Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Cosmic star-formation decline, the first galaxies and early black holes from the energy cascade of dark matter and gas

Zhijie Xu

zhijie.xu@pnnl.gov

zhijie.xu@pnnl.gov https://orcid.org/0000-0003-0459-4531

Article

Keywords:

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11556357/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.

Vol. **000**, 1–21 (2026) Preprint 29 October 2026 Compiled using LaTeX style file v3.0

Cosmic star-formation decline, the first galaxies and early black holes from the energy cascade of dark matter and gas

Zhijie (Jay) Xu,[1]⋆
[1] *Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99354, USA*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Three puzzles in galaxy evolution follow from the energy cascade in haloes and galaxy centres: the steep decline of cosmic star formation since $z \simeq 2$, the excess of UV-bright galaxies at $z > 10$, and over-massive early black holes. Haloes of fixed mass collapse at a fixed density contrast until a crossover at $a_* \simeq 0.5$–0.64, then grow denser. Gas instead radiates its random energy; if it keeps a fixed contrast, the cascade rate shared by all galaxy cores rises as $(1+z)^{5/2}$. Like haloes, galaxies have two levels: cores share a rate and boundaries a contrast, so the rate at the effective radius lies below the core rate by $F \propto M^{-2/3}$, about 1 in spheroids and $\sim 10^2$ in low-mass discs. (i) *Late decline.* Infalling gas reaches the galaxy only if the core radiates its heat fast enough. After a_* this cooling fraction falls as $(1+z)^{5/4}$, steepening the local slope of the star-formation density from 1.5 to 2.7 with no added parameter. (ii) *First galaxies.* Core densities at a given scale are the same in every host, so the threshold for feedback-free star formation is crossed in all haloes at once. A sharp onset at $z \simeq 10.5$ fits the JWST densities and luminosity functions better than bursty or feedback-free alternatives ($\Delta\mathrm{AIC} \geq 42$). (iii) *Black holes.* Luminosity is limited by the pressure of the dense core gas, which rises with redshift, rather than by electron scattering. The resulting ceiling, $L/L_{\rm Edd} \propto (1+z)^2 M_{\rm BH}^{-1/5}$, encloses all but 24 of 1.5×10^5 quasars. Because the core sets the black hole, $M_{\rm BH}/\sigma^5$ stays constant, while dynamical masses at the effective radius of disc-dominated hosts lie a factor $F \simeq 70$–90 below the core level, nearly cancelling the rise of the core rate, $(1+z)^{5/2} \simeq 88$ at $z \simeq 5$. This is why JADES black holes lie 1–2 dex above the local relation with stellar mass but near those with dispersion and dynamical mass. Against local disc hosts measured alike, the JADES hosts lie higher in σ^3/R_e by +1.18 dex at fixed σ and +1.98 dex at fixed dynamical mass, as predicted.

Key words: galaxies: high-redshift – galaxies: evolution – galaxies: star formation – quasars: supermassive black holes – dark matter

1 INTRODUCTION

Three observations describe the growth of galaxies and black holes over cosmic time. The cosmic star-formation density peaks at $z \simeq 1.9$ and falls by a factor of 8.9 to $z = 0$, locally as $(1+z)^{2.7}$ (Madau and Dickinson 2014; Walter *et al.* 2020). At $z > 9$, JWST finds UV luminosity densities that decline only slowly to $z \simeq 12$ and faster beyond (Donnan *et al.* 2024; McLeod *et al.* 2026). And broad-line AGN at $4 < z < 11$ host black holes 1–2 dex above the local relation with stellar mass, yet close to the local relations with velocity dispersion and dynamical mass (Maiolino *et al.* 2024a; Pacucci and Loeb 2024), some radiating far above the Eddington limit (Suh *et al.* 2025; Sun *et al.* 2026).

Each observation poses a question that current models answer only with a fitted threshold. (i) *Why does star formation decline so steeply?* The rate at which baryons accrete onto haloes falls by only a factor of 2.7–4.0 between $z = 1.9$ and 0 (Section 3), so the gas reaching galaxies must fall behind the gas reaching haloes by a further factor of 2–3, while the depletion time of molecular gas stays nearly constant (Decarli *et al.* 2019; Walter *et al.* 2020): the decline must come from gas supply rather than star formation efficiency. Slowing halo growth, feedback, virial shock heating in massive haloes (Dekel and Birnboim 2006) and gas depletion (Tacconi *et al.* 2018) can each be tuned to the curve, but none fixes, without calibration on galaxies, when the extra suppression begins and how steep it is. (ii) *Why are galaxies at $z > 10$ more abundant than expected?* Models whose efficiency depends on halo mass but not on time (Tacchella *et al.* 2018; Harikane *et al.* 2023) reproduce the star-formation density to $z \simeq 10$ but fall short at $z \simeq 10$–13, and, with a fixed light output per unit star formation, predict a faster decline of UV-bright galaxies than observed (Donnan *et al.* 2024). Feedback-free starbursts, bursty star formation, a top-heavy initial mass function, little dust, or progressively younger stellar populations each reproduce the densities, but do not explain why efficient star formation begins near $z \simeq 10$ or in which haloes. Each must also add few ionizing photons (Muñoz *et al.* 2024). (iii) *Why are early black holes so massive and so luminous?*

⋆ E-mail: zhijie.xu@pnnl.gov; zhijiexu@hotmail.com

© 2026 The Authors

UHZ1, $4 \times 10^7 M_\odot$ at $z = 10.1$ (Bogdán *et al.* 2024), is out of reach of Eddington-limited growth from a light seed, and LID-568 radiates at $L/L_{\rm Edd} \simeq 40$ (Suh *et al.* 2025), above slim-disc saturation (Watarai *et al.* 2000). Revised masses may move some objects towards local relations (Rusakov *et al.* 2026; Sun *et al.* 2026), but two questions remain: what limits black-hole luminosity if not the Eddington limit, and why hosts look normal in dispersion and dynamical mass but not in stellar mass.

Why one framework. All three problems depend on how fast gas in galaxy centres radiates its energy and contracts, but each compares that rate against a different benchmark. After $z \simeq 2$ accreted gas cools too slowly compared with the growth of its halo. At $z > 10$ it collapses faster than supernova feedback can respond. Around early black holes it is dense enough to hold in radiation above the Eddington limit. The usual explanations treat the three separately, each with its own fitted threshold: an epoch at which the late decline steepens, a redshift and a halo mass above which the early boost acts, and a cap on the Eddington ratio, a fixed multiple of $L_{\rm Edd}$. If all three are governed by a single gas clock, compared with a halo clock measured in dark-matter simulations, each fitted threshold is replaced by a prediction. The epoch becomes the crossover of the halo clock. The halo mass drops out, because a density threshold is crossed in all haloes at once. The fixed cap becomes a ceiling on the Eddington ratio that scales as $(1+z)^2 M_{\rm BH}^{-1/5}$, rising with redshift. Timing and scalings are then predicted rather than fitted, and the three problems test one another.

This work. We build that clock from the energy cascade, the transfer of kinetic energy from large to small scales at a rate $\varepsilon = \sigma^3/r$ (Xu 2026a, 2025). Three ideas emerge. *Clocks.* Dark-matter haloes of fixed mass collapse at a fixed density contrast until a crossover a_*, when their accumulated virial energy overtakes their bulk-flow energy, and grow denser afterwards. Gas radiates its random energy instead of accumulating it; we hypothesize that gas keeps a fixed contrast at all epochs. The two clocks tick together until a_* and part afterwards. *One rate per epoch.* The cores of galaxies, defined dynamically by their gas cascade rather than photometrically, share a single core cascade rate at a given epoch, $\varepsilon_b(z) \propto (1+z)^{5/2}$, measured in 177 local bulges over four decades in mass. *Levels.* A cascade rate depends on where it is measured: as a halo has a plateau rate in core and a lower virial rate at boundary, a galaxy has its core rate and a rate at its effective radius lower by a structural factor F. At a given epoch cores share one rate, whereas galaxy and halo boundaries share one contrast. Comparing like with like resolves an apparent conflict with the JWST black-hole hosts, which we verify against local hosts measured in the same way.

Applied to the three problems, these ideas give (i) a cooling fraction that falls as $(1+z)^{5/4}$ after a_*, with onset and exponent set by the clocks; (ii) a density threshold crossed in every halo at the same epoch; and (iii) a luminosity ceiling $L/L_{\rm Edd} \propto (1+z)^2 M_{\rm BH}^{-1/5}$, a constant $M_{\rm BH}$–σ relation and a ratio $M_{\rm BH}/M_b$ that evolves with z following the level of the host mass. The clocks fix timing and exponents in (i); the onset density, boost size and switch width in (ii), the ceiling normalization in (iii) and the value of ε_b itself are fitted.

The paper is organized as follows. Section 2 introduces the cascade, its levels and its two clocks, and states and tests the hypothesis that gas has a fixed contrast. Sections 3, 4 and 5 cover the three problems, ordered from strongest to weakest evidence. Section 6 concludes. The appendices give the derivations (Appendices A–C) and the model details, notation, data and code (Appendix D–F).

2 CASCADES, LEVELS AND CLOCKS

2.1 The cascade rate and its two levels

A cascade passes energy through a range of scales without accumulating it, like water through a pipe of changing width: the flow rate is the same in every section, while the speed adjusts to the width. Structures of size r and one-dimensional dispersion σ pass on their specific energy σ^2 within their dynamical time r/σ, so the transfer rate per unit mass is

$$\varepsilon \equiv \frac{\sigma^2}{r/\sigma} = \frac{\sigma^3}{r}. \tag{1}$$

In a steady cascade this flux is the same at every scale, so ε is independent of r and the dispersion adjusts to the scale, $\sigma^2 = (\varepsilon r)^{2/3}$. Velocity scaling alone does not prove a cascade, so Table F2 grades every ingredient by its evidence: a measured energy flux (F), a measured scaling (S), a calibration (C), or a hypothesis tested here (H).

Two levels. For matter in virial equilibrium with the mass m_r inside r, $\sigma^2 = \gamma G m_r/r$, and Eq. (1) becomes a profile,

$$\varepsilon(r) = \left(\tfrac{4\pi}{3}\gamma G \bar\rho_r\right)^{3/2} r^2, \tag{2}$$

with $\bar\rho_r$ the mean density within r. Where $m_r \propto r^{5/3}$ (or $\sigma^2 \propto r^{2/3}$) the profile is flat: this plateau is the *cascade rate* of the structure, set by the energy supplied to it.

At the boundary R of a structure of mass M, the profile gives the *virial rate*, the specific energy of the structure divided by its crossing time:

$$\varepsilon_{\rm v} \equiv \frac{(\gamma GM/R)^{3/2}}{R} = \frac{\sigma^2}{t_{\rm dyn}} = \left(\tfrac{4\pi}{3}\gamma G\bar\rho\right)^{3/2} R^2 \propto a^{-5/2}, \tag{3}$$

where $\bar\rho$ is the mean density within R. At fixed density contrast the mean density scales as $\bar\rho \propto a^{-3}$, and at fixed mass the radius then scales as $R \propto a$, a fixed comoving size. The virial rate, $\sigma^3/R \propto a^{-5/2}$, is therefore set by the cosmic clock alone, for any mass, profile or cosmology (Appendix A3); it applies to dark matter before a_* and, under our hypothesis, to gas at all epochs. Within a structure, $\varepsilon(r)$ falls from the plateau to the boundary by the level ratio F, which for a halo is set by its concentration (Fig. 1). Rates can therefore be compared only at the same level, a rule used throughout the paper.

2.2 Dark matter in haloes

The dark-matter cascade. For the specific peculiar kinetic and potential energies of all dark matter, the Layzer–Irvine equation $dE_y/dt + H(2K_p + P_y) = 0$ (Irvine 1961; Layzer 1963) admits in matter domination the solution

$$K_p = -\varepsilon_u t, \qquad P_y = \tfrac{7}{5}\,\varepsilon_u t, \tag{4}$$

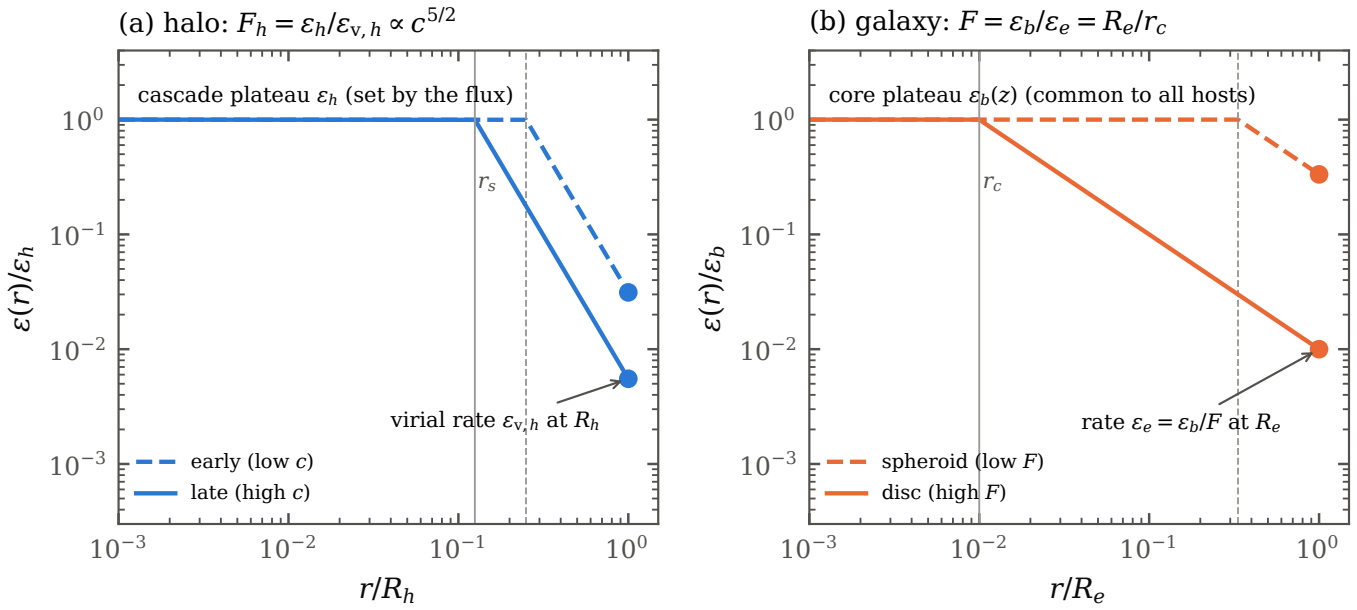


Figure 1. The two levels of a cascade rate (schematic). The rate $\varepsilon(r) = \sigma^3/r$ (Eq. 2) is flat on a plateau and falls beyond its edge (vertical lines) to the value at the boundary (dots); the drop is the level ratio. (a) A dark-matter halo: the plateau ε_h is set by the energy flux $|\varepsilon_u|$. Beyond the scale radius r_s the enclosed mass converges and $\varepsilon \propto r^{-5/2}$, so $F_h = \varepsilon_h/\varepsilon_{v,h}$ grows with concentration as $c^{5/2}$ (Eq. 11); c grows slowly with time (dashed: early, solid: late). (b) A galaxy: the core rate $\varepsilon_b(z)$ is common to all hosts at a given epoch (Eq. 13). Beyond the core radius r_c the dispersion is roughly flat and $\varepsilon \propto r^{-1}$, so the rate at the effective radius is $\varepsilon_e = \varepsilon_b/F$ with $F = R_e/r_c$: of order unity for spheroids (dashed) and 10^2–$10^{2.5}$ for low-mass discs (solid; Section 2.5).

with constant $\varepsilon_u < 0$: kinetic energy is fed to virialized structures at a constant rate. The equation alone does not select this solution; N-body simulations do. They show $K_p \propto t$ and $P_y = -1.4K_p$, and they measure the energy flux directly, through halo mass and radially inside haloes, at $|\varepsilon_u| \simeq (2\text{–}5) \times 10^{-7}\,\mathrm{m^2\,s^{-3}}$ (Xu 2021). Inside a halo of mass m_h the scale-independent cascade (plateau in Fig. 1a) has the rate

$$\varepsilon_h(m_h, z) = |\varepsilon_u| \left(\frac{m_h}{m_h^*}\right)^{2/3}, \tag{5}$$

equal to $|\varepsilon_u|$ in haloes of the characteristic mass m_h^* defined below. The plateau fixes the inner density profile, $\rho \propto \varepsilon_h^{2/3} G^{-1} r^{-4/3}$, inside the scale radius r_s (Fig. 2b).

Two mass scales and the crossover. The velocity of a particle in a halo is the sum of the motion of the halo as a whole, carried by the large-scale bulk flow, and its orbit inside the halo. The first has a dispersion σ_h^2 nearly independent of halo mass; the second has the virial dispersion $\sigma_v^2 \propto m_h^{2/3}$. Small haloes are dominated by the bulk flow, large ones by their orbits, and the two are equal at *characteristic mass*,

$$\sigma_v^2(m_h^*) = \sigma_h^2(m_h^*). \tag{6}$$

A second scale, the *cutoff mass* M^*, marks the exponential tail of the mass function: haloes near M^* hold most of the virial energy and haloes above it are rare. The crossover a_* is the epoch at which the mean virial energy of halo particles overtakes their mean bulk-flow energy, equivalently at which M^* reaches m_h^* (Fig. 2a). Before a_* characteristic haloes are rare peaks that collapse self-similarly, at a fixed contrast. After a_* they sit inside larger structures that are still collapsing, accrete pre-processed matter and keep a fixed comoving size, so haloes of fixed mass grow denser than the background.

The clock exponent. We describe how the internal dynamics of a component of fixed mass scales with the expansion by one exponent θ,

$$\sigma^2 \propto a^{-\theta}, \quad R \propto a^{\theta}, \quad \rho \propto a^{-3\theta}, \quad t_{\rm dyn} \propto a^{3\theta/2}. \tag{7}$$

$\theta = 1$ is collapse at fixed contrast: internal time-scales keep pace with cosmic time, $t \propto a^{3/2}$. $\theta = 1/2$ means contrast growing as $a^{3/2}$. Dark matter has $\theta = 1$ before a_* and $\theta = 1/2$ after (Appendix A1), and in both regimes

$$\sigma_v^2 \propto m_h^{2/3} a^{-\theta}, \qquad R_h \propto m_h^{1/3} a^{\theta}, \qquad m_h^* \propto a^{3\theta}, \tag{8}$$

so $m_h^* \propto a^3$ before a_* and $\propto a^{3/2}$ after. The cutoff mass grows faster, as $M^* \propto a^{15/4}$ before a_* and $\propto a^3$ after it in Illustris-1-Dark (Fig. 2a). Three diagnostics (the energy crossing, M^* reaching m_h^* and the onset of a universal power spectrum $P(k) \propto k^{-1}$) place the crossover at $a_* \simeq 0.5$ in Illustris-1-Dark, with $m_h^*(z=0) = 1.2 \times 10^{13}\mathrm{M}_\odot$ (Xu 2026a) (the Illustris-1-Dark value of Fig. 2a).

The crossover from the primordial spectrum. Before a_* the bulk energy is that of the linear flow, which in matter domination grows as $\langle K_{ph} \rangle \propto A_s c^2 a/a_{\rm eq}$, while the virial energy grows as $|\varepsilon_u| t \propto a^{3/2}$. The faster-growing virial energy catches up at (Appendix A2)

$$a_* = \left[\frac{3}{8}\,\beta_V^2\,\Omega_m^{3/2}\,\frac{A_s}{\Omega_r}\,\frac{c^2 H_0}{|\varepsilon_u|}\right]^2, \tag{9}$$

where β_V, of order unity, absorbs the shape of the transfer function. The crossover is thus fixed by the primordial

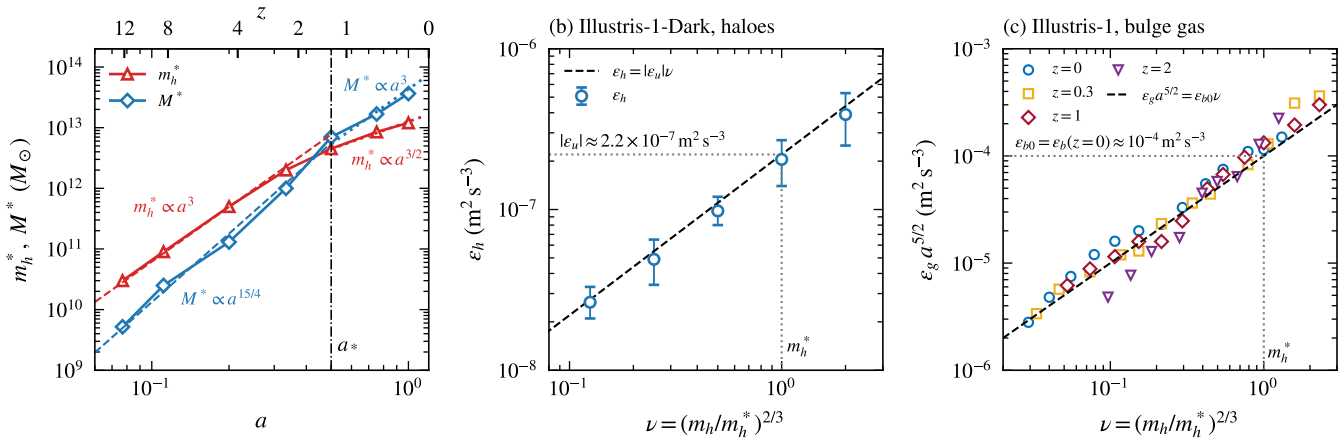


Figure 2. Measurements in Illustris that anchor the clocks. (a) Characteristic mass m_h^* and cutoff mass M^* against a in Illustris-1-Dark. Dashed: $m_h^* \propto a^3$ and $M^* \propto a^{15/4}$ before the crossover $a_* \approx 0.5$ (dash-dotted); dotted: $m_h^* \propto a^{3/2}$ and $M^* \propto a^3$ after it; least-squares slopes of the plotted points: $a^{2.74}$ and $a^{3.59}$ before, $a^{1.27}$ and $a^{2.87}$ after. (b) Plateau rate of dark-matter haloes against $\nu = (m_h/m_h^*)^{2/3}$, following $\varepsilon_h = |\varepsilon_u|\nu$ with $|\varepsilon_u| \approx 2.2 \times 10^{-7}\,\mathrm{m^2\,s^{-3}}$ (the value labelled in panel b; a least-squares fit to the five plotted points gives $2.0 \times 10^{-7}\,\mathrm{m^2\,s^{-3}}$). (c) Cascade rate of halo gas in Illustris-1 ($r \simeq$ 1–10 kpc), scaled as $\varepsilon_g a^{5/2}$, against ν at $z = 0$–2: all redshifts collapse onto $\varepsilon_g a^{5/2} = \varepsilon_{b0}\nu$ with $\varepsilon_{b0} \approx 10^{-4}\,\mathrm{m^2\,s^{-3}}$ (Eq. 12); fit at $z \le 2$: $1.08 \times 10^{-4}\,\mathrm{m^2\,s^{-3}}$, scatter 0.12 dex; the $z = 4$ points of the author's data file, not plotted, lie 0.4 dex lower. Dotted lines in (b) and (c) mark the rates at the characteristic scale m_h^*.

amplitude A_s, the radiation content that sets the equality epoch, and the cascade rate. For Planck 2018 parameters (Planck Collaboration *et al.* 2020), $|\varepsilon_u| = 3.7 \times 10^{-7}\,\mathrm{m^2\,s^{-3}}$ and $\beta_V = 1$, it gives $a_* = 0.64$ ($z_* = 0.56$); evaluating both energies in ΛCDM instead of matter domination lowers this to 0.54. We adopt $a_* = 0.64$ as fiducial and $a_* = 0.50$, measured in Illustris-1-Dark, as the alternative, and in Section 3 infer a_* from galaxies.

