

To see or not to see a $z \sim 13$ galaxy, that is the question

Targeting the [C II] 158 μm emission line of HD1 with ALMA

M. Kaasinen¹, J. van Marrewijk¹, G. Popping¹, M. Ginolfi^{1,2}, L. Di Mascolo^{3,4,5}, T. Mroczkowski¹, A. Concas¹,
C. Di Cesare^{6,7,8}, M. Killi^{9,10}, and I. Langan^{1,11}

¹ European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany
e-mail: melanie.kaasinen@eso.org

² Dipartimento di Fisica e Astronomia, Università di Firenze, Via G. Sansone 1, 50019 Sesto Fiorentino (Firenze), Italy

³ Department of Physics, University of Trieste, Via Tiepolo 11, 34131 Trieste, Italy

⁴ INAF – Osservatorio Astronomico di Trieste, Via Tiepolo 11, 34131 Trieste, Italy

⁵ IFPU – Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy

⁶ Dipartimento di Fisica, Sapienza, Università di Roma, Piazzale Aldo Moro 5, 00185 Roma, Italy

⁷ INFN, Sezione di Roma I, Piazzale Aldo Moro 2, 00185 Roma, Italy

⁸ INAF/Osservatorio Astronomico di Roma, Via di Frascati 33, 00078 Monte Porzio Catone, Italy

⁹ Cosmic Dawn Center (DAWN), Jagtvej 128, 2200 Copenhagen N, Denmark

¹⁰ Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, 2100 Copenhagen Ø, Denmark

¹¹ Univ. Lyon, Univ. Lyon1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon (CRAL) UMR5574,
69230 Saint-Genis-Laval, France

Received 29 September 2022 / Accepted 16 January 2023

ABSTRACT

Context. Determining when the first galaxies formed remains an outstanding goal of modern observational astronomy. Theory and current stellar population models imply that the first galaxies formed at least at $z = 14$ – 15 . But to date, only one galaxy at $z > 13$ (GS-z13-0) has been spectroscopically confirmed.

Aims. The galaxy ‘HD1’ was recently proposed to be a $z \sim 13.27$ galaxy based on its potential Lyman break and tentative [O III] 88 μm detection with ALMA. We hereby aim to test this scenario with new ALMA Band 4 observations of what would be the [C II] 158 μm emission if HD1 is at $z \sim 13.27$.

Methods. We carefully analyse the new ALMA Band 4 observations and re-analyse the existing ALMA Band 6 data on the source to determine the proposed redshift.

Results. We find a tentative 4σ feature in the Band 4 data that is spatially offset by $1''.7$ and spectrally offset by 190 km s^{-1} from the previously reported 3.8σ [O III] 88 μm feature. Through various statistical tests, we demonstrate that these tentative features are fully consistent with both being random noise features.

Conclusions. We conclude that we are more likely to be recovering noise features than both [O III] 88 μm and [C II] 158 μm emission from a source at $z \sim 13.27$. Although we find no credible evidence of a $z \sim 13.27$ galaxy, we cannot entirely rule out this scenario. Non-detections are also possible for a $z \sim 13$ source with a low interstellar gas-phase metallicity or ionisation parameter and/or high gas density. Moreover, the new continuum and line upper limits provide no strong evidence for or against a lower-redshift scenario. Determining where and exactly what type of galaxy HD1 is, will now likely require JWST/NIRSpec spectroscopy.

Key words. galaxies: evolution – galaxies: high-redshift – galaxies: individual: HD1 – galaxies: ISM – techniques: interferometric

1. Introduction

Determining when the first galaxies formed remains an outstanding goal of modern observational astronomy, with important implications for theories of structure formation, baryonic physics, and cosmology. Simulations (e.g. Ceverino et al. 2018; Graziani et al. 2020; Pallottini et al. 2022) and observations (e.g. Hashimoto et al. 2018; Roberts-Borsani et al. 2020) imply that the first galaxies were already present at $z \sim 14$ – 15 , but confirming (or rejecting) their existence remains a challenge. At the time of first submitting this manuscript, the highest-redshift galaxy to have been confirmed was GN-z11 at $z \sim 11$ (Oesch et al. 2014, 2016; Jiang et al. 2021). At the time of publication, four $z > 10$ galaxies have spectroscopically confirmed with the James Webb Space Telescope (JWST), the highest-

redshift galaxy being GS-z13-0 at $z = 13.2$ (Curtis-Lake et al. 2022; Robertson et al. 2022).

To identify potential first galaxies, two main photometric methods have been employed. The first method has been to search for galaxies with a ‘break’ (or ‘drop’) in emission bluewards of Lyman alpha ($\text{Ly}\alpha$), resulting from the absorption of UV photons by the neutral hydrogen in the early Universe (e.g. Oesch et al. 2016). For $z > 10$ galaxies, this break lies in the near-infrared bands. Using the Hubble Space Telescope (HST), many such Lyman break candidates have been discovered (e.g. Oesch et al. 2014; Calvi et al. 2016; Stefanon et al. 2019) and, since the release of the first JWST/NIRCam data, many more, even higher redshift $z > 10$ candidates have been identified in this way (e.g. Castellano et al. 2022; Harikane et al. 2023; Naidu et al. 2022a,b; Atek et al. 2023; Donnan et al. 2023;

Yan et al. 2023). For the second main photometric pre-selection method, candidates have been identified based on their red *Spitzer* IRAC [3.6]–[4.5] μm colour (referred to as the ‘IRAC excess’), proposed to arise from intense [O III] $\lambda\lambda$ 4959,5007 \AA and H β emission within the 4.5 μm band-pass (Roberts-Borsani et al. 2016). Both of these photometric techniques are highly efficient in amassing high-redshift candidates, but additional spectroscopic observations are still required to confirm the proposed redshifts.

Spectroscopically confirming galaxies beyond $z \sim 7$ remains a challenge. The Ly α emission line, typically used at lower redshifts, is significantly attenuated for galaxies in the Epoch of Reionisation (EoR) by the high columns of neutral hydrogen surrounding the galaxies. Rest-frame UV lines are also challenging to observe, as shown in the follow-up of the rest-frame UV emission from GN-z11 with Keck (Jiang et al. 2021). Prior to JWST, one of the most efficient methods of spectroscopically confirming high-redshift candidates has been via their [O III] 88 μm and [C II] 158 μm line emission, which both remain well within reach of the Atacama Large sub-/Millimetre Array (ALMA) even for galaxies at $z > 10$. Using ALMA, a few $z \sim 8$ –9 galaxies with *Spitzer* and Hubble photometry have already been spectroscopically confirmed (Laporte et al. 2017; Hashimoto et al. 2018; Tamura et al. 2019; Bakx et al. 2020). For some of these confirmed $z = 8$ –9 galaxies, the best fit models to their rest-frame-optical to near-infrared (NIR) photometry imply that a significant stellar mass must already have been present at $z = 14$ –15 (e.g. Hashimoto et al. 2018; Roberts-Borsani et al. 2020). The existence of such large stellar masses at these redshifts is also corroborated by simulations of first galaxies (e.g. Ceverino et al. 2018; Pallottini et al. 2022). While multiple ALMA DDT programmes have now been conducted with the aim of confirming/rejecting $z > 12$ galaxy candidates identified using HST or JWST photometry, there are no conclusive results yet (Bakx et al. 2023; Popping 2023; Yoon et al. 2022). Thus, the interstellar medium (ISM) of these youngest of galaxies is yet to be observed with ALMA.

Early last year, Harikane et al. (2022) presented tantalising evidence for a potential $z = 13.27$ galaxy. The authors first created a catalogue of potential $z \sim 13$ candidates, by searching the photometric data sets of the COSMOS and XDS fields, performing deblending corrections and searching for a Lyman break in the *H* band. After removing any foreground interlopers, Harikane et al. (2022) identified two potential $z \sim 13$ Lyman Break Galaxies (LBGs). They followed up the first candidate, referred to as HD1, with ALMA Band 6 observations using four frequency tunings to cover the redshift range of $12.6 < z < 14.3$ for the targeted [O III] 88 μm emission line. From these observations, a tentative 3.8σ line detection was reported at ≈ 237.8 GHz, which, if interpreted as [O III] 88 μm emission, would correspond to a redshift of $z = 13.27$. If confirmed, HD1 would be the highest redshift galaxy observed as of yet, existing at only ~ 320 Myr after the Big Bang. The existence of such a galaxy would imply little to no evolution in the UV luminosity function of galaxies from $z = 8$ to $z = 13$ (see Bowler et al. 2020; Harikane et al. 2022), in contrast to predictions by galaxy formation models for the evolution of the UV luminosity function of galaxies at $z > 10$ (e.g. Yung et al. 2019; Behroozi et al. 2020). As potentially one of the earliest known galaxies to form, HD1 thus presents an important benchmark for galaxy formation models and cosmological simulations.

The tentative line detection and photometric redshift constraints for HD1 leave a large degree of ambiguity. As demonstrated in Harikane et al. (2022), the spectral energy distribution

of HD1 can also be described by a $z \sim 4$ galaxy with a very low star formation rate. In that case, the 3.8σ line emission may instead correspond to the CO(9–8) or CO(10–9) line emission. Alternatively, the 3.8σ feature may simply be the result of random noise features. To distinguish between these possible scenarios, we conducted ALMA Band 4 observations, covering the frequency that would correspond to the [C II] 158 μm emission of the proposed $z = 13.27$ galaxy. Not only would a [C II] 158 μm detection confirm the redshift of HD1, it would also provide useful constraints on its ISM properties (e.g. ionisation state and metallicity).

In this paper, we describe these new ALMA observations and their implications. We present the ALMA Band 4 observations and data reduction as well as a consistent re-analysis of the Band 6 data obtained by Harikane et al. (2022) in Sect. 2. We present the data analysis in Sect. 3 and summarise the implications of our work in Sect. 5. Throughout this paper, we use a flat Λ -CDM concordance model ($H_0 = 70.0 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.30$).

2. ALMA observations and data reduction

2.1. The new Band 4 observations

To confirm or reject the proposed redshift of HD1, we obtained ALMA Band 4 observations through the ALMA Director’s Discretionary Time (DDT) programme (2021.A.00008.S, PI: G. Popping), targeting the [C II] 158 μm emission that would correspond to the tentative [O III] 88 μm detection. We used a setup with four spectral windows (SPWs) centred on 131.287 GHz, 133.164 GHz, 143.206 GHz, and 145.087 GHz. All SPWs except the one centred on 133.164 GHz were observed in Time Division Mode (TDM) with a bandwidth of 1.875 GHz (to obtain a continuum measurement), whereas the other one was set up using 240 channels with a 7.812 kHz spectral resolution and 1.875 GHz bandwidth (to target the [C II] 158 μm line). The observations were centred on RA $10^{\text{h}}01^{\text{m}}51.31^{\text{s}}$, Dec $+02^{\text{d}}32^{\text{m}}50.0^{\text{s}}$ and were carried out on March 3 and 6, 2022. The total on-source time was 4.9 h, with an average PWV of 4.746 mm.

