The Astrophysical Journal, 994:57 (11pp), 2025 November 20 https://doi.org/10.3847/1538-4357/ae0292
© 2025. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

Witnessing the Violent, Merger-driven Formation of an Extremely Massive Galaxy 1.7 Gyr after the Big Bang

L.K. Pitchford[1,2], J. Cairns[3], D. Farrah[4,5], D.L. Clements[3], E. Hatziminaoglou[6,7,8], I. Pérez-Fournon[7,8], and L. Wang[9,10]

[1] Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843-4242, USA; lura.k.pitchford@gmail.com
[2] George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University, College Station, TX 77843-4242, USA
[3] Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK
[4] Department of Physics and Astronomy, University of Hawaii, 2505 Correa Road, Honolulu, HI 96822, USA
[5] Institute for Astronomy, 2680 Woodlawn Drive, University of Hawaii, Honolulu, HI 96822, USA
[6] ESO, Karl-Schwarzschild-Str. 2, 85748 Garching bei München, Germany
[7] Instituto de Astrofísica de Canarias, C/Vía Láctea, s/n, E-38205 San Cristóbal de La Laguna, Tenerife, Spain
[8] Universidad de La Laguna, Dpto. Astrofísica, E-38206 San Cristóbal de La Laguna, Tenerife, Spain
[9] SRON Netherlands Institute for Space Research, Landleven 12, 9747 AD, Groningen, The Netherlands
[10] Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands
Received 2024 September 4; revised 2025 August 14; accepted 2025 August 30; published 2025 November 13

Abstract

We combine near-infrared imaging in two bands from the Hubble Space Telescope with archival observations of molecular gas to study SDSS J160705.16+533558.6 (J1607), an extremely luminous broad-line quasar at $z = 3.65$ that is also bright in the submillimeter. Via subtraction of the quasar point-spread function, we show that its host galaxy is massive, with a stellar mass of $(5.8 \pm 3.0) \times 10^{11}\,M_{\odot}$, making it comparable to giant early-type galaxies (ETGs) at $z \sim 0$. If the supermassive black hole in the quasar is accreting at the Eddington limit, then its mass is $3.5 \times 10^{9}\,M_{\odot}$, which is also consistent with local massive ETGs. The host has an extremely high star formation rate of $4300 \pm 500\,M_{\odot}\,\mathrm{yr}^{-1}$ and a molecular gas mass of $(2.4 \pm 0.9) \times 10^{10}\,M_{\odot}$. The quasar has two companions: one at a projected separation of 11 kpc with a stellar mass of $(7.9 \pm 5.0) \times 10^{10}\,M_{\odot}$ but no detected molecular gas and one 6 kpc further away in the same direction with a molecular gas mass of $(2.6 \pm 1.3) \times 10^{10}\,M_{\odot}$ but no detected stellar emission. Since neither companion shows evidence for active galactic nucleus activity, this may represent merger-driven quenching, in which the dynamics of the merger strip molecular gas from infalling galaxies. Overall, irrespective of whether the host is merging with the companions, these properties mark J1607 as forming what will become an extremely massive ($\sim 10^{12}\,M_{\odot}$) galaxy by $z = 0$.

Unified Astronomy Thesaurus concepts: Quasars (1319); Active galaxies (17); High-redshift galaxies (734)

1. Introduction

Among the most striking results from Herschel was the discovery of a population (2–3 deg^{-2}) of extremely luminous star-forming galaxies, with star formation rates (SFRs) exceeding about $2000\,M_{\odot}\,\mathrm{yr}^{-1}$, at $z \gtrsim 2$ (C. D. Dowell et al. 2014; V. Asboth et al. 2016; F. Gao et al. 2021; L. Wang et al. 2021). Examples include SPT 0346–52 at $z = 5.66$ (forming stars at a rate of $4500\,M_{\odot}\mathrm{yr}^{-1}$; J. Ma et al. 2016), HFLS3 and SPT 0311–58 at $z = 6.34$ and $z = 6.90$ (both at $2900\,M_{\odot}\,\mathrm{yr}^{-1}$; D. A. Riechers et al. 2013; D. P. Marrone et al. 2018), ADFS-27 at $z = 5.66$ ($2400\,M_{\odot}\,\mathrm{yr}^{-1}$; D. A. Riechers et al. 2017), HXMM01 at $z = 2.31$ ($2000\,M_{\odot}\,\mathrm{yr}^{-1}$; H. Fu et al. 2013), and NA.v1.489 at $z = 2.69$ ($1900\,M_{\odot}\,\mathrm{yr}^{-1}$; H. Nayyeri et al. 2017). However, this population has proved challenging to study. There are no direct analogs at low redshifts, since virtually all sources at $z < 1$ have SFRs at least an order of magnitude lower than those seen in the extreme starburst population. Galaxy formation models either do not contain them at all or do so at space densities an order of magnitude or more below what is observed (C. M. Baugh et al. 2005; C. G. Lacey et al. 2010; C. Gruppioni et al. 2011; Q. Guo et al. 2011; C. C. Hayward et al. 2013). The trigger mechanisms and evolutionary pathways for extreme starburst galaxies thus remain unclear.

Although they are challenging to study, extreme starbursts are suspected to play an important role in the formation of massive quiescent galaxies. Such quiescent systems are already observed by $z \sim 2$ (e.g., D. Szomoru et al. 2012; K. E. Whitaker et al. 2012; S. Toft et al. 2014), making extreme starbursts plausible antecedents for them (S. Wellons et al. 2015; R. Davé et al. 2017). There is also evidence of a tight spread in stellar ages in (some) quiescent galaxies at $z \sim 2$, meaning that all of the stars in these galaxies must have formed at roughly the same time (e.g K. Glazebrook et al. 2017). The formation histories of low-redshift massive quiescent galaxies are, however, complex and not yet fully understood (e.g., R. M. McDermid et al. 2015; D. Farrah et al. 2023), so their links with high-redshift extreme starbursts have yet to be clarified.

Some insights into extreme starbursts can be gained from the most actively star-forming galaxies at low redshift. These galaxies, the ultraluminous infrared galaxies (ULIRGs; galaxies with infrared luminosities above $10^{12}\,L_{\odot}$), typically have SFRs a factor of 4–20 times lower than those seen in high-redshift extreme starbursts. ULIRGs at $z \lesssim 0.2$ are almost always mergers (e.g., D. L. Clements et al. 1996; D. Farrah et al. 2001). At higher redshifts, some ULIRGs are found in merging systems (e.g., M. Zamojski et al. 2011; R. J. Ivison et al. 2012), while others are not (J. Melbourne et al. 2009;

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.

T. Wiklind et al. 2014). The merger fraction among ULIRGs likely declines with increasing redshift (e.g., J. S. Kartaltepe et al. 2012), though it is not known if this also applies to the most luminous starbursts. ULIRGs at $z \lesssim 0.2$ virtually always harbor active star formation, and in about half of them, the star formation is accompanied by a luminous active galactic nucleus (AGN; A. Efstathiou et al. 2022; D. Farrah et al. 2022). At higher redshifts, the SFRs associated with ULIRGs are higher (e.g., M. J. Michałowski et al. 2017; B. Gullberg et al. 2018; K. Małek et al. 2018), and many are again accompanied by an AGN (D. Farrah et al. 2016, 2017). A small number of quasars at $z \gtrsim 1$ with SFRs exceeding $2000\,M_{\odot}\,\mathrm{yr}^{-1}$ are known (e.g., L. K. Pitchford et al. 2016), and rapidly star-forming galaxies adjacent to quasars at very high redshift have been identified (R. Decarli et al. 2017). However, to date, case studies of extreme starbursts at high redshift have focused on systems for which there is no evidence for a luminous AGN (H. Fu et al. 2013; D. A. Riechers et al. 2013, 2017; J. Ma et al. 2016; H. Nayyeri et al. 2017; D. P. Marrone et al. 2018; R. Pavesi et al. 2018). Only a few well-studied extreme starbursts also harbor AGN (e.g., Y. Shao et al. 2019).

Case studies of high-redshift extreme starbursts that also host luminous AGN are thus of value in helping place this population within the context of galaxy assembly. We here present such a study of SDSS J160705.16+533558.6 (hereafter J1607), a system previously shown to harbor both a broad-line luminous quasar with $L_{\mathrm{AGN}} \sim 10^{14}\,L_{\odot}$ and a starburst with $L_{\mathrm{SB}} \sim 10^{13.5}\,L_{\odot}$ (D. L. Clements et al. 2009) at a redshift of 3.65 (J. K. Adelman-McCarthy et al. 2008). Submillimeter observations further show a candidate companion, an ∼10 kpc emission region that is offset from the quasar by about 1.″5 (∼11 kpc) to the northwest. This companion is consistent with a merger (D. L. Clements et al. 2009). To describe the components of this system (quasar, host, and potential companion) and determine what role, if any, the $z > 3$ luminous starburst population might play in assembling local ellipticals, we have obtained Wide Field Camera 3 (WFC3) imaging from the Hubble Space Telescope (HST) in F110W and F160W. Extracting the host requires removing the quasar through subtracting the point-spread function (PSF). To effectively remove the quasar, we create two types of PSFs, empirical and synthetic, and compare the resultant subtractions. These PSF-subtracted images allow us to better describe the host galaxy and its companion. We combine the galaxies' properties with archival data to set the whole system within the broader context of how obscured luminous starbursts are triggered and evolve.

