presented in Yung et al. (2022), we refer the reader to that work (and the accompanying paper series) for discussion related to the construction of the lightcone and the predicted galaxies. Here, we focus on the caveats most closely related to the new results presented in this work.

Like many SAMs, the SC SAMs compute galaxy properties based on an input merger history. Thus, these models typically do not account for the effects of environment on scales larger than halos, such as the tidal interaction of halos in a high density environment (see these studies that account

for impact from environment: Buck et al. (2019); Ayromlou et al. (2021); Kuschel et al. (2022)). Furthermore, the modelling of satellite galaxies also has some significant known issues and limitations. The SC SAM does not model the tidal or ram pressure stripping of satellites (which would impact their hot and cold gas, stellar bodies, and dark matter halos), and is known to produce satellite populations that are 'over-quenched' relative to observations (Kimm et al. 2009). A related issue is that the SC SAMs do not use the information on sub-halo evolution from the N-body simulations, but use a simple semi-analytic model to estimate the time required for satellite orbits to decay until they merge with the central galaxy, as well as the tidal stripping and destruction of satellites. Although the model has been tuned to produce good agreement with the satellite conditional mass functions and fractions (see Appendix C), the predicted radial distribution of satellites does not match predictions from e.g. UniverseMachine. As a result, we have re-assigned the radial positions of the surviving satellites as described in Section 2. In addition, in this work, in order to increase the dynamic range of our lightcones, we made use of halo merger histories constructed using the extended Press-Schechter formalism (see Section 2). These do not capture the known correlation between halo large scale environment (clustering) and mass accretion history, leading to small differences in the predicted clustering of galaxy populations selected by stellar mass or luminosity.

We also add that the sample selection throughout this work is idealized, but these catalogues provide a platform that can be used to include more realistic observational effects. While the effects of IGM extinction and ISM dust attenuation are included in the modelling pipeline, the photometry presented in this work does not include nebular emission from ionized H_{II} regions. It has been shown that strong emission line features in star forming galaxies can have an effect on the photometric redshift estimates from broad-band fluxes (e.g. Finkelstein et al. 2011, 2022c; Larson et al. 2018). In addition, we assumed that the observed magnitude includes the total intrinsic luminosity, while in real galaxy surveys, the extracted photometry for extended objects depends in a complex way on the galaxy size and light profile, the depth and S/N of the image, and the method used to extract the photometry. In a companion paper, Bagley et al. (in preparation) add observational effects to these lightcones to more realistically explore the role of foreground contamination on the selection of high-redshift Lyman-break galaxies. We add that gravitation lensing is another caveat that affects the recovered LF through magnification bias, and modulates the angular correlation function, but is outside of the scope of both this work and its companion work. Also, in preparation for the *JWST* Cosmic Evolution Early Release Survey (CEERS), we are creating noiseless mock images which then have *JWST* instrument effects injected (Bagley and the CEERS Collaboration, in preparation). Similar procedures can be used to generate realistic mock images for the *Roman Space Telescope* and other facilities.

6 SUMMARY AND CONCLUSIONS

We presented a suite of five 2-deg^2 lightcones that spans $0 < z < 10$ and resolves galaxy populations down to $\log(M_*/M_\odot) > 7$. We provide predictions for galaxies based on dark matter halos sourced from a cosmological N-body simulation and the physically motivated and well-established Santa Cruz semi-analytic model with configurations detailed in the *Semi-analytic forecasts for JWST* project. Taking advantage of the large area and large number of sample galaxies, we present forecasts that highlight the potential of future surveys that will be orders of magnitude wider than current generation observations of comparable depth. In this work, we focus on predictions that utilise the spatial distribution and redshift evolution for galaxies available in these lightcones.

We summarize our main conclusions below.

(i) We present the predicted evolution of object counts and number density as a function of redshift, as well as the contribution of these galaxies to the cosmic SFR.

