Second edition
of Cosmology in the Alps conference

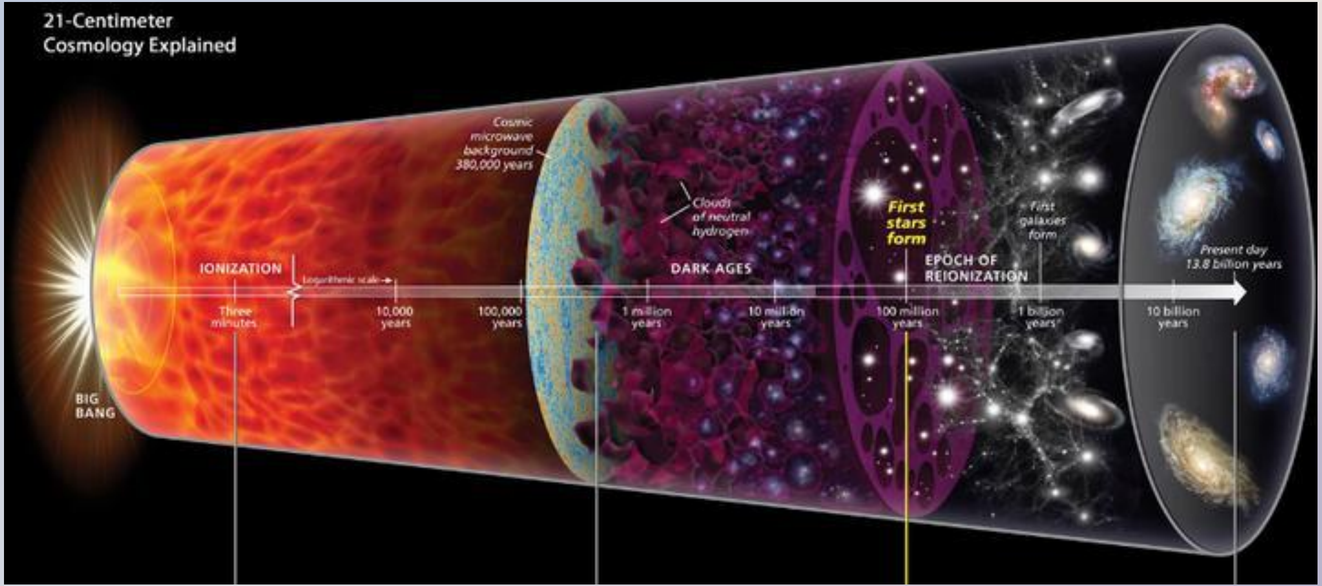
16 - 20 MARCH 2026
Les Diablerets, Hotel Glacier

Linking Galaxy Evolution and the 21-cm Signal: Joint Machine Learning Constraints on the Epoch of Reionization

Anshuman Tripathi (1)*, Aniket Kumar(2), Abhirup Datta(1)

1) Department of Astronomy, Astrophysics and Space Engineering
Indian Institute of Technology, Indore, India
2) Department of Physical Sciences IISER Mohali, India

The first billion years!



Epoch of reionization (EoR): weakly constrained

- ★ First sources of radiation?
- ★ How reionization happened?

Lyman α: large optical depth, absorbed at high neutral fraction of hydrogen.

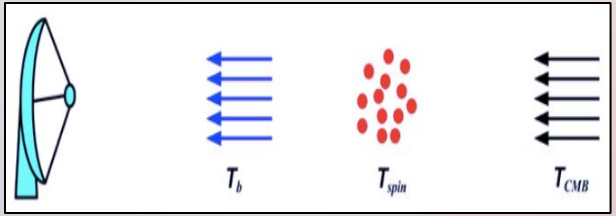
21 cm emission

Credits: Kelly/Discover magazine

Observation of 21-cm Signal

$$\delta T_b(\nu) = \frac{T_s - T_\gamma}{1+z}(1 - \exp^{-\tau_{\nu_0}})$$

$$\delta T_b \approx 27(1 - x_i)\left(\frac{\Omega_{b,0}h^2}{0.023}\right)\left(\frac{0.15}{\Omega_{m,0}h^2}\frac{1+z}{10}\right)^{1/2}\left(1 - \frac{T_\gamma}{T_s}\right)$$

