Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **541,** 2070–2092 (2025) https://doi.org/10.1093/mnras/staf747
Advance Access publication 2025 May 7

Probing the co-evolution of Supermassive Black Holes and their hosts from scaling relations pairwise residuals: dominance of stellar velocity dispersion and host halo mass

Francesco Shankar [1]★ Mariangela Bernardi,[2] Daniel Roberts [1] Miguel Arana-Catania [3] Tobias Grubenmann,[1,4] Melanie Habouzit,[5,6] Amy Smith,[1] Christopher Marsden,[1] Karthik Mahesh Varadarajan,[1] Alba Vega Alonso Tetilla [1] Daniel Anglés-Alcázar [7] Lumen Boco,[8] Duncan Farrah [9,10] Hao Fu [1,11] Henryk Haniewicz,[1] Andrea Lapi [12] Christopher C. Lovell [13] Nicola Menci,[14] Meredith Powell [15] and Federica Ricci [16]

[1]*School of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK*
[2]*Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd St, Philadelphia, PA 19104, USA*
[3]*Digital Scholarship at Oxford, University of Oxford, Oxford OX1 3BG, UK*
[4]*School of Computing, Engineering and the Built Environment, Edinburgh Napier University, Edinburgh EH14 1DJ, UK*
[5]*Department of Astronomy, University of Geneva, Chemin d'Ecogia, CH-1290 Versoix, Switzerland*
[6]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[7]*Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA*
[8]*Institut für Theoretische Astrophysik, Zentrum für Astronomie, Universität Heidelberg, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany*
[9]*Department of Physics and Astronomy, University of Hawai'i at Mānoa, 2505 Correa Road, Honolulu, HI 96822, USA*
[10]*Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Dr., Honolulu, HI 96822, USA*
[11]*Center for Astronomy and Astrophysics and Department of Physics, Fudan University, Shanghai 200438, China*
[12]*SISSA, Via Bonomea 265, I-34136 Trieste, Italy*
[13]*Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Road, Portsmouth PO1 3FX, UK*
[14]*INAF – Osservatorio Astronomico di Roma, via Frascati 33, I-00078 Monte Porzio, Italy*
[15]*Leibniz-Institut fur Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[16]*Dipartimento di Matematica e Fisica, Università Roma Tre, via della Vasca Navale 84, I-00146 Roma, Italy*

Accepted 2025 April 29. Received 2025 March 25; in original form 2024 December 30

ABSTRACT

The correlations between supermassive black holes (SMBHs) and their host galaxies still defy our understanding from both the observational and theoretical perspectives. Here, we perform pairwise residual analysis on the latest sample of local inactive galaxies with a uniform calibration of their photometric properties and with dynamically measured masses of their central SMBHs. The residuals reveal that stellar velocity dispersion σ and, possibly host dark matter halo mass $M_{\rm halo}$, appear as the galactic properties most correlated with SMBH mass, with a secondary (weaker) correlation with spheroidal (bulge) mass, as also corroborated by additional machine learning tests. These findings may favour energetic/kinetic feedback from active galactic nuclei (AGNs) as the main driver in shaping SMBH scaling relations. Two state-of-the-art hydrodynamic simulations, inclusive of kinetic AGN feedback, are able to broadly capture the mean trends observed in the residuals, although they tend to either favour $M_{\rm sph}$ as the most fundamental property, or generate too flat residuals. Increasing AGN feedback kinetic output does not improve the comparison with the data. In the Appendix, we also show that the galaxies with dynamically measured SMBHs are biased high in σ at fixed luminosity with respect to the full sample of local galaxies, proving that this bias is not a by-product of stellar mass discrepancies. Overall, our results suggest that probing the SMBH–galaxy scaling relations in terms of total stellar mass alone may induce biases, and that either current data sets are incomplete, and/or that more insightful modelling is required to fully reproduce observations.

Key words: black hole physics – galaxies: fundamental parameters – galaxies: nuclei – (galaxies:) quasars: supermassive black holes – galaxies: structure.

★ E-mail: f.shankar@soton.ac.uk

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

1 INTRODUCTION

Supermassive black holes (SMBHs) appear to be ubiquitous in the cores of local galaxies measured with sufficient resolution. Their masses $M_{\rm bh}$ correlate with their host galaxy physical properties such as stellar velocity dispersion σ and bulge $M_{\rm sph}$ or total galaxy stellar mass $M_{\rm gal}$ (e.g. Ferrarese & Ford 2005; Beifiori et al. 2012; Kormendy & Ho 2013; McConnell & Ma 2013; Graham 2016; Saglia et al. 2016). The very existence of these correlations suggests a degree of co-evolution between SMBHs and their hosts. A number of physical processes such as gas accretion, cold flows, fly-bys, mergers, or secular instabilities may have all contributed to this co-evolution by promoting star formation and stellar mass growth in the host galaxies while triggering gas accretion and SMBH mergers on to the central SMBHs (e.g. Granato et al. 2004; Cattaneo et al. 2006; Shankar et al. 2012; Menci et al. 2014; Fontanot et al. 2020), although SMBHs may have also preceded the formation of their galactic hosts, at least at redshifts $z \gtrsim 5$ (e.g. Hu et al. 2022; Ding et al. 2023; Kokorev et al. 2023; Maiolino et al. 2024; Matthee et al. 2024; Pacucci et al. 2023; Bogdán et al. 2024; Greene et al. 2024; Inayoshi & Ichikawa 2024; Li et al. 2025).

Theoretical models suggest that SMBHs will eventually follow scaling relations between SMBH mass and their host galaxy properties, although the degree of coeval evolution depends on the type of physical processes regulating the mass growth of the two systems (e.g. Somerville & Davé 2015; Byrne et al. 2023). Accreting SMBHs shining as active galactic nuclei (AGNs) could launch powerful winds (e.g. Cano-Díaz et al. 2012; Farrah et al. 2012; Cicone et al. 2014; Carniani et al. 2016; Fiore et al. 2017; Musiimenta et al. 2023) and/or jets that may potentially halt star formation in the host galaxies via removal/heating of the gas, self-regulating black hole growth, predicting a steep and tight correlation with stellar velocity dispersion $M_{\rm bh} \propto \sigma^{\alpha}$, with $\alpha \sim 3$–5 (e.g. Silk & Rees 1998; Cavaliere & Vittorini 2000; King 2003; Granato et al. 2004; Di Matteo, Springel & Hernquist 2005; Fabian 2012; Sijacki et al. 2015; Menci et al. 2023), or even a correlation with the potential well of the host galaxy of the type $M_{\rm bh} \propto M_{\rm gal}\sigma^2$ (e.g. Hopkins et al. 2007a), although strong episodes of AGN feedback are not necessarily a strict pre-requisite to yield tight scaling relations between SMBHs and their hosts (e.g. Granato et al. 2004; Anglés-Alcázar, Özel & Davé 2013). SMBH mergers could have also played a significant role in shaping SMBH–galaxy scaling relations. Repeated galaxy and SMBH mergers could induce a linear relation preferentially between SMBH mass $M_{\rm bh}$ and host galaxy stellar mass $M_{\rm gal}$ or $M_{\rm sph}$ (e.g. Jahnke & Macciò 2011a), modulating their scatter and redshift evolution (e.g. Robertson et al. 2006; Hirschmann et al. 2010), or generating breaks in particular in the $M_{\rm bh}$–$M_{\rm gal}$ relation due to the impact of dry mergers, or in general mergers being more frequent at high masses and/or in specific galaxy subsamples (e.g. Graham 2023a, b, and references therein). The combined and different impact of supernova and AGN feedback, respectively, dominating below and above the characteristic mass of $M_{\rm gal} \sim 3 \times 10^{10}\,{\rm M}_{\odot}$ can also generate breaks in the SMBH scaling relations (e.g. Cirasuolo et al. 2005; Shankar et al. 2006; Fontanot, Monaco & Shankar 2015), albeit mounting evidence from observational and theoretical works points to a non-negligible role of AGN activity even in low-mass galaxies (e.g. Penny et al. 2018; Arjona-Galvez, Di Cintio & Grand 2024; Bichang'a et al. 2024; Mezcua & Domínguez Sánchez 2024).