All data processing and calibration were performed with CASA, version 6.2 (McMullin et al. 2007). We used the calibrated measurement sets generated by the observatory for the imaging of the observations. Imaging was performed with the `tclean` task, adopting natural weighting. We create both a continuum map (using the SPWs observed in TDM) and spectral cube (using the SPW centred at 133.16 GHz). No continuum emission is detected. The continuum root-mean-square (rms) noise level is $5.2 \mu\text{Jy beam}^{-1}$ (yielding a 3σ upper limit of $15.6 \mu\text{Jy beam}^{-1}$). For the spectral cube at 133.16 GHz, the rms is $127.3 \mu\text{Jy beam}^{-1}$ per 78 MHz. The continuum and spectral cube image have a beam of $2''.78 \times 2''.16$ and $2''.91 \times 2''.27$, respectively. The final data cube used in the analysis reported here spans a frequency range of 132.22–134.08 GHz, corresponding to a redshift range of $z = 13.17$ –13.37. The CLEANed Band 4 continuum map showed several off-centre point sources. For the spectral cube, these were removed through `CASA-imcontsub`¹ (because `CASA-uvcontsub`² is less optimised for removing off-centre co-detections).

¹ <https://casa.nrao.edu/docs/taskref/imcontsub-task.html>

² <https://casa.nrao.edu/docs/taskref/uvcontsub-task.html>

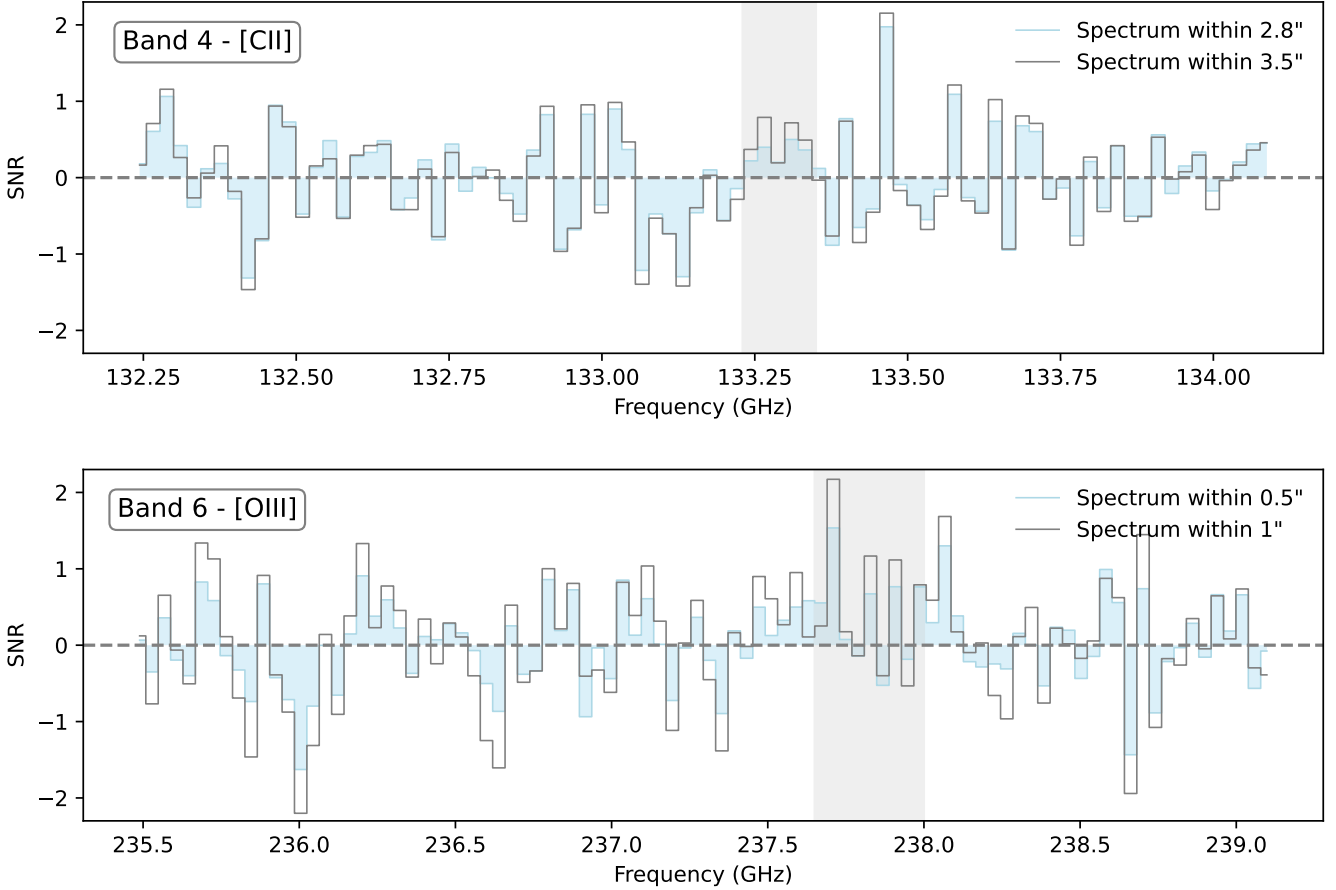


Fig. 1. Integrated aperture spectra for HD1 around the expected frequency of the [C II] 158 μm (top panel) and [O III] 88 μm (bottom panel) emission lines. To easily compare the spectra extracted within apertures of different size, spectra are presented in terms of signal-to-noise ratio rather than actual flux density. For the [C II] 158 μm emission we adopt apertures of 2''.8 and 3''.5, corresponding to a similar size to the beam and 1.25 times the beam of the ALMA band 4 data, respectively. In the bottom panel we adopt an aperture of 0''.5 and 1''.0, corresponding to roughly the beam size and the aperture adopted in [Harikane et al. \(2022\)](#) for the same data, respectively. The grey shaded area marks the tentative [C II] 158 μm feature (top panel), and the location of the [O III] 88 μm line feature presented in [Harikane et al. \(2022\)](#).

2.2. Revisiting the ALMA Band 6 observations

As stated in [Harikane et al. \(2022\)](#), the ALMA Band 6 DDT observations (2019.A.00015.S, PI: A. K. Inoue) are observed with a spectral scan setup that should target the [O III] 88 μm emission line for a redshift range of $12.6 < z < 14.3$. The observations span the frequency range 222–250 GHz. For the full observational details we refer to [Harikane et al. \(2022\)](#).

We tested whether manual flagging could significantly alter the tentative [O III] 88 μm like feature visible in the image-plane (Fig. 5 in [Harikane et al. \(2022\)](#)). Because antenna DA41 had slightly elevated noise levels over the entire time domain, we tested the impact of completely flagging data from this antenna and found no clear difference between the moment-0 maps generated using the initial, CASA version 6.3 pipeline reduced and calibrated data and those where we flagged antenna DA41. In this work, we therefore use the calibrated data reinstated using the ALMA pipeline generated scripts provided by the observatory.

We image the data using the `tclean` task, adopting natural weighting and creating cubes with channels of 50 km s^{-1} width. The sensitivity per channel is $0.3 \text{ mJy beam}^{-1}$, where the beam FWHM is $0''.51 \times 0''.87$. We find no evidence of any continuum emission. For the continuum, the typical rms noise level is $8.0 \mu\text{Jy beam}^{-1}$, yielding a 3σ upper limit of $24.0 \mu\text{Jy beam}^{-1}$.

3. Data analysis

3.1. Inspecting the [C II] 158 μm and [O III] 88 μm images and spectra

We searched for the [C II] 158 μm (and [O III] 88 μm) emission line from HD1, with the aim of securely confirming its spectroscopic redshift. Thus, we created aperture integrated spectra around the spatial location of HD1, focusing on the spectral range where the [C II] 158 μm and [O III] 88 μm emission lines are expected, given the redshift proposed by [Harikane et al. \(2022\)](#). If HD1 is indeed at $z = 13.27$, the [C II] 158 μm emission line should be located at a frequency of 133.18 GHz. In the top panel of Fig. 1, we present the integrated spectrum of HD1 around this frequency within apertures of 2''.8 and 3''.5 diameter (corresponding to approximately the beam size and 1.25 times the beam size). No emission line is visible around 133.18 GHz for either aperture, indicating that the [C II] 158 μm emission line is not detected at the expected frequency and location. However, there does appear to be a tentative feature centred at a frequency of ~ 133.27 GHz (indicated by the grey shading in Fig. 1), particularly for the larger aperture. We refer to this as the ‘tentative [C II] 158 μm feature’ in the remainder of the text and describe it again later in this section.

In the bottom panel of Fig. 1, we aim to reproduce the [O III] 88 μm detection reported in [Harikane et al. \(2022\)](#), by re-

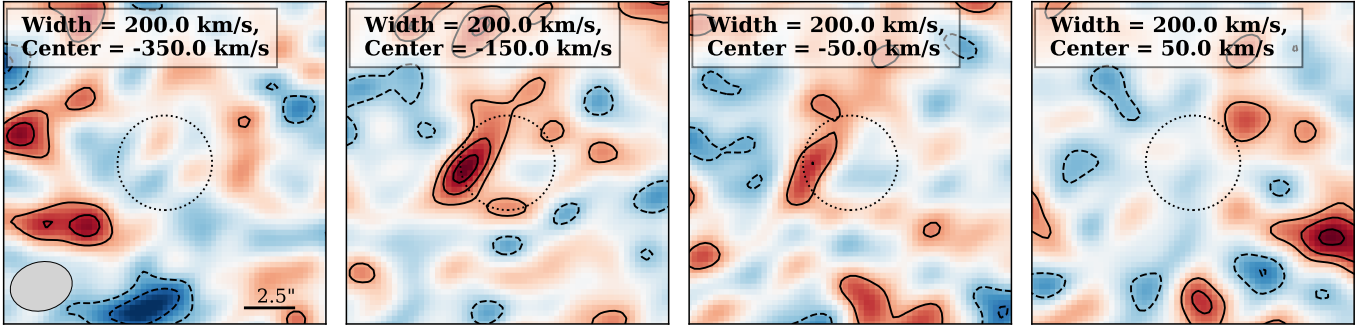


Fig. 2. ALMA Band 4 continuum-subtracted moment-0 maps of HD1. All maps are based on naturally weighted data. Each panel shows a moment-0 map collapsed over the same same integration width of 200 km s^{-1} , but centred at different frequencies, as indicated at the top of each panel. The central frequencies are indicated in km s^{-1} with respect to the reference frequency $\nu = 133.2 \text{ GHz}$, the frequency of the expected [C II] line. For all panels, the contours are drawn at $-3.5, -2.5, -1.5, 1.5, 2.5, 3.5\text{-}\sigma$ and have rms values of $8.32, 8.25, 8.44, \text{ and } 8.88 \text{ mJy beam}^{-1}$ (from left to right). The synthesised beam FWHM is indicated by the ellipse in the lower left and the image scale is shown on the lower right in the left panel. The dashed circle has a $2''.2$ radius.

