NFYK15014E Gravitational Dynamics and Galaxy Formation

Reports Compilation

Kushagra Shrivastava

Included Reports:

Report 1: Measuring Mass and Sizes of Galaxies

Report 2: Dark Matter Halos, UVLFs and SMFs

Report 3: SED fitting using BAGPIPES

Report 4: Tracing the ISM of GN20

Report 5: Galaxy Kinematics and Dynamics

Report 6: Reionized Bubbles around Primordial Galaxies

Report 7: Physics of the Intergalactic Medium

Report 1: Measuring Mass and Sizes of Galaxies

Kushagra Shrivastava

The following outputs have been produced using `Python 3.12.3` running on Ubuntu's Windows Subsystem for Linux. This was done to run the `pysersic` package locally on a Windows machine without the need for `Google Collab`.

Output 1: Plotting the 2-D Sérsic Profile

The Sérsic profile as a function from the center of the galaxy is given by,

$$\Sigma(r) = \Sigma_e \exp\left\{-\kappa\left[(r/r_e)^{1/n} - 1\right]\right\} \tag{1}$$

where r_e is the half-light radius, Σ_e is the surface brightness at r_e, and $\kappa \approx 1.9992n - 0.3271$ such that half of the total flux is within r_e. To further ensure that the brightness at $r = 0$ is the same for all profiles (which is arbitrarily set to 1),

$$\Sigma(0) = 1 \Rightarrow 1 = \Sigma_e \exp(\kappa) \Rightarrow \Sigma_e = \exp(-\kappa) \tag{2}$$

the above normalisation is followed. The plot of some theoretical combinations of n and r_e is shown in Figure 1.

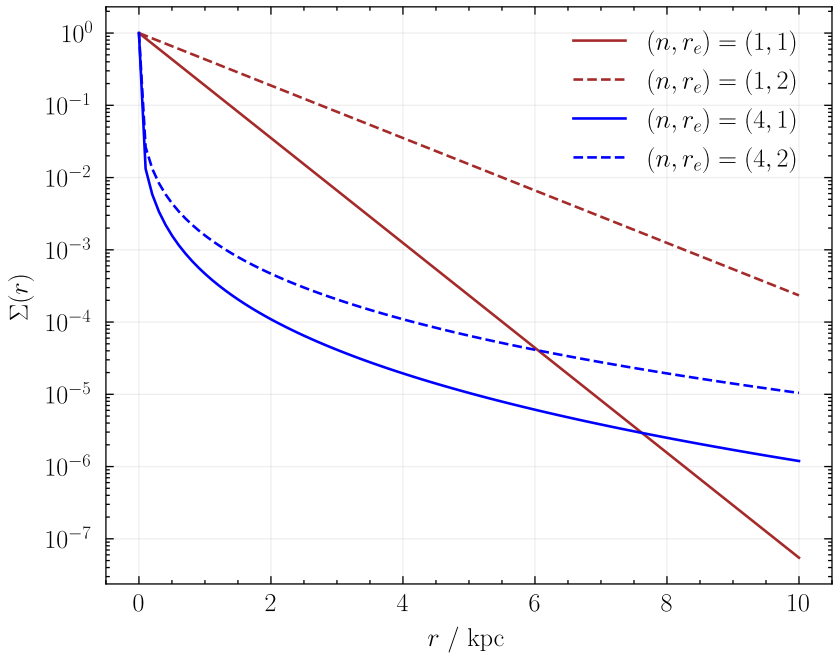

Figure 1: Theoretical Sérsic profiles for different index and half-light radius

Output 2: Plotting the JWST images

Using the `plot_image()` function in `pysersic`, the raw images (after masking) are plotted in Figure 2. The PSF is the same as all images are taken using NIRCam F227W filter.

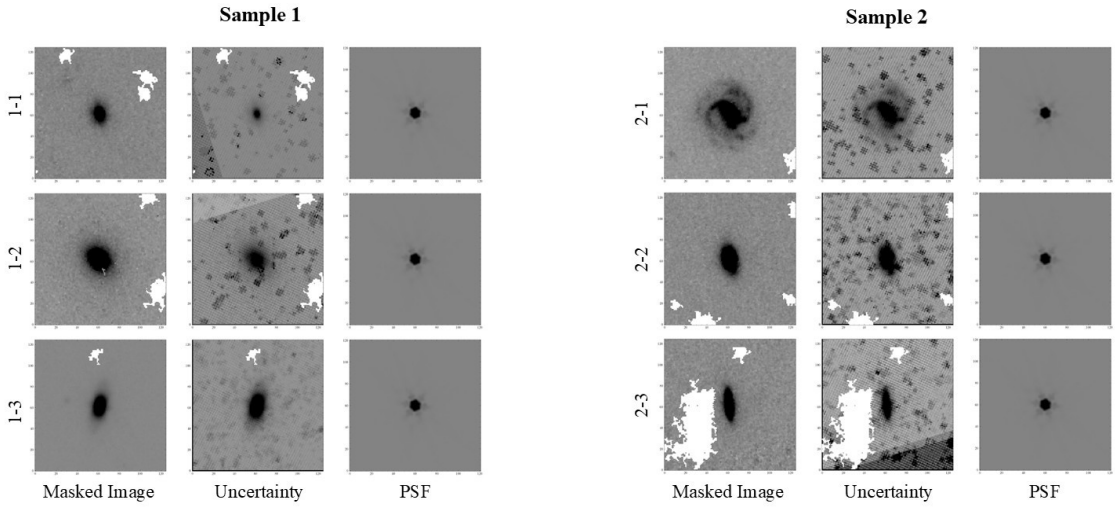

Figure 2: Masked, uncertainty, and point spread function (PSF) images for both samples plotted using `plot_image()` in `pysersic`. The PSF remains the same because as we used data from the same JWST sensor.

Visually, only Galaxy 2-1 appears distinct from the rest because of its spiral arms. The rest do not show any other structure visually. This serves as a motivation to perform surface brightness fitting to reveal the structure of the galaxy.

Output 3: Plotting the fitting residuals

The Sérsic profile was fitted on the galaxies with the procedure described in the example notebook provided in class. After the Bayesian inference and fitting, the original image, model image, and the residuals are shown in Figure 3.

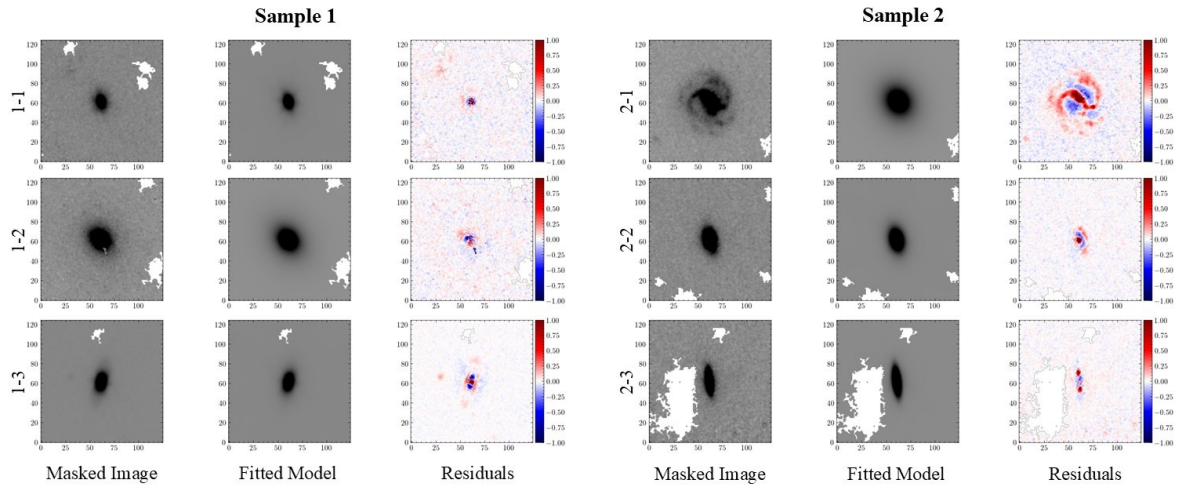

Figure 3: Masked image, Sérsic profile model, and residuals for all the galaxies after `pysersic` fitting. Notice how the residuals reveal the structure of the galaxy as compared to a elliptical profile.

The fitting in Sample 1 galaxies is relatively better than Sample 2, however all samples have significant residuals. The worst fitting is on Galaxy 2-1, which is clearly a spiral galaxy and the Sérsic profile cannot account for the arms. This is expected, as the 2-D

profile describes a flux distribution of an elliptical region. Therefore, the residuals serve as a display of how different is the structure of a galaxy as compared to a simple ellipse. The residuals for galaxies that appear elliptical showcases that they do not follow a perfect elliptical flux profile, but rather something different especially near the core.

Output 4: Fitting results and conversion to kpc scale

Out of the 7 parameters (flux, effective radius, Sérsic index, x and y coordinates of the galactic center, eccentricity, and the angular orientation), we are only concerned with the Flux, effective radius, and Sérsic index. The best-fitted values are shown in Table 1. The best fit value was defined to be the median, with the 16th and the 84th percentile values serving as the lower and upper error bounds respectively to 1-σ.

However, the calculated units for the effective radius is in pixels which needs to be converted to kpc. Given that the galaxies were at a redshift $z = 2.3$, the cosmological transverse distance per arcminute can be calculated directly using `astropy.cosmology.WMAP9.kpc_proper_per_arcmin(z)`. Let this distance be R_{cosmic}. This provides us with the transverse distance subtended by 1 arcminute at that redshift. Furthermore, it is known that each pixel of NIRCam F227W filter covers an angle of 0.04 arcseconds.Therefore, the best fitted values of r_e were converted to kpc using,

$$r_{e,\text{kpc}} = r_{e,''/\text{Pixels}} \times 0.04 \times \frac{R_{\text{cosmic}}}{60} \tag{3}$$

The final values are shown in Table 1.

Output 5: Stellar mass calculation

To obtain the stellar mass, first the best-fitted flux must be converted to the apparent magnitude in the AB magnitude scale using,

$$m_{\text{apparent}} = -2.5 \log_{10} \left(\frac{F_{\text{best-fit}}}{3631 \text{ Jy}} \right) \tag{4}$$

where 3631 Jy is the zero-magnitude flux in the scale. It must be noted that the units of the best fitted flux are 10 nJy, which were converted to Jy. Then, the cosmological luminosity distance to the galaxies, $D(z)$, was calculated using `astropy.cosmology.WMAP9.luminosity_distance(z)`, with $z = 2.3$. The output was in units of Mpc, which was converted to parsecs. Thereafter, the absolute magnitude was calculated using the distance-modulus,

$$M_{\text{absolute}} = m_{\text{apparent}} - 5 \log_{10} \left(\frac{D(z)}{10 \text{ pc}} \right) \Rightarrow m_{\text{apparent}} - 5 \log_{10} \left(\frac{18840.639 \times 10^6 \text{ pc}}{10 \text{ pc}} \right) \tag{5}$$

The luminosity of the galaxy, in terms of solar luminosity, was calculated using the absolute magnitude difference relation,

$$L_{\text{calculated}} = 10^{-0.4(M_{\text{absolute}} - 5.17)} \tag{6}$$

Lastly, the given stellar-mass to luminosity ratio at $z = 2.3$ was used to calculate the stellar mass.

$$M_* = L_{\text{calculated}} \times 10^{-0.27 \times z - 0.28} \approx 0.12560 \; L_{\text{calculated}} \tag{7}$$

Do note that this is a crude approximation, as by using the stellar-mass to luminosity ratio, we assume the total luminosity of the galaxy is coming from stars, which is not accurate. The calculated stellar mass for each galaxy is shown in Table 1. The error bounds were calculated by propagating the error in flux to apparent magnitude, then to absolute magnitude, then to luminosity and lastly to stellar mass.

Sample #	Flux / 10 nJy	$\log_{10}(M_* \ / \ M_\odot)$	r_e / Pixels	r_e / kpc	Sérsic Index
1-1	699 ± 7	11.002 ± 0.004	3.30 ± 0.07	1.11 ± 0.02	3.4 ± 0.1
1-2	1484^{+10}_{-11}	11.329 ± 0.003	11.3 ± 0.1	3.80 ± 0.04	2.42 ± 0.04
1-3	1686 ± 6	11.384 ± 0.002	4.84 ± 0.03	1.62 ± 0.01	1.94 ± 0.03
2-1	1573 ± 12	11.354 ± 0.003	14.5 ± 0.1	$4.87^{+0.05}_{-0.04}$	1.23 ± 0.03
2-2	553^{+5}_{-6}	10.900 ± 0.004	7.6 ± 0.1	2.55 ± 0.05	$1.03^{+0.06}_{-0.05}$
2-3	385 ± 3	10.743 ± 0.003	$8.30^{+0.07}_{-0.06}$	2.78 ± 0.02	$0.655^{+0.007}_{-0.003}$

Table 1: The best-fitted Flux, r_e in pixels, and Sérsic index, along with the calculated r_e in kpc and Stellar Mass, M_* in $M_\odot$. The error corresponds to the 16th and 84th percentile difference from the median. If only a single error value is reported, it means the lower and upper error bounds are the same. All errors are reported to 1 significant figure.

Output 6: Comparing results with Wel et al. (2014).

Using the $z = 2.25$ best-fit relation from Wel et al. (2014) as it is the closest to $z = 2.3$ for our galaxies, the best-fit lines are over-plotted against our samples as shown in Figure 4. Error bars are excluded from the points as they are not visible at the given scale in a log-log plot.

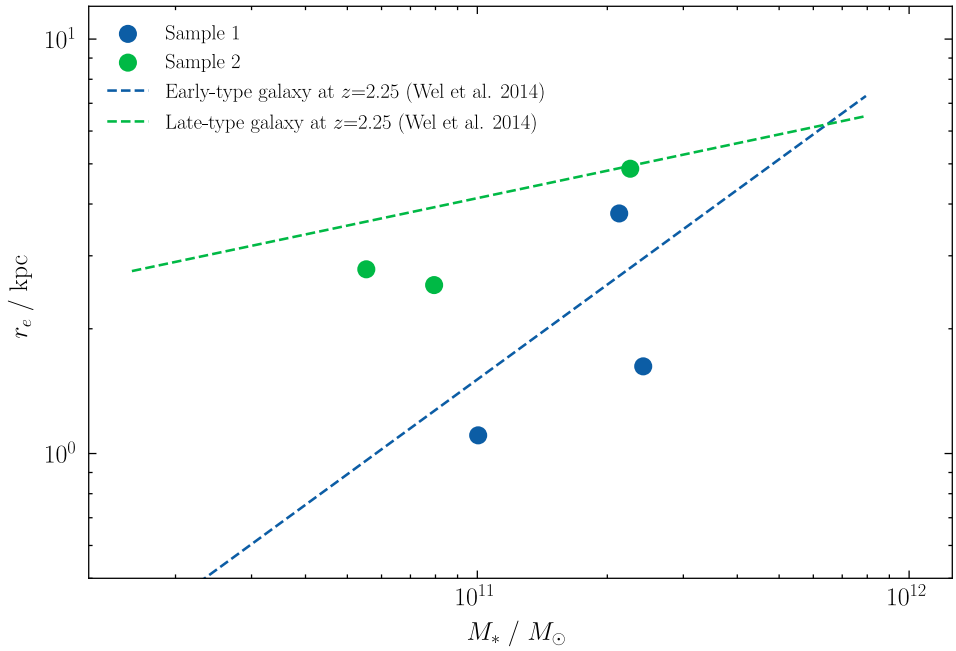

Figure 4: Log-log scatter plot of effective radius in kpc against stellar mass in solar masses for the galaxies in the two samples. The best fit lines are adopted from Wel et al. (2014) relation for effective radius against stellar mass.

Output 7: Discussion on size and Sérsic index

Galaxies in Sample 1 generally have a smaller effective radius than Sample 2 (except Galaxy 1-2) implying that half of the brightness is of Sample 1 galaxies is contained within a smaller radius. Sample 1 galaxies also have a larger Sérsic index than Sample 2 galaxies, and thus have an elliptical shape as compared to Sample 2's disk like shape. Overall, Galaxies in Sample 1 are smaller and have a more ellipse like shape, but galaxies in Sample 2 are larger and have a more disk like shape.

Output 8: Classification of the samples

Sample 1 is more likely to be of early-type galaxies, as the data seems to follow the early-type galaxies best fit line in Figure 4. Furthermore, ETGs are expected to be elliptical and smaller in shape, which sample 1 satisfies as explained in the previous section.

However, the above conclusion cannot be made with a high confidence because there is a lot of inherent scatter between the galaxies. As shown in Figure 5 of Wel et al. (2014), some ETGs lie along the LTG's best fit and vice-versa. Therefore, it is still possible that Sample 2 might be of ETGs. The provided data is insufficient to conclusively classify the samples.

Report 2: Dark Matter Halos, UVLFs, and SMFs

Kushagra Shrivastava

Introduction

This report explores the theoretical and observational connection between galaxy formation in dark matter halos. Using the `MassFunction` class in the `hmf` package in `python`, we analyse properties of HMFs. We use these to model the UVLF and the SMF, which are observable parameters. We conclude that the model-observation prediction for UVLF is generally in alignment, with underpredictions at lower redshifts due to dust, and at very high redshifts due to unexplained mechanisms. For the SMF, we observe generally good model-observation prediction at lower redshifts, however it fails completely to account for the observations at higher redshift. We end with a toy model which includes redshift evolution in the SMF model, to show it shows slightly better agreement with SMF data.

To make the code easier to work with, we also used the following helper functions to account for h (defined as $h = H_0/100$ km s^{-1} Mpc^{-1}) in the `MassFunction` class.

```
def getting_halo_mf(z,Mmin,Mmax,dlog10m=0.005):
    '''
    Helper function to create Halo Mass Functions.
    Cosmological parameters are left as default Flat Lambda-CDM Cosmology.
    '''
    return halomass.MassFunction(z=z, Mmin=Mmin, Mmax=Mmax,dlog10m=dlog10m)

def physical_halomass(hmf):
    '''
    Converts halomass to physical units of M_\odot
    '''
    return hmf.m / hmf.cosmo.h

def physical_dndm(hmf):
    '''
    Converts comoving number density to physical units of Mpc^-3 M_\odot^-1
    '''
    return hmf.dndm * (hmf.cosmo.h)**4

def physical_dndlog10m(hmf):
    '''
    Converts comoving log10 number density to physical units Mpc^-3
    '''
    return hmf.dndlog10m * (hmf.cosmo.h)**3
```

Part 1: The Halo Mass Function

Answer 1: Using the helper functions created before, the Halo Mass Functions (HMFs) in the mass range of 10^6 $M_\odot$ to 10^{18} $M_\odot$ were compiled.

```
redshifts = np.arange(0,17,2)
halo_mass_functions = [getting_halo_mf(i,6,18) for i in redshifts]
```

Answer 2: With the HMFs, the comoving number density of halos, dn/dM_h, was plotted as a function of Halo Mass, M_h, as shown in Figure 1.

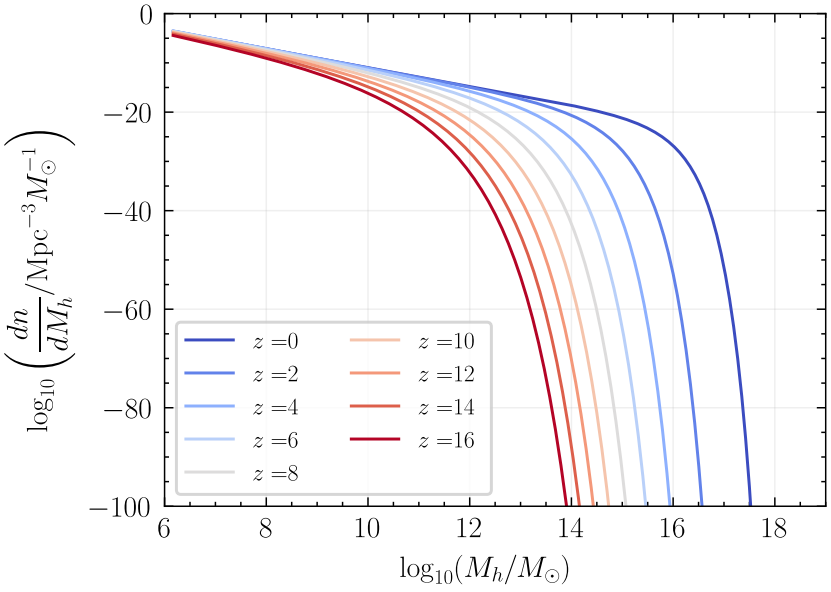

Figure 1: Comoving Number Density of Halos against Halo Mass at different redshifts. The abundance of higher mass halos at lower redshifts is due to their formation through mergers of smaller mass halos.

Answer 3: Lower mass halos (especially with $M_h < 10^{10}$ $M_\odot$) are more abundant at all redshifts as seen in Figure 1.

Answer 4: The evolution of a higher mass halo is much more stronger, as at a fixed M_h, the halo abundance increases with decreasing redshift. This is because higher mass halos are formed through mergers of lower mass halos, and more mergers can take place as cosmological time passes. This further explains Answer 3, as lower mass halos does not undergo such development, and remain abundant across redshifts.

Answer 5: To estimate distribution of nearby Dark Matter Halos, Figure 2 is plotted to zoom in on the concerned region. For our Milky Way Galaxy, we can assume $z \approx 0$ and $M_h = 10^{12}$ $M_\odot$ corresponds to $dn/dM_h \approx 10^{-14.75}$ Mpc^{-3} $M_\odot^{-1}$. Therefore,

$$dn \approx 10^{12} \times 10^{-14.75} = 10^{-2.75}\ \mathrm{Mpc}^{-3} \approx 1.8 \times 10^{-3}\ \mathrm{Mpc}^{-3} \tag{1}$$

which seems like pretty scarce. However this comment cannot be made as a standalone comment, without comparing it with halo numbers at different masses. Therefore, we compare it with the M_h of dwarf galaxies at $z = 0$,

$$dn \approx 10^{10} \times 10^{-11} = 10^{-1}\ \mathrm{Mpc}^{-3} \approx 0.1\ \mathrm{Mpc}^{-3} \tag{2}$$

which is 56× higher. Therefore Dwarf galaxy sized Halos are a lot more common than Milky Way sized halos.

With this information, we can further estimate the average spatial distance between galaxies. Assuming a grid-like structure, i.e., there is one galaxy at the center of each halo and the halo covers a cube area, the separation of galaxies n given a certain volume, can be written as,

$$n = \left(\frac{1}{dn}\right)^{1/3} \tag{3}$$

Therefore,

$$n_{\text{Milky Way}} = \left(10^{2.75}\right)^{1/3} \text{ Mpc} \approx 8.25 \text{ Mpc} \tag{4}$$

$$n_{\text{Dwarf Galaxies}} = (0.1)^{1/3} \text{ Mpc} \approx 0.464 \text{ Mpc} \tag{5}$$

The same analysis can also be applied to halos of the same mass at $z = 15$. M_h of $10^{12}\ M_\odot$ and $10^{10}\ M_\odot$ corresponds to dn/dM_h of 10^{-30} and 10^{-16} Mpc$^{-3}\ M_\odot^{-1}$ respectively. Therefore,

$$dn_{10^{12}} \approx 10^{12} \times 10^{-30} = 10^{-18} \text{ Mpc}^{-3} \tag{6}$$

$$dn_{10^{10}} \approx 10^{10} \times 10^{-16} = 10^{-6} \text{ Mpc}^{-3} \tag{7}$$

The number density of low-mass halos is 10^{15}× greater than higher mass halos, which is much higher than the 56× calculated at $z = 0$. The reason is the same as explained in Answer 4, because the higher-mass halos are formed through mergers of lower-mass halos, and thus are present in greater abundance at lower redshifts.

Furthermore, $z = 15$ is also the range at which JWST observes. Assuming it probes a comoving volume of 10^6 Mpc^3, the number of Milky Way mass halos it can observe is,

$$n_{\text{MW mass halos}} \approx 10^6 \times dn_{10^{12}} = 10^6 \times 10^{-18} = 10^{-12} \tag{8}$$

Therefore, we should expect to see almost no MW mass halos in that volume.

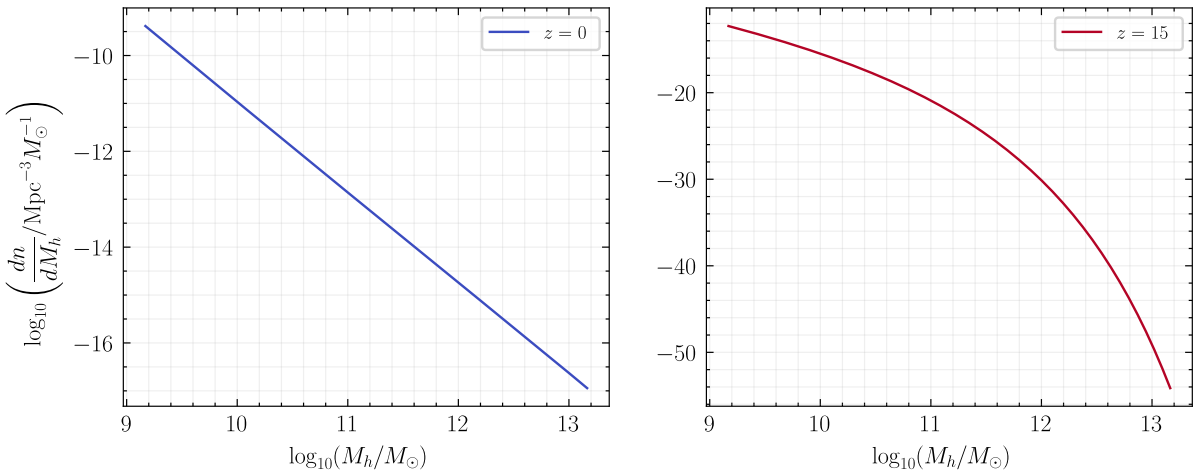

Figure 2: Comoving Number Density of Halos against Halo Mass at $z = 0$ (left) and $z = 15$ (right) redshifts to estimate features of Dark Matter Halos.

Part 2: Star Formation in Dark Matter Halos

Answer 1: The halo mass accretion rate is given by the following equation,

$$\dot{M}_h = 0.03 \times 10^{-9} \times M_h \times \left(\frac{M_h}{10^{12}}\right)^{0.14} \times (1+z)^{2.5} \tag{9}$$

where M_h is the halo mass and z is the redshift. It was implemented in `python` as,

```
def halo_mass_accretion(Mh, z):
    return 0.03e-9 * Mh * (Mh/10**12)**0.14 * (1+z)**2.5
```

As the relation is directly proportional to a positive power of the redshift, accretion is more powerful at the early cosmic times.

Answer 2: The instantaneous star formation efficiency is given by a double power law of the form,

$$\epsilon_{\rm inst}(M_h) = \frac{2\epsilon_0}{\left(\frac{M_h}{M_c}\right)^{-a} + \left(\frac{M_h}{M_c}\right)^{b}} \tag{10}$$

where M_c is the critical mass for efficiency to peak. In `python`, the function was implemented as,

```
def instantaneous_star_formation(Mh, e0=0.18, Mc=10**11.64, a=0.69, b=0.65):
    return 2*e0 / ((Mh/Mc)**(-a) + (Mh/Mc)**(b))
```

The values of the free parameters are the same as provided in the exercise.

Answer 3: $\epsilon_{\rm inst}(M_h)$ is plotted against halo mass, as shown in Figure 3. The efficiency peaks at approximately $10^{11.25}$ $M_\odot$ by looking at the graph.

Answer 4: Before the peak, at the lower mass end, the efficiency is lower due to feedback process such as Type IIa supernovae, photoionisation, etc. These processes increases the temperature of the DM halo, which increases the time to cool the gas and thus preventing gravitational collapse to form protostars. At the higher mass end, Active Galactic Nuclei (AGN) form in the halos which cause outflow of gases due to its relativistic jets and radiation pressure. Thus, less gas is available for stellar formation and efficiency decreases.

Answer 5: The star formation rate (SFR) is given by,

$$SFR(M_h, z) = \epsilon_{\rm inst}(M_h) \times f_b \times \dot{M}_h \tag{11}$$

where $f_b \approx 0.16$ is the cosmic baryon fraction. The same function was implemented in `python` as,

```
def stellar_formation_rate(Mh,z, fb = 0.16):
    return fb * halo_mass_accretion(Mh,z) * instantaneous_star_formation(Mh)
```

Answer 6: From the SFR computed at different values,

$$SFR(10^{10}, 5) \approx 5.87 \times 10^{-2}\ M_\odot\ {\rm yr}^{-1} \tag{12}$$

$$SFR(10^{10}, 15) \approx 6.82 \times 10^{-1}\ M_\odot\ {\rm yr}^{-1} \tag{13}$$

star formation is faster at early cosmic times. As stated in answer 1, the mass accretion is more powerful at that cosmic time, which is because of the higher density in the cosmic web, leading to higher star formation.

