Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **523,** 3874–3884 (2023) https://doi.org/10.1093/mnras/stad1652
Advance Access publication 2023 June 2

A synthetic *Roman Space Telescope* High-Latitude Time-Domain Survey: supernovae in the deep field

Kevin X. Wang,[1] Dan Scolnic,[1] M. A. Troxel,[1★] Steven A. Rodney,[2] Brodie Popovic,[1] Caleb Duff,[2] Alexei V. Filippenko,[3] Ryan J. Foley,[4] Rebekah Hounsell,[5,6] Saurabh W. Jha,[7] David O. Jones,[4†] Bhavin A. Joshi,[8] Heyang Long,[9] Phillip Macias,[4] Adam G. Riess,[8,10] Benjamin M. Rose[1] and Masaya Yamamoto[1]

[1]*Department of Physics, Duke University, Durham, NC 27708, USA*
[2]*Department of Physics and Astronomy, University of South Carolina, SC 29208, USA*
[3]*Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA*
[4]*Department of Astronomy & Astrophysics, University of California, Santa Cruz, CA 95064, USA*
[5]*University of Maryland, Baltimore County, Baltimore, MD 21250, USA*
[6]*NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA*
[7]*Department of Physics and Astronomy, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA*
[8]*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*
[9]*Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[10]*Space Telescope Science Institute, Baltimore, MD 21218, USA*

Accepted 2023 May 30. Received 2023 May 30; in original form 2022 April 18

ABSTRACT

NASA will launch the *Nancy Grace Roman Space Telescope* (hereafter *Roman*) in the second half of this decade, which will allow for a generation-defining measurement of dark energy through multiple probes, including Type Ia supernovae (SNe Ia). To improve decisions on survey strategy, we have created the first simulations of realistic *Roman* images that include artificial SNe Ia injected as point sources in the images. Our analysis combines work done on *Roman* simulations for weak gravitational lensing studies as well as catalogue-level simulations of SN Ia samples. We have created a time series of images over 2 yr containing ∼1050 SNe Ia, covering a 1 deg^2 subarea of a planned 5 deg^2 deep survey. We have released these images publicly for community use along with input catalogues of all injected sources. We create secondary products from these images by generating coadded images and demonstrating recovery of transient sources using image subtraction. We perform first-use analyses on these images in order to measure galaxy detection efficiency, point source detection efficiency, and host-galaxy association biases. The simulated images can be found here at https://roman.ipac.caltech.edu/sims/SN_Survey_Image_sim.html.

Key words: software: simulations – transients: supernovae.

1 INTRODUCTION

The *Nancy Grace Roman Space Telescope* (hereafter *Roman*) is a NASA mission that will make advances into a wide range of astrophysics, and should be able to make an unprecedented measurement of the nature of dark energy with multiple probes (Spergel et al. 2015). The project is currently in Key Design Phase C (final design and fabrication) and most of the specifications are now set.[1] The 2.4 m space-based telescope is optimized for observations in the near-infrared (NIR), with a Wide Field Instrument (WFI) that can accommodate eight filters spanning a wavelength range of 0.48–2.3 μm and has a large field of view (FoV) of 0.281 deg^2.

Because of its location at L2, large aperture size, and FoV, *Roman* will be extremely well suited for studies of cosmological growth and expansion on large scales. The extragalactic survey will be divided into two main parts, with a wide-area High-Latitude Survey (HLS) that will enable studies of weak lensing and galaxy clustering, as well as a smaller area High-Latitude Time-Domain Survey. As originally formulated in the *Roman* Science Requirement Document,[2] the top goal of the time-domain survey is to study the expansion history of the Universe from redshift $z \approx 0.1$ to $z \approx 2$. Current samples of Type Ia supernovae (SNe Ia) used for cosmological analyses (Scolnic et al. 2021) contain of the order of 1500 objects, though the number of SNe Ia with well-sampled light curves (cadence better than once every 10 rest-frame days per filter) above $z > 1$ is only ∼20. *Roman* has the ability to discover and measure over 10 000 SNe Ia, with a significant fraction of these at $z > 1$ (Hounsell et al. 2018). In this work, we create the first simulated images of the time-domain survey to better prepare for the future mission.

★ E-mail: michael.troxel@duke.edu
† NASA Einstein Fellow.
[1]https://roman.gsfc.nasa.gov/science/WFI_technical.html
[2]https://science.nasa.gov/science-pink/s3fs-public/atoms/files/Kruk_APAC_Oct2018_r2.pdf

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

Table 1. Detector effects; they are discussed in more detail by Troxel et al. (2021). GALSIM is open-source software used to simulate images (Rowe et al. 2015) and is discussed in Section 2.2.2.

Effect	Description	Relevant functions within the GALSIM module
Background light	Zodiacal and stray light and thermal emission of the telescope	Uses roman.getSkyLevel
Photon count noise	Poisson noise on photon counts	Uses noise.PoissonNoise
Reciprocity failure	Response to high flux is larger than that to low flux	Uses roman.roman_detectors.addReciprocityFailure
Electron quantization	Electron count is rounded to nearest integer	Uses detectors.quantize
Dark current	Poisson noise caused by thermal electrons	Uses DeviateNoise on PoissonDeviate
Classical non-linearity	Non-linear charge-to-voltage response	Uses detectors.applyNonlinearity with roman.NLfunc
Interpixel capacitance	Pixels are affected by charge on neighbouring pixels	Uses roman.applyIPC
Read noise	Noise from electronics as charge is converted to digital values	Gaussian noise with standard deviation from roman.read_noise
e to ADU	Electron counts are converted to ADUs	Currently applies gain of 1 since actual gain is not yet known
ADU quantization	ADU rounded to nearest integer	Uses detectors.quantize

Precision measurements of SNe Ia across this redshift range are critical for understanding the equation-of-state parameter for dark energy, $w = P/\rho$, where P is the pressure and ρ is the energy density. A common parametrization of w combines a constant with a term that depends linearly on the scale factor (a), such that $w = w_0 + w_a(1 - a)$ (Linder 2003). The recent Pantheon + analysis of a composite sample containing ∼1500 SNe Ia found $w_0 = -0.841^{+0.066}_{-0.061}$ and $w_a = -0.65^{+0.28}_{-0.32}$ (after including constraints from measurements of the cosmic microwave background and baryon acoustic oscillations), which shows that currently there is only a very weak constraint on the evolution of dark energy (Brout et al. 2022). While current systematic constraints are of comparable magnitude to the statistical constraints, both can be improved with a larger sample and redshift range.

Although at present there is not a single defined survey strategy for the SN key project in terms of cadence, depth, filter choices, etc., we rely here on the initial study done by Hounsell et al. (2018, hereafter H18) that evaluated multiple survey strategies and performed catalogue-level simulations with SNANA (Kessler et al. 2009) to predict final constraints on dark-energy parameters from the programmes in terms of the figure of merit, first introduced for dark-energy studies by the Dark Energy Task Force (Albrecht et al. 2006). For simplicity in this analysis, we choose a single, representative survey strategy from those presented by H18. With the image simulations presented here, we can measure detection efficiencies and other survey characteristics that can be used for the less computationally expensive catalogue-level simulations.

To enable the image simulations, we rely on the work done by Troxel et al. (2021, hereafter T21), which created realistic simulated *Roman* images that covered 6 deg^2 following a reference survey design for the High-Latitude Imaging Survey (HLIS). They simulated multiple effects that can cause systematic bias in cosmological analysis, such as background light, Poisson noise, reciprocity failure, dark current, classical non-linearity, and interpixel capacitance. These detector effects are described in detail by T21 and are also summarized in Table 1. They include both stars and galaxies in their images. In this study, we can use the same architecture for our simulated images where we overlay SNe Ia as stellar point sources on top of galaxy images and analyse time series of images.

