

Multiphase Gas Structure in JWST Little Red Dots

Seok-Jun Chang

`sjchang@MPA-Garching.MPG.DE`

MPA <https://orcid.org/0000-0002-0112-5900>

Max Gronke

<https://orcid.org/0000-0003-2491-060X>

Jorryt Matthee

Institute of Science and Technology Austria <https://orcid.org/0000-0003-2871-127X>

Charlotte Mason

University of Copenhagen <https://orcid.org/0000-0002-3407-1785>

Article

Keywords:

Posted Date: October 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11486004/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations:

There is **NO** Competing Interest.

Figure 15 is not available with this version

Multiphase Gas Structure in JWST Little Red Dots

Seok-Jun Chang^{1*}, Max Gronke², Jorrit Matthee³,
Charlotte Mason^{4,5}

^{1*}Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1,
85748, Garching b. München, Germany.

²Astronomisches Rechen-Institut, Zentrum für Astronomie, Universität
Heidelberg, Mönchhofstraße 12-14, 69120, Heidelberg, Germany.

³Institute of Science and Technology Austria (ISTA), Am Campus 1,
3400, Klosterneuburg, Austria.

⁴Cosmic Dawn Center (DAWN).

⁵Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200,
Copenhagen, Denmark.

*Corresponding author(s). E-mail(s): sjchang@mpa-garching.mpg.de;

Abstract

JWST has uncovered a population of compact high-redshift galaxies, ‘little red dots’ (LRDs), exhibiting broad Balmer emission lines that often show narrow absorption features and strong Balmer breaks — spectral signatures attributed to dense gas surrounding rapidly growing black holes, but whose small-scale structure remains unknown. A striking feature of LRD spectra is that the $H\beta$ absorption is systematically deeper than $H\alpha$, yet in any homogeneous gas this ordering is physically impossible. Here we model the $H\alpha$ and $H\beta$ line profiles using radiative transfer simulations for 11 LRDs with JWST sensitive high resolution spectroscopy, and independently constrain the $n = 2$ hydrogen column density, $N_{\text{HI},n=2}$, from the Balmer-break continuum. We find that the $N_{\text{HI},n=2}$ from $H\alpha$ in a homogeneous medium cannot reproduce the observed $H\beta$ absorption in several objects, and that this $N_{\text{HI},n=2}$ typically falls two orders of magnitude below the lower bound set by the Balmer break. Both tensions are simultaneously resolved by a multiphase medium of dense clumps within diffuse hot ionized gas. This multiphase gas structure is a necessary ingredient for any viable LRD model, regardless of the nature of the central engine.

Keywords: galaxies: high-redshift; radiative transfer; galaxies: active; JWST

JWST has revealed compact sources at $z \sim 4\text{--}9$, ‘little red dots’ (LRDs), with red optical continua, broad Balmer emission lines of several thousand km s^{-1} , Balmer breaks, and Balmer absorption [1–8], the latter seen in $\sim 60\%$ of LRDs with deep spectra [9–11]. They are widely interpreted as rapidly accreting black holes in dense gas, but their nature remains debated: models invoke dense gas cocoons [12–16], super-Eddington accretion [17, 18], and dust-free black-hole atmospheres [19, 20]. In most of these scenarios, scattering of Balmer lines through dense gas plays a central role in shaping the observed spectra [21], but thus far, observed Balmer line profiles have not been confronted with radiative transfer models with scattering physics.

Interpreting these absorption features requires a model for how Balmer photons propagate through the gas. Previous LRD analyses fit the profiles with multi-component parametric functions [1, 12, 22, 23] or with Sobolev-approximation line transfer [15], both assuming simple absorption or moderate optical depth. At the column densities implied by the Balmer break ($N_{\text{HI},n=2} \gtrsim 10^{17} \text{ cm}^{-2}$), however, the line-center optical depth of the Balmer lines exceeds 10^3 : photons scatter thousands of times before escaping, and their profiles depend on the gas geometry, frequency redistribution, and branching between decay channels. We previously modeled the full scattering behavior of the Balmer lines in LRDs with the Monte Carlo radiative-transfer code RT-SCAT [21]. Here we apply it to fit the observed $\text{H}\alpha$ and $\text{H}\beta$ profiles of LRDs.

Balmer absorption and Balmer-break strength are empirically correlated across the LRD population [9], and physically linked: $\text{H}\alpha$, $\text{H}\beta$, and the Balmer continuum all arise from hydrogen in the $n = 2$ level. The $n = 2$ column density, $N_{\text{HI},n=2}$, therefore sets all three, so they form a single consistency test rather than independent diagnostics. This constraint has not been imposed in previous modeling of LRD spectra. We therefore start from the simplest assumption that the same $n = 2$ H I gas produces all three features, and ask what structure this gas must have.

Applying this test reveals a puzzle. In LRD spectra, the $\text{H}\alpha$ absorption trough is typically *weaker* than that of $\text{H}\beta$ [9, 23], which is impossible in a homogeneous gas: the $\text{H}\alpha$ line cross-section is ≈ 5.7 times larger than that of $\text{H}\beta$ [21], so at the same $N_{\text{HI},n=2}$, $\text{H}\alpha$ absorption is stronger than $\text{H}\beta$, or both lines are saturated. Moreover, homogeneous Balmer-line fits give $N_{\text{HI},n=2} \sim 10^{14} \text{ cm}^{-2}$, three orders of magnitude below the $\gtrsim 10^{17} \text{ cm}^{-2}$ required by the Balmer break (left panels of Figure 1). No single $N_{\text{HI},n=2}$ in a homogeneous medium satisfies all three observables.

We apply this test to the 11 LRDs from the broad-line sample of [9] with high signal-to-noise NIRSpec spectra of both $\text{H}\alpha$ and $\text{H}\beta$. A homogeneous single-phase gas fails it, whereas a multiphase medium of optically thick $n = 2$ H I clumps largely removes the line-break $N_{\text{HI},n=2}$ tension. This line-break comparison relies on the shared-gas assumption, but the multiphase model also reproduces $\text{H}\alpha$ and $\text{H}\beta$ at a consistent $N_{\text{HI},n=2}$, a line-fitting result independent of that assumption.

$\text{H}\alpha$ and $\text{H}\beta$ modeling.

We model the $\text{H}\alpha$ and $\text{H}\beta$ line profiles of all 11 LRDs with RT-SCAT for two gas configurations (see Methods): (i) a *homogeneous* single-phase medium with uniform $N_{\text{HI},n=2}$, and (ii) a *clumpy* medium with the same total $N_{\text{HI},n=2}$ in optically thick

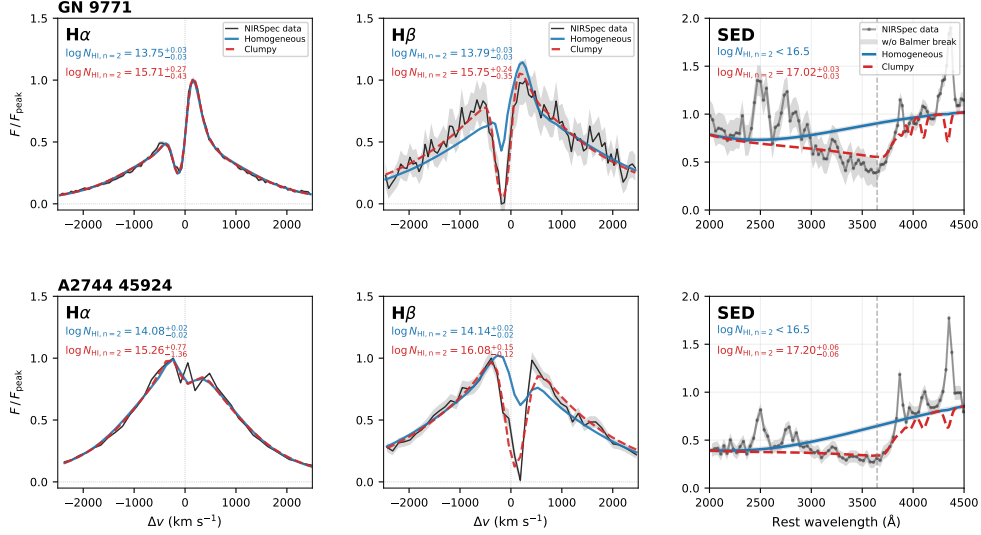


Fig. 1 $H\alpha$, $H\beta$, and Balmer-break fits for two representative LRDs. The top and bottom panels show fitting results of GN 9771 and A2744 45924, respectively. The columns show $H\alpha$ (left), $H\beta$ (middle), and the rest-frame SED around the Balmer break (right). Black curves represent the observational data, NIRSpect for $H\alpha$ & $H\beta$ and PRISM for the SED; grey shading indicates the $\pm 1\sigma$ uncertainty. The blue solid and red dashed lines are the best fits for the homogeneous and clumpy models, respectively. Fit in two steps: $H\alpha$ alone first, then $H\beta$ and the Balmer-break SED fits with $N_{\text{HI},n=2}$ prior from $H\alpha$ fit (see Methods). The median $N_{\text{HI},n=2}$ with its 16th–84th percentile range is annotated in each panel. Homogeneous SED fits that cannot produce the Balmer break (84th percentile of $N_{\text{HI},n=2}$ below $10^{16.5} \text{ cm}^{-2}$, i.e. Balmer-continuum optical depth $\tau_{\text{BB}} < 0.5$) are shown without uncertainties and labeled $\log N_{\text{HI},n=2} < 16.5$ (see Methods). For both objects, $N_{\text{HI},n=2}$ from $H\alpha$ in the homogeneous model are too low to reproduce the observed $H\beta$ or the Balmer break, while the clumpy model provides $N_{\text{HI},n=2}$ values of $H\alpha$ consistent with both. Fits for the remaining 9 objects are shown in Extended Data Figures 3 and 4.

clumps embedded in optically thin inter-clump gas, described by a covering factor f_c (the mean number of clumps along a sightline) and a per-clump column $N_{\text{HI,cl}} = 3N_{\text{HI},n=2}/4f_c$. In both models, the lines are emitted by a central source and scattered by the surrounding $n = 2$ H I gas, after subtraction of the continuum and narrow host emission (see Methods for the source model, host subtraction, and priors). We first fit $H\alpha$ alone and use its $N_{\text{HI},n=2}$ posterior as a prior for the $H\beta$ fit, testing whether the $H\alpha$ -preferred column also reproduces $H\beta$.

In the homogeneous model, the fixed $H\alpha/H\beta$ scattering cross-section ratio links the two profiles. The $H\alpha$ absorption requires $N_{\text{HI},n=2} \sim 10^{14} \text{ cm}^{-2}$, at which the predicted $H\beta$ absorption is too weak to match the observed depth (Fig. 1). In the clumpy model, two mechanisms break this coupling. First, $H\alpha$ photons scatter off the *surfaces* of optically thick clumps and escape through the inter-clump channels[24–28] without penetrating the high- $N_{\text{HI},n=2}$ interiors. Because the competing $3p \rightarrow 1s$ (Ly β) decay is immediately reabsorbed by ground-state hydrogen, $H\alpha$ behaves as a near-perfect resonance line[21]: photons remain $H\alpha$ across many scatterings, as for Ly α , and $H\alpha$ absorption stays weak even at $N_{\text{HI},n=2} \sim 10^{16} \text{ cm}^{-2}$. Second, each

$H\beta$ scattering has a $\approx 26\%$ probability of decaying through the $n = 3$ level ($BR_3 \approx 0.26$ [21]), converting the photon into $P\alpha$ and $H\alpha$. After ~ 4 scatterings, $\sim 70\%$ of $H\beta$ photons are removed, producing strong absorption even where $H\alpha$ escapes freely. The clumpy model therefore reproduces both lines at similar $N_{HI,n=2}$. Extended Data Fig. 2 illustrates how both profiles depend on f_c at fixed $N_{HI,n=2}$.

Figure 1 shows two representative LRDs, GN 9771 and A2744 45924, observed at high and medium NIRSpect resolution. The homogeneous model (blue) cannot reproduce the observed $H\beta$ absorption in either object, whereas the clumpy model (red) fits both lines at $N_{HI,n=2} \approx 10^{15.5} \text{ cm}^{-2}$. Extended Data Figures 3 and 4 show the fits for the remaining 9 objects.

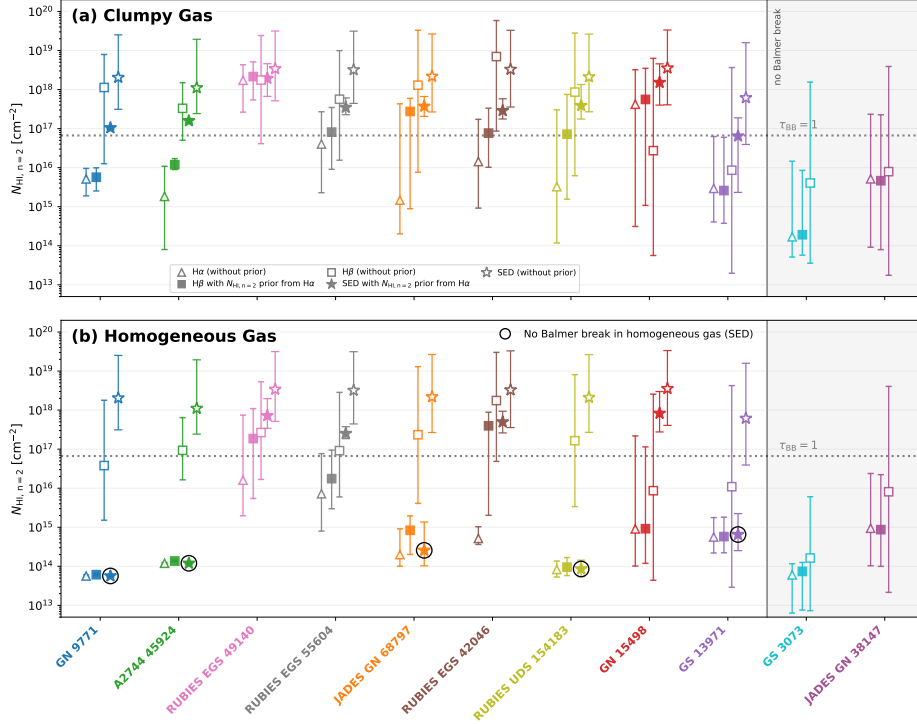


Fig. 2 $N_{HI,n=2}$ consistency across $H\alpha$, $H\beta$, and the Balmer break, for all 11 LRDs. (a) Clumpy model. (b) Homogeneous model. For each object, triangles show the $H\alpha$ fit; filled squares and filled stars show the $H\beta$ fit and the Balmer-break SED fit, with a $N_{HI,n=2}$ prior from the $H\alpha$ fit. Hollow squares and hollow stars show the same $H\beta$ and SED fits without this prior. Error bars are 68 percent credible intervals. The horizontal dotted line marks the column for Balmer-continuum optical depth $\tau_{BB} = 1$ ($N_{HI,n=2} \approx 6 \times 10^{16} \text{ cm}^{-2}$). Black circles in (b) mark homogeneous SED fits that cannot produce the Balmer break (84th percentile of $N_{HI,n=2}$ below $10^{16.5} \text{ cm}^{-2}$, i.e. $\tau_{BB} < 0.5$): 5 of the 9 break-detected objects. The shaded region marks GS 3073 and JADES GN 38147, which show no Balmer break. For the clumpy model, $N_{HI,n=2}$ of the $H\alpha$ fits and $H\beta$ and SED fits with prior cluster together for most objects, and the SED fit with the prior reaches $\tau_{BB} = 1$ for all 9 break-detected objects. For the homogeneous model, $N_{HI,n=2}$ from $H\alpha$ is too low to reach the break threshold for 6 of the 9 objects.

SED modeling with Balmer break.

Under the shared-gas assumption, the Balmer-continuum opacity provides an independent lower bound on $N_{\text{HI},n=2}$. The Balmer-continuum photoionisation cross-section at the series limit is $\sigma_{\text{bc}} \approx 1.5 \times 10^{-17} \text{ cm}^2$ [29], giving unit continuum optical depth at $N_{\text{HI},n=2} \approx 6 \times 10^{16} \text{ cm}^{-2}$. An observed Balmer break therefore requires $N_{\text{HI},n=2} \gtrsim 10^{17} \text{ cm}^{-2}$.

We fit the SED of each LRD with a three-component model: a power-law UV continuum and an optical blackbody continuum, with a Balmer break and high- n Balmer-series absorption around 3646 Å (see Methods). For 6 of the 9 break-detected objects, the homogeneous $N_{\text{HI},n=2}$ from H α lies more than two orders of magnitude below this bound, so the homogeneous model cannot produce the break (Extended Data Figs. 3 and 4; see Methods), whereas the clumpy SED fits with the H α prior reproduce it for all 9.

$N_{\text{HI},n=2}$ tension between Balmer lines and break

Figure 2 compares the $N_{\text{HI},n=2}$ estimates from H α , H β , and SED. We consider a model to fail for a given object if its $N_{\text{HI},n=2}$ from H α lies more than a factor of 100 below the Balmer-break threshold, or the SED fit with the H α prior does not reach it (see Methods and Extended Data Table 3). Of the 9 objects with a detected Balmer break, the clumpy model passes for all 9, and the homogeneous model passes for 3 (RUBIES EGS 55604, RUBIES EGS 49140, and GN 15498); for the latter two, the homogeneous H α absorption is at the line center, where its depth depends on the host-line subtraction (see Methods). GS 3073 and JADES GN 38147 show no Balmer break and are excluded from the pass/fail count (shaded region in Fig. 2). The H β fits with the $N_{\text{HI},n=2}$ prior show the same pattern. At $N_{\text{HI},n=2}$ from the H α fit, the clumpy model fits H β clearly better than the homogeneous model for GN 9771, A2744 45924, and RUBIES UDS 154183 (see Fig. 1 and Extended Data Fig. 3), and the two fit equally well for the other objects. Full $N_{\text{HI},n=2}$ posteriors for all three probes, with and without the $N_{\text{HI},n=2}$ prior from the H α fit, are shown in Extended Data Fig. 5 and Fig. 6.

