Juodžbalis, I, Conselice, CJ, Singh, M, Adams, N, Ormerod, K, Harvey, T, Austin, D, Volonteri, M, Cohen, SH, Jansen, RA, Summers, J, Windhorst, RA, D'Silva, JCJ, Koekemoer, AM, Coe, D, Driver, SP, Frye, B, Grogin, NA, Marshall, MA, Nonino, M, Pirzkal, N, Robotham, A, Ryan Jr, RE, Ortiz III, R, Tompkins, S, Willmer, CNA and Yan, H

 EPOCHS VII: discovery of high-redshift (6.5 (6.5 < z <12) AGN candidates in JWST ERO and PEARLS data

Article

For more information please contact researchonline@ljmu.ac.uk

EPOCHS VII: discovery of high-redshift ($6.5 < z < 12$) AGN candidates in JWST ERO and PEARLS data

Ignas Juodžbalis [1]★ Christopher J. Conselice,[1] Maitrayee Singh,[1] Nathan Adams [1],

Katherine Ormerod,[1] Thomas Harvey [1], Duncan Austin,[1] Marta Volonteri,[2] Seth H. Cohen,[3] Rolf A. Jansen,[3] Jake Summers,[3] Rogier A. Windhorst,[3] Jordan C. J. D'Silva,[4,5] Anton M. Koekemoer,[6] Dan Coe,[6,7,8] Simon P. Driver,[4] Brenda Frye,[9] Norman A. Grogin,[6] Madeline A. Marshall,[5,10] Mario Nonino,[11] Nor Pirzkal,[6] Aaron Robotham,[4] Russell E. Ryan Jr,[6] Rafael Ortiz III,[3] Scott Tompkins,[3] Christopher N. A. Willmer[9] and Haojing Yan[12]

[1] *Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[2] *Institut d'Astrophysique de Paris, Sorbonne Université, CNRS, UMR 7095, 98 bis bd Arago, F-75014 Paris, France*
[3] *School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA*
[4] *International Centre for Radio Astronomy Research (ICRAR) and the International Space Centre (ISC), The University of Western Australia, M468, 35 Stirling Highway, Crawley, WA 6009, Australia*
[5] *ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Mount Stromlo Rd, Australian Capital Territory, 2611 AU, Australia*
[6] *Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[7] *Association of Universities for Research in Astronomy (AURA) for the European Space Agency (ESA), STScI, Baltimore, MD 21218, USA*
[8] *Center for Astrophysical Sciences, Department of Physics and Astronomy, The Johns Hopkins University, 3400 N Charles St. Baltimore, MD 21218, USA*
[9] *Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ, 85721-0009, USA*
[10] *National Research Council of Canada, Herzberg Astronomy & Astrophysics Research Centre, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada*
[11] *INAF-Osservatorio Astronomico di Trieste, Via Bazzoni 2, 34124 Trieste, Italy*
[12] *Department of Physics and Astronomy, University of Missouri, Columbia, MO 65211, USA*

ABSTRACT

We present an analysis of a sample of robust high-redshift galaxies selected from the 'blank' fields of the Prime Extragalactic Areas for Reionization Science (PEARLS) survey and Early Release Observations (ERO) data from JWST with the aim of selecting candidate high-redshift active galactic nuclei (AGN). Sources were identified from this parent sample using a threefold selection procedure, which includes spectral energy distribution (SED) fitting to identify sources that are best fitted by AGN SED templates, a further selection based on the relative performance of AGN and non-AGN models, and finally morphological fitting to identify compact sources of emission, resulting in a purity-oriented procedure. Using this procedure, we identify a sample of nine AGN candidates at $6.5 < z < 12$, from which we constrain their physical properties as well as measure a lower bound on the AGN fraction in this redshift range of 5 ± 1 per cent. As this is an extreme lower limit due to our focus on purity and our SEDs being calibrated for unobscured Type 1 AGN, this demonstrates that AGN are perhaps quite common at this early epoch. The rest-frame UV colours of our candidate objects suggest that these systems are potentially candidate obese black hole galaxies (OBGs). We also investigate *Chandra* and VLA maps of these areas from which we calculate detection limits. Of note is a $z = 11.9$ candidate source exhibiting an abrupt morphological shift in the reddest band as compared to bluer bands, indicating a potential merger or an unusually strong outflow.

Key words: galaxies: active – galaxies: high-redshift – quasars: supermassive black holes.

1 INTRODUCTION

The origin and evolution of supermassive black holes remains an active area of research in astrophysics. One of the major problems is that predicted masses for black hole seeds, which are expected to form from population III stars at $z = 10 - 50$ (Gao et al. 2007), are too low to grow to the sizes observed at lower redshifts. For

example, $z = 7.5$ quasars host black holes with masses in excess of $10^9 M_\odot$ (Yang et al. 2020b), which are difficult to form unless super-Eddington accretion takes place (Pezzulli, Valiante & Schneider 2016). Other formation models, such as the direct collapse of gas clouds, stellar collisions in dense clusters, and collapsing primordial density fluctuations, similarly lack conclusive observational evidence to distinguish them from one another as summarized in Volonteri, Habouzit & Colpi (2021).

A few candidate direct collapse black holes have been identified to date, pre-JWST. This includes candidates identified by Pacucci

★ E-mail: ignas.juodzbalis@gmail.com

et al. ([2016](#)) in the GOODS-S region of the CANDELS survey (Grogin et al. [2011](#); Koekemoer et al. [2011](#)), which are very faint objects. Thus, a next generation instrument, with increased survey depth is likely to identify more of such candidates (Nabizadeh et al. 2023 in preparation). The recent launch of the JWST has given us such an instrument and presents an excellent opportunity to start investigating the formation of young central massive black holes and start testing the validity of current black hole formation models by direct observations. While most black hole seeds are expected to have formed between redshifts of $z = 14$ and $z = 30$ (Volonteri, Haardt & Madau [2003](#); Yue et al. [2014](#)), which lies somewhat beyond the expected capabilities of the telescope, JWST may be able to detect accreting black hole seeds up to $z = 10 - 12$ (Trinca et al. [2022](#)). It is also important to note that some black hole seed formation models, for instance Bellovary et al. ([2011](#)) and Trenti, Stiavelli & Shull ([2009](#)) predict their formation to occur, albeit at a significantly diminished pace up to the end of reionization at $z \sim 5$, giving further credence to the idea that JWST surveys may be able to validate our current understanding on the origin of these objects.

Currently, JWST efforts in tracing the evolution of galaxies have been focused largely on the detection of high-redshift star-forming galaxies (Adams et al. ([2023b](#)), Rodighiero et al. ([2023](#)), Finkelstein et al. ([2022](#)), and Naidu et al. ([2022b](#))) and morphological evolution of galaxies (Ferreira et al. ([2022](#)), Treu et al. ([2022b](#)), and Fudamoto, Inoue & Sugahara ([2022](#))). The first year of observation also yielded three spectroscopically confirmed active galactic nuclei (AGN). Two at $z \approx 5$ (Onoue et al. [2022](#); Kocevski et al. [2023](#)) and one at $z \approx 8$ by Larson et al. ([2023](#)), as well as two candidates, one at $z \approx 12$ by Fudamoto et al. ([2022](#)) and one at $z \approx 11$ by Maiolino et al. ([2023](#)). More recent studies by Furtak et al. ([2023](#)); Labbe et al. ([2023](#)); Matthee et al. ([2023](#)) also hint at the presence of a significant population of partially obscured AGN at $z < 7$. However, no search aimed explicitly at searching for high-redshift ($z > 7$) AGN candidates has been carried out in detail using imaging – a combination of both spectral energy distributions and morphology/structure.

Furthermore, finding forming black holes through AGN is an important process that needs to be done photometrically, if possible, to find objects that can be followed up with Near Infrared Spectrograph (NIRSpec) spectroscopy. This will be critical for determining the demographics of early AGN as well as 'naked' black holes that might exist at early times. If these systems can be found photometrically, in principle, this would allow for large samples and demographics of this population to be characterized.

In this work, we present the results for such a search for candidate AGN sources using JWST *Near-Infrared Camera* (NIRCam) imaging data (Rieke et al. [2022](#)). We identify a small sample of excellent candidates for being high-redshift AGN. We discuss in this paper our method for finding these objects and give a description of the properties of these potential early/young AGN or black holes, and provide a pathway to use our methods to find further larger samples of such objects.

This paper is organized as follows – Section [2](#) contains a brief description of the data and the reduction process. Section [3](#) presents a discussion of AGN identification methods used and checks of their validity, Section [4](#) presents an overview of the properties of the selected sources. The findings of this paper are summarized in Section [5](#). Where appropriate we adopt a standard cosmology of $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, all reported magnitudes use the AB system.

2 DATA

2.1 Sample and data reduction

The galaxy sample from which we carry out this analysis derives from the Early Release Observations alongside the PEARLS GTO Survey fields: El Gordo, MACS-0416, and the North Ecliptic Pole Time Domain Field (NEP-TDF) (Windhorst et al. [2023](#)). This data set is from the EPOCHS sample which are, a re-redcution and analysis of these fields to create a homogeneous reduction and ultimately a catalogue, see Adams et al. ([2023a](#)) and Conselice et al. (in preparation). This sample results from processing these data ourselves at all steps using our own refined methods that maximizes our detection of faint galaxies and the accuracy of the photometry. This paper is part VII in this series, with other forthcoming papers on the star formation rate (SFR), stellar populations, morphologies, and stellar masses of this sample. This paper is our first look at finding AGN within distant galaxies and can serve as a first look at how to approach this problem.