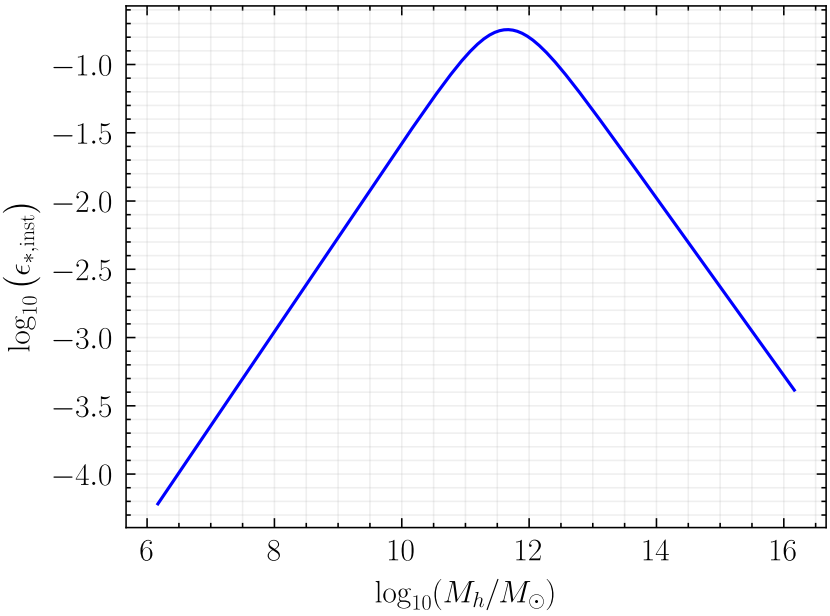

Figure 3: Instantaneous star formation efficiency as a function of halo mass. Notice that the peak occurs very close to the mass of the milky way halo, which is 10^{12}. The graph has a slope on either side of the peak due to feedback processes suppressing stellar formation.

Part 3: The Luminosity function

Part 3.1: Theoretical Predictions

Answer 1: To compute the UV magnitude, we first compute the UV-luminosity of the halo using,

$$L_{UV} = \frac{SFR}{\kappa_{UV}} \tag{14}$$

where $\kappa_{UV} = 1.15 \times 10^{-28}\ M_\odot\ \mathrm{yr}^{-1}\ \mathrm{erg}^{-1}\ \mathrm{s}$. Following which, we can calculate the magnitude in AB scale by,

$$M_{UV} = 51.63 - 2.5 \log_{10}(L_{UV}) \tag{15}$$

The same function was implemented in `python` as

```
def M_UV_sfr(sfr, kappa_UV = 1.15 * 10**-28):
    L_UV =  sfr/ kappa_UV
    return 51.63 - 2.5*np.log10(L_UV)
```

Answer 2: The previous function can also be further abstracted by calculating the SFR from the halo mass using the functions created in Part 2.

```
def M_UV_galaxy(Mh, z):
    return M_UV_sfr(stellar_formation_rate(Mh,z))
```

Answer 3: A plot of UV magnitude against Halo Mass is shown in Figure 4. At a fixed mass, Halos at higher redshifts (early cosmic time) are more luminous because they have more young stars. This is due to higher SFR, as explained in Part 2 Answer 6.

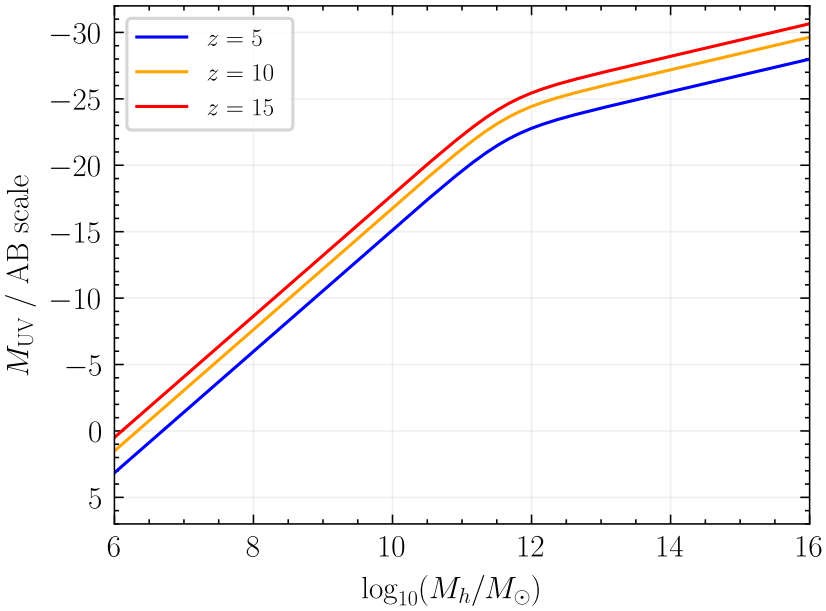

Figure 4: Plot of UV magnitude against Halo Mass for different redshifts. We can notice the halos at higher redshifts are more luminous.

Answer 4: The UV Luminosity Function (UVLF) is defined as $\phi = dn/dM_{UV}$. Using the chain rule, we can perform a change of variable to calculate it from the halo mass function,

$$\phi = \frac{dn}{d\log_{10} M_h} \left| \frac{d\log_{10} M_h}{dM_{UV}} \right| \tag{16}$$

The Jacobian was calculated numerically using `numpy.gradient` function in `python`,

```
def UVLF(M_min, M_max, z):
    hmf = getting_halo_mf(Mmin=M_min,Mmax=M_max,z=z)
    mass = physical_halomass(hmf)
    dndlog10m = physical_dndlog10m(hmf)
    muv = M_UV_galaxy(mass,z)
    return (mass, muv, dndlog10m * np.abs(np.gradient(np.log10(mass), muv)))
```

Answer 5 and 6: Figure 5 shows the UVLF plotted against the UV-magnitude at different redshifts. As observed here, the normalisation decreases with redshift as the entire curve moves down. In contrast, the faint end slope across redshifts evolve with a similar small gradient. Lastly, we observe that between -22.5 and -27.5 magnitude, the UVLF's gradient becomes much steeper. Lower redshifts make this transition before higher redshifts.

Answer 7: We also observe that at a constant magnitude, the UVLF decreases with increasing redshift (before the turn-off regime where there is a transitionary period). To understand this, we can relook at Figure 4 and observe that for a constant magnitude, the halo mass is lower with increasing redshift. At a lower halo mass, the star formation efficiency decreases, which results in fewer younger stars leading to a less bright UV magnitude. Thus the comoving number density of galaxies per unit UV magnitude decreases.

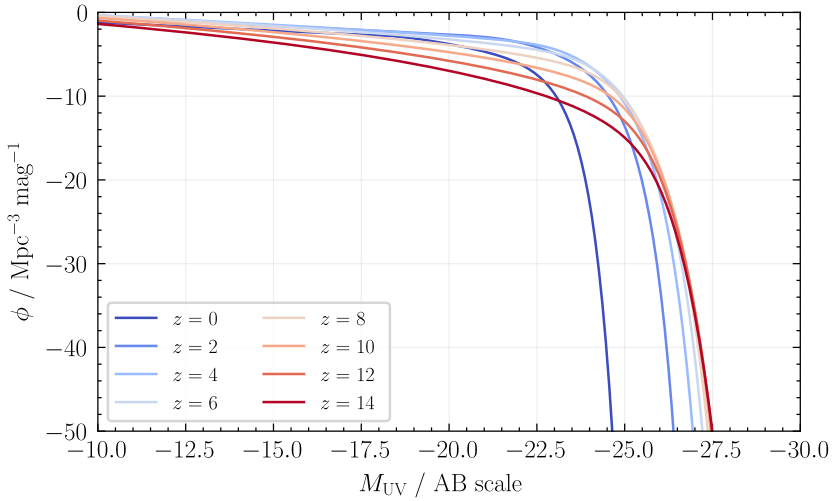

Figure 5: Plot of UVLF against M_{UV} at different redshifts.

Part 3.2: Observations

Answer 1, 2, and 3: The observed UVLF from Bouwens et al. 2021 and the model UVLF against the UV-magnitude is shown in Figure 6. A `coolwarm` color map was used to match the redshifts to their corresponding colours for easy visualisation. The model generally holds well for the observed data, however it starts to overpredict the UVLF at the brighter end for the lower redshifts. This might be because as redshift decreases, the intergalactic medium becomes enriched with dust, which attenuates the UV light, causing extinction. The current UVLF model does not account for such extinction, and therefore it overpredicts.

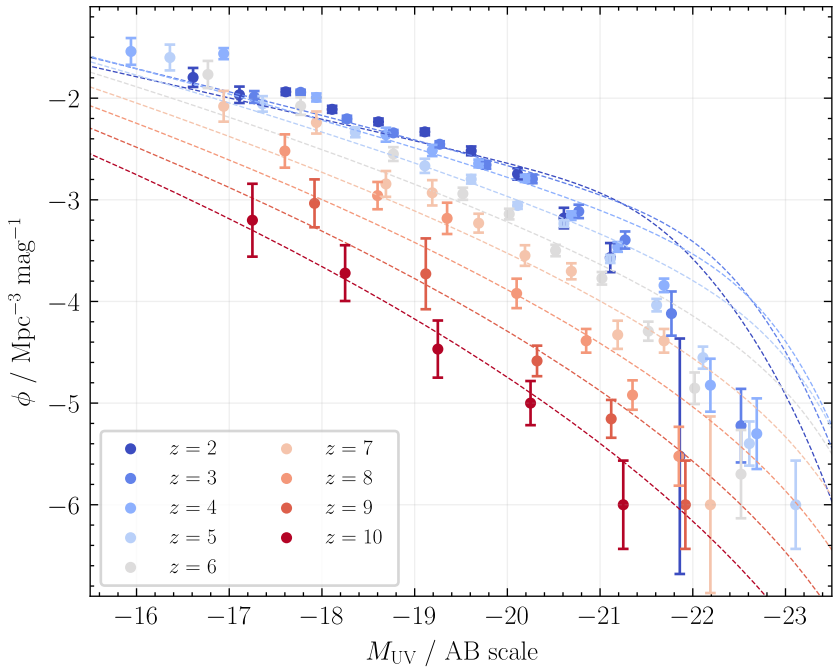

Figure 6: Plot of model UVLF (dashed lines) and observed UVLF (Bouwens et al. 2021) against M_{UV} at different redshifts. Notice how the model generally follow the observations, however, it overpredicts the UVLF at the brighter end for lower redshifts.

Answer 4: The observed UVLF at even higher redshifts from Whitler et. al. (2025) and Donnan et. al. (2024) with the model UVLF against UV magnitudes is plotted in Figure 7. The model severely underpredicts the UVLF. This suggests our current understanding of halo evolution and their stellar formation at very high redshifts is incorrect. There must be some unknown mechanism which is driving a higher stellar formation in halos at that cosmic epoch.

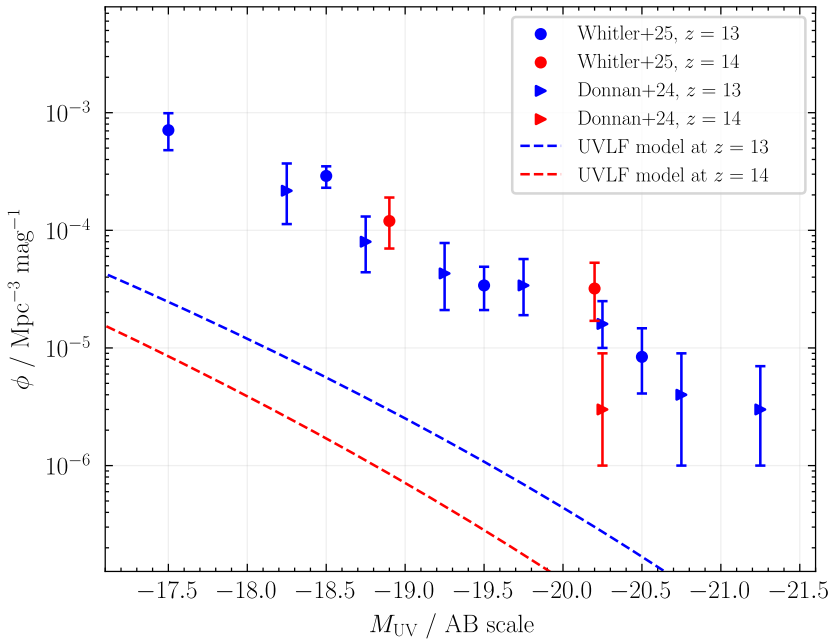

Figure 7: Plot of model UVLF (dashed lines) and observed UVLF (Whitler et al. 2025 and Donnan et al. 2024) against M_{UV} at different redshifts. Notice how the model severely underpredicts the UVLF, suggesting our understanding of galaxy formation at very high redshifts needs to be improved.

Part 4: Computing the Stellar Mass Function from the Halo Mass Function

Part 4.1: Stellar Mass Function at Constant Redshift

Answer 1: SHMR is the Stellar-to-Halo Mass Relation, which explains what proportion of a halo mass will be converted to stellar mass over its lifetime. Its functional form is,

$$M_* = \epsilon_*(M_h) \times f_b \times M_h \tag{17}$$

where M_h is the halo mass, $\epsilon_*(M_h)$ is the integrated star formation efficiency, i.e., the efficiency throughout the lifetime of a star, and f_b is the cosmic baryonic fraction, which is the fraction of mass in the universe that is baryonic (in our context, we assume the same fraction is valid in our dark matter halos). This expression can be derived by integrating equation 11 over time.

Answer 2: The integrated star formation efficiency takes the form of a double power

law, given by,

$$\epsilon_*(M_h) = \frac{2\epsilon_0}{\left(\frac{M_h}{M_c}\right)^{-\alpha} + \left(\frac{M_h}{M_c}\right)^{\beta}} \tag{18}$$

where ϵ_0 is the peak efficiency, M_c is the critical halo mass at which that peak efficiency is reached, and α and β are the parameters controlling the low and high end slope before and after the peak efficiency respectively. Visually, it has the same shape as the instantaneous star formation efficiency shown in Figure 3. Therefore, as M_h increases, the efficiency increases until it reaches the peak, and then stars decreasing again.

Answer 3 and 4: The Stellar Mass Function, $dn/d\log M_*$, can be obtained using the SHMR and the Halo Mass Function through a change of variables as shown below,

$$\frac{dn}{d\log M_*} = \frac{dn}{d\log M_h}\left|\frac{d\log M_h}{d\log M_*}\right| \tag{19}$$

This was implemented in `python` as shown below.

```
def eps_func(Mh, e0, Mc, alpha, beta):
    return 2*e0 / ((Mh/Mc)**(-alpha)+(Mh/Mc)**(beta))

def Mstar_from_Mhalo(Mhalo,dndlog10m, e0, Mc, alpha, beta,fb=0.16):
    Mstar =  eps_func(Mhalo, e0, Mc, alpha, beta)*fb*Mhalo
    jacb = np.gradient(np.log10(Mhalo),np.log10(Mstar))
    return Mstar, dndlog10m * np.abs(jacb)
```

Answer 5: The graph of model HMF, model SMF and observed SMF (Shuntov et. al. 2025) against mass is shown in Figure 8 at $z = 2.25$.

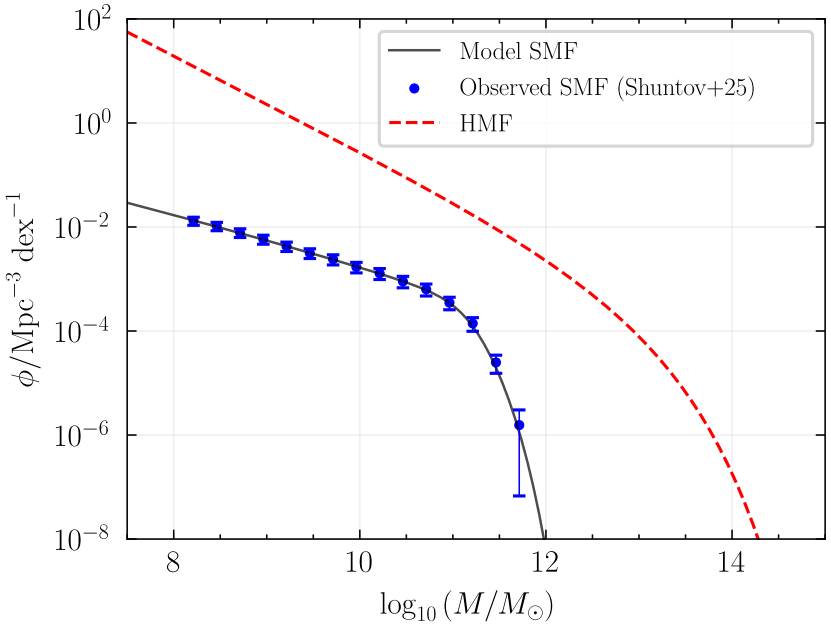

Figure 8: Plot of model SMF, Observed SMF (Shuntov et. al. 2025) and the HMF against the mass at $z = 2.25$. The model fits the observational data well, but the SMF is significantly below the HMF.

Answer 6: The best-fit SHMR values (same as quoted in class) were found to be $\epsilon_0 = 0.12$, $M_c = 10^{12.25}$, $\alpha = 1.1$, and $\beta = 0.45$. While changing these parameters, it was

observed that changing two of them together might lead to them canceling each other out. To mathematically find the degeneracies, we take the limits of equation 18. In case when $M_h >> M_c$,

$$\left(\frac{M_h}{M_c}\right)^{-\alpha} \approx 0 \Rightarrow \epsilon_* \approx 2\epsilon_0 M_c^{\beta} M_h^{-\beta} \tag{20}$$

In the case when $M_h << M_c$,

$$\left(\frac{M_h}{M_c}\right)^{\beta} \approx 0 \Rightarrow \epsilon_* \approx 2\epsilon_0 M_c^{-\alpha} M_h^{\alpha} \tag{21}$$

Therefore, changing any of the above combination of parameters, at the same M_h, would lead to cancellation of the net effect, rendering them degenerate.

Answer 7: The main reason why the SMF is much lower than the HMF is due to low efficiency of stellar formation. The larger difference between SMF and HMF at the lower and higher mass ends can be explained using physical processes already specified in Part 2 Answer 4 (Type IIa SNe and photoionisation for the lower end, and AGNs for the higher end both hinder stellar formation).

Answer 8: The halos around the mass of 10^{11} $M_\odot$ seem to be the most efficient as they have the lowest gap between the HMF and the SMF at that mass. This might be because this mass is a nice tradeoff point between the increased mass and not so strong AGN, leading to the most efficient stellar formation.

Part 4.2: Redshift Evolution of the Stellar Mass Function

Answer 9, 10, and 11: The observed SMF from Shuntov et al. 2025 across all the given redshift data bins is shown in Figure 9. The middle value of the redshift bin was used in the fitting. For example, if the bin was in the interval (z_1, z_2), the fitting redshift was $z_{\text{fit}} = (z_1 + z_2)/2$.
Answer 12: Between redshifts of 1.1 and 3, the non-evolving SHMR SMF model fits the data really well. This is to be expected because our best fit parameters were from $z = 2.25$. At redshifts greater than 5.5, the model completely fails as it cannot capture the observations at all.

Furthermore, in the low mass regime of less than 10^9 $M_\odot$, the model underpredicts the SMF at both the low and high redshifts (except the region mentioned in Answer 12 where it correctly predicts). This might be because the feedback caused by the low mass processes might not be as severe as expected. For example, the timescale for Type IIa SNe to occur in low-mass halos might be longer than expected, as the star's initial mass may not be as high as predicted. Therefore, the Type IIa SNe might not be as frequent as expected and therefore, the model underpredicts.

On the other hand, in the high mass regime of greater than 10^{11} $M_\odot$, the model overpredicts in the low redshift regime, and underpredicts in the high redshift regime. The overprediction may be attributed to attenuation of UV light due to intergalactic dust. The underprediction may be due to AGN feedback. As observed, the underprediction gets worse as the redshift increases. This implies that the AGN feedback efficiency has been decreasing over cosmic time. Their feedback must have been significant during the early cosmic epochs which makes our models completely fail.

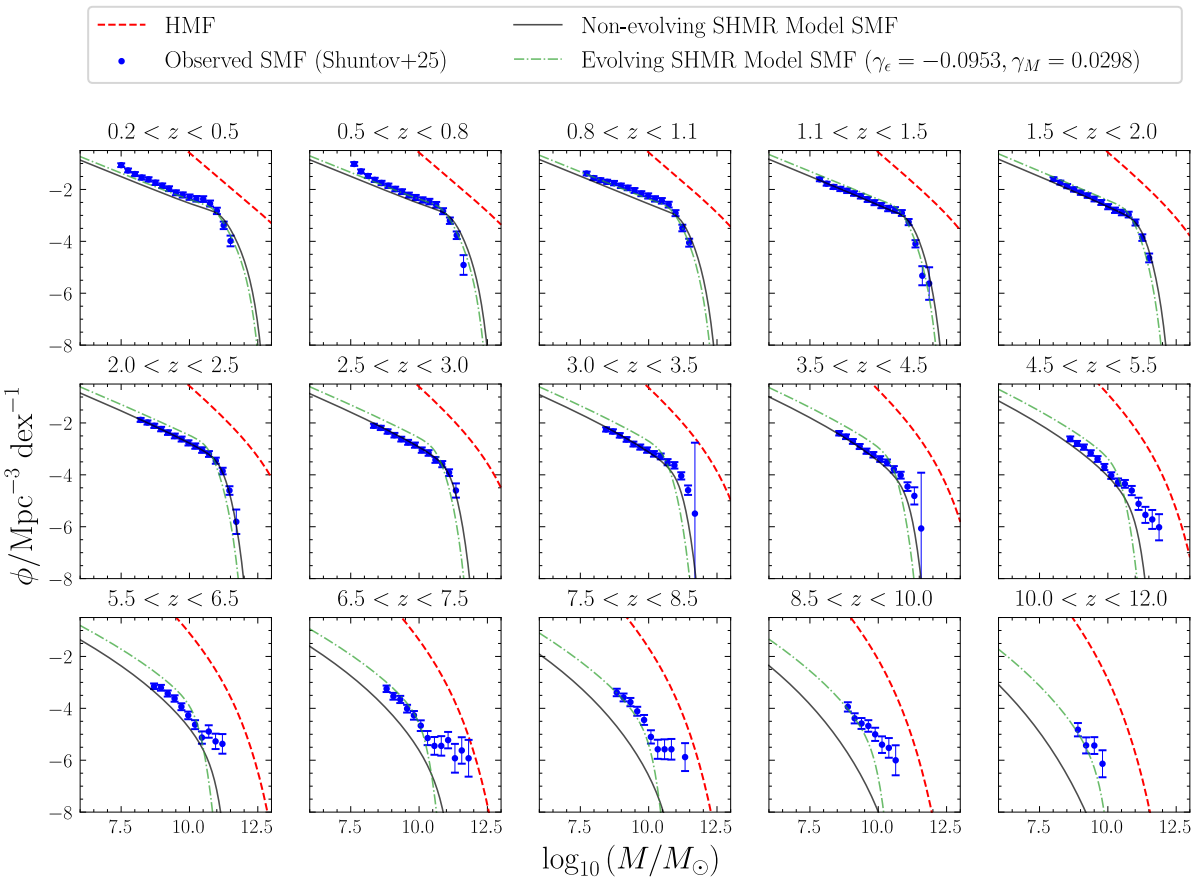

Figure 9: Plot of non-evolving model SMF, evolving model SMF, Observed SMF (Shuntov et. al. 2025) and the HMF against the mass at various redshifts. While neither fitting is great across all redshifts, the evolving-model SMF seems to fit the observations better especially at higher redshifts.

Lastly, as predicted by the JWST, we observe brighter than expected galaxies at redshifts greater than 5, given by the rightward position of the observed datapoint as compared to the model. The SHMR may be changed to account for evolution at different redshifts, as shown in the next part.

Answer 13: To account for the evolving SHMR due to redshift, the following parameterization was considered,

$$\epsilon_0(z) = \epsilon_{0,0} \times (1+z)^{\gamma_\epsilon} \tag{22}$$

$$M_c(z) = M_{c,0} + 10^{\gamma_M \cdot z} \tag{23}$$

where γ_ϵ and γ_M were fitting parameters. The values of the other SHMR parameters were not changed and kept constant from the previous parts. We then modified our code to calculate the χ^2 value across all the redshifts bin given by,

$$\chi^2 = \sum_{\text{all } z} \left(\frac{\phi_{\text{observed}} - \phi_{\text{interpolated}}}{\sigma_{\phi,\ \text{observed}}} \right)^2 \tag{24}$$

The interpolated value was calculated by interpolating the model to the mass of the observed ϕ value using `numpy.interp`. After this calculation, the χ^2 value was minimized using the Nelder-Mead simplex method in `scipy.optimize.minimize`. The best fit parameter values were $\gamma_\epsilon = -0.0953$ and $\gamma_M = 0.0298$. The evolving model results are

shown in Figure 9 with the green dot-dash line.

While the evolving model does not perform significantly better at the lower redshifts, it seems to be performing better than the non-evolving model at the higher redshifts. That being said, this was an arbitrarily simple model just to account for redshift. A better model with theoretical backing to justify the redshift may perform better.

Report 3: SED fitting using BAGPIPES

Kushagra Shrivastava

In the following report, `bagpipes V.1.3.4` was used with `numpy V.1.26.4`. The versions were manually downgraded in a `python V.3.12.2` virtual environment to account for the issues with the new `numpy` update. (Note: As of 20 February 2026, `bagpipes V.1.3.5` doesn't require the above setup as it automatically accounts for everything.)

Part 1: Exploring Model Spectra and Photometry

Figure 1 shows the model spectrum and photometry with SFH prescription of Burst and dust A_V of 0.1 and 2.0. The photometry points present only at the higher wavelength end corresponds to the filters available with JWST.

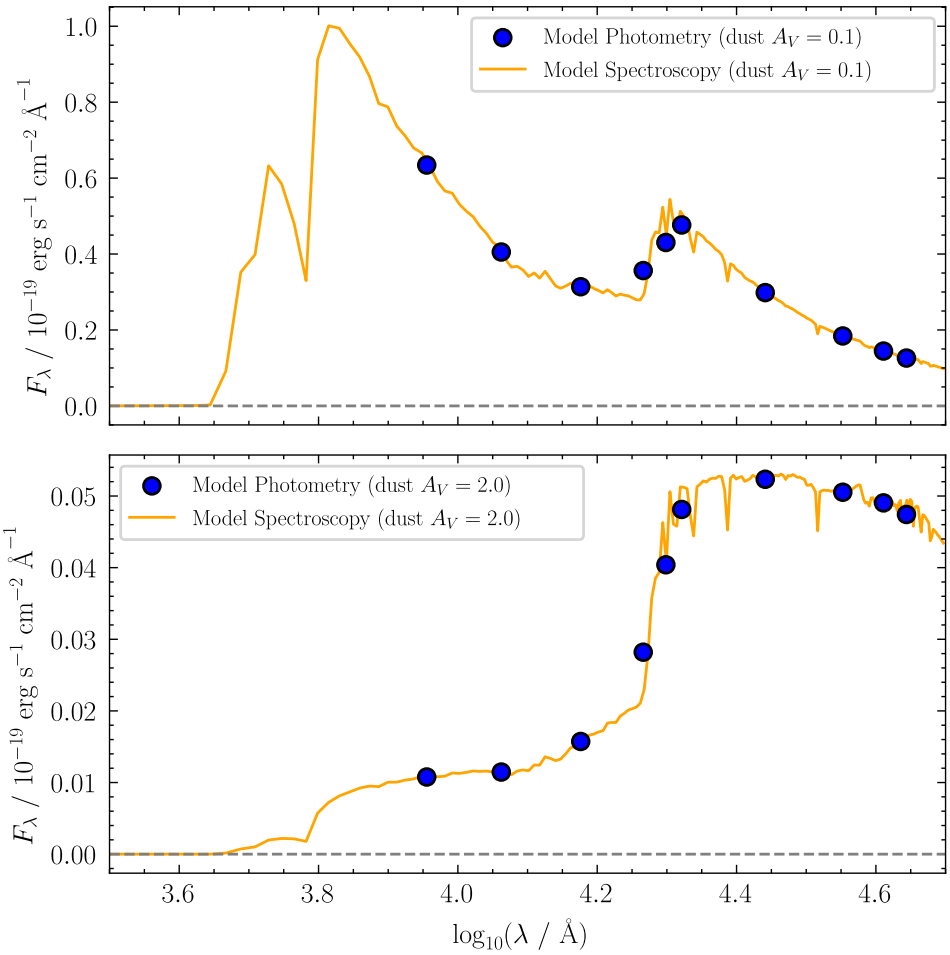

Figure 1: Model spectrum and photometry plot using Burst SFH for different dust A_V. We can see clear difference between the two spectrums, as with a higher A_V, the spectral flux density decreases. The decrease is larger at shorter wavelengths.