Despite intense theoretical and observational work undertaken in recent decades, the origin, shape, and evolution of the SMBH–galaxy scaling relations remain largely poorly understood. Current calibrations of these relations are usually based on samples of inactive SMBHs with robust dynamical mass measurements limited to around a hundred sources (e.g. Saglia et al. 2016; Sahu, Graham & Davis 2019a), orders of magnitude smaller than the samples adopted to study the scaling relations controlling the structural and dynamical properties of galaxies (e.g. Bernardi et al. 2011a, b; Marsden et al. 2022; Figueira et al. 2024). The calibrations of the SMBH scaling relations based on AGN sources are extracted from significantly larger samples than those characterizing dynamically measured inactive SMBHs, but still provide different results, sometimes showing significant systematic offsets with respect to the local relations (e.g. Reines & Volonteri 2015a; Shankar et al. 2019; Farrah et al. 2023). The pairwise residual analysis has been put forward by a series of seminal papers as a powerful tool to dissect the most fundamental relations between SMBH mass and their host galaxy properties, adding critical insight into the physical processes responsible for generating these scaling relations. The correlations between residuals are, in fact, an efficient way of determining whether one variable Y is directly dependent on another variable X or whether the dependence originates from a third variable Z. If Y depends exclusively on X, then the residuals of the correlations with X should be uncorrelated. The residual analysis is characterised by at least three key features that make this an ideal method to study scaling relations, as demonstrated by extensive Monte Carlo simulations (e.g. Bernardi et al. 2007; Hopkins et al. 2007a; Shankar et al. 2016; Shankar, Bernardi & Sheth 2017; Marsden et al. 2020): (1) it clearly identifies the dependence of SMBH mass on a single galactic variable, while fixing another one, thus avoiding overinterpreting the dependence of SMBH on a variable in a direct relation; (2) it is as statistically robust as a direct scaling relation, providing equivalent, if not tighter, constraints on the slopes; (3) it is less affected by potential biases that could distort the normalization of the direct scaling relations. Residuals are thus ideal for unveiling the strength of any underlying dependence of SMBH mass on, e.g. stellar velocity dispersion or effective radius at fixed total/bulge stellar mass, therefore setting stringent constraints on the most fundamental galactic property linked to SMBH. Alternatively, it may unveil the existence of any SMBH 'fundamental plane', where the SMBH mass is linked to two (or more) galactic properties in equal measure (e.g. Hopkins et al. 2007b; Saglia et al. 2016).

Previous work based on residual analysis applied to the local SMBH dynamical samples available at the time, indicated a non-negligible correlation of SMBH mass on both stellar velocity dispersion and galactic stellar or even dynamical mass, possibly supporting the view of an underlying dependence of SMBH on the galactic potential well of the type $M_{\rm bh} \propto M_{\rm gal}\sigma^2$ (e.g. Hopkins et al. 2007a; Iannella, Greco & Feoli 2021). Other analyses carried out on more recent and/or more uniform SMBH galaxy samples with detailed Monte Carlo quantification of the statistical uncertainties, still showed a hint for a possible SMBH fundamental plane relation, but with a significantly stronger and steeper dependence on stellar velocity dispersion and a weaker and flatter correlation with total host galaxy stellar mass (e.g. Bernardi et al. 2007; Shankar et al. 2016; Marsden et al. 2020). Residual analysis performed by Shankar et al. (2019) on uniform local samples of type 1 AGN from van den Bosch et al. (2015) and Ho & Kim (2014) confirmed these findings.

Some preliminary tests performed on theoretical models inclusive of AGN feedback struggled to reproduce the strong residual with stellar velocity dispersion at fixed galaxy stellar mass (Barausse et al. 2017; Menci et al. 2023). Barausse et al. (2017) showed that their semi-analytic model, once tuned to reproduce the (mean) correlation of black hole mass with velocity dispersion, was not able to simultaneously account for the correlation with stellar mass, in line with what was also found by Sijacki et al. (2015) in the

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

Illustris simulation, a possible signature of biases in the SMBH–galaxy scaling relations. In addition, the residual analyses performed by Barausse et al. (2017) on their semi-analytic model and the Horizon–AGN hydrodynamic simulation (Dubois et al. 2016), both showed a weak correlation with galaxy stellar mass at fixed stellar velocity dispersion. Menci et al. (2023), more recently, added a new physical treatment of AGN-driven winds in their semi-analytic model of galaxy formation, with outflow expansion and mass outflow rates depending on AGN luminosity, halo circular velocity, and gas fractions. From pairwise residual analysis they still found that the model predicts equally strong correlations of SMBH mass with stellar mass, stellar velocity dispersion, and host halo mass, at variance with some of the empirical results discussed above.

The purpose of this paper is twofold. On the one hand, we will be revisiting the pairwise residuals of SMBH scaling relations, making use of the latest SMBH–galaxy sample from Sahu et al. (2019a) with uniform measurements of the host galaxy and bulge stellar masses, along with other photometric properties. On the other hand, we will also compare with two state-of-the-art hydrodynamic simulations TNG and Simba with different implementations of SMBH growth and AGN feedback recipes and efficiencies. We will show that the residuals on the new data clearly point to stellar velocity dispersion, and even possibly to host halo mass, as the dominant host properties linked to SMBH, in line with some of the previous claims. For completeness, we will also compare the residuals outputs with the predictions from a variety of machine learning regression algorithms that largely confirm the results of the residuals, although not providing the same level of information. In Appendix A, we will also revisit and confirm the existence of a bias between the local SMBH galactic sample and the larger comparison sample of local galaxies.

In what follows, wherever relevant, we will adopt a reference cosmology with $h = 0.7$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. The theoretical models may adopt slightly different choices of cosmological parameters, but these differences are small and do not affect any of the results presented in the next sections. Differences in stellar mass estimates between models and data may be present, but they do not affect the residuals analysis, as further discussed below.

