From a biased perspective: quasars, mergers, and planet-forming discs

Pizzati, E.

4 "Little red dots" cannot reside in the same dark matter halos as comparably luminous unobscured quasars

Abstract

The James Webb Space Telescope (JWST) has uncovered a new population of candidate broad-line AGN emerging in the early Universe, named "little red dots" (LRDs) because of their compactness and red colors at optical wavelengths. LRDs appear to be surprisingly abundant ($\approx 10^{-5}\,\mathrm{cMpc}^{-3}$) given that their inferred bolometric luminosities largely overlap with those of the UV-luminous quasars identified at high z in wide-field spectroscopic surveys. In this work, we investigate how the population of LRDs and/or other UV-obscured AGN relates to the one of unobscured, UV-selected quasars. By comparing their number densities, we infer an extremely large and rapidly evolving obscured:unobscured ratio, ranging from $\approx 20:1$ at $z \approx 4$ to $\approx 2300:1$ at $z \approx 7$, and possibly extending out to very high ($\approx 10^{47}\,\mathrm{erg\,s}^{-1}$) bolometric luminosities. This large obscured:unobscured ratio is incompatible with the UV-luminous duty cycle measured for unobscured quasars at $z \approx 4-6$, suggesting that LRDs are too abundant to be hosted by the same halos as unobscured quasars. This implies that either (a) the bolometric luminosities of LRDs are strongly overestimated or (b) LRDs follow different scaling relations than those of UV-selected quasars, representing a new population of accreting SMBHs emerging in the early Universe. A direct comparison between the clustering of LRDs and that of faint UV-selected quasars will ultimately confirm these findings, and shed light on key properties of LRDs such as their host mass distribution and duty cycle. We provide a mock analysis for the clustering of LRDs and show that it is feasible with current and upcoming JWST surveys.

Published in: **EP**, Joseph F Hennawi, Joop Schaye, Anna-Christina Eilers, Jiamu Huang, Jan-Torge Schindler, Feige Wang, *"Little Red Dots" cannot reside in the same dark matter halos as comparably luminous unobscured quasars*, Monthly Notices of the Royal Astronomical Society, Volume 539, Issue 4, June 2025, Pages 2910–2925, doi.org/10.1093/mnras/staf660. Reprinted here in its entirety.

4.1 Introduction

The connection between the quasar phenomenon and the accretion of material onto a supermassive black hole (SMBH) was first hypothesized to account for the extraordinary luminosity inherent to quasar activity (e.g., Salpeter 1964; Zel'dovich & Novikov 1967; Lynden-Bell 1969). According to this picture, most of the accreting material contributes to growing the mass of the SMBH, but a small fraction of this material (known as the *radiative efficiency*) is converted into energy and radiated away, giving rise to the quasar phenomenon.

The argument first proposed by Soltan (1982) embeds this connection into a cosmological context: integrating the total energy emitted by quasars over all cosmic time and assuming a standard radiative efficiency of $\approx 10\%$, one finds that the mass that has been accreted on black holes per unit of comoving volume up until today is comparable to the total mass density of the SMBHs we observe in the local Universe. This implies that SMBHs grew their mass while, at the same time, they were shining as active luminous quasars.

Extensions of this argument have been employed to relate the growth of black holes to quasar activity at different cosmic times (e.g., Yu & Tremaine 2002; Shankar et al. 2010a). While specific assumptions vary, these arguments are all based on the key idea that the bulk of black hole growth in the Universe is traced by the evolving demographic properties of luminous quasars. Wide-field optical spectroscopic surveys such as the Sloan Digital Sky Survey (SDSS, York et al. 2000) and the 2dF QSO redshift survey (2QZ, Croom et al. 2004) examined the properties of UV-luminous, type-1 quasars, and consistently showed that quasar activity peaks around $z \approx 2$ and declines rapidly towards higher redshifts (e.g., Richards et al. 2006; Kulkarni et al. 2019).

UV-luminous quasars, however, are not the whole story. The radiation emitted from accreting SMBHs can be obscured by intervening dust and gas, resulting in a diverse population of Active Galactic Nuclei (AGN) whose emission properties vary greatly across the electromagnetic spectrum (e.g., Padovani et al. 2017). A general dichotomy exists, however, between unobscured AGN/quasars, exhibiting a UV-optical continuum from the accretion disk, and obscured/reddened AGN whose UV emission is partly (or completely) extincted by the dust that surrounds the SMBH. Whether this obscuration results from a viewing-angle effect (Antonucci 1993; Urry & Padovani 1995) or signifies a distinct "dust-enshrouded" population (Sanders et al. 1988; Hopkins et al. 2005) has been hotly debated. Nevertheless, decades of AGN censuses across the electromagnetic spectrum (optical, X-ray, mid-IR, radio) have allowed us to map the contribution of UV-obscured AGN activity as a function of redshift and AGN luminosity (e.g., Ueda

et al. 2003, 2014; Merloni et al. 2014; Aird et al. 2015; Glikman et al. 2018) The resulting consensus is that a significant fraction ($\approx 20-80\%$) of AGN can be obscured in the UV, even at quasar-like (intrinsic) luminosities ($L_{\rm bol} \gtrsim 10^{45}{\rm erg\,s^{-1}}$), and that this fraction evolves mildly with redshift. Studies that include the contribution of obscured AGN environments to the total SMBH growth budget (e.g., Hopkins et al. 2007b; Shen et al. 2020) support the general picture outlined by the Soltan argument, pointing to a radiative efficiency for accretion on SMBHs close to $\approx 10\%$, and indicating that the bulk of SMBH growth took place during cosmic noon ($z \approx 1-3$).

While a multi-wavelength exploration of AGN activity is possible at $z \lesssim 3$, our understanding of black hole growth and accretion in the high-redshift Universe ($z \gtrsim 4$) has been informed almost exclusively by the population of UV-luminous, type-1 quasars detected by optical/NIR wide-field surveys up to $z \approx 7.5$ (e.g., Fan et al. 2023). This population is commonly assumed to trace the underlying evolution of AGN/SMBH activity (including UV-obscured sources) at high z by simply extrapolating the obscuration properties of quasars from low/intermediate redshifts (e.g., Shen et al. 2020). Whether this extrapolation is reliable and can offer an unbiased view of SMBH growth and AGN activity in the first billion years of the Universe is currently unclear. Several simulations (e.g., Ni et al. 2020; Vito et al. 2022; Bennett et al. 2024) and observations (Vito et al. 2018; Circosta et al. 2019; D'Amato et al. 2020; Gilli et al. 2022), for example, have suggested a rapid evolution of the obscuration properties of quasars/AGN in the early Universe, due to the presence of high-column density gas within the innermost regions of their host galaxies.

The advent of the James Webb Space Telescope (JWST) marks a huge step forward in the study of AGN activity and SMBH growth in the early Universe. JWST has the sensitivity to go beyond the UV-selected quasar population that has been studied for decades (e.g., Fan et al. 2023). Indeed, early results are already causing a seismic shift in our understanding of AGN populations at high z: photometric and spectroscopic JWST surveys are uncovering surprisingly large samples of faint AGN candidates at $z \approx 4-10$ (e.g., Harikane et al. 2023; Maiolino et al. 2024; Übler et al. 2023; Kocevski et al. 2023; Kokorev et al. 2023; Scholtz et al. 2023; Matthee et al. 2024b; Greene et al. 2024; Bogdán et al. 2024; Kocevski et al. 2024; Mazzolari et al. 2024; Furtak et al. 2024; Taylor et al. 2024). Although selection methods vary, the most reliable candidates are identified via broad Hα or Hβ lines. These lines can be used to infer AGN luminosities of $L_{\rm bol} \gtrsim 10^{44-45}\,{\rm erg\,s^{-1}}$ and black hole masses of $M_{\rm BH} \gtrsim 10^{6-7}\,{\rm M_\odot}$ These masses and luminosities vastly extend the range of AGN properties that we can probe at high z, offering key insights on the co-evolution of SMBHs and their host galaxies (e.g., Inayoshi et al. 2022; Pacucci et al. 2023), the contribution of AGN to hydrogen reionization (e.g., Maiolino et al. 2024; Dayal et al. 2024; Madau

et al. 2024), and potentially also on SMBH seeding/growth models (e.g., Pacucci & Loeb 2022; Li et al. 2024).

Yet, relating this new population of JWST AGN to the one of UV-selected high-z quasars has proven challenging. Even though they generally resemble standard, type-1 quasars at rest-frame optical wavelengths, JWST broad-line AGN appear to be much more abundant than what was expected by extrapolating the quasar luminosity function (QLF) to faint UV luminosities (Harikane et al. 2023). It is currently unclear whether QLF studies have been strongly underestimating the number of faint UV quasars that are present at high z (e.g., Giallongo et al. 2019), or whether the AGN population revealed by JWST using broad optical lines presents substantially different properties from those of UV-selected, type-1 quasars, as also suggested by their peculiar Spectral Energy Distribution (SED) features such as X-ray weakness (Maiolino et al. 2024; Lambrides et al. 2024a) and (tentative) lack of variability (Kokubo & Harikane 2024). Upcoming JWST surveys will probe the properties of these broad-line AGN in the rest-frame UV, providing key insight into their nature and allowing a direct comparison to the UV-selected quasar population.

Interestingly, however, some of the AGN revealed by JWST are even more remarkable: a significant fraction of them ($\gtrsim 20\%$; Harikane et al. 2023; Taylor et al. 2024) show a steep red continuum in the rest-frame optical pointing to moderate dust reddening values of $A_V \approx 1-4$ (Kokorev et al. 2024a; Greene et al. 2024). When correcting for the attenuation of dust to the continuum and/or broad-line emission, these obscured/reddened AGN have inferred bolometric luminosities of $L_{\rm bol} \approx 10^{45-46}\,{\rm erg\,s^{-1}}$ and SMBH masses up to $\approx 10^{7-8}\,{\rm M_\odot}$ (Greene et al. 2024; Kocevski et al. 2024; Harikane et al. 2023). Hence, they largely overlap in luminosity and SMBH mass with the population of UV-selected, type-1 quasars revealed in pre-JWST surveys (Fan et al. 2023; Matsuoka et al. 2022). This is incredibly surprising, since these UV-luminous quasars with comparable luminosities (and redshifts) were selected from wide-field 1400 deg^2 deep imaging surveys probing volumes of $\approx 10^{10}$ cMpc^3 (Matsuoka et al. 2022), whereas JWST AGN are identified in surveys of not more than $\approx 300-600$ arcmin^2 probing a volume not greater than $\approx 10^6-10^7$ cMpc^3 (Matthee et al. 2024b; Kokorev et al. 2024a). Such a massive difference indicates that these AGN may be tracing a new population of broad-line, obscured sources[1] that are far more abundant than

[1]Standard AGN classifications (e.g., Padovani et al. 2017) divide low-z quasars in type-1 (showing broad emission lines in their spectra) and type-2 (showing only narrow emission lines). Type-2 quasars are generally identified with obscured sources whose broad lines are extincted by dust. Even though their continuum is heavily reddened at optical and UV wavelengths, JWST AGN are always revealed by broad optical lines, and hence they officially belong to the type-1 quasar category. While examples of type-1, reddened quasars exist at low redshifts, they are rare compared to the global quasar population (Wang et al., in prep.), making the interpretation of these new JWST AGN sources even more challenging.

comparably luminous UV-unobscured quasars. According to this picture, our understanding of SMBH growth and quasar/AGN activity at high-z – which was entirely based on the demographic properties of UV-luminous quasars – needs to be thoroughly revised to account for this new, large AGN population that is in place in the early Universe (e.g., Inayoshi & Ichikawa 2024; Li et al. 2025a).

As shown by Greene et al. (2024), the reddened broad-line AGN in JWST surveys tend to have a characteristic v-shaped SED, with the red continuum in the rest-frame optical transitioning to relatively blue colors in the rest-frame UV. While the physical origin of this SED shape is currently unclear (e.g., Killi et al. 2024; Li et al. 2025a; Wang et al. 2024; Kokorev et al. 2024b; Inayoshi & Maiolino 2025), several studies have exploited these peculiar SED features and applied specific color and compactness cuts to NIRCam photometry to isolate obscured broad-line AGN photometrically (e.g., Labbe et al. 2025; Pérez-González et al. 2024; Kokorev et al. 2024a; Kocevski et al. 2024; Akins et al. 2024). By applying similar photometric selections, Greene et al. (2024) and Kocevski et al. (2024) have proved that a large fraction of the selected sources ($\gtrsim 70-80\%$) is indeed comprised of reddened, high-redshift ($z \approx 4-8$), broad-line AGN. Sources selected using these methods have become known as "Little Red Dots" (LRDs henceforth; Matthee et al. 2024b) because of their compactness and peculiar colors in NIRCam imaging. We note that this term has been used in the literature to refer to samples obtained following different spectroscopic and photometric criteria. Here, with the term "Little Red Dots" we refer to the above-mentioned population of candidate broad-line AGN that are red at optical wavelengths, and hence have quasar-like inferred bolometric luminosities and black hole masses. We include in our analysis both spectroscopic (Greene et al. 2024) and photometric (Kokorev et al. 2024a) samples: while the latter may be subject to a significant degree of contamination (e.g., Taylor et al. 2024), their number densities agree well with the ones from spectroscopy (Greene et al. 2024)[2]. We mention the caveat, however, that even for spectroscopically confirmed broad-line LRDs, the presence of an accreting SMBH and the nature of the observed SED are still heavily debated (e.g., Durodola et al. 2024; Li et al. 2025a; Pérez-González et al. 2024; Ananna et al. 2024; Yue et al. 2024b; Maiolino et al. 2024; Kokubo & Harikane 2024; Baggen et al. 2024; Inayoshi & Maiolino 2025). In the following, we *assume* that LRDs are obscured, broad-line AGN, and examine the consequences of the large obscured:unobscured ratio at quasar-like bolometric luminosities that is implied by this assumption. We defer the reader to Sec. 4.5 for a discussion on the nature of LRDs and the conclusions we can draw from our results. There, we will also examine how the general population of faint (unobscured)

[2]On top of that, a moderate degree of contamination does not impact the main conclusions of our analysis (see Sec. 4.5 for further discussion).

broad-line AGN revealed by JWST (e.g., Harikane et al. 2023; Maiolino et al. 2024; Taylor et al. 2024) fits in the discussion presented in this work.