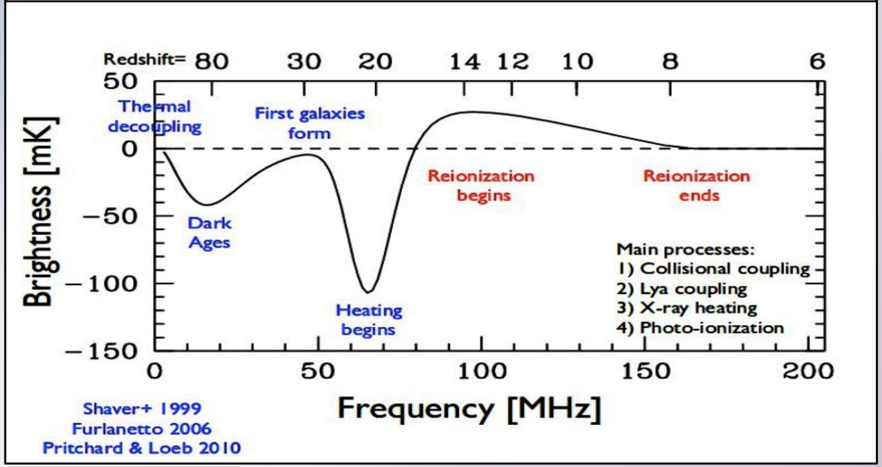


Field 1959; Pritchard & Loeb 2012

Observing type: Single radiometer

Signal: Global 21-cm Signal

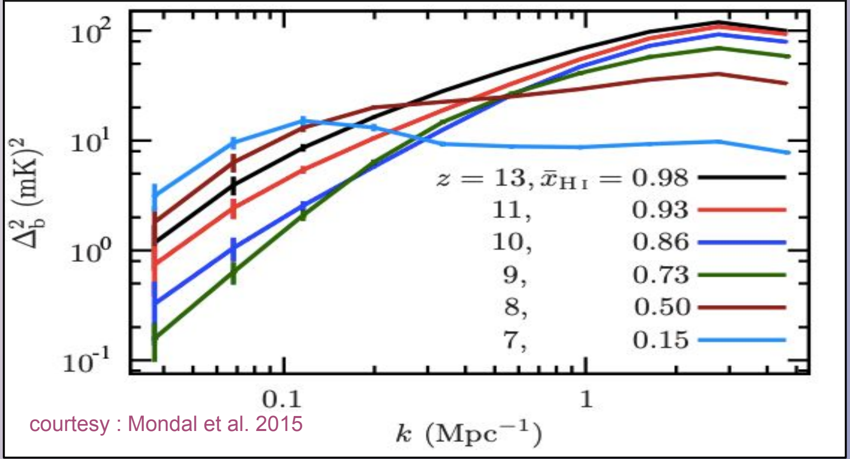
Experiments: EDGES, SARAS, BigHorns, LEDA, REACH



