**Astronomy & Astrophysics**

Consistency of JWST black hole observations with NANOGrav gravitational wave measurements

John Ellis[1,2,3] , Malcolm Fairbairn[2] , Gert Hütsi[3] , Juan Urrutia[3,4,*], Ville Vaskonen[3,5,6] , and Hardi Veermäe[3]

[1] Theoretical Physics Department, CERN, Geneva, Switzerland
[2] King's College London, Strand, London, WC2R 2LS, UK
[3] Keemilise ja Bioloogilise Füüsika Instituut, Rävala pst. 10, 10143 Tallinn, Estonia
[4] Departament of Cybernetics, Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia
[5] Dipartimento di Fisica e Astronomia, Università degli Studi di Padova, Via Marzolo 8, 35131 Padova, Italy
[6] Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Via Marzolo 8, 35131 Padova, Italy

ABSTRACT

JWST observations have opened a new chapter in supermassive black hole (SMBH) studies, stimulating discussion of two puzzles: the abundance of high-z SMBHs and the fraction of dual active galactic nuclei (AGNs). We argue that the answers to these puzzles may be linked to an interpretation of the data on the nanohertz gravitational waves (GWs) discovered by NANOGrav and other pulsar timing arrays as SMBH binaries whose evolution is driven by interactions with their environments down to $O(0.1\,\mathrm{pc})$ separations. We show that the stellar mass–black hole mass correlations found in JWST data and in low-z inactive galaxies are similar, and present a global fit to these data, excluding low-z AGNs. Matching the NANOGrav and dual-AGN data requires that binary evolution due to environmental effects at separations below $O(1\,\mathrm{kpc})$ be rapid on cosmological timescales. According to this interpretation, the SMBHs in low-z AGNs are the tip of the iceberg of a local SMBH population in mainly inactive galaxies. This interpretation is consistent with the 'little red dots' observed with JWST being AGNs, and would favour the observability of GW signals from black hole binaries in LISA and decihertz GW detectors.

Key words. quasars: supermassive black holes – cosmology: theory – early Universe

1. Introduction

There has been some surprise at the number of supermassive black holes (SMBHs) discovered with JWST at high redshifts, z (see Übler et al. 2023; Larson et al. 2023; Harikane et al. 2023; Bogdan et al. 2024; Ding et al. 2023; Maiolino et al. 2024a; Yue et al. 2023). Prior expectations for the number of SMBHs in the early Universe were largely based on empirical relationships between the SMBH mass–halo mass ratio and the SMBH mass–stellar mass ratio, such as that found by Reines & Volonteri (2015, hereafter RV15) in the local (low-z) population of active galactic nuclei (AGNs). From a statistical analysis, Pacucci et al. (Pacucci et al. 2023, hereafter P23) conclude that there was a $>3\sigma$ discrepancy between the SMBH mass/stellar mass ratios, $M_{\rm BH}/M_*$, of a subset of the high-z JWST data and this fit to local AGN data. We note, however, that RV15 also analysed a sample of local inactive galaxies (IGs), finding higher values of $M_{\rm BH}/M_*$. Moreover, larger values of $M_{\rm BH}/M_*$ had also been found by Izumi et al. (2021, hereafter I21) in a compilation of high-z SMBHs observed previously by JWST.

Another surprise in the JWST data was the discovery of several close pairs in a sample of AGNs (Perna et al. 2023), whereas cosmological simulations generally predicted at least an order of magnitude fewer dual AGNs. This observation was used in Padmanabhan & Loeb (2024) to study SMBH binary evolution. JWST observations have also revealed a population of 'little red dots' (LRDs) in the redshift range $z \sim 2-11$ that are interpreted as previously hidden AGNs that trace black hole (BH) growth in the early Universe (Labbé et al. 2023b; Akins et al. 2023; Barro et al. 2024; Labbé et al. 2023a; Matthee et al. 2024; Kocevski et al. 2024).

In parallel with the emergence of JWST data, pulsar timing arrays (PTAs) have detected the existence of a stochastic gravitational wave background (SGWB) in the nanohertz range (Agazie et al. 2023c; Antoniadis et al. 2024; Reardon et al. 2023; Xu et al. 2023), whose most plausible astrophysical interpretation is gravitational wave (GW) emission by SMBH binaries. Analysing their data under this assumption and guided by astrophysical priors, the NANOGrav Collaboration (Agazie et al. 2023b) found that their data were best fit by a $M_{\rm BH}/M_*$ ratio that was significantly higher than that found in the RV15 analysis of local AGNs. In parallel, in Ellis et al. (2024a) we found – using the extended Press-Schechter formalism (Press & Schechter 1974; Bond et al. 1991) to estimate the halo merger rate and assuming a fixed probability, $p_{\rm BH}$, for a halo merger to yield a BH merger – that a good fit to the NANOGrav data was possible with the $M_{\rm BH}/M_*$ ratio found by RV15 in their analysis of local IGs. It was also found in Ellis et al. (2024a) that using the RV15 fit to local AGNs significantly degraded the fit to the NANOGrav data: $\Delta\chi^2 = 43$ compared to the fit based on local IGs (see Fig. 9 of Ellis et al. 2024a and the discussion in the accompanying text, as well as the analysis below).

It has been suggested recently that the $M_{\rm BH}/M_*$ relationship may evolve with z (Pacucci & Loeb 2024), or that selection biases and measurement uncertainties (Li et al. 2024) may account for much of the tension between the JWST high-z SMBH data and the data on local AGNs considered in RV15.

* Corresponding author; juan.urrutia@kbfi.ee

However, we took the prima facie evidence at face value to be able to compare the information on SMBHs provided by the high-z SMBH data and NANOGrav measurements. The puzzle of the masses of JWST BHs may be linked to the somewhat unexpected strength of the GW signal reported by NANOGrav and other PTAs (Agazie et al. 2023a), and can be reframed as a question as to why both datasets correspond better to the data on local IGs than to local AGNs.

In this paper we quantify in more detail the consistency between (i) the NANOGrav nanohertz GW data and (ii) the JWST data and previous data on high-z SMBHs, discussing both the SMBH mass/stellar mass ratio and the fraction of dual AGNs in the JWST data. The paper is organised as follows: In Sect. 2 we analyse the BH mass-stellar mass relations found in recent high-z data from JWST and other observations, showing that they are quite consistent with data on low-z IGs. In Sect. 3 we show that the high-z SMBH data are also consistent with our analysis of the NANOGrav data on the nanohertz GW background (Ellis et al. 2024a). In Sect. 4 we show that our interpretation of the NANOGrav and JWST data can also accommodate the unexpectedly high fraction of dual AGNs observed with JWST (Perna et al. 2023), if the hardening timescale due to halo-halo interactions is short on cosmological timescales. We also show that our interpretation of the JWST and NANOGrav data is consistent with the population of LRDs reported recently (Matthee et al. 2024; Kocevski et al. 2024), assuming that they are AGNs triggered during mergers. We present our conclusions in Sect. 5.

2. Analysis of the BH mass–stellar mass relation

We considered the 33 high-z AGNs observed by JWST compiled in the appendix. In addition, we also considered a sample of 21 high-z AGNs observed previously, extracted from the observations compiled in I21. Both samples are shown in Fig. 1. The errors in the BH masses were reported in I21, but not the uncertainties in the stellar masses: we estimated them to be approximately 0.3 dex. The figure also shows the local IG and AGN measurements compiled in RV15. For clarity, we do not show the error bars for these points, but we take the uncertainties reported in RV15 into account in the following analysis.