If a huge obscured LRD population is indeed present at high redshifts, the first question that awaits to be answered is: how does this population compare to that of comparably luminous, UV-selected quasars in terms of SMBH mass and accretion rate, host environments, and evolution history? Are LRDs standard, actively accreting quasars whose emission is attenuated by intervening dust and gas, or do they represent a different evolutionary stage in the accretion history of SMBHs? Are UV-luminous quasars and LRDs drawn from the same population of halos/galaxies?

In this work, we take a first step towards answering these questions by studying the properties of quasars and LRDs in terms of their number density and large-scale environment/host halo mass. In particular, we argue that the extreme abundance of LRDs/obscured AGN is at odds with the duty cycle of UV-luminous quasar activity at $z \approx 4-6$ inferred from the combination of quasar clustering and luminosity function measurements (Shen et al. 2007; Eilers et al. 2024; Pizzati et al. 2024a,b). This indicates that LRDs cannot be drawn from the same population of dark matter halos as UV-selected quasars, notwithstanding that quasars and LRDs have the same inferred bolometric luminosities and SMBH masses. Hence, provided that these luminosities and masses are indeed correct, LRDs would need to obey fundamentally different scaling relations than the ones holding for quasars, as the same SMBH masses are linked to smaller host halo masses. Possibly, this points to the fact that LRDs represent a different evolutionary stage in the accretion history of SMBHs at early cosmic time.

In order to support these conclusions and unveil the accretion history and large-scale environment of LRDs, measuring the clustering of these sources is key. Here, we suggest that a convincing measurement of the duty cycle and host halo mass of LRDs can be obtained by using NIRCam/WFSS observations of LRD fields and measuring the cross-correlation between LRDs and [O III] line emitters, with a similar setup and strategy to current JWST programs targeting UV-luminous, high-z quasars, such as EIGER (Kashino et al. 2023; Eilers et al. 2024) and ASPIRE (Wang et al. 2023). Using the methodology developed in previous work (Pizzati et al. 2024a,b, Chapters 2-3), we provide a mock analysis for these clustering measurements and discuss the prospect of undertaking this measurement with current and future JWST programs.

The paper is structured as follows. In Sec. 4.2, we compare the abundance of LRDs/obscured AGN with the one of the UV-luminous high-z quasar population, inferring a large and rapidly evolving obscured:unobscured ratio at $z \approx 4-8$. Sec. 4.3 studies the implications of this large ratio in terms of host dark matter halo populations, and points to clustering studies as a way to determine the nature of LRDs. Sec. 4.4 provides a mock analysis of this

clustering measurement. The results are discussed and summarized in Sec. 4.5.

4.2 The staggeringly high abundance of UV-obscured AGN implied by little red dots

In this section, we compare the luminosity function of the UV-luminous, unobscured population of quasars to that of the new population of UV-obscured "Little Red Dots" (LRDs) uncovered in JWST surveys. Our goal is to study the abundance of these two populations across cosmic time, and infer an estimate of the AGN obscured fraction at different redshifts.

To this end, we use bolometric luminosities as a way to probe the intrinsic radiation emitted by the different quasar/AGN populations prior to any obscuration effects. The bolometric luminosities of UV-luminous, type-1 quasars can be easily inferred from their UV-continuum absolute magnitude by assuming standard bolometric correction factors that are available in the literature (e.g., Richards et al. 2006; Runnoe et al. 2012a; Shen et al. 2020). In this work, we use the relation between the M_{1450} absolute magnitude and the bolometric luminosity $L_{\rm bol}$ presented in Runnoe et al. (2012a)[3]. While other bolometric correction factors may return slightly different results because of the choices made for the quasar SED and the parametrization of the UV-bolometric relation, the uncertainty in the bolometric correction for UV-selected, type-1 quasars is relatively small and has little impact on our conclusions.

Estimating the intrinsic bolometric luminosity of the LRD population, instead, is much more challenging. While bolometric luminosities are easy to constrain for UV-selected quasars because one directly probes the big-blue-bump (where the bulk of the emission comes out, Sanders et al. 1989), dust obscuration prevents a direct determination of the LRD luminosities from their UV emission. For low-z, dust-obscured quasars, it is usually possible to constrain the radiation reprocessed by dust in the mid-IR with Spitzer (e.g., Lacy et al. 2015). However, this is currently not a viable option for LRDs, as they appear to manifest only at high z and the bulk of their expected mid-IR emission is redshifted to wavelengths of $\approx 70\mu$m, which are not accessible from the ground and are only probed by shallow surveys (e.g., Herschel). The only option that remains available for estimating the bolometric luminosities of LRDs is to use the emission in the optical continuum and/or broad optical lines and convert that to a bolometric luminosity using some scaling relations (e.g., Richards et al. 2006; Runnoe et al. 2012b), which are however

[3]The bolometric correction for $\lambda = 1450$ Å is $\log_{10} L_{\rm iso}/{\rm erg\,s^{-1}} = 4.745 + 0.910 \log_{10} \lambda L_\lambda/{\rm erg\,s^{-1}}$. $L_{\rm iso}$ refers to the bolometric luminosity computed under the assumption of isotropy, and it is related to the observed bolometric luminosity $L_{\rm bol}$ through the relation $L = 0.75\, L_{\rm iso}$.

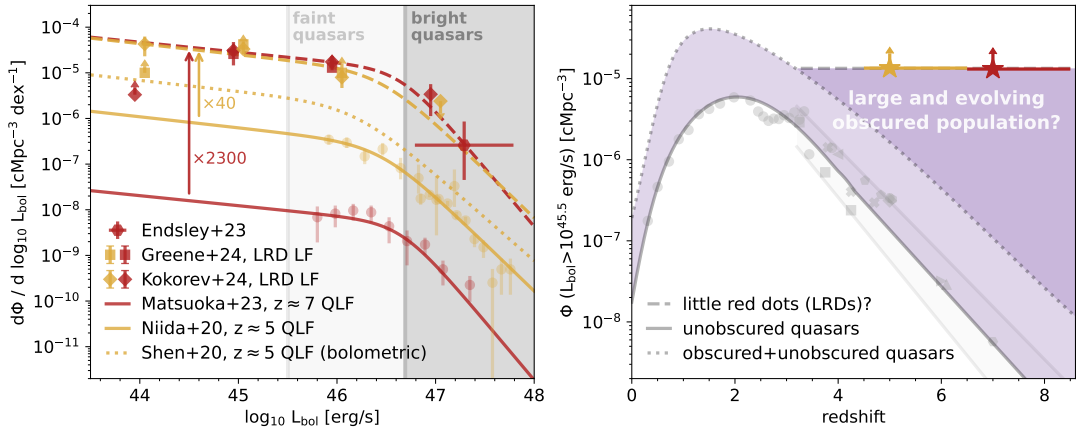

Figure 4.1: *Left*: Luminosity function (LF) of UV-selected quasars, expressed in terms of bolometric luminosities, compared to the bolometric LF of Little Red Dots (LRDs) at different redshifts. Solid lines show the fits to the unobscured quasar luminosity functions (QLFs) at $z \approx 5$ (Niida et al. 2020; golden color) and $z \approx 7$ (Matsuoka et al. 2023; red). Data points for these QLFs are also shown as circles. The bolometric QLF compiled by Shen et al. (2020) at $z \approx 5$ is shown with a dotted line. Bolometric LFs for LRDs are shown with square (Greene et al. 2024) and diamond (Kokorev et al. 2024a) symbols. Golden (red) symbols refer in this case to the redshift range $4.5 < z < 6.5$ ($6.5 < z < 8.5$). The number density implied by the single source identified by Endsley et al. (2022, 2023) at $z \approx 7$ (see main text) is shown as a red hexagon. Vertical arrows show by how much the QLF fits (solid lines) need to be rescaled to match the LRD LF. Dashed lines show the rescaled QLFs: the $z \approx 5$ ($z \approx 7$) QLF is rescaled by a factor of 40 (2300). The light (dark) grey shaded region highlights the luminosity range $10^{45.5}\,\mathrm{erg\,s^{-1}} < L_{\rm bol} < 10^{46.5}\,\mathrm{erg\,s^{-1}}$ ($L_{\rm bol} > 10^{46.5}\,\mathrm{erg\,s^{-1}}$). *Right:* Evolution of the number density of quasar/AGN (above the luminosity threshold $L_{\rm bol} > 10^{45.5}\,\mathrm{erg\,s^{-1}}$) with redshift. Gray points show the number densities obtained by integrating individual fits to the unobscured QLFs above the luminosity threshold (partly adapted from the compilation in Schindler et al. 2023). The solid line shows an evolutionary model for the unobscured quasar number density obtained by smoothly interpolating between the fit of Kulkarni et al. (2019) at $z < 4$, and an exponential decline $\Phi \propto 10^{-kz}$, with $k = 0.7$, at higher redshifts. The gray shaded area and the two lines at $z > 3$ are meant to bracket our uncertainty on the number density of high-z unobscured quasars. Dotted lines show the number density evolution predicted by the bolometric QLF of Shen et al. (2020) (see their "global fit B"). Colored star symbols show the number density for the LRDs obtained by integrating the rescaled QLFs from the left panel. The flat evolution of the LRD number density implied by the data points is highlighted with a horizontal dashed line, while the (light+dark) purple-shaded areas show the AGN obscured:unobscured ratio inferred from LRDs and low-z multi-wavelength observations.

fairly uncertain. Even more relevantly, one has to properly account for the effects of dust obscuration on the observed optical emission. Current estimates of the bolometric luminosities for the LRD population (e.g., Greene et al. 2024; Kokorev et al. 2024a; Akins et al. 2024) rely on the assumption that the optical continuum of LRDs is dominated by dust-reddened AGN radiation and use the slope of the SED in the optical continuum to infer the amount of obscuration in place. However, this continuum emission could be contaminated by radiation from the host galaxy: disentangling the contributions of the central SMBH and the stellar light to the SED of LRDs is currently a hotly debated problem (e.g., Durodola et al. 2024; Li et al. 2025a; Pérez-González et al. 2024; Baggen et al. 2024; Inayoshi & Maiolino 2025). As mentioned before, here we simply assume that bolometric luminosity estimates for LRDs are correct. A discussion on how our results are impacted by uncertainties in the bolometric luminosities of LRDs can be found in Sec. 4.5.

In the left panel of Figure 4.1, we show the luminosity function of UV-luminous, unobscured quasars (expressed in terms of bolometric luminosities) at two sample redshifts of $z \approx 5$ (Niida et al. 2020; golden solid line and points) and $z \approx 7$ (Matsuoka et al. 2023; red solid line and points). These luminosity functions can be compared to the bolometric luminosity functions of LRDs measured by Greene et al. (2024) (squares) and Kokorev et al. (2024a) (diamonds)[4]. Golden (red) symbols refer to the redshift range $4.5 < z < 6.5$ ($6.5 < z < 8.5$). This plot highlights the strikingly different abundance of LRDs compared to the UV-luminous quasar population. As also mentioned in the introduction, this difference reflects the fact that LRDs are common in the small fields ($\approx 300 - 600$ arcmin^2) probed by JWST surveys, whereas unobscured quasars are notoriously rare and can be sampled only by wide-field surveys of ≈ 2000 deg^2.

