

Documentation GRS simulation products

Anahita Alavi, James Colbert

April 2022

1 Simulation of Grism Data

(Section 1 is from Wang et al. 2022)

We have developed a pipeline to simulate grism spectroscopic observations for the *Roman* HLSS. *Roman* grism enables spectroscopy at $R = 460\lambda/\mu\text{m}$ for $\lambda = 1 - 1.93\mu\text{m}$. Using the pipeline, we have produced simulated data products that are representative of the planned HLSS data, for use by the Science Investigation Team (SIT) and the community. Fig.1.3 illustrates an outline of the pipeline and products.

The simulations cover an area of 4 deg^2 over the redshift range of $0 < z < 3$. By incorporating the survey parameters of the planned HLSS, such as detection limits, exposure times, roll angles, and dithering, the simulation products have been created to resemble as closely as possible the future observations. In addition to the grism slitless spectroscopy, we also simulate the direct imaging that is used to locate the spectra.

1.1 Motivation behind the grism simulation

Slitless grism spectroscopic surveys, including HLSS, can suffer from incompleteness and low purity for a variety of reasons. Sources that we fail to detect because their fluxes are close to the detection limit of the survey will result in incompleteness. Furthermore, some sources may be missed due to their flux and/or emission lines blending with those of nearby objects. In addition to completeness, the fraction of sources with correctly identified emission lines, known as purity, is also a crucial parameter in galaxy redshift surveys. We have designed the mock slitless spectroscopy observations so that they can be used to measure the completeness and purity for the *Roman* HLSS (see HLSS6 science requirements in Wang-2022).

1.2 Simulation inputs

To create a realistic simulation of the sky, the pipeline uses input galaxies and stars with a wide variety of physical properties. We use the aXeSIM software (Kuemmel et al. 2007) to generate a synthetic grism spectrum for each source. AXeSIM is a simulation software package developed in support of *Hubble Space Telescope* slitless spectroscopy modes. It can, however, also simulate spectral images for other slitless spectroscopic instruments, including the *Roman* Wide Field Instrument (WFI). AXeSIM requires tables of sources with information about their coordinates, brightness, shapes, redshifts

and spectra. In addition, aXeSIM requires configuration files describing the trace and dispersion solutions for the instrument.

Here we describe the assumptions we made for the creation of the input tables of sources and more details on how we simulate the grism spectra.

- Galaxies: We use the *Roman* HLSS 4 deg² galaxy mock catalog with SEDs as the input catalog, from which we select galaxies spanning the redshift range of 0 to 3. The galaxy mock catalog is simulated using the semi-analytic model (SAM) of galaxy formation, Galacticus (Benson 2012), which adds galaxies to cosmological dark matter only N-body simulations. We refer the reader to Wang et al. (2022) for a detailed description of this mock catalog.

The input galaxy catalog includes galaxy coordinates, sizes, brightness, and redshifts – we use these to create the galaxy object tables that later become the inputs to the grism simulation machinery.

- Coordinates : The sky coordinates (Ra,Dec) in the input catalog are converted to the detector coordinates (X,Y) on each Sensor Chip Assembly (SCA).
- Inclination angle: We assume random inclination angle for each galaxy.
- Size: The input catalog provides galaxy sizes in terms of disk and spheroidal radius. These sizes, along with the random inclination angle, are used to calculate the galaxy morphology, in terms of semi-major (a) and semi-minor (b) axes. To avoid including very thin galaxies in our sample, we limited our sources to having a major to minor axis ratio $a/b = 1-4$ (e.g., see Fig. 11 in odewahn-1997). It is noteworthy that we had to rescale the sizes to improve agreement with observations. Using data from the WFC3 Infrared Spectroscopic Parallel survey (WISP, atek-2010,Colbert-2013), we examined the size distribution of galaxies in the galaxy mock catalog. We compared the normalized distributions of the semi-major sizes of mock galaxy catalog and WISP galaxies. To make the comparison, we only used mock galaxies down to the same magnitude limit as the WISP sub-sample (i.e., H-band = 25 AB magnitude). Although the normalized size distribution of the mock catalog has the same peak as the WISP galaxies, the mock catalog contains more large galaxies. In order to make the distributions match, we need to rescale the size of large mock galaxies, semi-major > 2.5 pixel (i.e., 0.275), by a factor of 1/2. We apply the same rescaling factor to the semi-minor sizes in the mock galaxy catalog. Incorporating these rescaled sizes, we then use two-dimensional Gaussian profiles as image templates to model the light profiles of galaxies in our final images.
- Brightness: We apply a magnitude cut of $m = 28$ AB magnitude in the *Roman* H-band filter (F158). This value corresponds to the 5σ point source detection limit for a 1 hour exposure time in the H-band direct imaging.
- SED: In addition to the galaxy properties, we also use the SEDs that accompany the input catalog, as the spectral template for each object. We

