The Astrophysical Journal Letters, 946:L39 (7pp), 2023 April 1
https://doi.org/10.3847/2041-8213/acc322

Broad Emission Lines in Optical Spectra of Hot, Dust-obscured Galaxies Can Contribute Significantly to JWST/NIRCam Photometry

Jed McKinney[1], Luke Finnerty[2], Caitlin M. Casey[3], Maximilien Franco[3], Arianna S. Long[3,12], Seiji Fujimoto[1,4,5,12], Jorge A. Zavala[6], Olivia Cooper[1], Hollis Akins[1], Alexandra Pope[7], Lee Armus[8], B. T. Soifer[9], Kirsten Larson[10], Keith Matthews[9], Jason Melbourne[9], and Michael Cushing[11]

[1] Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard Stop C1400, Austin, TX 78712, USA; jed.mckinney@austin.utexas.edu
[2] Department of Physics & Astronomy, 430 Portola Plaza, University of California, Los Angeles, CA 90095, USA
[3] Department of Astronomy, The University of Texas at Austin, Austin, TX, USA
[4] Cosmic Dawn Center (DAWN), Jagtvej 128, DK-2200 Copenhagen N, Denmark
[5] Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen Ø, Denmark
[6] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[7] Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
[8] IPAC, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[9] Division of Physics, Math, and Astronomy, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[10] AURA for the European Space Agency (ESA), Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[11] Department of Physics and Astronomy, University of Toledo, 2801 West Bancroft Street, Toledo, OH 43606, USA

Received 2022 December 20; revised 2023 February 27; accepted 2023 March 8; published 2023 April 4

Abstract

Selecting the first galaxies at $z > 7 - 10$ from JWST surveys is complicated by $z < 6$ contaminants with degenerate photometry. For example, strong optical nebular emission lines at $z < 6$ may mimic JWST/NIRCam photometry of $z > 7$–10 Lyman-break galaxies (LBGs). Dust-obscured $3 < z < 6$ galaxies in particular are potentially important contaminants, and their faint rest-optical spectra have been historically difficult to observe. A lack of optical emission line and continuum measures for $3 < z < 6$ dusty galaxies now makes it difficult to test their expected JWST/NIRCam photometry for degenerate solutions with NIRCam dropouts. Toward this end, we quantify the contribution by strong emission lines to NIRCam photometry in a physically motivated manner by stacking 21 Keck II/NIRES spectra of hot, dust-obscured, massive ($\log M_*/M_\odot \gtrsim$ 10–11) and infrared (IR) luminous galaxies at $z \sim$ 1–4. We derive an average spectrum and measure strong narrow (broad) [O III]$_{5007}$ and Hα features with equivalent widths of 130 ± 20 Å (150 ± 50 Å) and 220 ± 30 Å (540 ± 80 Å), respectively. These features can increase broadband NIRCam fluxes by factors of $1.2 - 1.7$ (0.2–0.6 mag). Due to significant dust attenuation ($A_V \sim 6$), we find Hα+[N II] to be significantly brighter than [O III]+Hβ and therefore find that emission-line dominated contaminants of high $-z$ galaxy searches can only reproduce moderately blue perceived UV continua of $S_\lambda \propto \lambda^\beta$ with $\beta > -1.5$ and $z > 4$. While there are some redshifts ($z \sim 3.75$) where our stack is more degenerate with the photometry of $z > 10$ LBGs at $\lambda_{\rm rest} \sim$ 0.3–0.8 μm , redder filter coverage beyond $\lambda_{\rm obs} > 3.5$ μm and far-IR/submillimeter follow-up may be useful for breaking the degeneracy and making a crucial separation between two fairly unconstrained populations, dust-obscured galaxies at $z \sim$ 3–6 and LBGs at $z > 10$.

Unified Astronomy Thesaurus concepts: High-redshift galaxies (734); Luminous infrared galaxies (946); Lyman-break galaxies (979)

1. Introduction

A major objective baked into the design of JWST is detecting the light from the first galaxies residing at ultrahigh redshifts ($z > 10$). Delivering on its promise, more than 30 galaxy candidates with photometric redshift solutions favoring $z > 10$ were identified within the first months of publicly available data (Bradley et al. 2022; Donnan et al. 2023; Finkelstein et al. 2022a; Harikane et al. 2023; Naidu et al. 2022a; Yu-Yang Hsiao et al. 2022). Assessing the fidelity of these samples is critical, particularly because the statistics assuming current $z > 10$–15 candidates are real may or may not violate ΛCDM predictions (Boylan-Kolchin 2022; Labbe et al. 2022; Maio & Viel 2022; Naidu et al. 2022a).

Spectroscopic confirmation is needed to verify these redshifts. However, some early attempts at spectroscopic follow-up using facilities such as ALMA have yielded upper limits or tentative low-signal-to-noise ratio (SNR) detections (e.g., Bakx et al. 2023; Fujimoto et al. 2022; Kaasinen et al. 2023; Yoon et al. 2022). JWST/NIRSpec has proven capable of spectroscopically detecting the rest-frame optical light from galaxies up to $z \sim$ 9–10 (Carnall et al. 2023; Roberts-Borsani et al. 2022), but this might not be well suited for rapidly validating redshifts in statistical samples. A complimentary approach, born from similar rest-frame optical colors of $z > 10$ Lyman-break galaxies (LBGs), dusty galaxies (Howell et al. 2010; Casey et al. 2014), and in some cases similar optical/near-IR apparent magnitudes (Zavala et al. 2023), is to use far-IR/submillimeter follow-up observations of cold dust continuum (Zavala et al. 2023) and/or far-IR cooling lines (Fujimoto et al. 2022) to identify or rule out $z < 6$ IR-luminous galaxies lurking within $z > 10$ candidate catalogs.

Dusty sources have posed a problem to the fidelity of high $-z$ galaxy catalogs since they were selected from Hubble Space

---

[12] NASA Hubble Fellow.