(ii) We predict that the uncertainty due to field to field variance for deep surveys ($m_{\rm F184,lim} \sim 29$) can be reduced to about $\sim 10\%$ at $z \sim 6$–10 for a 1000 arcmin^2 survey. For brighter limits ($m_{\rm F184,lim} \sim 24$), the uncertainty due to field-to-field variance for such a survey is 25–50 per cent at this survey area.

(iii) We predict that for magnitude limits typical of moderate depth *Roman* surveys ($m_{\rm F184,lim} \sim 28$), the uncertainty on two-point correlation function estimates in 0.5 deg^2 fields at $z \sim 4$–4.5 for separation ~ 10 to 100 arcsecond due to field-to-field variance is ~ 15 per cent. Surveys with an area of a few deg^2 should be sufficiently large for field-to-field variance to be a sub-dominant source of uncertainty.

(iv) We compared our predictions with two other lightcone simulations from the literature that populated halos with galaxies using empirical models. Our physics-based models predict a significantly higher stellar mass to halo mass relation at $z \sim 4$ compared with the empirical models, although all three models match the observed rest-UV luminosity function. All three models make similar predictions for the angular correlation functions at $z \sim 4$–9.

The results presented in this paper are intended as just a few examples of the predictions that can be extracted from these lightcones. All of the lightcones are made publicly available so that the community can exploit them for numerous additional applications.

DATA AVAILABILITY

The data underlying this paper are available in the Data Product Portal hosted by the Flatiron Institute at https://www.simonsfoundation.org/semi-analytic-forecasts/ and the Flatiron Institute Data Exploration and Comparison Hub (Flahub, https://flathub.flatironinstitute.org/group/sam-forecasts).

REFERENCES

Andrews B. H., Martini P., 2013, ApJ, 765, 140
Ayromlou M., Kauffmann G., Yates R. M., Nelson D., White S. D., 2021, MNRAS, 505, 492
Bagley M. B., et al., 2022, arXiv:2211.02495
Barrera M., et al., 2022, arXiv:2210.10419
Beckwith S. V. W., et al., 2006, AJ, 132, 1729
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013b, ApJ, 763, 18
Behroozi P. S., Wechsler R. H., Conroy C., 2013c, ApJ, 770, 57
Behroozi P. S., Wechsler R. H., Lu Y., Hahn O., Busha M. T., Klypin A., Primack J. R., 2014, ApJ, 787
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Behroozi P., et al., 2020, MNRAS, 499, 5702
Benson A. J., 2010, New Astron., 17, 175
Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697
Bernyk M., et al., 2016, ApJS, 223, 9
Bhowmick A. K., Somerville R. S., Di Matteo T., Wilkins S., Feng Y., Tenneti A., 2020, MNRAS, 496, 754
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Boselli A., Cortese L., Boquien M., 2014, A&A, 564, A65
Bouwens R. J., et al., 2015, ApJ, 803, 34
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

Buck T., Macciò A. V., Dutton A. A., Obreja A., Frings J., 2019, MNRAS, 483, 1314

Calette A. R., Avila-Reese V., Rodríguez-Puebla A., Hernández-Toledo H., Papastergis E., 2018, Rev. Mex. Astron. y Astrofísica, 54, 443

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Casey C. M., et al., 2022, arXiv:2211.07865

Chabrier G., 2003, PASP, 115, 763

Coil A. L., Mendez A. J., Eisenstein D. J., Moustakas J., 2017, ApJ, 838, 87

Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486

Costantin L., et al., 2022, arXiv:2208.00007

Cowley W. I., Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 2018, MNRAS, 474, 2352

Dayal P., Ferrara A., Dunlop J. S., Pacucci F., 2014, MNRAS, 445, 2545

Dayal P., Mesinger A., Pacucci F., 2015, ApJ, 806, 67

Drakos N. E., et al., 2022, ApJ, 926, 194

Duncan K., et al., 2014, MNRAS, 444, 2960

Dunlop J. S., et al., 2021, JWST Proposal ID 1837. Cycle 1 GO/Treasury

Ellis R. S., et al., 2013, ApJ, 763, L7

Endsley R., Behroozi P., Stark D. P., Williams C. C., Robertson B. E., Rieke M., Gottlöber S., Yepes G., 2020, MNRAS, 493, 1178