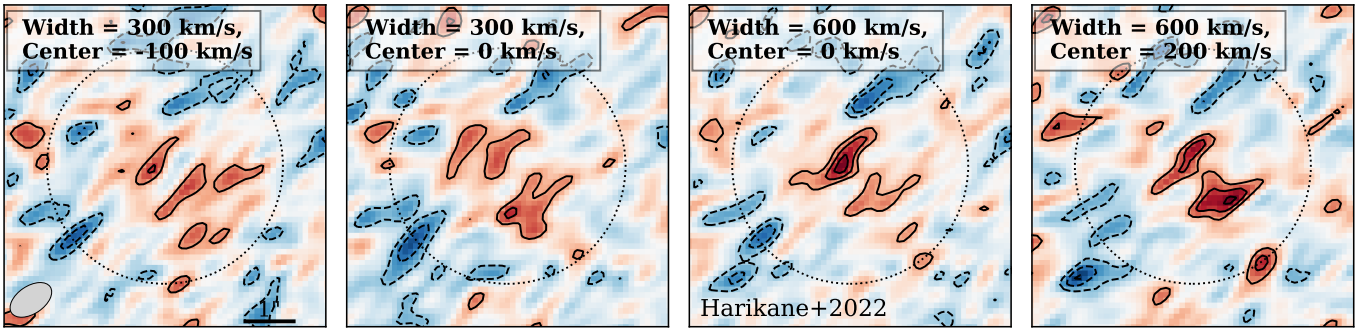


Fig. 3. ALMA Band 6 moment-0 maps of HD1. All maps are based on naturally weighted data. Each panel shows a moment-0 map created using a different integration width and central frequency. The integration width and central frequency are provided in km s^{-1} with respect to the reference frequency, $\nu = 237.8 \text{ GHz}$, as reported in Harikane et al. (2022), at the top of each panel. For each panel, the contours are drawn at $-3.5, -2.5, -1.5, 1.5, 2.5, 3.5\text{-}\sigma$. From left to right, the moment-0 maps have rms values of $38.3, 37.4, 51.5, \text{ and } 52.5 \text{ mJy beam}^{-1}$. The synthesised beam FWHM is indicated by the ellipse in the lower left and the scale of the image is shown on the lower right in the left panel. The dashed circle has a $2''.2$ radius.

senting the integrated spectrum calculated within multiple apertures around the location of HD1. We adopted apertures of $0''.5$ and $1''$ around HD1, corresponding to about twice the beam size of the data and an aperture identical to the one adopted in Harikane et al. (2022), respectively. We reproduced the feature at $\sim 237.8 \text{ GHz}$ when adopting an aperture of 1 arcsec (i.e. approximately six consecutive channels with a positive flux, though most of them with $S/N < 1$). However, we found that this feature was less pronounced when we adopted the $0''.5$ aperture.

To further investigate the tentative features visible in the spectra and test their robustness, we created moment-0 maps focusing on various spectral ranges around the location of the expected emission lines (varying the central frequency and width of the velocity range over which channels are collapsed). Figure 2 shows four, naturally weighted, continuum-subtracted moment-0 maps integrated over a width of 200 km s^{-1} around the expected [C II] $158 \mu\text{m}$ redshifted frequency (133.18 GHz) and location of HD1. We find a tantalising $\sim 4\sigma$ feature, offset by $\sim -150 \text{ km s}^{-1}$ from the expected frequency and spatially offset $1''.8$ from the location of HD1 reported in Harikane et al. (2022, second panel). However, similar features are found throughout the data cube (see for example the right panel in Fig. 2).

We were able to reproduce the findings of Harikane et al. (2022) when following the same procedures to create

moment-0 maps (collapsed over 600 km s^{-1}), as shown in the third panel of Fig. 3. However, features of similar significance to the one found by Harikane et al. (2022) are present when slightly shifting the central frequency ($\sim 200 \text{ km s}^{-1}$) from the original reported central frequency (right panel of Fig. 3). Moreover, integrating over a narrower velocity width (the left two panels, consistent with other high-redshift studies, e.g. Hashimoto et al. 2018) yields no clear positive signal at the reported location of HD1. The multitude of tentative features found in the moment-0 maps of both the Band 4 and 6 data just by shifting the central frequency and/or width over which channels are collapsed, illustrates the difficulty in interpreting tentative detections. A more robust analysis of the noise properties of the data is therefore necessary to provide actual meaning to these features.

3.2. Noise properties of the ALMA data

In this section, we describe the tests we performed to determine whether the tentative line- and image-plane detections, in both the Band 4 and 6 data, could be attributed to structured noise or a non-Gaussian noise distribution. First, we compared the distribution of the pixel values (in flux density per beam) for different moment-0 maps, created with the Band 4 and 6 data (see Sect. 3.1). The left panels of Figs. 4 and 5 show the inverse cumulative distribution function (inverse CDF) of the signal-free

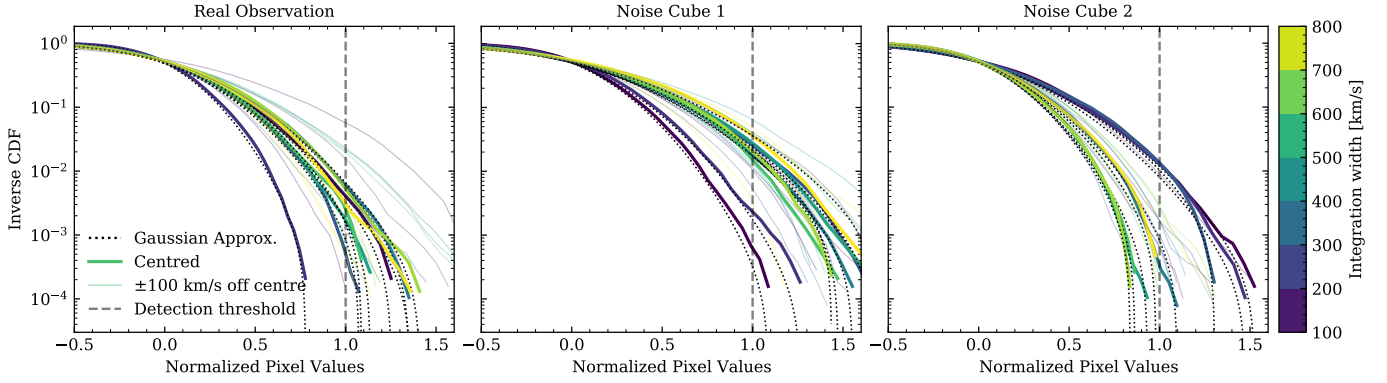


Fig. 4. Inverse cumulative distribution of pixel values outside of a $5.2''$ aperture mask around the location of HD1, normalised by the brightest pixel value within that mask for different moment-0 maps of the ALMA Band 4 observations. Each moment-0 map is integrated with varying velocities widths (as indicated by the colourbar) and varying central frequencies that are off centred by $\pm 100 \text{ km s}^{-1}$ from the supposed [C II] $158 \mu\text{m}$ line (which corresponding moment-0 maps are highlighted in bold). The left panel shows the real Band 4 observations, while the middle and right panel show the pixel-distribution of different pure noise observations (see Sect. 3.2). The thin dotted black lines are the assumed Gaussian distribution of the background fluctuations, estimated by computing the mean and standard deviation over the pixel values. Note that the brightest pixel value within the $5.2''$ aperture can be interpreted as the peak of the surface brightness of a source. Thus, distributions of pixels values that fall left of the vertical dashed line are indicative of a (tentative) detection within the aperture. For further clarification, see Sect. 3.2.

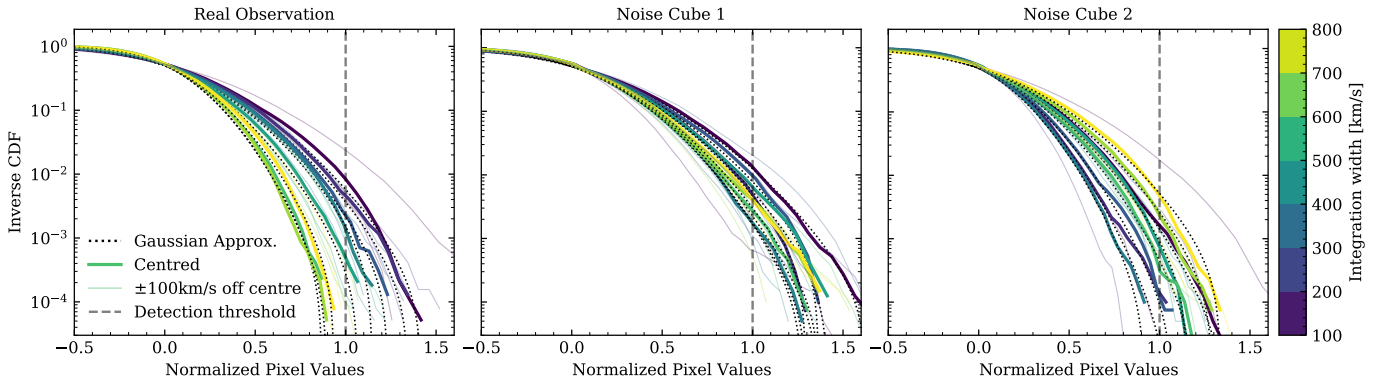


Fig. 5. Same as Fig. 4 but for Band 6 and outside of a $2.2''$ aperture mask.

pixel values for the Band 4 and Band 6 observations, respectively. We took the signal-free pixel values to be the pixel values that fall outside of a $2.2''$ radius for the Band 6 observations and a $5.2''$ radius for the Band 4 observations, centred on the phase reference position. Each line in Figs. 4 and 5 indicates a unique moment-0 map that is integrated at a certain central frequency and over a certain integration width (with the latter indicated by the colourbar). In Fig. 4, the ‘centred’ lines, highlighted in bold, correspond to the velocity offset shown in the second panel in Fig. 2 (-150 km s^{-1}). For Fig. 5, the bold lines correspond to the pixel distribution in the moment-0 maps that are centred at the reference frequency of the tentative [O III] $88 \mu\text{m}$ detection as reported in Harikane et al. (2022), and shown here in the third panel of Fig. 3. To guide the eye, the thin dotted black lines show the assumed Gaussian distribution of the background fluctuations, estimated by computing the mean and standard deviation over the pixel values per integration width from the highlighted moment-0 maps.

We found only minor deviations from the assumed Gaussian distribution (solid colour versus dotted black lines in Fig. 3), which could be caused by large-scale noise correlation (due to side lobes and beam smearing of the primary beam), imperfect continuum subtraction, gridding and weighting artefacts resulting from the discrete Fourier sampling of the visibilities, and/or minor atmospheric absorption. If the pixel distributions over-

shoot the Gaussian assumptions, these values would be associated with a higher significance than when compared to the actual underlying distribution. However, we did not detect any clear deviations that would indicate strong non-Gaussianity in the moment-0 maps.