To estimate the strength of the balmer break, we use the $D4000$ index defined as,

$$D4000 = \frac{\sum_{4000}^{4100} F_\lambda}{\sum_{3850}^{3950} F_\lambda} \tag{1}$$

For the spectrum shown in Figure 1, the index was 1.70 and 2.10 for dust A_v 0.1 and 2.0 respectively. This is expected, as with the increased A_V, we expect to observe an overall decrease in spectral flux density, but more so at shorter wavelengths (corresponding to UV region). Therefore we notice an increase in the D4000 index.

The same change is expected in the UVJ diagram as shown in Figure 2. Dust will decrease the spectral flux density, and therefore increase the magnitude. As the attenuation is greater at shorter wavelengths, U magnitude increases the most, followed by V, and followed by J. Therefore, both $U - V$ and $V - J$ is expected to increase, which is also seen in the figure. Therefore, the galaxy will be classified as Star Forming at the lower attenuation and Quiescent at higher attenuation.

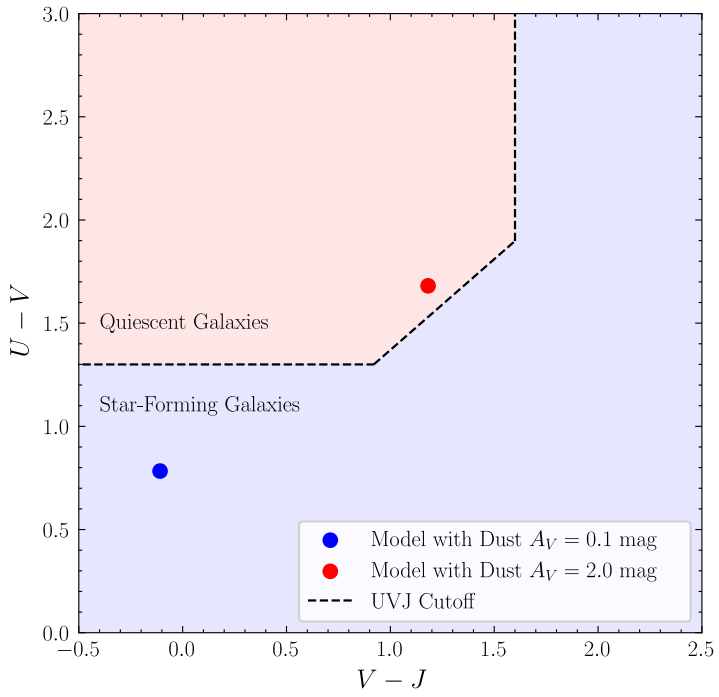

Figure 2: Model UVJ values at different dust attenuation. Increasing the magnitude of attenuation causes the galaxy to move top-right in the diagram.

SFH Prescription	D4000 index
Exponential	1.39
Burst	1.70
Double Power Law	1.51
Constant	1.40
Delayed	1.45

Table 1: D4000 index calculated for different SFH Prescriptions with all other model values kept constant.

Now, keeping the dust A_v constant at 0.1, and then changing the prescription for the Star-Formation Histories (SFH), we can recalculate the D4000 index as shown in Table

1 and the UVJ diagram as shown in Figure 3. Based on the prescription, the galaxy's type is different. This is expected as different SFH prescriptions start star formation at different times and at different rates. Therefore at the same redshift, we expect to observe these galaxies with a different population of stars and rate of star formations. This further showcases the importance of including SFH in galaxy modelling.

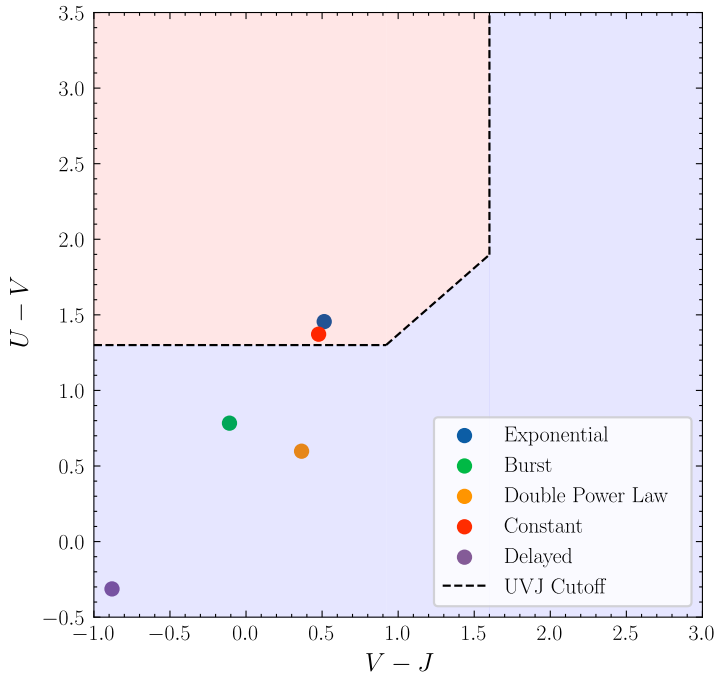

Figure 3: Model UVJ values at different SFH prescriptions. Notice that the galaxy is star-forming or quiescient depending on its SFH. This is expected, as SFH dictates when a star will be formed.

Part 2: Exploring Real JWST Data

After importing the .fits photometry and spectra of galaxy 8290, we first converted the spectral flux density to units of erg s^{-1} cm^{-2} $Å^{-1}$. This was done directly by using `astropy.units` along with `spectral_density` to force it to convert Hz^{-1} to $Å^{-1}$. After this conversion, both the spectra and photometry results were plotted as shown in Figure 4.

We notice that the Photometry and Spectra are generally well in agreement, except at the low wavelength where it appears slightly higher than spectroscopy and at higher wavelengths where it appears slighty lower. This difference could be due to inherent errors in spectroscopy itself (as noticed by the large error bars) or by low sensitivity of the photometry as photometry averages out the flux and provides the central wavelength.

Furthermore, the Balmer Break is very prominent in the spectrum at about 21000 Å. Given that the rest frame wavelength of the Balmer Break is about 4000 Å, the redshift can be estimated as,

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}} \approx \frac{21000 - 4000}{4000} = 4.25 \tag{2}$$

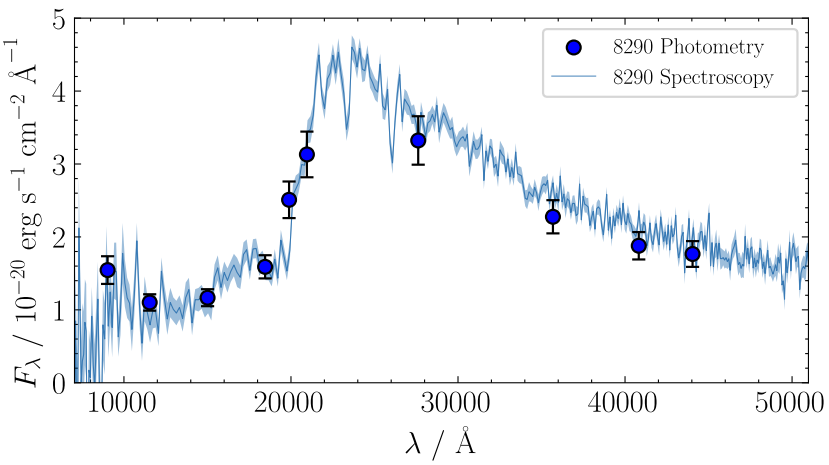

Figure 4: Photometry and Spectra of Galaxy 8290. The Balmer Break is noticeable at about 21000 Å. Using this feature, we can estimate the redshift to be 4.25.

Given we have estimated the redshift, we can now also estimate the AB magnitude in the rest-frame. To do this, we convert the observed wavelengths to rest-frame wavelengths using $\lambda_{\text{rest}} = \lambda_{\text{emit}}/(1+z)$. After which we found the closest wavelength to 5500 Å, which was $\approx$ 5498 Å. Using the flux at this wavelength in units of Jy, we obtained the AB magnitude of 23.9 using the formula that defines the AB magnitude system.

Similarly, we calculated the D4000 index of 1.47 by summing the flux in bins (rest wavelength) according to Equation 1.

The spectrum shown in 4 does not have the same resolving power at all wavelengths. To visualise the resolving power difference, we plotted Figure 5 using the resolving power directly from the third column of `jwst_nirspec_prism_disp.fits`.

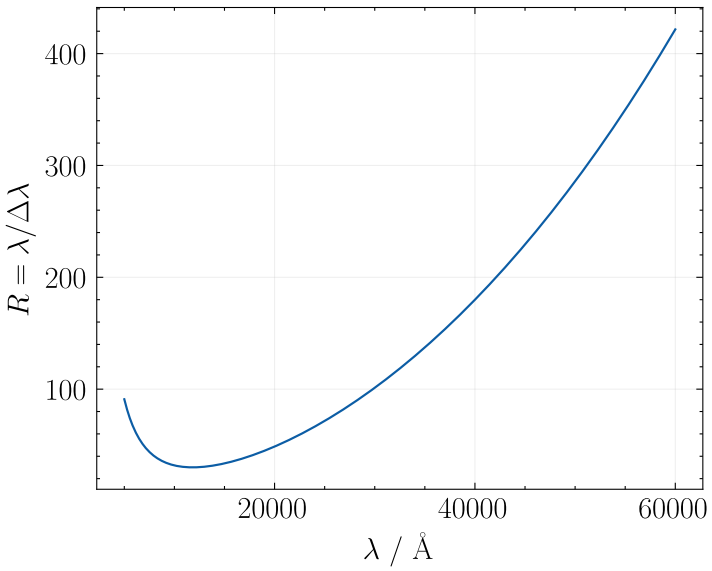

Figure 5: Resolving power of the NIRCam PRISM. Notice how the power is variable, and much higher at longer wavelengths.

Part 3: SED modelling of NIRCam Photometry

Importing the Photometry for Galaxy 8290, setting the SFH to delayed/double powerlaw, and turning off calibration, R curve, and noise, the best-fit photometry model is plotted in Figure 6 and Figure 7. For delayed SFH, the fit is alright, with the SED model underfitting in the middle wavelengths. The estimated photometric redshift is $4.03^{+0.12}_{-0.23}$, which is similar to our estimate from earlier within the error bounds. For double powerlaw SFH, the fit is much better, with the SED fitting almost all points well. The photometric redshift is 4.23 ± 0.11, which is closer to our earlier estimate. Table 2 summarises everything from the fitting.

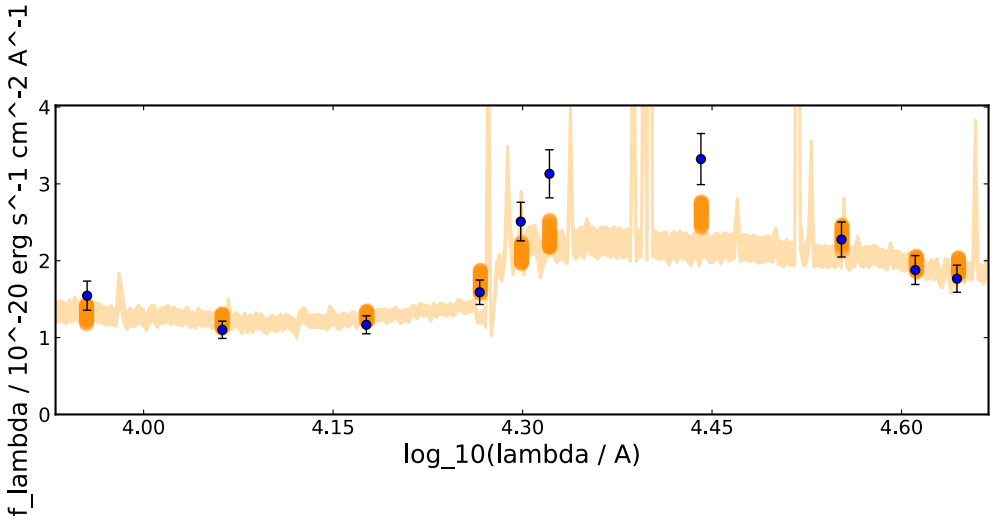

Figure 6: Best fit model using delayed SFH and only photometric data for Galaxy 8290. The fit is alright, with the SED underpredicting in the middle wavelengths. The ugly graph is a result of being unable to extra data or force latex to work.

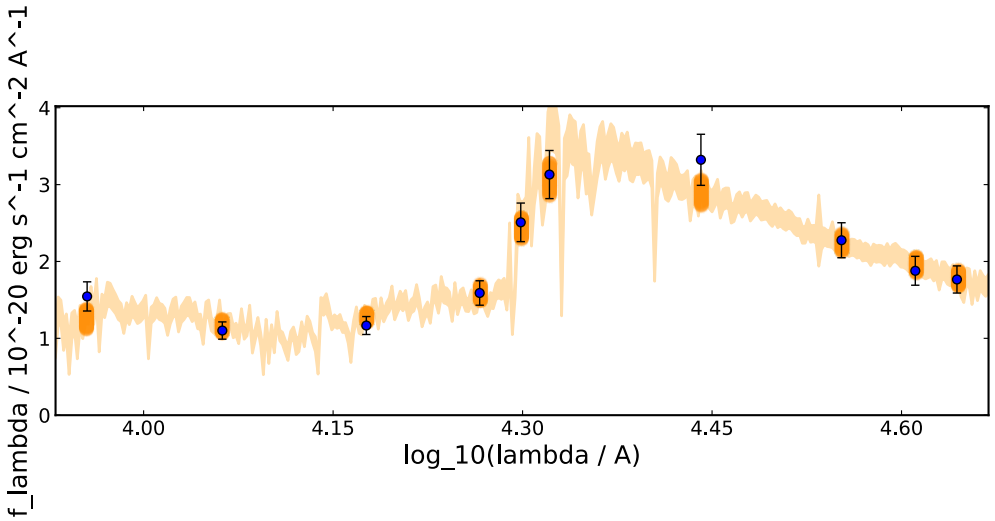

Figure 7: Best fit model using double powerlaw SFH and only photometric data for Galaxy 8290. The fit is better than delayed SFH, with the SED underpredicting only around two points. The ugly graph is a result of being unable to extra data or force latex to work.

Using the `posterior.get_advanced_quantities()` method, we can calculate the UVJ values. The plot is shown in Figure 8. While the scatter is quite a lot, the median with

1-σ error bars imply the galaxy is Quiescent and Star Forming for Double Power Law and Delayed SFH respectively. The same conclusion is supported by the sSFR criterion (the values are quoted in Table 2).

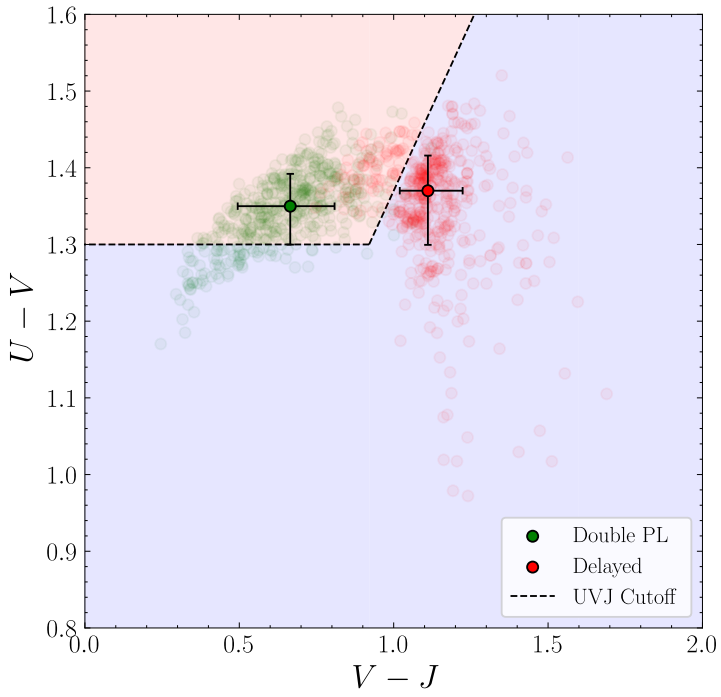

Figure 8: UVJ plot for delayed and star-forming SFH. The transparent scatter are the individual UVJ values, and the solid point is the median with 1-σ error bars.

Furthermore, we can use the Bayesian Evidence (BE) to compare the two models, and Double PL's BE is greater than Delayed, implying a Double PL is the preferred model.

For this preferred model, the t_{form} and t_{quench} are 0.625 and 1.24 Gyr respectively. This shows the quenching took place in a relatively short period of time, implying the quenching process must have been an extremely fast one.

SFH	z	BE	Type (UVJ)	t_{age} / Gyr	sSFR / Gyr^{-1}	Type (sSFR)
Delayed	$4.03^{+0.12}_{-0.23}$	440	SF	1.53	2.14	SF
Double PL	4.23 ± 0.11	447	Q	1.44	1339	Q

Table 2: This table summarizes the values from photometry-only fitting of 8290.BE, SF and Q is Bayesian Evidence, Star-Forming, and Quiescient respectively.

Part 4: Spectral Fitting of a NIRCam PRISM spectrum

Following the same order as previous part, Figure 9 and 10 shows the best-fitted spectroscopic SED for delayed and double PL SFH respectively. Clearly, the double PL SFH fit is much better than the delayed SFH. The Bayesian Evidence is -551 and -319 respectively for delayed and double PL SFH, clearly implying the latter is a better fit.

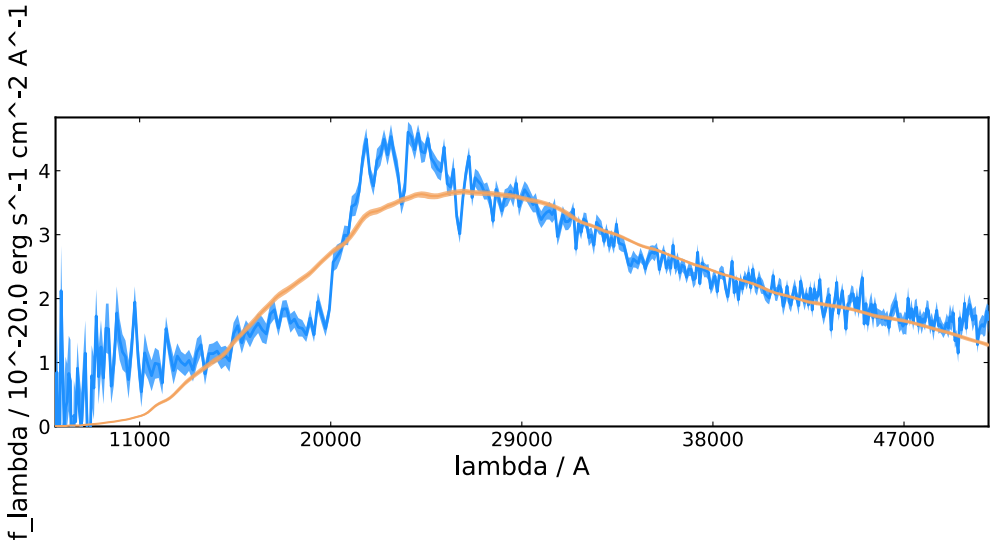

Figure 9: Best fit model using delayed SFH and only spectroscopic data for Galaxy 8290. The fit is bad, with the SED only accounting for the higher wavelength region. The ugly graph is a result of being unable to extra data or force latex to work.

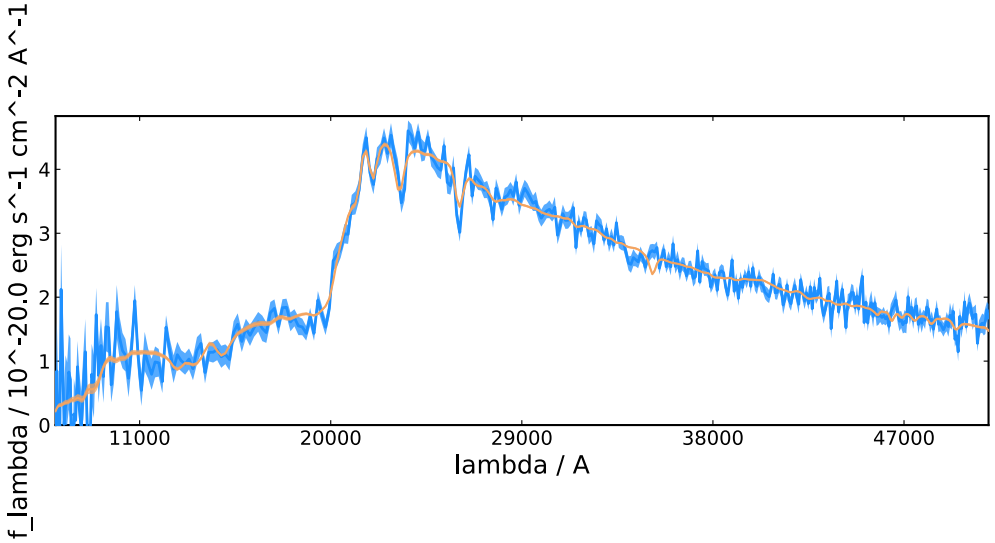

Figure 10: Best fit model using double PL SFH and only spectroscopic data for Galaxy 8290. The fit is very good, with the SED matchingthe observations very well. The ugly graph is a result of being unable to extra data or force latex to work.

The spectroscopic redshift using double PL SFH is $4.356^{+0.004}_{-0.005}$, which is within the error bounds compared to the photometric redshift, but with much smaller errors. We perform a comparison of the different values between photometric and spectroscopic fitting with double PL SFH is shown in Table 3. The spectroscopic fit have smaller error bounds. The SFR is effectively 0 for both cases, further supporting this is a Quiescent galaxy.

Overall, spectroscopic fitting seems to be a better method to account for SED parameters as it gives lower errors. However, photometry is better to obtain colour related parameters, such as magnitudes. Ideally, a combination of both should be the best. The parameters obtained through photometry, such as SFR and stellar mass, can be propogated over to other calculations such as UVLFs.

Lastly, the t_{form} and t_{quench} for the double PL SFH is 0.88 and 0.98 Gyr respectively. This is an even smaller interval of Star Formation, as compared to the photometric fit. This implies the physical mechanism should be extremely abrupt, to create and quech in such a short timeperiod.

Parameter	Photometric	Spectroscopic
Redshift	4.23 ± 0.11	$4.356^{+0.004}_{-0.005}$
$\log_{10}(M_{\text{star}} \ / \ M_{\odot})$	10.34 ± 0.05	$10.56^{+0.07}_{-0.08}$
SFR / $M_{\odot}$ yr^{-1}	$0.03^{+2}_{-0.03}$	$(2.3 \times 10^{-26})^{1.5\times10^{-6}}_{2.3\times10^{-26}}$

Table 3: Double PL SFH fitted SED for 8290 parameters. The spectroscopic errors are smaller, and the SFR is effectively 0 for both cases.

Part 5: Implications

Comoving volume

The survey area of 160 arcmin^2 corresponds to a solid angle of,

$$\Omega_{\text{survey}} = 160 \times \left(\frac{\pi}{60 \times 180}\right)^2 \text{ sr} \approx 1.35 \times 10^{-5} \text{ sr} \tag{3}$$

Assuming we know the comoving distance d_4 and d_5 at redshift 4 and 5 respectively, the comoving volume V_{survey} will be the proportional volume of the shell created between redshifts 4 and 5. As the volume of a sphere is $\frac{4}{3}\pi r^3$,

$$V_{\text{survey}} = \frac{\Omega_{\text{survey}}}{4\pi} \times \frac{4\pi}{3}\left(d_5^3 - d_4^3\right) = \frac{\Omega_{\text{survey}}}{3}\left(d_5^3 - d_4^3\right) \tag{4}$$

Using the `astropy` module in `python` and the `Planck18` cosmology to calculate the above,

$$V_{\text{survey}} \approx 4.86 \times 10^5 \text{ Mpc}^3 \tag{5}$$

Number density based on assumptions

Based on the assumption that in the survey area we only found two quiescent galaxies, the observed number density is,

$$n_{\text{survey}} = \frac{2}{V_{\text{survey}}} \approx 4.12 \times 10^{-6} \text{ Mpc}^{-3} \tag{6}$$

SMF modelling and number densities

Figure 11 shows the plotting of the given Schechter function, ϕ, in log-space. Therefore, the function being plotted is,

$$\phi = \frac{dn}{d\log M_*} \tag{7}$$

To obtain the number density, we can integrate the SMF over mass. Given that we are working in log-space, this implies,

$$n = \int_{\log_{10}(M_{\text{min}})}^{\infty} \phi(\log_{10} M)\ d\log_{10} M \tag{8}$$

Using `scipy.integrate.quad`, we can integrate from the lower limit to a large number (300), as using `np.inf` would equate to 10^{∞} in linear-space and python can't compile. The results are shown in the table 4.

$\log_{10}(M_{\min}\ /\ M_{\odot})$	$n\ /\ 10^{-6}\ \mathrm{Mpc}^{-3}$
9.5	4.06
10.0	3.42
10.5	1.65

Table 4: Calculated number densities obtained by integrating the SMF in log-space.

Comparing the number density obtained through integration with the one calculated with assumption in previous sub-part, we can say that the minimum total stellar mass will be $10^{9.5}\ M_{\odot}$.

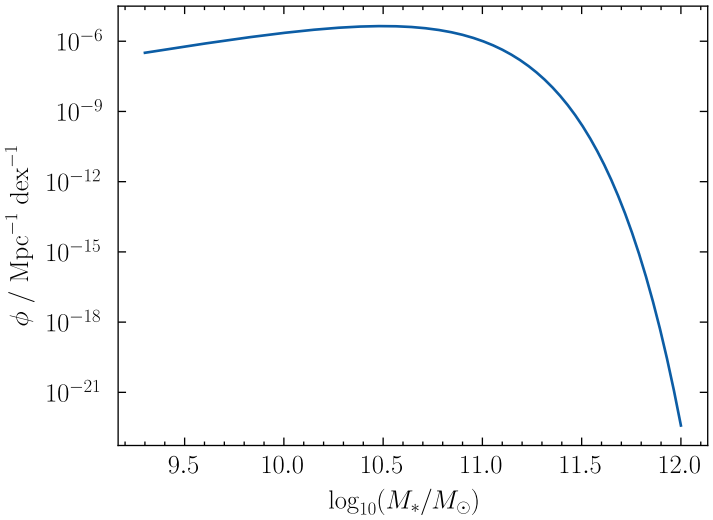

Figure 11: SMF using the given Schechter Function in a log-log space. We notice an increase in the SMF until it hits a knee, and then drops exponentially.

Stellar Mass Density

To obtain the Stellar Mass Density, we can perform similar integration to equation 8 in log-space, in the form of,

$$\rho = \int_{\log_{10}(M_{\min})}^{\infty} M^2 \ln(10) \phi(\log_{10} M)\ d\log_{10} M \tag{9}$$

Starting from a lower limit of $10^9\ M_{\odot}$, $\rho = 1.34 \times 10^5\ M_{\odot}\ \mathrm{Mpc}^{-3}$.

Part 6: Extensions

Using the double PL SFH, we also fit both the spectrum and photometry of 6620 and 8777 as well. Figure 12 shows the fitted SED for 8777. The fit is not very accurate, as we can observe in the spectra. The difficulty here might be because the drop in intensity at the Balmer break is not as significant. In fact the best fit does not even have a visible Balmer Break. One reason why this is the case because there is noise or contamination from other light sources, which increases the flux throughout the spectrum, reducing the

contrast of the features that should be observed.

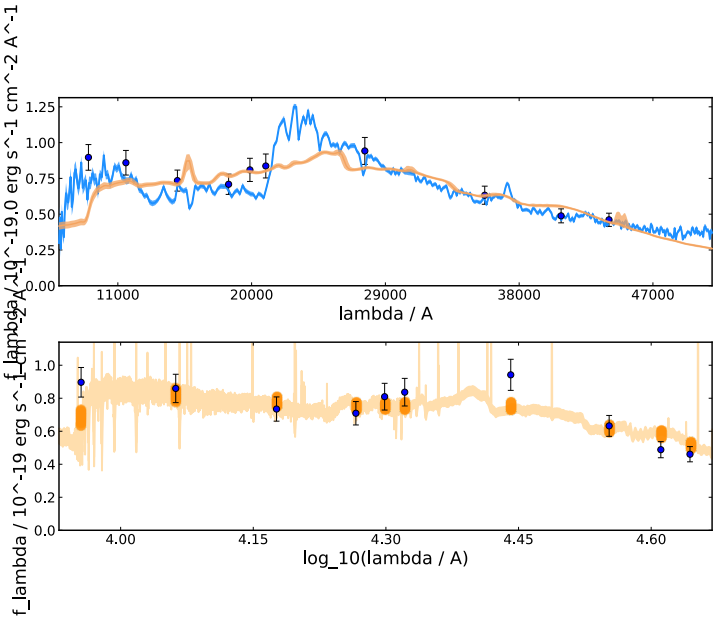

Figure 12: Best-fit and observational data SED for galaxy 8777.