This paper is structured as follows. In Section 2, we introduce J1607 and describe the HST observations. In Section 3 and Section 4, we describe the PSF creation and subtraction methods, respectively. We make use of two different sets of PSFs, the empirical PSFs being built from a nearby star and the synthetic PSFs being created in TINY TIM. We then present our results in Section 5, place them into context with other galaxies in Section 6, and present our conclusions in Section 7. Throughout this work, we assume $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\Lambda} = 0.7$, and $\Omega_{\mathrm{M}} = 0.3$.

2. HST Observations

Given its high redshift, we observed J1607 using the HST WFC3 IR channel, which has a pixel scale of 0.″13, with two orbits each of the F110W (Wide *YJ*) and F160W (WFC3 *H*) filters as part of Program ID 15200 on 2018 January 28 and 2018 February 10. We used a four-point BOX-MIN dither pattern with default point and line spacings of 0.″572 and 0.″365, respectively. We also used SPARSE sampling so that the nucleus of the quasar could not saturate. Because the submillimeter companion is located more than 0.″5 from the quasar, we did not request a separate PSF observation but instead build a PSF from a nearby star as described in Section 3.1. Each dither had an exposure time of 702.9 s for a total on-source time of $2811.8\,\mathrm{s\,orbit}^{-1}$. We then use DRIZZLEPAC (S. L. Hoffmann et al. 2021) to combine each individual calibrated (*_flt*) exposure. The HST observations were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute (STScI) and can be accessed via doi:10.17909/c8wq-7923.

3. PSF Creation

To extract the flux of the star-forming companion and search for the underlying host, we must first subtract the quasar. Similar PSF subtraction methods have been applied to other high-redshift systems in the past, with varying levels of success (see, e.g., T. Goto et al. 2009; M. Mechtley et al. 2012, 2016; V. Marian et al. 2019, 2020; M. A. Marshall et al. 2020). We subtract the PSF using two methods, which differ only in how the PSFs are created. We refer to these PSFs as the empirical and synthetic PSFs. Because diffraction spikes are artifacts caused by the structure of the telescope, the empirical PSFs are likely to better model them than are the synthetic PSFs. As such, we expect the empirical PSF subtractions to provide more accurate subtractions of these features. We therefore rely on the empirical PSFs even though the synthetic PSFs are more widely used in the literature (e.g., P. G. van Dokkum et al. 2008; A. van der Wel et al. 2011; C. Villforth et al. 2014) as they should provide a better overall subtraction (see also, e.g., M. Mechtley et al. 2016).

Comparing the empirical PSFs (and the resultant subtractions) to the synthetic also allows us to test the standard approach, i.e., TINY TIM, against something that is more accurate, the empirical approach, and shows that an empirical approach can be applied even in the absence of numerous stars in the field. While some studies require multiple stars distributed over the whole field of view to accurately sample the PSF (e.g., E. Glikman et al. 2015), we find that a single star within about 20″ of the target provides an excellent sampling of the effective HST optics and WFC3 detector. Once we have our best subtraction, we compare the results to those from a more complex fitting algorithm, GALFIT (C. Y. Peng et al. 2011).[11]

3.1. Empirical PSF Creation

The empirical PSFs are derived using the Python PHOTUTILS package (L. Bradley et al. 2021), which builds the effective PSF first described by J. Anderson & I. R. King (2000). We derive an effective PSF for each of our orbits from the background-subtracted images using a single, unsaturated star in our field that is relatively close to the quasar (within 20″). The near-infrared (NIR; F110W–F160W) color of our PSF star differs from that of our quasar by 0.1 mag. We are mostly concerned with removing the

[11] http://users.obs.carnegiescience.edu/peng/work/galfit/galfit.html

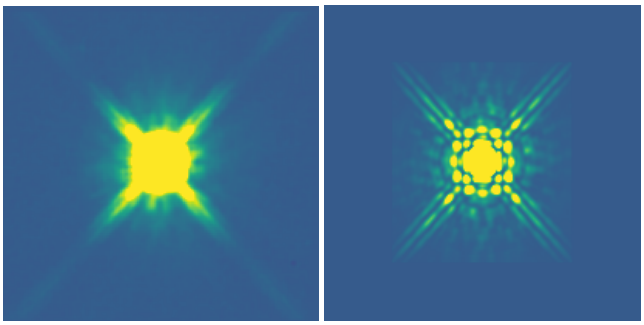


Figure 1. Empirical (left) and synthetic (right) PSFs for F110W. Each image is 51 × 51 pixels (∼6.″6 × 6.″6). As the empirical PSF is created using a star in the same field as J1607, it better describes the telescope optics at the time of observation than does the synthetic PSF.

diffraction spikes as one passes through the companion, so we do not expect the slight color mismatch to affect our results. We set the size of each PSF to be 51 × 51 pixels (∼6.″6 × 6.″6) in the native scale in order to include the diffraction spikes and to have the brightest pixel be the central pixel. Because our empirical PSFs should be unaffected by, e.g., spatial variations across the field or time variations due to the instrument focus, we expect the empirical PSFs to better describe (and later remove) the diffraction spikes when compared against the synthetic PSFs (see Section 3.2). The normalization radius is set to 4 pixels to enclose the central region of the quasar, but not the star-forming companion, in order to scale the PSF to match the flux of the quasar. We then oversample by a factor of 4 to make subpixel shifts in the position of the PSFs relative to each of the calibrated images in order to improve the accuracy of the final PSF subtraction. This is done because the WFC3 native scale is significantly undersampled. The use of the four-point dither for our observations does mitigate this, but the allowance of subpixel shifts provides an even more accurate PSF; an oversampling of four is chosen as J. Anderson (2016) finds it to work well for WFC3/IR images. To ensure that the brightest pixel is the central pixel, each axis is made odd, i.e., the oversampled PSF is 205 × 205 pixels in size. The empirical PSF is scaled such that the sum of its pixels is the square of the oversampling factor. One of the empirical PSFs is provided in the left panel of Figure 1.

3.2. Synthetic PSF Creation

The synthetic PSF models are generated using the STScI program TINY TIM, as it has long been the standard synthetic PSF-modeling software for HST observers. TINY TIM is split into three programs. TINY1 requests a series of inputs from which to generate a PSF model, including the coordinates of the source on the detector grid, the filter used for the observations, the diameter of the PSF in arcseconds, and the desired oversampling of the PSF model. TINY2 then uses these inputs to generate a slightly oversampled and undistorted model PSF. The WFC3 and Advanced Camera for Surveys instruments do suffer from significant distortion that is not corrected for by the internal optics (J. E. Krist et al. 2011), however, as they are located away from the optical axis of the HST instrument. Therefore, TINY3 applies a geometric distortion to the PSF generated by TINY2 to account for this effect.

Table 1
Flux Density Comparisons for the Empirical and Synthetic PSF Subtractions

HST Band	Component	Empirical Flux Density (μJy)	Synthetic Flux Density (μJy)
F110W	PSF	241.3 ± 7.2	276.9 ± 8.4
	Host	10.9 ± 9.9	35.7 ± 12.6
	Companion	3.1 ± 0.1	3.3 ± 0.1
F160W	PSF	280.5 ± 8.4	320.3 ± 7.2
	Host	16.7 ± 12.4	51.7 ± 14.9
	Companion	3.3 ± 0.1	3.3 ± 0.1

Note. The empirical and synthetic flux densities denote which PSF was used in the subtraction.

We generate a synthetic PSF model for each filter centered on the rough position of the quasar on the detector array, with a diameter of 4″. Additionally, we follow the method of J. Biretta (2012, 2014) and edit the TINY TIM Cold Mask parameters to their improved values to better describe artifacts such as diffraction spikes. We generate the final synthetic PSFs oversampled by a factor of 4 to match that of the empirical PSFs. Each model PSF produced by TINY TIM is scaled such that the sum of its pixels is 1. One of the synthetic PSFs is shown in the right panel of Figure 1.

4. PSF Subtraction

Once we have the empirical and synthetic PSFs, we subtract them from each individual calibrated image (corresponding to each dither position) prior to drizzling them together. We employ the same basic method for the subtractions. Once the PSF fluxes are scaled to match the quasar flux, we oversample the full field by a factor of 4, the same oversampling factor used to create the PSFs. To help determine the best alignments between the PSFs and the quasar images, we manually inspect a grid of oversampled quasar cutouts with the same size as the PSFs and with centers in a 10 × 10 pixel box around the oversampled position corresponding to the native scale center. Once the best visual alignment between the PSF and the oversampled quasar cutout is found, we subtract the two. We then paste the PSF-subtracted cutout back into the original image and regrid the image back into the native pixel scale of the observations. We repeat this for each dither for both orbits before drizzling together the eight frames to make a single image for each of the filters.

4.1. Empirical PSF Subtraction

The empirical PSFs are normalized to the square of the oversampling factor, so we must scale the PSF using the ratio of the quasar flux to the unscaled PSF flux. This results in the PSF flux densities given in the first and fourth rows of the first column of Table 1. Errors are estimated using the differences in flux density from each dithered exposure.