Two halo levels. A halo has two levels: the plateau ε_h inside the scale radius r_s, and the virial rate $\varepsilon_{\mathrm{v},h} = V_h^3/R_h$ at the boundary, with $V_h^2 = Gm_h/R_h$. At fixed halo mass the two evolve differently. The plateau follows the cascade flux through $m_h^* \propto a^{3\theta}$ (Eqs 5 and 8). The boundary follows the clock, $V_h^2 \propto a^{-\theta}$ and $R_h \propto a^{\theta}$:

$$\varepsilon_{\mathrm{v},h} = \frac{(V_h^2)^{3/2}}{R_h} \propto \frac{a^{-3\theta/2}}{a^{\theta}} = a^{-5\theta/2}. \tag{10}$$

Both rates scale as $m_h^{2/3}$, so their ratio depends on time but not on mass. The ratio is set by the concentration $c \equiv R_h/r_s$. The cascade mass budget gives $\varepsilon_h \propto (Gm_h)^{3/2} r_s^{-5/2}$, which is $\varepsilon_{\mathrm{v},h}$ with r_s in place of R_h (Appendix A3). Hence

$$F_h \equiv \frac{\varepsilon_h}{\varepsilon_{\mathrm{v},h}} \propto c^{5/2} \propto a^{\theta/2}, \qquad c \equiv \frac{R_h}{r_s} \propto a^{\theta/5}, \tag{11}$$

where F_h is the level ratio. The virial rate at boundary falls faster than the plateau. The halo makes up the difference by slowly becoming more concentrated, as $a^{1/5}$ before a_* and $a^{1/10}$ after, at every mass.

2.3 Gas in galaxy centres: one rate for all cores

Halo gas. In Illustris-1, gas in haloes has $\sigma_g^2(r) \propto r^{2/3}$ over $r \simeq$ 1–10 kpc, a scale-independent rate σ_g^3/r, and is in virial equilibrium with the enclosed mass of all components (Xu 2025). Its cascade rate depends on host mass and redshift,

$$\varepsilon_g(m_h, z) = \varepsilon_b(z) \left(\frac{m_h}{m_h^*}\right)^{2/3}, \quad \varepsilon_b(z) = \varepsilon_{b0}(1+z)^{5/2}, \tag{12}$$

with $\varepsilon_{b0} \simeq 10^{-4}\,\mathrm{m^2\,s^{-3}}$ (Fig. 2c). The two factors of Eq. (12) come from the two ends of the halo-gas plateau. At its outer end the gas sits in the potential of its halo and is driven in proportion to the depth of that halo relative to a characteristic one, $\varepsilon_h/|\varepsilon_u| = (m_h/m_h^*)^{2/3}$ (Eq. 5). At its inner end it joins the core. The mass dependence thus comes from the dark matter and the normalization from the core, $\varepsilon_b \propto (1+z)^{5/2}$. Equivalently, $\varepsilon_g/\varepsilon_h = \varepsilon_b/|\varepsilon_u|$: at a given epoch, dissipation raises the gas plateau above the dark-matter plateau by the same factor in every halo.

Bulges: one rate for all hosts (H2). Inside the halo gas lies the baryon-dominated core of the galaxy. By *core* we mean the region in which the gas cascade runs at a single rate ε_b, of radius r_c ($\simeq R_b$ for a bulge). This is a dynamical definition, unrelated to the photometric cores of core-Sérsic galaxies (10–500 pc) or to cored dark-matter profiles. Its size grows with mass, $r_c \propto M^{3/5}$ at fixed ε_b: from tens of pc in low-mass discs to several kpc in massive spheroids, where the core extends to the effective radius. Cores behave differently from the halo gas. For the 177 local galaxies of Table B1 of Xu (2025), the bulge dispersion σ_b and bulge mass M_b are compiled from the literature, and the bulge radius is the virial radius they imply, $R_b \equiv GM_b/(3\sigma_b^2)$ (Marconi and Hunt 2003), so that $\gamma_r = 1/3$ by construction. The virial rate, $\varepsilon_{\mathrm{v},b} = \sigma_b^3/R_b = 3\sigma_b^5/(GM_b)$, is log-normal with mean $10^{-4.14}\,\mathrm{m^2\,s^{-3}}$ and dispersion 0.52 dex, and over 3.9 dex in bulge mass it does not rise with mass: $\varepsilon_{\mathrm{v},b} \propto M_b^{-0.09\pm0.05}$, whereas the halo-gas rate would scale as $M_b^{2/3}$ if $M_b \propto m_h$ (Fig. 3a,b).

The rate does not depend on host type either: Seyferts,

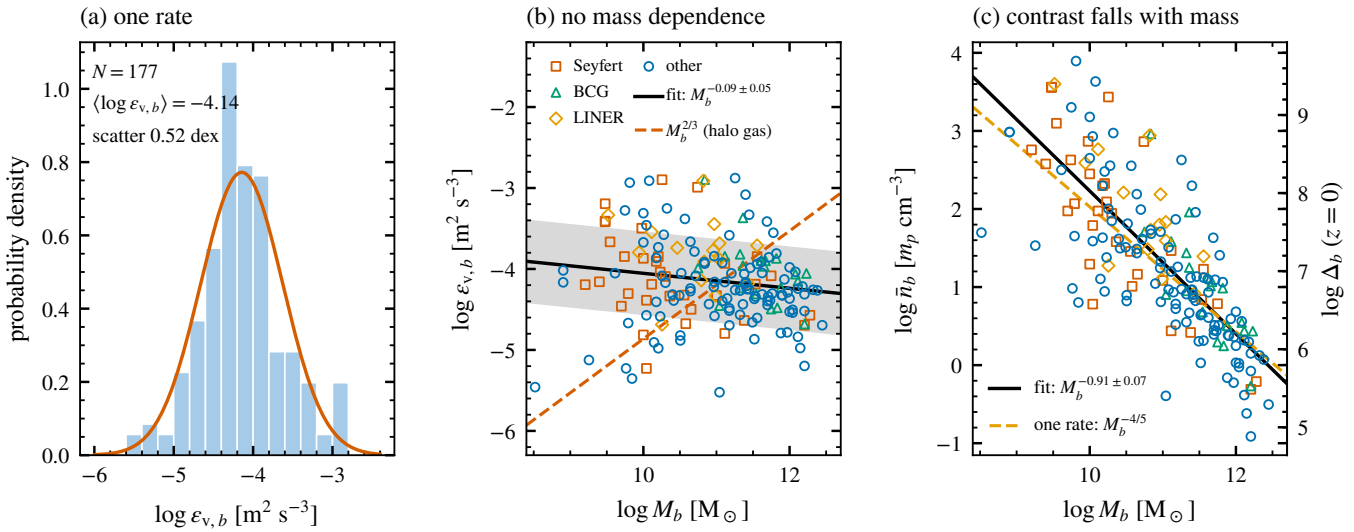


Figure 3. One cascade rate for all bulges: the 177 local galaxies of Table B1 of Xu (2025). (a) Distribution of the bulge virial rate $\varepsilon_{\mathrm{v},b} = \sigma_b^3/R_b$ with a Gaussian of the same mean and dispersion in $\log \varepsilon_{\mathrm{v},b}$. (b) $\varepsilon_{\mathrm{v},b}$ against bulge (dynamical) mass for Seyferts, brightest cluster galaxies, LINERs and other hosts. Black: least-squares fit with bootstrap error; band: ± the dispersion; red dashed: the halo-gas scaling $M_b^{2/3}$ normalized at the median. (c) Mean density within R_b (right axis: contrast Δ_b over the mean matter density at $z = 0$) against M_b. Black: fit; orange dashed: $M_b^{-4/5}$, the scaling required by one common rate (Eq. A10).

brightest cluster galaxies and other hosts have median rates within 0.25 dex of one another, and the 15 LINERs lie 0.4 dex higher. A rate shared by all bulges fixes their structure. Through the mass budget $M_b = \gamma_r^{-1}\varepsilon_b^{2/3}G^{-1}R_b^{5/3}$ it requires $\sigma_b \propto M_b^{1/5}$, $R_b \propto M_b^{3/5}$ and $\bar\rho_b \propto M_b^{-4/5}$. A common contrast would instead give $M_b^{1/3}$, $M_b^{1/3}$ and M_b^0. Table B1 gives $\sigma_b \propto M_b^{0.18\pm0.01}$, close to the common rate. Since $R_b = GM_b/(3\sigma_b^2)$, the sizes and densities follow from this one slope, $R_b \propto M_b^{0.64\pm0.02}$ and $\bar\rho_b \propto M_b^{-0.91\pm0.07}$ (Fig. 3c), and are not separate tests. Bulges of different mass are thus different cuts through one profile, $\rho \propto \varepsilon_b^{2/3}G^{-1}r^{-4/3}$, and at a given epoch (Fig. 1 for $F = 1$)

$$\varepsilon_{\mathrm{v},b} \simeq \varepsilon_b(z) \qquad \text{in every host.} \tag{13}$$

We call this the common-rate hypothesis, H2. It is measured at $z = 0$ for bulges and cores; it does *not* hold for whole galaxies at their effective radius R_e (Section 2.5). The rate where the bulge meets the halo-gas plateau, in characteristic haloes, is the Illustris value ε_{b0}; the two determinations, 10^{-4} and $10^{-4.14}\,\mathrm{m^2\,s^{-3}}$, agree within the scatter.

Two nested cascades. Each galaxy therefore holds two cascades: the halo gas on 1–10 kpc scales, with a rate $\varepsilon_g \propto m_h^{2/3}$ set by its halo, and the core inside it, with the common rate $\varepsilon_b(z)$. Their ratio, $\varepsilon_g/\varepsilon_{\mathrm{v},b} = \nu \equiv (m_h/m_h^*)^{2/3}$, is unity in characteristic haloes of mass m_h^*, where one cascade runs from the halo gas into the centre. At fixed host mass ν grows towards high redshift as m_h^* falls. What sets the common value of ε_b is not derived here. The working assumption is that gas contracts dissipatively inside the halo gas plateau until its rate reaches a value fixed by the background density, through fixed contrast, rather than by the host.

2.4 Fixed contrast in time (H1)

Why gas should not cross over. Dark matter changes clock at a_* because, being collisionless, it cannot radiate: the random energy delivered by the cascade accumulates (Eq. 4) until it exceeds the bulk energy. Gas also receives random energy from the cascade but radiates it within about a dynamical time, so energy flows through the gas instead of accumulating. We therefore hypothesize that each core keeps its density contrast in time (H1). A core of fixed mass $M_b = \frac{4\pi}{3}\Delta_b\bar\rho R_b^3$ with constant contrast Δ_b has $R_b \propto a$, $\sigma_b^2 \propto a^{-1}$, $\theta_b = 1$, and a virial rate (Appendix B1)

$$\varepsilon_{\mathrm{v},b} = \gamma_r^{3/2}G^{3/2}M_b^{2/3}\left(\tfrac{4\pi}{3}\Delta_b\bar\rho\right)^{5/6} \propto (1+z)^{5/2} \tag{14}$$

in any cosmology. Here M_b is the total gravitating mass within R_b.

H1 and H2 are distinct and compatible. H1 concerns time: a given core keeps its contrast Δ_b. H2 concerns hosts at one epoch: all cores share one rate. Neither implies the other. Since $\varepsilon_{\mathrm{v}} \propto M^{2/3}\Delta^{5/6}$ (Eq. 14), H2 requires the contrast to fall with mass as $\Delta_b \propto M_b^{-4/5}$, which Table B1 shows (Fig. 3c); a contrast independent of mass would instead give $\varepsilon_{\mathrm{v},b} \propto M_b^{2/3}$, excluded by 2.6 dex (Fig. 3b), the rise $M_b^{2/3}$ predicts across the 3.9 dex (3.0 dex relative to the fitted slope). Both H1 and H2 can hold if $\Delta_b = C\,M_b^{-4/5}$ with C fixed in time, and the common rate then evolves as $\varepsilon_b \propto (1+z)^{5/2}$.

Cores share a rate, boundaries share a contrast. H1 applies to every level of a gas structure, so the core and the boundary at R_e both keep their contrast and their ratio F is fixed in time. Across masses the two levels differ. Cores share a rate, so their contrast falls with mass. Galaxy boundaries, like halo boundaries, share a contrast: their mean density is the same multiple of the cosmic mean at every mass.

For the 46 local disc hosts of Appendix D, the mean density within R_e scales as $M^{-0.04\pm0.09}$ over $10^{8.6}$–$10^{11.6}\mathrm{M}_\odot$, at $\Delta \simeq 10^{6.6}$, with $R_e \propto M^{0.35\pm0.03}$ and $\sigma_e \propto M^{0.33\pm0.02}$, close to the $M^{1/3}$ of a common contrast. The rate at R_e therefore rises with mass, $\varepsilon_e \propto M^{0.63\pm0.08}$, against $M^{2/3}$ expected, so $F = \varepsilon_b/\varepsilon_e \propto M^{-2/3}$. This is the size–virial radius relation, $R_e \propto R_{200}$ at $0 < z < 3$ (Kravtsov 2013; Huang *et al.* 2017), read as a contrast: with $M(< R_e) \simeq 0.02\, m_h$, the disc boundary inherits the mass-independent contrast of its halo. Spheroids, whose core extends to R_e, follow the core instead: their contrast scales as $M^{-0.78\pm0.09}$ for the 100 early-type hosts of the same compilation. Table F3 summarizes both cascades at both levels.

A typical core. For $\varepsilon_{\mathrm{v},b} = 10^{-4}\,\mathrm{m^2\,s^{-3}}$, $M_b = 10^{11}\mathrm{M}_\odot$ and $\gamma_r = 1/3$, the core has $R_b = 3.1$ kpc, $\sigma_b = 213\,\mathrm{km\,s^{-1}}$ and $t_b = R_b/\sigma_b = 14$ Myr (22 Myr for $\gamma_r = 1$). Its mean density, $31\, m_p\,\mathrm{cm^{-3}}$, is a contrast $\Delta_b \simeq 2\times10^7$, five orders of magnitude above dark matter haloes: gas can radiate and contract, while dark matter cannot.

Is fixed contrast allowed? Two checks show that it is, though not that it is required. Both test the balance hypothesis of "cosmic quenching" in Xu (2025): gas in virial equilibrium cools within about a crossing time, $\sigma^2/\varepsilon = r/\sigma$, and so radiates its cascade input where it arrives, at the rate $\varepsilon_{\mathrm{v},b}$. This is also why gas, unlike dark matter, does not accumulate random energy and has no crossover.

Energy budget. For net energy loss Γ per unit mass (radiated minus received), the Layzer–Irvine equation is $dE_y/dt + H(2K_p + P_y) = -\Gamma$ (Appendix B2). Virial equilibrium, $2K_p + P_y = 0$, removes the expansion term and gives $E_y = -K_p$. For energies $\propto a^x$, $dE_y/dt = xHE_y$, so

$$\Gamma = x\, H\, K_p. \tag{15}$$

Exact balance ($\Gamma = 0$) means $x = 0$: a core of fixed physical size and constant rate. A core at fixed contrast ($x = -1$) expands with the universe and loosens, so it must gain a little energy with $\Gamma < 0$. With $K_p = \frac{3}{2}\sigma_b^2$ and $t_b = R_b/\sigma_b$,

$$\frac{|\Gamma|}{\varepsilon_{\mathrm{v},b}} = \tfrac{3}{2}\, H t_b = (1.5\text{–}2.3)\times10^{-3}\,: \tag{16}$$

radiation removes all but 0.2 percent of the cascade input, which is Γ. Balance thus sets the level of ε_b, and a residual of order Ht_b sets its clock x. Since $Ht_b \ll 1$, any clock with $|x| \sim 1$ allows fixed contrast but does not require it.

Cooling. Gas at temperature T and hydrogen density n_H radiates a power per unit volume $\propto n_\mathrm{H}^2\Lambda(T)$, where $\Lambda(T)$ is the cooling function of collisional-ionization equilibrium (Sutherland and Dopita 1993). Per unit gas mass the radiative rate is therefore $\propto n_\mathrm{H}\Lambda(T)$ (Eq. B5). If the cascade is balanced by radiation, this rate must equal $\varepsilon_{\mathrm{v},b}$. In value: at $\sigma_b = 210\,\mathrm{km\,s^{-1}}$ ($kT = 0.28$ keV, $\Lambda = 1.7\times10^{-23}\,\mathrm{erg\,cm^3\,s^{-1}}$ at $Z = 0.3Z_\odot$), radiating $10^{-4}\,\mathrm{m^2\,s^{-3}}$ requires $n_\mathrm{H} \simeq 0.11\,\mathrm{cm^{-3}}$, typical of hot gas in the inner few kpc of massive ellipticals (Mathews and Brighenti 2003). In evolution: at fixed contrast, $n_\mathrm{H} \propto (1+z)^3$ and $T \propto (1+z)$, so the radiative rate scales as $(1+z)^{3+s}$, with $s = d\ln\Lambda/d\ln T$. Matching $(1+z)^{5/2}$ requires $s \simeq -1/2$ (Appendix B3), which is close to the slope at bulge virial temperatures (Fig. B1): for $\sigma_b = 150$–$300\,\mathrm{km\,s^{-1}}$, s runs from -1.2 to -0.1, and over $0 < z < 3$ the radiative rate evolves as $(1+z)^{2.3}$–$(1+z)^{3.0}$ for $\sigma_{b0} = 150$–$250\,\mathrm{km\,s^{-1}}$, bracketing 5/2. Cooler gas ($\sigma_{b0} \le 120\,\mathrm{km\,s^{-1}}$, $s \le -1.5$) cannot stay in balance, so the check applies to hot core gas, the gas that matters for cooling in massive haloes.

Comoving permanence. With a common rate the profile of every core is $\rho \propto \varepsilon_b^{2/3} G^{-1} r^{-4/3}$, and with $\varepsilon_b \propto a^{-5/2}$

$$\frac{\rho}{\bar\rho} \propto a^3\, a^{-5/3}\, r^{-4/3} = \left(\frac{r}{a}\right)^{-4/3}\,: \tag{17}$$

plotted as contrast against comoving radius, the inner profile of a core does not change with time and is the same in every host. Dark matter at m_h^* has the same $-4/3$ profile fixed in physical radius instead (Xu 2023, 2026a): gas has comoving permanence where dark matter has physical permanence. Table F3 collects the clocks and the levels.

2.5 Levels inside galaxies

The rate at the effective radius. Galaxies have levels, as haloes do. The common rate of H2 belongs to the core, but most galaxies, and all disc-dominated ones, extend far beyond it. At the effective radius the rate is $\varepsilon_e = \sigma_e^3/R_e$, with σ_e the dispersion within R_e. Outside the core, rotation curves and stellar dispersions are roughly flat, so $\varepsilon \propto r^{-1}$ and $m_r \propto r$. The virial rate at R_e and ratio of levels F

$$\varepsilon_e = \frac{\varepsilon_b(z)}{F}, \qquad F \simeq \frac{R_e}{r_c} \simeq \frac{M_\mathrm{dyn}}{M_\mathrm{core}}, \tag{18}$$

with M_dyn and M_core the dynamical masses within R_e and r_c. In a spheroid the core extends to R_e and $F \simeq 1$; this is the level Table B1 measures. A disc has $F \gg 1$. What sets the ratio F is the contrast at each level, $\Delta = 3M/(4\pi R^3\bar\rho)$, through $\varepsilon_\mathrm{v} \propto M^{2/3}\Delta^{5/6}$ (Eq. 14). Cores share a rate, so their contrast falls as $M_b^{-4/5}$. Disc boundaries share a contrast, $\Delta_e \simeq 10^{6.6}$ at every mass (Section 2.4), so $\varepsilon_e \propto M^{2/3}$ and $F \propto M^{-2/3}$. Under H1 both contrasts are fixed in time, so ε_e and ε_b both rise as $(1+z)^{5/2}$ and F is redshift independent.

Calibration. We calibrate F for local star-forming galaxies from integrated ionized-gas dispersions (Barat *et al.* 2019) and late-type sizes (van der Wel *et al.* 2014; Shen *et al.* 2003) (Appendix D): $F = 135$–400 at stellar mass $M_* = 10^{7.5}$–$10^9\mathrm{M}_\odot$ (lower limits, because the dispersions there sit at the instrumental floor), 51–147 at $10^{9.5}\mathrm{M}_\odot$, 15–37 at $10^{10}\mathrm{M}_\odot$ and 5–9 at $10^{10.5}\mathrm{M}_\odot$. Low-mass discs thus sit 10^2–$10^{2.5}$ below the core rate. Stellar dispersions within R_e of 46 local disc hosts (Appendix D) give $F \simeq 102$, 47, 21 and 10 at $M_* = 10^9$, $10^{9.5}$, 10^{10} and $10^{10.5}\mathrm{M}_\odot$: the two calibrations agree below $10^{9.5}\mathrm{M}_\odot$. The tests below compare like with like and do not depend on this. Compared at the same level, star-forming galaxies at $z > 1$ follow fixed contrast (Fig. 4; Table 1): at fixed stellar mass ε_e rises by $+1.10 \pm 0.12$ dex for extreme emission-line galaxies at $z \simeq 1.7$ (Maseda *et al.* 2014) (fixed contrast: +1.08) and by $+1.98 \pm 0.12$ dex for the JADES broad-line AGN hosts at $z \simeq 5.2$ (Maiolino *et al.* 2024a) (+1.98); non-AGN galaxies at $z \simeq 6.2$ (de Graaff *et al.* 2024) rise by $+1.46 \pm 0.19$ dex, short of +2.14. Together the non-AGN samples give $\varepsilon_e \propto (1+z)^{2.08\pm0.19}$ at fixed M_*, close to fixed contrast and far below the halo-gas law, Eq. (12).