$N_{\text{HI},n=2}$ and Balmer break strength relation

The Balmer-break color, which is the flux ratio at 5500 Å and 3600 Å, $f_{\nu,5500}/f_{\nu,3600}$, is a direct measure of the break strength [9, 19, 30]. Figure 3 shows the break color as a function of $N_{\text{HI},n=2}$ from the H β fit with the prior from the H α fit, which is blind to the break, in both clumpy and homogeneous models. We use H β because H α alone is only weakly sensitive to $N_{\text{HI},n=2}$ in the clumpy model (Methods); the H α -only values are shown for comparison. The clumpy model shows a positive correlation with the break strength (Pearson $r = +0.64$, Spearman $\rho = +0.62$), as expected if the clumpy $N_{\text{HI},n=2}$ faithfully estimates the true column, whereas the homogeneous model shows no such trend ($r = 0.01$, $\rho = -0.05$). This suggests that only the clumpy model links $N_{\text{HI},n=2}$ from the Balmer lines to the observed Balmer break.

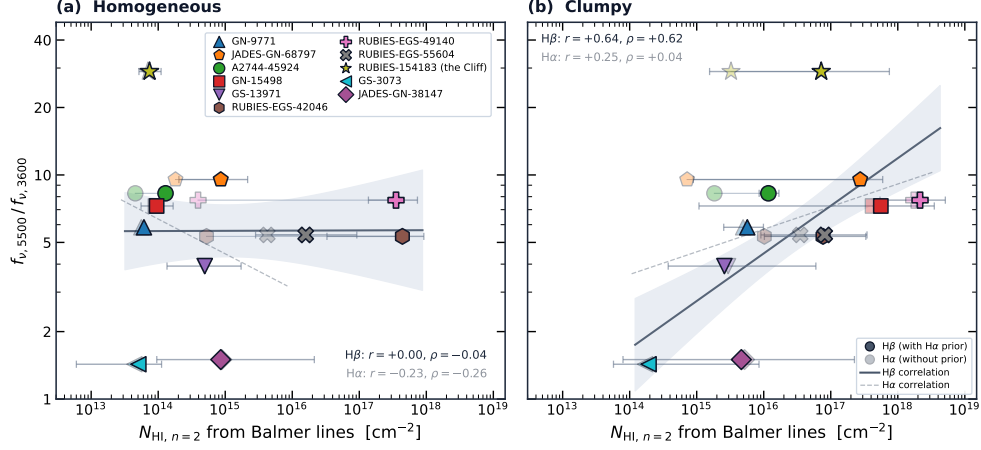


Fig. 3 Balmer-break color versus $n = 2$ neutral hydrogen column density for all 11 LRDs. The Balmer-break color $f_{\nu,5500}/f_{\nu,3600}$ measures the Balmer-break strength, with larger values indicating a stronger break [9]. The left and right panels show it as a function of $N_{\text{HI}, n=2}$ from the homogeneous and clumpy fits. Each object has a unique marker shape and the same color as in Fig. 2. Opaque markers show $N_{\text{HI}, n=2}$ from the $\text{H}\beta$ fit with the prior from the $\text{H}\alpha$ fit, and faded markers show the $\text{H}\alpha$ fit without the prior. The solid (dashed) lines represent the correlation between $\log(f_{\nu,5500}/f_{\nu,3600})$ versus $\log N_{\text{HI}, n=2}$ from $\text{H}\beta$ ($\text{H}\alpha$) fit. Pearson (r , in log-log space) and Spearman (ρ) coefficients are given in each panel. For $\text{H}\beta$, the homogeneous model shows no correlation ($r = +0.01$, $\rho = -0.05$), while the clumpy model shows a positive one ($r = +0.64$, $\rho = +0.62$, $p \approx 0.04$). For $\text{H}\alpha$ alone, neither model shows a significant correlation (clumpy $r = +0.25$, $\rho = +0.04$; homogeneous $r = -0.06$, $\rho = -0.11$); the clumpy slope is positive but not significant.

Relation to current LRD models.

The tension in $N_{\text{HI}, n=2}$ between the Balmer lines and the Balmer break — reinforced by the correlation with Balmer-break strength — shows that a multiphase gas structure (i.e., the clumpy model) is required to explain both features through the same $n = 2$ populated H I gas. The multiphase gas constraint established here is agnostic to the nature of the central engine and applies across the range of currently proposed LRD models. In the dense gas cocoon scenario [12, 15], the constraint implies that the cocoon must be fragmented on small scales rather than smooth. In the gas-enshrouded black hole and dense accretion disc models [18, 19], the gas atmosphere must similarly be inhomogeneous to satisfy the joint Balmer-line and Balmer-break constraints. Black holes can accrete well above the Eddington rate, through radiatively inefficient slim discs or from dense surrounding gas [31–34]. In quasi-stars, such accretion proceeds inside a massive envelope [35], a picture since applied to LRDs [16, 18, 36]. These models treat the envelope as smooth, although its locally super-Eddington outer layers are expected to be unstable and inhomogeneous [36]. If the break instead arises from a supermassive star or an evolved stellar population [30, 37, 38], it decouples from the gas, though the Balmer absorption still requires a multiphase $n = 2$ H I medium. Our result does not discriminate between these scenarios but identifies a necessary ingredient, strong small-scale inhomogeneity in the $n = 2$ H I gas, that any viable model must accommodate. The broad line wings do not enter this test. Whether they arise

from Thomson scattering in an inner ionized region [21] or the broad-line region itself [17], in our model they only set the incident line profile; the Balmer absorption and break form in the $n = 2$ H I gas outside the emission region, which must be multiphase in either case. The H α and H β profiles alone also require multiphase structure, independently of whether the break arises from the same gas.

Multiphase gas structure and formation

The clumpy geometry required by our modeling can be interpreted physically as a multiphase medium: dense, cool, $n = 2$ enhanced H I clumps with covering factor $f_c \sim 5\text{--}60$ and individual column densities $N_{\text{HI,cl}} \sim 10^{15}\text{--}10^{17} \text{ cm}^{-2}$, coexisting with a diffuse, optically thin interclump medium (see Extended Data Fig. 1). The total column density $N_{\text{HI},n=2}$ and f_c are constrained by the Balmer-line modeling, but the large-scale morphology, whether a discrete cloud ensemble, porous shell, or fragmented disc, is not determined by the present data alone. A clumpy neutral medium has also been considered a viable explanation for the Ly α profile of the lensed LRD QSO1 [39].

Such multiphase structure is common around AGNs. Broad emission lines are attributed to small, dense ($n_{\text{H}} \sim 10^{10} \text{ cm}^{-3}$) clouds with a volume filling fraction of $\sim 10^{-6}$, confined by a hotter medium or magnetic fields [40, 41]; smooth line profiles require $\gtrsim 10^6\text{--}10^8$ such emitters [42, 43], and internal turbulence of $\sim 10^3 \text{ km s}^{-1}$ is allowed in the line-emitting gas [44]. Outflowing absorbers, from broad and mini broad absorption line regions near the accretion disc [45–47] to kpc-scale FeLoBAL and associated absorbers [48–50], likewise resolve into multiple narrow components. Our fits are consistent with this picture: for N_{cl} clumps of radius r_{cl} within a region of size R , $f_c \sim N_{\text{cl}}(r_{\text{cl}}/R)^2$, so $f_c \sim 5\text{--}60$ implies $N_{\text{cl}} \sim 10^5\text{--}10^6$ for $r_{\text{cl}}/R \sim 10^{-2}$, and more for smaller clumps¹, comparable to BLR and BAL cloud numbers. However, BLR clouds are photoionized and emit the broad lines, whereas the clumps in our model are a neutral scattering medium with a substantial $n = 2$ population.

Such clumps can form and survive in this environment. Turbulent multiphase gas fragments into many small structures rather than remaining monolithic, and mixing at the interfaces with the hot gas induces rapid cooling, so the cool phase grows rather than being destroyed [52, 53]. Since only a small fraction of the hydrogen is in $n = 2$, the total clump columns implied by $N_{\text{HI,cl}} \sim 10^{15}\text{--}10^{17} \text{ cm}^{-2}$ exceed both the fragmentation scale of rapidly cooling gas ($\sim 10^{17} \text{ cm}^{-2}$ [54]) and the minimum column for survival in a hot wind ($\sim 10^{18} \text{ cm}^{-2}$ [55]). The clump population implied by our fits is thus a natural outcome of cooling in a turbulent medium, although the data do not constrain its formation history.

Cool clumps alone do not explain their large $n = 2$ population. Multiple scatterings of Ly α in optically thick clumps can pump $n = 2$ near their surfaces [56–58], and collisional excitation may contribute in dense gas; comparing the observed Ly α and H α fluxes can constrain their relative importance [39], and a detailed calculation will be presented in a companion paper.

One circumnuclear setting for such gas is a thermally unstable disc wind, in which cool clumps condense from a hot outflow. The densities invoked in previous LRD

¹The radiative transfer depends on f_c and $N_{\text{HI,cl}}$ but not on clump size for small volume filling factors [25, 51], so r_{cl} and N_{cl} are not individually constrained.

studies, $n_{\text{H}} \sim 10^9\text{--}10^{11} \text{ cm}^{-3}$ [13, 19], overlap with those of broad-line region clouds ($n_{\text{H}} \sim 10^8\text{--}10^{10} \text{ cm}^{-3}$), which are commonly associated with clumpy disc winds [45, 59]. We therefore suggest, cautiously, that the multiphase $n = 2$ H I clumps in LRDs may represent an early BLR-like or proto-BLR medium formed during rapid black hole growth. In this picture, what distinguishes LRDs from ordinary broad-line AGNs is not the presence of clumps but their ionization state. Shielded by the dense inner gas of a super-Eddington flow, the clumps stay neutral. As the accretion rate decreases and the disc thins [18, 31, 60], reduced shielding could induce ionization of these clumps, eventually giving rise to a standard BLR, which would also explain why LRD-like spectra are rare at lower redshift [61, 62]. The present data establish the presence of strongly inhomogeneous $n = 2$ H I gas, but do not uniquely determine its disc-wind origin or identify the clumps as proto-BLR clouds.

Origin of the red optical continuum

The multiphase gas constraint does not by itself determine the origin of the red optical continuum in LRDs. Proposed origins include a super-Eddington photosphere [19, 63], supermassive stars [37, 38], a transmitted AGN SED [13, 64], and nebular continuum emission [15, 65]. The nebular case is directly testable, because there the continuum and the hydrogen lines both arise from recombination, whereas the other models place the continuum in a region separate from the line emission. Since H α and the nebular continuum both scale with the emission measure, a pure-recombination spectrum from low-density gas (where collisions are negligible) that is optically thin in the Balmer lines has an intrinsic H α equivalent width (EW) of $\sim 5,000\text{--}10,000 \text{ \AA}$ regardless of geometry (Extended Data Fig. 7), up to 10–100 times larger than the observed LRD values. Reproducing the observed continuum from nebular gas alone therefore requires either substantial H α suppression, relaxing one of these two assumptions, or a separate, non-nebular continuum component.

Relaxing the optically thin assumption, bound-free absorption from $n = 3$ has been suggested to suppress H α [15]. Resonance scattering lengthens the $n = 3$ column an H α photon samples relative to the continuum, but our RT calculations give a path-length enhancement of only $\sim 2\text{--}30$ (Extended Data Fig. 8)², two to four orders of magnitude short of what is required (see Methods). The $n = 3$ column needed instead predicts near-saturated Paschen-line absorption, which is not seen in LRDs with Paschen-band observations and clear Balmer absorption [66–68].

In denser gas, collisional de-excitation of $n = 3$ offers a second route. Because a trapped H α photon revisits $n = 3$ repeatedly before escaping, this channel becomes accessible at $n_e \sim 10^9\text{--}10^{11} \text{ cm}^{-3}$, within the range invoked in previous LRD studies [13, 19], although the observed H β survival caps how optically thick the emission region can be [21] (see Methods). Both channels predict the same Paschen-line signature, so we cannot determine which, if either, resolves the puzzle. Either way, the EW argument requires H α suppression or a continuum source distinct from the line-emitting gas, regardless of whether the $n = 2$ medium is clumpy.

²Dust along this path could likewise extinguish H α relative to the continuum, though direct evidence for dust in this gas is lacking (see Methods).

Caveats, robustness, and observational predictions

The line-break consistency test assumes that the same $n = 2$ H I gas produces both the Balmer-line absorption and the Balmer break. If the break instead formed in a separate continuum-emitting region, as in photospheric or stellar models of the red continuum, the line-break comparison alone would not require clumpiness. That region would, however, need an $N_{\text{HI},n=2}$ that happens to track the Balmer-line absorption, which the suggestive correlation between the $\text{H}\beta$ -fit $N_{\text{HI},n=2}$ and break strength disfavors but cannot exclude. The line-profile fits alone do not rely on this assumption: at $N_{\text{HI},n=2}$ from the $\text{H}\alpha$ fit, the homogeneous model cannot reproduce the observed $\text{H}\beta$ absorption in several objects, whereas the clumpy model reproduces both lines.

Host-galaxy subtraction, the unresolved $\text{H}\beta$ core at $\mathcal{R} \sim 1000$, and an $N_{\text{HI},n=2}$ - f_c degeneracy when no clear absorption is detected affect the exact parameters of individual objects (see Methods), but do not explain why the homogeneous model fails to fit both lines while the clumpy model succeeds.

Our radiative transfer treats $N_{\text{HI},n=2}$ as a free parameter and neglects collisional processes. In practice, the $n = 2$ and $n = 3$ populations are set jointly by $\text{Ly}\alpha$ pumping, collisional excitation and de-excitation, and line trapping, all of which depend on gas density; coupling these to the radiative transfer would translate the fitted $N_{\text{HI},n=2}$ and f_c into physical densities and clump sizes.

The multiphase model makes several testable predictions. Higher Balmer lines share the branching loss of $\text{H}\beta$: excited np states with $n \geq 4$ can decay through intermediate levels $2 < n' < n$, emitting Paschen, Brackett, or higher-series photons instead of the original Balmer photon. $\text{H}\gamma$ and $\text{H}\delta$ should therefore show absorption comparable to $\text{H}\beta$, not intermediate between $\text{H}\alpha$ and $\text{H}\beta$: the transition is a step at $\text{H}\alpha$, the only near-perfect resonance line, rather than a gradual hierarchy. Saturated $\text{H}\gamma$ absorption has already been invoked for the broad $\text{H}\gamma$ profile of GN 9771 [22], and deep high-resolution spectroscopy now resolves $\text{H}\alpha$ - $\text{H}\delta$ absorption in individual LRDs [69]; a test across the sample awaits higher signal-to-noise, as these lines are intrinsically weaker. For the LRDs whose $\text{H}\beta$ cores remain unresolved or have insufficient signal-to-noise, high-resolution NIRSpec follow-up can test whether they show the same $\text{H}\alpha$ - $\text{H}\beta$ discrepancy already unambiguous in GN 9771 ($\mathcal{R} = 2700$). Finally, Paschen-band spectroscopy would test both $\text{H}\alpha$ suppression channels discussed above at once: near-saturated Paschen-line absorption is expected if either operates, whereas its absence would favor a non-nebular contribution to the red optical continuum.

Acknowledgements. This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with programs #1181, 1216, 4233, 5664, 8204 and 9433. Support for program #5664 was provided by NASA through a grant from the Space Telescope Science Institute. The authors acknowledge the RUBIES team, led by co-PIs A. de Graaff and G. Brammer, and the #9433 team, led by co-PIs R. Maiolino and F. D'Eugenio, for developing their observing programs with a zero-exclusive-access period. SJC acknowledges support from the ERC synergy grant 101166930 - RECAP. JM acknowledges funding by

the European Union (ERC, AGENTS, 101076224). MG thanks the European Union for support through ERC-2024-STG 101165038 (ReMMU). CAM acknowledges support by the European Union ERC grant RISES (101163035), Carlsberg Foundation (CF22-1322), and VILLUM FONDEN (37459). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNRF140. The authors used an AI assistant (Claude, Anthropic) for building the fitting pipeline and language editing. All analyses, interpretations, and conclusions were made and verified by the authors, who take full responsibility for the content.

Declarations

- **Funding:** S.-J.C. discloses support for this work from the ERC synergy grant 101166930 (RECAP). J.M. discloses support from the European Union (ERC, AGENTS, 101076224). M.G. discloses support from the European Union (ERC-2024-STG 101165038, ReMMU). C.M. discloses support from the European Union ERC grant RISES (101163035), the Carlsberg Foundation (CF22-1322), and VILLUM FONDEN (37459); the Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNRF140.
- **Competing interests:** The authors declare no competing interests.
- **Ethics approval and consent to participate:** Not applicable.
- **Consent for publication:** Not applicable.
- **Data availability:** The JWST NIRSpec spectra used in this work are described in [9]. This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope, obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with programs #1181, 1216, 4233, 5664, 8204 and 9433 (see Extended Data Table 1). The radiative-transfer model grids and MCMC posterior chains are available from the corresponding author upon reasonable request and will be deposited in a public repository upon acceptance (DOI to be confirmed upon acceptance).
- **Materials availability:** Not applicable.
- **Code availability:** The RT-SCAT radiative-transfer code is described in [21, 28]. Posterior sampling uses the affine-invariant ensemble sampler EMCEE [70]. The MCMC fitting pipeline built around these tools is available from the corresponding author upon reasonable request.
- **Author contributions:** S.-J.C. led the project, performed the radiative-transfer simulations, analyzed the simulated data, and modeled the Balmer lines and SED. M.G. contributed to building the fitting pipeline and to the scientific interpretation of the fitting results and their connection to the multiphase nature of the gas. J.M. provided the observed spectra for the sample of 11 LRDs, contributed to the interpretation of the data and to connecting the modeling results to LRD observations, and offered insight into how the results compare with other proposed LRD models. C.M. contributed to the observational interpretation of early-universe and LRD observations and to checking the models against recent observational results.