2 DATA

2.1 Observational data

In this work, we make use of the sample of local galaxies with dynamical mass measurements of their central SMBHs by Sahu et al. (2019c), which in total comprises around 150 galaxies. Of these, following Sahu, Graham & Hon (2023), we retain 73 early-type galaxies (ETGs) and 28 late-type galaxies (LTGs), which have uniformly calibrated 3.6 μm Spitzer photometry and detailed galaxy modelling and decompositions performed by Savorgnan & Graham (2016), Davis, Graham & Cameron (2019), and Sahu et al. (2019b). From these two samples, we further remove four ETGs, NGC1194, NGC1316, NGC5018, and NGC5128, which are classified as mergers by Kormendy & Ho (2013), and two LTGs, NGC4395 and NGC6926, which do not have secure SMBH mass measurements. Our final SMBH sample is reported in Tables C1 and C2. We convert all galaxy AB total absolute magnitudes to stellar masses adopting as magnitude of the sun $M_{3.6\mu m,\odot} = 6.02$ and a constant mass-to-light ratio at 3.6 μm of $(M/M_\odot)/(L/L_\odot) = 0.6$ as in Sahu et al. (2019c). We note that, as accurately described by Forbes et al. (2017), the mass-to-light ratio at 3.6 μm is highly insensitive to metallicity, and only weakly to stellar ages, with a small uncertainty of ±0.1 dex, further supporting our choice of a constant conversion between luminosity and stellar mass. We also stress that the exact value chosen for the (constant) mass-to-light ratio is irrelevant to the residual analysis. To each galaxy we associate a stellar velocity dispersion σ and associated error from the Hyperleda catalogue (Paturel et al. 2003), with values corrected to a common aperture of 0.595 kpc/h. Following Sahu et al. (2019b), we assign a 0.2 mag error to magnitudes (N. Sahu, private communication), which translates into a 0.08 dex error in total galaxy stellar mass when ignoring any error in mass-to-light ratio. In addition, we will also study residuals against stellar spheroidal mass,[1] effective radius R_e, and Sérsic index n, all taken from table A1 in Sahu et al. (2020). As no specific errors have been reported for effective radii and Sérsic indices, we assign to these galactic properties typical uncertainties of 10 per cent, noticing that moderate variations to these errors have minimal impact on the pairwise residuals. In Section 5 we also explore the residuals on the subsample of 41 galaxies from Sahu et al. (2023) with dark matter halo mass measurements from Marasco et al. (2021). We assign to all halo masses an error of 0.24 dex, which is the typical error reported by Marasco et al. (2021) for most of the galaxies in their sample. Some ETGs have quoted uncertainties significantly larger than 0.24 dex, but if included in the residuals, would generate unrealistic values of the Pearson correlation coefficient (larger than unity). In addition, Monte Carlo simulations based on random extractions of the data points, without considering any measurement error but only fitting the raw distribution of points at each iteration, provide very similar results to the full residual analysis inclusive of the assumed statistical uncertainties. We also performed various stability tests by recalibrating the Hyperleda stellar velocity dispersion at different apertures following the prescriptions from both Bernardi et al. (2017) and Cappellari et al. (2006) finding that the residuals are stable against increasing the aperture up to twice the spheroidal effective radius quoted in Sahu et al. (2020).

2.2 Hydrodynamical simulations

We compare the data described in Section 2.1, in terms of both correlations and residuals, with the predictions from state-of-art hydrodynamical simulations that incorporate AGN feedback in two different ways and thus allow to probe the efficacy of these models in setting the scaling relations between SMBHs and their host galaxies. More specifically, we make use of two simulations: TNG50-1[2] (e.g. Pillepich et al. 2018; Springel et al. 2018; Nelson et al. 2019), and Simba M50N512[3] (Davé et al. 2019). Comparison between observed and predicted SMBH scaling relations, sometimes also considering pairwise residuals, have been carried out previously by adopting specific semi-analytic and hydrodynamical models inclusive of both thermal and kinetic AGN feedback modes (e.g. Hopkins et al. 2007a; Barausse et al. 2017; Habouzit et al. 2021a, 2022a, b; Menci et al. 2023), albeit with previous versions of the SMBH data.

TNG50-1 is the smallest, highest resolution simulation of IllustrisTNG with a volume of $(51.7\ \mathrm{cMpc})^3$ containing 2160^3 dark matter particles (of mass $m_{\rm DM} = 4.5 \times 10^5\ M_\odot$) and 2160^3 gas cells (with initial baryon mass of $m_{\rm gas} = 8.5 \times 10^4\ M_\odot$). This cosmological volume is evolved using AREPO (Springel 2010) from $z = 127$ to $z = 0$ with 100 publically available snapshots for $z \leq 20$. The cosmological parameters of TNG50-1 are: $\Omega_M = 0.3089$, $\Omega_b = 0.0486$,

[1] In this work we use the word 'spheroid' as synonym of 'bulge'.
[2] https://www.tng-project.org
[3] http://simba.roe.ac.uk

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

$\Omega_\Lambda = 0.6911$, $h = 0.6774$, $n_s = 0.9667$, and $\sigma_8 = 0.8159$. We will simply refer to this simulation as TNG from here onwards.

Simba M50N512 is the median box size of the Simba simulations with a volume of $(50h^{-1}$ cMpc$)^3$ containing 512^3 dark matter particles (of mass $m_{\rm DM} = 1.2 \times 10^7$ M$_\odot$) and 512^3 gas elements (with initial baryon mass of $m_{\rm gas} = 2.88 \times 10^6$ M$_\odot$). This cosmological volume is evolved using GIZMO (Hopkins 2015) from $z = 100$ to $z = 0$ with 151 publically available snapshots for $z \leq 20$. The cosmological parameters of Simba are $\Omega_M = 0.3$, $\Omega_b = 0.048$, $\Omega_\Lambda = 0.7$, $h = 0.68$, $n_s = 0.97$, and $\sigma_8 = 0.82$. We will simply refer to this simulation as Simba from here onwards. We stress that the slight difference in cosmological parameters adopted in the simulations and the reference data does not alter any of our comparison tests in what follows.

It is relevant to briefly note here the main key features characterizing the accretion on to the central SMBH and AGN feedback in the two reference hydrodynamical simulations (e.g. Habouzit et al. 2021b). In TNG, the accretion on to the SMBH, which follows the Bondi–Hoyle–Lyttleton formalism (but without any ad hoc α-boost as included in other cosmological simulations), is kernel-weighted over neighbouring cells, ensuring that each accretion episode closely reflects the physical conditions of the gas in the central region of the host. The TNG simulation includes both thermal AGN feedback, where energy is deposited in the surroundings of the accreting SMBH, and kinetic AGN feedback, where momentum is injected in the SMBH surroundings with a direction that is chosen at random at each event, thus generating after a few events a nearly isotropic kinetic feedback. The Simba simulation also includes the Bondi–Hoyle–Lyttleton model but only for hot gas above 10^5 K, while below this temperature the accretion is torque-limited, originating from the gas inflow rate driven by gravitational instabilities to the accretion disc of the SMBH, following the formalism developed by Hopkins & Quataert (2011) and Anglés-Alcázar et al. (2015). The AGN outflows in Simba are bipolar along the angular momentum vector of the stellar disc, and particles are ejected randomly from the black hole accretion kernel with a velocity that varies according to the value of the Eddington ratio, mimicking the effects of radiative- and jet-mode AGN winds (see Davé et al. 2019 for full details).

To pin down the putative impact of AGN feedback, we also make use of the Cosmology and Astrophysics with MachinE Learning Simulations (CAMELS),[4] which is a suite of different hydrodynamical simulations (e.g. Villaescusa-Navarro et al. 2021, 2023). The simulations in CAMELS are organized into different *suites*. We use the TNG and Simba suites. The TNG suite is based on the same subgrid physics as the original IllustrisTNG simulation. Likewise, the Simba suite is based on the same subgrid physics as the Simba simulation. All simulations within CAMELS have a volume of $(25$ cMpc$/h)^3$ with 256^3 dark matter particles and 256^3 gas resolution elements. The simulations span a range from redshift $z = 127$ to $z = 0$. The cosmological parameters for the simulations are $\Omega_b = 0.049$, $h = 0.6711$, $n_s = 0.9624$, $w = -1$, $M_\nu = 0.0$ eV, $\Omega_k = 0.0$. Each suite in the CAMELS repository (TNG and Simba in our case) is further split into different *sets*. In this paper, we used the *Extreme* (EX) sets, and in particular we focus on the EX1 set which represents simulation runs with very efficient AGN feedback with respect to the baseline simulations, namely with the AGN wind outflow rate increased by a factor of 100.