A local test at the JADES mass. We compare the JADES hosts with 18 local disc and S0 hosts of broad-line AGN at the same mass (Appendix D), measured the same way: dispersion, whole-galaxy R_e without the nucleus, and $M = 5\sigma^2 R_e/G$. At fixed dispersion ($\sigma = 60$–$100\,\mathrm{km\,s^{-1}}$, nine local discs) the JADES hosts lie $+1.18 \pm 0.12$ dex higher in σ^3/R_e, against $\frac{3}{2}\log(1+z) = +1.19$ for a contrast common in time and across masses (at fixed contrast $\varepsilon_e \propto (1+z)^{3/2}\sigma^2$); at fixed dynamical mass they lie $+1.98 \pm 0.13$ dex higher, against $+1.98$. Equivalently, the local hosts have $F = 68$ (20–90), the value the clock requires of the JADES hosts.

Consequence. At $z \simeq 5$ the core rate is $(1+z)^{5/2} \simeq 88$ times its local value, and the rate at R_e of a low-mass disc lies a factor $F \simeq 70$–90 below its core. The two nearly cancel, so such a host has the σ^3/R_e of a local bulge. This is not evidence against the clock: it is a rise in time offset by a fall from core to disc, the same fall seen in local discs of that mass. The cores themselves (tens of pc) are unresolved at $z > 4$, so there the core rate is tested only through black holes (Section 5.2).

2.6 Six rates in dark matter and gas

The two cascades define six rates, each tied to a level (Tables 2 and F3). *Dark matter.* The cascade flux $|\varepsilon_u|$ is constant in time. It sets the halo plateau, $\varepsilon_h = |\varepsilon_u|(m_h/m_h^*)^{2/3}$, which at fixed halo mass m_h falls as $a^{-2\theta}$ because m_h^* grows. The halo boundary lies below the plateau by the level ratio $F_h \propto c^{5/2}$, which grows with time but not with mass, so its virial rate $\varepsilon_{\mathrm{v},h} = V_h^3/R_h = \varepsilon_h/F_h$ falls as $a^{-5\theta/2}$.

Gas. The core rate $\varepsilon_b = \varepsilon_{b0}(1+z)^{5/2}$ is the same in all hosts at a given epoch. Its value, $\varepsilon_{b0} \simeq 10^{-4}\,\mathrm{m^2\,s^{-3}}$, is measured in local bulges and matches radiative cooling of hot core gas at the observed densities (Section 2.4). It equals the core virial rate $\varepsilon_{\mathrm{v},b} = \sigma_b^3/R_b$, because the core boundary lies at the edge of its plateau. The galaxy boundary lies below the core by the level ratio $F \propto M^{-2/3}$, fixed in time, so $\varepsilon_e = \sigma_e^3/R_e = \varepsilon_b/F$. On 1–10 kpc scales, the gas of the halo cascades at $\varepsilon_g = \varepsilon_b(m_h/m_h^*)^{2/3}$, taking its mass dependence from the halo and its normalization from the core, $\varepsilon_g/\varepsilon_h = \varepsilon_b/|\varepsilon_u|$ (Fig. 2c). Since ε_g is measured from all gas within r, including the galaxy's own, the two rates overlap where R_e lies in this range: at $z \simeq 0$ the disc level lies a factor of about 4 below the halo-gas plateau, $\log \varepsilon_e/\varepsilon_g = -0.64 \pm 0.07$ (Appendix D).

The two outer boundaries share a contrast across masses, so $\varepsilon_{\mathrm{v},h}$ and ε_e both scale as $M^{2/3}$; the core shares a rate and does not depend on mass.

Each application below reads one level. Late quenching compares the two boundary rates, ε_e and $\varepsilon_{\mathrm{v},h}$. The onset of efficient star formation and the luminosity ceiling read the core rate ε_b. The host relations read ε_b or ε_e, according to the radius at which the host mass is measured. The halo-gas rate ε_g links the two cascades but is not used by the applications.

3 THE LATE DECLINE OF STAR FORMATION

The first application compares the dark-matter and gas clocks after they part, $a > a_*$.

Cooling fraction. Gas accreted by a halo is heated by infall to a specific energy $\sim V_h^2$, which it must radiate within the halo dynamical time $t_h = R_h/V_h$ to reach the galaxy. The specific power this requires is $V_h^2/t_h = V_h^3/R_h = \varepsilon_{\mathrm{v},h}$, the virial rate of the halo. The power available is the rate at which the galaxy's gas cascade radiates, measured where accreted gas joins the galaxy: its boundary rate $\varepsilon_e = \sigma_e^3/R_e$. Both rates are boundary rates, and both boundaries share a contrast across masses (Section 2.5). We take their ratio as the cooling fraction (Appendix B4):

$$f_{\mathrm{cool}} = \frac{\varepsilon_e}{\varepsilon_{\mathrm{v},h}} \propto \left(\frac{M}{m_h}\right)^{2/3} \left(\frac{\Delta_e}{\Delta_h}\right)^{5/6} \propto a^{-\frac{5}{2}(1-\theta)}. \tag{19}$$

Since $M(< R_e) \simeq 0.02\,m_h$ at all masses (Appendix D) and Δ_e is the same at all masses, f_{cool} is the same for all haloes at a given epoch.

Before a_* ($\theta = 1$) halo and galaxy both keep fixed contrasts, and f_{cool} is constant. After a_* ($\theta = 1/2$) the halo's contrast grows as $\Delta_h \propto a^{3/2}$ while the galaxy's stays fixed: the halo becomes denser relative to the galaxy it feeds. Accreted gas then has more energy to lose, $V_h^2/\sigma_e^2 \propto a^{1/2}$, and less time to lose it, $t_h/t_e \propto a^{-3/4}$, so $f_{\mathrm{cool}} \propto (1+z)^{5/4}$. Normalized to unity at a_*,

$$Q(z) = \min\left[1, \left(\frac{1+z}{1+z_*}\right)^q\right], \qquad q = \tfrac{5}{2}(1-\theta) = \tfrac{5}{4}. \tag{20}$$

Q has no parameter fitted to galaxies: z_* is the crossover measured in N-body simulations, and q follows from the two clock exponents.

Star-formation density. We use the standard halo-accretion model, in which a halo forms stars at an efficiency f_* times its baryonic inflow $f_b\dot{m}_h$ (Tacchella *et al.* 2018; Harikane *et al.* 2018),

$$\rho_{\mathrm{SFR}}(z) = f_b \int_{m_{\mathrm{min}}} f_*(m_h, z)\, \dot{m}_h\, \frac{dn}{d\ln m_h}\, d\ln m_h, \tag{21}$$

where $f_b = \Omega_b/\Omega_m$, m_{min} is the atomic-cooling mass (Barkana and Loeb 2001), $\dot{m}_h$ the mean accretion rate of Fakhouri *et al.* (2010) and $dn/d\ln m_h$ the Sheth and Tormen (1999) mass function (Appendix E). The efficiency is usually a double power law in halo mass,

$$f_*^{\mathrm{std}} = \frac{2f_{\mathrm{pk}}}{(m_h/M_c)^{-\beta_*} + (m_h/M_c)^{\gamma_*}}. \tag{22}$$

Harikane *et al.* (2018) find $\beta_* = 1.1$, $\gamma_* = 0.3$ and $M_c = 10^{11.35}\mathrm{M_\odot}$ at $z \simeq 4$–7, with $SFR/\dot{m}_h = 0.017$ at the peak, i.e. $f_{\mathrm{pk}} \simeq 0.11$ of the baryonic inflow. The efficiency rises at low mass, where stellar feedback expels gas from shallow potentials, and falls at high mass, where shock heating and active nuclei keep gas hot (Dekel and Birnboim 2006). It depends on halo mass alone (Tacchella *et al.* 2018; Harikane *et al.* 2018), or, to match the efficiency at $z \simeq 0$, acquires a factor $(1+z)^{-0.6}$ at all epochs (Harikane *et al.* 2018).

We adopt a form that rises and falls in the same way, with different slopes,

$$f_* = \frac{\min[f_0\,(m_h/10^{12}\mathrm{M_\odot})^{2/3},\, f_{\mathrm{max}}]}{1 + m_h/M_q}\, Q(z). \tag{23}$$

Its low-mass slope, 2/3, is not fitted: it is the mass dependence of the halo cascade rate, and at fixed redshift it coincides with energy-driven feedback, $f_* \propto v_c^2$. It saturates at

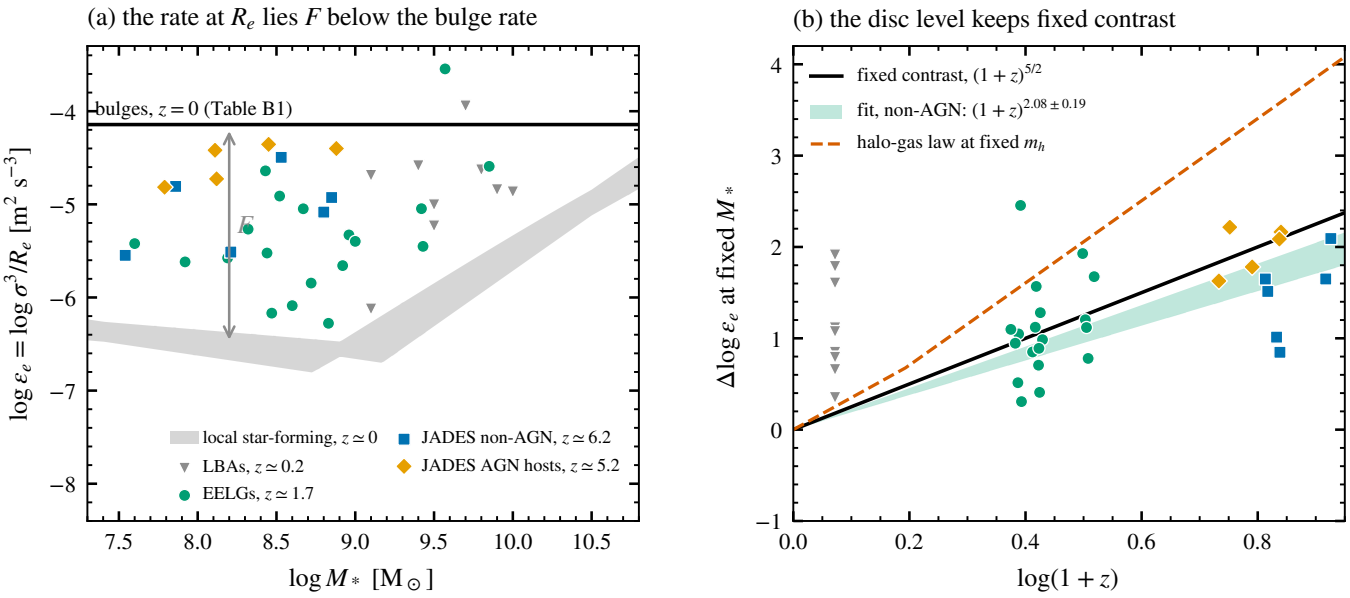


Figure 4. The rate at the effective radius. (a) $\varepsilon_e = \sigma_e^3/R_e$ against stellar mass, with integrated ionized-gas dispersions: local star-forming galaxies (grey band: range over two dispersion apertures and two size relations; Appendix D), local Lyman-break analogues (Gonçalves *et al.* 2010), extreme emission-line galaxies at $z \simeq 1.7$ (Maseda *et al.* 2014), JADES non-AGN galaxies at $z \simeq 6.2$ (de Graaff *et al.* 2024) and JADES broad-line AGN hosts at $z \simeq 5.2$ (Maiolino *et al.* 2024a). Black: mean bulge rate at $z = 0$ (Table B1); the arrow marks F for a local disc of $10^{8.2}$M$_\odot$. (b) Offset of ε_e from local galaxies of the same stellar mass against redshift. Black: fixed contrast, $(1+z)^{5/2}$; band: fit through the origin to the non-AGN samples; red dashed: the halo-gas law at fixed halo mass ($a_* = 0.64$).

Table 1. Offsets of the rate at R_e from local star-forming galaxies of the same stellar mass (Δ_{M_*}) or dynamical mass ($\Delta_{M_{\rm dyn}}$, same estimator at all redshifts), in dex, with standard errors; predictions for fixed contrast and for the halo-gas law. References: [a]Maseda *et al.* (2014); [b]de Graaff *et al.* (2024); [c]Maiolino *et al.* (2024a); [d]Gonçalves *et al.* (2010).

Sample	N	$\bar{z}$	Δ_{M_*}	$\Delta_{M_{\rm dyn}}$	$\frac{5}{2}\log(1+z)$	halo gas
Lyman-break analogues[d]	10	0.18	$+1.09 \pm 0.16$	$+1.23 \pm 0.15$	$+0.18$	$+0.25$
EELGs, $z \simeq 1.4$–2.3[a]	19	1.72	$+1.10 \pm 0.12$	$+0.81 \pm 0.11$	$+1.08$	$+1.75$
JADES, non-AGN[b]	6	6.24	$+1.46 \pm 0.19$	$+1.08 \pm 0.18$	$+2.14$	$+3.66$
JADES AGN hosts[c]	5	5.20	$+1.98 \pm 0.12$	$+1.79 \pm 0.07$	$+1.98$	$+3.36$

Table 2. The six cascade rates (m^2 s^{-3}): dark matter above the line, gas below. $\theta = 1$ before a_* and 1/2 after (Eq. 8); $m_h^* \propto a^{3\theta}$. Level ratios: $F_h \propto c^{5/2}$ (time, not mass), $F \propto M^{-2/3}$ (mass, not time). Evidence codes as in Table F2.

Rate	Meaning	Expression	Evidence
$\lvert\varepsilon_u\rvert$	cascade flux	(2–5) $\times 10^{-7}$, constant; 3.7×10^{-7} adopted	F
ε_h	halo plateau	$\lvert\varepsilon_u\rvert(m_h/m_h^*)^{2/3} \propto m_h^{2/3} a^{-2\theta}$	F, S
$\varepsilon_{{\rm v},h}$	halo boundary, R_h	$\varepsilon_h/F_h \propto m_h^{2/3} a^{-5\theta/2}$	S
ε_g	halo gas, 1–10 kpc	$\varepsilon_b(m_h/m_h^*)^{2/3} \propto m_h^{2/3} a^{-5/2-2\theta}$	C ($z \lesssim 2$)
ε_b	galaxy core, $= \varepsilon_{{\rm v},b}$	$\varepsilon_{b0}(1+z)^{5/2} \propto a^{-5/2}$; $\varepsilon_{b0} \simeq 10^{-4}$	S; H
ε_e	galaxy boundary, R_e	$\varepsilon_b/F \propto M^{2/3} a^{-5/2}$	S; H

$f_{\rm max}$ and falls as m_h^{-1} above the quenching mass M_q. The essential difference is $Q(z)$. The three parameters are fitted by least squares in $\log \rho_{\rm SFR}$ to the curve of Madau and Dickinson (2014) (MD14) at 81 redshifts in $0 \le z \le 8$; the fit of the same functional form by Walter *et al.* (2020) (W20) serves as a systematic. For $a_* = 0.64$ the fit gives $f_0 = 0.27$, $M_q = 7.2 \times 10^{12}$M$_\odot$ and $f_{\rm max}$ at its bound of unity, so that f_* peaks at $\simeq 0.5$ near 7×10^{12}M$_\odot$. Because the fit is to a smooth curve rather than to measurements, the rms measures only how closely the model follows the curve; it is not a likelihood and does not rank models statistically.

The local index: a parameter-free test of q. Near $z = 0$, Q multiplies the unquenched density by $[(1+z)/(1+z_*)]^q$ whenever $z_* > 0$, so at fixed efficiency parameters the local index $d\ln\rho_{\rm SFR}/d\ln(1+z)|_{z=0}$ rises by exactly q, whatever a_* and the form of f_*; refitting the efficiency changes this by at most 0.05. Slowing halo growth alone gives a local index of 1.50. The rising contrast of haloes relative to the gas they feed adds $q = 5/4$, giving 2.71–2.73 (Table 3), against 2.69 for the MD14 fit and 2.86 for the W20 fit. The halo model supplies 1.5 of the observed index and the parting of the clocks supplies the remaining 5/4.

Table 3. The late decline (Fig. 5): fits of Eq. (21) to MD14 at $0 \le z \le 8$. N_p: fitted parameters (a_* is fixed by the N-body crossover); rms: deviation from the MD14 curve, in dex; local: $d\ln\rho_{\rm SFR}/d\ln(1+z)$ at $z=0$; slope: least-squares slope in $\ln(1+z)$ at $0 \le z \le 1$; $z_{\rm peak}$: redshift of maximum $\rho_{\rm SFR}$. Benchmarks modify Eq. (23) with $Q \equiv 1$. [a]As in Harikane *et al.* (2018), Eq. 25. [b]Power-law form of Harikane *et al.* (2018), Eq. 26 ($\zeta = -0.6$ there), with a free exponent. [c]Dekel and Birnboim (2006). [d]Madau and Dickinson (2014), [e]Walter *et al.* (2020). [f]Eq. (22) with the slopes of Harikane *et al.* (2018), $\beta_* = 1.1$, $\gamma_* = 0.3$; $f_{\rm pk}$ and M_c fitted.

Model	N_p	rms	local	slope	$z_{\rm peak}$
constant efficiency[a]	3	0.078	1.50	1.53	2.1
$f_* \propto (1+z)^\zeta$, $\zeta = 0.66$[b]	4	0.062	1.95	1.79	1.9
$M_q \propto (1+z)^\delta$, $\delta = 4.1$[c]	4	0.059	3.44	2.68	1.9
Q, $q = 5/4$, $a_* = 0.64$	3	0.062	2.73	2.33	2.0
Q, $q = 5/4$, $a_* = 0.50$	3	0.055	2.71	2.69	1.9
Harikane slopes[f]	2	0.116	1.37	1.42	2.4
$\times Q$, $a_* = 0.64$	2	0.108	2.61	2.25	2.3
MD14 fit[d]			2.69	2.55	1.86
W20 fit[e]			2.86	2.70	1.73

Shape of the efficiency. The mass-integrated density does not constrain the shape of f_*. The fit to Eq. (23) places the peak efficiency at $\simeq 10^{12.9}\mathrm{M}_\odot$, about ten times higher in mass than abundance matching and clustering find (Moster *et al.* 2013; Harikane *et al.* 2018), because in this halo model only efficient star formation in massive haloes, whose accretion peaks late, moves the density peak to $z \simeq 1.9$. With the slopes of Harikane *et al.* (2018) and a fitted turnover, at $10^{11.75}\mathrm{M}_\odot$, the peak moves to $z \simeq 2.3$ (Table 3). The quench term is insensitive to this choice: in every efficiency we tested, including both of these, a free double power law and the redshift-dependent form of Harikane *et al.* (2018), Q raises the local index by 1.22–1.25, i.e. by q. With the Harikane slopes the MD14 slope gives $a_* \simeq 0.55$ and the W20 slope $a_* \lesssim 0.50$, within the range below.

Comparison with benchmarks. Table 3 compares the quench term with three benchmarks that use the same halo model: (i) a constant efficiency, as in Harikane *et al.* (2018), who argue that halo accretion alone sets the shape of the late decline; (ii) a power-law redshift dependence of the efficiency, $f_* \propto (1+z)^\zeta$, the form of Harikane *et al.* (2018) with a free exponent; and (iii) a quenching mass evolving as $M_q \propto (1+z)^\delta$, our power-law form of the idea that cold streams feed haloes above the shock-heating mass at $z \gtrsim 2$ (Dekel and Birnboim 2006). The first adds no parameter, the other two one exponent each. All fit the curve about equally well (rms 0.055–0.078 dex), so the curve alone does not discriminate. The local index does. Constant efficiency gives 1.50, too shallow by more than one unit. The fit prefers $\zeta = +0.66$, an efficiency rising with redshift, opposite to the $\zeta = -0.6$ of Harikane *et al.* (2018), and still gives only 1.95. The evolving quenching mass reproduces the slope over $0 < z < 1$ (2.68) but overshoots the local index (3.44), because its suppression steepens towards $z = 0$. The quench term gives 2.71–2.73, against 2.69–2.86, with no added parameter.

The crossover from galaxies. The local index does not depend on a_*, but the slope over $0 < z < 1$ does, because

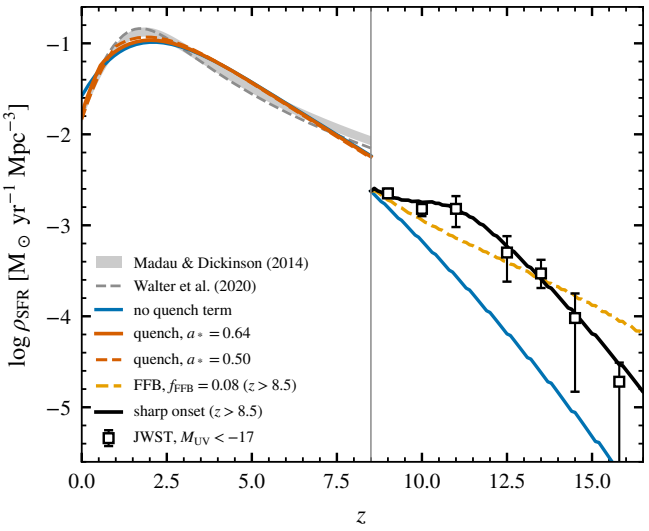


Figure 5. Cosmic star-formation density (Salpeter IMF). Grey band and dashed line: the MD14 and W20 fits. Squares: JWST densities to $M_{\rm UV} = -17$ (Table F1). At $z < 8.5$: the halo model without the quench term (blue) and with it, for $a_* = 0.64$ and 0.50 (red). At $z > 8.5$: constant efficiency (blue), feedback-free starbursts with fitted $f_{\rm FFB}$ (Eq. 28) and the sharp onset of Section 4 (black). Models are integrated over all haloes at $z < 8.5$ and over haloes with SFR $\ge 0.315\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ ($M_{\rm UV} < -17$ for the MD14 UV conversion) at $z > 8.5$, the limit of the JWST densities, hence the step.