We use the diffraction spikes to ensure the proper alignment for each dither. The best drizzled subtractions for each filter are shown in the middle column of Figure 2. When compared to the unsubtracted drizzle products (the left column of the same figure), we see that the empirical PSFs are reasonably successful in removing the diffraction spikes.

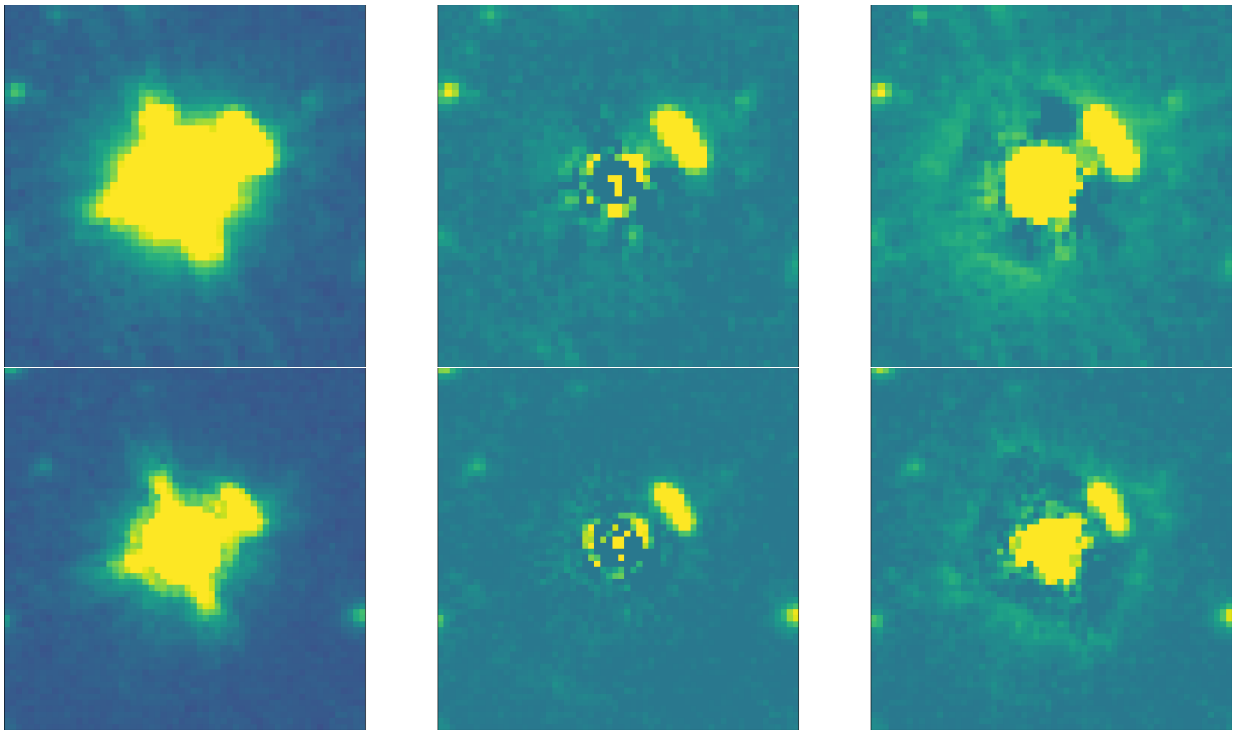


Figure 2. PSF subtraction results comparison. Each cutout is 60 × 60 pixels (7.″8 × 7.″8). The left column shows the unsubtracted images. The subtracted images using the empirical PSFs and the synthetic PSFs are shown in the middle and right columns, respectively. The top row shows the F110W images, while those for F160W are in the bottom row. In each image, north is up, and east is to the left. The scaling has been chosen to emphasize the companion galaxy; the two sets of subtracted images have a different scaling to the unsubtracted images. The empirical PSFs do a better job of removing the diffraction spikes.

4.2. Synthetic PSF Subtraction

As the synthetic PSF is already scaled such that the sum of all pixel values is 1, we can simply multiply the PSF by the quasar flux. Doing so gives the flux densities in the first and fourth rows of the second column of Table 1.

In subtracting the synthetic PSF, however, we cannot rely as heavily on the diffraction spikes to aid us in the alignment, as the synthetic PSFs are unable to accurately reproduce these features. Thus, we must also use the central quasar region to help align the PSFs to the images. Using the unsubtracted images, we find that the Airy ring peaks at about 5 pixels from the central pixel. We allow this ring to have a width of 2 pixels and use it in combination with the diffraction spikes to align the PSFs. The results from the synthetic PSF subtractions for each filter are shown in the right column of Figure 2. It is clear when comparing these to the empirical-PSF-subtracted images that the synthetic models are unable to account for the diffraction spikes. When considering the residuals, those from the empirical subtraction are smaller by more than a factor of 3. Because the synthetic models do not recreate the PSF as accurately as do our empirical models, we use the synthetic PSF subtractions instead to provide upper limits. We do also note the existence of empirical WFC3/IR PSFs created by STScI. However, because the user is unable to define the size of the PSF, these PSFs present similar issues to the synthetic PSFs in that we cannot fully remove the diffraction spikes from the image. Further, as the diffraction spikes change based on the telescope optics at the time of the observations, our empirical PSFs still most accurately define these features.

4.3. GALFIT PSF Subtraction

As a final comparison, we run GALFIT with our empirical PSFs, as it is a more sophisticated algorithm than our subtraction method. However, we try to match our initial method as much as possible. We therefore set up GALFIT in the following way. We continue to work on the individual dithers but provide a cutout of the region immediately surrounding the quasar to be fit. As we have subtracted the background already, we do not allow GALFIT to fit its own background (i.e., we input our background-subtracted images, not just the calibrated images from the HST pipeline). To ensure that the central pixels alone do not dictate the fit, we mask the central 3 × 3 pixels, as well as any bright pixels outside of the central region that GALFIT might otherwise try to fit.

We fit three separate components in GALFIT: one PSF and two Sérsic components (the quasar host and optical companion). The cutout is centered on the quasar, which then also defines the center starting point for both the PSF and the host components. The starting PSF magnitude is determined based on the total system magnitude. Since the quasar outshines its host, the host magnitude starting point is an order of magnitude fainter. We further set constraints on both the effective radius and the Sérsic index, with the former taking on values between 1 and 6 pixels and the latter taking on values between 1 and 5. For the companion, the starting center, magnitude, axis ratio, and position angle are based on a combination of the original, unsubtracted images (for the center, axis ratio, and position angle) and the remnant companion in our initial subtractions (for the magnitude). They are all, however, still allowed to vary.

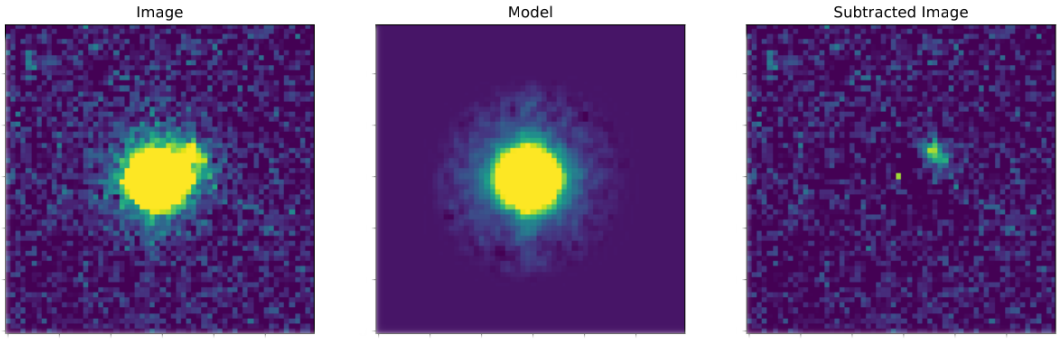


Figure 3. An example of the HSC PSF subtraction in the *y* band. Each image is 60 × 60 pixels. Only the companion galaxy remains in the residual image.

5. Results

5.1. HST Photometry

After subtracting the PSF, we extract the photometry of the host and the companion. As mentioned above, the Airy ring peaks at ∼5 pixels and is taken to have a width of 2 pixels. Therefore, to determine the flux density of the host, we use a circular aperture of radius 4 pixels, and the results are given in the second and fifth rows of Table 1 for the two sets of PSF subtractions.[12] We use an elliptical aperture centered on the companion to obtain its flux density for the two sets of subtractions; these are given in the third and sixth rows of Table 1.

There are two main sources of error in our flux densities: the calibration error, which is ∼1% for the wideband WFC3 filters (J. Mack et al. 2021), and the error that we introduce in completing the PSF subtraction. We find the former to dominate for the companion and the latter to dominate for the host. To obtain an estimate for the subtraction error that we are introducing, we allow the centers to change from the best values (i.e., the centers that best subtract the diffraction spikes). We take the standard deviation of those flux densities where the diffraction spikes appear aligned upon visual inspection to then be the subtraction-induced error. For the host, we additionally include the uncertainties on the PSFs themselves. We do not have a strong detection of the host in either filter.

