

Measurements of the $z = 4 - 10$ X-ray Luminosity Function: the high space density of moderate-luminosity, obscured AGN

C. L. Barlow-Hall,^{1*} and J. Aird,^{1,2}

¹*Institute for Astronomy, The University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*

²*School of Physics & Astronomy, University of Leicester, University Road, Leicester LE1 7RJ, UK*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Supermassive Black Holes (SMBHs) are often theorised to undergo significant growth within the early Universe, however, the X-ray Luminosity Function (XLF), often used as the principal tracer of the SMBH accretion density, lacks observational constraints above $z \sim 6$, until now. We present new measurements of the $z = 4 - 10$ XLF at intermediate luminosities, taking advantage of recent deep near-infrared imaging from UltraVISTA that enables us to identify galaxies and AGN at high redshifts within which we identify X-ray sources using the Chandra COSMOS data. We first performed a cross-match to a deep Chandra source list, for which the X-ray sensitivity can be accurately quantified, before exploiting the available X-ray data further through direct extraction of X-ray counts at the positions of COSMOS2020 galaxies. With the resulting $z = 4 - 10$ X-ray AGN sample, comprised of 21 blind detections and a further 11 directly extracted detections, we have measured the early space density of AGN, at moderate-luminosities where the majority of early SMBH growth occurred. These measurements reveal **higher space-densities than expected**, based on the extrapolation of XLF models from lower redshifts. Whilst our measured space densities at $z = 4 - 5$ are consistent with model predictions, at $z = 5 - 7$ we find space densities of the order of $10\times$ the extrapolated model predictions and could be as high as $220\times$ the model extrapolations at $z = 7 - 10$. In addition, we find evidence that a large fraction of the early AGN population are heavily obscured, with an obscured fraction of $0.982_{-0.008}^{+0.007}$; correcting for this obscuration further increases the measured space densities. Comparing to recent JWST results, these measurements begin to bridge the gap between the bright-end of the quasar luminosity function and the latest JWST observations of very early, low-luminosity AGN, indicating that a larger fraction of the first galaxies likely play host to a rapidly growing SMBH than previously thought.

Key words: keyword1 – keyword2 – keyword3

1 INTRODUCTION

Supermassive Black Holes (SMBHs) are some of the most violent and esoteric objects in the Universe, yet despite the consensus that they played a key role in shaping the Universe we see today (e.g. Somerville & Davé 2015; Heckman & Best 2014; Kormendy & Richstone 1995) their formation and subsequent growth remains a very open question. When SMBHs undergo rapid mass accretion they produce strong emissions across a wide range of wavelengths, making them observable at the centre of their host galaxies at a given wavelength (e.g. see Padovani et al. 2017) as what are known as Active Galactic Nuclei (AGN). The necessity of feedback, due to these AGN, to limit the growth of high mass galaxies is clearly demonstrated by Cosmological Hydrodynamical simulations (e.g. Somerville & Davé 2015), whilst studies such as those of Boyle & Terlevich (1998); Ueda et al. (2014); Aird et al. (2015) have shown the evolution of the rate of Black Hole mass density growth closely maps the trend observed in the rate of stellar density build up. It is thus accepted that SMBHs co-evolve with their host galaxy (see Heckman & Best 2014), undergoing periods of AGN activity,

with most galaxies expected to play host to a SMBH (Rees 1984; Kormendy & Ho 2013). Nevertheless, how the seeds of these black holes form in the very early Universe and which processes drive their early growth to produce the SMBHs that we observe remains unclear (e.g. see Inayoshi et al. 2020, for a review).

Observations of AGN at high redshift indicate that they contain SMBHs with masses of the order of $10^{8-9}M_{\odot}$ (e.g. Zubovas & King 2021) indicating that significant mass growth, of several orders of magnitude, must occur within the early Universe and thus during the epoch of initial stellar mass assembly of galaxies. Reaching the observed masses from even the most massive black hole seeds in the available cosmic time requires near-continuous growth at the highest possible rates (i.e. \sim the Eddington limit Trakhtenbrot et al. 2016; Orofino et al. 2018; Zhu et al. 2022). As noted by Sharma & Sharma (2024, and references therein), however, no SMBH can be expected to accrete at a constant or smoothly increasing rate, and thus even higher levels of mass growth must occur during active periods.

The growth of AGN across cosmic time, as a population, is traced by the Quasar¹ Luminosity Function (QLF); a measure of the space

* E-mail: c.barlow-hall@roe.ac.uk

¹ In this paper, we use the term “Quasar” to refer to all radiatively efficient

density of AGN at a given redshift and luminosity (Page et al. 1996; Boyle et al. 2000; Kalfountzou et al. 2014), and typically has the general form of a double power law. The QLF has been measured using Optical/Ultra Violet (UV) (e.g. Ren & Trenti 2021; Ross et al. 2013; McGreer et al. 2013; Schmidt 1968; Boyle et al. 2000; Fan et al. 2001), Infrared (IR) (e.g. Lacy et al. 2015; Assef et al. 2011) and X-ray selected AGN (e.g. Aird et al. 2010, 2015; Ueda et al. 2014; Gilli et al. 2009; Pouliaxis et al. 2024; Georgakakis et al. 2015; Miyaji et al. 2001; Ananna et al. 2019), all of which show the general trend of a rising space density from $z \sim 0$ up to cosmic noon before it declines again above $z \sim 3$ (as found by Schmidt et al. 1995; Aird et al. 2010; Vito et al. 2014; Aird et al. 2015). X-ray surveys are particularly useful for determining the QLF, known as the X-ray Luminosity Function (XLF), thanks to the uncontaminated AGN selection using X-rays and the well defined sensitivity of X-ray surveys (Padovani et al. 2017). X-ray emissions from an AGN are also a close tracer of the instantaneous growth of the SMBH, as they are produced by a hot electron corona very close to the SMBH (see Brandt & Alexander 2015, and references therein). Notably multiple studies have found that the XLF can probe the regime of X-ray bright, optically normal galaxies, thus identifying lower luminosity and obscured AGN that would otherwise go undiscovered at rest-frame UV/Optical (see e.g. Brandt & Alexander 2015; Brandt & Hasinger 2005; Lusso et al. 2023, for reviews). For the most heavily obscured AGN, when the line-of-sight column density reaches the Compton-thick regime (equivalent hydrogen column density, $N_{\text{H}} \gtrsim 10^{24} \text{ cm}^{-2}$) even hard X-ray emission will be substantially suppressed and the emergent X-ray spectrum is dominated by a Compton “reflection” component (e.g. Ikeda et al. 2009; Brightman & Nandra 2011); while such Compton-thick sources may still be detected in sufficiently deep X-ray imaging it becomes more challenging to infer their intrinsic luminosities and thus measure the overall XLF including these populations (see e.g. Buchner et al. 2015; Ananna et al. 2019).

Many studies have used parametric models to describe the evolution of the XLF, mainly applying evolutionary trends to the standard double power-law which characterises the population, and constrained these models using X-ray surveys (such as e.g. Page et al. 1996; Gilli et al. 2007; Ueda et al. 2014; Aird et al. 2015). However, due to both the X-ray flux limit and the drop in the space density toward higher redshifts, samples of X-ray selected AGN at high redshift remain very limited (e.g. Luo et al. 2017 identified just two $z > 5$ sources in the *Chandra* Deep Field South 7 Ms sample, while Marchesi et al. 2016 identified four $z > 6$ sources within the $\sim 2 \text{ deg}^2$ *Chandra* COSMOS-Legacy survey). For faint sources we need deep observations, while as the space density is known to decrease toward higher redshifts a large survey area is required. As such, samples of X-ray selected AGN at $z \gtrsim 3$ remain relatively small and parametric models of the evolution of the XLF toward higher redshifts remain poorly constrained at $z \gtrsim 5$ (see Vito et al. 2014; Ueda et al. 2014; Vito et al. 2018; Georgakakis et al. 2015, for examples of XLF fit to $3 \leq z \leq 6$). Most recently Pouliaxis et al. (2024) extended their model of the XLF out to $z \sim 6$ based on X-ray selected AGN from a combination of *Chandra* and XMM-Newton imaging in order to reach the depth and survey area required at these redshifts, yet still lacked sources at higher redshifts.

Despite the limited X-ray detections of AGN above $z = 6$, recent

observations with JWST have revealed many AGN candidates out to $z \sim 10$, predominantly identified through detection of broad H α and H β emission lines (e.g. Kocevski et al. 2024; Maiolino et al. 2024; Greene et al. 2024; Matthee et al. 2024) as well as the detection of high-ionisation lines and application of narrow-line diagnostics (e.g. Scholtz et al. 2025; Arevalo Gonzalez et al. 2025). Approximately 20% (Hainline et al. 2025) of the AGN candidates found by JWST are the so-called Little Red Dots, which appear very red in colour with a characteristic V-shaped spectral energy distribution (SED) and compact morphology (Matthee et al. 2024; Greene et al. 2024; Kocevski et al. 2024). Whilst X-rays provide a fairly unambiguous signature for the identification of AGN, however the vast majority of these JWST AGN lack X-ray detections (see Bogdán et al. 2024, for a notable exception). Some studies, such as those of Maiolino et al. (2025a), have suggested this apparent X-ray weakness is due to a higher density of obscuring gas within these high redshift sources (see also Inayoshi et al. 2024) or the geometry of the system (Madau & Haardt 2024), with Brooks et al. (2024) finding evidence for this potentially X-ray obscuring material.

With the advent of JWST the samples of very high redshift galaxies and of AGN within high redshift galaxies is increasing, indicating a far higher space density of AGN at $z > 6$ than extrapolated models previously predicted, including the XLF models. It should be noted, however, that JWST is imaging only very small survey areas and probing the faint-end of the QLF, a regime where current XLF models are very poorly constrained. Recently, Barlow-Hall et al. (2023) placed constraints upon the bright-end of the XLF at $z \sim 6$ based on a large compilation of *Swift* X-ray telescope observations (the Extragalactic Serendipitous *Swift* Survey, ExSeSS; see Delaney et al. 2023), finding a steep bright-end slope consistent with the bright-end of the model XLF extrapolations (see also Wolf et al. 2021, for constraints from *eROSITA*). However, the space density of AGN within the regime between the faint sources found by JWST and the bright sources found by large sky surveys remains unconstrained. Here we utilise the COSMOS field, providing a relatively large survey area with deep X-ray imaging, in order to constrain the moderate luminosity XLF and bridge the gap in AGN space density measurements from large area shallow surveys that find the brightest sources to the small area deep fields which are able to probe the large numbers of very faint sources.

The *Chandra* X-ray Observatory (Weisskopf et al. 2002) is one of the most sensitive X-ray observatories to date, probing fluxes of $\leq 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ in the 0.5–2 keV energy band in the deepest observations, a factor of ≥ 100 deeper than achieved in ExSeSS. In order to obtain this depth of imaging, *Chandra* has observed a much smaller area of sky compared to *Swift* and other less sensitive survey telescopes. However, the depth granted by *Chandra* is crucial to probing moderate luminosity AGN at high redshifts. Although this X-ray data existed prior to the work presented here, it is with recent deep NIR Ultra VISTA imaging across large areas that allows for the identification of high redshift galaxies and AGN (along with reliable photometric redshift measurements) enabling new catalogues such as COSMOS2020 (Weaver et al. 2022) to cover the $z = 4 - 10$ regime (see also Bowler et al. 2020; Finkelstein et al. 2022; Donnan et al. 2023b, 2024, which have found an excess of high-redshift galaxies only visible in NIR observations). Thus, we are now able to identify reliable high-redshift counterparts to the faintest X-ray sources revealed in the *Chandra*-COSMOS Legacy imaging.

The form of the COSMOS2020 catalogue and the X-ray data used in this work are detailed in §2. The cross-matching of COSMOS2020 and X-ray sources performed in order to obtain the blind source sample is outlined in §2.3, and §2.4 details the creation of the blind and

AGN i.e. that capture the bulk of SMBH mass assembly (Alexander & Hickox 2012), regardless of luminosity or obscuration level, and thus the QLF, ideally, should capture the overall SMBH accretion density.

extracted sample through the direct extraction of X-ray counts for all COSMOS2020 sources. We then detail the measurement of the XLF using these X-ray AGN samples in §3, and test the extremes of these measured constraints arising from the differing photometric redshift estimates from COSMOS2020 (§3.4). The impact of obscuration on the high redshift XLF is then investigated in §4. Finally, the implications of these new constraints on the XLF are discussed in §5 and we summarise our findings in §6.

Given the extent of details on the sample and method of XLF measurement, the casual reader may wish to skip to §4.3, and the discussion in §5. More specifically the comparison between our final X-ray constraints on the bolometric luminosity function of AGN and recent JWST measurements are shown in Figure 15 and the measured evolution of the Black Hole Accretion Rate Density is shown in Figure 16. Throughout this work we adopt a Flat Λ CDM cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$, and all errors given are the 1σ uncertainties on the values.

2 THE COSMOS FIELD DATA SETS

The COSMOS field, centred at $10^{\text{h}}00^{\text{m}}27^{\text{s}}.92 + 02^{\circ}12'03''.50$ (J2000), was originally defined by Scoville et al. (2007) for large-area deep (0.05 arcsec) Hubble Space Telescope (HST) imaging in order to address a number of science goals; from large dark matter distributions and galaxy assembly to early galaxies and the evolution of AGN. This $\sim 2 \text{ degree}^2$ equatorial field was chosen as its location ensured it was visible to all astronomical instruments, at the time, thus allowing for detailed multi-wavelength studies of the sources within the field. The COSMOS field also benefits from having a relatively low and uniform Galactic extinction, compared to other equatorial fields, whilst being devoid of bright X-ray, UV and radio sources and thus reducing the contamination of the signals of faint sources within the field. Following its definition, the COSMOS field has been surveyed using a wide range of facilities spanning the full wavelength range (including *HST*, *Spitzer*, GALEX, *XMM-Newton*, *Chandra*, Subaru, ALMA, VLA, VISTA, ESO-VLT, UKIRT, NOAO, CFHT, and more), with deep observations performed by almost all major astronomical facilities (see e.g. Sanders et al. 2007; Capak et al. 2007; Taniguchi et al. 2015; Le Fèvre et al. 2020). The wealth of data available within the COSMOS field makes it incredibly important to extragalactic astronomy. Crucially for the identification of galaxies out to very high redshifts, the COSMOS field has very deep imaging at Near and Mid-Infrared wavelengths from the VISTA telescope (McCracken et al. 2012; Moneti et al. 2023; Ashby et al. 2018).

2.1 UV-to-IR photometry and photometric redshifts from COSMOS2020

Since COSMOS was first defined, several photometric catalogues of this field have been publicly released as deeper data has become available (Capak et al. 2007; Ilbert et al. 2009, 2013; Muzzin et al. 2013; Laigle et al. 2016). The COSMOS2020 catalogue of Weaver et al. (2022) is the most up to date release, including deep UV (Sawicki et al. 2019) and ultra deep optical imaging (Aihara et al. 2019) as well as Near IR data from the UltraVISTA survey DR4 (Moneti et al. 2023; McCracken et al. 2012). For COSMOS2020, all legacy datasets used in previous iterations of the COSMOS catalogue (e.g. COSMOS2015: Laigle et al. 2016) have also been reprocessed to align with newly available astrometry from Gaia DR1 (Gaia Collaboration et al. 2016). The legacy imaging has also been resampled to a pixel

scaling of 0.15 arcsec and stacked images aligned to the COSMOS Tangent point ($10^{\text{h}}00^{\text{m}}27.92^{\text{s}} + 02^{\circ}12'03''.50$). Weaver et al. (2022) identified sources within the wealth of deep COSMOS data using two different detection methods; implementing SExtractor (Bertin & Arnouts 1996) as in Laigle et al. (2016), and THE FARMER software of Weaver et al. (2023) which utilises THE TRACTOR photometry modelling code (see Lang et al. 2016, for details). These two versions of the COSMOS catalogue are referred to as the CLASSIC Catalogue and THE FARMER Catalogue, respectively.

For this work we use THE FARMER catalogue, as it has been found to have a higher completeness than the Classic Catalogue (e.g. Shuntov et al. 2022). It also provides more accurate photometry for the faint sources (which will comprise the majority of the high redshift sample we will use in this paper), is more robust against source blending and is able to identify a higher density of $z > 6$ sources (see e.g. Brinch et al. 2023; Sillassen et al. 2022).

To create THE FARMER catalogue, source detection is performed on a χ^2 detection image that combines the i -, z -, Y - J - H - and K_s -band imaging and is generated using the SWarp package (Szalay et al. 1999). Parametric models of the source shape are then used to perform photometry across all wavebands. THE FARMER software measures the photometric fluxes by scaling one of five parametric models (Point Source, Simple Galaxy, ExpGalaxy, DevGalaxy and Composite Galaxy²) assigned to each source (see Weaver et al. 2023, for details). Thus, this method of measuring the photometry does not require PSF homogenisation or the use of fixed apertures and the best quality photometry is not degraded to the lowest quality resolution achieved. As these models cannot be applied to saturated bright stars or sources contaminated by stellar halos, Weaver et al. (2022) adopt the HSC-SSP PDR2 (Coupon et al. 2018) bright-star masks in order to mask out all stars with a G-band magnitude of 18 or less in the Gaia DR2 star catalogue. In order to prevent the introduction of inhomogeneities to the constraints on the models for photometric extraction, THE FARMER catalogue is also limited to the UltraVISTA region such that there is coverage in all wavelength bands used. After applying these cuts, THE FARMER photometric catalogue of COSMOS sources obtained by Weaver et al. (2022) covers 1.792 deg^2 of sky and contains a total of 964 506 sources, all determined to have reliable detections.

Weaver et al. (2022) determine the photometric redshifts of their sources using both the EAzy (Brammer et al. 2008) and LePhare³ (Arnouts et al. 2002; Ilbert et al. 2006) photometric redshift codes, with the photometric measurements corrected for Galactic extinction based on the dust maps of Schlafly & Finkbeiner (2011). We take the LePhare photometric redshift measurements as the redshifts of our final galaxy catalogue, as LePhare accounts for a possible AGN component in the photometric redshift determination. LePhare is also found to have a lower outlier percentage when comparing the photometric redshift estimates to available spectroscopic measurements (Weaver et al. 2022). LePhare is used to determine the photometric redshift and physical parameters of each source by fitting with a range of templates including 19 elliptical/spiral galaxy templates from Polletta et al. (2007), 12 from models of blue star-forming galaxies and 2 galaxy templates with exponentially declining star

² Point source models use the PSF directly, SimpleGalaxy uses a fixed exponential light profile and ExpGalaxy uses an exponential light profile parametrised by the source's image properties. The DevGalaxy model is similar to the ExpGalaxy, but uses the de Vaucouleurs light profile, whilst the CompositeGalaxy model combines both the ExpGalaxy and DevGalaxy models.

³ <http://www.cfht.hawaii.edu/~arnouts/lephare.html>

formation histories (Bruzual & Charlot 2003; Onodera et al. 2012). LePhare is also used to perform fits using AGN templates consisting of both unobscured sources where the AGN dominates the light at most wavelengths and obscured sources with significant host galaxy contributions (see Salvato et al. 2009, 2011, for details), as well as stellar templates from Pickles (1998). For sources at redshifts of $z > 10$, the observed wavelength of the Lyman break will be shifted outside the wavelength range of the deep near-IR photometry in the COSMOS2020 catalogue and thus such extremely high-redshift sources are likely to be undetected. As such, the redshift grid used by LePhare and thus our sample is limited to a maximum redshift of $z = 10$.

Finally, we restrict the catalogue to only include sources that lie within the footprint of the *Chandra* X-ray imaging, as deep X-ray data is not available for sources outside this region. The final, filtered catalogue consists of 816 944 optical/IR selected sources.

2.2 Chandra Imaging

The *Chandra* X-ray Observatory is the highest angular resolution X-ray telescope to date. With an on-axis resolution of 0.5 arcsec, *Chandra* can achieve the sensitivity needed to find moderate-luminosity high-redshift X-ray sources.

The Chandra-COSMOS (C-COSMOS) program of Elvis et al. (2009) was the first contiguous survey of the COSMOS field and provided an unprecedented area and depth of X-ray imaging at the time. Utilising a strategy of overlapping pointings, with offsets of 8 arcmin around the survey centre at $10^{\text{h}}00^{\text{m}}24^{\text{s}} + 02^{\circ}10'55''$, provided a near-uniform sensitivity across the centre of the COSMOS field. The final C-COSMOS survey covered $\sim 0.9 \text{ deg}^2$ to an effective exposure of 160 ks over the central $\sim 0.5 \text{ deg}^2$ and 80 ks over the remaining $\sim 0.4 \text{ deg}^2$. Later, Civano et al. (2016) combined the C-COSMOS survey with 2.8 Ms of additional *Chandra* imaging (performed over four periods: Nov 2012-Jan 2013, Mar 2013-July 2013, Oct 2013-Jan 2014 and Mar 2014) to produce the *Chandra* COSMOS-Legacy Survey, providing uniform coverage of the 1.7 deg^2 COSMOS/HST field to an effective exposure of 160ks and a total area coverage of $\sim 2.2 \text{ deg}^2$, 80% of which has a resolution of 2-4 arcsec.