By directly comparing the luminosity functions of UV-luminous quasars and LRDs, we can quantify the different abundances of these two populations as a function of their luminosity. Interestingly, we find that the *shape* of

[4]The Greene et al. (2024) luminosity function is obtained from a small sample of spectroscopically-confirmed broad-line LRDs in the UNCOVER field (Bezanson et al. 2024). The work of Kokorev et al. (2024a) applies the photometric selection suggested by Labbe et al. (2025) and Greene et al. (2024) to a larger sample of JWST blank fields, identifying 260 AGN candidates in ≈ 640 arcmin^2 of JWST imaging. While several other LRD luminosity functions have been published in the literature (see e.g., Matthee et al. 2024b; Kocevski et al. 2024; Lin et al. 2024), none of these are based on *unattenuated* bolometric luminosities. Accounting for the effect of dust attenuation is key if our goal is to compare the luminosities of LRDs to the ones of UV-luminous quasars. The only exception is the recent work of Akins et al. (2024), who also published an LRD bolometric luminosity function corrected for obscuration effects. However, their photometric selection differs significantly from the one presented in Greene et al. (2024) and Kokorev et al. (2024a), and hence we do not include their sample in the analysis. We note however that they find even larger number densities for LRDs, which would strengthen our conclusion on the presence of a large obscured high-z AGN population.

the LRD luminosity function resembles the one of the UV-luminous quasar luminosity function (QLF) at both redshifts. Indeed, if we scale up the Niida et al. (2020) fit to the $z \approx 5$ QLF by a factor of ≈ 40, we get a good match to the LRD luminosity function in the redshift range $4.5 < z < 6.5$. This suggests that LRDs may constitute a new, obscured population of accreting SMBHs at $z \approx 5$, outnumbering unobscured quasars by ≈40:1 at all luminosities. Similar – but even more extreme – conclusions can be drawn at $z \approx 7$. In this case, the fit to the Matsuoka et al. (2023) QLF needs to be scaled up by a factor of ≈ 2300 to match the LRD luminosity function at $6.5 < z < 8.5$, implying an even larger obscured:unobscured ratio, roughly independent of luminosity.

We note that care must be taken to extend these conclusions to a large range of bolometric luminosities. Most LRDs have inferred (dust-corrected) bolometric luminosities in the range $\approx 10^{44-46}\,\mathrm{erg\,s^{-1}}$. The faintest high-$z$ unobscured quasars identified in wide field surveys have luminosities of $\approx 10^{45.3}\,\mathrm{erg\,s^{-1}}$ (e.g., Matsuoka et al. 2022). Hence, a proper comparison between LRD and quasar number densities can be carried out only for the *bright* population of LRDs with $L_{\mathrm{bol}} \approx 10^{45.5-46.5}\,\mathrm{erg\,s^{-1}}$. At lower bolometric luminosities, the UV-luminous QLFs are only based on extrapolations; hence, conclusions on the obscured fraction of faint ($L_{\mathrm{bol}} \lesssim 10^{45}\,\mathrm{erg\,s^{-1}}$) AGN are only tentative. At very bright luminosities of $L_{\mathrm{bol}} \approx 10^{47}\,\mathrm{erg\,s^{-1}}$, the number density of UV-luminous quasars is very well constrained (e.g., Schindler et al. 2023). Very bright LRDs, on the other hand, are hard to find in the small field of views (FoVs) probed by JWST surveys and the only constraints we have on their number density come from the work of Kokorev et al. (2024a) (see also Akins et al. 2024), which is however only based on photometry with no spectroscopic confirmation.

Interestingly, signs of a large obscured AGN population at high bolometric luminosities ($L_{\mathrm{bol}} \gtrsim 10^{47}\,\mathrm{erg\,s^{-1}}$) come from different data. Using multi-wavelength observations in mid-/far-IR, sub-mm, and radio, Endsley et al. (2022, 2023) (see also Lambrides et al. 2024b) discovered an extremely luminous ($L_{\mathrm{bol}} = (2.0 \pm 0.2) \times 10^{47}\,\mathrm{erg\,s^{-1}}$) obscured, radio-loud quasar at $z = 6.83$ in just $1.5\,\mathrm{deg}^2$ of COSMOS imaging, and argued for an extremely large obscured:unobscured ratio of $\sim 2000 : 1$. We can get an estimate of the number density implied by this source by simply computing the total comoving volume in the COSMOS field for the redshift range $6.6 < z < 6.9$ (in which the source was photometrically selected; see Endsley et al. 2022). We get a volume of $3.8 \times 10^6\,\mathrm{cMpc}^3$ and a number density of $2.6 \times 10^{-7}\,\mathrm{cMpc}^{-3}$. For reference, we add this source to the luminosity function plot of Fig. 4.1 (left), by assuming a 1 dex bin in bolometric luminosity centered on the quasar's measured L_{bol}. Upper and lower limits are computed assuming Poisson statistics for a single source (see Gehrels 1986). Despite the large uncertainties, this source supports the existence of

a large obscured population at the bright end of the QLF compatible with the one found for LRDs.

In what follows, we will consider two separate hypotheses: (a) there is a large obscured AGN/quasar population at bolometric luminosities $L_{\rm bol} \approx 10^{45} - 10^{46}\,{\rm erg\,s^{-1}}$ (i.e., at the faint end of the quasar luminosity function; light-grey shaded area in the left panel of Fig. 4.1); (b) this large obscured population extends to very large bolometric luminosities of $L_{\rm bol} \approx 10^{47}\,{\rm erg\,s^{-1}}$ (dark-grey shaded area). While the former is supported by a fairly large sample of LRDs that have been argued to overlap in luminosity with the faint quasar population (e.g., Greene et al. 2024; Matthee et al. 2024b; Lin et al. 2024; Taylor et al. 2024; Schindler et al, in prep.), the latter is currently based only on a handful of sources (i.e., the photometrically-selected LRDs in Kokorev et al. 2024a; Akins et al. 2024 and the obscured quasars from Endsley et al. 2022, 2023; Lambrides et al. 2024b) and thus it is only tentative (see Sec. 4.5 for further discussion).

In the right panel of Figure 4.1, we show how the quasar/AGN number density evolves with redshift by integrating the QLF above a bolometric luminosity threshold of $L_{\rm bol} = 10^{45.5}\,{\rm erg\,s^{-1}}$ (light grey vertical line in the left panel). The cosmological evolution of the UV-luminous, type-1 quasar population has been analyzed in the recent work of Kulkarni et al. (2019). The solid grey line in Fig. 4.1 (right) shows their best-fitting model at $z < 4$. For higher redshifts, the Kulkarni et al. (2019) model is very uncertain and does not agree well with the data. For this reason, at $z > 4$ we assume that the cosmic number density of unobscured high-z quasars declines exponentially as $\Phi(z) \propto 10^{-kz}$, and set $k = 0.7$ for our fiducial model (Schindler et al. 2023). We then smoothly interpolate between the fit of Kulkarni et al. (2019) at $z < 4$ and this exponential decrease at higher redshift. Together with this global evolution model, we also show individual (gray) points obtained by integrating local fits to the QLFs above the luminosity threshold (fits are taken from Yang et al. 2016; Akiyama et al. 2018; McGreer et al. 2018; Matsuoka et al. 2018; Schindler et al. 2019; Kulkarni et al. 2019; Niida et al. 2020; Onken et al. 2022; Pan et al. 2022; Schindler et al. 2023; Matsuoka et al. 2023). Overall, these individual data points agree with the global evolutionary model, but a significant spread is present due to uncertainties in the QLF measurements (especially at the faint end, $L_{\rm bol} \lesssim 10^{46}\,{\rm erg\,s^{-1}}$). To quantify this uncertainty, we plot two gray lines corresponding to different exponential declines of the quasar number density, $k = 0.65$ and $k = 0.78$ (e.g., Wang et al. 2019; Matsuoka et al. 2023); these two lines are normalized at $z = 4$ to twice and half of the fiducial model, respectively.

Together with the measurements for the UV-luminous quasar number density, we show (Fig. 4.1, right panel) with a dotted line the model for the evolution of the AGN *bolometric* number density from Shen et al. (2020). This work employs multi-wavelength observations (from X-rays to mid-IR) to

include the contribution of all quasars/AGN to the number density budget. In particular, by exploiting X-ray observations at $0 < z < 3$ (e.g., Ueda et al. 2003, 2014; Merloni et al. 2014; Aird et al. 2015), they include a model for AGN obscuration, and account for the obscured fraction of quasars/AGN in their luminosity function estimates. As mentioned in the introduction, observations generally constrain the AGN obscured fraction only at $z \lesssim 3$, so the Shen et al. (2020) model is effectively extrapolating the behaviour of the AGN obscured populations from cosmic noon to the high z Universe. Nonetheless, the work of Shen et al. (2020) represents our best guess (prior to JWST observations) for how the global AGN/SMBH population evolves as a function of redshift. By comparing the number density of UV-selected quasars (solid grey line in the right panel of Fig. 4.1) with the number density from Shen et al. (2020) (which includes obscured sources), we can estimate the obscured:unobscured ratio of AGN as a function of redshift. The same ratio can be studied as a function of intrinsic luminosity by considering the UV-luminous and the bolometric QLFs at a single redshift. As an example, we do this in the left panel of Fig. 4.1 by showing the Shen et al. (2020) predictions for the bolometric QLF at $z = 5$ with a golden dotted line. In general, the obscured:unobscured ratio implied by comparing the bolometric (Shen et al. 2020) to the UV (Kulkarni et al. 2019) QLFs evolves moderately with redshift and luminosity, ranging from $\approx$ a few: 1 up to $\approx 20:1$ for the case of high redshift and low bolometric luminosity. We note that these values are inevitably very uncertain, as the method employed here is subject to the exact parametrizations employed by Kulkarni et al. (2019) and Shen et al. (2020) for their respective QLFs. Nevertheless, we present this comparison between UV-selected and bolometric models to outline the conventional wisdom on AGN/quasar populations that is being challenged by the new population of LRDs/broad line AGN uncovered in JWST surveys.

The number density evolution of LRDs can be estimated by integrating their bolometric luminosity functions in the left panel of Fig. 4.1 above the same $L_{\rm bol}$ threshold of $10^{45.5}\,{\rm erg\,s^{-1}}$ employed before (vertical light grey line). In practice, given that the rescaled UV QLFs (dashed lines in the left panel of Fig. 4.1) are good fits to the LRD bolometric luminosity functions, we can simply rescale the unobscured quasar number density obtained at $z = 5$ and $z = 7$ to get the LRD number densities at the same redshifts. We show as colored star symbols (Fig. 4.1, right panel) the LRD number densities obtained after this rescaling. Following Greene et al. (2024), we plot these symbols as lower limits.

As argued before, the AGN number density implied by JWST observations of LRDs is surprisingly large and non-evolving. To highlight this behavior, we plot (Fig. 4.1, right panel) a horizontal dashed line for $z \gtrsim 3$ corresponding to the abundance $\Phi_{\rm LRD} \approx 1.3 \times 10^{-5}\,{\rm cMpc^{-3}}$. At $z \gtrsim 6$, this abundance is many orders of magnitude higher than the one measured for unobscured quasars, implying that our general understanding of SMBH

accretion and quasar activity in the early Universe may need to be deeply revised. Inayoshi & Ichikawa (2024) (see also Akins et al. 2024) have already examined the challenges that these LRD number densities pose to our paradigm of SMBH growth as well as the co-evolution of SMBHs and galaxies. In this work, we focus on the consequences of the large and rapidly evolving AGN obscured fraction that can be inferred by comparing LRDs to unobscured quasars. In Fig. 4.1 (right), we show with a light purple shading the region between the unobscured quasar evolution model and the bolometric (obscured+unobscured) model of Shen et al. (2020). A darker shading highlights the dramatic increase in the obscured fraction at $z \gtrsim 4$ that is needed to match LRD measurements.

Dividing the LRD number density, $\Phi_{\rm LRD}$ (which, to a first approximation, is not evolving with redshift), by the number density of UV-luminous quasars (solid grey line in the left panel of Fig. 4.1), we infer an obscured:unobscured ratio that increases from $r_{\rm obsc} \approx 20^{+20}_{-10} : 1$ at $z = 4$ to $r_{\rm obsc} \approx 2300^{+3500}_{-1400} : 1$ at $z = 7$. In the following section, we will also make use of the obscured:unobscured ratio at $z = 6.25$, which is $r_{\rm obsc} \approx 815^{+1600}_{-545} : 1$. The uncertainties on these obscured ratios are computed by considering the grey shaded area (and grey lines) in Fig. 4.1 (right), and are meant to quantify the scatter (coming from systematics in the QLF modeling) between different number density measurements for the unobscured quasar population. Given the challenges with interpreting and contextualizing LRD measurements, we currently do not attempt to model uncertainties for the LRD population, and defer to Sec. 4.5 for a discussion of the significance of our results.

4.3 Little red dots and UV-selected quasars: do they belong to the same population?

From the analysis performed in the previous section, we concluded that: (a) LRDs imply the existence of a large and rapidly evolving obscured AGN population (at redshifts $z \approx 4-7$ and bolometric luminosities $L_{\rm bol} \approx 10^{45}-10^{46}\,{\rm erg\,s^{-1}}$) which outnumbers UV-luminous quasars by several orders of magnitude (Fig. 4.1, right); (b) there is tentative evidence (Fig. 4.1, left) that this obscured population extends to even higher bolometric luminosities ($L_{\rm bol} \approx 10^{47}\,{\rm erg\,s^{-1}}$). In this section, we examine the implications of these findings in the context of AGN host dark matter halo masses and duty cycles.