References

- [1] Matthee, J. *et al.* Little Red Dots: An Abundant Population of Faint Active Galactic Nuclei at $z \sim 5$ Revealed by the EIGER and FRESCO JWST Surveys. *Astrophysical Journal* **963**, 129 (2024).
- [2] Greene, J. E. *et al.* UNCOVER Spectroscopy Confirms the Surprising Ubiquity of Active Galactic Nuclei in Red Sources at $z > 5$. *Astrophysical Journal* **964**, 39 (2024).
- [3] Kocevski, D. D. *et al.* Hidden Little Monsters: Spectroscopic Identification of Low-mass, Broad-line AGNs at $z > 5$ with CEERS. *Astrophysical Journal Letters* **954**, L4 (2023).
- [4] Kokorev, V. *et al.* A census of photometrically selected little red dots at $4 < z < 9$ in JWST blank fields. *Astrophysical Journal* **968**, 38 (2024).
- [5] Setton, D. J. *et al.* Little Red Dots at an Inflection Point: Ubiquitous “V-Shaped” Turnover Consistently Occurs at the Balmer Limit. *arXiv e-prints* arXiv:2411.03424 (2024).
- [6] Juodžbalis, I. *et al.* JADES - the Rosetta stone of JWST-discovered AGN: deciphering the intriguing nature of early AGN. *Monthly Notices of the Royal Astronomical Society* **535**, 853–873 (2024).
- [7] Labbe, I. *et al.* An unambiguous AGN and a Balmer break in an Ultraluminous Little Red Dot at $z=4.47$ from Ultradeep UNCOVER and All the Little Things Spectroscopy. *arXiv e-prints* arXiv:2412.04557 (2024).
- [8] Hviding, R. E. *et al.* RUBIES: A spectroscopic census of little red dots. All point sources with v-shaped continua have broad lines. *A&A* **702**, A57 (2025).
- [9] Matthee, J. *et al.* The Engine and its Flows: Little Red Dot spectra are shaped by the column densities of their gas envelopes. *arXiv e-prints* arXiv:2603.17667 (2026).
- [10] Juodžbalis, I. *et al.* An (in)complete NIRSpect census of Balmer absorption in Type 1 AGN – radiation-driven outflows in little red dots, quasars and variable stars. *arXiv e-prints* arXiv:2607.19057 (2026).
- [11] Davis, K. *et al.* OCEANS of Absorption: High-resolution NIRSpect Spectroscopy Reveals Diverse Balmer-line Absorption in Little Red Dots. *arXiv e-prints* arXiv:2606.00258 (2026).
- [12] Rusakov, V. *et al.* JWST’s little red dots: an emerging population of young, low-mass AGN cocooned in dense ionized gas. *arXiv e-prints* arXiv:2503.16595 (2025).

- [13] Inayoshi, K. & Maiolino, R. Extremely Dense Gas around Little Red Dots and High-redshift Active Galactic Nuclei: A Nonstellar Origin of the Balmer Break and Absorption Features. *Astrophysical Journal Letters* **980**, L27 (2025).
- [14] Kido, D., Ioka, K., Hotokezaka, K., Inayoshi, K. & Irwin, C. M. Black hole envelopes in Little Red Dots. *MNRAS* **544**, 3407 (2025).
- [15] Sneppen, A. *et al.* Inside the cocoon: a comprehensive explanation of the spectra of Little Red Dots. *arXiv e-prints* arXiv:2601.18864 (2026).
- [16] Naidu, R. P. *et al.* Little Red Dots as Intermediate Mass, Super-Eddington Engines: Insights from Type IIⁿ Supernovae and the 1837-1856 Great Eruption of η Carinae. *arXiv e-prints* arXiv:2606.30711 (2026).
- [17] Madau, P. & Maiolino, R. Little red dots as obscured little blue dots: A super-Eddington unification model. *A&A* (2026).
- [18] Chon, S., Hirano, S., Ishiyama, T., Chang, S.-J. & Springel, V. Overmassive black holes and little red dots naturally form in simulations. *Nature* **657**, 621 (2026).
- [19] Naidu, R. P. *et al.* A “Black Hole Star” Reveals the Remarkable Gas-Enshrouded Hearts of the Little Red Dots. *arXiv e-prints* arXiv:2503.16596 (2025).
- [20] de Graaff, A. *et al.* A remarkable ruby: Absorption in dense gas, rather than evolved stars, drives the extreme Balmer break of a little red dot at $z = 3.5$. *Astronomy & Astrophysics* **701**, A168 (2025).
- [21] Chang, S.-J., Gronke, M., Matthee, J. & Mason, C. Impact of resonance, Raman, and Thomson scattering on hydrogen line formation in Little Red Dots. *Monthly Notices of the Royal Astronomical Society* **545**, 1–21 (2026).
- [22] Torralba, A. *et al.* The warm outer layer of a little red dot as the source of [Fe II] and collisional Balmer lines with scattering wings. *A&A* **707**, A75 (2026).
- [23] D’Eugenio, F. *et al.* BlackTHUNDER strikes twice: Balmer-line absorption in an overmassive Little Red Dot at $z = 7.04$. *MNRAS* **547**, 1 (2026).
- [24] Neufeld, D. A. The transfer of resonance-line radiation in clumpy media. *Astrophysical Journal Letters* **370**, L85 (1991).
- [25] Chang, S.-J., Yang, Y., Seon, K.-I., Zabludoff, A. & Lee, H.-W. Radiative Transfer in Ly α Nebulae. I. Modeling a Continuous or Clumpy Spherical Halo with a Central Source. *Astrophysical Journal* **945**, 100 (2023).
- [26] Gronke, M., Dijkstra, M., McCourt, M. & Oh, S. P. From Mirrors to Windows: Lyman-alpha Scattering around Galaxies in the Presence of Multiphase Gas. *Astrophysical Journal Letters* **833**, L26 (2016).

- [27] Gronke, M., Dijkstra, M., McCourt, M. & Oh, S. P. Resonance line transfer in multiphase clumpy media. *Astronomy and Astrophysics* **607**, A71 (2017).
- [28] Chang, S.-J. & Gronke, M. Probing cold gas with MgII and Ly α radiative transfer. *Monthly Notices of the Royal Astronomical Society* **532**, 3526–3555 (2024).
- [29] Karzas, W. J. & Latter, R. Electron Radiative Transitions in a Coulomb Field. *Astrophysical Journal Supplement Series* **6**, 167 (1961).
- [30] Wang, B. *et al.* RUBIES: Evolved Stellar Populations with Extended Formation Histories at $z \sim 7-8$ in Candidate Massive Galaxies Identified with JWST/NIRSpec. *ApJL* **969**, L13 (2024).
- [31] Abramowicz, M. A., Czerny, B., Lasota, J. P. & Szuszkiewicz, E. Slim Accretion Disks. *Astrophysical Journal* **332**, 646–658 (1988).
- [32] Ohsuga, K., Mori, M., Nakamoto, T. & Mineshige, S. Supercritical Accretion Flows around Black Holes: Two-dimensional, Radiation Pressure-dominated Disks with Photon Trapping. *ApJ* **628**, 368–381 (2005).
- [33] Jiang, Y.-F., Stone, J. M. & Davis, S. W. A Global Three-dimensional Radiation Magneto-hydrodynamic Simulation of Super-Eddington Accretion Disks. *ApJ* **796**, 106 (2014).
- [34] Inayoshi, K., Haiman, Z. & Ostriker, J. P. Hyper-Eddington accretion flows on to massive black holes. *MNRAS* **459**, 3738 (2016).
- [35] Begelman, M. C., Rossi, E. M. & Armitage, P. J. Quasi-stars: accreting black holes inside massive envelopes. *MNRAS* **387**, 1649 (2008).
- [36] Begelman, M. C. & Dexter, J. Little Red Dots as Late-stage Quasi-stars. *ApJ* **996**, 48 (2025).
- [37] Chisholm, J. *et al.* Little Red Dots as Globular Clusters in Formation. *arXiv e-prints* arXiv:2602.15935 (2026).
- [38] Chon, S., Hirano, S., Chen, K.-J. & Springel, V. Formation of Heavy Seed Black Holes and Little Red Dots-like Compact Clusters in Metal-enriched Star-forming Regions. *arXiv e-prints* arXiv:2609.00437 (2026).
- [39] Tang, M. *et al.* SPURS: Evidence for Clumpy Neutral Envelopes and Ionized IGM Surrounding Little Red Dots in Abell 2744 from Ultra-Deep Rest-UV Spectroscopy. *arXiv e-prints* arXiv:2604.03563 (2026).
- [40] Rees, M. J. Magnetic confinement of broad-line clouds in active galactic nuclei. *MNRAS* **228**, 47P (1987).

- [41] Netzer, H. in *Active Galactic Nuclei: Basic Physics and Main Components* (eds Alloin, D., Johnson, R. & Lira, P.) *Physics of Active Galactic Nuclei at all Scales*, Vol. 693 of *Lecture Notes in Physics* 1 (Springer, 2006).
- [42] Arav, N., Barlow, T. A., Laor, A. & Blandford, R. D. Keck high-resolution spectroscopy of Mrk 335: constraints on the number of emitting clouds in the broad-line region. *MNRAS* **288**, 1015 (1997).
- [43] Dietrich, M., Wagner, S. J., Courvoisier, T. J. L., Bock, H. & North, P. Structure of the broad-line region of 3C 273. *A&A* **351**, 31 (1999).
- [44] Bottorff, M., Ferland, G., Baldwin, J. & Korista, K. Observational Constraints on the Internal Velocity Field of Quasar Emission-Line Clouds. *ApJ* **542**, 644 (2000).
- [45] Elvis, M. A Structure for Quasars. *ApJ* **545**, 63–76 (2000).
- [46] Hamann, F. *et al.* A high-velocity narrow absorption line outflow in the quasar J212329.46–005052.9. *MNRAS* **410**, 1957 (2011).
- [47] Hamann, F. *et al.* Extreme-velocity quasar outflows and the role of X-ray shielding. *MNRAS* **435**, 133 (2013).
- [48] Faucher-Giguère, C.-A., Quataert, E. & Murray, N. A physical model of FeLoBALs: implications for quasar feedback. *MNRAS* **420**, 1347 (2012).
- [49] D’Odorico, V., Cristiani, S., Romano, D., Granato, G. L. & Danese, L. Chemical abundances in QSO host galaxies and environments from narrow absorption line systems. *MNRAS* **351**, 976 (2004).
- [50] Finn, C. W. *et al.* A compact, metal-rich, kpc-scale outflow in FBQS J0209–0438: detailed diagnostics from HST/COS extreme UV observations. *MNRAS* **440**, 3317 (2014).
- [51] Hansen, M. & Oh, S. P. Lyman α radiative transfer in a multiphase medium. *MNRAS* **367**, 979 (2006).
- [52] Gronke, M., Oh, S. P., Ji, S. & Norman, C. Survival and mass growth of cold gas in a turbulent, multiphase medium. *Monthly Notices of the Royal Astronomical Society* **511**, 859–876 (2022).
- [53] Gronke, M. & Schneider, E. E. Simulations of multi-phase gas in and around galaxies. *Living Reviews in Computational Astrophysics* **12**, 2 (2026).
- [54] McCourt, M., Oh, S. P., O’Leary, R. & Madigan, A.-M. A characteristic scale for cold gas. *Monthly Notices of the Royal Astronomical Society* **473**, 5407–5431 (2018).

- [55] Gronke, M. & Oh, S. P. The growth and entrainment of cold gas in a hot wind. *MNRAS* **480**, L111 (2018).
- [56] Dijkstra, M., Sethi, S. & Loeb, A. 3-cm Fine Structure Masers: A Unique Signature of Supermassive Black Hole Formation via Direct Collapse in the Early Universe. *Astrophysical Journal* **820**, 10 (2016).
- [57] Dijkstra, M., Gronke, M. & Sobral, D. Ly α Signatures from Direct Collapse Black Holes. *Astrophysical Journal* **823**, 74 (2016).
- [58] Seon, K.-i. & Kim, C.-G. Ly α Radiative Transfer: Monte Carlo Simulation of the Wouthuysen-Field Effect. *Astrophysical Journal Supplement Series* **250**, 9 (2020).
- [59] Murray, N., Chiang, J., Grossman, S. A. & Voit, G. M. Accretion Disk Winds from Active Galactic Nuclei. *ApJ* **451**, 498 (1995).
- [60] Inayoshi, K. Little Red Dots as the Very First Activity of Black Hole Growth. *Astrophysical Journal Letters* **988**, L22 (2025).
- [61] Kocevski, D. D. *et al.* The Rise of Faint, Red Active Galactic Nuclei at $z \lesssim 4$: A Sample of Little Red Dots in the JWST Extragalactic Legacy Fields. *ApJ* **986**, 126 (2025).
- [62] Lin, R. *et al.* Discovery of Local Analogs to JWST’s Little Red Dots. *ApJL* **980**, L34 (2025).
- [63] Liu, H., Jiang, Y.-F., Quataert, E., Greene, J. E. & Ma, Y. The Balmer Break and Optical Continuum of Little Red Dots from Super-Eddington Accretion. *ApJ* **994**, 113 (2025).
- [64] Killi, M. *et al.* Deciphering the JWST spectrum of a ‘little red dot’ at $z \sim 4.53$: An obscured AGN and its star-forming host. *A&A* **691**, A52 (2024).
- [65] Ando, M., Harikane, Y., Katz, H., Inayoshi, K. & Tanaka, T. S. The UV Side of Little Red Dots: Red, Compact, and Iron-Enhanced Rest-UV Emission with a Strong Downturn around Ly α . *arXiv e-prints* arXiv:2606.03522 (2026).
- [66] Brazzini, M. *et al.* Ruling out dominant electron scattering in Little Red Dots’ Rosetta Stone using multiple hydrogen lines. *MNRAS* **544**, L167 (2025).
- [67] Kokorev, V. *et al.* The Deepest GLIMPSE of a Dense Gas Cocoon Enshrouding a Little Red Dot. *ApJ* **1004**, 153 (2026).
- [68] Kiyota, T. *et al.* ATLAS. III. Dust Around Little Red Dots: Hydrogen Line Ratios beyond Dust-free Non-Case B Models. *arXiv e-prints* arXiv:2608.10832 (2026).
- [69] Wang, B. *et al.* Balmer Absorption Series and Broad Metal Lines in Two

- Luminous Little Red Dots. *arXiv e-prints* (2026).
- [70] Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: The MCMC hammer. *PASP* **125**, 306–312 (2013).
 - [71] Cooper, R. A. *et al.* The Physical Conditions of Low-Mass Galaxies at $z = 3.5 - 7.0$ from JWST Spectroscopy: The Behaviour of Spectral Line Ratios with Burstiness. *arXiv e-prints* arXiv:2609.01205 (2026).
 - [72] Cameron, A. J. *et al.* JADES: Probing interstellar medium conditions at $z \sim 5.5-9.5$ with ultra-deep JWST/NIRSpec spectroscopy. *A&A* **677**, A115 (2023).
 - [73] Laor, A. Evidence for Line Broadening by Electron Scattering in the Broad-Line Region of NGC 4395. *ApJ* **643**, 112–119 (2006).
 - [74] Draine, B. T. *Physics of the Interstellar and Intergalactic Medium* Princeton Series in Astrophysics (Princeton University Press, Princeton, NJ, 2011).
 - [75] Osterbrock, D. E. & Ferland, G. J. *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* 2nd edn (University Science Books, Sausalito, CA, 2006).
 - [76] Chang, S.-J. *et al.* Escape of Resonantly Scattered $\text{Ly}\beta$ and $\text{H}\alpha$ from Hot and Optically Thick Media. *JKAS* **51**, 5 (2018).
 - [77] Neufeld, D. A. The Transfer of Resonance-Line Radiation in Static Astrophysical Media. *Astrophysical Journal* **350**, 216 (1990).
 - [78] Dijkstra, M. $\text{Ly}\alpha$ Emitting Galaxies as a Probe of Reionisation. *Publications of the Astronomical Society of Australia* **31**, e040 (2014).

Methods

Sample selection and spectroscopy

Our sample of 11 LRDs is drawn from the broad $\text{H}\alpha$ -selected sample of [9], which compiled high-quality JWST NIRSpec spectra at $z \approx 3.5-7$ from multiple JWST survey programs. Both NIRSpec PRISM/CLEAR and higher-resolution grating (G235M or G395H) spectra are available for all objects; we use the PRISM spectra for the SED continuum fitting and the grating spectra for the Balmer-line profile modeling. We select the 11 objects with sufficient signal-to-noise in both $\text{H}\alpha$ and $\text{H}\beta$ for joint line-profile fitting; the sample is listed in Extended Data Table 1. Full details of the sample selection and data reduction are described in [9].

Balmer-line radiative-transfer modelling

We model the $\text{H}\alpha$ and $\text{H}\beta$ line profiles using RT-SCAT, a Monte Carlo radiative transfer code for line formation by scattering processes previously applied to $\text{Ly}\alpha$ [25] and

Mg II emission[28]; the RT algorithm and clumpy-medium treatment are described in detail therein. The Balmer-line implementation is described in [21], which is our previous work on Balmer-line radiative transfer for resonance and Thomson scattering³. The scattering geometry consists of a central isotropic source surrounded by two concentric spherical regions (Fig. 15 of [21]; illustrated for both models in Extended Data Fig. 1): an inner, fully ionised H II region ($r < 0.01 R$) responsible for Thomson scattering, characterised by electron temperature $T_e = 10^4$ K and Thomson optical depth τ_{Th} ; and an outer neutral H I region ($0.01 R < r < R$) with a significantly populated $n = 2$ level, responsible for Balmer-line resonance scattering. Photons emitted by the central source first traverse the Thomson-scattering H II layer, which broadens the line wings symmetrically, before entering the outer H I layer where resonance scattering shapes the line core and absorption features. We consider two configurations for the outer H I sphere: a *homogeneous* sphere with uniform $n = 2$ H I column density $N_{\text{HI},n=2}$, and a *clumpy* sphere characterised by the same total column density $N_{\text{HI},n=2}$ distributed in spherical optically-thick clumps of fixed radius $r_{\text{cl}} = 10^{-3} R$, embedded in an optically thin inter-clump medium; the individual clump column density is $N_{\text{HI,cl}} = \frac{3}{4} N_{\text{HI},n=2}/f_c$ (the factor $3/4$ accounts for the mean chord length through a sphere); the covering factor f_c is the mean number of clumps along a line of sight from the central source to the outer edge of the sphere[28].