Figure 13 shows the fitted SED for 6620. The photometry is much higher than the spectroscopic flux. This may be because the net amount of light entering a spectrograph is possibly lower than a photometry sensor. Therefore, the net flux observed is much higher in photometry. This reason may be true also because the shape of both graphs seem similar, with just vertical displacement between the two. Compared to the other two galaxies, we need to figure out a method to account for the difference in spectral flux density. Therefore, a direct fitting of spectra or photometry data will not provide accurate values.

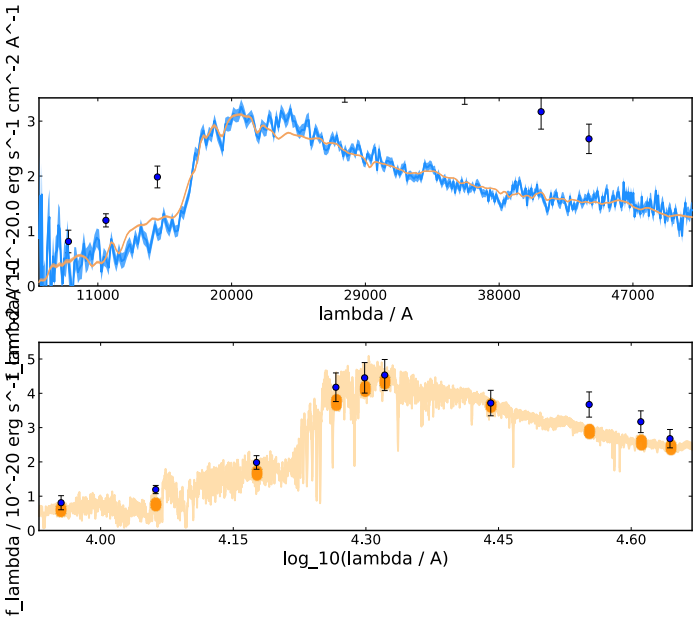

Figure 13: Best-fit and observational data SED for galaxy 6620. Most interesting feature is the difference in flux between the photometric and spectroscopic data.

Report 4: Tracing the ISM of GN20

Kushagra Shrivastava

Part 1: General Characteristics

To answer questions in this section, we create a `FlatLambdaCDM` object from `astropy.cosmology` with $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_m = 0.3$, and then use the built in functions to calculate the values.

Answer 1: At $z = 4.05$, the age of the universe was 1.494 Gyr with a look-back time of 11.97 Gyr. The numbers are different due to expansion of the universe and finite speed of light.

Answer 2: The luminosity distance of GN20 is 3.638×10^4 Mpc.

Answer 3: Milky Way's SFR and Stellar Mass is approximately 2 $M_\odot$ yr^{-1} and $5 \times 10^{10} M_\odot$, which is 1500 and 8 times less than GN20. Therefore, GN20 was a much more active galaxy as compared to our milky way.

Answer 4: The time taken by GN20 to double its stellar mass is simply,

$$t_{\text{double } M_*} = \frac{M_*}{SFR} \approx 1.333 \times 10^8 \text{ yr} \tag{1}$$

Furthermore, assuming that it has maintained a constant SFR, its current stellar mass would be,

$$M_{*,\text{ current}} = M_{*,\text{ observed}} + \text{SFR} \times t_{\text{lookback}} \approx 3.632 \times 10^{13}\ M_\odot \tag{2}$$

Part 2: Spectroscopy and derivation of Observables

Answer 1: Assuming a redshift of 4.05, the observed frequency is,

$$\nu_{\text{obs}} = \frac{\nu_{\text{rest}}}{1+z} = \frac{115.3}{1+4.05} \approx 22.83 \text{ GHz} \tag{3}$$

Using a Gaussian model defined by the equation,

$$F(\nu) = A \exp\left(-\frac{(\nu-\mu)^2}{2s^2}\right) \tag{4}$$

with mean μ and standard deviation s, the spectrum of CO(J=1-0) line from GN20 is shown in Figure 1. The best fit values for A, μ and s are 0.2890, 22.84, and 2.102×10^{-2} respectively. The errors are a few magnitudes lower than the values, and are not quoted for brevity.

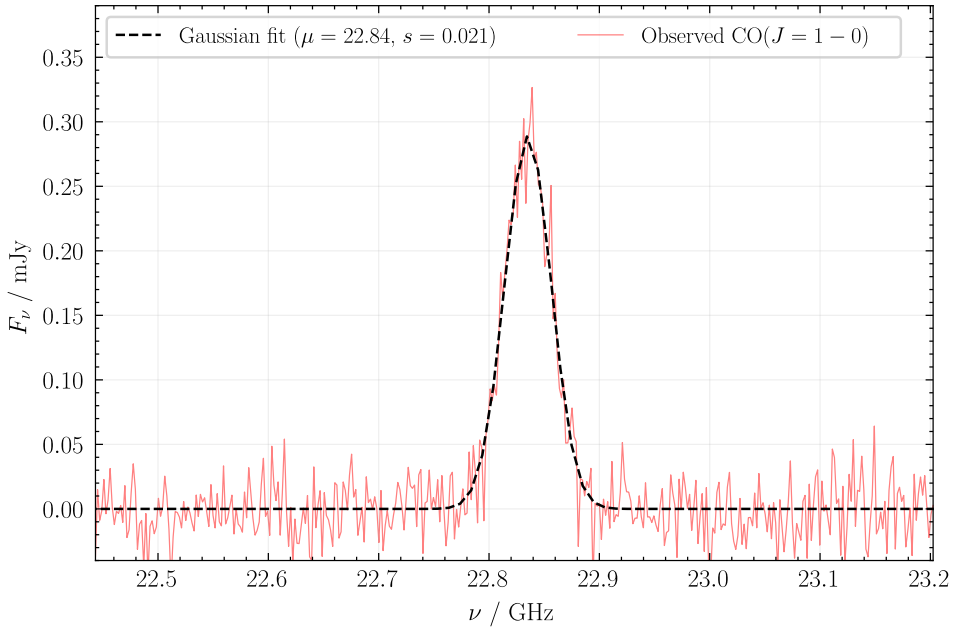

Figure 1: Spectrum of CO($J = 1 - 0$) line from GN20 obtained using the VLA, along with a fitted gaussian model.

Answer 2: The central frequency is given by μ calculated above, which is 22.84 GHz. To find the redshift, we can simply compare it with the rest wavelength,

$$z = \frac{\nu_{\text{rest}}}{\nu_{\text{obs}}} - 1 \approx \frac{115.3}{22.84} - 1 \approx 4.05 \tag{5}$$

which is the same as stated in the question.

Answer 3: Another noticeable feature is that the emission line is broadened. More than thermal broadening, Doppler broadening caused by the rotation of the galaxy plays a dominant role. Different parts of the galaxy, moves at different relative velocity from us, and in different quantities. Assuming symmetric rotation above the galaxy's axis, we expect to see a symmetrically broadened emission line.

We then convert the frequency to the rotational velocity of the galaxy using the formula, $V = c(\nu_0 - \nu)/\nu_0$. The resulting graph is shown in Figure 2.

The measured width, σ, is approximately 276 km s^{-1}. To find the velocity integrated flux density, we integrate the gaussian analytically,

$$\begin{aligned} I_{CO} = \int_{-\infty}^{\infty} A \exp\left(-\left(\frac{\nu - \mu}{\sqrt{2}s}\right)^2\right) d\nu &= \sqrt{2\pi} A\sigma \\ &\approx 199.8 \text{ mJy km s}^{-1} \\ &= 0.1998 \text{ Jy km s}^{-1} \end{aligned} \tag{6}$$

Answer 4: The Full-Width at Half Maximum of a gaussian function is approximately,

$$\text{FWHM} \approx 2\sqrt{2\ln 2}\sigma \approx 650 \text{ km s}^{-1} \tag{7}$$

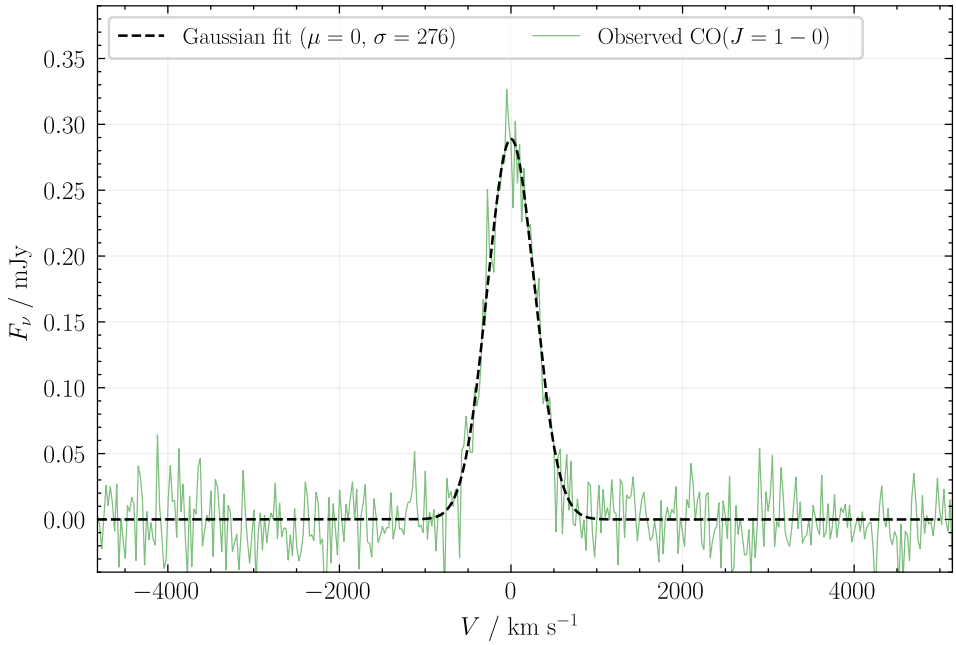

Figure 2: Spectrum of CO($J = 1 - 0$) line from GN20 obtained using the VLA, along with a fitted gaussian model in units of line-of-sight velocity of the galaxy.

Answer 5: Using the given equation, we estimate the luminosity temperature of CO,

$$L'_{\rm CO} = 3.25 \times 10^7 \frac{(S_{\rm CO}\Delta V) \times D_L^2}{\nu_{\rm obs}^2 (1+z)^3} \approx 1.28 \times 10^{11} \text{ K km s}^{-1} \text{ pc}^2 \tag{8}$$

Part 3: The molecular gas reservoir

Answer 1: First we estimate the molecular hydrogen mass using the luminosity temperature calculated in the previous part,

$$M_{\rm H_2} = \alpha_{\rm CO} \times L'_{\rm CO} = \begin{cases} 0.8 \times L'_{\rm CO}, & \text{for Starbursts} \\ 4.5 \times L'_{\rm CO}, & \text{for Normal} \end{cases} = \begin{cases} 1.02 \times 10^{11}\ M_\odot \\ 5.76 \times 10^{11}\ M_\odot \end{cases} \tag{9}$$

Given the assumption that gas is mostly found in atomic, molecular, and ionised form, and $M_{\rm H_2} >> M_{\rm HI} + M_{\rm ion}$,

$$M_{\rm gas} \approx M_{\rm H2} = \begin{cases} 1.02 \times 10^{11}\ M_\odot\ , \text{for } \alpha_{\rm CO} = 0.8 \\ 5.76 \times 10^{11}\ M_\odot\ , \text{for } \alpha_{\rm CO} = 4.5 \end{cases} \tag{10}$$

Answer 2: Given we have calculated the gas mass, and have been provided with the stellar mass, only dust mass is further required. Provided the known ratio, $M_{\rm dust} \approx M_{\rm gas}/100$, we can get the total mass,

$$M_{\rm bar} = M_{\rm stars} + M_{\rm gas} + M_{\rm dust} = \begin{cases} 5.03 \times 10^{11}\ M_\odot, \text{ for } \alpha_{\rm CO} = 0.8 \\ 9.82 \times 10^{11}\ M_\odot, \text{ for } \alpha_{\rm CO} = 4.5 \end{cases} \tag{11}$$

Answer 3: Assuming there is only gas and stellar mass components to the total mass within the effective radius and nothing else (dust, dark matter, etc),

$$M_{\mathrm{tot},R_{\mathrm{eff}}} = \frac{M_*}{2} + \frac{M_{\mathrm{H2}}}{2} \Rightarrow \alpha_{\mathrm{CO}} = \frac{1}{L'_{\mathrm{CO}}} \left(2M_{\mathrm{tot},R_{\mathrm{eff}}} - M_*\right) \approx 1.56 \tag{12}$$

Answer 4: The M_{bar} as stated in equation 11 is slightly more than double the M_{tot}. This simply implies that the rest of the mass is outside the effective radius, which is approximately double of what is inside.

Answer 5: To find how long GN20 can maintain its current SFR, we use the gas depletion scale,

$$\tau_{\mathrm{dep}} = \frac{1}{\mathrm{SFE}} = \frac{M_{\mathrm{H2}}}{\mathrm{SFR}} \approx 0.0341 \text{ Gyr} \tag{13}$$

Answer 6: The galaxy will continue to exist, however no new star-formation will take place unless gas is injected from the Intergalactic Medium. The population of stars will get concentrated towards older stars away from the Main Sequence. GN20 will become quenched and redder over time.

Part 4: Far-Infrared Emission

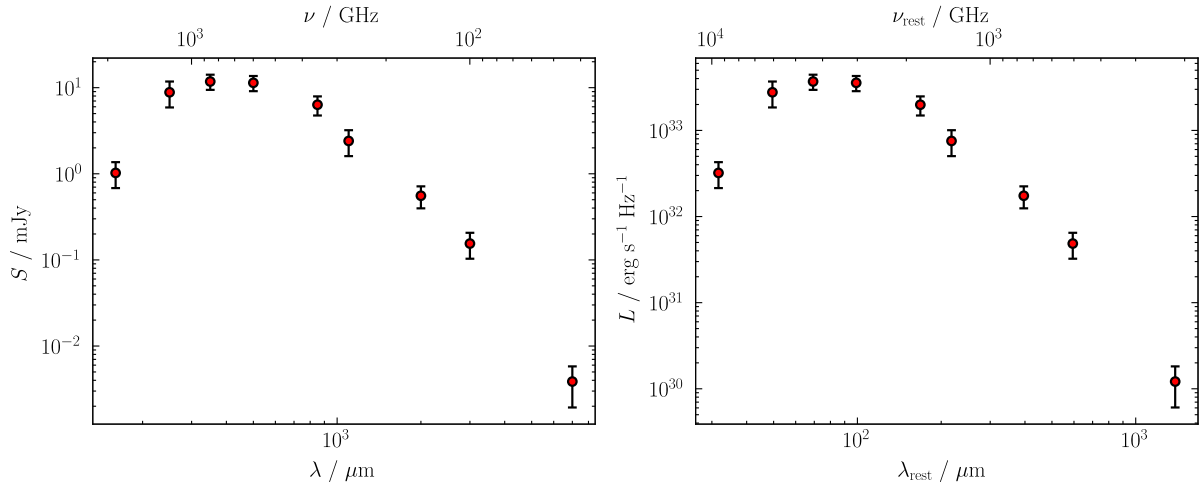

Figure 3: **(Left)** SED from photometry observations for GN20, with both wavelength and frequency plotted on the same graph. **(Right)** Observations converted to the rest-frame luminosity, wavelength, and frequency. We observe that the photometric points resemble a black-body emission curve.

Figure 3 shows the SED from photometric observations of GN20. The left panel shows the observed values, whereas the right panel shows the rest-frame luminosity,frequency, and wavelength. The luminosity was converted to the rest-frame using the following conversion,

$$L_{\mathrm{rest}} = \frac{4\pi D_L^2}{1+z} S_{\mathrm{obs}} \times 10^{-26} \text{ erg s}^{-1} \text{ Hz}^{-1} \tag{14}$$

where D_L is the luminosity distance in cm.

Answer 1: The SED resembles a black-body radiation curve. Therefore, the luminosity seen in Figure 3 Right can be best fitted using a Modified Black Body (MBB) Model, with the analyticaly equation,

$$L_\nu = A \times \frac{\nu^{4.8}}{\exp\left(\frac{h\nu}{k_B T_{\rm dust}}\right) - 1} \tag{15}$$

where A is a normalisation constant and $T_{\rm dust}$ is the dust temperature. $k_B = 1.38 \times 10^{-16}$ erg K^{-1} and $h = 6.62 \times 10^{-27}$ erg s. A and $T_{\rm dust}$ was fitted using `scipy.optimize.curve_fit`. Figure 4 shows the observed data with the best-fit MBB model. The best-fit values are shown below, with the errors being equal to 1 standard deviation.

$$A = (2.0 \pm 0.3) \times 10^{-25} \text{ erg s}^{-1} \text{ Hz}^{-5.8} \tag{16}$$

$$T_{\rm dust} = 38 \pm 1 \text{ K} \tag{17}$$

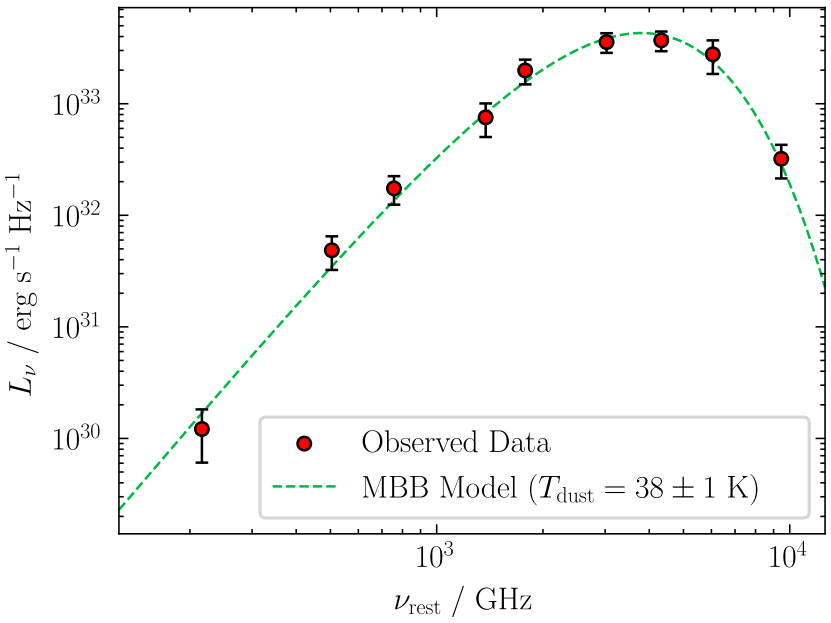

Figure 4: GN20 SED with a MBB Model best-fit. The best-fit temperature is 38 ± 1 K.

Answer 2: As $T_{\rm dust} = 38 \pm 1$ K, we can verify this value using the Wien's Displacement Law. The peak frequency should be,

$$\lambda_{\rm peak} = \frac{b}{T_{\rm dust}} \Rightarrow \nu_{\rm peak} = \frac{c}{b} T_{\rm dust} = 3900 \pm 100 \text{ GHz} \tag{18}$$

From the best fit model, $\lambda_{\rm peak} = 3806$ GHz, which is in agreement with the estimate from the Wien's Displacement Law within the error bounds.

Answer 3: At 757 GHz, the luminosity from the best fit model is 1.36×10^{32} erg s^{-1} Hz^{-1}. We can estimate the dust mass at this frequency using,

$$M_{\rm dust} = \frac{L_\nu}{4\pi \kappa_\nu B_\nu(T_{\rm dust})} \tag{19}$$

where $B_v(T_{\rm dust})$ is the Planck function also evaluated at 757 GHz. Plugging in the values, we obtain,

$$M_{\rm dust} \approx 8.6 \times 10^8 \ M_\odot \tag{20}$$

Answer 4: Now, we can calculate the gas-to-dust ratio. For this ratio for use a variety of $\alpha_{\rm CO}$.

$$\delta_{\rm gd} = \begin{cases} 119, \text{if } \alpha_{\rm CO} = 0.8 \\ 669, \text{if } \alpha_{\rm CO} = 4.5 \end{cases} \tag{21}$$

Answer 5: To find the total luminosity in the infrared range, we can integrate equation 15. Using `scipy.integrate.quad`, the obtained value is,

$$L_{\rm IR} = \int_{c/(1000\times10^{-6}\ \rm Hz)}^{c/(8\times10^{-6}\ \rm Hz)} L_\nu \ d\nu \approx 1.95 \times 10^{46} \text{ erg s}^{-1} \approx 5.09 \times 10^{12} \ L_\odot \tag{22}$$

Answer 6: Finally, we can calculate the SFR of GN20 using,

$$\text{SFR} = 1.49 \times 10^{-10} \times L_{\rm IR} \approx 759 \ M_\odot \ \text{yr}^{-1} \tag{23}$$

Part 5: CMBR effect on Dust Observations

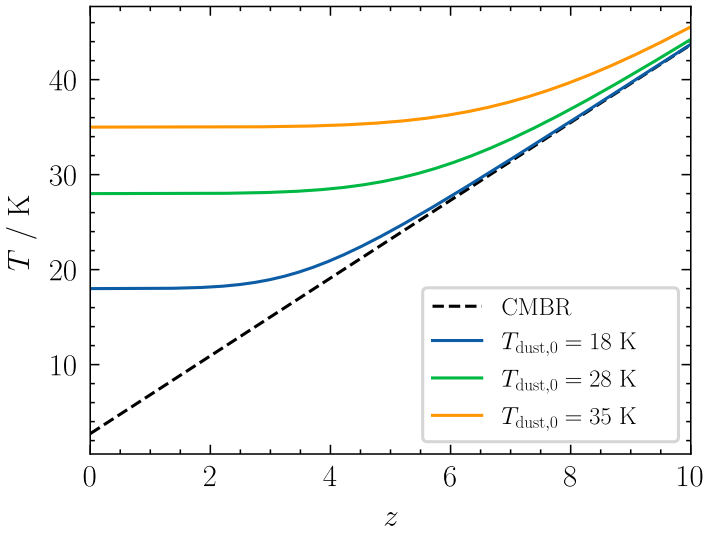

Figure 5: Temperature variation of dust grains accounting for CMBR and redshift. The temperature is more sensitive at smaller values of $T_{\rm dust,0}$.

In this part, we calculate the effect of the CMBR on our observations. The CMBR and dust temperature adjusted can be approximated using,

$$T_{\rm CMB}(z) = 2.73 \times (1+z) \tag{24}$$

$$T_{\rm dust}(z) = \left\{ T_{\rm dust,0}^{4+\beta} + 2.73^{4+\beta} \left[(1+z)^{4+\beta} - 1 \right] \right\} \tag{25}$$

The plots for the above equations are shown in Figure 5.

Answer 1: If the intrinsic temperature is low, we expect to see a non-negligible effect of CMBR at lower redshifts. As the intrinsic temperature increases, the effect becomes significant only at higher redshifts.

Now, we plot the ratio of the observed flux against CMBR, as shown in Figure 6.

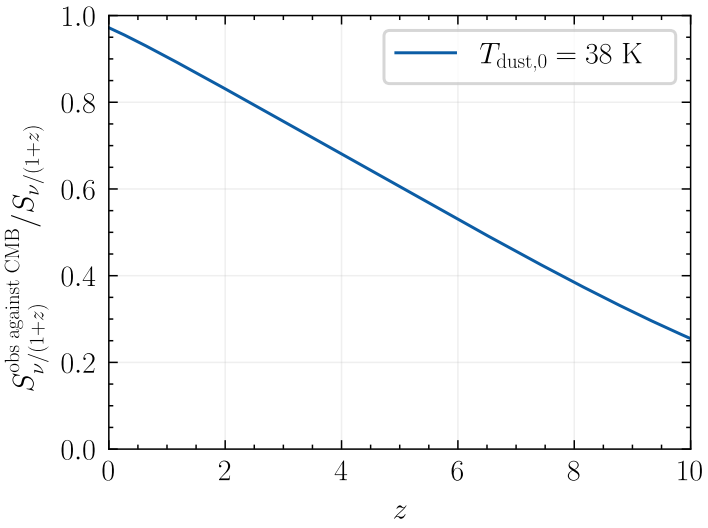

Figure 6: Fraction of flux observed against the CMBR. As the redshift increases, we observe an even lower flux.

Answer 2: At about a redshift of 4, we only observe 0.7 of the intrinsic flux density, and thus we miss about 30%. This was calculated at an observation wavelength of 3mm, which lies in the microwave regime. As CMBR lies in the microwave regime as well, these wavelengths will be the most affected.

Now using a new model defined by,

$$L_\nu^{\text{obs against CMB}} = 4\pi\kappa_\nu M_{\text{dust}} \left\{B_\nu(T_{\text{dust}}(z)) - B_\nu(T_{\text{CMB}}(z))\right\} \tag{26}$$

we can refit our data.

Answer 3: Visually, the fit appears similar to the previous one. Furthermore given the error on the photometry, the final temperature is also the same within the error bounds. However, let us ignore the error for a while. The new $T_{\text{dust}}^{\text{CMB}} = 37.7$ K as compared to 38.4 K from the non-CMBR fit. Also, the mass of dust, $M_{\text{dust}} \approx 9.5 \times 10^8\ M_\odot$, which is larger than the previous mass. This mass is much close to the expected $10^9\ M_\odot$.

Answer 4: The fraction of the true flux density missed is the same as the intrinsic flux density missed, as stated in Answer 2.

Report 5: Galaxy Dynamics

Kushagra Shrivastava

Part 1: From Mass Profiles to Velocity Curves

Under the assumption of no pressure support and a spherical symmetry (at least valid at a first-order for thin disk systems), the relationship between circular velocity and mass profiles can be written as,

$$v_c(R) = \sqrt{\frac{GM(R)}{R}} \tag{1}$$

where $M(R)$ is the mass enclosed within a radius R. The different velocity profiles can be created by assuming different mass distributions.

Point-like: In this distribution, the whole mass is concentrated at the center of the galaxy. For such a distribution, it would be unphysical to define any velocity at $R = 0$ kpc. For $R > 0$, the velocity profile simply becomes,

$$v_{\text{point-like}}(R) = \sqrt{\frac{GM_*}{R}} \tag{2}$$

where M_* is the total mass.

Constant Density: If the density is constant, ρ_0, we can use the volume of a sphere to estimate the mass distribution (again, simple assumption of galaxies being spherical).

$$M(R) = \rho_0 V(R) = \frac{4}{3}\pi R^3 \rho_0 \Rightarrow v_{\text{cost. }\rho}(R) = \sqrt{\frac{4}{3}G\pi\rho_0 R} \tag{3}$$

Therefore, the circular velocity and radius are directly proportional. Now, assuming the density is constant upto a certain radius R_0, after which it becomes 0, we expect to see a Keplerian profile beyond R_0. Therefore,

$$v_{\text{cost. }\rho}(R) = \begin{cases} \sqrt{\frac{4}{3}G\pi\rho_0 R}, & \text{for } R \leq R_0 \\ \sqrt{\frac{GM_*}{R}}, & \text{for } R > R_0 \end{cases} = \begin{cases} \sqrt{\frac{GM_*}{R_0^3}}R, & \text{for } R \leq R_0 \\ \sqrt{\frac{GM_*}{R}}, & \text{for } R > R_0 \end{cases} \tag{4}$$

Exponential Profile: Suppose the brightness of a galaxy can be modeled as a Sérsic Profile with $n = 1$. Then, the Mass profile can be written as,

$$M(R) = M_* \left[1 - e^{-R/R_d}\left(1 + \frac{R}{R_d}\right)\right] \tag{5}$$

where R_d is the scale radius. Therefore the velocity profile becomes,

$$v_{\text{exp}}(R) = \sqrt{\frac{GM_*}{R}\left[1 - e^{-R/R_d}\left(1 + \frac{R}{R_d}\right)\right]} \tag{6}$$

Flat-curve: In order to produce a flat rotation curve, we need a density profile which eliminates the radius dependence of the velocity profile.

$$v_{\text{flat}}(R) = \sqrt{\frac{G}{R}\left(\frac{4}{3}\pi R^3\right)\rho(R)} = \sqrt{\frac{4\pi G}{3}R^2\rho(R)} \tag{7}$$

To remove the R dependence,

$$R^2\rho(R) \propto \text{constant} \Rightarrow \rho(R) \propto \frac{1}{R^2} \Rightarrow \rho(R) = \frac{A}{R^2} \tag{8}$$

where A is a constant. This further implies that the mass profile is,

$$M(R) = \frac{4}{3}\pi A R \tag{9}$$

For plotting purposes, let us assume we obtain the total mass of the galaxy within 15 kpc. Therefore,

$$A = \frac{3}{4\pi}M_* \times \frac{1}{15\text{ kpc}} \Rightarrow v_{\text{flat}}(R) = \sqrt{\frac{GM_*}{15\text{ kpc}}} \tag{10}$$

The velocity profiles for the abovementioned mass profiles are shown in Figure 1.