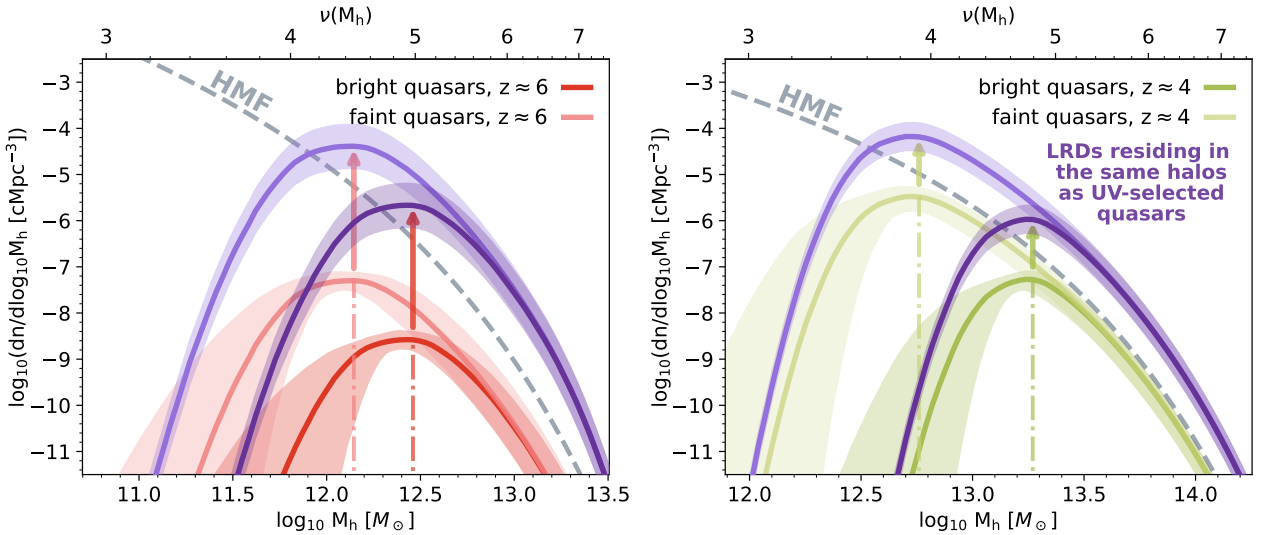

Figure 4.2: Quasar-host mass functions (QHMFs) at $z \approx 6$ (*left panel*) and $z \approx 4$ (*right*) as a function of the (sub)halo mass, $M_{\rm h}$, and the peak height, $\nu(M_{\rm h})$ (see text for definitions). Darker (lighter) colors represent QMHFs obtained by setting a bolometric luminosity threshold corresponding to that of bright (faint) quasars, i.e., $L_{\rm bol} = 10^{46.7}\,{\rm erg\,s^{-1}}$ ($L_{\rm bol} = 10^{45.5}\,{\rm erg\,s^{-1}}$). Median and 1σ uncertainties (obtained by randomly sampling the posterior distributions shown in Pizzati et al. 2024a,b), are represented with solid lines and shaded areas, respectively. The dashed-dotted lines highlight the median values of the QHMF distributions. The (sub)halo mass functions (HMFs) at the respective redshifts are plotted with dashed grey lines in both panels. Purple colors show the QHMFs distributions when scaled up by the obscured:unobscured ratios ($r_{\rm obsc}$) derived in Sec. 4.2, and represent the host mass distribution of LRDs under the hypothesis that they are drawn from the same halo population as UV-selected quasars. Dark (light) purple is associated with bright (faint) quasar bolometric luminosities. The colored arrows represent the values of $r_{\rm obsc}$ by which the QHMFs are scaled up. The (purple) shaded regions represent the effect of the uncertainties on $r_{\rm obsc}$ (see Sec. 4.2). The host mass distributions for LRDs overshoot the HMFs at the massive end, implying that LRDs are too abundant to reside in the same dark matter halos as comparably luminous, unobscured quasars.

4.3.1 The host dark matter halos and duty cycles of high-z unobscured quasars and their luminosity dependence

Determining which halos can host quasar activity as a function of cosmic time is one of the main questions in the field, as it is key to embedding quasars in the structure formation picture: this sheds light on the processes governing SMBH accretion and growth, as well as the co-evolution between SMBHs and their host halos/galaxies. In this context, quasar clustering measurements have been widely used to estimate the masses of the halos hosting UV-luminous quasars at different redshifts (Porciani et al. 2004; Croom et al. 2005; Porciani & Norberg 2006; Shen et al. 2007; Ross et al. 2009; Eftekharzadeh et al. 2015; Arita et al. 2023; Eilers et al. 2024). The idea behind these measurements is straightforward: according to the ΛCDM cosmology, the clustering of any populations of objects increases with the masses of the dark matter halos they reside in (e.g. Kaiser 1984; Bardeen et al. 1986; Mo & White 1996).

As pointed out by, e.g., Martini & Weinberg (2001); Haiman & Hui (2001), determining the quasars' characteristic host halo masses can also give us insight into their accretion history. Suppose that – as routinely assumed – all massive halos host a SMBH at their center. The *duty cycle* of quasar activity determines what fraction of these SMBHs, on average, are active as UV-luminous quasars at any given moment. By comparing the number density of potential quasar hosts – obtained from quasar clustering measurements – to the observed unobscured quasar number density, one can constrain this *UV-luminous quasar duty cycle*. Given the connection between quasar activity and SMBH accretion and growth, the quasar duty cycle offers a direct view into the growth mode of SMBHs at a given cosmic epoch.

In Pizzati et al. (2024a,b), we developed a method to constrain the UV-luminous quasar duty cycle ($\varepsilon_{\rm QSO}$) as well as the mass distribution of the (sub)halos that host unobscured quasars (the so-called "*quasar host mass function*"; QHMF) by simultaneously fitting the clustering of quasars and their luminosity function. The method builds on a conditional luminosity function (CLF) framework, which links in a statistical sense the population of dark matter subhalos to that of quasars (e.g., Yang et al. 2003; Ren et al. 2020). We employ a description for the CLF based on an empirical relation between the quasar bolometric luminosity, $L_{\rm bol}$, and the host (sub)halo mass, $M_{\rm h}$, with log-normal scatter, σ. This relation is also normalized by an active

fraction, $f_{\rm on,UV}$, which accounts for the fact that not all quasars are actively accreting and UV-luminous at a given time:

$$\begin{aligned}\mathrm{CLF}(L_{\rm bol}|M_{\rm h})\,\mathrm{d}L_{\rm bol} &= \\ &= \frac{f_{\rm on,UV}}{\sqrt{2\pi}\sigma}\,\exp\left(\frac{(\log_{10}L_{\rm bol}-\log_{10}L_{\rm c}(M_{\rm h}))^2}{2\sigma^2}\right)\mathrm{d}\log_{10}L_{\rm bol}.\end{aligned} \tag{4.1}$$

We assume a power-law $L_{\rm c}(M_{\rm h})$ relation, parametrized by a slope, γ, and a normalization $L_{\rm ref}$. In terms of logarithmic quantities:

$$\log_{10}L_{\rm c}(M_{\rm h}) = \log_{10}L_{\rm ref} + \gamma\,(\log_{10}M_{\rm h} - \log_{10}M_{\rm ref})\,, \tag{4.2}$$

with $M_{\rm ref}$ fixed to $\log_{10}M_{\rm ref}/\mathrm{M}_\odot = 12.5$.

By fitting the quasar clustering and the QLF at any given redshift, we have enough information to constrain the quasar luminosity-halo mass relation (γ and $L_{\rm ref}$), its intrinsic scatter (σ), and the active fraction of quasars ($f_{\rm on,UV}$) – see Table 4.1. Once these quantities are known, the QHMF can be obtained by statistically assigning quasars to subhalos and selecting only the subhalos whose quasars are brighter than some luminosity threshold, $L_{\rm thr}$ (which is usually set according to observations):

$$\mathrm{QHMF}(M_{\rm h}|L_{\rm bol}>L_{\rm thr}) = \mathrm{HMF}(M_{\rm h})\int_{L_{\rm thr}}^{\infty}\mathrm{CLF}(L_{\rm bol}|M_{\rm h})\,\mathrm{d}L_{\rm bol}, \tag{4.3}$$

where HMF stands for the (sub)halo mass function. A comparison between the QHMF and the HMF can then return the value of the UV-luminous quasar duty cycle, $\varepsilon_{\rm QSO}$:

$$\varepsilon_{\rm QSO} = \frac{\int_{M_{\rm med}}^{\infty}\mathrm{QHMF}(M|L_{\rm bol}>L_{\rm thr})\,\mathrm{d}M}{\int_{M_{\rm med}}^{\infty}\mathrm{HMF}(M)\,\mathrm{d}M}. \tag{4.4}$$

The lower integration limit is set to the median value[5] of the QHMF, $M_{\rm med}$ (see, e.g., Ren et al. 2020). For more details on the parametrization employed for the CLF and the definition of the various quantities at play, we refer the reader to Sec. 2 in Pizzati et al. (2024a) and Sec. 2 and Appendix A in Pizzati et al. (2024b).

The framework developed in these works builds on large-volume, dark-matter-only cosmological simulations. In particular, Pizzati et al. (2024b) uses the new FLAMINGO-10k simulation (part of the FLAMINGO project, Schaye et al. 2023; Kugel et al. 2023), which evolves 10080^3 cold dark matter (CDM) particles and 5600^3 neutrino particles in a box size of $L = 2.8\,\mathrm{cGpc}$

[5]The median of the QHMF is defined as the halo mass $M_{\rm med}$ satisfying the relation $\int_{M_{\rm med}}^{\infty}\mathrm{QHMF}(M_{\rm h}) = 0.5\int_0^{\infty}\mathrm{QHMF}(M_{\rm h})$.

Table 4.1: Constraints (median values and 16th-84th percentiles) on the parameters describing the conditional luminosity function (CLF; eq. 4.1) of quasars at $z \approx 4$ and $z \approx 6$. Taken from Pizzati et al. (2024a,b).

Redshift	σ	$\log_{10} L_{\rm ref}$ [erg s^{-1}]	γ	$f_{\rm on}$ [%]
$z \approx 4$	$0.20^{+0.13}_{-0.08}$	$45.2^{+0.3}_{-0.3}$	$2.00^{+0.22}_{-0.23}$	51^{+32}_{-31}
$z \approx 6$	$0.55^{+0.37}_{-0.31}$	$46.45^{+0.79}_{-1.35}$	$3.17^{+0.32}_{-0.34}$	$3.9^{+21}_{-3.2}$

assuming the "3x2pt + all" cosmology from Abbott et al. (2022)[6]. The model includes subhalos, which are found using the upgraded Hierachical Bound-Tracing (HBT+) code (Han et al. 2012, 2018). Subhalo masses, $M_{\rm h}$, are defined as peak bound masses[7].

In the analysis performed in Pizzati et al. (2024a), we applied this framework to the quasar auto-correlation functions measured by Eftekharzadeh et al. (2015) ($z \approx 2.5$) and Shen et al. (2007) ($z \approx 4$) using wide-field spectroscopic surveys such as SDSS (York et al. 2000) and BOSS (Ross et al. 2013). In particular, we showed that the $z \approx 4$ clustering measurements of Shen et al. (2007) imply a characteristic host halo mass for quasars of $\log_{10} M_{\rm h}/{\rm M}_\odot \approx 13.3$, corresponding to a very large UV-luminous quasar duty cycle of $\varepsilon_{\rm QSO} = 33^{+34}_{-23}$ %. In Pizzati et al. (2024b), we extended the framework to interpret the quasar-galaxy cross-correlation function recently measured by Eilers et al. (2024) at $z = 6.25$. This work exploited the JWST NIRCam wide-field slitless spectroscopic mode to pick up [O III] emitting galaxies in quasars fields, and inferred the clustering of quasars by measuring the cross-correlation function between quasars and [O III] emitting galaxies in conjunction with the auto-correlation function of these galaxies. By simultaneously fitting these two quantities, Pizzati et al. (2024b) found a characteristic host mass for $z \approx 6$ quasars of $\log_{10} M_{\rm h}/{\rm M}_\odot \approx 12.5$, lower than the one found at $z \approx 4$ and in line with results at $z \approx 2.5$.

However, when converting these host halo masses into peak heights[8], $\nu(M_{\rm h})$ – which measure how rare the large-scale over-density fluctuations are in the original linear field – we find that quasar clustering measurements at $z \approx 4$ and $z \approx 6$ point to similar values of $\nu \approx 4 - 6$. This implies that high-z, UV-luminous quasars seem to live in similarly biased and over-dense

[6]The cosmology parameters are: $\Omega_{\rm m} = 0.306$, $\Omega_{\rm b} = 0.0486$, $\sigma_8 = 0.807$, $\rm H_0 = 68.1\ km\,s^{-1}\,Mpc^{-1}$, $n_{\rm s} = 0.967$; the summed neutrino mass is 0.06 eV.

[7]In practice, we compute the mass of each (sub)halo by summing up the mass of all its bound particles and consider the largest mass that a (sub)halo has had across cosmic history.