Specifically, the NIRCam data used for source detection and photometry originated from the CEERS (PID: 1345, PI: S. Finkelstein, see also Bagley et al. [2022](#)), GLASS (PID: 1324, PI: T. Treu, Treu et al. [2022a](#)), SMACS 0723 (PID: 2736, PI: K. Pontopp-idan, Pontoppidan et al. [2022](#)), and NGDEEP (PID: 2079, PIs: S. Finkelstein, Papovich and Pirzkal, Bagley et al. [2023](#)) public surveys. We also include GTO targets of El-Gordo, MACS 0416, and NEP from the Prime Extragalactic Areas for Reionization Science (PEARLS) survey (PI: R. Windhorst & H.Hammel, PID: 1176 & 2738, (Windhorst et al. [2023](#))).

NIRCam filter sets used by these surveys were largely similar, with all of them utilizing some combination of the seven wide (F090W, F115W, F150W, F200W, F277W, F356W, and F444W) and one medium width (F410M) filter.

The reduction procedure applied to all unprocessed JWST data products is described in detail by Adams et al. ([2023a](#)) and in Conselice et al. (in preparation), and can be summarized as follows: (i) initial processing is carried out using version 1.8.2 of the JWST pipeline (Bushouse et al. [2022](#)) and v0995 of the Calibration Reference Data System (CRDS), which were most recent in the first half of 2023. (ii) Wisps and artefacts from F150W and F200W are subtracted using a template set in between stages 1 and 2 of the pipeline. (iii) 1/f noise correction, derived by Chris Willott,[1] is applied after stage 2 of the pipeline. (iv) A 2D sky subtraction is run on each of the individual calibrated frames before stacking in stage 3 of the pipeline. (v) After stage 3 of the pipeline, the final F444W image was matched to a *Gaia*-derived WCS using available *Gaia* stars in the NIRCam imaging, and all other images in the other bands were then aligned with the new F444W image to ensure consistent source positions. The processed images have a final resolution of 0.03 arcsec pixel^{-1}.

2.2 Initial high-redshift catalogue construction

Source detection and measurement from the processed science images was carried out using SEXTRACTOR Bertin & Arnouts ([1996](#)), with the key configuration parameters taken from Table [1](#) in Adams et al. ([2023b](#)). We used the F444W band for initial source detection. Using this fluxes of the detected sources were then measured in each band using 0.32 arcsec diameter apertures, with corrections applied

---

[1] https://github.com/chriswillott/jwst

Table 1. A summary of average 5σ detection depths of each NIRCam filter for the surveys examined in this work. It should be noted that due to bright foreground objects and edge noise effects these values varied by up to 1 mag across the images and thus were examined individually for each candidate object. Bands that did not have data for a particular field are marked with '–'.

Band	CEERS	GLASS	NGDEEP	SMACS-0723	MACS 0416	El Gordo	NEP
F090W	–	28.90	–	28.80	28.70	28.30	28.50
F115W	28.75	28.95	29.65	–	28.65	28.30	28.50
F150W	28.60	28.50	29.50	28.75	28.50	28.15	28.60
F200W	28.80	28.80	29.50	28.90	28.65	28.40	28.70
F277W	28.95	29.15	29.80	29.25	28.95	28.80	29.00
F356W	29.05	29.30	29.75	29.40	29.10	28.85	29.10
F410M	28.35	–	–	–	28.65	28.35	28.45
F444W	28.60	29.50	29.55	29.15	28.90	28.70	28.75

derived from Point Spread Function (PSF) models taken from Perrin et al. (2012). Detection depths were calculated individually for each source by placing 0.32 arcsec diameter apertures in empty spaces of the image, then picking 200 nearest apertures for each source and calculating the total background rms across all of them. The 5σ detection depth was estimated as 5 times the background rms in magnitudes. A summary of average 5σ depths is provided in Table 1. This is further described in Adams et al. (2023a) and Conselice et al. (in preparation).

Initial source redshifts were constrained by photometrically fitting the SEXTRACTOR catalogue sources with both the LEPHARE and EAZY Brammer, van Dokkum & Coppi (2008) codes. All detected sources were run through these SED fitting tools in order to provide a photometric redshift estimate. Both EAZY and LEPHARE were run with a minimum 5 per cent flux error to account for uncertainties in the calibration of the NIRCam detector.

The LEPHARE code was run using the BC03 stellar population synthesis (SPS) template set (Bruzual & Charlot 2003) with both exponentially decaying and constant star formation histories (SFHs). We include templates with 10 characteristic time-scales between $0.1 < \tau < 30$ Gyr, and 57 different ages between 0 and 13 Gyr, with fixed metallicities $Z = \{0.2, 1.0\} \, Z_\odot$. The redshift range is allowed to vary between $0 < z < 25$, and dust extinction is varied from $0 < E(B - V) < 3.5$ to account for potential dusty lower-z contaminants (e.g. Naidu et al. 2022a; Zavala et al. 2023). Attenuation from the inter-galactic medium (IGM) follows the treatment derived in (Madau 1995) and nebular line emission is modelled internally by LEPHARE.

EAZY was run with the 12 default Flexible Stellar Population Synthesis (fsps) templates (tweak_fsps_QSF_v12_v3), which model a range of stellar ages, dust extinction, and metallicities (Conroy & Gunn 2010), along with 6 additional templates from Larson et al. (2022). These templates were designed to better reproduce the blue colours and β slopes of high-z galaxies. Some high-z galaxies have been shown to have high-equivalent width (EW) emission lines, which are more accurately modelled by the FSPS templates, which can boost photometric measurements by as much as a magnitude. This EAZY template set will be referred to as FSPS+Larson hereafter.

The selection criteria for constructing the robust high-redshift galaxy catalogue can be summarized as follows, although see Conselice et al. (in preparation) and (Adams et al. 2023a):

(i) The candidate object must have a higher than 5σ detection in the first two bands redwards of the fitted Lyman break and $< 3\sigma$ detections in bands containing the break.

(ii) The integrated probability density function (PDF) within 10 per cent of the best fit redshift must contain at least 60 per cent of the probability.

(iii) Less than 10 per cent of the PDF must lie in the $z < 5$ range; secondary peak solutions, if present, must have a maximum lower than 50 per cent of the primary peak to avoid Lyman–Balmer break confusion.

(iv) Candidate must have $\chi^2_R < 6$ to be considered 'good' or $\chi^2_R < 3$ for a 'robust' classification.

(v) The PDF criteria must be satisfied by both codes and their photometric redshifts have to be consistent within 3σ.

The above procedure is discussed in depth by Adams et al. (2023a).

A total of 214 high-redshift ($6.5 < z < 12$) sources were identified using our criteria and were further analysed for the presence of an AGN component. It should be noted here that the lowest redshift available in CEERS and NGDEEP surveys was 8.5 instead as these fields use F115W as the bluest available band and we did not incorporate the *Hubble Space Telescope (HST)* imaging into our selection.

3 CANDIDATE AGN SELECTION

In the following section, we describe how our AGN candidates were found using our methods based on the full EPOCHS sample described earlier. This involves several step including the initial discovery of the objects through a series of photometric redshift codes and tests see (Adams et al. 2023a) and the previous section for further details. We then carry out a further analysis examining the likelihood that these systems are dominated by emission from black holes to construct our final sample.

3.1 SED fitting

This work seeks to identify robust candidate AGN. The easiest ones to identify using imaging are those of the unobscured, Type 1, variety, where the immediate surroundings of the black hole are capable of outshining its host galaxy. Our methods are designed to identify these particular AGN candidates and subsequently are not expected to produce a complete sample of the AGN population in the data covered by the EPOCHS sample.

We begin our selection by refitting our sources from the robust galaxy catalogue using EAZY with a set of SED templates for direct collapse black hole (DCBH) hosts from Nakajima & Maiolino (2022) added on top of the FSPS+Larson set (see Section 2.2). These templates are tuned for unobscured, intermediate mass ($10^5 - 10^6 \, \mathrm{M_\odot}$) active black holes, which may reasonably be expected to make up a
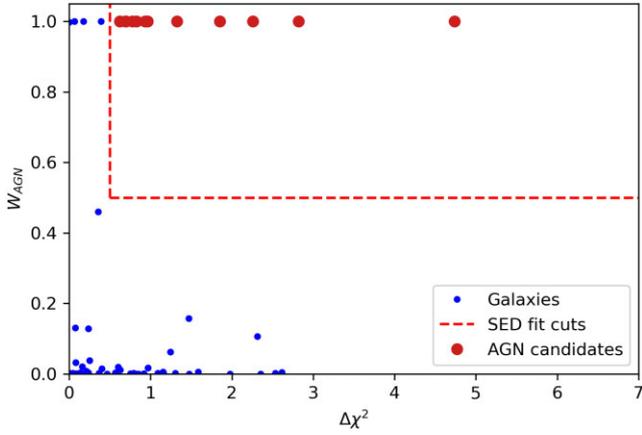
Figure 1. Graph illustrating the separation of our pre-selected candidates from the rest of the sample in the parameter space. The weights of the Nakajima templates (W_{AGN}) are plotted on the y-axis, the *x*-axis shows the difference between the χ_R^2 values given by the Nakajima and FSPS+Larson sets ($\Delta\chi^2$). Only sources with $\chi_R^2 < 3$ given by the Nakajima templates are plotted.

significant fraction of high-redshift AGN. This gives us an AGN+star formation set of templates from which we can find galaxies that match various combinations of these templates. The continuum shape of these SEDs is characterized chiefly by their UV power-law slopes (α), and the so called Big-Bump temperatures (T_{bb}). We adopt the full range of values for both, with $\alpha = -1.2, -1.6, -2.0$, and $T_{bb} = 5 \times 10^4, 1 \times 10^5, 2 \times 10^5$ K. We fix the ionization parameter $\log U$ to –0.5 and metallicity Z to 0.014 as reasonable choices for pristine, high-redshift environments (Sarmento, Scannapieco & Pan 2017). These parameters are otherwise difficult to constrain using SED fitting. Thus, the additional set of SEDs consists of nine templates with the aforementioned parameters and will be referred to as the 'Nakajima' set hereafter.