Telescope (HST) extragalactic deep fields (e.g., Dunlop et al. 2007). HST samples at $z \sim 6$–7 were contaminated by $z \sim 2$ dusty star-forming galaxies (DSFGs). Now, $z \sim 3$–6 dusty galaxies may be contaminating $z > 7$–10 JWST samples. A contributing factor to this contamination is that both populations have similar, and uncertain, number densities: DSFGs at $z \sim 3$–6 have volume number densities $n \sim 10^{-5}$–$10^{-6}\,\mathrm{Mpc}^{-3}$ (Koprowski et al. 2017; Michałowski et al. 2017; Rowan-Robinson et al. 2018; Dudzevičiūtė et al. 2020; Gruppioni et al. 2020; Long et al. 2022; Manning et al. 2022), similar to early measurements of bright $z > 10$ LBGs (Bouwens et al. 2022; Finkelstein et al. 2022a; Harikane et al. 2023; Naidu et al. 2022a). Disentangling these two populations is therefore also crucial for reducing uncertainties in their respective number densities, which are currently inflated by sample purity (e.g., Bouwens et al. 2022).

Of particular concern within the rest-frame optical/near-IR spectra of IR-luminous galaxies is the relative contribution of strong, narrow, and broad emission lines to broadband filter fluxes, which could mask red continuum slopes produced by dust attenuation. Strong nebular lines can change JWST/NIRCam colors (Zackrisson et al. 2008; Schaerer & de Barros 2009; Stark et al. 2013; Wilkins et al. 2013, 2020, 2022). This may be a promising tool for pseudospectroscopy of lower-redshift galaxies using narrowband filters but in this particular instance is a source of potential population confusion for broadband high-redshift surveys. Indeed, some of the hottest and most luminous dusty galaxies at $z > 1$ exhibit very high rest-frame optical line equivalent widths (EWs; Finnerty et al. 2020). These arise from a combination of low dust-attenuated continuum levels with bright lines emergent from less obscured regions, as well as ionized outflows driven by active galactic nuclei (AGNs). To what extent do these strong lines contribute to JWST photometry?

In this Letter, we take an empirically grounded approach and quantify the contamination from emission lines from hot, dust-obscured galaxies to JWST/NIRCam photometry. In Section 2 we describe Keck II/NIRES observations of a sample of four luminous IR galaxies ($\log L_{\mathrm{IR}}/L_{\odot} \sim 12.5$) and 17 hot, dust-obscured galaxies (DOGs; $\log L_{\mathrm{IR}}/L_{\odot} > 13$) at $z \sim 1$–4, which we stack to derive an average optical spectrum (Section 3). We quantify the contribution of strong and broad optical emission lines to NIRCam fluxes in Section 4 and discuss their impact on distinguishing between such sources at $z < 6$ and $z > 10$ LBGs. Section 5 summarizes our main findings. Throughout this work we adopt a ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2. Sample and Data

Hot DOGs were originally selected from WISE photometry as W1W2-dropouts and include the most luminous galaxies in the universe (Eisenhardt et al. 2012; Wu et al. 2012). This extreme population is experiencing a rapid phase of both supermassive black hole and stellar mass assembly (Eisenhardt et al. 2012) and are mostly found at $z \sim 2$–3 with $\log L_{\mathrm{IR}}/L_{\odot} \geqslant 13$ (Wu et al. 2012; Assef et al. 2015; Tsai et al. 2015) and hot dust temperatures $T_{d \sim 100}$ K (Wu et al. 2012), which are ∼40–60 K warmer than the dust temperatures found in $\log L_{\mathrm{IR}}/L_{\odot} \sim 12$ galaxies at $z \sim 1$–4 (Chapman et al. 2005; Magnelli et al. 2012; Swinbank et al. 2014; Drew & Casey 2022). Most hot DOGs exhibit strong ionized outflows in optical spectroscopy (Wu et al. 2018; Finnerty et al. 2020; Jun et al. 2020), as implied by broad-line components likely driven by radiative AGN feedback (Wu et al. 2018). These sources are rare with only ∼1000 over the full WISE all-sky survey (Cutri et al. 2012).

The were previously described in Finnerty et al. (2020). In brief, we obtained simultaneous *JHK* spectra at $R \sim 2700$ with Keck II/NIRES (Wilson et al. 2004) and reduced the data using SPEXTOOL (Cushing et al. 2004). Flux calibration was performed by comparing the integrated flux with K' photometry (see Finnerty et al. 2020 for details). Our stacked spectrum uses the 17 sources with detections of [O III], Hβ and/or [N II], and Hα.

In addition to the hot DOGs, we include in our analysis previously unpublished Keck II/NIRES spectra of four $z \sim 1$–2 galaxies with $\log M_*/M_{\odot} \sim 11$ and $\log L_{\mathrm{IR}}/L_{\odot} \sim 12.5$: GS 3 ($z = 0.544$, R.A./decl. = 03:32:08.66, −27:47:34.4), GS 7 ($z = 1.042$, R.A./decl. = 03:32:26.49, −27:40:35.7), GN 1 ($z = 1.432$, R.A./decl. = 12:36:45.83, +62:07:54.0), and GN 40 ($z = 1.609$, R.A./decl. = 12:36:49.65, +62:07:38.6). These targets were selected for existing Spitzer/InfraRed Spectrograph mid-infrared spectroscopy and bright IRAC Ch. 1 photometry from a 24 μm selected parent sample (Kirkpatrick et al. 2012, 2015). GS 3, GS 7, and GN 40 are mid-IR AGNs (Kirkpatrick et al. 2015), and GN 1 is a composite source with a mid-IR AGN fraction of 50%. These sources have L_{IR} on average 1 order of magnitude lower than those of the hot DOGs. We reduce the data for this subsample following the exact same procedures as the hot DOGs described in Finnerty et al. (2020). [O II], [O III], Hβ and [N II], Hα are individually detected with EWs on average lower than those of the hot DOGs but within their range.

3. Analysis

In this work we compute synthetic photometry in broadband JWST/NIRCam filters for the rest-frame optical spectrum of our stacked spectrum redshifted at $z = 1$–9. While we do not claim our sample to be definitively representative of *all* dusty systems owing to the extreme nature of hot DOGs, the final stack is derived from empirical data with no modeling required. While a more detailed study exploring a range of continuum templates with added nebular lines is warranted, such analysis is beyond the scope of this work given the current lack of constraint on rest-frame optical spectra of dusty galaxies beyond $z > 4$–5. To supplement our analysis of very luminous, massive hot DOGs, we also compute synthetic photometry for the average DSFG spectrum from Casey et al. (2017). The Casey et al. (2017) stack is constructed from Keck/MOSFIRE spectra of 20 LIRGs and ULIRGs with $\langle z \rangle = 2.1$, a more typical IR-bright galaxy population selected from ground-based single-dish submillimeter surveys (Casey et al. 2013).

