The Astrophysical Journal, 1011:46 (25pp), 2026 November 1 https://doi.org/10.3847/1538-4357/aeaa36
© 2026. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

PANORAMIC: The Dawn of Massive Quiescent Galaxies. I. Number Density and Cosmic Variance from 1000 arcmin^2 NIRCam Imaging

Zhiyuan Ji[1], Christina C. Williams[1,2], Peter Behroozi[1], Andrea Weibel[3], Christian Kragh Jespersen[4], Pascal A. Oesch[3,5], Rachel Bezanson[6], Katherine E. Whitaker[5,7], Jenny E. Greene[8], Gabriel Brammer[9,10], Pratika Dayal[11,12,13], Ivo Labbé[14], Sinclaire M. Manning[7], Pierluigi Rinaldi[15], Mengyuan Xiao[3], and Yunchong Zhang[6]

[1] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA; zhiyuanji@arizona.edu
[2] NSF–DOE Vera C. Rubin Observatory/NSF NOIRLab, 950 North Cherry Avenue, Tucson, AZ 85719, USA
[3] Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland
[4] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[5] Cosmic DAWN Center, Niels Bohr Institute, University of Copenhagen, Jagtvej 128, København N, DK-2200, Denmark
[6] Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA
[7] Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
[8] Department of Astrophysical Sciences, 4 Ivy Lane, Princeton University, Princeton, NJ 08540, USA
[9] Cosmic Dawn Center (DAWN), Denmark
[10] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, København N, DK-2200, Denmark
[11] Canadian Institute for Theoretical Astrophysics, 60 St George Street, University of Toronto, Toronto, ON M5S 3H8, Canada
[12] David A. Dunlap Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada
[13] Department of Physics, 60 St George Street, University of Toronto, Toronto, ON M5S 3H8, Canada
[14] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Melbourne, VIC 3122, Australia
[15] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
Received 2026 April 3; revised 2026 September 17; accepted 2026 September 18; published 2026 October 26

Abstract

We measure the number density and field-to-field variance of massive quiescent galaxies at $z \sim 3$–8 using the JWST/NIRCam pure-parallel imaging survey PANORAMIC together with archival observations, covering an area of 0.28 deg^2 ($\sim$1000 arcmin^2) in at least six filters. We identify quiescent galaxy candidates at $z \gtrsim 3$ with $M_* \gtrsim 10^{10}\,M_\odot$, comprising 101 galaxies in a gold sample of high-confidence candidates and 137 in a more inclusive silver sample. We measure their evolving comoving number density, finding (1.5 versus 3.1) $\times\ 10^{-5}\,\mathrm{Mpc}^{-3}$ at $z = 3$–4 for the gold and silver samples, respectively, and a decline by more than a factor of 20 by $z \sim 6$. Comparisons with empirical models and cosmological simulations show that widely used frameworks underpredict the abundance of massive quiescent galaxies at $z \gtrsim 4$ by $\gtrsim$1 dex, indicating that current implementations of early star formation, feedback, and quenching do not produce enough early quenched systems. With 34 independent sight lines, we present the first direct empirical measurement of field-to-field variance for quiescent galaxies at $z > 3$, finding a high cosmic variance of $\sigma_{\mathrm{CV}} \approx 0.7 \pm 0.3$. This exceeds predictions from abundance-matched mock catalogs, suggesting that early quiescent galaxies are more strongly clustered, and more likely to be found near one another or in more biased regions, than expected in current galaxy-formation models. Any successful model for the emergence of early massive quiescent galaxies must reproduce both their abundance evolution and their imprint on the large-scale distribution.

Unified Astronomy Thesaurus concepts: Galaxy formation (595); Galaxy evolution (594); Galaxy quenching (2040); High-redshift galaxies (734)

1. Introduction

The formation of massive quiescent galaxies remains a major theoretical challenge in astrophysics. Stellar archeology indicates that their stars formed in an extreme and rapid burst, followed by rapid truncation of star formation before $z > 3$ (D. Thomas et al. 2005, 2010). Such early assembly implies star formation efficiencies in massive galaxies that were more than an order of magnitude higher in the first Gyr (M. Xiao et al. 2024, 2025; A. de Graaff et al. 2025c) than those typical today (A. Saintonge & B. Catinella 2022). Meanwhile, their lower-redshift descendants remain physically distinct from the bulk of the star-forming population, typically exhibiting compact, centrally concentrated morphologies and passive evolution over billions of years (P. G. van Dokkum et al. 2008; P. F. Hopkins et al. 2010; M. Y. Grudić et al. 2019).

Decades of work at $z \lesssim 3$ have sought to understand how mass, morphology, and local environment shape the quenching of these systems (e.g., Y.-j. Peng et al. 2010; Y. Peng et al. 2015; Z. Ji et al. 2018; C. K. Jespersen et al. 2025a), yet no complete theoretical picture has emerged. Massive quiescent galaxies exhibit a diversity of quenching timescales (S. Belli et al. 2019; A. C. Carnall et al. 2020), likely linked to the disappearance of cold molecular gas and dust (C. C. Williams et al. 2021; K. E. Whitaker et al. 2021; R. Bezanson et al. 2024). But the empirical signatures of the relevant quenching mechanisms remain elusive, perhaps because they are short-lived and most visible only during the quenching phase itself (M. Park et al. 2023; S. Belli et al. 2024; L. Bugiani et al. 2025). As time passes, other unrelated processes such as merging and rejuvenation (e.g., Z. Ji & M. Giavalisco 2022, 2023;

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

C. Woodrum et al. 2022; K. A. Suess et al. 2023, 2025) complicate the forensic reconstruction of quenching mechanisms. These difficulties motivate direct searches for massive quiescent galaxies during the epoch when they first formed and quenched, i.e., $z \gtrsim 3$. Further, studies of quiescent galaxies at $z > 3$ present a complementary test of the early formation histories inferred from lower-redshift quenched systems.

Pre-JWST, candidate galaxies were confirmed with ground-based facilities out to $z_{\rm spec} = 3.7$, revealing old stellar populations with extremely early formation times (M. Kriek et al. 2016; K. Glazebrook et al. 2017; C. Schreiber et al. 2018; J. Antwi-Danso et al. 2025). However, the limited wavelength coverage from the ground ($\lambda < 2.5\,\mu$m) prevents accurate stellar population modeling beyond $z > 4$, where longer-wavelength spectroscopy is needed to measure the Balmer break and age-sensitive absorption features. The launch of JWST (J. P. Gardner et al. 2023) transformed this field by enabling efficient spectroscopic confirmation of quiescent galaxies at $z > 3$. JWST has now confirmed quiescent galaxies at $z \sim 4$–5 (A. C. Carnall et al. 2024; T. Kakimoto et al. 2024; A. de Graaff et al. 2025c; T. Nanayakkara et al. 2025; Y. Zhang et al. 2025; W. M. Baker et al. 2026; M. L. Hamadouche et al. 2026; K. Ito et al. 2026), with one known example at $z = 7.3$ (A. Weibel et al. 2025a). The improvement in measured formation and quenching timescales enabled by JWST is critical to the development of theories about their formation pathways.

A key missing piece of information is how the abundance of massive quiescent galaxies evolves with redshift, since this provides a direct statistical test of cosmological models for rapid growth and quenching. Unfortunately, massive ($M_* \gtrsim 10^{10}\,M_\odot$) quiescent galaxies at $z > 3$ are extremely rare: current estimates predict only one to two detections at $z > 4$ in a typical $\sim$200 arcmin^2 pencil-beam survey (e.g., F. Valentino et al. 2023). Multiple JWST studies have reported abundances that exceed simulation predictions, but these results are based on only two to six independent sight lines (A. C. Carnall et al. 2023; F. Valentino et al. 2023; S. Alberts et al. 2024a; A. S. Long et al. 2024; W. M. Baker et al. 2025b; E. Merlin et al. 2025; T. A. Russell et al. 2025; Y. Zhang et al. 2025; S. D. Stevenson et al. 2026). Taken together, these studies suggest that current measurements at $z > 3$ are strongly affected by field-to-field variance, likely reflecting strong clustering of this highly biased population (e.g., F. Valentino et al. 2023). Robust constraints on their abundance therefore require surveys that probe both a substantially larger area and many independent sight lines. Yet JWST is inefficient for contiguous wide-field imaging because of large slews, overheads, and the small field of view of the Near Infrared Camera (NIRCam; M. J. Rieke et al. 2023).

Fortunately, JWST can obtain NIRCam imaging in pure-parallel mode, enabling extremely efficient wide-area imaging while naturally sampling independent sight lines, the most effective strategy for characterizing the abundance and spatial distribution of rare, highly biased populations (e.g., C. L. Steinhardt et al. 2021; A. C. Trapp et al. 2022). This strategy has strong precedent from Hubble Space Telescope (HST) parallel programs such as BoRG and HIPPIES (M. Trenti et al. 2011; H. Yan et al. 2011; L. D. Bradley et al. 2012; V. Calvi et al. 2016), which surveyed $\sim$0.4 deg^2 in three to four filters, discovered bright $z \sim 6$–10 galaxies too rare for pencil-beam surveys, and reduced cosmic variance by observing more than 300 independent sight lines (e.g., T. Morishita et al. 2018; J. S. Bridge et al. 2019; S. Rojas-Ruiz et al. 2020; T. Morishita 2021; G. Roberts-Borsani et al. 2022; M. B. Bagley et al. 2024). Although limited in ancillary data, these HST parallel surveys remain an important legacy dataset: some of their brightest $z \sim 6$–9 Lyman-break galaxies have now been spectroscopically confirmed with JWST (G. Roberts-Borsani et al. 2025; S. Rojas-Ruiz et al. 2025), and the parallels provided essential context for interpreting the unexpectedly high abundance of bright galaxies at $z > 10$ (S. Rojas-Ruiz et al. 2026).

NIRCam pure parallels are significantly more powerful. Owing to the dichroic design, they provide 2–3$\times$ the filter coverage of HST parallels (T. Morishita et al. 2025; C. C. Williams et al. 2025). To date, they have nearly doubled the area with $\geqslant$six NIRCam filters (required for robust $z > 4$ selection) and increased the number of blank-field independent sight lines by $\gtrsim$6$\times$. These capabilities enable new constraints on the abundance and clustering of quiescent galaxies at $z > 4$, previously limited by small areas and cosmic variance. These data provide the first characterization of the clustering of this highly biased population, informing the galaxy–halo connection (e.g., B. E. Robertson 2010; A. Weibel et al. 2026a) and the role of large-scale environment in early quenching. Such constraints are essential for interpreting high abundances in rare, small-area samples (C. K. Jespersen et al. 2025b), and clustering measurements are key to understanding quiescent galaxy evolution in a cosmological context (C. K. Jespersen et al. 2025a).

In this paper, we present results using pure-parallel imaging from the PANORAMIC survey (C. C. Williams et al. 2025) combined with archival imaging from JWST Cycles 1–3 programs. The data comprise a total of 1000 arcmin^2 of NIRCam imaging across an unprecedented 34 independent sight lines. The paper is organized as follows. In Section 2, we describe the pure-parallel and archival imaging data. In Section 3, we describe the selection of our $z > 3$ quiescent galaxy sample. In Section 4, we provide our measurement of the abundance evolution of quiescent galaxies at $z > 3$, utilizing 34 independent sight lines (a sixfold increase over previous studies, enabling the most robust mitigation of cosmic variance to date). Finally, in Section 5, we provide a first characterization of the clustering properties of $z > 3$ quiescent galaxies by measuring the empirical field-to-field variance, which can be compared to simulations to gain insight into the underlying dark matter halo population that quiescent galaxies trace. We jointly interpret our abundance and clustering constraints, and their implications for the early quenching of massive galaxies, in Section 6.

Throughout this paper, we use "environment" in a purely descriptive sense, to denote the large-scale spatial context traced by cosmic variance, rather than any specific quenching mechanism. We adopt the AB magnitude system and the ΛCDM cosmology with Planck Collaboration et al. (2020) parameters, i.e., $\Omega_m = 0.315$ and $h = H_0/(100\,{\rm km\,s^{-1}\,Mpc^{-1}}) = 0.673$.

2. Data

This work is based on publicly available imaging from JWST/NIRCam and HST. To ensure robust photometric redshift constraints, we restrict our analysis to regions with coverage in at least six NIRCam filters (F115W, F150W, F200W, F277W, F356W, F444W; C. C. Williams et al. 2025).

We compute the total sky coverage of all NIRCam imaging considered in this study while carefully accounting for NIRCam's interdetector and intermodule gaps as well as the complex survey geometry. The resulting total survey area is 0.28 deg^2, corresponding to approximately 1000 arcmin^2, the largest to date for a JWST study of $z > 3$ quiescent galaxies.

Our dataset combines JWST/NIRCam Cycle 1 pure-parallel observations from the PANORAMIC survey (C. C. Williams et al. 2025) with archival JWST imaging from the following extragalactic legacy fields.

1. The Extended Groth Strip (EGS) field, with data from CEERS (ERS-1345; PI: Finkelstein; S. L. Finkelstein et al. 2025), GO-2279 (PI: Naidu), DDT-2750 (PI: Arrabal-Haro), and GO-2234 (PI: Bañados).
2. The Ultra Deep Survey (UDS) and COSMOS fields, with data from PRIMER (GO-1837; PI: Dunlop), COSMOS-Web (GO-1727; PIs: Kartaltepe & Casey; C. M. Casey et al. 2023), GO-1810 (PI: Belli), GO-1840 (PI: Álvarez-Márquez), and DDT-6585 (PI: Coulter).
3. The GOODS fields, with data from JADES (GTO-1180, GTO-1181, PI: Eisenstein; GTO-1210, PI: Luetzgendorf; D. J. Eisenstein et al. 2026), FRESCO (GO-1895; PI: Oesch; P. A. Oesch et al. 2023), CONGRESS (GO-3577; PIs: Egami & Sun), GO-2079 (PI: Finkelstein), JEMS (GO-1963; PIs: Williams, Maseda & Tacchella; C. C. Williams et al. 2023), JADES Origins Field (GO-3215; PI: Eisenstein; D. J. Eisenstein et al. 2025), GTO-1264 (PI: Robledo), GO-4762 (PI: Fujimoto), and SAPPHIRES (GO-6434; PI: Egami).
4. The A2744 cluster, with data from UNCOVER (PIs: Labbé & Bezanson; R. Bezanson et al. 2024), GLASS (ERS-1324; PI: Treu; T. Treu et al. 2022), DDT-2756 (PI: Chen), All the Little Things (GO-3516; PIs: Naidu & Matthee; R. P. Naidu et al. 2024), and Medium Bands, Mega Science (GO-4111; PI: Suess; K. A. Suess et al. 2024).
5. The JWST North Ecliptic Time-Domain Field, with data from PEARLS (GTO-2738; PIs: Windhorst & Hammel; R. A. Windhorst et al. 2023).

All imaging mosaics were consistently reduced using the software GRIZLI (G. Brammer 2023; F. Valentino et al. 2023), starting from level 2 calibrated data products retrieved from the Mikulski Archive for Space Telescopes (MAST).

The photometric catalog used in this study is from A. Weibel et al. (2026b; to which we refer the reader for details), constructed from uniformly reduced, point-spread function (PSF)–homogenized mosaics following the method described in A. Weibel et al. (2024) and C. C. Williams et al. (2025). Source detection was performed on inverse-variance-weighted stacks of the NIRCam Long-Wavelength images (F277W, F356W, and F444W) with the software SExtractor (E. Bertin & S. Arnouts 1996) in its dual-image mode. Fluxes were measured in $r = 0\farcs16$ circular apertures[16] on PSF-matched mosaics of all available JWST and HST bands and scaled to total using Kron apertures with additional PSF-wing corrections. Empirical PSFs were derived from isolated stars where available; otherwise, model PSFs from WEBBPSF (M. D. Perrin et al. 2014) were used after being rotated to match the position angle of the observations.

In Appendix A, we compare our photometry in GOODS-S with the latest public release from the JADES team (B. D. Johnson et al. 2026; B. E. Robertson et al. 2026) and find excellent agreement between the two measurements, demonstrating the robustness of our photometric procedure. To avoid mixing catalogs released by different surveys, we adopt our own photometric catalog from A. Weibel et al. (2026b) across the entire sample, ensuring uniform photometric measurements across all fields.

3. Sample Selection

Our parent sample comprises $\sim 10^6$ sources after removing artifacts and those with compromised photometry using `junk_flag` and `hugekron_flag` from A. Weibel et al. (2026b).[17] Running PROSPECTOR spectral energy distribution (SED) fitting (Section 3.2) to the full catalog is computationally prohibitive.[18] We therefore first identified quiescent candidates using established color-selection techniques, followed by PROSPECTOR fitting and visual inspection, a strategy similar to other high-z quiescent galaxy studies (e.g., Z. Ji & M. Giavalisco 2022; W. M. Baker et al. 2025a).

Our final $z \sim 3$–8 quiescent sample comprises a gold subsample (minimal classification uncertainty) and a more inclusive silver subsample. Their spatial distributions are shown in Figure 1. Below, we describe the selection procedure in detail.

3.1. Initial Selection

The goal of this initial step is to select quiescent candidates at $z > 3$ as broadly and comprehensively as possible. To this end, our selection *combines* candidate quiescent galaxies identified using two commonly employed color-selection techniques: (1) NIRCam-based observed colors and (2) rest-frame colors. Observed colors are effective at isolating broad spectral features (e.g., the Balmer break) with minimal assumptions about galaxy SED templates, but with a fixed filter set, our observed color selection only works within a specific redshift range. In contrast, selecting quiescent galaxies based on rest-frame colors can, in principle, be applied at any redshift. However, deriving rest-frame colors depends on the robustness of SED template fitting (e.g., EAzY; G. B. Brammer et al. 2008). As each method has its own strengths and limitations, we decided to combine quiescent galaxies selected by either method to construct an initial comprehensive sample of $z > 3$ quiescent candidates.

We select quiescent galaxies at $z \geqslant 3$ using NIRCam observed colors with the selection developed by A. S. Long et al. (2024). Those authors proposed two selection boxes: one with conservative color thresholds that minimize contamination from dusty star-forming galaxies but may miss dusty quiescent systems and a more inclusive "red selection wedge"

[16] The typical half-light size of $z > 3$ massive quiescent galaxies is $\lesssim$0.5 kpc (Z. Ji et al. 2026), corresponding to $\lesssim\farcs1$. Using an $r = \farcs16$ circular aperture—approximately the FWHM of the F444W PSF—captures most of the light from these quiescent galaxies while maintaining high-S/N photometry. This is essential for obtaining accurate color measurements and, consequently, reliable SED fitting. Missing light is accounted for through our aperture correction based on the Kron flux.

[17] The A. Weibel et al. (2026b) catalog has a `use_phot` flag, which—in addition to flagging sources with `junk_flag` or `hugekron_flag`—flags sources with `stellar_flag = True`. We intentionally retained sources with `stellar_flag` in the parent sample, as quiescent galaxies at $z > 3$ have typical effective radii $\lesssim\farcs1$ (Z. Ji et al. 2026), which is about half the F444W angular resolution.

[18] With our current PROSPECTOR setup, a single galaxy requires $\sim$10–15 hr.