While the companion flux densities are relatively consistent, the host flux densities from the synthetic subtractions are higher than those from either of the empirical subtractions. GALFIT gives the lowest host flux densities. For F110W, the GALFIT host flux density is less than a factor of 2 smaller than our flux density, but given the uncertainties (on both the GALFIT subtraction and our approach), it is entirely consistent. F160W provides a less constrained fit to the host from GALFIT but again is consistent within the uncertainties. Here, the flux density is nearly a factor of 5 lower, but the uncertainty is larger and nearly encompasses the full range of host flux densities from our initial PSF subtraction. Thus, more sophisticated subtraction routines do not make a difference for the relevant photometry. They do, however, introduce more free parameters that we cannot adequately constrain with the available data, and so we proceed, albeit cautiously, with our weak detection of the host but clear detection of the companion based off of our simplistic empirical PSF subtractions.

[12] We did also allow the host to encompass the full central region (radius of 6 pixels), but the flux densities changed by at most 5%; thus, the decision to exclude the Airy ring does not significantly change our results.

5.2. Hyper Suprime-Cam Subaru Strategic Program Photometry

In addition to the WFC3 data, we extract optical/NIR ancillary data for J1607 from the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP) Public Data Release 1 (PDR1; H. Aihara et al. 2018a). The ELAIS-N1 field is covered as part of the HSC-SSP Deep layer and includes photometry in the *g*, *r*, *i*, *z*, and *y* bands, reaching limiting magnitudes of 26.8, 26.6, 26.5, 25.6, and 24.8, respectively. Similar to the HST observations, due to the close proximity of the quasar host galaxy and the star formation companion, we must subtract the emission from the quasar in order to obtain photometry on the companion. Although the size of the PSF does change slightly between each of the HSC bands, it ranges from about $0\farcs3$ to $0\farcs4$, so we can easily separate the companion from the central quasar. We use GALFIT to complete the PSF subtraction of the quasar, first using model PSFs produced as part of the HSC-SSP PDR2 (H. Aihara et al. 2019) and, as a comparison, by fitting a Sérsic model with GALFIT. We find that the two models both produce good and comparable PSF subtractions except in the *i* band, where subtracting the HSC-SSP PSF model leaves significant residuals toward the center of the quasar compared to the Sérsic model. The difference in photometry estimated from the two sets of PSF-subtracted images is negligible. An example of the HSC PSF subtraction is given in Figure 3.

We estimate the photometry of the offset companion in each of the HSC bands using an elliptical aperture centered on the companion with the semimajor axis aligned with its elongated axis. We then increase the size of the semiminor axis of the aperture and record the aperture sum. We find that, for each band, the aperture sum increases as the semiminor axis increases to include more of the flux from the region of the companion until the region in which the emission from the quasar host has been subtracted is reached. In these regions, the PSF subtraction typically leaves small, negative residuals, and so the aperture flux starts to decrease. We therefore use the aperture sum from the semiminor axis, which corresponds to the maximum aperture sum, as this aperture should contain the maximum flux from the region of the companion without contaminating this flux with negative residuals from the PSF subtraction. We additionally visually inspect the resulting apertures to ensure proper placement. We then convert the aperture sum value to an AB magnitude using the standard

conversion factors in the HSC-SSP headers. In order to estimate the errors on our magnitudes in each band, we combine the rms noise in the image, the Poissonian error on the aperture sum (which measures counts per second), and the standard 1% calibration error (H. Aihara et al. 2018b) in quadrature before converting into an AB magnitude.

Though detected in the *r*, *i*, *z*, and *y* bands, the emission from the companion is undetected in the *g* band. In order to obtain upper limits on the *g*-band magnitude, we place another elliptical aperture in the same region as in the other bands but with a semiminor axis equal to the average of the semiminor axes of each of the apertures used for the other bands (the semiminor axis sizes in the different bands are all typically within ∼1 pixel of each other, so the average value should be comparable). We then extract the aperture sum in this region and convert this into a *g*-band magnitude.

5.3. Stellar Mass and SFR

Once we have our photometric data for the host galaxy and the companion, we determine the best-fit spectral energy distribution (SED) models using Code Investigating GALaxy Emission (CIGALE; D. Burgarella et al. 2005; S. Noll et al. 2009; M. Boquien et al. 2019; G. Yang et al. 2022). CIGALE fits galaxy SEDs from the far-UV to the radio and estimates their physical properties. For purposes of describing the quasar host and the companion, we focus on two such properties: SFR and stellar mass.

Given the lack of photometry available for both the host and the companion individually, we first use CIGALE to describe the total system (i.e., the quasar, its host, and the companion). All of the flux densities used in the fit are provided in Table 2. We assume a delayed star formation history with an optional constant burst. We allow the main stellar population age to vary between 250 and 1400 Myr with the lower bound chosen to be consistent with the highest-redshift galaxies currently observed by JWST (Y. Harikane et al. 2025). The age of the burst is allowed values between 10 and 50 Myr. We additionally use the stellar population synthesis model of G. Bruzual & S. Charlot (2003) with the initial mass function of G. Chabrier (2003) and a solar metallicity.

To describe the AGN, we use the smooth torus models of J. Fritz et al. (2006). We also fit the AGN component using the SKIRTOR (M. Stalevski et al. 2012, 2016) module, which describes a clumpy two-phase torus. The resultant AGN luminosity is lower for the clumpy torus but by less than a factor of 2. All other values from the SED fits were consistent with each other within the errors. We therefore proceed with the fit assuming the smooth torus, but our results would be largely unchanged if we instead assumed a clumpy torus.

Lastly, we rely on the models of D. A. Dale et al. (2014) to describe the dust emission. As with the AGN component, we also tested the other dust emission modules (e.g., C. M. Casey 2012; B. T. Draine et al. 2014). We found that they each produced reasonable fits to the data, as evidenced by their (reduced) chi-squared values. Though the B. T. Draine et al. (2014) module has more parameters, it does not significantly improve the fit, so we adopt the simpler D. A. Dale et al. (2014) module. Here, the AGN fraction is set to zero since we have included a separate AGN component.

The allowed parameter values for each of the modules defined above are provided in Table 3. The resultant SED fit is given in Figure 4. This gives an SFR (averaged over the last 10 Myr) of $(4500 \pm 500)\, M_{\odot}\, \mathrm{yr}^{-1}$. If we were to consider the SFR averaged over the last 100 Myr, the total system would instead show an SFR of $(2200 \pm 900)\, M_{\odot}\, \mathrm{yr}^{-1}$. However, the burst age from the SED fit is roughly 30 Myr, so we proceed with the 10 Myr SFR rather than the 100 Myr one as it better describes the current state of the system. The fit further gives a stellar mass of $(5.8 \pm 3.0) \times 10^{11}\, M_{\odot}$.

When looking at the probability distribution function (PDF) of the main stellar population age, it has a double-peaked profile at smaller values and a local maximum at higher values. Thus, to test the effects of the main population age on the SFR and stellar mass, we run two additional fits: one where the age is allowed to vary between 250 and 750 Myr (i.e., the double-peaked part of the PDF) and one where the age varies between

Table 2
Flux Densities Used in SED Fitting

Telescope/Survey	Wavelength (μm)	Flux Density (μJy/mJy)
CFHT[a]	0.4 (*u*)	4.44 ± 0.03
	0.9 (*z*)	208.36 ± 0.10
HSC-SSP[a,b]	0.5 (*g*)	60.31 ± 0.03
		0.56[d]
	0.6 (*r*)	181.82 ± 0.08
		2.08 ± 0.02
	0.8 (*i*)	162.03 ± 0.07
		2.64 ± 0.03
	0.9 (*z*)	***2.23 ± 0.02***
	1.0 (*y*)	207.27 ± 0.22
		2.28 ± 0.03
HST	1.1 (F110W)	244.9 ± 7.3
		3.1 ± 0.1
	1.6 (F160W)	287.1 ± 8.5
		3.3 ± 0.1
UKIDSS[a]	2.2 (*K*)	365.6 ± 1.4
IRAC[a]	3.6	356.8 ± 0.7
	4.5	416.1 ± 1.3
	5.8	661.8 ± 12.8
	8.0	1337.1 ± 4.6
MIPS	24[a]	5.7 ± 1.1[e]
	70[c]	22.2 ± 4.4[e]
PACS[a]	100	58.0 ± 17.8
	160	44.4 ± 26.0
SPIRE[a]	250	65.3 ± 1.9
	350	55.1 ± 2.9
	500	28.6 ± 7.1
SCUBA[c]	850	31.0 ± 6.2[e]
MAMBO[c]	1200	7.3 ± 0.9
SMA	1300 (230 GHz)	***3.13 ± 0.36***

Notes. The values in bold italics are used to describe the submillimeter companion, while the rest describe the total system (quasar + host + companion). The first part of the table (up to and including the IRAC data) provides the flux densities in μJy, while the second part gives them in mJy.
[a] Taken from the HELP catalog (R. Shirley et al. 2019).
[b] See Section 5.2 for info on extracting the companion's flux densities.
[c] D. L. Clements et al. (2009) and references therein.
[d] This represents an upper limit.
[e] To be conservative, we adopted a 20% uncertainty instead of the catalog value.