3.1. Stacking

Prior to stacking the data, we convert the observed wavelength range of each spectrum to the rest frame with spectroscopic redshifts derived from optical lines with low errors ($\Delta z \sim 10^{-3}$; Finnerty et al. 2020). Next, we rebin the spectra to a common wavelength grid corresponding to the lowest rest-frame spectral resolution ($R \sim 6400$). Finally, we calculate the sigma-clipped mean continuum flux from line-free regions, which we use to normalize each spectra in the stack.

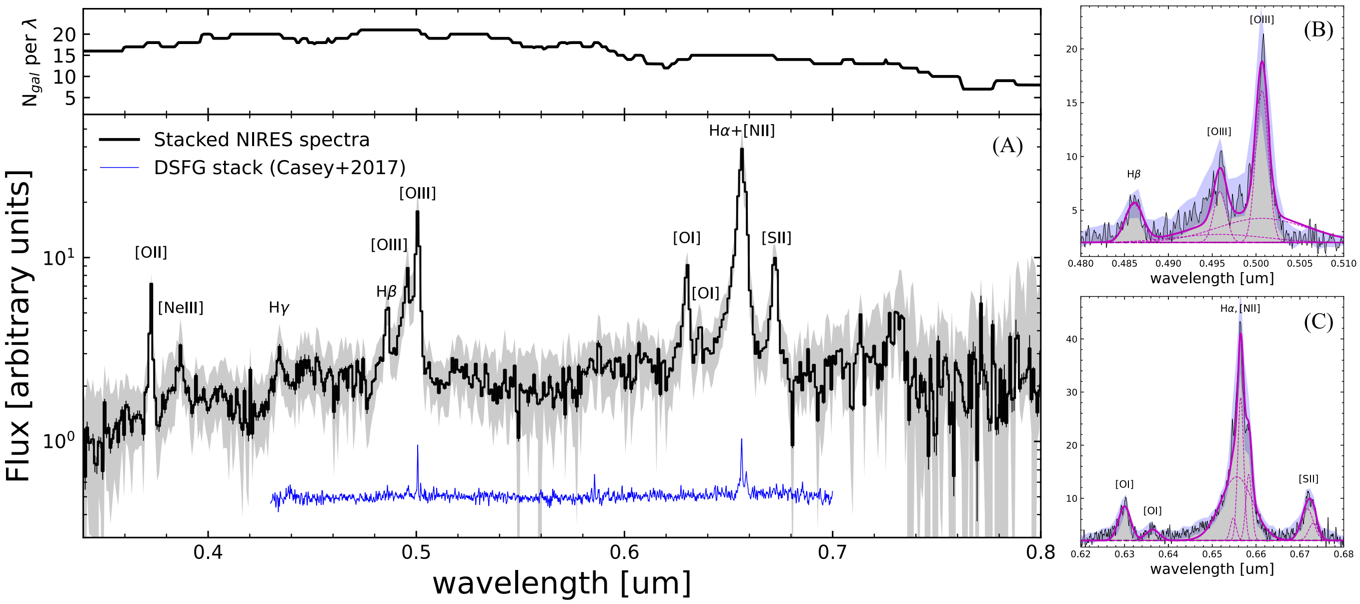

Figure 1. Stacked rest-frame optical spectrum of $z \sim 1$–4 IR-luminous galaxies detected with Keck II/NIRES. (A) The mean-weighted stacked spectrum (black). Shaded gray errors correspond to 16th–84th percentiles on the flux density derived from the bootstrapped stack distribution per wavelength. The upper panel gives the number of galaxies included in the stack as a function of wavelength. On average, 70% (>14) of the sample is represented at $\lambda_{\rm rest} = 0.34$–$0.80$ μm. We compare against the stacked (N =20) continuum-normalized DSFG spectrum from Casey et al. (2017; blue). Panels (B) and (C) show zoomed-in views on the [O III], Hβ and [N II], Hα, [S II], [O I] features, respectively. We measure broad and narrow components as expected from the individual spectra (Finnerty et al. 2020). Line fits are shown in purple with solid lines indicating the total line+continuum fit and dashed lines for individual line profiles. Sixteenth and 84th percentiles derived from 1000 bootstrapped stacks are shown in blue.

We tested multiple stacking procedures and found that a mean noise-weighted continuum stack produced the cleanest continuum and highest line signal-to-noise ratios (SNRs). In the stack, the input spectrum is first normalized by its sigma-clipped mean flux and then weighted by the spectral uncertainty per channel. This ensures that the stack is not dominated by particularly noisy spectral regions and/or the brightest spectra. As the goal of this analysis is the relative contribution of strong emission lines to photometry, we are not concerned with absolute normalization of the spectrum. To quantify the uncertainty on the continuum and line profiles, we repeat the stacking analysis 1000 times, using in each iteration 21 random samples of the input spectra with replacement ("bootstrapping"). From the bootstrapped uncertainties we determine that our final stacked spectrum is reliable at $\lambda_{\rm rest} = 0.34$–$0.8$ μm.

The final stacked spectrum is shown in Figure 1. We detect strong [O III], Hβ, [N II], and Hα emission lines, as well as [S II], [O II], [O I], [Ne III], and Hγ. The [O I]$_{6300}$/Hα line ratio is 0.5 ± 0.1, which is on the high end of the distribution measured for Seyfert galaxies in the Swift-BAT AGN Spectroscopic Survey (Koss et al. 2017). EWs for the strong lines around the [O III], Hβ and [N II], Hα complexes are listed in Table 1. Quoted uncertainties correspond to the standard deviation of EWs measured for 100 realizations of the spectrum perturbed by the spectral uncertainty per channel and assuming a 10% error on the continuum (uncertainties increase by a factor of 2.3 assuming a 20% error on the continuum).

We also compare our stacked spectrum to the stack from Casey et al. (2017) derived from 20 MOSFIRE spectra of DSFGs.[13] The stack of Casey et al. (2017) exhibits narrower emission lines than our spectrum and does not contain the broad outflow signatures found in the hot DOG rest-frame optical spectra (Wu et al. 2018; Finnerty et al. 2020).