Artificial SN injection has been used in many studies (Pain et al. 1996; Sullivan et al. 2006), including by the Dark Energy Survey (DES; Kessler et al. 2019) and LSST DESC (LSST Dark Energy Science Collaboration 2020). As discussed by Kessler et al. (2019), DES used artificial sources to track failures of their image-processing pipeline, to characterize their efficiency, and finally to trace biases in recovered photometry all the way to biases in the recovered cosmology (Brout et al. 2019). LSST DESC has similar goals with their simulated sky survey called 'Data Challenge 2' (DC2; LSST Dark Energy Science Collaboration 2020) and the artificial transients have been first analysed by Sánchez et al. (2021). In this paper, we utilize infrastructure for characterizing the relation between SN Ia properties and host-galaxy properties added as part of the SNANA package (Popovic, Scolnic & Kessler 2020).

There are a number of time-domain analyses that can be done with a large image simulation, including tests of the impact of saturation on the bright end and recovery near the cores of galaxies. We provide the code used to perform the analysis, which may be used as a base for more complicated tests.[3] One such example is the work in Rubin et al. (2021), which establishes a forward-modelling photometry pipeline to be used on *Roman* images.

The overview of this paper is as follows. In Section 2, we describe our methodology for constructing simulated images. Section 3 presents the details of generating and including SNe Ia in images. We provide analyses in Section 4, such as measuring how often an SN Ia is likely to be mistakenly associated with the wrong host galaxy and measuring the SN Ia detection efficiencies as a function of magnitude. The details about processing images in preparation for analysis, particularly for creating difference images and detecting point sources, are also provided in this section. In Section 5, we discuss the implications of this work for design of the *Roman* SN Ia cosmology survey, and describe future work.

2 METHODS

The flowchart in Fig. 1 shows an overview of the image-creation process, which will be described in more detail in this section.

2.1 Roman characteristics

The photometric system and other characteristics of the telescope are based on *Roman* Cycle 7 specifications. We briefly summarize useful characteristics here. There are eight imaging filters of the WFI: R062, Z087, Y106, J129, H158, F184, K213, and W149 (a very wide filter). These filters will respectively be abbreviated in the

[3] https://github.com/KevinXWang613/SNAnalysisCode

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

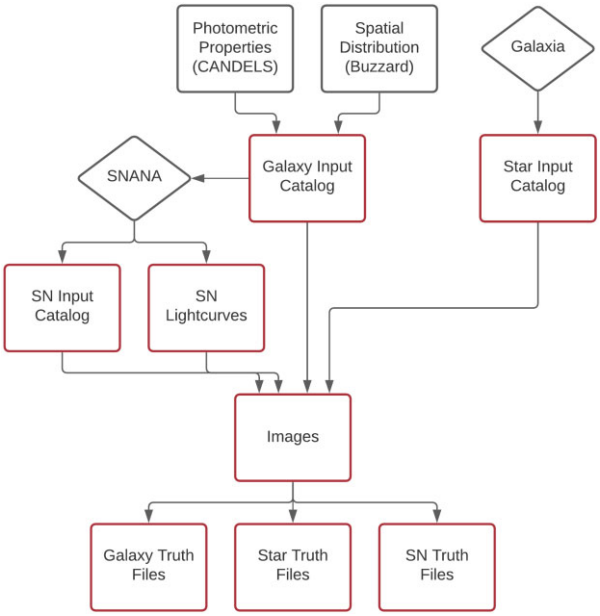

Figure 1. Flowchart depicting the process through which output products are created. Boxes labelled in red are the output files that are included in the data release.

text as *R*, *Z*, *Y*, *J*, *H*, *F*, *K*, and *W*. In μm, the central wavelengths of these filters are, (respectively) 0.620, 0.869, 1.060, 1.293, 1.577, 1.842, 2.125, and 1.464, and the widths in μm of the transmission bands are 0.280, 0.217, 0.265, 0.323, 0.394, 0.317, 0.35, and 1.030, in the same order.[4] The spatial resolution of the imaging component of the WFI is $\sim 0.11''$ pixel^{-1}, and each HgCdTe sensor chip array (SCA) image has a usable pixel grid of 4088 × 4088. The read noise is Gaussian with a standard deviation of 8.5e^{-}.

A total of 18 SCAs make up the *Roman* camera, arranged into a 6 × 3 array to generate an effective FoV of 0.281 deg2. The fill factor across the FoV is ∼88 per cent, owing to gaps between the SCAs. The simple pointing strategy we employ here (described below in Section 2.3) does not include a dither strategy to deal with these chip gaps. Bright objects will further limit the usable field, as twelve diffraction spikes due to the *Roman* secondary mirror support struts will be visible for most stars.

2.2 Image simulations from T21

The simulation engine used for this work is the publicly available code developed for the *Roman* HLIS, as described by T21.[5] Here, we explain the object creation (galaxies/stars) and image creation separately.

2.2.1 Object creation: galaxies and stars

The properties of our simulated galaxies are summarized in Table 2. As shown there, we use three main inputs to generate astrophysical objects for the simulated images: the spatial distribution of galaxies (from the Buzzard simulation), stellar population properties for the galaxies assigned to each position (from 3DHST catalogues), and foreground stars and dust [from the Galaxia Milky Way (MW) simulation].

The input galaxy distribution catalogue is taken from one realization of the Buzzard simulation (DeRose et al. 2019), to introduce realistic spatial clustering. The Buzzard simulation provides the angular coordinates on the sky for ∼2.7 million galaxies, across the full 5 deg^2 of our simulated *Roman* deep field. Each position is then assigned a random set of galaxy properties drawn from the galaxy-properties catalogue, which is based on the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS).

We do not use the same CANDELS catalogue as T21; instead, our galaxy properties are derived from a catalogue of observed galaxies from the CANDELS/3DHST fields (Brammer et al. 2012). This catalogue is a subset of the v4.1 catalogue produced by the 3DHST team (Momcheva et al. 2016), which contains ∼208 000 objects. Our version of the catalogue has been modified in two key ways. First, we have selected the subset of objects that are identified as galaxies and have estimates for redshift, stellar mass, and star formation rate (SFR), and also have a well-defined spectral energy distribution (SED) model. These selection cuts leave the sample with ∼112 000 galaxies, which we use as an input library. Secondly, we provide coefficients from Brammer et al. (2008) that can be used for future studies to create simulated spectra for any galaxy in the catalogue from template SEDs provided with Brammer et al. (2008) that will match the colours/properties of the CANDELS catalogue. As we take galaxy colours directly from observed galaxies, host-galaxy extinction is already applied to the simulated galaxies.

The redshift and mass estimates are described by Skelton et al. (2014). For the input redshift in our simulations, we adopt the spectroscopic redshift z from the CANDELS catalogue when available, and the photometric redshift otherwise. The mass estimates (M, in units of $\log_{10}(M/\mathrm{M}_\odot)$) are derived from fits to the observed SEDs using the FAST code (Kriek et al. 2018), and then corrected for overestimation of broad-band flux due to emission-line contamination, following Whitaker et al. (2014). We adopt the SFRs (in units of $(\log_{10}\,\mathrm{M}_\odot)\,\mathrm{yr}^{-1}$) derived from the combined ultraviolet and NIR flux measurements, as used by Whitaker et al. (2014). Fig. 2 shows the distribution of these galaxies across z, M, and SFR. Finally, we generate photometric magnitudes for each galaxy by integrating the SED of each 3DHST galaxy, multiplied by the transmission functions of the *Roman* bands used in the deep-field survey (*YJHF*).[6] For this purpose, we use SED fits generated by the 3DHST team with the EAZY code (Brammer et al. 2008).

As summarized in Table 2, the following true object properties are assigned to each simulated galaxy: (1) redshift; (2) sky position in right ascension (RA) and declination (Dec) from the Buzzard catalogue; (3) photometric properties (consistent *YJHF* magnitudes); (4) base full width at half-maximum (FWHM) size as circular object; (5) intrinsic shape components (defined using the distortion definition $|e| = (a^2 - b^2)/(a^2 + b^2)$, $e_1 = |e| \cos 2\beta$, $e_2 = |e| \sin 2\beta$) drawn from a Gaussian distribution of width 0.27 (truncated at ±0.7 to prevent impossible shapes); (6) random rotation angle; (7) ratio of fluxes in each of the three galaxy components: (a) de Vaucouleurs bulge, f_b, (b) exponential disc, f_d, and (c) star-forming knots (at most 25 per cent of disc flux), f_k – created using GALSIM's RandomKnots function; (8) stellar-population parameters (mass, SFR).