For this study we use the deep *Chandra* imaging provided by the *Chandra* COSMOS-Legacy survey (Civano et al. 2016). We adopt the custom reanalysis of these *Chandra* data following the data reduction and source detection process detailed in (Laird et al. 2009; Nandra et al. 2015; Kocevski et al. 2018) to produce the clean images and blind source catalogue used in this study and in previous works (Aird et al. 2017, 2018, 2019; Laloux et al. 2023, 2024). Performing this reanalysis provides access to all resulting data products, required for this work, and ensures that we can control and quantify the sensitivity of the *Chandra* imaging when producing the source catalogues.

The actual data reduction of the *Chandra* COSMOS-legacy data used in this work was performed by Georgakakis et al. (2017) using the pipeline described by Nandra et al. (2015) and Kocevski et al. (2018). Here we will briefly summarise the reduction of the *Chandra* imaging, the subsequent source identification, and the methods used to determine the sensitivity and assess the X-ray completeness.

2.2.1 Data Reduction and Initial Source Detection

The raw data from *Chandra* is cleaned and processed using the CIAO (*Chandra* Interactive Analysis of Observations) suite v.4.8 (Fruscione et al. 2009). Correcting for the telescope dithering, event files are then cleaned by removing Hot Pixels, image streaking and Cosmic-Ray

afterglows. Anomalous peaks in the observed X-ray background are identified from a light curve of the measured background as periods when the rate is ≥ 1.4 times the quiescent background (as given by the count rate limit required for an excess variance in the light curve of 0), and subsequently cut from the event file. With all anomalous pixels and events removed, the number of events (photon detections) of a given measured energy (0.5–7 keV, 0.5–2 keV, 2–7 keV and 4–7 keV for Full, Soft, Hard and Ultra Hard images respectively) at each pixel are summed across all observations to produce the final image file. For a given energy band the corresponding effective exposure map is generated, accounting for exposure time, detector quantum efficiency, and effective area (weighted by energy assuming a power-law spectrum of photon index $\Gamma = 1.4$), as well as bad pixels, chip gaps and edges, and the dithering motion of the telescope. Finally the Point Spread Function (PSF) at each pixel position is determined through simulations of *Chandra*'s High Resolution Mirror Assembly (HRMA), performed using the MARX Simulator of Wise (1997), then weighted by the effective exposure of each individual observation to generate a PSF map providing the radius corresponding to an enclosed energy fraction (EEF) of 70% and 90%.

Once the raw data is processed a source detection method is applied to the resulting image, in order to produce a blind⁴ catalogue of X-ray sources. Initial source detections are performed on the full resolution images, using the WAVDETECT source detection code from the CIAO suite (Fruscione et al. 2009). This initial detection is done with a very low significance threshold, given by a false probability threshold of 10^{-4} , and limited to the image region where the exposure is $\geq 10\%$ of the maximum exposure. This low significance threshold prevents real sources being missed, however there will be a number of spurious sources also introduced as a result. For all candidate sources the total photon counts, T , are measured in an aperture of radius given by the 70% EEF of the PSF. The local background level is measured in a 100 pixel wide annulus of inner radius $1.5\times$ the 95% EEF PSF, at each source position, then re-scaled to the source detection area to estimate the expected background in the source aperture, b . For each source, we can then calculate the net, effective exposure-corrected count rate,

$$c = \frac{T - b}{p_{70} t_{\text{exp}}} \quad (1)$$

where the factor of $p_{70} = 0.7$ is applied in order to correct for the choice of an aperture extraction radius corresponding to the 70% EEF at the source position and t_{exp} is the effective exposure. We also calculate the false probability, $P(\geq T|b)$, that gives the probability that the source is a spurious detection due to a fluctuation of the background, using the equation,

$$P(\geq T|b) = 1 - P(< T|b) = 1 - \sum_{i=0}^{T-1} \frac{b^i}{i!} e^{-b}. \quad (2)$$

A threshold false probability value of $\lesssim 4 \times 10^{-6}$ is then applied to the candidate sources, in order to remove those believed to be spurious detections and leave only the reliable source sample.

In order to search for potentially missed sources, all reliable detections are masked out and a second pass of this detection method is performed. The background and false probability of each candidate source are subsequently recalculated, and the same false probability threshold is applied once more. As there is negligible difference in the detections list following this second detection stage, the process

⁴ By ‘‘blind catalogue’’ we mean based only on the X-ray data and without using any a priori knowledge of source positions from other wavelengths.

is not run again. For sources within 5 pixels of one another only the source with the highest measured counts is kept, in order to prevent duplicated sources within the final catalogue. Once the source list is found, the background map is created by removing all source counts from the image and replacing based on the local background (see Georgakakis et al. 2008).

This source detection process is carried out on the images from all four energy bands (Full, Soft, Hard and Ultra Hard). The resulting reliable source lists are then merged and for any sources that are not independently detected in all bands the missing counts, background and exposure values are extracted from the corresponding imaging and maps at the source position. Count rates, c , in a given band may be converted to fluxes assuming a certain spectral model. However, in this work we convert directly from count rates to rest-frame 2–10 keV luminosity (L_X) for a given redshift.

In this work we use the catalogue resulting from this source extraction process, containing a total of 3 627 X-ray sources, referred to as the blind source catalogue. The images, background maps, exposure maps and PSF maps generated through the above reduction process are also used in §2.4 for the extraction of X-ray information at the positions of all sources in the reduced COSMOS2020 catalogue.

2.2.2 The Chandra Sensitivity across the COSMOS field

The X-ray Luminosity Function (XLF) determination requires the sensitivity of the X-ray survey to be known and quantified. We do this by generating a sensitivity map (following the process detailed in Georgakakis et al. 2008), determining the number of counts required to produce a significant detection, i.e. a detection above the specified false probability threshold given the expected background count rate, at each pixel position in the image. From the sensitivity map, we then determine the area curve; the total area of sky where the available Chandra imaging is sensitive to sources above a given X-ray flux threshold.

For any given pixel position in an image there is small probability that a random fluctuation will result in sufficient counts for a spurious detection (i.e. due to a fluctuation of the counts from the background) or result in a faint source appearing brighter and thus being detected when otherwise it would not be. The probability that the total *observed* counts exceeds the given detection threshold (L counts, determined for a given false probability threshold using equation 2) is thus described by the cumulative probability,

$$P(\geq L|c, b) = 1 - P(< L|c) = \frac{1}{\Gamma(L)} \int_0^\lambda e^{-t} t^{L-1} dt \quad (3)$$

where $\Gamma(L)$ is the gamma function and $\lambda = b + ct_{\text{exp}}p_{70}$ where $ct_{\text{exp}}p$ is the expected number of counts from a source with count rate c . The area curve is the sum of the cumulative probability distributions across all pixels in the image, determined for a given source count rate, multiplied by the sky area of a pixel.

Any masking applied to the Chandra imaging will reduce the sky area for our study. Thus, prior to the area curve calculation, we apply the bright-star mask and limit the survey area to the UltraVISTA region (see §2.2.1).

Sensitivity maps, and corresponding area curves, are calculated for the Full and Hard band imaging for both a false probability threshold of 4×10^{-6} , corresponding to the blind source catalogue detailed above, and for the less stringent threshold adopted when considering the X-ray data at known COSMOS2020 source positions (see §2.3 below). Figure 1 shows the resulting area curves, as given by a false probability threshold of 4×10^{-6} , for the Full and Hard energy

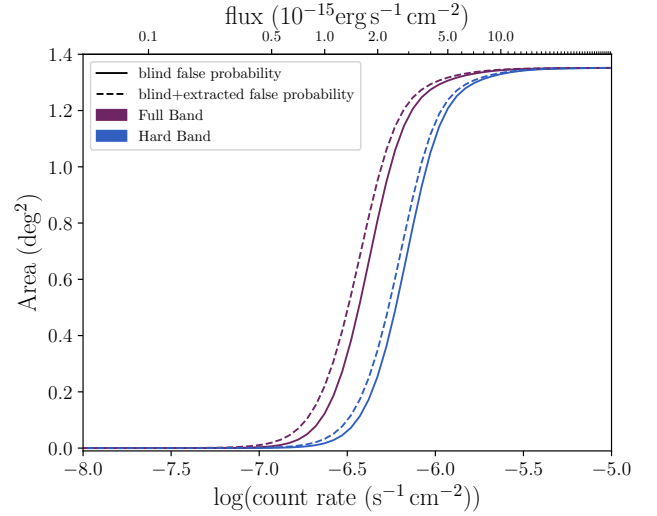


Figure 1. Hard band (2–7 keV, blue) and full band (0.5–7 keV, purple) area curves, giving the sky area coverage as a function of the effective area-corrected count rate, determined with the blind sample false probability threshold of 4×10^{-6} (solid line) and the blind+extracted sample threshold of 3.38×10^{-5} (dashed line). All use of these area curves is based on the count rate (converted to intrinsic X-ray luminosities at a given z under differing spectral assumptions); however, the upper axis shows the corresponding 2–10 keV and 0.5–10 keV observed-frame fluxes for the hard and full bands (respectively) using a fixed conversion that assumes a power-law spectrum with $\Gamma = 1.4$.

bands used for the XLF determination (§3 and §4). From these area curves, the full survey area of our X-ray catalogue is 1.35 deg^2 , with a 10% area coverage at a count rate of $2.41 \times 10^{-7} \text{ s}^{-1} \text{ cm}^{-2}$ and $3.89 \times 10^{-7} \text{ s}^{-1} \text{ cm}^{-2}$ for the Full and Hard band imaging respectively (corresponding to a flux of $1.02 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$ and $1.66 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$ respectively, for a power law spectrum of $\Gamma = 1.4$).

2.3 X-ray and redshift catalogue cross-matching

In order to perform redshift-dependent studies of the X-ray AGN population, we identify counterparts in the filtered COSMOS2020 sample (see §2.1) to sources within the blind X-ray catalogue (detailed in §2.2.1), producing a sample of X-ray sources with optical/IR photometry. The counterparts to our blind catalogue are determined using the Bayesian Cross-Matching code NWay (Buchner et al. 2021).

As detailed in Salvato et al. (2018), NWay determines the probability (p_i) of each multi-wavelength source being the true counterpart to the X-ray detection and the probability that the true counterpart of each X-ray source is in the multi-wavelength sample (p_{any}), given the separation between the X-ray source and potential counterparts, positional uncertainties and total area of sky. Weighting is then applied to these probabilities, given specified (or internally generated) magnitude and colour priors for the cross-match.

Before identifying counterparts to our blind X-ray sample, any X-ray sources within the bright-star masks applied to the COSMOS2020 catalogues—or within a distance from the mask corresponding to the X-ray positional uncertainty—are removed from the sample, in order to ensure the final cross-matched sample contains only X-ray sources for which, with reasonable confidence, the optical/IR counterparts do not lie within the masked regions of the COSMOS2020 sample.

Apertures of radius $r = 0.82$ arcsec, given by the mean positional uncertainty of all blind X-ray sources, are placed at the X-ray source positions, and all sources for which this aperture intersects the bright star mask regions are removed from the X-ray sample producing a sample of 2792 X-ray sources. In addition, as the COSMOS2020 sample used in this study is limited to the UltraVISTA sky area, we similarly limit the blind X-ray source sample to this area. Thus the resulting area of both our masked COSMOS2020 and blind X-ray source samples is 1.35 degrees², containing 816 944 and 2 408 sources respectively.

We consider potential COSMOS2020 counterparts, to the blind X-ray sources, within a 10 arcsec search radius of each X-ray position. This search radius exceeds 3 times the maximum X-ray positional uncertainty of the sample and thus ensures that $> 99\%$ of the true counterparts are included (we note that the typical offset between an X-ray source position and the final, best counterpart is typically ~ 0.47 arcsec, i.e. much lower than this initial search radius).

We perform NWay cross-matches using a number of internally generated priors. These priors are based on the magnitudes in the UltraVISTA *Ks*-band, IRAC channel 1 and IRAC channel 2, chosen as they are good tracers of AGN activity at high redshifts (see e.g. Reines & Comastri 2016) as well as the IRAC channel 1-channel 2 colour. For each run, we adopt the counterpart to each X-ray source with the highest p_i value as the best match and apply a threshold on p_{any} to select X-ray sources with a reliable in the COSMOS2020 catalogue.

In order to assess the fraction of incorrect associations from a given NWay run, and determine the appropriate p_{any} threshold, we repeat the cross-matching process with NWay using a catalogue of fake X-ray sources with randomised positions. The resulting catalogue of fake sources assigned COSMOS2020 counterparts provides an estimate of the incorrect associations that will be present in the sample of cross-matched blind X-ray sources.

The fake sample is generated by applying a random positional shift, of 20 – 120, arcsec to the positions of all sources in the unmasked blind X-ray sample. The same masking that is applied to the real sample of X-ray sources is also applied to this fake sample, in order to remove any fake sources that fall within the masked regions of the COSMOS2020 sample, where the photometry is unreliable. Following the same approach as for the real X-ray sample, we assign counterparts to the fake X-ray sample with the magnitude priors generated, by Nway, during the cross-matching of the real X-ray sample to COSMOS2020. The resulting sample is then, similarly, limited to only the potential counterparts with the largest p_i value, reducing the sample to only the best cross-matches identified for each fake X-ray source.

Using the real and fake cross-matched samples, we assess the quality of the resulting sample of counterparts to our blind X-ray sources. We determine the completeness (the fraction of the real sample which have a correctly matched counterpart) of the sample of counterparts to the sample of real X-ray sources, as given by,

$$\text{Completeness} = \frac{N(p_{\text{any}} > p_{\text{any,threshold}})}{N_{\text{total}}} \quad (4)$$

where N_{total} is the total number of X-ray sources in the sample, and $N(p_{\text{any}} > p_{\text{any,threshold}})$ is the number of matched sources with p_{any} (the probability that any of the potential counterparts in the COSMOS2020 sample are the true counterpart) greater than a given threshold value ($p_{\text{any,threshold}}$). The fraction of matched sources that are likely to be incorrectly matched, known as the false positive fraction, is similarly calculated from the fake match sample by taking

the ratio of the number of matches to the fake sample with p_{any} exceeding the same $p_{\text{any,threshold}}$ to the total number of X-ray sources.

The completeness and false positive fraction are calculated for a range of $p_{\text{any,threshold}}$ values, for each of the matched samples resulting from different combinations of magnitude and colour based priors. We find the cross-match performed with the *Ks*-band and IRAC colour magnitude priors produces the lowest False Positive Fraction of all tested combination of magnitude priors, while also achieving a Completeness of over 0.9. Thus, we select the cross-match performed with the combination of a *Ks*-band magnitude and IRAC channel1-channel2 colour magnitude priors for the final cross-matched sample.

For this study we choose a false positive fraction of 10%, as an acceptable failure rate of the cross-matching. This corresponds to a sample completeness of 0.97 and $p_{\text{any,threshold}} = 0.388$. When considering the high-redshift samples only, given by the best fit photometric redshift values (as defined in §3.1 below), this $p_{\text{any,threshold}}$ corresponds to a False Positive fraction of $1.0 \times 10^{-3}\%$, $3.3 \times 10^{-4}\%$ and $1.1 \times 10^{-4}\%$ for the samples with photometric redshifts of $z \geq 4$, $z \geq 5$ and $z \geq 6$ respectively. These very small False Positive Fractions in the high-redshift samples are due to a much lower surface density of high-redshift sources compared to the full COSMOS2020 sample. As such, a random alignment between an X-ray source position and a high-redshift galaxy, which is subsequently falsely identified as the best counterpart to the X-ray source, is very unlikely to occur.

2.4 X-ray counts extraction at galaxy positions

High-redshift AGN appear observationally very faint, thus there are likely to be fainter X-ray AGN within the high-redshift galaxy sample that fall just below the X-ray detection thresholds adopted in the creation of the blind sample. As such these very faint AGN will not be present within the blind source catalogue. In order to further exploit the available X-ray data of high-redshift AGN, and push the limits of this study to the fainter X-ray regime, we extract X-ray counts directly from the cleaned Chandra imaging (detailed in §2.2.1) at the positions of the COSMOS2020 sources within our sample.

Prior to extracting X-ray counts at the COSMOS2020 positions, any COSMOS2020 sources in our sample that lie within 2 times the 90% EEf radius for the Chandra PSF of a bright X-ray source⁵ position are removed. This ensures that no counts due to the emission from a bright, nearby source (identified through the blind X-ray detection) are erroneously included when attempting to identify additional, fainter X-ray sources.

For all sources within the COSMOS2020 sample, with those near bright X-ray sources removed, the 70% EEf radius for the Chandra PSF at each source position is identified from the PSF maps generated in §2.2.1. At the position of each COSMOS2020 source in our sample we place apertures of radius given by the 70% EEf radius, and measure the total counts and total background within these apertures from the Chandra image and background map respectively⁶. Similarly, the average effective exposure within the aperture, for each source, is determined from the exposure map. The net, effective exposure-corrected count rates are calculated as in equation 1.

For each extracted count rate of the COSMOS2020 sources, the

⁵ For these purposes, we assume all sources in the blind X-ray source sample count as “bright X-ray sources”.

⁶ The value of pixels within the source aperture are only counted if their centre lies within the source aperture. No additional weighting to the pixel values is applied.

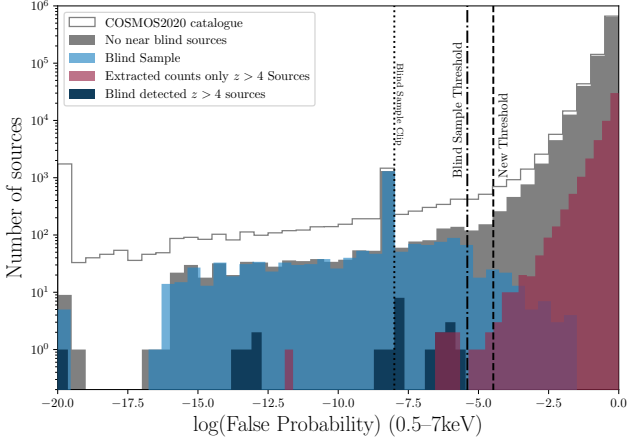


Figure 2. Histograms of the false probabilities of sources from the full-band Chandra-COSMOS data, showing values directly extracted from the Chandra imaging or the blind detections where available. The full sample of COSMOS2020 sources is shown by the grey outline, with those remaining once sources around the blind X-ray detected sources are removed shown in solid grey. The blind only sample, obtained using the full-band false probability threshold of 4×10^{-6} , is shown in blue (the histogram extends beyond the 4×10^{-6} threshold due to the detection of sources in the soft, hard or ultrahard bands, but not in the full-band shown here). The high redshift sources within this sample, for both the blind only sources (navy) and the newly extracted sources only (pink) are shown. The updated threshold applied to the extracted sample is indicated by the dashed line, with the threshold of 4×10^{-6} shown by the dot-dashed line, for comparison. For this histogram, we set a minimum false probability of 10^{-20} , resulting in a peak of sources where all those with a false probability below this are set to a value of 10^{-20} . Similarly any sources with extremely low (but non-zero) false probabilities within the blind detected source catalogue are set to a false probability of 1×10^{-8} (as indicated by the thin dotted line) in order to distinguish them from a false probability of 0, this results in an artificial peak in source numbers at this false probability.

false probability is calculated using equation 2. As can be seen by the pink histograms in figure 2, the frequency of sources increases exponentially towards a false probability of 1, i.e the majority of the extracted count rates for our COSMOS2020 sources are consistent with being produced by random fluctuations in the background.

In order to identify those sources for which the extracted count rates are unlikely to be due purely to a random background fluctuation, and thus obtain the sample of X-ray detections, we apply a false probability threshold to the extracted sample, as was done for the blind X-ray catalogue. However, as only positions of known COSMOS2020 sources are considered, and thus where a galaxy or AGN was known to be a priori, a less stringent threshold than was used in the formation of the blind catalogue can be applied. This enables fainter X-ray sources to be identified within the Chandra imaging, and increases the size of the sample of high-redshift, X-ray detected AGN without introducing a significant number of additional false X-ray sources.

The threshold false probability is determined iteratively. For a range of false probability thresholds, the fraction of the X-ray detected sample which have been falsely detected (the False Fraction) is calculated, using the following relation,

$$\text{FalseFraction} = \frac{\lambda_{\text{false}}}{N(f_{\text{prob}} < f_{\text{prob,threshold}})} \quad (5)$$

where $N(f_{\text{prob}} < f_{\text{prob,threshold}})$ is the number of X-ray de-

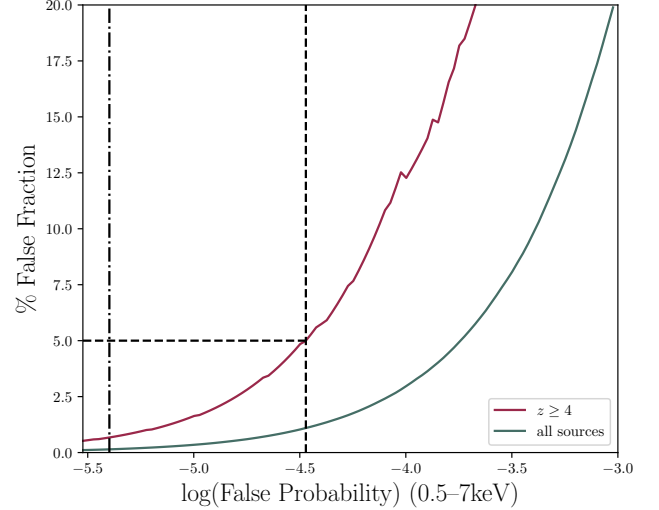


Figure 3. The change in false fraction with increasing false probability threshold, shown for all sources (sage) and $z \geq 4$ sources only (pink) in the full-band blind+extracted sample. The blind catalogue false probability fraction of 4×10^{-6} is shown by the dot-dashed line, with the new threshold false probability for the blind+extracted sample (3.38×10^{-5}) shown by the dashed line, corresponding to a False Fraction of 5% with respect to the $z \geq 4$ sources only (shown by the horizontal dashed line for reference).

tected sources in the sample, with f_{prob} less than the threshold, $f_{\text{prob,threshold}}$, and λ_{false} is the corresponding expected number of false detections, given by,

$$\lambda_{\text{false}} = N_{\text{gal}} f_{\text{prob,threshold}} \quad (6)$$

where N_{gal} is the total number of sources in the sample that X-ray data was extracted for. This false fraction calculation is carried out for both the full near-IR-selected source sample and the sample of sources with a best fit photometric redshift of $z \geq 4$ (see section 3.1 for details on the redshift values). The resulting False Fraction (shown in figure 3) can be seen to rise at lower False Probability Threshold values for the $z \geq 4$ sample (shown in blue) compared to the full sample (shown in red). Thus, for a given False Probability Threshold a higher proportion of the $z \geq 4$ sources are expected to be false detections compared to applying the same threshold to the entire galaxy sample. This occurs as, while the sample size of $z \geq 4$ sources is smaller than for the full galaxy sample (i.e. N_{gal} and thus λ_{false} is lower), the count rates of true X-ray sources at higher redshifts are typically lower than amongst the full galaxy sample and thus the number satisfying a given $f_{\text{prob,threshold}}$ is lower, decreasing the denominator in Equation 5 and increasing the false fraction.