Table 1
Strong Optical Emission-line Characteristics in Our Stacked Spectrum

Line	EW (Å)	FWHM (km $\rm s^{-1}$)
Hβ	45 ± 12	1450
H$\alpha_{\rm narrow}$	222 ± 27	730
H$\alpha_{\rm broad}$	540 ± 80	4000
[O III]$_{5007}$	127 ± 19	1000
[O III]$_{4959}$	43 ± 8	1000
[O III]$_{5007,\rm broad}$	144 ± 49	7300
[O III]$_{4959,\rm broad}$	48 ± 32	7300
[O I]$_{6300}$	109 ± 20	1600
[O I]$_{6363}$	38 ± 13	1600
[N II]$_{6548}$	35 ± 8	730
[N II]$_{6583}$	102 ± 17	730
[S II]$_{6716}$	103 ± 18	1200
[S II]$_{6730}$	43 ± 11	1200
A_V (H$\alpha_{\rm narrow}$ /Hβ)	6 ± 1	
A_V (H$\alpha_{\rm tot}$ /Hβ)	10 ± 1	

3.2. Synthetic Photometry

To test the effect of emission lines on JWST/NIRCam photometry, we subtract strong spectral features from the stack to produce a line-free continuum spectrum. To do so, we subtract the Gaussian model fits from the stack. These lines include all shown in Figures 1(B)–C) and include the range of velocity components required to fit individual hot DOGs, namely: broad [O III] and Hα emission, narrow [O I], [O III], [S II], Hα, and Hβ (Finnerty et al. 2020). We do not mask out [O II], [Ne III], and Hγ in this exercise as their lower EWs

---

[13] Available at http://www.as.utexas.edu/~cmcasey/downloads.html.

correspond to significantly less increase in broadband fluxes relative to [O III]+Hβ and [N II]+Hα. Following Finnerty et al. (2020), we assume narrow and broad profiles across different lines arise from the same kinematic components. This amounts to fixing [N II] widths to that of the corresponding Hα component. We also fix the [N II]λ6548, λ6584 Å ratio to 0.338 and the [O III]λ4959, λ5007 Å ratio to 0.335. In addition to the line-free stacked spectrum, we also compute a broad-line-only spectrum (continuum+broad line emission).

We calculate synthetic JWST/NIRCam photometry using the filter response profiles provided by the JWST User Documentation. We convolve each filter with both the stacked spectrum and line-subtracted stack for a range in redshift at $z = 1$–9 in steps of $\Delta z = 0.05$. We then take the ratio of filter flux between the line stack and line-subtracted (or broad-line-only) stack to infer the increase in flux attributed to emission lines as a function of redshift.

3.2.1. *Composite DSFG Spectrum from Casey et al. (2017)*

As both a check against our stack and a test for systems at lower $L_{\rm IR}$ than the hot DOGs, we repeat our synthetic photometry calculations for the composite DSFG spectrum from Casey et al. (2017). We scale their continuum-subtracted Hα flux in their stack to the equivalent of 100 $M_\odot$ yr^{-1} in star formation rate using the $F_{\rm H\alpha}$ calibration of Murphy et al. (2011). As the change in flux density due to nebular emission is a function of the relative strength between lines and continuum, we add the scaled DSFG spectrum to the empirically derived rest-frame 0.1–1 μm mean DSFG spectral energy distribution (SED) from Casey et al. (2014). For the line-free calculation we simply mask Hα, [O III], and Hβ from the stack prior to performing synthetic photometry, equivalent to computing fluxes for the continuum DSFG SED without adding the lines.

4. Results and Discussion

The results of our synthetic photometry are shown in Figure 2, which gives the flux ratio between our fiducial and line-subtracted stacked spectrum and the DSFG stack from Casey et al. (2017). For the former, we also show the increase in flux separated between the narrow and broad velocity components. On average, strong narrow+broad rest-frame optical lines increase NIRCam fluxes by factors of ∼1.2–1.7, with corresponding change in apparent magnitude by 0.2–0.6. The maximal increase in flux occurs when any particular wide-band filter is centered on the strong [N II]+Hα complex. The [O III] and Hβ lines collectively increase the wide-band flux maximally by ∼20%. Their broad components and those of [N II] and Hα increase synthetic flux densities by 1.2× on average, accounting for ∼66% of the boost for [O III]+Hβ and 25% for [N II]+Hα. Medium-band filters are more affected by the presence of strong emission lines and can be dominated by factors of ∼2.5 (1 mag) by emission lines when redshifted to the line's rest wavelength. For example, the F410M flux is increased by a factor of 2.5 at $z = 5.5$. In fact, $z = 5$ is a special regime where a boost in flux density is seen for all NIRCam long wavelength (LW) filters. While we do not show the increase in flux attributed to the relatively weaker [O II] line in Figure 2, this effect is maximally ∼10% if we mask the line following the methods outlined for [O III]+Hβ and [N II]+Hα.

The strong lines in the DSFG stack from Casey et al. (2017) that we have scaled to an Hα star-formation rate of 100 $M_\odot$ yr^{-1} (see Section 3.2.1) increase broadband fluxes by up to a factor of ∼1.5. Such boosting occurs over similar ranges in redshift and to the same degree as found for the narrow line components in the hot DOG stack. This demonstrates that significant line contamination can be present in the NIRCam photometry for IR-luminous galaxies more normal than the relatively extreme hot DOGs.

Given extreme levels of attenuation in massive dust-obscured galaxies at high redshift, their rest-frame optical spectra contain a combination of significantly reddened continuum with $\lesssim$5% of the total unobscured light escaping from the least obscured regions (Chapman et al. 2005; Howell et al. 2010). With a combination of strong lines emergent from less obscured regions on top of the very red continuum, ∼0.1–1 μm photometry of dusty galaxies can mimic that of ultrahigh-redshift LBG candidates in large surveys (Fujimoto et al. 2022; Zavala et al. 2023). In both scenarios, the inferred galaxy properties do not need to be particularly extreme relative to galaxy populations at their respective epochs. For example, Zavala et al. (2023) find that a $z > 10$ LGB candidate lies on the main sequence at $z = 5.1$ with modest mass and with an implied absolute magnitude similar to other optically faint $4 < z < 6$ DSFGs (Wang et al. 2019). Similarly, current $z > 10$ catalogs report absolute magnitudes around −19 mag close to L_* (Donnan et al. 2023; Finkelstein et al. 2022a). Thus, it can be difficult to rule out redshift solutions on the basis of inferred properties alone as these tend to be within reasonable limits of rather poorly characterized populations to begin with.