The H α and H β resonance cross-sections at line center differ by a factor of ≈ 5.7 (the ratio of oscillator strengths times wavelength; [21] equations 3 and 4), so that H α photons experience substantially higher opacity than H β at the same $N_{\text{HI},n=2}$. In the clumpy case, H α photons scatter predominantly off the near surface of each clump and escape without penetrating to high column density (*surface scattering*), whereas H β photons penetrate more deeply and are subject to a branching decay at the $4p$ level ($\text{BR}_{n=3} \approx 0.26$; [21] Table 3 and Fig. 4) that removes them from the H β line before re-emission. This mechanism allows the clumpy model to simultaneously reproduce both lines with a single $N_{\text{HI},n=2}$, including cases where H β absorption is deeper than H α . The asymmetry is atomic in origin: the $3p \rightarrow 1s$ (Ly β) decay is immediately reabsorbed by ground-state hydrogen, so a scattered H α photon is ultimately re-emitted as H α , whereas the $4p$ level also decays to $3s$ and $3d$, which cascade through $2p$ and emit Pa α and H α instead of H β . With $\text{BR}_3 \approx 0.26$ per scattering, $1 - (1 - \text{BR}_3)^4 \approx 70\%$ of H β photons are converted after four scatterings.

The free parameters of the homogeneous model are $N_{\text{HI},n=2}$, the radial velocity v_{exp} , the intrinsic emission-line width σ_{emit} , the random velocity dispersion σ_{ran} , and the Thomson optical depth τ_{Th} . The clumpy model retains $N_{\text{HI},n=2}$ (total column density) and adds the covering factor f_c ; the per-clump column density $N_{\text{HI,cl}} = \frac{3}{4} N_{\text{HI},n=2}/f_c$ is a derived quantity. In the clumpy model, σ_{ran} is the peculiar velocity dispersion of individual clumps about the constant outflow at v_{exp} , rather than a Doppler width of the bulk medium as in the homogeneous case; the random motion within clumps is the thermal motion, as their temperature is at $T = 10^4$ K. Parameter ranges are given in Extended Data Table 2. The intrinsic emission width is restricted to $\sigma_{\text{emit}} \in [100, 300]$ km s⁻¹; larger values would attribute the broad line wings to

³S.-J. Chang, the lead author of this work, is a developer of RT-SCAT for resonance scattering and its application to Balmer lines

intrinsic emission rather than Thomson scattering, removing the physical motivation for the two-component geometry of [21]. The total grid comprises 40,130 homogeneous and 181,212 clumpy model spectra. The $N_{\text{HI},n=2}$ lower bound of 10^{12} cm^{-2} represents optically thin gas, where $\tau_{0,\text{H}\alpha} \approx 0.3$ and $\tau_{0,\text{H}\beta} \approx 0.06$ (equations 3 and 4 of [21]). Thus, the gas is effectively transparent in both lines, and the fit is insensitive to $N_{\text{HI},n=2}$ below this value.

The continuum is first removed from each spectrum using the continuum model of [9], derived from a spectral decomposition of the full JWST NIRSpec grating spectrum. The narrow host-galaxy emission in H α and H β is then subtracted. For each object, the [O III] $\lambda 5007$ profile is used as a host template for both Balmer lines: $F_{\text{host,H}\beta}(v) = F_{[\text{O III}]}(v)/R_{\text{H}\beta}$ and $F_{\text{host,H}\alpha}(v) = F_{[\text{O III}]}(v)/R_{\text{H}\alpha}$, where $R_{\text{H}\beta} \equiv F_{[\text{O III}]} / F_{\text{H}\beta,\text{host}}$ and $R_{\text{H}\alpha} \equiv F_{[\text{O III}]} / F_{\text{H}\alpha,\text{host}}$ are the host [O III]-to-Balmer line ratios. Both are free parameters. For $R_{\text{H}\beta}$ we adopt a Gaussian prior on $\log_{10} R_{\text{H}\beta}$ centered on the median [O III]/H β ratio of high- z galaxies at the redshift of each object [71], with a width of 0.15 dex [72], truncated so that the subtracted host does not drive the observed H β and H α line cores below zero beyond their observational uncertainties. The host Balmer decrement, $R_{\text{H}\beta}/R_{\text{H}\alpha}$, is allowed to vary between 2.6 and 3.3 (Case B with mild departures). A third free parameter, Δv , is a single velocity offset shared by both lines between the [O III]-derived host/narrow-line-region zero point and the true center of the LRD’s own broad-line/clumpy-gas structure; it has a Gaussian prior centered at zero with width set in quadrature from that object’s own [O III] line width and the instrumental resolution.

We infer posterior distributions for all free parameters using Markov chain Monte Carlo (MCMC) sampling via the affine-invariant ensemble sampler EMCEE[70]. At each proposed parameter vector, the model line profiles are evaluated at continuous parameter values via a principal-component-analysis emulator trained on the pre-computed RT-SCAT grid, allowing the sampler to explore the full parameter space between grid points. Each χ^2 is evaluated over the velocity windows $|v| \leq 3000 \text{ km s}^{-1}$ (H α) and $|v| \leq 2000 \text{ km s}^{-1}$ (H β) using an optimal linear normalisation $A = (\sum_i O_i M_i \sigma_i^{-2}) / (\sum_i M_i^2 \sigma_i^{-2})$ sensitive only to line profile shape. The per-pixel uncertainty is $\sigma_i^2 = \sigma_{\text{obs},i}^2 + \sigma_{\text{jig},i}^2$, where σ_{obs} is the pipeline error and $\sigma_{\text{jig},i} = \sigma_{\text{MAD}} \sqrt{F_{\text{ref}} |f_i|}$ is a conservative allowance for residual spectral structure not captured by it. Here σ_{MAD} is the scatter of each line’s wings ($1500 < |v| < 5000 \text{ km s}^{-1}$) about a symmetric exponential profile, commonly used to describe the broad wings of LRD Balmer lines [12, 73], and F_{ref} is the median wing flux at $3000 < |v| < 5000 \text{ km s}^{-1}$. Model spectra are convolved to the instrumental resolution prior to comparison ($\mathcal{R} = 2700$ for G235H/G395H, $\mathcal{R} = 1000$ for G235M/G395M).

We fit H α and H β in two sequential stages, taking H α as the baseline because it has the highest signal-to-noise ratio of the Balmer lines in our sample and its profile most directly encodes the scattering geometry.

- *Stage 1:* we fit the H α profile alone with flat priors on all radiative transfer parameters (Extended Data Table 2). Because H β photons that decay through $n = 3$ are partly re-emitted as H α , the model H α profile includes this converted component, weighted by the intrinsic H α /H β flux ratio of the source, $F_{\text{H}\alpha}/F_{\text{H}\beta}$, a free parameter with a Gaussian prior centered on the observed $F_{\text{H}\alpha}/F_{\text{H}\beta}$ of each object. We

use a kernel density estimate of the resulting posterior on $\log N_{\text{HI},n=2}$ as the prior for the next stage.

- *Stage 2:* we fit the $\text{H}\beta$ profile alone with this prior imposed on $\log N_{\text{HI},n=2}$. All other radiative transfer parameters retain flat priors and are refit independently of Stage 1. This tests whether the $N_{\text{HI},n=2}$ from the $\text{H}\alpha$ fit in Stage 1 is also consistent with the observed $\text{H}\beta$ profile.

For each stage, we run 128 walkers for 50,000 steps in Stage 1 and 64 walkers for 20,000 steps in Stage 2 per object per model, initialized at random positions within the prior ranges; the first 5,000 steps are discarded as burn-in, and marginalized 68% credible intervals are derived from the remaining samples. Full spectral fits for all objects are shown in Extended Data Fig. 3 and 4. Marginalized $\log N_{\text{HI},n=2}$ posteriors are shown in Extended Data Fig. 5 and 6. Fitted parameters for all 11 objects are listed in Extended Data Tables 4 and 5.

Three effects limit the constraints for individual objects. Host subtraction can affect the line cores, which is why the host Balmer contributions are treated as free parameters with observational priors rather than fixed values. At $\mathcal{R} \sim 1000$, the $\text{H}\beta$ core is unresolved, which broadens the $N_{\text{HI},n=2}$ posteriors. In the clumpy model, $\text{H}\alpha$ is only weakly sensitive to $N_{\text{HI},n=2}$ because photons scatter off the clump surfaces, so $N_{\text{HI},n=2}$ is partly degenerate with f_c and its credible intervals are typically wider than in the homogeneous model (~ 2 versus ~ 1 dex; Fig. 2 and Extended Data Tables 4 and 5), particularly when no clear absorption feature is detected. $\text{H}\beta$ remains sensitive to $N_{\text{HI},n=2}$; the Stage 2 fit therefore combines the constraints of both lines and, for most objects, narrows the $N_{\text{HI},n=2}$ interval, which is why we use it in Fig. 3.

Balmer-break SED modelling

We model the rest-frame SED of each LRD with three spectral components. The first is a UV power-law continuum, $f_\lambda \propto \lambda^{\alpha_{\text{UV}}}$. The second is an optical continuum component, modeled as a blackbody with effective temperature T_{BB} . In practice, bound-free (Balmer and Paschen recombination continua) and free-free emission from the ionized gas can also contribute to the optical continuum alongside the blackbody component; T_{BB} should therefore be understood as an effective fitting parameter rather than a physical temperature. The third component is a Balmer break applied to the optical continuum blueward of the Balmer limit $\lambda_{\text{BE}} = 3646 \text{ \AA}$, with bound-free optical depth scaling as $\tau_{\text{bc}}(\lambda) = \tau_0 (\lambda/\lambda_{\text{BE}})^3$, where $\tau_0 = N_{\text{HI},n=2} \sigma_{\text{bc}}$ and $\sigma_{\text{bc}} = 1.5 \times 10^{-17} \text{ cm}^2$ [29]. The UV power-law component is left unattenuated, while only the optical continuum component is subject to this attenuation. This differential treatment is a modeling choice motivated by the fit itself: the optical continuum must arise from gas that shares the same $n = 2$ H I absorbing column responsible for the Balmer break and Balmer-line absorption, whereas the UV power-law component does not — consistent with it originating from a source not co-located with the $n = 2$ H I clumps (e.g., unobscured host-galaxy light). Broad emission lines are excluded from the fit by adopting continuum windows that avoid all major rest-frame emission features (see below).

The fit proceeds in three sequential stages.

- *Stage 1:* the UV power-law amplitude and slope α_{UV} are determined from emission-line-free UV windows (rest-frame 1300–1420, 2000–2300, and 3400–3620 Å)[9].
- *Stage 2:* the Stage 1 UV SED is subtracted from the observed spectrum and the residual optical continuum is fitted with the blackbody component in line-free optical windows (3655–4000, 4230–4280, 5100–5800, 6000–6270, and 7200–8000 Å)[9]; each window receives equal weight regardless of pixel count to prevent the fit being dominated by wide wavelength regions.
- *Stage 3:* with the UV and optical SEDs frozen from Stages 1 and 2, we attenuate the combined continuum with the Balmer-continuum optical depth ($\propto \nu^{-4}$) described above blueward of 3646 Å, plus a sum of Voigt-profile line optical depth from the higher Balmer series (upper level $n = 3$ –100, each weighted by its oscillator strength and total spontaneous decay rate) redward of 3646 Å, sharing a common Doppler width v_{ran} . Both terms are evaluated over the window 3000–3800 Å, which brackets the Balmer limit on both sides and is sensitive to both the depth of the break and the level of the unattenuated optical continuum just redward of 3646 Å. We mask 3700–3760 Å, where [O II] $\lambda 3727$ blends with the high- n Balmer lines at PRISM resolution. The two free parameters are $N_{\text{HI},n=2}$, which sets both optical depth terms, and $v_{\text{ran}} \in [10, 500] \text{ km s}^{-1}$. Rather than a flat prior, $\log N_{\text{HI},n=2}$ is given the same prior constructed from the H α fit posterior, while v_{ran} has a flat prior. The flux uncertainties include a residual-scatter term of the same form as for the Balmer lines, measured from the residuals about the Stage 1/2 continuum. We sample the posterior with EMCEE (32 walkers, 12,000 steps, the first 2,000 discarded), with the Stage 1/2 continuum held fixed.

We quantify the relative prominence of the optical continuum component through the optical-excess parameter

$$f_{\text{exc}} = \frac{\langle f_{\text{obs}} \rangle_{\text{opt}} - \langle f_{\text{UV}} \rangle_{\text{opt}}}{\langle f_{\text{UV}} \rangle_{\text{opt}}}, \quad (1)$$

where $\langle f_{\text{obs}} \rangle_{\text{opt}}$ and $\langle f_{\text{UV}} \rangle_{\text{opt}}$ are the mean observed flux density and mean UV power-law model flux density averaged over the Stage-2 optical windows, respectively. We claim a Balmer-break detection only when two criteria are satisfied simultaneously: (i) $f_{\text{exc}} \geq 1$, ensuring that a genuine red continuum component exists against which a break can be measured; and (ii) $\Delta\text{BIC} > 0$, computed with the pipeline flux uncertainties, meaning the Stage-3 model is preferred over the Stage-2 model without attenuation. Of the 11 LRDs in our sample, 9 satisfy both criteria. The remaining two — GS 3073 ($f_{\text{exc}} = 0.36$) and JADES GN 38147 ($f_{\text{exc}} = 0.63$) — have UV-dominated SEDs in which the optical continuum is too faint relative to the inner-disc UV to anchor a reliable Balmer-break measurement. These two non-detections do not imply an absence of $n = 2$ H I gas for SEDs.

Because both the Balmer break and the higher Balmer series lines are included in the Stage 3 fit, we constrain $N_{\text{HI},n=2}$ and v_{ran} directly, rather than providing only a saturation-limited lower bound on $N_{\text{HI},n=2}$ as a continuum-only fit would. For a small number of objects, the fit remains prior-dominated (see flags in Extended Data Table 3). Best-fit spectra for all objects are shown in the third column of Extended

Data Fig. 3 and 4 (for homogeneous fits that cannot produce the break, see Fig. 1 caption); full posterior distributions are shown in Extended Data Fig. 5 and 6, with numerical values listed in Extended Data Table 3. Full posterior distributions from both stages for GN 9771 and A2744 45924 are shown in Extended Data Figs. 9 and 10.

Line-break consistency test

For each of the 9 LRDs with a Balmer-break detection, we test whether the $N_{\text{HI},n=2}$ from the $\text{H}\alpha$ fit can also produce the Balmer break. A model fails if either (i) the median $\text{H}\alpha$ $N_{\text{HI},n=2}$ lies more than a factor of 100 below the Balmer-break threshold, $N_{\text{HI},n=2} \approx 6 \times 10^{16} \text{ cm}^{-2}$ ($\tau_{\text{BB}} = 1$), i.e. $\tau_{\text{BB}} < 0.01$, for which the break depth would be below one percent; or (ii) the 84th percentile of the SED fit with the $\text{H}\alpha$ prior lies below $10^{16.5} \text{ cm}^{-2}$ ($\tau_{\text{BB}} = 0.5$), i.e. the fit cannot produce the break. Failures under (i) are robust to the posterior width for 3 of the 6 homogeneous objects that fail it (GN 9771, A2744 45924, and RUBIES UDS 154183): the 84th percentile of the $\text{H}\alpha$ $N_{\text{HI},n=2}$ also lies more than two orders of magnitude below the threshold. Of the 9 detected objects, the clumpy model passes for all 9, and the homogeneous model passes for 3 (RUBIES EGS 55604, RUBIES EGS 49140, and GN 15498). For RUBIES EGS 49140 and GN 15498, the homogeneous $\text{H}\alpha$ trough sits at the line center ($v_{\text{exp}} \approx 10\text{--}20 \text{ km s}^{-1}$), where its depth depends on the host subtraction, and the $\text{H}\alpha$ $N_{\text{HI},n=2}$ posteriors are broad (68% intervals spanning 2.6 and 3.3 dex). Results are summarized in Fig. 2 and Extended Data Table 3.

At $N_{\text{HI},n=2}$ from the $\text{H}\alpha$ fit (Stage 2), the clumpy model reproduces the $\text{H}\beta$ absorption feature, whereas the homogeneous model predicts absorption that is too weak, for GN 9771, A2744 45924, and RUBIES UDS 154183 (Fig. 1 and Extended Data Fig. 3); the reduced χ^2 of the clumpy fit is lower by factors of 1.4–3.2 (Extended Data Tables 4 and 5). For the other objects, the two fits are comparable (χ^2 ratios within 1.0–1.04). For RUBIES EGS 42046, however, the homogeneous $\text{H}\beta$ fit shows $\log N_{\text{HI},n=2} \approx 17.6$, nearly three orders of magnitude above its $\text{H}\alpha$ value, so the tension there is between the two lines rather than with the break (Fig. 2b); the clumpy fits agree within their uncertainties.

Marginalized posterior distributions for all probes ($\text{H}\alpha$, $\text{H}\beta$, and SED, each with and without the $\text{H}\alpha$ -derived prior) are shown in Extended Data Fig. 5 and 6 for clumpy and homogeneous models, respectively.

Nebular continuum and $\text{H}\alpha$ intrinsic equivalent width

Extended Data Fig. 7 quantifies the equivalent width (EW) of $\text{H}\alpha$ that a purely nebular optical continuum would imply, before any optically thick $\text{H}\alpha$ radiative-transfer effects. This sets the baseline against which the observed LRD $\text{H}\alpha$ EWs, and any proposed $\text{H}\alpha$ suppression, must be judged.

We compute the hydrogen nebular continuum as the sum of free-free emission and the free-bound (radiative recombination) continua to $n = 2$ and $n = 3$,

$$j_{\lambda}^{\text{neb}} = j_{\lambda}^{\text{ff}} + j_{\lambda}^{\text{fb}}(n=2) + j_{\lambda}^{\text{fb}}(n=3), \quad (2)$$

evaluated over rest-frame 1500–8000 Å. Free-bound recombination to $n \geq 4$ contributes $\lesssim 3$ percent of the continuum in this range and is neglected. Recombination coefficients are from the tabulation of [74]. Two photon $2s \rightarrow 1s$ emission is omitted: at the densities invoked for LRD line-emitting gas in previous modeling ($n_e \gtrsim 10^5 \text{ cm}^{-3}$, well above the $2s$ critical density), collisional $2s \leftrightarrow 2p$ mixing redirects these decays into $\text{Ly}\alpha$; adding a two-photon term computed at $n_e = 10^9 \text{ cm}^{-3}$ changes $\text{EW}_{\text{H}\alpha}$ by < 0.01 percent at every temperature tested. The gas is taken to be photoionized and in Case B recombination [75], with bound-level populations taken in the low-density, collisionless limit apart from the $2s \leftrightarrow 2p$ mixing noted above, and $\text{H}\alpha$ is treated as optically thin in this baseline emissivity calculation.

