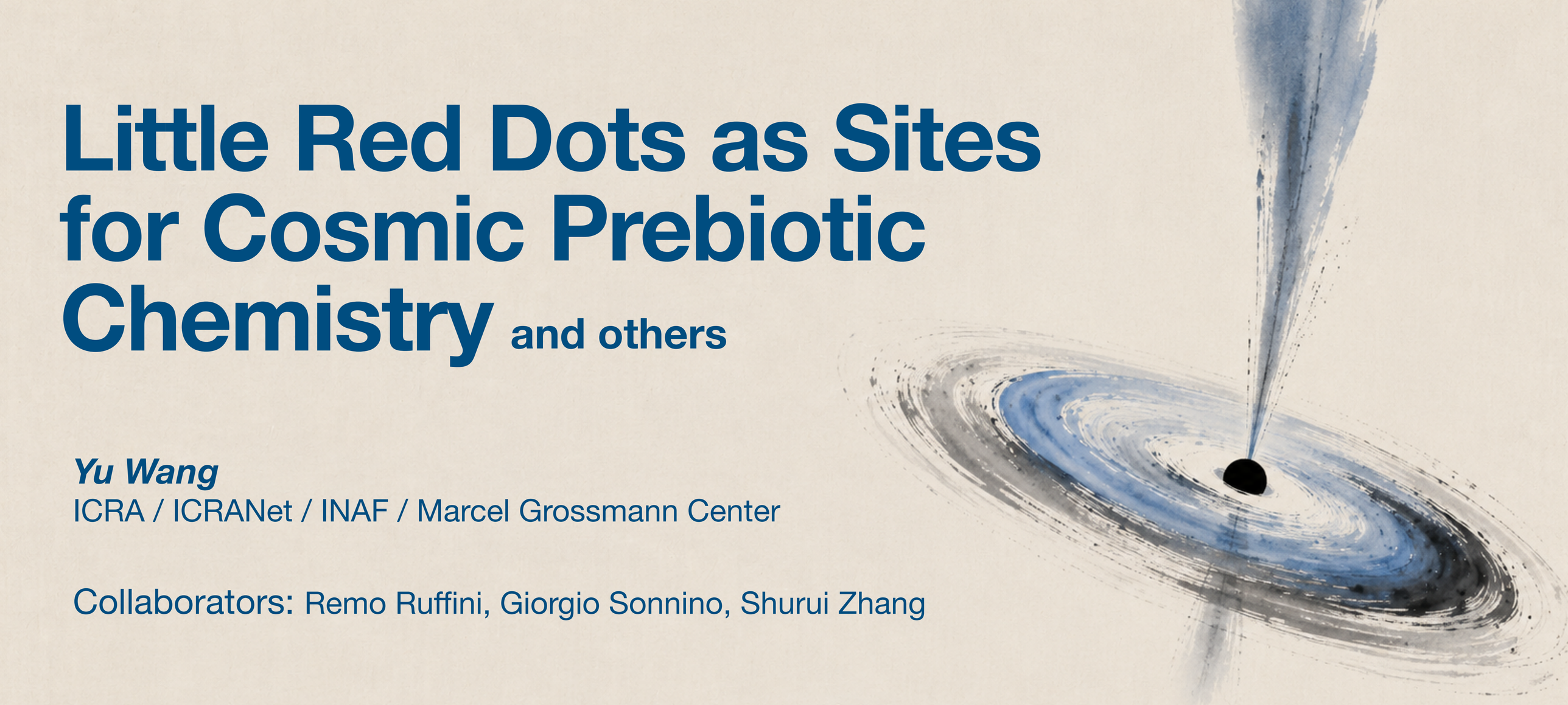
Little Red Dots as Sites for Cosmic Prebiotic Chemistry and others

Yu Wang
ICRA / ICRANet / INAF / Marcel Grossmann Center

Collaborators: Remo Ruffini, Giorgio Sonnino, Shurui Zhang

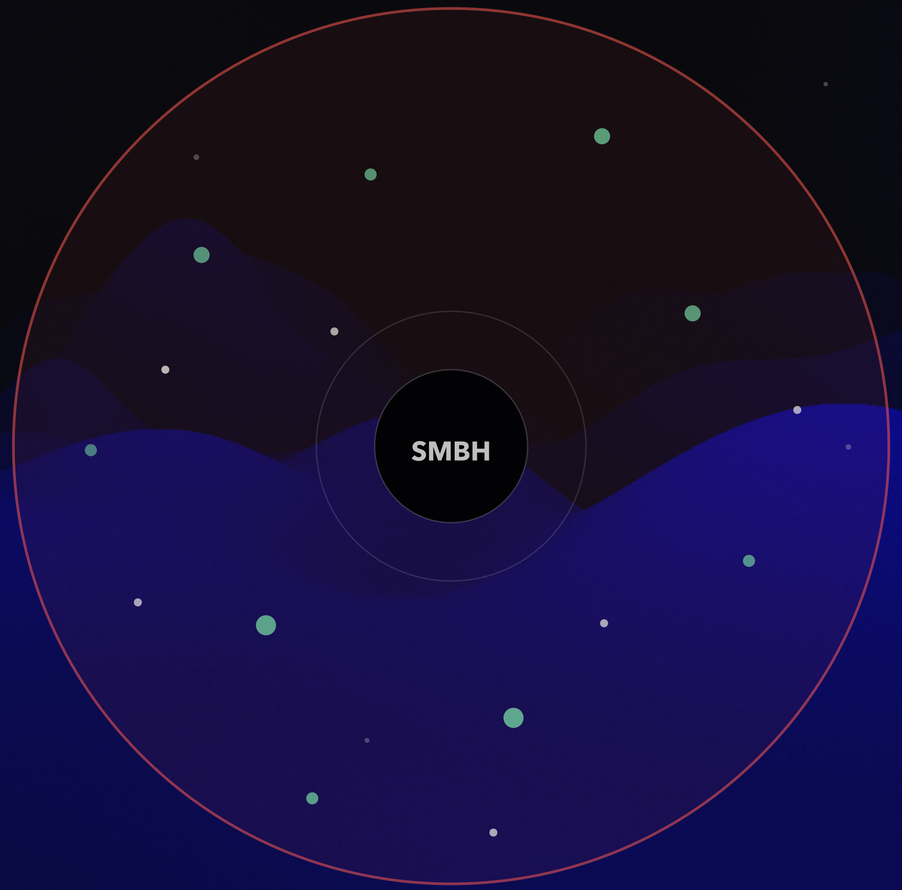


Jun 15–19 2026, Yerevan. High Energy Astrophysics and Cosmology in the era of all-sky surveys

Electromagnetically quiet galactic cores may shelter prebiotic chemistry

The paper places the Milky Way CMZ and JWST little red dots in the same physical frame: a million-solar-mass black hole does not necessarily erase its molecular reservoir. If UV/X-ray output is low, temperature is low, and cold dust survives, complex organic chemistry can accumulate.

black hole scale	chemical carrier	radiation field
$10^5 - 10^7$ **Msun**	**cold dust + molecular gas**	**weak UV / X-ray**



Cold molecular reservoir

Remo Ruffini and Yu Wang, ApJL, 2026

COMPARISON

LRD vs Milky Way CMZ

The table's quantities are recast as paired physical states: scale, mass reservoir, thermal window, and radiation field.

Physical quantity	 **Milky Way CMZ** local benchmark around Sgr A* — CMZ value	 **Little Red Dot** compact early-universe core — LRD value	
Radius overlap on ~100 pc scales	~100-250 pc	~30-200 pc	
Black-hole mass LRD spans broader range	4×10^6 Msun	$10^5 - 10^7$ **Msun**	
Stellar mass LRD range is wider	~10^7 Msun	$10^5 - 10^8$ Msun	
Molecular gas reservoir can reach CMZ scale	~10^7 Msun	$10^4 - 10^6$ **Msun**	inferred
Dust mass LRD has upper limit	~10^5 Msun	< 10^5 Msun	
Dust temperature cold-grain chemistry window	15-40 K	**< 50 K**	inferred
Gas temperature same nominal range	50-200 K	**50-200 K**	inferred
Ultraviolet subdominant in both	Subdominant	Subdominant	
X-ray LRD is undetected	**Weak**	**Undetected**	

CHEMISTRY WINDOW

Weak radiation plus 10-50 K dust opens a chemistry window

This does not prove an origin of life. It identifies a testable environment where molecules can form, survive, and later be redistributed.

1 Shield high energy

Weak accretion or internal obscuration reduces UV/X-ray destruction.

2 Preserve ice mantles

Dust near 10-50 K supports surface reactions and hydrogenation.

3 Accumulate organics

The CMZ cloud G+0.693-0.027 already shows nitrile precursors.

4 Seed early galaxies

If LRDs are common, they may contribute to early chemical enrichment.

Local anchor: the Milky Way CMZ

G+0.693-0.027 lies about 100 pc in projection from the Galactic Center. It is cold, dense, star-formation quiet, and rich in complex organics.