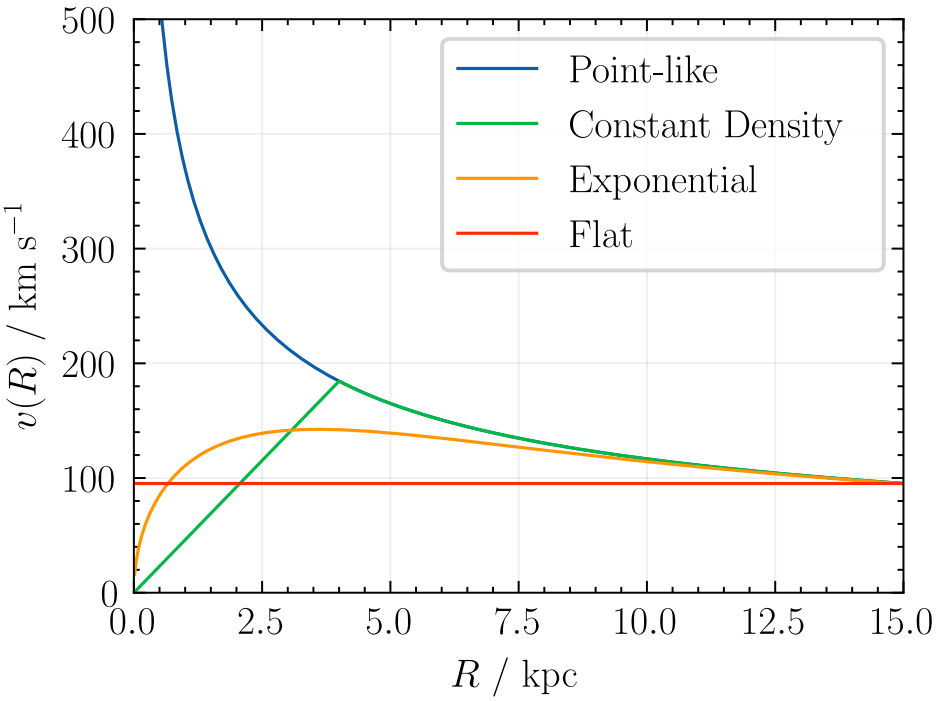

Figure 1: Plots of the different velocity profiles arising from different mass profiles where $M_* = 10^{10.5}\ M_\odot$, $R_0 = 4$ kpc, and $R_d = 2$ kpc. We notice how all profiles converge at larer radii.

The similarity across all the profiles is that they converge at the same value at larger radii. However, they are very different at smaller radii.

The constant density and exponential profiles are similar, with the main difference being the sudden kink present in the constant density profile, whereas the exponential profile

have a softer turnover. The Point-like profile is the least realistic, given it reaches extremely high circular velocities at small radii.

For the flat profile, if we see it even when we no longer can account for the visible starlight, it implies some other massive factor must be contributing to the rotation. This massive factor, would be called Dark Matter.

Part 2: Mapping Velocity Fields

Now, we extend the argument from the previous part to a 2 dimensional analog. Suppose the galaxy is circular, defined by a xy-cartesian plane. The x_{gal} and y_{gal} satisfy the relation,

$$x_{\text{gal}}^2 + y_{\text{gal}}^2 = R^2 \tag{11}$$

where R is the radial distance from the center of the galaxy. Let us assume that the center of the galaxy lies at $(0,0)_{\text{gal}}$. Now we make a further assumption that the sky is defined by another xy-cartesian plane, and the x-axis of both the planes are the same (this ensures that the semi-major axis of the projection later on always lies on the x-axis).

Let the inclination be i, and therefore we can project the galactic plane on the sky place,

$$x_{\text{sky}} = x_{\text{gal}} \tag{12}$$

$$y_{\text{sky}} = y_{\text{gal}} \cos(i) \tag{13}$$

$$\sqrt{x_{\text{sky}}^2 + \left(\frac{y_{\text{sky}}}{\cos(i)}\right)^2} = R \tag{14}$$

Given we have the position mapping, we can now map the velocity. In the galactic plane,

$$v_{\text{gal}}(R,\theta) = v(R) \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \tag{15}$$

Projecting the velocity on the plane perpendicular to the sky plane (i.e., the plane on our line of sight),

$$v_{\perp,\text{sky}}(R,\theta,i) = v(R)\sin(i) \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \tag{16}$$

The line-of-sight velocity observed will be in our direction. Therefore,

$$v_{\text{los}}(R,\theta,i) = v(R)\sin(i)\cos(\theta) \tag{17}$$

$$\Rightarrow v_{\text{los}}(x_{\text{sky}}, y_{\text{sky}}, i) = v\left(\sqrt{x_{\text{sky}}^2 + \left(\frac{y_{\text{sky}}}{\cos(i)}\right)^2}\right) \sin(i) \frac{x_{\text{sky}}}{\sqrt{x_{\text{sky}}^2 + \left(\frac{y_{\text{sky}}}{\cos(i)}\right)^2}} \tag{18}$$

Using equations 18 and 14, we created the maps shown in Figure 2 with `bilinear` interpolation in `matplotlib.pyplot.imshow`.

The Constant Density profile showcases a linear distribution of velocity along the x-axis, with the contours being straight in the y-axis. This is expected as everything moves

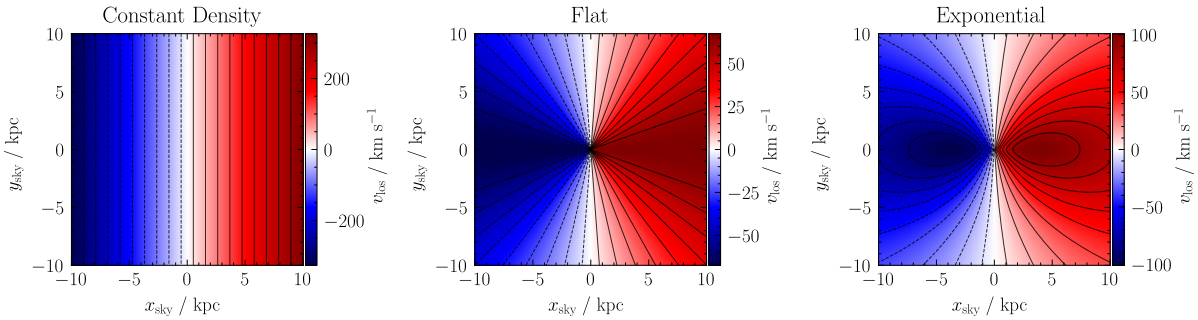

Figure 2: Plots of the different velocity profiles arising with velocity contours with inclination of 45°.

together for a solid body rotation.

The Flat rotation curve simply have straight lines from the center of the galaxy. This is expected as the line-of-sight velocity along each radius should be the same. It is different from the Constant Density profile because the motion is not a solid-body rotation but different distances from the center moves at a different speed.

The Exponential profile has closed iso-velocity contours, because in an Exponential profile the velocity increases and then decreases again. Therefore along each R, we expect to see two points with the same velocity. However closed curves become larger because of the effect of inclination and the azimuthal angle. Furthermore, the $R_{\rm max}$ occurs at 3.59 kpc which is less than $R_d = 2$ kpc. This implies that the scale radius is not the same as the maximum velocity radius.

Part 3: Deriving Rotation Curves from Kinematic Data

Answer 1: The systemic velocity of NGC 3198 is 660.7 km s^{-1}. Therefore, the redshift is,

$$z \approx \frac{v}{c} \approx 2.2 \times 10^{-3} \tag{19}$$

As the systemic velocity is much smaller than the speed of light. The above approximation is enough. If we were to use the relativistic formula,

$$z_{\rm rel} = \sqrt{\frac{1+v/c}{1-v/c}} - 1 \approx 2.2 \times 10^{-3} \tag{20}$$

The fractional difference between z and $z_{\rm rel}$ is merely 0.1%.

Answer 2: Using the given velocity data file, `NGC_3198_NA_MOM1_THINGS.FITS`,we obtain the line-of-sight velocity $v_{\rm los}$ by accounting for the systematic velocity, $v_{\rm los} = v_{\rm given} - v_{\rm sys}$

After the velocity, we convert the pixels to kpc using the pixel scale 1.5″/px and the redshift calculated in the previous part. Under a `FlatLambdaCDM` cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_m = 0.3$, the conversion scale is calculated using `kpc_proper_per_arcmin`. Therefore,

$$\text{pixel-to-kpc} = \frac{1.5}{60} \times \text{arcmin-to-kpc}(z) \tag{21}$$

To make the center of the galaxy as $(0, 0)$ in the concerned plane, we use the central pixels provided in the FITS file `CRPIX1` and `CRPIX2` and scale everything accordingly. The velocity profile with isovelocity contours is shown in Figure 3. The velocity profile is expected, as one region approaching towards us and one moves away from us. The isovelocity contours have the closest resemblance to a exponential profile, but without the closed contours. In the regions away from the center, the contours resemble the flat profile.

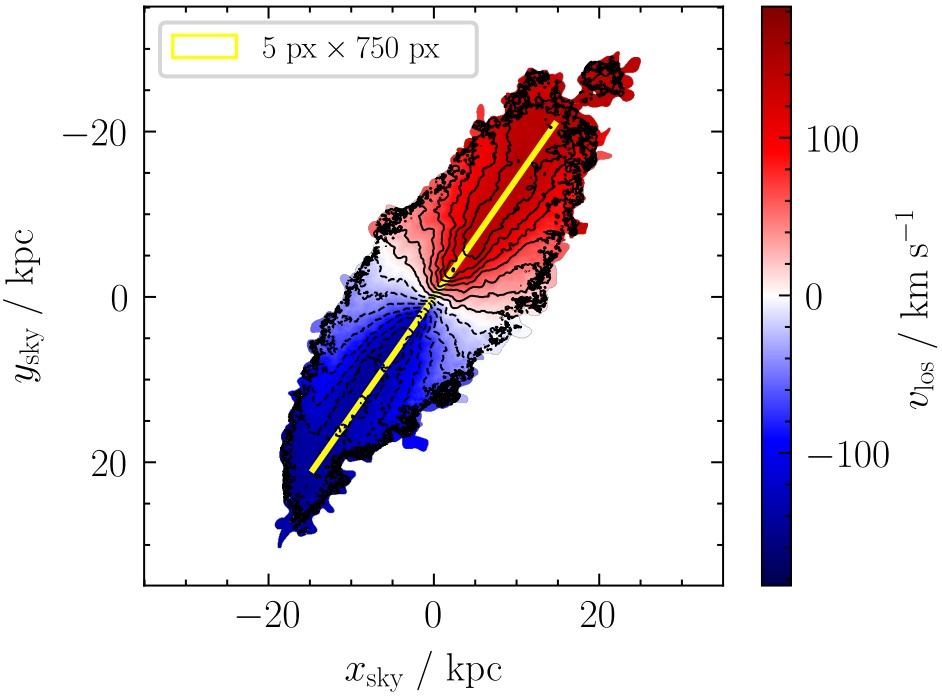

Figure 3: Velocity profile of NGC3198 corrected for systematic velocity, with overplotted isovelocity contours. The pixels are scaled to kpc. The long-slit with dimensions of 5 px by 750 px used for the long-slit spectroscopy is also shown here.

Answer 3: Using the `RectangularAperture` from the `photoutils.aperture` we create a 5 px by 750 px wide slit about the major axis of the galaxy. The rotation angle was the provided Position Angle of the galaxy. The length of the slit was chosen such that it does not expand in the regions where the isovelocity contours were 'turbulent'. The slit is overplotted in Figure 3.

Within the slit, we create a mask such that only pixels that are fully inside the slit are accounted by (using the `center` method in the `to_musk` function), and everything else is set to `NaN`. Using the center of the slit, we wrap the receding and approaching part and then concatenate them. After extraction, the circular velocities were calculated by accounting for the inclination, $v_c = v_{\rm los}/\sin i$. We do not need to worry about the position angle as we are extracting only about the major axis. Lastly, the median, $16^{\rm th}$ and $84^{\rm th}$ percentiles were calculated. The percentiles were used to calculate the 1-σ errorbars. To calculate the distance from the center, the same pixel-to-kpc ratio was used. The rotation curve is shown in Figure 4.

We observe an increasing velocity away from the center, and then at about 10 kpc it

becomes a flat curve.

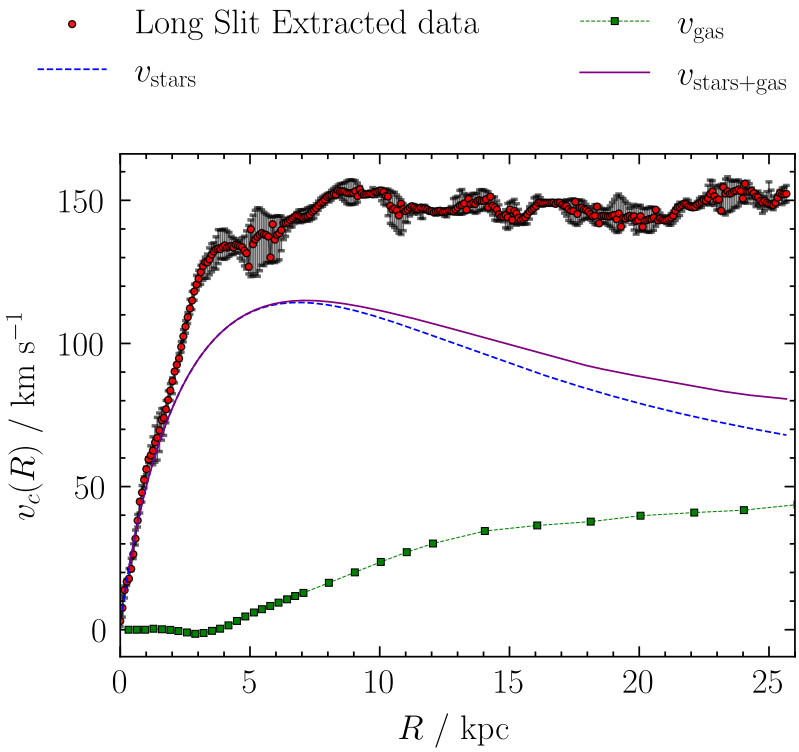

Figure 4: Rotation curve of NGC3198. The observed data is calculated from long slit spectroscopy, and we notice a flat curve at high radius. The stars, gas, and stars+gas contributions are also shown, and they are unable to account for the flat part of the observed rotation curve.

Answer 4 and 5: Using the thin disk approximation, the star contribution was estimated,

$$v_{\rm stars}(R) = \sqrt{4\pi G\Sigma_0 R_d y^2 \left[I_0(y)k_0(y) - I_1(y)k_1(y)\right]} \tag{22}$$

with $y = R/2R_d$, $\Sigma_0 = M_*/(2\pi R_d^2)$, M_* is stellar mass, and I_i and $k_i$are the modified bessel functions of the first and second kind respectively. The obtained velocity is shown in Figure 4. After about 2 kpc, the velocity becomes much less than the observed velocity.

After using the given velocities from the gas-component (shown in Figure 4), we can calculate the total velocity from the stars and gas,

$$v_{\rm stars+gas} = \sqrt{v_{\rm stars}^2 + v_{\rm gas}^2} \tag{23}$$

Note that the velocity contributions were squared then added because we are adding the gravitational potential from each of them. Even the combined contribution, shown in Figure 4 cannot account for the observed velocity.

Answer 6: The dark matter fraction ($f_{\rm DM} = 1 - (v_{\rm bar}/v_c)^2$) was calculated at different radii, and the results are shown in Table 1. Here $v_{\rm bar} = v_{\rm stars+gas}$. As the radius increases, the dark matter fraction increases. This implies there is a greater fraction of dark matter as compared to the baryons away from the center of the galaxy.

R	$1R_d$	$3R_d$	$5R_d$
f_{DM}	0.42	0.46	0.57

Table 1: Dark Matter Fraction calculated at different scale radii, where $R_d = 3.2$ kpc.

Part 4: Dynamical Constraints on Dark Matter Distributions

Answer 1: From the velocities obtained in the previous part, we can extract the Dark Matter velocity,

$$v_{\rm DM} = \sqrt{v_C^2 - v_{\rm stars+gas}^2} \tag{24}$$

The same was applied to the errors. For this part, instead of 16th and 84th percentile, the average of the both errors were used to simplify curve fitting.

The NFW profile is a density profile,

$$\rho_{\rm NFW}(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2} \tag{25}$$

where ρ_s and r_s are scale density and radius respectively. Integrating over a volume, we obtain the mass destribution,

$$M_{\rm NFW}(r) = 4\pi\rho_s r_s^3 \left[\ln\left(1+\frac{r}{r_s}\right) - \frac{r/r_s}{1+r/r_s}\right] \tag{26}$$

We can rewrite this using virial radius and virial mass. We will be fitting the concentration, c, and the virial mass, M_{200}, later on. Therefore,

$$M_{200} = \frac{4\pi}{3} r_{200}^3 200 \left(\frac{3H_0^2}{8\pi G}\right) \Rightarrow r_{200} = \left(\frac{GM_{200}}{100H_0^2}\right)^{1/3} \tag{27}$$

$$r_s = r_{200}/c \tag{28}$$

Substituting r_{200} in the mass profile,

$$M_{200} = 4\pi\rho_s \frac{GM_{200}}{100H_0^2c^3}\left[\ln(1+c) - \frac{c}{1+c}\right] \Rightarrow \rho_s = \frac{25H_0^2c^3}{\pi G\left(\ln(1+c) - (c/(1+c))\right)} \tag{29}$$

Given that r_s and ρ_s are now known using the fitting parameters c and M_{200}, the density profile can also be calculated. The fitting parameters will be calculated using the circular velocity model,

$$v_{\rm DM}(r) = \sqrt{\frac{GM_{\rm NFW}(r)}{r}} \tag{30}$$

The fitting was performed using the χ^2-minimization with `minimize` from `scipy.optimize`. The reduced-χ^2 parameter was then calculated using,

$$\chi_R^2 = \frac{\chi_{\rm min}^2}{n-f} \tag{31}$$

where $\chi_{\rm min}^2$ is the minimized value from the optimisation, n is the number of datapoints being fitted, and f is the parameters being fitted (2 in this case). The best fit curve is shown in the Figure 5. The best-fit values are $c = 5.2$ and $M_{200} = 7.6 \times 10^{11}\ M_\odot$. Comparing to Blok et al. (2018), where $c = 5.1$ and $M_{200} = 3.0 \times 10^{11}\ M_\odot$ (note that

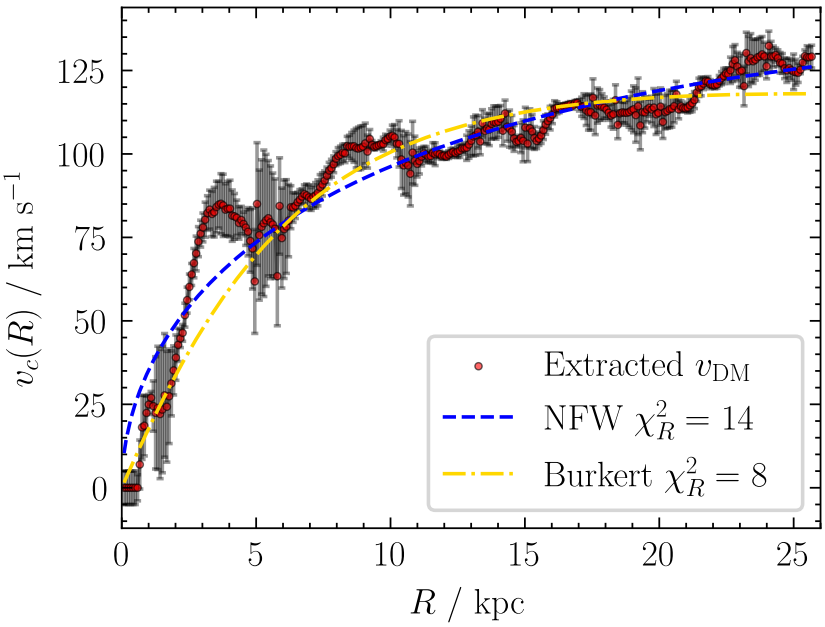

Figure 5: Rotation curve of NGC3198 with only Dark Matter contribution. The observed data is calculated from the quantities calculated in the previous part. The best fit values for NFW and Burkert profiles are $c = 5.2$, $M_{200} = 7.6 \times 10^{11}$ $M_\odot$,$r_0 = 7.9$ kpc, and $\rho_0 = 1.9 \times 10^7$ $M_\odot$ kpc^{-3}. The χ^2_R is similar for both models.

V_{200} is shown in the research paper, but we calculated the M_{200} here). The values are similar to each other. The reduced χ^2_R values are close to 1, with Burkert being a better fit. This conclusion is also supported by visual observation of the graph.

Answer 2: We also fitted the Burkert profile defined by,

$$\rho_{\rm Burkert}(r) = \frac{\rho_0}{(1 + r/r_0)(1 + (r/r_0)^2)} \tag{32}$$

where the best fit parameters are r_0 and ρ_0 respectively. Using the same method in the previous part, the model was fit to the data as shown in Figure 5. The best fit values are $r_0 = 7.9$ kpc and $\rho_0 = 1.9 \times 10^7$ $M_\odot$ kpc^{-3}. Using the best fit values for both profiles, the density profiles are shown in Figure 6.

The NFW profile shows a exponential increase as we go towards the center, which is the cusp of the profile. The Burkert profile is flat showcasing the core of the profile. Both profiles diverge significantly across all values, but especially at the smaller radii and very high radii.

Answer 3 and 4: Using the same procedure as before, NGC0247 and NGC3109 dark matter velocities were fitted with the NFW and Burkert profiles. The plots are shown in Figure 7 and Figure 8. The best fit parameters are shown in Table 2.

The NFW profiles fits NGC0247 much better than NGC3109, while the Burkert profile fits NGC3109 much better as compared to NGC0247. This implies that the dark matter distribution is cusped in NGC0247 and cored in NGC3109. The difference in the dark matter densities are due to the fact that NGC0247 is a spiral galaxy, while NGC3109 is a dwarf galaxy.

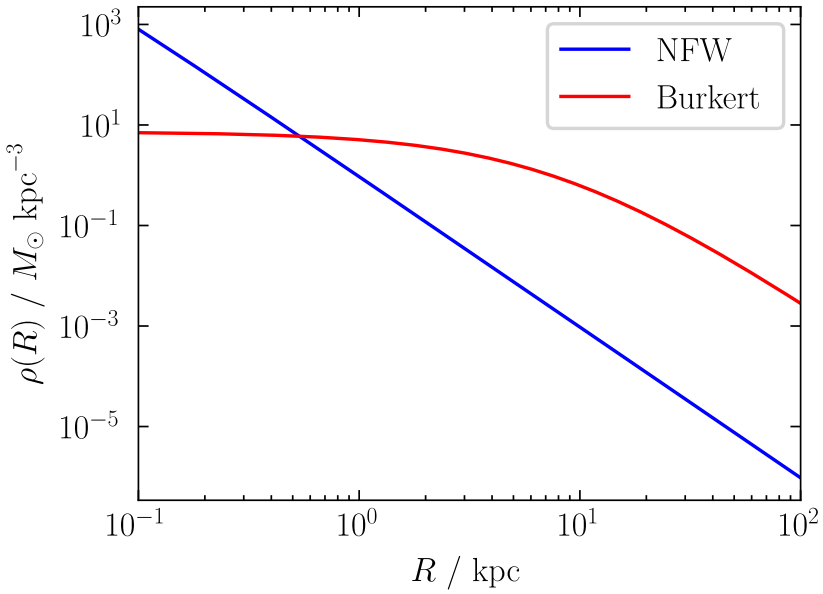

Figure 6: Density profiles using NGC3198's best fit parameters. The NFW profile shows a exponential increase as we go towards the center, which is the cusp of the profile. The Burkert profile is flat showcasing the core of the profile.

Galaxy	c	M_{200} / $M_\odot$	r_0 / kpc	ρ_0 / $M_\odot$ kpc^{-3}
NGC0247	5.1	$10^{11.4}$	3.2	$10^{7.6}$
NGC3109	0.34	$10^{13.7}$	5.1	$10^{7.3}$

Table 2: Best fit values for different DM profiles

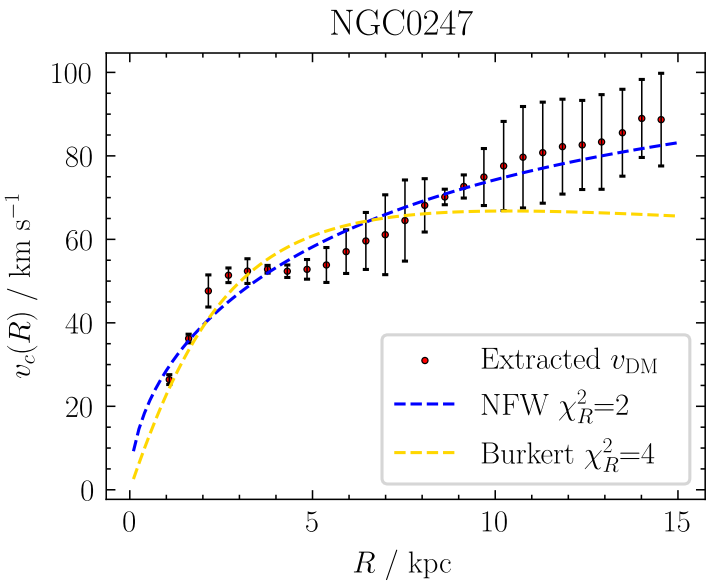

Figure 7: Rotation curve of NGC0247 with only Dark Matter contribution. The observed data is calculated from the quantities calculated in the previous part. The NFW profile fits it much better.

Answer 5: To calculate the total Dark Matter mass, we took M_{200} for the NFW profile and the maximum calculated mass value for the Burkert profile. The NFW profile was

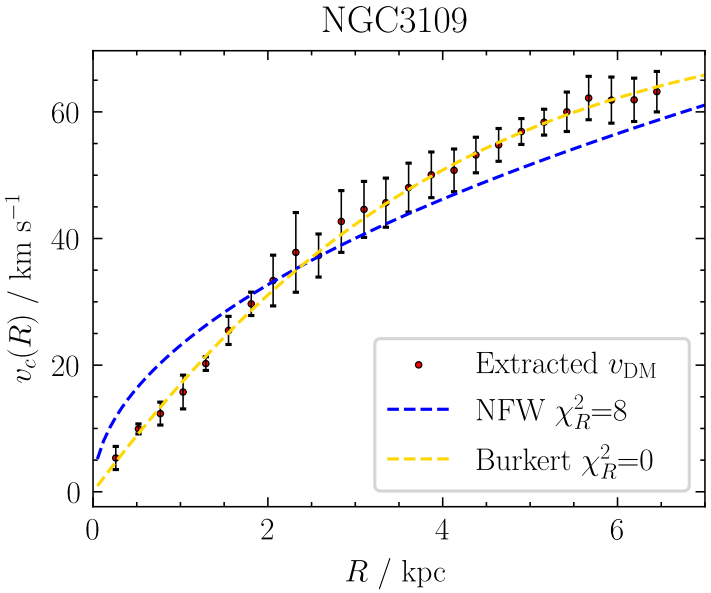

Figure 8: Rotation curve of NGC3109 with only Dark Matter contribution. The observed data is calculated from the quantities calculated in the previous part. The Burkert profile fits it much better.

used for NGC0247 and Burkert for NGC3109. Therefore, $M_{\rm DM,NGC0247} = 2.6 \times 10^{11} M_\odot$ and $M_{\rm DM,NGC3109} = 5.0 \times 10^{13}\ M_\odot$. The stellar to halo mass fraction and baryon fraction are shown in 3. Given the much smaller $f_{\text{star-halo}}$ for NGC3109 compared to NGC0247, it implies the formers star formation efficiency is very low. This is expected, as it is a drawf galaxy and likely not star forming, whereas the former is a spiral galaxy which is likely star-forming. Furthermore, the baryon fraction is much smaller than the cosmic baryon fraction 0.16. This showcases the missing baryon problem is very real in these galaxies, especially in NGC3109.