Observing type: Radio Interferometers

Signal : HI 21cm Power Spectrum, Tomography of IGM

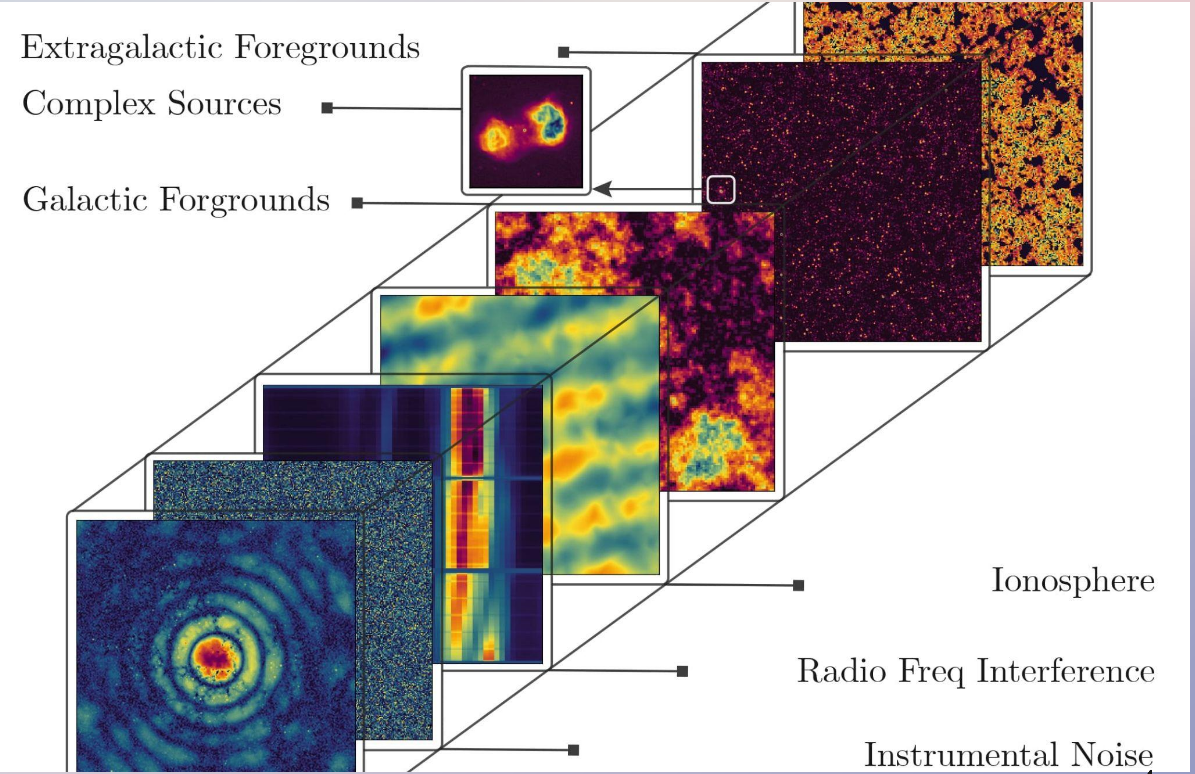
Experiments : LOFAR, GMRT, HERA, MWA, SKA



courtesy : Mondal et al. 2015

A challenging experiment

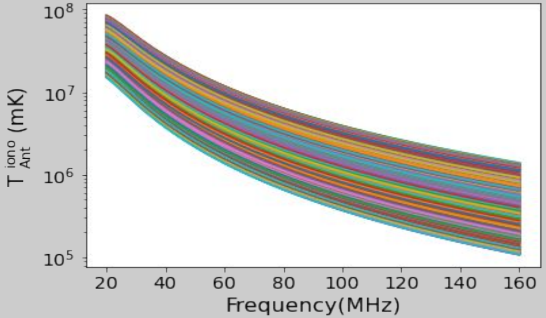
- ❑ Astrophysical foreground
- ❑ Earth's Ionosphere
- ❑ Radio Frequency Interference (RFI)
- ❑ Thermal Noise
- ❑ Instrumental Systematics



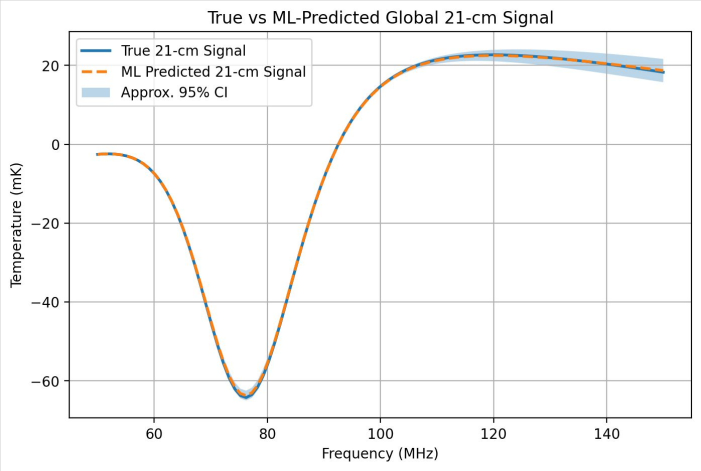
Chokshi et al., 2024

ANN Framework for Global 21cm

Global 21cm Signal
+
Foreground
+
Ionospheric effects
+
Thermal Noise



Developed an **ANN-based framework** to rapidly infer **21-cm signal parameters** directly from **composite sky spectra**, including foregrounds and instrumental systematics.