Q acts only below z_*. Refitting the efficiency for each a_*, the model slope is 2.33 for $a_* = 0.64$, 2.48 for 0.60, 2.63 for 0.55 and 2.69 for 0.50, the largest value reached. The MD14 slope, 2.55, gives $a_* = 0.58$; the W20 slope, 2.70, lies just above the largest model value and bounds $a_* \lesssim 0.50$. Galaxies therefore place the crossover at $a_* \simeq 0.50$–0.58, in agreement with $a_* \simeq 0.5$ measured in Illustris-1-Dark and with the ΛCDM estimate from the primordial spectrum, 0.54 (Eq. (9)). That galaxy surveys recover a crossover quantity measured in dark-matter is the central result of the paper.

4 THE FIRST GALAXIES

The second application uses the common core rate at early times, before the clocks part.

Data and deficit. We use UV luminosity densities integrated to $M_{\rm UV} = -17$ in seven bins at $9 \le z \le 15.8$ (Donnan *et al.* 2024; McLeod *et al.* 2026) and 28 luminosity-function points at $z = 9$–12.5 (Donnan *et al.* 2024), converted with $\mathcal{K}_{\rm UV} = 1.15 \times 10^{-28}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}/(\mathrm{erg\,s}^{-1}\,\mathrm{Hz}^{-1})$ (Salpeter; Table F1). To compare like with like, the model density at $z > 8.5$ is Eq. (21) integrated only over haloes brighter than the same limit, SFR $\ge 0.315\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, and the model luminosity function is $\phi(M_{\rm UV}) = (dn/d\ln m_h)\,|d\ln m_h/dM_{\rm UV}|$. With the efficiency fitted at $z \le 8$, the model falls short by 0.15 dex at $z = 9$ and by 1.0–1.1 dex at $z \ge 13.5$ (blue curve in Fig. 5).

What the data require. Matching the observed cumulative abundances to halo abundances gives the efficiency the

data need. It stays modest, $f_* \leq 0.12$, but at fixed halo mass ($4\times10^{10}\mathrm{M}_\odot$) it rises by a factor of 1.7 between $z = 10$ and 11, far faster than a smooth trend such as $(1+z)^{0.66}$, which gives 1.06. The rise is not confined to massive haloes: it occurs in all haloes counted ($m_h \gtrsim 10^{10}\mathrm{M}_\odot$ at $z \simeq 10$ for $M_{\rm UV} < -17$; see "No mass gate" below), over a short interval of time.

A sharp onset from the common core rate. Star formation escapes supernova regulation when the free-fall time of the star-forming gas is shorter than the supernova delay. This happens above $n_{\rm th} \simeq 3 \times 10^3\,\mathrm{cm}^{-3}$, the threshold for feedback-free starbursts (FFB; Dekel *et al.* 2023). In the core cascade the density at a given scale, $\rho \propto \varepsilon_b^{2/3} G^{-1} r^{-4/3}$, depends only on the common rate, so it is the same in every host, and at fixed comoving scale it rises as $(1+z)^3$ (Eq. 17). We use the core level because the threshold concerns the densest star-forming gas, and because galaxies at $z \gtrsim 9$ are compact, with $F \simeq 1$. With $n = n_0(1+z)^3$ for star-forming gas, the threshold is crossed in every halo at the same epoch,

$$1 + z_{\rm on} = \left(\frac{n_{\rm th}}{n_0}\right)^{1/3}. \tag{24}$$

We take the efficiency of Eq. (23), with the parameters fitted at $z \leq 8$ and $Q = 1$. For the haloes counted, $m_h \simeq 10^{10}$–$10^{11}\mathrm{M}_\odot \ll M_q$, so only the low-mass law $f_* = f_0(m_h/10^{12}\mathrm{M}_\odot)^{2/3}$ enters. We multiply Eq. (23) by a density switch,

$$f_* \to f_*\left[1 + (\mathcal{B}-1)\,S\right], \qquad S = \frac{1}{1+(n_{\rm th}/n)^p}. \tag{25}$$

S rises from 0 below the onset redshift to 1 above it (in cosmic time the boost fades as the core gas dilutes), over a width set by p. Once it is on, every halo forms stars $\mathcal{B}$ times more efficiently. The boost keeps the mass dependence $f_* \propto m_h^{2/3}$: supernovae still regulate star formation, but more weakly, because more gas collapses within one free-fall time ($\sim 1\,\mathrm{Myr}$) before the first supernovae explode ($\sim 3\,\mathrm{Myr}$). The three fitted parameters are $\mathcal{B} = 4.4$, $n_0 = 2.0\,\mathrm{cm}^{-3}$ and $p = 6.0$, giving $z_{\rm on} = 10.5$; S rises from 10 to 90 per cent between $z = 9.1$ and 11.9.

No mass gate. We test whether the boost must be confined to massive haloes by multiplying the switch by a mass gate,

$$f_* \to f_*\left[1 + (\mathcal{B}-1)\,S\,\mathcal{G}\right], \quad \mathcal{G} = \frac{1}{1+(M_{\rm g}/m_h)^2}, \tag{26}$$

which is close to 0 well below the gate mass $M_{\rm g}$ and close to 1 well above it. We try the two mass scales of the dark-matter clock: $M_{\rm g} = q_h m_h^*(z)$, above which haloes have decoupled from the bulk flow, and $M_{\rm g} = q_M M^*(z)$, above which haloes are rare, fitting q_h or q_M. The fit drives both gates to zero: χ^2 is unchanged and the information criterion rises only by the penalty for the extra parameter (Table 4). At 95 per cent confidence the gate mass lies below $9 \times 10^9\mathrm{M}_\odot$ at $z = 10$, beneath the smallest haloes counted. The physics agrees. Feedback-free bursts also need a surface density $\Sigma \gtrsim 10^3$–$10^4\mathrm{M}_\odot\,\mathrm{pc}^{-2}$, so that winds and radiation cannot disrupt them; in the models of Dekel *et al.* (2023) this sets their mass threshold. In the core cascade it is met in every core: the surface density of all matter within the core radius is $\Sigma \simeq 5 \times 10^5\,(M_b/10^{11}\mathrm{M}_\odot)^{-1/5}\mathrm{M}_\odot\,\mathrm{pc}^{-2}$ at $z = 11$, and it rises inwards, as $r^{-1/3}$, and towards lower core masses. Both conditions therefore reduce to the density threshold, which does not depend on mass. This holds for the haloes probed by $M_{\rm UV} < -17$, $m_h \gtrsim 10^{10}\mathrm{M}_\odot$ at $z \simeq 10$; a gate at lower masses is not constrained.

Comparison with the literature. We compare three published explanations of the excess, each implemented in the same halo model: (i) *redshift-independent efficiency*: the efficiency depends on halo mass but not on time, so the abundance at $z > 9$ is set by halo growth alone (Tacchella *et al.* 2018; Harikane *et al.* 2023). We represent it by Eq. (23) as fitted at $z \leq 8$, which has $Q = 1$ at these redshifts and therefore no time dependence; it has no free parameter here; (ii) *bursty star formation* (Sun *et al.* 2023; Shen *et al.* 2023), in which $\log \mathrm{SFR}$ at fixed halo mass scatters with dispersion $\sigma_{\rm SFR}$ (dex) about the median, so that

$$\phi_{\rm b}(M_{\rm UV}) = \int \phi(M')\,\frac{e^{-(M_{\rm UV}-M')^2/[2(2.5\sigma_{\rm SFR})^2]}}{\sqrt{2\pi}\,2.5\sigma_{\rm SFR}}\,dM', \tag{27}$$

with $\sigma_{\rm SFR}$ fitted and constant in time; the scatter moves galaxies from the steep faint side of the function to the bright side; (iii) *feedback-free starbursts*, in which the efficiency rises to $f_{\rm FFB}$ above the threshold mass $M_{\rm FFB} = 10^{10.8}[(1+z)/10]^{-6.2}\mathrm{M}_\odot$ (Dekel *et al.* 2023; Li *et al.* 2024),

$$f_* \to (1-w)\,f_* + w\,f_{\rm FFB}, \qquad w = \left[1 + e^{-\log_{10}(m_h/M_{\rm FFB})/0.15}\right]^{-1}, \tag{28}$$

with the threshold and transition width of Li *et al.* (2024) and either $f_{\rm FFB} = 0.2$, the lower end of their range 0.2–1, or a fitted value giving $f_{\rm FFB} = 0.08$. Our implementation keeps their threshold but not their non-FFB efficiency, SFR scatter, dust correction, stellar-mass-based UV conversion or halo mass function, so it tests the timing and mass dependence of the feedback-free efficiency rather than their full model.

All free parameters are fitted jointly to the densities ($\chi^2_{\rm J}$) and the luminosity functions ($\chi^2_{\rm LF}$), and models are ranked by AIC $= \chi^2_{\rm J} + \chi^2_{\rm LF} + 2k$, with k the number of fitted parameters. The two data sets share galaxies, so their sum counts some information twice and overstates every $\Delta\chi^2$, by up to a factor of two; halving all χ^2 values still leaves the smallest margin at $\Delta\mathrm{AIC} = 19$. The sharp onset fits best (Table 4, Fig. 6). A constant scatter of $\sigma_{\rm SFR} = 0.40$ dex ($\sigma_{\rm UV} = 1.0$ mag) still misses both the densities and the shape. Because the densities to $M_{\rm UV} = -17$ are dominated by faint galaxies, where the luminosity function is shallow, scatter raises them little, whereas the onset brightens all haloes alike (Appendix E). Shen *et al.* (2023) find that the scatter must instead grow from 0.75 mag at $z \simeq 9$ to 2.0 mag at $z \simeq 12$, itself a redshift trend that needs explaining. In our implementation the feedback-free efficiency $f_{\rm FFB} = 0.2$ overpredicts the abundance at all luminosities; with $f_{\rm FFB} = 0.08$ fitted it approaches the densities but not the shape (Fig. 5). The fitted boost raises the efficiency of 10^{10}–$10^{11}\mathrm{M}_\odot$ haloes to 0.05–0.25 once the switch is complete (0.21 at $10^{11}\mathrm{M}_\odot$ at $z \simeq 11.5$), below the 0.2–1 expected for fully feedback-free bursts. The data therefore favour an increase in efficiency that is timed by density and acts in all haloes.

The bright end. The slow evolution of the brightest galaxies is the hardest feature to reproduce. From $z = 9$ to 12.5 the abundance at $M_{\rm UV} = -20.25$ falls by a factor of 2.0 in

Table 4. Explanations of the $z > 9$ excess (Sheth–Tormen mass function, $a_* = 0.64$). k: parameters fitted jointly to the seven JWST densities ($\chi^2_{\rm J}$) and 28 luminosity-function points ($\chi^2_{\rm LF}$). ΔAIC is relative to the sharp onset. $R_{\rm br}$: ratio of $\phi(M_{\rm UV} = -20.25)$ at $z = 9$ and 12.5 (observed 2.0; Donnan *et al.* 2024). τ_e for $f_{\rm esc} = 0.1$. FFB rows are our implementation of the threshold of Li *et al.* (2024). The gated rows use Eq. (26); their gate parameters are driven to zero.

Model	k	$\chi^2_{\rm J}$	$\chi^2_{\rm LF}$	ΔAIC	$R_{\rm br}$	τ_e
constant efficiency	0	98.9	107.4	173	44.9	0.039
bursty, $\sigma_{\rm SFR} = 0.40$ dex	1	28.2	71.7	68	24.4	0.039
FFB threshold, $f_{\rm FFB} = 0.2$	0	76.1	233.0	276	2.3	0.040
FFB threshold, $f_{\rm FFB} = 0.08$	1	18.3	55.4	42	9.1	0.039
sharp onset, Eq. (25)	3	2.3	25.1	0	4.7	0.040
+ gate $q_h m_h^*$	4	2.3	25.1	2.0	4.7	0.040
+ gate $q_M M^*$	4	2.3	25.1	2.0	4.7	0.040

the data (Donnan *et al.* 2024), against 24–45 for constant efficiency or constant-scatter bursty star formation. The sharp onset gives 4.7, within about 1.5σ of the data. Its sharpness achieves this: the boost is weak at $z = 9$ and full by $z \simeq 12$, so bright galaxies become more luminous as the abundance of their host haloes falls. With a softer switch, $p = 2$, and the other parameters refitted, the ratio becomes 9.1 and the total χ^2 rises from 27 to 48. The feedback-free threshold with $f_{\rm FFB} = 0.2$ reproduces the ratio (2.3), but only by overpredicting every other bin. The boost hardly changes the photon budget of reionization: with $f_{\rm esc} = 0.1$ it gives $\tau_e = 0.040$, against 0.039 without it (Table 4), because few haloes have formed by $z \simeq 10$; matching the Planck value 0.054 needs $f_{\rm esc} \simeq 0.28$. Beyond $z \simeq 12$ the switch is saturated and the density follows the halo mass function: the model predicts $\log \rho_{\rm SFR}(M_{\rm UV} < -17) = -5.0$ at $z = 17$ and -6.5 at $z = 20$ ($\mathrm{M_\odot\,yr^{-1}\,Mpc^{-3}}$), against -4.3 and -5.1 for the feedback-free threshold with fitted $f_{\rm FFB}$.

5 EARLY SUPERMASSIVE BLACK HOLES

The third application applies the core cascade around supermassive black holes.

5.1 A luminosity ceiling

A black hole sits deep inside the cascade of its core, so its surroundings are set by the common core rate $\varepsilon_b(z)$, not by the rate at the effective radius. The Eddington limit, $L_{\rm Edd} = 4\pi G c M_{\rm BH}/\kappa$, with $\kappa = \sigma_T/m_p \simeq 0.4\,\mathrm{cm^2\,g^{-1}}$ the electron-scattering opacity, balances radiation against gravity for each unit mass of smooth, static gas: the radius cancels, and neither the amount nor the motion of the gas enters.

Gas in a core is instead clumpy and turbulent, with dispersion $\sigma(r)$, and is bound by all the mass within r. Radiation then pushes on it as a whole, delivering a force of about L/c, and can clear the gas that confines and feeds the black hole only if this exceeds the force holding that gas in, the dynamical pressure times the area, $4\pi r^2 \rho\,\sigma^2$. This is the momentum balance behind the $M_{\rm BH}$–σ relation in outflow models (King 2003). Unlike the Eddington balance, it depends on radius: in the cascade profile, $\rho \propto r^{-4/3}$ and $\sigma \propto r^{1/3}$, so the confining force grows as $r^{4/3}$. We evaluate it at the sphere

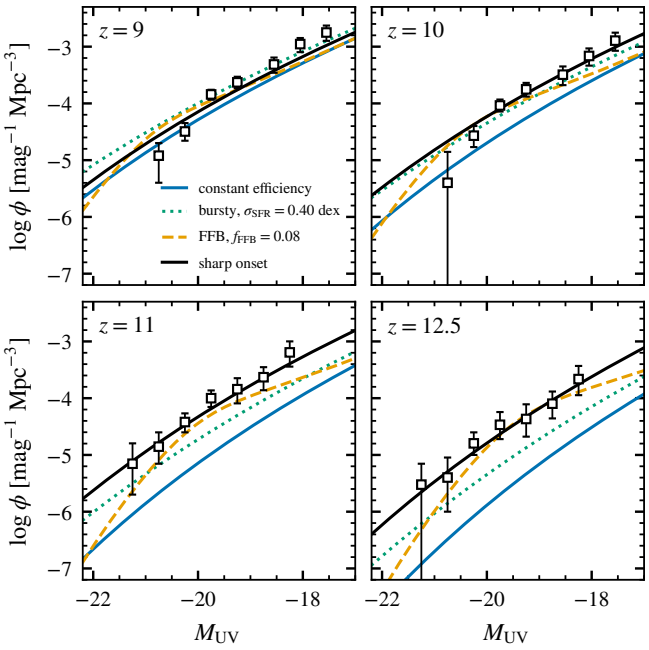


Figure 6. UV luminosity functions at $z = 9$–12.5 (squares; Donnan *et al.* 2024) against constant efficiency, bursty star formation (Eq. 27, fitted $\sigma_{\rm SFR}$), the feedback-free threshold (Eq. 28, fitted $f_{\rm FFB}$) and the sharp onset (Eq. 25; $\mathcal{B} = 4.4$, $n_0 = 2.0\,\mathrm{cm^{-3}}$, $p = 6.0$); $a_* = 0.64$.

of influence, where the profile encloses the black-hole mass, $M_{\rm BH} = \sigma_\bullet^5/(\gamma_r G \varepsilon_b)$. Setting L/c equal to the force there gives (Eq. (C4) in Appendix C1; Xu 2025)

$$L_X = x_r\,\varepsilon_b^{4/5} G^{-1/5} c\, M_{\rm BH}^{4/5},$$
$$\frac{L_X}{L_{\rm Edd}} = \gamma_r^{6/5} x_r\,\tau_\bullet \simeq 1.07\,\frac{x_r}{15}\,(1+z)^2 \left(\frac{M_{\rm BH}}{10^8 \mathrm{M_\odot}}\right)^{-1/5}, \qquad (29)$$

for $\varepsilon_{b0} = 10^{-4}\,\mathrm{m^2\,s^{-3}}$, where $\tau_\bullet$ (effective optical depth) is the weight of the mass within the sphere of influence in units of the Eddington force (Eq. C4). The factor $(1+z)^2 = [(1+z)^{5/2}]^{4/5}$ is the gas clock: denser core gas confines radiation more strongly. Because the core rate is common, the ceiling depends only on $M_{\rm BH}$ and z, not on the host. Individual cores scatter about the common rate by 0.5 dex in $\varepsilon_{{\rm v},b}$, and since $L_X \propto \varepsilon_b^{4/5}$, this spreads the ceiling of individual objects by 0.4 dex, a factor of 2.5.

High-redshift black holes lie below the ceiling, most of them 1–2 dex below (Table 5); only LID-568 approaches it. Type 1 SDSS quasars, corrected for incompleteness and mass errors, reach at most $L/L_{\rm Edd} \simeq 3$ (Kelly and Shen 2013), against a ceiling of 0.9–17 for 10^9–$10^{9.5}\mathrm{M_\odot}$ at $1 \le z \le 4$; and the highest ratios appear first, as $(1+z)^2$ requires. At low redshift the ceiling for 10^9–$10^{10}\mathrm{M_\odot}$ falls to 0.4–1.5 at $z = 0.5$ and 0.2–0.7 at $z = 0$, a sharp prediction for low-redshift quasars. Values below $L_{\rm Edd}$ require dusty gas, which absorbs the radiation fully; for dust-free gas, which is optically thin there ($\tau_\bullet \simeq 0.1$–0.2), the limit is $\max(L_{\rm Edd}, L_X)$.

A universal envelope. Because $\varepsilon_b(z)$ is common at a given epoch, the ceiling can be tested on the whole population. In terms of $\ell \equiv L/(M_{\rm BH}\varepsilon_b)$ and $\omega \equiv (GL/c^5)^{1/4}$ (Xu 2025),

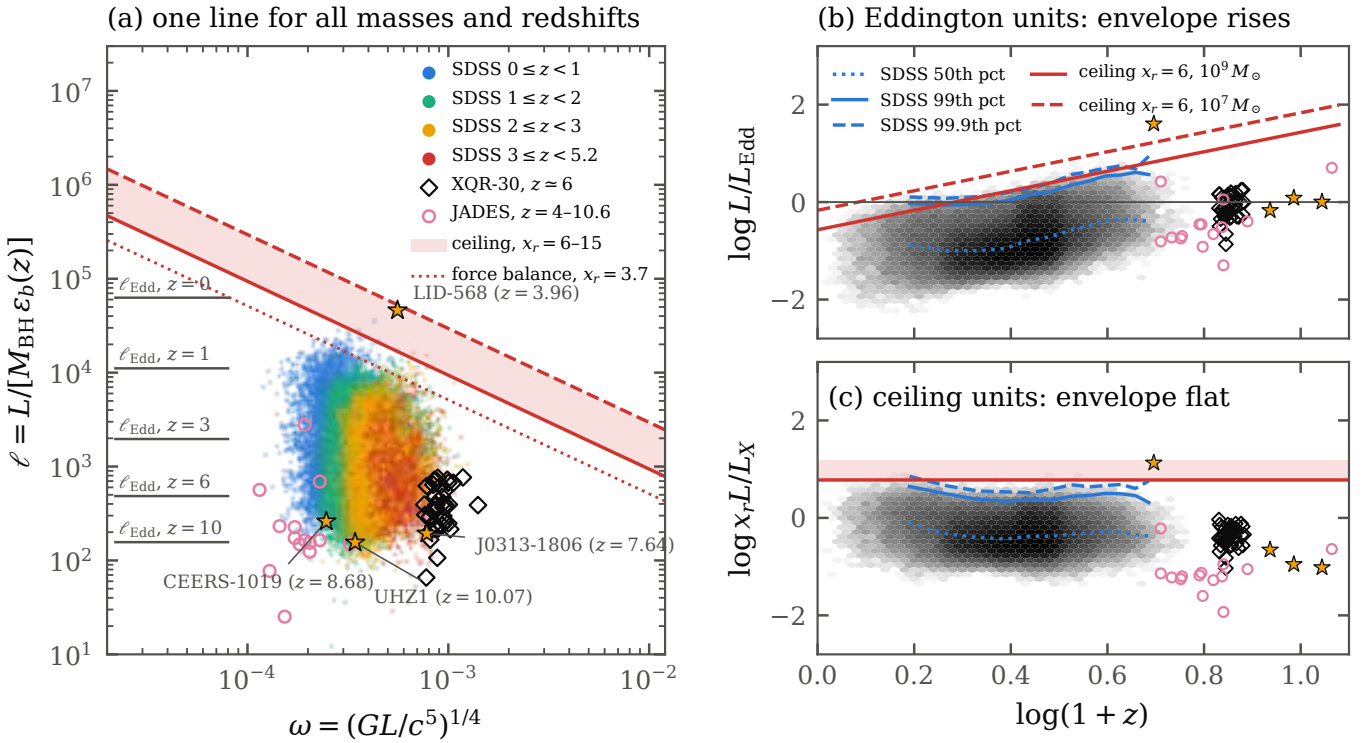


Figure 7. The luminosity ceiling as a universal envelope for 1.5×10^5 black holes. (a) The ℓ–ω plane, $\ell = L/[M_{\rm BH}\varepsilon_b(z)]$ and $\omega = (GL/c^5)^{1/4}$: SDSS quasars coloured by redshift (25 per cent subsample; Chen *et al.* 2019; Shen *et al.* 2011), XQR-30 quasars (Mazzucchelli *et al.* 2023), JADES black holes (Maiolino *et al.* 2024a) and single objects (stars). Shaded: the ceiling, Eq. (30), for $x_r = 6$ (solid; envelope of the population) to 15 (dashed; LID-568); dotted: the force-balance value $x_r = 3.7$. Horizontal segments: the Eddington limit at $z = 0$–10, which moves with redshift while the ceiling does not. (b) Eddington ratio against $\log(1+z)$ (grey: SDSS density; blue: SDSS 50th, 99th and 99.9th percentiles; symbols as in a), with the ceiling for $x_r = 6$ at $M_{\rm BH} = 10^9$ and $10^7 {\rm M}_\odot$. (c) The same objects in ceiling units, $(\ell\omega)^{4/5} = x_r L/L_X$: the envelope that rises in (b) stays flat to within 0.3 dex.