Galaxy	$f_{\text{star-halo}}$	f_b
NGC0247	0.0051	0.007
NGC3109	2.5×10^{-6}	1.3×10^{-5}

Table 3: Calculated fractions for different galaxies.

Report 6: Reionized Bubbles around Primordial Galaxies

Kushagra Shrivastava

Part 1: Recombination Lines in Astrophysical Nebulae

Answer 1.1: To calculate the Rydberg constant, let us consider the Hydrogen Atom transition from energy level 2 to 1. In the classical Bohr's model, this will correspond to an energy release of,

$$\Delta E_{2\to1} = E_1 - E_2 = \frac{-13.6}{1^2} - \frac{-13.6}{2^2} = -10.2 \text{ eV} \tag{1}$$

Using the Planck frequency-energy quantization we can calculate the photon wavelength that will be released due to this transition,

$$-\Delta E_{2\to1} = \frac{hc}{\lambda} \Rightarrow \frac{1}{\lambda} = \frac{-\Delta E_{2\to1}}{hc} = \frac{10.2 \times 1.6 \times 10^{-19}}{6.626 \times 10^{-34} \times 299792458} \approx 8.22 \times 10^6 \text{ m}^{-1} \tag{2}$$

Finally using the definition of the Rydberg constant,

$$\frac{1}{\lambda} = R_H \left(\frac{1}{1^2} - \frac{1}{2^2}\right) \Rightarrow R_H = \frac{8.22 \times 10^6}{0.75} \approx 1.1 \times 10^7 \text{ m}^{-1} \tag{3}$$

Answer 1.2: Using the value of R_H calculated before, we calculate the principal wavelength lines for different series using,

$$\lambda = \frac{1}{R_H}\left(\frac{1}{m^2} - \frac{1}{n^2}\right)^{-1} \tag{4}$$

The principal wavelengths are presented in Table 1.

Line	Wavelength / nm
Lyman (Lyα)	121.6
Balmer (Hα)	656.5
Paschen (Paα)	1875.6

Table 1: Primary lines wavelengths for different transitions in the Hydrogen atom.

The lines mentioned above will be redshifted by the time they reach us. The relation between the observed wavelength and the emitted wavelength is,

$$z = \frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} - 1 \tag{5}$$

where z is the redshift. The Low-resolution PRISM on JWST observes in the wavelength range of 0.60 to 5.30 μm, and the G235M grating in the range of 1.66 to 3.07 μm. Using equation 5 we can calculate the minimum and maximum redshift corresponding to the minimum and maximum range of the instruments. The calculated values are shown in Table 2.

Line	PRISM	G235M
Lyα	3.9 - 42.6	12.7 - 24.2
Hα	0* - 7.1	1.5 - 3.7
Paα	0* - 1.8	0* - 0.6

Table 2: Redshift observation range for different Hydrogen lines using PRISM and G235M gratings on the NIRSpec instrument on JWST. *If the redshift range starts from 0, it implies the lower limit of the grating is less than the line rest wavelength.

To compare with the corresponding transitions for the helium ion (He^+), we first need to calculate the Rydberg constant for He II. Given the electronic structure is the same, we can use the Rydberg formula. As $Z = 2$ for He II, $R_{\text{He II}} = Z^2_{\text{He II}} R_H = 4R_H$.

Therefore,

$$\lambda_{\text{He II}} = \frac{1}{4R_H}\left(\frac{1}{m^2} - \frac{1}{n^2}\right)^{-1} = \frac{\lambda_H}{4} \tag{6}$$

Therefore each observed electronic transition has a wavelength quarter to that of Hydrogen's. Therefore, $2 \to 1$, $3 \to 2$, and $4 \to 3$ transitions are 30.4, 164.1, and 468.9 nm respectively.

Answer 1.3: By energy conservation, the energy of the resulting photon (assuming the event immediately resulted in a ground state hydrogen) is,

$$E_\gamma = E_{\text{initial}} - E_{\text{final}} = K_e - E_{\text{H, n=1}} = K_e + 13.6 \text{ eV} \tag{7}$$

As the kinetic energy of an electron is continumm, the energy of a photon is,

$$E_\gamma \geq 13.6 \text{ eV} \tag{8}$$

As calculated here the energy of these photons are greater than 13.6 eV, and therefore, these photons can immediately reionise other neutral atoms in the region, 'on-the-spot'. Therefore the Case-B recombination's on-the-spot approximation is valid, as we do not have many hydrogen atoms in the ground state. The Hydrogen atoms are either ionised, or stay in an excited state and then cascade down to $n = 1$ state.

Answer 1.4: To check whether the H II region would be ionisation or density bounded, we calculate the Strömgren radius using,

$$R_S = \left(\frac{3\dot{N}_{\text{ion}}}{4\pi n_H^2 \alpha_{\text{rec}}}\right)^{1/3} \tag{9}$$

where $\alpha_{\text{rec}} = 2.54 \times 10^{-13}(T_e/10^4)^{-0.8}$. As the radius of H II region is $r = 2$ pc, the region would be ionisation bounded if $r > R_S$ and density bounded if $r < R_S$. As the Strömgren radius is parametrized by the electron temperature, the transition temperature

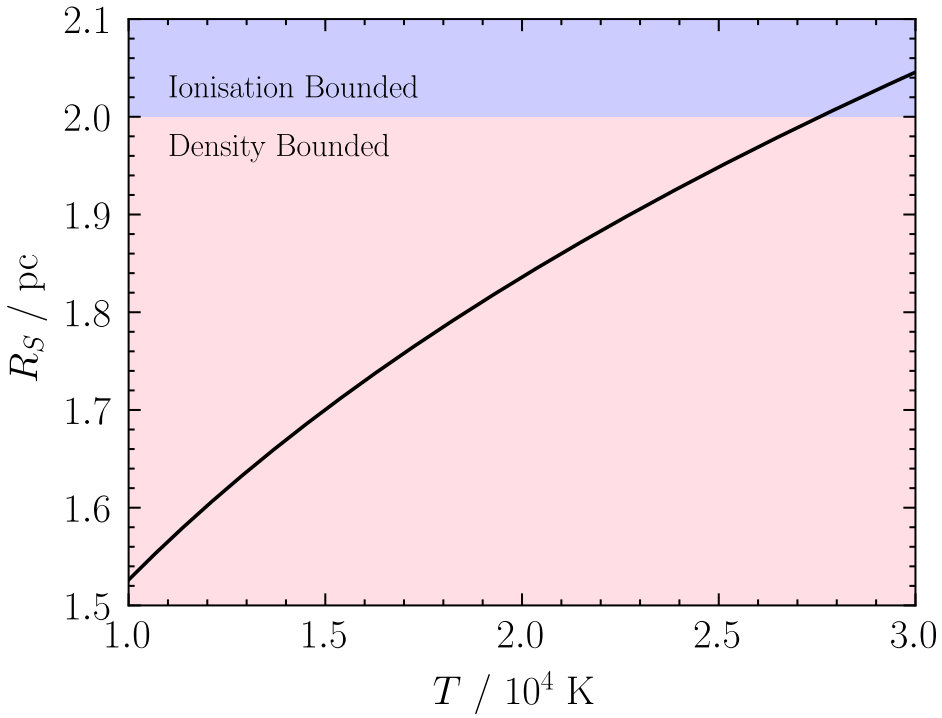

Figure 1: Strömgren radius, R_S, for a H II region parameterized by different electron temperature, T_e. The temperature where the region changes from density to ionisation bounded is $\approx 2.76 \times 10^4$ K.

occurs when $r = R_S$, which is $T_{e,\text{transition}} \approx 2.76 \times 10^4$ K. This was calculated using `scipy.optimize.root_scalar`. Figure 1 shows the Strömgren radius for 10000 K $\leq T_e \leq$ 30000 K.

Further assuming that in this H II region, 2/3 of all recombination emits results in the emission of a Lyα photon, we can estimate the Lyα luminosity of the cloud by,

$$L_{\text{Ly}\alpha} = \dot{N}_{\text{Ly}\alpha} E_{\text{Ly}\alpha} = \frac{2}{3} \dot{N}_{\text{ion}} E_{\text{Ly}\alpha} \approx 1.09 \times 10^{31} \text{ W} \tag{10}$$

where the energy of a Lyα is from equation 1. Suppose this cloud has a MS O-type star, the bolometric luminosity of this star is given by the Stephan-Boltzmann Law (assuming a perfect blackbody with $\epsilon = 1$),

$$L_{\text{bol}} = \sigma A T^4 = \sigma (4\pi (10 R_\odot)^2) T^4 \approx 8.83 \times 10^{31} \text{ W} \tag{11}$$

where we assume the radius of the O-type star is 10 times that of the sun. Comparing the ratio,

$$\frac{L_{\text{Ly}\alpha}}{L_{\text{bol}}} = \frac{1.09 \times 10^{31}}{8.83 \times 10^{31}} \approx 0.12 \tag{12}$$

The luminosity from Lyman α emission is about 12% of the bolometric luminosity of the central star.

Part 2: Measuring Distances in an Expanding Universe

Answer 2.1.a: Suppose the time evolution of an ionised bubble radius is only affected by the cosmic expansion. Therefore,

$$\frac{\mathrm{d}A}{dt} = 3H(z)A \tag{13}$$

where $A = R_{\mathrm{ion}}^3$. Solving for A,

$$\begin{aligned}
\int \frac{1}{A} dA &= 3 \int H(z)\ \mathrm{d}t \\
\ln(A) &= 3 \int \frac{\dot{a}(t)}{a(t)}\ \mathrm{d}t \\
\ln(A) &= 3 \ln[a(t)] + C \\
A &= e^C [a(t)]^3 \\
R_{\mathrm{ion}}(t) &= e^{C/3} a(t)
\end{aligned}$$

where C is an arbitrary constant. Suppose at the current time where $a = 1$, the radius is R_0, therefore,

$$R_{\mathrm{ion}}(t) = \frac{a(t)}{R_0} \tag{14}$$

This is expected, as cosmic expansion is the only factor affecting the radius of the bubble, so distances will scale with the scale factor. Now to estimate the e-folding time, let us assume the Hubble's parameter is a constant at $H(z = 7)$. Therefore the solution to the above equation becomes,

$$\begin{aligned}
\ln(A) &= 3H(z = 7)t + C \\
A &= C' \exp[3H(z = 7)t] \\
R_{\mathrm{ion}}(t) &= C'' \exp[H(z = 7)t]
\end{aligned} \tag{15}$$

The e-folding time is the time required for the quantity to increase by a factor of e. Therefore,

$$\begin{aligned}
\frac{eR_{\mathrm{ion}}(t)}{R_{\mathrm{ion}}(t)} &= \frac{C'' \exp[H(z = 7)t']}{C'' \exp[H(z = 7)t]} \\
e &= \exp[H(z = 7)(t' - t)] \\
t_e &= \frac{1}{H(z = 7)} \approx 1.14 \text{ Gyr}
\end{aligned} \tag{16}$$

Comparing this with the age of the universe at $z = 7$ following the Planck collaboration cosmology in `astropy.cosmology.Planck18.age(7)`, which is 0.76 Gyr. Therefore, the cosmic expansion is not a significant contributor to the evolution of the bubble as the e-folding time is much greater than the age of the universe at that time.

Answer 2.1.b: If the cosmic expansion and recombination is neglected, the ionised bubble radius evolves with the equation,

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{3 f_{\mathrm{esc,\ LyC}} \dot{N}_{\mathrm{ion}}}{4\pi \bar{n}_H} \Rightarrow R_{\mathrm{ion}}(t) = \left(\frac{3 f_{\mathrm{esc,\ LyC}} \dot{N}_{\mathrm{ion}}}{4\pi \bar{n}_H} t \right)^{1/3} \tag{17}$$

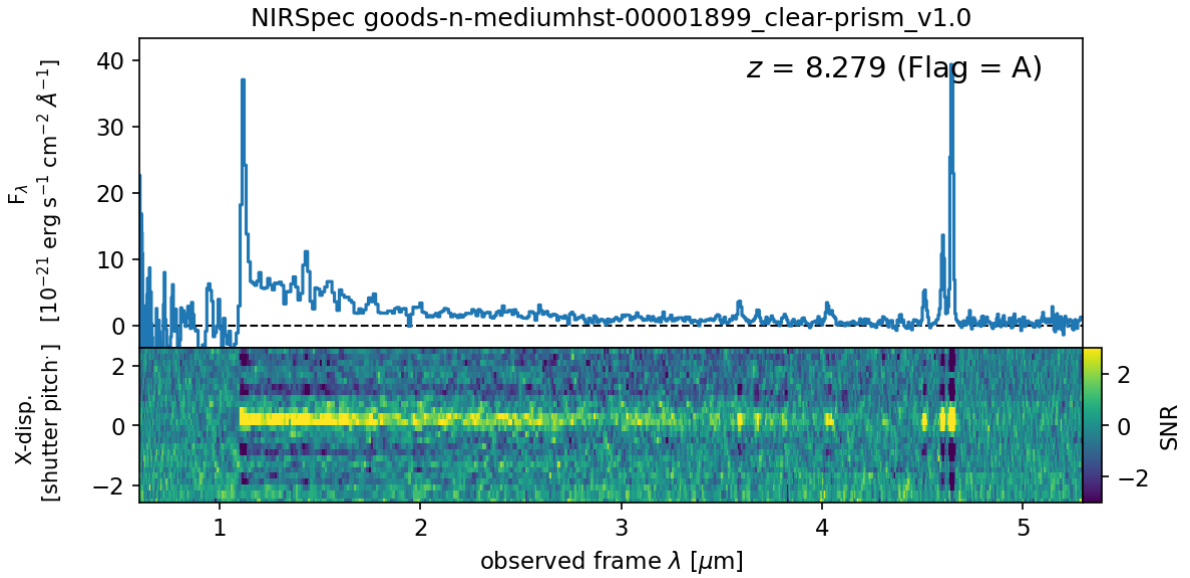

Figure 2: Source ID 1899's PRISM preview obtained directly from MAST. The top plot shows a 1D spectrum extracted from the bottom 2D spectrum showing the SNR per pixel.

following the assumptions of Mason and Gronke 2020 that $f_{\text{esc, LyC}}$ and $\dot{N}_{\text{ion}}$ are constant.

Answer 2.1.c: To figure out the radius at which the recombination events balance out the ionising output, we can assume that the rate of change of radius is 0 and ignore the cosmic expansion. Therefore,

$$R_{\text{ion, stable}} = \left(\frac{3 f_{\text{esc, LyC}} \dot{N}_{\text{ion}}}{4\pi \bar{n}_H^2 C_{\text{HII}} \alpha_B} \right)^{1/3} \tag{18}$$

using $f_{\text{esc, LyC}} = 0.1$, $\dot{N}_{\text{ion}} = 10^{54}$ s^{-1}, $C_{\text{HII}} = 3$, and $\alpha_B = (2/3) \times \alpha_{\text{rec}}(T_e = 20000 \text{ K})$. The mean hydrogen density is known at $z = 0$, and thus we can account for the density at a different redshift by,

$$\bar{n}_H = \frac{\bar{n}_H(z=0)}{a^3} = 1.88 \times 10^{-7} (1+z)^3 \text{ cm}^{-3} \tag{19}$$

Plugging in all the values, $R_{\text{ion, stable}}(z = 7) \approx 0.67$ Mpc and $R_{\text{ion, stable}}(z = 12) \approx 0.25$ Mpc.

Part 3: JADES Spectroscopic Galaxy Survey

Part 3.1: Extracting spectroscopic data from ID 1899's spectrum

Using the MAST system, we download the NIRSpec data from the JADES survey. Figure 2 shows the preview of the PRISM disperser for source ID 1899.

As noticed in Figure 2, the 2D spectrum shows a central high SNR trace, with negative SNR traces on either side. The reason for the low SNR traces is because of nodding in the Fixed Slit spectroscopy, where each slit is made out of out three shutters. The object is placed into each of these three shutters, and the exposures are centered about the central shutter. After this, the sky background is subtracted from the other shutter

areas. Therefore, we end up obtaining negative SNR traces on either sides due to the subtracted sky (equivalent to a negative source image) (Jakobsen, P. et al. 2022).

From the FITS file for the three different gratings (PRISM, G140M, and G395M), we plot the 1-D spectra of ID 1899 as shown in Figure **??**.

We can notice several emission lines in the spectra above. In NIRSpec, the longest observable wavelength is 5.30 μm in the PRISM grating. As stated in Table 2, the highest redshift observable for the Hα line is 7.1. From literature the redshift of ID 1899 is greater than 7 and therefore, the Hα line is out of observation range. Comparing with the Hβ line, the highest observable redshift is,

$$z_{\mathrm{H}\beta} = \frac{5.30 \times 10^{-6}}{4862.71 \times 10^{-10}} - 1 \approx 9.9 \tag{20}$$

which is less than the literature cited redshift value of ID 1899. Hence this line can be used to measure the spectroscopic redshift.

Other than being able to measure the redshift, the precision is also an important quantity to account for. Suppose the emission lines can be fit with a Gaussian profile, therefore, the FWHM of the fit and the standard deviation of the fit are related by,

$$\mathrm{FWHM} = 2\sqrt{2\ln 2}\sigma_\lambda \tag{21}$$

Comparing this with the expected resolution,

$$2\sqrt{2\ln 2}\sigma_\lambda = \frac{\lambda}{1000} \tag{22}$$

where λ is the central wavelength from the gaussian. To estimate the standard deviation in redshift, we propagate the error,

$$z = \frac{\lambda}{\lambda_{\mathrm{rest}}} - 1 \Rightarrow \sigma_{z+1} = \frac{\sigma_\lambda}{\lambda_{\mathrm{rest}}} \Rightarrow \sigma_\lambda = \lambda_{\mathrm{rest}}\sigma_{z+1} \tag{23}$$

Combining with the FWHM calculation,

$$\lambda_{\mathrm{rest}}\sigma_{z+1} = \frac{\lambda}{2000\sqrt{2\ln 2}} \Rightarrow \frac{\sigma_{z+1}}{z+1} \approx 4.25 \times 10^{-4} \tag{24}$$

In our redshift measurements, we expect to see a fractional precision on the order of 10^{-4}.

Now, given the high resolution of the G395M spectrum, we fit a gaussian profile to the emission lines (namely Hβ, [OIII]4960, and [OIII]5008) to obtain the spectroscopic redshift of the source. There are two sources which change the observed wavelength – cosmological redshift and local Doppler redshift. Therefore,

$$\lambda_{\mathrm{obs,cosmo}} = \lambda_{\mathrm{rest}}(z+1) \tag{25}$$

$$\lambda_{\mathrm{obs,full}} = \lambda_{\mathrm{obs,cosmo}}\left(\frac{v}{c}\right) = \lambda_{\mathrm{rest}}(z+1)\left(1+\frac{v}{c}\right) \tag{26}$$

under the assumption that local velocities are much smaller than the speed of light. Furthermore, the expected errors in observed wavelengths can be propagated from the error in velocity.

$$\sigma_\lambda = \frac{\lambda}{c}\sigma_v \tag{27}$$

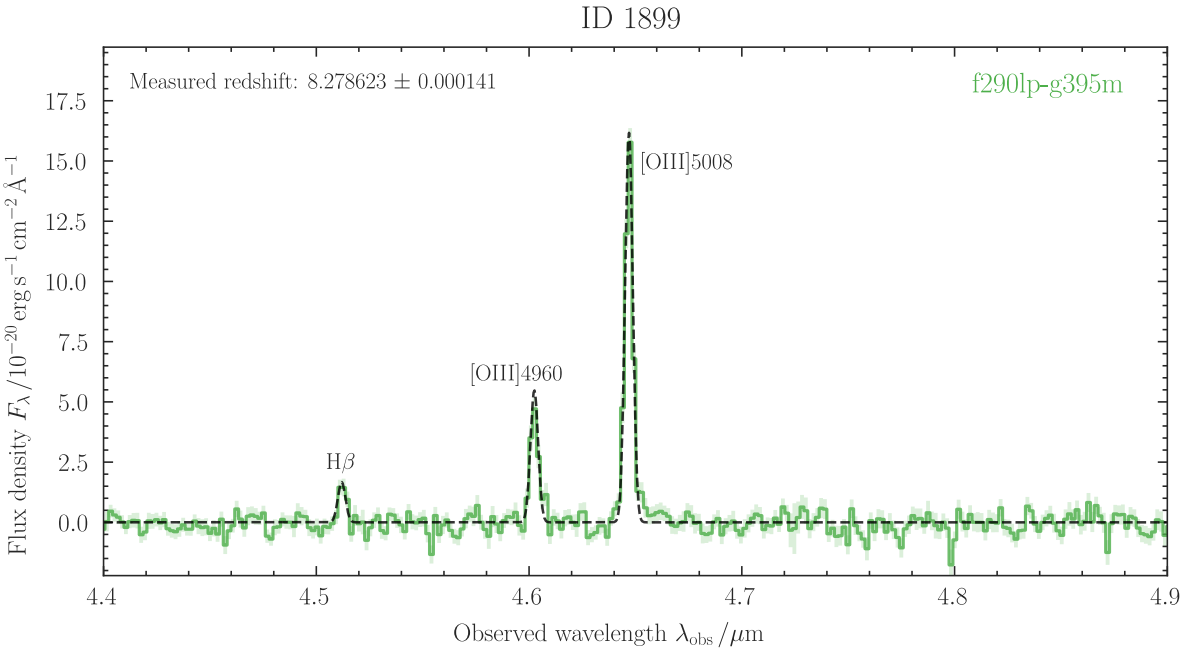

Figure 3: Least squares fitting of the emission lines observed in the G395M grating for source ID 1899. The best-fit spectroscopic redshift is 8.278623 ± 0.000141, and the fit is good with $\chi^2_R \approx 1.1$.

However, the width of the line not just depends on the intrinsic broadening by the star, but also by the broadening due to the instrument. As calculated in equation 22, the instrumental broadening will be,

$$\sigma_{\text{instrument}} = \frac{\lambda}{R(2\sqrt{2\ln 2})} \tag{28}$$

Convolving both the Gaussians, we expect to see the dispersion add quadratically. Therefore, the final model dispersion will be,

$$\sigma_{\text{full}} = \sqrt{\sigma^2_\lambda + \sigma^2_{\text{instrument}}} \tag{29}$$

Therefore using the `lmfit` package, we fit the emission line profiles to our data. While all the parameters are free, the ratio between the amplitudes of [OIII]5008 and [OIII]4960 is 2.98. Therefore amplitude of [OIII]5008 is not a free parameter. Figure 3 shows the spectrum along with the best-fit gaussian profiles. The measured redshift is 8.278623 ± 0.000141, and the model fit is really well, as $\chi^2_R \approx 1.1$.

Using the expected error as in equation 24, the uncertainty in redshift should be about 0.0039, whereas the one obtained from the fitting is 0.00014, which is one order of magnitude lower than original estimate. The error may be lower because the best fit is sampling multiple points, which will reduce the overall calculated standard deviation.

Given we know the systematic redshift, we use it in our fitting model again but with the G140M grating for the Lyman break. The R was kept the same, i.e., $R = 1500$, and the fitting was performed between 1.12 μm to 1.14 μm. Following the same procedure as before, Figure 4 shows the best fit Lyman α emission line. The Lyman alpha flux and the velocity offset calculated are $F_{\text{Ly}\alpha} \approx (710.9 \pm 71.5) \times 10^{-20}$ erg s^{-1} cm^{-2} and

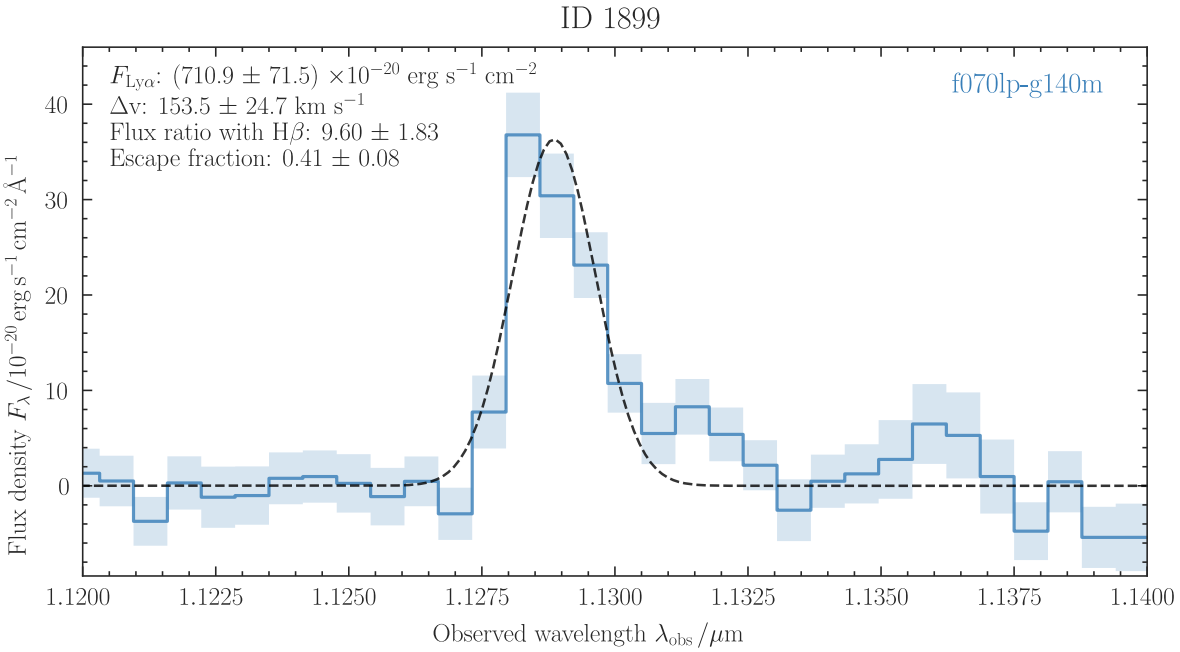

Figure 4: Least squares fitting of the Lyman break observed in the G140M grating. The best fit and calculated parameters are shown above. The fit is good, with $\chi^2_R \approx 1.4$.

$\Delta v \approx 153.5 \pm 24.7$ km s^{-1}.

We then calculate the observed ratio of Lyman α and Hβ luminosities,

$$\frac{L_{\mathrm{Ly}\alpha}}{L_{\mathrm{H}\beta}} = \frac{F_{\mathrm{Ly}\alpha}}{F_{\mathrm{H}\beta}} \approx 9.60 \pm 1.83 \tag{30}$$

where the error was calculated through error propagation, and the Hβ flux was obtained during the spectroscopic redshift calculation. The luminosity distance is not calculated as it cancels out during the ratio calculation.

Given that we know Hβ mostly escapes the ionising medium, we can estimate the Lyα escape fraction by,

$$f_{\mathrm{esc,\ Ly}\alpha} = \frac{(L_{\mathrm{Ly}\alpha}/L_{\mathrm{H}\beta})_{\mathrm{obs}}}{(L_{\mathrm{Ly}\alpha}/L_{\mathrm{H}\beta})_{\mathrm{theoretical}}} \approx 0.41 \pm 0.08 \tag{31}$$

Part 3.2: Extracting spectroscopic data from ID 10013682's spectrum

Following the same procedure as in the previous part, data from ID 10013682 was analysed as well. Figure 5 shows the 1-D spectrum across all the three gratings. Unlike ID 1899, ID 10013682's G395M grating has a low SNR (as also evident by our inability to visually distinguish the Hβ and [OIII]4960 emission lines from the noise). Therefore, PRISM was used to estimate the spectroscopic redshift. Unfortunately, the resolution is an order of magnitude lower than G395M, so a higher uncertainty was expected. For the fit we assumed $R = 350$. Figure 6 and 3 shows the PRISM and G140M fits, and Table 7 shows all the calculated values. Both fits are worse than the previous part, but statistically still significant as $\chi^2_R \approx 0.6$ for PRISM fit and $\chi^2_R \approx 2.3$ for G140M fit.