Each galaxy model is built with GALSIM from the truth parameters for the object in each *Roman* bandpass. The models are drawn in

[4] https://roman.gsfc.nasa.gov/science/WFI_technical.html
[5] https://github.com/matroxel/roman_imsim
[6] See https://roman.gsfc.nasa.gov/science/Roman_Reference_Information.html

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

Table 2. Properties of simulated objects in the Galaxy and star input catalogues. All properties are discussed in Section 2.2.1.

Category	Property	Source	Reference
Location	RA, Dec	Buzzard	DeRose et al. (2019)
Redshift and photometry	z, Y, J, H, F mags, FWHM size	3DHST SEDs +EAZY fits	Skelton et al. (2014) Brammer, van Dokkum & Coppi (2008)
Stellar population	SFR, M	3DHST catalogues	Skelton et al. (2014), Whitaker et al. (2014), Momcheva et al. (2016)
Morphology	a, b, $\theta_{\rm rot}$, $f_{\rm b}$, $f_{\rm d}$, $f_{\rm k}$	Random draws	
MW properties	RA, Dec, and Y-F of MW stars, MW extinction	Galaxia	Sharma et al. (2011)

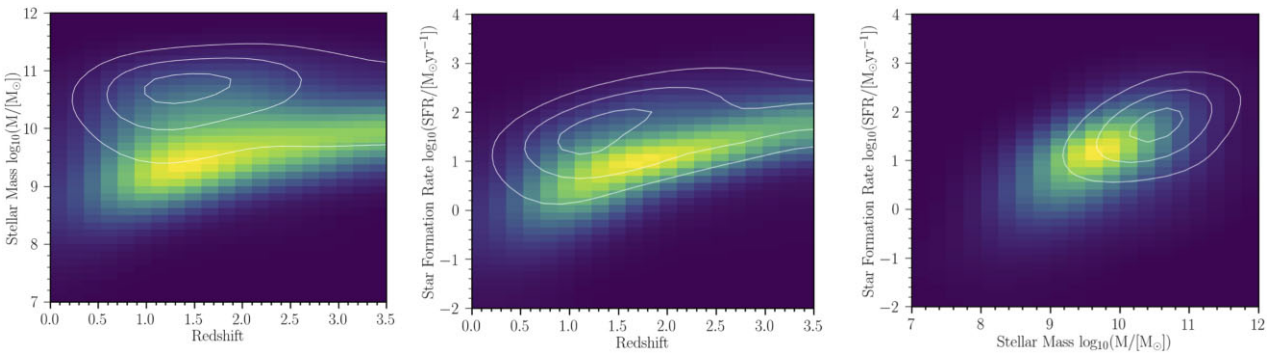

Figure 2. Stellar population properties of the galaxy catalogue ($z < 3.5$) used as input for *Roman* High-Latitude Time-Domain Survey simulations. Shading in each panel shows the distribution of the galaxy catalogue with logarithmic scaling. Overlaid contours show a representative sampling of SN Ia host galaxies, drawn according to the specific-SFR (sSFR)-based model from Andersen & Hjorth (2018). Contours mark 30 per cent, 60 per cent, and 90 per cent levels. Panel (a) presents the stellar mass versus redshift distribution; (b) shows SFR versus redshift; and (c) plots SFR against stellar mass. The SN Ia host-galaxy population is strongly skewed to higher mass and more actively star-forming galaxies at all redshifts.

dynamically sized square arrays of pixels (hereafter referred to as stamps). The size of each stamp is defined to include at least 99.5 per cent of the flux.

Following T21, we simulate the positions and magnitudes of input stars in the MW in *Roman* passbands using the galaxy simulation Galaxia (Sharma et al. 2011). Stars are simulated to mag 27 in the V band, and the fluxes are converted to that measured with *Roman* bandpasses using the stellar SED of α Lyrae derived from *Hubble Space Telescope (HST)* CALSPEC as packaged with GALSIM. The star distribution has a mean stellar density of $\sim$2.5 arcmin^{-2}. Extinction is added according to the Galaxia simulation using a formula based on the Schlegel, Finkbeiner & Davis (1998) extinction maps.

2.2.2 *Image creation*

To create the image, the focal plane of the telescope is arranged at a given RA, Dec location with a given position angle. Each SCA has a fixed relationship with the focal plane, allowing its footprint to be determined from the information about the centre point. Each image consists of a single SCA and is made up of 4088 $\times$ 4088 pixels. We identify the subset of galaxies and stars from the input catalogues that fall within that footprint, build a model for each object, and add them as stamps on to the SCA image. A cutout of one of the images is shown in Fig. 3.

T21 simulated images over 6 deg^2 in the H158 filter but did not include transient sources and did not structure those simulations for time-domain analysis. In this analysis, we focus more on the time component in order to build a temporal series of images, and also inject additional SNe Ia as star-like point sources on to the images. Still, the architecture of the simulations is the same, and we make top-level modifications to build the High-Latitude Time-Domain Survey.

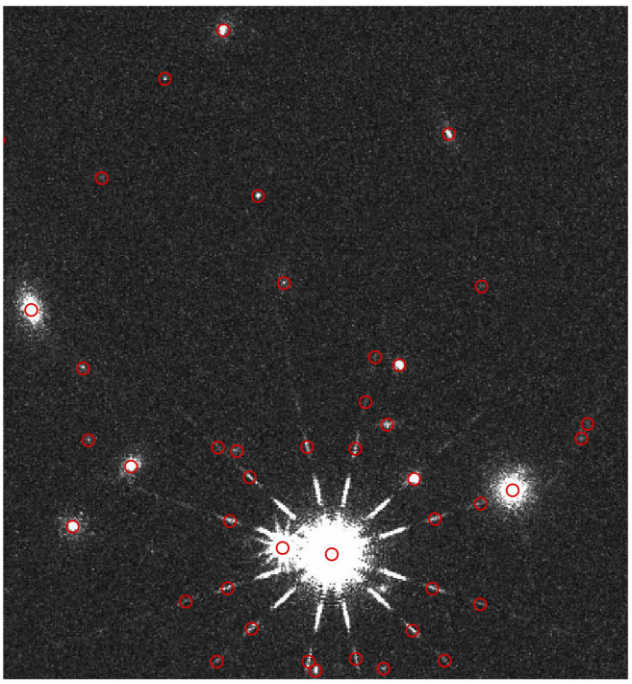

Figure 3. An example cutout ($0.75' \times 0.75'$) of a simulated image of one SCA that is $7.5' \times 7.5'$ across. The red circles show objects detected by Source Extractor that pass a signal-to-noise ratio (S/N) threshold. Diffraction spikes of the stars can be seen.

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

The images in the simulation are rendered using the GALSIM software package (Rowe et al. 2015). We use the galsim.roman module in GALSIM release v2.3.2., which has a number of *Roman*-specific implementation details (Kannawadi et al. 2016).[7]

Each image is processed through a series of modifications, simulating the physical effects that would happen in the detector such as dark current, detector non-linearity, and read noise. We do not include the effects of persistence between images. The details of these steps are given by T21. This module includes a high-resolution image of the *Roman* spider pattern (i.e. the obscuration by the struts and camera in the pupil plane) for each filter and SCA that the point-spread function (PSF) model is generated from. Due to limitations in processing speed building the PSF model from this highly sampled pupil plane mask, we downgrade the original pupil plane mask image to a coarser binning (a downsampling of factor 8 for all objects but the brightest stars) to produce a fast, but still sufficiently accurate, approximation of the correct PSF model. For the brightest saturated stars, this produces visual artefacts in the simulated diffraction, and so for visual purposes we use a smaller downsampling factor: a binning of 4 or 2 pixels is used for stars brighter than magnitude 15 and 12, respectively.