For further analysis, we used the CAESAR[5] package to extract the necessary information from all the simulations except for TNG, which uses SubFind (Springel et al. 2001) catalogues. In particular, we extracted the following information: Black Hole mass, stellar mass, stellar velocity dispersion, stellar half-mass radius, central or satellite galaxy flags, host halo mass, and bulge-to-total ratio. In what follows, we will only retain the central galaxies in the simulations with a central black hole and a stellar mass $M_{\rm gal} > 10^{10}$ M$_\odot$, to cover a stellar mass range similar to the data. In the CAESAR catalogue, velocity dispersions are calculated for each particle mass type, gas, stellar, dark matter, black hole, baryonic, and total mass. For each mass type the velocity dispersion in each coordinate axis is computed as the standard deviation of the linear momentum along the chosen axis, divided by the mean mass of the particles

$$\sigma_x = \sqrt{\frac{\sum_{i=0}^{N}(m_i v_{i,x} - \langle m v_x \rangle)^2}{N \langle m \rangle^2}}. \quad (1)$$

From equation (1) we calculate the 3D stellar velocity dispersion, $\sigma_{\rm 3D}$, from which, by making the common assumption that the orbits are isotropic, we derive the 1D stellar velocity dispersion $\sigma_{\rm 1D} = \sigma_{\rm 3D}/\sqrt{3}$. By default, CAESAR calculates the stellar velocity dispersion from all stellar particles associated with a galaxy, and therefore, they are not strictly expressed in the same aperture as in the data. We have checked, however, that restricting the calculation of σ in the simulations to a radial distance from the centre equal to the half-mass radius, or even to the aperture of the Hyperleda data base, has modest effects on σ with variations contained within $\lesssim 0.1$ dex, in the range of interest to this work. Another possible aperture correction could be applied to σ following the full Jeans modelling as detailed in Appendix C of Marsden et al. (2022). However, this treatment would imply convolving with the full stellar profile of the simulated galaxies and may affect some of the residual analysis when considering velocity dispersions. In what follows, we will thus only show results using the default $\sigma_{\rm 1D}$ from CAESAR, which is in line with the methodology also followed by Thomas et al. (2019). For TNG, we extracted the same information from the SubFind catalogues and auxiliary files, as we have done from the CAESAR catalogues for the other simulations. However, due to the velocity dispersion listed in the SubFind catalogue being computed for all particles associated with a galaxy, we recomputed the velocity dispersion for the stellar particles only using equation (1). Finally, we also verified that, in the region of overlap $\sigma \lesssim 100$ km s^{-1}, the stellar velocity dispersions computed from equation (1) are consistent, within 0.1 dex, with those in the iMaNGA mock catalogue (Nanni et al. 2023), inclusive of all relevant observational effects.

3 METHOD

The main aim of this work is to compute the pairwise residuals between black hole mass and a variety of galactic properties from the latest data and models. The pairwise residuals are a technique specifically designed to isolate the underlying dependence of one variable Y from the variable Z while keeping a third variable X 'fixed'. This step is achieved by computing the correlation between the residuals:

$$\Delta(Y|X) \propto \Delta(Z|X), \quad (2)$$

[4] https://camels.readthedocs.io

[5] https://caesar.readthedocs.io

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

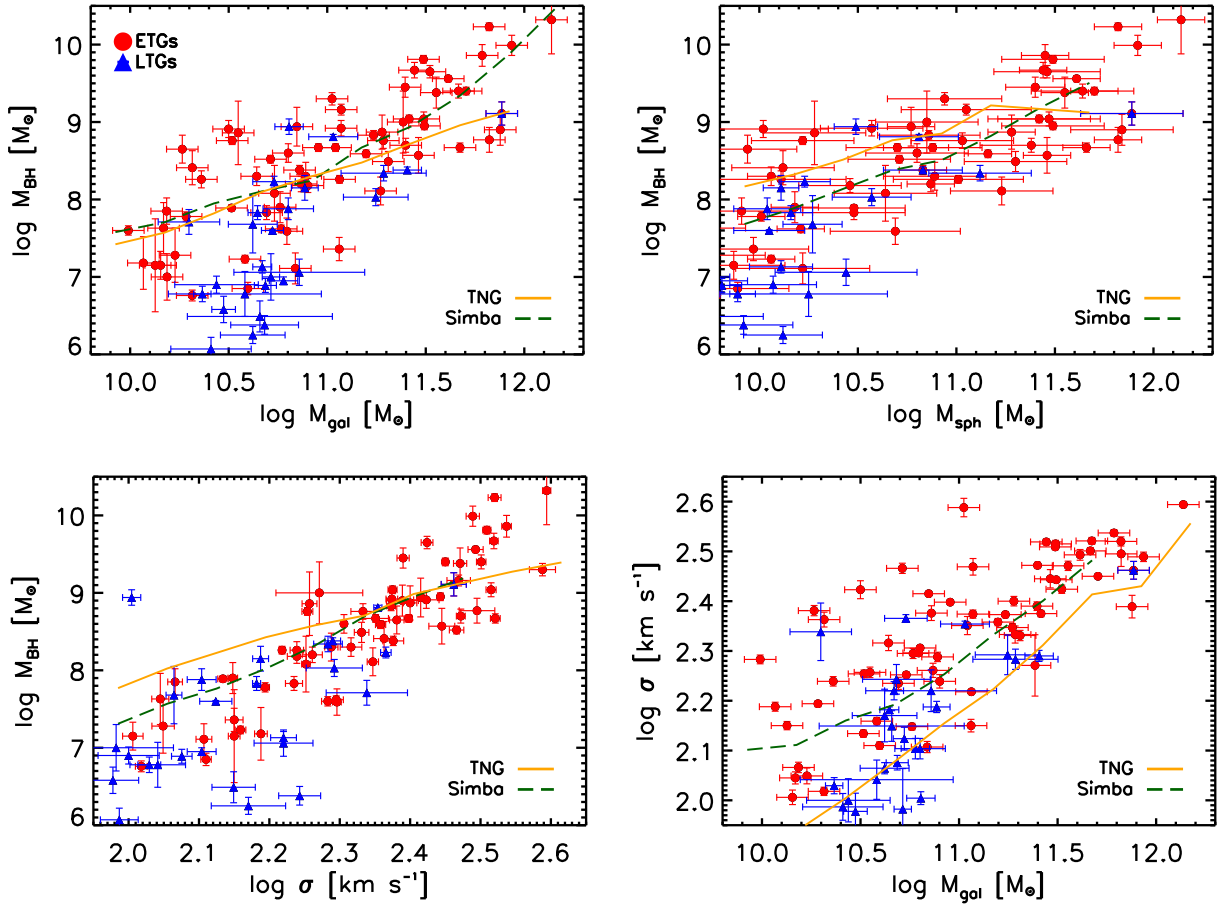

Figure 1. Top left: Scaling relation between SMBH mass $M_{\rm bh}$ and total galaxy stellar mass $M_{\rm gal}$ for local galaxies with dynamically measured masses of their central SMBHs. Top right: Identical to the top left panel but where only the bulge component of LTGs in the data is included. Bottom left: Corresponding scaling relation between SMBH mass $M_{\rm bh}$ and galaxy stellar velocity dispersion σ (see the text for details) for local galaxies. Bottom right: Correlation between σ and galaxy stellar mass $M_{\rm gal}$. Red circles and blue triangles in all panels refer, respectively, to early-type and LTGs. All the data are from Sahu et al. (2019a). Solid, orange and green, dashed lines show the predictions from the TNG and Simba simulations, respectively. The simulations tend to align with the data in the $M_{\rm bh}$–$M_{\rm gal}$ plane but significantly depart from them in the $M_{\rm bh}$–σ one.