We thus adopt the $f_{\text{prob,threshold}}$ value that corresponds to a False Fraction of 5% in the $z \geq 4$ sample, giving a threshold of 3.38×10^{-5} . This threshold is then applied to our combined sample of 882856 extracted and blind X-ray sources, to obtain our final sample of 2591 X-ray AGN containing 21 sources at $z \geq 4$.

3 INITIAL MEASUREMENTS OF THE XLF

In this section, we present our initial measurements of the XLF of AGN at $z = 4 - 10$. In §3.1 we define our primary sample of high-redshift X-ray detected AGN. We present measurements of the

XLF based purely on the blind X-ray sample in §3.2 and from the combination of the blind and extracted X-ray samples in §3.3. In §3.4 we consider the impact of differing photometric redshift estimates on our XLF measurements. Throughout this section, we neglect the impact of intrinsic obscuration on the estimated luminosities of our X-ray AGN and the resulting XLF measurements, deferring such studies to §4 below.

3.1 The High-Redshift sample and X-ray Luminosities

Following the processes detailed in §2, we obtained a sample of 2591 *Chandra* X-ray sources with counterparts in the FARMER version of the COSMOS2020 photometric redshift catalogue comprised of 2265 detections from the blind search of *Chandra* Legacy imaging and an additional 326 sources obtained through direct extraction of X-ray counts from the *Chandra* images at known source positions. Here we summarise this X-ray sample, the high-redshift cut and properties used in the following sections.

To define our sample of high-redshift sources, we adopt the LePhare photometric redshift estimates based on fits performed by Weaver et al. (2022). However, due to poor imaging or faint sources, a few of the sources in the parent near-IR–selected sample lack photometric measurements. For these sources LePhare was unable to obtain a fit based on the available photometry (these sources are marked as failed fits by LePhare). To ensure only sources with reliable photometric redshift measurements are included in our high- z sample, we remove the 261 sources for which LePhare was unable to reliably fit a photometric redshift measurement (due to a lack of photometric measurements, contamination by nearby sources or edge effects in the imaging)⁷. Following this the X-ray source sample is comprised of 2035 blind X-ray detected sources and 295 directly extracted sources, for a total sample size of 2330.

As detailed above (in §2.1), the LePhare code was used to perform redshift fits using galaxy templates and again using templates with an AGN spectral component. For sources with multiple redshift measurements from LePhare the better fitting photometric redshift (that with the smaller χ^2), either using galaxy or AGN spectral energy distribution templates, is taken to be the *best* estimate of the source redshift. As such 1239 of our X-ray sources have best-fitting photometric redshifts from galaxy spectral models and 1091 sources have best-fitting photometric redshifts based on AGN templates. Applying a redshift cut of $z_{\text{best}} \geq 4.0$ to this sample we obtain the primary high-redshift X-ray sample of 32 sources, with which we perform initial measurements of the observed X-ray Luminosity Function (see §3.2 and §3.3 below).

For all sources within the high-redshift samples, the measured net count rate in a given band (i.e. the full 0.5–7 keV and hard 2–7 keV bands) is converted directly into a rest-frame 2–10 keV luminosity (denoted as L_X). Following Aird et al. (2015), we initially assume an X-ray spectral model consisting of a power-law with a photon index $\Gamma = 1.9$, a reflection component modelled using the pexrav model (Magdziarz & Zdziarski 1995, which assumes an infinite plane reflector) of relative strength 1.0 and take the Galactic absorption ($N_{\text{H,Gal}}$)

⁷ This corresponds to a failed fraction of 10%. Whilst it would be unphysical for all 261 failed fits to correspond to high redshift sources ($\sim 8\times$ more sources than our primary sample), this failed fraction suggest there may be ~ 3 sources missed from our final $z \geq 4$ primary sample. If so our final measurements of the space density would be higher than currently, and thus further indicative of a larger space density than predicted by model extrapolations (see also section 5).

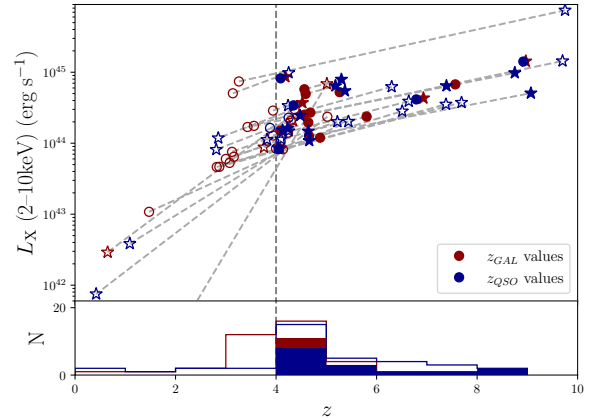


Figure 4. Rest-frame 2–10 keV X-ray luminosity against redshift, obtained from the full band count rate, for all sources included in our high-redshift samples. Sources are shown at redshifts based on both the galaxy template (z_{GAL}) and AGN template (z_{QSO}) photometric redshift fits (red and blue respectively). Sources that would leave the high redshift sample given either their galaxy or AGN redshift fit are indicated by empty points. For each source the best-fitting photometric redshift is shown as a star. The corresponding number of sources with z_{GAL} or z_{QSO} redshift is shown by the histogram, for the most conservative sample (solid) and the most optimistic sample (empty).

to be that of the COSMOS field (average of $10^{20.25} \text{cm}^{-2}$, HI4PI Collaboration et al. 2016).⁸ This model is then folded through the response function of *Chandra* to determine the conversion between the net, effective-exposure corrected count rate measured with *Chandra* (c from Equation 1) and the rest-frame 2–10 keV luminosity (L_X) of each source.

$$L_X = c \times \kappa(z, N_{\text{H,Gal}} = 10^{20.25}, \Gamma = 1.9), \quad (7)$$

where $\kappa(z, N_{\text{H,Gal}}, \Gamma = 1.9)$ is the redshift-dependent conversion factor. To calculate these conversion factors we adopt the Cycle 16 ACIS-I on-axis response function, corresponding to the approximate middle of the Legacy survey observing period. Variations of this response over the observation period of the C-COSMOS and COSMOS Legacy surveys (Cycles 9–17), mainly due to degradation of the sensitivity at low energies due to radiation damage and molecular contamination of the detector (see e.g. Grant et al. 2024; O’Dell et al. 2007; Cameron et al. 2002,⁹), as well as changes in relative response at different off-axis angles are captured by the effective exposure maps. We note that the effective exposure maps—and thus the count rates, c —are calculated assuming a simpler $\Gamma = 1.4$ spectral model (see Section 2.2.1), but our conversion factors incorporate the differing spectral assumptions when converting from a count rate directly into luminosities.

We determine L_X for the sources in our high-redshift samples using both z_{gal} and z_{QSO} . As can be seen in Figure 4, sources can move

⁸ We note that the spectral model used here is relatively simple and does not correctly model intrinsically obscured X-ray sources. To investigate the impact of obscuration on the XLF measurements in Section 4, we also use a more complex spectral model from Baloković et al. (2018) which includes the impact of both photoelectric absorption and Compton scattering of high energy photons by a dusty “torus” component.

⁹ See also The Chandra Proposers’ Observatory Guide v27.0 <https://cxc.harvard.edu/proposer/POG/html/chap6.html>

in and out of our redshift range depending on the choice of photometric redshift, as well as changing their estimated L_X (although the change in the number of sources at a given redshift has a more significant effect on our XLF measurements). In §3.4 we investigate the impact of this uncertainty on the measured XLF.

3.2 XLF measurements from the blind source catalogue

For our initial measurements of the XLF, we first consider only the more secure (lower false probability threshold) blind X-ray detected sources in the primary (i.e. $z_{\text{best}} > 4$) sample, consisting of 21 sources. Our sample size remains too small to perform a secure fit to a parametric model of the XLF. Thus, we divide these 21 $z \geq 4$ sources into redshift bins of $z = 4-5$, $5-7$ and $7-10$, corresponding to the highest redshift where parametric XLF models have been constrained in prior studies, the epoch of cosmic reionisation and extremely high redshifts, respectively. The sources within each of these redshift ranges are then split into 0.5 dex wide luminosity bins from $\log L_{X(2-10\text{keV})} \text{ erg s}^{-1} = 43-46$, which allows the shape of the X-ray Luminosity Function to be seen across each redshift range from binned measurements.

We estimate the space density of AGN, in the form of the XLF ($\phi_{\text{obs}}(L_{X,i}, z_j)$), in each luminosity ($L_{X,i}$) and redshift (z_j) bin, using the $n_{\text{obs}}/n_{\text{mdl}}$ method of Miyaji et al. (2001). This method accounts for changes in the X-ray sensitivity and the shape of the underlying XLF within our comparatively broad redshift and luminosity bins, providing a more robust estimate of the XLF compared to more simplistic methods (e.g. $1/V_{\text{max}}$ methods) given the low source numbers within our sample.

To obtain binned XLF measurements, the predicted value of the XLF, as given by a pre-existing parametric model evaluated at the centre of each bin, is scaled by the ratio of the number of observed sources in that bin to the number of sources predicted by the model:

$$\phi_{\text{obs}}(L_{X,i}, z_j) = \phi_{\text{mdl}}(L_{X,i}, z_j) \frac{n_{\text{obs}}}{n_{\text{mdl}}} \quad (8)$$

where n_{obs} is the observed number of sources in each bin. The parameterised XLF model, $\phi_{\text{mdl}}(L_{X,i}, z_j)$, is taken to be the value of the Luminosity Dependent Density Evolution (LDDE) XLF model of Georgakakis et al. (2015) at the centre of the $z - L_X$ bin (i.e. at $L_{X,i}$ and z_j). Despite requiring a parametric model from previous XLF studies, the $n_{\text{obs}}/n_{\text{mdl}}$ method is only marginally sensitive to the specific choice of model; the larger uncertainties due to our low source numbers dominates here. Finally, the predicted number of sources in the bin, n_{model} , is given by

$$n_{\text{mdl}} = \int_{z_1}^{z_2} \int_{\log L_{X_1}}^{\log L_{X_2}} \phi(L_X, z) A(c|L_X, z) \frac{dV_{\text{co}}}{dz} d \log L_X dz, \quad (9)$$

where $\phi(L_X, z)$ is the parametrised XLF model that varies across the $L_X - z$ bin, $\frac{dV_{\text{co}}}{dz}$ is the differential comoving volume at a given redshift, z , and $A(c|L_X, z)$ is the sky area imaged by Chandra to a given count rate, c , corresponding to a given 2–10 keV rest-frame luminosity, L_X , and redshift, as determined in §2.2.2 with a false probability threshold of 4×10^{-6} (as used in the blind sample creation). For each redshift and luminosity interval, error bars are given by the 1σ equivalent uncertainties (i.e. 68.3% confidence interval) on the number of sources based on Poisson uncertainties in the source numbers (Gehrels 1986) folded through the calculation of the observed XLF. Where there are no observed sources in a bin (i.e. $N_{\text{obs}} = 0$) we place an upper limit on the XLF based on the upper 1σ -equivalent limit for a sample size of 0. The resulting observed, binned XLF measurements are shown in Figure 5.

For comparison, we also plot the parametric models of Georgakakis et al. (2015) (LDDE model), Ueda et al. (2014) and Pouliaxis et al. (2024) that are based on fits to data at redshifts of $z = 3-5$, $z = 0.002-5$ and $z = 3-6$ respectively, extrapolated here when necessary to the central redshift of each bin; $z = 4.5$, $z = 6$ and $z = 8.5$ respectively. The uncertainties in these models are obtained through Monte Carlo simulations using the quoted XLF model parameter values and their uncertainties.¹⁰

At $z = 4-5$ our observed XLF measurements are consistent with the parametric models, which have been fit to data including this redshift range. However, as the redshift increases, our measurements of the XLF can be seen to remain at high space densities, with only a slight decrease from the values at $z = 4-5$ towards $z = 7-10$, whilst the extrapolated models predict a clear drop in the space density. These results indicate a less steep evolution in the space density of AGN in the early Universe than has been predicted from parametric models, with more sources found at higher redshifts than previously expected.

Despite the obvious trends revealed by our measurements, there remain large uncertainties due to the small number of sources. To improve the constraints further, in the next sub-section we combine the full sample of blind detections and directly extracted sources to improve our constraints.

3.3 XLF measurements from both blind and extracted X-ray samples

In order to fully exploit the deep X-ray data available from the *Chandra* COSMOS-Legacy survey, and improve the statistics of our measurements of the high-redshift AGN X-ray Luminosity Function, we extracted X-ray counts at the positions of all COSMOS2020 sources. A sample of X-ray AGN is then determined from the combination of the blind X-ray detected sources and the directly extracted X-ray measurements (see §2.4 for details of the sample creation). By using the known positions of sources based on pre-existing, multi-wavelength imaging, we are able to increase our sample of high-redshift sources and include lower significance detections without introducing substantial contamination due to false-positive detections. Following the process detailed above (in section 3.2), we obtain the blind+extracted sample measurements of the XLF using this new X-ray AGN catalogue.

The rest-frame 2-10 keV luminosity (L_X) of each source within this combined catalogue is based on the X-ray count rates from the blind source catalogue, where available; otherwise the directly extracted X-ray measurements are used. We measure the binned XLF based on this combined sample of $z \geq 4$ X-ray detected sources as described above for the blind X-ray detected source sample only. As a less stringent false probability threshold was used in the creation of the blind+extracted X-ray sample ($f_{\text{prob,thresh}} = 3.38 \times 10^{-5}$), the area curve for the *Chandra* imaging is recalculated (see §2.2.2 for details of the calculation). This new area curve is used to determine the predicted number of sources (Equation 9). The resulting measurements of the high-redshift XLF are shown in figure 6.

With the less stringent false probability threshold possible for this combined sample, we have improved the sensitivity of the X-ray observations and increased the area of sky visible by Chandra at lower luminosities. This new sensitivity has reduced the upper limits on the

¹⁰ We note that propagating the quoted uncertainties in each parameter when extrapolating these prior XLF measurements will not fully capture co-variances between parameters.

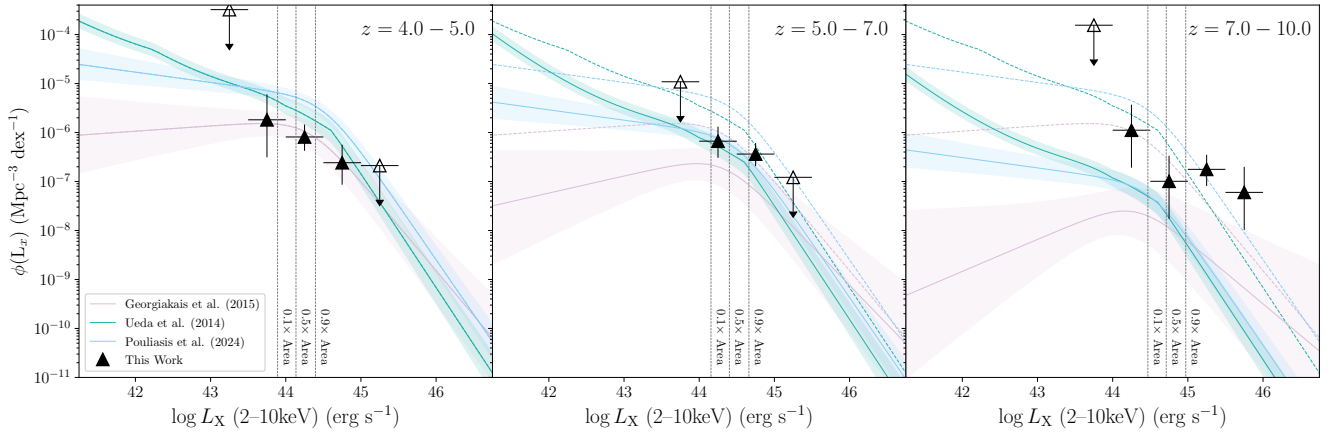


Figure 5. Binned measurements of the high redshift XLF of AGN (black triangles), based on the cross-matched blind source sample and the best-fitting photometric redshift estimates (i.e. the primary high-redshift sample). At $z = 4 - 5$ (left hand panel) our measurements are consistent with the model XLF of Georgakakis et al. (2015) (pink), Ueda et al. (2014) (green) and Pouliaxis et al. (2024) (blue), which were fit to data within the same redshift range. At $z = 5 - 7$ our measured space densities are higher than predicted by the extrapolated model of Georgakakis et al. (2008) but consistent with Ueda et al. (2014) and Pouliaxis et al. (2024), while at $z = 7 - 10$ (right hand panel) we measure significantly higher space densities than predicted by any of the model extrapolations. The luminosities corresponding to the *Chandra* sensitivity being achieved over 10%, 50% and 90% of the total survey area are indicated by the vertical dashed lines. Upper limits are provided by the upper Poisson 1σ uncertainty on 0 sources, as given by Gehrels (1986). For comparison the XLF parametric models evaluated at $z = 4.5$ (i.e. as shown in the $z = 4 - 5$ to panel) are shown in the $z = 5 - 7$ and $7 - 10$ panels by the dashed pink, green and blue lines.

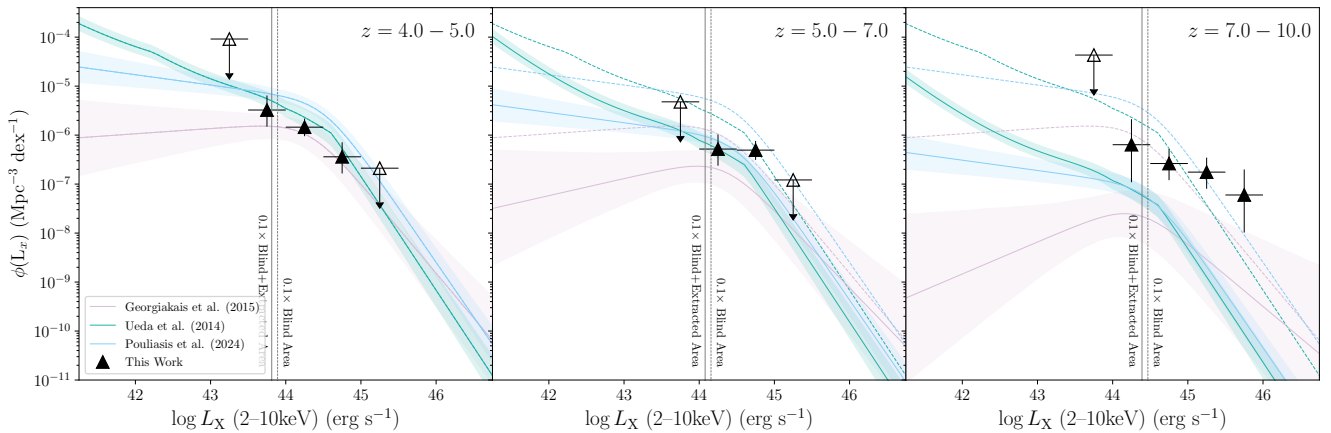


Figure 6. As for Figure 5, with sources from both the blind source detection and directly extracted counts. Consistent with models at $z = 4 - 5$, our binned XLF measurements begin to hint at a higher than predicted space density at $z = 5 - 7$ whilst clearly indicating the space density of AGN is higher than predicted by extrapolations of parametric models at $z = 7 - 10$.

space density at the faint-end of our observed XLF measurements, allowing for tighter constraints on the faint-end of the high-redshift XLF.

Due to the larger sample size used here, we have increased the number of sources at each redshift interval, and thus reduced the uncertainty in the observed XLF measurements. As can be seen in figure 6, the space density at $z = 4 - 5$ increases slightly from that in figure 5, to be above that of Georgakakis et al. (2015), whilst remaining consistent with all extrapolated models shown. These improved observed XLF measurements remain significantly higher than the model predictions at $z = 7 - 10$ and continue to indicate a higher space density at $z = 5 - 7$ than previously predicted. We also note the relatively flat slope of our $z = 7 - 10$ measurements over the

full $L_X \approx 10^{44-46}$ erg s^{-1} range, with no evidence of a break at $L_X \sim 10^{44.5}$ erg s^{-1} seen in the model extrapolations.