Distant hypothesis: LRDs

No X-ray detection, red spectra, and compact scale make LRDs candidate quiet cores. Direct molecular-line evidence is the next key test.

CMZ is the observable benchmark; LRDs are the early-universe analog candidates.

Life is assembled through cooperation among many parts of the Universe, much like an Airbus aircraft assembled from components produced in factories across the world and transported to a final assembly site.

Molecular clouds may manufacture some ingredients. Comets and small bodies may preserve and transport them. Planetary surfaces then provide the conditions for further chemistry for cells.

Not only biomolecules, but also the elements in our bodies come from different parts of the Universe. For example, nickel in our bodies may have originated in supernova explosions.

Speculation ***A living body can reflect the entire history of cosmic evolution.***

CMB temperature increases with redshift. So the universe had to expand and cool then dust-surface chemistry could become favorable at z=15~20 (T~50K). Remarkably, the same epoch when the first galaxies and dense structures began to form.

Speculation ***The universe did not only provide the chemical conditions for forming life; it also began building the galaxies where to host life.***

LRD reframes the origin of life from an “accidental event on Earth” into a “universal chemical process on cosmic scales.”

Speculation ***It is as if the universe were maximizing the probability for life to emerge.***

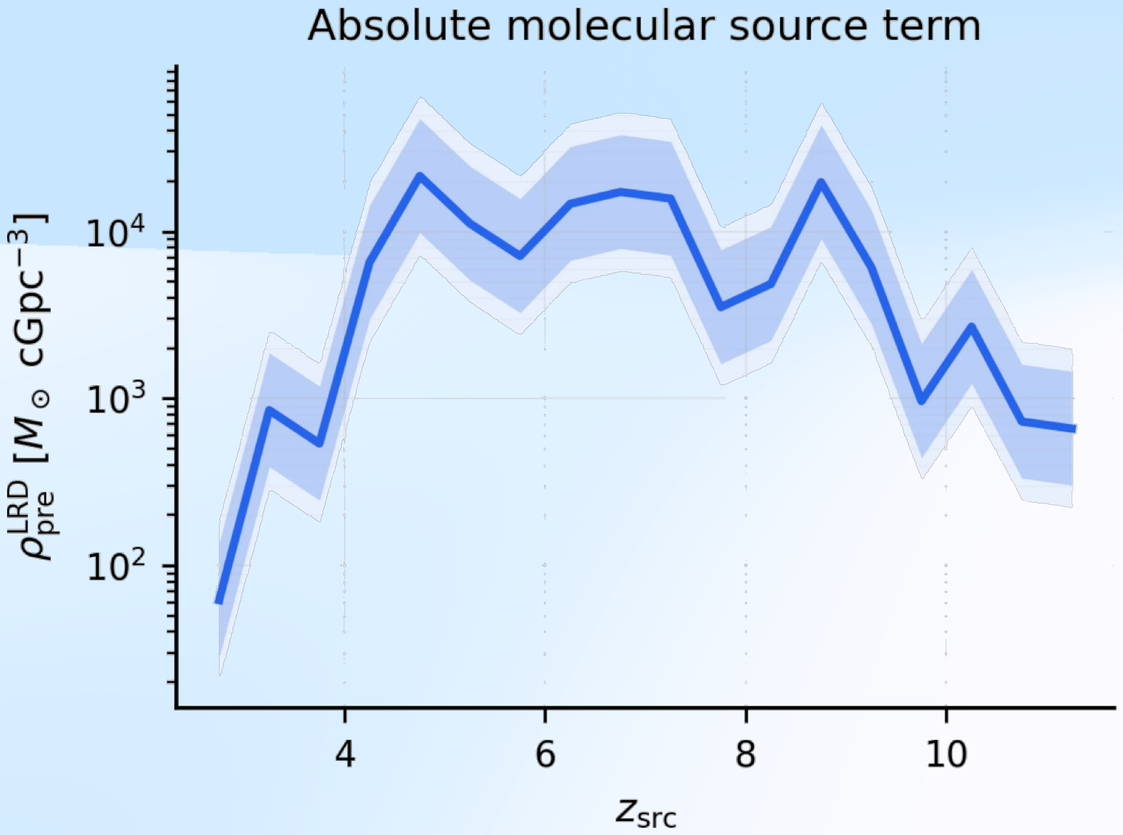
SOURCE TERM

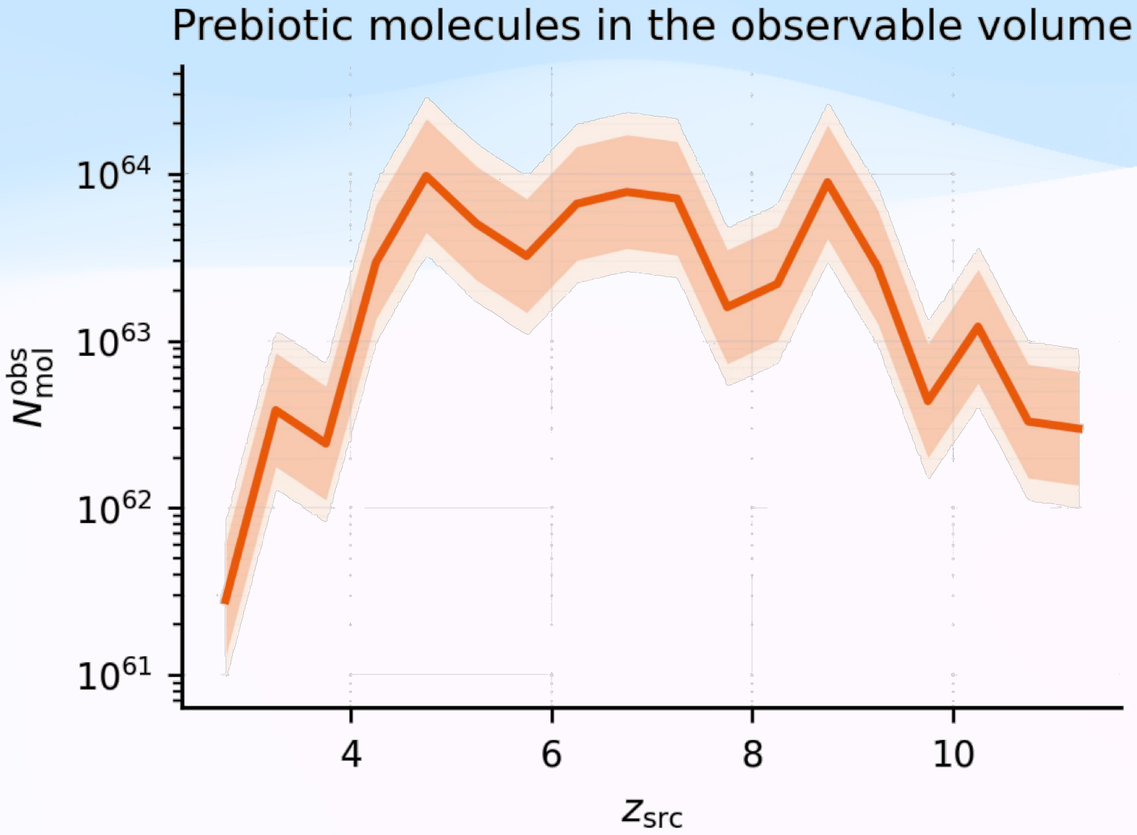
LRDs make the early molecular inventory large enough to matter

The first result is a timing result: prebiotic feedstock can be prepared before planet systems assemble.