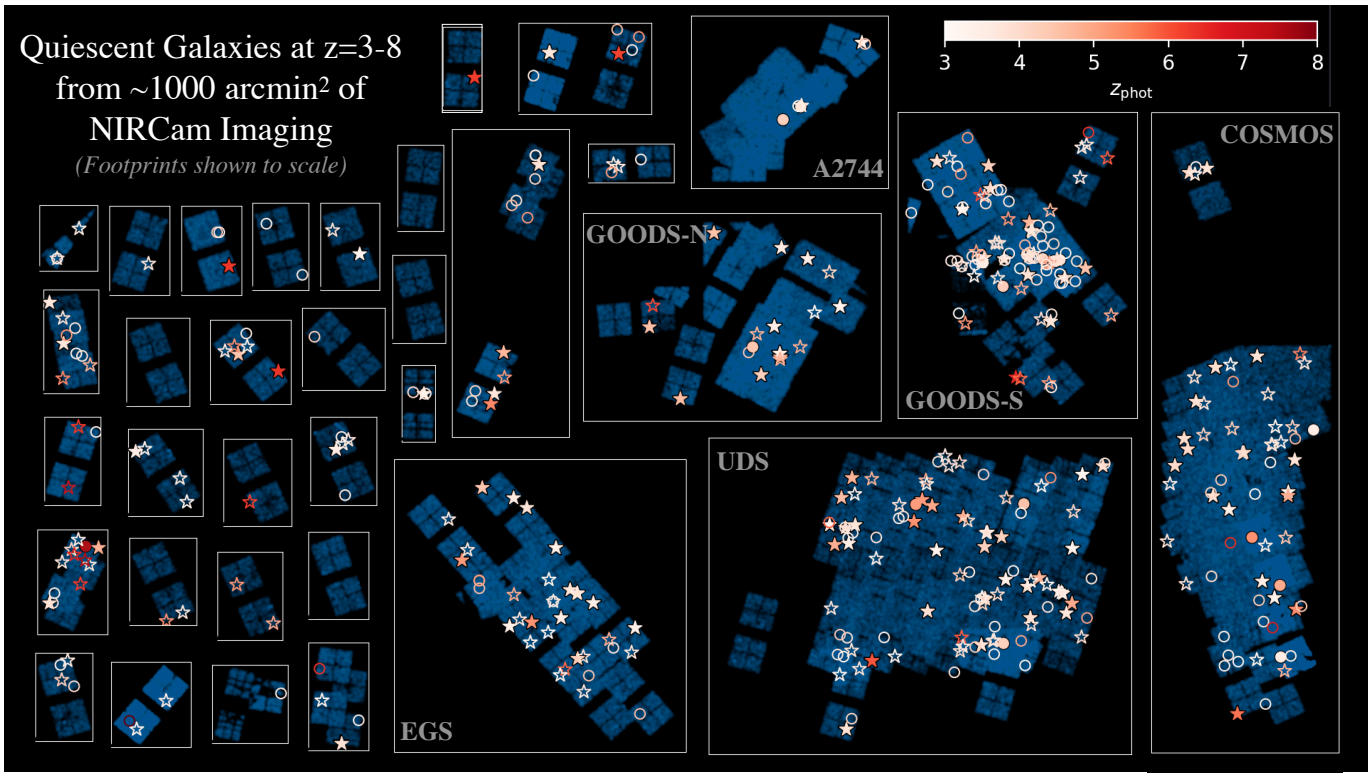

Figure 1. Quiescent galaxies at $z \sim 3$–8 (a total of 406) identified in this work from ~1000 arcmin^2 of JWST/NIRCam imaging. Galaxies in the gold and silver samples are shown as filled and open symbols, respectively (see Section 3). Massive systems with $M_* \geqslant 10^{10}\,M_\odot$—the main focus of this study (101 in the gold sample and 137 in the silver sample)—are shown as stars, while less massive ones are shown as circles. Footprints separated by $<1^\circ$ are combined. The single NIRCam pointings come from the pure-parallel program PANORAMIC (C. C. Williams et al. 2025). Because our selection requires at least six-filter NIRCam coverage, in the COSMOS field, only the regions around the PRIMER survey are included. The strong impact of cosmic variance (i.e., spatial clustering) on the distribution of $z \geqslant 3$ quiescent galaxies is visually evident in this plot.

(see the Appendix of A. S. Long et al. 2024). In light of recent discoveries of early massive quiescent galaxies with significant dust reservoirs (Z. Ji et al. 2024a; D. J. Setton et al. 2024; J. C. Siegel et al. 2025), we chose to use the red selection wedge defined as

$$\begin{aligned}
&(\mathrm{F150W} - \mathrm{F277W}) < 1.5 + 6.25 \times (\mathrm{F277W} - \mathrm{F444W})\\
&(\mathrm{F150W} - \mathrm{F277W}) > 1.15 - 0.5 \times (\mathrm{F277W} - \mathrm{F444W})\\
&(\mathrm{F150W} - \mathrm{F277W}) > -0.6 + 2.0 \times (\mathrm{F277W} - \mathrm{F444W}).
\end{aligned} \tag{1}$$

For the rest-frame color-selection method, we adopt the rest-frame *UVJ* technique (R. J. Williams et al. 2009). However, the original *UVJ* selection box was defined for $z < 2$, while quiescent galaxies at $z > 3$ are typically younger and thus exhibit bluer colors (e.g., S. Belli et al. 2019; A. C. Carnall et al. 2023; F. Valentino et al. 2023; W. M. Baker et al. 2025a). To account for this, we use an extended *UVJ* selection that expands the original box to more completely capture $z > 3$ quiescent candidates (J. Antwi-Danso et al. 2023; F. Valentino et al. 2023; W. M. Baker et al. 2025b) as follows:

$$\begin{aligned}
&(U - V) > 0.70\\
&(V - J) < 1.83\\
&(U - V) > 0.88 \times (V - J) + 0.26.
\end{aligned} \tag{2}$$

Our criteria closely follow those in F. Valentino et al. (2023), who expanded the original *UVJ* selection box by 0.23 mag. The only difference is the $(U - V)$ threshold: F. Valentino et al. (2023) adopted $(U - V) > 1$, whereas we further relax it to $(U - V) > 0.70$, consistent with spectroscopically confirmed high-redshift quiescent galaxies with young stellar ages (e.g., M. Park et al. 2023; A. Weibel et al. 2025a).[19] We also note that W. M. Baker et al. (2025a) proposed an alternative high-z *UVJ* box (lime dotted line in Figure 2); it allows $(U - V) < 0.7$, but all of their spectroscopically confirmed quiescent galaxies have $(U - V) \gtrsim 0.7$. Our *UVJ* selection is therefore also very similar to that of W. M. Baker et al. (2025a).

We required a signal-to-noise ratio $(\mathrm{S/N}) \geqslant 5$ in the F150W, F277W, and F444W filters to ensure high-quality color measurements and, consequently, reliable SED-fitting results. We imposed a cut of EAzY-derived $z_{\rm phot} \geqslant 3$ from A. Weibel et al. (2026b) and then applied the aforementioned color-selection methods to identify quiescent galaxy candidates. In addition, although they constitute only a small fraction (≈2.8%), some sources in the A. Weibel et al. (2026b) catalog have $z_{\rm phot} = -1$, indicating unreliable EAzY fits that may result from the incomplete set of SED templates used in their EAzY configuration. To ensure we did not overlook $z > 3$ quiescent galaxies due to such failed fits, we also included sources with $z_{\rm phot} = -1$ that satisfy the red selection wedge (Equation (1)).[20]

[19] A. Weibel et al. (2025a) did not report rest-frame colors for the RUBIES $z_{\rm spec} = 7.29$ quiescent galaxy; we measure $(U - V) = 0.9 \pm 0.1$.

[20] Because these sources lack reliable EAzY solutions, the rest-frame *UVJ* selection could not be applied.

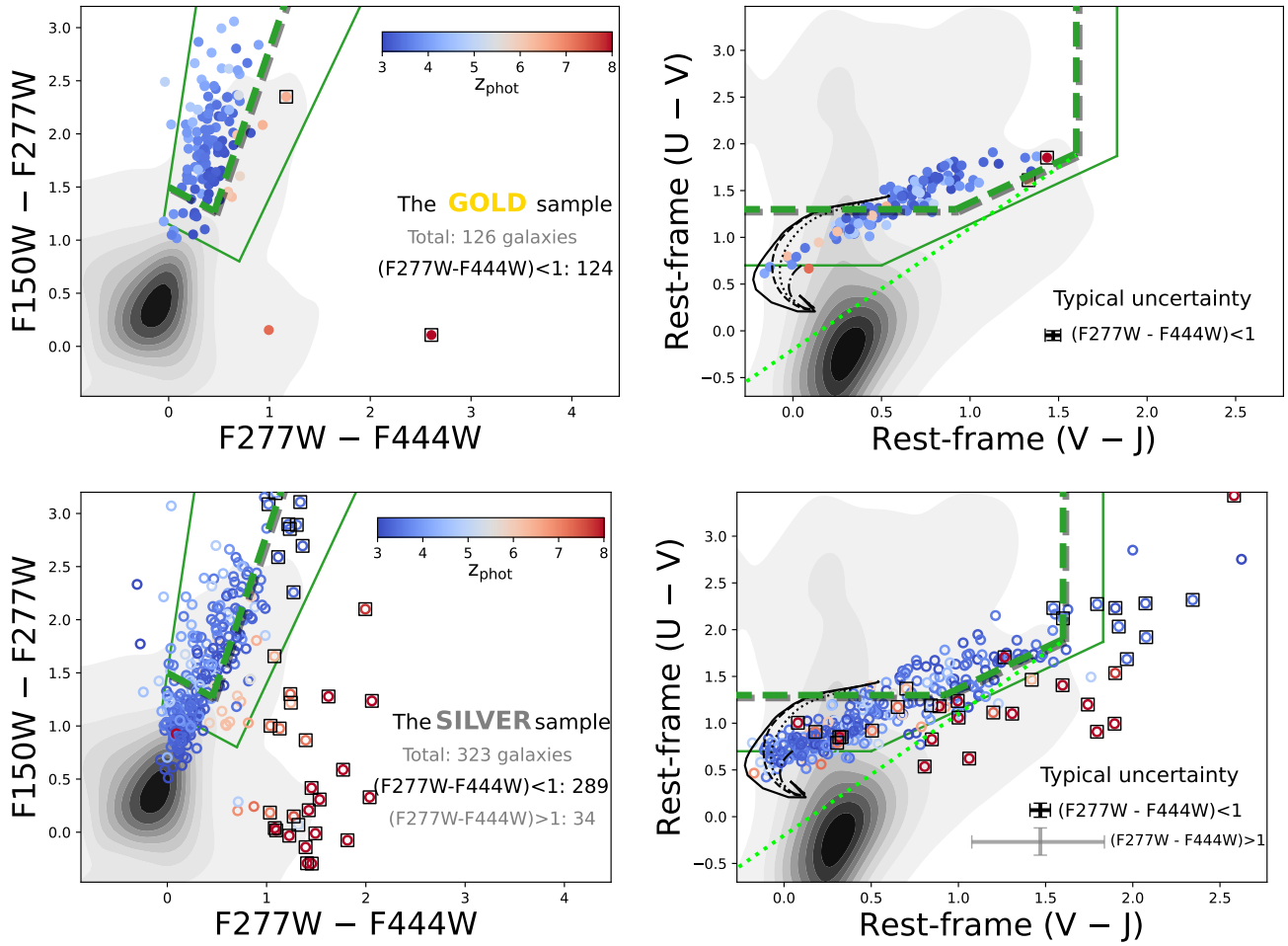

Figure 2. The distribution of $z \sim 3$–8 quiescent galaxies in the observed F150W − F277W versus F277W − F444W (left) and the rest-frame $U - V$ versus $V - J$ (right) color planes. The upper and lower rows show the gold and silver samples, respectively (Section 3.2). Gray contours show all galaxies in the photometric catalog of A. Weibel et al. (2026b) with $z_{\rm phot} > 3$ and S/N > 5 in F150W, F277W, and F444W. The green solid lines represent the red selection wedge (A. S. Long et al. 2024) and the extended *UVJ* selection criteria (which are combined for our initial selection). The green dashed lines represent more conservative selection criteria adopted in this study (Section 3.1). The lime dotted line represents the selection from W. M. Baker et al. (2025a). Sources removed for red colors, i.e., (F277W − F444W) > 1, are marked with open squares. Right panels show typical rest-frame color uncertainties (quantified as the range spanned by the three SED fittings described in Section 3.2). We show these uncertainties separately for silver sample galaxies with F277W − F444W > 1 and F277W − F444W < 1. The black solid, dashed, dotted, and dashed–dotted lines in the *UVJ* plot represent the 0.02–1 Gyr evolutionary tracks of the FSPS stellar population models with $E(B - V) = 0.3$ and a delayed-τ SFH, where $\tau = 10$, 50, 100, and 500 Myr, respectively.

Our initial selection described above yielded 2633 galaxies in total, of which 2148 and 815 satisfied the NIRCam-based observed color selection and the rest-frame *UVJ* selection, respectively. As we will discuss in Section 3.3, these two selections are highly consistent with each other at $z > 3$. Nonetheless, combining the two methods allowed us to identify robust quiescent galaxies that would have been missed if only one method had been applied.

3.2. Refined Selection

In this section, we use SED fitting and visual inspection to refine our quiescent galaxy selection, excluding contaminants and sources whose emission is likely nonstellar. We used the SED-fitting framework PROSPECTOR (B. D. Johnson et al. 2021) in a similar configuration to Z. Ji et al. (2024b). Accurately measuring the star formation rates (SFRs) of quiescent galaxies remains challenging (C. Conroy 2013), as the inferred values are highly sensitive to the assumed star formation history (SFH). We therefore performed SED fitting using three SFH models: a delayed-τ SFH and two nonparametric SFHs, one with a continuity prior and one with a bursty continuity prior. The detailed assumptions adopted in our SED analysis and their comparison are presented in Appendix B.

With the SED-fitting results in hand, we refined the initial selection of quiescent galaxies by applying a redshift-dependent cut on specific SFR (sSFR),

$$\mathrm{sSFR} \leqslant \frac{0.2}{t_H(z)}, \tag{3}$$

which has been commonly used in recent studies of high-redshift quiescent galaxies (e.g., C. Pacifici et al. 2016; A. C. Carnall et al. 2023; W. M. Baker et al. 2025a). To account for the systematic differences in SFR estimates across the SED models, we define two categories.

1. *Gold sample*. A galaxy is assigned to this group if its sSFR satisfies the above threshold in all three SED fits.
2. *Silver sample*. The sSFR threshold is satisfied in one/two of the three fits.

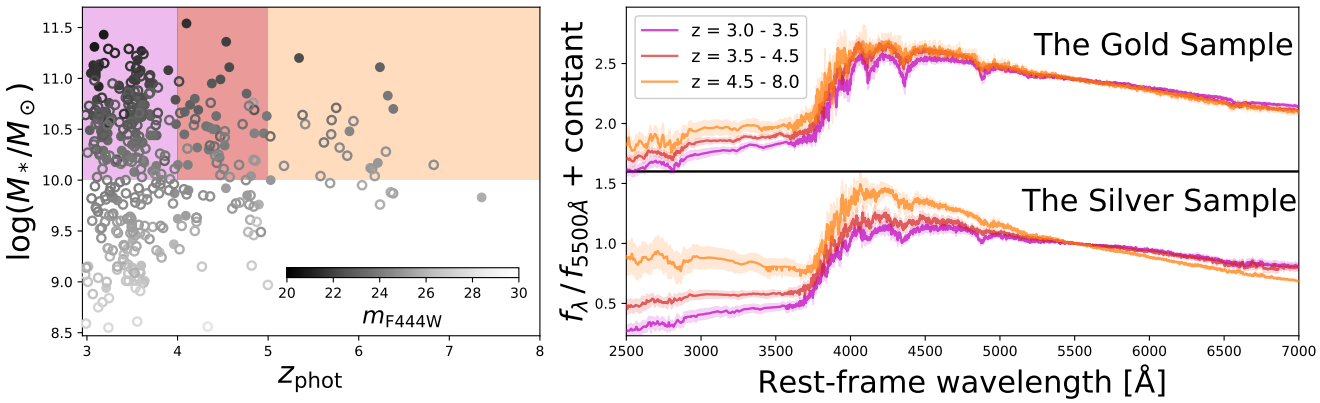

Figure 3. Left: M_* versus z_{phot} for our final samples of quiescent galaxies. The gold sample is shown as filled circles and the silver sample as open circles. This study focuses on massive systems with $M_* > 10^{10}\,M_\odot$. Right: median best-fit SEDs of $>10^{10}\,M_\odot$ quiescent galaxies in three redshift bins: $3 \leqslant z < 3.5$ (magenta), $3.5 \leqslant z < 4.5$ (red), and $z \geqslant 4.5$ (orange). The shaded regions represent the 1σ uncertainties, derived by bootstrapping the galaxy samples in each redshift bin 200 times.

We perform a final visual inspection of the image cutouts for all galaxies in the gold and silver samples. This step enabled us to remove objects with clear imaging artifacts that were not flagged by automated pipelines (e.g., bright PSF spikes or regions strongly affected by cosmic-ray hits). In total, we excluded 20 sources through this procedure. Appendix C presents detailed notes on the removed sources. While silver candidates that satisfy the sSFR threshold in only one versus two of the SED fits may have different levels of reliability, the currently available data are too limited to quantify this difference robustly.

3.2.1. Removal of Sources with Very Red Colors

Finally, we excluded sources from the gold and silver samples with very red colors of F277W − F444W $\geqslant 1$. This decision is motivated by the recent discovery of the so-called little red dots (LRDs; e.g., D. D. Kocevski et al. 2023; J. E. Greene et al. 2024; J. Matthee et al. 2024). Their physical origin remains uncertain, but they are likely dominated by a nonstellar power source that can generate strong Balmer breaks, in some cases appearing similar to quiescent galaxies in photometry (e.g., C. C. Williams et al. 2024; A. de Graaff et al. 2025b; I. Labbe et al. 2025; R. P. Naidu et al. 2026). Because of this ambiguity, their inclusion in quiescent galaxy samples is potentially problematic. Comprehensively identifying LRDs with photometry alone remains challenging. Photometric selections recover only about half of spectroscopically confirmed LRDs, with the remainder missed due to faint rest-UV fluxes, bluer rest-optical colors, or uncertain redshifts (R. E. Hviding et al. 2025). This highlights the incompleteness of purely photometric criteria, though a conservative cut of F277W − F444W $\geqslant 1$ has been demonstrated to remove $\gtrsim$60% of spectroscopic LRDs (Section 5.1 of R. E. Hviding et al. 2025).

As Figure 2 shows, the F277W − F444W > 1 population (open squares) clusters preferentially at $z_{phot} \sim 7$, because the Balmer break redshifts into the F277W − F444W color. At this redshift, using broadband photometry alone, it is difficult to distinguish a rising or flat stellar continuum from strong emission line boosts by, e.g., LRDs. Removing red sources may also remove legitimate $z \gtrsim 7$ quiescent galaxies. We discuss this issue in detail in Section 6.2.

We note that removing sources with F277W − F444W > 1 alone does not eliminate all possible LRD-like contaminants from the final sample (R. E. Hviding et al. 2025). Cross-matching our final catalog with the A. deGraaff et al. (2026) compilation of LRDs shows that three such sources remain in the baseline sample, including the Cliff (A. de Graaff et al. 2025b). Although this represents only a very small (<1%) fraction of the total, it shows that there is some degeneracy between quiescent galaxies and LRDs in terms of the photometry. We refer the reader to A. Weibel et al. (2026, in preparation) for a photometric search of sources like the Cliff and further discussion of the degeneracy with quiescent galaxies. Cross-matching with their sample suggests a small potential contamination of <5% of our quiescent sample and reveals four more spectroscopically confirmed LRDs with broad lines. In the analysis that follows, we remove all seven confirmed broad-line LRD sources from our sample and note that this has no significant impact on our results.

3.3. Final Samples of Quiescent Galaxies

Our final $z > 3$ sample consists of 118 galaxies in the gold sample and 288 galaxies in the silver sample, spanning a stellar-mass range of $10^{8.5}$–$10^{11.5}\,M_\odot$ (Figure 3). This study focuses on the massive population with $M_* \gtrsim 10^{10}\,M_\odot$, where our sample is highly complete, typically 2–5 mag brighter than the corresponding 5σ depths in the shallowest surveys considered here. In this mass regime, we have 101 and 137 galaxies in the gold and silver samples, respectively. The full catalog of 238 massive quiescent galaxy candidates identified in this work can be found on Zenodo: doi:10.5281/zenodo.22132642.

We cross-matched our final samples with robust, spectroscopically confirmed ones from the literature, finding 32 matches at $z > 3$. The details are presented in Appendix D; here, we briefly summarize the key conclusions: 29 spectroscopic confirmations are recovered by our final selection (19 in the gold and 10 in the silver samples). This inclusion rate of $29/32 = 90.6\%$ indicates that our photometric selection is highly complete at the approximately 90% level. Three spectroscopic confirmations are missed by our selection. Two galaxies ($2/32 = 6.25\%$) were excluded because their sSFRs are above our threshold (although still 0.3–0.5 dex below the star-forming main sequence). One galaxy (3.13%)

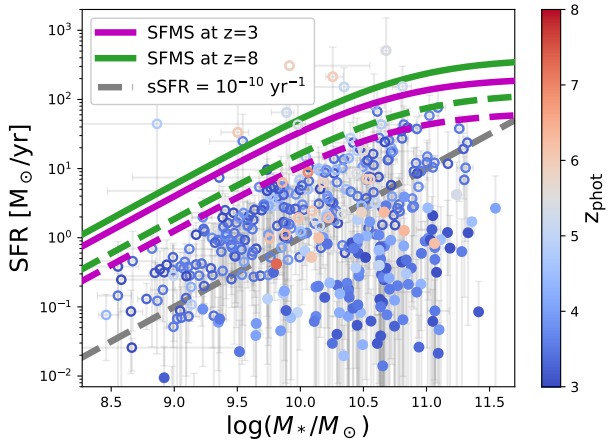

Figure 4. SFR versus M_*. Quiescent galaxies in the gold and silver samples are shown as filled and open circles, respectively. Stellar masses and SFRs are plotted as the mean from the three SED fittings described in Section 3.2, with error bars indicating the full range spanned by these fittings. The magenta and green solid lines show the star-forming main sequence at $z = 3$ and $z = 8$ from P. Popesso et al. (2023), while the corresponding dashed lines mark the relation offset by −0.5 dex. The gray dashed line marks sSFR $= 10^{-10}$ yr^{-1}.

was contaminated by bright neighbors causing problematic segmentation in the automatic pipelines. We chose not to add these three missed galaxies back into our final samples, as doing so would complicate the selection function and make our results less straightforward to compare to future studies. Including them would not affect any of our conclusions.