3.4 Impact on the XLF measurements from differing photometric redshift estimates

Our adopted photometric redshifts were obtained by fitting different combinations of templates to the observed Spectral Energy Distributions (SEDs) of sources in the COSMOS2020 catalogue. The choice of templates and fitting process can impact the measured redshift, in particular the inclusion of an AGN component to the SED can drastically alter the resulting redshift measurement. The LePhare code was used to determine a redshift for each source in the COSMOS2020 catalogue both with and without an AGN component in the templates;

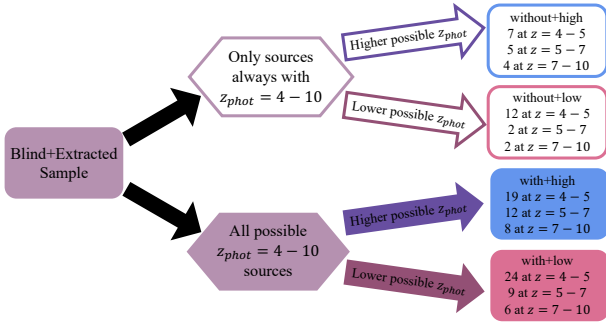


Figure 7. Flow chart illustrating the creation of the four photometric-redshift version dependant high-redshift samples, from our blind+extracted sample, with+low (solid pink), with+high (solid blue), without+low (empty pink) and without+high (empty blue). The number of sources in each redshift interval are shown for the four sample variants.

these fits for all sources in our high-redshift samples are shown in Appendix A. As such, all sources within our blind+extracted X-ray sample have two possible photometric redshifts, with the choice of SED fit having the potential to dramatically alter the measured XLF.

As can be seen in Figure 4, the photometric redshift of a source when found with AGN templates can alter whether that source falls within the high-redshift sample, compared to its photometric redshift as given by galaxy-only templates. The different fits can also move sources between our redshift bins and alter the estimated L_X . For the previous measurements of the XLF, in §3.2 and §3.3, we used the redshift of the best (lowest χ^2) template fit, as described in §3.1. Here we investigate the impact of the different photometric redshifts in order to obtain the most conservative and most optimistic values for the XLF measurements based on our X-ray sample.

Considering both forms of the photometric redshift, from the galaxy-only templates and the AGN templates, we generate four different high-redshift samples. The creation of these samples is summarised in Figure 7. Some of the sources within our X-ray sample only fall into the range of $z = 4 - 10$ with either their galaxy or their AGN photometric redshift fits, but not the other. Thus, we consider here sources where *either* the galaxy photometric redshift, z_{gal} , or the AGN photometric redshift, z_{QSO} , is ≥ 4.0 , thus creating an inclusive sample of 39 possible $z \geq 4.0$ sources hereafter referred to as the “with” sample. In addition, we create a conservative high-redshift sample of 16 sources where we require that *both* z_{gal} and z_{AGN} are ≥ 4.0 (even if the redshift values themselves differ), and remove any sources that *might* be at lower redshifts (referred to as the “without” sample). For the sources where both z_{gal} and z_{AGN} places them within the $z = 4 - 10$ redshift range, we also consider the cases where the redshift of each source is taken to be the smallest of the two redshift values and the case where it is taken to be the largest value, hereafter referred to as the “low” and “high” samples respectively (that can be defined from both the “with” and “without” samples). The resulting distribution of source numbers for the four samples (with+low, with+high, without+low, without+high) are shown in figure 7.

The observed XLF measurements are determined for all four high-redshift sample variants following the process detailed in §3.2. The without sample can be expected to produce the most conservative XLF measurements; the without+high sample producing the lowest XLF measurements in the $z = 4 - 5$ range and the without+low

giving the smallest measurements in the $z = 7 - 10$ redshift range. Conversely, there will be a maximum XLF measurement in the $z = 4 - 5$ and $z = 7 - 10$ ranges when determined from the with+low and with+high samples respectively. The four forms of the XLF measurements are shown in figure 8.

At $z = 4 - 5$ the different XLF measurements all remain consistent with the models, whilst the most optimistic values suggest the XLF tends to higher space densities at the faint end. Towards higher redshifts, however, even the most conservative measurement indicates a higher space density of AGN than predicted by the extrapolations of models.

4 OBSCURED AGN AND OBSCURATION-CORRECTED MEASUREMENTS OF THE XLF

Obscuring material in both the host galaxy and torus of dusty material within the AGN will absorb lower energy, soft X-ray photons produced by the corona more readily than higher energy, hard photons, preventing them from escaping the system and thus reducing the observed soft-band counts from a source. As such, a more heavily obscured AGN will appear to produce relatively more counts in the hard-band than in the soft-band, as more low energy photons are absorbed. Thus, despite the same initial production of soft and hard photons, an obscured source will have a stronger signal at harder X-rays relative to its observed emissions in softer bands (see Hickox & Alexander 2018). In the early Universe, as fewer stars have formed, galaxies are expected to have a higher gas content (see e.g. Carilli & Walter 2013, for review) and thus early SMBHs are likely to be more heavily obscured due to galaxy-wide gas. This higher gas content may also increase the obscuration due to the torus around the SMBH as it is funnelled into the central regions of the galaxy, fuelling the growth of the SMBH that is expected to occur. Studies such as that of Vito et al. (2018) have found the obscured fraction of AGN within their samples to increase with redshift, consistent with these expectations.

A decrease in observed counts at Soft and Full band energies, due to obscuring material around the AGN, results in the intrinsic luminosity of sources being under estimated and can result in an AGN not being observed above the X-ray background. Thus, increased obscuration may result in AGN which have lower observed luminosities and thus reduce the measured space density of AGN. At the higher redshifts probed in this study, our observed X-ray bands probe relatively high rest-frame energies, mitigating the impact of obscuration, but sufficiently high levels of obscuration will still affect our observed count rates (and thus inferred luminosities), particularly for column densities in the Compton-thick regime. In this section we investigate the effect of obscuration in the high-redshift blind+extracted sample on our measurements of the X-ray Luminosity Function (see §4.1 and §4.2). We will then perform a simple correction to the measurements, in order to account for any obscuration present in our sample (see §4.3).

4.1 The Hard-Band XLF

Obscured sources are typically more easily detected at hard-band (2–7 keV) energies than in the soft-band (0.5–2 keV). However, the measurements of the XLF we performed above were done using the full-band (0.5–7 keV) only in order to maximise our sample size and sensitivity. Given the high redshifts of our sample the observed hard-band imaging from Chandra will be probing very high energies, of > 10 keV, and should thus be relatively immune to obscuration (for $N_{\text{H}} < 10^{24}$) compared to the full-band observations. Thus, a

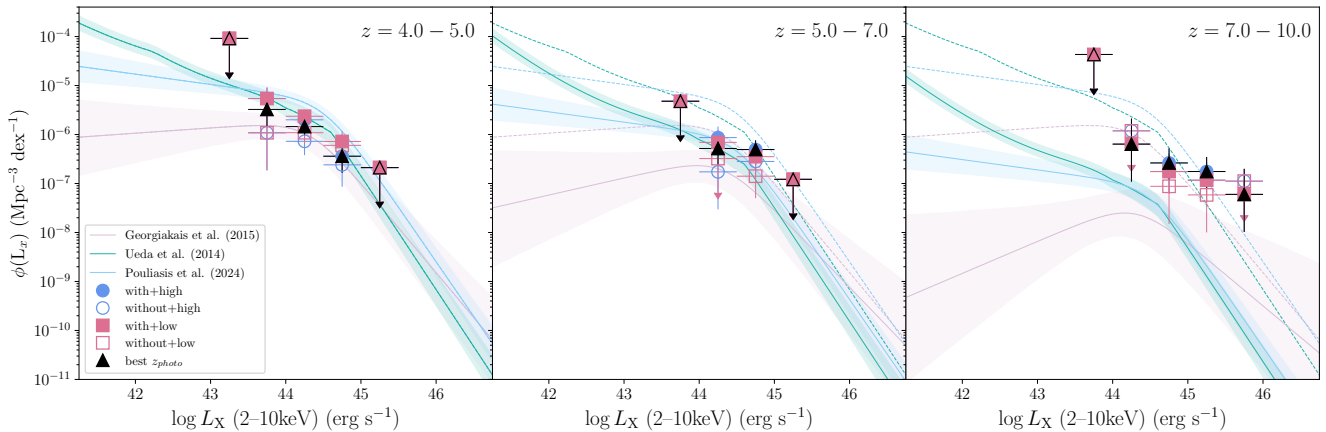


Figure 8. Binned XLF measurements at $z \geq 4$ determined from the blind and extracted sample using either the lowest (pink) or the highest (blue) possible photometric redshift, with (solid symbols) and without (empty symbols) sources for which only one photometric redshift measurement, given by the galaxy or AGN redshift fits, is $z \geq 4$. For comparison, the measured XLF from the best redshift values (as found in §2.4) is shown by the black points. Parametric models from Ueda et al. (2014), Georgakakis et al. (2015) and Pouliaxis et al. (2024) (green, pink and blue lines respectively) are shown extrapolated to the centre of each redshift range (solid) or evaluated at a fixed $z = 4.5$ (dashed). Even for the lowest measured XLF, providing the most conservative result, we find a higher space density than predicted by the parametric models at $z = 7 - 10$ and consistent with minimal evolution since $z \sim 4$.

first step to assess whether obscuration in our high-redshift sources has a significant effect upon the observed XLF measurements is to calculate the observed XLF from source detections made in the hard-band, which provides a more reliable tracer of L_X at the expense of sensitivity and sample size.

Using the sample of blind detected and hard-band extracted sources, obtained in §2.4 with a false probability threshold of 1.74×10^{-5} corresponding to a 5% False Fraction in the $z \geq 4$ sources, we calculate the observed XLF measurements given by the hard-band count rates. The measured hard-band count rates of our sources are converted directly to a rest-frame 2–10 keV X-ray luminosity (L_X , as in §3.1) and the measured XLF given by these hard-band observations is determined following the process detailed in §3.2. For these measurements, the sensitivity in the hard-band is calculated using the false probability threshold of 1.74×10^{-5} (see §2.2.2).

The resulting measurements of the hard-band XLF are shown by the blue diamonds in figure 9, with the original full-band measurements shown by the purple circles for comparison. The hard-band XLF measurements indicate a higher space density of AGN, in the early Universe, than is seen with the full-band measurements. This disparity between the full and hard band measurements can be seen to increase with redshift. This is indicative of an increasingly obscured population towards higher redshifts, as has been indicated by prior X-ray studies from Vito et al. (2018); Peca et al. (2023) and in recent discoveries of large obscured populations of AGN by JWST (Yang et al. 2023; Maiolino et al. 2024; Greene et al. 2024) which have generally remained undetected at X-ray wavelengths. With our hard-band XLF measurements indicating significant obscuration within the high-redshift sample, our full-band measurements of the XLF can be expected to be significantly affected by this obscuration. However, obscuration may vary between sources rather than affecting the sample in a uniform manner. As such, in order to further analyse the obscuration within our sample and the impact on our XLF measurements, we must determine the hardness ratios of each source in the full-band sample.

4.2 Hardness Ratios

A good indication of the amount of obscuration there is in an AGN, by material surrounding the black hole and accretion disk, is given by the ratio of the measured soft and hard counts of the source, known as the hardness ratio. Fewer soft-band photons will escape in a more obscured AGN and hence it will have a more positive hardness ratio than a less obscured AGN of the same intrinsic spectrum.

For all sources within our high-redshift blind+extracted sample (summarised in section 3.1) we determine their Hardness Ratios, following the Bayesian Estimation method detailed in Park et al. (2006). As detailed in §2.4, the total hard and soft-band counts of each source are taken to be the 70% PSF measurements obtained in the blind catalogue creation, where available, or extracted directly from the Chandra imaging at the position of the COSMOS2020 source. The posterior distribution of possible values of the Hardness Ratio, for each source in the sample, is then determined allowing for the Poisson nature of the hard and soft band counts. As such the Hardness Ratio of sources only detected in one energy band can still be determined and constraints can be placed on the level of obscuration of such a source. For this calculation we take the form of the Hardness Ratio (HR) to be,

$$HR = \frac{H - S}{H + S}, \quad (10)$$

where H and S are the total number of source counts measured in the hard and soft-bands, respectively. The Hardness Ratio for each source within our high redshift sample is thus calculated from the measured total counts in the hard and soft bands, with the corresponding background and effective exposure (that re-weights all HRs to account for the telescope vignetting and combination of Chandra observations from different epochs), using a Gaussian quadrature algorithm (as detailed in Appendix A.2 of Park et al. 2006).

The resulting hardness ratios for all $z \geq 4$ sources within our blind+extracted sample are shown in figure 10. We find a weak trend in the overall hardness of the population, with the majority of $z > 7$ sources lying at relatively high hardness ratios ($HR \gtrsim -0.25$). As can be seen in figure 10, there is significant scatter across the redshift

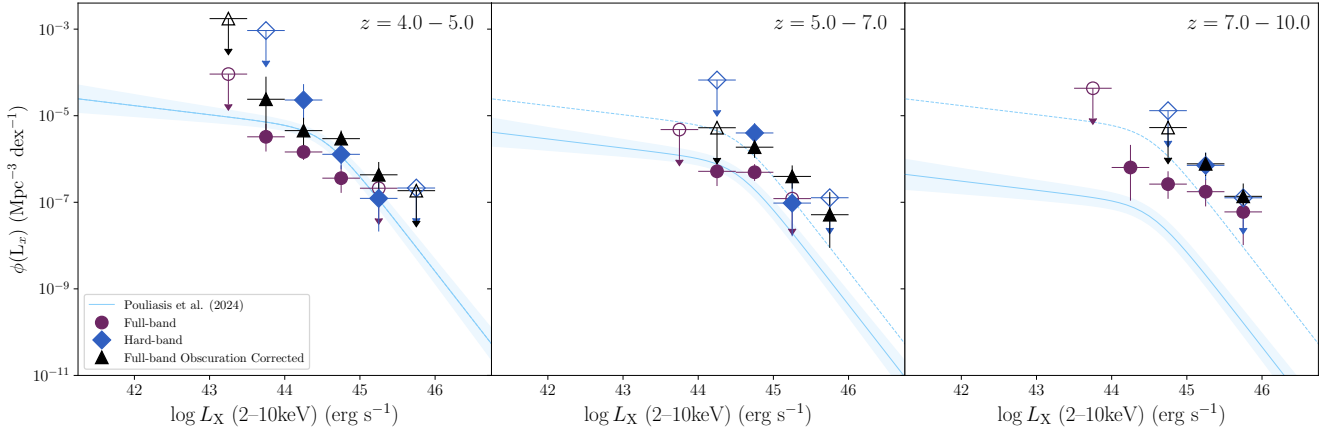


Figure 9. Full band obscuration corrected XLF measurements (black triangles; determined in §4.1) for the redshift bins of $z = 4 - 5$, $5 - 7$ and $7 - 10$. The uncorrected full band (purple circles; from figure 6), hard band (blue diamonds) and Pouliaxis et al. (2024) model extrapolations are shown for comparison. Our corrected measurements can be seen to be more consistent with the hard band measurements than the uncorrected case, indicating an even higher space density of AGN at these high-redshifts than is predicted by the model extrapolations.

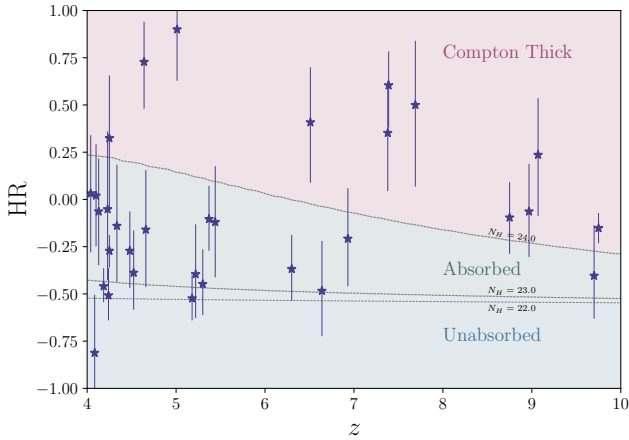


Figure 10. The Hardness Ratio (HR) with respect to the best-fitting redshift of all sources within our blind+extracted sample. The dotted lines indicated the expected value of HR based on the Borus spectral model (Baloković et al. 2018) with a column density of $N_{\text{H}} = 10^{22} \text{cm}^{-2}$, 10^{23}cm^{-2} and 10^{24}cm^{-2} , and are used to divide our sample into corresponding unabsorbed, absorbed and Compton Thick regions. It can be seen that at $z > 7$ the majority of sources lie within the Compton Thick region, however there remains a broad distribution of HR in the sample.

range of this study, indicating that a single obscuration correction applied to the XLF measurements would be inaccurate. However, it should be noted that the conversion between Hardness Ratio and the obscuration, N_{H} , of a source is dependent upon the redshift. This is due to the shift in the X-ray spectrum towards lower energies with redshift, and as such two sources of the same N_{H} but different redshifts, will have different HR values.

Previously our spectral models assumed negligible intrinsic obscuration, using the spectral model of Aird et al. (2015). Here, in order to account for the effect of obscuration on our observed source counts we employ the Borus spectral model of Baloković et al. (2018), which includes an obscuring torus component. Using this Borus model (folded through the Chandra response at approximately the

middle of the survey period as in section 3.1 to convert count rates into intrinsic rest-frame 2–10 keV luminosities)¹¹ we determine the Hardness Ratio across the redshift range of $z = 4 - 10$ corresponding to constant obscuration (i.e. constant values of N_{H}). With increasing redshift, the observed hardness ratio for a source of given N_{H} decreases, with a more rapid drop at higher N_{H} values (see figure 10). For obscuration below $N_{\text{H}} = 10^{23} \text{cm}^{-2}$ the Hardness Ratio remains roughly constant, with the track of $N_{\text{H}} = 10^{22} \text{cm}^{-2}$ showing no redshift evolution. As such we take unabsorbed¹² sources to be those of $N_{\text{H}} < 10^{23} \text{cm}^{-2}$ (given by their measured HR), while we define “absorbed” sources as those with Hardness Ratios corresponding to $10^{23} \text{cm}^{-2} \leq N_{\text{H}} \leq 10^{24} \text{cm}^{-2}$ and Compton-thick sources as those with Hardness Ratios corresponding to $N_{\text{H}} > 10^{24} \text{cm}^{-2}$, as indicated by the coloured regions in figure 10.

With Hardness Ratios determined for all sources in the blind+extracted sample, we split the sample into unabsorbed, absorbed and Compton-thick based on their measured HR. As can be seen in figure 10, there are fewer unabsorbed sources (those that lie below the $\log N_{\text{H}} = 23$ line on the HR- z plot) with increasing redshift, which is consistent with a more obscured population of AGN at higher redshifts. The resulting numbers of sources in each redshift range ($z = 4 - 5$, $z = 5 - 7$ and $z = 7 - 10$) and within each obscuration range (unabsorbed, absorbed and Compton-thick) are shown in Figure 11. The majority of our sources can be seen to fall within the absorbed and Compton-thick N_{H} ranges, with the majority of sources found to be absorbed (see also Vito et al. 2014).

It is important to note that the number of detected sources with different levels of obscuration will be strongly affected by selection biases. In particular, more obscured sources produce fewer observable counts and thus if they had the same *intrinsic* space density as less obscured sources they would be under-represented in observed samples, although the extent of the effect will depend on the ob-

¹¹ The degradation of Chandra’s sensitivity across its lifetime has negligible impact upon our results (see Barlow-Hall 2024), and are incorporated into the exposure maps used to obtain source counts.

¹² This interval will potentially contain some sources with a low level of absorption, however for the purpose of this work they shall all be assumed to be unabsorbed, as in Vito et al. (2018).