Table 3
Allowed Model Parameters in Total System CIGALE Fit to the SED

Parameter	Value
Star Formation History (`sfhdelayedbq`)	
Main stellar population *e*-folding time (Myr)	100, 500, 1000
Main stellar population age (Myr)	250–1400; increments of 50
Burst age (Myr)	10, 20, 30, 40, 50
Ratio of SFR after/before	0.1, 0.5, 1, 2, 5, 10, 25
Dust Attenuation (`dustatt_modified_starburst`)	
Amplitude of the UV bump	3.0
Dust Emission (`dale2014`)	
AGN fraction	0.0
Alpha slope	0.0625, 1.0, 2.0, 3.0, 4.0
AGN (`fritz2006`)	
Torus outer-to-inner radii ratio	60
Optical depth at 9.7 μm	6.0
Torus dust exponent for density gradient along radial direction	−1.0, −0.5, 0.0
Torus dust exponent for density gradient along polar direction	0.0, 6.0
Opening angle of torus (deg)	60, 100, 140
Angle between equatorial axis and line of sight	89.99
AGN fraction	0.0, 0.1, 0.3, 0.5, 0.7, 0.9, 0.99

Note. The angle between the equatorial axis and the line of sight is chosen as J1607 is a type 1 quasar. The same parameters are used when fitting the companion, though the AGN component is no longer considered.

800 and 1400 Myr (i.e., the increase in the PDF to higher values). The first run, equivalent to a younger system, gives an SFR of $4400 \pm 600\,M_{\odot}\,\mathrm{yr}^{-1}$ and a stellar mass of $(5.5 \pm 3.2) \times 10^{11}\,M_{\odot}$. An older system gives an SFR of $4700 \pm 400\,M_{\odot}\,\mathrm{yr}^{-1}$ and a stellar mass of $(6.1 \pm 2.8) \times 10^{11}\,M_{\odot}$. Given the uncertainties, each of the fits are consistent with one another, so we proceed with the fit covering the full range of ages.

We next consider the CIGALE-predicted attenuated stellar flux densities in our two HST filters and compare them to our empirical-PSF-subtracted values (here the sum of the host and companion values in Table 1). The CIGALE flux densities end up being a factor of 2–3 higher than expected, but the uncertainties on our flux densities are large enough that the two are just consistent within 2.3σ for F110W and 1.7σ for F160W.

Though we do not have enough data available to fully constrain the respective SEDs of either the companion or the host, we do attempt to describe the former using what is available. Because the companion does not contain a bright central source, it has more photometry than does the host; we have both the WFC3 and the HSC-SSP data. Several cutouts are provided in Figure 5, which show that we are unable to separate the quasar host and companion in all images other than those from HSC-SSP (and of course the HST images in Figure 2).

We additionally include data from new observations with the Submillimeter Array (SMA; P. T. P. Ho et al. 2004) at 230 GHz, corresponding to 1.3 mm. As older submillimeter observations showed the peak of the submillimeter emission to come from the companion rather than the host (D. L. Clements et al. 2009), we attribute all of the submillimeter flux to it. This is not to say that the host galaxy of the quasar has no dust; it just implies that the host has less dust than does the companion. We leave a more detailed description of the SMA observations to future work (J. Cairns et al. 2025, in preparation).

In fitting the SED, we use the same parameters as those provided in Table 3, neglecting the AGN component. The best fit to the SED for the companion is shown in Figure 6. From the fit, we estimate an SFR (again averaged over the last 10 Myr) of $(180 \pm 40)\,M_{\odot}\,\mathrm{yr}^{-1}$ and a stellar mass of $(7.9 \pm 5.0) \times 10^{10}\,M_{\odot}$.

Since we only have the two HST points for the host, CIGALE does not prove useful in describing the SFR or stellar mass of the galaxy. We can, however, roughly estimate these values as our total system is comprised of the quasar/its host galaxy and its companion. Thus, the difference (in, e.g., SFRs) between the total system and the companion gives rough estimates of the host's properties. This gives an SFR of $(4300 \pm 500)\,M_{\odot}\,\mathrm{yr}^{-1}$ and a stellar mass of $(5.0 \pm 3.1) \times 10^{11}\,M_{\odot}$. The SFRs and stellar masses are summarized in Table 4.

There is a possibility that the quasar light is being scattered by the interstellar medium. In such a scenario, some of the quasar light would still be contributing to the host flux, even after the PSF has been subtracted. As the quasar exhibits broad lines (and therefore is unobscured), we do not expect the host to have a comparatively high amount of obscuration. If both the AGN and its host are relatively unobscured, the contribution of the AGN in the host residuals is likely negligible. As we cannot address this potential issue further, we proceed (albeit cautiously) with the above SFR and stellar mass values for the host given the limited data available.

5.4. AGN Luminosity

From the fit to the total system SED, we extract an AGN luminosity of $(11.6 \pm 0.6) \times 10^{13}\,L_{\odot}$. This is about a factor of 2 higher than that from D. L. Clements et al. (2009), but we have incorporated newer observations taken from the Herschel Extragalactic Legacy Project (HELP; R. Shirley et al. 2019) catalog. We additionally fit the fluxes to the type 1 quasar templates of M. Polletta et al. (2007) and again obtain a value within about a factor of 2, $L_{\mathrm{AGN}} \sim 6.1 \times 10^{13}\,L_{\odot}$.

If we assume Eddington-limited accretion, this gives a soft lower limit on the black hole mass of $\log(M_{\mathrm{BH}}\,[M_{\odot}]) = 9.55$. If we instead adopt the redshift-dependent Eddington ratio (i.e., the fraction of the Eddington limit at which the AGN is actually accreting) of F. Shankar et al. (2013), which is 0.41 at $z = 3.65$, we obtain a black hole mass estimate of $\log(M_{\mathrm{BH}}\,[M_{\odot}]) = 9.93 \pm 0.30$. This agrees within the uncertainties with the virial black hole mass estimate of S. Rakshit et al. (2020), $\log(M_{\mathrm{BH}}\,[M_{\odot}]) = 9.74 \pm 0.08$, which was found using the C IV emission properties of J1607 and the M. Vestergaard & B. M. Peterson (2006) relation.

Considering both the black hole mass and the stellar mass of the host, we find that J1607 is consistent with other quasars at similar redshifts (J. Kormendy & L. C. Ho 2013). It is additionally comparable to the local elliptical population (e.g., J. Kormendy & L. C. Ho 2013; A. E. Reines & M. Volonteri 2015).

6. Discussion

J1607 is a case study of an extremely luminous early-stage merger 1.7 Gyr after the Big Bang. It harbors both a rapidly accreting supermassive black hole (SMBH) and two separate star-forming progenitors. Among the small number of $z > 3$ luminous galaxies that have been studied in detail (e.g., J. Ma et al. 2015; D. A. Riechers et al. 2017; R. Pavesi et al. 2018;

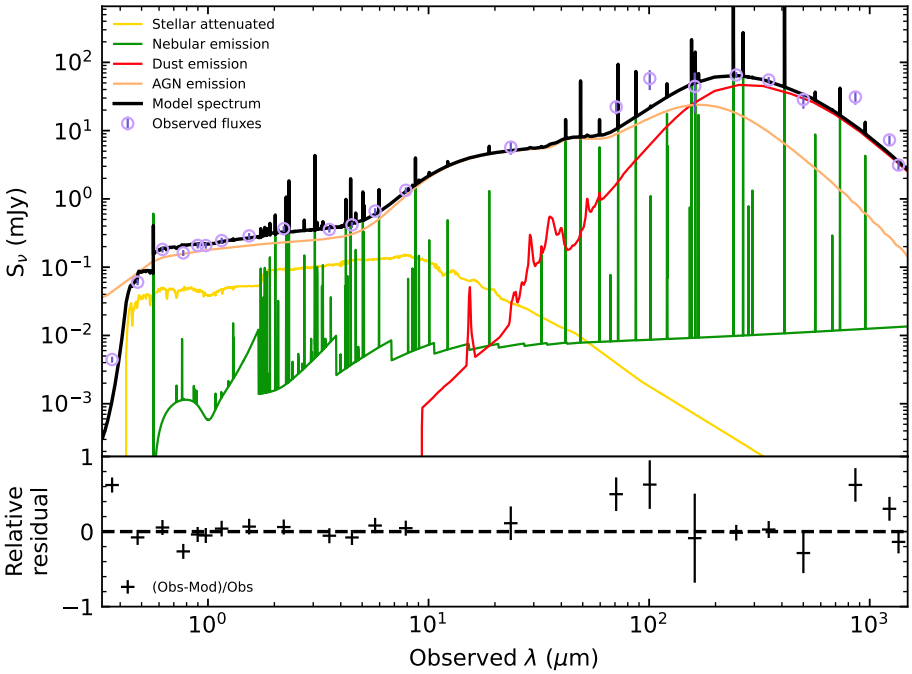


Figure 4. SED fit for the total system. The total fit is depicted by the solid black curve. The orange line describes the AGN component, while the red line shows the starburst component. For a comparison of the stellar fluxes between the CIGALE fit, our PSF subtractions, and the GALFIT PSF subtractions, see Sections 5.1 and 5.3.