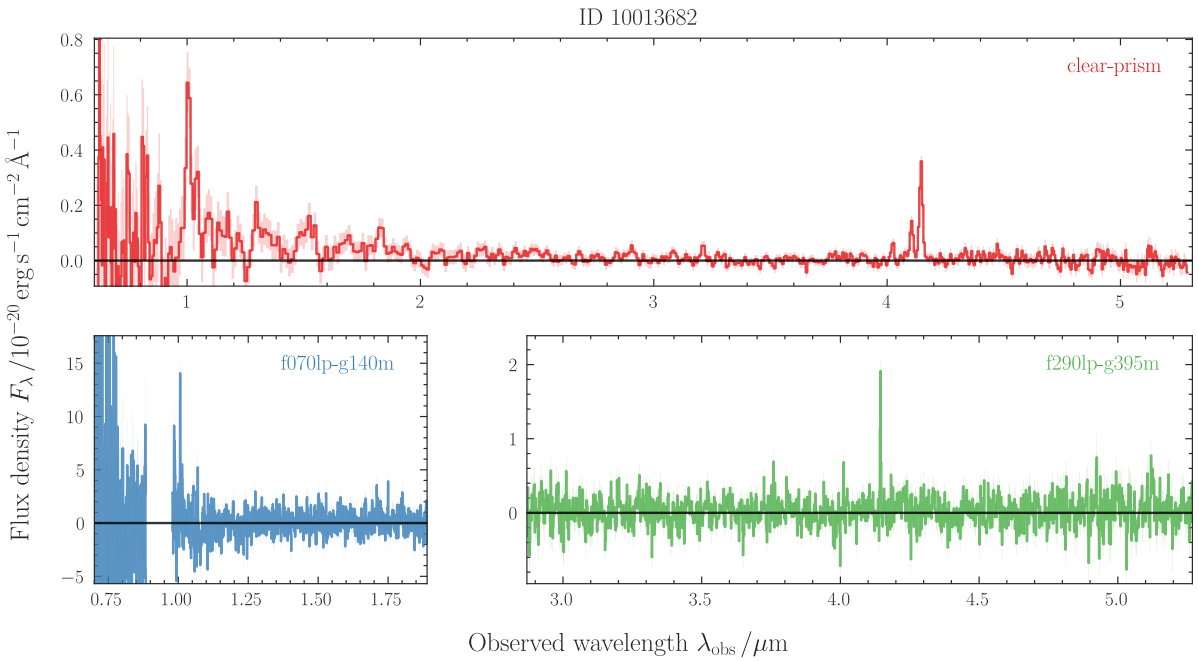

Figure 5: 1-D spectra from Source ID 10013682 through different JWST NIRSpec gratings. The SNR in the G395M grating is low (can be visually confirmed as Hβ and [OIII]4960 lines cannot be discerned away from the noise). The spectroscopic redshift was thus calculated using PRISM.

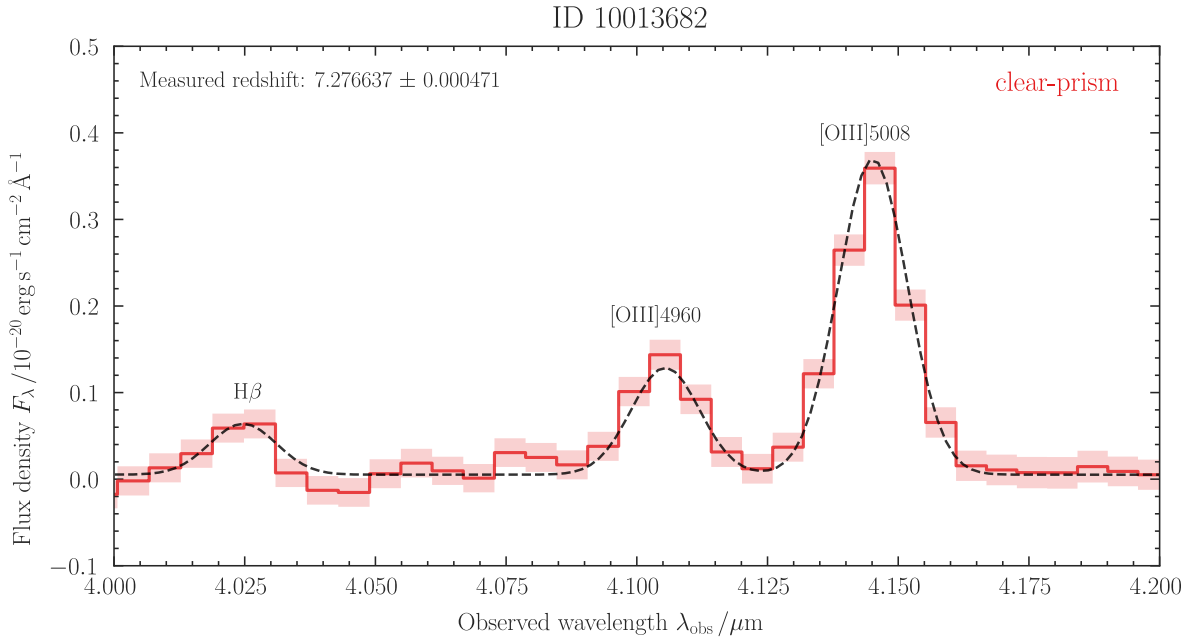

Figure 6: Least squares fitting of the emission lines observed in the PRISM grating for source ID 10013682. The best-fit spectroscopic redshift is 8.278623 ± 0.000141, and the fit is good with $\chi^2_R \approx 0.6$.

We note that the redshift is in full alignment with comparison with the literature. While the $\Delta v_{\mathrm{Ly}\alpha}$ is different, it still falls under the error bounds. The escape fraction, however, is greater than 1, which is unphysical. However this problem has already been addressed by Saxena, Aayush et al. 2023. They state a higher than 1 fraction as well if measured using PRISM. However, if you measure using the grating, we observe a 0.68 ± 0.20 which

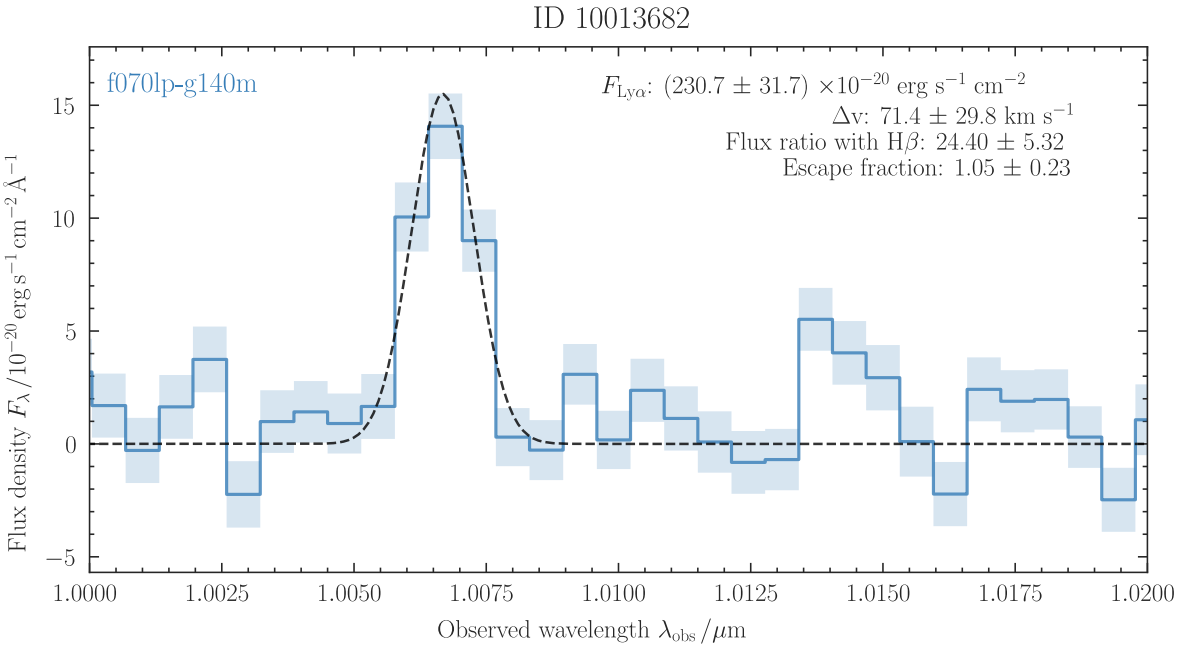

Figure 7: Least squares fitting of the Lyman break observed in the G140M grating. The best fit and calculated parameters are shown above. The fit is good, with $\chi^2_R \approx 2.3$.

Parameters	This work	Witstok, Joris et al. 2024
RA / °	–	+53.16746
Dec / °	–	-27.77201
z_{spec}	7.276637 ± 0.000471	7.276
M_{UV}	–	-16.86 ± 0.28
$\Delta v_{\mathrm{Ly}\alpha}$ / km s^{-1}	71.4 ± 29.8	178 ± 100
$f_{\mathrm{esc,\ Ly}\alpha}$	1.05 ± 0.23	0.67 ± 0.18

Table 3: Calculated and literature values for Source ID 10013682.

is close to the value calculated by Witstok, Joris et al. 2024. Therefore our result is not surprising.

Part 3.3: Calculate Ionised Bubble Size

In this section, we used the `lymana_absorption` package. It was cloned from its git repository, and a reproducing code was ran to ensure it works properly.

Using the in-built function to calculate the opacity of the IGM, `tau_IGM`, we estimate the transmission of the Lyα photons at different bubble radii. The transmission against different radii is shown in Figure 8. Furthermore, as we know the escape fraction, we can estimate at which value of ionic radii does the transmission equated the escape fraction through interpolation performed by `numpy.interp`. The calculated ionic radii is $2.639^{+4.322}_{-1.274}$ pMpc. This value with error bounds is also shown in Figure 8.

The errors stated here were calculated using the escape fraction error stated by Witstok, Joris et al. 2024. We also estimated the error by solely considering the uncertainty on the velocity offset of ±100 km s^{-1}. We take limits of the velocity offset and average the two values of R_{ion} obtained to estimate the error. This method gives us 2.639±0.109pMpc.

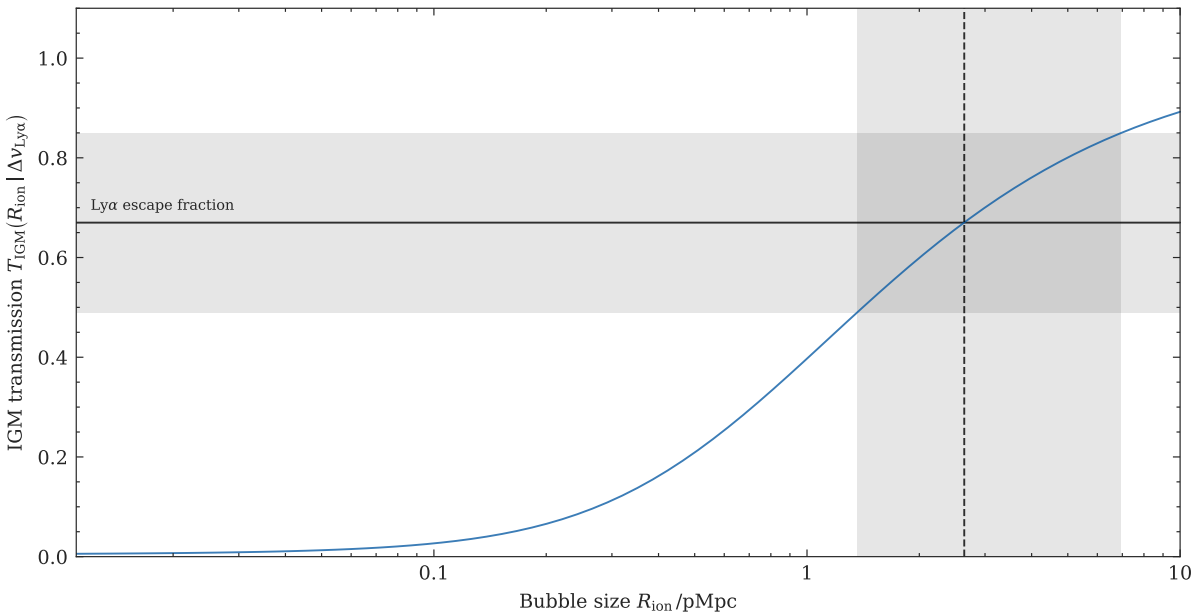

Figure 8: IGM tranmission against Bubble size for source ID 10013982. We find the lower limit on the Bubble size by equating the known escape fraction with the transmission. The error bounds on $R_{\rm ion}$ are due to error in escape fraction alone. As described, if we simply consider the limits of the velocity offset and average them out, we obtain 2.639 ± 0.109 pMpc

It must be noted that the calculated value is just a lower limit, because we assume there is no other material in the line-of-sight through the IGM. Presence of dust inside the galaxy, or backscattering the photons away from the line-of-sight within the galaxy, can also reduce the amount of Lyα photons reaching us.

Finally, we calculate how long such a region would take to form, assuming values from the literature and 100% escape fraction. We use equation 17 to estimate the time taken to create such a medium. Rewriting,

$$T = \frac{4\pi R_{\rm ion}^3 \bar{n}_H}{3\dot{N}_{\rm ion}} \tag{32}$$

$\bar{n}_H$ can be calculated using equation 19 as we already know the spectroscopic redshift. To obtain the ionising rate, we make use of the relation,

$$\dot{N}_{\rm ion} = \xi_{\rm ion} L_{\nu,\rm UV} \tag{33}$$

$\xi_{\rm ion}$ is known from literature and Luminoisty density can be calculated from the absolute magnitude quoted in the literature through converting between flux density. Therefore,

$$F_\nu = 10^{-\frac{M_{\rm UV} + 48.6}{2.5}} \tag{34}$$

$$L_\nu = 4\pi (10\ \text{pc})^2 F_\nu \tag{35}$$

where distance is 10 pc since we have the absolute magnitude. Plugging in $\xi_{\rm ion} = 10^{25.66}$ and $M_{\rm UV} = -17$, we obtain it would take a time period of 61 Gyr to create this bubble! This is much greater than the current age of the universe, let alone the age of the universe at that redshift (0.7 Gyr). Therefore, it is impossible that this galaxy alone could create an ionisation bubble of this size at that point in our universe.

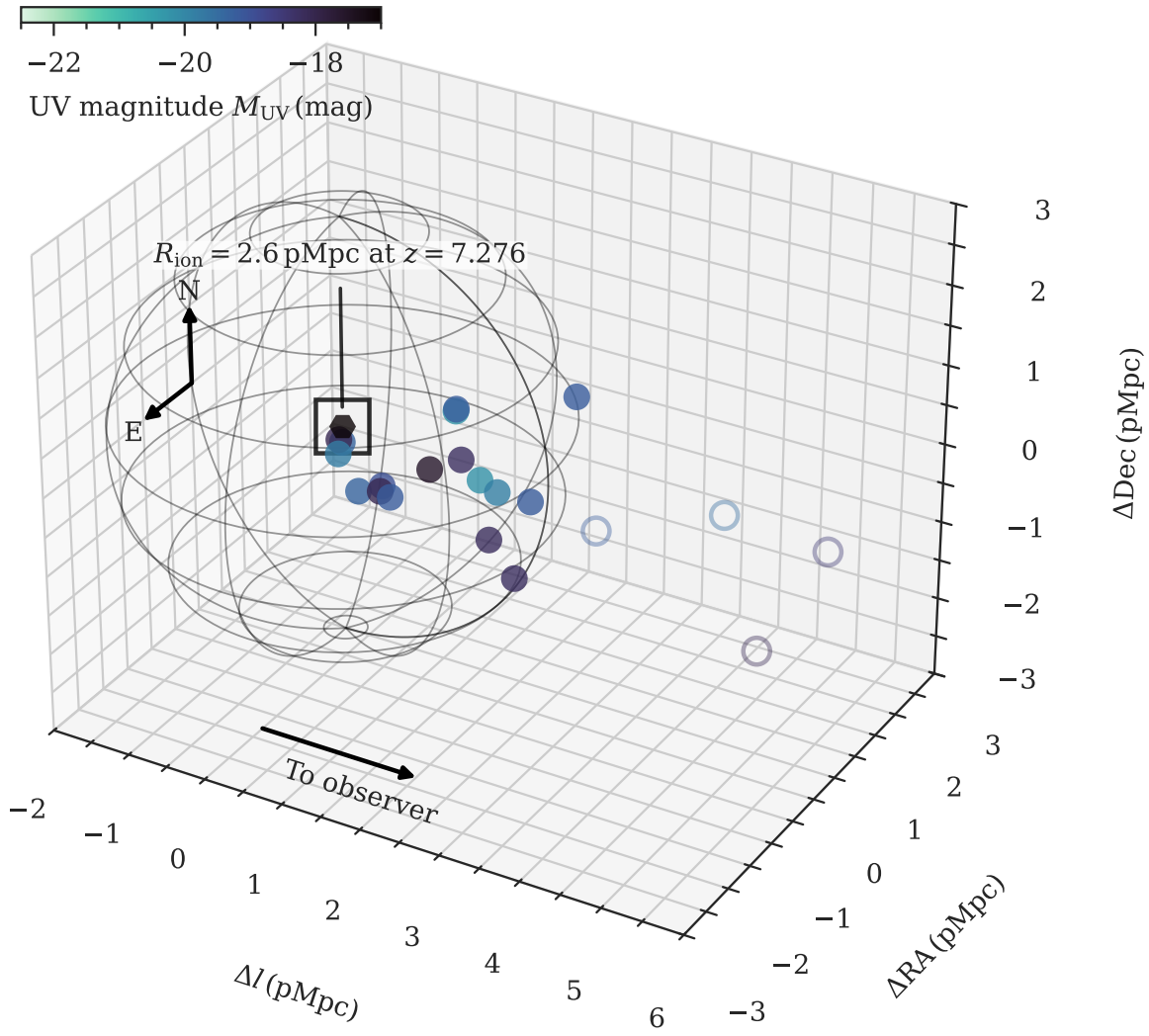

Figure 9: Neighbourhood of source ID 10013682 plotted using data from the FRESCO survey.

Part 3.4: Visualise the environment of the ionised bubble

To observe the region around our source ID 10013682 galaxy, we use the FRESCO survey galaxies to find neighbours within the ionising bubble. We obtain the LoS separation using the comoving distance between the source and the neighbouring galaxy, and the sky separation using their angular position on the sky. We set a maximum cut of about 6 pMpc to remove any galaxy beyond this region. Following which, we obtain the 3D distance between them using the pythagoras theorem as,

$$D_{\rm 3D} = \sqrt{D_{\rm Los}^2 + D_{\rm sky}^2} \tag{36}$$

If a galaxy's 3D separation is less than or equal to the radius of the ionic bubble, it is considered in the calculation of the total ionisation rate, which was calculated as,

$$\dot{N} = \sum_{\text{galaxies in the ionising bubble}} \left(1.67 \times 10^{54} 10^{\dfrac{M_{UV}+20}{-2.5}} \left(\frac{912}{1500} \right)^{\beta_{\rm UV}+2} \right) \tag{37}$$

where $\beta_{\rm UV}$ and M_{UV} are obtained from the FRESCO catalog data. The final calculated value is 13.60×10^{54} s^{-1}.

Plugging this value into equation 32, we obtain 0.56 Gyr, which is acceptable within the age of the universe. Therefore, it appears that including the ionising potential of the neighbouring galaxies accounts for the size of the ionic radius.

References

Jakobsen, P. et al. (2022). "The Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope - I. Overview of the instrument and its capabilities". In: *A&A* 661, A80. DOI: 10.1051/0004-6361/202142663. URL: https://doi.org/10.1051/0004-6361/202142663.

Mason, Charlotte A. and Max Gronke (Nov. 2020). "Measuring the properties of reionized bubbles with resolved Lyα spectra". In: 499.1, pp. 1395–1405. DOI: 10.1093/mnras/staa2910. arXiv: 2004.13065 [astro-ph.GA].

Saxena, Aayush et al. (2023). "JADES: Discovery of extremely high equivalent width Lyman-ssion from a faint galaxy within an ionized bubble at z = 7.3". In: *A&A* 678, A68. DOI: 10.1051/0004-6361/202346245. URL: https://doi.org/10.1051/0004-6361/202346245.

Witstok, Joris et al. (2024). "Inside the bubble: exploring the environments of reionisation-era Lyman-tting galaxies with JADES and FRESCO". In: *A&A* 682, A40. DOI: 10.1051/0004-6361/202347176. URL: https://doi.org/10.1051/0004-6361/202347176.

Report 7: Physics of the Intergalactic Medium

Kushagra Shrivastava

Overview

The distribution of Baryonic and Dark Matter plays a significant role in seeding galaxy formation. In this report, we analyse the distribution and observation of gas and dark matter in the Inter Galactic Medium (IGM) through Lyman-alpha (Lyα) forests. Section 1 analyses the hydrodynamical cosmological RAMSES simulation, Section 2 calculates the Lyα forest spectrum obtained through the simulations and compares it with observational GRB data, Section 3 accounts for the physical factors, mainly photoionization, that affects the density of neutral hydrogen in the IGM and thus the Lyα forest flux, and lastly, Section 4 focuses on toy-models to account for sources of the said photoionization.

Part 1: Analysing the RAMSES Simulation

In this section, we analyse the gas and dark matter density obtained from the RAMSES simulation. Throughout this report, ρ_b and ρ_m represents the gas and dark matter density respectively. Furthermore, $\bar{\rho}_b$ and $\bar{\rho}_m$ represent the averages. While the absolute values do not differ significantly, the density fluctuations do. We denote these fluctuations as,

$$\delta_{b,m} = \frac{\rho_{b,m} - \bar{\rho}_{b,m}}{\bar{\rho}_{b,m}} \tag{1}$$

The simulation box is 40 cMpc h^{-1} on each side. We arbitrarily use a cartesian coordinate system and assume each edge lies along a coordinate axis. To visualize the LSS present in the simulation, we plot a 2D slice of $\rho/\bar{\rho}$ at $Z = 20$ cMpc h^{-1} as shown in Figure 1.

The ratio is color-mapped in a log-space to account for the severe overdensity in the dense regions. In the case of dark matter distribution, there were 37 points with a density of 0 g cm^{-3}. To prevent issues in the log-space, we reassign the density ratio as 10^{-20}. This number was chosen such that it is close, but multiple magnitudes lower than the second lowest density ratio (4.5×10^{-9}).

We also plot a 1D skewer, which is a straight line passing through $(X, Y) = (20, 20)$ cMpc h^{-1} parallel to the Z-axis, as shown in Figure 2.

Across Figures 1 and 2 we notice the Dark Matter and Gas follow each other structures, at least on large scales. The large yellow regions in the 2D plots are generally the same across both, and the 1D skewer also show the same. Furthermore, we plot the gas to dark matter density ratios, and observe a positive linear relation with a Pearson Correlation Coefficient of 0.874.

$X-Y$ plane RAMSES simulation slices at $Z = 20$ cMpc h^{-1}

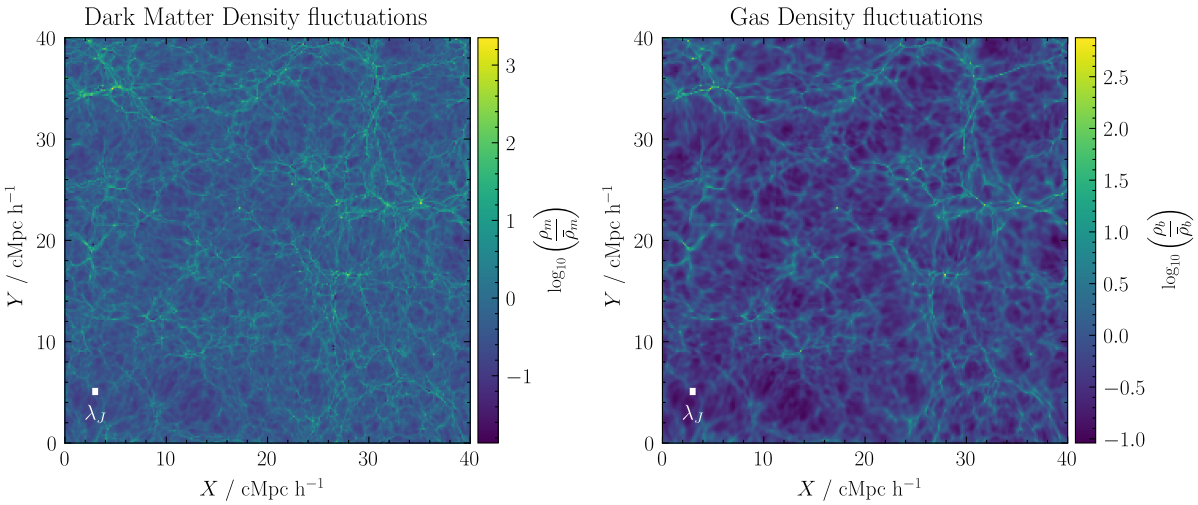

Figure 1: 2D slice of the RAMSES simulation. We notice the LSS in the Dark Matter and Gas slices to follow largely the same structure. Furthermore, the white cube represents the cosmological Jeans length in the horizontal direction. On that scale, we notice that the Gas Density fluctuations appears more smooth than the Dark Matter Density fluctuations.

Other than the structure, we can visually observe the smoothing of in the Gas Density ratio slice. At small scales, the gas motion is affected by hydrodynamic gas pressure, with an equilibrium between the gravitational force and the pressure force. The critical length for this equilibrium is called the Cosmological Jeans Length, $\lambda_{\rm J}$. At the redshift of the RAMSES simulation ($\approx$4), $\lambda_{\rm J} \approx 110$ pkpc. In terms of simulation units,

$$\lambda_{\rm J} = 110 \text{ pkpc} = \frac{110}{1000}(1+z)h \approx 0.372 \text{ cMpc h}^{-1} \approx 2.38 \text{ px.} \tag{2}$$

This length is superimposed on the 2D slices in Figure 1. It is small, but visually, on that scale we observe that the density ratio field is blurrier (and thus smoother) in the Gas Density slice.