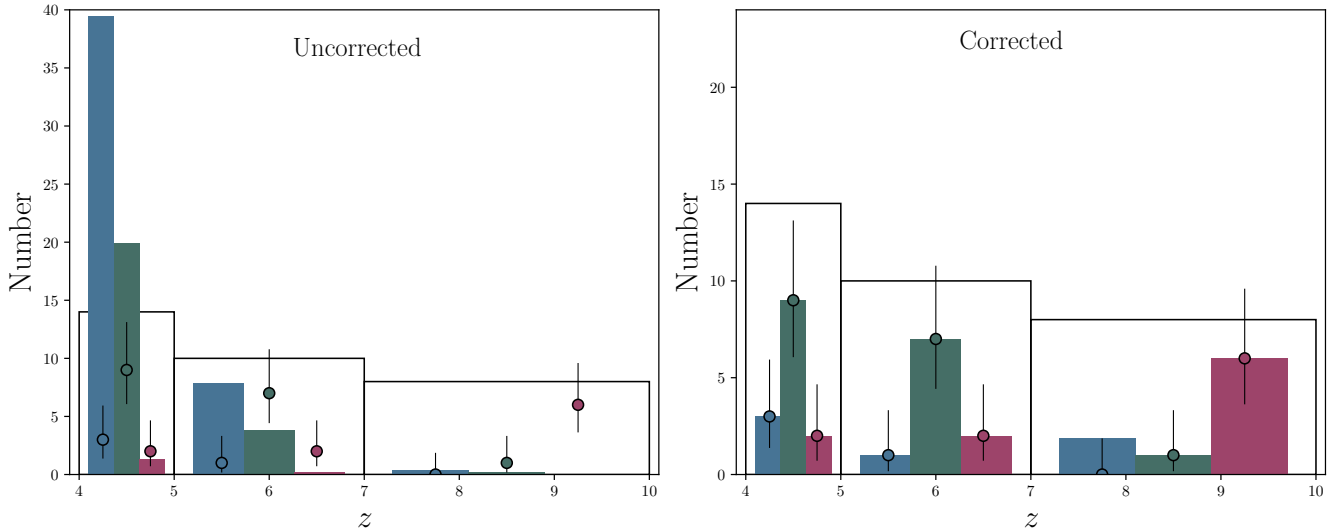


Figure 11. Histograms of model predicted source numbers (left) and corrected source numbers (right; see section 4.3) for unabsorbed (blue), absorbed (green) and Compton-thick (red) samples, with the observed number of sources shown by the points. Open histograms show full sample of observed sources before N_{H} binning, within each redshift interval. It should be noted that within the highest redshift bin ($z = 4 - 10$) the unabsorbed source number is corrected to the upper limit on 0 sources, as given by [Gehrels \(1986\)](#), and as such is above the observed number of sources.

served energy band. To explore this effect further, we determine the predicted number of sources in each redshift and obscuration bin, assuming all sources in a given obscuration bin have a space density given by the PDE XLF model from [Pouliasis et al. \(2024\)](#). To convert an intrinsic rest-frame 2–10 keV luminosity to an observed full-band count rate, we adopt the *Borus* spectral model evaluated at $N_{\text{H}} = 10^{22.5} \text{cm}^{-2}$, $10^{23.5} \text{cm}^{-2}$ and $10^{24.5} \text{cm}^{-2}$, corresponding to the approximate centre of each obscuration interval ($N_{\text{H}} < 10^{23}$, $10^{23} \text{cm}^{-2} \leq N_{\text{H}} \leq 10^{24} \text{cm}^{-2}$ and $N_{\text{H}} > 10^{24} \text{cm}^{-2}$ respectively). The corresponding count-rate is used to determine the area in equation 9 and folded through the XLF model to produce an expected number of sources in our survey area. We thus calculate the predicted number of sources within each redshift interval, summed across all possible luminosities¹³ for a given obscuration level. The resulting N_{H} dependent model predictions are shown by the coloured histograms in the left panel of figure 11.

The predicted number of sources overall decreases with redshift, as sources appear fainter at higher redshifts and thus are less likely to be detected, as well as due to the decrease in space density toward higher redshift in the [Pouliasis et al. \(2024\)](#) XLF model. The predicted numbers also drop rapidly with increasing N_{H} due to the suppression of the observed counts from such sources; only the brightest (and thus rarest) examples are expected to be detected. However, in the *observed* samples we find more absorbed ($N_{\text{H}} \approx 10^{23-24} \text{cm}^{-2}$, green data points in Figure 11, left) than unabsorbed ($N_{\text{H}} \lesssim 10^{23} \text{cm}^{-2}$, blue data points) at $z = 4 - 5$ and $z = 5 - 7$. At $z = 7 - 10$ we do not detect any unabsorbed sources, and the majority of the detected sources appear to be heavily obscured (Compton-thick), despite the strong selection biases *against* finding such sources.

The hardness ratios measured for our sample, and the comparison of predicted to observed source numbers at different levels of

obscuration, clearly indicates higher intrinsic space density of obscured sources at high redshifts and the need for corrections to the parametrisation of the XLF calculation to include both an N_{H} and redshift dependence.

4.3 Obscuration Corrections to the XLF

The study of the obscuration in the blind and extracted sample of high-redshift AGN, described above, indicates a significant redshift-dependent population of obscured sources, which is not accounted for in the model XLF. Thus, a correction to account for the disparity between the observed numbers of obscured sources and the predicted numbers (shown in the left hand plot of figure 11) needs to be applied to scale up the XLF and thus provide an accurate characterisation of the space density of high- z AGN.

With the broad distribution in Hardness Ratio with redshift, a correction factor that depends on both redshift and N_{H} is required to suitably account for obscuration within our sample. Given the limited size and dynamic range of our samples, we shall assume that the same correction factor applies at all L_{X} (i.e. we assume no luminosity dependence in the obscured fraction and that the underlying shape of the XLF is the same across all levels of obscuration). We determine a scale factor, $\gamma(N_{\text{H},\text{bin}}, z_{\text{bin}})$, that must be applied to our fiducial XLF model, $\phi_{\text{fidcl}}(L_{\text{X}}, z)$ (given by the [Pouliasis et al. 2024](#) PDE model) to recover a corrected model for the N_{H} and z bin,

$$\phi_{\text{mdl,corr}}(L_{\text{X}}, z | N_{\text{H}}) = \gamma(N_{\text{H},\text{bin}}, z_{\text{bin}}) \phi_{\text{fidcl}}(L_{\text{X}}, z). \quad (11)$$

This scale factor is defined such that we recover the observed number of sources in a given $N_{\text{H}}-z$ bin after folding through the appropriate area curve, i.e. such that $n_{\text{obs}}(N_{\text{H},\text{bin}}, z_{\text{bin}}) = n_{\text{mdl}}(N_{\text{H},\text{bin}}, z_{\text{bin}} | \phi_{\text{mdl,corr}})$. As in equation 9 for the calculation of the expected number of sources given an XLF model, we can write this as

¹³ We take the minimum luminosity of the range of possible X-ray luminosities as that given by the count rate observable at 0.1% of the total survey area.

$$\begin{aligned}
 n_{\text{mdl}}(N_{\text{H,bin}}, z_{\text{bin}} | \phi_{\text{mdl,corr}}) &= \\
 \iint \left[\gamma(N_{\text{H,bin}}, z_{\text{bin}}) \phi_{\text{fidcl}}(L_X, z) A(c | L_X, z, N_{\text{H,bin}}) \right] \frac{dV_{\text{co}}}{dz} dz d \log L_X \\
 &= \gamma(N_{\text{H,bin}}, z_{\text{bin}}) n_{\text{mdl}}(N_{\text{H,bin}}, z_{\text{bin}} | \phi_{\text{fidcl}}) \quad (12)
 \end{aligned}$$

with sky area $A(c | L_X, z, N_{\text{H,bin}})$ at an observed count rate, c , that depends on the redshift (z), rest-frame 2–10 keV luminosity (L_X) and obscuring column density ($N_{\text{H,bin}}$) and is calculated assuming the Borus spectral model (Baloković et al. 2018) and the *Chandra* response. The correction factor can thus be written as,

$$\begin{aligned}
 \gamma(N_{\text{H,bin}}, z_{\text{bin}}) &= \frac{n_{\text{mdl}}(N_{\text{H,bin}}, z_{\text{bin}} | \phi_{\text{mdl,corr}})}{n_{\text{mdl}}(N_{\text{H,bin}}, z_{\text{bin}} | \phi_{\text{fidcl}})} \\
 &= \frac{n_{\text{obs}}(N_{\text{H,bin}}, z_{\text{bin}})}{n_{\text{mdl}}(N_{\text{H,bin}}, z_{\text{bin}} | \phi_{\text{fidcl}})}, \quad (13)
 \end{aligned}$$

where $n_{\text{obs}}(N_{\text{H,bin}}, z_{\text{bin}})$ is the observed number of sources within a given obscuration range ($N_{\text{H,bin}}$) for a given redshift bin, and $n_{\text{mdl}}(N_{\text{H,bin}}, z_{\text{bin}} | \phi_{\text{fidcl}})$ is the expected total number of sources (across all L_X) in an $N_{\text{H}}-z$ bin assuming our fiducial XLF model. We calculate the scale factor for each z bin for the unabsorbed, absorbed and Compton-thick N_{H} bins, assuming the shape and redshift evolution of the XLF remains that of the fiducial model within the bin. The obscured column density is taken to be $N_{\text{H}} = 10^{22.5}$, $10^{23.5}$ and $10^{24.5} \text{ cm}^{-2}$, for unabsorbed, absorbed and Compton-thick sources, respectively, as a representative value for the conversion to count rates. Figure 12 shows the fiducial and corrected XLF models at $z = 4 - 5$, $5 - 7$ and $7 - 10$, for unabsorbed, absorbed and Compton-thick sources. As previously in our measurements of the XLF, in the case of no observed sources, we take the Poisson upper limit from Gehrels (1986) on 0 sources in place of the observed number.

Applying the resulting redshift and obscuration dependent correction factors, $\gamma(N_{\text{H,bin}}, z_{\text{bin}})$, to the predicted numbers scales them to agree with the observed number of sources in each redshift and obscuration interval, as shown in the right hand plot of figure 11.

Our corrected XLF model matches the observed binned XLF well across all three redshift ranges and for the different levels of obscuration (see Figure 12). With increasing obscuration, an increasingly positive shift in the normalisation of the fiducial model is required to agree with the observations (in contrast, the fiducial model XLF over predicts the space density of unabsorbed sources and the corrected model is shifted down). It should be noted that at $z = 7 - 10$ there are no unabsorbed sources within our sample, and as such the model can only be corrected to the upper limit on the binned XLF, as given by the Poisson uncertainty of Gehrels (1986), and the shape traced by these limits depends mainly on the sensitivity and area coverage of the *Chandra* data used here.

Having obtained the correction factors and confirmed they suitably adjust the model XLF for the obscuration observed, we can obtain the obscuration-corrected, binned measurements of the XLF across all N_{H} using the entire full-band blind+extracted sample. Equation 8 is updated to

$$\begin{aligned}
 \phi_{\text{obs,corr}}(L_{X,i}, z_j) &= \phi_{\text{mdl,corr}}(L_{X,i}, z_j) \frac{n_{\text{obs}}(L_{X,i}, z_j)}{n_{\text{mdl}}(L_{X,i}, z_j)} \\
 &= (\gamma_{\text{unabs}} + \gamma_{\text{abs}} + \gamma_{\text{ctk}}) \phi_{\text{fidcl}}(L_{X,i}, z_j) \times \\
 &\quad \frac{n_{\text{obs}}(L_{X,i}, z_j)}{n_{\text{mdl,unabs}} + n_{\text{mdl,abs}} + n_{\text{mdl,ctk}}} \quad (14)
 \end{aligned}$$

z	γ_{unabs}	γ_{abs}	γ_{ctk}
4–5	$0.10^{+0.08}_{-0.05}$	$0.38^{+0.19}_{-0.13}$	$0.9^{+1.2}_{-0.6}$
5–7	$0.5^{+0.4}_{-0.3}$	$1.1^{+0.8}_{-0.5}$	4^{+6}_{-3}
7–10	< 4.73	$5.7^{+13}_{-0.5}$	217^{+130}_{-86}

Table 1. Obscuration corrections to the Pouliaisis et al. (2024) PDE XLF model for unabsorbed (γ_{unabs}), absorbed (γ_{abs}) and Compton-thick (γ_{ctk}) sources determined here.

where γ_{unabs} , γ_{abs} and γ_{ctk} are the obscuration correction factors for the z bin, for the N_{H} bins of unabsorbed, absorbed and Compton-thick sources respectively (see table 1 for values). The predicted number of unabsorbed ($n_{\text{mdl,unabs}}$), absorbed ($n_{\text{mdl,abs}}$) and Compton-thick ($n_{\text{mdl,ctk}}$) sources in each luminosity–redshift bin are calculated using equation 12 but with the corresponding integration limits, and summed to recover the overall predicted number of sources in the L_X-z bin. The resulting observed XLF measurements are shown by the black triangles in figure 9, with the full-band blind+extracted XLF before absorption correction (from §3.3) and the hard-band blind XLF (from §4.1) plotted for comparison.

With the obscuration correction, our Full-band measurements more closely follow those of the Hard-band XLF measurements, showing a higher space density across all redshift ranges than the uncorrected Full-band measurements. At $z = 4 - 5$ these corrected measurements remain broadly consistent with the prior model predictions, as in section 3.3. However, even at $z = 5 - 7$ our measurements suggest a higher space density than the extrapolated models, and the models drastically under-predict the sample observed at $z = 7 - 10$.

Across the full redshift range of our sample ($z = 4 - 10$) and within each redshift bin ($z = 4 - 5$, $5 - 7$ and $7 - 10$) we use our obscuration correction factors (γ_{unabs} , γ_{abs} and γ_{ctk}) to determine the intrinsic fraction of obscured (i.e. $N_{\text{H}} > 10^{23}$) AGN, f_{obsc} ,

$$f_{\text{obsc}} = \frac{\gamma_{\text{abs}} + \gamma_{\text{ctk}}}{\gamma_{\text{unabs}} + \gamma_{\text{abs}} + \gamma_{\text{ctk}}}, \quad (15)$$

and the intrinsic fraction of Compton-thick AGN, f_{ctk} ,

$$f_{\text{ctk}} = \frac{\gamma_{\text{ctk}}}{\gamma_{\text{unabs}} + \gamma_{\text{abs}} + \gamma_{\text{ctk}}}. \quad (16)$$

Uncertainties are determined through random posterior draws of possible “true” values of the fraction. From this the fraction of obscured and Compton-thick AGN across the redshift range of $z = 4 - 10$ is $0.982^{+0.007}_{-0.008}$ and $0.91^{+0.03}_{-0.03}$ respectively.

Across the three redshift bins ($z = 4 - 5$, $5 - 7$ and $7 - 10$) used in this work, we find the fraction of obscured ($N_{\text{H}} > 10^{23}$) sources increases with redshift with the fraction of Compton Thick sources becoming increasingly dominant towards earlier times; increasing from $0.66^{+0.15}_{-0.13}$ at $z = 4 - 5$ to $0.80^{+0.13}_{-0.04}$ at $z = 5 - 7$ and $0.97^{+0.02}_{-0.01}$ at $z = 7 - 10$. This can be seen in figure 13. Whilst the measured fraction of Compton Thick sources is roughly consistent with previous results (Pouliaisis et al. 2024; Vito et al. 2018), this fraction rapidly increases with redshift as found by Gilli et al. (2022) who suggest such levels of obscuration may be due to the high gas content throughout the host galaxies of early AGN.

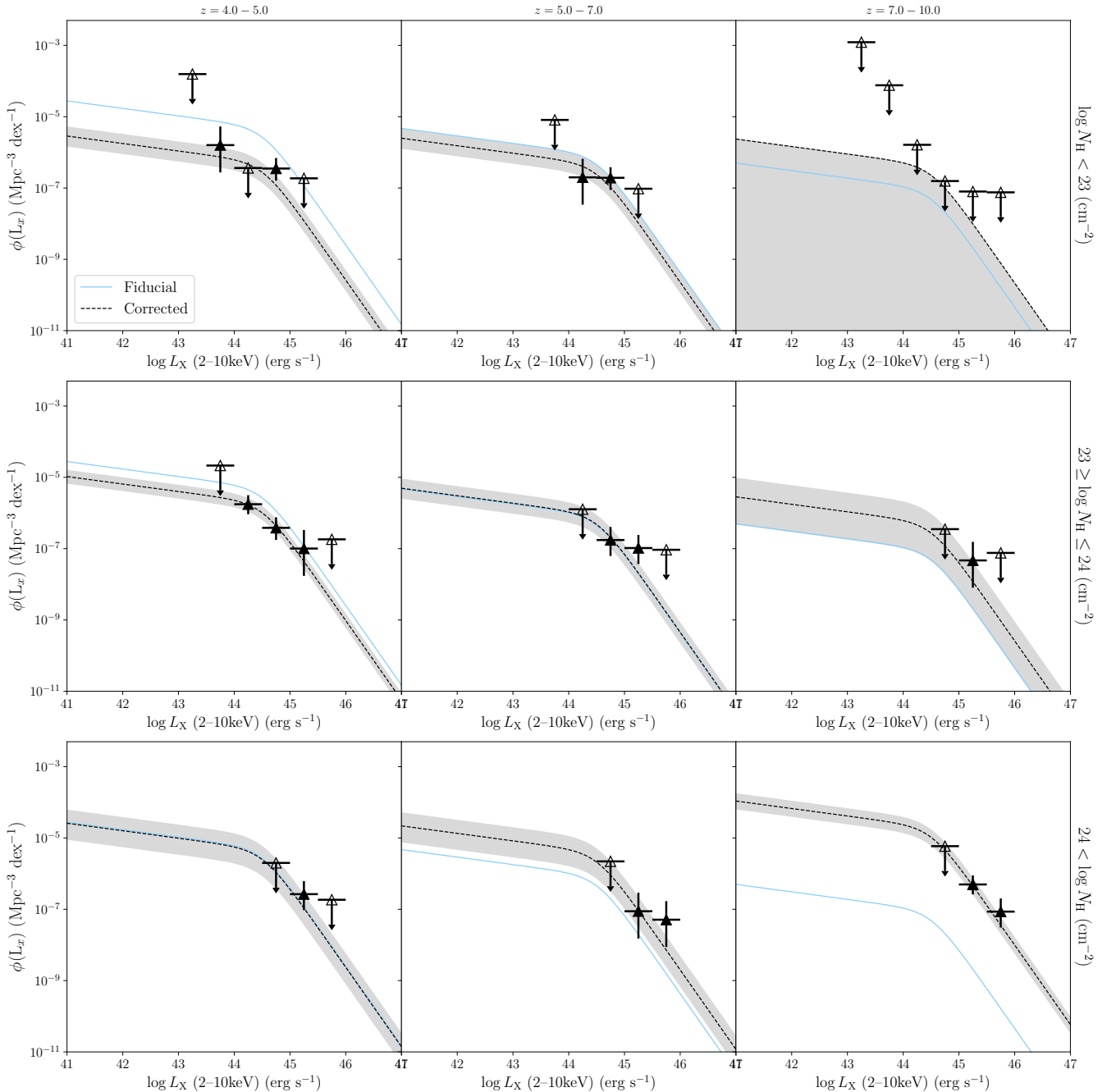


Figure 12. The [Pouliasis et al. \(2024\)](#) XLF model before (solid line) and after (dashed line) correction to our obscuration binned measurements, for $N_{\text{H}} < 10^{23}$ (top row), $10^{23} \text{cm}^{-2} \leq N_{\text{H}} \leq 10^{24} \text{cm}^{-2}$ (middle row) and $N_{\text{H}} > 10^{24} \text{cm}^{-2}$ (bottom row). Our measured XLF are shown for comparison by the black triangles, with limits given by [Gehrels \(1986\)](#) where no sources exist within our sample (open triangles). The Obscuration correction can be seen to reproduce the shape of the XLF well, as seen with our blind+extracted measurements. For the case of $N_{\text{H}} < 10^{23}$ at $z = 7 - 10$ the blind+extracted sample contains no sources, thus the obscuration correction can only provide an upper limit on the number of sources, based on the survey area of the sample.

5 DISCUSSION

5.1 Comparisons to previous measurements of the XLF at high redshift

The AGN X-ray Luminosity Function (XLF) has previously only been measured to $z \sim 6$ (e.g. [Georgakakis et al. 2015](#); [Pouliasis et al. 2024](#); [Vito et al. 2018](#)), to which parametric models were then

fitted, due to the limitations on high-redshift X-ray surveys (i.e. the small area observable to the required depth with current telescopes, observed decline in space density of AGN with increasing redshifts, and the difficulty of identifying the counterparts to X-ray AGN at high redshifts) and the XLF beyond $z = 6$ has remained poorly constrained. With new deep near-IR imaging from UltraVISTA, incorporated within the COSMOS2020 multi-wavelength catalogue, it

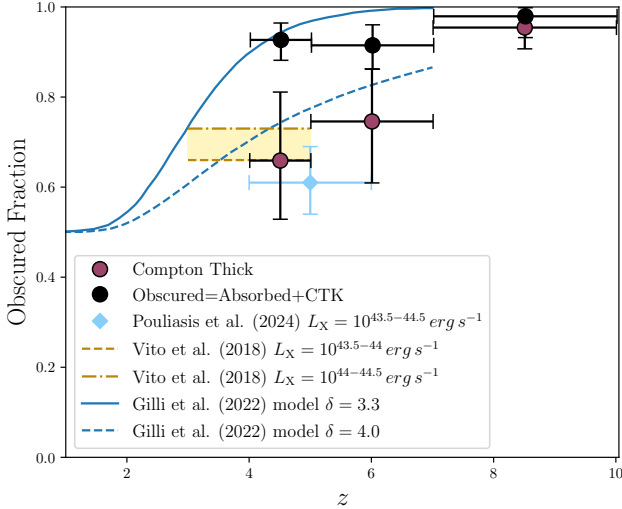


Figure 13. Fraction of the blind+extracted sample that are Compton Thick (pink points) and the fraction that are obscured (black points), across the full luminosity range of the sample. Previous observational measurements of the obscured fraction from Vito et al. (2018) for $L_X = 10^{43.5-44.5} \text{ erg s}^{-1}$ (dashed) and $10^{44-44.5} \text{ erg s}^{-1}$ (dot-dashed) are shown by the gold region and the measurement from Pouliaxis et al. (2024) is shown by the red diamond. Our Compton Thick fraction is consistent with these observational measurements at $z = 4 - 5$, but shows a clear increase with redshift. The high obscured fraction measured from our blind+extracted sample is indicative of the large fraction of obscured sources in the early Universe which has been predicted by the models of Gilli et al. (2022) based on the expected column densities of gas within the host galaxies of high- z AGN.

is possible to measure redshifts for sources at $z > 6$ that we have matched with X-ray sources in the *Chandra*-COSMOS Legacy imaging and used to identify additional lower significance X-ray sources. We have thus obtained measurements of the XLF at moderate luminosities out to $z = 10$ that demonstrate substantially higher space densities of early X-ray AGN than predicted by the extrapolations of prior parametric models.