After re-fitting, we derive the weighting for each template in the fit via the following equation

$$W_i = a_i \left(\sum_j a_j \right)^{-1}, \tag{1}$$

where a_i is the linear coefficient of the i-th template as defined in Brammer et al. (2008). From these, we define the W_{AGN} parameter as $W_{AGN} = \sum W_i$ for all AGN templates in the set. This parameter thus serves as an indication of the relative weight of AGN versus non-AGN templates in the best fit for each source.

Sources were then selected according to the following criteria:

(i) $\chi_R^2 < 3$, consistent with the 'robust' classification from Section 2.2.

(ii) Nakajima templates having $W_{AGN} > 0.5$ in the fit, ensuring that a candidate is mostly fitted by an AGN model.

(iii) The new χ_R^2 value is lower by at least 0.5 than the one given by FSPS+Larson set to ensure that the fit improvement provided by adding the AGN models is not the result of an expanded parameter space.

(iv) Redshift given by the Nakajima templates consistent with other redshift estimates as the location of the Lyman break should be insensitive to the nature of emission.

The above procedure resulted in 12 sources being selected from the initial sample. This selection is illustrated by Fig. 1. The AGN

candidates are strongly separated from the rest of the sample along the W_{AGN} axis, with most sources concentrated either at 1 or 0. This is likely because EAZY prefers single template models instead of mixed templates. Therefore, this parameter does not necessarily correspond to a physical AGN fraction, but it remains useful for further selection of strong candidates.

Afterwards, as part of our refined method for finding AGN, we then use CIGALE (Yang et al. 2020a) to narrow our selection. Before fitting our SEDs to these templates we increase the SEXTRACTOR measured flux errors such that they have a floor, or lower limit values, which represent a 5 per cent error. We then set the 3σ upper limits with 1σ errors in bands which contain fluxes consistent with a 3σ background fluctuation. Upper limits of 1σ with 1σ error were set in bands containing negative flux measurements.

We modelled the AGN component of the CIGALE templates using the SKIRTOR continuum emission models (Stalevski et al. 2016) with a varying rest-frame UV slope (α) as the key shape parameter. This was chosen to largely overlap with the α values from the Nakajima set. The allowed viewing angles were $30°$ and $70°$ in order to consider both obscured and unobscured emissions. The stellar emission was modelled using an initial mass function from Salpeter (1955) together with the BC03 templates, and a delayed exponential SFH, with stellar ages ranging from 500 to 100 Myr, and $0.5 < \tau < 2$ Gyr. This range is narrower than the one used with LEPHARE due to the need to simplify the parameter space to allow for more AGN models and the high-redshift nature of the fitted galaxies already being confirmed by the previous selection steps.

The stellar and gas metallicities in our fit were sampled from the range $Z = \{0.2, 1.0\} Z_\odot$, consistent with what was used within the LEPHARE fitting. The nebular emission was modelled with the ionization parameter using values of –1.0, –1.5, and –2.0. The dust extinction for the stellar component was modelled using the Calzetti dust attenuation law, assuming $0 < E(B - V) < 0.9$. AGN polar dust extinction was assumed to follow the SMC curve, taken from Prevot, M. L. et al. (1984), with extinction values ranging from $0 < E(B - V) < 0.8$. The values not listed were kept to CIGALE defaults.

The relative performance of AGN versus non-AGN models was established by running CIGALE with two groups of SED templates, the first one with $f_{AGN} = 0$, while the second ranges from $0.1 < f_{AGN} < 1$. This f_{AGN} parameter quantifies the ratio of observed infrared luminosity of the AGN component to the total observed infrared luminosity of the source. The average performance of the two template sets was quantified using a parameter P_{AGN}:

$$P_{AGN} = \frac{N_{AGN}\left(\chi_R^2 < \chi_{lim}^2\right)}{N_{GAL}\left(\chi_R^2 < \chi_{lim}^2\right)} \times \frac{N_{GAL}}{N_{AGN}}, \tag{2}$$

where $N_{AGN}(\chi_R^2 < \chi_{lim}^2)$ is the number of AGN models giving $\chi_R^2 < \chi_{lim}^2$ and $N_{GAL}(\chi_R^2 < \chi_{lim}^2)$ is the number of galaxy models satisfying the same criterion, N_{AGN} and N_{GAL} are the total number of AGN and galaxy templates fitted. We use this ratio to normalize the number of models as otherwise one type would dominate over the other. In cases where $N_{GAL}(\chi_R^2 < \chi_{lim}^2) = 0$ and $N_{AGN}(\chi_R^2 < \chi_{lim}^2) \neq 0$, P_{AGN} was set to equal 99, if no models gave $\chi_R^2 < \chi_{lim}^2$, P_{AGN} was set to equal 0. Sources with $P_{AGN} > 1$ were selected for further morphological and structural analysis.

The value of χ_{lim}^2 was fixed by using different values χ_{lim}^2 for equation (2) to classify a sample of known AGN and likely non-AGN sources, minimizing the number of misclassifications. For the known AGN sample we use three objects in total – the two spectroscopic AGN from Kocevski et al. (2023), CEERS 1670 and CEERS 3210, and one from Larson et al. (2023), CEERS 1019, as
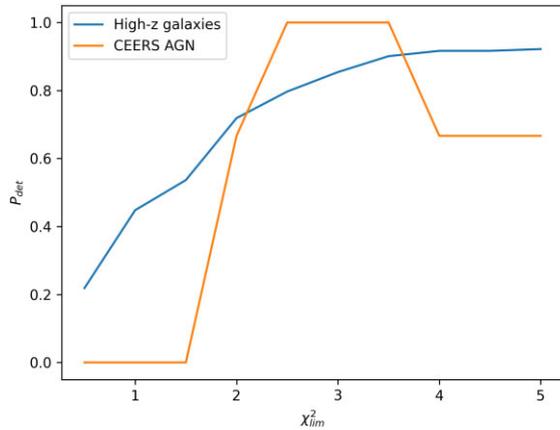
Figure 2. A graph showing the dependence on the probability of classifying a source as an AGN (P_{det}), and the value of χ^2_{lim} for the samples of CEERS AGN (orange line) and high-z galaxies (blue line). It can be seen that $\chi^2_{lim} = 2.5$ gives the maximum chance of correctly classifying an AGN while minimizing the chance of incorrect classification. The probability of incorrect classification remains relatively high, however, at 80 per cent.

these low mass sources are more likely to be representative of the ultra high-redshift AGN population. None of these sources end up in our robust catalogue due to them having $>3\sigma$ detections in the F115W band where the Lyman-break is located, causing them to fail one of the robust redshift criteria in Section 2.2. In fact objects such as these objects require *HST* data for reliable classification to ensure that there is a Lyman-break and not a Balmer-break within the 'drop' filter. However, we use our measured photometry from the original SEXTRACTOR catalogue for fitting this object. We also note that CEERS 3210 is obscured, while CEERS 1670 is a classic, more evolved, Type 1 AGN, thus the Nakajima templates, calibrated for AGN hosted by pristine early environments, do not reproduce their photometry well. The CEERS 1019 object, while having a relatively flat continuum has a strong OIII line visible in the F444W band, which is not captured well by any SED templates used in our fitting. This results in a fit which does not imply an AGN as the templates we use do not have line emission this strong. This reveals that even stronger line emission from AGN should be implemented in future AGN template models. This source also has a continuum strongly influenced by stellar emission, see discussion in Larson et al. (2023).

The non-AGN high-redshift galaxy sample was taken from the original robust galaxy catalogue by removing all galaxies that satisfied the EAZY selection criterion and were not classified by us as AGN using our methods. We run these galaxies through our procedure and the results of this are shown in Fig. 2 as the blue line. We find that some of these galaxies do have a high AGN fraction, and thus it remains possible, if not likely, that some of these systems are in fact AGN. However, using our methods we are more certain to find a pure selection of AGN as also shown by the orange line.

As can be seen in Fig. 2, $\chi^2_{lim} = 2.5$ gives the highest probability of correctly classifying an AGN, however, 80 per cent of the remaining galaxy sample has $P_{AGN} > 1$, and while some of these sources may harbour obscured AGN akin to CEERS 3210, such a high AGN fraction is unlikely as the fraction of dust reddened AGN at $4 < z < 7$ was estimated to be about 10 per cent by Harikane et al. (2023). Therefore, this method can only be used to exclude dubious sources as its purity is too low for standalone use. Nevertheless, a further two sources are excluded from the sample by this method.

This is our main method for finding AGN. It is however important to note that this is not a unique method, and other methods using photometry and SED fitting can be developed. However, our method does provide a way for finding a sample with a high probability of being AGN. It is also important to note that the template set with which we constrain most of our sources was produced using unobscured AGN models, thus our search is inherently biased towards Type 1 AGN in unevolved low metallicity environments.