The intrinsic rest-frame EW is

$$\text{EW}_{\text{H}\alpha}^{\text{int}} = \frac{L_{\text{H}\alpha}}{L_{\lambda, \text{neb}}(\lambda_{\text{H}\alpha})}, \quad (3)$$

with $L_{\text{H}\alpha} = n_e n_p \alpha_{\text{H}\alpha}^{\text{eff}}(T) h\nu_{\text{H}\alpha} V$. Both the line and the continuum luminosity scale with the emission measure $\int n_e n_p dV$, so the electron density and emitting volume cancel in the ratio and $\text{EW}_{\text{H}\alpha}^{\text{int}}$ depends only on temperature, through the atomic coefficients. The adopted fiducial density ($n_e = n_p = 10^9 \text{ cm}^{-3}$) is therefore immaterial to the result; it is motivated by previous LRD modeling rather than by the present Balmer radiative-transfer fits. We evaluate $\text{EW}_{\text{H}\alpha}^{\text{int}}$ on a grid of $T = 10^3$ – 10^5 K.

Extended Data Fig. 7 is built from this calculation. The left panel shows the nebular continuum at $T = 5000$, 10^4 , and 2×10^4 K, each normalized to unity at rest-frame 5500 Å, with $\text{H}\alpha$ and $\text{H}\beta$ drawn on top as Gaussian emission lines (velocity dispersion 500 km s^{-1}) whose integrated fluxes are set self-consistently from $\text{EW}_{\text{H}\alpha}^{\text{int}}$ at that temperature and the Case B ratio $\text{H}\alpha/\text{H}\beta = 2.86$. The right panel plots $\text{EW}_{\text{H}\alpha}^{\text{int}}$ against T ; the shaded band marks the representative observed LRD range, $\text{EW}_{\text{H}\alpha} \sim 100$ – 2000 Å [9].

Over the plotted grid $\text{EW}_{\text{H}\alpha}^{\text{int}}$ is of order 10^3 – 10^5 Å, and equals $\approx 1.4 \times 10^4$, 9.1×10^3 , and 5.7×10^3 Å at $T = 5000$, 10^4 , and 2×10^4 K — roughly 10–100 times larger than the observed LRD $\text{H}\alpha$ EWs. A nebular-dominated optical continuum therefore over-predicts the $\text{H}\alpha$ equivalent width, and reconciling the two requires either substantial $\text{H}\alpha$ suppression or a separate, non-nebular optical-continuum component. Because both the $\text{H}\alpha$ line and the nebular continuum scale with the emission measure, this equivalent width is set by atomic physics alone — independent of density, geometry, and the number of emitting zones — and it stays $\gtrsim 10^3$ Å across the whole temperature range, so a multi-temperature mixture cannot evade it either. Any model that reproduces the observed, lower $\text{H}\alpha$ equivalent widths from nebular gas — including radiative-transfer treatments of dense LRD envelopes — must therefore invoke one of the two routes above: additional suppression of $\text{H}\alpha$, or an optical continuum that is not purely recombination emission. Although the red color of the nebular continuum is set by fixed atomic thresholds (the Balmer edge at 3646 Å and the Paschen edge at 8204 Å), its slope and break strength still depend on temperature; the continuum shape is not temperature-independent.

The predicted EW also assumes that $\text{H}\alpha$ is not collisionally suppressed. Collisional de-excitation of the $n = 3$ level becomes significant only near its critical density

$n_{\text{crit}} = A/q$ [74], where A is the Einstein A coefficient and q the collisional de-excitation rate coefficient. Averaged over the $n = 3$ sublevels the two radiative channels are $A_{32} \approx 4.4 \times 10^7 \text{ s}^{-1}$ ($\text{H}\alpha$) and $A_{31} \approx 5.6 \times 10^7 \text{ s}^{-1}$ ($\text{Ly}\beta$), so with $q \approx 1.6 \times 10^{-7} \text{ cm}^3 \text{ s}^{-1}$ at $T = 10^4 \text{ K}$ the critical density is $n_{\text{crit}} \approx 6 \times 10^{14} \text{ cm}^{-3}$ if both photons escape. In gas optically thick to the Lyman series, the $3 \rightarrow 1$ decay does not remove the excitation: the $\text{Ly}\beta$ photon is reabsorbed locally, since the Lyman lines scatter off the ground state and are more optically thick still, and the atom is returned to $n = 3$, so that a photon alternates in identity between $\text{Ly}\beta$ and $\text{H}\alpha$ until it escapes as $\text{H}\alpha$ [76]. The radiative loss rate from $n = 3$ is then set by A_{32} alone and $n_{\text{crit}} \approx 3 \times 10^{14} \text{ cm}^{-3}$, varying by less than a factor of two over $T = 3000\text{--}2 \times 10^4 \text{ K}$. Including the collisional $3 \rightarrow n > 3$ channels, which depopulate $n = 3$ without necessarily destroying the $\text{H}\alpha$ photon, lowers this to $\sim 4 \times 10^{13} \text{ cm}^{-3}$. All of these lie well above the densities invoked for LRD line-emitting gas in previous studies, $n_{\text{H}} \sim 10^9\text{--}10^{11} \text{ cm}^{-3}$ [13, 19], and above the $10^6\text{--}10^{10} \text{ cm}^{-3}$ adopted where the optical continuum is attributed to nebular emission [15].

The collisional process becomes more significant if the emission region is locally optically thick for $\text{H}\alpha$. Considering scattering, the critical number density is given by

$$n_{\text{crit}} = \frac{A_{32}}{qN_{\text{scat}}}, \quad (4)$$

where N_{scat} is the number of scatterings in the local emission region. For a photon injected at line centre $N_{\text{scat}} \sim \tau_0$, so that $\tau_0 \sim 10^3\text{--}10^4$ at $N_{\text{HI},n=2} \sim 10^{15.5}\text{--}10^{16.5} \text{ cm}^{-2}$ places n_{crit} at $\sim 10^{10}\text{--}10^{11} \text{ cm}^{-3}$, within the density range of the photoionization models above. Equivalently, a trapped $\text{H}\alpha$ photon feeds the $n = 3$ population rather than escaping [74]. The $\text{Ly}\beta$ channel amplifies this further: each $\text{Ly}\beta$ scattering is an absorption $1s \rightarrow 3p$ followed by re-emission, and is therefore itself a visit to $n = 3$, while $n_2/n_1 \ll 1$ makes the $\text{Ly}\beta$ optical depth far exceed that of $\text{H}\alpha$. With the $3p$ branching ratio of 0.88 into $\text{Ly}\beta$ and 0.12 into $\text{H}\alpha$, a photon undergoes of order eight $\text{Ly}\beta$ scatterings for each conversion into $\text{H}\alpha$, so the number of $n = 3$ visits exceeds the $\text{H}\alpha$ scattering count above by a comparable factor, lowering n_{crit} to $\sim 10^9\text{--}10^{10} \text{ cm}^{-3}$. In addition, the population of the $n = 3$ level is enhanced by trapping, and the Paschen line also becomes optically thick.

The emission region cannot be arbitrarily optically thick, however, and the $\text{H}\beta$ line bounds it. The line-center cross-sections scale as $f\lambda$, giving $\sigma_0(\text{H}\beta)/\sigma_0(\text{H}\alpha) \approx 0.14$, so a region with $\tau_{0,\text{H}\alpha} \sim 10^3\text{--}10^4$ has $\tau_{0,\text{H}\beta} \sim 10^2\text{--}10^3$. At such depths, $\text{H}\beta$ is efficiently removed by $4p \rightarrow 3$ branching and converted to $\text{Pa}\alpha$ and $\text{H}\alpha$ [21], so the region would destroy most of its own $\text{H}\beta$. Requiring instead $\tau_{0,\text{H}\beta} \lesssim 10$, so that $\text{H}\beta$ survives, caps $\tau_{0,\text{H}\alpha}$ at ~ 70 and raises the density needed for collisional de-excitation to $n_e \gtrsim 4 \times 10^{12} \text{ cm}^{-3}$, above the values quoted above. The observed Balmer ratios therefore bound this channel, unless the observed $\text{H}\beta$ originates in gas other than the region invoked to suppress $\text{H}\alpha$.

Collisions also act in the opposite direction. Collisional excitation out of $n = 2$ populates $n = 3$ and adds $\text{H}\alpha$ beyond its recombination value, overtaking recombination once $n_2/n_p \gtrsim 3 \times 10^{-6}$ at $T = 10^4 \text{ K}$ (4×10^{-5} at 5000 K), a ratio independent of

density. In a gas with the enhanced $n = 2$ populations inferred here, collisional excitation therefore raises the intrinsic $H\alpha$ equivalent width and widens the discrepancy, while only collisional de-excitation relieves it. Modeling the balance requires solving the level populations in an optically thick medium, which the present calculation does not attempt.

Understanding the collisional processes and radiative transfer in the local emission region is therefore important for the origin of the LRD continuum and for line formation. The present work models the scattering region, but the emission region that produces the Balmer lines and the recombination continuum warrants the same treatment. If that region is locally optically thick and dense enough for collisional processes to operate, the densities $n_H \sim 10^{10}\text{--}10^{11} \text{ cm}^{-3}$ suggested for LRDs [13, 19] are directly relevant to how the lines form.

Path length enhancement of resonance scattered $H\alpha$

Extended Data Fig. 8 measures how much extra H I column an $H\alpha$ photon samples because of resonance scattering, and converts that into an effective $n = 3$ bound-free optical depth, to test whether $H\alpha$ destruction during scattering can lower the intrinsic EW established above.

We use RT-SCAT [21] in an idealized configuration that isolates the resonance scattering path: a central point source emitting monochromatically at $H\alpha$ line center; a static outer H I sphere ($v_{\text{exp}} = 0 \text{ km s}^{-1}$, $\sigma_{\text{ran}} = 0 \text{ km s}^{-1}$) at $T = 10^4 \text{ K}$, giving a thermal Doppler parameter $b \approx 12.9 \text{ km s}^{-1}$; Thomson scattering disabled ($\tau_{\text{Th}} = 0$); and an optically thin interclump medium. Both the homogeneous and clumpy geometries are run, for $H\alpha$ and (identically) $H\beta$, with $N_{\text{HI},n=2} = 10^{13}\text{--}10^{20} \text{ cm}^{-2}$ and, for the clumpy model, covering factor $f_c \in \{1, 2, 5, 10, 20, 50, 100\}$ at fixed clump radius $r_{\text{cl}} = 10^{-3}R$. Each configuration follows 10^5 photons; for $H\alpha$ the escaping fraction is unity (pure scattering, no true absorption), which we monitor as a consistency check, while $H\beta$ photons are additionally removed by $4p \rightarrow 3$ conversion. Because the source is monochromatic and the medium is static, this calculation isolates the longest path available under these assumptions; an intrinsic line width, bulk motion, or clump velocity dispersion would each be expected to shorten it, so the resulting enhancement is approximately an upper limit.

For each escaping photon, we accumulate the path-integrated $n = 2$ column,

$$N_{\text{HI},n=2}^{\text{path}} = \sum_i n_{\text{HI},n=2,i} \Delta s_i, \quad (5)$$

summed over all propagation segments through $n = 2$ gas, and define the enhancement factor

$$\mathcal{E}_{\text{path}} \equiv \frac{\langle N_{\text{HI},n=2}^{\text{path}} \rangle}{N_{\text{HI},n=2}}, \quad (6)$$

the ensemble mean over all escaping photons. The mean, rather than the median, is the relevant statistic here because an absorption optical depth is additive along a photon's path. Extended Data Fig. 8 plots $\mathcal{E}_{\text{path}}$ against $N_{\text{HI},n=2}$; the upper axis gives the $H\alpha$ line-center optical depth $\tau_{0,H\alpha} = \sigma_0 N_{\text{HI},n=2}$ with $\sigma_0 \approx 5 \times 10^{-13} \text{ cm}^2$ at $T = 10^4 \text{ K}$.

Across the covering factors recovered from the H α fits ($f_c \approx 3$ –90), $\mathcal{E}_{\text{path}} \approx 2$ –30 over seven decades of column in both geometries, and we adopt a representative value $\mathcal{E}_{\text{path}} \sim 8$ (homogeneous: 2.0–28.9 over the full grid, ≈ 8 near the fitted columns; clumpy at $f_c = 20$, the sample-median covering factor: ≈ 3 –8). This is two to four orders of magnitude below the naive expectation that a resonantly trapped photon traverses of order τ_0 radial columns before escaping ($\tau_{0,\text{H}\alpha} \sim 10^4$ – 10^6 at the fitted columns). The medium is instead in the regime where photons escape after frequency redistribution — within the optically thick Doppler core at lower columns (a random walk of many short steps), or into the optically thin damping wing at higher columns (a few long flights) — with the accumulated path growing only as $\simeq (a\tau_0)^{1/3}$, where a is the Voigt damping parameter [77, 78]. In the clumpy models $\mathcal{E}_{\text{path}}$ falls as f_c decreases: at fixed total column, fewer clumps are each more optically thick, and a photon that strikes one is re-emitted from its thin surface layer rather than crossing it. Below $f_c \approx 3$ the medium is porous — some H α photons escape without intersecting any clump — and $\mathcal{E}_{\text{path}} < 1$.

Extended Data Fig. 8 measures path enhancement, not H α destruction. To connect the two we form the path-integrated $n = 3$ column,

$$N_{n=3}^{\text{path}} = \frac{n_3}{n_2} \mathcal{E}_{\text{path}} N_{\text{HI},n=2}, \quad (7)$$

taking for n_3/n_2 its LTE (Boltzmann) value at the temperature of the cool absorbing layer, computed from the level statistical weights and the $n = 2$ –3 excitation energy: $n_3/n_2 \approx 0.03$ –0.25 for $T \approx 5000$ – 10^4 K. This is an overestimate — resonant Ly α trapping enhances the $n = 2$ population and lowers the true n_3/n_2 below its LTE value — so the $\tau_{\text{bf,eff}}$ derived below is an upper limit. The effective H α bound-free optical depth is

$$\tau_{\text{bf,eff}}(\text{H}\alpha) = \sigma_{\text{bf},n3}(\text{H}\alpha) N_{n=3}^{\text{path}}, \quad (8)$$

with $\sigma_{\text{bf},n3}(\text{H}\alpha) \approx 1 \times 10^{-17} \text{ cm}^2$ — the Paschen-edge threshold cross-section ($2.1 \times 10^{-17} \text{ cm}^2$ at 8206 Å) scaled by $(\nu_3/\nu_{\text{H}\alpha})^3 \approx 0.5$ [29]. At the fitted columns ($N_{\text{HI},n=2} \sim 10^{17} \text{ cm}^{-2}$) this gives $\tau_{\text{bf,eff}} \approx 0.2$ –2; even at this LTE upper limit the H α reduction is at most ~ 85 percent, short of the factor ~ 10 –100 the EW discrepancy requires.

An H α photon that photoionizes hydrogen from $n = 3$ is not necessarily lost, since a subsequent recombination can regenerate H α . Writing the destruction probability per $n = 3$ photoionization as ϵ_{dest} , so that $\tau_{\text{dest,eff}} \simeq \epsilon_{\text{dest}} \tau_{\text{bf,eff}}$, the estimate above takes the conservative maximum $\epsilon_{\text{dest}} = 1$ (every photoionization permanently removes one H α photon). Since the reduction falls short of the required factor ~ 10 –100 even in this limit, allowing for recombination recycling ($\epsilon_{\text{dest}} < 1$) does not change the conclusion.

Producing a factor-of-10 H α suppression would instead require $N_{n=3} \sim 3 \times 10^{16} \text{ cm}^{-2}$. The same $n = 3$ column implies a Paschen continuum break of optical depth $\tau \sim 0.6$ and a Pa α line-center optical depth $\sim 10^3$ – 10^4 (the bound-bound to bound-free cross-section ratio $\sigma_{\text{bb}}(\text{Pa}\alpha)/\sigma_{\text{bf},n3} \approx 10^4$), i.e. saturated Paschen line absorption. Saturated Paschen-line absorption is absent in the LRDs with measured Paschen lines, even though these same objects show clear Balmer absorption [66, 67]; the weaker continuum break is not yet constrained and could be partly filled in by Paschen recombination continuum. The same constraint applies to the collisional route

above, but for a different reason: bound–free suppression requires a large $n = 3$ population as a prerequisite, whereas collisional de-excitation under line trapping produces one as a consequence. Both therefore predict the same Paschen signature, and its depth scales with the claimed suppression, reaching $\tau_{\text{Pa}\alpha} \approx 0.94 \tau_{\text{H}\alpha}$ at 10^4 K in the fully thermalized limit. Paschen features lie redward of rest-frame $1 \mu\text{m}$ and are difficult to measure at these redshifts, so this is not a proof; but it shows that resolving the EW puzzle through $n = 3$ bound–free suppression of $\text{H}\alpha$ carries an observational cost, saturated Paschen-line absorption that is not observed, which a separate optical-continuum component does not, and that Paschen-band spectroscopy is a clean discriminating test [15].

Dust along this same enhanced path could similarly extinguish $\text{H}\alpha$ relative to the continuum, independent of the bound–free mechanism above. However, current LRD models generally invoke dust-poor or dust-free line-emitting gas, and direct evidence for dust in this medium is lacking, so we do not pursue this further.

Extended Data

Table 1 LRD sample. Objects sorted by redshift. $z_{[\text{OIII}]}$: systemic redshift from the $[\text{O III}]_{4960,5008}$ doublet. Grating: NIRSpec disperser used for Balmer-line profile fitting. $\mathcal{R}$: approximate spectral resolution at $\text{H}\beta$. $\text{S/N}(\text{H}\alpha)$, $\text{S/N}(\text{H}\beta)$: integrated-line signal-to-noise measured from our flux-calibrated spectra. PID: JWST programme ID. PI: proposal principal investigator(s).