$\sim 10^{64}$

molecules before pathway losses



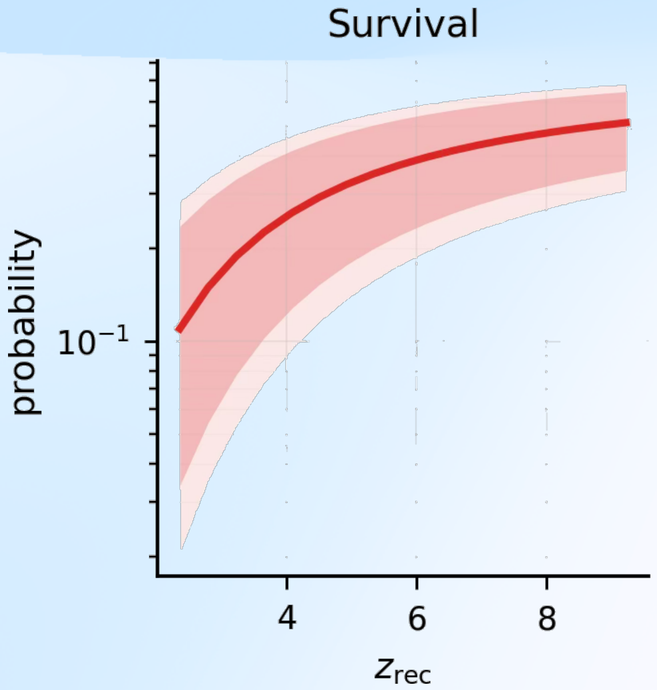


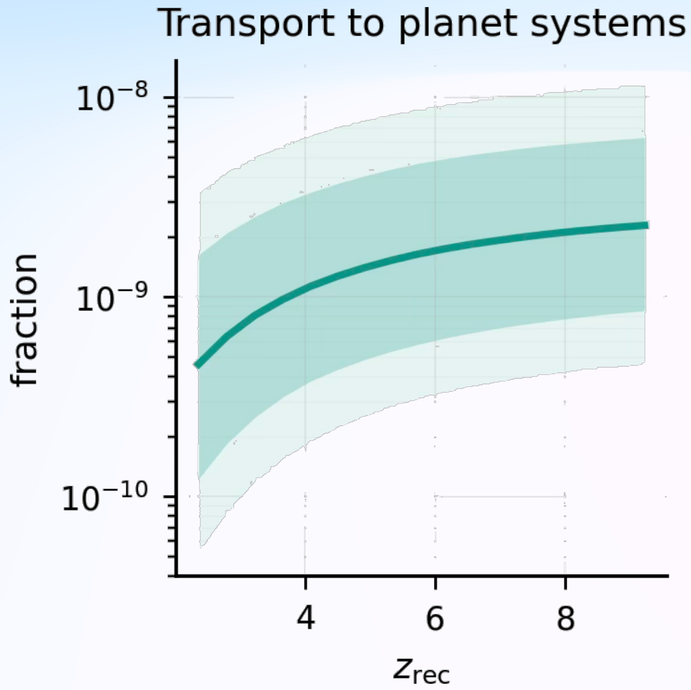
Yu Wang, Remo Ruffini and Giorgio Sonnino, 2026

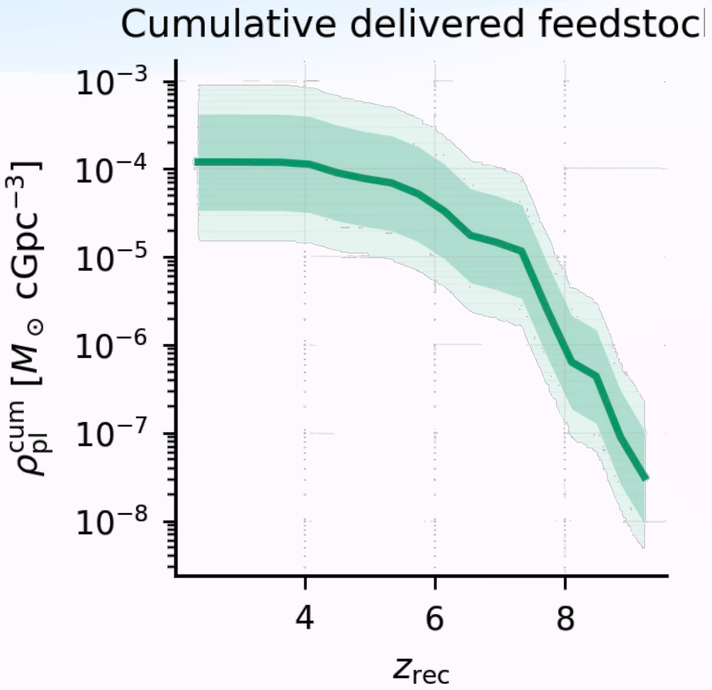
PATHWAY

The pathway is a sequence of physical filters

Material is counted only after ejection, survival, recapture, transfer into planet systems, and a finite dose.







ENDPOINT

Nearly all material is lost. Tens of planets remain.

This is a conditional count, but it is not zero.

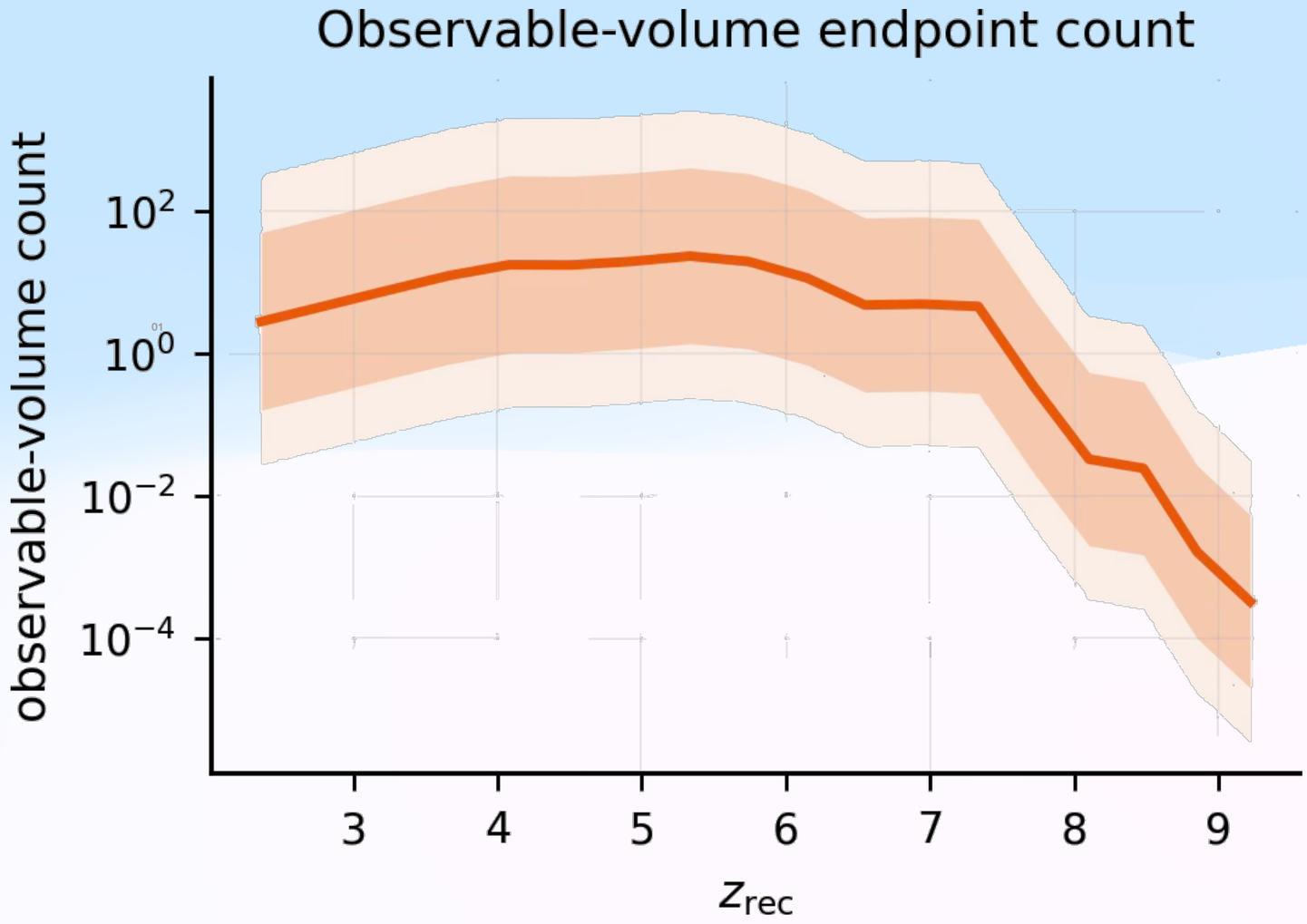
$\sim 10^{64}$

molecules before losses

ejection, travel, recapture, planet transfer, finite dose

~ 23

peak median planets



Growth of Little Red Dots

OBSERVATIONAL ANCHOR

LRDs are compact, black-hole dominated, and now locally anchored by an observed Egg.

Observed Egg

J1025+1402

Cool metal-enriched envelope dominates the emission.

X-ray Dot Transient

transition candidate

LRD-like optical continuum plus luminous X-rays.

Direct BH mass

Abell 2744-QSO1

Hα dynamics give $\log(M_{\rm BH}/M_{\odot}) \approx 7.7$; little room for an extended stellar component and dark matter halo.

Weak X-ray / radio

LRD branch signature

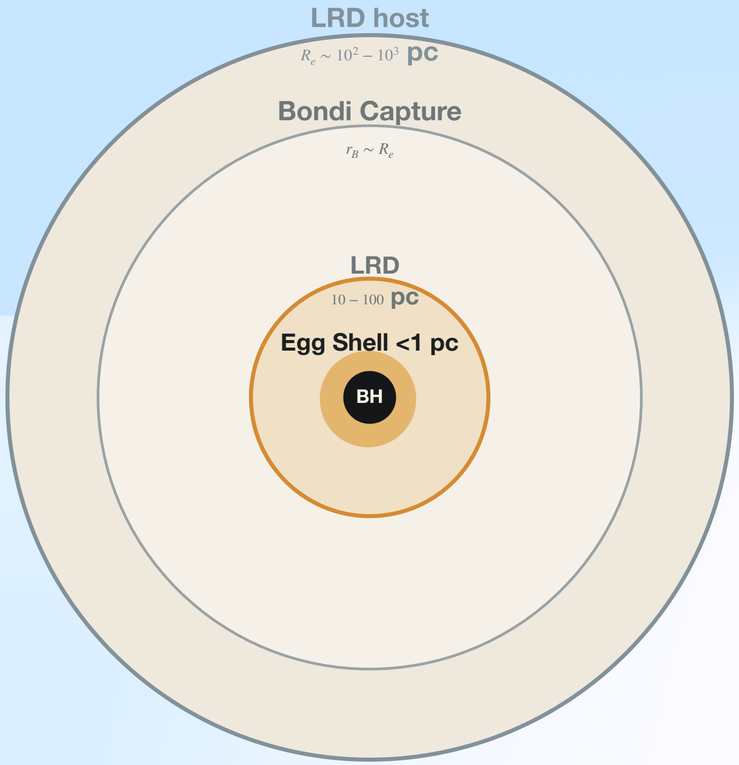
Envelope-dominated reprocessing suppresses coronal and jet output.