2.3 Specific survey strategy

H18 proposed various strategies for the *Roman* extragalactic survey. They created a realistic catalogue-level simulation of the survey, for the history of observations as well as the library of detected SN light curves. The inputs toward SNANA that we utilize here are the sequence of date and filter of each observation.

We choose to follow the imaging: High-z strategy from H18 with updates described in a new reference survey by Rose et al. (2021). The strategy has a deep tier and a medium/wide tier, which covers 5 deg^2 and 21 deg^2, respectively. For this analysis, we focus solely on the deep tier owing to computational limitations. The survey in the deep tier uses filters *YJHF* (0.7–2.0 μm), and the exposure times are 300, 300, 300, and 900 s for those filters, respectively. The number of SNe Ia predicted by this strategy is ∼8000 using cut-offs for generation at $0.05 < z < 3.0$. No SNR cuts are applied at this stage.

For the creation of simulated images, we specify additional details of the observing strategy. We choose the position to be in the Akari Deep Field South so that it is in a continuous viewing zone for longer observing seasons; the centre position of the fields are chosen at 71° RA, −53° Dec. The total survey spans 2 yr with an observation cadence of 5 d, such that there are 146 visits to the SN field. H18 did not need to specify the exact slewing strategy, as it was assumed that all SNe could be discovered in the survey area. Here, we detail the chosen observation sequence for a 5 deg^2 field, which is illustrated in Fig. 4.

(i) One orient of the telescope angle each month, and the angle is rotated by 30 degree every 30 d.

(ii) For a given orient, there are 16 individual pointings that tile the field. The tiling is done such that there is one central row of 6 pointings and two parallel rows of 5 pointings.

(iii) Pointings are chosen to minimize gaps and the well-covered area is 5 deg^2, whereas the largest extent of any of the chips is 8 deg^2.

(iv) The sequence of pointings goes from south-east to north-east to north to south to south-west to north-west, using a 'snake' pattern that minimizes distance between successive pointings (see animated version).

[7]https://github.com/GalSim-developers/GalSim

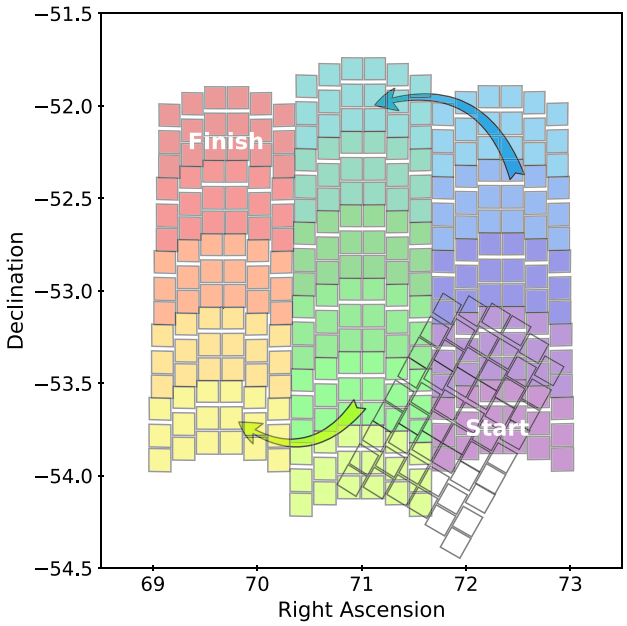

Figure 4. A portion of the pointing strategy. The pointing sequence for a single visit begins in the lower right (the purple WFI footprint, labelled 'Start') and proceeds up and down the columns as indicated by the arrows, ending in the upper left (red footprint labelled 'Finish'). The subsequent visit is rotated by 30 degree, and the first three pointings of that visit are shown as unfilled regions in the lower right. An animation of this pointing strategy is available in the online version.

(v) Within each column, the distance between pointings is ∼24′, and the movement between columns is ∼49′.

(vi) The pointing sequence is done in one filter, and once completed, the filter is switched and the sequence is done backward until each filter has been used.

For each 'snake' pattern, the areas covered by SCAs do not overlap, so each SN Ia will show up only in one SCA.

3 INCLUDING SUPERNOVAE IN IMAGES

Following H18, we use the same SN Ia rates (SNRs) model (Graur et al. 2014; Rodney et al. 2014), light-curve model (Guy et al. 2010), SN Ia population characteristics (Scolnic & Kessler 2016), and area covered. In order to assign SNe Ia to host galaxies, we follow the process outlined by Popovic et al. (2020), defining properties of each simulated SN Ia based on the properties of its simulated host galaxy. This allows us to propagate host-galaxy selection biases as well as specific correlations between SN Ia properties and host-galaxy properties which can be used in a broader cosmological analysis of simulated data. Additionally, we vary the SNRs based on the host-galaxy properties, as discussed below.

3.1 Matching galaxy catalogues to SN Ia catalogues and injection of SNe Ia on to images

To project SNRs on to galaxy properties, we use the piecewise linear model from Andersen & Hjorth (2018) which relates sSFR to specific SNR (sSNR). We note that the SNRs prescribed here are relative at a given redshift, and the overall SNRs with redshift have no galaxy dependence as given in Graur et al. (2014) and Rodney et al. (2014). The relations are shown in Fig. 5. Here, the

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

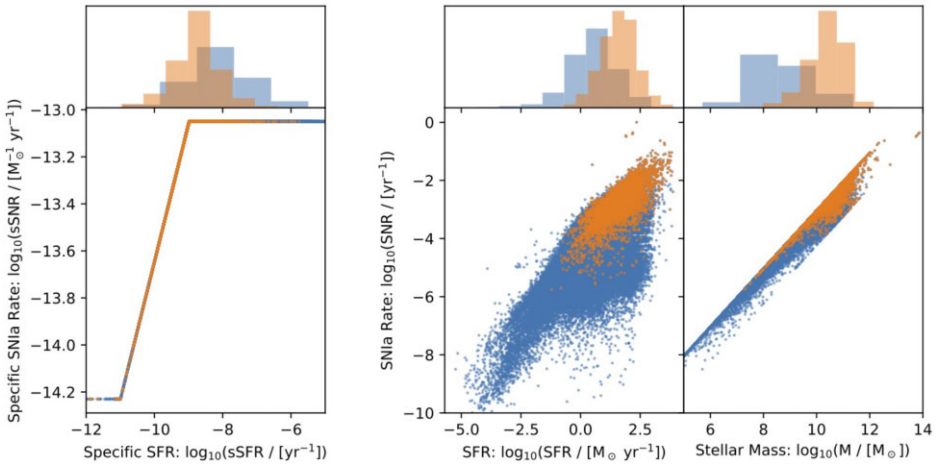

Figure 5. Application of an SNR model to the input galaxy catalogue. Each lower panel plots all 112 000 galaxies in blue, overlaid with a selection of 5000 'SN Ia host galaxies' in orange. Smaller panels above show normalized histograms for both populations. The host-galaxy selection process follows the Andersen & Hjorth (2018) sSFR-based model to assign an SNR prediction to each galaxy in the catalogue. That model is illustrated in the left pair of panels, plotting the sSNR against the sSFR, with logarithmic scaling on both axes. The two pairs of panels on the right show the total SNR for every galaxy, computed by multiplying the sSNR by the total stellar mass. The middle panels plot this SNR against the SFR. The far right-hand panels plot SNR against the stellar mass.

piecewise function is used so that there are not unphysically high SNRs from galaxies that are outside of the range $-12 < \mathrm{sSFR} < -8$, given as $\log_{10}(M_{star}\, M_{\mathrm{galaxy}}^{-1})\,\mathrm{yr}^{-1}$, where we have well-calibrated SNR models.

Multiplying the sSNR by the total stellar mass, we derive the absolute SNR of each galaxy in the input catalogue (see the far right-hand panels in Fig. 5). We then normalize these SNR values to get the relative probability of hosting an SN Ia for each galaxy.