The work performed in this paper relies on photometric redshift measurements. As investigated in section 3.4, the choice of templates used in the photometric redshift fitting adds an uncertainty to the XLF measurements. As such we determined the uncertainty range in our XLF measurements due to the uncertainties intrinsic to photometric redshifts, shown in figure 8. We note that even in our most pessimistic scenario based on the differing photometric redshift solutions, we still find a significantly higher space density of X-ray AGN at $z = 7 - 10$ than expected based on prior model extrapolations. We also find evidence that a large fraction of the early X-ray AGN population is obscured; applying obscuration corrections to the full-band XLF measurements (see §4) further increases our measurements of the space density of high- z X-ray AGN. We now combine our obscuration corrected XLF and photometric redshift uncertainties, in order to determine the most optimistic and most conservative measurements of the XLF from our sample, which we then compare with constraints from prior studies.

For our most conservative estimate, we take the lower bound of the uncertainty range shown in figure 8 when considering the impact of photometric redshift uncertainties (N.B. these measurements do *not* include any correction for obscuration). We then take the upper uncertainty on the obscuration-corrected measurements as our most

optimistic, upper limit. The resulting range in our observed XLF measurements, at the redshifts of $z = 4 - 5$, $5 - 7$ and $7 - 10$, are shown in figure 14.

We note that applying the obscuration corrections (as in section 4.3) whilst also considering the range of possible redshifts for each source (as in section 3.4) would provide a more comprehensive uncertainty range of our XLF measurements. However, as there is significant obscuration across the full redshift range of the sample, this would increase the disparity between the model predictions and the most optimistic space densities at $z = 7 - 10$, whilst the most conservative measurements of the XLF would remain the value of the lower uncertainties found in section 3.4 for the uncorrected XLF. As such, with the obscuration corrections only applied to the best redshift measurements, the results shown in figure 14 may still be relatively conservative, rather than showing strict upper limits on the high-redshift AGN space density. We also recall that a number of X-ray sources in the blind source catalogue either lack reliable counterparts or have reliable counterparts but lack reliable photometric redshift estimates. Some of these X-ray sources may fall within our redshift range and would increase our upper limit further still.

In Figure 14 we compare these bounds on our XLF measurements with both parametric model extrapolations. Within the redshift range of $z = 4 - 5$, at which previous XLF models have been fitted to survey data, we find our measurements of the XLF to be consistent with the parametric models (as is to be expected) and thus suitably representative of the AGN population at the end of cosmic reionisation (see left hand panel of figure 14). However, the upper limits on our measurements, which account for the obscuration within the sample, suggest that the space density of fainter AGN could be substantially higher than predicted based on prior model extrapolations. Our higher space density estimates are due to the potentially substantial fraction of the population at these redshifts that are obscured or Compton-thick ($\geq 90\%$) that may not be fully accounted for in the prior studies or their parametrisations.

At higher redshifts (centre panel of figure 14), our measurements remain broadly consistent with the parametric model extrapolations at $z = 5 - 7$. Whilst the most conservative measurements are clearly consistent with the model predictions, our most optimistic values indicate that the true space density of AGN may be substantially higher. We also compare here with direct constraints on the bright end of the XLF by Barlow-Hall et al. (2023) using *Swift* (ExSeSS) and by Wolf et al. (2021) with eROSITA (eFEDs, determined for a narrower redshift interval of $z = 5.7 - 6.4$). We find that these while these previous estimates are consistent with our limits, our best estimate (solid black square) and upper bound indicate that the space density of $L_X \sim 10^{45-46} \text{ erg s}^{-1}$ AGN at $z \sim 5 - 7$ could be up to a factor ~ 100 higher than that measured by Barlow-Hall et al. (2023) and Wolf et al. (2021). It should be noted that the measurement of Barlow-Hall et al. (2023) will be incomplete, in particular for obscured AGN, and should thus be considered a lower limit, which could explain the substantial discrepancy with our new *Chandra*/COSMOS measurements.

Despite the agreement between our blind+extracted measurements and the extrapolations of parametric models out to $z = 7$, the extrapolated models can be seen to under-predict the observed XLF at $z = 7 - 10$ (see right hand plot of 14). Although the most conservative measurements of the observed XLF at this very high redshift is within 1σ of the Georgakakis et al. (2015) LDDE model predictions, the AGN space-density in the very early universe is likely much higher than previously expected. Although previous studies (e.g. Brusa et al. 2009; Vito et al. 2018) suggest the need for a high-redshift decline in

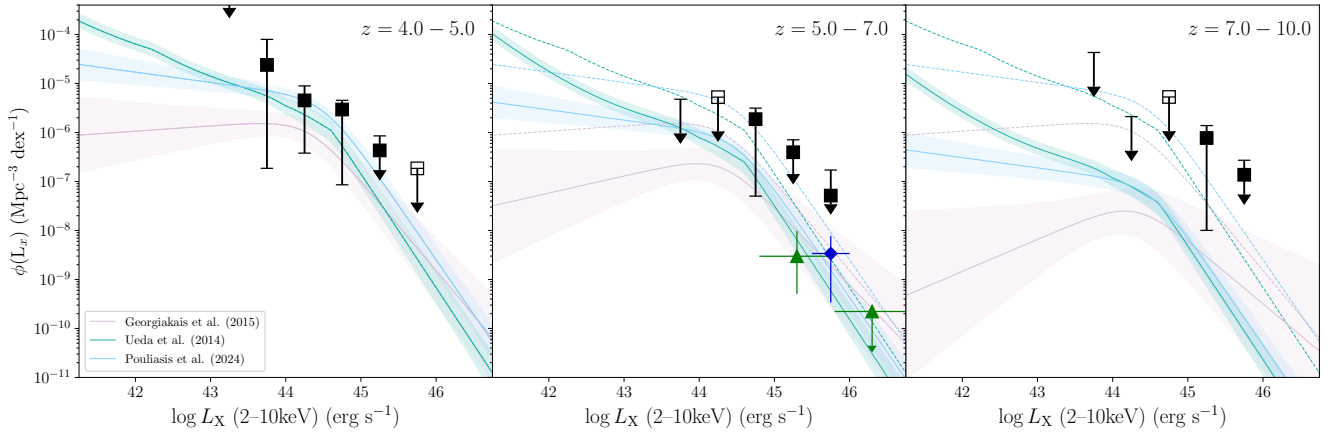


Figure 14. The bounds of the XLF measured in this paper, with lower bounds given by the lower bounds of the photometry dependent XLF determined in §3.4 and the upper bounds given by the upper uncertainty on the obscuration corrected XLF measurements (see section 4.3). Extrapolated models of XLF fit to redshift $z \sim 3 - 6$ data are plotted at the central redshift of each bin (i.e. $z = 4.5$, $z = 6$ and $z = 8.5$) for comparison, with the dotted lines indicating the space density given by the models at $z = 4 - 5$. The bright-end XLF measurements of Barlow-Hall et al. (2023) and Wolf et al. (2021) are shown by the green triangles and the blue diamond respectively.

the AGN space density, the results we have presented here indicate that there is no such decline in the AGN space density at $z \gtrsim 7$.

The space density of AGN we have determined here indicates a large number of actively accreting supermassive black holes within the very early Universe, which *are* detectable even with current X-ray telescopes, even if the high levels of obscuration (becoming predominantly Compton-thick at the highest redshifts that we probe) mean that their observable X-ray flux is strongly suppressed relative to their intrinsic luminosities. Our findings suggest that significant black hole growth may be occurring at early times, which we investigate further in §5.3. This early AGN population within the first galaxies would likely also have a significant impact upon the evolution of galaxies in the early Universe. In addition, as contamination of galaxy observations by AGN will significantly impact any measurements of galaxy properties, we stress the importance of accounting for the impact of AGN light on any measurements of galaxy properties at $z \gtrsim 4$ (as is accounted for in e.g. de Graaff et al. 2025; Trinca et al. 2024; Bowler et al. 2020).

5.2 Comparison with the bolometric QLF and new AGN populations identified with JWST

New, sensitive IR imaging enabled by JWST (along with follow-up spectroscopy of faint sources) has enabled the identification of surprising large populations of moderate- to low-luminosity AGN at redshifts above $z \sim 4$ (e.g. Greene et al. 2024; Matthee et al. 2024; Kokorev et al. 2024; Larson et al. 2023), with some studies claiming to have found early AGN out to $z \sim 10$ (Kovács et al. 2024). These early JWST results indicate a relatively high AGN occupation fraction within the first galaxies, with the majority of JWST studies pointing towards unexpectedly high AGN space densities within the early Universe (see e.g. Greene et al. 2024; Maiolino et al. 2024).

In order to compare our X-ray measurements of the $z \gtrsim 4$ AGN space density to the JWST findings, we convert our measurements of the XLF into estimates of the bolometric Quasar Luminosity Function (QLF), adopting the luminosity-dependent bolometric corrections of Shen et al. (2020). Applying the bolometric correction to the full-band, obscuration-corrected XLF measurements ob-

tained above (see section 4.3), we compare our new QLF estimates at $z = 5 - 7$ to measurements based on JWST data by Maiolino et al. (2024) (from UV luminosity function measurements at $z \sim 4 - 6$) and Greene et al. (2024) (see figure 15). Greene et al. (2024) performed their measurements of the QLF on the space density of Little Red Dots, based on photometric selection of Red Compact sources and spectroscopic broad lines, and thus may not capture all low-luminosity AGN. In contrast, Maiolino et al. (2024) use a less stringent selection of AGN, based simply on the identification of a broad $H\alpha$ or $H\beta$ line, and thus measure a higher space density than Greene et al. (2024) at low-luminosities. In addition to the JWST measurements, we also show the previous constraints on the bright-end QLF based on X-ray measurements from Barlow-Hall et al. (2023) and from a bolometric conversion of the Wolf et al. (2021) measurements of the XLF (converted using the Shen et al. 2020, bolometric correction).

As can be seen in figure 15, our obscuration-corrected measurements of the AGN space density follow a consistent trend between the early JWST results and X-ray observations at the bright-end of the luminosity function. As noted by Scholtz et al. (2025), the majority of AGN detected by JWST are found to have no detected X-ray emission even when image stacking is applied (see Maiolino et al. 2025b; Yue et al. 2024). This lack of X-ray detection points towards a potentially very high obscuration fraction (of the order of $100\times$ the fraction of unobscured sources, as in Maiolino et al. 2025b, and references there in), as is consistent with the obscured population inferred from our blind+extracted sample here ($\sim 220\times$, given by the obscuration correction to the model XLF). Despite this, other studies with JWST have successfully identified high- z AGN with X-ray detections (Kocevski et al. 2024; Bogdán et al. 2024; Napolitano et al. 2025). Pouliaxis et al. (2024) argued that X-ray selection is likely to miss all but a small fraction of the heavily obscured AGN population; precisely those sources being found by JWST. With our blind+extracted sample pushing the limit of current X-ray imaging, and probing higher luminosity AGN than the JWST imaging of studies such as Maiolino et al. (2024); Greene et al. (2024), it is likely that the AGN found by JWST are simply lower (intrinsic) luminosity equivalents to the population that we have identified here, which also

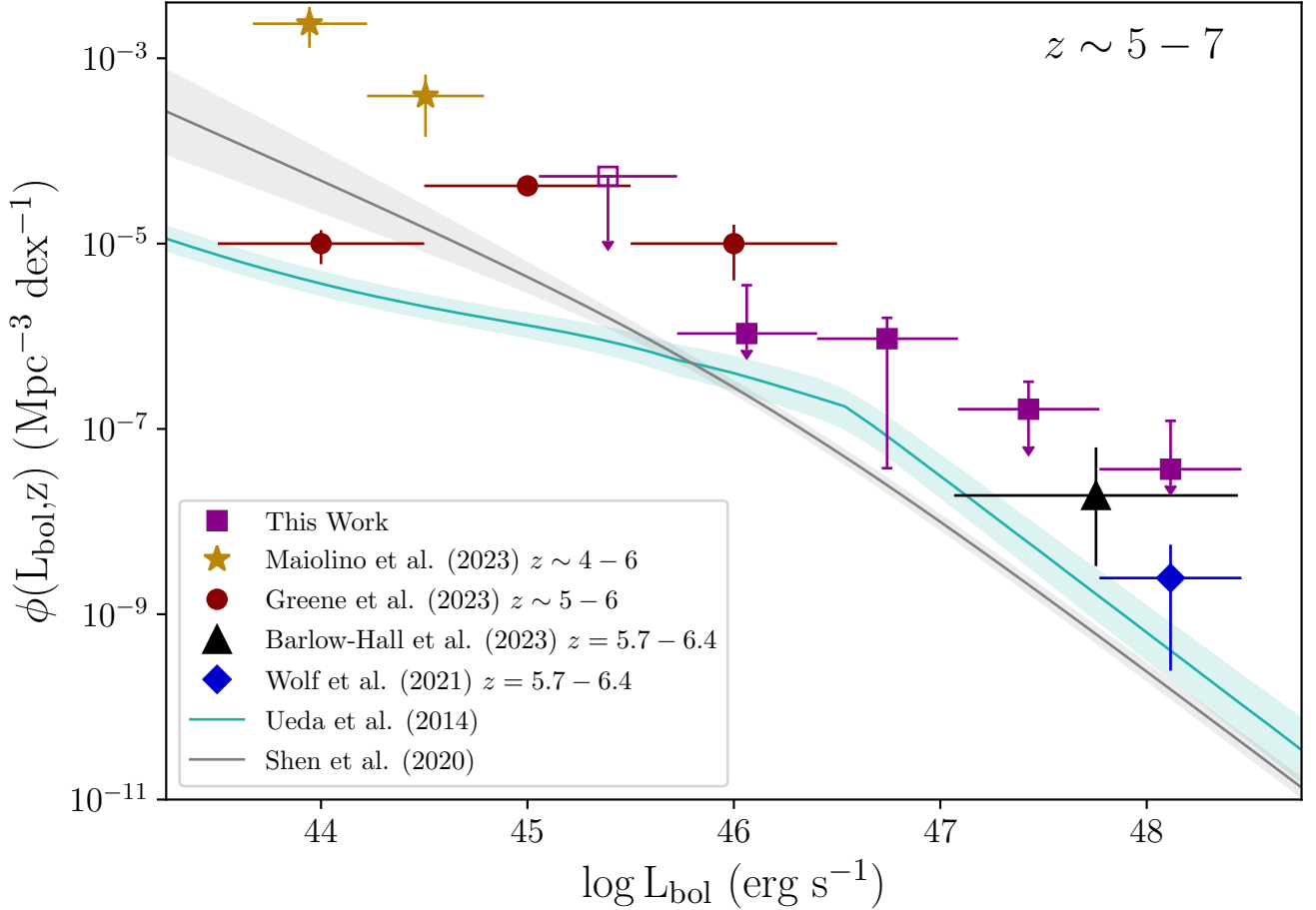


Figure 15. Our estimates of the bolometric QLF (purple squares) based on our final XLF measurements and converted to bolometric estimates based on the corrections from Shen et al. (2020). The solid squares indicate our best estimates, while the uncertainties show the range from the maximum obscuration-corrected measurement to the minimum unobscured measurements (as shown in figure 14, with arrows indicating conservative upper limits given the range of possible photometric redshifts). Bright-end $z = 5.7 - 6.4$ X-ray measurements are shown from the results of Barlow-Hall et al. (2023) (black triangle) and Wolf et al. (2021) (blue diamond). The JWST measurements at $z = 5 - 6$ from Greene et al. (2024) are shown by the red circles, and the $z = 4 - 6$ UVLF measurements of Maiolino et al. (2024), similarly converted to bolometric luminosities using the conversions of Shen et al. (2020), are shown by the gold stars. For comparison, the bolometric QLF model of Shen et al. (2020) and bolometric converted XLF model of Ueda et al. (2014) are plotted at $z = 5.7 - 6.4$ (grey and turquoise respectively). Our constraints on the QLF can be seen to bridge the gap between the latest JWST observations and the bright-end X-ray measurements, as well as being consistent with substantially higher space densities of low-to-moderate luminosity AGN in the early Universe than predicated by extrapolations of previous models.

exhibit high (often Compton-thick) levels of obscuration that suppresses their X-ray emission and makes them difficult to detect even in the very deepest X-ray imaging (or stacks). In order to investigate this further, spectra and JWST imaging of our high redshift sample are required to determine their physical nature and compare their properties with the population of compact AGN (known as Little Red Dots, e.g. Kocevski et al. 2024) found by JWST.

The moderate luminosity measurements of the AGN space density we have obtained in this work, clearly point towards a reconciliation of the JWST AGN population and those identified through X-ray surveys. With improved redshift measurements from the mid-IR imaging of JWST, even fainter X-ray sources could be identified within the Chandra Deep Fields (following the method detailed in section 2.4) and thus place further constraints on the faint-end of the XLF, where the majority of the AGN population resides. In order to truly reconcile these measurements, the impact of Compton-thick levels of

obscuration needs to be carefully taken into account. Only with even deeper X-ray observations than currently feasible with Chandra do we stand a chance of identifying JWST AGN at X-ray wavelengths, assuming they have some intrinsic X-ray emission due to the AGN, requiring the next generation of X-ray observatories such as AXIS (Reynolds et al. 2023; Cappelluti et al. 2024).

5.3 The Implications on Early Black Hole Growth

Our measurements indicate a substantially higher space density of moderate-luminosity AGN at $z \gtrsim 6$ than expected given earlier X-ray studies and in line with recent JWST findings. Previous measurements of X-ray selected AGN out to redshifts of $z \sim 6$ found a decline in the space density of AGN with increasing redshift from which studies such as Vito et al. (2014); Aird et al. (2015) determined the Black Hole Accretion Density (BHAD) to drop more rapidly with

redshift than the Star Formation Rate Density (SFRD) of galaxies (see also e.g. [Pouliasis et al. 2024](#); [Zhang et al. 2023](#); [Habouzit et al. 2021](#)). This trend suggests the build-up of stellar mass within galaxies may have occurred before their central SMBH underwent significant accretion-driven mass build-up. Using our XLF measurements we determine the corresponding BHAD in each of our redshift bins ($z = 4 - 5$, $5 - 7$ and $7 - 10$) following the process detailed in [Yang et al. \(2021\)](#) and [Pouliasis et al. \(2024\)](#), whereby

$$\begin{aligned} \text{BHAD} &= \frac{1 - \epsilon}{\epsilon c^2} \int L_{\text{bol}} \times \phi_{\text{bol}}(L_{\text{bol}}, z) d \log L_{\text{bol}} \\ &= \frac{1 - \epsilon}{\epsilon c^2} \int L_{\text{bol}} \times \phi_{\text{corr}}(L_X, z) \frac{d \log L_X}{d \log L_{\text{bol}}} d \log L_{\text{bol}} \quad (17) \end{aligned}$$

where we take the radiative efficiency, ϵ , to be 0.1 (as in [Pouliasis et al. 2024](#)), $\phi_{\text{corr}}(L_X, z)$ is our obscuration-corrected XLF determined in §4, and the $\frac{d \log L_X}{d \log L_{\text{bol}}}$ accounts for the change of variables between X-ray and bolometric luminosities. As our model assumes a declining XLF within each redshift bin, we calculate the BHAD across each redshift bin, for the bolometric luminosity range of $43 < \log L_{\text{bol}} < 49$ then average over the redshift bin to obtain the BHAD at the given redshifts.

Our resulting BHAD measurements (see figure 16) show an increase with redshift from the end of cosmic reionisation ($z \sim 5$) out to $z \gtrsim 10$. Our measurements are consistent with the BHAD that is estimated based on AGN populations identified with JWST ([Yang et al. 2023](#)) and in particular the high BHAD inferred from the large populations of Little Red Dots identified with JWST out to $z \sim 10$ ([Akins et al. 2025](#), red hexagons in figure 16). Unlike the X-ray measurements of [Vito et al. \(2014\)](#); [Aird et al. \(2015\)](#), the BHAD found here indicates that significant Black Hole mass assembly may have occurred *prior* to stellar mass build up. As such, BHs may build up a significant fraction of their mass before the stellar mass of their host galaxy is assembled, as has been postulated in the work of [Guetzoyan et al. \(2025\)](#); [Terrazas et al. \(2024\)](#). Mass measurements of early SMBHs also find over-massive SMBHs (e.g. [Maiolino et al. 2024](#); [Kocevski et al. 2024](#)), i.e. they are more massive than their host galaxies, which requires this BH mass build-up to precede that of the host galaxy's stellar mass.