The new JWST SMBH observations have been interpreted by P23 as inconsistent with the local mass relationship, prompting, for example, the proposals that the SMBH-stellar mass relationship evolves with redshift (Pacucci & Loeb 2024) or that selection biases and measurement uncertainties in the high-z population may be largely responsible for the apparent tension with the local relationship (Li et al. 2024). As mentioned earlier, rather than pursuing these hypotheses, we recall the tension between different local active and inactive SMBH populations RV15 and emphasise the consistency of the JWST measurements with the local IG measurements as well as the other high-z measurements.

As commonly done in the literature, we assumed a power-law form for the relation between the SMBH mass and the stellar mass of the host galaxy, with a log-normal intrinsic scatter. The probability distribution of the SMBH mass, m, for a given stellar mass, M_*, is then parametrised by the magnitude, a, and logarithmic slope, b, of the mean and the width, σ, of the distribution:

$$\frac{\mathrm{d}P(m|M_*,\boldsymbol{\theta})}{\mathrm{d}\log_{10}m} = \mathcal{N}\left(\log_{10}\frac{m}{M_\odot}\middle| a + b\log_{10}\frac{M_*}{10^{11}M_\odot}, \sigma\right), \quad (1)$$

where $\boldsymbol{\theta} = (a, b, \sigma)$ and $\mathcal{N}(x|\bar{x}, \sigma)$ denotes the probability density of a Gaussian distribution with mean $\bar{x}$ and variance σ^2.
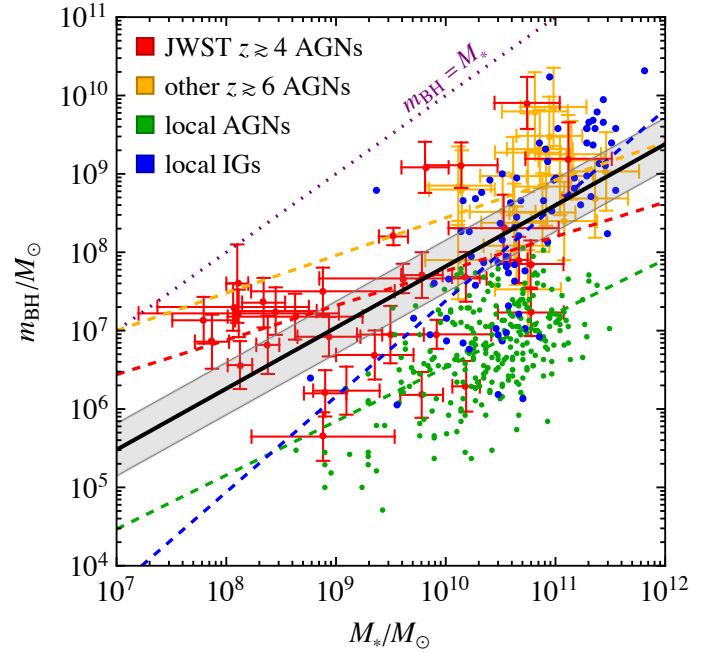
Fig. 1. Measurements and fits of the BH mass–stellar mass relation. The data points with errors are from JWST measurements (red; see the appendix) and other high-z measurements (yellow) from I21. The local AGN measurements (green) and the local IG measurements (blue), both from RV15, are shown without errors, but the errors are taken into account in the analysis. The dashed lines show the best fits to each of the datasets separately, and the solid black line shows the best fit combining the high-z data and the local IG data. The grey band shows the scatter in the latter fit.

We fitted the parameters a, b, and σ to the measured data. The likelihood function is

$$\mathcal{L}(\boldsymbol{\theta}) \propto \prod_j \int \mathrm{d}\log_{10}m \, \mathrm{d}\log_{10}M \, \frac{\mathrm{d}P(m|M,\boldsymbol{\theta})}{\mathrm{d}\log_{10}m}$$
$$\times \mathcal{N}(\log_{10}m|\log_{10}m_j)\mathcal{N}(\log_{10}M|\log_{10}M_j), \quad (2)$$

where the product is over the data points, m_j and M_j denote the mean measured BH and stellar masses, and we assumed that the posteriors of the BH and stellar mass measurements are log-normal and uncorrelated. We used flat priors for a, b, and σ with $b > 0$.

As seen in Fig. 1, the high-z AGN measurements are largely consistent with each other and with the fit to the local IGs, though the latter prefers a slightly stronger tilt. The posteriors in a and b found in a global fit that includes all observations except the local AGNs, marginalised over σ, are shown in the upper panel of Fig. 2, and the posteriors of the fits to individual datasets are shown in the appendix. The lower panel of Fig. 2 displays uncertainty ellipses for the individual datasets as well as that for the global fit (shown in grey) that omits the local AGN data (shown in green). The global best fit is

$$a = 8.6, \quad b = 0.8, \quad \sigma = 0.8, \quad (3)$$

and the 68% CL parameter ranges are indicated by the vertical dashed lines in the upper panel of Fig. 2. The fit to the local AGN population prefers, instead, much lower a values but has a slope b similar to that in the global fit. We note that the linear relation ($b = 1$) is compatible with our global fit at the 68% CL.
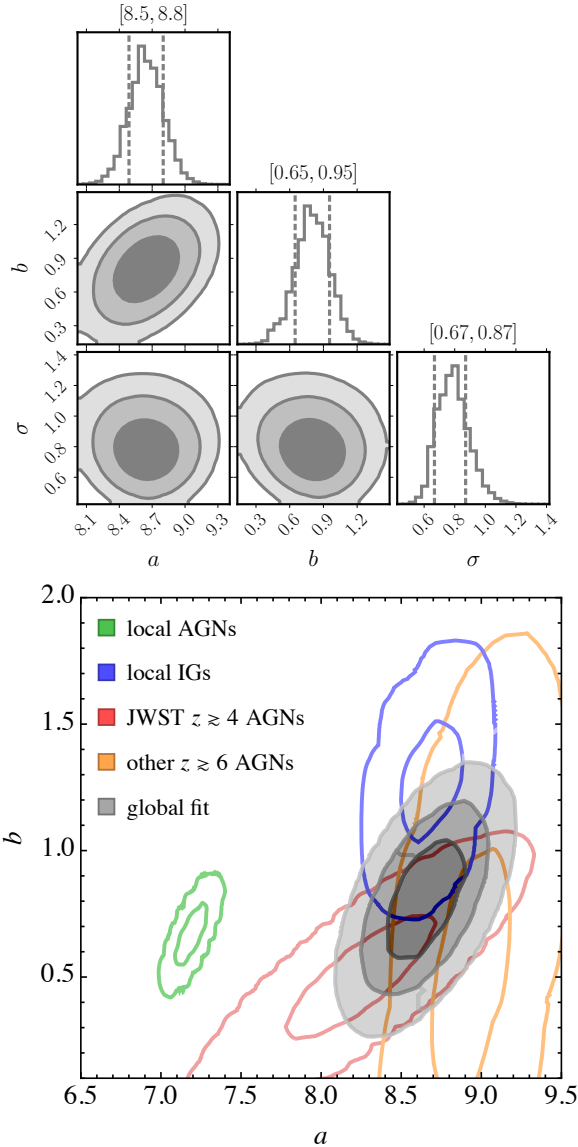
Table 1. Overlaps Eq. (4) between the posteriors obtained in the fits to different datasets.