refer the reader to Sec 3.2 in Yun et al. (2022) for a detailed description of how these SEDs and their various components, continuum and emission lines, are constructed. We truncate the SEDs to span an observed wavelength range of $\lambda = 10000 - 20000 \text{ \AA}$ to match the wavelength range of *Roman* WFI grism observations.

- Stars: For a realistic sky scene, we also incorporate stars in our simulation. We use the model presented in chang-2010 for the number of stars per unit of apparent magnitude, m , within a certain solid angle ($\Delta\Omega$) as follows:

$$\phi(m) = \Delta\Omega \int \psi(M, r) n(r) r^2 dr \quad (1)$$

where $\psi(M, r)$ is the local stellar luminosity function taken from Just-2015. The stellar density profile, $n(r)$, is based on a model presented in Juric-2008 as the sum over disk and spheroidal halos. It is a good approximation to neglect the angular variation in stellar density over 4 deg^2 . Note that we only use a single stellar type of G0V for the stellar spectral templates for simplicity. In order to model starlight profiles, we use PSF images as the image templates. Models of the PSFs are constructed using the WebbPSF Python package from STScI perrin-2014, which transforms models of the telescope and instrument optical state into PSFs taking into account detector pixel scales, rotations, filter profiles, and point source spectra. In order to optimize both the code run time and the final spatial resolution simultaneously, we made the PSFs with an oversample factor of 2 inside a $364 \times 364 \text{ pixel}^2$ box.

- Roll angle and dither: In HLSS, spectra will be obtained at different roll angles. This will help resolve possible ambiguity in the correct source to associate with pixels illuminated by overlapping spectra, and thus mitigate the incompleteness caused by spectral confusion and emission lines obscured by nearby neighbors. The simulated products assume 4 different roll angles: 0, 5, 170, and 175 degrees. In addition, to cover the gap between the detectors, we include a two point dither pattern, $\Delta X(\Delta Y) = 0$ & $\Delta X(\Delta Y) = 1/4$ SCA size (where "SCA" denotes the area covered by one H4RG detector), at each roll angle.
- Background and exposure time: For the grism observation background we use a single value of 0.57 e/sec/pix , based on a zodiacal background calculation representative of the HLSS sky coverage. The direct imaging filter has a narrower bandpass, so for the direct image simulations we use a lower value of 0.25 e/sec/pix for the background. Note that these are typical values and that the sky may vary within the HLSS area. For the simulated grism and direct imaging exposures, we use 301 sec and 141 sec exposure times, respectively, which are the current survey parameters of the Reference HLSS.

1.3 Simulation overview

To run the full simulation, we first divide the 4 deg^2 area into 15 regions each covering, at most, one *Roman* field of view. The pipeline simulates one of these regions at a

time, with all WFI 18 detectors in their proper configurations. The pipeline first builds a large input object table for each region that includes all the sources contained within that region. Source coordinates in this input table are in the focal plane local coordinate system of the WFI. Next, the pipeline creates 18 smaller lists associated with 18 detectors and converts the source coordinates to the local detector coordinate system ((X,Y) in SCA coordinates, see above). The input list for each individual detector is big enough that it will not miss any sources even when the detector FOV is rotated.