Table 5. Eddington ratios against the ceiling, Eq. (29) with $x_r = 15$ (divide by 2.5 for $x_r = 6$). Literature values are approximate. [a]Suh *et al.* (2025); [b]Sun *et al.* (2026); [c]Larson *et al.* (2023); [d]Bogdán *et al.* (2024); [e]Maiolino *et al.* (2024b).

Object	z	$M_{\rm BH}$ [M$_\odot$]	$L/L_{\rm Edd}$	Ceiling
LID-568[a]	3.97	7×10^6	~ 40	45
BH* stack[b]	4–9	10^4–10^5	5–50	107–676
CEERS-1019[c]	8.68	9×10^6	~ 1	163
UHZ1[d]	10.07	4×10^7	~ 1	158
GN-z11[e]	10.60	1.6×10^6	~ 5	330

Eq. (29) becomes one line for every mass and redshift,

$$\ell\,\omega = x_r^{5/4}, \qquad \frac{L}{L_X} = \frac{(\ell\omega)^{4/5}}{x_r}, \tag{30}$$

whereas the Eddington limit, $\ell_{\rm Edd} = (4\pi Gc/\kappa)/\varepsilon_b(z)$, moves down by $(1+z)^{5/2}$. Figure 7 places 151 789 black holes on this plane: 141 828 SDSS-IV quasars with virial masses from Hβ, Mg II or C IV (Chen *et al.* 2019) and the bolometric corrections of Shen *et al.* (2011); a random 9902 of the DR7 catalogue (Shen *et al.* 2011); 41 quasars at $z \simeq 6$ (Mazzucchelli *et al.* 2023); 14 JADES black holes at $z = 4$–10.6 (Maiolino *et al.* 2024a); and four single objects.

From $z = 0.4$ to 4.3 the 99th-percentile envelope of $L/L_{\rm Edd}$ rises by 0.6 dex in the raw bins; at fixed mass it rises as $(1+z)^{1.59\pm0.08}$, and at fixed redshift it falls as $M_{\rm BH}^{-0.36}$ (Figure 7b). In ceiling units, with percentiles recomputed from L/L_X, the same envelope varies by only 0.3 dex, with residual slopes of -0.20 ± 0.12 in $\log(1+z)$, consistent with zero, and -0.18 in $\log M_{\rm BH}$. Virial masses scattered by 0.3–0.4 dex bias ratios with $M_{\rm BH}$ in the denominator towards a negative mass slope of $\simeq -0.3$, which can account for the residual. Dividing out the gas clock thus removes almost all of the redshift evolution of the envelope. At fixed mass (10^8–$10^{9.5}$M$_\odot$) the envelope of $L/L_{\rm Edd}$ rises as $(1+z)^{2.15\pm0.15}$, the slope of the ceiling, but the median and 16th percentile rise as steeply (1.94, 1.92), so the flux limit of SDSS shifts the distribution: this agreement is consistent with the ceiling rather than a test of it.

The population also fixes the normalization: $\ell\omega = 10$ ($x_r = 6.3$) lies above all but 24 of the 151 730 SDSS quasars (0.016 per cent, as expected from the 0.3–0.4 dex scatter of virial masses and the 0.5 dex scatter of $\varepsilon_{{\rm v},b}$), $x_r = 10.1$ lies above all of them, and the 99th and 99.9th percentiles lie at $x_r = 2.7$ and 4.5, at the top of the force-balance range. Only LID-568 requires $x_r \simeq 13.5$. We adopt $x_r = 6$–15. Quasars at $z \simeq 6$ and the JWST black holes lie 1–2 dex below it, because $(1+z)^2$ raises the ceiling faster than their Eddington ratios.

Growth. At a time-averaged fraction $f_{\rm duty}$ of the ceiling, $\dot{M}_{\rm BH} = f_{\rm duty}(1-\eta_r)L_X/(\eta_r c^2)$, so $M_{\rm BH}^{1/5}$ grows linearly in $\int(1+z)^2 dt$ (Appendix C2): growth is polynomial, not exponential, and front-loaded. For $\eta_r = 0.1$ and $x_r = 15$, seeds of

$10^2\mathrm{M}_\odot$ at $z = 25$ or $10^4\mathrm{M}_\odot$ at $z = 20$ reach GN-z11, CEERS-1019, UHZ1, GHZ9 (Kovács *et al.* 2024), Abell 2744-QSO1 (Furtak *et al.* 2024) and J0313−1806 (Wang *et al.* 2021) for $f_{\rm duty} = (1.0\text{–}4.4) \times 10^{-3}$, equivalent to a steady Eddington ratio of 0.7–2.0, whereas Eddington-limited growth from the same seeds reaches only 9×10^4 and $3.5 \times 10^6\mathrm{M}_\odot$ by $z = 10$. Light seeds are therefore not excluded. The calculation shows feasibility; it does not predict duty cycles, seed abundances or mass functions.

5.2 Black-hole–host relations

An identity and a condition. Eliminating R_b between $\sigma_b^2 = \gamma_r G M_b/R_b$ and $\varepsilon_{\mathrm{v},b} = \sigma_b^3/R_b$ gives $M_b = \sigma_b^5/(\gamma_r G \varepsilon_{\mathrm{v},b})$. For relations $M_{\rm BH} = A\sigma_b^5 = BM_b$ the ratio of the coefficients is therefore an identity of host structure (Xu 2025),

$$\frac{B}{A} = \gamma_r G\, \varepsilon_{\mathrm{v},b}, \tag{31}$$

valid when σ_b, M_b and $\varepsilon_{\mathrm{v},b}$ refer to the same structure. With the Table B1 (Xu 2025) zero points, $A_0 = 10^{-3.14}\mathrm{M}_\odot\,(\mathrm{km\,s^{-1}})^{-5}$ and $B_0 = 0.0026$, it returns $\varepsilon_0 = 10^{-4.09}\,\mathrm{m^2\,s^{-3}}$, the common cascade rate of gas.

The identity fixes the ratio B/A; the cascade supplies the missing condition. The dissipation power of the gas of a virialized core, with gas mass $f_g M_b$, is

$$P_g = \varepsilon_{\mathrm{v},b}\, f_g M_b = \frac{f_g\, \sigma_b^5}{\gamma_r G}, \tag{32}$$

independent of $\varepsilon_{\mathrm{v},b}$ and of redshift at fixed σ_b. Any self-regulation in which an output proportional to $M_{\rm BH}$ balances this power gives $M_{\rm BH} \propto \sigma_b^5$ with a constant coefficient. For a coupled fraction η_c of the Eddington luminosity, $\eta_c L_{\rm Edd} = P_g$,

$$A = \frac{f_g\, \kappa}{4\pi\gamma_r \eta_c\, G^2 c}, \tag{33}$$

the form of Silk and Rees (1998), here obtained from the cascade power. The Table B1 coefficient requires $\eta_c \simeq 5 \times 10^{-4}$ for $f_g = 0.1$ (Appendix C3), of order σ/c, the energy coupling of an outflow that carries the momentum L/c (King 2003); the cascade fixes the form and constancy of A, not its value. With A constant, Eq. (31) makes B follow the rate of the level whose mass is used:

$$B_{\rm core} = \gamma_r G A\, \varepsilon_b(z) = B_0(1+z)^{5/2}, \quad B_{\rm dyn} = \frac{B_{\rm core}}{F}, \tag{34}$$

where $B_{\rm dyn}$ uses a dynamical mass measured at R_e. Figure 8 tests the three statements in turn.

Constant A locally. In Table B1, the 91 ellipticals and S0 bulges with black-hole masses give $\Delta\log A = (-0.27 \pm 0.10)\,\Delta\log\varepsilon_{\mathrm{v},b}$, close to constant A and far from the -1 of constant B (Fig. 8a); equivalently $M_{\rm BH} \propto \sigma^{4.25\pm0.36} R_b^{0.28\pm0.12}$. The six most compact early types (NGC 1277, NGC 1332, NGC 4342, Mrk 1216, NGC 3998, NGC 6861) are the cleanest case: their rates are +1.09 dex above the mean, their coefficients A are normal ($\Delta\log A = -0.07$), and their black holes are over-massive relative to the bulge by the same factor ($\Delta\log B = +1.07$). They are over-massive because their bulges are dense, not because their black holes are anomalous; in the clock, such rates were common at $z_f \simeq 1.7$ (NGC 1277: $z_f \simeq 3.2$), consistent with their being relics. The 37 spiral bulges behave differently, $\Delta\log A = (-1.14\pm0.17)\,\Delta\log\varepsilon_{\mathrm{v},b}$ ($M_{\rm BH} \propto \sigma^{0.6\pm0.7} R_b^{0.8\pm0.2}$): their black holes track bulge mass rather than dispersion, the known behaviour of pseudobulges (Kormendy and Ho 2013).

Constant A in time. Where selection is weak, A does not evolve: the $M_{\rm BH}$–σ_* relation is unchanged to $z \simeq 1$ (Shen *et al.* 2015), and quiescent hosts lie on it at $z = 2.09$ ($\Delta\log A = -0.31$; Ito *et al.* 2025) and $z = 4.66$ (-0.12; GS-9209, Carnall *et al.* 2023) (Fig. 8b). Flux-selected AGN lie above it, as luminosity selection favours black holes above the relation: the JADES hosts by +0.39 (relation of Kormendy and Ho 2013) to +0.76 dex (Table B1 zero point), and $z \simeq 6$ quasars by $+1.65 \pm 0.14$ dex with [C II] dynamics (Neeleman *et al.* 2021). That offset reflects the tracer, not an evolving A: the two $z \simeq 6.2$–6.4 quasars with stellar absorption-line dispersions lie on the relation, $\Delta\log A = -0.13 \pm 0.41$ and -0.22 ± 0.45 (Onoue *et al.* 2025; Liu *et al.* 2026). Constant A applies to hosts regulated by their core cascade. Abell 2744-QSO1, with $M_{\rm BH} = 10^{7.7}\mathrm{M}_\odot$ and a narrow-line dispersion below $22\,\mathrm{km\,s^{-1}}$ where regulation would need $150\,\mathrm{km\,s^{-1}}$, lies more than 4 dex above it (Juodžbalis *et al.* 2026); The Cliff at $z = 3.55$ lies +1.1 to +2.6 dex above it, depending on the black-hole mass and the dispersion adopted (Ivey *et al.* 2026). In QSO1 the black hole dominates the mass; in The Cliff, for its fiducial mass, its sphere of influence ($\simeq 70$ pc) exceeds the effective radius ($\simeq 39$ pc). In neither is there a host cascade yet.

B follows the level measured. At fixed A, B measures the cascade rate of the structure whose mass is used (Fig. 8c, right axis). Four data sets agree with Eq. (34).

(1) *Spheroids are consistent with the core track.* Since $B = \gamma_r G A\, \varepsilon_{\mathrm{v},b}$, a constant A gives $B = M_{\rm BH}/M_b \propto \varepsilon_b \propto (1+z)^{5/2}$. For broad-line AGN at $z = 1$–2, the ratio of black-hole mass to spheroid stellar mass evolves as $(1+z)^{1.96\pm0.55}$ after selection correction (Bennert *et al.* 2011), and as $(1+z)^{2.09\pm0.30}$, or $\simeq (1+z)^{1.6}$ after selection correction, with bulge masses estimated from Sérsic indices (Ding *et al.* 2020). The ratio to total stellar mass evolves more slowly, as $(1+z)^{1.15\pm0.15}$ and $(1+z)^{0.55\pm0.15}$: it adds the disc, read at R_e, to the core, and has no single predicted exponent.

(2) *Compact quiescent spheroids carry high rates.* Of the 39 quiescent galaxies with stellar dispersions at $z = 1.5$–4.7 (Belli *et al.* 2017; Forrest *et al.* 2022, and references therein) 34 have σ_*^3/R_e above the local zero point ε_0; the other five, all at $z < 1.75$, lie at most 0.14 dex below it. The 13 galaxies at $z = 2.7$–4.0 exceed it by $+1.04 \pm 0.12$ dex, like the local relics (+1.09), and 0.5 dex below the core track (+1.57), as expected if R_e encloses more than the core. GS-9209 at $z = 4.66$ shows the same shortfall: +1.45 dex against +1.88 predicted. Against local early types of the same stellar mass, their rates rise as $(1+z)^{1.72\pm0.14}$ (local reference: the 100 early-type hosts of Appendix D, $M_* = 10^{-0.15} L_K$), close to the disc level, $(1+z)^{2.08\pm0.19}$, rather than as $(1+z)^{5/2}$. Since $F \propto M^{-2/3}$ with M the dynamical mass within R_e, the lower $M_{\rm dyn}/M_*$ found at $z \gtrsim 3$ (Forrest *et al.* 2022) raises F at fixed stellar mass and lowers the exponent; at fixed dynamical mass, 5/2 should be recovered. Compact relics have the highest B at $z = 0$ because $B = \gamma_r G A\, \varepsilon_{\mathrm{v},b}$ and their rates are the highest locally (+1.09 dex): a normal black hole for its σ in a very small spheroid. Their rates equal the core track at

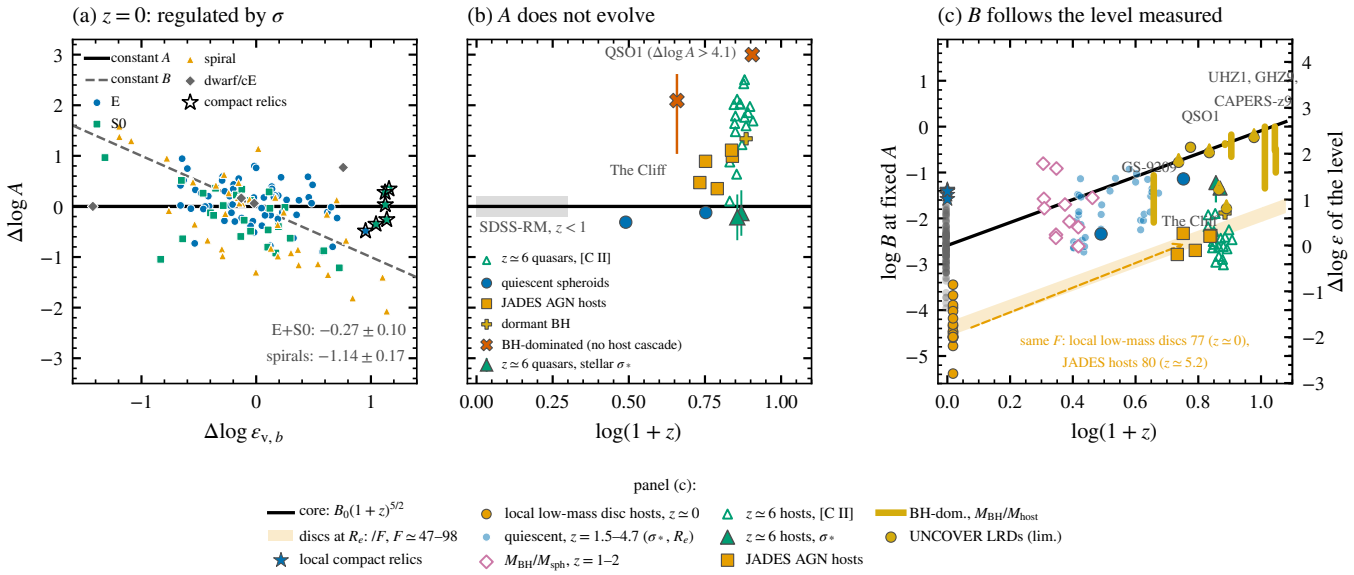


Figure 8. Black-hole–host relations. (a) Local regulation (Table B1): offset of $A = M_{\mathrm{BH}}/\sigma^5$ against the offset of the bulge rate $\varepsilon_{\mathrm{v},b}$. Ellipticals and S0 bulges follow constant A (solid; slopes at bottom right), spiral bulges constant B (dashed); compact relics (stars) have normal A at high $\varepsilon_{\mathrm{v},b}$. (b) Offset of A against redshift (zero point of Table B1). Quiescent spheroids (circles) and the two $z \simeq 6$ quasar hosts with stellar dispersions (filled triangles) lie on the local relation (grey band: SDSS-RM limit at $z < 1$, Shen *et al.* 2015); JADES AGN hosts (squares), [C II] quasar hosts (open triangles) and the dormant black hole (plus) lie above it; black-hole-dominated systems without a host cascade (crosses: The Cliff, and QSO1, a lower limit plotted at the top edge) lie far above. (c) B at fixed A, i.e. $B_0\,\varepsilon/\varepsilon_0$ with $\varepsilon_0 = 10^{-4.09}\,\mathrm{m^2\,s^{-3}}$ (Eq. 31; right axis: offset of the rate of the level measured), against redshift. Solid: the core track $B_0(1+z)^{5/2}$; band: the disc level at R_e, lower by $F \simeq 47$–98. Grey: Table B1 galaxies at $z = 0$. The compact relics (Trujillo *et al.* 2014; Ferré-Mateu *et al.* 2017) lie above the core track at $z = 0$, at its value for $z \simeq 1.7$; quiescent galaxies at $z = 1.5$–4.7 (small circles) and quiescent black-hole hosts (large circles, as in b) lie about 0.5 dex below it. Local low-mass disc hosts (orange circles at $z \simeq 0$) fill the band and have the same F as the JADES hosts (dashed arrow); [C II] quasar hosts lie near the band, those with stellar dispersions between band and core. Black-hole-dominated systems (bars: $M_{\mathrm{BH}}/M_{\mathrm{host}}$; circles: little red dots with host limits) reach the core track. $M_{\mathrm{BH}}/M_{\mathrm{sph}}$ at $z = 1$–2 (open diamonds) has no σ and is not corrected to fixed A.

$z_f \simeq 1.7$, consistent with cores that formed early and have not grown since.

(3) *Disc-dominated hosts sit at the disc level.* F does not evolve at fixed mass and structure, so a host measured at R_e lies at $B_e = B_0(1+z)^{5/2}/F$. The JADES hosts have $F \simeq 80$, like local low-mass disc hosts with stellar dispersions ($F \simeq 77$; Section 2.5; here F is measured from $\varepsilon_0 = 10^{-4.09}\,\mathrm{m^2\,s^{-3}}$, so the same local hosts have $F = 68$ when measured from $\varepsilon_{b0} = 10^{-4.14}\,\mathrm{m^2\,s^{-3}}$ as in Section 2.5). At $z \simeq 5$, $(1+z)^{5/2} \simeq 88$ nearly cancels F, which is why they appear to follow the local M_{BH}–M_{dyn} relation. The tracer selects the level: the $z \simeq 6$ quasar hosts give $F \simeq 120$ with [C II], which samples the gas disc, but $F \simeq 6$–8 with stellar dispersions, which sample the inner region. At fixed dynamical mass (median $10^{10.5}\mathrm{M}_\odot$) these hosts lie $+1.42 \pm 0.09$ dex above local discs, and $+1.25 \pm 0.08$ dex above local disc hosts with stellar dispersions, against $+2.16$ for fixed structure at $z \simeq 6.3$: their F is $\simeq 0.75$–0.9 dex larger than that of local discs of the same mass (Appendix D).

(4) *Where the black hole dominates, B reaches B_{core}.* The dormant black hole at $z = 6.68$ has $M_{\mathrm{BH}}/M_* \simeq 0.4$ (Juodžbalis *et al.* 2024), against predicted $B_{\mathrm{core}} = 0.42$; QSO1 at $z = 7.04$ has $M_{\mathrm{BH}}/(M_{\mathrm{BH}} + M_*) > 0.22$–0.67 (Juodžbalis *et al.* 2026), against 0.47; GHZ9, CAPERS-LRD-z9 and UHZ1 at $z \simeq 9$–10 have $M_{\mathrm{BH}}/M_{\mathrm{host}} \simeq 0.05$–1 (Napolitano *et al.* 2025; Taylor *et al.* 2025; Bogdán *et al.* 2024), against 0.87–1.07; and The Cliff has $M_{\mathrm{BH}}/M_{\mathrm{dyn}} = 0.008$–0.087, against 0.11. Among six little red dots at $z = 4.5$–8.5 with host limits (Chen *et al.* 2025), the most black-hole-dominated reach 0.8–0.9 B_{core}. The largest excess, +0.2 dex, is UNCOVER-38108, the only one with a detected host.

Two predictions follow. First, B_{core} reaches unity at $z \simeq 10$: a black hole on the local A relation would then outweigh its core, so at higher redshift black holes either lie below that relation or dominate their cores. Second, disc-dominated hosts measured at R_e lie a factor $(1+z)^{5/2}/F$ from the local, spheroid-calibrated M_{BH}–M_{dyn} relation: low-mass discs ($F \simeq 47$–98) should lie 0.2–1.2 dex below it at $z \simeq 1$–3 and cross it at $z \simeq 4$–5, massive discs ($F \simeq 10$) already at $z \simeq 1.5$.

6 CONCLUSIONS

(i) *Clocks.* Dark-matter haloes of fixed mass collapse at fixed density contrast until the crossover $a_* \simeq 0.5$–0.64, measured in N-body simulations and estimated from the primordial spectrum, and grow denser afterwards. Gas radiates its random energy; we take it to keep a fixed contrast (H1), so the cascade rate common to all galaxy cores at a given epoch (H2, measured in 177 local bulges over 3.9 dex in mass) rises as $(1+z)^{5/2}$. H1 and H2 are distinct and compatible: the core contrast falls with mass as $M_b^{-4/5}$ and is fixed in time, whereas the contrast at the boundary of a disc, like that of a halo, is the same for all masses.