The vast majority of our $z > 3$ quiescent candidates lie within or close to the quiescent regions in both the observed NIRCam color space and the rest-frame *UVJ* plane, supporting the robustness of our extended color criteria (green solid lines in Figure 2). A more conservative color selection (green dashed lines in Figure 2) would miss a substantial fraction of the population. We also find that relying on a single method is incomplete: using only the NIRCam red selection wedge would miss 5% of the gold sample and 26% of the silver sample, while using *UVJ* (from PROSPECTOR) alone would miss 3% and 10%,[21] respectively. Finally, the silver sample is systematically bluer and lies closer to the selection boundaries than the gold sample, consistent with the silver criteria admitting more recently quenched or less securely quiescent systems, whereas the gold sample selects the most robustly quiescent galaxies.

In the SFR–M_* plane (Figure 4), our quiescent samples lie well below the star-forming main sequence. All galaxies in the gold sample are more than 1 dex below the main sequence, while the silver sample, despite slightly higher inferred sSFRs, also lies systematically below it, with the vast majority offset by more than 0.5 dex (at least 131/137, >95%). Only a very small subset of the full silver sample falls closer to the main sequence (about 10/288, about 3.5%), indicating that contamination by star-forming systems is minimal. Taken together, these trends demonstrate that our photometric selection identifies a robust population of quiescent galaxies at $z > 3$, with the gold sample representing the cleanest, most strongly quenched subset.

[21] These numbers become much higher if the rest-frame *UVJ* colors are from EAzY, highlighting the large uncertainties in EAzY rest-frame *UVJ* estimates. A similar finding is reported in W. M. Baker et al. (2025b).

Figure 5 shows the gold-sample quiescent galaxy candidates at $z \sim 6$ that are still pending spectroscopic confirmation. The photometric coverage provides good sampling of the Balmer break, tracing the SED on both the blue and red sides of rest-frame 4000 Å. The blue-side SEDs resemble those of poststarburst or quiescent galaxies, while at longer wavelengths, these objects do not show the unusually red continuum behavior typically associated with LRDs. Together, these features strengthen the interpretation of these systems as high-redshift quiescent candidates, although spectroscopic confirmation is still required.

4. Number Density of Massive Quiescent Galaxies at Redshift $z \sim 3$–8

Here we present our constraints on the cosmic number density of quiescent galaxies with $M_* \geqslant 10^{10}\,M_\odot$ at $z \geqslant 3$, based on 1000 arcmin^2 of NIRCam imaging. Our measurements carefully account for both random and systematic uncertainties.

4.1. Method

To compute the cosmic number density of photometrically selected quiescent galaxies, many previous studies simply bin quiescent galaxies by redshift and divide by that bin's survey volume (e.g., A. C. Carnall et al. 2023; F. Valentino et al. 2023; S. Alberts et al. 2024b; W. M. Baker et al. 2025a). While these studies often account for redshift uncertainties by Monte Carlo resampling of the $z_{\rm phot}$ probability distribution function (PDF), uncertainties in sSFR are typically ignored. Although some earlier works have incorporated the sSFR PDF when defining robust quiescent samples (e.g., A. C. Carnall et al. 2020, 2023), the classification of a galaxy as quiescent has remained deterministic in those studies.

For this study, we combine data from multiple surveys that differ in imaging coverage and depth and hence the quality of the inferred physical properties. To handle these variations consistently and to propagate measurement uncertainties more robustly, we adopt a probabilistic framework. We make use of the full joint posterior distribution of $z_{\rm phot}$ and sSFR derived from our PROSPECTOR SED fitting for each quiescent galaxy to compute their cosmic number density.

For the ith galaxy in the sample, rather than assigning a fixed classification based on a single best-fit value, we evaluate the probability of it being quiescent at redshift z_k as

$$P_i^{\rm Q}(z_k) = P_i(z_k) \times \int_0^{+\infty} \mathcal{S}({\rm sSFR},\, z_k)\, P_i({\rm sSFR}|z_k)\, d({\rm sSFR}), \quad (4)$$

where $P_i(z_k)$ is the redshift posterior probability density, $P_i({\rm sSFR}|z_k)$ is the conditional sSFR posterior from our PROSPECTOR SED fitting, and $\mathcal{S}({\rm sSFR}, z)$ is the selection function for quiescent galaxies:

$$\mathcal{S}({\rm sSFR}, z) = \begin{cases} 1 & \text{if sSFR} < \frac{0.2}{t_H(z)}, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

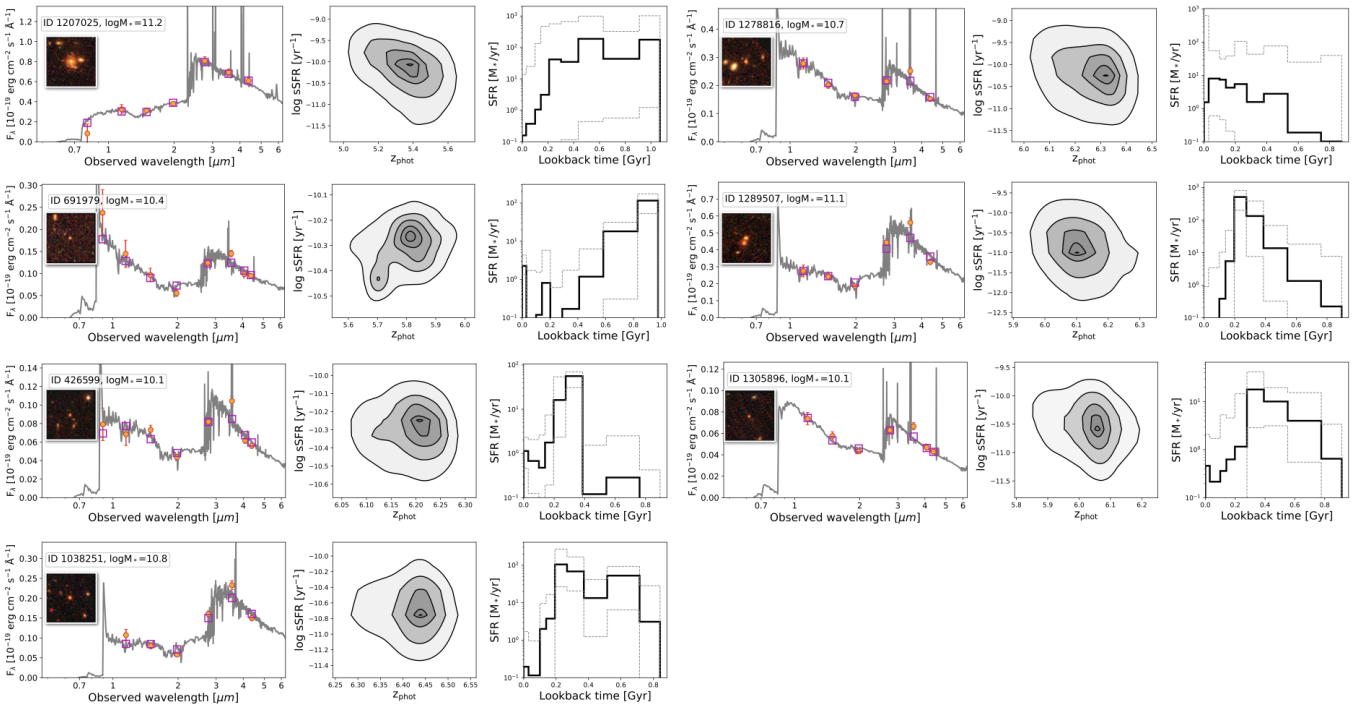

Figure 5. Candidates of massive quiescent galaxies at $z \sim 6$ in our gold sample that have not yet been spectroscopically confirmed. For each object, we show the JWST image cutout and observed SED with the best-fit model (left), the posterior constraints on photometric redshift and sSFR (middle), and the inferred SFH (right) from the continuity SFH SED fitting.

Summing over the full sample, the expected number of quiescent galaxies in the redshift bin $[z_k, z_k + \Delta z)$ is

$$N_{\rm Q}(z_k) = \sum_{i=1}^{N_{\rm gal}} \int_{z_k}^{z_k+\Delta z} P_i(z) \times \left[\int_0^{0.2/t_H(z)} P_i({\rm sSFR}|z)\, d({\rm sSFR}) \right] dz, \quad (6)$$

and the corresponding number density is

$$n_{\rm Q}(z_k) = \frac{N_{\rm Q}(z_k)}{\Delta V(z_k)}, \quad (7)$$

where $\Delta V(z_k)$ is the comoving volume of the redshift bin, computed from the survey geometry and cosmology.

This approach naturally incorporates the varying uncertainties of individual measurements, making it well suited for complex samples drawn from multiple surveys. Because the galaxy classification above is probabilistic and our selection function for quiescent galaxies depends on both redshift and sSFR, we can exploit the full joint posterior distribution of $z_{\rm phot}$ and sSFR to assess whether some quiescent candidates still have a non-negligible probability of being nonquiescent. This joint posterior depends on the adopted SED assumptions, so we discuss robustness among different SED assumptions in Section 4.2.

Moreover, this framework can be applied to larger samples to estimate the number of quiescent galaxies that may be missed by our selection (Section 3.3) and, in turn, to evaluate potential systematics in the number-density measurements. We applied this approach to the full set of 2633 galaxies from the initial color-based selection (Section 3.1). We find that the inferred number density of quiescent galaxies would increase by a factor of $\approx$1.5 relative to the gold+silver results reported below in Section 4.2.

Finally, for the uncertainty in the number-density measurements, we consider three sources of error,

$$\sigma^2_{\rm total} = \sigma^2_{\rm measurement} + \sigma^2_{\rm Poisson} + \sigma^2_{\rm CV}, \quad (8)$$

where $\sigma_{\rm measurement}$ is estimated as the 16th–84th percentile range from 200 bootstrap resamplings of our sample, $\sigma_{\rm Poisson}$ accounts for Poisson shot noise, and $\sigma_{\rm CV}$ accounts for cosmic variance, which can be calculated as

$$\sigma_{\rm CV} = \tilde{\sigma}_{\rm CV} \cdot n_Q, \quad (9)$$

where $\tilde{\sigma}_{\rm CV}$ is fractional.

To estimate $\tilde{\sigma}_{\rm CV}$, we use the code developed by C. K. Jespersen et al. (2025b), which is based on the B. P. Moster et al. (2011) cosmic-variance calculator, combined with constraints from the UNIVERSEMACHINE simulations (P. Behroozi et al. 2019). This code computes cosmic variance for 0.5 dex stellar-mass bins in each redshift bin. To obtain the cosmic variance for the stellar-mass bin at $M_* \geqslant 10^{10}\, M_\odot$, we follow F. Valentino et al. (2023) and weight the contribution of each 0.5 dex mass bin to the total cosmic variance by the quiescent galaxy stellar-mass function of W. M. Baker et al. (2025a).[22] We compute the cosmic variance for each field separately and then combine them following B. P. Moster et al. (2011),

$$\tilde{\sigma}^2_{\rm CV} = \frac{\sum_i A_i^2\, \tilde{\sigma}^2_{{\rm CV},i}}{\left(\sum_i A_i\right)^2}, \quad (10)$$

[22] The highest-redshift bin of the W. M. Baker et al. (2025a) stellar-mass function (their Table 2) is $z = 4$–5. For $z > 5$, we adopt the same function as for $z = 4$–5.

Table 1
Cosmic Number Density of Quiescent Galaxies with $M_* \geqslant 10^{10}\,M_\odot$ at $z \geqslant 3$

	Redshift	N No.	Number Density 10^{-5} Mpc^{-3}	$\tilde{\sigma}_{\rm CV}$ Fractional
Gold	(3.0, 3.5)	36	1.81 ± 0.53	0.06
	(3.5, 4.0)	37	1.52 ± 0.48	0.08
	(4.0, 5.0)	21	0.54 ± 0.39	0.12
	(5.0, 6.0)	2	0.09 ± 0.07	0.19
	(6.0, 8.0)	5	0.07 ± 0.06	0.59
Gold+silver	(3.0, 3.5)	97	3.14 ± 0.71	0.06
	(3.5, 4.0)	74	2.35 ± 0.61	0.08
	(4.0, 5.0)	44	1.10 ± 0.54	0.12
	(5.0, 6.0)	14	0.20 ± 0.11	0.19
	(6.0, 8.0)	9	0.10 ± 0.07	0.59

Note. The number density and uncertainty (including the cosmic-variance error) are derived from our probabilistic approach (Section 4.1), using the posteriors assuming the delayed-τ SFH. Using posteriors from alternative SED fits does not change the results (Section 4.2).

where A_i is the area of the ith field. Although the cosmic-variance error is large for individual fields, as expected for high-mass, high-z galaxies in small fields (C. L. Steinhardt et al. 2021), it becomes relatively small when combining our total sky coverage of 1000 arcmin^2.

4.2. Results

Using the probabilistic framework described above, we measured the cosmic number density of massive quiescent galaxies in both the gold and gold+silver samples across the redshift range $z = 3$–8. The results are presented in Table 1 and Figures 6 and 7.

For the gold sample, the number density drops by a factor of $\gtrsim$20 from $z = 3$ to $z > 6$, decreasing from approximately 2×10^{-5} to 7×10^{-7} Mpc^{-3}. At such low abundances (i.e., $\sim 10^{-6}$ Mpc^{-3}), a sky coverage of $\approx$400 arcmin^2 is required to detect a single galaxy over $z = 5$–6, underscoring the necessity of wide-area surveys to robustly constrain the number density of the earliest massive quiescent galaxy populations.

As shown in Figure 6, the number densities derived from the probabilistic approach, using posteriors from different SFH fittings (gold regions with different line-style edges), are highly consistent with each other for galaxies in the gold sample. They are also in excellent agreement with those obtained from a much simpler deterministic approach (gold star symbols), indicating that

$$\int_0^{0.2/t_H(z)} P_i(\mathrm{sSFR}|z)\, d(\mathrm{sSFR}) \approx 1, \tag{11}$$

i.e., nearly all of the PDF mass lies in the quiescent solution, regardless of the SED assumptions. This demonstrates the very high purity of the gold sample, with the galaxies in it having a very high confidence of being genuinely quiescent (also see Section 3.3 and Figure 4).

When the silver sample is included in our analysis, we observe an increase in the number density of quiescent galaxies (gray hatched regions in Figure 6). The magnitude of this increase depends on the SED assumptions. Relative to the gold sample alone, the number density of the combined gold and silver samples increases by a factor of $\approx$1.6 when using $z_{\rm phot}$–sSFR posteriors from the delayed-τ and continuity SFH fittings and by a factor of $\approx$2 when using posteriors from the bursty continuity fitting, thereby highlighting the systematic biases introduced by different SED assumptions. Importantly, the increase persists across all SFH priors. This is consistent with many silver objects carrying a high posterior probability of being quiescent under our SED assumptions. Together with their broadly quiescent-like colors (Section 3.3), this is indicative that the silver sample contains a real quiescent component, even if its purity cannot be established without independent spectroscopic confirmations.

Our measurements are in broad agreement with previous studies before the launch of JWST. Figure 6 shows that our new constraints are consistent with the results of J. R. Weaver et al. (2023; blue triangles) based on the pre-JWST COSMOS dataset. This is likely due to the fact that massive quiescent galaxies at $z > 3$ are bright (typically 23–24 mag in F444W), and thus they were detected and easily characterized by Spitzer/IRAC at 3–4 μm (J. R. Weaver et al. 2022).

Our constraints are consistent with earlier JWST studies based on much smaller NIRCam sky coverage (green symbols in Figure 6). We also measure the number density of gold-sample quiescent galaxies separately in each subfield. Figure 8 shows evidence for an enhanced abundance of quiescent systems in EGS at $z = 3$–3.5, consistent with S. Jin et al. (2024). In the $z = 3.5$–4 bin, we find an excess in GOODS-S, in agreement with W. M. Baker et al. (2025a; their Figure 6), who reported a similar overabundance using the photometrically selected quiescent sample from Z. Ji et al. (2024a). At $z = 4$–5, we find tentative evidence for an excess of quiescent galaxies in GOODS-N and UDS. Although direct number-density measurements are not yet available for these two fields, we note the recent discovery of two spectroscopically confirmed massive quiescent galaxies at $z \sim 4.6$ in PRIMER/UDS from the EXCELS survey (A. C. Carnall et al. 2024). The two galaxies lie within $<$1 Mpc of each other, hinting at the presence of large-scale structure in UDS at this redshift (A. C. Carnall et al. 2024; C. K. Jespersen et al. 2025a).

4.3. Comparing with Model Predictions

A central goal of this paper is to place our measured number densities of massive quiescent galaxies at $z > 3$ in the broader context of galaxy-formation models. The redshift evolution of the abundance provides a direct, population-level constraint on the mechanisms that assemble massive systems and then rapidly shut down their star formation.

Throughout this section, we implicitly assume that the stellar masses and SFRs inferred from the observations are directly comparable to the corresponding quantities reported by models. In practice, however, these quantities are not guaranteed to be defined or measured in an identical way: the models provide "intrinsic" physical properties, whereas observational estimates depend on SED-fitting assumptions and can suffer systematic offsets relative to the underlying values. Quantifying the impact of such systematics is nontrivial and would ideally require end-to-end synthetic observations built from the models and analyzed with the same pipelines used for the data. Such synthetic datasets are not yet available for the full set of models considered below, and producing them is beyond the scope of this work. With this caveat in mind, we present comparisons to different sets of models in which quenching emerges from a range of (often

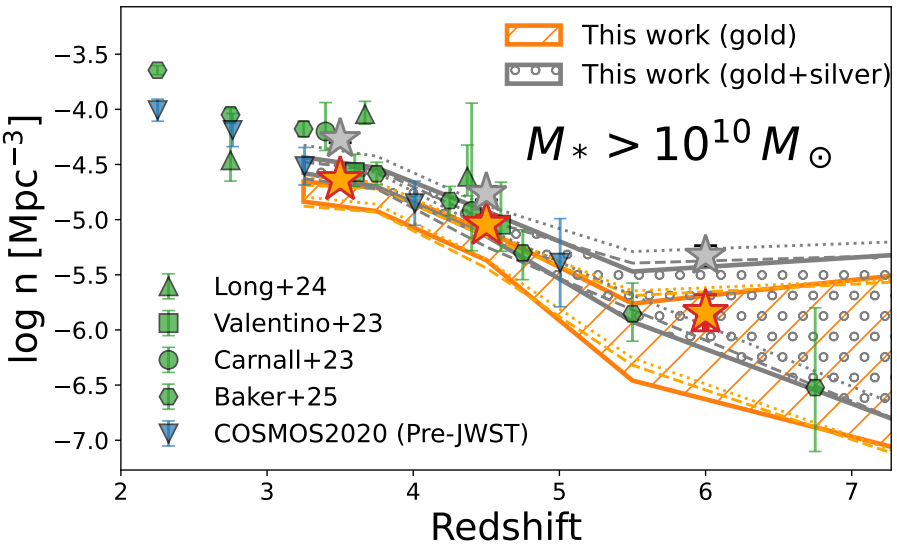

Figure 6. Cosmic number density of $z \geqslant 3$ quiescent galaxies with $M_* \geqslant 10^{10}\,M_\odot$ and sSFR $< 0.2/t_{\rm H}$ (Section 3). Hatched regions show the measurements (1σ range) obtained with the probabilistic approach (including the cosmic-variance error) described in Section 4.1, which fully accounts for the posterior distributions of redshift and sSFR. The regions with solid, dashed, and dotted edges correspond to the results based on the delayed-τ, continuity, and bursty continuity SFH fittings, respectively. The large star symbols indicate the measurements from a simple, deterministic approach, where the number density is calculated directly from galaxy samples binned according to their best-fit $z_{\rm phot}$, with only Poisson noise included. For comparison, we show measurements based on smaller JWST surveys, including F. Valentino et al. (2023; their Table 5), A. C. Carnall et al. (2023; their Table 3, robust subsample), A. S. Long et al. (2024; their Table 2, NIRCam color selection), and W. M. Baker et al. (2025a; their Table 1). We also show pre-JWST measurements from J. R. Weaver et al. (2023) based on COSMOS2020 (other pre-JWST studies are not shown for clarity, but they find generally consistent results; e.g., C. Schreiber et al. 2018; E. Merlin et al. 2019; A. C. Carnall et al. 2020).

unresolved subgrid) prescriptions for gas accretion, star formation, and feedback.