[8]The peak height $\nu(M_{\rm h}, z)$ is formally defined as $\nu(M_{\rm h}, z) = \delta_{\rm c}/\sigma(M_{\rm h}, z)$ – with $\delta_{\rm c} \approx 1.69$ being the critical linear density for spherical collapse and $\sigma^2(M_{\rm h}, z)$ the variance of the linear density field smoothed on a scale $R(M_{\rm h})$; we compute $\nu(M_{\rm h}, z)$ using the python package COLOSSUS (Diemer 2018, see Sec. 5 in Pizzati et al. 2024b).

environments, corresponding to $(4-6)\sigma$ peaks in the initial linear density field (see also, e.g., Costa 2024). Due to the rapid decline of the unobscured quasar number density with redshift (solid gray line in the right panel of Fig. 4.1), these similar environments lead to very different values for the quasar UV-luminous duty cycles at $z \approx 4$ and $z \approx 6$: while UV-luminous $z \approx 4$ quasars are sufficiently abundant to occupy a large fraction of the coeval $\nu \approx 4-6$ halos, at $z \approx 6$ quasars are so rare that the same occupation fraction drops by more than an order of magnitude, with an implied duty cycle of $\varepsilon_{\rm QSO} = 0.9^{+2.3}_{-0.7}\%$.

We report the inferred values of the parameters describing the CLF and the $L_{\rm c}(M)$ relation (eq. 4.1-4.2) at $z \approx 4$ and $z \approx 6$ in Table 4.1. Further discussion on the comparison between quasar clustering results at these two redshifts can be found in Sec. 5 of Pizzati et al. (2024b) (see also Eilers et al. 2024). We mention the caveat, however, that the strong clustering measured at $z \approx 4$ is rather surprising and it is yet to be fully accounted for by any evolutionary models of quasar activity (Pizzati et al. 2024a, and references therein). Additionally, several other studies (e.g., Timlin et al. 2018; He et al. 2018; García-Vergara et al. 2019) have also attempted to measure quasar clustering at $z \approx 4$, challenging the exceptionally strong clustering inferred by Shen et al. (2007). Nevertheless, the Shen et al. (2007) measurement remains the most robust, as it is based on a large sample of spectroscopically-selected quasars. Future spectroscopic surveys (such as DESI, Yang et al. 2023) will further refine these measurements and provide more stringent constraints on the quasar auto-correlation function up to $z \approx 5$. Here, we take the Shen et al. (2007) result at face value, but stress the fact that our conclusions for $z \approx 4$ and $z \approx 6$ are completely independent.

In Fig. 4.2, we show the QHMFs obtained by our model at $z \approx 4$ and $z \approx 6$, together with HMFs at the respective redshifts. As discussed above, the QHMF can be obtained only once a bolometric luminosity threshold for quasars has been set. Both quasar clustering measurements on which our work is based (Shen et al. 2007; Eilers et al. 2024) focus on very bright unobscured quasars with $L_{\rm bol} \approx 10^{47}\,{\rm erg\,s^{-1}}$, with the work of Shen et al. (2007) extending down to slightly fainter objects of $L_{\rm bol} > 10^{46.7}\,{\rm erg\,s^{-1}}$. For consistency (see also Appendix D of Pizzati et al. 2024b), we show our $z \approx 6$ QHMF results setting the same bolometric luminosity threshold employed by Shen et al. (2007) at $z \approx 4$ (i.e., $L_{\rm bol} = 10^{46.7}\,{\rm erg\,s^{-1}}$). The QHMFs obtained in this way are plotted in Fig. 4.2 with red ($z \approx 6$) and green ($z \approx 4$) lines, and labeled as "bright quasars" as they only refer to the bright end of the unobscured quasar population.

Fainter, unobscured quasars are found at both $z \approx 4$ and $z \approx 6$ down to $L_{\rm bol} \approx 10^{45.3}\,{\rm erg\,s^{-1}}$ (Akiyama et al. 2018; Kulkarni et al. 2019; Matsuoka et al. 2022). However, the clustering of this fainter population is still largely unconstrained in the high-z Universe. A first attempt at measuring the clustering of $z \approx 6$ faint quasars was made by Arita et al. (2023): despite the

Table 4.2: Constraints (median values and 16th-84th percentiles) on the UV-luminous active fraction $f_{\rm on,UV}$ (coming from clustering measurements; see Pizzati et al. 2024a,b) and on the obscured:unobscured ratio for LRDs, $r_{\rm obsc}$ (from abundance arguments; see Sec. 4.2), at $z = 4$ and $z = 6.25$. The product $f_{\rm on,UV} r_{\rm obsc}$ exceeds unity at both redshifts, which is unphysical. In the last columns, we also report the median mass, $M_{\rm med}^{\rm (faint)}$ for the halos hosting faint unobscured quasars (see Fig. 4.2) and the number density of halos above this mass, $n_{\rm h}(> M_{\rm med}^{\rm (faint)})$, to be compared with the LRD number density of $\Phi_{\rm LRD} \approx 1.3 \times 10^{-5}\,{\rm cMpc}^{-3}$.

Redshift	$\log_{10} f_{\rm on,UV}$	$\log_{10} r_{\rm obsc}$	$\log_{10} f_{\rm on,UV} \cdot r_{\rm obsc}$	$f_{\rm on,UV} \cdot r_{\rm obsc}$	$\log_{10} M_{\rm med}^{\rm (faint)}/{\rm M}_\odot$	$n_{\rm h}(> M_{\rm med}^{\rm (faint)})$	$\Phi_{\rm LRD}/n_{\rm h}(> M_{\rm med}^{\rm (faint)})$
$z = 6.25$	$-1.40^{+0.83}_{-0.74}$	2.9 ± 0.5	$1.5^{+1.0}_{-0.9}$	32^{+284}_{-28}	≈ 12.15	$\approx 7.2 \times 10^{-7}\,{\rm cMpc}^{-3}$	≈ 18
$z = 4$	$-0.29^{+0.21}_{-0.41}$	1.3 ± 0.3	$1.0^{+0.4}_{-0.5}$	10^{+15}_{-7}	≈ 12.76	$\approx 1.6 \times 10^{-6}\,{\rm cMpc}^{-3}$	≈ 8.1

large uncertainties at play, these authors find a relatively large characteristic host halo mass of $M_{\rm h} = 7^{+11}_{-6} \times 10^{12}\,{\rm M}_{\odot}$ (but see Appendix C of Pizzati et al. 2024b, where it is shown that different assumptions on the quasar correlation function make these constraints much weaker). The relatively large inferred host mass for the faint quasar population would be in line with results at lower redshift ($z \lesssim 2.5$), which generally predict little to no dependence of quasar clustering with bolometric luminosity (e.g., Shen et al. 2009; Eftekharzadeh et al. 2015).

As our model is based on an empirical relation between quasar luminosities and (sub)halo masses, it can be used to predict the clustering of faint unobscured quasars at high redshift. With light-colored lines in Fig. 4.2, we plot the predictions for the "faint quasars" QHMFs at the two redshifts of interest. These QHMFs are obtained by lowering the bolometric luminosity threshold, $L_{\rm thr}$ in eq. 4.3, down to $L_{\rm bol} = 10^{45.5}\,{\rm erg\,s^{-1}}$. We note that such a low bolometric luminosity threshold implies that the results are sensitive to the relation between *faint* quasar luminosities and host halo masses. This relation is based on the extrapolation of our CLF parametrization down to low $L_{\rm bol}$, and it currently lacks support by constraints on the clustering of faint unobscured quasars. However, our fitting framework matches the unobscured QLF over the entire range of magnitudes, from the very bright to the very faint end, with a minimal number of parameters. Therefore, while faint quasar clustering studies will ultimately test our predictions, the QHMFs shown in Fig. 4.2 for faint quasars represent our best knowledge of how faint quasars populate the host halo mass spectrum, and are informed by our current understanding of unobscured quasar demographics.

At $z \approx 6$ (left panel of Fig. 4.2), we predict that the "faint quasars" QHMF peaks at $\log_{10} M_{\rm h}/{\rm M}_{\odot} \approx 12.15$, with a rather large spread in the host mass distribution (0.5 dex at 1 standard deviation). This implies a very mild dependence of clustering on bolometric luminosity, as a change of ≈ 1 dex in $L_{\rm bol}$ results in a change of ≈ 0.3 dex in the median of the host mass distribution, $M_{\rm med}$. This mild dependence is driven by two factors: a steep $L_{\rm bol} - M_{\rm h}$ relation and a large scatter around this relation[9]. These results are in broad agreement with clustering studies at low redshift, which find little to no dependence of clustering strength on luminosity (Croom et al. 2005; Myers et al. 2006; Shen et al. 2009) and attribute that to a large scatter in quasar luminosities at fixed halo mass (e.g., Adelberger & Steidel 2005; Lidz et al. 2006).

The strong clustering measured for bright quasars at $z \approx 4$ implies a slightly different dependence of quasar clustering on luminosity, with ≈ 1 dex in $L_{\rm bol}$ corresponding to ≈ 0.5 dex in $M_{\rm med}$. Such a luminosity dependence is a consequence of the large duty cycle measured for bright quasars: if

[9]The slope of the $L_{\rm bol} - M_{\rm h}$ relation and its scatter are directly constrained by a combination of the quasar clustering strength and the *shape* of the quasar luminosity function (Pizzati et al. 2024a).

these quasars occupy a large fraction of the available massive halos, fainter quasars will inevitably need to reside in less massive hosts. In practice, this is achieved in our model with a small predicted scatter for the $z \approx 4$ $L_{\rm bol} - M_{\rm h}$ relation (also found by White et al. 2008; Wyithe & Loeb 2009; Shankar et al. 2010b). The slightly different dependence of clustering on luminosity at the two redshifts considered, while interesting, has little impact on the conclusions presented in this work: at both redshifts, faint quasars also live in massive halos corresponding to highly biased environments which trace back to rare $\approx 4\sigma$ fluctuations in the linear density field.

4.3.2 Connecting the UV-luminous duty cycle to the AGN obscured population

Having described current constraints on the duty cycle and host mass distribution of UV-luminous, unobscured quasars, we turn our attention to the large population of LRDs/obscured AGN discussed in Sec. 4.2. The most general question connected to this obscured high-z population is how it fits into our understanding of SMBH accretion/AGN activity across the history of the Universe. In this context, determining whether LRDs and UV-selected quasars are drawn from the same population of dark matter halos can offer key insights into the nature of these sources. According to AGN unification models (e.g., Antonucci 1993; Padovani et al. 2017), the diversity of AGN emission across the electromagnetic spectrum can be entirely explained by a viewing-angle effect: the intrinsic emission from a quasar/AGN varies for different lines of sight because of, e.g., dust and gas obscuration. The natural consequence of this model is that all types of AGN (irrespective of their observed SEDs) share the same *intrinsic* properties, such as the bolometric luminosity, SMBH mass, and host halo mass distributions. Hence, if LRDs fit into this AGN unification picture, we expect them to reside in the same halos as comparably luminous UV-selected quasars. However, several studies at low z have challenged this AGN unification scenario by showing that obscured (type-2 or reddened type-1) quasars live in different dark matter halos than those of UV-luminous, type-1 quasars (e.g., Hickox et al. 2011; Allevato et al. 2014; Petter et al. 2023; Córdova Rosado et al. 2024). According to these studies, obscured quasars/AGN represent a different stage in the co-evolution between accreting SMBHs and their host galaxies/halos. Analogously, LRDs could also represent a different evolutionary phase in the accretion history of SMBHs. If that is the case, the host halo mass distribution of LRDs could be different than the one of unobscured quasars, even when matching their bolometric luminosities and SMBH masses. An obvious consequence of this hypothesis is that LRDs would be described by very different scaling relations (e.g., SMBH mass-halo/galaxy mass) than those in place for UV-luminous quasars, as identical SMBH masses would correspond to very different host halo/galaxy masses.

In this work, we point out that an indirect answer to whether LRDs and UV-selected quasars reside in the same dark matter halos comes from current constraints on the clustering of quasars at $z \approx 4-6$ (Sec. 4.3.1). From these constraints, we conclude that LRDs and unobscured quasars cannot be drawn from the same host halo distribution. Hence, their different SED properties reveal fundamental differences in their scaling relations. Our argument is simple: clustering measurements determine the host mass distribution of unobscured quasars; if LRDs followed the same distribution, the large obscured fraction derived in Sec. 4.2 implies that LCDM cosmology would not produce enough halos at these masses to accommodate this abundant population.

The argument can be visualized in Fig. 4.2: using dark (light) purple lines, we show the QHMFs of bright (faint) quasars scaled up by the obscured:unobscured ratios, $r_{\rm obsc}$, determined in Sec. 4.2 (plotted with colored arrows for reference). These obscured ratios are independent of bolometric luminosities, and increase rapidly with redshift from $r_{\rm obsc} \approx 20^{+20}_{-10} : 1$ at $z = 4$ to $r_{\rm obsc} \approx 815^{+1600}_{-545} : 1$ at $z = 6.25$. By multiplying the QHMF by $r_{\rm obsc}$, we are effectively computing the host mass distribution for LRDs/obscured AGN under the hypothesis that they reside in the same kind of halos as UV-luminous quasars. At both $z = 6.25$ (left panel in Fig. 4.2) and $z = 4$ (right), the host halo mass distributions for LRDs exceed the respective halo mass functions (HMFs). This is unphysical: cosmology sets hard (and well-constrained) limits on the number of (sub)halos that are available as quasar hosts as a function of mass, and the LRD number densities appear to be incompatible with these limits[10].