(ii) *Levels.* Like a halo, a galaxy has levels: the rate at its effective radius lies below the core rate by $F = R_e/r_c$, about 1 in spheroids and 10^2–$10^{2.5}$ in low-mass discs. Compared at the same level, star-forming galaxies at $z \leq 7$ follow fixed contrast, $(1+z)^{2.08\pm0.19}$, and the JADES broad-line hosts lie above local disc hosts measured alike by +1.18 dex at fixed σ and $+1.98 \pm 0.13$ dex at fixed dynamical mass, as predicted (+1.19, +1.98).
(iii) *Late decline.* After a_* the fraction of accreted gas able to cool falls as $(1+z)^{5/4}$. With no added parameter this raises the local index of the star-formation density from 1.5 to 2.7, against 2.69–2.86 observed, where benchmarks with one more parameter give 1.95 and 3.44. The slope at $0 < z < 1$ places the crossover at $a_* \simeq 0.50$–0.58.
(iv) *First galaxies.* Because the core rate is common to all mass, a density threshold for efficient star formation is crossed in all haloes at once. A boost of the efficiency by a factor of 4.4 that switches on sharply at $z \simeq 10.5$ fits the JWST densities and luminosity functions far better than bursty or feedback-free alternatives ($\Delta\mathrm{AIC} \geq 42$), needs no mass gate and leaves the photon budget of reionization unchanged. It predicts $\log \rho_{\mathrm{SFR}} = -6.5$ at $z = 20$, 1.4 dex below feedback-free starbursts.
(v) *Black holes.* The core cascade bounds black-hole luminosities as $L/L_{\mathrm{Edd}} \propto (1+z)^2 M_{\mathrm{BH}}^{-1/5}$; with one normalization this encloses all but 24 of 1.5×10^5 quasars, and admits light seeds; the luminous-quasar shortfall and early quiescence remain (see Limits). M_{BH}/σ^5 is constant, and M_{BH}/M_b follows the level at which the host mass is measured: the core track $(1+z)^{5/2}$ for compact relics, compact quiescent spheroids and black-hole-dominated systems at $z \simeq 7$–10, and the disc level for the JADES hosts, which is why they appear local. Quasar hosts with stellar dispersions lie on the local M_{BH}–σ relation at $z \simeq 6$.
(vi) *Limits.* The onset density, boost and switch width, the ceiling normalization and the M_{BH}–σ coupling are fitted, and the common core rate is measured, not derived. The model does not explain early quiescence at $z = 3$–5 or the luminous quasar hosts at $z \simeq 6$, and does not model the dust that may lower the brightest galaxies at $z = 9$. A decisive test is the inflow efficiency at fixed halo mass in hydrodynamic simulations, which should be $\propto (1+z)^{\frac{5}{4}}$ after a_*.

AUTHOR CONTRIBUTIONS

Z.J. Xu performed the conception, data analysis, and writing.

FUNDING

This research was supported by the Laboratory Directed Research and Development at Pacific Northwest National Laboratory (PNNL). PNNL is a multiprogram national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under contract no. DE-AC05-76RL01830.

DATA AVAILABILITY

All code, input data, intermediate results and figure scripts are released with this paper (Appendix F); one command rebuilds every number, figure, table and the manuscript.

REFERENCES

P. Madau and M. Dickinson, ARA&A **52**, 415 (2014).
F. Walter, C. Carilli, M. Neeleman, *et al.*, ApJ **902**, 111 (2020).
C. T. Donnan, R. J. McLure, J. S. Dunlop, *et al.*, MNRAS **533**, 3222 (2024), 2403.03171 .
D. J. McLeod, J. S. Dunlop, R. J. McLure, C. T. Donnan, *et al.*, arXiv e-prints (2026), 10.48550/arXiv.2604.16666, 2604.16666 .
R. Maiolino, J. Scholtz, E. Curtis-Lake, *et al.*, A&A **691**, A145 (2024a), 2308.01230 .
F. Pacucci and A. Loeb, ApJ **964**, 154 (2024), arXiv:2401.04159 [astro-ph.GA] .
H. Suh, J. Scharwächter, E. P. Farina, *et al.*, Nature Astronomy **9**, 271 (2025).
W. Q. Sun, R. P. Naidu, H. Liu, *et al.*, arXiv e-prints (2026), 10.48550/arXiv.2609.09274, 2609.09274 .
R. Decarli, F. Walter, J. Gónzalez-López, *et al.*, ApJ **882**, 138 (2019).
A. Dekel and Y. Birnboim, MNRAS **368**, 2 (2006).
L. J. Tacconi, R. Genzel, A. Saintonge, *et al.*, ApJ **853**, 179 (2018).
S. Tacchella, S. Bose, C. Conroy, *et al.*, ApJ **868**, 92 (2018).
Y. Harikane, M. Ouchi, M. Oguri, *et al.*, ApJS **265**, 5 (2023).
J. B. Muñoz, J. Mirocha, J. Chisholm, *et al.*, MNRAS **535**, L37 (2024).
Á. Bogdán, A. D. Goulding, P. Natarajan, *et al.*, Nature Astronomy **8**, 126 (2024).
K.-y. Watarai, J. Fukue, M. Takeuchi, and S. Mineshige, PASJ **52**, 133 (2000).
V. Rusakov, D. Watson, G. P. Nikopoulos, *et al.*, Nature **649**, 574 (2026), arXiv:2503.16595 [astro-ph.GA] .
Z. Xu, "Universal spectrum and scaling laws for dark matter halo mass functions and density profiles," Research Square preprint rs-8012958 (v3) (2026a).
Z. Xu, MNRAS **536**, 3554 (2025), arXiv:2501.07608 [astro-ph] .
W. M. Irvine, *Local Irregularities in a Universe Satisfying the Cosmological Principle*, Ph.D. thesis, Harvard University (1961).
D. Layzer, The Astrophysical Journal **138**, 174 (1963).
Z. Xu, arXiv e-prints (2021), 10.48550/arXiv.2109.09985, arXiv:2109.09985 .
Planck Collaboration, N. Aghanim, Y. Akrami, *et al.*, Astronomy and Astrophysics **641**, A6 (2020), 1807.06209 .
A. Marconi and L. K. Hunt, ApJ **589**, L21 (2003), astro-ph/0304274 .
A. V. Kravtsov, ApJL **764**, L31 (2013), 1212.2980 .
K.-H. Huang, S. M. Fall, H. C. Ferguson, A. van der Wel, N. Grogin, A. Koekemoer, S.-K. Lee, P. G. Pérez-González, and S. Wuyts, ApJ (2017), 10.3847/1538-4357/aa62a6, 1701.04001 .
R. S. Sutherland and M. A. Dopita, ApJS **88**, 253 (1993).
W. G. Mathews and F. Brighenti, ARA&A **41**, 191 (2003).
Z. Xu, Scientific Reports **13**, 16531 (2023), arXiv:2210.01200 [astro-ph.CO] .
D. Barat, F. D'Eugenio, M. Colless, L. Cortese, B. Catinella, S. Oh, *et al.*, MNRAS **487**, 2924 (2019), 1905.12637 .
A. van der Wel, M. Franx, P. G. van Dokkum, *et al.*, ApJ **788**, 28 (2014).
S. Shen, H. J. Mo, S. D. M. White, M. R. Blanton, G. Kauffmann, W. Voges, *et al.*, MNRAS **343**, 978 (2003), astro-ph/0301527 .
M. V. Maseda, A. van der Wel, H.-W. Rix, E. da Cunha, C. Pacifici, I. Momcheva, *et al.*, ApJ **791**, 17 (2014), 1406.3351 .

A. de Graaff, H.-W. Rix, S. Carniani, K. A. Suess, S. Charlot, E. Curtis-Lake, *et al.*, A&A **684**, A87 (2024).
T. S. Gonçalves, A. Basu-Zych, R. Overzier, D. C. Martin, D. R. Law, D. Schiminovich, *et al.*, ApJ **724**, 1373 (2010), 1009.4934 .
Y. Harikane, M. Ouchi, Y. Ono, S. Saito, P. Behroozi, S. More, K. Shimasaku, J. Toshikawa, Y.-T. Lin, M. Akiyama, J. Coupon, Y. Komiyama, A. Konno, S.-C. Lin, S. Miyazaki, A. J. Nishizawa, T. Shibuya, and J. Silverman, PASJ **70**, S11 (2018), arXiv:1704.06535 [astro-ph.GA] .
R. Barkana and A. Loeb, Phys. Rep. **349**, 125 (2001).
O. Fakhouri, C.-P. Ma, and M. Boylan-Kolchin, MNRAS **406**, 2267 (2010).
R. K. Sheth and G. Tormen, MNRAS **308**, 119 (1999).
B. P. Moster, T. Naab, and S. D. M. White, MNRAS **428**, 3121 (2013), 1205.5807 .
A. Dekel, K. C. Sarkar, Y. Birnboim, N. Mandelker, and Z. Li, MNRAS **523**, 3201 (2023).
G. Sun, C.-A. Faucher-Giguère, C. C. Hayward, *et al.*, ApJL **955**, L35 (2023).
X. Shen, M. Vogelsberger, M. Boylan-Kolchin, *et al.*, MNRAS **525**, 3254 (2023).
Z. Li, A. Dekel, K. C. Sarkar, *et al.*, A&A **690**, A108 (2024).
A. King, ApJ **596**, L27 (2003).
B. C. Kelly and Y. Shen, ApJ **764**, 45 (2013).
R. L. Larson, S. L. Finkelstein, D. D. Kocevski, *et al.*, ApJL **953**, L29 (2023).
R. Maiolino, J. Scholtz, J. Witstok, *et al.*, Nature **627**, 59 (2024b).
Z.-F. Chen, S.-X. Yi, T.-T. Pang, Z.-G. Chen, R.-J. Gui, Z.-W. Wang, X.-H. Mo, and T.-F. Yi, ApJS **244**, 36 (2019).
Y. Shen, G. T. Richards, M. A. Strauss, P. B. Hall, D. P. Schneider, S. Snedden, D. Bizyaev, H. Brewington, V. Malanushenko, E. Malanushenko, D. Oravetz, K. Pan, and A. Simmons, ApJS **194**, 45 (2011).
C. Mazzucchelli, M. Bischetti, V. D'Odorico, C. Feruglio, J.-T. Schindler, M. Onoue, E. Bañados, G. D. Becker, F. Bian, S. Carniani, R. Decarli, A.-C. Eilers, E. P. Farina, S. Gallerani, S. Lai, R. A. Meyer, S. Rojas-Ruiz, S. Satyavolu, B. P. Venemans, F. Wang, J. Yang, and Y. Zhu, A&A **676**, A71 (2023).
O. E. Kovács, Á. Bogdán, P. Natarajan, *et al.*, ApJL **965**, L21 (2024).
L. J. Furtak, I. Labbé, A. Zitrin, *et al.*, Nature **628**, 57 (2024).
F. Wang, J. Yang, X. Fan, *et al.*, ApJL **907**, L1 (2021).
J. Silk and M. J. Rees, A&A **331**, L1 (1998).
Y. Shen, J. E. Greene, L. C. Ho, W. N. Brandt, K. D. Denney, *et al.*, ApJ **805**, 96 (2015), arXiv:1502.01034 [astro-ph.GA] .
I. Trujillo, A. Ferré-Mateu, M. Balcells, A. Vazdekis, and P. Sánchez-Blázquez, ApJL **780**, L20 (2014), 1310.6367 .
A. Ferré-Mateu, I. Trujillo, I. Martín-Navarro, A. Vazdekis, M. Mezcua, M. Balcells, and L. Domínguez, MNRAS (2017), 10.1093/mnras/stx171, 1701.05197 .
J. Kormendy and L. C. Ho, ARA&A **51**, 511 (2013), arXiv:1304.7762 [astro-ph.CO] .
K. Ito, T. S. Tanaka, K. Shimasaku, M. Ando, M. Onoue, *et al.*, MNRAS **538**, 1501 (2025), arXiv:2408.08492 [astro-ph.GA] .
A. C. Carnall, R. J. McLure, J. S. Dunlop, D. J. McLeod, V. Wild, F. Cullen, *et al.*, Nature **619**, 716 (2023), arXiv:2301.11413 [astro-ph.GA] .
M. Neeleman, M. Novak, B. P. Venemans, F. Walter, R. Decarli, *et al.*, ApJ **911**, 141 (2021), arXiv:2102.05679 [astro-ph.GA] .
M. Onoue, X. Ding, J. D. Silverman, Y. Matsuoka, T. Izumi, *et al.*, Nature Astronomy **9**, 1541 (2025), 2409.07113 .
W. Liu, E. P. Farina, X. Fan, R. Decarli, M. Onoue, *et al.*, ApJ (2026), 10.3847/1538-4357/ae9307, 2606.03037 .
I. Juodžbalis, C. Marconcini, F. D'Eugenio, R. Maiolino, A. Marconi, H. Übler, J. Scholtz, *et al.*, Nature **653**, 1017 (2026), arXiv:2508.21748 [astro-ph.GA] .
L. R. Ivey, F. D'Eugenio, R. Maiolino, Y. Isobe, I. Juodžbalis, *et al.*, MNRAS **550**, stag1220 (2026), 2604.09177 .
V. N. Bennert, M. W. Auger, T. Treu, J.-H. Woo, and M. A. Malkan, ApJ **742**, 107 (2011), arXiv:1102.1975 [astro-ph.CO] .
X. Ding, J. Silverman, T. Treu, A. Schulze, M. Schramm, S. Birrer, *et al.*, ApJ **888**, 37 (2020), arXiv:1910.11875 [astro-ph.GA] .
S. Belli, A. B. Newman, and R. S. Ellis, ApJ **834**, 18 (2017), 1608.00608 .
B. Forrest, G. Wilson, A. Muzzin, D. Marchesini, M. C. Cooper, *et al.*, ApJ **938**, 109 (2022), arXiv:2208.04329 [astro-ph.GA] .
I. Juodžbalis, R. Maiolino, W. M. Baker, S. Tacchella, J. Scholtz, F. D'Eugenio, J. Witstok, *et al.*, Nature **636**, 594 (2024), arXiv:2403.03872 [astro-ph.GA] .
L. Napolitano, M. Castellano, L. Pentericci, C. Vignali, R. Gilli, *et al.*, ApJ **989**, 75 (2025), arXiv:2410.18763 [astro-ph.GA] .
A. J. Taylor, V. Kokorev, D. D. Kocevski, H. B. Akins, F. Cullen, M. Dickinson, S. L. Finkelstein, *et al.*, ApJ **989**, L7 (2025), arXiv:2505.04609 [astro-ph.GA] .
C.-H. Chen, L. C. Ho, R. Li, and M.-Y. Zhuang, ApJ **983**, 60 (2025), 2411.04446 .
Z. Xu, "Structure-motivated phantom-crossing dark energy: joint constraints from DESI DR2, supernovae and growth," Research Square preprint rs-11371459 (v1) (2026b).
P. Tozzi and C. Norman, ApJ **546**, 63 (2001).
R. C. E. van den Bosch, ApJ **831**, 134 (2016), 1606.01246 .
V. F. Baldassare, C. Dickey, M. Geha, and A. E. Reines, ApJL **898**, L3 (2020), 2006.15150 .
J. E. Greene, L. C. Ho, and A. J. Barth, ApJ **688**, 159 (2008), 0810.1972 .
S. J. Kimbrell, A. E. Reines, Z. Schutte, J. E. Greene, and M. Geha, ApJ **911**, 134 (2021), 2103.06289 .
D. J. Eisenstein and W. Hu, The Astrophysical Journal **496**, 605 (1998), astro-ph/9709112 .

APPENDIX A: THE DARK-MATTER CLOCK

A1 The halo clock and the characteristic mass

Consider haloes of fixed mass m_h whose mean physical density evolves as $\rho_h \propto a^{-3\theta}$. Mass conservation fixes their radius and virial equilibrium fixes their dispersion,

$$R_h = \left(\frac{3m_h}{4\pi\rho_h}\right)^{\frac{1}{3}} \propto m_h^{\frac{1}{3}} a^{\theta}, \quad \sigma_v^2 \simeq \frac{Gm_h}{R_h} \propto m_h^{\frac{2}{3}} a^{-\theta}. \tag{A1}$$

Dividing ρ_h by the mean matter density, $\bar{\rho} \propto a^{-3}$, gives the contrast at fixed mass,

$$\Delta_h \equiv \frac{\rho_h}{\bar{\rho}} \propto a^{3(1-\theta)}, \tag{A2}$$

constant for $\theta = 1$ and growing as $a^{3/2}$ for $\theta = 1/2$ after a_*; this enters the cooling fraction (Appendix B4). Simulations show that at the characteristic mass the virial dispersion scales as $\sigma_v^2(m_h^*) \propto m_h^{*1/3}$ (Xu 2026a). Inserting Eq. (A1),

$$m_h^{*2/3} a^{-\theta} \propto m_h^{*1/3} \quad \Longrightarrow \quad m_h^* \propto a^{3\theta}. \tag{A3}$$

Before a_* characteristic haloes collapse at fixed contrast, $\theta = 1$, so $m_h^* \propto a^3$ ($\propto t^2$ in matter domination). After a_* they keep a fixed comoving size, $R_h^* \propto a$; along the characteristic track Eqs (A1) and (A3) give $R_h^* \propto m_h^{*1/3} a^{\theta} \propto a^{2\theta}$, so a fixed comoving size requires $\theta = 1/2$ and then $m_h^* \propto a^{3/2}$ ($\propto t$). Along the track $R_h^* \propto a$ after a_*, while at fixed mass $R_h \propto m_h^{1/3} a^{1/2}$.

A2 The crossover from the primordial spectrum

Bulk energy. In linear theory a density mode δ_k moves matter with velocity amplitude $v_k = aHf\delta_k/k$, where $f = d\ln D/d\ln a$. In matter domination $f = 1$, $\delta \propto a$ and $aH \propto a^{-1/2}$, so $v \propto a^{1/2}$ and the bulk kinetic energy grows as $v^2 \propto a$. A mode of wavenumber k enters the horizon at a_k, where $k = a_k H(a_k)/c$, with contrast of order $A_s^{1/2}$, and grows as a/a_k afterwards. Its velocity is $v_k \simeq aH\, A_s^{1/2}(a/a_k)/k$; using $aH \propto a^{-1/2}$ to write $k = aH(a/a_k)^{1/2}/c$,

$$v_k^2 \simeq A_s\, c^2\, \frac{a}{a_k}. \tag{A4}$$

Later-entering modes move more slowly, and modes entering before equality grow little until $a_{\rm eq}$, so the variance is dominated by modes entering at equality, $a_k \simeq a_{\rm eq} = \frac{\Omega_r}{\Omega_m}$:

$$\langle K_{ph}\rangle = \tfrac{1}{4}\beta_V^2 A_s c^2\, \frac{\Omega_m}{\Omega_r}\, a, \tag{A5}$$

where β_V absorbs the shape of the transfer function and is calibrated on simulations ($\beta_V = 0.84$ for Illustris-1-Dark and 1.7 for Virgo; Xu 2026b).

Virial energy and crossing. The cascade feeds kinetic energy to virialized matter at the constant rate $|\varepsilon_u|$ (Eq. 4), so $\langle K_{pv}\rangle = |\varepsilon_u| t$ with $t = 2a^{3/2}/(3H_0\Omega_m^{1/2})$ in matter domination. The ratio $\langle K_{pv}\rangle/\langle K_{ph}\rangle \propto a^{1/2}$ grows monotonically, so the two cross once:

$$\begin{aligned} &\tfrac{1}{4}\beta_V^2 A_s c^2\, \frac{\Omega_m}{\Omega_r}\, a_* = \frac{2|\varepsilon_u|}{3H_0\Omega_m^{1/2}}\, a_*^{3/2} \\ &\implies a_* = \left(\frac{3}{8}\,\beta_V^2\,\Omega_m^{3/2}\,\frac{A_s}{\Omega_r}\,\frac{c^2 H_0}{|\varepsilon_u|}\right)^2, \end{aligned} \tag{A6}$$

which squared is Eq. (9). Dimensions check: $c^2H_0/|\varepsilon_u|$ is $(\mathrm{m^2\,s^{-2}})(\mathrm{s^{-1}})/(\mathrm{m^2\,s^{-3}})$, dimensionless. For $\Omega_m = 0.311$, $h = 0.677$, $A_s = 2.1\times10^{-9}$, $\Omega_r = 9.1\times10^{-5}$, $|\varepsilon_u| = 3.7\times10^{-7}\,\mathrm{m^2\,s^{-3}}$ and $\beta_V = 1$ the right-hand side is 0.80, so $a_* = 0.64$.

Sensitivity and dark energy. Equation (A6) gives $a_* \propto \beta_V^4|\varepsilon_u|^{-2}A_s^2\Omega_m^3H_0^2\Omega_r^{-2}$. The range $|\varepsilon_u| = (2\text{–}5)\times10^{-7}\,\mathrm{m^2\,s^{-3}}$ alone moves a_* by a factor of $(5/2)^2 \simeq 6$, so the measured crossover, $a_* \simeq 0.5$ in Illustris-1-Dark, is the better anchor. At $a \simeq 0.6$ the cosmological constant is not negligible: replacing growth factor, growth rate, Hubble rate and age by their ΛCDM values lowers the bulk energy at $a = 0.64$ by 20 per cent and the age by 8 per cent, and moves the crossing to $a_* = 0.54$ for the same β_V and ε_u. This assumes that the virial energy keeps growing as $|\varepsilon_u|t$ once Λ matters and that β_V carries over.

A3 Rate profiles and the two levels

The profile. With $\sigma^2 = \gamma G m_r/r$ and $m_r = \frac{4\pi}{3}\bar\rho_r r^3$, Eq. (1) becomes

$$\begin{aligned} \varepsilon(r) &= \frac{\sigma^3}{r} = \frac{1}{r}\left(\gamma\frac{Gm_r}{r}\right)^{\frac{3}{2}} \\ &= \left(\tfrac{4\pi}{3}\gamma G\bar\rho_r\right)^{\frac{3}{2}} r^2 = (\gamma G m_r)^{\frac{2}{3}}\, t^{-\frac{5}{3}}, \end{aligned} \tag{A7}$$

with $t = r/\sigma$. A flat profile needs $m_r \propto r^{5/3}$. At fixed contrast and fixed comoving scale, $\bar\rho_r \propto a^{-3}$ and $r \propto a$, so $\varepsilon \propto a^{-9/2}a^2 = a^{-5/2}$ for any mass and profile.