The pipeline then runs aXeSIM to simulate the images and spectra. We note that aXeSIM has some limitations that we had to solve for these simulations. First, aXeSIM runs one detector at a time. Therefore, we parallelize the pipeline so that all 18 detectors run simultaneously. Second, aXeSIM disperses light only along the x-axis. However, *Roman* WFI images will be dispersed along the y-axis. To overcome this limitation, we combine a set of 90 degree rotations in both the input files and final products. Third, aXeSIM does not account for instrument distortion from sky coordinates to detector coordinates. We add distortion to the X-Y pixel coordinates in our input object tables. We warn, however, that the final output images do not have the correct WCS coordinates. In order to determine the correct sky position of each source, the user should refer to the input catalog.

The final simulation data products include essentially all of the effects that need to be considered in evaluating plans to extract spectra. However, a few limitations were necessary.

- In the final simulation products, all detector-level artifacts have already been removed. Therefore, the images do not contain cosmic rays, bad pixels, saturation, or dark currents. Additionally, the products are already flat-fielded.
- These simulations do not include the 0th and 2nd order spectra, the wavelength dependent PSF, or changes to sky background within the HLSS coverage.
- In these simulations, we had to exclude some very narrow ranges of redshifts from our final products. The WFI sensitivity curve that we used for the simulation goes slightly bluer (9000 Å) than the blue wavelength cut of our spectral templates at 10000 Å. As aXeSIM extrapolates the SEDs to the blue edge of the sensitivity curve, we would see unrealistic broad spectral features for galaxies with strong emission lines that are redshifted to 10000 Å. Thus, we had to exclude such redshifts from our samples. These narrow redshift ranges are $0.517 < z < 0.531$, $0.993 < z < 0.997$, $1.012 < z < 1.016$, $1.053 < z < 1.057$ and $1.676 < z < 1.687$.

Our final products will be publicly released on the IPAC website.¹ The full 4 deg² simulation products are provided in four different roll angles and two different dithering patterns. Each pair of roll angle and dithering includes 18 simulated direct images and 18 simulated 2D slitless dispersed images (i.e., grism spectra). Fig.1.3 displays an example of our simulated direct and dispersed images for detector number 1 (i.e., SCA1) at two different roll angles of 0 and 5 degree.

¹<https://roman.ipac.caltech.edu/>

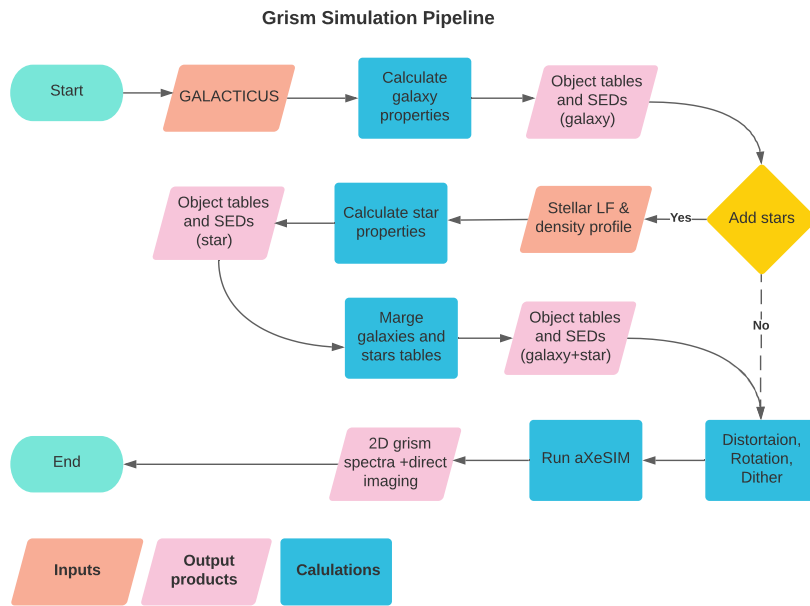


Figure 1: Illustration of the steps in our *Roman* grism observation simulations. In this flow chart, we display the input catalogs and output products with orange and pink parallelograms, respectively. We also represent various calculation processes in blue rectangles. As described in Section 1, in addition to the input galaxy mock catalog produced using Galacticus, the pipeline allows the addition of stars to the simulated images.