Finally, we normalised the distribution of pixel values by the brightest pixel value within the above described apertures. In the case of a real, unresolved source at the center of the map, this pixel value would be interpreted as the estimated peak source flux. Therefore, in Figs. 4 and 5, any CDF that ends at values lower than unity (leftwards of the dashed vertical line) implies a possible detection of a source within the central aperture, based on Gaussian statistics. Both left panels in Figs. 4 and 5 show that the tentative [C II] $158 \mu\text{m}$ and [O III] $88 \mu\text{m}$ image-plane detections are only viable (left of the 1.0 mark) when integrated over a unique spectral range (200 and $\sim 500\text{--}700 \text{ km s}^{-1}$, respectively), which could give a physical insight into the velocity structure of HD1. We test how likely these features are to be real in the rest of this section and the following section.

To effectively analyse the noise features, we also created jack-knifed measurement sets. By randomly inverting the amplitudes of half of the visibilities and re-imaging each cube as for the original moment-0 maps, we created pure noise realisations, for which any possible signals were canceled out by the binning. This jack-knifing technique removed any true sources, while

preserving the noise structure of the data. We repeated this jack-knifing procedure 10 times to generate 10 mock image cubes per ALMA Band, using different random seeds for the initial inversion in the uv -plane. The original and mock, jack-knifed data cubes have near equivalent noise properties and are all well-approximated by Gaussian distributions, as shown in the middle and right panel of Figs. 4 and 5.

In Figs. 4 and 5, we highlight two jack-knifed measurement sets per Band. The central panels show realisations wherein all the brightest pixel values fall outside of the central region ($2''2$ for the Band 6 and $5''2$ for the Band 4 data) for all moment-0 maps, which can be interpreted as a non-detection if the observations are treated as real. The right panels show two corresponding jack-knifed moment-0 maps per Band (with noise properties identical to the real observations) for which the brightest pixel value falls within the central region for specific integration widths when making the moment-0 maps. These noise peaks would thus be interpreted as a (tentative) detection when compared with the approximated Gaussian noise distribution. The middle and right panels that show the jack-knifed data cubes reveal very similar CDFs to the real data, with many pixels exceeding the source flux. The close match between the source-free jack-knifed cubes and real data for both Bands indicates that the tentative [C II] $158\ \mu\text{m}$ and [O III] $88\ \mu\text{m}$ lines are consistent with noise fluctuations.

3.3. Likelihood of detecting two, correlated, noise peaks

We tested the likelihood of finding two tentative detections in each data set, in the case that neither corresponds to a real source. In other words, we estimated how likely we are to obtain emission-line-like features at the same redshift and location on the sky if neither are actually line emission from the source. To this end, we applied two line-finding algorithms to both the real and jack-knifed data sets: (1) the python-based line-finding algorithm used by Béthermin et al. (2020) (see their Sect. 6) and Ginolfi et al. (2022) and (2) the algorithm Findclumps, described in Walter et al. (2016), which is available as part of the python library, Interferopy (Boogaard et al. 2021)³. For the real data, we used the dirty image cubes and replicated the noise properties in the jack-knifed cubes as described in Sect. 3.2. We ensured that all data cubes have the same spectral resolution of $50\ \text{km s}^{-1}$ but kept the spatial resolution of the original imaged cubes (i.e. cell sizes of $0''.1$ and $0''.25$, and beam sizes of $0''.51 \times 0''.87$ and $2''.91 \times 2''.27$, for the Band 6 and 4 data respectively). The frequency ranges of the data are the same as those shown in Fig. 1; the Band 6 and 4 data sample similar redshift ranges for [O III] $88\ \mu\text{m}$ and [C II] $158\ \mu\text{m}$ emission ($z = 13.19\text{--}13.41$ and $z = 13.17\text{--}13.37$ respectively).

The first line-finding algorithm that we adopted loops over all channels, excluding those at the edge of the cube (which have a higher rms), and, for each channel, creates moment maps of $200\text{--}800\ \text{km s}^{-1}$ velocity width. For each of these moment maps, the algorithm searches for pixels for which the absolute value exceeds that of the specified noise threshold (i.e. to find both positive and negative peaks), and for each of these central pixels, a spectrum is generated using a square aperture covering an area equivalent to that of the beam. If the S/N of the spectral feature, re-binned over the width of the moment-0 map, also exceeds the specified threshold, the feature is listed as a ‘peak’. Findclumps works slightly differently. Although it also collapses the cube over various user-specified ‘kernels’, in our case

integration widths of $200\text{--}800\ \text{km s}^{-1}$, it then runs SExtractor (Bertin & Arnouts 1996) on the 2D image to identify ‘clumps’ in emission above the specified S/N. Findclumps never generates a spectrum to perform an additional S/N cut over, but uses only the 2D image.

Using both algorithms, we searched for lines with a velocity width between 200 and $800\ \text{km s}^{-1}$. We first performed this search for the entire cubes to test the significance of the tentative lines via the line fidelity, defined as,

$$\text{fidelity}(S/N) = 1 - \frac{N_{\text{neg}}(S/N)}{N_{\text{pos}}(S/N)} \quad (1)$$

where $N_{\text{pos}}(S/N)$ and $N_{\text{neg}}(S/N)$ are the number of positive and negative line candidates with a given S/N, respectively (see Sect. 3 of Walter et al. 2016) and the S/N is that of the peak in the velocity-integrated moment-0 image. In Fig. 6, we show the mean fidelity and number of detections per S/N bin (averaged over all integration widths) for the Band 4 and 6 data, as found with Findclumps. Using the first line-finding algorithm gave almost identical results. Following Walter et al. (2016), we consider a fidelity threshold of 60% as the minimum significance for a line detection (the vertical dashed line). The S/N corresponding to this fidelity threshold are 5.3 and 5.2, for the Band 4 and 6 data respectively. We repeated the same exercise for the jack-knifed data cubes, which are equivalent to pure noise cubes, any potential source having been nullified by the jack-knifing (see Sect. 3.2). We took the mean over all line widths and all 10 cubes. The mean S/N at a fidelity of 60% for both the Band 4 and 6, jack-knifed cubes is 5.3 ± 0.2 . Conversely, the mean fidelity of the Band 4 and 6 jack-knifed cubes, at the S/N of the tentative features, are 0.06 ± 0.04 and 0.09 ± 0.05 respectively. The combined fidelity is then $= 0.5 \pm 0.4\%$. Thus, the significance of the tentative features, both on their own and in combination, is too low for them to be considered real.

Next, we tested the significance of the line detections given the expected source position. We searched for features within $\sim 2''.9$ of the expected source position, equivalent to $\leq 10\ \text{kpc}$ at $z = 13.27$. This search region is motivated by the possibility of finding close companions as well as components within the exact same source. We based the S/N thresholds applied to both algorithms on the tentative [O III] $88\ \mu\text{m}$ and [C II] $158\ \mu\text{m}$ lines. Since these features have a minimum moment-0 S/N threshold of 3.8σ (the [C II] $158\ \mu\text{m}$ feature actually being at 4σ) and spectral S/N threshold of 3.1σ , we applied these cutoffs to the line-finding algorithms. We removed all duplicates (i.e. noise peaks with any spatial or spectral overlap), by keeping the component with the highest S/N. In this way, we generated a list of potential, positive and negative $\geq 3.8\sigma$ features (defined via the moment-0 S/N) of velocity width $200\text{--}800\ \text{km s}^{-1}$, for each cube.

Applying the first line-finding algorithm to the real data cubes, we found one positive and one negative $\geq 3.8\sigma$ feature within $10\ \text{kpc}$ of the expected source position in the Band 4 data cube. We found three positive and eight negative $\geq 3.8\sigma$ features in the Band 6 data cube. Similarly, applying Findclumps to the Band 4 data cube, we found one positive and one negative $\geq 3.8\sigma$ feature within the $2''.9$ aperture. We also found five positive and nine negative $\geq 3.8\sigma$ features in the Band 6 data cube within this aperture. The slightly greater number of features found by Findclumps in the Band 6 data is due to the fact that Findclumps does not perform the additional S/N cut on the spectrum. All lists included the line candidates for the proposed $z = 13.27$ source (by design). The higher number of $\geq 3.8\sigma$ features in the Band 6 data is mainly due to the difference in spatial

³ <https://interferopy.readthedocs.io/en/latest/>

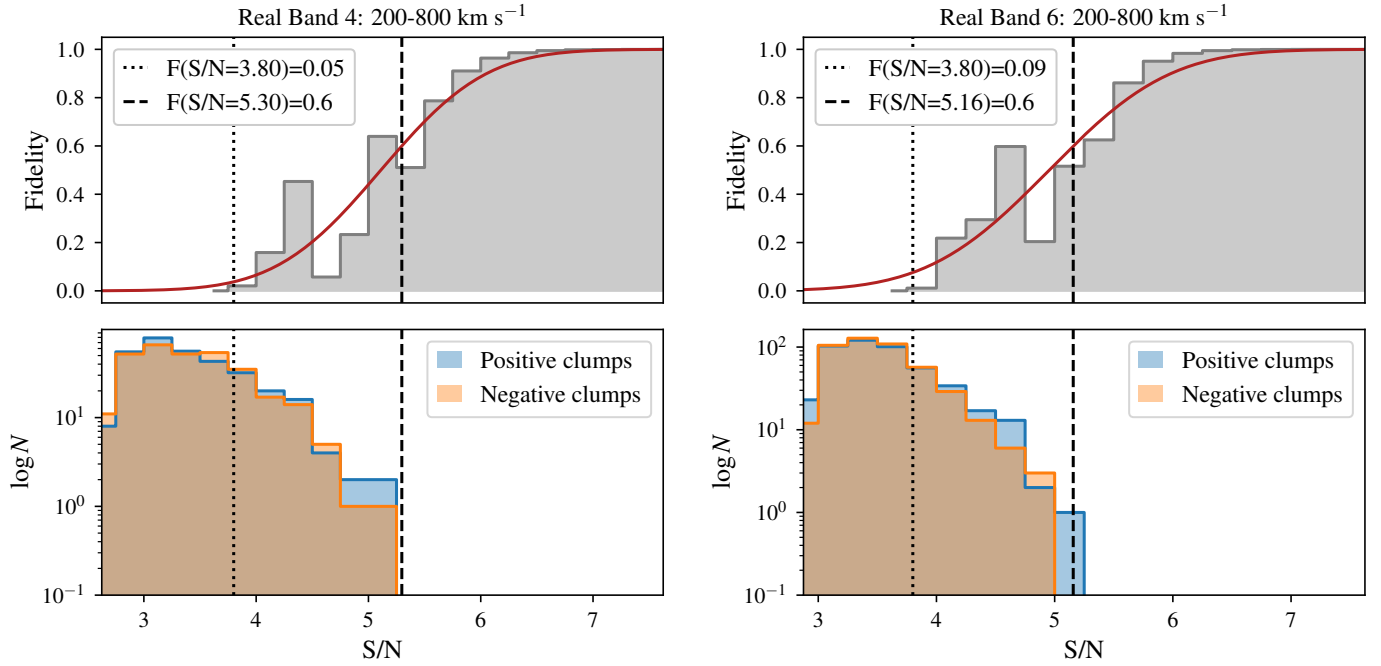


Fig. 6. Fidelity (top) and number of peaks found (bottom) as a function of S/N for both the Band 4 data (left) and Band 6 data (right) using `Findclumps`. The dashed vertical line marks a fidelity of 0.6, beyond which it is more likely to select peaks that are positive than negative. The dotted vertical lines mark the S/N threshold of the tentative [O III] 88 μm and [C II] 158 μm features. The red line in top panels is an error function fit to the fidelity histogram. This functional form has no underlying physical significance but was found empirically to provide a good fit (see [Walter et al. 2016](#)).