Endsley R., et al., 2022, MNRAS, 512, 4248

Euclid Collaboration et al., 2022a, A&A, 662, A92

Euclid Collaboration et al., 2022b, A&A, 662, A112

Finkelstein S. L., Cohen S. H., Moustakas J., Malhotra S., Rhoads J. E., Papovich C., 2011, ApJ, 733

Finkelstein S. L., et al., 2015a, ApJ, 810, 71

Finkelstein S. L., et al., 2015b, ApJ, 814, 95

Finkelstein S. L., et al., 2017, JWST Proposal ID 1345. Cycle 0 Early Release Science

Finkelstein S. L., et al., 2021, JWST Proposal ID 2079. Cycle 1 GO/Treasury

Finkelstein S. L., et al., 2022a, arXiv:2211.05792

Finkelstein S. L., et al., 2022b, ApJ, 928, 52

Finkelstein S. L., et al., 2022c, ApJL, 940, L55

Gabrielpillai A., Somerville R. S., Genel S., Rodriguez-Gomez V., Pandya V., Yung L. Y. A., Hernquist L., 2022, MNRAS, 517, 6091

Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41

Garcia-Argumanez A., et al., 2022, arXiv:2207.14062

Gardner J. P., et al., 2006, Space Science Reviews, 123, 485

Gnedin N. Y., Kravtsov A. V., 2011, ApJ, 728, 88

Grogin N. A., et al., 2011, ApJS, 197, 35

Harikane Y., et al., 2016, ApJ, 821, 123

Harikane Y., et al., 2018, PASJ, 70, 1

Harikane Y., et al., 2022a, ApJS, 259, 20

Harikane Y., et al., 2022b, ApJ, 929, 1

Henriques B. M., White S. D., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663

Henriques B. M. B., Yates R. M., Fu J., Guo Q., Kauffmann G., Srisawat C., Thomas P. A., White S. D. M., 2020, MNRAS, 491, 5795

Hu W., et al., 2019, ApJ, 886, 90

Illingworth G. D., et al., 2013, ApJS, 209, 6

Ivezić Ž., et al., 2019, ApJ, 873, 111

Kakos J., Primack J. R., Rodríguez-Puebla A., Tejos N., Yung L. Y. A., Somerville R. S., 2022, MNRAS, 514, 1857

Kannan R., et al., 2022a, arXiv:2210.10066

Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022b, MNRAS, 511, 4005

Kartaltepe J., et al., 2021, JWST Proposal ID 1727. Cycle 1 GO/Treasury

Katsianis A., Tescari E., Blanc G., Sargent M., 2017a, MNRAS, 464, 4977

Katsianis A., et al., 2017b, MNRAS, 472, 919

Kaviraj S., et al., 2017, MNRAS, 467, 4739

Kimm T., et al., 2009, MNRAS, 394, 1131

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Koekemoer A. M., et al., 2011, ApJS, 197, 36

Koekemoer A. M., et al., 2013, ApJS, 209, 3

Koekemoer A. M., et al., 2019, arXiv:1903.06154

Kuschel M., et al., 2022, arXiv:2205.12169

Lacey C., Cole S., 1993, MNRAS, 262, 627

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Larson R. L., et al., 2018, ApJ, 858, 94

Larson R. L., et al., 2022, arXiv:2211.10035

Leung G. C. K., et al., 2022, in prep

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388

McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184

McLeod D. J., McLure R. J., Dunlop J. S., 2016, MNRAS, 459, 3812

Moster B. P., Somerville R. S., Newman J. A., Rix H.-W., 2011, ApJ, 731, 113

Munshi F., Brooks A. M., Applebaum E., Christensen C. R., Quinn T., Sligh S., 2021, ApJ, 923, 35