To quantify the parameter space where confusion between these populations is significant, we fit an LBG template to the synthetic NIRCam flux densities derived from our stacked spectrum. We first normalize the stack to a continuum flux on the order of ∼10 nJy over $\lambda_{\rm obs} = 2$–3.5 μm and assume it to be undetected in F115W and F150W. This represents a plausible scenario given the relative filter depths of JWST Cycle 1 extragalactic deep fields (Bagley et al. 2022; Casey et al. 2022; Finkelstein et al. 2022b) and is similar to CEERS-93316 (Donnan et al. 2023) and CEERS2_2159 (Finkelstein et al. 2022b), a $z = 16.4$ LBG candidate selected from CEERS (Bagley et al. 2022; Finkelstein et al. 2022a). CEERS-93316 has a tentative 2.6σ SCUBA-2 detection (Zavala et al. 2023) and environmental evidence (Naidu et al. 2022b), both indicating a possible lower-redshift solution at $z \sim 4.8$.

Figure 3 (left) shows the 2D posterior distribution in redshift and UV slope β for LBG template fits to our stacked spectrum's synthetic NIRCam flux densities. We repeat the fitting analysis 1000 times after perturbing the input spectrum by the spectral uncertainty and in three redshift ranges for the stack: $3.5 < z < 4$, $4 < z < 5$, and $5 < z < 6$. The cumulative EW of Hα+[N II] is greater than EW([O III]+Hβ) by a factor of ∼3, which precludes LBG fits with $\beta < -1.5$ when the stack is redshifted to $z > 4$. This is because both features fall within a broadband filter, and so the strong lines do not mask the red continuum in the stack. At $3.5 < z < 4$, F277W picks up the strong Hα+[N II] emission while [O III]+Hβ is missed by F200W. This produces degenerate photometry with $z \sim 16$ LBGs. At $z \sim 16$, the Lyman break falls halfway between F200W, mimicking the red slope of the dusty galaxy stack, while the very blue continuum mimics the F277W flux density of the line-contaminated stack. In summary, the hot DOG stack can reproduce very blue UV slopes $\beta \sim -2.5$ for $z_{\rm stack} \sim 3.5$–4 but not for $z_{\rm stack} > 4$. This supports the purity of the very blue

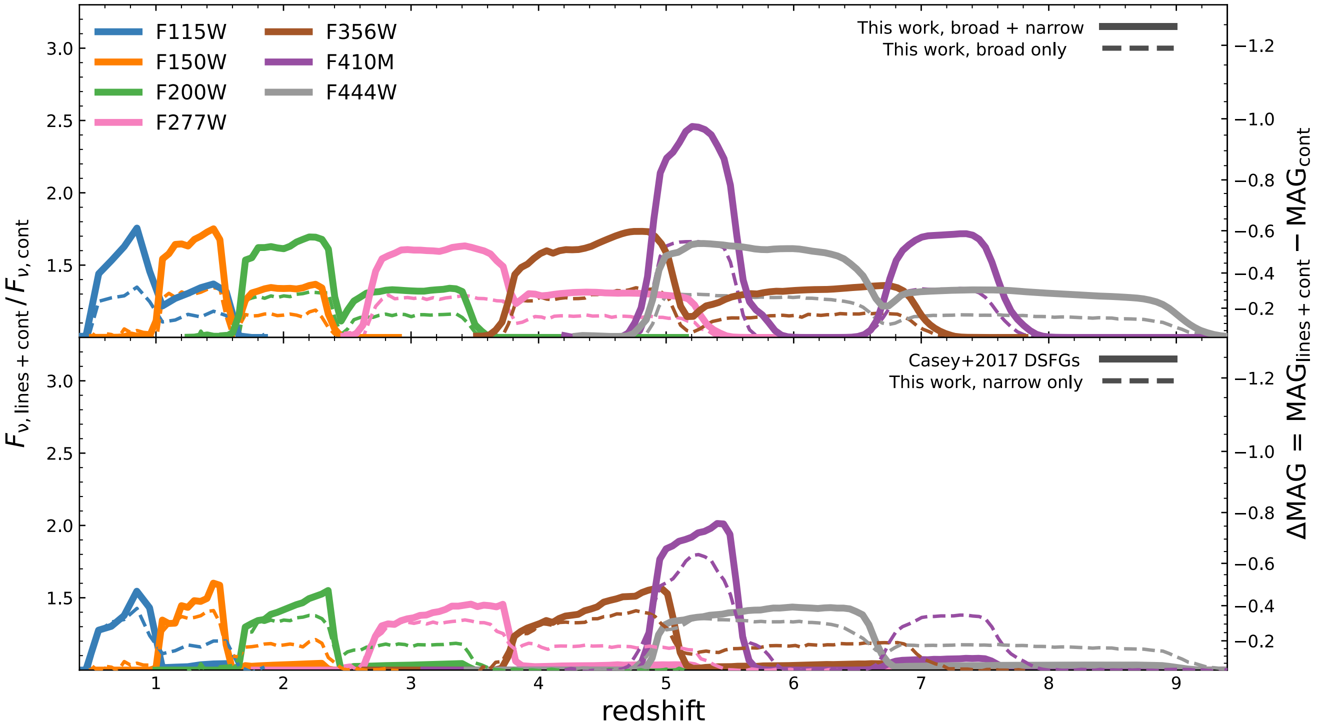

Figure 2. Increase in JWST/NIRCam flux by strong rest-frame emission lines for the average SED of hot, dust-obscured and IR-luminous galaxies at $\lambda_{\mathrm{rest}} = 0.34$–$0.8\ \mu$m as a function of redshift. (Top) Solid lines account for broad and narrow velocity components, whereas dashed lines include only the broad component. Maximally, strong nebular emission lines can boost the broadband flux at ∼25%–80% ($|\Delta$ mag$| = 0.2$–0.6) from $z \sim 1$–8. Medium-band filters such as F410M can be boosted by up to a factor of 2.5 when they overlap with strong emission lines at $z \sim 5.5$. The increase in flux attributed to velocity-broadened features is ∼25% on average ($|\Delta$ mag$| = 0.2$). The double-peak effect for a given filter arises from the Hα complex first passing through, followed by [O III]+Hβ. (Bottom) Increase in flux attributed to strong line emission for the average DSFG spectrum of Casey et al. (2017) scaled to an Hα star formation rate of 100 $M_{\odot}$ yr^{-1} (solid) and the narrow velocity component in our stack (dashed). The increase in flux by strong Hα+[N II] in the DSFG stack is consistent with the narrow line component for these lines in the hot DOG stack. At $z \sim 5$, all of the NIRCam LW filters are boosted by ∼20%–100% for hot DOGs and DSFGs.