Dark matter haloes. Inside the scale radius r_s the cascade profile is $\rho = \beta_r\varepsilon_h^{2/3}G^{-1}r^{-4/3}$, and integrating it over the halo gives a mass budget (Xu 2021)

$$m_h = K_h\,\varepsilon_h^{2/3}G^{-1}r_s^{5/3}, \tag{A8}$$

with K_h a profile constant, so $\varepsilon_h = K_h^{-3/2}(Gm_h)^{3/2}r_s^{-5/2}$. The virial rate has the same form with the boundary in place of the scale radius, $\varepsilon_{{\rm v},h} = (Gm_h)^{3/2}R_h^{-5/2}$, and their ratio depends on concentration alone,

$$F_h = \frac{\varepsilon_h}{\varepsilon_{{\rm v},h}} = K_h^{-3/2}\left(\frac{R_h}{r_s}\right)^{5/2} = K_h^{-3/2}c^{5/2}. \tag{A9}$$

At fixed mass $\varepsilon_h \propto m_h^{*-2/3} \propto a^{-2\theta}$ and $\varepsilon_{{\rm v},h} \propto R_h^{-5/2} \propto a^{-5\theta/2}$, so $F_h \propto a^{\theta/2}$ and $c \propto a^{\theta/5}$, in agreement with the scale radius of the cascade theory, $r_s \propto m_h^{1/3}m_h^{*4/15}$, for which $c \propto a^\theta m_h^{*-4/15} \propto a^{\theta/5}$.

Gas cores. The core cascade has the same structure with the common rate. Inside its edge, $\sigma^2 = \gamma_r G m_r/r$ and $\sigma^3/r = \varepsilon_b$ give $m_r = \gamma_r^{-1}\varepsilon_b^{2/3}G^{-1}r^{5/3}$, the same function of r in every host. A core of mass M_b is the part of this profile within R_b,

$$\begin{aligned} &R_b \propto \varepsilon_b^{-2/5}M_b^{3/5}, \qquad \sigma_b \propto \varepsilon_b^{1/5}M_b^{1/5}, \\ &\Delta_b \propto \frac{M_b}{\bar\rho R_b^3} \propto \varepsilon_b^{6/5}\bar\rho^{-1}M_b^{-4/5}, \end{aligned} \tag{A10}$$

the scalings measured in Table B1. At fixed mass, $\varepsilon_b \propto a^{-5/2}$ gives $R_b \propto a$ and Δ_b constant in time. Outside a core with flat dispersion, $m_r \propto r$ and $\varepsilon \propto r^{-1}$, which gives Eq. (18).

APPENDIX B: GAS COOLING FRACTION

B1 The virial rate at fixed contrast

A structure of mass M and contrast Δ has $R = (\frac{3M}{4\pi\Delta\bar\rho})^{1/3}$. Inserting this in Eq. (3),

$$\varepsilon_{\rm v} = (\gamma GM)^{3/2}R^{-5/2} = \gamma^{3/2}G^{3/2}M^{2/3}\left(\tfrac{4\pi}{3}\Delta\bar\rho\right)^{5/6}. \tag{B1}$$

Dimensions check: $(G\rho)^{1/2}$ is an inverse time and $(GM)^{2/3}(G\rho)^{5/6}$ is $\mathrm{m^2\,s^{-3}}$. At fixed mass and contrast only $\bar\rho \propto (1+z)^3$ changes, so $\varepsilon_{\rm v} \propto (1+z)^{5/2}$, Eq. (14), with no cosmological parameter. Across cores at fixed time, the measured $\Delta_b \propto M_b^{-4/5}$ cancels the factor $M_b^{2/3}$, which is why $\varepsilon_{{\rm v},b}$ does not depend on mass.

B2 Energy budget

For a gas component with net energy loss Γ per unit mass the Layzer–Irvine equation becomes (Xu 2025)

$$\frac{dE_y}{dt} + H\,(2K_p + P_y) = -\Gamma, \qquad E_y = K_p + P_y. \tag{B2}$$

Take $K_p = \mathcal{K}a^x$ and $P_y = -\mathcal{P}a^x$ with $\mathcal{K},\mathcal{P} > 0$. With $\dot a = aH$, $dE_y/dt = x(\mathcal{K}-\mathcal{P})Ha^x$ and $H(2K_p + P_y) = (2\mathcal{K} - \mathcal{P})Ha^x$, so

$$\Gamma = [\mathcal{P}(1+x) - \mathcal{K}(2+x)]\,Ha^x. \tag{B3}$$

For virialized gas, $2K_p = -P_y$, i.e. $\mathcal{K} = \mathcal{P}/2$, and the bracket reduces to $x\mathcal{P}/2$, which is Eq. (15). Fixed contrast has $\sigma^2 \propto$

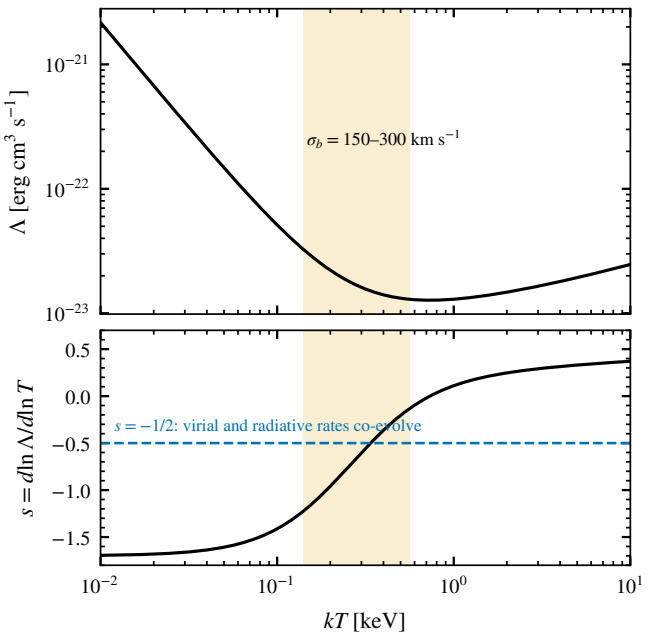


Figure B1. Cooling function of collisional-ionization-equilibrium plasma at $Z = 0.3Z_\odot$ (top) and its logarithmic slope (bottom), from the fit of Tozzi and Norman (2001) to Sutherland and Dopita (1993) (Eq. B6). Shaded: virial temperatures of cores with $\sigma_b = 150$–$300\,\mathrm{km\,s^{-1}}$ ($\mu = 0.6$). The virial and radiative rates evolve alike if $s = -1/2$ (dashed).

a^{-1}, so $x = -1$ and $\Gamma = -H|P_y|/2 = -HK_p$: a net gain of HK_p. Relative to the virial rate, with $K_p = \frac{3}{2}\sigma_b^2$,

$$\frac{|\Gamma|}{\varepsilon_{\mathrm{v},b}} = \frac{\frac{3}{2}H\sigma_b^2}{\sigma_b^3/R_b} = \tfrac{3}{2}\,Ht_b, \tag{B4}$$

which is 1.5×10^{-3} for $t_b = 14$ Myr and 2.3×10^{-3} for 22 Myr ($H = H_0$).

B3 Cooling consistency

For gas of fixed mass with radius $r \propto a^{\theta_g}$, $n \propto a^{-3\theta_g}$, $T \propto \sigma^2 \propto a^{-\theta_g}$ and the virial rate scales as $a^{-5\theta_g/2}$. The radiative rate per unit mass is $\propto n\Lambda(T) \propto a^{-\theta_g(3+s)}$, so the two keep a fixed ratio for any θ_g only if $s = -1/2$. Per unit gas mass,

$$\varepsilon_{\mathrm{rad}} = \frac{n_e}{n_\mathrm{H}}\frac{n_i}{n_\mathrm{H}}\frac{X\,n_\mathrm{H}\,\Lambda(T)}{m_p}, \tag{B5}$$

with $n_e/n_\mathrm{H} = 1.165$, $n_i/n_\mathrm{H} = 1.080$ and $X = 0.71$ for $Z = 0.3Z_\odot$, and $kT = \mu m_p\sigma^2$ with $\mu = 0.6$. We use the fit of Tozzi and Norman (2001) to Sutherland and Dopita (1993),

$$\Lambda = \left[C_1(kT)^{-\alpha} + C_2(kT)^{\beta} + C_3\right]\times10^{-22}\,\mathrm{erg\,cm^3\,s^{-1}}, \tag{B6}$$

with kT in keV, $C_1 = 8.6\times10^{-3}$, $C_2 = 5.8\times10^{-2}$, $C_3 = 6.3\times10^{-2}$, $\alpha = 1.7$ and $\beta = 0.5$. At $\sigma_b = 210\,\mathrm{km\,s^{-1}}$ ($kT = 0.28$ keV), $\Lambda = 1.7\times10^{-23}\,\mathrm{erg\,cm^3\,s^{-1}}$, and $\varepsilon_{\mathrm{rad}} = 10^{-4}\,\mathrm{m^2\,s^{-3}}$ requires $n_\mathrm{H} = 0.11\,\mathrm{cm^{-3}}$.

B4 The cooling fraction

Demand and supply. Gas accreted by a halo is heated by infall to $\sim V_h^2$ and must radiate it within $t_h = R_h/V_h$, a specific power $V_h^2/t_h = \varepsilon_{\mathrm{v},h}$. Once it joins the galaxy, at R_e, it radiates at $\varepsilon_e = \sigma_e^3/R_e$ per unit mass. The cooling fraction compares the two:

$$f_{\mathrm{cool}} = \frac{\varepsilon_e\,t_h}{V_h^2} = \frac{\varepsilon_e}{\varepsilon_{\mathrm{v},h}}. \tag{B7}$$

Only its time dependence is used; in absolute terms the ratio exceeds unity, and Q is normalized at a_*.

In terms of contrasts. Applying Eq. (B1) to the galaxy boundary ($\sigma_e^2 = \gamma_e GM/R_e$) and the halo boundary ($\gamma = 1$), the factors of G and $\bar\rho$ cancel:

$$f_{\mathrm{cool}} = \gamma_e^{3/2}\left(\frac{M}{m_h}\right)^{2/3}\left(\frac{\Delta_e}{\Delta_h}\right)^{5/6}. \tag{B8}$$

Both boundaries share a contrast across masses, and $M/m_h \simeq 0.02$, so f_{cool} does not depend on mass. In time, Δ_e is constant (H1) and $\Delta_h \propto a^{3(1-\theta)}$, so

$$f_{\mathrm{cool}} \propto a^{-\frac{5}{2}(1-\theta)} = \begin{cases}\mathrm{const}, & a < a_*,\\ a^{-5/4}, & a > a_*.\end{cases} \tag{B9}$$

Energy and time. Writing each virial rate as $\sigma^2/t_{\mathrm{dyn}}$ splits Eq. (B7) into two factors,

$$f_{\mathrm{cool}} = \frac{\sigma_e^2}{V_h^2}\times\frac{t_h}{t_e}, \qquad t_e = \frac{R_e}{\sigma_e}. \tag{B10}$$

The first compares the energy the galaxy's gas carries with the energy the accreted gas must lose, $\propto(\Delta_e/\Delta_h)^{1/3}$; the second compares the time available with the galaxy's own dynamical time, $\propto(\Delta_e/\Delta_h)^{1/2}$. After a_* they fall as $a^{-1/2}$ and $a^{-3/4}$, giving $q = 1/2 + 3/4 = 5/4$.

Alternatives. Each variant changes one ingredient of Eq. (B7). (i) The energy factor alone gives $q = 1/2$. (ii) The core rate in place of the boundary rate, $f_{\mathrm{cool}} = \varepsilon_{\mathrm{v},b}/\varepsilon_{\mathrm{v},h}$, gives the same q but a factor $m_h^{-2/3}$, because cores share a rate rather than a contrast. (iii) The halo plateau in place of the halo virial rate, $f_{\mathrm{cool}} = \varepsilon_e/\varepsilon_h \propto a^{-5/2+2\theta}$, has no constant phase before a_*. (iv) An evolving galaxy contrast, $\Delta_e \propto a^x$, gives $q = \frac{5}{4} - \frac{5}{6}x$. The disc level measured at fixed stellar mass, $(1+z)^{2.08\pm0.19}$, corresponds to $x \simeq 0.5$ and $q \simeq 0.8$; at fixed dynamical mass the JADES hosts follow $(1+z)^{5/2}$, i.e. $x = 0$ (Section 2.5).

APPENDIX C: BLACK HOLES

C1 The luminosity ceiling

Confining force. Gas within radius r, of enclosed mass m_r, mean density $\bar\rho_r = 3m_r/(4\pi r^3)$ and dispersion σ_r, with $\sigma_r^2 = \gamma_r G m_r/r$, exerts on the sphere of that radius the force

$$4\pi r^2\bar\rho_r\sigma_r^2 = \frac{3m_r\sigma_r^2}{r} = \frac{3\sigma_r^4}{\gamma_r G}. \tag{C1}$$

In the core cascade, $\sigma_r^3/r = \varepsilon_b$, so $m_r = \sigma_r^5/(\gamma_r G\varepsilon_b)$. The sphere of influence is where $m_r = M_{\mathrm{BH}}$: $\sigma_\bullet^5 = \gamma_r G\varepsilon_b M_{\mathrm{BH}}$, at $r_\bullet = \sigma_\bullet^3/\varepsilon_b$.

Ceiling. Radiation delivers momentum at the rate L/c, and clears the confining gas only if this exceeds the force there:

$$\frac{L_X}{c} = \frac{3\sigma_\bullet^4}{\gamma_r G} \implies L_X = 3\gamma_r^{-1/5}\,\varepsilon_b^{4/5} G^{-1/5} c\, M_{\rm BH}^{4/5}, \tag{C2}$$

which is Eq. (29) with $x_r = 3\gamma_r^{-1/5}$ (W, as required). Using the local density, $\rho(r) = (4\pi r^2)^{-1} dm_r/dr = \frac{5}{9}\bar{\rho}_r$ for $m_r \propto r^{5/3}$, instead gives $x_r = \frac{5}{3}\gamma_r^{-1/5}$.

Relation to the Eddington limit. For electron scattering, $\kappa = 0.04\,{\rm m^2\,kg^{-1}}$, the Eddington luminosity is $L_{\rm Edd} = 4\pi G c M_{\rm BH}/\kappa$, i.e. $L_{\rm Edd}/M_{\rm BH} = 6.3\,{\rm m^2\,s^{-3}}$, so that

$$\frac{L_X}{L_{\rm Edd}} = \frac{x_r\,\varepsilon_b^{4/5}\,\kappa}{4\pi G\,(G M_{\rm BH})^{1/5}}, \tag{C3}$$

which is 1.07 for $\varepsilon_b = 10^{-4}\,{\rm m^2\,s^{-3}}$, $M_{\rm BH} = 10^8 {\rm M}_\odot$ and $x_r = 15$. With $\sigma_\bullet^2 = \gamma_r G M_{\rm BH}/r_\bullet$, the confining force of Eq. (C2) is $3\gamma_r G M_{\rm BH}^2/r_\bullet^2$, and Eq. (C3) becomes

$$\begin{aligned} \frac{L_X}{L_{\rm Edd}} &= \gamma_r^{6/5} x_r\,\tau_\bullet = 3\gamma_r\tau_\bullet, \\ \tau_\bullet &\equiv \frac{\kappa M_{\rm BH}}{4\pi r_\bullet^2} \simeq 0.27\,(1+z)^2 \left(\frac{M_{\rm BH}}{10^8 {\rm M}_\odot}\right)^{-1/5}. \end{aligned} \tag{C4}$$

Here $\tau_\bullet$ is the electron-scattering optical depth of the mass within the sphere of influence, spread over that sphere; equivalently, it is the weight of that mass in units of the Eddington force $L_{\rm Edd}/c$. The Eddington limit balances radiation and gravity on each unit of mass, $\kappa F/c = g$, which holds in smooth gas bound by the black hole alone. Eq. (C2) instead balances the radiation's total momentum, L/c, against the weight of all the confining gas. Radiation delivers that full momentum when the gas absorbs it: when it is optically thick and clumpy, so that radiation escapes between dense clumps, or dusty, so that each photon is absorbed once. For the force-balance normalization, $L_X = \tau_\bullet L_{\rm Edd}$: where $\tau_\bullet > 1$, at $z \gtrsim 1$ for $10^8 {\rm M}_\odot$, the gas is heavy enough to hold radiation above $L_{\rm Edd}$. Where $\tau_\bullet < 1$, as for massive black holes at low redshift, dust-free gas absorbs only a fraction $\tau_\bullet$ of the momentum, and the limit is $\max(L_{\rm Edd}, L_X)$; a ceiling below $L_{\rm Edd}$ then requires dust. The calibrated $x_r = 15$ raises L_X by a factor of 4 at fixed $\tau_\bullet$.

C2 Growth at a fraction of the ceiling

If a black hole accretes at a time-averaged fraction $f_{\rm duty}$ of the ceiling with radiative efficiency η_r, $\dot{M}_{\rm BH} = f_{\rm duty}(1-\eta_r)L_X/(\eta_r c^2) \propto M_{\rm BH}^{4/5}$, so $M_{\rm BH}^{1/5}$ grows at a rate independent of mass and, with $\varepsilon_b = \varepsilon_{b0}(1+z)^{5/2}$ (Xu 2025),

$$M_{\rm BH}^{1/5}(t) = M_0^{1/5} + \frac{f_{\rm duty}(1-\eta_r)x_r\varepsilon_{b0}^{4/5}}{5\eta_r\, c\, G^{1/5}} \int_{t_0}^{t} (1+z')^2 dt'. \tag{C5}$$

For comparison, the equivalent steady Eddington ratio is $\lambda_{\rm E} = \ln(M_{\rm BH}/M_0)\,t_S/(t-t_0)$ with Salpeter time $t_S = 50$ Myr for $\eta_r = 0.1$.

C3 Normalization of the $M_{\rm BH}$–σ relation

Derivation. A core of dispersion σ_b and radius R_b has mass $M_b = \sigma_b^2 R_b/(\gamma_r G)$ and cascade rate $\varepsilon_{{\rm v},b} = \sigma_b^3/R_b$. Its gas, a fraction f_g of the mass, dissipates the power

$$P_g = \varepsilon_{{\rm v},b} f_g M_b = \frac{f_g \sigma_b^5}{\gamma_r G}, \tag{C6}$$

in which the core radius cancels (Eq. 32). Equating a coupled fraction η_c of the Eddington luminosity to this power,

$$\eta_c \frac{4\pi G c}{\kappa} M_{\rm BH} = P_g, \tag{C7}$$

gives $M_{\rm BH} = A\sigma_b^5$ with A from Eq. (33), or equivalently

$$\eta_c = \frac{f_g\,\kappa}{4\pi\gamma_r G^2 c\, A}. \tag{C8}$$

APPENDIX D: LEVELS IN GALAXIES: CALIBRATION AND THE JWST HOSTS

Local reference. For local star-forming galaxies we use integrated ionized-gas dispersions from the SAMI survey, $\log\sigma = 0.48\log M_* - 2.98$ within 3 arcsec and $\log S_{0.5} = 0.40\log M_* - 2.14$ within R_e, each flat below its kinematic limit (26.4 and 22.4 km s^{-1}, reached at $M_* = 10^{9.17}$ and $10^{8.73}{\rm M}_\odot$; Barat *et al.* 2019, Table 2), and the late-type size–mass relations of van der Wel *et al.* (2014) and Shen *et al.* (2003). The four combinations bracket the reference $\varepsilon_e(M_*)$ (Table D1). Below the kinematic limits the true dispersions may be lower, so the reference rates are upper limits and F lower limits there.

Table D1. Local disc level: integrated gas dispersion, effective radius, rate at R_e and $F = \varepsilon_{b0}/\varepsilon_e$ with $\varepsilon_{b0} = 10^{-4.14}\,{\rm m^2\,s^{-3}}$ (Table B1), against stellar mass. †At the kinematic limit of SAMI: F is a lower limit.

$\log M_*$	$\sigma_{\rm gas}$ [km s^{-1}]	R_e [kpc]	$\log\varepsilon_e$	F
7.5	22–26†	1.1	−6.50 to −6.27	135–227
8.0	22–26†	1.3–1.5	−6.62 to −6.34	159–302
8.5	22–26†	1.6–2.0	−6.75 to −6.42	187–403
9.0	26–29†	1.8–2.7	−6.66 to −6.37	169–328
9.5	38–46	2.2–3.6	−6.31 to −5.85	51–147
10.0	66–72	2.7–4.8	−5.71 to −5.33	15–37
10.5	115	3.4–6.5	−5.12 to −4.84	5–9

High-redshift samples. Extreme emission-line galaxies at $z = 1.4$–2.3 have integrated [O III] or Hα dispersions and HST sizes (Maseda *et al.* 2014). For JADES non-AGN galaxies at $z = 5.5$–7.4 we recover the integrated dispersion from the published dynamical mass with $M_{\rm dyn} = 6\sigma^2 R_e/G$ (de Graaff *et al.* 2024). For the five JADES broad-line AGN hosts with dispersions, sizes and dynamical masses (Table 3 of Maiolino *et al.* 2024a; we exclude the dual-BLR candidate 73488 and ID 954, whose stellar mass is flagged) we remove the +0.175 dex gas-to-stellar correction applied by Maiolino *et al.* (2024a), so that all samples use gas dispersions. Lyman-break analogues at $z \simeq 0.2$, compact local starbursts, use $S_{0.5} = (\sigma^2 + 0.5 v_{\rm shear}^2)^{1/2}$ (Gonçalves *et al.* 2010); they lie $+1.09 \pm 0.16$ dex above normal local discs of the same mass, as expected for galaxies with smaller F. For matching at fixed dynamical mass we apply the same estimator, $M_{\rm dyn} = 6\sigma^2 R_e/G$, to the local reference and to every sample (the reference is evaluated over $7.0 \le \log M_* \le 11.6$ and held at its edge value below; extending it to $\log M_* = 6.8$ changes the JADES-host offset from +1.79 to +1.77).

Why bulge-level comparisons fail for the JADES hosts. Compared with local *bulges*, the JADES hosts have σ^3/R_e close to the local core rate, not the +1.98 dex the clock predicts. The two quantities sit at different levels. The dispersion is the integrated narrow-line width and R_e the half-light radius of the whole compact, disc-like galaxy (Sérsic $n \simeq 1$); at $z \simeq 5$ the common core profile would enclose $\sim 10^{10}\mathrm{M}_\odot$ within 150 pc, ten times the dynamical mass, so R_e lies far outside any core. Compared with local discs of the same mass, the hosts lie $+1.98 \pm 0.12$ dex higher at fixed M_* and $+1.79 \pm 0.07$ dex at fixed $M_{\rm dyn}$, against +1.98 predicted (Table 1).