Object	$z_{[\text{OIII}]}$	Grating	$\mathcal{R}$	$\text{S/N}(\text{H}\alpha)$	$\text{S/N}(\text{H}\beta)$	PID	PI
RUBIES UDS 154183	3.548	G235H	2700	49	7	9433	Maiolino
A2744 45924	4.464	G395M	1000	498	129	8204	Greene & Labbe
JADES GN 68797	5.039	G395M	1000	174	13	1181	Eisenstein
GN 15498	5.085	G395H	2700	26	5	5664	Matthee
RUBIES EGS 42046	5.276	G235M	1000	88	11	4233	De Graaff & Brammer
GS 13971	5.482	G395H	2700	57	10	5664	Matthee
GN 9771	5.535	G395H	2700	342	39	5664	Matthee
GS 3073	5.553	G395H	2700	242	98	1216	Lützgendorf
JADES GN 38147	5.867	G395M	1000	36	7	1181	Eisenstein
RUBIES EGS 49140	6.685	G235M	1000	57	17	4233	De Graaff & Brammer
RUBIES EGS 55604	6.983	G235M	1000	55	15	4233	De Graaff & Brammer

Table 2 Free parameters of the Balmer-line radiative-transfer models. $N_{\text{HI},n=2}$ is the total H I column density for both models; in the clumpy case, the individual clump column density is $N_{\text{HI},\text{cl}} = \frac{3}{4} N_{\text{HI},n=2}/f_c$ (derived). Dashes indicate the parameter does not apply to that model; ‘same’ denotes an identical range to the homogeneous model. T_e and r_{cl} are fixed.

Parameter	Homogeneous	Clumpy
$N_{\text{HI},n=2} [\text{cm}^{-2}]$	$10^{12}\text{--}10^{20}$	same
$v_{\text{exp}} [\text{km s}^{-1}]$	–400 to +400	same
$\sigma_{\text{emit}} [\text{km s}^{-1}]$	100–300	same
$\sigma_{\text{ran}} [\text{km s}^{-1}]$	50–150	same
τ_{Th}	0.5–3.0	same
f_c	—	1–100

Table 3 Balmer-break SED fitting results for all 11 LRDs. α_{UV} : UV power-law slope from Stage 1. T_{BB} : effective optical continuum temperature from Stage 2; this is a fitting parameter, not a physical temperature (see text). f_{exc} : optical-excess parameter (equation 1); these three columns do not depend on the H I velocity model. $\log N_{\text{HI},n=2}$ and v_{ran} are the Stage-3 high-n Balmer series fit values (see text), reported separately for the clumpy and homogeneous models since each is fit with its own model-specific H α -derived prior on $\log N_{\text{HI},n=2}$. Superscripts/subscripts give the 84th/16th percentiles of the Stage-1, Stage-2, and Stage-3 posteriors. The last two rows (GS 3073 and JADES GN 38147) did not satisfy the detection criterion ($f_{\text{exc}} < 1$); dashes indicate that the Stage-3 high-n Balmer series fit is not meaningful for these objects. For JADES GN 38147, the optical excess is too weak to constrain T_{BB} . ‡: the 84th percentile of $N_{\text{HI},n=2}$ is below $10^{16.5} \text{ cm}^{-2}$ ($\tau_{\text{BB}} = 0.5$), too low to produce the Balmer break; the median is given and v_{ran} is unconstrained.

Object	z	α_{UV}	T_{BB} (K)	f_{exc}	Clumpy		Homogeneous	
					$\log N_{\text{HI},n=2} (\text{cm}^{-2})$	$v_{\text{ran}} (\text{km s}^{-1})$	$\log N_{\text{HI},n=2} (\text{cm}^{-2})$	$v_{\text{ran}} (\text{km s}^{-1})$
GN 9771	5.535	$-0.87^{+0.05}_{-0.05}$	5344^{+114}_{-109}	$1.95^{+0.13}_{-0.13}$	$17.02^{+0.03}_{-0.03}$	400^{+67}_{-85}	$13.75^{\ddagger}$	—
A2744 45924	4.464	$-0.43^{+0.06}_{-0.06}$	4802^{+71}_{-70}	$2.65^{+0.19}_{-0.19}$	$17.20^{+0.06}_{-0.06}$	388^{+74}_{-88}	$14.08^{\ddagger}$	—
RUBIES EGS 49140	6.685	$-0.62^{+0.12}_{-0.11}$	4712^{+224}_{-209}	$2.80^{+0.31}_{-0.27}$	$18.29^{+0.38}_{-0.46}$	275^{+96}_{-80}	$17.86^{+0.44}_{-0.32}$	300^{+100}_{-87}
RUBIES EGS 55604	6.983	$-1.08^{+0.10}_{-0.10}$	4644^{+329}_{-300}	$2.27^{+0.23}_{-0.24}$	$17.54^{+0.24}_{-0.19}$	356^{+89}_{-94}	$17.40^{+0.17}_{-0.13}$	367^{+84}_{-93}
JADES GN 68797	5.039	$-0.70^{+0.18}_{-0.17}$	3396^{+89}_{-87}	$5.64^{+1.04}_{-0.90}$	$17.57^{+0.25}_{-0.26}$	151^{+129}_{-74}	$14.41^{\ddagger}$	—
RUBIES EGS 42046	5.276	$-0.25^{+0.16}_{-0.16}$	4913^{+343}_{-312}	$1.26^{+0.34}_{-0.28}$	$17.46^{+0.30}_{-0.21}$	342^{+105}_{-118}	$17.70^{+0.28}_{-0.28}$	330^{+111}_{-116}
RUBIES UDS 154183	3.548	$-1.13^{+0.61}_{-0.60}$	4106^{+90}_{-87}	$18.15^{+13.2}_{-7.70}$	$17.59^{+0.54}_{-0.34}$	211^{+129}_{-82}	$13.93^{\ddagger}$	—
GN 15498	5.085	$-1.80^{+0.20}_{-0.20}$	4255^{+196}_{-187}	$6.63^{+1.68}_{-1.36}$	$18.18^{+0.48}_{-0.58}$	269^{+142}_{-131}	$17.92^{+0.56}_{-0.48}$	278^{+138}_{-135}
GS 13971	5.482	$-1.65^{+0.11}_{-0.11}$	3680^{+251}_{-236}	$2.90^{+0.43}_{-0.40}$	$16.81^{+0.47}_{-1.44}$	200^{+188}_{-138}	$14.81^{\ddagger}$	—
GS 3073	5.553	$-1.61^{+0.02}_{-0.02}$	8699^{+1565}_{-1075}	$0.36^{+0.02}_{-0.02}$	—	—	—	—
JADES GN 38147	5.867	$-1.52^{+0.11}_{-0.11}$	—	$0.63^{+0.17}_{-0.16}$	—	—	—	—

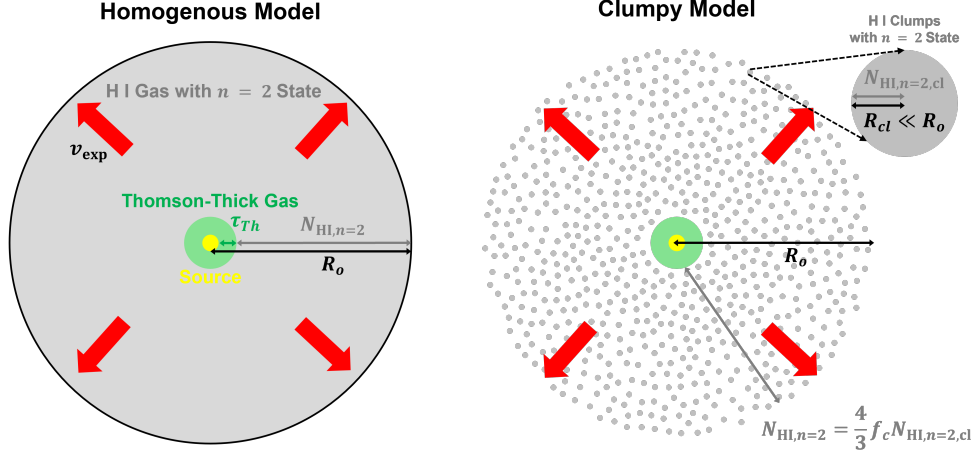


Fig. 1 Scattering geometry of the homogeneous (left) and clumpy (right) H I gas. A central isotropic source (yellow) emits $H\alpha$ and $H\beta$ emission with a Gaussian profile of width σ_{emit} . The source is surrounded by two concentric regions: an inner, fully ionized H II region (green, $r < 0.01 R_o$) with Thomson optical depth τ_{Th} , and an outer neutral H I region (grey, $0.01 R_o < r < R_o$) with a significantly populated $n = 2$ level and total column density $N_{HI,n=2}$ and a random motion σ_{ran} . Red arrows: bulk radial outflow velocity v_{exp} , constant with radius ($v(r) = v_{exp}$). Left: the H I region is a uniform-density sphere; this homogeneous-model geometry follows [21], Fig. 15. Right: the same total column $N_{HI,n=2}$ is instead distributed among discrete, optically thick clumps of radius $R_{cl} \ll R_o$ and column density $N_{HI,n=2,cl}$, related by $N_{HI,n=2} = \frac{4}{3} f_c N_{HI,n=2,cl}$, where f_c is the covering factor, which is the average number of clumps from the central source. Individual clumps additionally carry a random peculiar velocity dispersion σ_{ran} . The random motion within clumps is the thermal motion, as their temperature is at $T = 10^4$ K. This clumpy-medium treatment follows [25, 28].

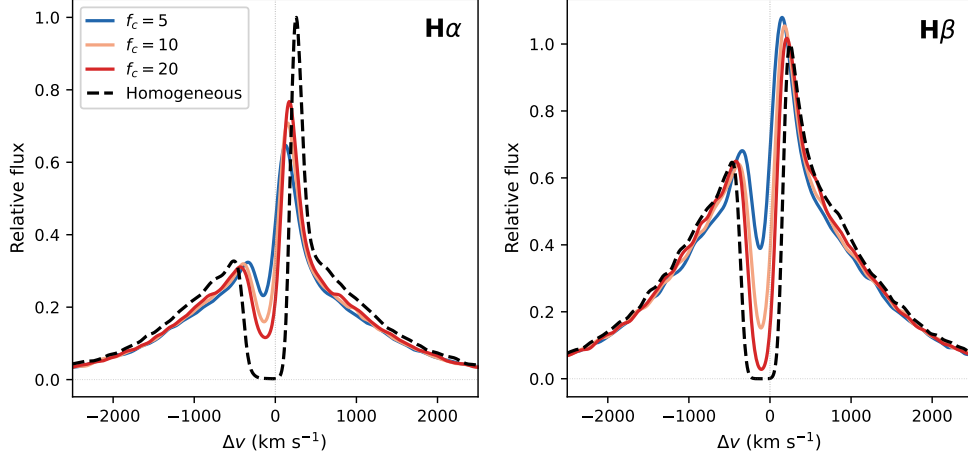


Fig. 2 Effect of covering factor f_c on $H\alpha$ and $H\beta$ profiles at fixed total column density. Simulated $H\alpha$ (left) and $H\beta$ (right) profiles from direct RT-SCAT Monte Carlo outputs (no emulator) at fixed representative parameters ($\log N_{\text{HI},n=2} = 17 [\text{cm}^{-2}]$, $v_{\text{exp}} = 100 \text{ km s}^{-1}$, $\sigma_{\text{emit}} = 200 \text{ km s}^{-1}$, $\sigma_{\text{ran}} = 100 \text{ km s}^{-1}$, $\tau_{\text{Th}} = 1.5$), for clumpy models with $f_c = 5, 10, 20$ and the homogeneous case. The primary contrast is between clumpy and homogeneous: all three clumpy models show dramatically weaker $H\alpha$ absorption than the homogeneous case at the same $N_{\text{HI},n=2}$, while $H\beta$ absorption is reduced far less. This asymmetric response — the porous clump geometry selectively suppresses $H\alpha$ while retaining substantial $H\beta$ absorption — may be the physical origin of the $H\alpha$ – $H\beta$ absorption discrepancy observed in LRDs. The three f_c values cluster together in both panels, indicating that the surface-scattering regime is reached robustly across the range of covering factors inferred from the data[26, 27].

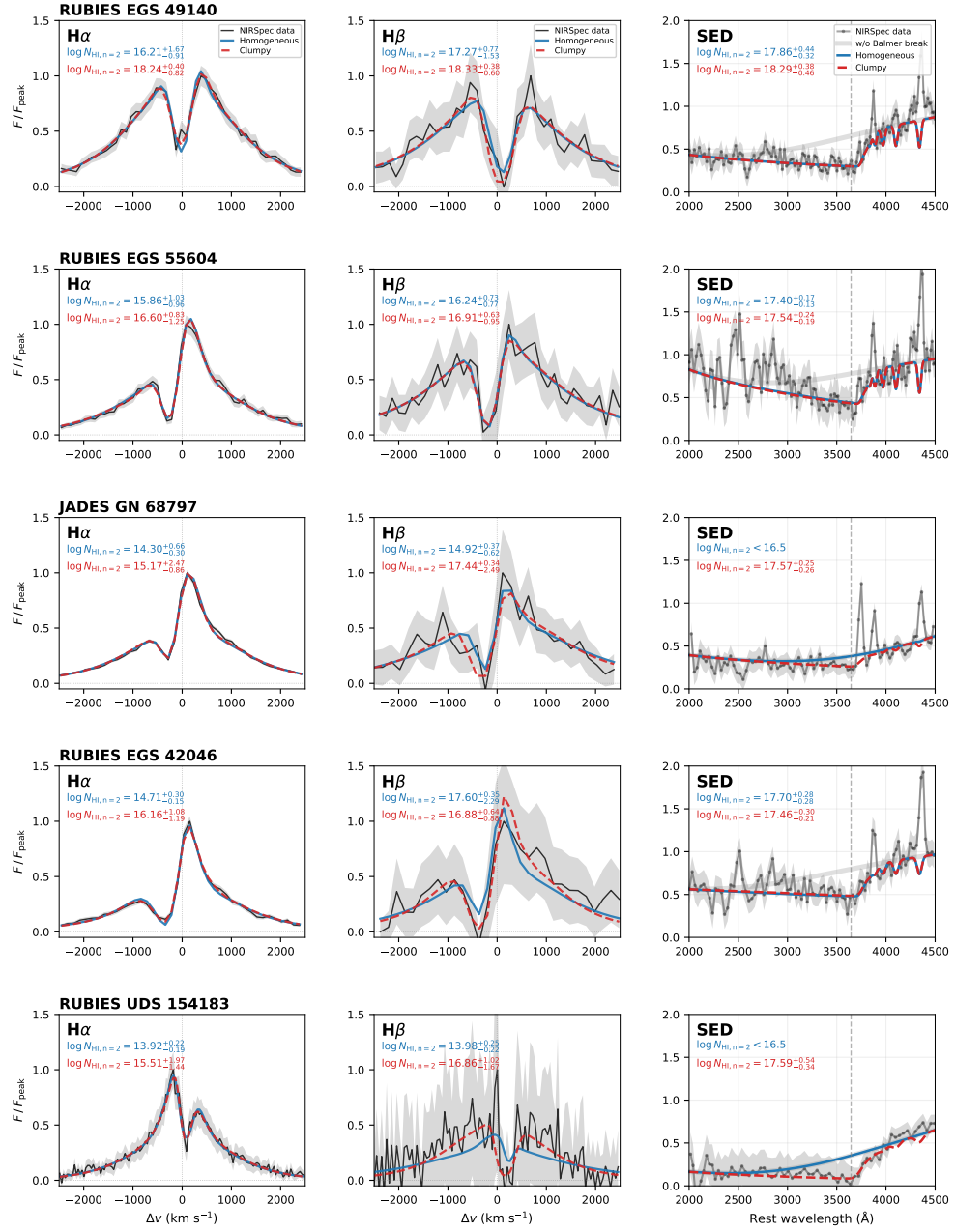


Fig. 3 H α , H β , and Balmer-break fits for the remaining 9 LRDs (1 of 2). Format and colors as in Fig. 1; GN 9771 and A2744 45924 are shown in Fig. 1. One row per object, in order: RUBIES EGS 49140, RUBIES EGS 55604, JADES GN 68797, RUBIES EGS 42046, and RUBIES UDS 154183. The remaining 4 objects are shown in Extended Data Fig. 4.