I. Juodžbalis et al. 2025, X. Lin et al. 2025, N. Lujan et al. 2025, R. E. Hviding et al. 2026

PHYSICAL PICTURE

Radius hierarchy: whole LRD capture, sub-pc Egg.

A black hole of LRD mass can influence the whole compact source, while the radiating envelope forms only after low-angular-momentum gas circularizes.



$$r_B = \frac{GM_{\rm BH}}{c_s^2} \simeq 430\,{\rm pc}\left(\frac{M_{\rm BH}}{5\times10^6\,M_\odot}\right)\left(\frac{c_s}{10\,{\rm km\,s^{-1}}}\right)^{-2}$$

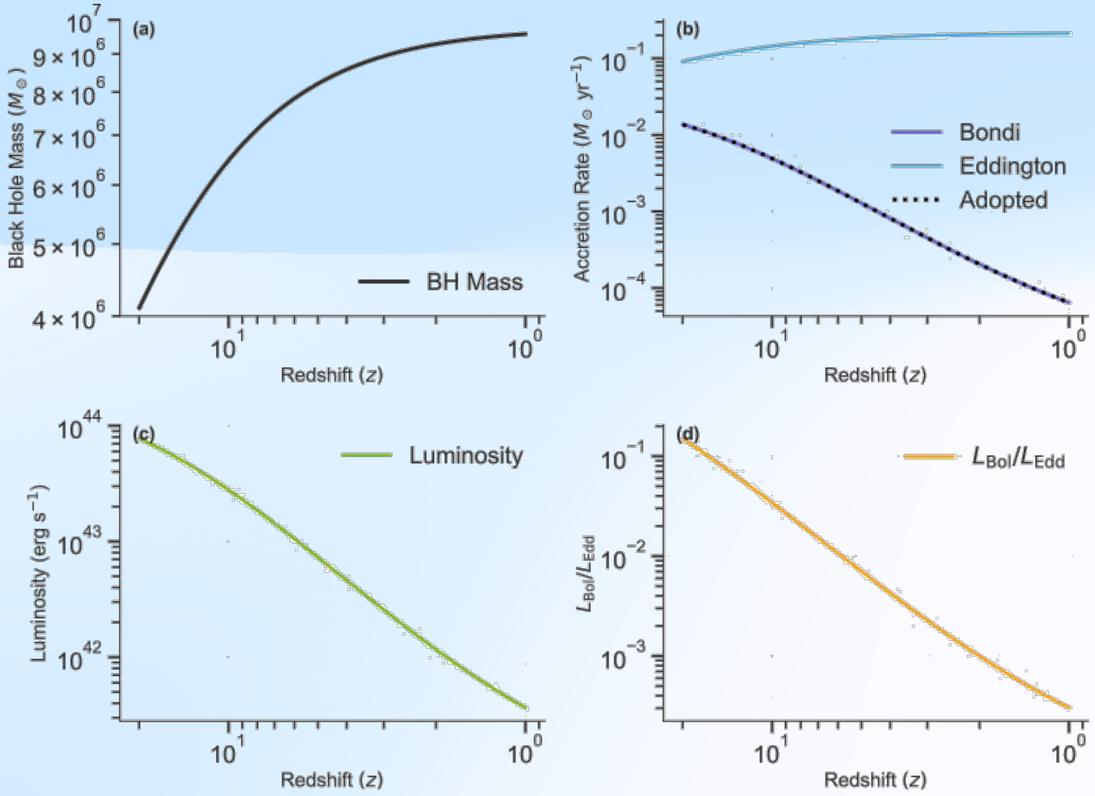
$$r_{\rm circ} = j^2/(GM_{\rm BH}) = r_B(v_\perp/c_s)^2 r_{\rm circ} \sim 4\times10^{-2}\,{\rm pc}$$

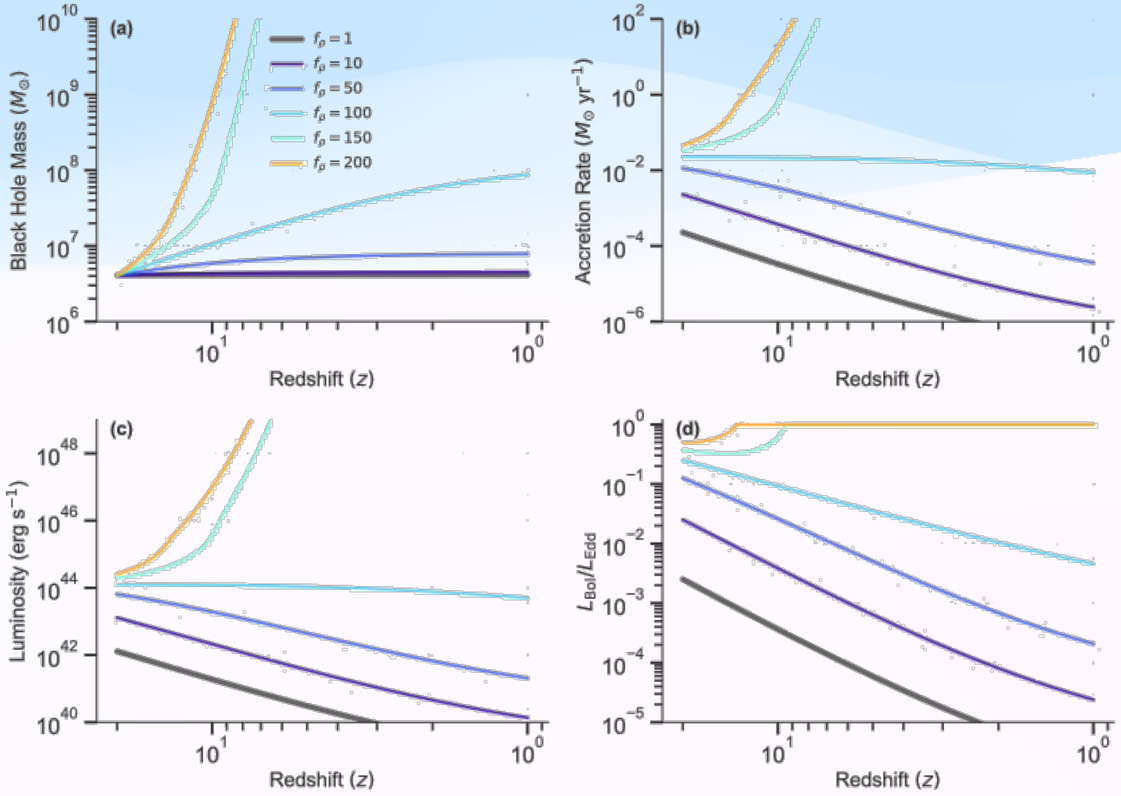
Interpretation: outside $r_{\rm circ}$ the flow is nearly radial; inside it, gas forms an optically thick Egg/photosphere that reprocesses accretion power.

Supply decline makes galaxy; dense environments reopen the AGN path.

The representative track shows supply-limited fading; the density sweep shows the same accretion law bifurcating into LRD and AGN outcomes.