	JWST $z \gtrsim 4$ AGNs	Other $z \gtrsim 6$ AGNs	Local AGNs	Local IGs	Global fit
JWST $z \gtrsim 4$ AGNs	1	0.07	10^{-6}	0.007	0.3
Other $z \gtrsim 6$ AGNs		1	10^{-9}	0.02	0.15
Local AGNs			1	10^{-9}	10^{-9}
Local IGs				1	0.08
Global fit					1

Notes. The cells of the table are colour-coded for clarity.

The overlap coefficients are tabulated in Table 1. We see numerically that the fit to the JWST data has better overlaps with the fits to the other high-z SMBHs and local IGs, as well as the global fit, than with the fit to local AGNs. The local AGN data overlap only slightly with the other high-z SMBHs, the local IGs, and the global fit.

3. Analysis of NANOGrav observations

We performed the analysis of the NANOGrav GW signal in the same way as in Ellis et al. (2023, 2024a,b). We summarise here the main features of the model: We related the SMBH merger rate to the halo merger rate, R_h, estimated from the Press-Schechter formalism (Press & Schechter 1974; Bond et al. 1991; Lacey & Cole 1993), as

$$
\frac{\mathrm{d}R_{\mathrm{BH}}}{\mathrm{d}m_1\mathrm{d}m_2} \approx p_{\mathrm{BH}} \int \mathrm{d}M_{\mathrm{v},1}\mathrm{d}M_{\mathrm{v},2} \frac{\mathrm{d}R_h}{\mathrm{d}M_{\mathrm{v},1}\mathrm{d}M_{\mathrm{v},2}}
$$
$$
\times \prod_{j=1,2} \frac{\mathrm{d}P(m_j|M_*(M_{\mathrm{v},j},z))}{\mathrm{d}m_j},
\tag{5}
$$

where p_{BH} characterises the BH merger efficiency and $M_*(M_{\mathrm{v},j},z)$ is the halo mass-stellar mass relation that we took from Girelli et al. (2020). We included environmental effects parametrised by a spectral index, α, and a reference frequency,

$$
t_{\mathrm{int}} = t_{\mathrm{GW}} \left[\frac{f_r}{f_{\mathrm{GW}}} \right]^{\alpha},
\tag{6}
$$

where f_r is the frequency of the emitted GW. The environmental effects drive the binary evolution at frequencies above f_{GW}, which we parametrised as $f_{\mathrm{GW}} = f_{\mathrm{ref}}(\mathcal{M}/M_{\odot})^{-\beta}$. We fixed $\beta = 5/8$, corresponding to binary evolution by gas infall (see the discussion below Eq. (12)).

In Ellis et al. (2024a) we found that the local IG fit also gives a much better fit to the NANOGrav GW observations than the local AGN fit, and we now develop this point further. Figure 3 displays in a corner plot the posterior distributions for the parameters of this fit: α, f_{ref} and p_{BH}. The posterior for α is compatible with models of environmental effects and the fit prefers large values of $f_{\mathrm{ref}} \gtrsim 10$ nHz. The preferred range of p_{BH} is $[0.16, 0.38]$. The best fit is achieved at $p_{\mathrm{BH}} = 0.37$, $f_{\mathrm{ref}} = 27$ nHz and $\alpha = 3.3$. In particular, the value $\alpha = 8/3$ expected for gas infall agrees well with the data. By fixing $\alpha = 8/3$, we find that the best fit is given by $p_{\mathrm{BH}} = 0.37$, $f_{\mathrm{ref}} = 30$ nHz.

In the upper panel of Fig. 4 we compare our global fit Eq. (3) with the NANOGrav SGWB data (Agazie et al. 2023c) and the best fit SMBH binary GW signals obtained with the BH mass-stellar mass relations derived in P23 and Li et al. (2024). The
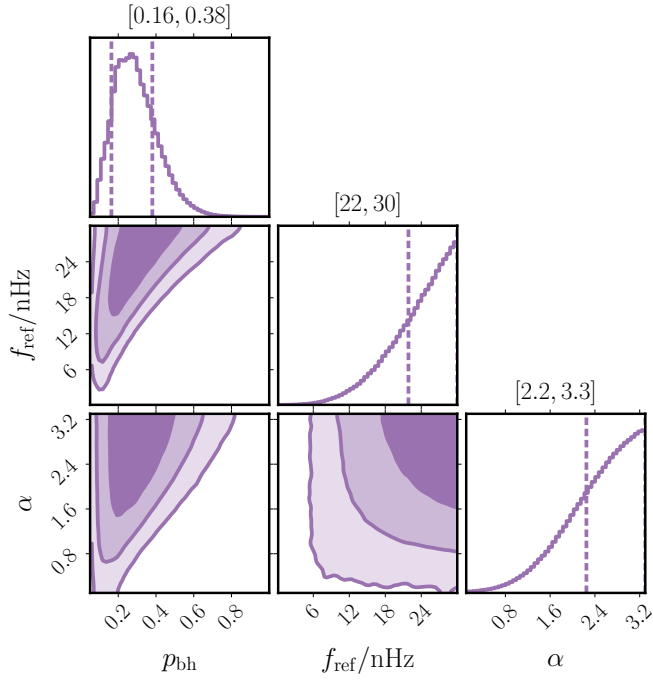
Fig. 2. *Upper panel*: corner plot showing the posteriors for the global fit to JWST (Übler et al. 2023; Larson et al. 2023; Harikane et al. 2023; Bogdan et al. 2024; Ding et al. 2023; Maiolino et al. 2024a; Yue et al. 2023) and previous high-z SMBH data from I21 as well as low-z RV15 data on SMBHs in IGs. *Lower panel*: 68, 95, and 99.8% CL contours (grey) of the posteriors for a and b defined in Eq. (1), obtained in a global fit that includes all observations except local AGNs by marginalising over σ. For comparison, the 68 and 99.8% CL of individual datasets are shown: JWST high-z measurements (red; Übler et al. 2023; Larson et al. 2023; Harikane et al. 2023; Bogdan et al. 2024; Ding et al. 2023; Maiolino et al. 2024a; Yue et al. 2023), other high-z measurements from I21 (yellow), local IGs (blue), and local AGNs from RV15 (green). The fits were performed using flat priors and including the uncertainties in the observations. Corner plots showing the posterior distributions for these fits are displayed in the appendix.

We quantified the overlaps between the posteriors for the fits to the different datasets i and j using the overlap coefficient

$$
O_{i,j} = \int \mathrm{d}\boldsymbol{\theta} \min \left[\mathcal{L}_i(\boldsymbol{\theta}), \mathcal{L}_j(\boldsymbol{\theta}) \right].
\tag{4}
$$

For example, for two Gaussian distributions of similar width σ with a mean separation of 4σ the overlap coefficient is $O \approx 0.05$.

Fig. 3. Posterior distributions of the parameters for a fit to NANOGrav GW data (Agazie et al. 2023c) including energy loss due to interactions with the environment.

Fig. 4. *Upper panel*: best-fit nanohertz GW signals found using the BH mass–stellar mass relations proposed in Pacucci & Loeb (2024), Li et al. (2024), and our global fit (3) compared to the signal measured by NANOGrav (grey). We imposed $\alpha = 8/3$, as expected from gas infall (see the discussion below Eq. (12)). *Lower panel*: chirp masses and redshifts of 100 realisations of the strongest binary contributing to the first bin of the upper panel for the scenarios proposed in Pacucci & Loeb (2024), and Li et al. (2024) compared with our analysis (Ellis et al. 2024a). Contours of $\log_{10} \Omega_{\mathrm{GW}}$ are shown for comparison.

data and the best-fit simulations are represented by 'violins' whose widths represent the posteriors as functions of the GW density: following Ellis et al. (2024a), we calculated the overlaps between the violins to assess the qualities of the fits. For both P23 and Li et al. (2024), the best fits saturate at $p_{\mathrm{BH}} = 1$ (i.e. every galactic merger leads to a GW signal), which is an indication that the generated SGWB is insufficient. Comparing the best-fit likelihoods, we find that the PTA GW background favours the global fit mass relation at higher than the 3σ CL compared to the scaling relations proposed in Pacucci et al. (2023) and Li et al. (2024).