ANN Model

Saved ANN Model

Test Data set (Signal + Foreground +Ionospheric effect + Thermal Noise)

	P1	P2	P3	P4	P5	P6
True	14.2000	6.7000	9.7700	3.3100	2.8300	2.8200
Predicted	14.1595	6.2139	9.7655	3.3256	3.3230	2.8217
Error	0.0198	0.5248	0.0166	0.0097	0.5393	0.0050

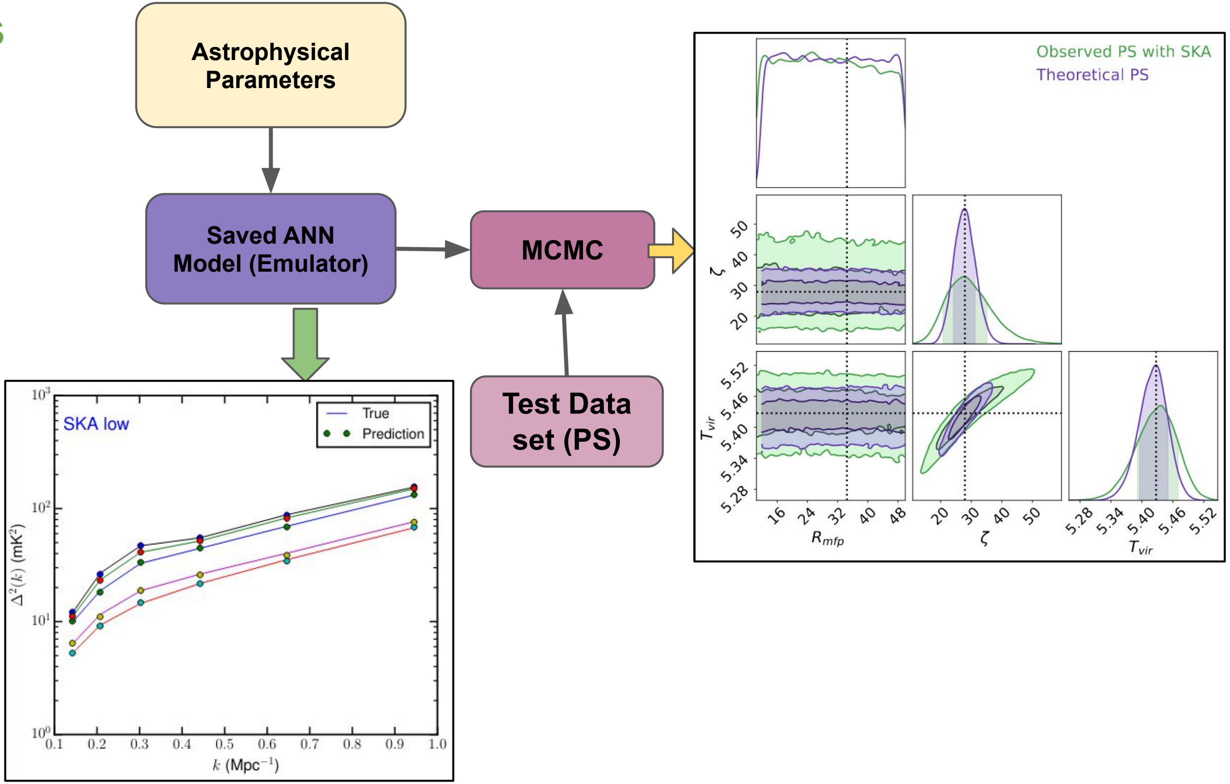
TARANG-21: Training-based AI for Recovery & Analysis of the Global 21-cm Signal

Building a signal statistics emulator

- Speed up the random walk in a large multi-dimensional parameter space by emulating the signal statistics instead of time consuming simulations

- Sample the parameter space with even smaller step size

- Compare efficiency of various statistics in constraining cosmic dawn parameters