Naab T., Ostriker J. P., 2017, ARA&A, 55, 59

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Oesch P. A., et al., 2013, ApJ, 773, 75

Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, ApJ, 855, 105

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713

Onions J., et al., 2012, MNRAS, 423, 1200

Ono Y., et al., 2018, PASJ, 70, 2

Overzier R., Lemson G., Angulo R. E., Bertin E., Blaizot J., Henriques B. M., Marleau G. D., White S. D., 2013, MNRAS, 428, 778

Pandya V., et al., 2020, ApJ, 905, 4

Papovich C., et al., 2016, ApJS, 224, 28

Pasquale B., et al., 2014, in Figueiro M., Lerner S., Muschaweck J., Rogers J., eds, Vol. 9293, International Optical Design Conference 2014. p. 929305, doi:10.1117/12.2177847

Pasquale B. A., et al., 2018, in Johnson R. B., Mahajan V. N., Thibault S., eds, Curr. Dev. Lens Des. Opt. Eng. XIX. SPIE, p. 107450K, doi:10.1117/12.2325859

Peebles P. J. E., 1980, in , Large-Scale Struct. Universe. Princeton University Press, Princeton, https://ui.adsabs.harvard.edu/abs/1980lssu.book.....P

Peeples M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, ApJ, 786, 54

Planck Collaboration 2016, A&A, 594, A13

Popping G., Somerville R. S., Trager S. C., 2014, MNRAS, 442, 2398

Popping G., Narayanan D., Somerville R. S., Faisst A. L., Krumholz M. R., 2019a, MNRAS, 482, 4906

Popping G., et al., 2019b, ApJ, 882, 137

Price-Whelan A. M., et al., 2018, AJ, 156, 123

Pullen A. R., et al., 2022, arXiv:2209.02497

Reback J., et al., 2022, pandas-dev/pandas: Pandas, doi:10.5281/zenodo.6408044, https://doi.org/10.5281/zenodo.6408044

Robitaille T. P., et al., 2013, A&A, 558, A33

Rodríguez-Puebla A., Behroozi P., Primack J., Klypin A., Lee C., Hellinger D., 2016, MNRAS, 462, 893

Rodríguez-Puebla A., Primack J. R., Avila-Reese V., Faber S. M., 2017, MNRAS, 470, 651

Rose C., et al., 2022, arXiv:2208.11164

Sinha M., Garrison L. H., 2020, MNRAS, 491, 3022

Skibba R. A., et al., 2015, ApJ, 807, 152

Smith A., Cole S., Baugh C., Zheng Z., Angulo R., Norberg P., Zehavi I., 2017, MNRAS, 470, 4646

Snyder G. F., Lotz J. M., Rodriguez-Gomez V., da Silva Guimarães R., Torrey P., Hernquist L., 2017, MNRAS, 468, 207

Snyder G. F., Pena T., Yung L. Y. A., Rose C., Kartaltepe J. S., Ferguson H. C., 2022, arXiv:2211.09677

Somerville R., 2005, New Astron. Rev., 49, 366

Somerville R. S., Davé R., 2015, ARA&A, 53, 31

Somerville R. S., Kolatt T. S., 1999, MNRAS, 305, 1

Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087

Somerville R. S., Lee K., Ferguson H. C., Gardner J. P., Moustakas L. A., Giavalisco M., 2004, ApJ, 600, L171

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, MNRAS, 423, 1992

Somerville R. S., Popping G., Trager S. C., 2015, MNRAS, 453, 4338

Somerville R. S., et al., 2021, MNRAS, 502, 4858

Song M., et al., 2016, ApJ, 825, 5

Spergel D., et al., 2013, arXiv:1305.5425

Spergel D., et al., 2015, arXiv:1503.03757

Springel V., 2005, MNRAS, 364, 1105

Stevans M. L., et al., 2018, ApJ, 863, 63

Stevans M. L., et al., 2021, ApJ, 921, 58

Switzer E. R., et al., 2021, Journal of Astronomical Telescopes, Instruments, and Systems, 7, 044004