where

$$\Delta(Y|X) \equiv \log Y - \langle \log Y | \log X \rangle \quad (3)$$

is the residual computed in the Y variable (at fixed X) from the log–log-linear fit of $Y(X)$ versus X, i.e. $\langle \log Y | \log X \rangle$. Equation (2) calibrates the degree of correlation between Y and Z once the correlations in X of both variables Y and Z are effectively 'factored out'; a perfectly null correlation between the $\Delta(Y|X)$ and $\Delta(Z|X)$ residuals would imply that an apparent correlation between Y and Z is actually a reflection of the underlying dependencies of Y and Z on X. In contrast, a strong correlation between $\Delta(Y|X)$ and $\Delta(Z|X)$ would instead imply a minor role of X in setting the correlation between Y and Z. Full details on the full pairwise residual analysis formalism which we follow in this work, inclusive of the treatment of errors in the variables and the calculation of the Pearson coefficient, are provided in Sheth & Bernardi (2012) and in appendix B of Shankar et al. (2017). For each pair of variables Y and Z, the correlation between their linear residuals (Equation 3) is computed a 100 times, and at each iteration 5 per cent of the objects are removed at random from the original sample. From the complete set of realizations, we then measure the mean slope and its standard deviation. At each iteration, when calculating in the observed data samples the intrinsic slopes and degree of correlations in the residuals, we also take into account the measurement error on each point. Similarly to what we performed with the observational data, for the simulated outputs we follow a stochastic iterative method to compute the residuals, where at each iteration we remove 5 per cent of the sample, and then calculate the mean slope and Pearson coefficient. However, in the latter case, when computing residuals we do not include observational statistical uncertainties, instead, we simply apply a linear fit to the simulated points to define the residual at each iteration. In addition, in what follows, we also compare the results from the pairwise residual analysis with the outputs extracted from Machine Learning (ML) regression algorithms. We provide details on the latter methods directly in Section 4.3.

4 RESULTS

4.1 Scaling relations

We start in Fig. 1 by providing a broad comparison between the scaling relations of local inactive galaxies with dynamically measured SMBHs, from here onwards simply defined as the 'SMBH sample', and the predictions from hydrodynamic simulations. For

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

the latter, we here use the mean relations predicted by the reference Simba and TNG simulations[6] (green, dashed and solid, orange lines, respectively), we will further discuss below the impact of adopting some variations of these simulations. The red circles and blue triangles in all panels refer to ETGs and LTGs, respectively. The top left panel of Fig. 1 shows the relation between SMBH mass as a function of total galaxy stellar mass $M_{\rm gal}$, while in the top right panel we only retain the stellar spheroidal component $M_{\rm sph}$ of both ETGs and LTGs. It is immediately interesting to note that in the data, in terms of total stellar mass, LTGs define a steeper $M_{\rm bh}$–$M_{\rm gal}$ relation than ETGs (top left panel), as already noted by Sahu et al. (2019a), while the correlation with SMBH mass becomes more linear when only the stellar spheroidal component is retained, or at least the steepening is shifted to lower stellar masses (top right panel).

Both simulations align with the ETG data, while the LTGs lie mostly below the simulations' predictions when total stellar mass is considered. It is relevant to note that this comparison is only qualitative as the models have been mostly calibrated on local galaxy stellar mass functions from, e.g. Bernardi et al. (2013), which have SDSS photometries and specific choices of mass-to-light ratios that may differ from the ones derived from the 3.6 μm photometry by Sahu et al. (2023). As discussed in Section 2.2, aperture corrections in the stellar velocity dispersions in the simulations are small and have a minor impact on the results in Fig. 1 as also noted by Marsden et al. (2022). Nevertheless, other factors may affect the comparison with the local scaling relations of SMBHs, in particular in the $M_{\rm bh}$–$M_{\rm gal}$ plane, such as the existence of possible biases, which we extensively discuss in Appendix A.

The left bottom panel of Fig. 1 reports the local SMBH sample in the $M_{\rm bh}$–σ plane, compared with predictions from the Simba and TNG simulations, similarly to the top panels. The TNG simulation tends to produce a slightly flatter relation at higher masses with respect to the distribution of the data, in line with what also found by Li et al. (2020, see their fig. 1), while the Simba simulation better aligns with the data. In the bottom right panel of Fig. 1, we report, for completeness, the relation between σ and $M_{\rm gal}$ in the data and in the two simulations. The Simba simulation again broadly aligns with the data, while the TNG tends to predict $\gtrsim 0.1$ dex systematically lower mean stellar velocity dispersions at fixed stellar mass, as expected given that the simulation is consistent with the $M_{\rm bh}$–$M_{\rm gal}$ relation but predicts somewhat lower velocity dispersions at fixed SMBH mass. We stress that alignment in the σ–$M_{\rm gal}$ plane only ensures internal self-consistency between the model and the SMBH data sample, but not necessarily alignment between the model and the entire local galaxy population if the SMBH sample is affected by selection bias, as suggested in Appendix A. In other words, current data sets on SMBHs with dynamical mass measurements are still too sparse to be used to securely discern among successful theoretical models. Additional spurious numerical effects may also affect the comparison between models and data. For example, the simulated central stellar velocity dispersions may be affected by resolution in some galaxies and/or being artificially inflated by the contribution of dark matter particles (e.g. Schaye et al. 2015). Overall, Fig. 1 proves that comprehensively and simultaneously reproducing the scaling relations of local SMBHs still represents a non-trivial task even for some of the current state-of-the-art SMBH cosmological models.

[6]When computing the predicted mean relations we only retain bins inclusive of at least five galaxies.

4.2 Residual analysis

In the previous Section 4.1, we provided a broad comparison between available local data on galaxies with SMBH mass dynamical mass measurement and the predictions from two state-of-the-art hydrodynamic simulations. We now go deeper into the study of the SMBH scaling relations by analysing the pairwise residuals, following the formalism described in Section 3 and in Shankar et al. (2017). Fig. 2 shows the residual analysis for the SMBH sample, distinguished in ETGs and LTGs as in the previous Figures. The left panels refer to residuals as a function of stellar velocity dispersion σ at fixed galaxy stellar mass $M_{\rm gal}$, while the bottom panels only include ETGs. All panels in Figs 2 and 3 uniformly use stellar velocity dispersion from the Hyperleda data base and thus we label it as σ_H. As discussed in Section 3, each residual plot reports the mean slope and standard deviation (magenta solid and dotted lines), along with its Pearson correlation coefficient r and associated error, which we report in the bottom right corner of each panel. In all panels we also include a direct linear fit to the residuals (black, dotted lines), which are usually very close to the outcome of the iterative method, proving that the outcomes extracted from the residuals are robust against random statistical fluctuations, despite the relatively modest size of the SMBH sample.