Tripathi A. et al. 2025, JCAP, arXiv:2502.20962

The UV Luminosity Function as a Cosmic Blueprint

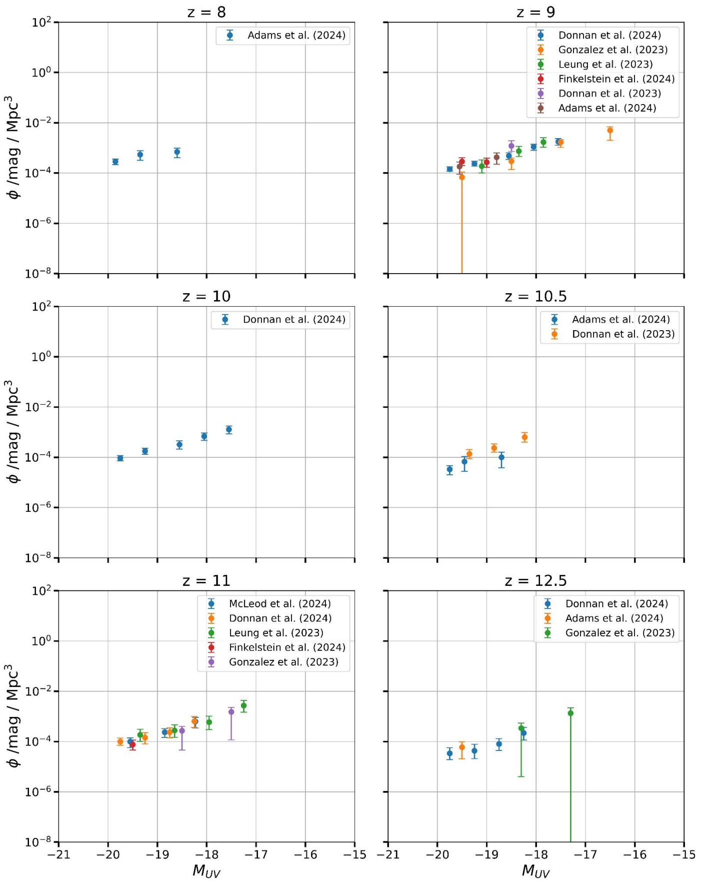
Direct Tracer of Star Formation: Rest-frame UV light is dominated by massive, short-lived O and B stars, providing a real-time snapshot of the cosmic Star Formation Rate (SFR).

The Reionization Photon Budget: Integrating the UVLF reveals the total number of ionizing photons escaping into the Intergalactic Medium (IGM) to drive the Epoch of Reionization.

Mapping Dark Matter Halos: The shape of the UVLF translates directly to the underlying Dark Matter Halo Mass Function, linking visible light to invisible mass.

Probing Feedback Physics: The faint-end slope and turnover mass (Mturn) tell us exactly where and how early star formation was suppressed by supernovae or inefficient cooling.

The JWST Revolution: Observations at z>10 test our fundamental assumptions about how rapidly the very first galaxies assembled.



Tripathi + Kumar et al. in prep

UV Luminosity Function (UVLF): Traces early galaxy populations and constrains star-formation and ionizing photon production efficiencies.

$$f_* = f_{*,10}\left(\frac{M_h}{10^{10}\,M_\odot}\right)^{\alpha_*}, \qquad M_* = f_*\left(\frac{\Omega_b}{\Omega_m}\right)M_h,$$

$$\dot{M}_* = \frac{M_*}{t_*\,H(z)^{-1}}$$

$$\log_{10}\left(\frac{L_{\rm UV}}{\rm erg\,s^{-1}\,Hz^{-1}}\right) = 0.4\,(51.63 - M_{\rm UV})$$

$$f_{\rm duty}(M_h) = \exp\left(-\frac{M_{\rm turn}}{M_h}\right)$$

$$\phi(M_{\rm UV}) = \left|\frac{{\rm d}M_h}{{\rm d}M_{\rm UV}}\right|\left[f_{\rm duty}(M_h)\frac{{\rm d}n}{{\rm d}M_h}\right]$$