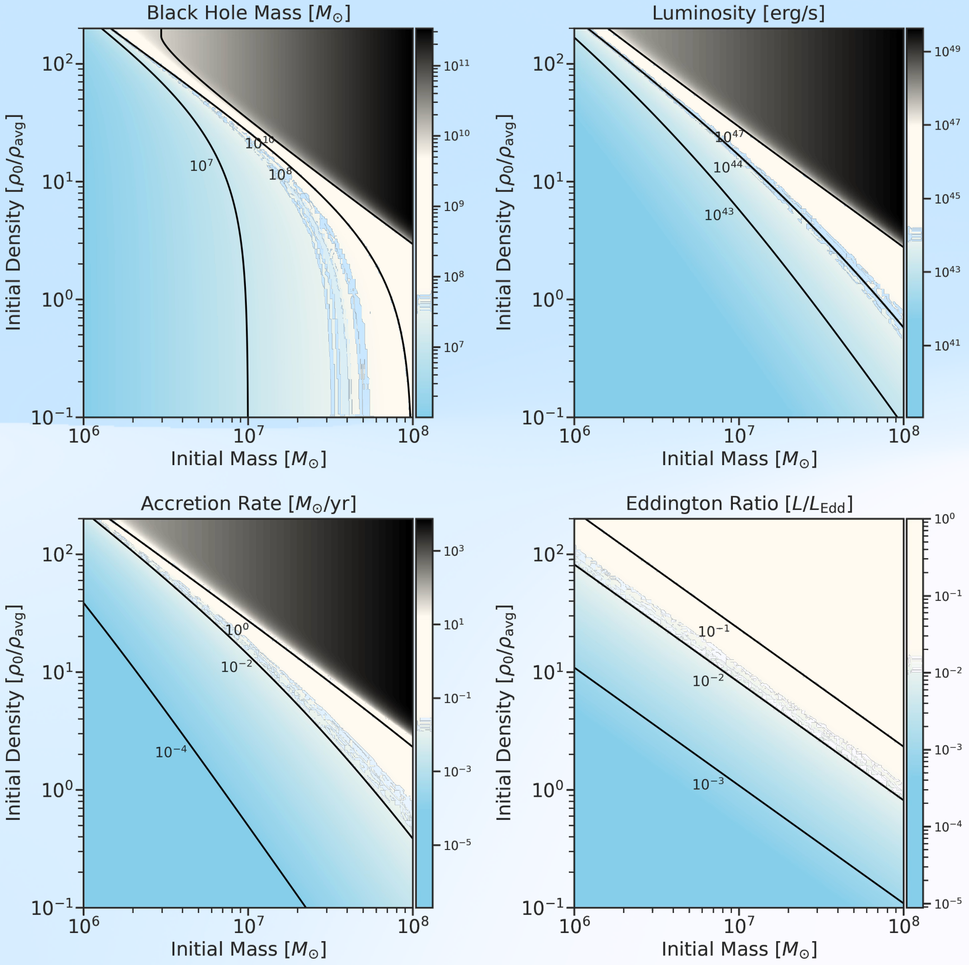
$$\rho(z) = f_\rho \rho_b(z), \quad \rho_b(z) = \Omega_b \rho_{\mathrm{crit},0}(1+z)^3 \qquad \dot{M}_{\mathrm{acc}} = \min(\dot{M}_{\mathrm{Edd}}, \dot{M}_{\mathrm{B}})$$

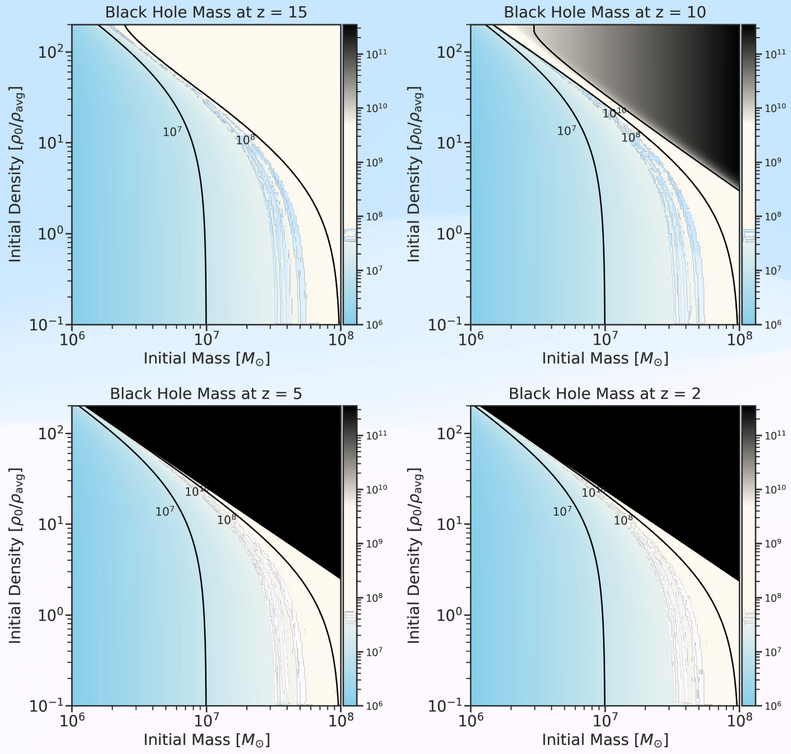




Parameter maps separate LRD/galaxy, transition, and AGN regimes.

Heatmaps show the regime growth from z=30 to z=10 (left); the redshift snapshots show how the mass develops with time (right).

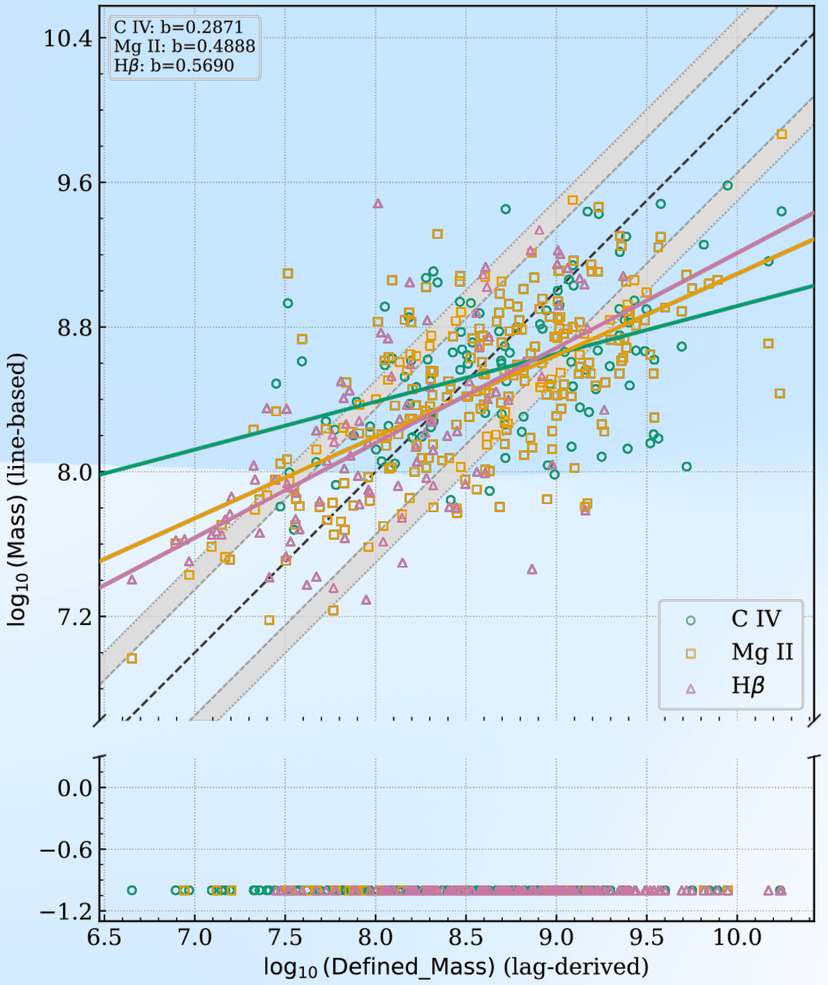


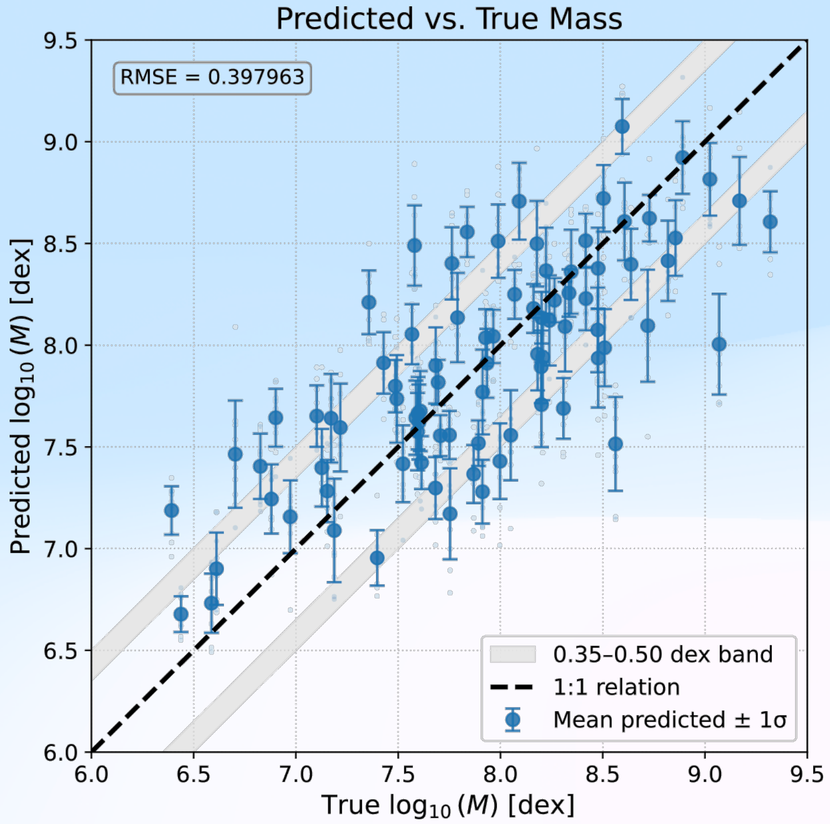


$$M_{\rm seed} = 10^{6-8}\,M_\odot, \quad f_\rho = 0.1\text{-}200, \quad z : 30 \rightarrow 10$$

Lower-left stays supply-limited and LRD-like / galaxies; upper-right reaches the rare high-supply AGN tail; the diagonal band is the transition region, where X-ray Dot-like objects belong.