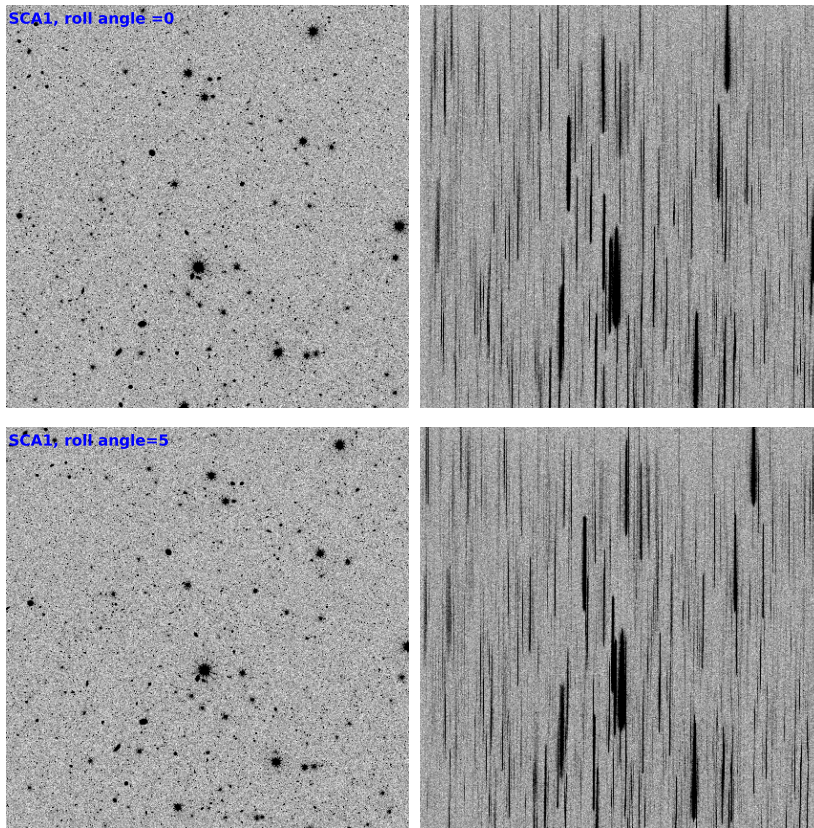


Figure 2: An example of simulation final products. This example displays a direct image on the left and a 2D slitless dispersed image on the right for detector number 1. The images are shown for two roll angles of 0 and 5 degree.

2 Simulation products

Simulation products are organized into 15 separate directories (i.e., named as FOV0, FOV1, ..., FOV14), each corresponding to a separate region covering, at most, one Roman field of view (see above). Figure 1 illustrates these 15 regions, which cover a total area of 4 deg². For each FOV, there are 4 sub-directories, each representing a roll angle of 0, 5, 170, and 175 degrees. It should be noted that these roll angles are measured around the central reference point of the focal plane. Additionally, each directory representing a roll angle has two sub-directories for two dithers of $\Delta X(\Delta Y) = 0$ & $\Delta X(\Delta Y) = 1022$ pixels. For each set of roll angle and dither, we release input catalogs, input SED templates, simulated direct images, and simulated grism spectra separately for 18 detectors. Below is a detailed description of these products.

IMPORTANT: For some FOVs, there are some detector sub-folders (e.g., for a specific roll angle and dither) that are empty and do not contain any simulated images. It means that no sources have been placed on that detector at those specific roll angles and dithers.

2.1 Input catalogs

Input catalogs are lists of all sources (e.g., galaxies and stars) and their properties that have been fed into the simulation pipeline (e.g., aXeSIM). The flowchart in Fig. 1 illustrates how the input catalog, or as we refer to it, the object table, is generated in the first half of the pipeline (top three rows in Fig. 1) and later is used as an input to aXeSIM. For each roll angle and dither pattern, we provide 18 input object tables for 18 detectors including all the sources that lie within the detector field of view. The object table for each detector is stored in ASCII format under that detector sub-directory. These input object tables follow the same format as the Model Object Table (MOT files) for the aXeSIM software (see section 5.8 in the aXeSIM manual). The following columns are included in our MOT files.