A growing number of studies have found substantial tension between theoretical predictions and the observed number densities of massive quiescent galaxies, particularly at $z \gtrsim 3$ (e.g., A. C. Carnall et al. 2023; F. Valentino et al. 2023; S. Alberts et al. 2024b; W. M. Baker et al. 2025b). We find a similar disagreement when comparing our measurements to a broad suite of state-of-the-art semianalytic models (SAMs) and cosmological hydrodynamical simulations (Figure 7; models listed therein).

Overall, the hydrodynamical simulations span a wide range of predictions, but several widely used frameworks underpredict the abundance of massive quiescent galaxies at $z \gtrsim 4$ by $\gtrsim$1 dex, with discrepancies that typically grow toward earlier epochs. The SAMs shown in Figure 7 generally predict higher abundances than the hydrodynamical simulations at $z \gtrsim 3$, and some (i.e., SHARK and GAEA; light blue solid and dashed lines in Figure 7) approach the lower envelope of our gold-sample constraints out to $z \sim 5$; nonetheless, the SAMs considered here still underpredict the abundance of our gold +silver sample at $z \gtrsim 4$. We return to this point in Section 6.1, where we discuss plausible explanations for why some SAMs appear to better reproduce the abundance of quiescent galaxies at $z > 3$ and the implications for the physical drivers of early quenching.

4.3.1. Empirical Models: UNIVERSEMACHINE

While the SAMs/simulations above are often used as tests of specific physical quenching pathways in galaxy-formation models, empirical models anchored to key observed correlations at lower redshifts (e.g., stellar-mass functions and the SFR–M_* relation) provide a deliberately "physics-light" but equally stringent baseline for the expected buildup of quiescent populations. In this case, consistency or inconsistency offers a direct diagnostic of whether relations calibrated at lower redshift can be reliably extrapolated into the $z > 3$ regime: agreement would suggest that the demographic link between mass growth and quenching inferred at later times already captures the essential ingredients of quenching at early epochs, whereas tension would imply that the high-z Universe requires additional evolution in quenching efficiency, timescales, or triggering channels beyond what is encoded by lower-z calibrations.

We compare our measurements to the UNIVERSEMACHINE empirical framework (P. Behroozi et al. 2019), which assigns SFHs to dark matter halos in a way that is statistically constrained by a wide range of observational demographic measurements across cosmic time. We generate mock realizations matched to the exact survey geometry of each pointing shown in Figure 1. To better sample the range of large-scale environments, we generate 16 light cones for each pointing. In each Monte Carlo realization, we randomly select one light cone per pointing, combine them to form a mock survey with the same total footprint as the data, apply the same quiescent selection (Section 3), and compute the number density of massive quiescent galaxies at $z > 3$. Repeating this procedure 10^4 times, we measure the median and mean predicted number densities and the 1σ scatter (taken as the 16th–84th percentile range) among realizations. The results are shown in the right panel of Figure 7.

UNIVERSEMACHINE broadly reproduces the rapid decline of the quiescent galaxy number density over $z = 3$–5. The median and mean predictions lie close to our observational constraints over $z \simeq 3.5$–5, while at $z \simeq 3.0$–3.5, the model tends to predict a higher abundance of massive quiescent systems than we infer. This is notable because UNIVERSEMACHINE includes a quenched-fraction constraint at $z \sim 3$–4 based on the UltraVISTA/COSMOS measurements of A. Muzzin et al. (2013), which are dominated by galaxies more massive than those in our sample, with $M_* \gtrsim 10^{11}\,M_\odot$. This difference is

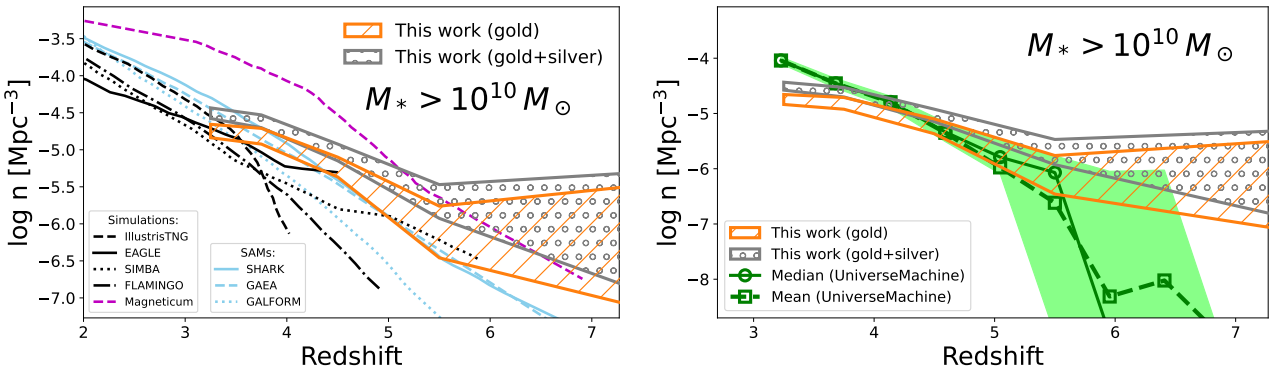

Figure 7. Comparison of the cosmic number density of quiescent galaxies with predictions from simulations and SAMs (left) and from the UNIVERSEMACHINE empirical framework (right). All models adopt the same stellar-mass cut ($\geqslant 10^{10}\,M_\odot$) and sSFR threshold ($<0.2/t_{\rm H}$) as the observations. Simulations include ILLUSTRISTNG (A. Pillepich et al. 2018), EAGLE (J. Schaye et al. 2015), SIMBA (R. Davé et al. 2019), FLAMINGO (J. Schaye et al. 2023), and MAGNETICUM (L. C. Kimmig et al. 2025; R.-S. Remus & L. C. Kimmig 2025); SAMs include SHARK (C. d. P. Lagos et al. 2024), GAEA (G. De Lucia et al. 2024), and GALFORM (C. G. Lacey et al. 2016). Simulation and SAM curves are taken from W. M. Baker et al. (2025a) and C. d. P. Lagos et al. (2025). UNIVERSEMACHINE predictions are described in Section 4.3.1; the light green band indicates the 1σ (16th–84th percentile) range, and the y-axis limits differ between panels to highlight the predictions at $z > 6$ from UNIVERSEMACHINE. To avoid overcrowding, we show only measurements from the probabilistic approach based on delayed-τ SED-fitting posteriors (shaded regions with solid edges in Figure 6).

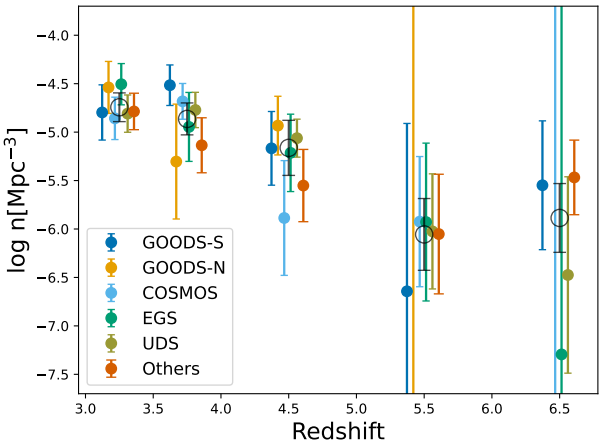

Figure 8. Similar to Figure 6 but showing the number density of quiescent galaxies in the gold sample measured in individual fields. The black open circles indicate the results obtained by combining all fields together. The redshift values for each field have been slightly displaced for clarity.

unlikely to be driven by cosmic variance: UltraVISTA covers $\simeq$1.62 deg^2, about 6 times the total area considered here, while our sample is distributed over 34 independent sight lines. The expected cosmic variance of our combined sample, $\sim$0.06–0.08 at $z \sim$ 3–4 (Table 1), is comparable to the $\sim$0.08–0.10 estimated for UltraVISTA (A. Muzzin et al. 2013). Moreover, our measurements are in good agreement with the more recent COSMOS2020 measurements of J. R. Weaver et al. (2023). Since both UltraVISTA and COSMOS2020 probe galaxies in the COSMOS field, the difference in their inferred quiescent abundances highlights the sensitivity of high-redshift quiescent galaxy measurements to photometric redshifts and SED-based stellar masses and SFRs. The tension at $z \simeq$ 3.0–3.5 may therefore largely reflect uncertainties in the earlier quenched-fraction constraints used to calibrate UNIVERSEMACHINE.

At higher redshifts, the UNIVERSEMACHINE predictions drop rapidly: the median and mean abundances decline sharply beyond $z \gtrsim 5$ and fall well below our inferred number densities, while the predicted 1σ scatter increases substantially at $z > 5$. We also observe that the mean prediction exceeds the median at $z \gtrsim 6$, indicating a skewed distribution of the UNIVERSEMACHINE predictions, in which most realizations contain very few (or no) such systems, with the mean boosted by a small tail of rare light cones that include one or more massive quiescent galaxies. Because UNIVERSEMACHINE is an empirically calibrated model rather than a fully physics-forward treatment of galaxy quenching, this comparison offers a particularly direct test of extrapolatability: the growing tension at $z \gtrsim 5$ implies that the onset and/or efficiency of quenching at the earliest times and/or the rapid mass growth required to populate the massive end by these epochs is not fully captured by lower-z-anchored empirical prescriptions. In other words, within the current UNIVERSEMACHINE extrapolation, massive quiescent galaxies at $z > 5$ are expected to be exceedingly rare, and matching our measured abundances would require substantially earlier and/or more efficient quenching (and/or its maintenance) than is implied by relations calibrated at lower redshift.

5. Cosmic Variance of Massive Quiescent Galaxies at Redshift $z \sim$ 3–8

Beyond the mean number-density evolution, the field-to-field variance in galaxy number density carries additional information about their connection to dark matter halo properties and large-scale structure. In pencil-beam surveys, these fluctuations are often dominated by cosmic variance, i.e., the clustering-driven excess (relative to Poisson) variance obtained by integrating the power spectrum weighted by the survey window function (J. A. Newman & M. Davis 2002; R. S. Somerville et al. 2004; M. Trenti & M. Stiavelli 2008; B. P. Moster et al. 2011; C. L. Steinhardt et al. 2021). With the multipointing geometry and sample size available here, we can, for the first time in the $z > 3$ regime, empirically quantify the spatial clustering of massive quiescent galaxies across independent sight lines. Translating the measured cosmic variance into a unique characteristic halo mass is nontrivial, since the derived clustering amplitude relies on nonlinear modes. Nevertheless, the observed field-to-field scatter provides a direct, model-independent constraint on the clustering strength (i.e., effective bias) of these systems (e.g., B. E. Robertson 2010).

5.1. Method

5.1.1. Cosmic-variance Estimations

We measure the fractional cosmic variance ($\sigma_{\rm CV}$) in the galaxy number counts per NIRCam pointing (N) using two complementary estimators. First, we perform a bootstrap variance-decomposition measurement following B. P. Moster et al. (2011). This method does not require assuming a functional form for the galaxy count distribution, but it can be biased high if a small number of outliers inflate the variance. The observed field-to-field variance in galaxy counts receives contributions from Poisson counting noise and cosmic variance,

$$\sigma_{\rm CV}^2 = \frac{\langle N^2\rangle - \langle N\rangle^2 - \langle N\rangle}{\langle N\rangle^2}, \quad (12)$$

where $\langle N\rangle$ is the mean number of galaxies per pointing and the Poisson contribution to the variance is $\sigma_{\rm P}^2 = \langle N\rangle$. Uncertainties on $\sigma_{\rm CV}$ are estimated from 10,000 bootstrap resamples of the independent pointings (with replacement), recomputing Equation (12) for each realization, and summarizing the resulting $\sigma_{\rm CV}$ distribution using the median and 16th/84th percentiles. In each bootstrap realization, the estimator above is applied to the independent PANORAMIC pointings shown in Figure 1, except for where two pointings are separated by less than 1°; in these cases, we randomly select one pointing from each region to preserve spatial independence. For the legacy fields (GOODS-S, GOODS-N, A2744, EGS, UDS, and COSMOS), whose footprints extend beyond a single NIRCam pointing, we instead randomly select a single NIRCam pointing within each field (random center and position angle) while requiring that the full footprint of the chosen pointing lies within the legacy-field coverage (the same procedure was used in A. Weibel et al. 2026a).

A more principled approach for quantifying the cosmic variance is to directly fit the galaxy number count distribution. We thus also estimate $\sigma_{\rm CV}$ using a Bayesian distribution-fitting approach that models the full field-to-field count distribution, thereby using all information in the observed distribution. Following C. K. Jespersen et al. (2025b), we model the underlying positive-definite distribution of counts with a gamma functional form, which allows the mean and variance to vary independently and recovers the Poisson-like limit as $\mathrm{Var}(N) \to \langle N\rangle$. As noted by C. K. Jespersen et al. (2025b), the choice of functional form (gamma versus negative binomial) is empirically negligible. To reduce sensitivity to rare, highly overdense pointings in the skewed count distribution, we define the likelihood using a negative binomial form evaluated on the binned number counts, which effectively downweights extreme overdensities compared to a pure Poisson likelihood (D. W. Hogg et al. 2010). We parameterize the model in terms of the mean $\mu \equiv \langle N\rangle$ and $\sigma_{\rm CV}$, using the variance relation implied by Equation (12):

$$\mathrm{Var}(N) = \mu + \mu^2\,\sigma_{\rm CV}^2. \quad (13)$$

We fit the data and sample the posterior using the Markov Chain Monte Carlo (MCMC) sampler `emcee` (D. Foreman-Mackey et al. 2013). Following A. Weibel et al. (2026a), we adopt a Jeffreys prior on the total variance, $p(\sigma) \propto \sigma^{-1}$, which favors small variance unless a larger variance is strongly supported by the data (H. Jeffreys 1946). Posterior constraints on μ and $\sigma_{\rm CV}$ are reported using the median and 16th/84th percentiles of the marginalized posteriors.

5.1.2. Constructing Abundance-matched Samples from UNIVERSEMACHINE

To compare our cosmic-variance measurements with theoretical expectations, we construct mock quiescent galaxy samples from the UNIVERSEMACHINE catalogs. As discussed in Section 4.3.1, applying the same quiescent galaxy selection used for our observations yields a UNIVERSEMACHINE quiescent number density at $z > 3$ that is inconsistent with the data. Because cosmic variance is a second-order statistic that depends on the tracer population, a comparison of field-to-field fluctuations is not readily interpretable when the underlying mean abundance (the first-order statistic) already differs between the data and the mocks.

We therefore construct abundance-matched UNIVERSEMACHINE samples by enforcing agreement in the mean number density. In each redshift bin, we rank UNIVERSEMACHINE galaxies with $M_* \geqslant 10^{10}\,M_\odot$ by their sSFR and select the N galaxies with the lowest sSFR, where N is chosen such that the resulting UNIVERSEMACHINE number density matches the observed quiescent galaxy abundance in the same redshift interval (Table 1). This procedure is equivalent to applying a redshift-dependent sSFR threshold, yielding samples with matched first-order statistics while largely preserving UNIVERSEMACHINE's predicted clustering for an abundance-matched population. The implied $\log(\mathrm{sSFR}/\mathrm{yr}^{-1})$ thresholds are -11.5, -9.5, -9.4, -9.3, and -8.9 from low to high redshift. Relative to the $0.2/t_{\rm H}$ criterion, these thresholds are lower in the lower-redshift bins (e.g., $\sim 10^{-10}\,\mathrm{yr}^{-1}$ at $z = 3.5$) and larger in the highest-redshift bin (e.g., $\sim 10^{-9.5}\,\mathrm{yr}^{-1}$ at $z = 7$). Similar to Section 4.3.1, in total, we generate 10^4 realizations of the mock observations and use them to derive the median and 16th/84th percentiles of the resulting N distribution.

Our choice of an sSFR-ranked abundance-matching scheme is motivated by two considerations. First, it is robust to potential systematic offsets in the absolute sSFR scale of UNIVERSEMACHINE at high redshift. UNIVERSEMACHINE assigns SFRs empirically by calibrating a flexible model to reproduce a broad set of ensemble observables over the redshift range where constraints exist (P. Behroozi et al. 2019). However, the quenched component of the SFR distribution is itself poorly constrained observationally, so UNIVERSEMACHINE fixes the lognormal parameters of the quiescent population to values based on $z \sim 0$ data. As a result, the absolute sSFR scale of quiescent galaxies in the model, particularly at high redshift, should be regarded as uncertain. A rank-based selection is therefore less affected by monotonic offsets and avoids selecting fundamentally different populations simply because the SFR scale differs between model and data. Second, and more importantly for our purposes, since cosmic variance depends on the tracer bias, matching the comoving number density provides a practical way to define samples that are more comparable in their typical host-halo population than a hard sSFR cut when the baseline abundances disagree. Finally, observational quiescent selections (e.g., sSFR-based) effectively isolate the low-sSFR tail at fixed stellar mass; our procedure provides a model-internal analog of

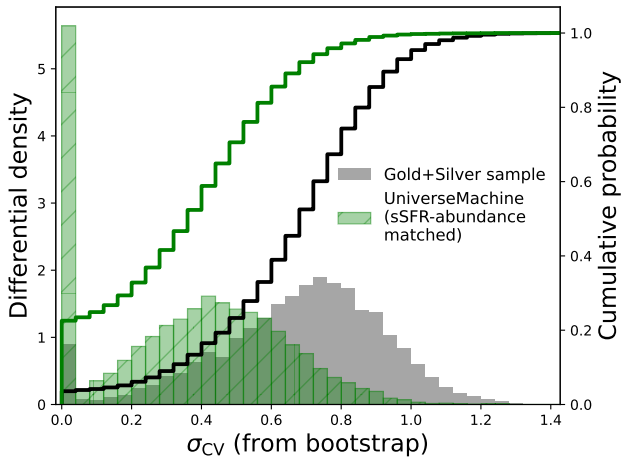

Figure 9. Comparison of bootstrap-inferred cosmic variance (Equation (12), on the scale of a single NIRCam pointing), σ_{CV}, between the observed gold +silver quiescent sample and the sSFR-abundance-matched UNIVERSEMACHINE mocks using the same number of sight lines. The filled histograms show the differential distributions of σ_{CV}, while the solid curves show the corresponding cumulative probabilities (right axis). The observations exhibit systematically larger field-to-field fluctuations than the mocks in two complementary senses: (i) the observed distribution is shifted to higher values, with a larger median σ_{CV} than UNIVERSEMACHINE, and (ii) the UNIVERSEMACHINE distribution includes substantial probability mass at $\sigma_{CV} \simeq 0$, with ∼25% of bootstrap realizations returning $\sigma_{CV} = 0$. See Section 5.2 for discussion of the interpretation of the low-σ_{CV} tail in the UNIVERSEMACHINE mocks.

this selection while preserving UNIVERSEMACHINE's predicted clustering for an abundance-matched population.

5.2. Results

We first report the cosmic-variance constraints for our quiescent galaxy samples. The gold sample alone lacks sufficient statistics for a robust cosmic-variance measurement (see Appendix E), so in what follows, we focus on the combined gold+silver sample. For the gold+silver sample, we measure $\sigma_{CV} = 0.71^{+0.20}_{-0.28}$ using the bootstrap variance-decomposition estimator (Figure 9) and $\sigma_{CV} = 0.65^{+0.25}_{-0.30}$ from the Bayesian MCMC distribution-fitting method (the top row of Figure 10). We note that the agreement between these two approaches does not generally imply that the two estimators are equally robust: the bootstrap variance-decomposition estimator is more sensitive to outlier fields, and our sample likely does not contain an unusually extreme field.