From all panels, it is evident that the residual with stellar velocity dispersion is significant, with a correlation coefficient of $r \sim 0.66$–0.75, stronger for the ETG subsample. The residual with galaxy stellar mass at fixed σ is instead less significant, in line with what also found by Bernardi et al. (2007) and Shankar et al. (2016), quantitatively demonstrating the fundamental importance of stellar velocity dispersion in driving the local SMBH–galaxy scaling relations. We repeat the exercise on the analysis of the residuals substituting total with spheroid stellar mass in Fig. 3. In this case we find, when considering both ETGs and LTGs (top panels), tentative evidence for a 'fundamental plane' of SMBHs, with SMBH mass being correlated with both stellar velocity dispersion σ and spheroid stellar mass $M_{\rm sph}$ in roughly equal strength. The residuals point to a relation of the type $M_{\rm bh} \propto \sigma^{2.2} M_{\rm sph}^{0.8}$, which is reminiscent of what predicted by some hydrodynamic models (Hopkins et al. 2007a), as a consequence of AGN feedback coupling SMBH mass with the binding energy of the host. Hopkins et al. (2007b) also found evidence of a fundamental plane-type relation in the local SMBH samples available at the time, although they were mostly referring to total galaxy stellar mass.

The bottom panels of Fig. 3 show that, when considering the bulge stellar mass of only ETGs, the correlation with stellar velocity dispersion becomes even stronger, while the one with bulge mass becomes flatter, suggesting a tilted fundamental plane, as also put forward by Hopkins et al. (2007b). However, in Fig. B1 we report the results of the residuals with respect to σ and $M_{\rm sph}$ extracted from the Saglia et al. (2016) SMBH sample,[7] which confirms a strong and steep dependence of SMBH mass on stellar velocity dispersion, close to $M_{\rm bh} \propto \sigma^4$, but a significantly weaker correlation with spheroid stellar mass, therefore not supporting the existence of a fundamental plane-type relation, as also noted in Shankar et al. (2016). The existence of a fundamental plane for SMBHs is thus not yet unambiguously confirmed, and in any case it is mostly evident when the *spheroid* and not the total stellar mass are considered.

[7]We checked that the residuals from the Saglia et al. (2016) sample still yield similar results even when restricting to the galaxies in common with Sahu et al. (2023) and adopting their SMBH masses and stellar velocity dispersions.

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

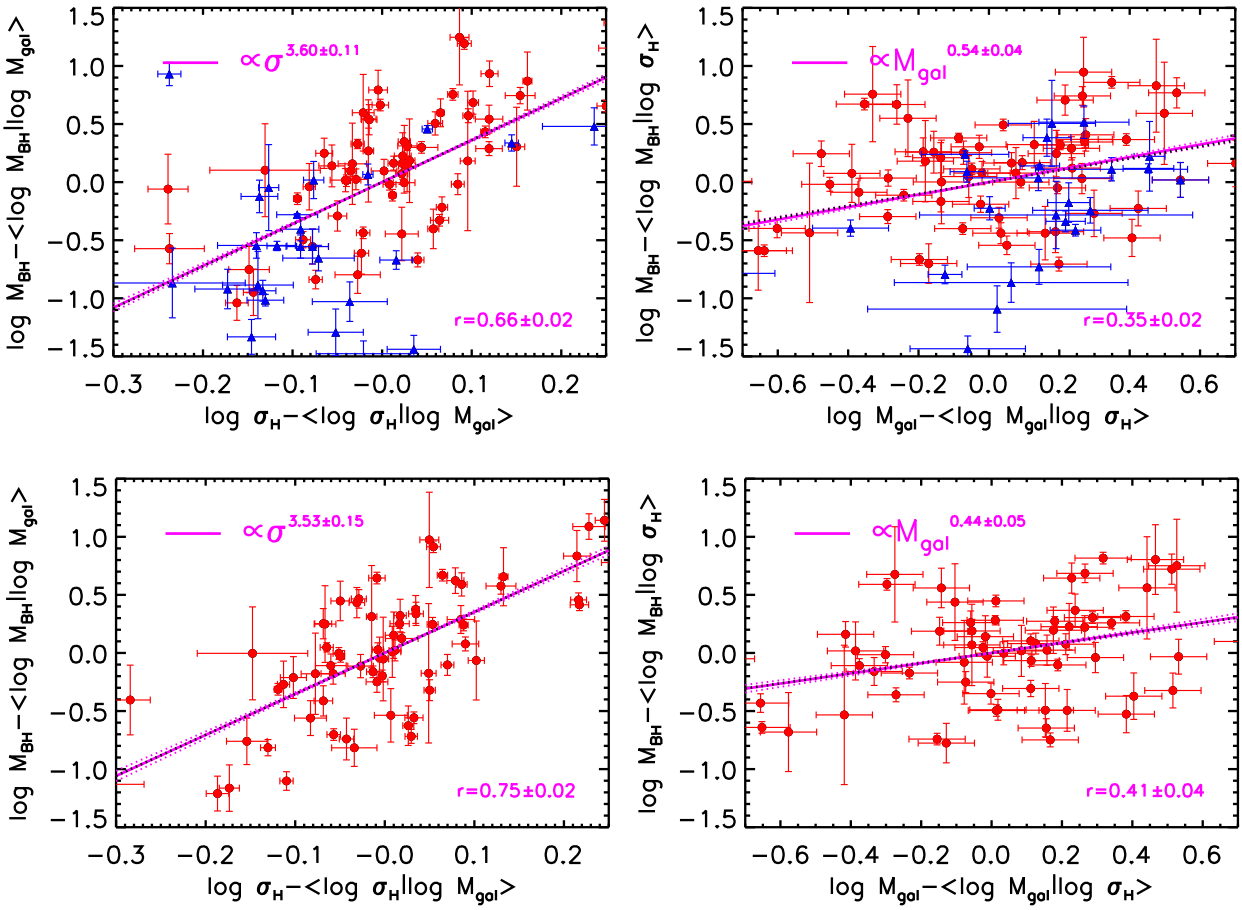

Figure 2. Residual analysis for the local sample of galaxies with dynamically measured SMBH mass from Fig. 1. Left panels: Residuals as a function of stellar velocity dispersion σ at fixed galaxy stellar mass $M_{\rm gal}$. Right panels: Residuals as a function of galaxy stellar mass at fixed stellar velocity dispersion. Top panels use the full galaxy sample, including LTGs, while bottom panels include only ETGs. The residual with stellar velocity dispersion is very significant, while the one with stellar mass is negligible at fixed σ, fully confirming the primary role of σ compared to total galaxy stellar mass.

In Fig. B2, we also show the pairwise residuals applied to the Sahu et al. (2023) sample between SMBH mass and spheroid Sérsic index n (left) and spheroid half-light effective radius R_e (right) at fixed stellar velocity dispersion (top) and spheroidal mass (bottom). When fixing stellar velocity dispersion, mild and weak residual correlations are found with n and R_e, but these completely disappear when calculating the residuals at fixed spheroid mass $M_{\rm sph}$ (bottom panels), further indicating that residual correlations between the SMBH mass and the structural properties of the host galaxy (or at least its spheroidal component) are mostly induced by the underlying correlation between SMBH mass and stellar spheroidal mass, in line with the analysis carried out by Shankar et al. (2017). We note that the anticorrelation with R_e in the lower, right panel is only apparently strong possibly induced by some random error, as the data points are uniformly scattered with negligible correlation, as also indicated by the direct fit to the residuals (black, dotted line). In summary, the residual analysis applied to the latest local SMBH data sample with dynamical mass measurements continues to support the dominance of stellar velocity dispersion over total or even spheroidal stellar mass as a more fundamental galactic property related to SMBH mass, further supporting the view that σ is a key property in the evolution of galaxies (e.g. Bernardi et al. 2011a, b; Bluck et al. 2020).