Tacchella S., et al., 2022, ApJ, 927, 170

Toshikawa J., et al., 2018, PASJ, 70, 1

Trenti M., Stiavelli M., 2008, ApJ, 676, 767

Virtanen P., et al., 2020, Nat. Methods, 17, 261

Vogelsberger M., et al., 2020, MNRAS, 492, 5167

Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435

Wilkins S. M., et al., 2022a, arXiv:2204.09431

Wilkins S. M., et al., 2022b, MNRAS, 517, 3227

Williams C. C., et al., 2018, ApJS, 236, 33

Wold I. G. B., et al., 2019, ApJS, 240, 5

Wold I. G. B., et al., 2022, ApJ, 927, 36

Yang X., Mo H. J., Van Den Bosch F. C., Zhang Y., Han J., 2012, ApJ, 752

Yang S., Somerville R. S., Pullen A. R., Popping G., Breysse P. C., Maniyar A. S., 2021, ApJ, 911, 132

Yang S., Popping G., Somerville R. S., Pullen A. R., Breysse P. C., Maniyar A. S., 2022, ApJ, 929, 140

Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., 2019a, MNRAS, 483, 2983

Yung L. Y. A., Somerville R. S., Popping G., Finkelstein S. L., Ferguson H. C., Davé R., 2019b, MNRAS, 490, 2855

Yung L. Y. A., Somerville R. S., Popping G., Finkelstein S. L., 2020a, MNRAS, 494, 1002

Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., Venkatesan A., Behroozi P., Ferguson H. C., 2020b, MNRAS, 496, 4574

Yung L. Y. A., Somerville R. S., Finkelstein S. L., Hirschmann M., Davé R., Popping G., Gardner J. P., Venkatesan A., 2021, MNRAS, 508, 2706

Yung L. Y. A., et al., 2022, MNRAS, 515, 5416

Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, ApJ, 771, L19

Zheng Z.-Y., et al., 2017, ApJ, 842, L22

van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22

APPENDIX A: SUMMARY OF AVAILABLE COLUMNS IN THE LIGHTCONES

In Table A1, we summarize the predicted quantities available in the mock catalogues. In addition to basic coordinate information, such as RA, Dec, and redshift, we provide a comprehensive set of halo properties and galaxy physical properties. We also provide rest- and observed-frame photometry for a wide variety of space- and ground-based telescopes, including Galaxy Evolution Explorer (GALEX), Sloan Digital Sky Survey (SDSS), Advanced Camera for Surveys (ACS) and Wide Field Camera 3 (WFC3) on-board *Hubble*, Infrared Array Camera (IRAC) on-board *Spitzer*, United Kingdom Infrared Telescope (UKIRT), Visible and Infrared Survey Telescope for Astronomy (VISTA), Dark Energy Camera (DE-Cam) of the Dark Energy Survey, *JWST* NIRCam broad- and medium-bands, *Roman* WFI, Euclid Observatory, and Rubin Observatory (formerly known as LSST).

While we present photometry for a large collection of space- and ground-based telescopes in this unified catalogue to maximize convenience, we add that the development and verification of some of these results are from past studies. Here we give a general summary for the references of past studies. Reference Somerville et al. (2021, for $z = 0$ to 6) and Yung et al. (2019a,b, for $z = 4$ to 10) for predicted physical properties of galaxies. Reference this work for *Roman*, *Euclid*, *Rubin* (labelled as LSST), and DECam. Reference Yung et al. (2019a, 2022) for *JWST* photometry. Reference Somerville et al. (2021) for standard rest-frame filters, *HST*, *Spitzer*, SDSS, GALEX, UKIRT, VISTA, and other filters used in the CANDELS survey.