We can quantify this by considering the UV-luminous active fraction, $f_{\rm on,UV}$, which is a parameter in our CLF model (see eq. 4.1) and is closely related to the UV-luminous duty cycle (Pizzati et al. 2024a). The parameter $f_{\rm on,UV}$ represents the fraction of SMBHs that are actively accreting and unobscured at the same time. If we multiply this UV-luminous active fraction by the obscured:unobscured ratio $r_{\rm obsc}$, we are effectively computing an "obscured" active fraction (i.e., the fraction of halos hosting actively accreting LRDs/obscured AGN). The physical limit set by the number of available sub(halo) hosts can be then rephrased as $f_{\rm on,UV} \cdot r_{\rm obsc} < 1$. In Table 4.2, we report the values of $f_{\rm on,UV}$ and $r_{\rm obsc}$ and of their product at the two redshifts of interest, $z = 6.25$ and $z = 4$. We find that, despite the large uncertainties at play, these products are significantly larger than unity, with a value of $f_{\rm on,UV} \cdot r_{\rm obsc} \approx 10$ at $z = 4$ and $f_{\rm on,UV} \cdot r_{\rm obsc} \approx 36$ at $z = 6.25$. Coming back to the visual representation in Fig. 4.2, the product

[10]Note that this argument is valid only for a maximum occupation fraction of unity. Given that we model the distribution of all subhalos, however, it is natural to assume that each subhalo can host at most one accreting SMBH at its center.

$f_{\rm on,UV} \cdot r_{\rm obsc}$ represents the maximum ratio between the scaled-up QHMFs (purple lines; see also eq. 4.3) and the HMFs (dashed lines).

An even simpler way to present this argument is to consider the median mass values, $M_{\rm med}$, for, e.g., the faint-quasar QMHFs (Fig. 4.2, light-colored lines). In the right-hand columns of Table 4.2, we report these $M_{\rm med}^{\rm (faint)}$ values at the two redshifts of interest, together with the number density of halos above these mass thresholds, $n_{\rm h}(> M_{\rm med}^{\rm (faint)})$. When compared to the number density of LRDs, $\Phi_{\rm LRD} \approx 1.3 \times 10^{-5}\,{\rm cMpc}^{-3}$ (which is approximately constant with redshift; see Sec. 4.2), these number densities are a factor of ≈ 8.1 (≈ 18) smaller at $z = 4$ ($z = 6.25$). This confirms the fact that LRDs are simply too numerous to live in the same (sub)halos as UV-luminous quasars. As discussed in Sec. 4.3.1, these halo masses correspond to similar environments at $z \approx 4$ and $z \approx 6$ (i.e., $(4-6)\sigma$ peaks in the linear density field; Fig. 4.2). Since the number density of these environments is roughly constant with redshift (e.g., Tinker et al. 2008) (and so is $\Phi_{\rm LRD}$), LRDs outnumber their candidate host halos by similar factors at the two redshifts considered.

As a final note, we point out that our results are valid for any values of the quasar bolometric luminosities. Yet, in Sec. 4.2 and 4.3.1 we considered bright ($L_{\rm bol} > 10^{46.7}\,{\rm erg\,s^{-1}}$) and faint ($10^{45.5}\,{\rm erg\,s^{-1}} < L_{\rm bol} < 10^{46.7}\,{\rm erg\,s^{-1}}$) quasars separately because their properties are constrained differently. In particular, the abundance of obscured AGN is better constrained at faint bolometric luminosities by the large sample of LRDs with $L_{\rm bol} \approx 10^{46}\,{\rm erg\,s^{-1}}$; the evidence for an analog obscured population at large bolometric luminosities is instead only tentative (Sec. 4.2). On the other hand, the clustering of bright unobscured quasars has been directly measured (Sec. 4.3), but the QHMF and duty cycle for the faint quasar population is solely based on the extrapolation of our model to fainter bolometric luminosities – which constrained to match the faint end of the QLF. For this reason, the results presented lead to different conclusions depending on the bolometric luminosities considered. If a large obscured population is indeed present at $L_{\rm bol} \approx 10^{47}\,{\rm erg\,s^{-1}}$, then this is already in direct conflict with with constraints on the host masses and duty cycle of bright unobscured quasars (Shen et al. 2007; Eilers et al. 2024). A measurement of quasar clustering at the faint end of the QLF ($L_{\rm bol} \gtrsim 10^{45.5}\,{\rm erg\,s^{-1}}$), on the other hand, would provide support for our predictions for the properties of faint unobscured quasars, and will make it possible to directly compare the properties of UV-luminous quasars and LRDs at the same bolometric luminosities.

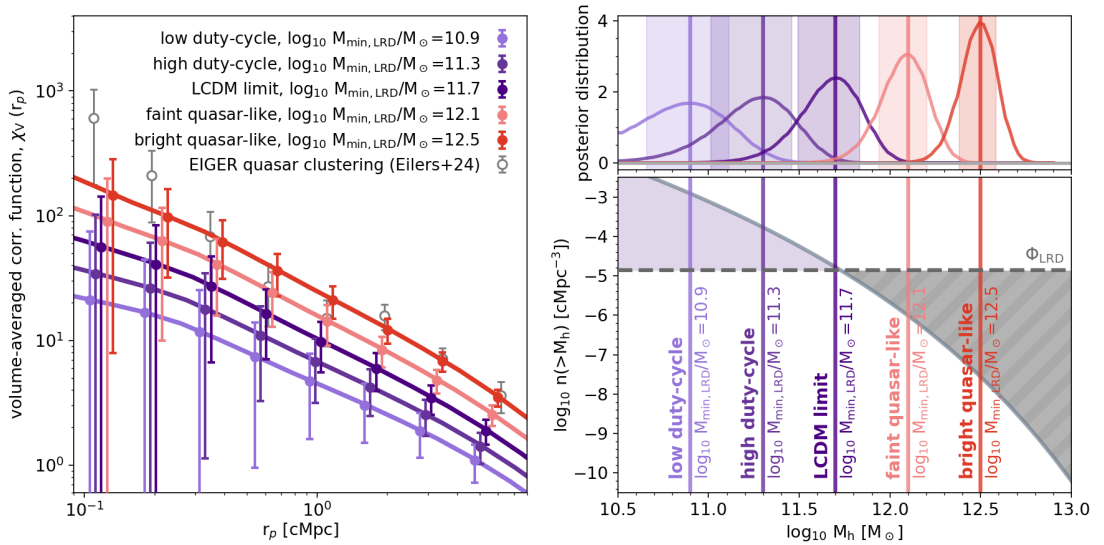

Figure 4.3: *Left:* Mock measurements (colored data points) for the LRD-galaxy cross-correlation functions obtained for different values of the minimum host mass for LRDs, $M_{\rm min,LRD}$. The measurements are obtained by putting together 10 different LRD fields, and extracting galaxy counts by setting a minimum host mass for galaxies (i.e., [O III] emitters) of $M_{\rm min,OIII} = 10^{10.56}\,{\rm M}_\odot$ and a background galaxy number density of $n_{\rm OIII} = 7.84 \times 10^{-4}\,{\rm cMpc}^{-3}$. The theoretical predictions for these cross-correlation functions are coming from the model of Pizzati et al. (2024b) and are shown as solid colored lines. Error bars are computed by assuming Poisson uncertainties on the galaxy number counts. Gray points refer to the UV-luminous quasar-galaxy cross-correlation function measurements from the EIGER survey (Eilers et al. 2024). *Top right:* Mock inference analysis for the LRD-galaxy cross-correlation function measurements, as a function of the minimum host LRD mass, $M_{\rm min,LRD}$. Values of $M_{\rm min,LRD}$ considered for the mock measurements are color-coded as in the other panels. The posterior distributions are obtained by computing the agreement between the mock measurements and the theoretical models for different minimum host LRD masses. Shaded regions show the 16th and 84th percentiles of their respective posterior distributions. *Bottom right:* Number density of $z = 6.25$ halos above $M_{\rm h}$, $n_{\rm h}(> M_{\rm h})$, as a function of halo mass $M_{\rm h}$ (solid gray line). The values of $M_{\rm min,LRD}$ considered in the analysis are highlighted with colored vertical lines. The dashed horizontal line corresponds to the LRD number density, $\Phi_{\rm LRD}$. The purple (gray) shaded area shows the region for which $M_{\rm h} < 10^{11.7}\,{\rm M}_\odot$ ($M_{\rm h} > 10^{11.7}\,{\rm M}_\odot$). In the purple region, $\Phi_{\rm LRD} < n_{\rm h}(> M_{\rm h})$ and hence the number of LRDs is less than the number of host halos available, whereas the gray region is unphysical as LRDs are too abundant for the number of host (sub)halos (assuming a maximum occupation fraction of unity).

4.4 The host mass and duty cycle of little red dots: a mock analysis

The indirect arguments presented in the previous section suggest that LRDs cannot live in the same dark matter halos as unobscured UV-luminous quasars, and hence – provided their bolometric luminosities are correctly estimated – they may constitute a fundamentally different population of accreting SMBHs. How do we determine this new population's host halo masses and duty cycle? In this section, we argue that this can be done using current (and upcoming) JWST observations.

Existing JWST programs such as EIGER (Kashino et al. 2023; Eilers et al. 2024) and ASPIRE (Wang et al. 2023) have already shown that the clustering of luminous, UV-selected quasars can be effectively measured using JWST NIRCam slitless spectroscopy to study the distribution of [O III] line emitting galaxies in the neighboring regions of the quasars. The same strategy can be applied to any other population of objects: the cross-correlation between this population and [O III] line emitters at a certain redshift can be measured, and the clustering of this said population can be inferred by simultaneously constraining the auto-correlation function of the [O III] line emitters.

In the following, we examine a simple proof-of-concept analysis that aims to measure the clustering of LRDs using JWST[11]. We focus here on $z = 6.25$, which is the redshift at which the clustering of UV-luminous quasars with [O III] emitters has already been measured by the EIGER survey (Eilers et al. 2024, see also Sec. 4.3.1). Following Eilers et al. (2024) (see also, e.g., Kaiser 1984; Martini & Weinberg 2001; Haiman & Hui 2001), we postulate that LRDs inhabit a fraction of all the (sub)halos whose mass is larger than some minimum mass threshold, $M_{\rm min,LRD}$[12]. This fraction is equal to the LRD duty cycle, $\varepsilon_{\rm LRD}$, and can be determined by comparing the LRD number density ($\Phi_{\rm LRD}$ in Sec. 4.2) to the abundance of halos with $M_{\rm h} > M_{\rm min}$. We note that we only consider LRDs with quasar-like bolometric luminosities (i.e., with the same bolometric luminosities as faint $z \approx 6$ quasars, $L_{\rm bol} > 10^{45.5}\,{\rm erg\,s^{-1}}$), as we are interested in matching LRDs and UV-luminous quasars in $L_{\rm bol}$ space.

[11] An alternative approach would be to directly measure the auto-correlation function of LRDs. Even though LRDs have a relatively high number density, however, measuring an autocorrelation function would require very large samples that are challenging to obtain given the small FoV of JWST.

[12] In other words, we do not model the LRD host mass distribution parametrically as described in Sec. 4.3 for unobscured quasars, but we assume that such a distribution can be obtained by rescaling the HMF above the minimum mass threshold $M_{\rm min,LRD}$. A more sophisticated parametrization would result in large degeneracies in the parameter space that could not be resolved by clustering measurements alone (e.g., Pizzati et al. 2024a; Muñoz et al. 2023).

We consider five different values of the minimum host (sub)halo mass for LRDs: $\log_{10} M_{\rm min,LRD}/{\rm M}_\odot = 10.9, 11.3, 11.7, 12.1, 12.5$. In the bottom right panel of Fig. 4.3, we put these values into context by showing the number density of $z = 6.25$ halos above $M_{\rm h}$, $n_{\rm h}(> M_{\rm h})$, as a function of halo mass (solid grey line); we highlight the values of $M_{\rm min,LRD}$ considered with colored vertical lines. By comparing the LRD number density ($\Phi_{\rm LRD}$, dashed horizontal line) to the integrated halo mass function ($n_{\rm h}(> M_{\rm h})$) for different minimum halo masses, we can directly relate the abundance of LRDs to that of available host dark matter halos. We find that the number of LRDs equals the number of host halos (i.e., the duty cycle is equal to unity) for a minimum host mass of $M_{\rm min,LRD} \approx 10^{11.7}\,{\rm M}_\odot$. Assuming that there can be only one LRD per (sub)halo, values of $M_{\rm min,LRD}$ above this threshold mass are unphysical. Values significantly lower than this threshold, on the other hand, imply a low duty cycle for LRDs, as only very few (sub)halos host LRDs at any given time.