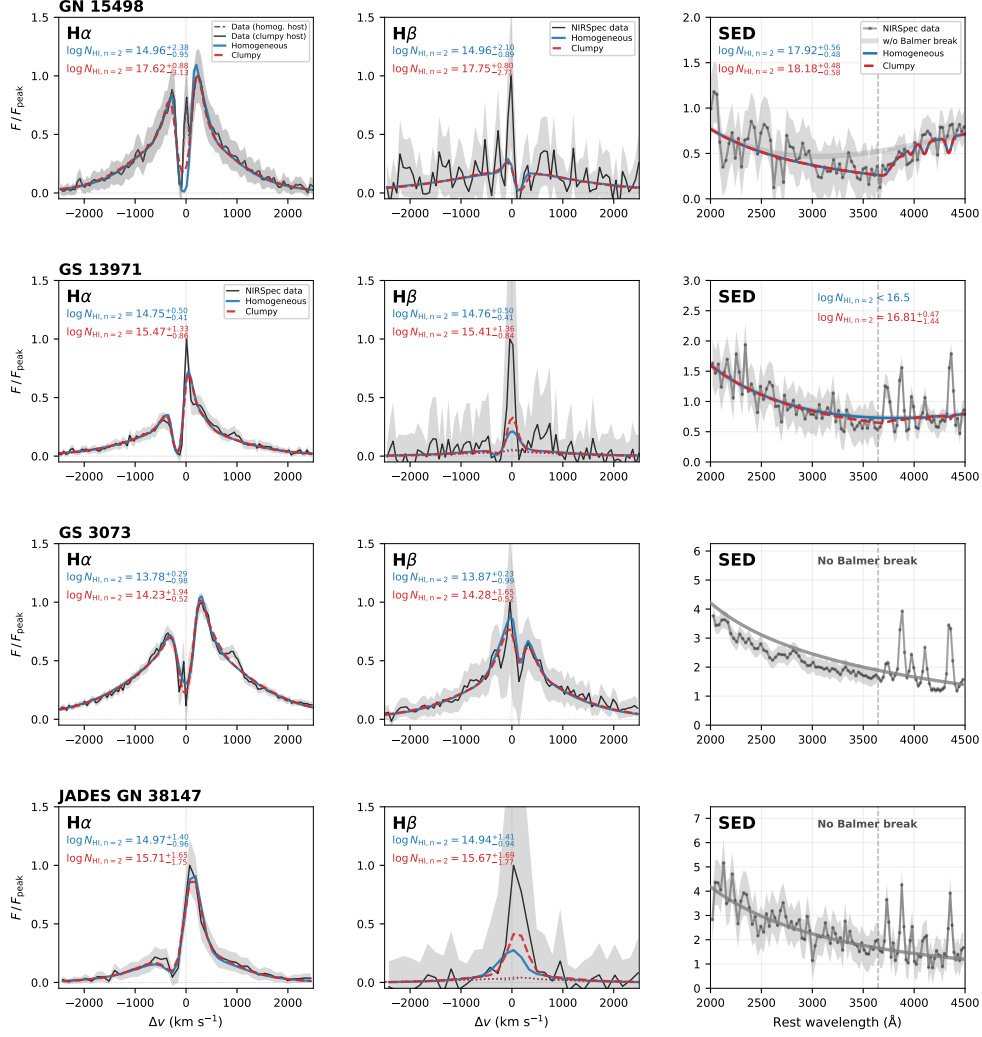


Fig. 4 H α , H β , and Balmer-break fits for the remaining 9 LRDs (2 of 2). Format and colors as in Fig. 1; the first 5 objects are shown in Extended Data Fig. 3. One row per object, in order: GN 15498, GS 13971, GS 3073, and JADES GN 38147. GS 3073 and JADES GN 38147 did not satisfy the Balmer-break detection criterion ($f_{\text{exc}} < 1$); their SED panels show only the continuum without a break and are labeled ‘No Balmer Break’. For GN 15498, the H α panel shows both observed spectra for the clumpy and homogeneous models, since the two diverge visibly for this object due to different [O III]-based subtraction.

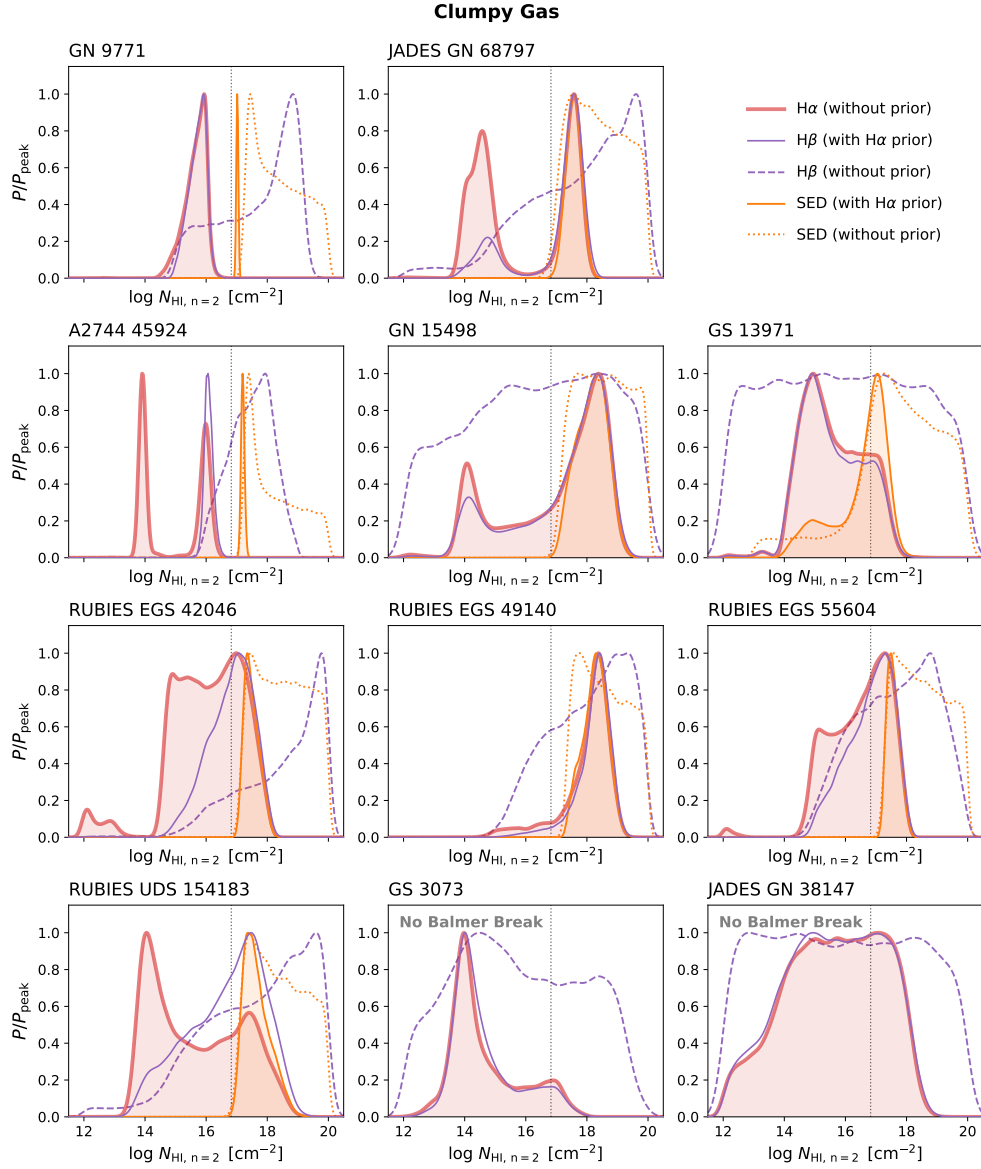


Fig. 5 $N_{\text{HI}, n=2}$ posteriors for all 11 LRDs, clumpy model. Each panel shows, for one object, the one-dimensional posterior probability density (peak-normalised) for $\log N_{\text{HI}, n=2}$ from five probes: the $\text{H}\alpha$ -only fit (solid); the $\text{H}\beta$ fit with the $\text{H}\alpha$ -derived $N_{\text{HI}, n=2}$ prior (solid) and without this prior (dashed); and the Balmer-break SED fit with the $\text{H}\alpha$ -derived prior (solid) and without this prior (dotted). The $\text{H}\alpha$ and SED-with-prior posteriors are shaded, and the vertical dotted line marks $\tau_{\text{BB}} = 1$. $\text{H}\alpha$ and $\text{H}\beta$ posteriors are from the full MCMC chains; SED posteriors are from the Stage-3 MCMC chains (see Methods). GS 3073 and JADES GN 38147 have no SED posteriors because these two objects have no Balmer break.

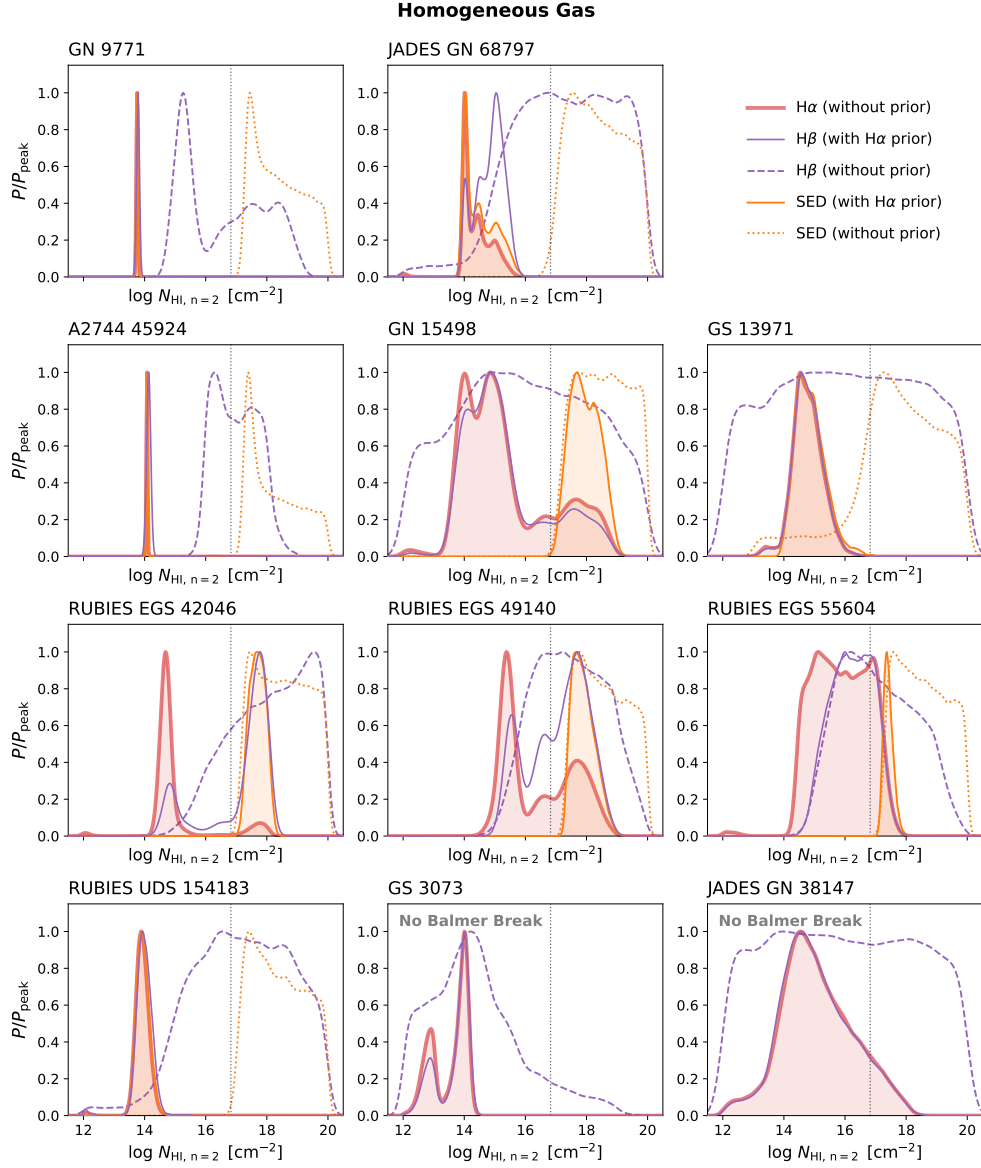


Fig. 6 $N_{\text{HI}, n=2}$ posteriors for all 11 LRDs, homogeneous model. As Extended Data Fig. 5, for the homogeneous radiative-transfer model.

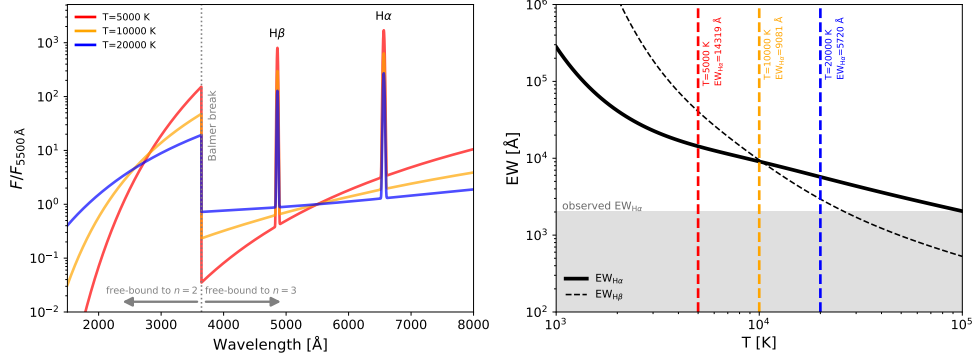


Fig. 7 Nebular continuum with H α and H β emission, and the equivalent-width cost of a nebular-dominated optical continuum. *Left:* the hydrogen nebular recombination continuum (free-free + free-bound; two-photon emission omitted) at $T = 5000, 10^4$, and 2×10^4 K, normalized to unity at rest-frame 5500 Å, with the H α and H β emission it implies drawn on top at the self-consistent single-zone Case B equivalent widths of the right panel (Gaussian, 500 km s $^{-1}$). *Right:* the intrinsic EW $_{H\alpha}$ implied if the nebular continuum alone supplies the observed optical continuum, versus T (electron density and emitting volume cancel in this ratio); vertical lines mark the three left-panel temperatures, the shaded band the representative observed LRD range, ~ 100 –2000 Å. The thin dashed curve shows the corresponding EW $_{H\beta}$, shown for consistency with the H β line drawn in the left panel and not as an independent constraint. Over the plotted range EW $_{H\alpha} \sim 10^3$ – 10^5 Å, one to two orders of magnitude above typical observed LRD values. See Methods.

Table 4 Homogeneous (smooth) model fitted parameters. Median and 68 percent credible intervals for all 11 LRDs, from the individual-fit chain (see Methods): for each object, the H α row is the H α -only fit; the H β row is the H β fit using the $N_{\text{HI},n=2}$ posterior of this H α fit as a prior. Object order as in Fig. 2. Italicized rows (GS 3073 and JADES GN 38147) have no Balmer-break detection; line-fitting results are included for completeness. v_{exp} : envelope expansion velocity (negative values indicate infall). σ_{emit} : intrinsic emission-line width. σ_{ran} : random velocity dispersion. τ_{Th} : Thomson optical depth. χ^2_r : χ^2 of the median posterior model per spectral bin ($|v| < 3000$ km s $^{-1}$ for H α , < 2000 km s $^{-1}$ for H β).

Object	Fit	$\log N_{\text{HI},n=2}$ (cm $^{-2}$)	v_{exp} (km s $^{-1}$)	σ_{emit} (km s $^{-1}$)	σ_{ran} (km s $^{-1}$)	τ_{Th}	χ^2_r
GN 9771	H α	$13.75^{+0.03}_{-0.03}$	173^{+6}_{-6}	227^{+8}_{-8}	118^{+4}_{-4}	$1.79^{+0.02}_{-0.02}$	1.30
	H β	$13.79^{+0.03}_{-0.03}$	353^{+28}_{-28}	135^{+30}_{-24}	71^{+15}_{-14}	$2.93^{+0.05}_{-0.13}$	1.25
A2744 45924	H α	$14.08^{+0.02}_{-0.02}$	-25^{+2}_{-2}	102^{+3}_{-1}	149^{+0}_{-1}	$1.88^{+0.01}_{-0.01}$	14.07
	H β	$14.14^{+0.02}_{-0.02}$	-256^{+12}_{-70}	295^{+4}_{-10}	123^{+6}_{-12}	$2.99^{+0.01}_{-0.04}$	5.38
RUBIES EGS 49140	H α	$16.20^{+1.67}_{-0.91}$	12^{+8}_{-7}	157^{+101}_{-44}	102^{+42}_{-43}	$1.81^{+0.14}_{-0.16}$	0.49
	H β	$17.27^{+0.77}_{-1.53}$	-73^{+82}_{-88}	200^{+70}_{-69}	109^{+26}_{-25}	$2.41^{+0.39}_{-0.49}$	0.29
RUBIES EGS 55604	H α	$15.86^{+1.03}_{-0.96}$	260^{+32}_{-24}	272^{+21}_{-42}	69^{+21}_{-13}	$1.71^{+0.09}_{-0.09}$	0.44
	H β	$16.25^{+0.73}_{-0.77}$	158^{+103}_{-90}	204^{+67}_{-70}	104^{+28}_{-26}	$2.50^{+0.33}_{-0.47}$	0.22
JADES GN 68797	H α	$14.30^{+0.66}_{-0.30}$	320^{+16}_{-15}	221^{+27}_{-31}	81^{+28}_{-21}	$1.80^{+0.05}_{-0.04}$	1.33
	H β	$14.92^{+0.37}_{-0.62}$	279^{+84}_{-112}	191^{+72}_{-64}	113^{+27}_{-37}	$2.54^{+0.32}_{-0.49}$	0.29
RUBIES EGS 42046	H α	$14.71^{+0.30}_{-0.15}$	331^{+24}_{-21}	144^{+41}_{-29}	141^{+7}_{-22}	$1.74^{+0.08}_{-0.08}$	0.82
	H β	$17.60^{+0.35}_{-2.29}$	303^{+69}_{-102}	193^{+71}_{-65}	118^{+22}_{-30}	$1.65^{+0.67}_{-0.57}$	0.19
RUBIES UDS 154183	H α	$13.92^{+0.22}_{-0.19}$	-57^{+20}_{-24}	280^{+14}_{-25}	131^{+13}_{-20}	$1.11^{+0.09}_{-0.08}$	0.23
	H β	$13.98^{+0.25}_{-0.22}$	-165^{+251}_{-131}	206^{+67}_{-73}	101^{+34}_{-34}	$2.43^{+0.40}_{-0.59}$	0.21
GN 15498	H α	$14.96^{+2.38}_{-0.95}$	21^{+52}_{-14}	273^{+20}_{-53}	95^{+40}_{-37}	$1.00^{+0.12}_{-0.16}$	0.30
	H β	$14.97^{+2.10}_{-0.89}$	-107^{+217}_{-150}	198^{+70}_{-68}	98^{+35}_{-33}	$2.38^{+0.44}_{-0.70}$	0.30
GS 13971	H α	$14.75^{+0.50}_{-0.41}$	119^{+27}_{-27}	275^{+18}_{-37}	59^{+13}_{-7}	$1.20^{+0.09}_{-0.09}$	0.32
	H β	$14.76^{+0.50}_{-0.42}$	44^{+243}_{-294}	201^{+67}_{-68}	100^{+35}_{-33}	$1.72^{+0.88}_{-0.82}$	0.09
<i>GS 3073</i>	H α	$13.78^{+0.29}_{-0.98}$	100^{+154}_{-25}	292^{+6}_{-12}	139^{+8}_{-58}	$1.77^{+0.07}_{-0.11}$	0.50
	H β	$13.87^{+0.23}_{-0.99}$	-54^{+218}_{-126}	254^{+33}_{-55}	104^{+32}_{-35}	$1.34^{+0.22}_{-0.21}$	0.18
<i>JADES GN 38147</i>	H α	$14.97^{+1.41}_{-0.96}$	232^{+95}_{-83}	155^{+80}_{-41}	80^{+36}_{-21}	$0.78^{+0.18}_{-0.16}$	0.27
	H β	$14.94^{+1.40}_{-0.94}$	1^{+269}_{-267}	203^{+67}_{-70}	99^{+35}_{-33}	$1.66^{+0.91}_{-0.81}$	0.08