3.2 Structural measures

In order to improve the purity of our sample, we apply morphological cuts to search for sources containing a point source. To do this, we use GALFIT, a 2D fitting algorithm that uses a Levenberg–Marquardt algorithm to find the optimum solution to a fit (Peng et al. 2002, 2010). We select sources which are best fit by a PSF, a combination of an extended Sersic profile and a PSF, or a single Sersic profile with half light radius less than the full width at half-maximum (FWHM) of the PSF used in the fitting process. We define the best-fitting model as the model with the lowest χ^2_v, which is defined by GALFIT as

$$\chi^2_v = \frac{1}{N_{DOF}} \sum_{x=1}^{n_x} \sum_{y=1}^{n_y} \frac{(f_{data}(x, y) - f_{model}(x, y))^2}{\sigma(x, y)^2}. \quad (3)$$

Within our fitting we use the SEXTRACTOR parameters for each object as the initial guesses, and run GALFIT three times for each object, modelling the source as an extended Sersic profile, a Sersic profile containing a PSF, and as just a pure PSF. For all fits, the ERR extension of the image, which is a measure of the noise of the image, is used as the input sigma image. Sources containing an AGN cannot be modelled accurately by a single Sersic fit, as the AGN appears as a distinct point source. However, determining the structures of sources with angular sizes similar to the PSF of the telescope is difficult and results should be interpreted with some caution (Dewsnap et al. 2023). As a result of this, we also select sources where the determined half light radius is less than the FWHM of the PSF. All object cutouts are from the F444W NIRCam image, where the PSF for this band has a FWHM of 4.83 pixels on our pixel scale, therefore sources with $R_e < 4.83$ pixels are selected as being a point source object. We use the F444W band for our fitting process, as this is closest to the rest-frame optical for each source, and keep this consistent throughout in order to model each source using the same parameters and constraints. Due to the faint magnitudes of these sources, we fix the Sersic index to a value of $n = 1$. Where multiple sources are fit simultaneously, the image positions of all objects are constrained to within ± 3 pixels, to ensure the correct sources are fit. We also visually inspect fits and residual images as a final quality check. An example of each fit is shown in in Fig. 3.

The final structural analysis result in two of our objects being classified as a PSF, five objects being classified as a Sersic profile with $R_e <$ the FWHM of the PSF, and four objects classified as a combined model of an extended Sersic profile containing a PSF. The remaining two objects are not morphologically classified as a pure AGN due to their radius being larger than that of the PSF. These could in fact still be AGN, but we are interested here in systems where the AGN dominates the light of the source. The classification of each object is given in Table 2, and properties, including radii for those fit as Sersic profiles, are given in Table 4.

We further check our results in the F277W band, where most of our sources have higher S/N ratios, and find that our classifications do not change. In particular, we check the sizes of our sources best fit by a single Sersic profile, and find that in general, we recover

(a) NEP-z7l

(b) S0723-z11c

(c) CEERS-z9d
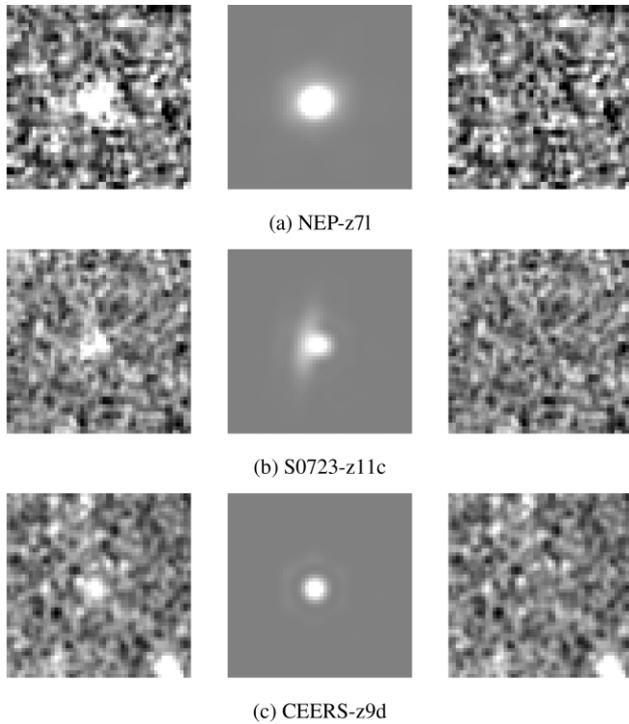
Figure 3. Top: an example morphological fit of NEP-z7l, which is best fit by a Sersic profile, which is compact enough to be classified as an AGN, with R_e < F444W PSF FWHM. Middle: S0723-z11c, best fit by a combined Sersic and PSF model whose centres do not coincide. This system is very unusual in structure and we discuss this object later in this paper. Bottom: CEERS-z9d, best fit by a PSF. The left panels show the data images, the centre panels show the PSF models, and the right panels show the residual images, created by subtracting the model image from the data image. The residual images show these models are good fits, as there are no bright areas of remaining light unaccounted for, or any dark areas over-fit by the model. All cutouts are 1.5' × 1.5'.

Table 2. A table summarizing the morphological classification of each object within our sample. To be classified as a Sersic fit, an object must have R_e < 4.83 pixels, the FWHM of the F444W PSF re-sampled to our pixel scale. Both objects that are not selected are rejected due to their best fit model having R_e larger than the PSF FWHM.

PSF	Sersic	Combined	Not selected
NEP-z9b	NEP-z7l	S0723-z11c	NGD-z8f
CEERS-z9d	NEP-z8ff	CEERS-z6c	NEP-z7m
–	CEERS-z8d	NEP-z6c	–
–	CEERS 1019	NGD-z8e	–
–	CEERS-z8g	–	–

them as having sizes smaller than the FWHM in F277W. We find one source as having $R_{e,F277W} \sim 1.04 \times$ FWHM, which could occur due to noise, or faint extended emission better detected in this band, and as such, we still classify this as a compact Sersic profile, small enough to be a PSF. The only source we do not recover in this way, is CEERS 1019, however this source has a very complex morphology, which we analyse further.

We find that the source discovered in Larson et al. (2023) is classified differently in our methods than in the discovery paper, where the object is found to be three components, with the central component best fit by a combination of a PSF and Sersic profile. Our combined fit of a Sersic profile and PSF has a marginally

higher χ_ν^2, therefore we select this object as a compact Sersic profile, small enough to be classified as a PSF. This source has a complex morphology due to likelihood of it being a merger, and thus we replicate the fitting process completed in the discovery paper, and model the source as a three component model, with two components fit by Sersic profiles, and a central PSF component. We find that this has a lower χ_ν^2 than our model fits, confirming our original findings that this object is compact enough to contain a point source. Our final classification information for the 12 sources selected from Nakajima templates is given in Table 3.

4 AGN SOURCE PROPERTIES

Using our selection procedure we identify a total of nine robust candidate sources out of a sample of 214. We also include the CEERS 1019 source from Larson et al. (2023) for the sake of comparison with our candidates, for a total sample of 10 high-redshift sources. Thus we estimate an AGN fraction at $6.5 < z < 12$ of 5 ± 1 per cent, assuming a Poisson counting error. Due to our investigation focusing on purity rather than completeness as well as being strongly biased towards Type 1 AGN, this value should be viewed as very much a lower bound estimate. This is still consistent with the 1–10 per cent observable AGN fractions derived from the FLARES simulation results by Kuusisto et al 2023 (in prep) and matches well with Harikane et al. (2023) finding of $\sim$5 per cent of galaxies at $4 < z < 7$ hosting low luminosity Type 1 AGN, potentially indicating that the AGN fraction does not evolve much during this epoch.

The f_{AGN} values for our sources were estimated by rerunning CIGALE with the same parameters as in Section 3.1, except the f_{AGN} parameter was varied over the full range of 0 to 1 in steps of 0.1. Physical values of R_e were measured by noting that the pixel scale of the images was 0.03 arcsec, and using angular diameter distances calculated at best fit redshifts, with both GALFIT and redshift errors contributing to the final uncertainties. The values of T_{bb} and α were taken from the best-fitting Nakajima templates, the model grid for these being too sparse to estimate meaningful uncertainties. We also measure the rest-frame absolute UV magnitudes M_{UV} by redshifting the best fit SED to $z = 0$ and convolving it with a top-hat between 1450 and 1550 Å in wavelength space, with the uncertainties provided by redshift errors. Photometric redshifts and their errors were taken from LEPHARE results. All physical properties measured for our candidate sources are presented in Table 4. The best fit SEDs for all our sources are shown in Fig. 4 and Fig. 5.

4.1 X-ray and radio limits

We check if any of our candidate sources present in the NEP and CEERS fields have X-ray detections by matching our final candidate catalogue with *Chandra* deep field point source catalogues from AEGIS-X survey of the Extended Groth Strip, overlapping the CEERS field Laird et al. (2009), and a deep X-ray survey of the JWST NEP field. The matching was carried out using 0.3 arcsec matching radii. However, none of our sources in CEERS and NEP fields appear to have X-ray detections in *Chandra* data. Thus, we use it to estimate upper limits on the full band (0.5 – 10 keV) X-ray luminosity of our sources.

For the AEGIS-X data, we take the 50 per cent completeness limit of 1.30×10^{-15} erg s^{-1} cm^{-2} from Laird et al. (2009) as our limiting flux. For the TDF survey we were able to determine a 3σ detection limit by checking the catalogue for the faintest sources that were detected at 3σ significance. This came out to 6×10^{-6} cps, using a conversion factor of 1 cps $= 2.842 \times 10^{-11}$ erg s^{-1} cm^{-2}, this

Table 3. Classification information for the 12 sources initially selected from fitting the Nakajima template set. The first column lists our catalogue ID, the following columns present the parameters used to classify each source as an AGN: χ^2_{Naka} - the value of χ^2 given by the Nakajima AGN template set, $\Delta\chi^2$ - the difference between best fit χ^2 of the Nakajima and FSPS+Larson template sets, P_{AGN} - the parameter defined in equation (2), χ^2_{GALFIT} - best fit χ^2 given by GALFIT and f_{PSF} - fraction of total flux contained in the compact component, $f_{PSF} = 1$ implies the best fit morphology is a PSF, $f_{PSF} = 0$ implies a Sersic profile best fit and $f_{PSF} < 1$ implies the best fit morphology is a Sersic profile containing a PSF. Column labelled R_e contains the the half-light radii in pixels for sources best fit by a Sersic profile, errors on these are standard GALFIT errors. Where a source is best fit by a combined model, this is taken from the half light radius of the extended Sersic profile, and for those best fit by a PSF, the radius is the upper limit given by the PSF FWHM. Second to last column contains the photometric redshifts given by LEPHARE. The final column indicates if the source was selected as an AGN.