Part 2: Simulating the Lyα forest spectrum

The structure we notice in the previous part can be observed in spectra of bright background objects. This feature is known as the Lyman-alpha (Lyα) forest. The Lyα photons gets absorbed by the neutral hydrogen atoms. In the expanding universe, we can write the optical depth of this Lyα absorption line as,

$$\tau_\alpha(\nu) = \int_0^{z_s} \sigma[\nu(1+z')] n_{\rm HI}(z') \left|\frac{{\rm d}l'}{dz'}\right| {\rm d}z' \tag{3}$$

The redshift derivative of proper distance is given by,

$$\left|\frac{{\rm d}l}{{\rm d}z}\right| = \frac{c}{H(z)(1+z)} \tag{4}$$

The absorption cross section can be approximated as a constant multiplied by Dirac-delta line functions.

$$\sigma[\nu(1+z')] \approx \sigma_\alpha \delta_D[\nu(1+z') - \nu_\alpha] = \sigma_\alpha \frac{\delta_D[z' - (\nu_\alpha/\nu) + 1]}{|\nu|} \tag{5}$$

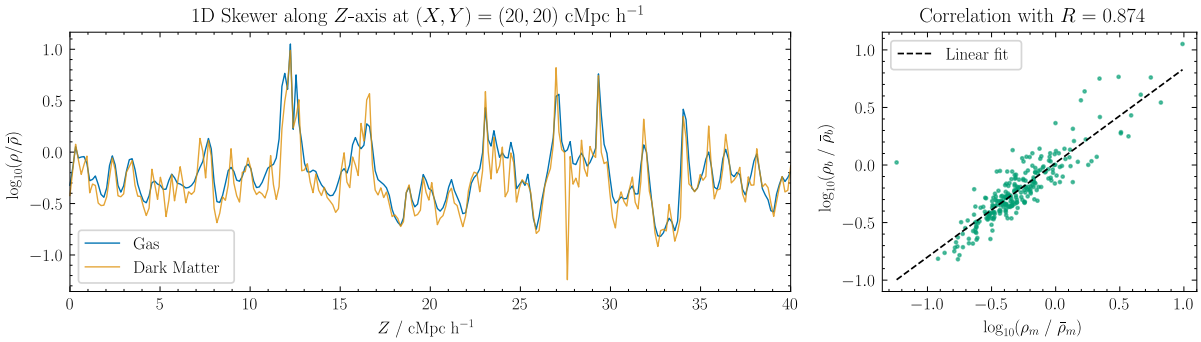

Figure 2: The left plot is the 1D skewer from the RAMSES simulation. Visually, the Gas and Dark Matter ratios follow each other closely. The right plot is a density ratio plot which further shows there is a linear correlation between the Dark Matter and Gas ratios.

where we used the summation identity of Dirac Delta functions. By definition of redshift,

$$\nu = \frac{\nu_\alpha}{1+z} \tag{6}$$

Therefore, we can rewrite Equation 3 as,

$$\tau_\alpha(\nu) = \int_0^{z_s} \sigma_\alpha \frac{\delta_D[z'-z]}{|\nu|} n_{\rm HI}(z') \frac{c}{H(z')(1+z')} \mathrm{d}z' = \frac{\sigma_\alpha c n_{\rm HI}(z)}{vH(z)(1+z)} = \frac{\sigma_\alpha c n_{\rm HI}(z)}{\nu_\alpha H(z)} \tag{7}$$

where we used the Dirac-Delta identity,

$$\int_a^b g(x')\delta_D[x'-x]\mathrm{d}x' = g(x), \text{ where } x' \in [a,b] \tag{8}$$

Under the Gunn-Peterson approximation, and the redshift the RAMSES simulation is based at, we can rewrite the optical depth approximation as,

$$\tau_\alpha^{\rm GP}(\lambda) \approx \frac{\sigma_\alpha \lambda_\alpha n_{\rm HI}(z)}{H_0 \Omega_m^{1/2} (1+z)^{3/2}} \tag{9}$$

The neutral hydrogen density can be rewritten compared to the total hydrogen density ratio.

$$n_{\rm HI} = x_{\rm HI} n_{\rm H} = \frac{x_{\rm HI}(1-Y)\rho_b}{m_H} \tag{10}$$

where $Y = 0.24$ is the helium mass fraction. With this dependence, we can directly convert the density from RAMSES simulation to the Lyman alpha optical depth. Therefore the transmitted flux can be written as,

$$F(\lambda) = \exp\left(-\frac{\sigma_\alpha \lambda_\alpha (1-Y)}{H_0 \Omega_m^{1/2} m_H} \frac{x_{\rm HI}\rho_b}{(1+z)^{3/2}}\right) \tag{11}$$

Along with the Flux, we arbitrarily choose the Z axis as the direction of observation, and thus the far end of the simulation is $z_{\rm sim} \approx 4$. We discretize the redshift at different slices along the Z axis using,

$$z_i = z_{\rm sim} - \frac{H_0 \Omega_m^{1/2} (1+z_{\rm sim})^{3/2} x_i}{c} \tag{12}$$

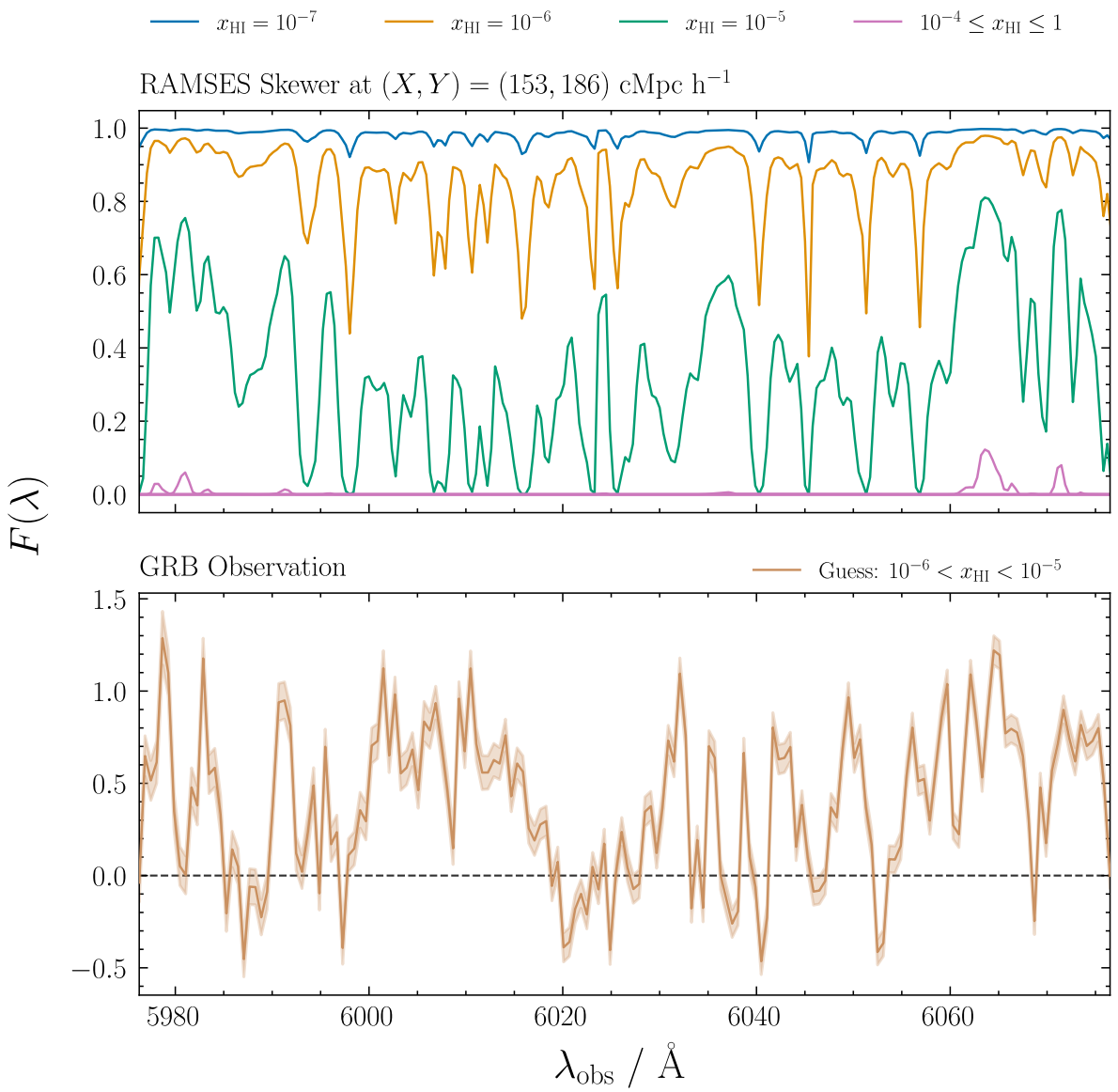

Figure 3: (Top) Transmitted flux through the RAMSES simulation at different neutral hydrogen density fractions. (Bottom) Observed Lyα spectrum from a GRB. Comparing both plots, we can expect the IGM neutral hydrogen fraction to be between 10^{-6} and 10^{-7}.

where $H_0 \approx 2.20 \times 10^{-18}$ s^{-1}, $c \approx 9.72 \times 10^{-15}$ Mpc s^{-1} and x_i is in pMpc. Units were chosen to maintain dimensional consistency. We converted the redshift to the redshifted Lyman alpha wavelengths,

$$\lambda_i = \lambda_\alpha (1 + z_i) \tag{13}$$

Figure 3 shows a random (X, Y) skewer from the simulation, with the flux and wavelengths calculated from equations 11 and 13 respectively. We alter the neutral hydrogen ratio, and we notice that as the amount of neutral hydrogen ratio increases, the absorption increases. After 10^{-4}, the flux becomes effectively 0. We compare the simulated spectrum, with observed GRB data also shown in Figure 3. Visually, the shape of the spectra does resemble the simulation, and we can notice that the neutral hydrogen fraction observed in the IGM medium is likely to be between 10^{-6} and 10^{-5}.

To perform another visual comparison, we plot density histograms for different neutral

hydrogen ratios as shown in Figure 4. This was done to randomly choose 200 skewers and plot their transmitted flux distributions, to account for statistical outliers in our choice of skewer in Figure 3. Again, as observed, the GRB distribution shows a neutral hydrogen ratio between 10^{-6} and 10^{-5}.

Part 3: Role of IGM Photoionisation in the Lyα forest spectrum

In the previous part, we arbitrarily changed the neutral hydrogen ratio. However, the evolution of the neutral hydrogen density is dependant on various factors. Suppose we only account for the cosmic expansion, then the redshift evolution of the neutral hydrogen density can be written as,

$$n_{\rm HI}(z) = n_{\rm HI}(0) \times (1+z)^3 \tag{14}$$

Using this, we rewrite equation 9 as

$$\tau_\alpha^{\rm GP} \approx 2.2 \times 10^5 x_{\rm HI} \left(\frac{1+z}{5}\right)^{3/2} \left(\frac{\Omega_b h^2}{0.022}\right) \left(\frac{\Omega_m h^2}{0.147}\right)^{-1/2} \tag{15}$$

However accounting for all the physical parameters, the neutral hydrogen density changes as,

$$\frac{{\rm d}n_{\rm HI}}{{\rm d}t} = \alpha(T) n_e n_{\rm HII} - \Gamma_{\rm HI} n_{HI} - C n_e n_{\rm HI} \tag{16}$$

where the first is increase in neutral hydrogen density due to recombination, the second term is reduction in neutral hydrogen density due to photoionization, and the third term is reduction in neutral hydrogen density due to collisional ionization between electrons and neutral hydrogen.

Suppose the photoionization and collisional ionization are comparable in strength. Therefore,

$$\Gamma_{\rm HI} n_{\rm HI} = C n_e n_{\rm HI} \Rightarrow \Gamma_{\rm HI} = C n_e \tag{17}$$

Typically, the density of the neutral hydrogen should be of the same order as the density of electrons. Furthermore, the collisional ionization coefficient should depend on the temperature of the IGM. Typically, $\Gamma_{\rm HI} \sim 10^{-12}$ s^{-1} and $n_e \sim 10^{-5}$ cm^{-3}. Therefore, $C \sim 10^{-7}$ cm^3 s^{-1}. However around 10^4 K, the value of C should be much lower. In order for collisional ionization to be more significant, the temperature of the IGM should be higher by a few magnitudes.

Therefore, we assume that the collisional ionization does not matter at such temperatures. Suppose we assume that recombination and the photoionization are in an equilibirum. Therefore, we can rewrite equation 9 as,

$$\tau_\alpha^{\rm GP} \approx 2.2(1+\delta_b) \left(\frac{\Gamma_{\rm HI}}{10^{-12}\ {\rm s}^{-1}}\right)^{-1} \left(\frac{T}{10^4\ {\rm K}}\right)^{-0.72} \left(\frac{1+z}{5}\right)^{9/2} \left(\frac{\Omega_b h^2}{0.22}\right) \left(\frac{\Omega_m h^2}{0.147}\right)^{-1/2} \tag{18}$$

where h is the reduced Hubble's constant. The two models are plotted in Figure 5 with $\Gamma_{\rm HI} \approx 2.5 \times 10^{-12}$ s^{-1} and $x_{\rm HI} \approx 10^{-6}$. The optical depth increases much faster with redshift if we consider photoionization.

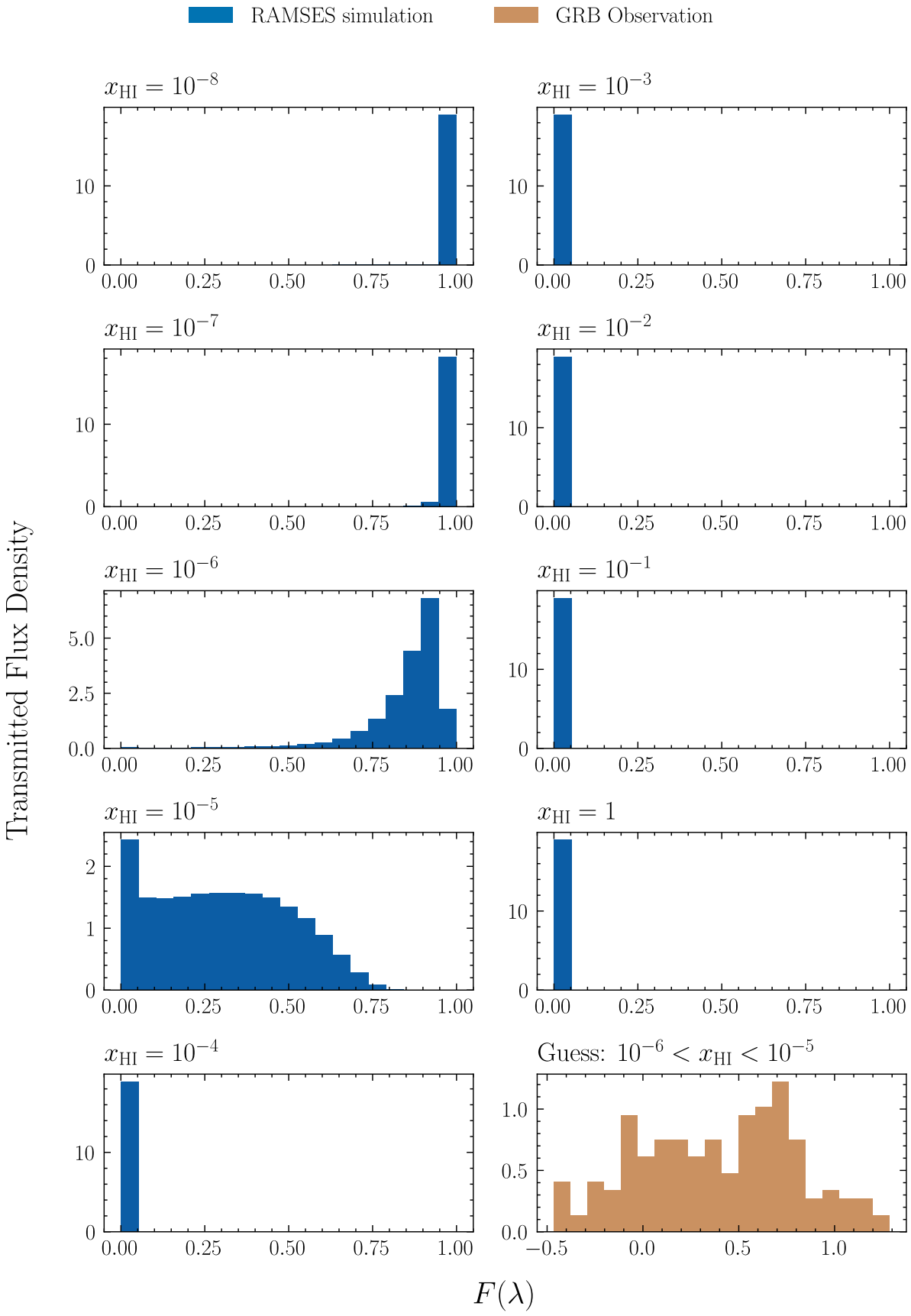

Figure 4: Transmitted Flux Density histograms for 200 random 1D skewers from the RAMSES simulation. Again, we notice the GRB histogram suggests a neutral hydrogen ratio between 10^{-6} and 10^{-5}.

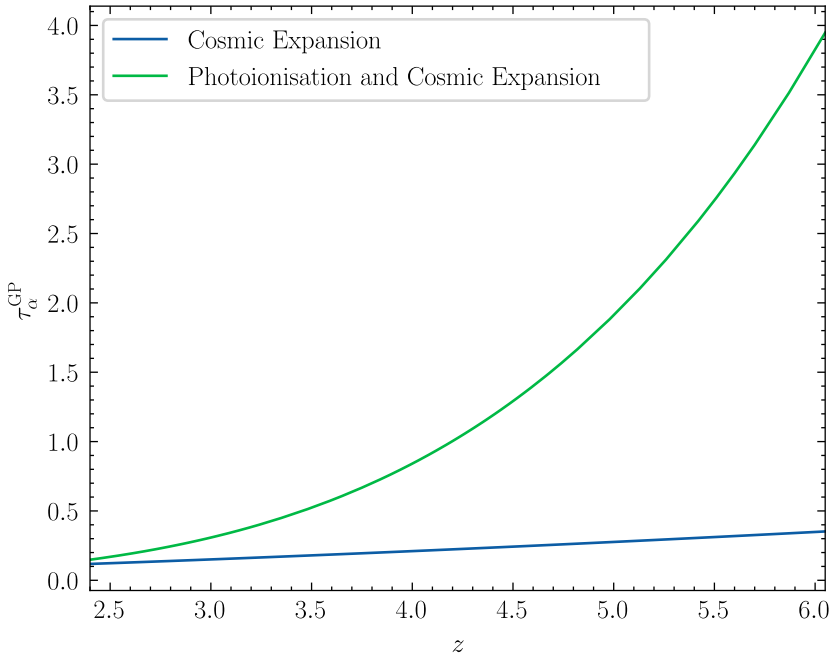

Figure 5: Comparison of different optical depth models. The Photoionization model increases the optical depth much faster with redshift.

Furthermore, we fit Songaila (2004) mean Lyα forest transmission ($\langle e^{-\tau_\alpha^{GP}} \rangle$) data. We perform χ^2 fitting to best fit the models. Using equation 15, the best fit parameter is x_{HI} and for equation 3, the best fit parameter is Γ_{HI}. The χ^2 is defined as,

$$\chi^2 = \left(\frac{F_{obs} - F_{model}}{\sigma_{model}} \right)^2 \tag{19}$$

and the minimization takes place using `scipy.optimize.minimize`. To improve stability, the variable parameter in the function is not the best fit parameters themselves, but the powers of the best-fit parameters. As the errors are asymmetric, we consider the σ_+ if $F_{model} > F_{obs}$, and σ_- otherwise. The best fit values are $\Gamma_{HI} \approx 2.57 \times 10^{-12}$ s^{-1} and $x_{HI} \approx 6.07 \times 10^{-6}$. Figure 6 shows the data and the best-fit models. Clearly, cosmic expansion alone does not explain the observed mean Lyα, and the photoionization model accounts for it very well.

Using the same methodology, we best fit the mean Lyα transmission from our RAMSES simulation to the GRB data shown in Figure 3. We keep the redshift as a constant at 4. To calculate the mean transmission of the simulation, we randomly choose 5000 1D skewers parallel to the Z-axis, and calculate the mean transmission for each skewer. Then we calculate the mean, of all the skewer-meaned transmissions. We chose 5000 random skewers out of the available 256^3 options, because we wanted to under sample to ignore the extremely high density skewers which might skew the average. Similar to the previous part, our best fit parameter was Γ_{HI}. As we were best-fitting one value, the error from the observation was propagated assuming each datapoint was independent. The mean-observed error was thus,

$$\langle \sigma_{obs} \rangle = \frac{\sqrt{\Sigma_i \sigma_i^2}}{N} \tag{20}$$

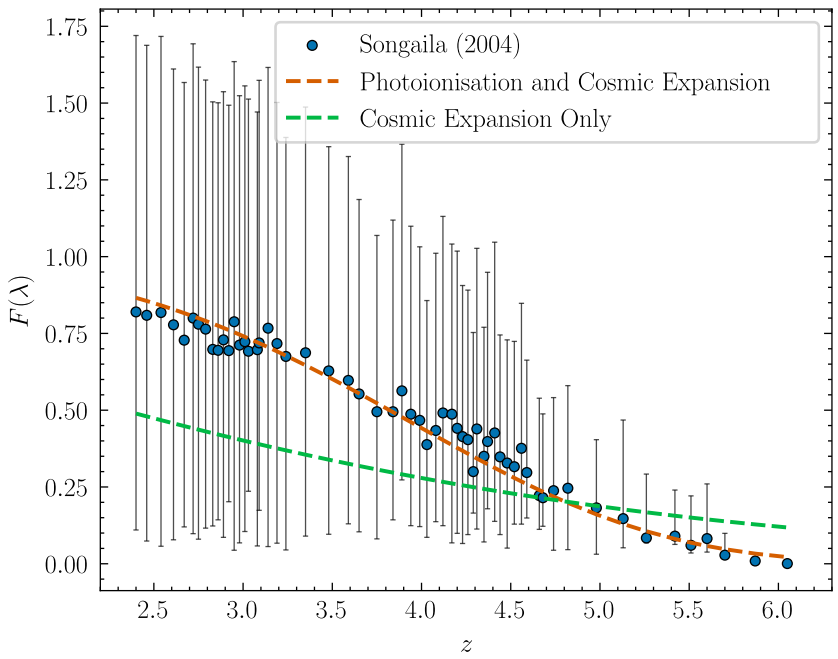

Figure 6: Best fit functions for Songaila (2004) dataset. The best fit values for Photoionisation and Cosmic Expansion model and Cosmic Expansion model only are $\Gamma_{\rm HI} \approx 2.57 \times 10^{-12}$ s^{-1} and $x_{\rm HI} \approx 6.07 \times 10^{-6}$ respectively.

where σ_i is the individual error, and N is the number of data points. The obtained best fit value is $\Gamma_{\rm HI} \approx 1.26 \times 10^{-12}$ s^{-1}.

Part 4: Sources to account for Photoionization

In the previous part, we randomly best-fitted the $\Gamma_{\rm HI}$ parameter. In this section, we account for the physical parameters that contribute to the photoionisation. The largest contributor is the ionizing radiation from the galaxies. As discussed previously in this course, the UVLF $\Phi(L)$ which is the number of galaxies per unit comoving volume per unit luminosity, can be used to quantify the total ionising radiation if we integrate over the luminosity. The photoionisation rate can be written as,

$$\Gamma_{\rm HI} = \frac{\sigma_L \lambda_{\rm mfp}}{h_p(\alpha_s + 3)} f_L (1+z)^3 \int_{L_{\rm UV}^{\rm min}}^{\infty} L_{\rm UV} \Phi(L_{\rm UV}, z) {\rm d}L_{\rm UV} \tag{21}$$

where $\sigma_L \approx 6.3 \times 10^{-18}$ cm^2 is a proportionality constant between the photoionization cross-section and the frequency dependance, $\lambda_{\rm mfp} \approx 100$ Mpc is the mean free path for reabsorption, h_p is the Plancks constant, $\alpha_s \approx 3$ is the UV spectral slope, and f_L is the fraction of ionizing luminosity per unit UV luminosity.

The UVLF can also be written in terms of absolute magnitude as a Schechter function,

$$\Phi(M_{\rm UV}) = \frac{\ln 10}{2.5} \phi_* 10^{-0.4(M_{\rm UV} - M^*_{\rm UV})(\alpha+1)} \exp\left(-10^{-0.4(M_{\rm UV} - M^*_{\rm UV})}\right) \tag{22}$$

with ϕ_*, $M^*_{\rm UV}$,α being best fit parameters from literature at $z = 4$.

We can recast the integral in equation 21 in units of absolute magnitude using the Luminosity Magnitude relation.

$$M_{\rm UV} = 51.6 - 2.5 \log_{10}(L_{\rm UV}) \tag{23}$$

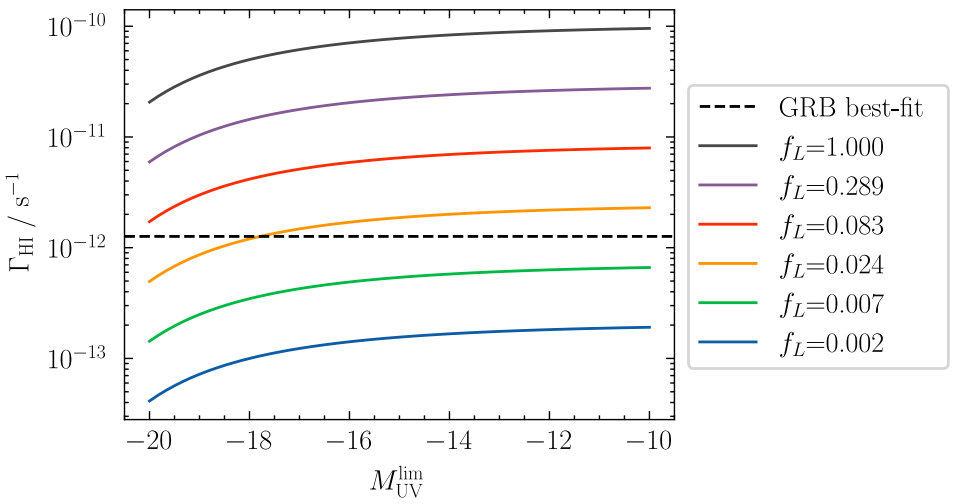

Figure 7: Photoionisation rate against limiting UV-magnitudes for different ionising luminosity fraction. From the

Therefore equation 21 can we written as,

$$\Gamma_{\rm HI} = \frac{\sigma_L \lambda_{\rm mfp}}{h_p(\alpha_s+3)} f_L (1+z)^3 \int_{\infty}^{M_{\rm UV}^{\rm min}} L_{\rm UV}(M_{\rm UV})\Phi(M_{\rm UV}){\rm d}M_{\rm UV} \tag{24}$$

This equation can be casted as a `Python` function as shown below with proper unit conversion to account for dimensional consistency.

```
import numpy as np
import astropy.units as u
from scipy.integrate import quad

def schechter_function_M_UV(M_UV, phi_star=1.46e-3, M_UV_star=-20.95,alpha=-1.72):
    first_term = (np.log(10)/2.5)*phi_star
    second_term = 10**(-0.4*(M_UV-M_UV_star)*(alpha+1))
    third_term = np.exp(-10**(-0.4*(M_UV-M_UV_star)))
    return first_term*second_term*third_term

def luminosity_magnitude_relation(M):
    return 10**((51.6-M)/(2.5)) # L is in erg s^-1 Hz^-1

def integral_helper(M):
    return luminosity_magnitude_relation(M)*schechter_function_M_UV(M)

def photoionisation_rate(f_L,M_lim,mfp=100,alpha_s=3,z=4):
    sigma_L = ((6.3e-18)*(u.cm**2)).to(u.Mpc**2).value #Mpc^2
    hp = 6.626196e-27 #erg s
    integral,_ = quad(integral_helper,-200,M_lim)
    return sigma_L*mfp*f_L*((1+z)**3)/(hp*(alpha_s+3))*integral
```

Using the above functions, we plot the values of $\Gamma_{\rm HI}$ against $M_{\rm UV}^{\rm lim}$ for different values of f_L as shown in Figure 7. The best-fit value of $\Gamma_{\rm HI}$ from the previous part is also plotted on the graph. That measured value implies the IGM for the GRB could've had a limiting magnitude in the range of -20 to -10, and an escape ratio of somewhere on the order of

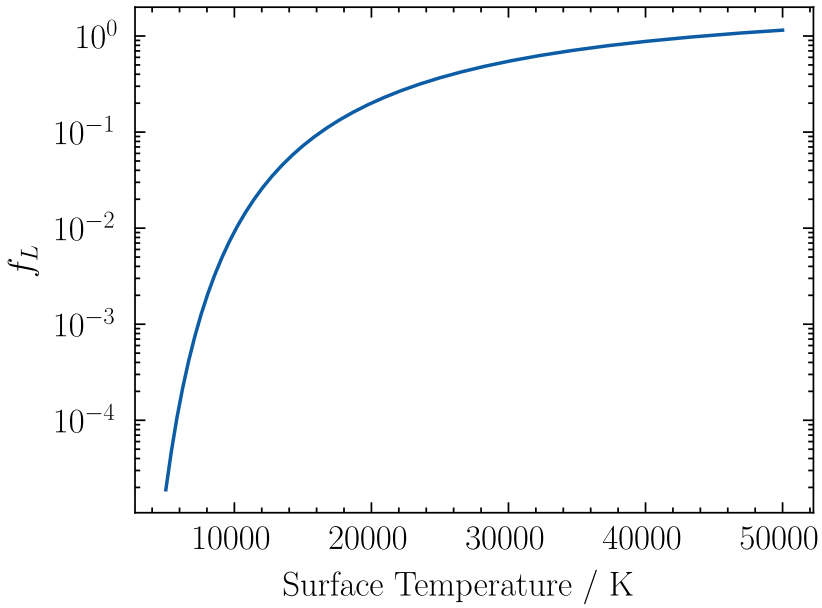

Figure 8: Escape fraction as a function of the stellar surface temperature. The stars are assumed to be perfect blackbodies, and the amount of emitted radiation is calculated using the Planck blackbody function.

10^{-2}. Suppose the limiting magnitude is -15, thus the value of f_L will be 0.0162 to account for the measured value of $\Gamma_{\rm HI}$. This was calculated using `scipy.optimize.fsolve`.

Now suppose the ionizing radiation is being emitted from stars of the same surface temperature. We use the Planck blackbody function, defined by,

$$B_\nu(T) = \frac{2h_p\nu^3}{c^2}\frac{1}{\exp\left(\dfrac{h_p\nu}{kT} - 1\right)} \tag{25}$$

to estimate the escape fraction. By definition,

$$f_L = \frac{L_{\rm ion}}{L_{\rm UV}} = \frac{B_{c/912\text{Å}}(T)}{B_{c/1500\text{Å}}(T)} \tag{26}$$

Figure 8 shows the different values of f_L across different Surface Temperatures. However, the current calculation of f_L is unphysical, because not every single photon that is emitted will contribute to the radiation. There will be a fraction which will be absorbed by the galaxy itself. Therefore, we recast the ratio as,

$$f_L = f_{\rm esc} f_L^{\rm intr} \tag{27}$$

where $f_{\rm esc}$ is the fraction of photons that do not get absorbed, and $f_L^{\rm intr}$ is the actual fraction that is emitted.

Suppose instead of -15, the limiting magnitude was -10. To account for the measured value of $\Gamma_{\rm HI}$, the value of $f_L \approx 0.0132$. Suppose the galaxies are all made of stars with a surface temperature of 30000 K, and thus $f_L^{\rm intr} \approx 0.546$. Therefore, $f_{\rm esc} \approx 0.0242$.