APPENDIX B: CODE SNIPPET FOR CALCULATING VOLUME IN A LIGHTCONE SLICE

This simple piece of code is very handy for calculating the comoving volume for a slice of the lightcone. This is essential to calculate the volume-averaged quantities presented in this work. We use the `comoving_volume` function from `astropy` to calculate the comoving volume of the whole universe at two given redshift `z1` and `z2`, and obtain the comoving volume of the universe between `z1` and `z2`. We then multiple the full-sky co-moving volume by the fraction of the surveyed area (in square arcminutes) over 4π steradian ≈ 148510656 arcmin^2 to obtain the volume within the surveyed area.

```
from astropy.cosmology import FlatLambdaCDM
cosmo = FlatLambdaCDM(Om0=0.307, Ob0=0.048, H0=67.8)

def volume_estimate(z1, z2, sqarcmin):
    volume_z1 = cosmo.comoving_volume(z1)
    volume_z2 = cosmo.comoving_volume(z2)
    sky_fraction = sqarcmin/148510656.

    vol = (volume_z2-volume_z1).value*sky_fraction
    return abs(vol) ### return volume in cMpc^3
```

APPENDIX C: SATELLITE GALAXY COMPARISON

In this Appendix, we provide a set of comparisons across the Santa Cruz SAM, UNIVERSEMACHINE, and DREaM for

Table A1. List of all predicted quantities available in the mock galaxy catalogues. We provide the columns (in the ASCII format catalogue 'lightcone.dat'), brief descriptions, and units associated with these quantities. For rest- and observed-frame, we list the intrinsic, unattenuated photometry filter sets and filter bands available (e.g. UV1500_rest, UV2300_rest, and UV2800_rest), and for each band we also provide unattenuated luminosity of the bulge (*_bulge, e.g. UV1500_rest_bulge) and with dust attenuation (*_dust, e.g. UV1500_rest_dust). The wide-field and ultradeep lightcones presented in Yung et al. (2022) share the exact same format and columns.

Category	Column	Quantity	Descriptions	Units
	0	halo_id_nbody	id copied from n-body simulations, unique with lightcone	
	1	gal_id	id assigned to galaxies, unique within each halo	
	2	gal_type	0 = central, 1 = satellite	
	3	z_nopec	cosmological redshift without peculiar velocity	
	4	redshift	cosmological redshift with peculiar velocity	
	5	ra	right ascension	Deg
	6	dec	declination	Deg
Halo Properties	7	m_vir	virial mass	10^{10} M$_\odot$
	8	V_vir	virial velocity	km s^{-1}
	9	r_vir	virial radius	Mpc
	10	c_NFW	concentration parameter in NFW profile	
	11	spin	spin parameter of halo	
	12	mstar_diffuse	mass of diffuse stellar population in halo	10^{10} M$_\odot$
	13	m_hot_halo	hot gas mass in halo	10^{10} M$_\odot$
	14	Z_hot_halo	hot gas metallicity in halo	Z$_\odot$
Galaxy Properties	15	v_disk	velocity of disk	km s^{-1}
	16	r_disk	radius of disk	kpc
	17	sigma_bulge	velocity dispersion of bulge	km s^{-1}
	18	rbulge	bulge radius	kpc
	19	mhalo	halo mass	10^{10} M$_\odot$
	20	mstar	stellar mass	10^{10} M$_\odot$
	21	mcold	cold gas mass	10^{10} M$_\odot$
	22	mbulge	bulge mass	10^{10} M$_\odot$
	23	mbh	black hole mass	10^{10} M$_\odot$
	24	maccdot	black hole accretion rate	M$_\odot$ yr^{-1}
	25	maccdot_radio	black hole accretion rate (radio mode)	M$_\odot$ yr^{-1}
	26	Zstar	stellar metallicity	Z$_\odot$
	27	Zcold	cold gas metallicity	Z$_\odot$
	28	mstardot	instantaneous SFR	M$_\odot$ yr^{-1}
	29	sfr_ave	SFR averaged over 100 Myr	M$_\odot$ yr^{-1}
	30	meanage	mass-weighted mean stellar age	Gyr
	31	tmerge	time elapsed since last merger	Gyr
	32	tmajmerge	time elapsed since last major merger	Gyr
	33	cosi	angle of inclination	
	34	tauV0	extinction optical depth in the *V*-band	
	35	maccdot_BH	black hole accretion rate	M$_\odot$ yr^{-1}
	36	sfr10myr	SFR averaged over 10 Myr	M$_\odot$ yr^{-1}
	37	mstarold	mass of stars older than 1 Gyr	Z$_\odot$
	38	ageold	mass-weighted mean age of stars older than 1 Gyr	Gyr
	39	zstarold	metallicity of stars older than 1 Gyr	Z$_\odot$
		filter set	available bands (all magnitudes are given in the AB system)	
Rest-frame photometry	40-48	{ }_rest	UV1500, UV2300, UV2800	
	49-72	{ }_rest	U, B, V, R, I, J, H, K	
Observed-frame photometry	73-78	galex_{ }	FUV, NUV	
	79-93	sdss_{ }	u, g, i, r, z	
	94-108	acs{ }	f435w, f606w, f775w, f814w, f850lp	
	109-123	wfc3{ }	f275w, f336w, f105w, f125w, f160w	
	133-141	UKIRT_{ }	J, H, K	
	142-147	irac_{ }	ch1, ch2	
	148-171	NIRCam_{ }	F070W, F090W, F115W, F150W, F200W, F277W, F356W, F444W	
	172-204	NIRCam_{ }	F140M, F162M, F182M, F210M, F250M, F335M, F360M, F410M F430M, F460M, F480M	
	205-216	Euclid_{ }	VIS, Y, J, H	
	217-240	Roman_{ }	F062, F087, F106, F129, F146, F158, F184, F213	
	241-258	LSST_{ }	u, g, r, i, z, y	
	259-276	DECam_{ }	u, g, r, i, z, Y	
	277-279	NEWFIRM_K_atm		
	280-294	VISTA_{ }	z, Y, J, H, Ks	