Fig. 4 shows the comparison between the mean residuals derived from the SMBH sample (magenta, long–dashed lines; top panels of Fig. 2) and those predicted from the reference simulations Simba and TNG (see Section 2.2). For the simulations, stellar velocity dispersions, which we label in the plots as σ_e to differentiate them from the ones from Hyperleda, are calculated using equation (1) over all the stellar particles associated to the host galaxy. The left panels are the residuals with stellar velocity dispersion σ at fixed galaxy stellar mass $M_{\rm gal}$, and the right panels plot the residuals with $M_{\rm gal}$ at fixed σ. We find that the Simba simulation provides both residuals as a function of σ and $M_{\rm gal}$ very close to the data, in particular the one with σ, although the significance of these correlations may not be as strong. The TNG simulation also provides a significant correlation with σ in reasonable agreement with the data. Both simulations interestingly predict a negligible correlation with stellar mass at fixed stellar velocity dispersion, even weaker than in the data. A stronger correlation between SMBH mass and stellar velocity dispersion may be induced by the direct dynamical coupling generated by the kinetic AGN feedback incorporated in both simulations, which we briefly discussed in Section 2.2, although this cannot be confirmed at this level of the analysis.

Fig. 5 shows the same residuals as predicted by the simulations but with total stellar mass replaced by spheroidal stellar mass $M_{\rm sph}$. In Simba (top panels), the dependence of residuals with $M_{\rm sph}$ at fixed σ are more significant than in the case of total stellar mass, while the dependence of SMBH mass with σ at fixed $M_{\rm sph}$ is noticeably reduced, with a low Pearson correlation coefficient of just $r \sim 0.2$. The TNG simulation continues to predict a significant correlation

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

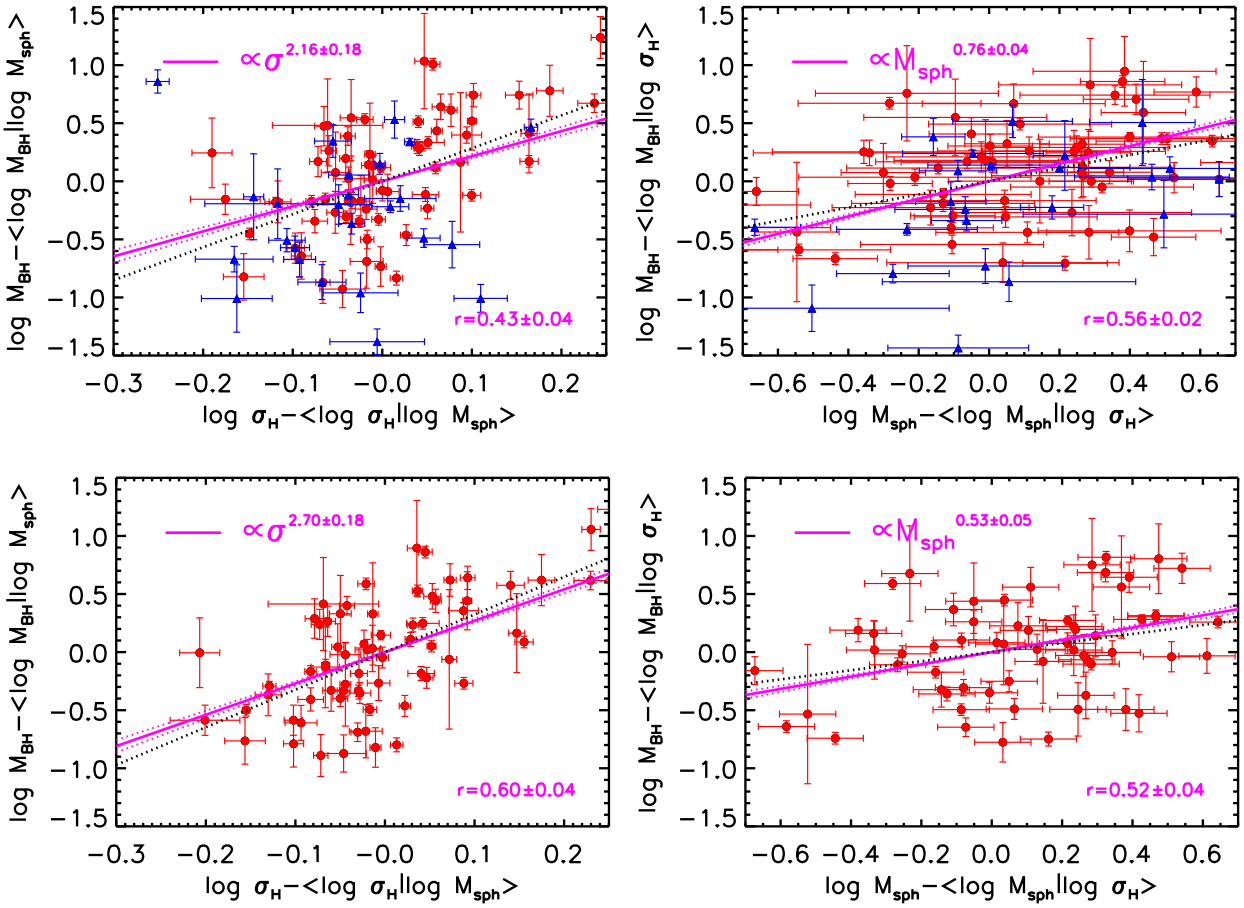

Figure 3. Same layout as Fig. 2 but only considering the spheroidal component of each galaxy. The residual with stellar velocity dispersion appears now equally significant to stellar mass, possibly hinting at the existence of a fundamental plane for the *spheroidal* component only.

with σ at fixed $M_{\rm sph}$, and also with $M_{\rm sph}$ at fixed σ, though flatter than in the Sahu et al. (2023) data. In other words, within the remit of the tests carried out in this work, the Simba simulation tends to point to stellar spheroidal mass as the most fundamental galactic property correlated with SMBH mass in central galaxies, while the TNG would still favour σ as the galactic property most correlated to SMBH mass. The residuals predicted by current state-of-the-art hydrodynamic simulations reported in Figs 4 and 5 depict SMBH–galaxy scaling relations that, although not yet fully aligned with what suggested by the current (and limited) data, are becoming increasingly closer to what observed, improving on previous comparisons (e.g. Barausse et al. 2017). These hydrodynamic simulations, however, do not suggest the existence of any dynamical fundamental plane of SMBHs, favouring a scenario in which the SMBH is closely correlated to only one single galactic variable, either $M_{\rm sph}$ or σ, for Simba and TNG, respectively (see also the discussion in Menci et al. 2023).