- **NUMBER:** This is an integer number assigned to each source and it will be referred as source ID. For each FOV directory, every source has the same ID regardless of the roll angle or dither pattern. As a result, these source IDs can be used to match sources from images with different roll angles and dithers.
- **X(Y)_IMAGE:** The x and y coordinates of each source in the local coordinates system of each detector. These x and y coordinates vary within the range of [0,4088].
- **A(B)_IMAGE:** The semi-major (minor) axis of each source in pixels.
- **THETA_IMAGE:** The position angle of the major axis. This angle is measured in a counter-clockwise direction from the y-axis. As discussed above, we assign random position angles to our input sources.
- **MAG_F158W :** An AB magnitude measured in the F158 Roman filter.

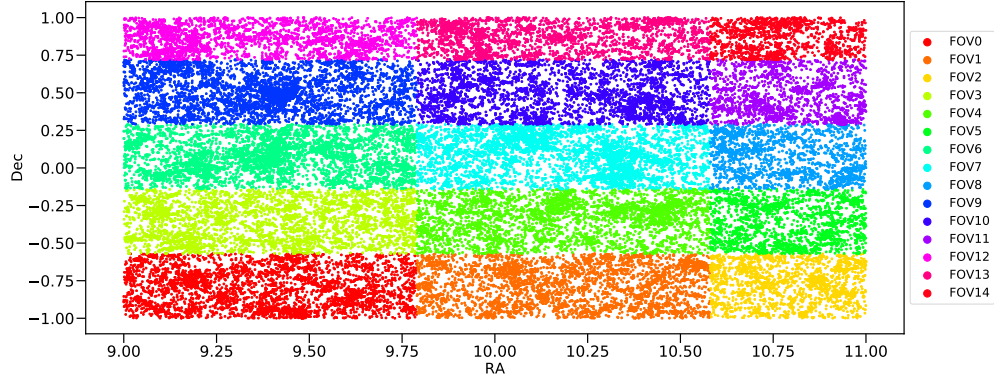


Figure 3: This figure illustrates the full 4 deg^2 area from GALACTICUS mock catalog included in our grism simulations. Each color specifies a separate region covering one *Roman* field of view. Each point in this plot shows a galaxy from our input mock catalog. We only include 1% of the sources in the full mock catalog to make this plot.

- **SPECTEMP**: A spectroscopic ID assigned to each source to identify its SED template. To facilitate source-SED matching, this spectroscopic template ID matches the source ID (i.e., column NUMBER).
- **Z**: The redshift of the source. We note that this column is set to zero for stars.
- **MODIMAGE**: This column specifies which image template should be used for each source. For galaxies where we use a two-dimensional Gaussian profile, this value is set to 0. For stars where we use PSF image to model the starlight profile, this value is set to 1. We note that we also include our PSF images with our data releases.

All of these columns were used as input to the aXeSIM software to simulate the images. However, there are two additional columns in our MOT files that have **not** been used by aXeSIM. These are sky coordinates of the sources (i.e., RA and DEC in deg). These sky coordinates correspond to the source position in the GALACTICUS simulation.

IMPORTANT: Axesim fails to transfer WCS information to output simulated image headers. We manually update some of the WCS information, including CRVAL, CRPIX, and CD matrix keywords, in the header of our output images. However, we note that our final simulated images do not contain the correct RA and DEC coordinates for each source (i.e., distortion). **One should use the input catalog, not the WCS position of each pixel as derived from the WCS in the header, to determine the distortion free WCS location of the objects. Users should only use the X and Y coordinates to match the sources in the catalogs and the images.**

2.2 Input SEDs

For each FOV directory, we also include a folder of SEDs for all sources found in that FOV and used in the simulation. Users can match sources and their SED files by using the source ID provided in the MOT catalog and also used in the SED file name. These SEDs are given in $\text{erg}/\text{sec}/\text{cm}^2/\text{\AA}$ units. Wavelengths are given in \AA .