263,604 Supermassive Black Hole Masse





1. Neural network predictions closely follow Reverberation Mapping masses with tight scatter. Making a population size level catalogue with high accuracy.

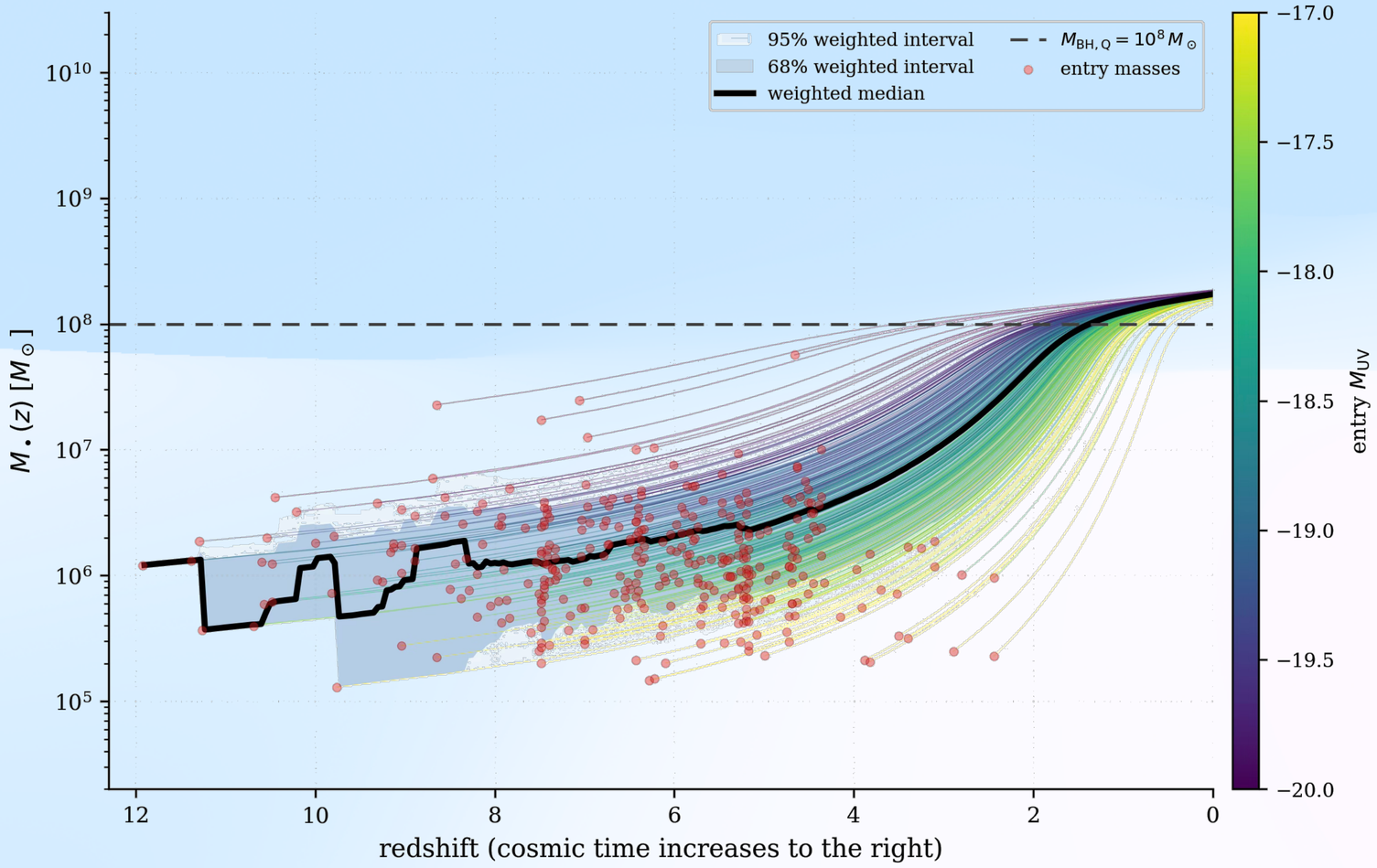
2. Neural network remains high accuracy for low ($< 10^{7.5} M_{\odot}$) and high ($> 10^{9} M_{\odot}$) mass, where virial estimators show strong biases.

3. Neural network predicts all quasar masses, while virial estimators fails when reliable emission lines are absent.

MINIMAL MODEL

Five parameters fit all LRD growth tracks

Goal: minimize freedom. Four abundance parameters plus one delay gate, each with a physical role.



Four baseline parameters

Values, physical meaning, and model role.

Eddington ratio

Population active-state Eddington-ratio scale; controls how strongly descendants grow once active.

Initial duty cycle

Effective occupancy of the luminous active channel before high-mass regulation.

Duty cycle evolution

One exponent for how the accretion environment evolves from the LRD epoch to lower redshift.

Percentage from LRD to AGN

Parent-population / completeness factor, not a probability; near unity means no large hidden correction.

One correction for delay

LRD Phase delay

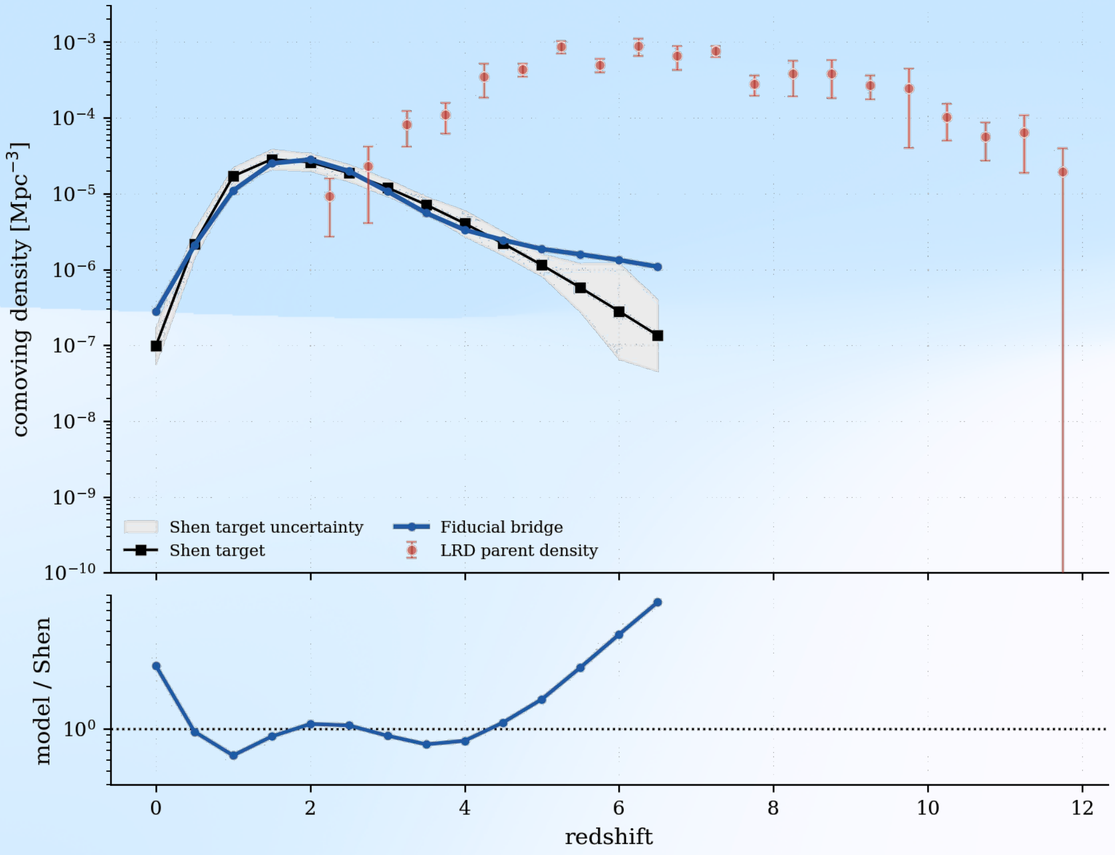
Clearing or maturation time before luminous unobscured AGN appear; Delta t~0.15 Gyr only smooths turn-on.