For the sSFR-abundance-matched UNIVERSEMACHINE mock sample, we infer systematically smaller cosmic variance. From bootstrapping (Figure 9), we find $\sigma_{CV} = 0.38^{+0.24}_{-0.38}$, where the 1σ lower bound is consistent with zero. Similarly, as shown in the bottom row of Figure 10, the MCMC fit yields $\sigma_{CV} = 0.31^{+0.28}_{-0.22}$. The posterior for σ_{CV} is strongly asymmetric, with its maximum toward very small values (bottom right panel). It is important to emphasize that this concentration of posterior mass near $\sigma_{CV} = 0$ does not imply that the UNIVERSEMACHINE galaxies are not clustered. Rather, it indicates that, for the same finite number of sight lines (34) as in our observational sample, a nonnegligible fraction of UNIVERSEMACHINE realizations do not produce a statistically measurable excess variance above Poisson noise (Appendix F provides a detailed discussion on the detectability of σ_{CV}). In the limit of an arbitrarily large number of sight lines, the intrinsic cosmic variance of the sSFR-matched UNIVERSEMACHINE sample is $\sigma_{CV} \sim 0.43$, which is lower than our current observational constraint at mild significance.

Before interpreting this difference, we first ask whether the observed excess σ_{CV} could simply arise from random observational uncertainty. Random redshift or quiescent-classification errors could broaden the observed field-to-field count distribution, thereby increasing the measured σ_{CV} above the intrinsic value. To assess the scale of this effect, we make a simple order-of-magnitude estimate. For the combined massive gold+silver sample used in our main cosmic-variance analysis, we have 238 galaxies across 34 independent sight lines, corresponding to a mean of $\mu \approx 238/34 = 7$ objects per field, while the measured excess variance is $\sigma_{CV} \approx 0.7 \pm 0.3$. If each potentially ambiguous source has an independent probability p of being scattered into or out of the quiescent class, then the additional fractional field-to-field variance from this random contamination is approximately $\sigma^2_{rand} \approx p(1-p)/\mu$. Previous work suggests that contamination rates of order ∼10% are plausible for JWST-era photometric quiescent selections, while values as high as ∼30% have been discussed for less clean selections (J. Antwi-Danso et al. 2023; S. D. Stevenson et al. 2026). Assuming this random component is independent of the intrinsic cosmic variance, subtracting it in quadrature implies $\sigma_{CV,int} \approx (\sigma^2_{CV} - \sigma^2_{rand})^{1/2}$, or ≈0.69 for $p = 0.1$ and ≈0.67 even for the conservative $p = 0.3$ case. Thus, purely random, field-independent classification noise is highly unlikely to dominate the observed excess scatter.

More generally, contamination need not act only as a random noise term. Because massive red/quiescent galaxies are generally among the most strongly clustered, highly biased galaxy populations (e.g., I. Zehavi et al. 2011), the inclusion of true interlopers (for example, due to redshift misassignments) would be expected to dilute rather than enhance the clustering signal. In that case, the inferred cosmic variance would be lowered by roughly a factor of $(1 + f_{interlopers})^2$ (see, e.g., C. C. Williams et al. 2011). Contamination of this kind therefore cannot explain an observed σ_{CV} that is larger than the UNIVERSEMACHINE prediction and may instead imply that our measurement is conservative. In addition, our field-to-field variance is estimated from a finite number (though the largest to date) of independent pointings and is therefore still sensitive to small number statistics and the presence or absence of rare overdensities. On the mock side, our sSFR-based abundance matching is intentionally rank-based and designed to match only the mean number density; it does not guarantee that the abundance-matched UNIVERSEMACHINE selection reproduces the same halo-occupation mix, satellite fraction, or assembly-dependent correlations as the real quiescent population at fixed abundance.

With these caveats in mind, we discuss the implications of our comparison. An observed cosmic variance that exceeds model predictions may indicate that the observed galaxies trace more strongly biased environments than assumed in the model. For example, this population may preferentially occupy rarer peaks or more massive halos or exhibit a stronger environmental dependence that enhances small-scale clustering.

To help narrow down the physical origin of the discrepancy, we perform an additional test: instead of abundance matching in sSFR, we match the observed mean

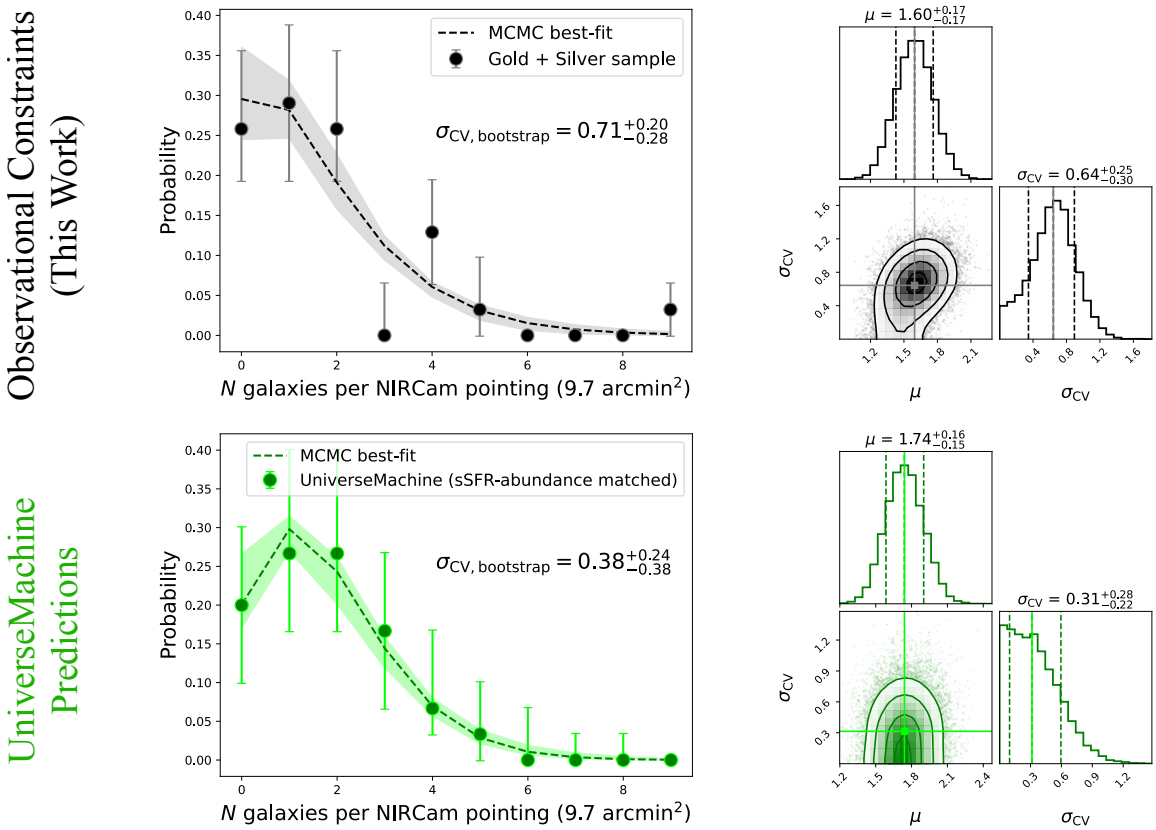

Figure 10. MCMC-inferred cosmic variance. Top: observational constraints for our gold+silver sample of massive quiescent galaxies at $z > 3$. The left panel shows the probability distribution of the number of galaxies per pointing, N (black points with uncertainties from bootstrap resampling). The dashed curve indicates the MCMC best-fit model (see Section 5.1.1), with the shaded band showing the posterior predictive 1σ range. The right panel shows the corner plot and marginalized posteriors. Bottom: same as the top row but for the sSFR-abundance-matched UNIVERSEMACHINE mock catalog (see Section 5.1.2), yielding a lower σ_{CV} relative to our observational constraints, consistent with our conclusions based on bootstrap-inferred cosmic variance (Figure 9).

number density using halo mass alone, independent of sSFR, by selecting the N most massive halos in each redshift bin. This selection should, if anything, maximize the expected field-to-field fluctuations in the mocks by isolating the most strongly biased tracers (i.e., the most massive halos), thus serving as the maximum variance predicted by the UNIVERSEMACHINE. This halo-mass-based abundance-matched sample has slightly higher halo masses,[23] $10^{12.5\pm0.2}\,M_\odot$, compared to $10^{12.1\pm0.4}\,M_\odot$ for the sSFR-abundance-matched sample. We find that the intrinsic cosmic variance of the halo-mass-matched UNIVERSEMACHINE sample is $\sigma_{CV} \sim 0.51$, slightly higher than for the sSFR-abundance-matched sample but still below our current observational constraint. This result suggests that the discrepancy is unlikely to be driven solely by the details of the sSFR-based selection and may instead point to additional environment-dependent effects in the real Universe. For example, the quenching probability may be enhanced via processes analogous to galaxy conformity (S. M. Weinmann et al. 2006). Because such effects are not calibrated in the UNIVERSEMACHINE empirical model above $z > 1$, if quenching depends partly, or even primarily, on large-scale environment rather than halo mass alone, this could result in a higher effective cosmic variance than predicted by UNIVERSEMACHINE.

[23] Here peak halo mass is used, which is defined as the maximum mass the (sub)halo ever reached over its history.

6. Discussion

6.1. Implications of the Joint Measurement of High Abundance and High Bias

Our key result is that massive quiescent galaxies at $z > 3$ are both unexpectedly abundant and strongly clustered, placing a stringent constraint on models of early quenching. These two observables probe different, but not necessarily independent, aspects of the quenching problem. In a downsizing framework, if the earliest-quenched and/or most massive systems preferentially arise in overdense regions, then the same processes that elevate the quenched number density may also enhance the effective bias of the population. In this sense, the abundance evolution constrains how many galaxies quench, while the clustering constrains where they reside. The joint requirement is therefore demanding: matching the abundance alone is insufficient if quenched systems populate the wrong halos or environments, and matching the clustering alone is not meaningful if the quenched population is too rare.

Our high inferred number densities imply that quenching must operate efficiently within the first ∼1–2 Gyr of cosmic time. More broadly, the abundance result indicates that the dominant mechanism must be able to rapidly suppress star formation and maintain low sSFRs despite the intense gas supply expected at these epochs (L. J. Tacconi et al. 2020). This poses a significant challenge for current models, since it requires not only early and efficient stellar-mass assembly but also equally rapid truncation of star formation. This early

quenching is also directly relevant to the massive quiescent population observed at later times. Stellar population studies of quiescent galaxies at $z \lesssim 2$ have long inferred that at least some of these systems formed the bulk of their stars at $z \gtrsim 3$, but their subsequent evolution makes it increasingly difficult to reconstruct the conditions under which they originally quenched. Our number-density measurements instead constrain the abundance of massive galaxies that were already quiescent at these early epochs and can therefore be compared with the early formation histories inferred for the oldest quiescent galaxies at lower redshift. In this way, the $z > 3$ population provides a direct link between the fossil record seen at later times and the epoch when the first massive galaxies actually became quiescent.

Complementing the abundance measurement, the clustering signal constrains how quenching is distributed across the large-scale density field. An elevated clustering would suggest that massive quiescent galaxies preferentially occupy more biased regions, linking the emergence of this population to the broader structure in which their host halos grow.

We measured the cosmic variance of massive quiescent galaxies at $z > 3$, finding a possible excess relative to abundance-matched mock galaxies from UNIVERSEMACHINE. At present, however, this tension is only at low significance ($\sim 1\sigma$) and, as noted in Section 6.3, will require more than 100 additional independent sight lines to establish robustly. If confirmed, this signal would point to a dependence on large-scale environment in early quenching. One useful framework for interpreting such a trend is galactic conformity, in which quenched systems are associated with an enhanced likelihood of quenched neighbors (e.g., S. M. Weinmann et al. 2006; G. Kauffmann et al. 2013; W. G. Hartley et al. 2015; L. Kawinwanichakij et al. 2016; A. M. Berti et al. 2017). Although the physical origin of conformity at low redshift may not directly carry over to the earlier Universe, a conformity-like signal at $z > 3$ would suggest that quenching is shaped by more than halo mass alone. Similarly, if galaxies systematically affect each other in a way that changes the local stellar-to-halo mass relation (in either direction), the clustering signal could be enhanced (J. F. Wu et al. 2024). The high bias could also be caused by unmodeled halo assembly effects (assembly bias; see, e.g., C. K. Jespersen et al. 2022; C.-Y. Chuang et al. 2024).

Taken together, the combination of high abundance and elevated clustering is highly restrictive. It favors scenarios in which the dominant high-redshift quenching pathways are both (i) rapid and effective enough to overcome strong gas replenishment and (ii) coupled to large-scale environment strongly enough to enhance the probability of quenching in the vicinity of already-quenched galaxies. This would be difficult to reconcile with quenching driven solely by internal heating within individual halos, since such a process would not naturally influence neighboring halos. Instead, it points to mechanisms that are both efficient and farther reaching.

The model comparison provides a suggestive clue to the nature of such mechanisms. Among the models considered here, the SAMs generally predict higher quenched abundances than the hydrodynamical simulations at $z \gtrsim 3$, and SHARK and GAEA lie closest to our measurements (Figure 7). A notable shared feature of these two models is the inclusion of an additional radiatively efficient active galactic nucleus (AGN) wind/outflow channel beyond low-accretion jet/radio heating. This suggests that a radiatively efficient, quasar-wind-like AGN feedback mode may be an important ingredient for producing the observed abundance of $z > 3$ quiescent galaxies in the stellar-mass range ($M_* \sim 10^{10}$–$10^{11}\,M_\odot$) probed by our sample. Its effectiveness likely depends on the central gas reservoir: when gas is strongly centrally concentrated, ejective feedback can couple more efficiently to the interstellar medium and more readily expel and/or strongly heat it (S. Belli et al. 2024; R. L. Davies et al. 2024). Because massive galaxies at earlier times are more compact (A. van der Wel et al. 2014; Z. Ji & M. Giavalisco 2022), their gas reservoirs may also be more centrally concentrated, making this channel particularly effective at high redshift.

The evidence for excess clustering in our $z > 3$ quiescent galaxy sample is also qualitatively consistent with a quasar-wind-driven picture, in which rapid quenching is linked to biased regions of large-scale structure. Ejective feedback can act on short timescales by expelling and/or strongly disrupting the cold-gas reservoir, while quasar triggering is expected to be enhanced in the densest peaks where early massive galaxies assemble. Intriguingly, Y. Zhu et al. (2025) recently reported observational evidence consistent with quasar radiative feedback influencing galaxy growth on intergalactic scales at $z = 6.3$, finding suppressed nebular emission in galaxies surrounding the extremely luminous quasar J0100+2802 out to ~1–2 Mpc, comparable to the size of a single NIRCam pointing. While the direct connection between suppressed nebular emission and the long-term quenching of massive systems remains to be established, this result provides a physically motivated example of how quasar episodes at the relevant epochs could plausibly contribute to environmentally correlated suppression in their surroundings. If this interpretation is correct, the quiescent galaxies identified here may also be spatially correlated with UV-luminous galaxies, as expected for the biased environments of rapidly growing high-redshift quasars (e.g., T. Costa et al. 2014, 2019). Measuring this cross-correlation requires a well-characterized selection function for the UV galaxy population across our heterogeneous fields, and we leave such an analysis to future work.

We emphasize, however, that we are not arguing that quasar-wind-driven feedback is uniquely required by our observations, nor that it is necessarily the dominant quenching mechanism in all massive quiescent galaxies at $z > 3$. Rather, we identify it as one plausible scenario that is qualitatively consistent with both the high abundance and the tentative excess clustering measured in our sample.

Other known quenching channels may also contribute, although they do not all naturally explain the possible environmental dependence suggested by our measurements. For example, supernova feedback can regulate star formation but is generally expected to become less effective in massive systems (A. Dekel & Y. Birnboim 2006). Major mergers can trigger strong starbursts, but current simulations do not generally find a direct connection between merger events and quenching (e.g., F. Rodríguez Montero et al. 2019; S. Kurinchi-Vendhan et al. 2024). Halo-scale quenching provides a more direct connection to environment, since the suppression of gas accretion is tied to halo mass and can produce an environmental dependence of quiescence (e.g., A. Cattaneo et al. 2006); however, cold streams are expected to continue penetrating massive hot halos efficiently at high redshift (A. Dekel & Y. Birnboim 2006; A. Dekel et al. 2009).

Importantly, these quenching channels are already included, with different implementations, in many of the models considered here (Section 4.3), yet most still underpredict the abundance of massive quiescent galaxies at $z > 3$ (Figure 7). Our results may therefore require either modifications to how these known mechanisms operate at early times, an additional quenching channel such as quasar-driven feedback, or both. Regardless, any successful model should reproduce not only the high abundance of early quiescent galaxies but also their possible preference for overdense environments.

Finally, we note that all references above to "environment" or "environmental dependence" are intended in the descriptive sense defined in Section 1, namely, the large-scale spatial context traced by cosmic variance, rather than as evidence for any specific physical quenching mechanism.

6.2. The Number Density of z ∼ 7 Quiescent Galaxies

JWST NIRSpec spectroscopy now enables absorption-line-anchored SFH inference for quiescent galaxies at $z \sim 4$, yielding strong constraints on formation epochs and quenching timescales. Several confirmed systems show short, intense formation followed by rapid early quenching (e.g., A. de Graaff et al. 2025a; T. Nanayakkara et al. 2025), including GS-9209 at $z = 4.658$ quenched by $z_{\rm quench} \approx 6.5$ (A. C. Carnall et al. 2023) and PRIMER-EXCELS-117560 and 109760 at $z \simeq 4.6$ with $z_{\rm quench} \approx 6.5$–7 (A. C. Carnall et al. 2024). If these inferences of $z_{\rm quench}$ are correct, some progenitors of the $z \sim 3$–5 quiescent population should already be quenched massive systems by $z \sim 7$. JWST has confirmed at least one such object at $z \simeq 7.3$ (A. Weibel et al. 2025a). Wide-area NIRCam surveys provide a direct test of whether this early quenched tail is already present at $z \sim 7$.

In our $\sim$1000 arcmin^2 NIRCam dataset, the baseline selection (Section 3.3) yields only one candidate with $M_* > 10^{10}\,M_\odot$ and $z_{\rm phot} > 6.5$ (Figure 3), suggesting a lower abundance than naive expectations from SFH archeology. However, our baseline sample intentionally excludes very red sources with F277W − F444W > 1 to mitigate contamination from LRDs. As mentioned in Section 3.2.1, these removed red objects cluster at $z_{\rm phot} \sim 7$ under our PROSPECTOR-based inference (assuming stellar dominated emission), raising the concern that the cut could also remove genuine $z \gtrsim 7$ quiescent galaxies. Indeed, R. E. Hviding et al. (2025) find that 60% of spectroscopically confirmed LRDs satisfy F277W − F444W > 1, so neither keeping nor removing them is safe with photometry alone: inclusion risks nonstellar contamination, while exclusion may suppress the population of interest.

We therefore reanalyze the removed red (F277W − F444W > 1) objects that are plausibly quiescent by our SFH criteria, requiring (i) $z_{\rm phot} > 6.5$ and $M_* > 10^{10}\,M_\odot$ in all three PROSPECTOR fits and (ii) a minimum $\log({\rm sSFR}/{\rm yr}^{-1}) < -9.5$ across the fits. This yields 12 galaxies. We refit these galaxies with the `agn_blue_sfhz_13` templates,[24] which include an LRD template from M. Killi et al. (2024). Our goal of this analysis is not classification but to test whether an LRD-like template can explain the photometry comparably well.

Figure 11 presents examples of this comparison. One source (the top row of Figure 11) shows negligible LRD contribution and is well fit by both PROSPECTOR and EAzY without requiring an LRD component; this is the spectroscopically confirmed $z \simeq 7.3$ quiescent galaxy from A. Weibel et al. (2025a), which was excluded from our initial selection because it has (F277W − F444W) = 1.1 > 1. For seven sources, the PROSPECTOR fit yields a comparable or smaller $\chi^2_{\rm tot}$ than the EAzY stellar+LRD solution, consistent with largely stellar SEDs. For the remaining 13 sources, the EAzY stellar+LRD solution improves $\chi^2_{\rm tot}$ substantially, suggesting plausible LRD-like contributions. We emphasize, however, that these $\chi^2_{\rm tot}$ differences are not decisive. With only 6–7 broad bands, distinct physical models can reproduce the same colors, and the EAzY test is incomplete because `agn_blue_sfhz_13` includes only a single LRD-like template and cannot capture the diversity of LRD SEDs and line properties. Nonetheless, we use this comparison as a pragmatic filter to define an LRD-disfavored subset.