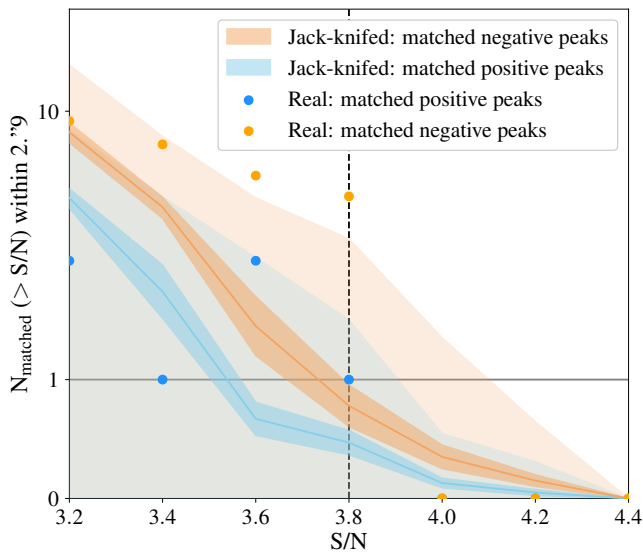


Fig. 7. Number of matched peaks, within 1000 km s^{-1} , above a certain S/N and within $2''.9$ of the expected position of HD1. The filled blue and orange circles correspond to the number of positive and negative matched peaks, respectively, in the real data cubes. The solid line, dark- and light shaded regions correspond to the mean, standard deviation and standard error about the mean for the matched pairs of jack-knifed data cubes. We find one matched positive pair in the real data (the tentative [O III] 88 μm and [C II] 158 μm features) at $S/N > 3.8$. Given the mean and standard deviation on the matched peaks in the jack-knifed data, one set of matched peaks in the real data is perfectly consistent with being noise.

resolution. The fidelity within this smaller search region is 0 for the Band 4 data but is negative for the Band 6 data.

Cross-matching the real Band 4 and 6 data cubes, we recovered only one matched pair of positive $\geq 3.8\sigma$ features within 10 kpc and 1000 km s^{-1} of each other, which correspond to the proposed [O III] 88 μm line at $z = 13.27$ (Band 6 cube) and tentative [C II] 158 μm line, offset by ~ 6 kpc and ~ -190 km s^{-1} ⁴. However, we also found three matched pairs of negative $\geq 3.8\sigma$ features within 10 kpc and 1000 km s^{-1} of each other, implying that the positive matched pair is unlikely to be real. The number of matched peaks are shown as the filled circles in Fig. 7.

We also restricted the line-finding procedure to find peaks within $\sim 2''.9$ of the source position for the jack-knifed data cubes (of which there are 10 each for the Band 4 and 6 data). Of the 10 jack-knifed Band 4 cubes, five (four) have at least one positive $\geq 3.8\sigma$ feature within 10 kpc of the expected source position (the centre of the cube), whereas three (four) have at least one negative feature of the same significance. The two quoted values are for the two different line-finding algorithms. Of the 10 jack-knifed Band 6 cubes, all have at least one positive $\geq 3.8\sigma$ feature within 10 kpc of the expected source position, with a mean of six (eight) such features per entire cube. Likewise, all jack-knifed Band 6 cubes have at least one negative $\geq 3.8\sigma$ feature, with a mean of seven (eight) such features per cube (with `Findclumps` finding the higher number of peaks).

We expanded this analysis, testing how likely we are to find two noise features (one each in the Band 4 and 6 data) that are offset by ≤ 10 kpc and ≤ 1000 km s^{-1} for $3.2 < S/N < 4.4$. The mean, standard error on the mean⁵, and standard deviation (over all matched pairs) are shown in Fig. 7. Of the 100 pairs of matched Band 4 and 6 noise cubes, 30 pairs have at least one

⁴ The slight difference in spectral offsets compared to the value in Sect. 3.2 is due to the fact that here we are measuring the separation of the central line frequency from the line-finding algorithm.

⁵ σ/\sqrt{N} , where σ is the standard deviation and N is the number of elements in the sample.

matching $\geq 3.8\sigma$ noise feature, with a maximum of six matched noise peaks per pair of Band 4 and 6 noise cubes. The mean number of corresponding pairs of $\geq 3.8\sigma$ emission features within 10 kpc and 1000 km s^{-1} is 0.5, with a standard error on the mean of 0.1 and standard deviation of 1. Repeating the exercise for the negative peaks, we found that 25 of the matched cubes have at least one matched pair of negative peaks, with a mean of 0.8 such features, a standard error on the mean of 0.2 and standard deviation of 1.8. The large standard deviation for the matched noise cubes is the result of the small initial search area about the source centre. The mean number of matched peaks implies that with the data in hand there is, on average, a 50% chance of detecting spatially and spectrally correlated $\geq 3.8\sigma$ noise peaks in both the Band 4 and 6 data. However, given the standard deviation, the likelihood is consistent with 100%. We found that the number of matched noise peaks drops sharply with S/N. The number of jointly detected noise peaks drops to 0 at a S/N threshold of ~ 4.4 . Thus, at $S/N > 4.4$ two matched positive features in the Band 4 and 6 data are certain to be real whereas at $S/N > 3.8$ they are fully consistent with both being noise peaks.

4. Exploring the potential redshift solutions

Since we find no evidence of any line emission from a $z \sim 13.3$ galaxy, we revisit the likelihoods of the potential redshift solutions using the existing photometry and new upper limits. By modelling the spectral energy distribution (SED) of HD1 using the SED modelling tool PANHIT, Harikane et al. (2022) found two potential redshift solutions, $z \gtrsim 13$ and $z \sim 4$, corresponding to either a Lyman or Balmer break bluewards of $1.8 \mu\text{m}$. Since the new ALMA Band 4 data analysed here do not confirm the $z \sim 13.3$ scenario, we further explored other potential redshift solutions. We repeated the SED-fitting analysis to test the likelihood of different redshift solutions and checked whether these solutions are consistent with our upper limits on the submillimetre line emission (the continuum upper limits already being included in the SED modelling). To explore the likelihood of different redshift solutions, we used both the Bayesian Analysis of Galaxies for Physical Inference and parameter ESTimation (Bagpipes; Carnall et al. 2018) and the high- z version of the Multi-wavelength Analysis of Galaxy Physical Properties (MAGPHYS; da Cunha et al. 2015; Battisti et al. 2019) SED-fitting codes. Using these, we modelled the existing optical-NIR photometry (presented in Harikane et al. 2022), the upper limits on the *Herschel* photometry⁶, from the PACS Evolutionary Probe (PEP, Lutz et al. 2011) and *Herschel* Multi-tiered Extragalactic Survey (HERMES, Oliver et al. 2012) and the upper limits on the ALMA Band 6 and 4 continuum obtained here (Sect. 2).

4.1. BAGPIPES fits

To model the SED of HD1 with Bagpipes, we followed a similar approach to Donnan et al. (2023). We applied the Salim et al. (2018) parameterisation to model the dust attenuation. For this, we centred the power-law deviation from the Calzetti et al. (2000) law to, $\delta = 0$, applied a Gaussian prior with a standard deviation, $\delta = 0.1$, and set a uniform prior on the 2175 \AA bump strength, B , from 0 to 5. We set the logarithm of the ionisation parameter to $\log(U) = -3$ and allowed the stellar metallicity, Z_* , to vary with a uniform prior from $0.01 < Z_*/Z_\odot < 3$. We explored a variety of star formation histories (SFHs): con-

stant, exponentially rising and double power law. For the constant SFH, we varied the time before observations at which stars began forming from 1 Myr to the age of the Universe.

To effectively exploit all the available spectral information, we incorporated the 14 upper limits in our modelling process in a way that is statistically meaningful in comparison to the three actual detections. We therefore extended the likelihood distribution function in Bagpipes to include an upper limit term that is consistent with the modified normal likelihood formalism for non-detections presented in Sawicki (2012). This implementation gave very similar results to setting all flux values to 0 in the standard version of Bagpipes. In this case, the $z > 13$ solution has a far higher probability than the $z \sim 4$ solution, at least when we provide the wide and uniform priors described previously.

The likelihood distributions of the fit parameters (assuming a constant SFH) are shown in Fig. 8. This results in two main redshift scenarios, $z \sim 0.3$ and $z \sim 14$, with the former appearing far more likely. To allow the nested sampling algorithm to even find the $z > 0.3$ solutions, we needed to use a large number of live points (5000). In other words, when we left large uniform priors of $A_V = 0-10$ and $B = 0-5$, we found that the most likely solution is neither of the two scenarios presented in Harikane et al. (2022), but instead a third scenario of a dwarf galaxy at $z \sim 0.2-0.4$, with stellar mass of $10^{8.3} M_\odot$, $A_V \sim 6-10$ and a stellar population age of a few Gyr.

The low-redshift scenario dominates for all SFHs, as long as we do not restrict the A_V . There is a difference of $\Delta z = 0.05$ between the dominant, low-redshift solutions for a constant, exponential and double power law SFH. When we restricted the optical extinction further, for example to $A_V = 0-5$, then the $z \sim 4$ solution started to become plausible. Further restrictions to A_V ruled out the low-redshift solution entirely, increasing the likelihood of the passive $z \sim 4$ solution. When we forced the redshift range to $z = 2-20$, but kept $A_V = 0-10$ then the $z \gtrsim 14$ solution dominated over that at $z \sim 4$. When restricting the redshift range in this way, the relative likelihood of the $z \sim 4$ and $z \sim 14$ solutions depended more strongly on the type of SFH and priors on the relevant ages. Given the range of redshift probability distributions that we found for different prior combinations, we are forced to conclude that with the available photometry there is no robust way to identify a single, best redshift fit.

4.2. MAGPHYS fits

We checked the Bagpipes results against MAGPHYS high- z , fitting for a range of potential redshift solutions, a subset of which are shown in Fig. 9. The redshift solutions that we chose to test are guided mainly by the results of the previous SED modelling performed by Harikane et al. (2022) and here, with Bagpipes. We also fit for $z \sim 6.78$, a redshift solution that would correspond to a velocity offset of 1030 km s^{-1} between the CO(16–15) and CO(9–8) lines that would be covered by the Band 4 and 6 observations.