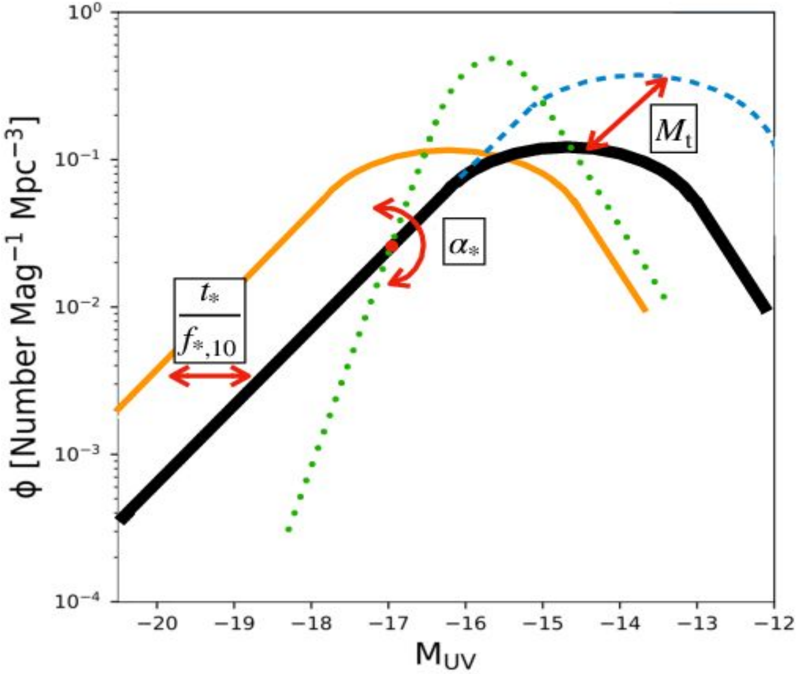


Fig: Schematic view of the influence of the model parameters on the UVLF. Figure adapted from Gillet et al. (2020).

21cmFAST_v4: Key Parameters Used to Model the JWST UV Luminosity Function

$f_{*,10}$: The baseline star formation efficiency for a $10^{10} M_{\odot}$ halo.

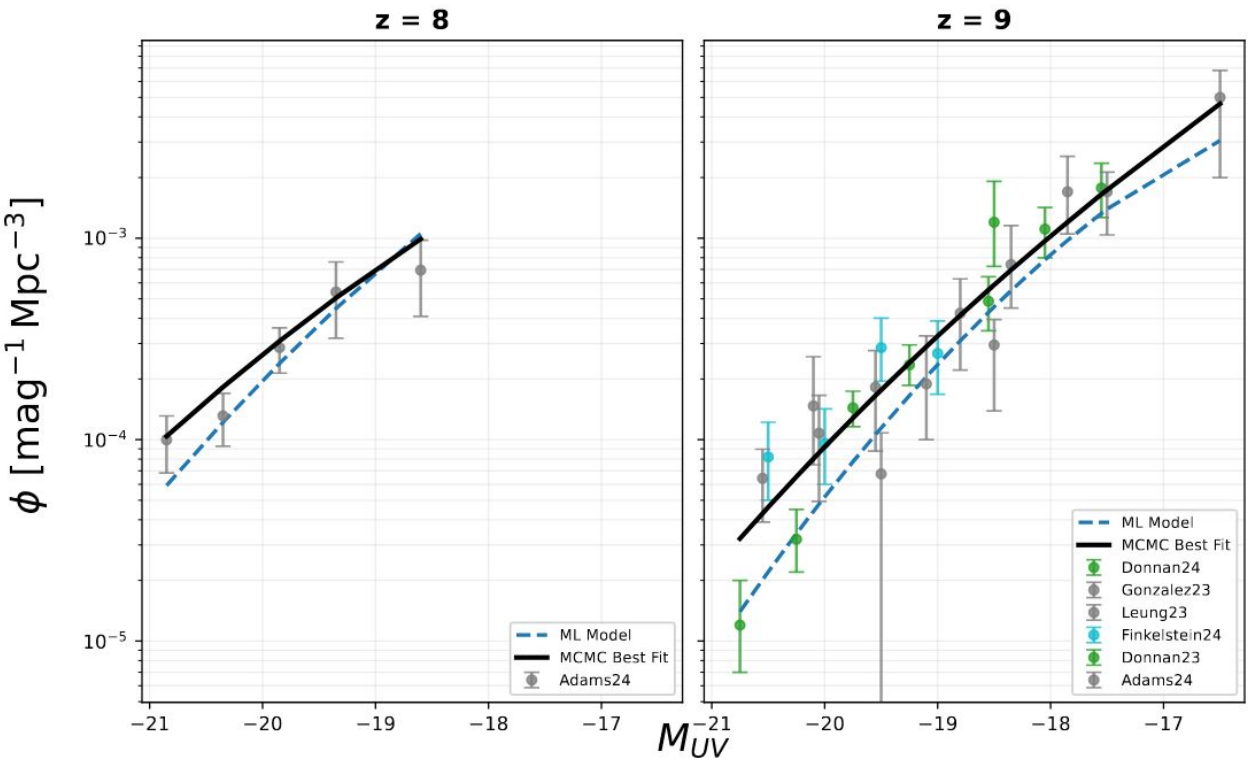
α_*: How steeply star formation efficiency scales with halo mass (mass-dependence).