Retaining only sources for which the EAzY stellar+LRD fit is not substantially better than PROSPECTOR, we obtain $N = 8$ plausible $z \sim 7$ massive quiescent candidates. Over $A = 0.28\,{\rm deg}^2$ and $z = 6.5$–8.5, this implies $n \simeq (1.9 \pm 1.4) \times 10^{-6}\,{\rm Mpc}^{-3}$ (Poisson error only). Systematics from SED modeling and residual LRD or interloper contamination likely dominate, so we treat this as indicative and do not include it in our primary number-density results (Section 4) pending spectroscopic confirmation.

Finally, we compare this provisional $z \sim 7$ abundance with expectations from SFH archeology at $z \sim 3$–5. The existence of GS-9209 and the two PRIMER-EXCELS systems, with inferred $z_{\rm quench} \sim 6.5$–7, implies that at least three massive quenched galaxies were already in place by $z \sim 7$. Although these objects come from targeted spectroscopy and therefore do not have a well-defined survey selection function, adopting a parent area of $A \simeq 400$ arcmin^2 for GOODS and UDS yields a crude lower limit of $n \gtrsim 2 \times 10^{-6}\,{\rm Mpc}^{-3}$, comparable to our LRD-disfavored estimate. Current wide-area photometric samples are therefore beginning to test whether the early quenched tail inferred from $z \sim 3$–5 spectroscopy is consistent with the observed $z \sim 7$ population. Robust progress, however, will require well-defined selection functions and larger, uniformly selected spectroscopic samples; we defer such a comparison to Y. Zhang et al. (2026).

6.3. Future Strategies for Constraining the Demographics of z > 3 Massive Quiescent Galaxies

Our study demonstrates the power of combining large survey areas over many independent sight lines to constrain the evolution of massive quiescent galaxies at $z > 3$ using NIRCam. With large areas and independent sight lines, we gain not only improved constraints on abundance evolution but also direct empirical information on clustering through cosmic variance. Together, these observables provide stronger insight into how early massive quiescent galaxies formed. In particular, our results highlight that environment may already play a role in the buildup of quiescent galaxies early in the Universe, if the tentative excess clustering inferred here is confirmed with future data.

For the cosmic-variance measurement, the key quantity is the number of independent sight lines, since the uncertainty on $\sigma_{\rm CV}$ decreases approximately as $1/\sqrt{n_{\rm field}}$ (see Appendix F). This makes JWST pure-parallel surveys especially powerful: they naturally produce many single-pointing fields distributed

[24] https://github.com/gbrammer/eazy-photoz/tree/master/templates/sfhz

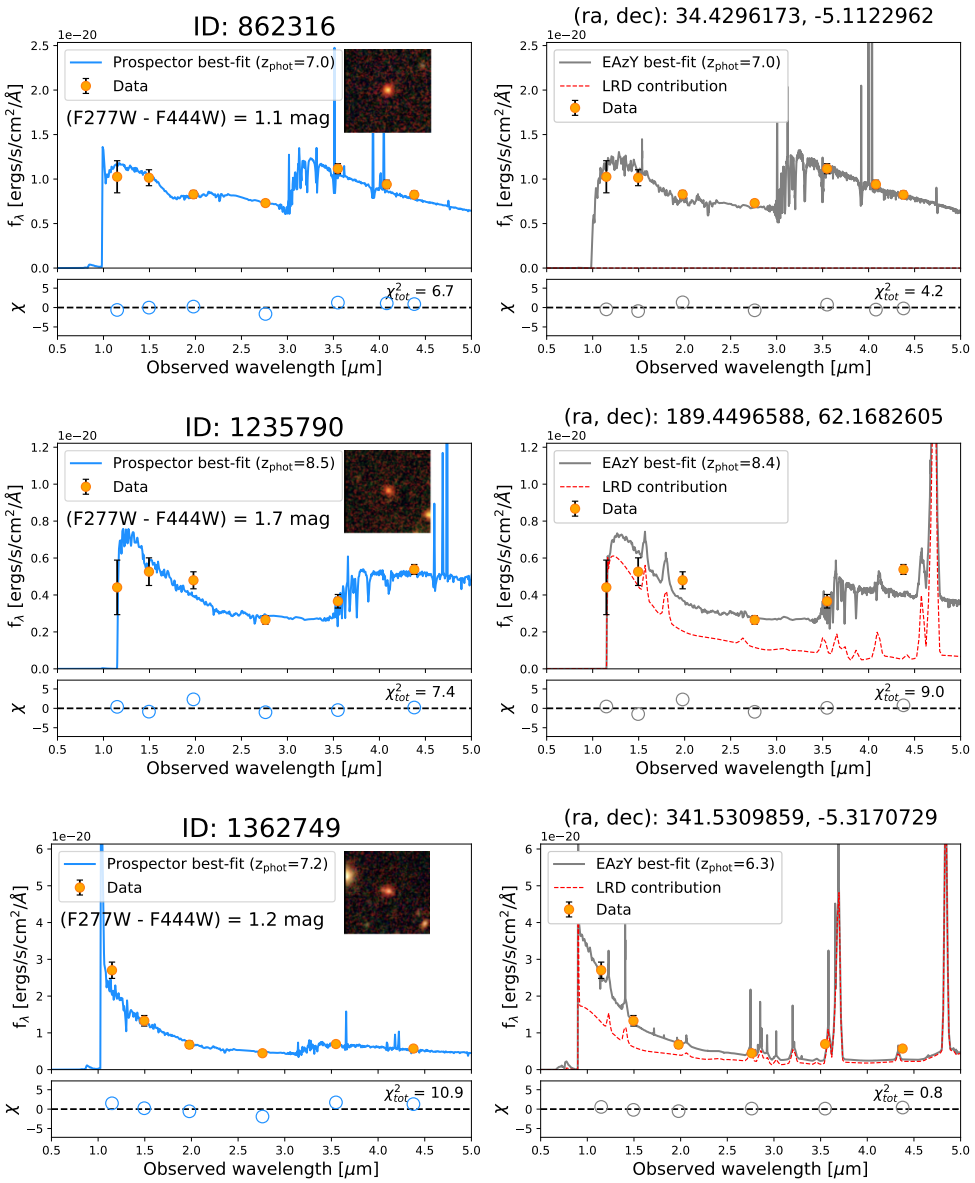

Figure 11. Examples from our reanalysis of the full sample of $z \sim 7$ quiescent candidates with (F277W − F444W) > 1 mag. For each source, the left panel shows the best-fitting PROSPECTOR SED model (no LRD component included) assuming a continuity SFH, with the inset showing the $2'' \times 2''$ NIRCam cutout, while the right panel shows the best-fitting EAzY template and the dashed red curve indicates the contribution from the LRD component (see Section 6.2). The lower panels display normalized residuals and the total χ^2 for each fit. The three rows are representative examples illustrating different outcomes of the reanalysis: from top to bottom, a case where an LRD component is not required, a case where fits with and without an LRD component provide comparably good descriptions of the data, and a case where including the LRD component yields a clearly improved fit.

across the sky, enabling direct measurements of field-to-field fluctuations. With 34 independent sight lines, we obtain a first estimate of the cosmic variance of the gold+silver sample and find only mild tension with abundance-matched UNIVERSEMACHINE mocks. Figure 12 shows how this test improves with increasing sight-line number: while the present dataset still yields broad uncertainties, adding ∼50–150 independent pointings, achievable within a single pure-parallel program with $\gtrsim$six NIRCam filters (J. Valenti 2026), would already tighten the constraint substantially. Importantly, the massive quiescent galaxies targeted here are several AB magnitudes brighter than typical pure-parallel depths.

Looking further ahead, the accumulation of hundreds of sight lines over multiple JWST cycles (e.g., analogous to superBoRG, which assembled $\gtrsim$300 independent HST parallel pointings collected over a decade; T. Morishita 2021) would not only tighten variance-based tests dramatically but could also begin to build statistical samples of candidate quiescent galaxies at $z \gtrsim 7$. This is difficult to achieve with wide-area near-infrared surveys such as Roman or Euclid, because the selection of these systems requires well-sampled photometry beyond 3 μm, which at present is only available with JWST.

We also emphasize that filter completeness is at least as important as depth for quiescent galaxy classification. For red, low-sSFR systems at $z > 3$, the main failure mode is often not missing faint sources but misclassifying them because the SED is poorly constrained around the Balmer/4000 Å break. In practice, continuous wavelength coverage across the short- and long-wavelength NIRCam channels is critical: removing F200W, for example, creates a gap between F150W and F277W that worsens photometric redshift stability and increases the model dependence of SED fitting. Future surveys

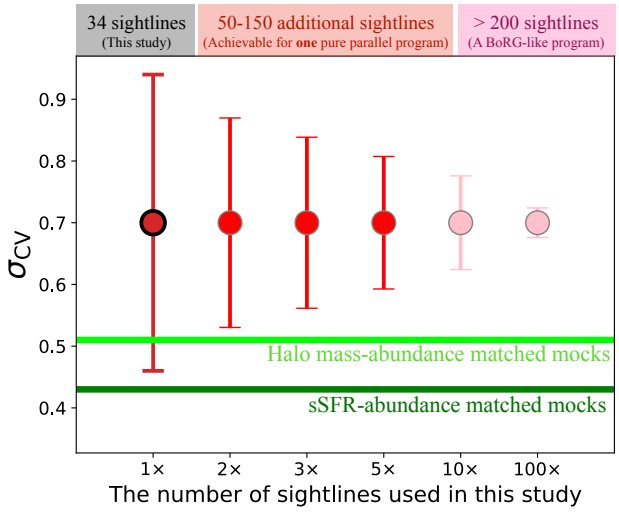

Figure 12. Expected constraints on the cosmic variance, σ_{CV}, as a function of the number of independent sight lines, in units of the current sample size. The points and error bars show the inferred σ_{CV} and its uncertainty assuming the same underlying value as measured in this work, with "1×" corresponding to the present dataset of 34 sight lines. The labels at the top indicate approximate observational regimes, from the current study to one pure-parallel JWST program and a BoRG-like multicycle survey. The horizontal green lines show the intrinsic cosmic variance of the UNIVERSEMACHINE sSFR-abundance-matched and halo-mass-abundance-matched mock samples (Sections 5.1.2 and 5.2). Larger numbers of sight lines would substantially improve the ability to distinguish the observed signal from model predictions.

optimized for quiescent galaxy demographics should therefore prioritize filter sets that minimize such gaps, even at some cost in single-band depth.

Finally, spectroscopy remains the bottleneck for turning candidates into robust constraints. Many JWST/NIRSpec confirmations come from heterogeneous programs that target relatively small, preselected samples (e.g., A. C. Carnall et al. 2024; T. Nanayakkara et al. 2024; W. M. Baker et al. 2025b). Their complex and nonuniform selection functions complicate attempts to derive an unbiased census. Recent JWST spectroscopy has also begun to address this limitation. In particular, RUBIES (A. de Graaff et al. 2025a) has explicitly attempted spectroscopic completeness to enable abundance based on spectroscopically confirmed samples (Y. Zhang et al. 2025). Nevertheless, cosmic variance remains a key limitation for any program that targets a small number of contiguous fields. A natural next step is therefore large spectroscopic programs that target uniformly selected samples across many independent sight lines, which can (i) quantify contamination rates in "silver"-like selections and (ii) directly measure quiescence diagnostics (e.g., Balmer absorption and continuum breaks) while robustly constraining the SFHs of these galaxies.

7. Summary

In this work, we present a wide-area JWST/NIRCam study designed to measure both the cosmic abundance and field-to-field variance of massive quiescent galaxies at $z \gtrsim 3$. Combining the JWST/NIRCam pure-parallel survey PANORAMIC and several legacy surveys in deep fields, we assemble the largest JWST/NIRCam imaging dataset to date for this purpose. Restricting to regions with $\geqslant$six NIRCam filters, we compile a $\sim$0.28 deg^2 ($\sim$1000 arcmin^2) dataset distributed across $\sim$34 independent sight lines. This scale enables number-density measurements that are less skewed by rare environments and less biased by small footprints, providing a direct empirical constraint on cosmic variance.

Using a carefully identified sample of quiescent galaxy candidates spanning $z \approx 3$–8 with $M_\star \geqslant 10^{10} M_\odot$, we infer number densities using a probabilistic approach. We find that the abundance of massive quiescent galaxies declines steeply with redshift: the conservative gold sample drops from $\sim 1.8 \times 10^{-5}$ Mpc^{-3} at $z = 3.0$–3.5 to $\sim 0.7 \times 10^{-6}$ Mpc^{-3} at $z \sim 6$, with the combined gold+silver sample yielding different absolute densities but a consistent evolutionary trend. These results show that massive quenched systems were already in place in the early Universe but became progressively less common toward $z \gtrsim 6$. At $z \sim 7$, we identify a small set of plausible candidates, implying a number density of order 10^{-6} Mpc^{-3}, but this estimate remains provisional because photometry alone cannot securely distinguish such systems from LRD/AGN contaminants.

We compare our measured number densities to predictions from a suite of SAMs and hydrodynamical simulations, finding that they underpredict the measured abundance of massive quiescent galaxies at $z \gtrsim 4$ by $\gtrsim$1 dex, highlighting persistent challenges in reproducing early quenching. We separately assess UNIVERSEMACHINE as an empirically calibrated baseline, anchored primarily at lower redshift and extrapolated to early times. The growing tension between UNIVERSEMACHINE and the observed abundance of massive quiescent galaxies at $z \gtrsim 5$ suggests that reproducing our measurements may require earlier and/or more efficient quenching (and/or maintenance of quiescence) than implied by relations calibrated at lower redshift.

The multiple-sight-line NIRCam coverage enables an empirical measurement of cosmic variance. We measure large field-to-field variance, with $\sigma_{CV} \approx 0.7 \pm 0.3$. We compare this observed cosmic variance to abundance-matched UNIVERSEMACHINE mocks, constructed to match the observed mean abundance. The mocks exhibit systematically smaller field-to-field variance, providing tentative evidence (large error bars) that the observed quiescent galaxy population is more strongly clustered (i.e., more highly biased) than UNIVERSEMACHINE predicts once the mean abundance is matched.

Taken together, our findings favor quenching pathways that are both rapid enough to overcome strong gas replenishment at early times and sufficiently coupled to large-scale environment to enhance the probability of quenching in biased regions. More broadly, these results suggest that quenching driven solely by internal heating within individual halos may be insufficient to explain the observed combination of high abundance and strong clustering at $z > 3$.

Finally, this work demonstrates the power of large-area, multiple-sight-line surveys for studying rare, highly biased populations in the early Universe. For massive quiescent galaxies at $z > 3$, increasing the number of independent sight lines is essential not only for reducing the influence of rare structures on abundance measurements but also for directly characterizing their large-scale distribution. The dataset presented here marks a first step in that direction: it shifts the $z > 3$ quiescent galaxy census from field-limited measurements to a variance-controlled census. We are still only at the beginning. As larger imaging and spectroscopic samples extend this approach across many more sight lines, they will turn early quenched galaxies into a powerful laboratory for understanding how the first massive galaxies formed, quenched, and traced structure in the young Universe.

Acknowledgments

This work is based in part on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with program JWST-GO-2514. Z.J., C.C.W., and K.E.W. acknowledge support for program JWST-GO-2514. Support for program JWST-GO-2514 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127. The work of C.C.W. is supported by NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00072, as well as from the Swiss National Science Foundation (SNSF) through project grant 200020_207349. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNRF140. This work was performed in part at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-2210452. P.D. warmly acknowledges support from an NSERC discovery grant (RGPIN-2025-06182).

Data Availability

The PANORAMIC data products, including the photometric catalogs, are hosted on MAST as a High-Level Science Product under doi:10.17909/fpzr-as35. The catalog of 238 massive quiescent galaxy candidates identified in this work can be found on Zenodo: doi:10.5281/zenodo.22132642.

Appendix A
Comparison of Photometry

To assess the quality of our photometric procedure described in Section 2, we compare the HST and JWST/NIRCam photometry derived from our pipeline for the initially selected sample galaxies (Section 3.1) in GOODS-S with the Kron photometry released by the JADES team (DR5; B. D. Johnson et al. 2026; B. E. Robertson et al. 2026). We note that the JADES imaging reduction and photometry were performed with a customized Guaranteed Time Observations pipeline and that their source detection and segmentation strategies differ somewhat from those used in the GRIZLI pipeline adopted in our study (Section 2). Nonetheless, as shown in Figure A1, our photometry is highly consistent with that from JADES within the uncertainties across both HST and NIRCam, indicating that, for the galaxy sample considered in this work, the results are not sensitive to the choice of pipeline. This consistency demonstrates the robustness of our measurements.

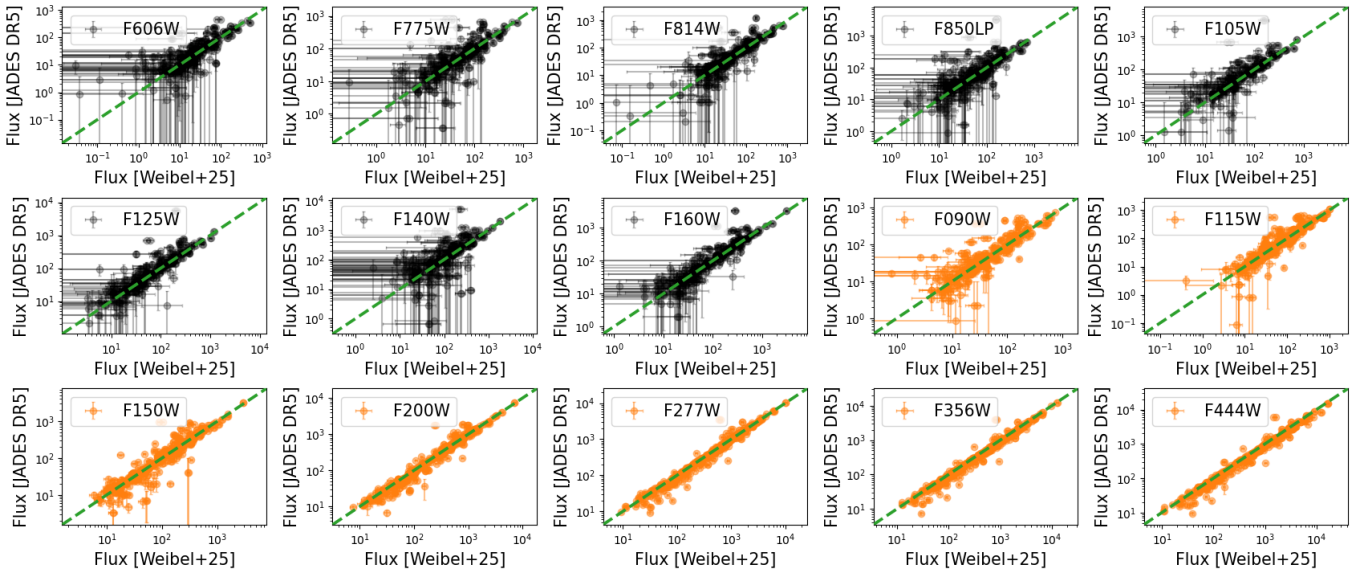

Figure A1. HST (black) and NIRCam (orange) photometric comparison for the initially selected galaxies in GOODS-S. The *x*-axis shows the photometry from our pipeline (Section 2), and the *y*-axis shows the measurements from the latest JADES DR5. The fluxes are given in units of nJy. The green dashed line marks the one-to-one relation. The two sets of photometry are highly consistent within the uncertainties.

Appendix B
SED Fitting Assumptions

We employed the FSPS stellar population synthesis code (C. Conroy et al. 2009; C. Conroy & J. E. Gunn 2010) with MIST stellar isochrone libraries (J. Choi et al. 2016; A. Dotter 2016) and MILES stellar spectral libraries (J. Falcón-Barroso et al. 2011). We applied the P. Madau (1995) intergalactic medium (IGM) absorption model and included the nebular emission model of N. Byler et al. (2017). Dust attenuation was modeled following S. Tacchella et al. (2022), in which the attenuation of nebular emission and young stellar populations was treated differently from that of old stellar populations (S. Charlot & S. M. Fall 2000). The redshift was fit as a free parameter with a flat prior of $z \in (0, 20)$. In this work, we carried out SED fitting with three SFH models that are commonly employed in the literature.