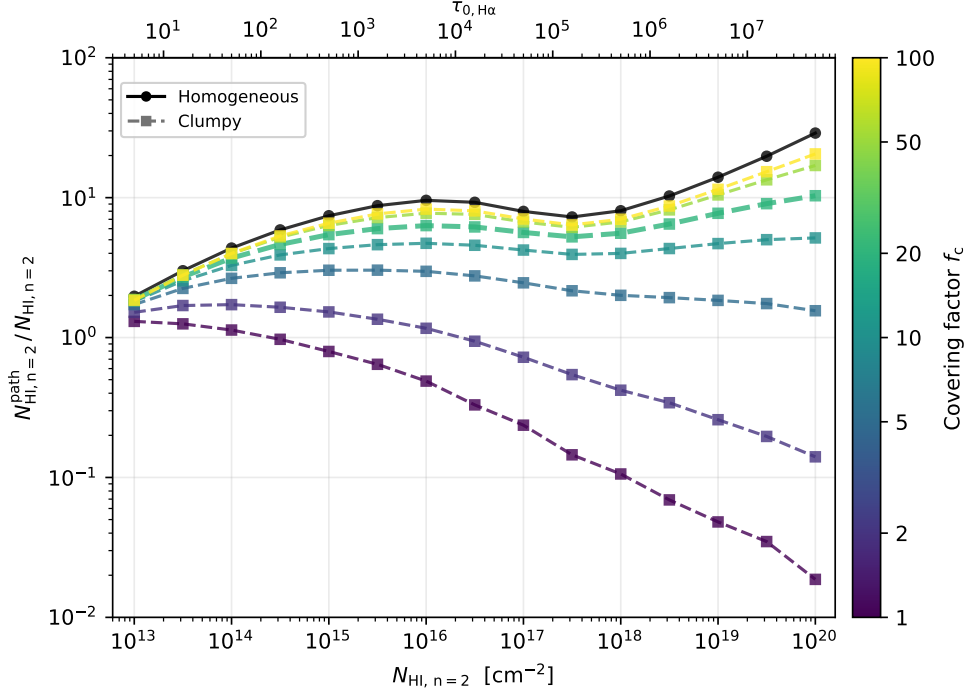


Fig. 8 The $n = 2$ neutral-hydrogen column sampled by a resonantly scattered $H\alpha$ photon exceeds the radial column by only a factor ~ 8 . Ensemble-mean path-integrated $n = 2$ column divided by the radial column $N_{\text{HI},n=2}$, from RT-SCAT Monte Carlo runs of 10^5 line-center $H\alpha$ photons per grid point through a static HI medium ($\tau_{\text{Th}} = 0$), for the homogeneous model (solid) and clumpy models with $r_{\text{cl}} = 10^{-3}R$ and covering factor $f_c = 1$ –100 (dashed, color-coded; $f_c = 20$, the sample-median value, drawn heavier). Upper axis: $H\alpha$ line-center optical depth $\tau_{0,H\alpha} = \sigma_0 N_{\text{HI},n=2}$, $\sigma_0 \simeq 5 \times 10^{-13} \text{ cm}^2$. The amplification is ≈ 2 –30 across seven decades of column in both geometries (representative value ~ 8) — two to four orders of magnitude below the $\sim \tau_0$ expected for a core-trapped photon. Converted to an $n = 3$ bound-free optical depth, this is short of the factor ~ 10 –100 suppression of $H\alpha$ the equivalent-width puzzle would require (see Methods). $H\beta$ (not shown): ratio below unity throughout.

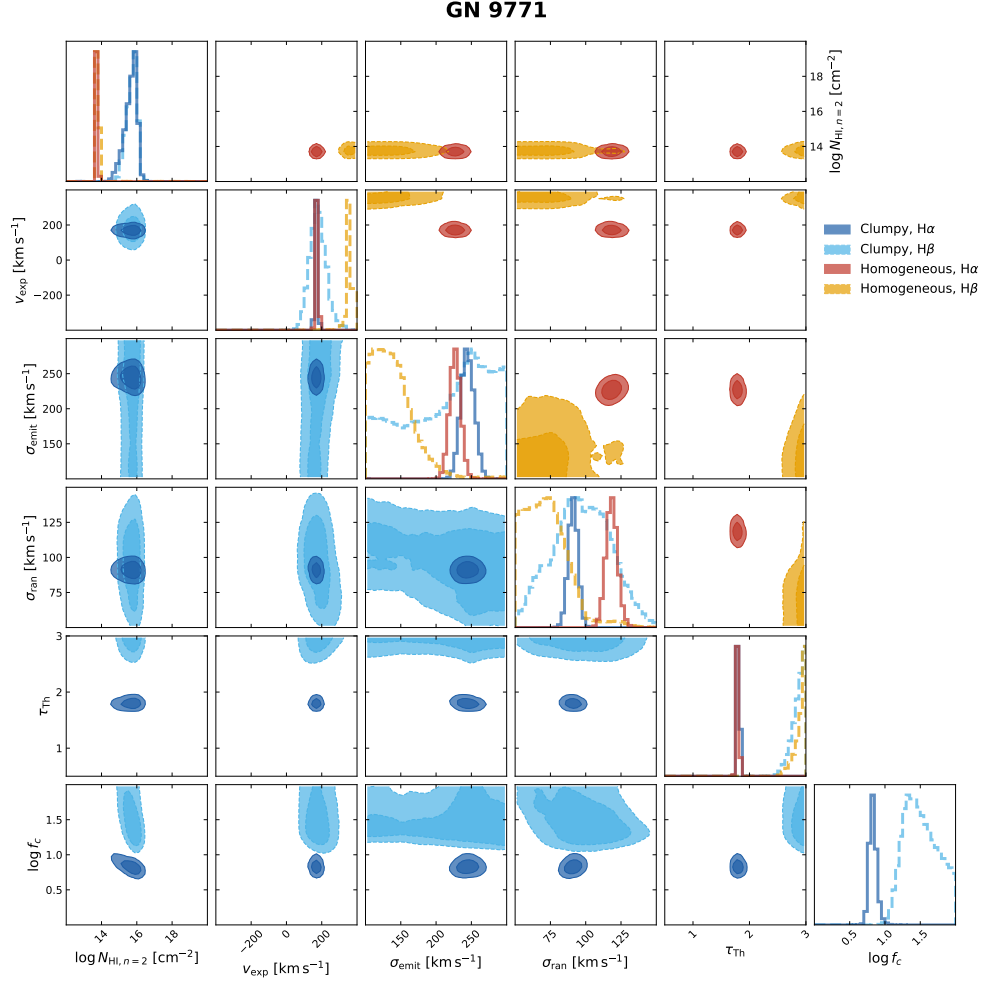


Fig. 9 Posterior distributions of the Balmer-line radiative-transfer fits for GN 9771. The lower-left triangle shows the clumpy model and the upper-right triangle the homogeneous model, with the axes of the upper triangle transposed. Dark blue and red show the H α -only fit (Stage 1), and light blue and orange show the H β -only fit (Stage 2), which uses the $N_{\text{HI},n=2}$ prior from the H α fit of the same model (see Methods). Filled contours mark the 68 and 95 percent credible regions; H β contours have dashed outlines. Diagonal panels show the one-dimensional marginalized posteriors, each normalized to unit peak (solid: H α , dashed: H β). The homogeneous model has no f_c , so its column is left empty. Axis ranges span the full prior ranges of Extended Data Table 2; the remaining nuisance parameters are marginalized over. With the H α -based prior, the H β fits of both models stay at $N_{\text{HI},n=2}$ from H α ($\log N_{\text{HI},n=2} \approx 15.8$ and 13.8 $[\text{cm}^{-2}]$ for clumpy and homogeneous). The homogeneous H β fit is pushed to the upper prior bounds of v_{exp} and τ_{Th} and still underpredicts the H β absorption (Fig. 1), whereas the clumpy H β fit reproduces it.

A2744 45924

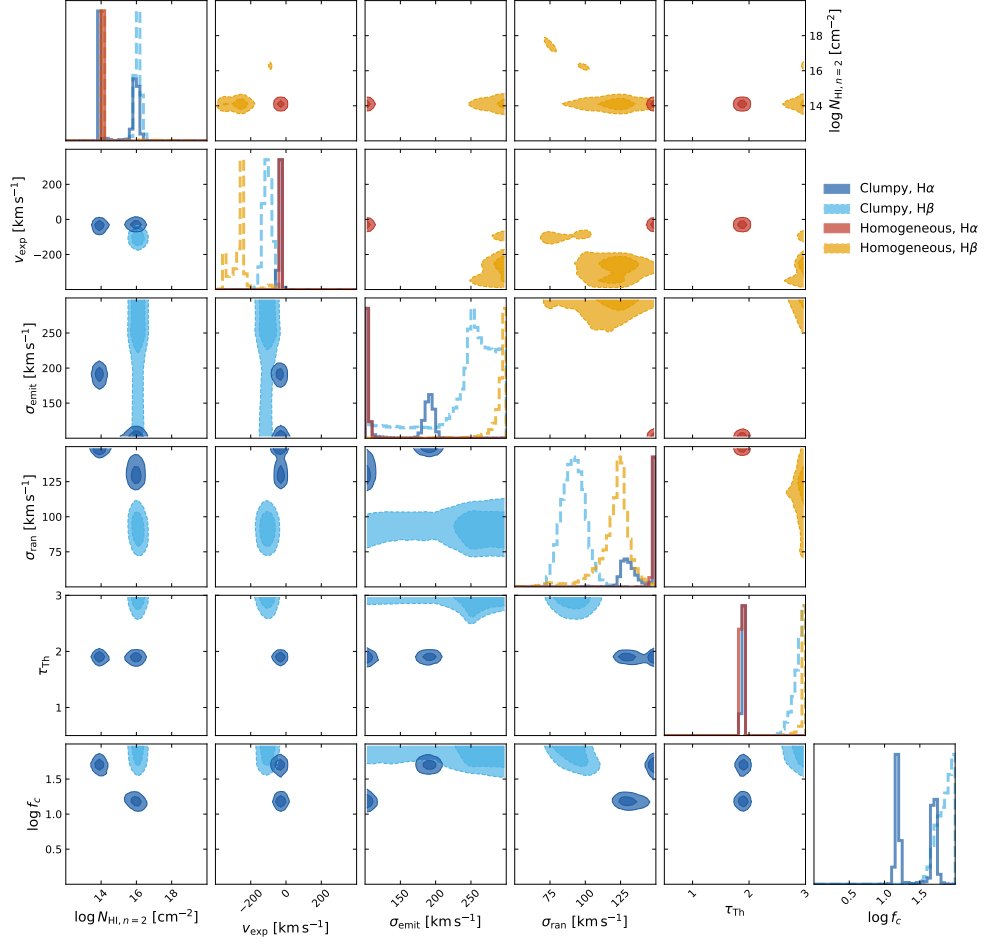


Fig. 10 Posterior distributions of the Balmer-line radiative-transfer fits for **A2744 45924**. Same as Extended Data Fig. 9. The clumpy H α posterior is bimodal, with modes at $\log N_{\text{HI},n=2} \approx 14$ and $16 \text{ [cm}^{-2}\text{]}$ that differ in f_c ; the clumpy H β fit selects the upper mode ($\log N_{\text{HI},n=2} \approx 16.1$). The homogeneous H β fit stays close to $N_{\text{HI},n=2}$ from H α ($\log N_{\text{HI},n=2} \approx 14.1$), is pushed to the upper τ_{Th} and σ_{emit} bounds, and still underpredicts the H β absorption (Fig. 1). For this object, the host Balmer decrement sits at its upper limit of 3.3 in both models.

Table 5 Clumpy model fitted parameters. As Extended Data Table 4, for the clumpy radiative-transfer model (adds the covering factor f_c , shown as $\log f_c$). All other columns and sources as in Extended Data Table 4.

Object	Fit	$\log N_{\text{HI},n=2} \text{ (cm}^{-2}\text{)}$	$\log f_c$	$v_{\text{exp}} \text{ (km s}^{-1}\text{)}$	$\sigma_{\text{emit}} \text{ (km s}^{-1}\text{)}$	$\sigma_{\text{ran}} \text{ (km s}^{-1}\text{)}$	τ_{Th}	χ^2_{r}
GN 9771	H α	$15.71^{+0.27}_{-0.43}$	$0.83^{+0.05}_{-0.05}$	169^{+6}_{-6}	244^{+10}_{-9}	91^{+3}_{-3}	$1.80^{+0.02}_{-0.02}$	1.15
	H β	$15.75^{+0.24}_{-0.35}$	$1.49^{+0.30}_{-0.22}$	178^{+6}_{-44}	229^{+48}_{-86}	97^{+21}_{-21}	$2.88^{+0.08}_{-0.14}$	0.39
A2744 45924	H α	$15.27^{+0.76}_{-1.37}$	$1.26^{+0.45}_{-0.09}$	-28^{+5}_{-12}	111^{+83}_{-10}	144^{+5}_{-17}	$1.89^{+0.01}_{-0.01}$	13.88
	H β	$16.08^{+0.15}_{-0.12}$	$1.87^{+0.10}_{-0.14}$	-106^{+23}_{-23}	257^{+28}_{-47}	92^{+8}_{-8}	$2.90^{+0.07}_{-0.12}$	2.13
RUBIES EGS 49140	H α	$18.24^{+0.40}_{-0.82}$	$1.50^{+0.31}_{-0.30}$	16^{+10}_{-9}	217^{+49}_{-80}	127^{+17}_{-30}	$1.79^{+0.13}_{-0.15}$	0.30
	H β	$18.33^{+0.38}_{-0.60}$	$1.56^{+0.30}_{-0.36}$	-84^{+89}_{-89}	199^{+72}_{-70}	117^{+22}_{-28}	$2.49^{+0.34}_{-0.46}$	0.29
RUBIES EGS 55604	H α	$16.60^{+0.83}_{-1.25}$	$1.75^{+0.18}_{-0.28}$	256^{+33}_{-30}	272^{+21}_{-70}	69^{+12}_{-11}	$1.77^{+0.10}_{-0.10}$	0.50
	H β	$16.91^{+0.63}_{-0.95}$	$1.57^{+0.29}_{-0.33}$	170^{+106}_{-92}	203^{+69}_{-70}	107^{+27}_{-27}	$2.53^{+0.31}_{-0.45}$	0.23
JADES GN 68797	H α	$15.17^{+2.47}_{-0.86}$	$1.47^{+0.45}_{-0.87}$	314^{+22}_{-33}	168^{+46}_{-51}	62^{+19}_{-9}	$1.84^{+0.07}_{-0.06}$	1.88
	H β	$17.44^{+0.33}_{-2.49}$	$1.40^{+0.39}_{-0.52}$	288^{+79}_{-117}	197^{+70}_{-66}	108^{+29}_{-34}	$2.45^{+0.38}_{-0.50}$	0.28
RUBIES EGS 42046	H α	$16.16^{+1.08}_{-1.19}$	$1.46^{+0.23}_{-0.19}$	347^{+28}_{-30}	116^{+25}_{-12}	135^{+10}_{-22}	$1.76^{+0.09}_{-0.08}$	0.73
	H β	$16.89^{+0.88}_{-1.44}$	$1.71^{+0.32}_{-0.40}$	313^{+63}_{-104}	201^{+66}_{-68}	123^{+19}_{-30}	$1.78^{+0.64}_{-0.54}$	0.20
RUBIES UDS 154183	H α	$15.51^{+1.97}_{-1.44}$	$1.20^{+0.46}_{-0.40}$	-53^{+17}_{-21}	272^{+20}_{-31}	105^{+23}_{-25}	$1.05^{+0.10}_{-0.12}$	0.24
	H β	$16.86^{+1.02}_{-1.66}$	$1.41^{+0.40}_{-0.57}$	-170^{+105}_{-110}	206^{+67}_{-73}	107^{+30}_{-34}	$2.05^{+0.62}_{-0.69}$	0.15
GN 15498	H α	$17.62^{+0.88}_{-3.13}$	$1.57^{+0.30}_{-0.37}$	31^{+34}_{-16}	254^{+34}_{-81}	95^{+29}_{-19}	$0.86^{+0.17}_{-0.19}$	0.26
	H β	$17.75^{+0.80}_{-2.71}$	$1.03^{+0.62}_{-0.63}$	-100^{+250}_{-161}	203^{+67}_{-71}	100^{+35}_{-34}	$2.40^{+0.42}_{-0.70}$	0.30
GS 13971	H α	$15.47^{+1.43}_{-0.86}$	$1.51^{+0.31}_{-0.31}$	122^{+24}_{-25}	275^{+18}_{-34}	54^{+6}_{-3}	$1.20^{+0.10}_{-0.10}$	0.33
	H β	$15.41^{+1.37}_{-0.84}$	$1.02^{+0.66}_{-0.67}$	29^{+252}_{-201}	200^{+68}_{-68}	99^{+34}_{-33}	$1.73^{+0.86}_{-0.82}$	0.09
GS 3073	H α	$14.23^{+1.94}_{-0.52}$	$1.33^{+0.50}_{-0.51}$	91^{+27}_{-22}	287^{+10}_{-43}	121^{+16}_{-15}	$1.78^{+0.06}_{-0.07}$	0.47
	H β	$14.28^{+1.65}_{-0.53}$	$1.01^{+0.65}_{-0.65}$	-31^{+141}_{-131}	246^{+39}_{-63}	102^{+33}_{-35}	$1.30^{+0.23}_{-0.23}$	0.18
JADES GN 38147	H α	$15.72^{+1.65}_{-1.75}$	$1.21^{+0.51}_{-0.64}$	251^{+89}_{-120}	139^{+63}_{-29}	86^{+36}_{-25}	$0.79^{+0.19}_{-0.17}$	0.30
	H β	$15.67^{+1.69}_{-1.77}$	$0.99^{+0.68}_{-0.65}$	-2^{+269}_{-272}	205^{+66}_{-71}	99^{+35}_{-34}	$1.69^{+0.89}_{-0.84}$	0.08