NIRCam samples of Cullen et al. (2023) and Topping et al. (2022), which predominantly have $\beta < -1.5$ and $7 < z < 14$.

As further demonstration of the confusion between our stack at $z \sim 3.5$–4 and $z \sim 16$ LBGs, we show in Figure 3 (right) the LBG fit to our stack's JWST/NIRCam flux densities. At $z_{\mathrm{stack}} = 3.75$ we find a best-fit LBG solution with $z = 16$ and UV spectral index $\beta = -2.4$. Lower-redshift ($z < 4$) solutions with red continuum slopes and flux densities dominated by strong emission lines should be considered when fitting the very blue ($\beta < -2$) spectral energy distributions (SEDs) of $z \sim 16$ candidates. The lower-redshift solutions could be ruled out with medium-band filters, longer wavelength sampling using NIRCam's redder filters, MIRI observations, and/or far-IR/submillimeter follow-up to detect cold dust continuum and fine-structure lines (Fujimoto et al. 2022). Measurements that strongly rule out UV spectral indices $\beta < -1.5$ and only allow lower $-z$ solutions at $z > 4$ should be particularly constraining against massive, IR-luminous interlopers with strong optical lines, provided they sample the SED with more than three filters.

At the core of this problem is the fact that $z \gtrsim 10$ LBGs and $4 < z < 6$ DSFGs appear to have comparable apparent magnitudes in JWST near-IR broadband photometry (Donnan et al. 2023; Naidu et al. 2022a; Zavala et al. 2023). For example, DSFGs can be as faint as 26.5 mag at $\lambda_{\mathrm{rest}} < 1\ \mu$m (Casey et al. 2017) and possibly fainter among the heavily obscured $z > 5$ population (Manning et al. 2022). WISE-selected hot DOGs at $z \sim 4$, on the other hand, tend to be much brighter, around 22–23 mag at 1.2 μm (Tsai et al. 2015), which makes this particular population less likely to contaminate ultrahigh-redshift candidate catalogs that include magnitude cuts. Future rest-frame optical spectroscopy of DSFGs is needed to test if strong lines in $z > 4$ DSFGs contribute to broadband fluxes as much as the hot DOGs do.

Based on the first analysis of JWST deep field observations at 5σ point-source depths at $\sim 28 - 29$ mag, the projected sky density of candidates at $z > 10$ is approximately 350 ± 120 deg^{-2} (Donnan et al. 2023; Finkelstein et al. 2022a; Harikane et al. 2023; Naidu et al. 2022a). Although preliminary, these source counts represent the population that could potentially be contaminated by low $-z$ dusty interlopers. In contrast, the sky density of luminous IR galaxies with $\log L_{\mathrm{IR}}/L_{\odot} > 12(12.5)$ and $z \sim 3$–4 is 400 deg^{-2} (100 deg^{-2}) (Casey et al. 2018; Zavala et al. 2021). If we assume the samples of Finnerty et al. (2020) and Casey et al. (2017) include a range of physically possible rest-frame optical properties for IR-bright galaxies ($\log L_{\mathrm{IR}}/L_{\odot} > 11$), then their similar number counts to ultrahigh-redshift LBG candidates may be reason to be concerned about contamination. The fainter dusty galaxy population with $\log L_{\mathrm{IR}}/L_{\odot} < 11$ are much more numerous based on the general shape of 1 mm number counts (Fujimoto et al. 2016; González-López et al. 2020) and may also be an important source of contamination as galaxies fainter in the IR are less likely to be significantly obscured in the rest-frame optical. Further spectroscopic follow-up is required to assess the purity of ultrahigh-redshift catalogs. In the meantime, F150W dropouts ($z > 10$) with $\beta \sim -2$ and no/poor SED constraint above NIRCam/F356W should be checked against possible intermediate-redshift dusty galaxy solutions.