M_{turn} : The minimum halo mass threshold where star formation is exponentially suppressed (usually due to feedback or inefficient gas cooling).

t_*: The star formation timescale (represented as a fraction of the Hubble time at that redshift).

σ_* : The new scatter parameter, representing the burstiness/stochasticity of the galaxies.

Bayesian vs ML

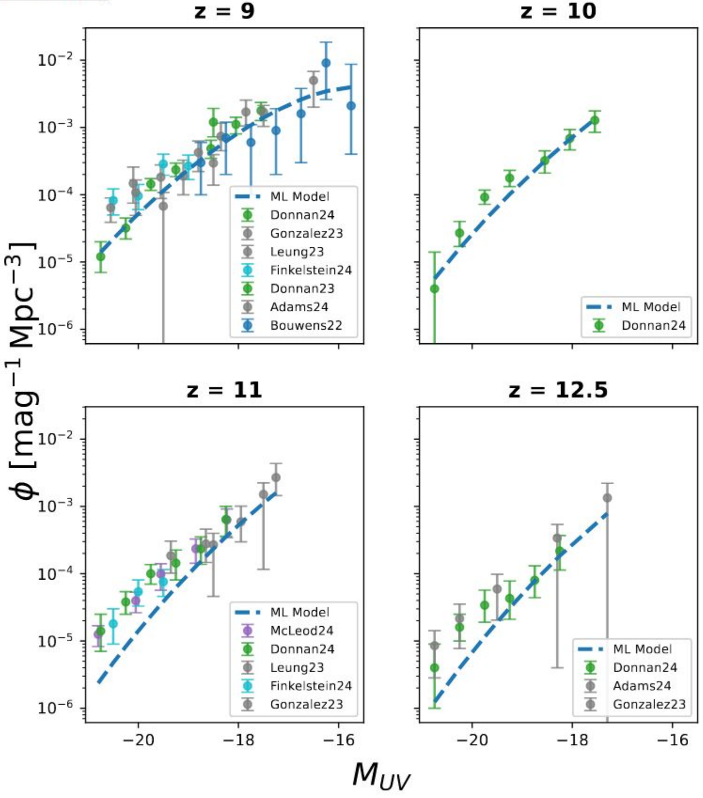


The Bayesian-based approach took a full day, while the machine learning-based approach completed in just about an hour.

Redshift (z)	$f_*, 10$	α_*	$M_{\rm turn}$	t_*	σ_*
$z = 9.0$	-1.168 ± 0.434	0.347 ± 0.092	9.870 ± 0.835	0.528 ± 0.017	0.482 ± 0.041
$z = 10.0$	-1.079 ± 0.093	0.287 ± 0.053	9.246 ± 0.201	0.497 ± 0.016	0.483 ± 0.003
$z = 11.0$	-0.939 ± 0.074	0.266 ± 0.054	8.901 ± 0.286	0.533 ± 0.019	0.543 ± 0.008
$z = 12.5$	-0.809 ± 0.490	0.317 ± 0.145	8.996 ± 0.522	0.494 ± 0.047	0.494 ± 0.013

We also calculated the chi-square values for this fit:

- For z=9, reduced chi-square = 1.80
- For z=10, reduced chi-square = 5.19
- For z=11, reduced chi-square = 2.90
- For z=12.5, reduced chi-square = 2.04