galaxies between $4.0 < z < 4.5$. Satellites are identified with `gal_id!=1` for the Santa Cruz SAM, `UPID!=-1` for UNIVERSEMACHINE, and `hostID!=-1` for DREaM. In Fig. C1, we show predicted UV luminosity functions and stellar mass functions for all galaxies and satellite galaxies presented in these lightcones. In the last panel, we also show the fraction of satellite galaxies as a function of stellar mass found in these models. The overall UV luminosity functions and stellar mass functions match fairly well between the three models (within the observational errors). The fraction of galaxies that are satellites as a function of stellar mass matches quite well between the Santa Cruz SAM and UNIVERSEMACHINE, while DREaM produces satellite fractions that are slightly lower at low masses.

In Fig. C2, we show histograms for the number counts of satellite galaxies, normalized to the volume of the lightcones, binned by their radial distances normalized to the virial radii of their host halos, $r/R_{\rm vir}$, for the Santa Cruz SAM and UNIVERSEMACHINE. These show reasonably good agreement.

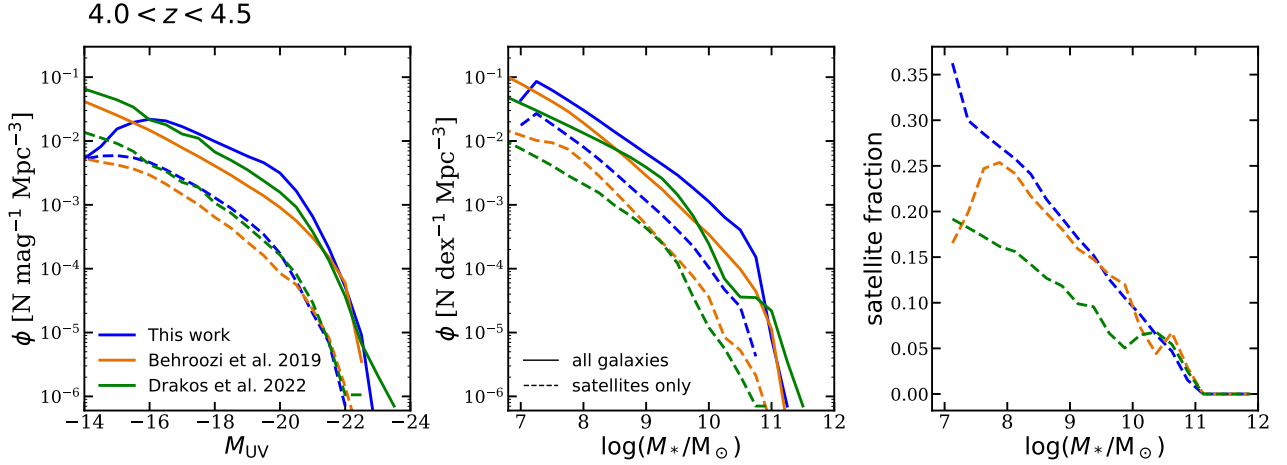This paper has been typeset from a TeX/LaTeX file prepared by the author.