We checked the best-fit SEDs for the different potential redshift solutions via the residuals and χ^2 values of the best-fit MAGPHYS models (see legend at top left of Fig. 9). The $z \sim 0.3$, $z \sim 4$ and $z > 13$ solutions, found with Bagpipes also provide a good fit when using MAGPHYS. The likelihood distributions of the stellar mass and A_V for each redshift solution, are similar to those found using Bagpipes. In contrast to the model results using Bagpipes, solutions yielding a redshift of $z > 13$ provide the best fit using MAGPHYS, in agreement with Harikane et al. (2022). This could be due to an array of factors, including the available model libraries or the different treatment

⁶ Available at <https://cosmos.astro.caltech.edu/page/herschel>

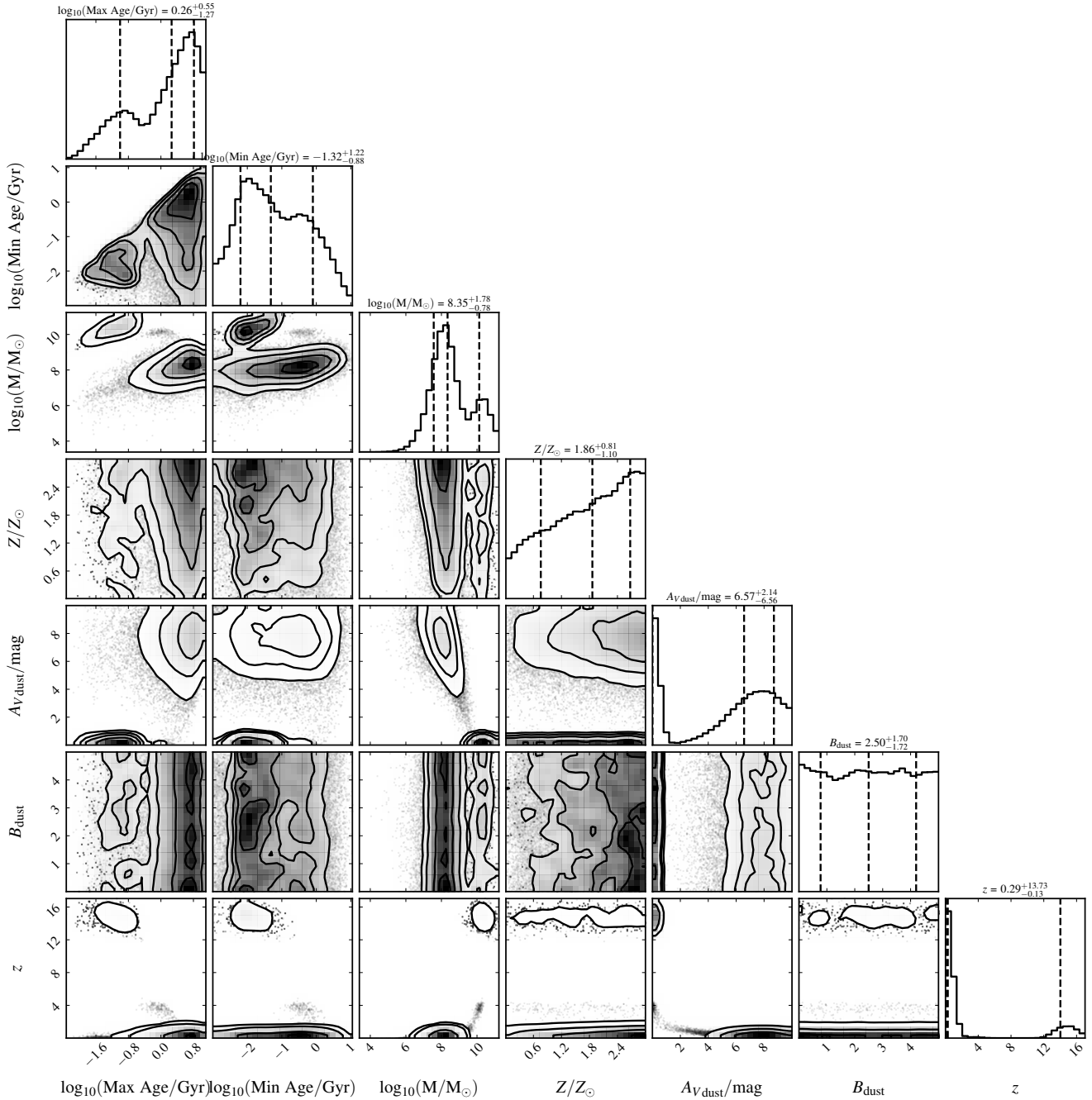


Fig. 8. Corner plot for the parameters fit to the photometry of HD1 using Bagpipes, assuming a constant SFH. From left to right (and top to bottom) posterior probability distributions are shown for the minimum and maximum age (i.e. the time since star formation switched off and on), the stellar mass (M/M_{\odot}), metallicity (Z/Z_{\odot}), optical extinction (A_V), 2175 Å bump strength (B) and redshift (z). With so few photometric detections, the 2175 Å bump strength and metallicity are unconstrained. As shown in the bottom right corner, the $z \sim 0.2$ – 0.3 solution dominates, with the $z \sim 14$ solution the next most probable. The low- z solution, which corresponds to a dwarf galaxy ($M_{\star} \sim 10^{8.3} M_{\odot}$) with a very high optical extinction ($A_V \sim 7.5$) also dominates for other assumed SFHs (not shown here). However, the relative probabilities of these solutions are highly dependent on the priors, as described in Sect. 4.3.

of dust attenuation and emission, which are not worth exploring here.

4.3. Potential redshift solutions from the SED fitting

Given the range of redshifts that could be fit using MAGPHYS and Bagpipes, we conclude there are many possible redshift (and hence SFR, stellar mass etc.) solutions that can be fit to the three photometric detections and 14 upper limits. Although the $z = 6.78$

solution, for example, definitely provides a poorer fit, there is no good way to distinguish between $z \sim 0.3$, $z \sim 4$ and $z > 13$ scenarios using the photometry alone. The SED fitting results are highly sensitive to any tiny colour offsets or calibration issues for the three photometric detections because the slope across these is too flat and they span a too narrow wavelength range for the best-fit solution to be trusted. For now, we conclude that there are three main possible sets of redshift solutions between which the current data cannot distinguish: (1) a dust-rich low-redshift

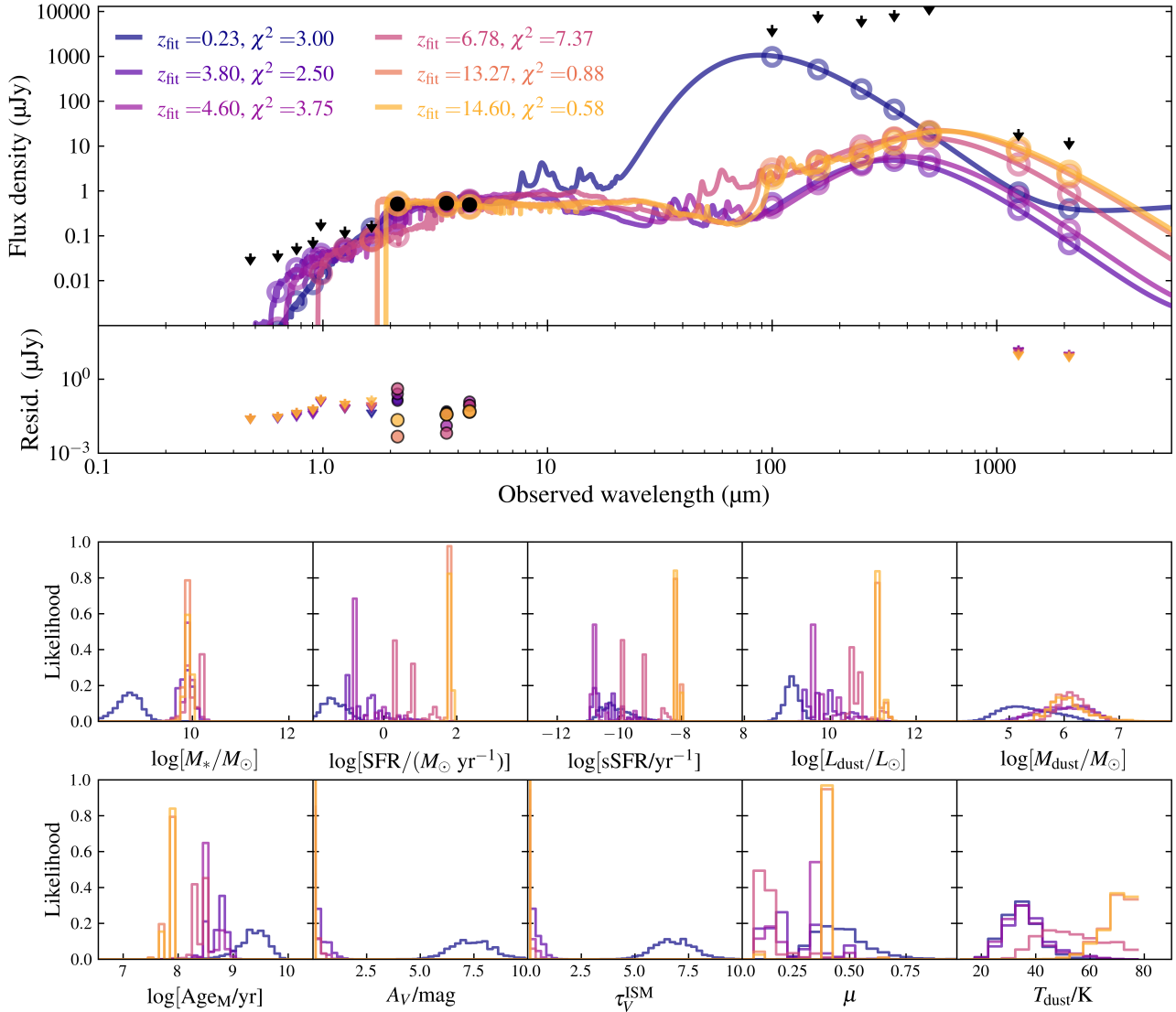


Fig. 9. Results of modelling the photometry for HD1 with MAGPHYS. Top panels: best-fit SEDs for different potential redshift solutions, indicated by the legend at the top left, with the residuals shown below. The $z \sim 3.8$ and $z = 4.6$ solutions correspond to a passive galaxy where the drop in flux is due to the Balmer break. The best intermediate-redshift fit to the limited available data is for $z = 3.8$. In this scenario, the Band 6 data would cover CO(10–9), which we do not expect to be detected for such a passive galaxy (see text). The highest redshift solution within the range probed by the Band 6 data ($12.6 < z < 14.3$) provides the best fit overall. The single redshift solution between $z = 4$ and 13 for which the two $\geq 3\sigma$ features could physically correspond to lines, $z = 6.76$, is a much poorer fit to the photometry. Bottom panels: likelihood distributions for the different fit parameters, colour-coded according to the legend in the top panel.

($z = 0.2$ – 0.3) dwarf galaxy, (2) a passive galaxy at $z = 3.6$ – 4.6 , for which the photometric break corresponds to the Balmer break, or, (3) a $z > 13$ galaxy, with a Lyman break bluewards of $2 \mu\text{m}$.