We show in the lower panel of Fig. 4 100 random realisations of the strongest binary contributing to the first NANOGrav bin, $f_{\mathrm{GW}} \in (2-4)$ nHz, for each of the three scenarios. In the case of Li et al. (2024), who advocated for a strong bias in the scaling relation, we can see that the SMBHs are too light, whereas in the case of Pacucci & Loeb (2024), who advocate for a steep z dependence, the heavy SMBH are merging mainly at $z > 1$, which also decreases the amplitude significantly.

Furthermore, in the lower panel of Fig. 4 we also see that BHs with $m_{\mathrm{BH}} < 10^9 \, M_\odot$ are very unlikely to be among the strongest sources contributing to the SGWB. Given that the mass and z range probed by NANOGrav coincides with the dynamically measured SMBH masses, the high merger efficiency necessary to explain the observed GW background implies that the local IG population is not significantly biased. The local AGNs lie in a lower mass range, and it is an open question whether that population is biased. However, given the consistency between the high-z AGN measurements and the local IG measurements and the fact that at high z the AGN fraction is close to unity but drops quickly at $z \lesssim 1$ (Aird et al. 2018; Georgakakis et al. 2017), it seems that the local AGN population reflects the intrinsic low-mass tail of the full BH mass-stellar mass relationship.
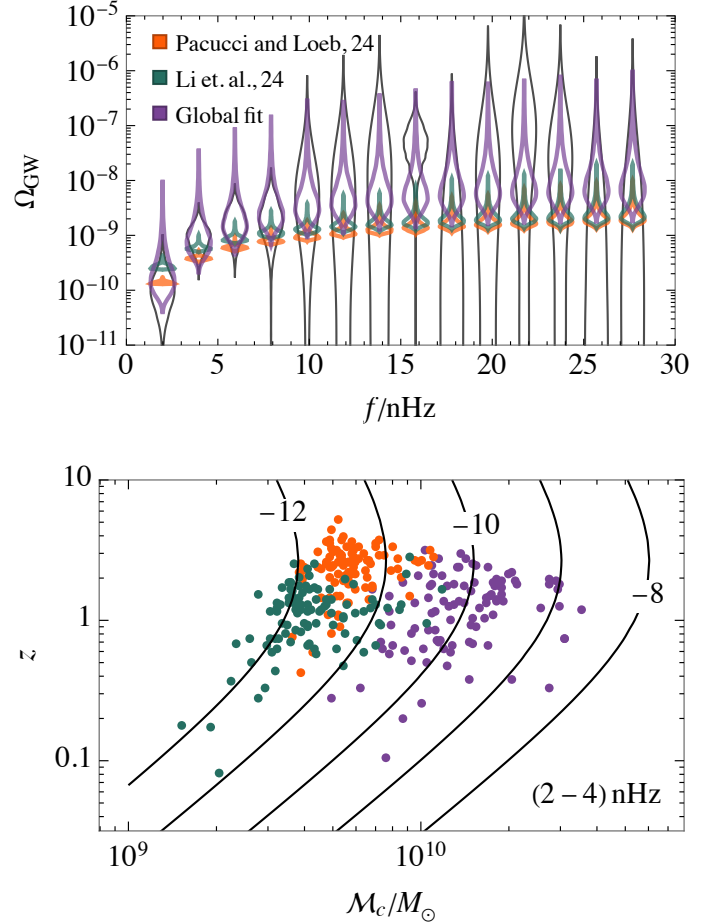
Comparing the local AGN population to the global fit in Fig. 1 we see that most of the local AGNs lie more than 2σ below the mean of the global fit. Hence, according to this analysis, the AGN fraction in the local SMBH population can only be $\lesssim 5\%$, in agreement with X-ray observations (Aird et al. 2018; Georgakakis et al. 2017). To conclude, in contrast to Li et al. (2024), our analysis indicates that the local AGN population represents the tip of the iceberg of the local SMBH population. As discussed in Ellis et al. (2024a,c), our interpretation of the NANOGrav and JWST data would favour the observability of GW signals from BH binaries in space-borne laser interferometers, for example LISA (Amaro-Seoane et al. 2017), TianQin (Wang et al. 2019), and Taiji (Ruan et al. 2020), and decihertz GW detectors, for example AION (Badurina et al. 2020), AEDGE (El-Neaj et al. 2020), and DECIGO (Kawamura et al. 2021).

4. Dual AGN fraction

We found in Ellis et al. (2024a,c) that the quality of the SMBH fit to the NANOGrav data is improved by allowing for the
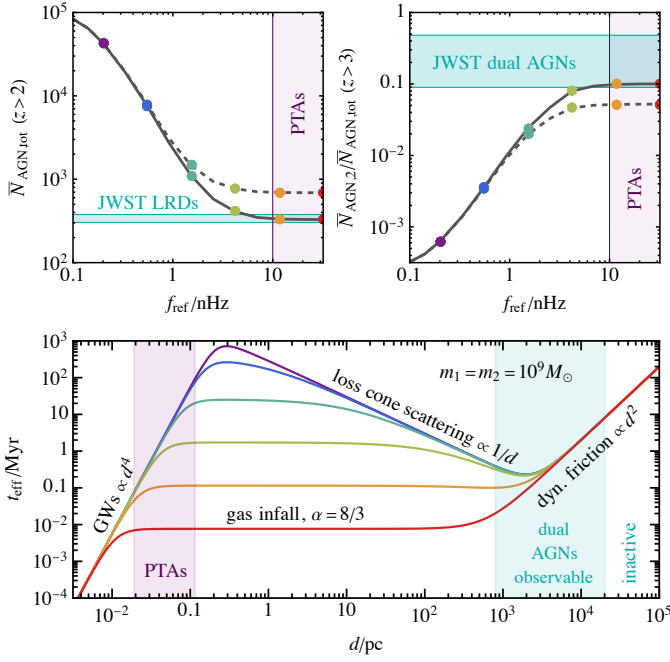
Fig. 5. *Upper panels*: expected total number of AGNs at $z > 2$ and the expected dual fraction at $z > 3$ as a function of the reference frequency, below which the binaries are driven by GW emission, assuming $p_{BH} = 0.2$, $\alpha = 8/3$, $A_{dyn} = 0.001$, and $A_{sh} = 1$. The dashed and solid curves are with and without galaxy mergers and where one of the galaxies does not contain an SMBH. *Lower panel*: illustration of the timescales of the environmental energy loss mechanisms of SMBH binaries as functions of their separations for the same parameter values as in the upper panels. The vertical bands illustrate the separation ranges probed by the PTAs and by the JWST dual AGN observations. The colour coding of the curves matches that shown in the upper panels.

evolution of the binary system to be affected by interaction with the environment as well as the emission of GWs. A similar conclusion was reached by the NANOGrav Collaboration (Agazie et al. 2023b). We discuss now the extent to which the unexpectedly high fraction of dual AGNs observed with JWST (Perna et al. 2023) can be accommodated within our analysis of high-z SMBH data and NANOGrav SGWB measurements including environmental effects.