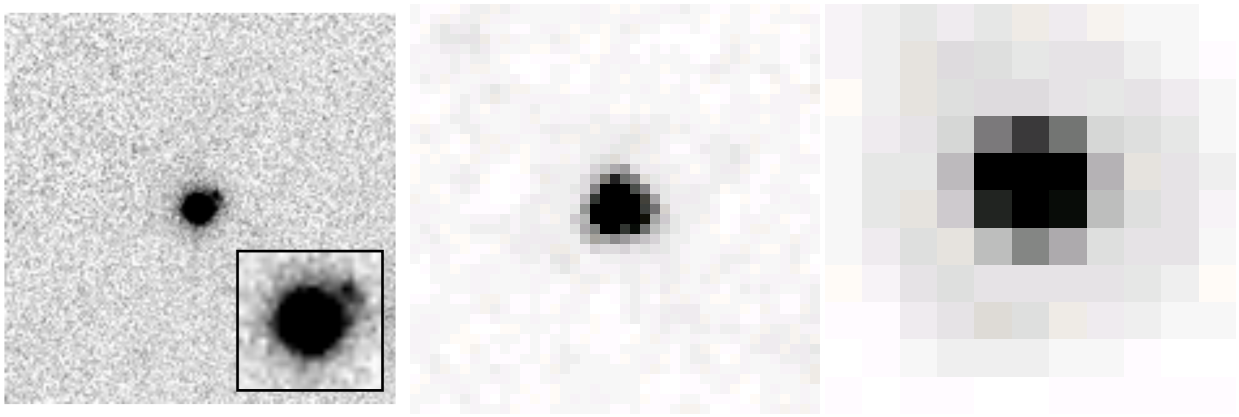


Figure 5. J1607 images. From left to right: HSC-SSP y, IRAC 3.6 μm, and MIPS 24 μm. All images are 25" on a side, with north up and east to the left. Within these cutouts, the companion is only visible in the HSC-SSP image, as seen by the zoom-in of the inner $2'' \times 2''$.

Y. Shao et al. 2019), very few have these properties (e.g., R. Decarli et al. 2017). J1607 thus offers a relatively unexplored window into early, luminous galaxy assembly. We first discuss the properties of the two galaxies (AGN host and submillimeter companion) in context with other galaxies. We then show one possible future path for the merger.

6.1. Comparison with Other High-redshift Galaxies

There exists a correlation between SFR and stellar mass among the majority of the star-forming galaxy population called the star-forming "main sequence" (e.g., K. G. Noeske et al. 2007; J. S. Speagle et al. 2014). This main sequence evolves with redshift; that is, at higher redshifts, main-sequence galaxies exhibit higher SFRs on average compared to their lower-redshift counterparts (e.g., D. Elbaz et al. 2011; G. Rodighiero et al. 2011; C. Schreiber et al. 2015). Starbursts lie above this main sequence as they are forming new stars at higher-than-average rates. The

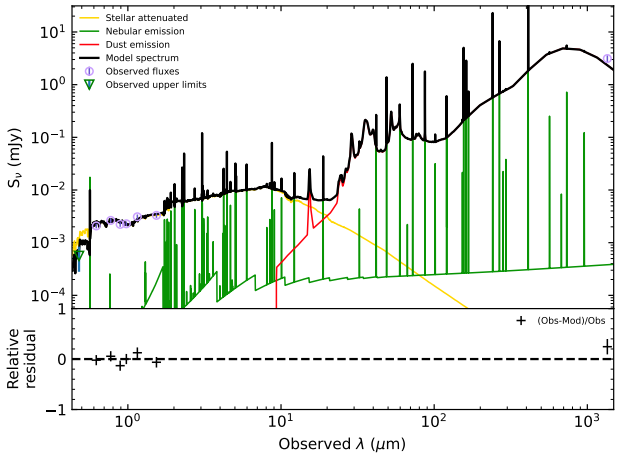


Figure 6. SED fit for the offset star formation companion. The total fit is depicted by the solid black curve, while the starburst component used to describe the submillimeter data is shown by the red line.

Table 4
System Properties from CIGALE SED Fits

Total System	
SFR ($M_\odot$ yr^{-1})	4500 ± 500
M_* (10^{11} $M_\odot$)	5.8 ± 3.0
Companion	
SFR ($M_\odot$ yr^{-1})	180 ± 40
M_* (10^{10} $M_\odot$)	7.9 ± 5.0
Host	
SFR ($M_\odot$ yr^{-1})	4300 ± 500
M_* (10^{11} $M_\odot$)	5.0 ± 3.1
L_{AGN} (10^{13} $L_\odot$)	11.6 ± 0.6
M_{BH} (10^9 $M_\odot$)	3.5 ± 0.2

Note. As we do not have enough photometry for the host to constrain its SED, we instead take the difference between the total system properties and the companion properties to describe the host. The SFRs are averaged over the last 10 Myr.

star formation observed in starbursts is often attributed to major mergers (E. Daddi et al. 2010), while main-sequence star formation is likely due to internal secular processes. This does not, however, preclude main-sequence galaxies from being involved in ongoing mergers (M. Sparre et al. 2015).

Assuming the main-sequence relation (Equation (28)) of J. S. Speagle et al. (2014) and the stellar mass resultant from the CIGALE fit, the total J1607 system lies above the main sequence by at least a factor of 2. If we assume that the total system is only the quasar host and the submillimeter companion, the host similarly lies above the main sequence at its assumed stellar mass. The companion, however, is consistent with the low end of the main sequence.[13] It is therefore plausible to assume that J1607 is a merger wherein the resultant burst in star formation in the quasar host causes it to lie above the main sequence, consistent with mergers triggering the most extreme starbursts at all epochs.

6.2. Galaxy Assembly through Mergers

Given its redshift and its active assembly of stellar and SMBH mass, J1607 is a candidate for being a major assembly episode of a massive, quiescent galaxy by $z \sim 2$. In this section, we constrain the properties of the quiescent galaxy that J1607 may signpost the assembly of and how these properties compare with the expected descendants of other far-infrared luminous galaxies. To do so, we consider the SFR of J1607 in context with its stellar mass and molecular gas mass. This requires an estimate of the molecular gas content of J1607 and suitable comparison populations.

We use two literature studies of the molecular gas in J1607. D. Iono et al. (2012) present a detection of ^{12}CO $J = 4$–3, a tracer of the gas in star-forming regions, with an estimated total gas mass of $M_{gas} = (1.1 \pm 0.2) \times 10^{11}\,M_\odot$. J. Fogasy et al. (2022) present a detection of the ^{12}CO $J = 1$–0 line, which traces cold gas. They split the CO emission into two main components: the quasar host and a companion that is

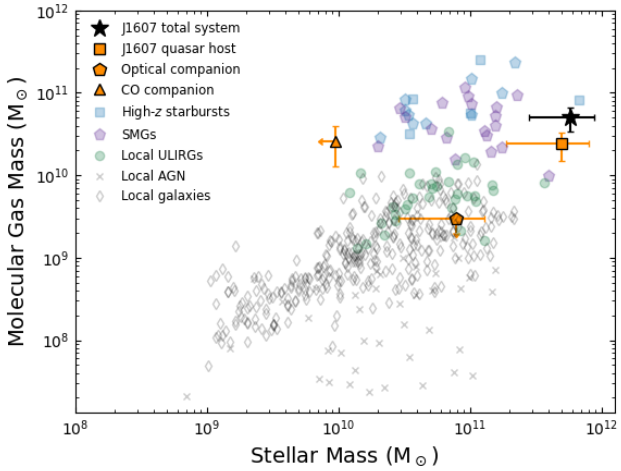


Figure 7. Molecular gas vs. stellar mass. We consider the total J1607 system (black star), as well as each of its components (shown in orange). We then compare to high-redshift starbursts (blue squares; D. A. Riechers et al. 2013, 2017; H. Nayyeri et al. 2017; D. P. Marrone et al. 2018; R. Pavesi et al. 2018), high-redshift SMGs (purple pentagons; H. Engel et al. 2010; M. S. Bothwell et al. 2013; H. Fu et al. 2013), local ULIRGs (green circles; D. Farrah et al. 2022), local AGN (black crosses; B. Husemann et al. 2022; I. Smirnova-Pinchukova et al. 2022), and the normal local galaxy population (black diamonds; A. Saintonge et al. 2017). The quasar host aligns with other high-redshift systems. Since the CO companion likely has a lower stellar mass than reported here, it is extremely gas-rich. The submillimeter companion is comparatively gas-poor, perhaps due to gas stripping over the course of the merger.

$\sim$16.8 kpc to the northwest of the quasar host. The gas masses for the host and the CO companion, respectively, are $(2.4 \pm 0.9) \times 10^{10}\,M_\odot$ and $(2.6 \pm 1.3) \times 10^{10}\,M_\odot$. We thus consider the following components of J1607: the system as a whole, the quasar host, the submillimeter companion (located 11 kpc to the northwest of the host), and the CO companion (located 17 kpc to the northwest of the host). As comparison samples,[14] we consider the following. First, we compile other hyperluminous infrared galaxies at high redshift with similar SFRs to J1607 (D. A. Riechers et al. 2013, 2017; H. Nayyeri et al. 2017; D. P. Marrone et al. 2018; R. Pavesi et al. 2018). Second, we include a collection of submillimeter bright galaxies (SMGs) at high redshift, which are infrared-luminous but typically have SFRs about a factor of 5 lower than in J1607 (H. Engel et al. 2010; M. S. Bothwell et al. 2013; H. Fu et al. 2013). Third is a sample of $z \sim 0$ ULIRGs (D. Farrah et al. 2022). Though they are about an order of magnitude less luminous than J1607, these ULIRGs are the most luminous sources in the nearby Universe and serve as a baseline from which to estimate possible trends with redshift. Finally, to extend the comparison with $z \sim 0$ systems to lower luminosities, we include local AGN (B. Husemann et al. 2022; I. Smirnova-Pinchukova et al. 2022) and normal galaxies (A. Saintonge et al. 2017), both of which are about an order of magnitude less luminous than the local ULIRG sample.