4.4. Consistency between the SED fits and line upper limits

We explored whether the upper limits on the line luminosities covered by the Band 4 and 6 data are consistent with the three redshift solutions inferred previously from the photometry. For this, we used the mean rms per 50 km s^{-1} channel, excluding the channels at the outskirts of the SPW, which is $0.319 \text{ mJy beam}^{-1}$ for the Band 6 data and $0.096 \text{ mJy beam}^{-1}$ for the Band 4 data. In each of the three cases, we assumed a line width of 200 km s^{-1} , applied a S/N threshold of 4, and assumed that the source is unresolved for both the Band 4 and 6 beam. The last assumption is motivated by the fact the maximum radius of the emission seen in the VISTA K_s image is $0''.7$, which is equivalent to the beam

size of the Band 6 data and significantly smaller than the beam of the Band 4 data. This gives upper limits of $127.7 \text{ mJy km s}^{-1}$ and $38.4 \text{ mJy km s}^{-1}$, for the Band 6 and 4 data respectively.

4.4.1. Low-redshift dwarf-galaxy scenario

We first explored the $z \sim 0.2$ – 0.3 scenario, which corresponds to a dust-rich but barely star-forming ($\text{SFR} \sim 10^{-2} M_{\odot} \text{ yr}^{-1}$) dwarf galaxy. In this case, a high visual extinction of $A_V \sim 7.5$ is required to explain the SED. For this scenario, the Band 6 observations would cover the frequency of HCN(3–2) for $z = 0.06$ – 0.20 and CO(3–2) for $z = 0.38$ – 0.56 whereas the Band 4 observations cover HCN(2–1) for $z = 0.32$ – 0.34 .

Since HCN is typically far fainter than CO (e.g. Vanzani et al. 2009), we first tested whether we could have detected CO(3–2), for a dwarf galaxy at $z \sim 0.4$. The ALMA data provide an upper limit on the CO(3–2) line luminosity of $11.7 \times 10^7 \text{ K km s}^{-1} \text{ pc}^2$.

The dust mass predicted by the MAGPHYS fit is $M_{\text{dust}} \sim 10^{5.8} M_{\odot}$. Supposing that the galaxy has a gas-to-dust mass ratio of 100, which is conservative for local dwarf galaxies (e.g. De Looze et al. 2017), the predicted gas mass is $6 \times 10^7 M_{\odot}$. Assuming the galaxy follows the mass-metallicity relation of McQuinn et al. (2020), the α_{CO} vs metallicity relation of Feldmann et al. (2012) and assuming a conservative, Milky Way outer disk CO(3–2)-to-CO(1–0) line luminosity ratio of 0.3 (Fixsen et al. 1999), the predicted CO(3–2) luminosity is $\sim 3 \times 10^6 \text{ K km s}^{-1} \text{ pc}^2$. This estimate is 37× lower than the upper limit on the line luminosity. Effects such as the variation in rms with frequency, minor changes in the dust-to-gas ratio or the galaxy actually having a higher metallicity cannot raise this value to more than the sensitivity. Thus, the lack of a line detection in the Band 4 and 6 data is perfectly consistent with the existence of a high-extinction, and low-SFR, dwarf galaxy at $z \sim 0.3$. Moreover, it may well be that the exact redshift is simply not covered by the observations. Although we cannot rule out this scenario, it has proven challenging to find an example of a low-redshift dwarf galaxy with such a high A_V in the literature.

4.4.2. Passive $z \sim 4$ galaxy scenario

Next, we tested the case where the drop in flux bluewards of the observed-frame $1.8 \mu\text{m}$, corresponds to a Balmer break (i.e. drop in flux bluewards of the rest-frame 3646 \AA). The potential redshift range consistent with the Balmer break solution is $z \sim 3.6\text{--}4.6$. For a subset of this range, $4.15 < z < 4.23$, the ALMA Band 4 observations would cover the frequency of the CO(6–5) line. The ALMA Band 6 observations would cover the frequency of CO(10–9) for $3.61 < z < 4.19$ and CO(11–10) for $4.07 < z < 4.71$.

We first considered the case of CO(6–5) emission from a $z = 4.15$ galaxy, for which the upper limit on the line luminosity is $7.1 \times 10^8 \text{ K km s}^{-1} \text{ pc}^2$. Converting the upper limit on the SFR from the best-fit solution found by Harikane et al. (2020), $\text{SFR} < 0.1 M_{\odot} \text{ yr}^{-1}$, to an IR luminosity via the relationship of Kennicutt (1998, assuming a Chabrier IMF), gives $L_{\text{IR}} = \text{SFR}/1.01 \times 10^{10} = 1 \times 10^9 L_{\odot}$. The best-fit value from MAGPHYS is $L_{\text{IR}} = 3.6 \times 10^9 L_{\odot}$. Given these L_{IR} vs. $L'_{\text{CO}(6-5)}$ relation of Liu et al. (2015), these L_{IR} estimates translate to a predicted $L'_{\text{CO}(6-5)} = (0.4\text{--}1.2) \times 10^7 \text{ K km s}^{-1} \text{ pc}^2$, which is at least a factor of six below the CO(6–5) detection threshold. Thus, the lack of a line detection remains perfectly consistent with a passive galaxy at $z \sim 4.15$.

We repeated the same exercise for CO(10–9) at $z \sim 3.85$. Using the L_{IR} values predicted from the SED fits and L_{IR} vs line luminosity relations from Liu et al. (2015), we found that the predicted value of $L'_{\text{CO}(10-9)} = (0.03\text{--}0.15) \times 10^7 \text{ K km s}^{-1} \text{ pc}^2$ is far lower than the 4σ line limit of $L'_{\text{CO}(10-9)} = 7.4 \times 10^8 \text{ K km s}^{-1} \text{ pc}^2$. The CO(11–10) line is even fainter and thus even less likely to be detected. Since neither the lines nor continuum covered by the Band 4 and 6 observations for the passive $z = 3.6\text{--}4.6$ galaxy solution should be detected at the sensitivity of our observations, we also cannot rule out (nor provide any stronger evidence for) this scenario.

4.4.3. Star-forming $z > 13$ galaxy scenario

Finally, we tested whether the lack of a significant line detection in the ALMA Band 4 or 6 data could rule out a galaxy at $z \sim 13.3$. In this case, the upper limits on the [O III] $88 \mu\text{m}$ and [C II] $158 \mu\text{m}$ line luminosity are $5.2 \times 10^8 \text{ K km s}^{-1} \text{ pc}^2$ and

$4.9 \times 10^8 \text{ K km s}^{-1} \text{ pc}^2$, respectively. In solar luminosity units, this corresponds to $L_{[\text{O III}]} < 6.4 \times 10^8 L_{\odot}$ and $L_{[\text{C II}]} < 1.1 \times 10^8 L_{\odot}$. In their Fig. 11, Harikane et al. (2020) presented the predicted ratio of the [O III] $88 \mu\text{m}$ and [C II] $158 \mu\text{m}$ luminosity with the SFR of galaxies based on CLOUDY (Ferland et al. 2017) ionisation modelling. From the SED fitting performed here (and by Harikane et al. 2022), the SFR for the $z \sim 13.3$ is $\sim 100 \text{ Myr}^{-1}$. Thus, the upper limits on these ratios are $\log [\text{O III}]/\text{SFR} < 6.7$ and $\log [\text{C II}]/\text{SFR} < 6.0$. This is fully consistent with an ISM metallicity of $\leq 0.2 Z_{\odot}$ and a density of $\sim 10^3 \text{ cm}^{-3}$. Moreover, Kohandel et al. (2023) recently conducted a more detailed study of simulated $z > 10$ galaxies, showing that they mostly fall below the local metal-poor [O III] $88 \mu\text{m}$ –SFR relation, mainly as a result of the low ionisation parameters ($U_{\text{ion}} < 10^{-3}$). Even their [O III]–brightest simulated galaxy fell below the detection limit for HD1. Thus, the lack of any significant ALMA Band 4 or 6 detection corresponding to the [C II] $158 \mu\text{m}$ and [O III] $88 \mu\text{m}$ emission lines, also does not rule out HD1 being a $z \sim 13.3$ galaxy. It could simply imply a low ISM metallicity, low ionisation parameter and/or high gas density. Alternatively, if HD1 were at $z > 14.3$ it would simply not have been covered by the Band 4 or 6 observations.

5. Summary

With this study, we aimed to confirm or reject the proposed redshift of $z = 13.27$ for the galaxy, HD1. We recover the 3.8σ feature in the Band 6 data found by Harikane et al. (2022) and find one 4σ feature in the Band 4 data that is spatially offset by $1''.7$ (6 kpc at $z \sim 13.27$) and spectrally offset by 190 km s^{-1} . We quantify the significance of these features through various statistical tests, revisit the SED modelling and test whether the upper limits on the line luminosities are consistent with the best fit parameters from the SED modelling. We find that the tentative features in the Band 4 and 6 data are fully consistent with being noise peaks and thus find no credible evidence for a $z \sim 13.3$ galaxy. However, even with the new line and continuum upper limits, we cannot rule out this scenario nor lower-redshift, passive-galaxy scenarios.

By extensively testing the ALMA data, we determined the likelihood that the spatially and spectrally matched features found in the Band 6 and 4 cubes are consistent with noise peaks. To this end, we first performed a jack-knifing analysis, creating mock noise cubes with the same noise properties as the data, but with any real signal nullified. We find that the pixel distribution of these cubes closely matches that of the real data, with a significant number of pixels exceeding the value of any putative $z = 13.27$ [O III] $88 \mu\text{m}$ or [C II] $158 \mu\text{m}$ emission. That is, features at a statistical significance $\geq 3.8\sigma$ (the significance of the proposed [O III] $88 \mu\text{m}$ feature) appear often in the ALMA data, and should be treated with caution (see also Vio & Andreani 2016; Vio et al. 2017, 2019). Moreover, by applying two line-finding algorithms to the real and jack-knifed data, we quantify the significance of the lines. We determine the line significance via the mean fidelity of the real and jack-knifed data, which is $6 \pm 4\%$ and $9 \pm 5\%$ for the putative [O III] $88 \mu\text{m}$ and [C II] $158 \mu\text{m}$ features, with a combined value of $0.5 \pm 0.4\%$. From this, and the significant number of positive and negative $\geq 3.8\sigma$ features detected within $2''.9$ (10 kpc at $z \sim 13.3$) of the source, we conclude that the putative features are likely noise peaks. Cross-correlating the positions of $\geq 3.8\sigma$ noise peaks in the jack-knifed cubes, we find a mean of 0.5 and standard deviation of 1 matched noise peaks (within $\leq 10 \text{ kpc}$ and $< 1000 \text{ km s}^{-1}$). Thus, even in combination, the $\geq 3.8\sigma$ features recovered in the Band 4 and 6

data are fully consistent with being noise peaks. The likelihood that such matched features are consistent with noise only drops to 0 at $S/N = 4.4$.