IMPORTANT: We note that these rest-frame SEDs are dust attenuated assuming a Calzetti attenuation curve and $A_v=1.6523$ for all sources. Therefore, the simulated images and grism spectra are also dust attenuated.

2.3 Input configuration files

Among the key input components of the axesim software is the configuration file, which describes the imprint of the spectrograph on the detector and contains essential parameters such as the description of the spectral trace and the dispersion solution together with their variations over the Field of View. We created 18 configuration files for 18 detectors. We have made these files available with our data releases.

2.4 More details about input sources: emission line fluxes, etc

Some users might be interested in further details about the simulated sources, in addition to the properties already included in the object tables. Therefore, we also provide the original GALACTICUS mock catalogs ² (see Section 3.2 Wang et al. 2022) that we have used to generate our object tables. For each of our 15 FOVs, we provide one original GALACTICUS catalog. The sky coordinates in the object tables (i.e., the last two columns in the MOT files) can be used to match the sources from the object tables and these original GALACTICUS catalogs. Since these original GALACTICUS catalogs contain significant amounts of information, we advise users that the files are large, on the order of 10 GB.

In particular, these GALACTICUS files include emission line luminosities for $H\alpha$, $H\beta$, $H\gamma$, [NII], [OII], and [OIII] emission lines. In this way, users can compare their line flux measurements from the simulated grism spectra with their original line flux values from GALACTICUS. However, certain points should be considered:

- The values included in the GALACTICUS catalogs are emission line luminosities at the rest frame. In order to compare these values with the measurements from the simulated grism images, one must convert these luminosities to emission line fluxes using the luminosity distance derived using the redshifts contained in the input catalogs. Users can also retrieve the redshift values from the original GALACTICUS catalogs by using the “lightconeRedshif” keyword. For

²We note that these are older versions of GALACTICUS mock catalogs. There have been recent updates in GALACTICUS mock catalogs.

this conversion, users need to choose a cosmology model as following (Using python astropy for instance):

```
from astropy.cosmology import FlatLambdaCDM
h = 0.6774
cosmo = FlatLambdaCDM(H0=100*h, Om0=0.3089, Tcmb0=2.725)
```

- In the GALACTICUS catalogs, emission line luminosities are intrinsic values, so dust attenuation is not considered. In order to compare these emission line values with line fluxes directly measured from simulated grism images, dust attenuation must be added. As previously explained, dust attenuation has been assumed to be constant $A_v = 1.6523$, with a Calzetti attenuation curve. Users can add a dust effect by following the steps below.

$$F_{\text{dust corrected}} = F_{\text{observed}} * 10^{0.4(A_v/R_v)*k(\lambda)} \quad (2)$$

with

$$k(\lambda) = 2.659(-1.857 + 1.040/\lambda) + R_v \quad \text{for } 0.63\mu\text{m} \leq \lambda \leq 2.20\mu\text{m}$$

$$2.659(-2.156 + 1.509/\lambda - 0.198/\lambda^2 + 0.011/\lambda^3) + R_v \quad \text{for } 0.12\mu\text{m} \leq \lambda \leq 0.63\mu\text{m} \quad (3)$$

where $R_v = 4.05$ for the Calzetti dust attenuation curve.

References

- [1] Atek, H., et al. 2010, ApJ, 723, 104
- [2] Benson, A. 2012, New Astronomy, 17, 175
- [3] Chang, C. 2010, The Astrophysical Journal, 724, 182
- [4] CColbert J., et al., 2013, The Astrophysical Journal, 779, 34
- [5] Juruić, A., 2008, ApJ, 673, 864
- [6] Just, A., 2015, MNRAS, 451, 149
- [7] Odewahn S. C. et al. 1997, AJ, 114, 2219
- [8] Perrin, M., et al., 2014, SPIE, 9143, 3XP
- [9] Wang, Y., 2022, ApJ, 928, 1