The next step is to match simulated SNe Ia with appropriate hosts. This begins with SNANA generating a set of simulated SNe Ia with redshifts in the range $0.06 < z < 2.99$. For each SN Ia, we then identify all galaxies within $\Delta z \pm 0.01$ as the subset of potential host galaxies. The redshift distributions of galaxies and SNe Ia are shown in Fig. 6. It is important to note that an SN Ia is generated before assigning it to a host galaxy, which has the consequence that the redshift distribution of generated SNe will not match the redshift distribution of the galaxies.

A single galaxy is drawn from that subset, using the SNR probabilities as weights. The resulting distribution of stellar mass and SFR for host galaxies is plotted as contours in Fig. 2. Selected host galaxies have on average higher stellar mass and SFR than galaxies overall. We also plot the magnitudes of the resulting host galaxies against redshift as contours in Fig. 6, which can be compared with the same relationship in all galaxies. The selection process using the weights causes host galaxies to be brighter on average than galaxies overall at any given redshift, an effect that can be seen directly in the images.

SNANA assigns an RA and Dec to the SN Ia using the selected host galaxy's intensity profile. The SN Ia is placed within the region that encloses 99 per cent of the host-galaxy flux density, weighted by the galaxy's Sersic profile (so the SN Ia distribution traces the host-galaxy light profile). Finally, the output from the SNANA simulation is a library of light curves for SNe Ia at specific redshifts and positions based on the host galaxy, as well as peak brightness date along with various other properties relevant to SNe Ia.

We inject the SNe Ia as point sources in the exact way that stars are included in the images. The flux of each SN Ia is defined per filter, in AB magnitudes. For our image simulations, we use the simulated magnitudes without noise added by SNANA because the image simulations contain their own noise. SNANA produces a truth file for each SN Ia light curve which can be used in downstream analyses.

The time of observation defined by SNANA is not identical to the time of observation for any given simulated image, owing to the slewing sequence. Therefore, we interpolate the SN Ia light curve to determine the flux used when injecting each SN Ia into an image. However, for this image set, the largest difference between the times created by SNANA and observation dates is 1/20 d, so there should not be a large effect.

In Fig. 7, we give an example time series of an SN Ia at redshift $z = 0.38$ injected near a galaxy in the Y band. The changing brightness of the SN Ia can be seen to match the variation in brightness from the generated light curve.

3.2 Image data products

The image data, galaxy input catalogues, and SN Ia light-curve catalogues are the three main products generated. These are publicly available.[8]

The images are created in a Flexible Image Transport System (FITS) format[9] and contain keyword information in their headers to make them compatible with *HST* MULTIDRIZZLE software. The headers are created using GALSIM's World Coordinate System (WCS) code and have WCS information in them. For each pointing, 18 separate image files are created, one for each SCA of the WFI.

While there are unchanging input catalogues for galaxy, star, and SN Ia photometry, an additional truth file is created for each image containing every galaxy drawn on the image, containing each object's ID from the original input catalogue, as well as their positions (both world coordinates and pixel positions). A similar truth file for stars and SNe Ia is created. We also include the magnitude drawn on the image in the truth file for SNe Ia.

[8] https://roman.ipac.caltech.edu/sims/SN_Survey_Image_sim.html
[9] https://fits.gsfc.nasa.gov/

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

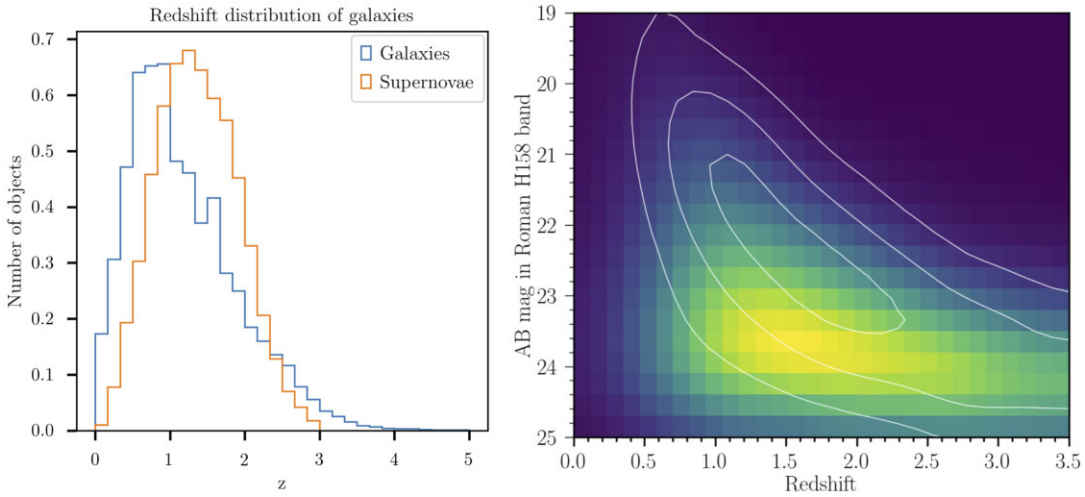

Figure 6. *Left*: The redshift distribution from the galaxy input catalogue (blue) that is used as input for the image simulations and for the library of host galaxies of the SNe Ia. The redshift distribution of the SNe Ia (orange) is capped at $z = 3$ because limits on detection efficiency mean it is not useful to generate higher redshift SNe Ia. *Right*: Distribution of H158 magnitudes versus redshift. Coloured shading shows all galaxies in the input galaxy simulation catalogue. White contours show a subset of galaxies selected as SN Ia hosts, based on the projected SNR in each galaxy. Contours mark 30 per cent, 60 per cent, and 90 per cent levels. The host galaxies are brighter on average than the overall galaxy population at every redshift.

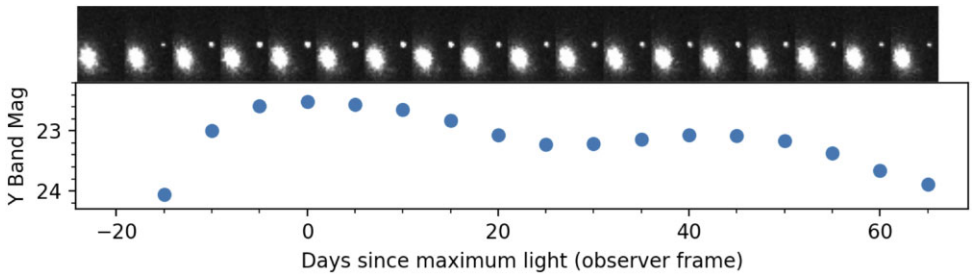

Figure 7. Simulated time series of an SN Ia at redshift $z = 0.38$ in intervals of five days with corresponding light curve. Magnitudes are ideal values before noise and error are added. In each of the images, the SN Ia is on the right with its host galaxy on the left. The images and light curve are aligned so that the points on the light curve match the image at the corresponding time.

4 ANALYSIS EXAMPLES

In this section, we describe a series of analyses that we did using our simulated images and the tools we used for each analysis.

4.1 Source detection and S/N of sources

We use Source Extractor (Bertin & Arnouts 1996) to detect objects on the images. Source Extractor allows for the automated detection of objects from images and outputs estimated properties for each of those objects, including position, shape, S/N, and other properties. It can be downloaded from its GitHub page.[10]

Fig. 3 shows one of our images, where we have highlighted objects detected by Source Extractor. We use default parameters for Source Extractor on our images except for the detection threshold (which is different between single images and coadds).

Using the images, we can also compare the S/N of point sources in the image to the S/N used in SNANA for catalogue-level simulations. We use Source Extractor to obtain the S/N for stars in the images by taking the inverse of the RMS uncertainty in magnitude for each star using a 5 pixel aperture diameter. We measure S/N for stars because they are more numerous than the SNe Ia and are not transient. However, because this does not include the contribution of galaxy noise, which is accounted for in the SNANA simulations, we add an approximate model as predicted by SNANA for a self-consistent comparison (on the level of 4 mmag). The results of the comparison are shown for *Y*-band images in Fig. 8. We find good agreement in S/N between the images and catalogue-level simulation across the entire magnitude range.