ID	χ^2_{Naka}	$\Delta\chi^2$	P_{AGN}	χ^2_{GALFIT}	f_{PSF}	R_e (px)	z	Selected?
NEP-z7l	0.96	2.8	1.3	5.093	0	2.87 ± 0.66	$8.2^{+0.5}_{-0.3}$	Yes
CEERS-z8i	1.91	0.70	1.6	1.178	0.19 ± 0.02	5.58 ± 0.67	$8.8^{+0.3}_{-0.3}$	Yes
NEP-z9b	1.04	1.9	1.4	10.61	1	<4.83	$9.5^{+0.7}_{-0.7}$	Yes
NEP-z8ff	0.55	2.3	1.4	0.801	0	2.34 ± 0.34	$8.7^{+0.2}_{-0.3}$	Yes
NEP-z7m	1.71	4.7	1.2	1.18	0	5.15 ± 0.44	$7.3^{+0.1}_{-0.1}$	No
NEP-z6c	1.83	0.62	1.2	0.887	0.17 ± 0.02	8.04 ± 1.14	$6.5^{+0.2}_{-0.2}$	Yes
CEERS-z8d	1.74	0.83	1.4	0.592	0	4.64 ± 0.42	$8.8^{+0.2}_{-0.2}$	Yes
S0723-z11c	0.77	0.93	1.3	0.939	0.36 ± 0.03	5.69 ± 1.1	$11.9^{+0.3}_{-0.4}$	Yes
NGD-z8e	2.75	1.3	0	2.273	0.34 ± 0.04	5.74 ± 1.85	$8.9^{+0.4}_{-0.3}$	No
CEERS-z8g	0.14	0.69	1.3	0.977	0	4.14 ± 0.62	$8.8^{+0.3}_{-0.3}$	Yes
NGD-z8f	1.39	0.78	0.9	2.594	0	8.09 ± 2.8	$8.9^{+0.3}_{-0.3}$	No
CEERS-z9d	1.38	0.96	1.5	0.738	1	<4.83	$9.0^{+0.3}_{-0.3}$	Yes

Table 4. A table summarizing the basic properties of our candidate sources. The first three columns contain the catalogue ID, separated from the field name by a dash, and right ascension and declination coordinates. Column 4 contains the photometric redshifts given by LEPHARE for our sources and the spectroscopic redshift for CEERS 1019, the uncertainty of which is not given due to being orders of magnitude lower than others. Columns 5 and 6 list big-bump temperatures and UV slopes given by the Nakajima template. Column 7, contains the best fit AGN fraction given by CIGALE. Radii measured in kpc are given in column 8. For sources best fit with a combined or Sersic model, this is taken from the half light radius of the extended Sersic profile, for those best fit by a PSF the radius is an upper limit given by the PSF FWHM. The last two columns contain absolute UV magnitudes, the one for CEERS 1019 was taken from literature, and X-ray luminosity limits in the 0.5–10 keV band for CEERS and NEP sources, respectively.

ID	RA	Dec.	z	T_{bb} (K)	α	f_{AGN}	R_e (kpc)	M_{UV}	L_X (erg s^{-1})
NEP-z7l	260.86441	65.81502	$8.2^{+0.5}_{-0.3}$	2×10^5	-2.0	0.1 ± 0.3	0.41 ± 0.11	$-19.26^{+0.07}_{-0.07}$	$<10^{43.0}$
CEERS-z8i	214.86438	52.77220	$8.8^{+0.3}_{-0.3}$	2×10^5	-1.6	0.8 ± 0.3	0.76 ± 0.12	$-19.16^{+0.15}_{-0.18}$	$<10^{44.0}$
NEP-z9b	260.68414	65.72350	$9.5^{+0.7}_{-0.7}$	2×10^5	-1.2	0.8 ± 0.3	< 0.624	$-19.50^{+0.11}_{-0.13}$	$<10^{43.2}$
NEP-z8ff	260.80899	65.83811	$8.7^{+0.2}_{-0.3}$	1×10^5	-1.2	0.2 ± 0.3	0.32 ± 0.06	$-19.70^{+0.11}_{-0.13}$	$<10^{43.1}$
NEP-z6c	260.81175	65.84165	$6.5^{+0.2}_{-0.2}$	5×10^4	-2.0	0.1 ± 0.3	1.32 ± 0.23	$-18.97^{+0.12}_{-0.14}$	$<10^{42.8}$
CEERS-z8d	215.01529	52.98497	$8.8^{+0.2}_{-0.2}$	1×10^5	-2.0	0.2 ± 0.3	0.63 ± 0.07	$-19.58^{+0.10}_{-0.11}$	$<10^{44.0}$
S0723-z11c	110.69756	-73.50050	$11.9^{+0.3}_{-0.4}$	2×10^5	-1.6	1.0 ± 0.3	0.63 ± 0.14	$-19.53^{+0.10}_{-0.10}$	$-$
CEERS-z8g	214.96892	52.87178	$8.8^{+0.3}_{-0.3}$	1×10^5	-1.2	0.3 ± 0.3	0.56 ± 0.10	$-19.14^{+0.14}_{-0.16}$	$<10^{44.0}$
CEERS-z9d	214.87038	52.80202	$9.0^{+0.3}_{-0.3}$	2×10^5	-1.2	0.5 ± 0.3	< 0.647	$-19.15^{+0.14}_{-0.16}$	$<10^{44.0}$
CEERS 1019	215.03539	52.89066	$8.68*$	$-$	$-$	0.1 ± 0.3	< 0.662	$-22.07^{+0.05}_{-0.05}$	$<10^{44.2}$

resulted in an upper limit of 1.7×10^{-16} erg s^{-1} for sources in the NEP survey fields. It should be noted that the X-ray catalogue for the NEP field did not have completeness estimates at the time of writing, thus this value may be an underestimate. The calculated X-ray luminosity limits are of order $10^{43} - 10^{44}$ erg s^{-1}. This places our sources at or below the characteristic X-ray luminosity of $\sim 10^{44}$ erg s^{-1} for AGN at $z = 4$–5 (Aird et al. 2015; Fotopoulou et al. 2016). However, our sources have low inferred stellar masses, so we probably would not expect the AGN to belong on the bright end of the luminosity function.

We check for radio detections by matching our candidates in the NEP field with the Very Large Array (VLA) catalogue for the same field, described by Hyun et al. (2023). Using 2 arcsec matching radii, as expected, no matches were found between our candidate sources, thus giving limiting fluxes of 10 μJy for all candidates in the NEP field, based on the flux cutoff in Hyun et al. (2023).

4.2 Near-infrared colours

In order to compare the photometry of our selected candidates with theoretical predictions for DCBHs, we adopt two sets of NIR colour
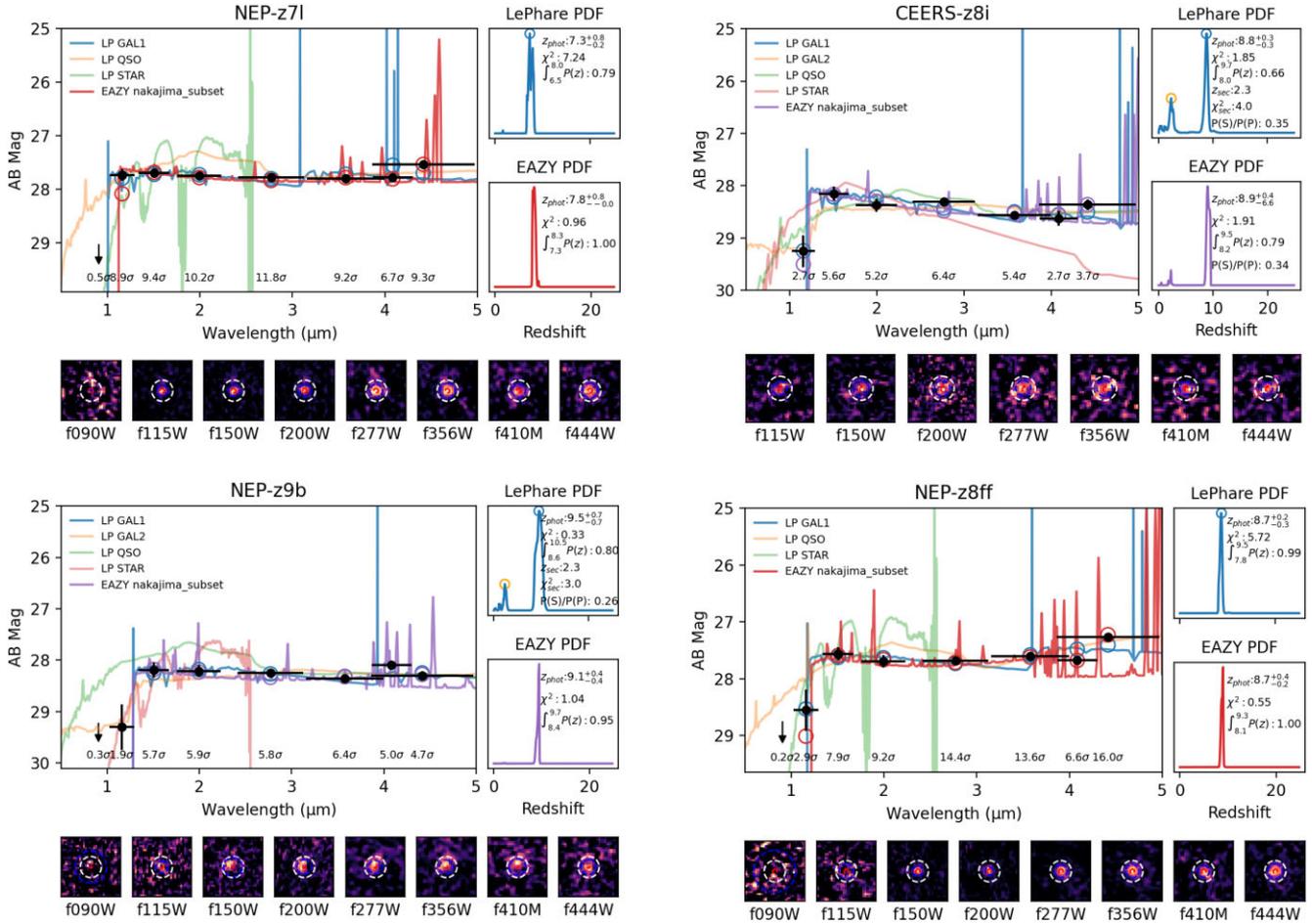
Figure 4. Best fit SEDs for the first 4 of our 9 candidates. The best-fitting LEPHARE galaxy SED is shown in blue, the best fit Nakajima SED in red for sources without a secondary solution identified by LEPHARE and purple for the others. PDFs obtained from both fits are shown to the right. The image cutouts of the source in all measurement bands are presented below the plots. The white circles show the position of the 0.32 arcsec apertures, the blue ones are the half light radii measured by SEXTRACTOR. The top left source, NEP-z7l, stands out as particularly well fitted by an AGN SEDs, with $\Delta\chi^2 > 2$ when compared to both FSPS+Larson and LEPHARE templates, with broad hydrogen line emissions explaining the observed excess in the F444W band. This source is also best fit by a PSF-like morphology. The general trend among our candidates is that broad AGN emission lines tend to better explain the slight excess fluxes in the red bands, leading to better χ^2 values. Image cutouts reveal most sources to either have a compact point-like nature or exhibit signs of a bimodal structure, indicating potential mergers or disturbed morphology.