FIT AND DELAY

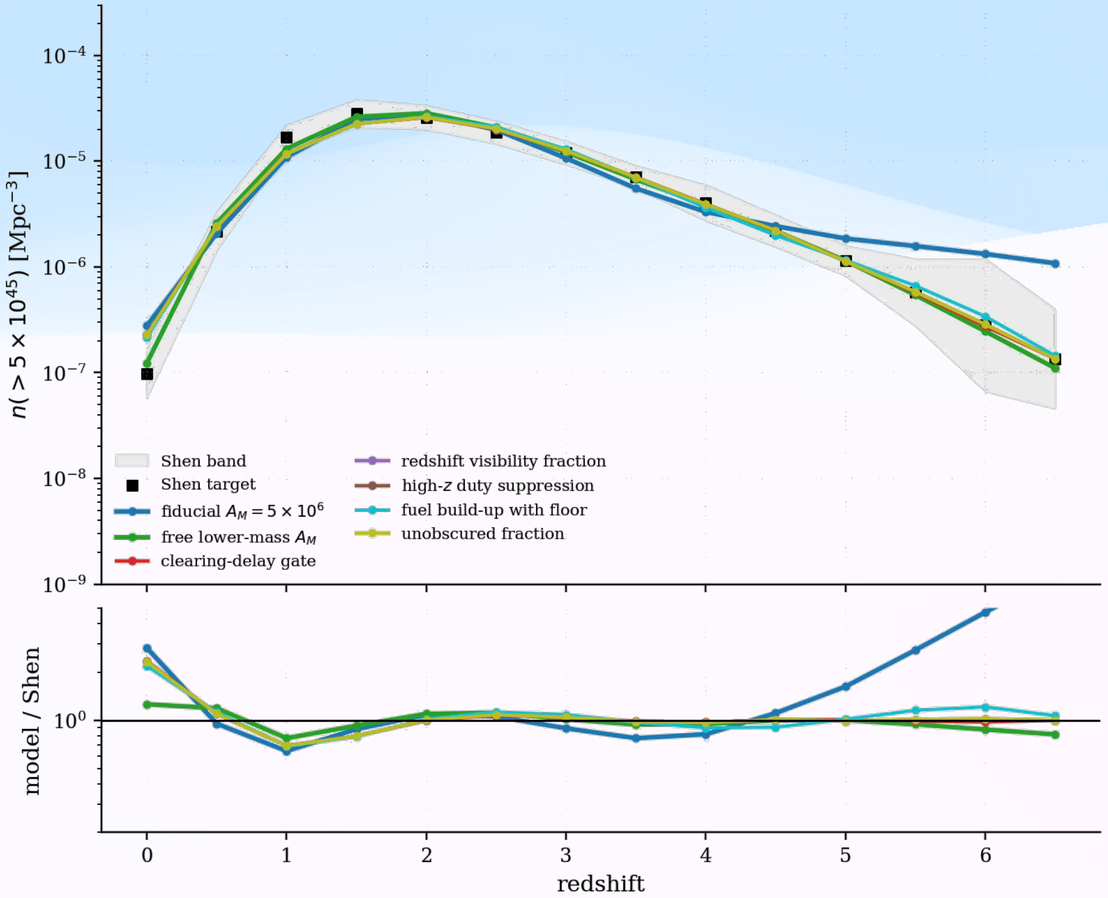
Good fit; delay fixes high-z timing

The baseline matches the main abundance scale. The delay gate suppresses the earliest visible descendants.

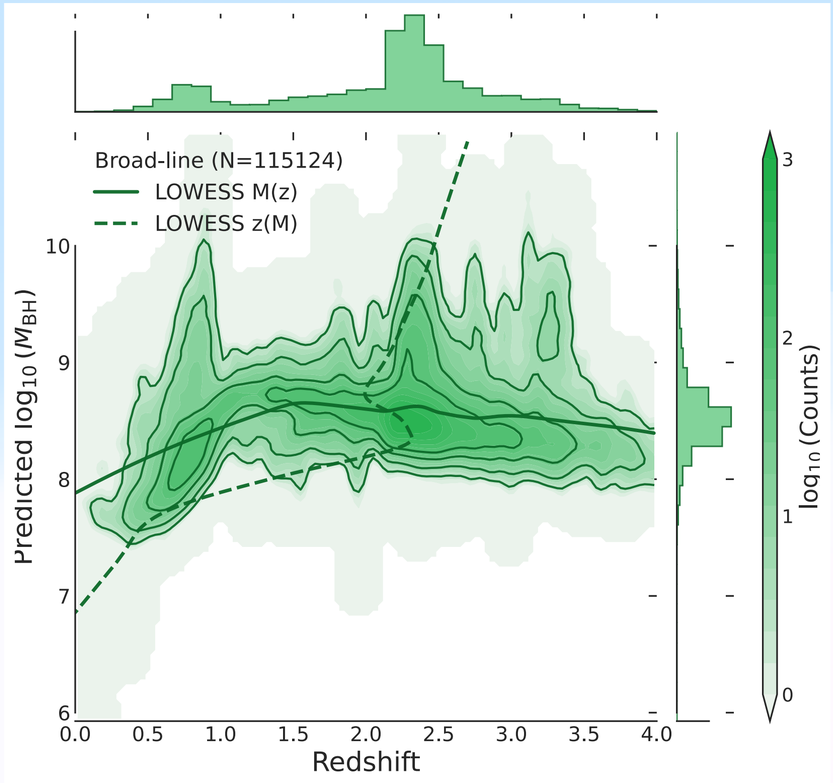
Baseline abundance fit



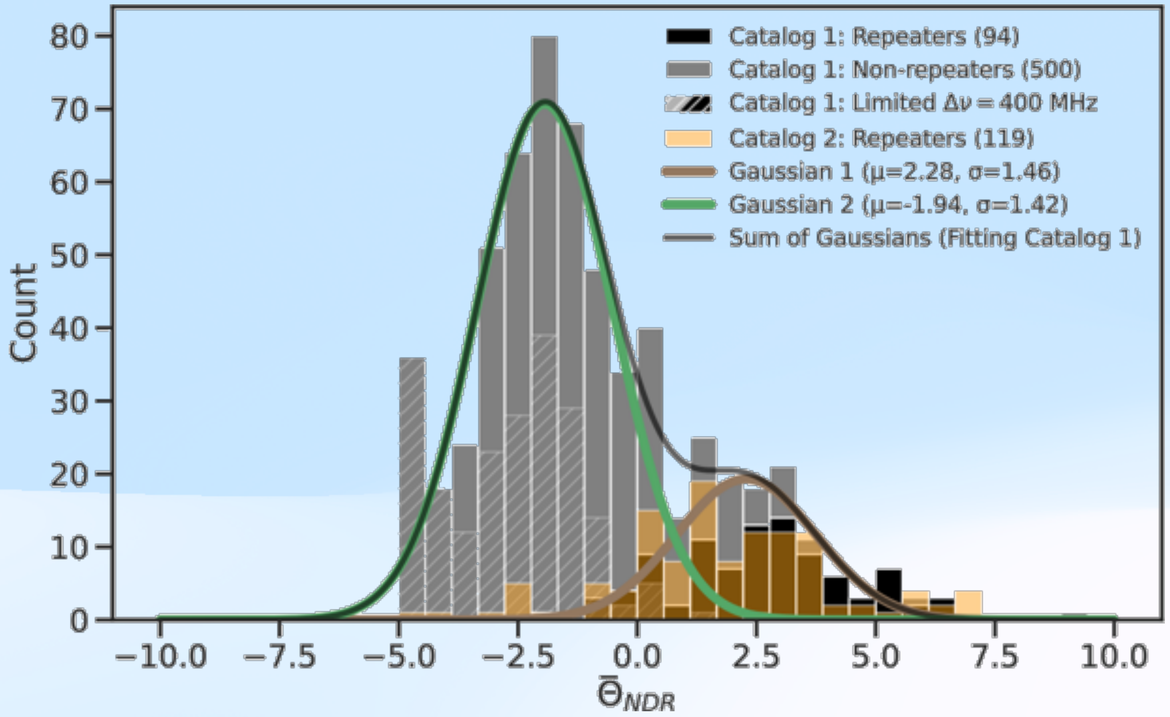
Delay / high-z diagnostic



Recent AI Works



Machine Phenomenology: A Simple Equation Classifying Fast Radio Bursts



Equation:

$$(\nu_p - \Delta\nu)^{0.3}\Delta t^{0.3} + \frac{\nu_p}{0.3\Delta\nu} + \alpha$$

Where ν_p is the peak frequency, $\Delta\nu$ is the frequency range, Δt is the duration, α is the spectral index.

Gaussian distributions of the classification by an equation derived by human-AI collaboration

Non repeating bursts are shown in grey and repeating bursts in black, and the overall distribution is modeled with two Gaussian components. The lower Gaussian contains most non repeaters, while the higher Gaussian is dominated by repeaters, giving a clear separation between the two populations. A small number of non repeaters fall within the repeater dominated region, suggesting that some may be unobserved repeaters. Repeaters from a newer Catalog, plotted in orange, align with the same high Gaussian, confirming the consistency of the classification. The leftmost bin is enhanced by the CHIME bandwidth limit at 400 MHz, which forces many wide bandwidth bursts into a single bin.