4.2 Stacking images and SN Ia host-galaxy association

Several processing steps need to be done in order to reach the full potential use of the images taken by *Roman*, including coadding images, discussed in this subsection, and difference imaging, discussed in the next. We have set up the process so that we can run the coaddition on an entire set of images, although there are currently still limitations in our mock coadd pipeline that limit how many images can be coadded together simultaneously.[11]

[10] https://github.com/astromatic/SExtractor

[11] Due to storage issues (how many images can be mapped pixel-by-pixel and stored in a 64-bit integer image bit-mask plane) as well as memory limitations when processing in parallel the coadd stage of the pipeline, we limit the number of overlapping images chosen to be coadded to 64. This is

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

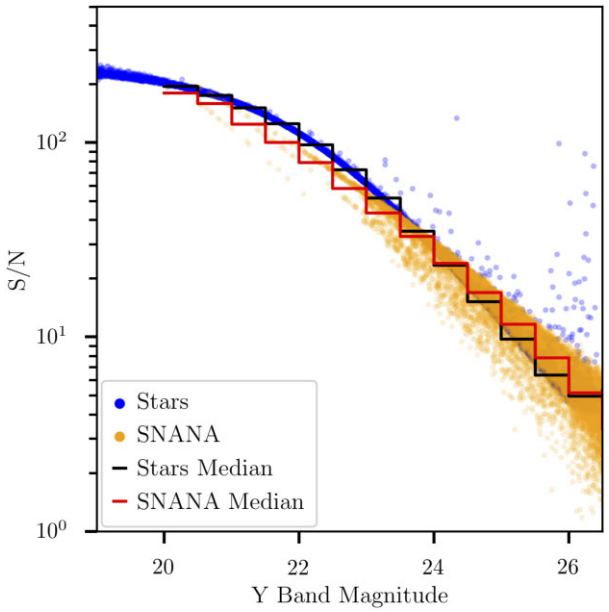

Figure 8. A comparison of S/N for stars detected that were injected in *Y* images (blue) and from our catalogue-level simulations with SNANA (orange). Objects are also binned by *Y*-band magnitude, and the median S/N per bin is shown in black for stars and red from SNANA. For SNANA, the magnitudes are simulated magnitudes, while for stars, the magnitudes are measured magnitudes from Source Extractor. The small number of stars (in blue) with high SNR at faint magnitudes are due to blending with nearby sources.

For these steps, we use the sndrizpipe data-processing pipeline,[12] a PYTHON implementation for *HST* data processing designed for *HST* SN surveys. The sndrizpipe package was originally developed for detection of SNe in the CANDELS and FrontierSN *HST* surveys (see Rodney et al. 2015a, b). It is built on the DRIZZLEPAC toolkit, a package of software tools developed by the Space Telescope Science Institute (STScI) for processing of data from *HST* and other observatories (Fruchter & Hook 2002).

In applying the sndrizpipe tools to the simulated *Roman* images, we first select a 'WCS reference image' – an arbitrarily selected image to define the WCS to which all other images will be registered. Next, we use the TweakReg tool to construct a catalogue of sources (stars and galaxies) in each individual exposure image and match them with the WCS reference image's source catalogue.

We can then combine all the images using the MULTIDRIZZLE algorithm implemented as the ASTRODRIZZLE PYTHON module.[13] We adopt an output pixel scale of 0.0575″ and set the ASTRODRIZZLE pixfrac parameter to 0.7. For now, we combine a limited set of images such that for the resulting coadded image, 10–15 single images contribute to most of its pixels. The current image simulations do not include cosmic rays, and so cosmic ray correction in the Drizzle process is turned off. Realistically simulating cosmic ray impacts would require simulating up-the-ramp fitting in the SCA measurements, which has not yet been implemented in these image simulations.

In the right-hand side of Fig. 9, we show an example of a stacked *Roman* image in *Y* band. A single image for the same location is shown for comparison. The increase in depth can be clearly seen both in the noise and as objects detected (marked in cyan) which are not in the single image.

We performed a basic galaxy detection efficiency analysis with both single images and the stacked images described above in *Y*. A simple classifier is used, where we consider an object to be detected if Source Extractor detects an object within 0.3″ of the truth value of its position. The results shown in the left-hand side of Fig. 10 display the increase in depth obtained from stacking images. For single images, the 50 per cent recovery threshold is around 24.9 mag. For coadded images, the 50 per cent recovery threshold is ∼26.0 mag, giving us about 1 mag of depth over the single images. The CANDELS catalogue goes above 30 mag, so we expect some fraction of galaxies to remain undetected. Although we do not expect to detect faint (e.g. 29 mag) objects, we do have a small percentage of false positives as a result of detections of unmasked features in the stellar diffraction spikes and of misidentification of overlapping detected bright galaxies near a faint background galaxy. This is due to the use of a simple matching algorithm by position on the sky, which is sufficient for the main results of the paper.

The same analysis can be performed on stars as well and is shown in the right-hand side of Fig. 10. Because stars and SNe Ia are both drawn as point sources on our images, these detection rates serve as useful proxies for potential SN Ia detection rates. As point sources, stars are easier to detect than galaxies, resulting in 50 per cent recovery at 26.1 and 27.3 mag for single images and for coadded images, respectively. Note that unlike stars, SNe will not be detected in full-stack coadds because they will have faded away in some images.

The *Roman* High-Latitude Time-Domain Survey will rely on obtaining redshifts of the host galaxies to use SNe Ia on a Hubble diagram. As discussed by Gupta et al. (2016) and Popovic et al. (2020) with regards to DES and Sloan Digital Sky Survey (SDSS), respectively, there may be misassignment of the true host galaxy of an SN. This can cause outliers on the Hubble diagram that cannot always be removed, and therefore cause a cosmological bias. This may be especially problematic for *Roman* owing to its large redshift range where fainter hosts will not be detected, as well as the possibility of blending (Melchior et al. 2018) of galaxies over the images.

What the image set allows us that cannot be done at the catalogue level is determining the galaxy detection limit, as some galaxies will be too faint to be detected with our algorithms. SNe Ia that belong to these galaxies may appear 'hostless' (Sullivan et al. 2006) or be assigned to a different galaxy.

Following Popovic et al. (2020), we estimate the rate of misassociation using our coadded images. Popovic et al. (2020) did this at the catalogue level by assigning the host based on the separation and properties of nearby galaxies. We tested a similar procedure here by using the Source Extractor isophote parameters to define the elliptical effective radius with our injected artificial sources on real images. Then, we follow the directional light radius method of Sullivan et al. (2006) to assign host galaxies, which uses the distance from the galaxy to the centre, as well as the orientation and shape of the galaxy. However, we found that simply using the distance between the SN Ia and galaxy position gave better association results for our images.

The results of the host-galaxy association analysis are shown in Fig. 11. There were 489 host galaxies in one sample of images. 411 host galaxies are detected and correctly identified, as represented by the blue points. Eight host galaxies are detected but are not the closest galaxies, potentially leading to misassociation; these

not a fundamental limit for *Roman*, but rather a limit of the current pipeline that was developed for this and related work using these image simulations.
[12]https://github.com/srodney/sndrizpipe
[13]https://github.com/spacetelescope/drizzlepac

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

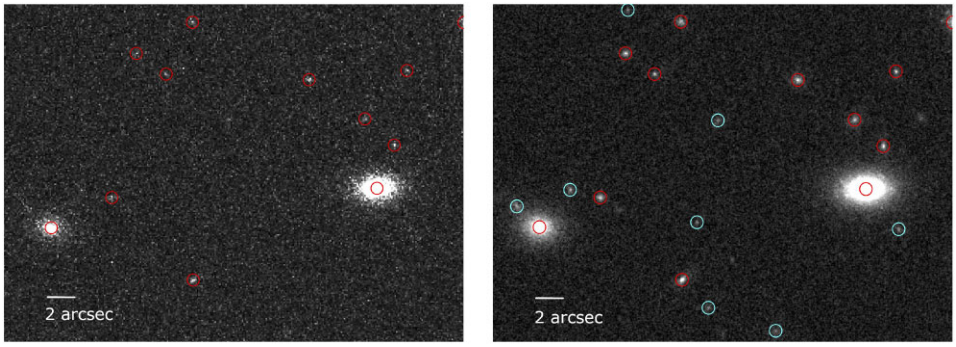

Figure 9. *Left*: A single image with objects detected using Source Extractor circled in red. *Right*: The same area in the sky in a coadded image made by combining 64 images. The actual pixel depth is 10–15 images for most areas on the coadd. Objects found only in the coadded image are circled in cyan while objects found in the coadded image that were also found in the single image are circled in red. There are no SNe Ia in this image.