The high hardness ratios that we measure for blind+extracted sample (see §4.3, above) indicates the majority of the moderate-luminosity AGN population at these redshifts is heavily obscured and may be almost entirely Compton-thick at $z \gtrsim 7$ (as indicated by the obscured and Compton-thick fractions in Figure 13). The trend we observed here is consistent with that predicted by [Gilli et al. \(2022\)](#) through extensive modelling of the the redshift evolution in the obscuration of AGN due to the column density of gas in the Interstellar Medium (ISM) of galaxies (as well as the obscuring torus associated with the AGN itself). [Gilli et al. \(2022\)](#) found the ISM density increases steeply with redshift as $(1+z)^{3.3-4}$ from a Milky Way-like ISM at $z = 0$ to $> 100\times$ higher column density at $z = 3$, and up to Compton-thick ISM densities at $z > 6$. The high fraction of Compton-thick sources found for our sample, $0.982^{+0.007}_{-0.008}$ for $z = 4 - 10$, thus appears to be consistent with a high abundance of gas within the early Universe, in agreement with the gas mass fractions expected within high redshift galaxies (see review of [Carilli & Walter 2013](#)). Thus, our results are indicative of highly obscured AGN due to the dense ISM of host galaxies, as has been proposed as the mechanism responsible for the lack of X-ray emissions in Little Red Dots. The high space density and obscuration fraction we have measured also indicates a substantial build-up of SMBHs within heavily obscured environments in the early Universe. Such highly abundant gas that may fuel both AGN and star formation within the first galaxies,

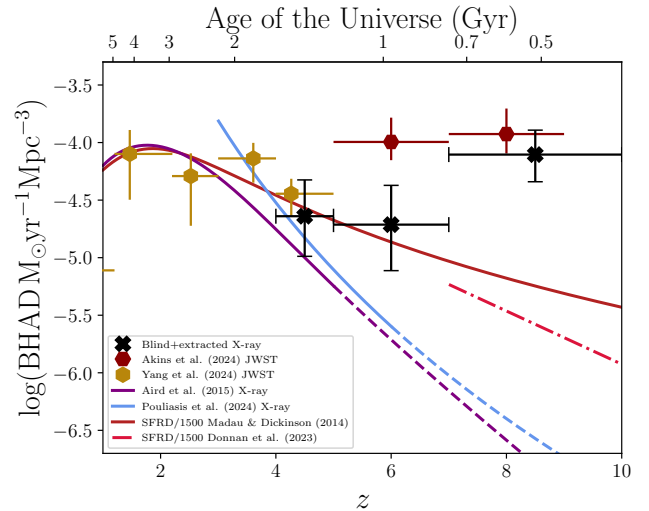


Figure 16. Evolution of the BHAD with redshift, as given by our measurements of the $z = 4 - 10$ XLF (black crosses). Previous X-ray based measurements by [Aird et al. \(2015\)](#) and [Pouliasis et al. \(2024\)](#) are shown by the purple and blue lines respectively and can be seen to follow a rapid decline with increasing redshift. For comparison we show the SFRD of [Madau & Dickinson \(2014\)](#), solid red), and a more recent estimate of the high redshift SFRD as measured by [Donnan et al. \(2023a\)](#), dot-dashed crimson), both rescaled by a factor of 1/1500 to compare with the BHAD. Our measurements of the BHAD can be seen to increase toward higher redshift and are consistent with recent JWST measurements based on mid-IR selected AGN ([Yang et al. 2023](#), shown by the gold hexagons) and Little Red Dots ([Akins et al. 2025](#), shown by the red hexagons). These measurements suggest early SMBHs underwent significant mass growth prior to the build-up of stellar mass within their host galaxies.

with early AGN activity potentially impacting the evolution of their host galaxies. Whilst assessing the mechanisms that may drive this early SMBH growth and any co-evolution with their host galaxies is beyond the scope of this work, our observation findings indicate the exciting prospects for the future of our understanding of galaxy formation within the early Universe.

6 CONCLUSIONS

Searches for high-redshift moderate-luminosity X-ray AGN, corresponding to the majority of the BH mass growth, have been limited by the lack of sufficiently deep near-IR data to reliably measure photometric redshifts as well as the sensitivity of current X-ray surveys. As such, the X-ray Luminosity Function (XLF) has been poorly constrained at high redshifts, with only a few studies fitting parametric models based on samples spanning to $z \sim 6$ (e.g. [Georgakakis et al. 2015](#); [Pouliasis et al. 2024](#)). In this work we have presented new observational measurements of the XLF at $z = 4 - 10$ by combining the deep NIR imaging from the UltraVISTA survey (incorporated within the COSMOS2020 photometric catalogue of [Weaver et al. 2022](#)) with the Chandra legacy imaging from [Civano et al. \(2016\)](#).

In order to measure the high-redshift XLF we have obtained a sample of $z \gtrsim 4$ X-ray detected sources within the COSMOS field, with photometric redshifts from the COSMOS2020 catalogue. We first performed a cross-match of blind X-ray detections in the Chandra COSMOS Legacy data with the COSMOS2020 catalogue, producing a sample of 21 (likely) high-redshift X-ray sources. We then

pushed the limits of the Chandra imaging by extracting X-ray counts at the positions of high-redshift sources within the COSMOS2020 catalogue, allowing us to increase our sample size with an additional 11 (likely) high-redshift X-ray sources that were not identified by the initial X-ray source detection procedure and pushing our sample to somewhat fainter flux limits.

We use our resulting samples of high-redshift AGN to measure the XLF at $z = 4 - 10$ following the $n_{\text{obs}}/n_{\text{mdl}}$ method of Miyaji et al. (2001). The XLF was measured for the blind sample only and then the blind+extracted sample using the best fit photometric redshifts of each source and subsequently testing the upper and lower bounds on the XLF measurements given the possible z_{photo} of our sources. We then investigated the effect of obscuration within the AGN sample on the measured XLF and applied an obscuration correction to our XLF measurements.

These new, improved XLF measurements at moderate luminosities place important constraints on the space density of AGN where previously only extrapolated predictions from XLF models at lower redshifts were available. Our conclusions are as follows:

(i) We have measured the XLF at redshifts of $z = 4 - 5$, $z = 5 - 7$ and $z = 7 - 10$, spanning a luminosity range of $L_X \approx 10^{43-46} \text{ erg s}^{-1}$, and placing upper limits on bins where no sources are found. These XLF measurements, using either the blind-only sample or the blind+extracted sample, are consistent with the XLF models at $z = 4 - 5$ (where the models were fit to pre-existing samples). However, our measurements indicate a higher space density of X-ray AGN toward higher redshifts than predicted by the model extrapolations. At $z = 7 - 10$ our measurements clearly show a space density closer to that of the models at $z = 4 - 5$, suggesting that any decline in the space density of AGN toward high redshift decline is far weaker than lower redshift measurements had previously indicated.

(ii) We also considered the impact of differing photometric redshift estimates for sources in our sample, and calculated the range of XLF measurements when considering the the maximum and minimum number of X-ray sources in our sample at each redshift and luminosity interval. The resulting measurements provide a more conservative uncertainty on our measured XLF, and remain consistent with the model predictions at $z = 4 - 5$ and $z = 5 - 7$. However, the upper bound of these limits indicate a potentially higher space density of AGN than previously expected at $z = 5 - 7$ and $z = 7 - 10$. Even our lower bounds hint at a higher space density of AGN than expected from the model extrapolations at $z = 7 - 10$, the highest redshifts investigated here.

(iii) Performing measurements of the XLF using only sources detected at hard X-ray energies (2–7 keV, corresponding to rest-frame energies $\gtrsim 10$ keV for the redshifts considered here), we found that obscuration likely has a significant impact on measurements of the XLF. The hard-band measurements indicate a much higher space density of AGN than our initial full-band measurements. The hardness ratios of our X-ray sample, determined through a Bayesian estimation method in order to account for the low photon counts, also indicate a significant fraction of our sample are likely to be heavily obscured or Compton-thick AGN.

(iv) Dividing the X-ray sample into unabsorbed, absorbed and Compton-thick AGN, based on their observed hardness ratios, we have estimated N_H - and redshift-dependent corrections to the XLF measurements. We find that the obscuration-corrected XLF, based on our full-band sample, indicates a substantially higher space density than the uncorrected measurements and thus further increases the discrepancy with the model extrapolations.

(v) From the obscuration corrections we calculate the intrinsic

fraction of obscured ($N_H = 10^{23-24} \text{ cm}^2$) and Compton-thick ($N_H > 10^{24} \text{ cm}^2$) AGN. We find the fraction of heavily obscured (absorbed and Compton-thick) sources increasingly dominates over the fraction of unobscured sources ($N_H \leq 10^{23} \text{ cm}^2$), with increasing redshift. The obscured fraction of AGN across the full redshift range of $z = 4 - 10$ is $0.982^{+0.007}_{-0.008}$. Notably the Compton-thick fraction of AGN dominates the population, rising from $0.66^{+0.15}_{-0.13}$ at $z = 4 - 5$ to $0.80^{+0.13}_{-0.04}$ and $0.97^{+0.02}_{-0.01}$ at $z = 5 - 7$ and $z = 7 - 10$ respectively.

(vi) Considering the full range of uncertainty in the XLF we have measured, we find that the space density of AGN in the early Universe *could* be substantially higher than previously expected. This larger population of early AGN could be expected to have a significant effect on the growth of the first galaxies.

(vii) Our obscuration-corrected XLF measurements indicate a rising space density towards lower luminosities, which appears to be consistent with the high space densities of lower luminosity AGN being discovered within deep JWST fields. Thus, our measurements point towards a reconciliation between X-ray measurements and the latest JWST findings at $z \geq 6$, and suggests that the apparent X-ray weakness of many of the high-redshift AGN found by JWST may be due to high levels of obscuration that suppress the X-ray emission.

(viii) Given our obscuration-corrected XLF measurements we calculate the corresponding Black Hole Accretion rate Density (BHAD) at $z = 4 - 5$, $z = 5 - 7$ and $z = 7 - 10$. These measurements show an mild increase BHAD toward higher redshift, at $z > 4$, and are consistent with recent estimates of high levels of black hole growth in the early Universe from JWST studies. Thus, substantial Black Hole growth may have occurred before the build-up of the bulk of stellar mass in such galaxies.

Whilst the results obtained here point towards a reconciliation between the new JWST AGN observations and bright-end X-ray measurements, further studies using larger area X-ray surveys and redshift measurements possible with both the mid-IR imaging of JWST and the larger sky areas covered by *Euclid* offer the potential to constrain the space density of AGN at fainter luminosities and to improve the constraints found here. Furthermore, our measurements of the XLF presents exciting possibilities for the population studies that will be possible with future X-rays telescopes, such as ESA's *NewAthena* X-ray Observatory mission Cruise et al. (2025) that will enable wide-area, deep X-ray surveys that will be highly effective at unveiling the prevalent population of high-redshift, heavily obscured, moderate-luminosity AGN out to $z \gtrsim 6$ suggested by our measurements.

ACKNOWLEDGEMENTS

The authors thank A. Georgakais and D. Alexander for their helpful comments. CLBH acknowledges support from an STFC PhD studentship. CLBH and JA acknowledge support from a UKRI Future Leaders Fellowship (grant codes MR/T020989/1 and MR/Y019539/1). For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission. We have benefited from the publicly available programming language PYTHON, including NUMPY & SCIPY (van der Walt et al. 2011; Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), ASTROPY (Astropy Collaboration et al. 2013) and the TOPCAT analysis program (Taylor 2013).

REFERENCES

- Aihara H., et al., 2019, *PASJ*, 71, 114
- Aird J., et al., 2010, *MNRAS*, 401, 2531
- Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, *MNRAS*, 451, 1892
- Aird J., Coil A. L., Georgakakis A., 2017, *MNRAS*, 465, 3390
- Aird J., Coil A. L., Georgakakis A., 2018, *MNRAS*, 474, 1225
- Aird J., Coil A. L., Georgakakis A., 2019, *MNRAS*, 484, 4360
- Akins H. B., et al., 2025, *ApJ*, 980, L29
- Alexander D. M., Hickox R. C., 2012, *New Astron. Rev.*, 56, 93
- Ananna T. T., et al., 2019, *ApJ*, 871, 240
- Arevalo Gonzalez F., et al., 2025, *arXiv e-prints*, p. [arXiv:2501.09585](https://arxiv.org/abs/2501.09585)
- Arnouts S., et al., 2002, *MNRAS*, 329, 355
- Ashby M. L. N., et al., 2018, *ApJS*, 237, 39
- Assef R. J., et al., 2011, *ApJ*, 728, 56
- Astropy Collaboration et al., 2013, *A&A*, 558, A33
- Baloković M., et al., 2018, *ApJ*, 854, 42
- Barlow-Hall C., 2024, Phd thesis, The University of Edinburgh
- Barlow-Hall C. L., Delaney J., Aird J., Evans P. A., Osborne J. P., Watson M. G., 2023, *MNRAS*, 519, 6055
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Bogdán Á., et al., 2024, *Nature Astronomy*, 8, 126
- Bowler R. A. A., Jarvis M. J., Dunlop J. S., McLure R. J., McLeod D. J., Adams N. J., Milvang-Jensen B., McCracken H. J., 2020, *MNRAS*, 493, 2059
- Boyle B. J., Terlevich R. J., 1998, *MNRAS*, 293, L49
- Boyle B. J., Shanks T., Croom S. M., Smith R. J., Miller L., Loaring N., Heymans C., 2000, *MNRAS*, 317, 1014
- Brammer G. B., van Dokkum P. G., Coppi P., 2008, *ApJ*, 686, 1503
- Brandt W. N., Alexander D. M., 2015, *A&ARv*, 23, 1
- Brandt W. N., Hasinger G., 2005, *ARA&A*, 43, 827
- Brightman M., Nandra K., 2011, *MNRAS*, 413, 1206
- Brinch M., et al., 2023, *ApJ*, 943, 153
- Brooks M., et al., 2024, *arXiv e-prints*, p. [arXiv:2410.07340](https://arxiv.org/abs/2410.07340)
- Brusa M., et al., 2009, *ApJ*, 693, 8
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Buchner J., et al., 2015, *ApJ*, 802, 89
- Buchner J., Salvato M., Budavafi T., Fotopoulou S., 2021, nway: Bayesian cross-matching of astronomical catalogs, *Astrophysics Source Code Library*, record ascl:2102.014
- Cameron R. A., Morris D. C., Virani S. N., Wolk S. J., Blackwell W. C., Minow J. I., O'Dell S. L., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, *Astronomical Society of the Pacific Conference Series Vol. 281, Astronomical Data Analysis Software and Systems XI*. p. 132
- Capak P., et al., 2007, *ApJS*, 172, 99
- Cappelluti N., et al., 2024, *Universe*, 10, 276
- Carilli C. L., Walter F., 2013, *ARA&A*, 51, 105
- Civano F., et al., 2016, *ApJ*, 819, 62
- Coupon J., Czakon N., Bosch J., Komiyama Y., Medezinski E., Miyazaki S., Oguri M., 2018, *PASJ*, 70, S7
- Cruise M., et al., 2025, *Nature Astronomy*, 9, 36
- Delaney J. N., Aird J., Evans P. A., Barlow-Hall C., Osborne J. P., Watson M. G., 2023, *MNRAS*, 521, 1620
- Donnan C. T., et al., 2023a, *MNRAS*, 518, 6011
- Donnan C. T., McLeod D. J., McLure R. J., Dunlop J. S., Carnall A. C., Cullen F., Magee D., 2023b, *MNRAS*, 520, 4554
- Donnan C. T., et al., 2024, *MNRAS*, 533, 3222
- Elvis M., et al., 2009, *ApJS*, 184, 158
- Fan X., et al., 2001, *AJ*, 122, 2833
- Finkelstein S. L., et al., 2022, *ApJ*, 940, L55
- Fruscione A., Burke D., Siemiginowska A., 2009, *Chandra News*, 16, 26
- Gaia Collaboration et al., 2016, *A&A*, 595, A2
- Gehrels N., 1986, *ApJ*, 303, 336
- Georgakakis A., Nandra K., Laird E. S., Aird J., Trichas M., 2008, *MNRAS*, 388, 1205
- Georgakakis A., et al., 2015, *MNRAS*, 453, 1946
- Georgakakis A., Aird J., Schulze A., Dwelly T., Salvato M., Nandra K., Merloni A., Schneider D. P., 2017, *MNRAS*, 471, 1976
- Gilli R., Comastri A., Hasinger G., 2007, *A&A*, 463, 79
- Gilli R., et al., 2009, *A&A*, 494, 33
- Gilli R., et al., 2022, *A&A*, 666, A17
- Grant C. E., Bautz M. W., Plucinsky P. P., Ford P. G., 2024, in den Herder J.-W. A., Nikzad S., Nakazawa K., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 13093, Space Telescopes and Instrumentation 2024: Ultraviolet to Gamma Ray*. p. 130931E ([arXiv:2406.18395](https://arxiv.org/abs/2406.18395)), doi:10.1117/12.3018498
- Greene J. E., et al., 2024, *ApJ*, 964, 39
- Guetzoan P., Aird J., Georgakakis A., Coil A. L., Barlow-Hall C., Hickox R. C., Rankine A. L., Terrazas B. A., 2025, *MNRAS*, 536, 79
- HI4PI Collaboration et al., 2016, *A&A*, 594, A116
- Habouzit M., et al., 2021, *MNRAS*, 503, 1940
- Hainline K. N., et al., 2025, *ApJ*, 979, 138
- Heckman T. M., Best P. N., 2014, *ARA&A*, 52, 589
- Hickox R. C., Alexander D. M., 2018, *ARA&A*, 56, 625
- Hunter J. D., 2007, *Computing in Science and Engineering*, 9, 90
- Ikeda S., Awaki H., Terashima Y., 2009, *ApJ*, 692, 608
- Ilbert O., et al., 2006, *A&A*, 457, 841
- Ilbert O., et al., 2009, *ApJ*, 690, 1236
- Ilbert O., et al., 2013, *A&A*, 556, A55
- Inayoshi K., Visbal E., Haiman Z., 2020, *ARA&A*, 58, 27
- Inayoshi K., Kimura S. S., Noda H., 2024, *arXiv e-prints*, p. [arXiv:2412.03653](https://arxiv.org/abs/2412.03653)
- Kalfountzou E., Civano F., Elvis M., Trichas M., Green P., 2014, *MNRAS*, 445, 1430
- Kocevski D. D., et al., 2018, *ApJS*, 236, 48
- Kocevski D. D., et al., 2024, *arXiv e-prints*, p. [arXiv:2404.03576](https://arxiv.org/abs/2404.03576)
- Kokorev V., et al., 2024, *ApJ*, 968, 38
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Kormendy J., Richstone D., 1995, *ARA&A*, 33, 581
- Kovács O. E., et al., 2024, *ApJ*, 965, L21
- Lacy M., Ridgway S. E., Sajina A., Petric A. O., Gates E. L., Urrutia T., Storrie-Lombardi L. J., 2015, *ApJ*, 802, 102
- Laigle C., et al., 2016, *ApJS*, 224, 24
- Laird E. S., et al., 2009, *ApJS*, 180, 102
- Laloux B., et al., 2023, *MNRAS*, 518, 2546
- Laloux B., et al., 2024, *MNRAS*, 532, 3459
- Lang D., Hogg D. W., Mykytyn D., 2016, The Tractor: Probabilistic astronomical source detection and measurement, *Astrophysics Source Code Library*, record ascl:1604.008
- Larson R. L., et al., 2023, *ApJ*, 953, L29
- Le Fèvre O., et al., 2020, *A&A*, 643, A1
- Luo B., et al., 2017, *ApJS*, 228, 2
- Lusso E., Valiante R., Vito F., 2023, in , *Handbook of X-ray and Gamma-ray Astrophysics*. p. 122, doi:10.1007/978-981-16-4544-0_116-1
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Madau P., Haardt F., 2024, *ApJ*, 976, L24
- Magdziarz P., Zdziarski A. A., 1995, *MNRAS*, 273, 837
- Maiolino R., et al., 2024, *A&A*, 691, A145
- Maiolino R., et al., 2025a, *Monthly Notices of the Royal Astronomical Society*, 538, 1921
- Maiolino R., et al., 2025b, *MNRAS*, 538, 1921
- Marchesi S., et al., 2016, *ApJ*, 827, 150
- Matthee J., et al., 2024, *ApJ*, 963, 129
- McCracken H. J., et al., 2012, *A&A*, 544, A156
- McGreer I. D., et al., 2013, *ApJ*, 768, 105
- Miyaji T., Hasinger G., Schmidt M., 2001, *A&A*, 369, 49
- Moneti A., et al., 2023, *VizieR Online Data Catalog: The fourth UltraVISTA data release (DR4) (Moneti+, 2019)*, *VizieR On-line Data Catalog: II/373*. Originally published in: 2012A&A...544A.156M
- Muzzin A., et al., 2013, *ApJS*, 206, 8
- Nandra K., et al., 2015, *ApJS*, 220, 10
- Napolitano L., et al., 2025, *A&A*, 693, A50
- O'Dell S. L., et al., 2007, in *UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XV*. SPIE, pp 19–30, doi:10.1117/12.734594, <https://www.spiedigitallibrary.org>.