Figure C1. UV luminosity functions (*left*) and stellar mass functions (*middle*) predicted from one of the 2-deg^2 lightcones, the DREaM mock catalogue, and a UNIVERSEMACHINE 2-deg^2 mock catalogue. We show distribution functions for all galaxies and satellite galaxies with solid and dashed lines, respectively. We also show the satellite fraction as a function of stellar mass (*right*).
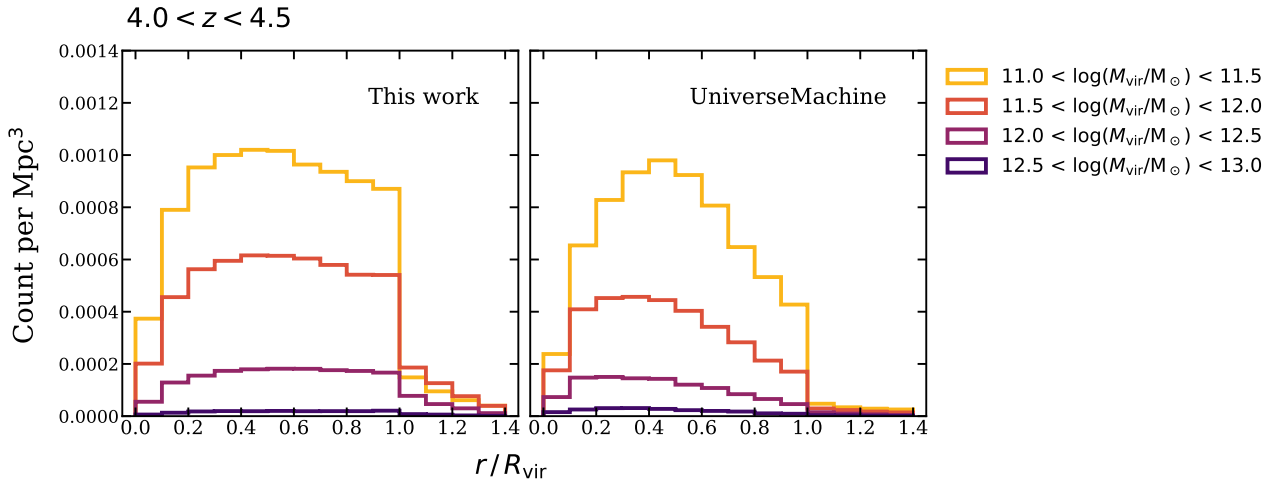

Figure C2. Histograms for the volume-averaged counts for predicted satellite galaxies binned by their radial distances normalized to the virial radii of their host halos, r/R_{vir}, in one of the 2-deg^2 lightcones presented in this work (*left*) and in the UNIVERSEMACHINE GOODS-S lightcone (*right*) in virial mass bins with $\Delta \log(M_{vir}/M_\odot) = 0.5$ for satellite galaxies encompassed in halos with mass $\log(M_{vir}/M_\odot) = 11.0$ to 13.0.