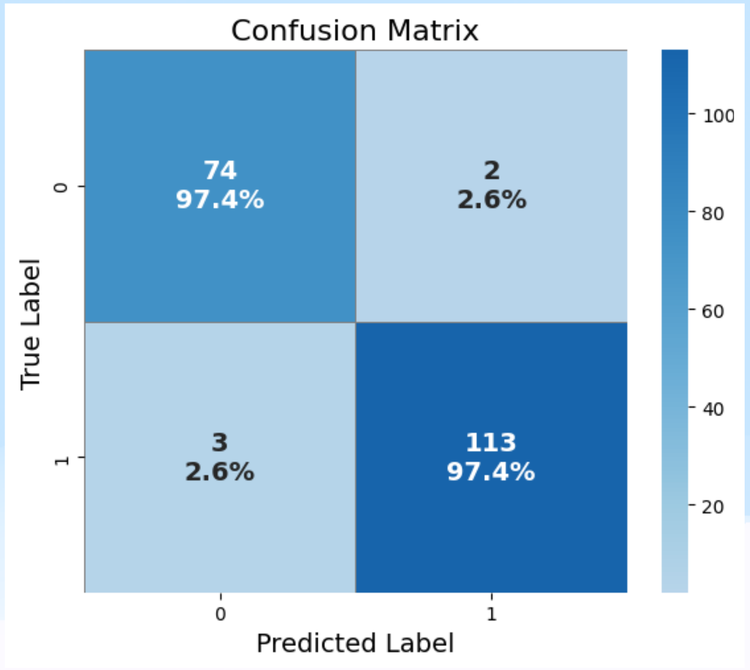
Large Language Models for Limited Noisy Data: A Gravitational Wave Identification Study

1. **Question:** Does LLMs offer real advantages over traditional neural networks for scientific data ?

2. **Prediction**: Patch tokenization suppresses local glitches, and self-attention captures long range coherence evolution. These properties align with the non-Gaussian and non-stationary data. And LLM works well with small dataset.

3. **Data**: Gravitational wave data provide a suitable test since signals show global coherent chirps, noise is complex, and only 90 observed events till O3 run.

4. **Test**

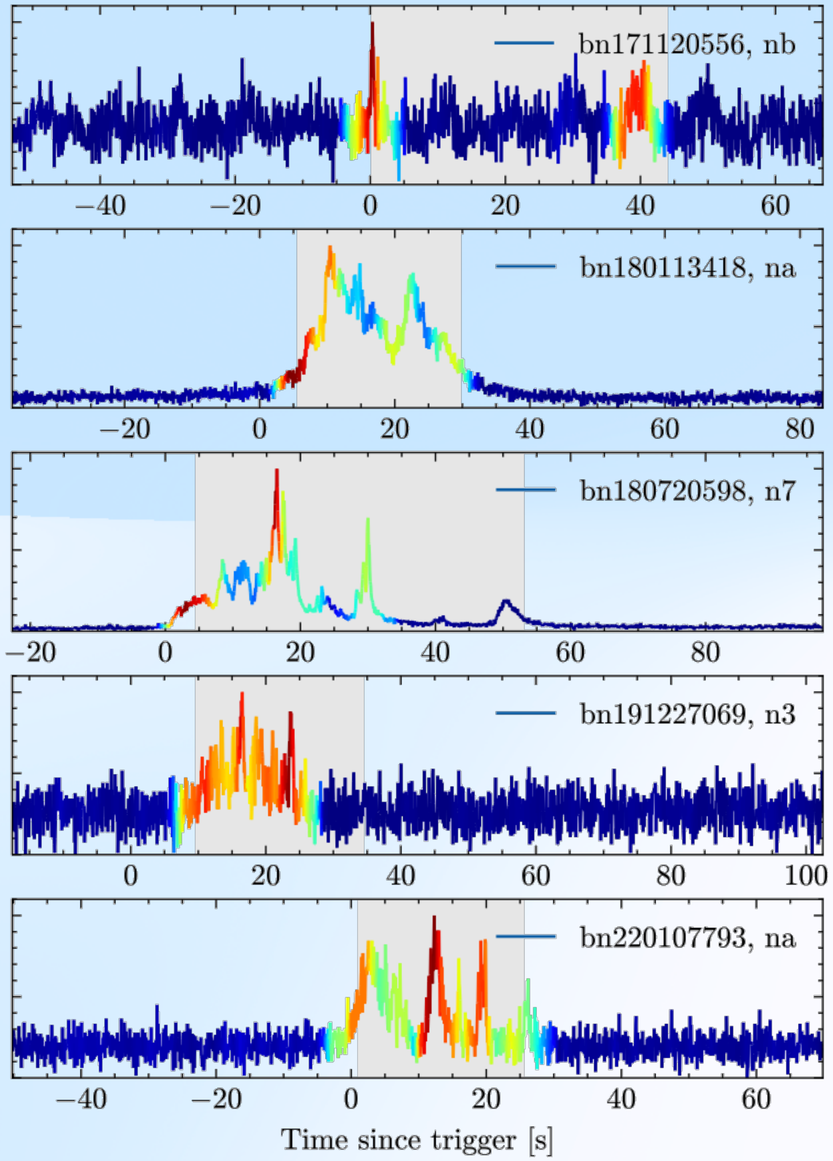


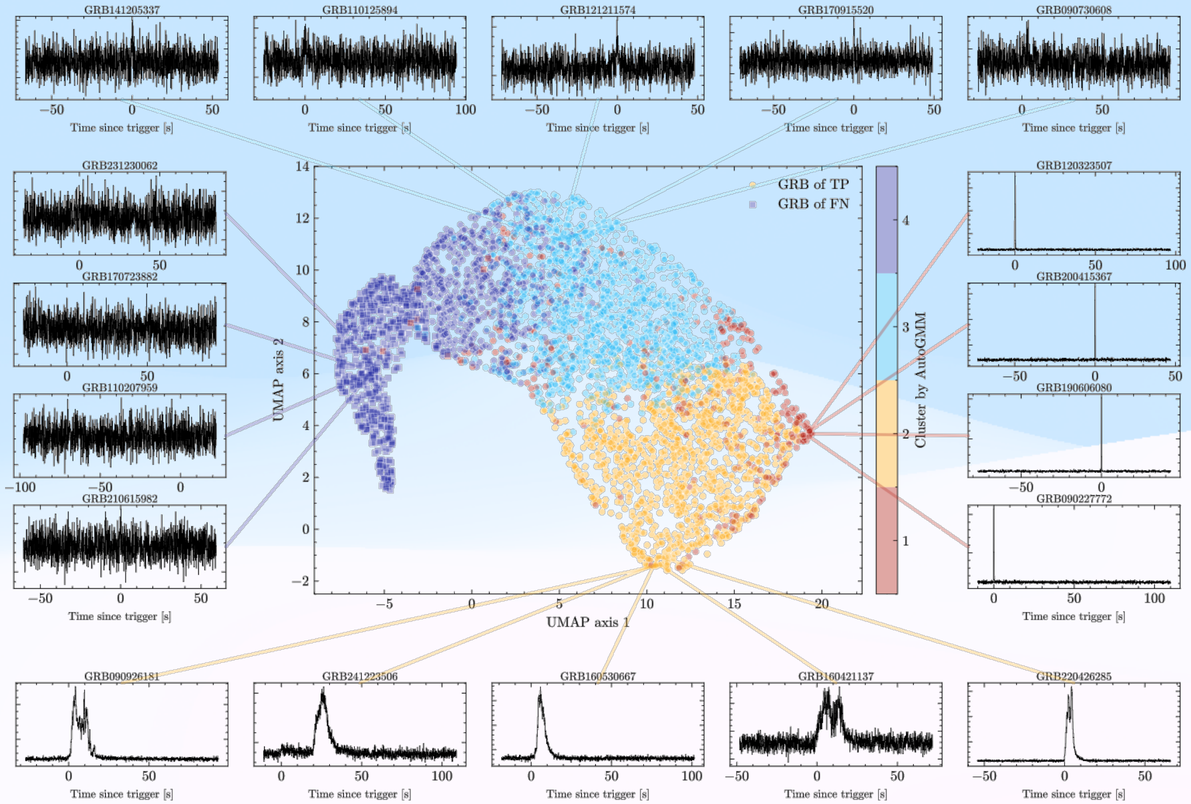
Identification performance of the finetuned LLM on LIGO observational data

After training on the observational 90 samples, an 8 billion parameter LLM reaches 97.4 percent recall for both signal and noise, converging within two epochs. Traditional networks depend on large synthetic datasets, but adding more than 500k simulated samples does not improve LLM performance.

LLMs provide an efficient and robust method for gravitational wave identification. The same advantages may apply to other astronomy with non-Gaussian and non-stationary noise, like radio pulsar and fast radio burst searches, dynamic spectra from low frequency interferometers, X-ray timing of accreting neutron stars and black holes, and high energy transient monitors.

Advancing Identification method of Gamma-Ray Bursts with Data and Feature Enhancement





Achieving more than 97% accuracy in GRB identification marks a clear improvement in high-energy transient astronomy. Even under difficult conditions, such as noisy detections and overlapping signals.

The clustering of GRBs using learned features offers a new way to classify bursts that may reflect different physical origins, for example, from neutron star mergers or from massive star collapse.

Thanks.

Science drives innovation. Innovation shapes the future.