- [org/conference-proceedings-of-spie/6686/668603/Managing-radiation-degradation-of-CCDs-on-the-Chandra-X-ray/10.1117/12.734594.full](https://doi.org/10.1117/12.734594)
- Onodera M., et al., 2012, *ApJ*, 755, 26
- Orofino M. C., Ferrara A., Gallerani S., 2018, *MNRAS*, 480, 681
- Padovani P., et al., 2017, *A&ARv*, 25, 2
- Page M. J., et al., 1996, *MNRAS*, 281, 579
- Park T., Kashyap V. L., Siemiginowska A., van Dyk D. A., Zezas A., Heinke C., Wargelin B. J., 2006, *ApJ*, 652, 610
- Peca A., et al., 2023, *ApJ*, 943, 162
- Pickles A. J., 1998, *PASP*, 110, 863
- Polletta M., et al., 2007, *ApJ*, 663, 81
- Pouliasis E., et al., 2024, *A&A*, 685, A97
- Rees M. J., 1984, *ARA&A*, 22, 471
- Reines A. E., Comastri A., 2016, *Publ. Astron. Soc. Australia*, 33, e054
- Ren K., Trenti M., 2021, *ApJ*, 923, 110
- Reynolds C. S., et al., 2023, in Sigmund O. H., Hoadley K., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 12678, UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XXIII. p. 126781E ([arXiv:2311.00780](https://arxiv.org/abs/2311.00780)), doi:10.1117/12.2677468
- Ross N. P., et al., 2013, *ApJ*, 773, 14
- Salvato M., et al., 2009, *ApJ*, 690, 1250
- Salvato M., et al., 2011, *ApJ*, 742, 61
- Salvato M., et al., 2018, *MNRAS*, 473, 4937
- Sanders D. B., et al., 2007, *ApJS*, 172, 86
- Sawicki M., et al., 2019, *MNRAS*, 489, 5202
- Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103
- Schmidt M., 1968, *ApJ*, 151, 393
- Schmidt M., Schneider D. P., Gunn J. E., 1995, *AJ*, 110, 68
- Scholtz J., et al., 2025, *A&A*, 697, A175
- Scoville N., et al., 2007, *ApJS*, 172, 1
- Sharma R., Sharma M., 2024, *MNRAS*, 531, 3287
- Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020, *MNRAS*, 495, 3252
- Shuntov M., et al., 2022, *A&A*, 664, A61
- Sillassen N. B., et al., 2022, *A&A*, 665, L7
- Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
- Szalay A. S., Connolly A. J., Szokoly G. P., 1999, *AJ*, 117, 68
- Taniguchi Y., et al., 2015, *PASJ*, 67, 104
- Taylor M., 2013, Starlink User Note, 253
- Terrazas B. A., Aird J., Coil A. L., 2024, *arXiv e-prints*, p. [arXiv:2411.08838](https://arxiv.org/abs/2411.08838)
- Trakhtenbrot B., et al., 2016, *ApJ*, 825, 4
- Trinca A., Schneider R., Valiante R., Graziani L., Ferrotti A., Omukai K., Chon S., 2024, *MNRAS*, 529, 3563
- Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, *ApJ*, 786, 104
- Virtanen P., et al., 2020, *Nature Methods*, 17, 261
- Vito F., Gilli R., Vignali C., Comastri A., Brusa M., Cappelluti N., Iwasawa K., 2014, *MNRAS*, 445, 3557
- Vito F., et al., 2018, *MNRAS*, 473, 2378
- Weaver J. R., et al., 2022, *ApJS*, 258, 11
- Weaver J., Zalesky L., Allen N., Taamoli S., 2023, The Farmer: Photometry routines for deep multi-wavelength galaxy surveys, Astrophysics Source Code Library, record ascl:2312.016
- Weisskopf M. C., Brinkman B., Canizares C., Garmire G., Murray S., Van Speybroeck L. P., 2002, *PASP*, 114, 1
- Wise M., 1997, *Chandra News*, 5, 22
- Wolf J., et al., 2021, *A&A*, 647, A5
- Yang J., et al., 2021, *ApJ*, 923, 262
- Yang G., et al., 2023, *ApJ*, 950, L5
- Yue M., et al., 2024, *ApJ*, 966, 176
- Zhang H., Behroozi P., Volonteri M., Silk J., Fan X., Hopkins P. F., Yang J., Aird J., 2023, *MNRAS*, 518, 2123
- Zhu Q., Li Y., Li Y., Maji M., Yajima H., Schneider R., Hernquist L., 2022, *MNRAS*, 514, 5583
- Zubovas K., King A., 2021, *MNRAS*, 501, 4289
- de Graaff A., et al., 2025, *A&A*, 697, A189
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science and Engineering*, 13, 22

APPENDIX A: COSMOS2020 LEPHARE REDSHIFT FITS OF THE $Z \geq 4$ BLIND+EXTRACTED SOURCES

In this appendix I provide plots of the redshift photometric fits of my sample of 39 blind+extracted $z \geq 4$ sources as performed by Weaver et al. (2022) using LePhare with both galaxy only (red) and AGN (blue) Spectral Energy Density (SED) models (see figure A1). The galaxy templates used by Weaver et al. (2022) are those of Polletta et al. (2007) Ilbert et al. (2009), Bruzual & Charlot (2003) and Onodera et al. (2012), and the AGN templates consist of a galaxy SED with a QSO component of varying dominance over the galaxy SED (see Salvato et al. 2009, for a detailed explanation of the creation of the AGN templates). For each source the probability distribution of the redshift with the galaxy template fits is shown in the inset plot with the redshift of the AGN fit shown by the blue line for comparison and the $z < 4$ region shaded. UltraVISTA Ks-band imaging for all of these sources is also shown, with the location of the source as given by the FARMER catalogue indicated by the orange cross, for which all sources can be seen to have a Ks-band detection as is required of my sample.

The shape of the SEDs of all my blind+extracted $z \geq 4$ X-ray sources are suitably reproduced by the LePhare fits performed by Weaver et al. (2022). For those sources which are within close proximity to a brighter object within the Ks-band image shown, the measured photometry at the blue end of the SEDs may be contaminated in certain filters resulting in the scatter of photometry in those sources. It can be seen that, by eye, there are cases where the galaxy only model provides a better fit to the measured photometry, the majority of which have the galaxy redshift (z_{GAL}) around $z \sim 4$. For a number of my sources the measured photometry rises at IR wavelengths, as does the AGN sed model, above that predicted by the galaxy only models. This trend is indicative of high-redshift AGN appearing as red sources with rising power-law slope, as is being seen by JWST (e.g. see Matthee et al. 2024; Maiolino et al. 2025b; Yue et al. 2024; Brooks et al. 2024).

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.

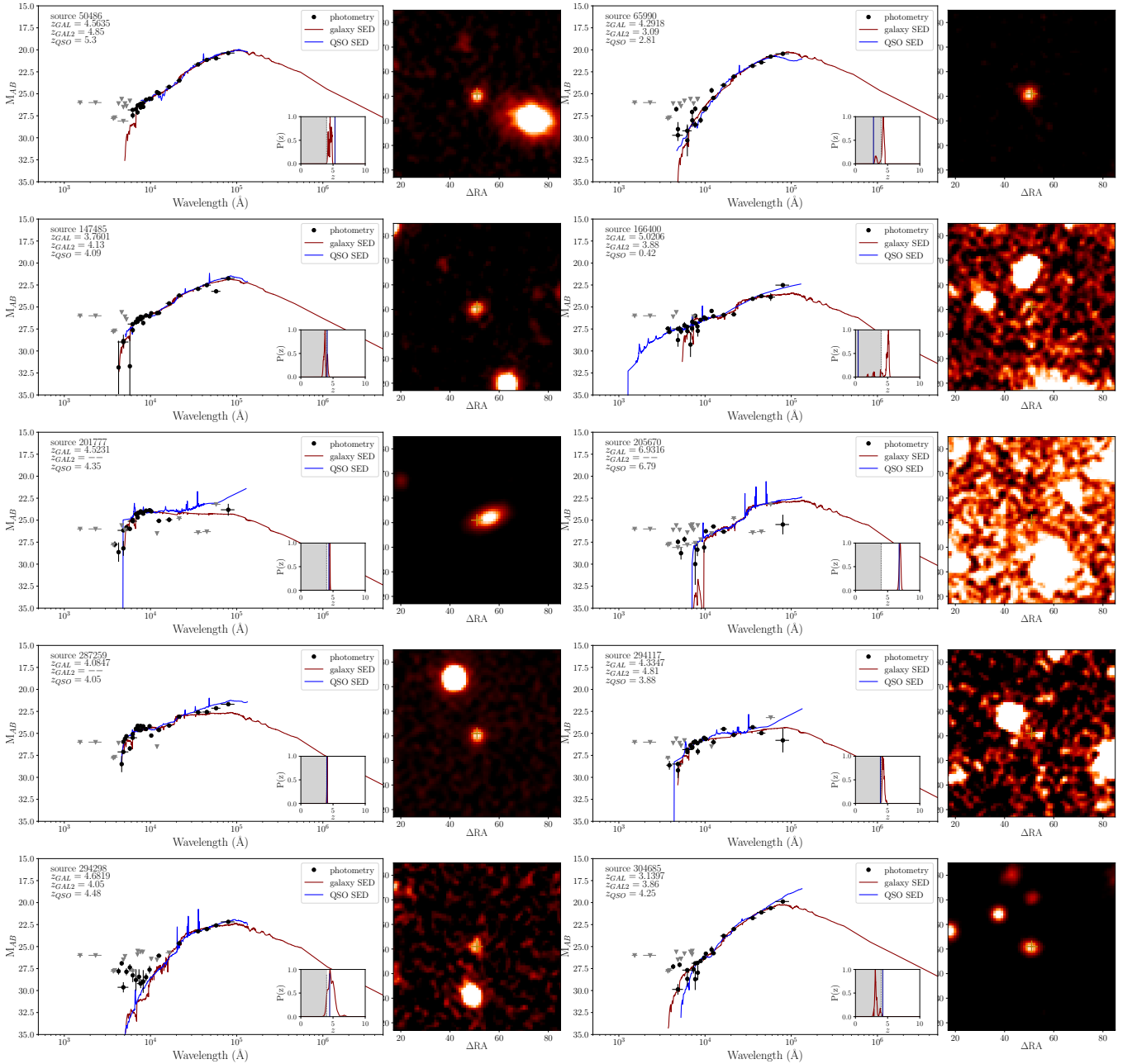


Figure A1. The LePhare photometric redshifts fits of all 39 potentially $z \geq 4$ sources in my blind+extracted sample, as given by galaxy templates (red) and AGN templates (blue). The photometric measurements of each source are shown by the black points, with the measurement in filters for which there are none detections plotted as an upper limit given by the filter depth (grey triangles). The spectral shape measured by the photometry of each source is reproduced by the fitted models. Some sources showing a slight rise in brightness at IR wavelengths above that predicted by the galaxy models, indicative of high redshift AGN being red sources with rising power-law slope (as has been found with JWST; [Matthee et al. 2024](#); [Maiolino et al. 2025b](#)) The probability distribution of possible redshifts based on the galaxy template fits is shown in the inset plot, with the best AGN redshift fit indicated by the blue line, and the region outside the redshift range of my blind+extracted sample shaded. The Ks-band UltraVISTA imaging of each source is shown, with the COSMOS2020 position indicated by the gold cross and a detection clearly seen in all sources which have been fit by both a galaxy and AGN template.

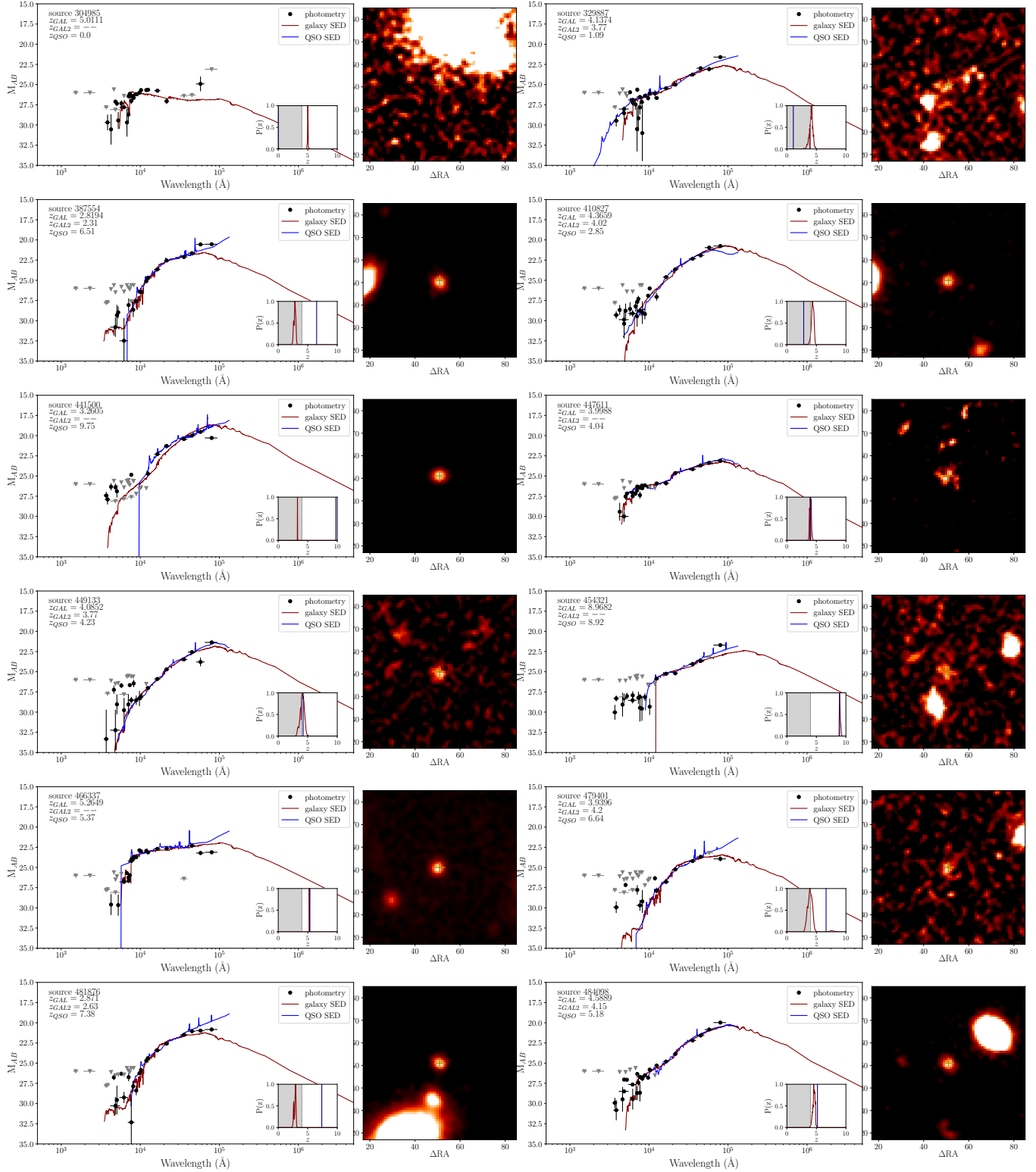


Figure A1 – continued

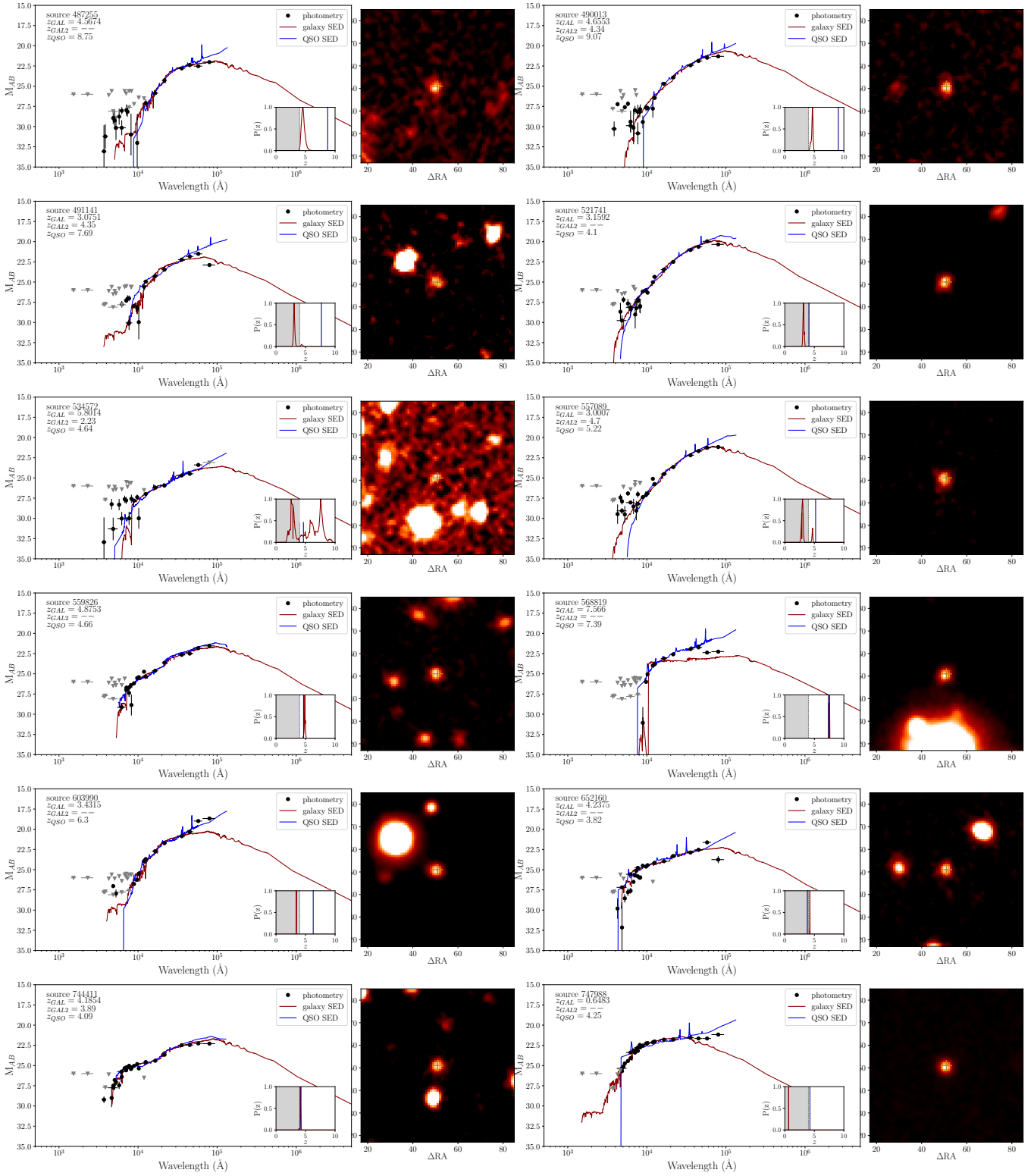


Figure A1 – continued

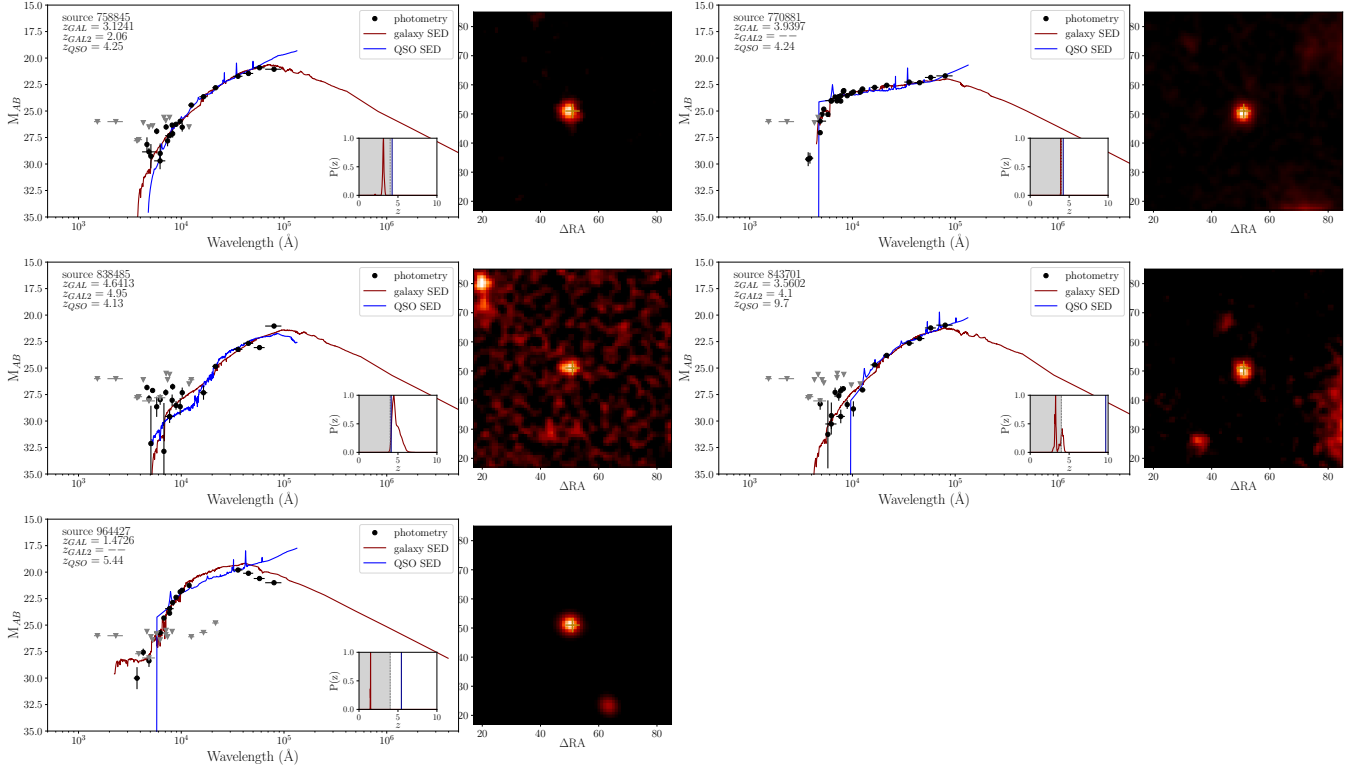


Figure A1 – continued