1. Parametric delayed-τ SFH—SFR(t) $\propto te^{-t/\tau}$.
2. Nonparametric SFH with the continuity prior (J. Leja et al. 2019) modeled as a piecewise step function composed of nine lookback time bins. The first two bins are fixed to 0–30 and 30–100 Myr; the last bin is fixed to $0.85t_{\rm H}$–$t_{\rm H}$, where $t_{\rm H}$ is the Hubble time at the epoch of observation; and the remaining six bins are evenly spaced in logarithmic time between 100 Myr and $0.85t_{\rm H}$. Changes in SFR between adjacent bins, $x = \log({\rm SFR}_{t_i}/{\rm SFR}_{t_{i+1}})$, are modeled as a Student's t-distribution with $\sigma = 0.3$, $\nu = 2$ (see J. Leja et al. 2019 for details), which strongly disfavors rapid changes in SFR over short timescales.
3. Nonparametric SFH with the bursty continuity prior. Similar to the continuity prior SFH but adopting $\sigma = 1$ in the Student's t-distribution, thereby permitting more bursty star formation (e.g., S. Tacchella et al. 2022).

In Figure B1, we compare $z_{\rm phot}$, M_*, SFR, and rest-frame $(U-V)$ and $(V-J)$ colors derived from the three SED runs for the initially selected sample. Except for SFR, all inferred properties are highly consistent across models. At the high-SFR end, all three SFH models produce consistent results; however, for galaxies with low sSFR, systematic offsets become apparent. The inferred SFR can differ substantially between models, underscoring the significant uncertainties associated with constraining the SFR of quiescent galaxies using photometric data alone.

We compare the goodness of fit among the three SFH models using $\chi^2_{\rm tot}$, defined as the χ^2 summed over all available photometric bands.[25] We find that approximately 60% of our quiescent sample favors—i.e., has the smallest $\chi^2_{\rm tot}$—the delayed-τ SFH, while roughly 25% and 15% favor the bursty continuity and continuity SFHs, respectively. These fractions are not sensitive to the number of available filters. This result is consistent with J. M. Helton et al. (2026), who reported that SED fitting with a delayed-τ SFH tends to better reproduce high-redshift galaxy colors than other SFHs.

We also note that the three SFH models do not contribute equally to the silver sample. Among silver galaxies with $M_* \geqslant 10^{10}\,M_\odot$, the bursty continuity fit satisfies the sSFR criterion (Equation (3)) for 84% of the sample, compared with 38% for the delayed-τ fit and 21% for the standard continuity fit. For 55% of the silver sample, the bursty continuity fit is the only one that satisfies the criterion. This trend does not change strongly with redshift, with corresponding bursty continuity fractions of 87%, 75%, and 81% at z = 3–4, 4–5, and 5–8, respectively. This difference mainly reflects the adopted SFH priors. The bursty continuity prior allows much larger changes in SFR between adjacent time bins and can therefore more easily produce a rapid decline in the most recent SFR, while the standard continuity prior disfavors such rapid changes. The delayed-τ model gives an intermediate case, since a sufficiently short τ can also produce a low present-day sSFR through the overall decline of the parametric SFH.

One concern is that our SED fitting does not include an AGN component, so inferred galaxy properties could be biased if AGN are present. Although we exclude LRD-like sources (Section 3.2.1), other AGN (e.g., X-ray and mid-IR) may remain. We estimate the impact using GOODS-S, where ultradeep Chandra data (B. Luo et al. 2017) and JWST/MIRI observations (S. Alberts et al. 2024a; G. H. Rieke et al. 2024) are available. Cross-matching with the GOODS-S AGN catalog of J. Lyu et al. (2024), which uses multiwavelength (X-ray to radio) criteria, we find that 4 of 42 ($\approx$9.5%) massive ($>10^{10}\,M_\odot$) galaxies in our gold+silver samples are classified as AGN. Using the AGN-inclusive SED fits from J. Lyu et al. (2024), we find that the rest-frame UV–near-IR SEDs of these four galaxies are still dominated by stellar emission. Prior work also suggests that for nonquasar AGN, omitting AGN components mainly adds scatter to M_* and introduces a $\sim$0.2 dex uncertainty in SFR, with no strong systematic bias (P. Santini et al. 2012; G. Yang et al. 2018; Z. Ji et al. 2022). This is unlikely to change our quiescent classification, since our galaxies lie $\gtrsim$0.5 dex below the star-forming main sequence (Figure 4). Given the low inferred AGN fraction ($\lesssim$10%) and the modest expected biases, residual AGN should have only a minor impact on our results.

[25] For the best-fit parameters, we adopt the median values of the posterior distribution. We also verify the results using the maximum-likelihood solution and find negligible differences.

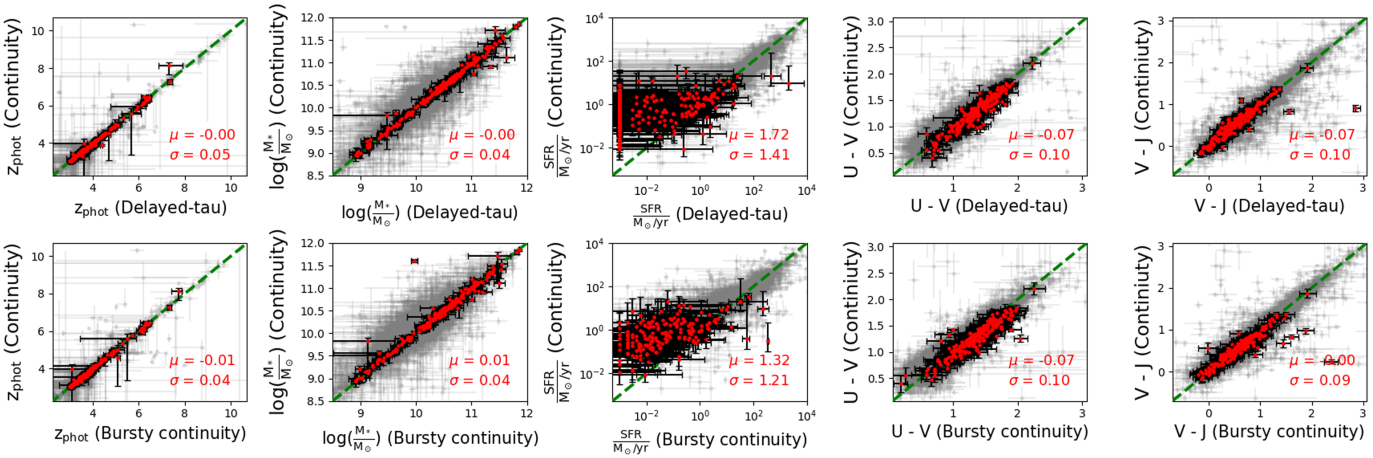

Figure B1. Comparisons of the inferred properties assuming three different SFHs (Section 3.2) for the initial sample of 2633 galaxies (Section 3.1). The green dashed line marks the one-to-one relation. For clarity, galaxies with SFR $< 10^{-3}\,M_{\odot}\,\mathrm{yr}^{-1}$ are all plotted at $10^{-3}\,M_{\odot}\,\mathrm{yr}^{-1}$. Galaxies with sSFR $\leqslant 10^{-10}\,\mathrm{yr}^{-1}$ from the nonparametric continuity prior are highlighted in red, showing the large systematic uncertainties in SFR measurements for quiescent galaxies when relying on photometric data alone. In each panel, we also label the median (μ) and standard deviation (σ) of y–x for these low-sSFR galaxies.

Appendix C
Removed Sources from Visual Inspection

A total of 20 sources were removed from the analysis based on our visual inspection of the imaging quality. The majority (13) of the defected sources are located in the PANORAMIC pointing J153500, where very bright sources in the F200W mosaic exhibited central flux deficits. This effect was likely caused by the short exposure times, which resulted in poorer-than-usual cosmic-ray rejection (C. C. Williams et al. 2025). Those authors updated the pixel-flagging procedure in J153500 to implement more aggressive cosmic-ray rejection, albeit at the cost of image quality for the brightest sources. As a result, some bright sources showed artificially reduced F200W flux in their central pixels, which in turn produced an apparent flux discontinuity between the F200W and F277W filters, mimicking the appearance of a Balmer break at $z \sim 6$ in the photometry. This effect led to the false identification of those 13 quiescent galaxy candidates in this pointing.

We note that our initial candidate selection following A. S. Long et al. (2024) did not rely on F200W (Equation (1)). Thus, the reduced F200W image quality in J153500 did not raise concerns about missing high-redshift candidates at that stage. For our refined selection based on SED fitting, however, a concern remains that legitimate $z > 3$ quiescent galaxies in this pointing might have been erroneously excluded owing to problematic F200W photometry. To assess this, we reran our SED analysis for the 32 initially selected sources in this pointing, excluding their F200W fluxes from the PROSPECTOR fitting. No new robust quiescent galaxies were identified. We are therefore confident that all qualified candidates in this pointing have been recovered, and since fainter candidates were not affected by this pixel flagging during processing, we include this pointing and its robust candidates in the subsequent analysis.

From this test with the five-filter NIRCam set (i.e., excluding F200W), we also found that the quality of the SED fitting decreases significantly. In particular, the large gap between F150W and F277W leads to an increase in the standard deviation of the $z_{\rm phot}$ measurement for our sample to 0.7 between the continuity and delayed-τ SFHs and to 0.5 between the continuity and bursty continuity SFHs—both substantially larger than the value ($\sim$0.05) found with the six-filter NIRCam set (Figure B1), which is the baseline requirement of this work (Section 2).

Appendix D
Cross-matching with Spectroscopically Confirmed Quiescent Galaxies at $z > 3$ from the Literature

We cross-matched our final samples (Section 3.3) with spectroscopically confirmed quiescent galaxies at $z \gtrsim 3$ from two recent studies: J. Antwi-Danso et al. (2025) and Y. Zhang et al. (2025). J. Antwi-Danso et al. (2025; their Table 3) compiled a list of quiescent galaxies at $z > 3$ spectroscopically confirmed through both ground-based and space-based observations (see references therein). In addition, Y. Zhang et al. (2025; their Table 2) reported quiescent galaxies from the RUBIES program confirmed with JWST/NIRSpec spectroscopy. Among the 79 galaxies reported, we found that 35 (after removing duplicates) lie at $z \geqslant 3$ and fall within our imaging coverage with six-band NIRCam photometry. Since the list of J. Antwi-Danso et al. (2025) includes a few galaxies from earlier studies based on low-S/N ground-based spectroscopy, we excluded ZF-COS-20133, ZF-COS-17779, and ZF-COS-19589, all of which were flagged as having uncertain spectroscopic confirmations in the original paper (C. Schreiber et al. 2018). After these removals, we were left with a total of 32 spectroscopically confirmed quiescent galaxies at $z > 3$ for cross-matching. A total of 29 are recovered by our selection procedure, with 19 in the gold sample and 10 in the silver sample. Four galaxies were not ultimately selected, and we examined the reasons for their exclusion.

1. 3D-EGS-26047 at $z_{\rm spec} = 3.234$ and 39138 at $z_{\rm spec} = 3.442$, both included in the sample of J. Antwi-Danso et al. (2025), were initially selected into our parent sample. In our three SED fits, both sources lie about 0.3–0.5 dex below the star-forming main sequence at $z \sim 3.5$. However, neither source is retained in our final sample, because both fail to satisfy the sSFR criterion adopted in Section 3.2, i.e., sSFR $\lesssim 10^{-10}\,\mathrm{yr}^{-1}$ at these redshifts.

   We also examine the available spectroscopic information for these two sources. For 3D-EGS-26047,

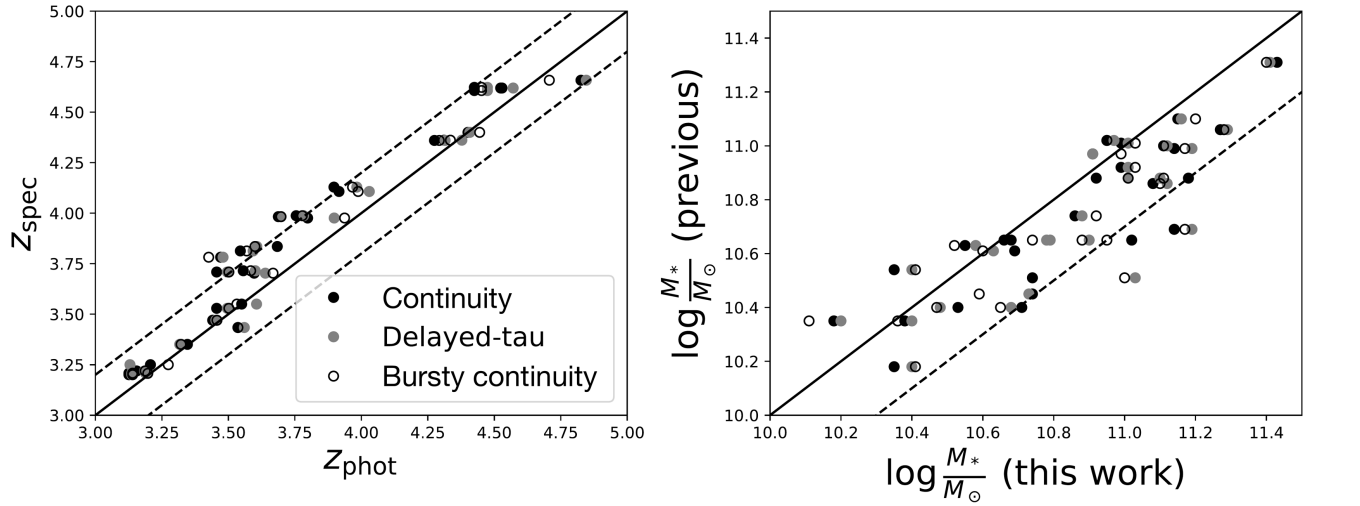

Figure D1. Comparison of our SED-derived properties with those reported in the literature for the 28 spectroscopically confirmed quiescent galaxies at $z > 3$. Results assuming the continuity, delayed-τ, and bursty continuity SFHs are shown as black and gray filled circles and black open circles, respectively. The black solid line indicates the one-to-one relation. Left: photometric versus spectroscopic redshifts, with the two black dashed lines marking $\Delta z = \pm 0.2$. Right: stellar masses from our SED fittings compared to those reported in the literature. The black dashed line marks $\Delta(\log M_*) = -0.3$ dex. We note that stellar-mass estimates in previous works are based on different methods and datasets and thus are not directly homogeneous.

the Keck/MOSFIRE spectrum is relatively noisy and does not cover Hα, but tentative detections of [O III] λ5007, Hβ, and [O II] λ3727 suggest a younger stellar population than is typical of quiescent systems. For 39138, JWST/NIRSpec prism spectroscopy shows a clear emission feature at the wavelength of Hα, although the low spectral resolution prevents the [N II] contribution from being separated. S. Jin et al. (2024) therefore conservatively estimated sSFR $= 10^{-9.9}\,\mathrm{yr}^{-1}$, placing this source very close to the selection boundary adopted here.

2. UNCOVER-18407 at $z_{\rm spec} = 3.97$ (D. J. Setton et al. 2024) was *not* initially selected, as it is marked with a `hugekron_flag` in the catalog of A. Weibel et al. (2026b). Sources with this flag were excluded from the outset (see the beginning of Section 3). Inspection of its image cutouts reveals four bright, very extended galaxies within $\lesssim 2''$ of its position, which likely caused problematic segmentation in the automatic pipelines.

Finally, for the 29 matched galaxies, we compare our SED-derived redshifts and stellar masses with those reported in the literature, as shown in Figure D1. For redshifts, our photometric values are in excellent agreement with spectroscopic measurements, with nearly all sources lying within $|\Delta z| \lesssim 0.2$. For stellar masses, we also find strong agreement, although our estimates are systematically higher by $\sim$0.1–0.2 dex on average. Such offsets are well known in the literature and are commonly attributed to systematic differences between SED modeling methods and assumptions (J. Leja et al. 2019, 2022; S. Lower et al. 2020; Z. Ji & M. Giavalisco 2022, 2023; C. K. Jespersen et al. 2026).

Appendix E
Cosmic-variance Analysis for the Gold Sample

Here we assess whether the gold sample alone provides sufficient statistics to constrain the cosmic variance on NIRCam-pointing scales. We apply the same two estimators used in the main text, namely, bootstrap variance decomposition and Bayesian distribution fitting, but restricting the analysis to the gold sample only. As Figure E1 shows, the inferred $\sigma_{\rm CV}$ using the bootstrap method frequently reaches the physical boundary at zero: $\sim$50% of the bootstrap realizations return $\sigma_{\rm CV} = 0$. This indicates that the small number of gold objects and the limited number of independent pointings prevent a measurable excess variance from being robustly detected. Consistent with this, the MCMC distribution-fitting approach yields a posterior for $\sigma_{\rm CV}$ that is strongly concentrated toward $\sigma_{\rm CV} = 0$, implying that the gold-sample counts distribution does not require super-Poisson dispersion. We therefore conclude that the gold sample by itself does not contain sufficient statistical power for a robust cosmic-variance measurement, and we adopt the combined gold +silver sample for the main cosmic-variance analysis presented in Section 5.

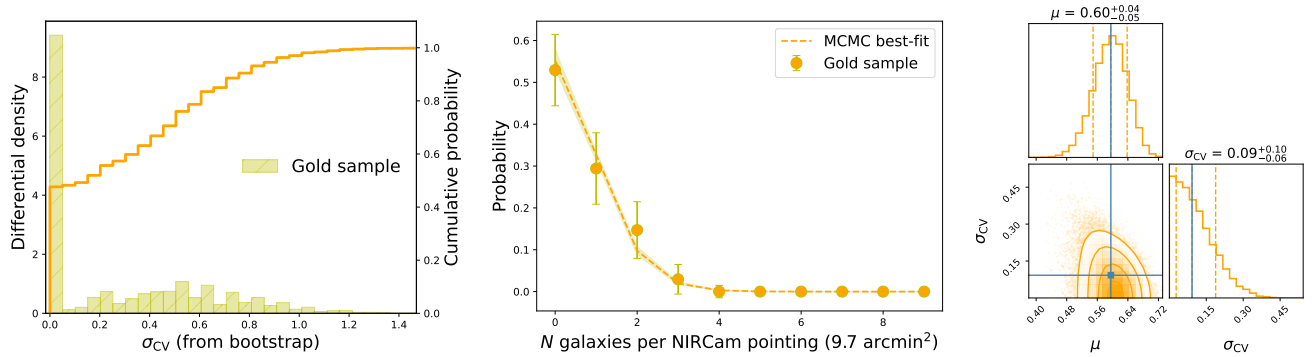

Figure E1. The left panel is similar to Figure 9 but shows the bootstrap-inferred $\sigma_{\rm CV}$ for the gold sample. Approximately 50% of the bootstrap realizations return $\sigma_{\rm CV} = 0$. The middle and right panels are similar to Figure 10 but show the MCMC-fitting-inferred $\sigma_{\rm CV}$ for the gold sample; the $\sigma_{\rm CV}$ posterior "piles up" at zero. Together, these results indicate that the gold sample alone lacks sufficient statistics for a robust cosmic-variance measurement.

Appendix F
Detectability of Cosmic Variance

Because the measured cosmic variance depends not only on the intrinsic clustering signal but also on the finite number of independent sight lines, it is useful to show explicitly how its detectability scales with survey design. For rare objects, the field-to-field variance is often dominated by Poisson fluctuations, making it difficult to distinguish a nonzero cosmic-variance term from noise. The purpose of this appendix is therefore not to provide a rigorous statistical estimator but rather to show the approximate scaling of the detectability of $\sigma_{\rm CV}$ with the mean number of objects per sight line and the number of independent fields.