We estimated the numbers of dual AGNs and objects that we identified as single AGNs assuming that the collisions of galaxies trigger AGN activity. This is in agreement with numerical simulations (Scudder et al. 2012; Callegari et al. 2011; Capelo et al. 2015) and low-z observations (Treister et al. 2012; Comerford et al. 2015; Donley et al. 2018; Koss et al. 2012), except for a small number of objects (Liu et al. 2018; Mićić et al. 2023). In this way, the AGN observations are linked to the NANOGrav observations and to the BH mass-stellar mass relation through the galaxy merger rate. Environmental effects determine the evolution of the separations of the components in dual systems; therefore, to estimate the numbers of dual AGNs and objects seen as single AGNs, we started by characterising the effective timescales of different environmental effects.

Several such effects are expected to contribute to the binary evolution in different ranges of the binary separation (Kelley et al. 2017), as illustrated in Fig. 5. The effective timescale for the binary evolution is defined as $t_{eff} \equiv |E|/\dot{E}$, where E is the binary total energy. The effective timescale can be expressed as

$$t_{eff}^{-1} = t_{GW}^{-1} + t_{env}^{-1}, \tag{7}$$

where t_{GW} is the timescale of the GW driven binary evolution and t_{env} is the timescale of the environmental effects. For a circular orbit the binary total energy is given by $E = -m_1 m_2/d$ where d denotes the separation of the BHs, so, in terms of the effective timescale, the separation evolves as

$$\dot{d} = -d\, t_{eff}^{-1}. \tag{8}$$

The dominant contribution to the binary evolution at small separations arises from GW emission, whose timescale is

$$t_{GW} = \frac{5d^4}{1024\eta M^3} \approx \frac{14\,\mathrm{Myr}}{\eta} \left[\frac{M}{10^9 M_\odot}\right]^{-3} \left[\frac{d}{0.1\,\mathrm{pc}}\right]^4, \tag{9}$$

where η and M denote the symmetric mass ratio and the total mass of the binary. However, other energy-loss mechanisms come into play at separations probed by PTAs and other observations.

At kiloparsec separations, dynamical friction drives the binary evolution (Kelley et al. 2017). Its timescale is (Binney & Tremaine 2008)

$$t_{dyn} \simeq \frac{20\,\mathrm{Myr}}{\ln\Lambda} \frac{\sigma}{200\,\mathrm{km/s}} \left[\frac{M}{10^9 M_\odot}\right]^{-1} \left[\frac{d}{\mathrm{kpc}}\right]^2, \tag{10}$$

where $\ln\Lambda$ is the Coulomb logarithm and σ is the velocity dispersion of stars.

Dynamical friction becomes inefficient at shorter distances when the binary binding energy becomes larger than the kinetic energy of the stars whose orbits pass close to the binary. The binary evolution is then driven by loss-cone scattering, whose timescale can be estimated as (Quinlan 1996)

$$t_{sh} \simeq \frac{\sigma}{H\rho d} \approx 300\,\mathrm{Myr}\frac{\sigma}{200\,\mathrm{km/s}} \left[\frac{\rho}{10 M_\odot/\mathrm{pc}^3}\right]^{-1} \left[\frac{d}{\mathrm{pc}}\right]^{-1}, \tag{11}$$

where $H \approx 15-20$ is the dimensionless hardening rate, and ρ and σ denote the mass density and the velocity dispersion of stars in the neighbourhood of the binary.

At intermediate scales between the GW-driven and loss-cone scattering-driven phases, we used the parametrisation Eq. (6) that is given in terms of the binary separation as

$$t_{int} = t_{GW} \left[\frac{d}{d_{ref}}\right]^{-\frac{3\alpha}{2}}. \tag{12}$$

The reference separation d_{ref} is related to the reference frequency used in Eq. (6) by $d_{ref}^3 = 8M/(\pi f_{ref}(\mathcal{M}/10^9 M_\odot)^{-\beta})^2$. We estimated the total timescale of the environmental effects as $t_{env} = (t_{int}^{-1} + t_{sh}^{-1})^{-1} + t_{dyn}$.

Effects that potentially determine the dynamics of binaries at intermediate scales include gas infall and viscous drag. The timescale of the binary evolution by gas infall can be estimated from the accretion timescale (Begelman et al. 1980), $t_{gas} \simeq m_1/\dot{m}_1$, where m_1 is the mass of the heavier BH, and the timescale of viscous drag has been estimated in Ivanov et al. (1999) to scale as, $t_{vd} \propto \sqrt{d}/M^{1/14}$. These correspond, respectively, to $(\alpha, \beta) = (8/3, 5/8)$ and $(\alpha, \beta) \approx (7/3, 3/4)$, while a transition directly from the loss-cone scattering to GW-driven evolution would correspond to $(\alpha, \beta) = (10/3, 2/5)$. For the PTA fit shown in Fig. 3 we fixed $\beta = 5/8$ after checking that the fit is not very sensitive to the value of β. We also find that the above values of $\alpha = 8/3$ and $\alpha = 10/3$ provide good fits to the data.

Given the estimates for the effective timescales, we can estimate the expected number of dual and single AGNs detected by JWST. As already noted, we assumed that AGN activity is triggered only during major mergers where the stellar mass of the heavier galaxy is not more than three times the stellar mass of the lighter one. Single AGNs are generated in major mergers in which one of the galaxies did not contain an SMBH. Moreover, dual AGNs are seen as single AGNs if one of the AGNs is not luminous enough or if their angular separation is too small to be resolved.

The expected number of detectable dual AGNs is

$$\bar{N}_{\text{AGN,2}} = f_{\text{sky}} \int d\lambda_2 \, p_{\text{det}}(m_1, z) p_{\text{det}}(m_2, z) p_{\text{dual}}(d, z), \tag{13}$$

where f_{sky} is the sky fraction being observed, p_{det} the probability that the BH is luminous enough to be detectable, and p_{dual} is the probability that a dual AGN is resolvable. Counting all mergers where at least one of the galaxies is occupied by an SMBH that becomes active during the merger and is sufficiently luminous to be detected, the expected number of detectable AGNs is

$$\begin{aligned}
\bar{N}_{\text{AGN,tot}} = \; & 2f_{\text{sky}} \int d\lambda_1 \, p_{\text{det}}(m, z) \\
& + 2f_{\text{sky}} \int d\lambda_2 \, p_{\text{det}}(m_2, z) \\
& - f_{\text{sky}} \int d\lambda_2 \, p_{\text{det}}(m_1, z) p_{\text{det}}(m_2, z).
\end{aligned} \tag{14}$$