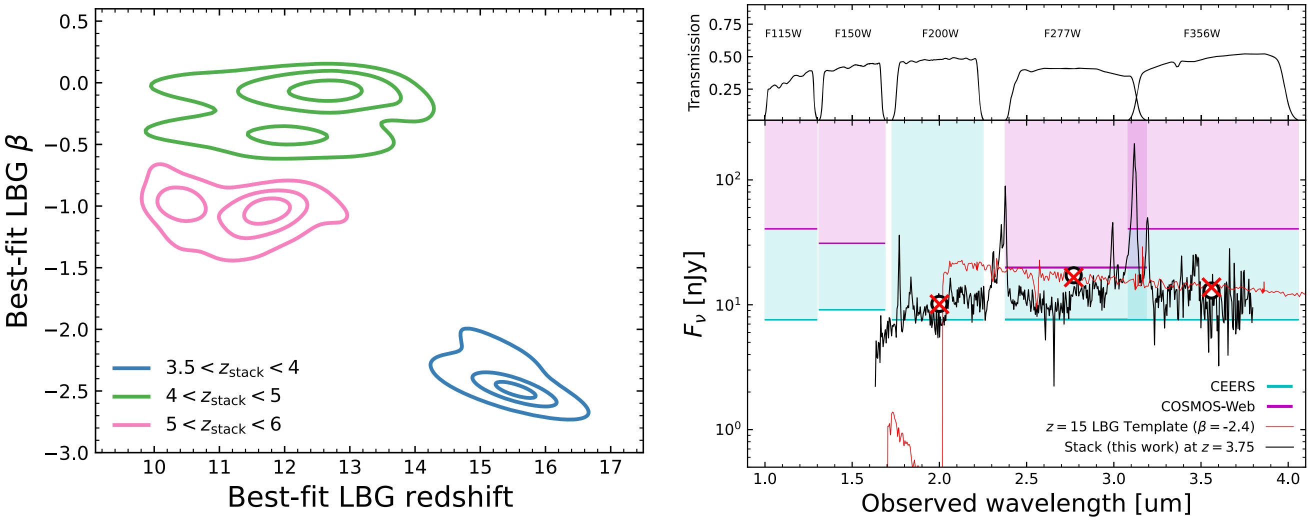

Figure 3. (Left) Allowed LBG redshift and UV slope β solutions when fitting the synthetic NIRCam flux densities of our stack redshifted to $z = 3.5$–4 (blue), $z = 4$–5 (green), and $z = 5$–6 (pink). Posterior contours are drawn at the 16th, 50th, and 84th percentiles. At $z > 4$, [N II]+Hα increase the NIRCam flux density more so than [O III]+Hβ, which does not allow the strong lines to mask the red continuum and therefore precludes LBG solutions with $\beta < -1.5$. At $3.5 < z < 4$, [N II]+Hα falls within F277W, while [O III]+Hβ is missed by F200W, allowing degenerate solutions with blue $\beta \leqslant -2$ LBGs at $z \sim 14$–17. (Right) Illustration of the degeneracy at $z \sim 15$ candidates and $z < 4$ dusty galaxies with strong rest-frame optical/emission lines. In this example, we redshift our stacked spectrum to $z = 3.75$ where strong line emission boosts the F277W filter flux by 60%. We then compute F200W, F277W, and F356W JWST/NIRCam photometry (black circles), assuming nondetections in F115W and F150W. We fit the synthetic photometry from the stack (black circles) with an LBG template (red line), deriving a photometric redshift of $z_{\rm phot} = 15$ and UV spectral index $\beta = -2.4$. Strong emission lines mask the red slope of the dusty template between F277W and F356W, and the SED is further confused with the Lyman break falling halfway between F200W. Such scenarios are possible given the relative filter depths of JWST Cycle 1 NIRCam extragalactic surveys in CEERS (blue; Bagley et al. 2022; Finkelstein et al. 2022b) and COSMOS-Web (pink; Casey et al. 2022) for example.

One promising check is to use JWST/MIRI to detect the emission at longer wavelengths and distinguish between the redder SEDs of dusty galaxies from bluer $z \gtrsim 10$ LBGs. As an example, we calculate MIRI photometry using both the low- and high-redshift SED solutions for CEERS-DSFG1 and CEERS-93316 from Zavala et al. (2023). The relative difference in flux between the low-z dusty galaxy SED and the high $-z$ LBG solution for both galaxies is 0.2–0.25 mag in MIRI/F560W and 0.35 in MIRI/F770W. The expected apparent magnitudes in all MIRI bands are 23–24 mag, well above the 5σ sensitivity thresholds of Cycle 1 JWST deep fields with MIRI coverage such as COSMOS-Web and CEERs (Casey et al. 2022; Finkelstein et al. 2022b).

5. Summary and Conclusion

In this Letter, we test the response of JWST NIRCam filters over broad rest-frame optical emission lines in the average spectrum of hot, dust-obscured galaxies at $z \sim 1$–4 and dusty, star-forming galaxies. As an empirical approach, we stack a sample of 21 IR-luminous galaxies with rest-frame optical spectra from Keck II/NIRES, for which we then compute synthetic photometry for $z = 1$–9. Our main results are as follows:

1. We measure broad rest-frame optical emission lines in the stack of $z \sim 1$–4 hot, dust-obscured galaxies. In particular, we measure [O III] and Hα EWs between 100 and 500 Å, which are high relative to normal star-forming galaxies at high redshift.
2. After masking out strong emission features from the spectrum, we measure synthetic NIRCam photometry with and without the lines. Narrow and broad components for [O III] and Hβ increase the measured filter flux by 30% and Hα + [N II] by 60% on average. Narrowband filters such as F410M can have their flux increased by a factor of 2–3 (0.7–1.2 mag).
3. Rest-frame optical photometry of dusty galaxies with strong nebular lines at $z \sim 3.5$–4 could be important contaminants in F150W dropout ($z > 10$ candidate) catalogs as the strong lines can help mask red UV spectral indices. However, UV spectral indices $\beta < -1.5$ are difficult for our stacked spectrum to reproduce for interloper redshifts $z > 4$.

Distinguishing between different galaxy populations with JWST imaging is a key first step toward testing various aspects of galaxy formation. While this work has focused just on JWST's NIRCam filters, the inclusion of deep MIRI photometry extending to longer wavelengths will add significant constraint on various redshift solutions to photometric fitting codes. In the absence of high SNR coverage in redder filters, far-IR/submillimeter follow-up can help identify dusty galaxies. On the near horizon, ToLTEC on the Large Millimeter Telescope (LMT) Alfonso Serrano will map multiple extragalactic fields (COSMOS, UDS, GOODS-S) down to the LIRG limit at 1.1, 1.4, and 2 mm as part of "The TolTEC Ultra-Deep Galaxy Survey," a public legacy program. These public data sets are well suited to quickly identify submillimter bright DSFG counterparts to JWST sources.

L.F. is a member of Student Researchers United (SRU-UAW). The data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. We wish to acknowledge the critical importance of the current and recent Maunakea Observatories crew, technicians, telescope operators, computer support, and office staff employees, especially during the challenging times presented by the COVID-19 pandemic. Their expertise, ingenuity, and dedication are indispensable to the continued successful operation of these observatories. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

ORCID iDs

Jed McKinney https://orcid.org/0000-0002-6149-8178
Luke Finnerty https://orcid.org/0000-0002-1392-0768
Caitlin M. Casey https://orcid.org/0000-0002-0930-6466
Maximilien Franco https://orcid.org/0000-0002-3560-8599
Arianna S. Long https://orcid.org/0000-0002-7530-8857
Seiji Fujimoto https://orcid.org/0000-0001-7201-5066
Jorge A. Zavala https://orcid.org/0000-0002-7051-1100
Olivia Cooper https://orcid.org/0000-0003-3881-1397
Hollis Akins https://orcid.org/0000-0003-3596-8794
Alexandra Pope https://orcid.org/0000-0001-8592-2706
Lee Armus https://orcid.org/0000-0003-3498-2973
B. T. Soifer https://orcid.org/0000-0002-8112-1132
Kirsten Larson https://orcid.org/0000-0003-3917-6460
Michael Cushing https://orcid.org/0000-0001-7780-3352