Evolution of Galaxy Properties

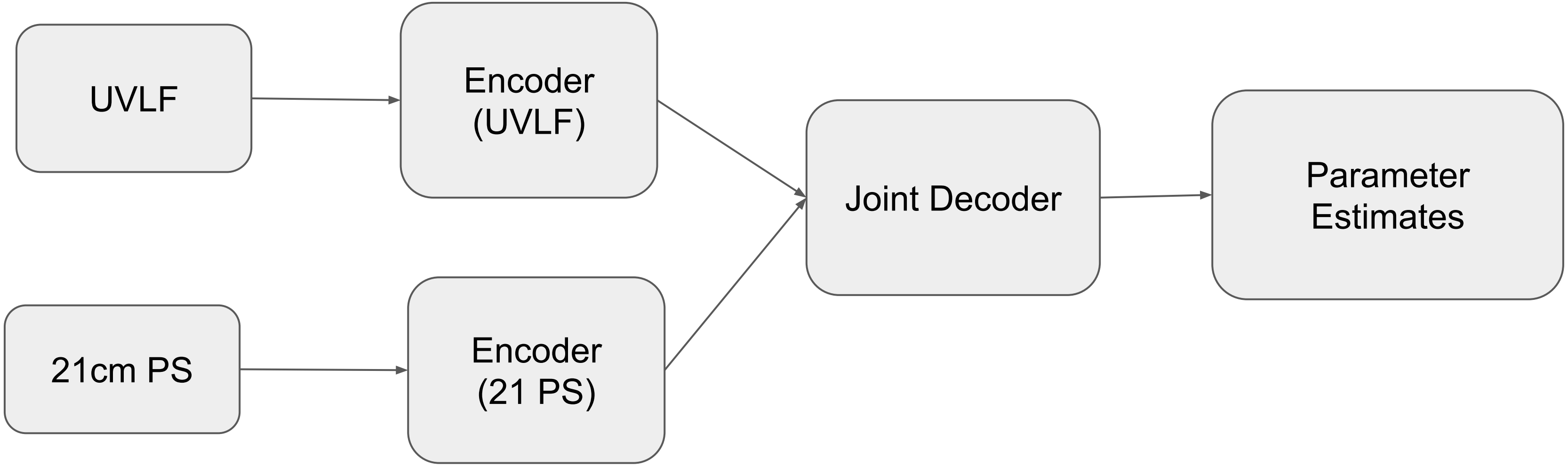
Increasing Star Formation Efficiency: The baseline efficiency (F*) steadily rises from −1.16 at z=9 to −0.81 at z=12.5. Early galaxies were remarkably efficient at converting gas into stars, which helps explain the excess of bright galaxies observed by JWST.

Active Star Formation in Smaller Halos: The turnover mass (M_{turn}) shifts down from ~9.87 at z=9 to ~8.90−9.00 at z>10. This suggests lower-mass dark matter halos were actively hosting star formation in the very early Universe, possibly before accumulated feedback mechanisms could heavily suppress them.

Consistent Burstiness: The stochastic scatter (σ*) remains highly stable around 0.5 across all measured redshifts. This confirms that early star formation was fundamentally bursty, requiring this scatter to properly fit the bright end of the UV Luminosity Function.

Stable Timescales and Scaling: Both the mass-scaling dependence (α*≈0.3) and the star formation timescale (t*≈0.5) show minimal evolution. The fundamental physics governing how long starbursts last and how efficiency scales with halo mass appear steady during this epoch.

Auto Encoder and Decoder



Joint Insight: Using UVLF constraints improves modeling of ionizing sources and refines predictions of the 21-cm signal from the IGM.

Joint UVLF–21-cm Analysis via ML

Case	$f_{*,10}$	α_*	$f_{\mathrm{esc},10}$	α_{esc}	t_*	M_{turn}	$\log_{10}(L_X)$
True Parameters	-2.39	0.302	-0.60	-0.686	0.87	8.72	39.34
21-cm PS Only	-0.88	0.322	-2.13	-0.184	0.50	8.99	40.94
21-cm PS + UVLF	-2.47	0.283	-1.13	-0.246	0.62	8.79	39.76

Ongoing work

Expand Training Data: Train the model on larger, more diverse simulated datasets to improve its overall accuracy and predictive performance.

Incorporate Updated Observations: Integrate the latest observational UVLF datasets (such as new JWST constraints) so the model reflects our most current understanding of early galaxies.

Thank You.