The merger rate and the evolution of the BH separation enter through

$$d\lambda_1 \equiv \frac{1}{1+z} \frac{dV_c}{dz} \frac{dn^1_{\text{AGN}}}{dm \, dd} dm \, dz \, dd \tag{15}$$

and

$$d\lambda_2 \equiv \frac{1}{1+z} \frac{dV_c}{dz} \frac{dn^2_{\text{AGN}}}{dm_1 \, dm_2 \, dd} dm_1 \, dm_2 \, dz \, dd, \tag{16}$$

where

$$\begin{aligned}
\frac{dn^1_{\text{AGN}}}{dm \, dd} = \; & \frac{1 - \sqrt{p_{\text{BH}}}}{|\dot{d}|} \int dM_{\text{v},1} dM_{\text{v},2} \frac{dR_h}{dM_{\text{v},1} dM_{\text{v},2}} \\
& \times p_{\text{AGN}}(M_{\text{v},1}, M_{\text{v},2}, d) \frac{dp_{\text{occ}}(m|M_*(M_{\text{v},2}, z))}{dm}
\end{aligned} \tag{17}$$

is the number density of AGNs formed in galaxy mergers where one of the galaxies did not contain an SMBH, and

$$\begin{aligned}
\frac{dn^2_{\text{AGN}}}{dm_1 \, dm_2 \, dd} = \; & \frac{1}{|\dot{d}|} \int dM_{\text{v},1} dM_{\text{v},2} \frac{dR_h}{dM_{\text{v},1} dM_{\text{v},2}} \\
& \times p_{\text{AGN}}(M_{\text{v},1}, M_{\text{v},2}, d) \prod_{j=1,2} \frac{dp_{\text{occ}}(m_j|M_*(M_{\text{v},j}, z))}{dm_j}
\end{aligned} \tag{18}$$

is the number density of dual AGNs. The probability that a galaxy with stellar mass M_* is occupied by an SMBH with mass m can be expressed as

$$\frac{dp_{\text{occ}}(m|M_*)}{dm} = \sqrt{p_{\text{BH}}} \frac{dP(m|M_*)}{dm}, \tag{19}$$

where the distribution dP/dm, given by Eq. (1), is normalised to 1, so $\sqrt{p_{\text{BH}}}$ can be interpreted as the SMBH occupation fraction,
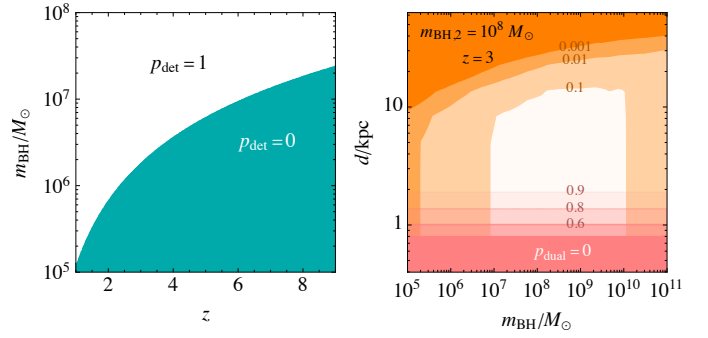
Fig. 6. *Left panel*: detectability probability, p_{det}, arising from the JWST sensitivity threshold on the flux. *Right panel*: dual identification probability, p_{dual}, at $z = 3$ arising from the JWST angular resolution, shown by the red regions. The orange regions show the probability that both galactic nuclei are active for $m_{\text{BH},2} = 10^8 M_\odot$ at $z = 3$.

which we took to be independent of M_*. The prefactor $1 - \sqrt{p_{\text{BH}}}$ in Eq. (17) gives the probability that one of the merging galaxies does not contain an SMBH. In this case, dynamical friction drives this single SMBH towards the centre of the galaxy that resulted from the merger. So, for Eq. (17) we estimated the timescale as $t_{\text{eff}} = t_{\text{dyn}}$. We assumed that the AGN activity starts when the distance between the galaxies becomes smaller than the half-stellar radius of the smaller galaxy. We approximated the half-stellar radius as being 10% of the virial radius of the halo:

$$\begin{aligned}
p_{\text{AGN}}(M_{\text{v},1}, M_{\text{v},2}, d) = \; & \Theta(0.1 \min_j R_{\text{v},j} - d) \\
& \times \Theta(3 \min_j M_{*,j} - \max_j M_{*,j}).
\end{aligned} \tag{20}$$

In the right panel of Fig. 6 the shades of orange show the dual AGN probability $dn^2_{\text{AGN}}/dn_{\text{tot}}$ where dn_{tot} is obtained from Eq. (18) with $p_{\text{AGN}} = 1$.

We estimated the detection probability, p_{det}, by setting a threshold on the flux:

$$p_{\text{det}}(m_{\text{BH}}, z) = \Theta\left(\frac{L}{4\pi D_L^2} - F_{\text{sens}}\right), \tag{21}$$

where D_L denotes the luminosity distance of the system. We approximated the AGN luminosity optimistically as $L \simeq L_{\text{Edd}}/3 \approx 4.3 \times 10^{43} (m_{\text{BH}}/10^6 M_\odot)$ erg/s and chose $F_{\text{sens}} \simeq 10^{-15}$ erg/s/cm^2, which corresponds to the dimmest source found in Perna et al. (2023). The threshold on the flux gives a redshift-dependent lower bound on the BH mass, as shown in the left panel of Fig. 6.

We estimated the probability, p_{dual}, by assuming that the duality of the AGNs is properly identified if its projected angular separation,

$$\theta(d, z) = \frac{d}{D_A(z)} \sqrt{1 - \sin^2 i \sin^2 \phi}, \tag{22}$$

where $D_A(z)$ denotes the angular diameter distance of the system and i and ϕ the inclination and phase of the dual, is larger than twice the angular resolution $\theta_{\text{res}} \approx 0.1''$ of JWST, and we averaged over the angles i and ϕ:

$$p_{\text{dual}}(d, z) = \frac{1}{4\pi} \int di \, d\phi \, \sin i \, \Theta(\theta(d, z) - 2\theta_{\text{res}}). \tag{23}$$

The probability p_{dual} is shown in the right panel of Fig. 6.

We write the timescales of dynamical friction and stellar loss-cone scattering as $t_{dyn} = 20\,\text{Myr}\,A_{dyn}[M/10^9 M_\odot]^{-1}[d/\text{kpc}]^2$ and $t_{sh} = 300\,\text{Myr}\,A_{sh}[d/\text{pc}]^{-1}$ (see Eqs. (10) and (11)), and study how the dual AGN fraction $\bar{N}_{AGN,2}/\bar{N}_{AGN,tot}$ and the total number of AGNs $\bar{N}_{AGN,tot}$ depend on the parameters A_{dyn}, A_{sh}, p_{BH}, f_{ref} and α. We considered the observations reported in Perna et al. (2023) of the GA-NIFS sample of 17 known AGNs at $z > 3$, of which $N_1 = 12$ were identified as single AGNs, $N_2 = 4$ as dual AGNs and one as a triple AGN. Based on these observations, we estimate that the dual fraction $N_2/(N_1 + N_2) = 0.25$. We also considered the JWST results of Kocevski et al. (2024) that cover 587.8 arcmin², corresponding to $f_{sky} \approx 4 \times 10^{-6}$, revealing 341 LRDs in the redshift range $2 < z < 11$. Interpreting these LRDs as AGNs, we used this as an observation of the total AGN number, $N_{tot} = 341$[1].

In the upper panels of Fig. 5 we show the dependences of the numbers of AGNs and fractions of dual AGNs on f_{ref} for $\alpha = 8/3$, corresponding to the gas infall timescale, $A_{sh} = 1$, $A_{dyn} = 0.001$ and $p_{BH} = 0.2$. For these parameters, the lower panel of Fig. 5 illustrates the timescales of the environmental energy loss mechanisms, including GW emission, gas effects, stellar loss-cone scattering and dynamical friction, as functions of the binary separation. The different curves in the lower panel correspond to different values of f_{ref} indicated by the coloured dots in the upper panel. The horizontal bands in the upper panel show the 95% CL bands of the JWST LRD and dual AGN observations calculated assuming the Poisson and binomial distribution, respectively. The value of p_{BH} is fixed so that, neglecting AGNs that may form in mergers where only one of the galaxies contains an SMBH, the expected total AGN number roughly matches the LRD observations for large f_{ref}. We see in the right panel that, in this benchmark case, the dual fraction is within the 95% range of the observed value for $f_{ref} > 6\,\text{nHz}$ if only the AGNs that may form in mergers where both of the galaxies contain an SMBH are considered. These values of p_{BH}, α and f_{ref} are within the 95% CL region of the NANOGrav fit, illustrating that it is possible to obtain a coherent description of the evolution of binary systems that explains both the NANOGrav GW signal and the JWST dual AGN and LRD observations.