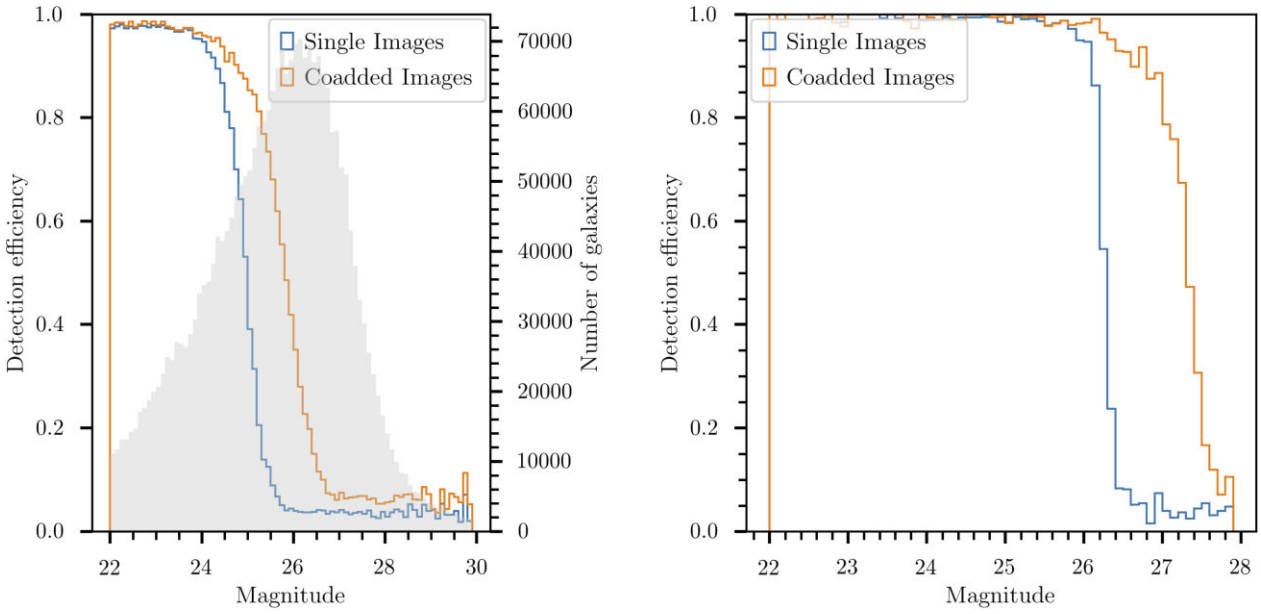

Figure 10. *Left*: A comparison of galaxy detection efficiency in both single images (blue) and coadded images (orange). A histogram of the magnitudes of the galaxies is shown in filled grey. *Right*: A comparison of star detection efficiency in both single images (blue) and coadded images (orange).

are shown as orange points. Furthermore, 70 host galaxies are not detected, which could also lead to misassociation. The redshift of the closest detected galaxy to the SN Ia position is given in green. The set of SNe Ia (and thus host galaxies) span a wide redshift range. At lower redshifts, performance is much better, with only 2 out of 98 galaxies below $z = 1$ and 8 out of 234 galaxies below $z = 1.5$ not detected.

A more sophisticated analysis could use the redshifts of the SN Ia and of the suspected host galaxies to exclude obvious redshift mismatches between the SN Ia and suspected host galaxy. If we assume that redshift differences greater than 0.2 would be detectable as outliers, indicated by the red lines in Fig. 11, we would correctly associate the eight misassociated galaxies mentioned above. Finally, we note that this number of SNe without identified hosts will be underestimated slightly if SNe are located in galaxies that are too faint to be included in the CANDELS catalogue used to generate galaxy properties; we note though that as shown in Fig. 5, we expect very few SNe to be in these types of galaxies, as the SFR and stellar mass are so low.

4.3 Difference imaging and validating simulated SN Ia detection efficiencies

Given a template image, we can subtract it from each of the individual epochs. As is typically done for space-based transient surveys, we do not include any convolution or other blurring of the image to match the PSF from epoch to epoch, because the PSF is presumed to be very stable across the time span of the imaging series (as discussed in T21). While there will be small variations in the PSF over time, mostly correlated with vibrational modes from movement of the telescope, we ignore these effects in this version of the simulation. The subtraction stage is therefore just a simple arithmetic operation at the end of the sndrizpipe pipeline. The result is a series of difference images in which the galaxy light has mostly been removed, and we are left with isolated point sources (the simulated SNe Ia) and residual noise. Fig. 12 shows a cutout of the observed image and difference image, centred on the location of an SN, which demonstrates that difference-imaging routines can work with these images.

As we do not have a full-scale pipeline to run this difference imaging on our set of simulated images, we work around that

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

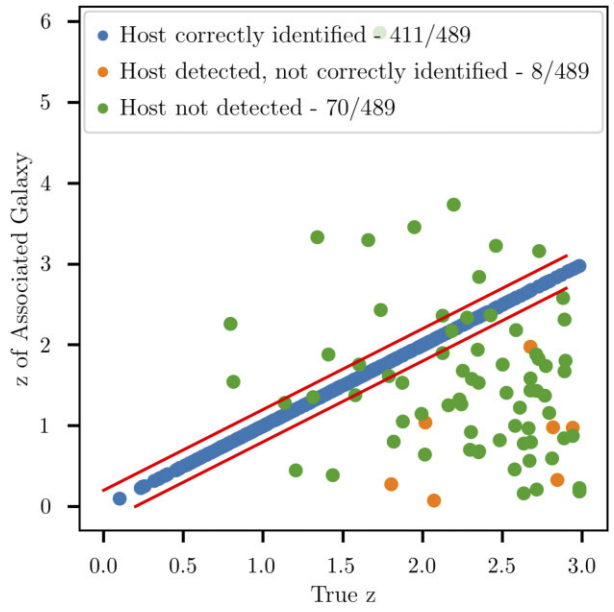

Figure 11. A scatterplot of the true host-galaxy redshift and the redshift of the closest galaxy detected from each SN. Blue shows correct assignments, orange indicates when the host is detected but another galaxy is assigned as the host, and green signifies when the host is not detected and another galaxy is assigned as host.