Luminous quasar hosts. For the 18 undisturbed [C II] hosts of Neeleman *et al.* (2021) we take $\sigma = (GM_{\rm dyn}/8R_e)^{1/2}$ from the published dynamical masses. At fixed $M_{\rm dyn}$ (median $10^{10.5}\mathrm{M}_\odot$ in our estimator) they lie $+1.42 \pm 0.09$ dex above local discs, against +2.16 for fixed contrast.

Local black-hole hosts at R_e. The local reference for the JADES test uses stellar dispersions and whole-galaxy radii, as for the high-redshift hosts. From the compilation of van den Bosch (2016) (dynamical, maser and reverberation masses; σ_e within R_e; 2MASS K-band R_e) we keep the 28 disc hosts on the $M_{\rm BH}$–σ relation ($|\Delta \log A| < 0.5$): Table B1 spirals, maser hosts, or galaxies with concentration $R_{80}/R_{20} < 4$. We add 18 low-mass disc and S0 hosts of broad-line AGN with stellar dispersions (Baldassare *et al.* 2020; Greene *et al.* 2008) whose total half-light radii we solve from HST bulge and disc decompositions (Kimbrell *et al.* 2021; Greene *et al.* 2008), excluding the nuclear point source. Every dynamical mass is $5\sigma^2 R/G$. For the JADES hosts we keep the +0.175 dex gas-to-stellar correction of Maiolino *et al.* (2024a) (for the dual-BLR candidate 10013704 we use the more massive BLR, $\log M_{\rm BH} = 7.50$; with the other, 5.65, the mean $\Delta \log A$ of the five hosts falls from +0.76 to +0.39 against the Table B1 zero point); without it the offset at fixed σ falls by 0.11–0.14 dex. Spheroids sort by compactness, not by type: M32 has $F = 0.7$, at the core level, whereas diffuse dwarf ellipticals have $F = 19$–116, like discs.

Three gas rates. The disc level can be compared with the halo-gas plateau of Eq. (12). With halo masses from the stellar-to-halo relation of Moster *et al.* (2013), the 46 local disc hosts have $\log \varepsilon_e/\varepsilon_g = -0.64 \pm 0.07$ (scatter 0.45 dex, no trend with halo mass; $M(<R_e)/m_h \simeq 0.02$): at $z \simeq 0$ the disc boundary sits a factor of about 4 below the halo-gas plateau, so $F \simeq 4/\nu = 4(m_h^*/m_h)^{2/3}$. The two rates then are different: at fixed structure F is constant, whereas the halo-gas law, tied to m_h^*, rises faster towards high redshift. Extreme emission-line galaxies at $z \simeq 1.6$ lie 1.3 dex and the JADES hosts 2.0 dex below the law, close to what fixed F implies; beyond $z \simeq 2$ the law requires gas faster than the halo can bind, and the disc level and the core are the only gas levels.

APPENDIX E: MODEL DETAILS

Halo model. Equation (21) is evaluated for Planck 2018 ($\Omega_m = 0.3111$, $\Omega_b = 0.0490$, $h = 0.6766$, $n_s = 0.9665$, $\sigma_8 = 0.8102$) on $10^5 \le m_h \le 10^{16}\mathrm{M}_\odot$ (400 logarithmic bins), with the linear power spectrum of Eisenstein and Hu (1998), the ΛCDM growth factor and the Sheth–Tormen mass function ($A = 0.3222$, $a = 0.707$, $p = 0.3$, $\delta_c = 1.686$), all implemented in the released code. The atomic-cooling floor is $m_{\rm min} = 10^8 h^{-1}\mathrm{M}_\odot[10/(1.98(1+z))]^{3/2}$ ($T_{\rm vir} = 10^4\,\mathrm{K}$, $\mu = 0.6$) and the mean accretion rate is $\dot m_h = 46.1\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}\,(m_h/10^{12}\mathrm{M}_\odot)^{1.1}(1+1.11z)E(z)$. The characteristic mass is $m_h^* = 1.2\times10^{13}\mathrm{M}_\odot\,a^{3/2}$ after a_* and $1.2\times10^{13}\mathrm{M}_\odot\,a_*^{3/2}(a/a_*)^3$ before; the cutoff mass is $M^* = m_h^*(a_*)(a/a_*)^3$ after a_* and $m_h^*(a_*)(a/a_*)^{15/4}$ before. Fits at $z \le 8$ minimize the mean squared deviation of $\log\rho_{\rm SFR}$ from MD14 at 81 redshifts (the local index is the derivative at $z = 0$; the slope is the least-squares slope over 101 redshifts uniform in $0 \le z \le 1$); for $a_* = 0.64$ the best parameters are $f_0 = 0.27$, $f_{\rm max} = 1$ (at the bound) and $M_q = 7.2\times10^{12}\mathrm{M}_\odot$. The baryon accretion rate onto all haloes above $m_{\rm min}$ peaks at $z \simeq 5.8$, declines at $z < 1$ as $(1+z)^{1.3}$, and falls by a factor of 4.0 from $z = 1.9$ to 0 (2.7 for haloes above $10^{11}\mathrm{M}_\odot$).

High-redshift fits. At $z > 8.5$ the efficiency of Eq. (23) is kept fixed ($Q = 1$) and modified by Eq. (25), (26), (27) or (28). All free parameters of a model are fitted jointly by minimizing $\chi^2_{\rm J} + \chi^2_{\rm LF}$, with the upper (lower) error of each point used when the model lies above (below) it; $\chi^2_{\rm J}$ uses the errors on $\log\rho_{\rm UV}$ and $\chi^2_{\rm LF}$ the linear errors on ϕ. Three luminosity-function points without a published lower error are given symmetric errors. Gate parameters are profiled: the 95 per cent bound is where χ^2 rises by 4 with the other parameters refitted.

Derived observables. Luminosity functions follow from $\phi(M_{\rm UV}) = (dn/d\ln m_h)|d\ln m_h/dM_{\rm UV}|$, with $M_{\rm UV} = 51.60 - 2.5\log_{10}[\mathrm{SFR}/\mathcal{K}_{\rm UV}]$ and $\mathrm{SFR} = f_b f_* \dot m_h$. For bursty star formation, ϕ is evaluated on a uniform grid of 0.01 mag and convolved with a Gaussian of width $2.5\sigma_{\rm SFR}$, which preserves the median SFR; its density to $M_{\rm UV} = -17$ is computed from the convolved function. Abundance matching (Section 4) equates the observed cumulative number density at each magnitude with $n(>m_h)$. Reionization integrates

$$\frac{dQ_{\rm H\,II}}{dt} = \frac{\dot n_{\rm ion}}{\bar n_{\rm H}} - \frac{Q_{\rm H\,II}}{t_{\rm rec}}, \qquad \dot n_{\rm ion} = \frac{f_{\rm esc}\xi_{\rm ion}\rho_{\rm SFR}}{\mathcal{K}_{\rm UV}}, \tag{E1}$$

from $z = 20$ over all haloes above $m_{\rm min}$, with $\bar n_{\rm H} = 1.9\times10^{-7}\,\mathrm{cm}^{-3}$, $t_{\rm rec}^{-1} = C\alpha_B\bar n_{\rm H}(1+Y/4X)(1+z)^3$, $C = 3$ and $\alpha_B = 2.6\times10^{-13}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$; τ_e includes singly ionized helium with hydrogen and doubly ionized helium at $z < 3$. For bursty star formation the mean SFR of the log-normal is used at $z > 8.5$.

Bulge and black-hole data. Bulge rates are σ_b^3/R_b for the 177 galaxies of Table B1 of Xu (2025); slopes are least-squares fits in $\log M_b$ with bootstrap errors (4000 resamples). Morphological classes (E, S0, spiral, dwarf) for the black-hole regressions are assigned from standard catalogues and released with the code. For the SDSS-IV quasars, virial masses and continuum luminosities are taken in the order Hβ with L_{5100} ($z < 0.8$), Mg II with L_{3000} ($z < 2.3$) and C IV with L_{1350}, with bolometric corrections 9.26, 5.15 and 3.81; envelopes are percentiles in redshift bins of width 0.3 between $z = 0.4$ and 4.3 (or mass bins of 0.25 dex), with bootstrap errors.

Table F1. JWST UV luminosity densities ($M_{\rm UV} < -17$, log of $\mathrm{erg\,s^{-1}\,Hz^{-1}\,Mpc^{-3}}$) and star-formation densities (log of $\mathrm{M_\odot\,yr^{-1}\,Mpc^{-3}}$). D24: Donnan *et al.* (2024); M26: McLeod *et al.* (2026). Model columns ($a_* = 0.64$): constant efficiency (no boost) and the sharp onset of Eq. (25).

z	Ref.	$\log \rho_{\rm UV}$	$\log \rho_{\rm SFR}$	no boost	onset
9	D24	$25.29^{+0.05}_{-0.05}$	-2.65	-2.80	-2.67
10	D24	$25.12^{+0.07}_{-0.08}$	-2.82	-3.17	-2.75
11	D24	$25.12^{+0.14}_{-0.20}$	-2.82	-3.56	-2.81
12.5	D24	$24.64^{+0.18}_{-0.32}$	-3.30	-4.19	-3.23
13.5	M26	$24.41^{+0.15}_{-0.16}$	-3.53	-4.62	-3.59
14.5	D24	$23.92^{+0.27}_{-0.81}$	-4.02	-5.06	-3.97
15.8	M26	$23.22^{+0.21}_{-1.28}$	-4.72	-5.72	-4.53

APPENDIX F: NOTATION, DATA AND CODE

Table F4 lists every symbol. The released package contains the input data (JWST densities and luminosity functions, Table B1 with morphological classes, the black-hole compilation, the high-redshift host and disc-level compilations, the local black-hole hosts at R_e and the further tests of the ceiling), a shared module for cosmology and the halo model, one script per analysis and figure, the JSON results from which every number in the text is taken, and a script that checks the numbers quoted in this paper against those results. Table F1 lists the JWST densities.

Table F2. Inputs and their evidence. F: energy flux measured in simulations; S: scaling measured in simulations or data; C: calibrated or fitted constant; H: hypothesis tested in this paper.

Symbol	Meaning	Value (source)	Evidence
ε_u	dark-matter cascade rate	$-(2\text{–}5)\times10^{-7}\,\mathrm{m^2\,s^{-3}}$ (Illustris, Virgo)	F
m_h^*	characteristic halo mass, Eq. (6)	$1.2\times10^{13}\mathrm{M_\odot}$ at $z=0$; $\propto a^{3\theta}$	S
M^*	cutoff mass of the mass function	$=m_h^*$ at a_*; $\propto a^{15/4}$ before, a^3 after	S
θ	dark-matter clock, Eq. (8)	1 before a_*, 1/2 after	S
a_*	crossover	0.5 (Illustris-1-Dark); 0.64 (Eq. 9); 0.50–0.58 (galaxies, Section 3)	S, C
ε_g	plateau rate of halo gas (1–10 kpc), Eq. (12)	$\varepsilon_b(z)(m_h/m_h^*)^{2/3}$ (Illustris-1, $0\le z\le2$)	S
$\varepsilon_b(z)$	common core rate, Eq. (13) (H2)	$10^{-4}(1+z)^{5/2}\,\mathrm{m^2\,s^{-3}}$; $10^{-4.14}$ and $\propto M_b^{-0.09\pm0.05}$ at $z=0$ (Table B1)	S
θ_b	gas clock: fixed contrast in time (H1)	1 at all epochs	H
Δ_b	core contrast	$\propto M_b^{-4/5}$ at fixed time; $\simeq2\times10^7$ at $M_b=10^{11}\mathrm{M_\odot}$	S
Δ_e	contrast at R_e of discs	$\propto M^{-0.04\pm0.09}$ at $z=0$ (46 local disc hosts); $\simeq10^{6.6}$	S
F	core-to-R_e rate ratio, Eq. (18)	$\simeq1$ spheroids; $\propto M^{-2/3}$ for discs: $\simeq102$ at $M_*=10^9\mathrm{M_\odot}$, 10 at $10^{10.5}\mathrm{M_\odot}$ (stellar dispersions)	S
γ_r	virial coefficient, $\sigma_b^2=\gamma_rGM_b/R_b$	1/3 (Table B1), $\simeq1$ (Illustris gas)	C
$n_{\rm th}$	density for star formation faster than supernova feedback	$3\times10^3\,\mathrm{cm^{-3}}$ (Dekel *et al.* 2023)	C
n_0, $\mathcal{B}$, p	onset density, boost size, sharpness, Eqs (24)–(25)	$2.0\,\mathrm{cm^{-3}}$, 4.4, 6.0 (fitted; $z_{\rm on}=10.5$)	C
x_r	normalization of the luminosity ceiling, Eq. (29)	6 (quasar envelope) to 15 (LID-568); 1.7–3.7 (force balance)	C
A	$M_{\rm BH}/\sigma_b^5$, Eq. (33)	$10^{-3.14}\mathrm{M_\odot}\,(\mathrm{km\,s^{-1}})^{-5}$ (Table B1); constant in time	C

Table F3. Two cascades, two levels each, at fixed mass (dark matter: halo of mass m_h; gas: galaxy of dynamical mass M, M_b for the core). $\theta=1$ before the crossover a_* and 1/2 after it (Eq. 8). Brackets: measured exponents (Illustris for haloes; Table B1 for cores; 46 local disc hosts with $M=5\sigma_e^2R_e/G$ for R_e, Appendix D). The plateaus are dual: the dark-matter plateau scales with mass and time, the gas core is common to all masses and scales with time. Both boundaries share a contrast across masses, so their rates scale as $M^{2/3}$. The ratio of the levels is dual as well: F_h changes with time but not with mass, F with mass but not with time. Evidence codes as in Table F2.

	Dark matter (halo)		Gas (galaxy)	
	plateau, $r<r_s$	boundary, R_h	core, $r\lesssim r_c\simeq R_b$	boundary, R_e
rate	$\varepsilon_h=\lvert\varepsilon_u\rvert(m_h/m_h^*)^{2/3}$	$\varepsilon_{{\rm v},h}=V_h^3/R_h=\varepsilon_h/F_h$	$\varepsilon_b=\varepsilon_{b0}(1+z)^{5/2}$	$\varepsilon_e=\sigma_e^3/R_e=\varepsilon_b/F$
clock exponent	$\theta=1$, then 1/2	$\theta=1$, then 1/2	1 (H1)	1 (H1)
mass dependence	$m_h^{2/3}$	$m_h^{2/3}$	none $[M_b^{-0.09\pm0.05}]$ (H2)	$M^{2/3}$ $[M^{0.63\pm0.08}]$
time dependence at fixed mass	${m_h^*}^{-2/3}\propto a^{-2\theta}$	$a^{-5\theta/2}$: $a^{-5/2}$, then $a^{-5/4}$	$a^{-5/2}$	$a^{-5/2}$ [JADES − local discs at fixed M: $+1.98\pm0.13$, predicted $+1.98$]
what is common	the flux $\lvert\varepsilon_u\rvert$ (along m_h^*)	the contrast	the rate, across hosts	the contrast, across masses
contrast in time	–	const, then $\propto a^{3/2}$	const	const
contrast across masses	–	$\Delta_h\simeq200$	$\Delta_b\propto M_b^{-4/5}$ $[M_b^{-0.91\pm0.07}]$, $\sim10^7$	$\propto M^0$ $[M^{-0.04\pm0.09}]$, $\sim10^{6.6}$
σ, R across masses	$\sigma^2(r)=(\varepsilon_h r)^{2/3}$	$V_h, R_h\propto m_h^{1/3}$	$M_b^{1/5}$, $M_b^{3/5}$ $[0.18\pm0.01$, $0.64\pm0.02]$	$M^{1/3}$, $M^{1/3}$ $[0.33\pm0.02$, $0.35\pm0.03]$
inner profile	$\rho\propto\varepsilon_h^{2/3}G^{-1}r^{-4/3}$,	steepening towards $r^{-5/2}$ outside r_s	$\rho\propto\varepsilon_b^{2/3}G^{-1}r^{-4/3}$, fixed in comoving r/a	$\sigma\simeq$ const, $\varepsilon\propto r^{-1}$ outside r_c
ratio of the levels	$F_h=\varepsilon_h/\varepsilon_{{\rm v},h}\propto c^{5/2}$ (Eq. 11): independent of mass, $\propto a^{\theta/2}$		$F=\varepsilon_b/\varepsilon_e\simeq R_e/r_c$: constant in time (H1), $\propto M^{-2/3}$; $\simeq1$ for spheroids, $\simeq77$ for discs of $M_*\simeq10^{9.4}\mathrm{M_\odot}$ at $z\simeq0$ and at $z\simeq5$	
evidence	F, S	S	S at $z=0$; H in time	S at $z=0$; H in time

Table F4. Notation. Rates are per unit mass ($\mathrm{m^2\,s^{-3}}$); the last column gives the defining equation or first use.

Symbol	Meaning	Units	Defined
$\varepsilon = \sigma^3/r$	cascade rate; $\varepsilon(r)$ its profile	$\mathrm{m^2\,s^{-3}}$	Eqs (1), (2)
$\varepsilon_{\rm v}$	virial rate at the boundary of a structure	$\mathrm{m^2\,s^{-3}}$	Eq. (3)
ε_u	dark-matter cascade rate (constant, < 0)	$\mathrm{m^2\,s^{-3}}$	Eq. (4)
ε_h, $\varepsilon_{{\rm v},h}$, F_h	halo plateau rate, halo virial rate, their ratio	$\mathrm{m^2\,s^{-3}}$; –	Eqs (5), (11)
ε_g, ν	halo-gas plateau rate; $\nu = (m_h/m_h^*)^{2/3} = \varepsilon_g/\varepsilon_b$	$\mathrm{m^2\,s^{-3}}$; –	Eq. (12)
$\varepsilon_b(z)$, ε_{b0}	common core (bulge) rate; its value at $z = 0$	$\mathrm{m^2\,s^{-3}}$	Eqs (12), (13)
$\varepsilon_{{\rm v},b}$	virial rate of a core, σ_b^3/R_b	$\mathrm{m^2\,s^{-3}}$	Eq. (14)
ε_e, F, r_c	rate at R_e; core-to-R_e ratio $F = \varepsilon_b/\varepsilon_e$; core radius	$\mathrm{m^2\,s^{-3}}$; –; kpc	Eq. (18)
ε_0	rate implied by the Table B1 zero points, $B_0/(\gamma_r G A_0)$	$\mathrm{m^2\,s^{-3}}$	Eq. (31)
θ, θ_b	clock exponent of dark matter; of gas cores	–	Eq. (7)
a_*, z_*	crossover scale factor and redshift	–	Eq. (9)
m_h, m_h^*, M^*	halo mass; characteristic mass; cutoff mass	$\mathrm{M_\odot}$	Eq. (6)
R_h, V_h, r_s, c	halo radius, circular velocity, scale radius, concentration	kpc, $\mathrm{km\,s^{-1}}$, kpc, –	Eq. (11)
M_b, R_b, σ_b, t_b	core (bulge) dynamical mass, radius, dispersion, R_b/σ_b	$\mathrm{M_\odot}$, kpc, $\mathrm{km\,s^{-1}}$, Myr	Eq. (14)
Δ_b, Δ_e, Δ_h	contrast of a core; at R_e; of a halo	–	Eqs (14), (A2)
γ, γ_r	virial coefficient, $\sigma^2 = \gamma G m/r$; its value for cores (1/3)	–	Eq. (2)
K_p, P_y, E_y, Γ	peculiar kinetic, potential, total energy; net loss	$\mathrm{m^2\,s^{-2}}$; $\mathrm{m^2\,s^{-3}}$	Eqs (4), (15)
A_s, β_V	primordial amplitude; velocity-integral factor	–	Eq. (9)
Λ, s	cooling function; $d\ln\Lambda/d\ln T$	$\mathrm{erg\,cm^3\,s^{-1}}$; –	Eq. (B6)
$f_{\rm cool}$, Q, q	cooling fraction; quench term; its exponent	–	Eqs (19), (20)
f_*, f_0, $f_{\rm max}$, M_q	star-formation efficiency; normalization, saturation, quenching mass	–, –, –, $\mathrm{M_\odot}$	Eq. (23)
$f_*^{\rm std}$, $f_{\rm pk}$, M_c, β_*, γ_*	double-power-law efficiency and its parameters	–	Eq. (22)
ζ, δ	exponents of the two benchmarks	–	Section 3
n, n_0, $n_{\rm th}$, $z_{\rm on}$	star-forming gas density; its value at $z = 0$; threshold; onset	$\mathrm{cm^{-3}}$	Eq. (24)
$\mathcal{B}$, p, S	boost size, sharpness, switch	–	Eq. (25)
$\mathcal{G}$, $M_{\rm g}$, q_h, q_M	mass gate, gate mass, gate mass in units of m_h^* or M^*	–, $\mathrm{M_\odot}$	Eq. (26)
$\sigma_{\rm SFR}$	dispersion of $\log$ SFR at fixed halo mass	dex	Eq. (27)
$f_{\rm FFB}$, $M_{\rm FFB}$, w	feedback-free efficiency, mass, weight	–, $\mathrm{M_\odot}$, –	Eq. (28)
$\mathcal{K}_{\rm UV}$, $\rho_{\rm SFR}$, $\rho_{\rm UV}$, ϕ	UV conversion; SFR and UV densities; luminosity function	see text	Eq. (21)
τ_e, $f_{\rm esc}$, $\xi_{\rm ion}$	Thomson depth; escape fraction; ionizing efficiency	–, –, $\mathrm{Hz\,erg^{-1}}$	Appendix E
L_X, x_r	luminosity ceiling; its normalization	W; –	Eq. (29)
ℓ, ω	$L/(M_{\rm BH}\varepsilon_b)$; $(GL/c^5)^{1/4}$	–	Eq. (30)
$\sigma_\bullet$, $r_\bullet$	dispersion and radius at the sphere of influence	$\mathrm{km\,s^{-1}}$, pc	Appendix C1
$f_{\rm duty}$, η_r, $\lambda_{\rm E}$	fraction of the ceiling; radiative efficiency; Eddington ratio	–	Eq. (C5)
A, B	$M_{\rm BH}/\sigma^5$; $M_{\rm BH}/M_b$	$\mathrm{M_\odot\,(km\,s^{-1})^{-5}}$; –	Eq. (31)
$B_{\rm core}$, $B_{\rm dyn}$	B for core masses; for dynamical masses at R_e	–	Eq. (34)
P_g, f_g, η_c, κ	core dissipation power; gas fraction; coupling; opacity	W; –; –; $\mathrm{m^2\,kg^{-1}}$	Eqs (32), (33)