Figure 5 shows that the expected number of systems seen as single AGNs would be much larger than observed if f_{ref} is small, because in that case, the binaries would spend a lot of time orbiting with $d \sim 1\,\text{pc}$ separation. This is related to the well-known final-parsec problem. Both the JWST and PTA observations prefer large values of f_{ref} for which the final-parsec problem is avoided. Then, the main contribution to the total number of AGNs comes from the same region as for dual AGNs, $d \gtrsim 1\,\text{kpc}$, where the effective timescale is largest. In this regime, both $\bar{N}_{AGN,2}$ and $\bar{N}_{AGN,tot}$ are proportional to the prefactor A_{dyn} of the dynamical friction timescale. For the benchmark shown in Fig. 5 we have fixed $A_{dyn} = 0.001$ so that the result is consistent with the LRD observations for values of p_{BH} consistent with the NANOGrav observations. Larger values of A_{dyn} or p_{BH} than those used in the above benchmark case would increase the

total AGN number while the dual fraction would remain roughly the same. Larger values of A_{sh} would instead decrease the dual fraction and increase the total AGN number.

5. Conclusions

We have presented a global analysis of the available data on high-z SMBHs detected with JWST (Übler et al. 2023; Larson et al. 2023; Harikane et al. 2023; Bogdan et al. 2024; Ding et al. 2023; Maiolino et al. 2024a; Yue et al. 2023) and in previous observations from I21, in conjunction with results from previous samples of SMBHs in low-z AGNs and IGs from RV15 and measurements of the SGWB background by NANOGrav (Agazie et al. 2023c), which are compatible with data from other PTAs (Agazie et al. 2023a). The JWST data are compatible with previous high-z SMBH data, and both are compatible with a sample of low-z IGs but incompatible with a low-z AGN sample (see Fig. 2 and Table 1). Moreover, we have shown that our global fit to the high-z SMBH and local IG data also fits the NANOGrav data on the nanohertz SGWB, whereas alternative interpretations of the high-z SMBH data that evoke an evolution of the M_{BH}/M_* relationship evolution with z (Pacucci & Loeb 2024) or selection biases and measurement uncertainties (Li et al. 2024) provide only poor fits to the SGWB data, as seen in Fig. 4.

We have also shown how the dual AGN fraction reported in Perna et al. (2023) can be accommodated within our global analysis. Our interpretation of the data on the dual AGN fraction (Perna et al. 2023) is related to the need for environmental effects to accommodate the spectral shape of the NANOGrav data (see Fig. 3 and the upper panel of Fig. 4 as well as Agazie et al. 2023b and Ellis et al. 2024a). Figure 5 shows that an extrapolation of these environmental effects to larger scales can accommodate the fraction of dual AGNs detected with JWST, can be seen in Fig. 5. Moreover, we also find that our global analysis is compatible with the abundance of LRDs in JWST observations reported recently (Matthee et al. 2024; Kocevski et al. 2024).

Observations of high-z SMBHs are still in their infancy, with much more data to come from JWST and other sources. Likewise, PTA observations of the nanohertz SGWB will develop rapidly in the near future. Doubtless, both sets of data will provide new puzzles. The central point of our paper is that these datasets should be considered together and that these puzzles may share common features. The resolution of these puzzles may also require inputs from other observational programmes, such as measurements of GWs at higher frequencies, interpreted within an integrated approach.

---

[1] The origin of the LRDs is debated because they are not detected in X-rays (Ananna et al. 2024). However, the non-detection of X-rays may be caused by high obscuration or intrinsic X-ray weakness of the AGNs (Maiolino et al. 2024b). Our objective is to determine whether the surprisingly high number density of LRDs can be accounted for under the assumption that they are AGNs whose nuclear activity is triggered during mergers. Determining, whether this mechanism could be responsible for triggering local AGN activity, exceeds the scope of this letter and is left for future work.

References

Agazie, G., Antoniadis, J., Anumarlapudi, A., et al. 2023a, ApJ, submitted [arXiv:2309.00693]
Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023b, ApJ, 952, L37
Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023c, ApJ, 951, L8
Aird, J., Coil, A. L., & Georgakakis, A. 2018, MNRAS, 474, 1225
Akins, H. B., Casey, C. M., Allen, N., et al. 2023, ApJ, 956, 61
Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017, arXiv e-prints [arXiv:1702.00786]
Ananna, T. T., Bogdán, A., Kovács, O. E., Natarajan, P., & Hickox, R. C. 2024, ApJ, submitted [arXiv:2404.19010]