The three components of J1607, as well as the comparison samples, are shown in the M_{gas}–M_* plane in Figure 7. Considering first J1607, the components of this system have an interesting spread in properties. The quasar host has a stellar mass of $(5.0 \pm 3.1) \times 10^{11}$ $M_\odot$ and a gas mass of $(2.4 \pm 0.9) \times 10^{10}\,M_\odot$, giving a gas mass fraction of $\sim$0.05.

[13] If we instead consider the high-redshift, uncalibrated description (Equation (30) of J. S. Speagle et al. 2014), the uncertainties are large enough such that both the host and the companion are consistent with the main sequence. The host is also consistent with the high end of the main sequence when considering the SFR averaged over the last 100 Myr.

[14] Where necessary, all masses have been converted assuming a Chabrier initial mass function for the stellar masses and a CO-to-H_2 conversion factor of 0.8 $M_\odot$ (K km s^{-1}pc^2)$^{-1}$ for the molecular gas masses.

The submillimeter companion is detected by HST but is not detected in CO (with the caveat that about 10% of the J. Fogasy et al. 2022 CO emission overlaps spatially with the submillimeter emission). This gives a stellar mass of $(7.9 \pm 5.0) \times 10^{10}\ M_{\odot}$ and an upper limit on the molecular gas mass of $\lesssim 3 \times 10^{9}\ M_{\odot}$. Its gas mass fraction is thus at most 0.04 (0.09 when considering the uncertainty on the stellar mass). In contrast, the CO companion is not detected in the HST imaging. It has a molecular gas mass of $(2.6 \pm 1.3) \times 10^{10}\ M_{\odot}$ and an upper limit on the stellar mass of $\lesssim 9.6 \times 10^{9}\ M_{\odot}$. This makes the CO companion's gas mass fraction at least 0.73 (0.58 when considering the large uncertainty on the gas mass), illustrating the diversity in gas mass fractions that is possible within the components of a galaxy at high redshift and highlighting the dynamically complex nature of high-redshift active galaxies.

Turning to comparisons between J1607 and other populations: J1607 as a whole aligns with the comparison hyperluminous starburst and SMG samples, though on the higher end in terms of stellar mass. The quasar host is also consistent with the other hyperluminous starbursts and SMGs. The submillimeter and CO emission, however, are not. The gas-poor submillimeter companion is in the same part of the $M_{\rm gas}$–M_{*} parameter space as the local ULIRG population, possibly even lower. The gas-rich CO companion is on the edge of the distribution of other high-redshift infrared-luminous galaxies, making it one of the most gas-rich objects known at high redshift. Neither the submillimeter companion nor the CO companion show any evidence for AGN activity (though we cannot rule it out). We therefore argue that the submillimeter companion is gas-poor because molecular gas has been tidally stripped from it due to the dynamics of the merger. This lends support to the idea of merger-induced gas stripping as a mechanism distinct from AGN feedback to quench star formation in high-redshift galaxies. A plausible scenario is one in which a companion galaxy infalling toward the quasar host had its gas largely stripped from it during infall, leading to the now spatially separated submillimeter companion and CO companion. Improved constraints on the velocity field of the gas and stars in J1607 would, however, be required to confirm or refute this idea.

We now consider possible futures for J1607. If all of the detected gas reservoir in J1607 acts as fuel for star formation with 100% efficiency during the current burst, then the depletion time for the gas, assuming the D. Iono et al. (2012) mass, is (24 ± 5) Myr. This is smaller than the depletion timescales for most SMGs (e.g., T. R. Greve et al. 2005; D. A. Riechers et al. 2011) but does align with other rapidly star-forming quasars (e.g., D. A. Riechers et al. 2006). The resulting total stellar mass is $(4\text{–}10) \times 10^{11}\ M_{\odot}$. This is comparable to the total stellar masses of massive early-type galaxies at $z \sim 0$ (e.g., C.-P. Ma et al. 2014; K. Mehrgan et al. 2019). Such systems also exist at $z \sim 2$ in small numbers (e.g., A. Cimatti et al. 2004; E. Daddi et al. 2005; I. Damjanov et al. 2009; S. Sherman et al. 2020). Considering instead the components of J1607, and assuming the CO companion does not merge, gives a similar result. Assuming the quasar host converts 100% of its gas to stars, then its depletion time is (6 ± 2) Myr, and its final total stellar mass is $(2\text{–}8) \times 10^{11}\ M_{\odot}$. A similar analysis for the submillimeter companion yields a final total stellar mass of $(3\text{–}13) \times 10^{10}\ M_{\odot}$. If the host and the submillimeter companion then merge, the resultant system would have a stellar mass in the range of $(2\text{–}9) \times 10^{11}\ M_{\odot}$. The consistency between the predicted final total stellar mass of J1607 and massive local early-type galaxies, along with the likely endpoints of major mergers as early-type systems, supports the idea that at least some massive ETGs locally complete the bulk of their assembly by $z \sim 2$.

7. Conclusions

We have presented HST WFC3 imaging data in F110W and F160W for J1607 at $z = 3.65$. We show that this system is comprised of three components: an AGN host galaxy and two companion sources at projected separations of 11 kpc and 17 kpc in the same direction. To investigate the properties of this system, we first remove the light from the AGN via a careful subtraction of the quasar PSF and then combine the HST imaging with archival data on the molecular gas content of the system. We test two methods for creating the PSFs, one built from a nearby star and one that is synthetic, and find that the former is better able to reproduce the diffraction spikes. We extract photometry for all three sources and combine it with archival data to fit the the SEDs with CIGALE to determine their physical properties. The AGN host SED is consistent with a stellar mass of $(5.0 \pm 3.1) \times 10^{11}\ M_{\odot}$ and an SFR of $(4300 \pm 500)\ M_{\odot}\ \mathrm{yr}^{-1}$. The closer companion SED gives a stellar mass of $(7.9 \pm 5.0) \times 10^{10}\ M_{\odot}$ and an SFR of $(180 \pm 40)\ M_{\odot}\ \mathrm{yr}^{-1}$. The AGN host is consistent with a very massive, extremely star-forming galaxy lying above the stellar mass–SFR main sequence at this epoch, while the closer companion likely lies on the main sequence. The SMBH mass of the AGN host, assuming Eddington-limited accretion, is $\sim 3.5 \times 10^{9}\ M_{\odot}$. The total stellar-mass-to-SMBH-mass ratio of the AGN host is consistent with the locally observed relation for massive early-type galaxies. Regardless of whether the closer companion merges with the AGN host and how long their star formation episodes last, these properties mark J1607 as being the primal assembly episode of a massive early-type galaxy.

The two companion sources present markedly different stellar and molecular gas properties. The closer companion is not detected in CO, giving an upper limit on its molecular gas mass of $\lesssim 3 \times 10^{9}\ M_{\odot}$. The more distant companion is clearly detected in CO, with a molecular gas mass of $(2.6 \pm 1.3) \times 10^{10}\ M_{\odot}$, but is not detected in our HST imaging, setting an upper limit on its stellar mass of $\lesssim 9.6 \times 10^{9}\ M_{\odot}$. Since neither companion shows evidence for AGN activity, and both lie in the same radial direction, we propose that this may signpost a form of merger-driven quenching. In this scenario, the dynamics of the merger, as well as or instead of AGN activity, are responsible for stripping molecular gas from the closer companion and thus inhibiting its star formation. This highlights the possibility that merger-driven feedback could play an important role in quenching star formation at high redshift.

Acknowledgments

This research is based on observations made with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. These observations are associated with program 15200. The Submillimeter Array is a joint project between the Smithsonian Astrophysical Observatory and the Academia Sinica Institute of Astronomy and Astrophysics and is funded by the Smithsonian Institution and

the Academia Sinica. The Herschel Extragalactic Legacy Project (HELP) is a European Commission Research Executive Agency funded project under the SP1-Cooperation, collaborative project, small or medium-scale focused research project, FP7-SPACE-2013-1 scheme, grant agreement No. 607254. D.L.C. and J.C. acknowledge support from STFC, in part through grant Nos. ST/K001051/1 and ST/V005359/1.

For the purpose of open access, the author(s) has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising.