Based on this discussion, we refer to the five different $M_{\rm min,LRD}$ cases considered in the following way (see Fig. 4.3): "low duty-cycle" ($M_{\rm min,LRD}/{\rm M}_\odot = 10^{10.9}\,{\rm M}_\odot$), "high duty-cycle" ($M_{\rm min,LRD}/{\rm M}_\odot = 10^{11.3}\,{\rm M}_\odot$), "LCDM limit" ($M_{\rm min,LRD}/{\rm M}_\odot = 10^{11.3}\,{\rm M}_\odot$), "faint quasar-like" ($M_{\rm min,LRD}/{\rm M}_\odot = 10^{12}\,{\rm M}_\odot$), "bright quasar-like" ($M_{\rm min,LRD}/{\rm M}_\odot = 10^{12.5}\,{\rm M}_\odot$). The first case ("low duty-cycle") corresponds to a duty cycle of $\varepsilon_{\rm LRD} \approx 1\%$, which is close to the duty cycle measured by Pizzati et al. (2024b) for UV-luminous quasar activity at the same redshift. In the second case, the implied LRD duty cycle increases to $\varepsilon_{\rm LRD} \approx 10\%$. The third case corresponds to the physical limit of a duty cycle of $\approx 100\%$. The last two cases, instead, would imply a duty cycle above unity, and correspond to host masses characteristic of UV-luminous quasars. Based on the discussion of Sec. 4.3.1, we associate the case $M_{\rm min,LRD} = 10^{12.1}\,{\rm M}_\odot$ to faint ($L_{\rm bol} \gtrsim 10^{45.5}\,{\rm erg\,s^{-1}}$) quasars – which have the same $L_{\rm bol}$ as LRDs –, while the larger mass of $M_{\rm min,LRD} = 10^{12.5}\,{\rm M}_\odot$ is close to the one found for luminous ($L_{\rm bol} \approx 10^{47}\,{\rm erg\,s^{-1}}$) unobscured quasars by Eilers et al. (2024).

The question we want to address here is whether we can use clustering measurements based on JWST slitless spectroscopy data to distinguish between these different $M_{\rm min,LRD}$ cases. We consider the following mock setup: JWST/NIRCam grism is used to image 10 different LRD fields. The distribution of [O III] line emitters in these fields can be employed to measure a LRD-galaxy cross-correlation function, from which the host mass and duty cycle of LRDs can be determined by exploiting the constraints on the galaxy-galaxy auto-correlation function (Eilers et al. 2024, Huang et al., in prep.).

In practice, we use the framework developed in Pizzati et al. (2024b), which outputs the cross-correlation function of any populations of objects that

are tracers of the underlying distribution of dark matter halos[13]. We employ this model to predict the LRD-galaxy cross-correlations for the different values of $M_{\rm min,LRD}$. [O III] line emitters are assumed to live in halos with a fixed threshold mass of $M_{\rm min,OIII} = 10^{10.56}\,{\rm M}_\odot$, which is set according to the results of Eilers et al. (2024) (see also Huang et al., in prep.). Based on these cross-correlation functions, we generate mock measurements by computing the expected number of galaxies as a function of the projected distance in each LRD field. The expected galaxy counts are obtained by setting a background galaxy number density of $n_{\rm OIII} = 7.84 \times 10^{-4}\,{\rm cMpc}^{-3}$, which is obtained by integrating the [O III] emitter luminosity function of Matthee et al. (2023) down to the threshold luminosity of $L_{\rm OIII,5008} = 10^{42}\,{\rm erg\,s^{-1}}$. We put together the 10 mock LRD fields and we compute the volume-averaged cross-correlation function, χ_V, by projecting the galaxy 3-d distributions over a comoving distance of $\pi_{\rm max} = 9.8\,{\rm cMpc}$, corresponding to a line-of-sight velocity of 1000 $\rm km\,s^{-1}$ at the redshift considered.

In the left panel of Fig. 4.3, we show the mock LRD-galaxy cross-correlation functions for different values of $M_{\rm min,LRD}$. We also show for reference the UV-luminous quasar-galaxy cross-correlation function measured by Eilers et al. (2024) by putting together 4 different quasars fields from the EIGER survey (Kashino et al. 2023). We note that, as also done in Eilers et al. (2024), the error bars we show are computed by considering only the contribution of Poisson uncertainties on the number counts. Other contributions to the error budget, such as cosmic variance or possible correlations between different data points, are neglected in this work and will be analyzed in a forthcoming study (Huang et al., in prep.).

The precision of our inference analysis is shown in the bottom left panel of Fig. 4.3. These posterior distributions are obtained by fitting the mock data with the LRD-galaxy cross-correlation function models obtained by varying the LRD mass threshold parameter, $\log_{10} M_{\rm min,LRD}/{\rm M}_\odot$. For each of these models, we compute the value of the χ^2 and plot in Fig. 4.3 the quantity $\exp(-\chi^2/2)$ (normalized to unity). By looking at the different posterior distributions, we learn that by putting together 10 LRD fields we can already constrain the values of $M_{\rm min,LRD}$ (and hence the characteristic host mass of LRDs) with an uncertainty of $\approx 0.1 - 0.3$ in $\log_{10} M_{\rm h}$. The posteriors are narrower and more peaked for larger $M_{\rm min,LRD}$. This follows from the fact that high-mass halos are more strongly clustered, and hence the clustering signal is stronger for large $M_{\rm min,LRD}$ (left panel of Fig. 4.3). In all cases considered, the uncertainty in $M_{\rm min,LRD}$ is small enough that,

[13]We use the FLAMINGO-10k large-volume cosmological simulation (Sec. 4.3) to build an analytical model for the cross-correlation function of any sets of halos with masses M_j and M_k, $\xi_h(M_j, M_k; r)$. An appropriate weighting scheme can then return the cross-correlation function between two different halo tracer populations. For more details on the model and the cosmological simulation employed, we refer the reader to Pizzati et al. (2024b).

in principle, it could be possible to tell apart the different scenarios. A larger number of LRD fields would be necessary, however, to reduce the uncertainties on $M_{\rm min,LRD}$ even further, and pinpoint its value even for the case of low $M_{\rm min,LRD}$.

The discussion presented here shows that, by measuring how galaxies cluster in LRD fields, it is indeed possible to determine whether LRDs live in the same dark matter halos as unobscured quasars (in agreement with, e.g., the AGN unification framework) or whether they are hosted by more common and less-biased environments, as it appears to be necessary given their large number density (see Sec. 4.3). In this latter case, measuring the host mass distribution of LRDs would also provide a way to quantify their duty cycle ($\varepsilon_{\rm LRD}$), which is a fundamental quantity that can help us to shed light on the accretion history of these enigmatic objects. A large value of $\varepsilon_{\rm LRD} \approx 10\%$ would suggest that LRDs have been actively accreting for a large fraction of cosmic time ($\gtrsim 100\,{\rm Myr}$), and hence – assuming a standard value for the radiative efficiency – that they would be able to build the relatively large black hole masses that have been inferred from their broad optical lines (up to $\gtrsim 10^8\,{\rm M}_\odot$; e.g., Greene et al. 2024; Kocevski et al. 2024). In particular, an accretion timescale of $\gtrsim 100\,{\rm Myr}$ corresponds to $\gtrsim 2t_{\rm S}$, where $t_{\rm S} \approx 45\,{\rm Myr}$ is the Salpeter time for exponential black hole mass growth (Salpeter 1964). This implies that LRDs are detectable above the observational luminosity threshold for at least a few Salpeter times, which is expected if the survey spans about one order of magnitude in luminosity. We point out that, for the same reason, large duty cycles of $\gtrsim 50\%$ are not to be expected, because they would imply that almost all LRDs shine above the observational threshold for a time that is much longer than the Salpeter timescale. Considering again a survey spanning about one order of magnitude in luminosity, a standard Eddington-limited growth that remains above the observational threshold for a time $t \gg 2t_{\rm S}$ would result in black holes that grow much more than one order of magnitude, and hence end up being more massive than what is actually observed. For this reason, while the threshold mass of $M_{\rm min,LRD} \approx 10^{11.7}\,{\rm M}_\odot$ represents a limit set by cosmological constraints on the number of available (sub)halos, black hole formation physics suggests an even more stringent limit on $M_{\rm min,LRD}$: if we require $\varepsilon_{\rm LRD} \lesssim 30\%$, this implies that $M_{\rm min,LRD}$ needs to be lower than $\approx 10^{11.5}\,{\rm M}_\odot$.

A very low LRD duty cycle $\varepsilon_{\rm LRD} \lesssim 1\%$, on the other hand, would also be puzzling, as it would raise the question of how to reconcile the large black hole masses measured for LRDs with their inherently sporadic activity. This is the same problem that has been brought up for the high-z UV-luminous quasar population, for which different methods generally infer low values of the quasar duty cycle and/or quasar lifetime (e.g., Khrykin et al. 2016, 2019; Eilers et al. 2018, 2020; Davies et al. 2018, 2019, 2020; Worseck et al. 2016, 2021; Ďurovčíková et al. 2024; Eilers et al. 2024) that appear to be in

tension with their large, $\gtrsim 10^9\,\mathrm{M}_\odot$ black hole masses. A possible solution to explain a low value of the duty cycle is super-Eddington accretion: if accretion on black holes takes place in short, radiatively inefficient bursts, then we expect a lower $\varepsilon_{\mathrm{LRD}}$ because the Salpeter timescale for black hole accretion becomes shorter than $\approx 45\,\mathrm{Myr}$. Interestingly, several studies have invoked super-Eddington accretion to explain the puzzling SED features of LRDs (e.g., Greene et al. 2024; Pacucci & Narayan 2024; Lambrides et al. 2024a). Measuring the clustering of LRDs and inferring their duty cycle would provide an independent way to support these claims.

Finally, if bright LRDs have large black hole masses ($\gtrsim 10^8\,\mathrm{M}_\odot$) but live in much smaller halos than UV-selected quasars, they need to obey fundamentally different scaling relations. Constraining the clustering of LRDs would provide insights into these relations: the lower the mass of the host halos, for instance, the more overmassive LRDs need to be with respect to the black hole mass-halo mass relation holding for unobscured quasars. We can also cast this in terms of the black hole mass-stellar mass relation – which has been extensively discussed in the recent literature (e.g., Pacucci et al. 2023; Yue et al. 2024a) – by converting halo masses to stellar masses using the relation provided by Behroozi et al. (2019). We find that halo masses in the range $M_{\mathrm{h}} \approx 10^{11} - 10^{11.5}\,\mathrm{M}_\odot$ correspond – at the redshift of interest – to stellar masses of $M_\star \approx 10^{8.4} - 10^{9.4}\,\mathrm{M}_\odot$. This implies that, assuming black hole mass measurements are not significantly overestimated, LRDs are highly overmassive with respect to the coeval black hole mass-stellar mass relation, as the ratio between black hole and galaxy masses would be in the range $\approx 10\% - 100\%$ (see also, e.g., Durodola et al. 2024).

4.5 Discussion and summary

In this work, we have examined how the new population of Little Red Dots (LRDs) revealed by JWST compares to the one of UV-luminous quasars that have been studied for decades using wide-field spectroscopic surveys (e.g. Fan et al. 2023). The basic observational evidence on which our work is based, is that a large fraction of LRDs exhibits broad emission lines in their spectra, whose properties directly point to the presence of AGN that are (at least partially) responsible for the observed emission (Greene et al. 2024; Kocevski et al. 2024). This, together with their very red colors at optical wavelengths, has led to the interpretation that LRDs could be standard, UV-luminous type-1 quasars whose radiation is (partially) obscured by intervening dust and gas. By correcting for the effects of this obscuration, it is possible to use broad lines to estimate the bolometric luminosities of the SMBHs accreting at the center of LRDs. Several works (e.g., Greene et al. 2024; Kokorev et al. 2024a; Akins et al. 2024) have shown that such (unattenuated) bolometric

luminosities extend up to $\approx 10^{46} - 10^{47}\,\mathrm{erg\,s^{-1}}$, well in the range that is characteristic of unobscured, type-1 quasars (Fig. 4.1, left panel).

Yet, the abundances of LRDs and UV-luminous quasars are remarkably different. In Fig. 4.1, we have directly compared the redshift evolution for the number density of UV-luminous quasars to the one for LRDs at the same bolometric luminosities. It is well-established that the abundance of unobscured quasars drops exponentially with increasing redshift (e.g., Richards et al. 2006; Schindler et al. 2023). Spectroscopic (Greene et al. 2024) and photometric (Kokorev et al. 2024a) surveys of LRDs, instead, find little to no evolution in their number density over a wide redshift range ($z \approx 4-8$), with an approximately constant value of $\Phi_{\mathrm{LRD}} \approx 1.3 \times 10^{-5}\,\mathrm{cMpc^{-3}}$ ($L_{\mathrm{bol}} > 10^{45.5}\,\mathrm{erg\,s^{-1}}$). By comparing the number density of LRDs to that of UV-luminous quasars as a function of redshift, we can estimate the obscured fraction of AGN implied by this LRD population. We infer a large and rapidly evolving obscured fraction, ranging from $\approx 20:1$ at $z \approx 4$ to $\approx 2300:1$ at $z \approx 7$.