Revisiting the SED modelling, we found three possible redshift solutions, $z \sim 0.2\text{--}0.3$, $z \sim 4$ and $z > 13$. These solutions all provide a good fit using `Bagpipes` and `MAGPHYS`. Applying these two different SED modelling tools, we find different likelihoods to [Harikane et al. \(2022\)](#) and [Fujimoto et al. \(2022\)](#), indicating that the results are highly sensitive to the input priors and underlying models. We caution against characterising the redshift probability using the existing photometry as the slope of the SED for the three photometric detections is too flat and they span too narrow of a wavelength range for the best-fit solution to be trusted. We test whether any of the three redshift solutions we found here can be ruled out by the upper limits on the lines covered by the ALMA observations. Even if it were within the redshift range for which the CO(3–2) would have been covered by the ALMA observations, we would not have detected the high-extinction and low-SFR dwarf galaxy that provided a good fit to the photometry. Likewise, the CO(6–5), CO(10–9) or CO(11–10) features that might have been covered for a passive $z \sim 4$ galaxy would not have been detected at the sensitivity of the observations. We also cannot rule out the $z > 13$ scenario. For $z = 13.17\text{--}13.37$, the redshift range covered by the ALMA Band 4 observations, it may simply be that the line luminosity is below the detection threshold due to a low ISM metallicity ($\leq 0.2 Z_{\odot}$), high gas density ($\geq 10^3 \text{ cm}^{-3}$) and/or low ionisation parameter ($U_{\text{ion}} \leq 10^{-3}$ as for the simulated $z > 10$ galaxies in [Kohandel et al. 2023](#)). Moreover, given the wavelength coverage, we cannot rule out HD1 simply being at $z > 13.37$ (i.e. the [C II] 158 μm line is not covered) or indeed $z > 14.3$ (in which case neither the [O III] 88 μm nor [C II] 158 μm line are covered by the existing ALMA data). Thus, we find that even with the addition of the ALMA Band 4 upper limits, the existing data cannot be used to constrain the redshift of HD1.

Determining the redshift and properties of HD1 will require at least one of the following: (1) deep, sub-arcsecond resolution photometry at $\sim 100 \mu\text{m}$ (as the upper limits provided by *Herschel* provide too little constraining power), (2) $\geq 10\times$ deeper observations at submillimetre wavelengths, or, (3) spectra covering the apparent break at $\sim 2 \mu\text{m}$ and flux bluewards of the break. For (1), there is currently no suitable observational facility and (2) may require significantly more ALMA time ($\sim 500 \text{ h}$). Thus, option (3) using deep JWST/NIRSpec observations, will likely be the most feasible way to determine the nature of HD1. Indeed, recently four $z > 10$ galaxies have been confirmed in this way ([Curtis-Lake et al. 2022](#)).

Put in a broader context, we find that a secure [O III] 88 μm or [C II] 158 μm detection with ALMA can be used to confirm a redshift, but for the typical sensitivities reached by ALMA in spectral scans, a non-detection cannot be used to rule out a redshift. On the other hand, once a high redshift has been confirmed, e.g. with JWST/NIRSpec spectroscopy, ALMA observations (including upper limits) may provide useful constraints on the metallicity and ionisation parameter of the ISM in the first galaxies. Looking to the future, we can expect significant advances to be provided by major new sub-/millimetre facilities, such as the Atacama Large Aperture Submillimetre Telescope (AtLAST; [Ramasawmy et al. 2022](#)). AtLAST will enable large, unbiased surveys of cosmological volumes in multiple bands, providing secure line identifications for large samples of high- z galaxies without the need for photometric pre-selection (e.g. [Algera et al. 2023](#)).

The main take-away message for the reader interested in confirming high-redshift galaxy candidates with ALMA, is to test the noise properties made in obtaining ALMA data products. All too often, we are willing to trust a $3\text{--}4\sigma$ peak at the position where we expect to find it. But without testing how likely this is to be a noise fluctuation, we may well be sorely mistaken.

Acknowledgements. We thank the referee for a constructive report that helped improve the analysis and discussion of the likely redshift of HD1. We also thank Dr. Elisabete da Cunha and Dr. Leindert Boogaard for helpful discussions that greatly improved the quality of the analysis and interpretation of the data herein. Our thanks also to the creators of the publicly available codes used herein; thanks to Dr. Elisabete da Cunha for `MAGPHYS`, Dr. Adam Carnall for `Bagpipes` and Drs. Mladen Novak, Leindert Boogaard and Romain Meyer for `Interferopy`. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2021.A.00008.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. Calibrated measurement sets were provided by the European ALMA Regional Centre network ([Hatziminaoglou et al. 2015](#)) through the calMS service ([Petry et al. 2020](#)). L.D.M. is supported by the ERC-StG ‘ClustersXCosmo’ grant agreement 716762. AtLAST has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 951815.

References

- Algera, H. S. B., Inami, H., Oesch, P. A., et al. 2023, *MNRAS*, 518, 6142
- Atek, H., Shuntov, M., Furtak, L. J., et al. 2023, *MNRAS*, 519, 1201
- Bakx, T. J. L. C., Tamura, Y., Hashimoto, T., et al. 2020, *MNRAS*, 493, 4294
- Bakx, T. J. L. C., Zavala, J. A., Mitsunashi, I., et al. 2023, *MNRAS*, 519, 5076
- Battisti, A. J., da Cunha, E., Grasha, K., et al. 2019, *ApJ*, 882, 61
- Behroozi, P., Arnroy, C., Wechsler, R. H., et al. 2020, *MNRAS*, 499, 5702
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- B  thermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, *A&A*, 643, A2
- Boogaard, L., Meyer, R. A., & Novak, M. 2021, <https://doi.org/10.5281/zenodo.5775603>
- Bowler, R. A. A., Jarvis, M. J., Dunlop, J. S., et al. 2020, *MNRAS*, 493, 2059
- Calvi, V., Trenti, M., Stiavelli, M., et al. 2016, *ApJ*, 817, 120
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
- Carnall, A. C., McLure, R. J., Dunlop, J. S., & Dav  , R. 2018, *MNRAS*, 480, 4379
- Castellano, M., Fontana, A., Treu, T., et al. 2022, *ApJ*, 938, L15
- Ceverino, D., Klessen, R. S., & Glover, S. C. O. 2018, *MNRAS*, 480, 4842
- Curtis-Lake, E., Carniani, S., Cameron, A., et al. 2022, *ApJ*, submitted [arXiv:2212.04568]
- da Cunha, E., Walter, F., Smail, I. R., et al. 2015, *ApJ*, 806, 110
- De Looze, I., Baes, M., Cormier, D., et al. 2017, *MNRAS*, 465, 3741
- Donnan, C. T., McLeod, D. J., Dunlop, J. S., et al. 2023, *MNRAS*, 518, 6011
- Feldmann, R., Gnedin, N. Y., & Kravtsov, A. V. 2012, *ApJ*, 747, 124
- Ferland, G. J., Chatzikos, M., Guzm  n, F., et al. 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385
- Fixsen, D. J., Bennett, C. L., & Mather, J. C. 1999, *ApJ*, 526, 207
- Fujimoto, S., Finkelstein, S. L., Burgarella, D., et al. 2022, *ApJ*, submitted [arXiv:2211.03896]
- Ginolfi, M., Piconcelli, E., Zappacosta, L., et al. 2022, *Nat. Commun.*, 13, 4574
- Graziani, L., Schneider, R., Ginolfi, M., et al. 2020, *MNRAS*, 494, 1071
- Harikane, Y., Ouchi, M., Inoue, A. K., et al. 2020, *ApJ*, 896, 93
- Harikane, Y., Inoue, A. K., Mawatari, K., et al. 2022, *ApJ*, 929, 1
- Harikane, Y., Ouchi, M., Oguri, M., et al. 2023, *ApJS*, 265, 5
- Hashimoto, T., Laporte, N., Mawatari, K., et al. 2018, *Nature*, 557, 392
- Hatziminaoglou, E., Zwaan, M., Andreani, P., et al. 2015, *The Messenger*, 162, 24
- Jiang, L., Kashikawa, N., Wang, S., et al. 2021, *Nat. Astron.*, 5, 256
- Kennicutt, R. C., Jr. 1998, *ARA&A*, 36, 189
- Kohandel, M., Ferrara, A., Pallottini, A., et al. 2023, *MNRAS*, 520, L16
- Laporte, N., Ellis, R. S., Boone, F., et al. 2017, *ApJ*, 837, L21
- Liu, D., Gao, Y., Isaak, K., et al. 2015, *ApJ*, 810, L14
- Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, *A&A*, 532, A90
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in *CASA Architecture and Applications*, eds. R. A. Shaw, F. Hill, & D. J. Bell, *ASP Conf. Ser.*, 376, 127
- McQuinn, K. B. W., Berg, D. A., Skillman, E. D., et al. 2020, *ApJ*, 891, 181

- Naidu, R. P., Oesch, P. A., Dokkum, P. V., et al. 2022a, *ApJ*, **940**, L14
- Naidu, R. P., Oesch, P. A., Setton, D. J., et al. 2022b, *ApJ*, submitted [arXiv:2208.02794]
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2014, *ApJ*, **786**, 108
- Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, *ApJ*, **819**, 129
- Oliver, S. J., Bock, J., Altieri, B., et al. 2012, *MNRAS*, **424**, 1614
- Pallottini, A., Ferrara, A., Gallerani, S., et al. 2022, *MNRAS*, **513**, 5621
- Petry, D., Stanke, T., Biggs, A., et al. 2020, *The Messenger*, **181**, 16
- Popping, G. 2023, *A&A*, **669**, L8
- Ramasawmy, J., Klaassen, P. D., Cicone, C., et al. 2022, in Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy XI, eds. J. Zmuidzinas, & J. R. Gao, *SPIE Conf. Ser.*, **12190**, 1219007
- Roberts-Borsani, G. W., Bouwens, R. J., Oesch, P. A., et al. 2016, *ApJ*, **823**, 143
- Roberts-Borsani, G. W., Ellis, R. S., & Laporte, N. 2020, *MNRAS*, **497**, 3440
- Robertson, B. E., Tacchella, S., Johnson, B. D., et al. 2022, *ApJ*, submitted [arXiv:2212.04480]
- Salim, S., Boquien, M., & Lee, J. C. 2018, *ApJ*, **859**, 11
- Sawicki, M. 2012, *PASP*, **124**, 1208
- Stefanon, M., Labbé, I., Bouwens, R. J., et al. 2019, *ApJ*, **883**, 99
- Tamura, Y., Mawatari, K., Hashimoto, T., et al. 2019, *ApJ*, **874**, 27
- Vanzi, L., Combes, F., Rubio, M., & Kunth, D. 2009, *A&A*, **496**, 677
- Vio, R., & Andreani, P. 2016, *A&A*, **589**, A20
- Vio, R., Vergès, C., & Andreani, P. 2017, *A&A*, **604**, A115
- Vio, R., Andreani, P., Biggs, A., & Hayatsu, N. 2019, *A&A*, **627**, A103
- Walter, F., Decarli, R., Aravena, M., et al. 2016, *ApJ*, **833**, 67
- Yan, H., Ma, Z., Ling, C., Cheng, C., & Huang, J.-S. 2023, *ApJ*, **942**, L9
- Yoon, I., Carilli, C. L., Fujimoto, S., et al. 2022, *ApJ*, submitted [arXiv:2210.08413]
- Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., Popping, G., & Davé, R. 2019, *MNRAS*, **483**, 2983