References

Assef, R. J., Eisenhardt, P. R. M., Stern, D., et al. 2015, ApJ, 804, 27
Bagley, M. B., Finkelstein, S. L., Koekemoer, A. M., et al. 2022, arXiv:2211.02495
Bakx, T. J. L. C., Zavala, J. A., Mitsuhashi, I., et al. 2023, MNRAS, 519, 5076
Bouwens, R., Illingworth, G., Oesch, P., et al. 2022, arXiv:2212.06683
Boylan-Kolchin, M. 2022, arXiv:2208.01611
Bradley, L. D., Coe, D., Brammer, G., et al. 2022, arXiv:2210.01777
Carnall, A. C., Begley, R., McLeod, D. J., et al. 2023, MNRAS, 518, L45
Casey, C. M., Chen, C.-C., Cowie, L. L., et al. 2013, MNRAS, 436, 1919
Casey, C. M., Cooray, A., Killi, M., et al. 2017, ApJ, 840, 101
Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2022, arXiv:2211.07865
Casey, C. M., Scoville, N. Z., Sanders, D. B., et al. 2014, ApJ, 796, 95
Casey, C. M., Zavala, J. A., Spilker, J., et al. 2018, ApJ, 862, 77
Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, ApJ, 622, 772
Cullen, F., McLure, R. J., McLeod, D. J., et al. 2023, MNRAS, 520, 14
Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, PASP, 116, 362
Cutri, R. M., Wright, E. L., Conrow, T., et al. 2012, Explanatory Supplement to the WISE All-Sky Data Release Products
Donnan, C. T., McLeod, D. J., Dunlop, J. S., et al. 2023, MNRAS, 518, 6011
Drew, P. M., & Casey, C. M. 2022, ApJ, 930, 142
Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, MNRAS, 494, 3828
Dunlop, J. S., Cirasuolo, M., & McLure, R. J. 2007, MNRAS, 376, 1054
Eisenhardt, P. R. M., Wu, J., Tsai, C.-W., et al. 2012, ApJ, 755, 173
Finkelstein, S. L., Bagley, M. B., Arrabal Haro, P., et al. 2022a, ApJL, 940, L55
Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2022b, arXiv:2211.05792
Finnerty, L., Larson, K., Soifer, B. T., et al. 2020, ApJ, 905, 16
Fujimoto, S., Finkelstein, S. L., Burgarella, D., et al. 2022, arXiv:2211.03896
Fujimoto, S., Ouchi, M., Ono, Y., et al. 2016, ApJS, 222, 1
González-López, J., Novak, M., Decarli, R., et al. 2020, ApJ, 897, 91
Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, A&A, 643, A8
Harikane, Y., Ouchi, M., Oguri, M., et al. 2023, ApJS, 265, 5
Howell, J. H., Armus, L., Mazzarella, J. M., et al. 2010, ApJ, 715, 572
Jun, H. D., Assef, R. J., Bauer, F. E., et al. 2020, ApJ, 888, 110
Kaasinen, M., van Marrewijk, J., Popping, G., et al. 2023, A&A, 671, A29
Kirkpatrick, A., Pope, A., Alexander, D. M., et al. 2012, ApJ, 759, 139
Kirkpatrick, A., Pope, A., Sajina, A., et al. 2015, ApJ, 814, 9
Koprowski, M. P., Dunlop, J. S., Michałowski, M. J., et al. 2017, MNRAS, 471, 4155
Koss, M., Trakhtenbrot, B., Ricci, C., et al. 2017, ApJ, 850, 74
Labbe, I., van Dokkum, P., Nelson, E., et al. 2022, arXiv:2207.12446
Long, A. S., Casey, C. M., Lagos, C. d. P., et al. 2022, arXiv:2211.02072
Magnelli, B., Saintonge, A., Lutz, D., et al. 2012, A&A, 548, A22
Maio, U., & Viel, M. 2022, arXiv:2211.03620
Manning, S. M., Casey, C. M., Zavala, J. A., et al. 2022, ApJ, 925, 23
Michałowski, M. J., Dunlop, J. S., Koprowski, M. P., et al. 2017, MNRAS, 469, 492
Murphy, E. J., Chary, R. R., Dickinson, M., et al. 2011, ApJ, 732, 126
Naidu, R. P., Oesch, P. A., Setton, D. J., et al. 2022b, arXiv:2208.02794
Naidu, R. P., Oesch, P. A., van Dokkum, P., et al. 2022a, ApJL, 940, L14
Roberts-Borsani, G., Treu, T., Chen, W., et al. 2022, arXiv:2210.15639
Rowan-Robinson, M., Wang, L., Farrah, D., et al. 2018, A&A, 619, A169
Schaerer, D., & de Barros, S. 2009, A&A, 502, 423
Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, ApJ, 763, 129
Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, MNRAS, 438, 1267
Topping, M. W., Stark, D. P., Endsley, R., et al. 2022, ApJ, 941, 153
Tsai, C.-W., Eisenhardt, P. R. M., Wu, J., et al. 2015, ApJ, 805, 90
Wang, T., Schreiber, C., Elbaz, D., et al. 2019, Natur, 572, 211
Wilkins, S. M., Coulton, W., Caruana, J., et al. 2013, MNRAS, 435, 2885
Wilkins, S. M., Lovell, C. C., Fairhurst, C., et al. 2020, MNRAS, 493, 6079
Wilkins, S. M., Vijayan, A. P., Lovell, C. C., et al. 2022, MNRAS, 517, 3227
Wilson, J. C., Henderson, C. P., Herter, T. L., et al. 2004, Proc. SPIE, 5492, 1295
Wu, J., Jun, H. D., Assef, R. J., et al. 2018, ApJ, 852, 96
Wu, J., Tsai, C.-W., Sayers, J., et al. 2012, ApJ, 756, 96
Yoon, I., Carilli, C. L., Fujimoto, S., et al. 2022, arXiv:2210.08413
Yu-Yang Hsiao, T., Coe, D., Abdurrouf, et al. 2022, arXiv:2210.14123
Zackrisson, E., Bergvall, N., & Leitet, E. 2008, ApJL, 676, L9
Zavala, J. A., Buat, V., Casey, C. M., et al. 2023, ApJL, 943, L9
Zavala, J. A., Casey, C. M., Manning, S. M., et al. 2021, ApJ, 909, 165