While this obscured fraction is mostly constrained at the bolometric luminosities for which a significant overlap between LRDs and unobscured quasars is present (i.e., $L_{\mathrm{bol}} \approx 10^{45} - 10^{46}\,\mathrm{erg\,s^{-1}}$), we find tentative evidence for it to extend to even larger bolometric luminosities ($L_{\mathrm{bol}} \gtrsim 10^{47}\,\mathrm{erg\,s^{-1}}$). There are two arguments in support of this evidence: (a) photometric observations (Kokorev et al. 2024a) constrain the *shape* of the LRD bolometric luminosity functions to closely resemble that of UV-luminous quasars (Nida et al. 2020; Schindler et al. 2023; Matsuoka et al. 2023), implying an obscured fraction that is constant with bolometric luminosity (Fig. 4.1, left panel); (b) recent observations of the COSMOS field have revealed candidate radio-loud AGN at $z \approx 7-8$ that are obscured in the UV (Endsley et al. 2022, 2023; Lambrides et al. 2024b); the very large bolometric luminosities of these sources ($L_{\mathrm{bol}} \approx 10^{47}\,\mathrm{erg\,s^{-1}}$) together with the small FoV of the observations, implies an AGN obscured fraction that is consistent to the one inferred for bright LRDs.

The large abundance of LRDs/obscured AGN has implications for their host halo masses. If obscuration were solely a viewing angle effect (e.g., Antonucci 1993), then we would expect LRDs to reside in the same environments as (equally bolometrically bright) UV-luminous quasars. Two decades of quasar clustering studies have constrained the masses of the dark matter halos hosting UV-luminous quasars at $0 \lesssim z \lesssim 6$ to be in the range $M_h \approx 10^{12} - 10^{13.5}\,\mathrm{M_\odot}$ (e.g., Porciani et al. 2004; Croom et al. 2005; Porciani & Norberg 2006; Shen et al. 2007, 2009; Ross et al. 2013; Eftekharzadeh et al. 2015; Arita et al. 2023; Eilers et al. 2024), with little to no dependence on quasar luminosity (e.g., Adelberger & Steidel 2005; Porciani & Norberg 2006; Shen et al. 2009). Several models have been put forward to interpret this host mass range in physical terms (e.g., Hopkins et al. 2007b; Fanidakis et al. 2013; Caplar et al. 2015). Whatever the reason for these characteristic

host masses, it is striking that the number density of available host halos at these masses drops very quickly below the measured abundance of LRDs as redshift increases. At $z \approx 6$, for example, LRDs (with $L_{\rm bol} > 10^{45.5}\,{\rm erg\,s^{-1}}$) are $\approx 5\times$ more abundant than $10^{12}\,{\rm M_\odot}$ halos (Fig. 4.3, top right panel) and can occupy all halos above the threshold mass of $\tilde{M}_{\rm h} > 10^{11.7}\,{\rm M_\odot}$. This implies that at these redshifts the host masses of LRDs are likely lower than the ones of UV-luminous quasars, even when matching them in $L_{\rm bol}$ space.

In Fig. 4.2, we have presented a quantitative analysis of this argument at the two redshifts for which we have constraints on the clustering of bright ($L_{\rm bol} \approx 10^{47}\,{\rm erg\,s^{-1}}$), high-$z$ unobscured quasars: $z = 4$ (Shen et al. 2007) and $z = 6.25$ (Eilers et al. 2024). We used the model developed in Pizzati et al. (2024a,b) to measure the UV-luminous quasar host mass functions (QHMFs) at these two redshifts. While these QHMFs are well-constrained by clustering measurements only for the bright quasar population, we can extend them to also include the contribution of faint ($L_{\rm bol} \gtrsim 10^{45.5}\,{\rm erg\,s^{-1}}$) quasars by using the empirical quasar luminosity-halo mass relations obtained by Pizzati et al. (2024a,b). These relations are fit to the faint end of the quasar luminosity function, and hence they correctly reproduce the demographic properties of the faint quasar population. While we find minor differences in the luminosity dependence of the QHMFs at the two redshifts considered, we reach a general fundamental conclusion that is valid for faint and bright sources alike: *the dark matter halos hosting UV-luminous quasars at* $z \gtrsim 4$ *are too rare to accommodate the large number density of LRDs.*

What are the implications of these findings? If LRDs live in more common and hence less biased halos than those of unobscured quasars, then they may represent an intrinsically different population of accreting SMBHs arising in the early Universe. This population may be tracing a distinct phase in the co-evolutionary sequence of SMBHs and galaxies, similarly to what has been argued for type-2/reddened quasars at low redshifts (e.g., Allevato et al. 2014; Córdova Rosado et al. 2024). In this scenario, the scaling relations between, e.g., black hole and halo/galaxy host masses need to be intrinsically different for LRDs and standard unobscured quasars, because similar black hole masses correspond to very different halo (and hence galaxy) masses. In particular, LRDs likely host SMBHs that are overmassive with respect to the coeval black hole-halo/galaxy mass scaling relations for unobscured quasars. Another possibility that has been put forward by several independent works to explain the enigmatic features of LRD SEDs (e.g., Greene et al. 2024; Pacucci & Narayan 2024; Lambrides et al. 2024a) is that LRDs are accreting at rates that are larger than the critical Eddington limit. In this latter case, LRDs could represent the early stages of black hole accretion and growth that are predicted by many theoretical models of SMBH evolution (e.g., Trinca et al. 2023; Li et al. 2024; Lupi et al. 2024). Interestingly, this would have direct implications for the clustering of LRDs, because a low duty cycle

(that is necessary for super-Eddington accretion) would only be possible if LRDs lived in very low mass halos ($M_{\rm h} \approx 10^{11}\,{\rm M}_{\odot}$ at $z \approx 6$; Sec. 4.4).

Alternatively, these results may be telling us that key properties of LRDs, such as their bolometric luminosities and the relative contribution of the central AGN and the host galaxy to their observed SEDs, have yet to be properly characterized. Indeed, the assumption on which our discussion is based, is that LRDs have the same bolometric luminosities as high-z UV-luminous quasars ($L_{\rm bol} \approx 10^{45} - 10^{47}\,{\rm erg\,s^{-1}}$). Currently, the bolometric luminosities of LRDs are estimated by their (dereddened) broad emission lines or by fitting AGN templates to their SEDs. In both cases, the resulting $L_{\rm bol}$ hinge on the assumption that the rest-frame optical continuum is dominated by AGN light (see e.g., Akins et al. 2024). If the contribution of the host galaxy to the rest-frame optical continuum emission (and possibly broad lines; see, e.g., Baggen et al. 2024) is non-negligible, then the inferred black hole masses and bolometric luminosities could change significantly. Several puzzling features of LRDs, such as their X-ray weakness (Ananna et al. 2024; Yue et al. 2024b; Maiolino et al. 2024) and (possibly) the lack of a hot dust torus (Wang et al. 2024; Pérez-González et al. 2024; Akins et al. 2024; Iani et al. 2024) and UV variability (Kokubo & Harikane 2024), point to the fact that LRD bolometric luminosities could be vastly overestimated. The presence of an evolved stellar population dominating (part of) the rest-frame optical is also suggested by the detection of a Balmer break in some LRD spectra (e.g., Wang et al. 2024; Kokorev et al. 2024b, but see Inayoshi & Maiolino 2025), although the large densities and stellar masses required to match the observed LRD luminosities remain a significant challenge to a purely stellar interpretation of LRD SEDs (e.g., Greene et al. 2024; Akins et al. 2024, but see Baggen et al. 2024). Regardless of the exact AGN contribution to these SEDs, if LRDs are not as bright as standard, UV-luminous quasars then they would naturally reside in lower mass halos, and they could easily be accommodated in the large number of $z \gtrsim 6$ host halos with masses of $M_{\rm h} \approx 10^{11} - 10^{11.5}\,{\rm M}_{\odot}$.

In this work, we have primarily focused on the population of LRDs whose inferred SMBH masses and bolometric luminosities largely overlap with those of UV-luminous quasars. However, JWST has uncovered a much larger population of AGN with broad optical (Hα or Hβ) lines, which are not necessarily reddened at optical wavelengths and hence do not respect the LRD selection criteria. Interestingly, the abundance of these broad-line AGN are even larger than the ones of LRDs: Maiolino et al. (2024), Harikane et al. (2023), and Taylor et al. (2024) find the number densities for these sources to be in the range $10^{-3} - 10^{-5}\,{\rm cMpc^{-3}mag^{-1}}$ ($4 \lesssim z \lesssim 7$). The intrinsic bolometric luminosities and SMBH masses of these broad-line AGN (that are not reddened in the rest-frame optical) are not as extreme as the ones of LRDs/reddened AGN (e.g., Harikane et al. 2023; Taylor et al. 2024). However, these sources can still reach UV magnitudes of $M_{\rm UV} \approx -22$

and bolometric luminosities of $L_{\rm bol} \approx 10^{45.5}\,{\rm erg\,s^{-1}}$, which are close to the ones of the faintest UV-selected quasars known at $z \gtrsim 4$ (Matsuoka et al. 2022). Given their number densities, these broad-line AGN overshoot the extrapolation of the UV-selected quasar luminosity functions by factors that are comparable to (or even higher than) those found for LRDs (Sec. 4.2). Hence, similar arguments to the ones presented in our analysis apply to this larger AGN population: their abundance suggests that they live in halos that are likely less massive than those of comparably luminous UV-selected quasars, implying that they obey fundamentally different scaling relations. While a proper comparison between UV-selected quasars and JWST AGN is only possible for the LRD population with large inferred bolometric luminosities and SMBH masses, it is interesting to investigate the host mass distributions, duty cycles, and scaling relations of this larger population of faint broad-line AGN.

Ultimately, a measurement of the clustering of LRDs and other broad-line AGN will constrain such properties and test the conclusions that we have drawn in this work. Recent arguments on the clustering of these objects rely on single detections of AGN in close proximity (Lin et al. 2024; Tanaka et al. 2024), on spectroscopic detections of galaxies in a single LRD field (Schindler et al. 2024), and on cross-correlating photometrically-selected galaxies and LRDs (Arita et al. 2025). In this work (Fig. 4.3), we have shown that a convincing measurement of LRD clustering can be made by using JWST NIRCam/WFSS observations of several LRD fields to extract a cross-correlation function between LRDs and [O III] line-emitting galaxies (see also Matthee et al. 2024a for recent results based on a similar approach). We have suggested that, by putting together ≈ 10 different fields, it is possible to infer the characteristic host halo mass of LRDs with an accuracy of $\log_{10} M_{\rm h} \approx 0.1 - 0.3$. In order to perform this kind of measurement, one would need to observe several fields containing LRDs using a NIRCam grism filter covering the [O III] doublet. Interestingly, such observations already exist for a fraction of the broad-line AGN in the sample of Matthee et al. (2024b): JWST surveys such as CONGRESS (GO3577) and GTO4540/GTO4549 are performing NIRCam/WFSS observations of the GOODS-N and GOODS-S fields, which contains ≈ 10 broad-line AGN from the Matthee et al. (2024b) sample. So a first step towards determining the clustering of these enigmatic sources at $z \gtrsim 5$ is already feasible with current data. Future JWST programs will be able to deliver the same kind of observations for samples of LRDs with quasar-like inferred bolometric luminosities and SMBH masses. By comparing the host halo masses resulting from these measurements to the different scenarios discussed in Sec. 4.4, it will be possible to get fundamental insights into the properties of these objects.

At the same time, the clustering of the faint, UV-luminous quasar population at high redshifts is also largely unconstrained. By using the same

strategy and targeting faint quasar fields with NIRCam/WFSS, it will also be possible to determine their clustering and host masses. This would test our model predictions (Fig. 4.2) and determine the luminosity dependence of quasar clustering at high-z, effectively constraining the scaling relation between the quasar bolometric luminosity and the host halo mass. Even more importantly, it would create a benchmark to which the LRD population can be effectively compared, allowing us to investigate the nature of quasar activity and SMBH populations in the early Universe.

Acknowledgements

We acknowledge helpful conversations with Junya Arita, Jorryt Matthee, Jacob Shen, Fengwu Sun, Marta Volonteri, Minghao Yue, and Ben Wang. We are grateful to the FLAMINGO-10k team for making their simulation available. We are also grateful to the ENIGMA group at UC Santa Barbara and Leiden University for discussion of an early version of this manuscript. JFH and EP acknowledge support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 885301). JTS is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 518006966. FW acknowledges support from NSF Grant AST-2308258. This work is partly supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860744 (BiD4BESt). This work used the DiRAC Memory Intensive service (Cosma8) at the University of Durham, which is part of the STFC DiRAC HPC Facility (www.dirac.ac.uk). Access to DiRAC resources was granted through a Director's Discretionary Time allocation in 2023/24, under the auspices of the UKRI-funded DiRAC Federation Project. The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.