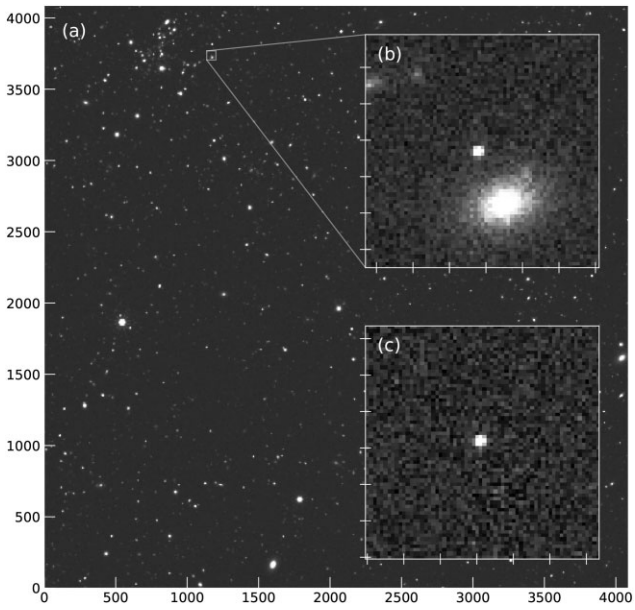

Figure 12. A simulated *Y*-band image for one *Roman* SCA. (a) The full frame of the simulated image, comprising 4088 × 4088 pixels, as indicated by the axis labels (roughly 7.5 arcmin on a side). (b) Zoom-in on a 10 arcsec × 10 arcsec region, centred on a bright SN Ia with magnitude 22.59. (c) The difference image of the same 10 arcsec × 10 arcsec region, generated by subtracting a single-epoch template image taken after the SN Ia has faded, with the same rotation and pointing centre.

limitation by analysing only SNe Ia that are isolated from their host galaxies. To ensure separation, we place a separation cut at 1.5 times the galaxy's semimajor axis and a host-galaxy brightness cut at 22 mag. We find that we can recover 50 per cent of the SNe Ia around an SN Ia magnitude of 24.5 mag and 0 per cent at 25 mag. While the results showcase the need for proper difference imaging, we are able to detect SNe Ia in our images, and the dropoff of efficiency for fainter magnitudes is expected owing to lower S/N. For future SNANA catalogue-level simulations, we can use the detection efficiency determined directly from the images, as there are likely features of the image analysis that cannot be fully modelled at the catalogue level.

5 CONCLUSIONS

This work is part of a larger ecosystem of simulations and analyses that will be used to optimize and prepare for the *Roman* High-Latitude Time-Domain Survey. One of the key elements of this work is that the catalogue-level simulations are explicitly tied to the image-level simulations. Follow-up analyses, like pixel-level simulations of prism spectroscopy with accurate SN Ia host SEDs, can utilize the images created here.

One main follow-up programme is on improving the difference-imaging pipeline. In recent surveys, typically two pipelines are written – one for fast turnaround of transient detection and another with more accurate control of systematic uncertainties. Examples of the former typically use image convolution and examples of the latter typically involve the scene-modelling approach. One particularly interesting analysis would be to examine whether the 'surface-brightness anomaly', which shows unaccounted scatter in the uncertainty estimates of SN Ia flux near bright galaxies, is seen in the *Roman* image differences (Kessler et al. 2015). A first use case would be to run the forward-modelling pipeline from Rubin et al. (2021) on the images created here, as that study has indicated the anomaly may not be present for photometry of *Roman* images.

Another follow-up program is to create joint simulations of LSST and *Roman* images. There is currently focus on joint processing for weak-lensing studies, but a similar effort would benefit SN Ia studies owing to increasing the wavelength and depth range. Studies of this kind could be particularly fruitful for transient detection probabilities that are highly dependent on pixel-level effects or non-standard image processing – e.g. transients near the cores of bright galaxies (including lensed SNe, nuclear transients, and quasars), as well as fast transients of various types, or very rare and faint transients like caustic-crossing stars and Population III SNe. These future studies can leverage recent data releases with more NIR coverage of transients for better modelling of the transients (e.g. Jones et al. 2022).

In summary, we have presented new simulated images for the *Roman Space Telescope* that contain injected SNe Ia. These images could be used to look at effects such as detection efficiency near bright galaxies as well as to calibrate assumptions used for catalogue-level simulations. We perform a variety of basic checks and analyses to demonstrate that SNe Ia were correctly drawn in the images and that the images can be useful for the aforementioned analyses. As a product of this analysis, we release 1 deg^2 of images over the full *Roman* High-Latitude Time-Domain Survey and present these at https://roman.ipac.caltech.edu/sims/SN_Survey_Image_sim.html. We encourage the community to use these images, and the code developed here can be further used for more simulations of different surveys/areas as well as studies of different transients.

ACKNOWLEDGEMENTS

Much of this work was supported by NASA under contract no. NNG17PX03C issued through the WFIRST Science Investigation Teams Programme. This work is based on an earlier work done by

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

the *Roman* Cosmology with the High-Latitude Survey Science Investigation Team (https://www.roman-hls-cosmology.space/), which was supported by NASA grant no. 15-WFIRST15-0008. DS was supported by DOE grant no. DE-SC0010007 and the David and Lucile Packard Foundation. MAT acknowledges support from NASA under JPL Contract Task 70-711320, 'Maximizing Science Exploitation of Simulated Cosmological Survey Data Across Surveys'. AVF is grateful for financial support from the Christopher R. Redlich Fund and many individual donors. This work is based on observations taken by the 3D-HST Treasury Program (GO 12177 and 12328) with the NASA/ESA HST, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract no. NAS5-26555. The material is based upon work supported by NASA under award no. 80GSFC21M0002. Support for DOJ was provided by NASA through the NASA Hubble Fellowship grant no. HF2-51462.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract no. NAS5-26555.

DATA AVAILABILITY

As stated in the paper, the simulated images are publicly released along with supporting files.[14] The code used to create the images is also publicly available,[15] as is the code used for the analysis.[16]

REFERENCES

Albrecht A. et al., 2006, preprint (arxiv:astro-ph/0609591)
Andersen P., Hjorth J., 2018, MNRAS, 480, 68
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Brammer G. B. et al., 2012, ApJS, 200, 13
Brout D. et al., 2019, ApJ, 874, 150
Brout D. et al., 2022, ApJ, 938, 110
DeRose J. et al., 2019, preprint (arXiv:1901.02401)
Fruchter A. S., Hook R. N., 2002, PASP, 114, 144
Graur O. et al., 2014, ApJ, 783, 28
Gupta R. R. et al., 2016, AJ, 152, 154
Guy J. et al., 2010, A&A, 523, A7
Hounsell R. et al., 2018, ApJ, 867, 23
Jones D. O. et al., 2022, ApJ, 933, 172
Kannawadi A., Shapiro C. A., Mandelbaum R., Hirata C. M., Kruk J. W., Rhodes J. D., 2016, PASP, 128, 095001
Kessler R. et al., 2009, PASP, 121, 1028
Kessler R. et al., 2015, AJ, 150, 172
Kessler R. et al., 2019, MNRAS, 485, 1171
Kriek M. et al., 2018, Astrophysics Source Code Library, record ascl:1803.008
Linder E. V., 2003, Phys. Rev. Lett., 90, 091301
LSST Dark Energy Science Collaboration, 2020, ApJS, 253, 31
Melchior P., Moolekamp F., Jerdee M., Armstrong R., Sun A. L., Bosch J., Lupton R., 2018, Astron. Comput., 24, 129
Momcheva I. G. et al., 2016, ApJS, 225, 27
Pain R. et al., 1996, ApJ, 473, 356
Popovic B., Scolnic D., Kessler R., 2020, ApJ, 890, 172
Rodney S. A. et al., 2014, AJ, 148, 13
Rodney S. A. et al., 2015a, AJ, 150, 156
Rodney S. A. et al., 2015b, ApJ, 811, 70
Rose B. M. et al., 2021, preprint (arXiv:2111.03081)
Rowe B. T. P. et al., 2015, Astron. Comput., 10, 121
Rubin D., Cikota A., Aldering G., Fruchter A., Perlmutter S., Sako M., 2021, PASP, 133, 064001
Sánchez B. et al., 2021, ApJ, 934, 96
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Scolnic D., Kessler R., 2016, ApJ, 822, L35
Scolnic D. et al., 2021, ApJ, 938, 113
Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, ApJ, 730, 3
Skelton R. E. et al., 2014, ApJS, 214, 24
Spergel D. et al., 2015, preprint (arXiv:1503.03757)
Sullivan M. et al., 2006, ApJ, 648, 868
Troxel M. A. et al., 2021, MNRAS, 501, 2044
Whitaker K. E. et al., 2014, ApJ, 795, 104

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.

[14] https://roman.ipac.caltech.edu/sims/SN_Survey_Image_sim.html
[15] https://github.com/matroxel/roman_imsim
[16] https://github.com/KevinXWang613/SNAnalysisCode

Downloaded from https://academic.oup.com/mnras/article/523/3/3874/7189910 by US Department of Engergy user on 27 August 2024

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society