Antoniadis, J., Arumugam, P., Arumugam, S., et al. 2024, A&A, 685, A94
Badurina, L., Bentine, E., Blas, D., et al. 2020, JCAP, 05, 011
Barro, G., Pérez-González, P. G., Kocevski, D. D., et al. 2024, ApJ, 963, 128
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307
Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton: Princeton University Press)
Bogdan, A., Goulding, A. D., Natarajan, P., et al. 2024, Nat. Astron., 8, 126
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Callegari, S., Kazantzidis, S., Mayer, L., et al. 2011, ApJ, 729, 85
Capelo, P. R., Volonteri, M., Dotti, M., et al. 2015, MNRAS, 447, 2123
Comerford, J. M., Pooley, D., Barrows, R. S., et al. 2015, ApJ, 806, 219
Ding, X., Onoue, M., Silverman, J. D., et al. 2023, Nature, 621, 51
Donley, J. L., Kartaltepe, J., Kocevski, D., et al. 2018, ApJ, 853, 63
Ellis, J., Fairbairn, M., Hütsi, G., et al. 2023, A&A, 676, A38
Ellis, J., Fairbairn, M., Hütsi, G., et al. 2024a, Phys. Rev. D, 109, L021302
Ellis, J., Fairbairn, M., Urrutia, J., & Vaskonen, V. 2024b, ApJ, 964, 11
Ellis, J., Fairbairn, M., Franciolini, G., et al. 2024c, Phys. Rev. D, 109, 023522
El-Neaj, Y. A., Alpigiani, C., Amairi-Pyka, S., et al. 2020, EPJ Quant. Technol., 7, 6
Georgakakis, A., Aird, J., Schulze, A., et al. 2017, MNRAS, 471, 1976
Girelli, G., Pozzetti, L., Bolzonella, M., et al. 2020, A&A, 634, A135
Harikane, Y., Zhang, Y., Nakajima, K., et al. 2023, ApJ, 959, 39
Ivanov, P. B., Papaloizou, J. C. B., & Polnarev, A. G. 1999, MNRAS, 307, 79
Izumi, T., Matsuoka, Y., Fujimoto, S., et al. 2021, ApJ, 914, 36
Kawamura, S., Ando, M., Seto, N., et al. 2021, Prog. Theor. Exp. Phys., 2021, 05A105
Kelley, L. Z., Blecha, L., & Hernquist, L. 2017, MNRAS, 464, 3131
Kocevski, D. D., Finkelstein, S. L., Barro, G., et al. 2024, arXiv e-prints [arXiv:2404.03576]
Koss, M., Mushotzky, R., Treister, E., et al. 2012, ApJ, 746, L22
Labbé, I., Greene, J. E., Bezanson, R., et al. 2023a, arXiv e-prints [arXiv:2306.07320]
Labbé, I., van Dokkum, P., Nelson, E., et al. 2023b, Nature, 616, 266
Lacey, C. G., & Cole, S. 1993, MNRAS, 262, 627
Larson, R. L., Finkelstein, S. L., Kocevski, D. D., et al. 2023, ApJ, 953, L29
Li, J., Silverman, J. D., Shen, Y., et al. 2024, arXiv e-prints [arXiv:2403.00074]
Liu, X., Guo, H., Shen, Y., Greene, J. E., & Strauss, M. A. 2018, ApJ, 862, 29
Maiolino, R., Scholtz, J., Curtis-Lake, E., et al. 2024a, A&A, 691, A145
Maiolino, R., Risaliti, G., Signorini, M., et al. 2024b, arXiv e-prints [arXiv:2405.00504]
Matthee, J., Naidu, R. P., Brammer, G., et al. 2024, ApJ, 963, 129
Mićić, M., Wells, B. N., Holmes, O. J., & Irwin, J. A. 2023, arXiv e-prints [arXiv:2310.09945]
Pacucci, F., & Loeb, A. 2024, arXiv e-prints [arXiv:2401.04159]
Pacucci, F., Nguyen, B., Carniani, S., Maiolino, R., & Fan, X. 2023, arXiv e-prints [arXiv:2308.12331]
Padmanabhan, H., & Loeb, A. 2024, A&A, 684, L15
Perna, M., Arribas, S., Lamperti, I., et al. 2023, arXiv e-prints [arXiv:2310.03067]
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Quinlan, G. D. 1996, New Astron., 1, 35
Reardon, D. J., Zic, A., Shannon, R. M., et al. 2023, ApJ, 951, L6
Reines, A. E., & Volonteri, M. 2015, ApJ, 813, 82
Ruan, W.-H., Guo, Z.-K., Cai, R.-G., & Zhang, Y.-Z. 2020, Int. J. Mod. Phys. A, 35, 2050075
Scudder, J. M., Ellison, S. L., Torrey, P., Patton, D. R., & Mendel, J. T. 2012, MNRAS, 426, 549
Treister, E., Schawinski, K., Urry, C. M., & Simmons, B. D. 2012, ApJ, 758, L39
Übler, H., Maiolino, R., Curtis-Lake, E., et al. 2023, A&A, 677, A145
Wang, H.-T., Jiang, Z., Sesana, A., et al. 2019, Phys. Rev. D, 100, 043003
Xu, H., Chen, S., Guo, Y., et al. 2023, arXiv e-prints [arXiv:2306.16216]
Yue, M., Eilers, A.-C., Simcoe, R. A., et al. 2023, arXiv e-prints [arXiv:2309.04614]

Appendix A: Compilation of JWST data and corner plots

In Fig. A.1 we present corner plots that exhibit the posteriors for the parameters a, b and σ that were introduced in (1) to characterise the BH mass-stellar mass relation for fits to various datasets discussed in the main text. These include the JWST high-z SMBH data compiled in Fig. A.2 (upper left), the previous high-z SMBH data compiled in I21 (upper right), the RV15 data on local AGNs (lower left), and the RV15 data on dynamically measured SMBH (lower right). The latter two fits agree with the fits found by RV15.
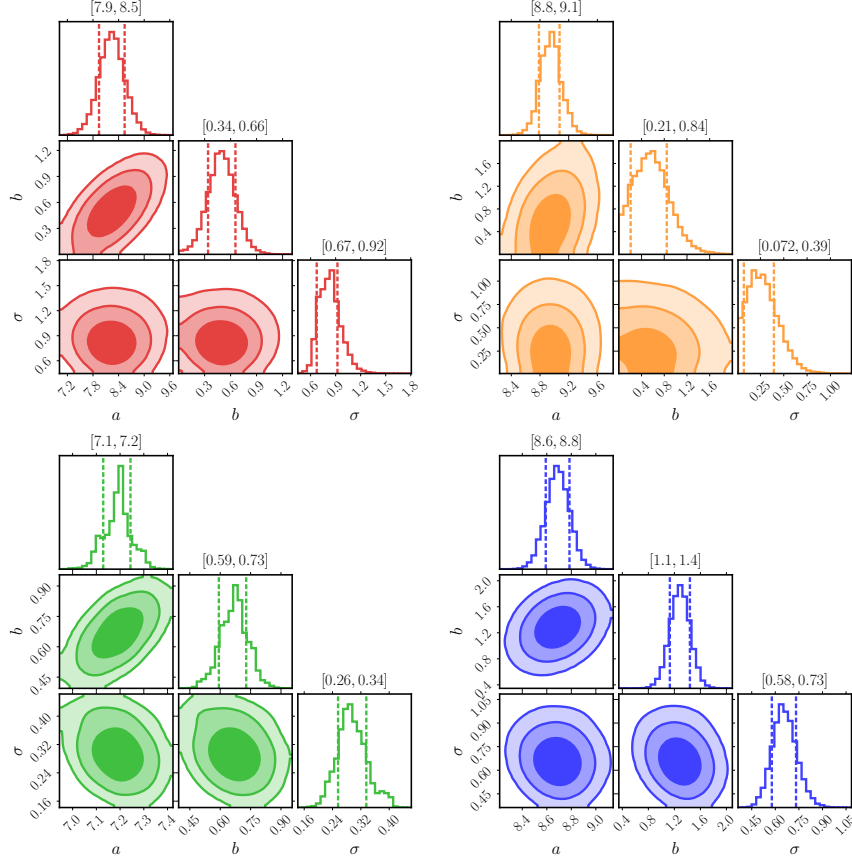

Fig. A.1. Posteriors for fits to the JWST high-z SMBH data compiled in Fig. A.2 *(upper left)*, previous high-z SMBH data compiled in I21 *(upper right)*, the RV15 data on local AGNs *(lower left),* and the RV15 data on the dynamically measured SMBH *(lower right)*.
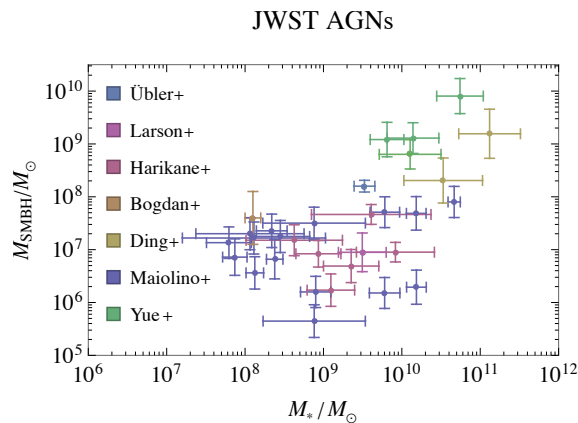

Fig. A.2. Scatter plot of the JWST data analysed in this paper. The following are the papers listed in the legend: Übler+ (Übler et al. 2023), Larson+ (Larson et al. 2023), Harikane+ (Harikane et al. 2023), Bogdan+ (Bogdan et al. 2024), Ding+ (Ding et al. 2023), Maiolino+ (Maiolino et al. 2024a), and Yue+ (Yue et al. 2023).