cuts. The first one consists of the 90 per cent purity cuts for an AGN number fraction (n_{AGN}) of 25 per cent from Table 1 of Goulding & Greene (2022), which were tailored to identify low mass BHs at $7 < z < 10$ accreting at an Eddington ratio of >0.1. The second set was adopted from Natarajan et al. (2017) and was derived for a hypothetical class of obese black hole galaxies (OBGs) at $z \sim 9$, which form after a DCBH acquires a stellar emission component.

Computing the colours using aperture corrected SEXTRACTOR magnitudes in each filter we found that our candidate sources have marginally flatter SEDs than the rest of the high-z galaxy sample, however, the overall colour difference is not substantial, as can be seen in Fig. 7. This same figure also shows that our sources are significantly bluer than the red predictions from Pacucci et al. (2016); Goulding & Greene (2022), likely due to differing assumed SED sets. The key difference seems to stem from Goulding & Greene (2022) assuming an $\alpha = -0.79$, derived by Constantin & Shields (2003) from low redshift narrow-line Seyfert 1 galaxies. We make a further comparison of the Nakajima SEDs with models used in Volonteri et al. (2023), which result in similar predicted colours to Goulding & Greene (2022). These models, hereafter referred to as

the Volonteri set, are explicitly parametrized by the black hole mass (M_{BH}) and the Eddington ratio (f_{edd}) and do not include nebular emission lines, unlike the Nakajima set. A comparison between the bluest and reddest SEDs possible from both model sets in the considered wavelength range is provided in Fig. 6. It can be seen from the figure that running the Volonteri models with lower M_{BH} results in bluer continuum shapes, however, the overall range of apparent slopes of Volonteri models is significantly redder than that of the Nakajima set. The likely reason for this is that the Volonteri models assume an $\alpha = -0.5$, following Thomas et al. (2016). These slopes differ significantly from the steeper values assumed by the Nakajima model, following Elvis, Risaliti & Zamorani (2002) results for low-redshift quasars. Thus a possible reason for our objects not matching the Goulding & Greene (2022) colour cuts is the differing assumptions of the underlying SED models.

It should also be noted, however, that the CEERS 1019 source is likewise not significantly differentiated from either the high-z galaxies or our AGN sample in the Goulding & Greene (2022) colour space. The GN-z11 source, while not in our photometric sample, has also been reported to have a blue ($\beta = -2.26 \pm 0.1$) rest-frame
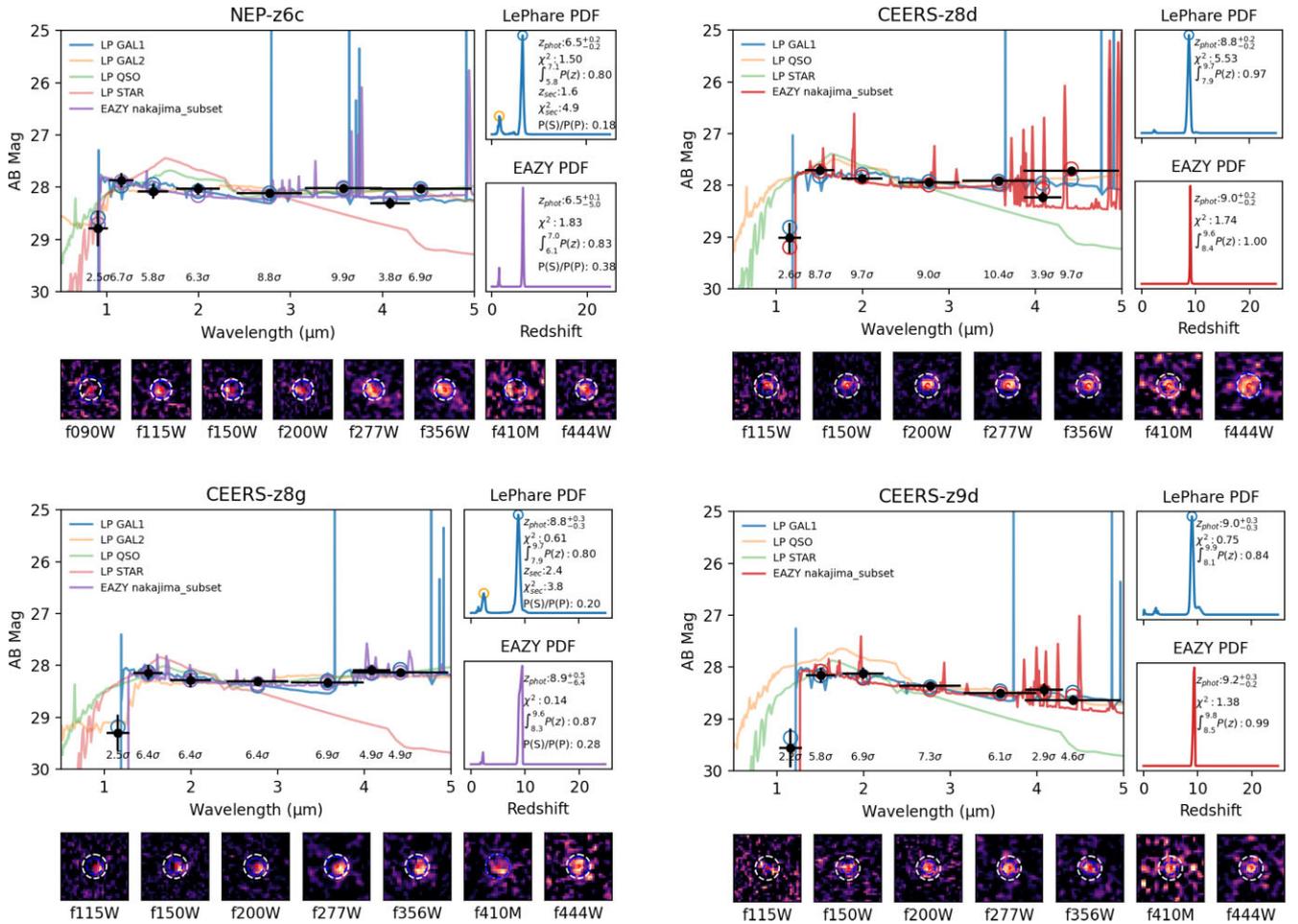
Figure 5. The other 4 candidate sources, excluding, S0723-z11c, which is discussed separately.
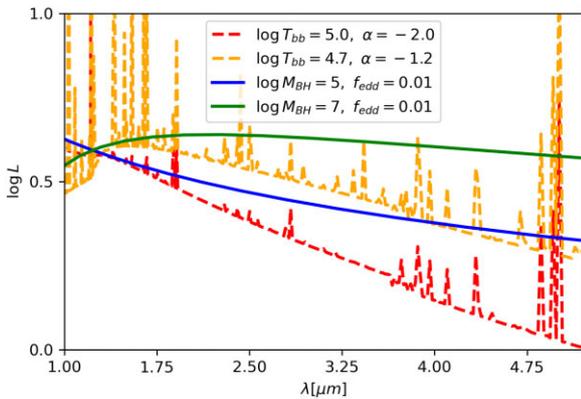

Figure 6. A comparison between the Nakajima SEDs (dashed lines) and AGN continuum SED models from Volonteri et al. (2023) (solid lines). The y-axis is scaled to the same continuum luminosity, all SEDs were redshifted to $z = 9$.

UV slope (Maiolino et al. 2023). These bluer than expected colours may also be partially attributed to some of our sources having a significant stellar component, as suggested by the AGN fractions given by CIGALE in Table 4.