ORCID iDs

L.K. Pitchford https://orcid.org/0000-0002-5206-5880
D. Farrah https://orcid.org/0000-0003-1748-2010
D.L. Clements https://orcid.org/0000-0002-9548-5033
E. Hatziminaoglou https://orcid.org/0000-0003-0917-9636
I. Pérez-Fournon https://orcid.org/0000-0002-2807-6459
L. Wang https://orcid.org/0000-0002-6736-9158

References

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2008, ApJS, 175, 297
Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, PASJ, 71, 114
Aihara, H., Arimoto, N., Armstrong, R., et al. 2018b, PASJ, 70, S4
Aihara, H., Armstrong, R., Bickerton, S., et al. 2018a, PASJ, 70, S8
Anderson, J. 2016, Empirical Models for the WFC3/IR PSF WFC3 2016-12, Space Telescope WFC Instrument Science Report
Anderson, J., & King, I. R. 2000, PASP, 112, 1360
Asboth, V., Conley, A., Sayers, J., et al. 2016, MNRAS, 462, 1989
Baugh, C. M., Lacey, C. G., Frenk, C. S., et al. 2005, MNRAS, 356, 1191
Biretta, J. 2012, TinyTIM Modeling of WFC3/IR Images WFC3 2012-13, Space Telescope WFC Instrument Science Report
Biretta, J. 2014, Improved TinyTIM Models for WFC3/IR WFC3 2014-04, Space Telescope WFC Instrument Science Report
Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Bothwell, M. S., Smail, I., Chapman, S. C., et al. 2013, MNRAS, 429, 3047
Bradley, L., Sipőcz, B., Robitaille, T., et al. 2021, astropy/photutils: v1.0.2, Zenodo, doi:10.5281/zenodo.4453725
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Casey, C. M. 2012, MNRAS, 425, 3094
Chabrier, G. 2003, PASP, 115, 763
Cimatti, A., Daddi, E., Renzini, A., et al. 2004, Natur, 430, 184
Clements, D. L., Petitpas, G., Farrah, D., et al. 2009, ApJL, 698, L188
Clements, D. L., Sutherland, W. J., McMahon, R. G., & Saunders, W. 1996, MNRAS, 279, 477
Daddi, E., Renzini, A., & Pirzkal, N. 2005, ApJ, 626, 680
Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686
Dale, D. A., Helou, G., Magdis, G. E., et al. 2014, ApJ, 784, 83
Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, ApJ, 695, 101
Davé, R., Rafieferantsoa, M. H., & Thompson, R. J. 2017, MNRAS, 471, 1671
Decarli, R., Walter, F., Venemans, B. P., et al. 2017, Natur, 545, 457
Dowell, C. D., Conley, A., Glenn, J., et al. 2014, ApJ, 780, 75
Draine, B. T., Aniano, G., Krause, O., et al. 2014, ApJ, 780, 172
Efstathiou, A., Farrah, D., Afonso, J., et al. 2022, MNRAS, 512, 5183
Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119
Engel, H., Tacconi, L. J., Davies, R. I., et al. 2010, ApJ, 724, 233
Farrah, D., Baloković, M., Stern, D., et al. 2016, ApJ, 831, 76
Farrah, D., Efstathiou, A., Afonso, J., et al. 2022, MNRAS, 513, 4770
Farrah, D., Petty, S., Croker, K. S., et al. 2023, ApJ, 943, 133
Farrah, D., Rowan-Robinson, M., Oliver, S., et al. 2001, MNRAS, 326, 1333
Farrah, D., Petty, S., Connolly, B., et al. 2017, ApJ, 844, 106
Fogasy, J., Knudsen, K. K., & Varenius, E. 2022, A&A, 660, A60
Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, MNRAS, 366, 767
Fu, H., Cooray, A., Feruglio, C., et al. 2013, Natur, 498, 338
Gao, F., Wang, L., Efstathiou, A., et al. 2021, A&A, 654, A117
Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, Natur, 544, 71
Glikman, E., Simmons, B., Mailly, M., et al. 2015, ApJ, 806, 218
Goto, T., Utsumi, Y., Furusawa, H., Miyazaki, S., & Komiyama, Y. 2009, MNRAS, 400, 843
Greve, T. R., Bertoldi, F., Smail, I., et al. 2005, MNRAS, 359, 1165
Gruppioni, C., Pozzi, F., Zamorani, G., & Vignali, C. 2011, MNRAS, 416, 70
Gullberg, B., Swinbank, A. M., Smail, I., et al. 2018, ApJ, 859, 12
Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101
Harikane, Y., Inoue, A. K., Ellis, R. S., et al. 2025, ApJ, 980, 138
Hayward, C. C., Narayanan, D., Kereš, D., et al. 2013, MNRAS, 428, 2529
Ho, P. T. P., Moran, J. M., & Lo, K. Y. 2004, ApJL, 616, L1
Hoffmann, S. L., Mack, J., Avila, R., et al. 2021, The Drizzlepac Handbook v2.0 (Baltimore, MD: STScI), 216.02
Husemann, B., Singha, M., Scharwächter, J., et al. 2022, A&A, 659, A124
Iono, D., Hatsukade, B., Kohno, K., et al. 2012, PASJ, 64, L2
Ivison, R. J., Smail, I., Amblard, A., et al. 2012, MNRAS, 425, 1320
Kartaltepe, J. S., Dickinson, M., Alexander, D. M., et al. 2012, ApJ, 757, 23
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Krist, J. E., Hook, R. N., & Stoehr, F. 2011, Proc. SPIE, 8127, 81270J
Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2010, MNRAS, 405, 2
Ma, C.-P., Greene, J. E., McConnell, N., et al. 2014, ApJ, 795, 158
Ma, J., Gonzalez, A. H., Vieira, J. D., et al. 2016, ApJ, 832, 114
Ma, J., Gonzalez, A. H., Spilker, J. S., et al. 2015, ApJ, 812, 88
Mack, J., Olszewksi, H., & Pirzkal, N. 2021, WFC3/IR Filter-Dependent Sky Flats WFC2 2021-1, Space Telescope WFC Instrument Science Report
Małek, K., Buat, V., Roehlly, Y., et al. 2018, A&A, 620, A50
Marian, V., Jahnke, K., Andika, I., et al. 2020, ApJ, 904, 79
Marian, V., Jahnke, K., Mechtley, M., et al. 2019, ApJ, 882, 141
Marrone, D. P., Spilker, J. S., Hayward, C. C., et al. 2018, Natur, 553, 51
Marshall, M. A., Mechtley, M., Windhorst, R. A., et al. 2020, ApJ, 900, 21
McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, MNRAS, 448, 3484
Mechtley, M., Jahnke, K., Windhorst, R. A., et al. 2016, ApJ, 830, 156
Mechtley, M., Windhorst, R. A., Ryan, R. E., et al. 2012, ApJL, 756, L38
Mehrgan, K., Thomas, J., Saglia, R., et al. 2019, ApJ, 887, 195
Melbourne, J., Bussman, R. S., Brand, K., et al. 2009, AJ, 137, 4854
Michałowski, M. J., Dunlop, J. S., Koprowski, M. P., et al. 2017, MNRAS, 469, 492
Nayyeri, H., Cooray, A., Jullo, E., et al. 2017, ApJ, 844, 82
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793
Pavesi, R., Riechers, D. A., Sharon, C. E., et al. 2018, ApJ, 861, 43
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W., 2011 GALFIT: Detailed Structural Decomposition of Galaxy Images, ascl:1104.010
Pitchford, L. K., Hatziminaoglou, E., Feltre, A., et al. 2016, MNRAS, 462, 4067
Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81
Rakshit, S., Stalin, C. S., & Kotilainen, J. 2020, ApJS, 249, 17
Reines, A. E., & Volonteri, M. 2015, ApJ, 813, 82
Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, Natur, 496, 329
Riechers, D. A., Carilli, L. C., Walter, F., et al. 2011, ApJL, 733, L11
Riechers, D. A., Leung, T. K. D., Ivison, R. J., et al. 2017, ApJ, 850, 1
Riechers, D. A., Walter, F., Carilli, C. L., et al. 2006, ApJ, 650, 604
Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJL, 739, L40
Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, ApJS, 233, 22
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2013, MNRAS, 428, 421
Shao, Y., Wang, R., Carilli, C. L., et al. 2019, ApJ, 876, 99
Sherman, S., Jogee, S., Florez, J., et al. 2020, MNRAS, 499, 4239
Shirley, R., Roehlly, Y., Hurley, P. D., et al. 2019, MNRAS, 490, 634
Smirnova-Pinchukova, I., Husemann, B., Davis, T. A., et al. 2022, A&A, 659, A125
Sparre, M., Hayward, C. C., Springel, V., et al. 2015, MNRAS, 447, 3548
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Stalevski, M., Fritz, J., Baes, M., Nakos, T., & Popović, L. Č 2012, MNRAS, 420, 2756
Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288
Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, ApJ, 749, 121
Toft, S., Smolčić, V., Magnelli, B., et al. 2014, ApJ, 782, 68
van der Wel, A., Rix, H.-W., Wuyts, S., et al. 2011, ApJ, 730, 38
van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJL, 677, L5
Vestergaard, M., & Peterson, B. M. 2006, ApJ, 641, 689
Villforth, C., Hamann, F., Rosario, D. J., et al. 2014, MNRAS, 439, 3342
Wang, L., Gao, F., Best, P. N., et al. 2021, A&A, 648, A8
Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, MNRAS, 449, 361
Whitaker, K. E., Kriek, M., van Dokkum, P. G., et al. 2012, ApJ, 745, 179
Wiklind, T., Conselice, C. J., Dahlen, T., et al. 2014, ApJ, 785, 111
Yang, G., Boquien, M., Brandt, W. N., et al. 2022, ApJ, 927, 192
Zamojski, M., Yan, L., Dasyra, K., et al. 2011, ApJ, 730, 125