Let N denote the number of galaxies per independent sight line, and let s_N^2 be the measured field-to-field sample variance across the $n_{\rm field}$ sight lines, which estimates the underlying variance Var(N). Similar to Equation (13), we write

$$\sigma_{\rm CV}^2 = \frac{{\rm Var}(N) - \langle N\rangle}{\langle N\rangle^2}, \tag{F1}$$

where $\langle N\rangle$ is the mean number of objects per sight line. Replacing the underlying variance by its sample estimator s_N^2,[26] the corresponding estimate becomes

$$\sigma_{\rm CV}^2 \approx \frac{s_N^2 - \langle N\rangle}{\langle N\rangle^2}. \tag{F2}$$

Assuming the field-to-field count distribution is Gaussian, the variance of the sample variance is

$${\rm Var}(s_N^2) \approx \frac{2\,{\rm Var}(N)^2}{n_{\rm field} - 1} \tag{F3}$$

(e.g., G. Casella & R. L. Berger 2002; M. H. DeGroot & M. J. Schervish 2012). Note that this assumption is adopted only to derive the approximate scaling relation; non-Gaussianity would modify the prefactor[27] but not the qualitative dependence on $n_{\rm field}$ and $\langle N\rangle$. The corresponding standard deviation of the sample variance is therefore

$$\delta s_N^2 \approx \sqrt{{\rm Var}(s_N^2)} \approx \sqrt{\frac{2}{n_{\rm field} - 1}}\,{\rm Var}(N). \tag{F4}$$

Propagating this uncertainty into $\sigma_{\rm CV}^2$ (Equation (F2)), we get

$$\delta(\sigma_{\rm CV}^2) \approx \frac{\delta s_N^2}{\langle N\rangle^2} = \frac{1}{\langle N\rangle^2}\sqrt{\frac{2}{n_{\rm field} - 1}}\,{\rm Var}(N). \tag{F5}$$

Therefore, the approximate S/N for detecting nonzero cosmic variance is

$${\rm S/N}(\sigma_{\rm CV}^2) = \frac{\sigma_{\rm CV}^2}{\delta(\sigma_{\rm CV}^2)} \approx \sqrt{\frac{n_{\rm field} - 1}{2}}\,\frac{\sigma_{\rm CV}^2}{1/\langle N\rangle + \sigma_{\rm CV}^2}, \tag{F6}$$

where we have used Equation (F1). Finally, because $\delta(\sigma_{\rm CV}^2) = 2\sigma_{\rm CV}\,\delta(\sigma_{\rm CV})$, we get ${\rm S}/N(\sigma_{\rm CV}) = 2\,{\rm S/N}(\sigma_{\rm CV}^2)$.

The expression above shows that, at fixed intrinsic $\sigma_{\rm CV}$, the detectability of cosmic variance increases approximately as $\sqrt{n_{\rm field}}$. This effect is especially important for rare populations, for which $\langle N\rangle$ is small. In that regime, the Poisson term dominates the field-to-field variance, so the uncertainty on the excess-variance estimate can exceed the intrinsic cosmic-variance signal. As a result, many realizations are statistically consistent with zero excess variance, and after imposing the physical condition $\sigma_{\rm CV} \geqslant 0$, they accumulate at $\sigma_{\rm CV} = 0$. The bootstrap distribution therefore develops a pileup at zero with a tail toward positive values, so summary statistics such as the median can be biased low when the number of sight lines is small. Increasing the number of independent sight lines reduces this effect by shrinking the uncertainty on the measured field-to-field variance. This is consistent with the main result of this work, in which the large number of independent fields is essential for empirically constraining the field-to-field variance of massive quiescent galaxies at $z > 3$. This also highlights that if the cosmic variance is high (especially if $\sigma_{\rm CV} > 1$), it is possible to detect it from a limited number of sight lines. For example, the fact that A. Weibel et al. (2026a) manage to detect cosmic variance at $z \sim 10$, even though the galaxy number densities are lower than in this work, is fully a result of their galaxy population having an intrinsically higher cosmic variance than the one investigated here.

[26] s_N^2 is the unbiased sample-variance estimator of the field-to-field variance, i.e., $\mathbb{E}[s_N^2] = {\rm Var}(N)$.

[27] If the field-to-field count distribution is a gamma function (C. K. Jespersen et al. 2025b), we can show that Equation (F3) becomes ${\rm Var}(s_N^2) \approx \left[\frac{2}{n_{\rm field}-1} + \frac{6}{n_{\rm field}}\left(\frac{1}{\mu} + \sigma_{\rm CV}^2\right)\right]{\rm Var}(N)^2$. This reduces to the Gaussian result in the limit of gamma $k \to \infty$, where $k = \mu/(1 + \mu\sigma_{\rm CV}^2)$, since the skewness and excess kurtosis then vanish and the distribution approaches a Gaussian.

ORCID iDs

Zhiyuan Ji https://orcid.org/0000-0001-7673-2257
Christina C. Williams https://orcid.org/0000-0003-2919-7495
Peter Behroozi https://orcid.org/0000-0002-2517-6446
Andrea Weibel https://orcid.org/0000-0001-8928-4465
Christian Kragh Jespersen https://orcid.org/0000-0002-8896-6496
Pascal A. Oesch https://orcid.org/0000-0001-5851-6649
Rachel Bezanson https://orcid.org/0000-0001-5063-8254
Katherine E. Whitaker https://orcid.org/0000-0001-7160-3632
Jenny E. Greene https://orcid.org/0000-0002-5612-3427
Gabriel Brammer https://orcid.org/0000-0003-2680-005X
Pratika Dayal https://orcid.org/0000-0001-8460-1564
Ivo Labbé https://orcid.org/0000-0002-2057-5376
Sinclaire M. Manning https://orcid.org/0000-0003-0415-0121
Pierluigi Rinaldi https://orcid.org/0000-0002-5104-8245
Mengyuan Xiao https://orcid.org/0000-0003-1207-5344
Yunchong Zhang https://orcid.org/0000-0001-6454-1699

References

Alberts, S., Lyu, J., Shivaei, I., et al. 2024a, ApJ, 976, 224
Alberts, S., Williams, C. C., Helton, J. M., et al. 2024b, ApJ, 975, 85
Antwi-Danso, J., Papovich, C., Esdaile, J., et al. 2025, ApJ, 978, 90
Antwi-Danso, J., Papovich, C., Leja, J., et al. 2023, ApJ, 943, 166
Bagley, M. B., Finkelstein, S. L., Rojas-Ruiz, S., et al. 2024, ApJ, 961, 209
Baker, W. M., Ito, K., Valentino, F., et al. 2026, A&A, 706, A91
Baker, W. M., Lim, S., D'Eugenio, F., et al. 2025a, MNRAS, 539, 557
Baker, W. M., Valentino, F., Lagos, C. d. P., et al. 2025b, A&A, 702, A270
Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143
Belli, S., Newman, A. B., & Ellis, R. S. 2019, ApJ, 874, 17
Belli, S., Park, M., Davies, R. L., et al. 2024, Natur, 630, 54
Berti, A. M., Coil, A. L., Behroozi, P. S., et al. 2017, ApJ, 834, 87
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bezanson, R., Labbe, I., Whitaker, K. E., et al. 2024, ApJ, 974, 92
Bradley, L. D., Trenti, M., Oesch, P. A., et al. 2012, ApJ, 760, 108
Brammer, G. 2023, grizli, v1.9.11, Zenodo, doi:10.5281/zenodo.8370018
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bridge, J. S., Holwerda, B. W., Stefanon, M., et al. 2019, ApJ, 882, 42
Bugiani, L., Belli, S., Park, M., et al. 2025, ApJ, 981, 25
Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44
Calvi, V., Trenti, M., Stiavelli, M., et al. 2016, ApJ, 817, 120
Carnall, A. C., Cullen, F., McLure, R. J., et al. 2024, MNRAS, 534, 325
Carnall, A. C., McLeod, D. J., McLure, R. J., et al. 2023, MNRAS, 520, 3974
Carnall, A. C., Walker, S., McLure, R. J., et al. 2020, MNRAS, 496, 695
Casella, G., & Berger, R. L. 2002, Statistical Inference (2nd ed.; Duxbury)
Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2023, ApJ, 954, 31
Cattaneo, A., Dekel, A., Devriendt, J., Guiderdoni, B., & Blaizot, J. 2006, MNRAS, 370, 1651
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Chuang, C.-Y., Jespersen, C. K., Lin, Y.-T., Ho, S., & Genel, S. 2024, ApJ, 965, 101
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Costa, T., Rosdahl, J., & Kimm, T. 2019, MNRAS, 489, 5181
Costa, T., Sijacki, D., Trenti, M., & Haehnelt, M. G. 2014, MNRAS, 439, 2146
Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827
Davies, R. L., Belli, S., Park, M., et al. 2024, MNRAS, 528, 4976
de Graaff, A., Brammer, G., Weibel, A., et al. 2025a, A&A, 697, A189
deGraaff, A., Hviding, R. E., Naidu, R. P., et al. 2026, MNRAS, 551, stag1567
de Graaff, A., Rix, H.-W., Naidu, R. P., et al. 2025b, A&A, 701, A168
de Graaff, A., Setton, D. J., Brammer, G., et al. 2025c, NatAs, 9, 280
DeGroot, M. H., & Schervish, M. J. 2012, Probability and Statistics (4th ed.; Addison-Wesley)
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Natur, 457, 451
De Lucia, G., Fontanot, F., Xie, L., & Hirschmann, M. 2024, A&A, 687, A68
Dotter, A. 2016, ApJS, 222, 8
Eisenstein, D. J., Johnson, B. D., Robertson, B., et al. 2025, ApJS, 281, 50
Eisenstein, D. J., Willott, C., Alberts, S., et al. 2026, ApJS, 283, 6
Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, A&A, 532, A95
Finkelstein, S. L., Bagley, M. B., Arrabal Haro, P., et al. 2025, ApJL, 983, L4
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gardner, J. P., Mather, J. C., Abbott, R., et al. 2023, PASP, 135, 068001
Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, Natur, 544, 71
Greene, J. E., Labbe, I., Goulding, A. D., et al. 2024, ApJ, 964, 39
Grudić, M. Y., Hopkins, P. F., Quataert, E., & Murray, N. 2019, MNRAS, 483, 5548
Hamadouche, M. L., Whitaker, K. E., Valentino, F., et al. 2026, arXiv:2602.02485
Hartley, W. G., Conselice, C. J., Mortlock, A., Foucaud, S., & Simpson, C. 2015, MNRAS, 451, 1613
Helton, J. M., Alberts, S., Rieke, G. H., et al. 2026, ApJ, 1006, 43
Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv:1008.4686
Hopkins, P. F., Murray, N., Quataert, E., & Thompson, T. A. 2010, MNRAS, 401, L19
Hviding, R. E., de Graaff, A., Miller, T. B., et al. 2025, A&A, 702, A57
Ito, K., Valentino, F., Brammer, G., et al. 2026, A&A, 710, A364
Jeffreys, H. 1946, RSPSA, 186, 453
Jespersen, C. K., Carnall, A. C., & Lovell, C. C. 2025a, ApJL, 988, L19
Jespersen, C. K., Cranmer, M., Melchior, P., et al. 2022, ApJ, 941, 7
Jespersen, C. K., Melchior, P., Spergel, D. N., et al. 2026, ApJ, 1002, 132
Jespersen, C. K., Steinhardt, C. L., Somerville, R. S., & Lovell, C. C. 2025b, ApJ, 982, 23
Ji, Z., & Giavalisco, M. 2022, ApJ, 935, 120
Ji, Z., & Giavalisco, M. 2023, ApJ, 943, 54
Ji, Z., Giavalisco, M., Kirkpatrick, A., et al. 2022, ApJ, 925, 74
Ji, Z., Giavalisco, M., Williams, C. C., et al. 2018, ApJ, 862, 135
Ji, Z., Williams, C. C., Rieke, G. H., et al. 2024a, arXiv:2409.17233
Ji, Z., Williams, C. C., Suess, K. A., et al. 2026, ApJ, 998, 239
Ji, Z., Williams, C. C., Tacchella, S., et al. 2024b, ApJ, 974, 135
Jin, S., Sillassen, N. B., Magdis, G. E., et al. 2024, A&A, 683, L4
Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, ApJS, 254, 22
Johnson, B. D., Robertson, B. E., Eisenstein, D. J., et al. 2026, arXiv:2601.15954
Kakimoto, T., Tanaka, M., Onodera, M., et al. 2024, ApJ, 963, 49
Kauffmann, G., Li, C., Zhang, W., & Weinmann, S. 2013, MNRAS, 430, 1447
Kawinwanichakij, L., Quadri, R. F., Papovich, C., et al. 2016, ApJ, 817, 9
Killi, M., Watson, D., Brammer, G., et al. 2024, A&A, 691, A52
Kimmig, L. C., Remus, R.-S., Seidel, B., et al. 2025, ApJ, 979, 15
Kocevski, D. D., Onoue, M., Inayoshi, K., et al. 2023, ApJL, 954, L4
Kriek, M., Conroy, C., van Dokkum, P. G., et al. 2016, Natur, 540, 248
Kurinchi-Vendhan, S., Farcy, M., Hirschmann, M., & Valentino, F. 2024, MNRAS, 534, 3974
Labbe, I., Greene, J. E., Bezanson, R., et al. 2025, ApJ, 978, 92
Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, MNRAS, 462, 3854
Lagos, C. d. P., Bravo, M., Tobar, R., et al. 2024, MNRAS, 531, 3551
Lagos, C. d. P., Valentino, F., Wright, R. J., et al. 2025, MNRAS, 536, 2324
Leja, J., Johnson, B. D., Conroy, C., et al. 2019, ApJ, 877, 140
Leja, J., Speagle, J. S., Ting, Y.-S., et al. 2022, ApJ, 936, 165
Long, A. S., Antwi-Danso, J., Lambrides, E. L., et al. 2024, ApJ, 970, 68
Lower, S., Narayanan, D., Leja, J., et al. 2020, ApJ, 904, 33
Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2
Lyu, J., Alberts, S., Rieke, G. H., et al. 2024, ApJ, 966, 229
Madau, P. 1995, ApJ, 441, 18
Matthee, J., Naidu, R. P., Brammer, G., et al. 2024, ApJ, 963, 129
Merlin, E., Fortuni, F., Calabró, A., et al. 2025, OJAp, 8, E170
Merlin, E., Fortuni, F., Torelli, M., et al. 2019, MNRAS, 490, 3309
Morishita, T. 2021, ApJS, 253, 4
Morishita, T., Mason, C. A., Kreilgaard, K. C., et al. 2025, ApJ, 983, 152
Morishita, T., Trenti, M., Stiavelli, M., et al. 2018, ApJ, 867, 150
Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJS, 206, 8
Naidu, R. P., Matthee, J., Katz, H., et al. 2026, Natur, 656, 329
Naidu, R. P., Matthee, J., Kramarenko, I., et al. 2024, arXiv:2410.01874
Nanayakkara, T., Glazebrook, K., Jacobs, C., et al. 2024, NatSR, 14, 3724
Nanayakkara, T., Glazebrook, K., Schreiber, C., et al. 2025, ApJ, 981, 78

Newman, J. A., & Davis, M. 2002, ApJ, 564, 567
Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023, MNRAS, 525, 2864
Pacifici, C., Kassin, S. A., Weiner, B. J., et al. 2016, ApJ, 832, 79
Park, M., Belli, S., Conroy, C., et al. 2023, ApJ, 953, 119
Peng, Y., Maiolino, R., & Cochrane, R. 2015, Natur, 521, 192
Peng, Y.-j., Lilly, S. J., KovaÄ, K., et al. 2010, ApJ, 721, 193
Perrin, M. D., Sivaramakrishnan, A., Lajoie, C.-P., et al. 2014, SPIE, 9143, 91433X
Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Popesso, P., Concas, A., Cresci, G., et al. 2023, MNRAS, 519, 1526
Remus, R.-S., & Kimmig, L. C. 2025, ApJ, 982, 30
Rieke, G. H., Alberts, S., Shivaei, I., et al. 2024, ApJ, 975, 83
Rieke, M. J., Kelly, D. M., Misselt, K., et al. 2023, PASP, 135, 028001
Roberts-Borsani, G., Bagley, M., Rojas-Ruiz, S., et al. 2025, ApJ, 983, 18
Roberts-Borsani, G., Morishita, T., Treu, T., Leethochawalit, N., & Trenti, M. 2022, ApJ, 927, 236
Robertson, B. E. 2010, ApJL, 716, L229
Robertson, B. E., Johnson, B. D., Tacchella, S., et al. 2026, arXiv:2601.15956
Rodríguez Montero, F., Davé, R., Wild, V., Anglés-Alcázar, D., & Narayanan, D. 2019, MNRAS, 490, 2139
Rojas-Ruiz, S., Bagley, M., Roberts-Borsani, G., et al. 2025, ApJ, 985, 80
Rojas-Ruiz, S., Finkelstein, S. L., Bagley, M. B., et al. 2020, ApJ, 891, 146
Rojas-Ruiz, S., Roberts-Borsani, G., Morishita, T., et al. 2026, ApJ, 1000, 285
Russell, T. A., Dobric, N., Adams, N. J., et al. 2025, MNRAS, 544, 4482
Saintonge, A., & Catinella, B. 2022, ARA&A, 60, 319
Santini, P., Rosario, D. J., Shao, L., et al. 2012, A&A, 540, A109
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Schaye, J., Kugel, R., Schaller, M., et al. 2023, MNRAS, 526, 4978
Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018, A&A, 618, A85
Setton, D. J., Khullar, G., Miller, T. B., et al. 2024, ApJ, 974, 145
Siegel, J. C., Setton, D. J., Greene, J. E., et al. 2025, ApJ, 985, 125
Somerville, R. S., Lee, K., Ferguson, H. C., et al. 2004, ApJL, 600, L171
Steinhardt, C. L., Jespersen, C. K., & Linzer, N. B. 2021, ApJ, 923, 8
Stevenson, S. D., Carnall, A. C., Leung, H.-H., et al. 2026, MNRAS, 545, staf2087
Suess, K. A., Beverage, A. G., Kriek, M., et al. 2025, ApJ, 993, 158
Suess, K. A., Weaver, J. R., Price, S. H., et al. 2024, ApJ, 976, 101
Suess, K. A., Williams, C. C., Robertson, B., et al. 2023, ApJL, 956, L42
Tacchella, S., Finkelstein, S. L., Bagley, M., et al. 2022, ApJ, 927, 170
Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, ARA&A, 58, 157
Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673
Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, MNRAS, 404, 1775
Trapp, A. C., Furlanetto, S. R., & Yang, J. 2022, MNRAS, 510, 4844
Trenti, M., Bradley, L. D., Stiavelli, M., et al. 2011, ApJL, 727, L39
Trenti, M., & Stiavelli, M. 2008, ApJ, 676, 767
Treu, T., Roberts-Borsani, G., Bradac, M., et al. 2022, ApJ, 935, 110
Valenti, J. 2026, JWST Pure Parallel Observing Scenarios, v1.0.0, Zenodo, doi:10.5281/zenodo.18462862
Valentino, F., Brammer, G., Gould, K. M. L., et al. 2023, ApJ, 947, 20
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJL, 677, L5
Weaver, J. R., Davidzon, I., Toft, S., et al. 2023, A&A, 677, A184
Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11
Weibel, A., de Graaff, A., Setton, D. J., et al. 2025a, ApJ, 983, 11
Weibel, A., Jespersen, C. K., Oesch, P. A., et al. 2026a, ApJ, 1005, 118
Weibel, A., Oesch, P. A., Barrufet, L., et al. 2024, MNRAS, 533, 1808
Weibel, A., Oesch, P. A., Williams, C. C., et al. 2026b, ApJ, 1002, 136
Weinmann, S. M., van den Bosch, F. C., Yang, X., & Mo, H. J. 2006, MNRAS, 366, 2
Whitaker, K. E., Williams, C. C., Mowla, L., et al. 2021, Natur, 597, 485
Williams, C. C., Alberts, S., Ji, Z., et al. 2024, ApJ, 968, 34
Williams, C. C., Giavalisco, M., Porciani, C., et al. 2011, ApJ, 733, 92
Williams, C. C., Oesch, P. A., Weibel, A., et al. 2025, ApJ, 979, 140
Williams, C. C., Spilker, J. S., Whitaker, K. E., et al. 2021, ApJ, 908, 54
Williams, C. C., Tacchella, S., Maseda, M. V., et al. 2023, ApJS, 268, 64
Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, ApJ, 691, 1879
Windhorst, R. A., Cohen, S. H., Jansen, R. A., et al. 2023, AJ, 165, 13
Woodrum, C., Williams, C. C., Rieke, M., et al. 2022, ApJ, 940, 39
Wu, J. F., Jespersen, C. K., & Wechsler, R. H. 2024, ApJ, 976, 37
Xiao, M., Oesch, P. A., Elbaz, D., et al. 2024, Natur, 635, 311
Xiao, M., Williams, C. C., Oesch, P. A., et al. 2025, A&A, 696, A156
Yan, H., Yan, L., Zamojski, M. A., et al. 2011, ApJL, 728, L22
Yang, G., Brandt, W. N., Vito, F., et al. 2018, MNRAS, 475, 1887
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59
Zhang, Y., de Graaff, A., Setton, D. J., et al. 2025, ApJ, 997, 252
Zhang, Y., Ji, Z., Bezanson, R., et al. 2026, ApJL, 1004, L35
Zhu, Y., Egami, E., Fan, X., et al. 2025, ApJL, 995, L5