A comparison of the AGN candidates the rest of the galaxies in the (F200W – F444W) colour band (Fig. 8) shows the relative flatness of
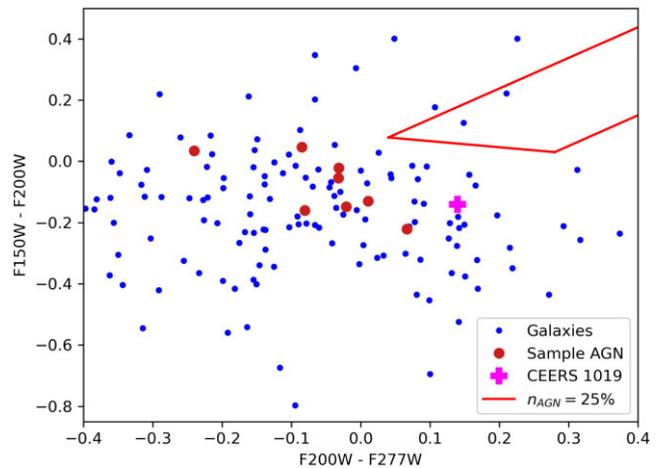
Figure 7. A plot of rest frame UV colours for our selected AGN sources, the remaining galaxies in our sample and CEERS 1019 which is a spectroscopically confirmed AGN at this redshift range Larson et al. (2022). The red lines denote the 90 per cent purity selection region for AGN sources, assuming 25 per cent of all galaxies host an AGN, given by Goulding & Greene (2022). S0723-z11c is excluded from this plot as it starts to drop out in F150W. As can be seen we do not find that our sources are found within this region of the plot.
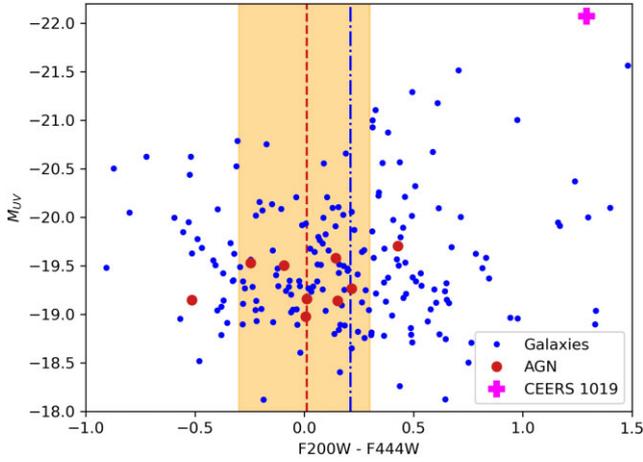
Figure 8. A comparison between our candidate AGN and the rest of the high-z galaxy sample plotted in terms of the (F200W – F444W) colour and the measured UV absolute magnitude. The red dashed line denotes the average colour for our candidate sample – the blue dot dashed line represents the average for galaxies that were not selected. The orange shaded region is the range of predicted OBG colours from Natarajan et al. (2017).

their spectra more clearly, with the average (F200W – F444W) colour being close to 0, while the same average for the high-z galaxies is at ∼0.2. The CEERS 1019 object appears redder in this figure, however, this is due to it possessing a strong O III line above an otherwise flat continuum (Larson et al. 2023). While the derived M_{UV} values do not differentiate our sources from the rest of the sample, it should be noted that 7 out of 9 of our candidates lie in the region –0.3 < F200W – F444W < 0.3, in line with predictions from Natarajan et al. (2017). Our magnitudes, however, are fainter by up to 1 mag than their predictions, assuming an optical bolometric correction $K_O = 5$ (Duras et al. 2020). It should be noted that this correction was derived from AGN at $z < 4$ and may not hold at the redshift range considered here. In general, much more work is needed to understand the SEDs and spectral characteristics of $z > 5$ AGN.

4.3 Masses and start formation rates

We adopt the SFRs from the parent sample of 214 sources. These SFRs were estimated by taking the average flux in the restframe 1450 and 1550 Å wavelength range, using it to calculate the UV luminosity, which, after dust corrections from Meurer, Heckman & Calzetti (1999), is converted into SFR using the procedure described in Madau & Dickinson (2014). Stellar masses were obtained by fitting the sample sources with BAGPIPES Carnall et al. (2018).

The above methods do not take into account potential AGN emissions, however, a comparison between our candidates and the parent sample may be useful in seeing if AGN may be efficiently identified by looking at outliers in an SFR versus stellar mass plot. Such a plot is presented in Fig. 9. As can be seen from the figure, calculating stellar masses and SFRs assuming purely stellar emission does not produce anomalous results for our candidates, likely due to their low UV luminosity. This low luminosity could be the result of low-black hole masses and accretion rates. However, an intriguing possibility is that some of our AGN may be the high-redshift counterparts of sources found by Labbe et al. (2023), which are characterized by a faint and relatively blue UV continuum and a bright, red rest-frame optical SED. At the redshifts considered in this paper the red continuum would mostly lie outside of the NIRcam
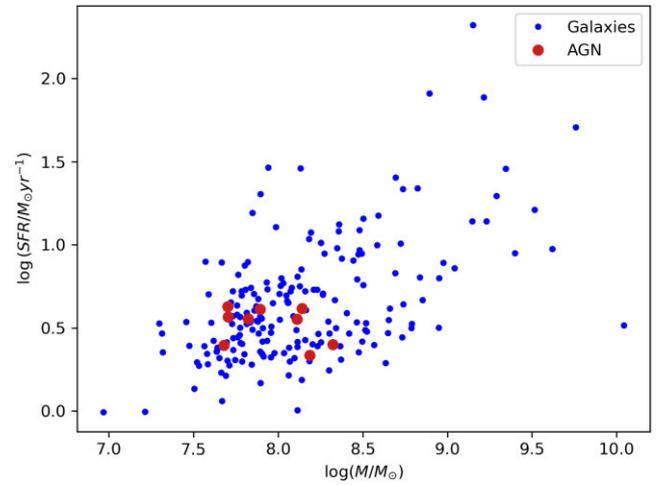
Figure 9. A plot comparing the SFRs and stellar masses, inferred by assuming stellar emissions, of our candidate AGN and the remaining galaxy sample. As can be seen, our sources do not show up as significant outliers in this plot.
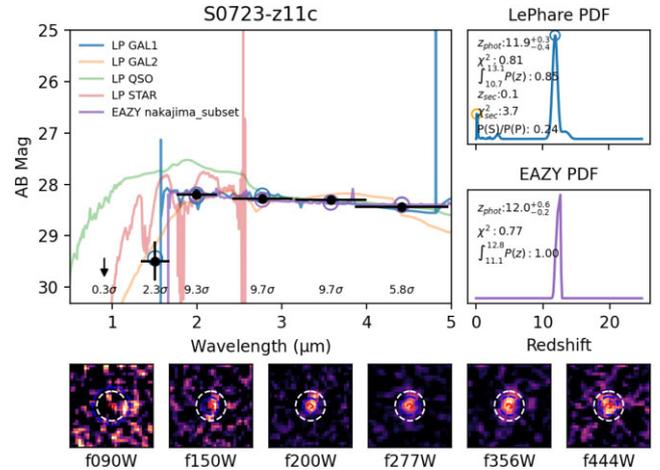

Figure 10. The best fit SEDs for the source S0723-z11c, our highest redshift AGN candidate. The plots are organized as in Fig. 4.

range. Thus deep mid-infrared observations are required to check this hypothesis.

4.4 An unusual object at $z = 12$

In terms of individual sources, S0723-z11c stands out as our candidate at the highest redshift of ∼12. As can be seen from Fig. 10, LEPHARE galaxy models give similar performance to the Nakajima set in terms of χ^2 values, however, the image cutouts presented in the same figure showcase a composite and complex nature of the source. The morphological fits in Fig. 3 identify the second component as a point source, contributing almost 40 per cent to the total emission in the F444W band. However, it is important to note that the apparent morphology changes somewhat drastically in this band with respect to others, as highlighted in Fig. 11. In order to better understand the complex morphology of this source, we run GALFIT across all bands in which the source had >5σ detections (F200W, F277W, F356W, and F444W), with the results summarized in Fig. 11. As can be seen from the figure, the source in each band is best fit by a combination of a Sersic profile and a PSF, with component locations being roughly
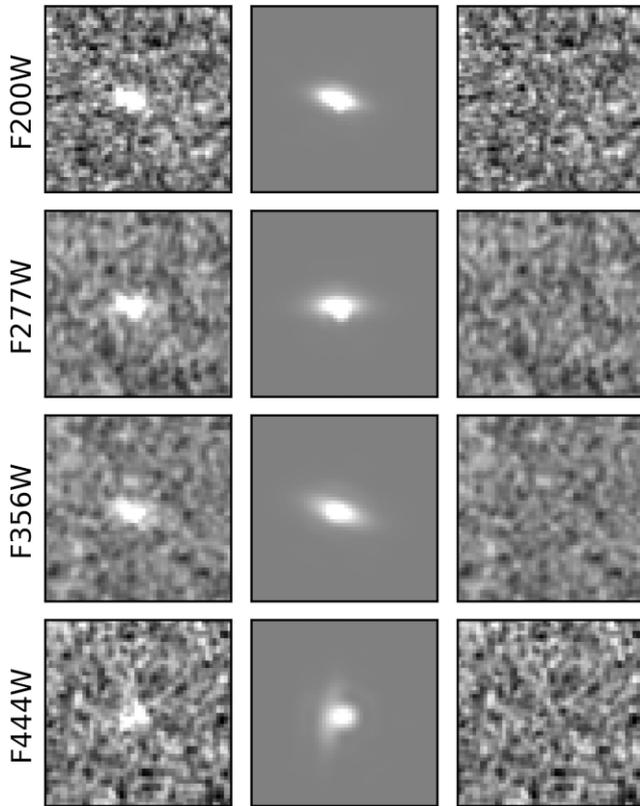
Figure 11. This figure shows the light in different wavebands for the unusual source S0723-z11c. We show a comparison between the actual data image (left), models obtained by GALFIT (centre), and the residual images (right), for S0723-z11c in the different bands in which the source is observed. Each cutout is 1.5' × 1.5'. The source is fit using a combined Sersic and PSF model in all bands.

consistent from F200W to F356W, with a rather abrupt location shift occurring in the F444W band. The consistent bimodal nature of this source along with the shift in the F444W band may point to a morphology disturbed by a merger event or a strong outflow.

A possible reason behind the abrupt nature of the shift between F356W and F444W band images is either line emission or a discontinuity in the continuum itself. Assuming a source redshift of $z = 12$, this emission feature should occur at rest-frame wavelengths of 300–383 nm. While this may be caused by either a NeV (3346 Å, 3426 Å) or OII (3727 Å) doublets, the spatially extended nature of this emission suggests that it may be due to a Balmer break, which in turn would indicate the presence of evolved stellar populations in the object. However, observations probing redwards of the F444W band or a spectroscopic followup is required to truly confirm the nature of this discontinuity in emission patterns.

5 CONCLUSIONS AND SUMMARY

In this paper, we have identified a population of high-redshift AGN candidates by utilizing a set photometric and morphological selection techniques. The basic idea behind our method and paper is to find systems that are much better fit by AGN templates with an active galaxy, or black holes, and are consistent with having a small point source that dominates the light of the galaxy. our methods is not meant to find complete samples of AGN or early black holes, but as a way to find the best candidates for further spectroscopy and detailed follow up.

Our parent samples originates from the EPOCHS sample of $z > 6.5$ galaxies whose luminosity function and selection is discussed in (Adams et al. 2023a). From this sample we refit the galaxies using a variety of model SEDs using the photometric redshift code EAZY. We use this method to isolate most of the sample sources, however, other steps in our pipeline – a combination of analysis of relative performance between AGN and non-AGN models and morphological fitting was also utilized in removing weak candidates. We are thus left with nine strong AGN sources that are likely emitting their light due to a central massive black hole.

We develop a new method of finding likely AGN through template fitting, such that we define a statistic P_{AGN} that reveals how likely an object is better fit by an AGN spectrum rather than a star forming one. It should also be noted that the reason for the P_{AGN} limited performance in isolation may be its implicit reliance on the general properties of the $z > 6$ AGN population being well known. New AGN templates are needed at these highest redshifts which will eventually be developed with the availability of more JWST spectroscopy of these objects. Our overall selection method may, however, provide a good way for searching for the highest redshift candidate AGN for spectroscopic followup using NIRSpec.

We find that the estimated AGN fraction in the interval of $6.5 < z < 12$ is 5 ± 1 per cent. However, our investigation was strongly biased towards Type 1 AGN, due to the initial set of SED templates not accounting for dust extinction, and calibrated for purity rather than completeness, thus this result only establishes a lower bound, which is nevertheless consistent with theoretical predictions.

We also find that rest-frame UV photometry of our candidates suggest that colour–colour cuts alone may not be sufficient to differentiate AGN from other galaxies at high redshifts – with SED and morphological fitting in conjunction with deep X-ray and spectroscopic observations being necessary for robust identification. However, colour–colour cuts may still be useful as a way to pre-select potential candidates, as evidenced by a large fraction of our sources lining up with bluer colours predicted for OBGs.

The presence of JWST-detectable AGN sources out to $z = 12$ alone suggests evidence in favour of the Direct Collapse Black Hole model (Trinca et al. 2022; Volonteri et al. 2023), while the photometric properties of our sample suggest evidence in favour of the OBG stage of galaxy formation and thus a type of 'naked' black holes existing in the early universe, however, follow up spectroscopy will be needed to confirm the nature of our objects and estimating their black hole masses in order to place more defined constraints on black hole seeding models.

HST. The Early Release Observations and associated materials were developed, executed, and compiled by the ERO production team: Hannah Braun, Claire Blome, Matthew Brown, Margaret Carruthers, Dan Coe, Joseph DePasquale, Nestor Espinoza, Macarena Garcia Marin, Karl Gordon, Alaina Henry, Leah Hustak, Andi James, Ann Jenkins, Anton Koekemoer, Stephanie LaMassa, David Law, Alexandra Lockwood, Amaya Moro-Martin, Susan Mullally, Alyssa Pagan, Dani Player, Klaus Pontoppidan, Charles Proffitt, Christine Pulliam, Leah Ramsay, Swara Ravindranath, Neill Reid, Massimo Robberto, Elena Sabbi, Leonardo Ubeda. The EROs were also made possible by the foundational efforts and support from the JWST instruments, STScI planning and scheduling, and Data Management teams. The effort of CEERS, NGDEEP, and GLASS teams in making their data public is hereby acknowledged. We would also like to thank Adi Zitrin, Rachana Bhatawdekar, and Nimish Hathi for their timely and useful comments.

DATA AVAILABILITY

Some of the data underlying this article is made available by Adams et al. (2023a). The remainder of the data set will be released together with Conselice et al. 2023. The catalogues of the sample discussed herein may be acquired by contacting the corresponding author.

REFERENCES

Adams N. J., et al., 2023a, preprint (arXiv:2304.13721)
Adams N. J., et al., 2023b, MNRAS, 518, 4755
Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, MNRAS, 451, 1892
Bagley M. B., et al., 2022, ApJL, 946, L12
Bagley M. B., et al., 2023, preprint (arXiv:2302.05466)
Bellovary J., Volonteri M., Governato F., Shen S., Quinn T., Wadsley J., 2011, ApJ, 742, L13
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, L1503
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bushouse H. et al., 2022, JWST Calibration Pipeline, Zenodo, https://doi.org/10.5281/zenodo 7325378 (2022).
Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, MNRAS, 480, 4379
Conroy C., Gunn J. E., 2010, Astrophysics Source Code Library, ascl:1010
Constantin A., Shields J. C., 2003, PASP, 115, 592
Dewsnap C., Barmby P., Gallagher S. C., Urry C. M., Ghosh A., Powell M. C., 2023, ApJ, 944, L137
Duras F., et al., 2020, A&A, 636, 73
Elvis M., Risaliti G., Zamorani G., 2002, ApJ, 565, L75
Ferreira L., et al., 2022, ApJ, 938, L2
Finkelstein S. L., et al., 2022, ApJL, 946, L13
Fotopoulou S., et al., 2016, A&A, 587, 142
Fudamoto Y., Inoue A. K., Sugahara Y., 2022, ApJ, 938, L24
Furtak L. J., et al., 2023, ApJ, 952, 142
Gao L., Yoshida N., Abel T., Frenk C. S., Jenkins A., Springel V., 2007, MNRAS, 378, 449
Goulding A. D., Greene J. E., 2022, ApJ, 938, L9
Grogin N. A., et al., 2011, ApJS, 197, 35

Harikane Y., et al., 2023, preprint (arXiv:2303.11946)
Hyun M., et al., 2023, ApJS, 264, 19
Kocevski D. D., et al., 2023, preprint (arXiv:2302.00012)
Koekemoer A. M., et al., 2011, ApJS, 197, 36
Labbe I., et al., 2023, preprint (arXiv:2306.07320)
Laird E. S., et al., 2009, ApJS, 180, 102
Larson R. L., et al., 2022, Spectral Templates Optimal for Selecting Galaxies at z>8 with JWST
Larson R. L., et al., 2023, preprint (arXiv:2303.08918)
Madau P., 1995, ApJ, 441, L18
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Maiolino R., et al., 2023, preprint (arXiv:2305.12492)
Matthee J., et al., 2023, Little Red Dots: an abundant population of faint AGN at $z \sim 5$ revealed by the EIGER and FRESCO JWST surveys preprint (arXiv:2306.05448)
Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, L64
Naidu R. P., et al., 2022a, preprint (arXiv:2208.02794)
Naidu R. P., et al., 2022b, ApJ, 940, L14
Nakajima K., Maiolino R., 2022, MNRAS, 513, 5134
Natarajan P., Pacucci F., Ferrara A., Agarwal B., Ricarte A., Zackrisson E., Cappelluti N., 2017, ApJ, 838, L117
Onoue M., et al., 2022, ApJL, 942, L17
Pacucci F., Ferrara A., Grazian A., Fiore F., Giallongo E., Puccetti S., 2016, MNRAS, 459, 1432
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097
Perrin M. D., Soummer R., Elliott E. M., Lallo M. D., Sivaramakrishnan A., 2012, in Clampin M. C., Fazio G. G., MacEwen H. A., Oschmann Jacobus M., J.eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference SeriesVol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham. p. 84423D
Pezzulli E., Valiante R., Schneider R., 2016, MNRAS, 458, 3047
Pontoppidan K., et al., 2022, ApJL, 936, L14
Prevot M. L., et al., 1984, A&A, 132, 389
Rieke M. J., et al., 2022, PASP, 135, 028001
Rodighiero G., Bisigello L., Iani E., Marasco A., Grazian A., Sinigaglia F., Cassata P., Gruppioni C., 2023, MNRAS, 518, 19
Salpeter E. E., 1955, ApJ, 121, L161
Sarmento R., Scannapieco E., Pan L., 2017, ApJ, 834, L23
Stalevski M., Ricci C., Ueda Y., Lira P., Fritz J., Baes M., 2016, MNRAS, 458, 2288
Thomas A. D., Groves B. A., Sutherland R. S., Dopita M. A., Kewley L. J., Jin C., 2016, ApJ, 833, L266
Trenti M., Stiavelli M., Shull J. M., 2009, ApJ, 700, L1672
Treu T., et al., 2022a, ApJ, 935, 110
Treu T., et al., 2022b, ApJL, 942, L28
Trinca A., et al.,2022, MNRAS, 519, 4753
Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, L559
Volonteri M., Habouzit M., Colpi M., 2021, Nat. Rev. Phys., 3, 732
Volonteri, 2023, MNRAS, 521, 241
Windhorst R. A., et al., 2023, AJ, 165, 13
Yang G., et al., 2020a, MNRAS, 491, 740
Yang J., et al., 2020b, ApJ, 897, L14
Yue B., Ferrara A., Salvaterra R., Xu Y., Chen X., 2014, MNRAS, 440, 1263
Zavala J. A. et al., 2023, ApJ, 943, L9

This paper has been typeset from a TeX/LaTeX file prepared by the author.