On the assumption that the $M_{\rm bh}$–σ relation and its residuals are mostly driven by AGN feedback (e.g. Silk & Rees 1998; Granato et al. 2004; Di Matteo et al. 2005; Robertson et al. 2006), we could test whether an increase in the AGN kinetic feedback could steepen and strengthen the relation with stellar velocity dispersion at fixed bulge mass, in better agreement with the data explored in this work. In the top and middle panels of Fig. 6 we report, in the same format as in Fig. 5, the predicted residuals of, respectively, the Simba and TNG models with an AGN kinetic outflow increased by a factor of 100, the so-called 'EX1' simulation runs in CAMELS (Anglés-Alcázar et al. 2021). As discussed in Section 2.2, the box of the EX simulations are significantly smaller than the reference ones, and thus the statistical results may be less robust. Nevertheless, it is intriguing to note that the residuals in both Simba and TNG are weakly affected by the increase in the AGN feedback kinetic output, generating a similar or even weaker correlations with σ at fixed $M_{\rm sph}$. More notably, as shown in the bottom panels of Fig. 6, the increase in AGN kinetic output in both simulations inevitably reduces the SMBH mass at fixed host galaxy stellar mass, more markedly for the Simba simulation, most probably induced by the self-regulation in SMBH growth, which is reduced proportionally to the availability of gas in the surrounding medium.

4.3 Machine learning algorithms

In the previous sections, we have used pairwise residuals to define the degree of linear correlation among several SMBH–galaxy scaling relations, finding that in the data σ appears as the most fundamental property linked to SMBH mass. The aim of this Section is to explore putative correlations among these variables by making use of a variety of distinct and complementary approaches, based on machine learning (ML) algorithms. We immediately stress that none of the correlation tests performed in this Section provides the wealth of information available from pairwise correlation residuals, which also output the slopes and *relative* strengths of the correlations.

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

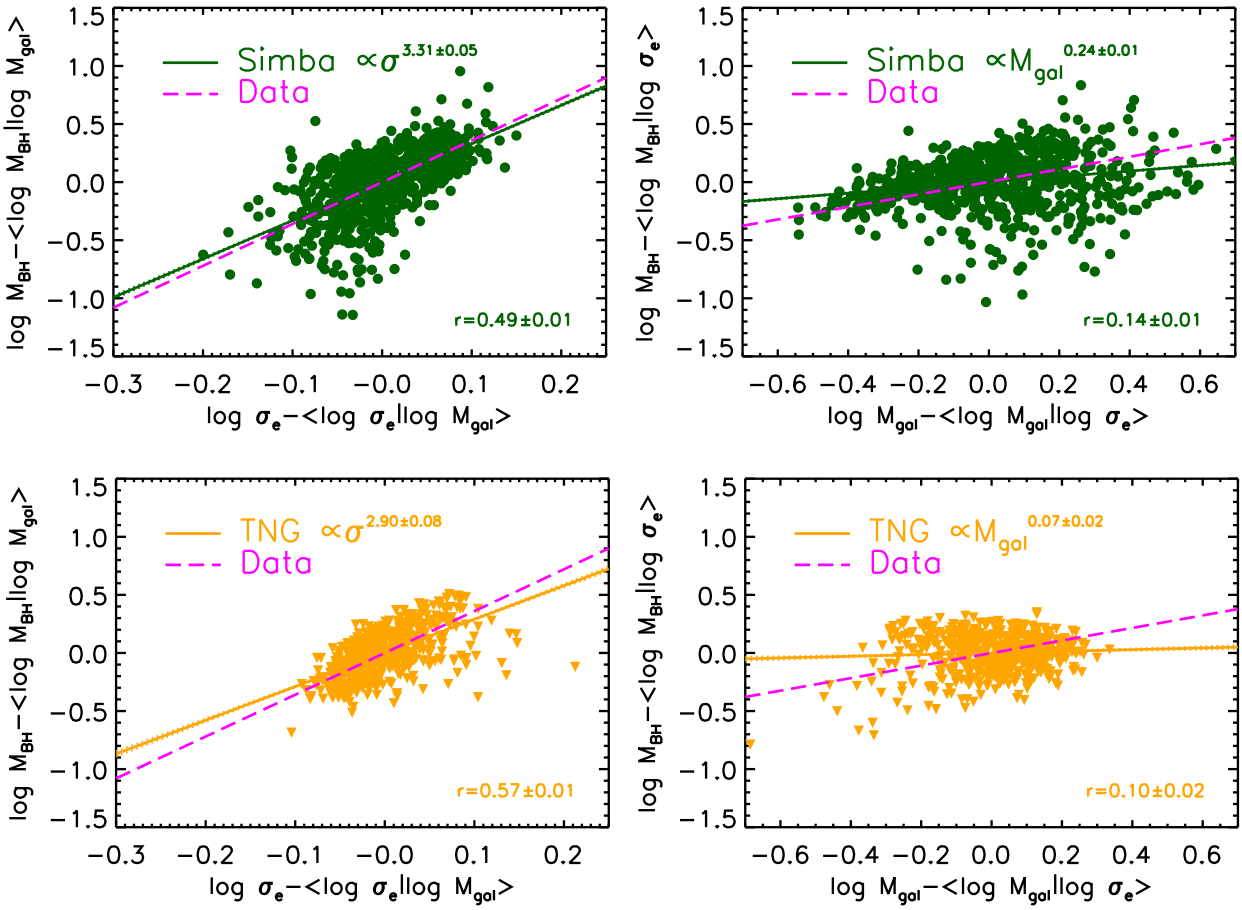

Figure 4. Residual analysis in the same format as in the top panels of Fig. 2, for the Simba (top panels) and TNG100 (bottom panels) simulations, as labelled, compared to the residuals extracted from the data (magenta dashed lines). Simulations predict weaker correlations with stellar velocity dispersion than in the data.

Nevertheless, the ones discussed here are useful complementary approaches to probe and confirm any degree of correlation.

In Fig. 7, we present the result of applying ML regression techniques to both the observed and simulated data. Each regression has been performed using only one variable in each case as a predictor of $M_{\rm bh}$. In this way, we can test the predictive power of each variable independently. As it is standard practice in ML regression analysis, we report in the plots the value of the coefficient of determination R^2 which indicates the quality of the fit by measuring the total variance of the outcome as indicated by the predictor

$$R^2(y, \hat{y}) = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}, \quad (4)$$

where y_i and $\bar{y}$ correspond to the measured and average data values, respectively, while $\hat{y}_i$ are the values predicted by a given model. The closer the coefficient of determination is to 1, the higher the correlation between the two variables, which indicates either a direct causal relationship between the variables or correlations through other intermediate variables. Values of R^2 close to zero or negative suggest, instead, no causal relationship between the variables. We note that the R^2 coefficient of determination is different from the Pearson correlation coefficient reported in the pairwise residuals. The two statistical indicators are in fact not directly comparable to each other for several reasons: (i) they are characterized in different ways, with r explicitly defined on both variables x and y, while R^2 explicitly defined on the y variable and only implicitly on x via the predicted value $\hat{y}$, which depends on x; (ii) the parameter r assumes a linear fit between x and y, while the ML regressions adopted here do not follow any particular functional relation between variables, allowing for any type of linear or non-linear relationship; (iii) the definition of r adopted in this work also includes measurement errors in both variables (see full formalism in Sheth & Bernardi (2012) and in Appendix B in Shankar et al. 2017). Despite these differences, both parameters provide a quantitative estimate of the degree of correlation strength among variables, and our aim here is only to test whether, within a given set of observational or numerical data, the two statistical indicators point to the same variables as being more or less correlated with the mass of the central SMBH.

For each predictor, we show the regression results using different techniques. This approach ensures the robustness of the results by confirming that they are independent of the details of any of the specific ML recipe adopted in the analysis. The selected algorithms include a wide and diverse set of methodologies, including linear and non-linear methods, Bayesian approaches, ensemble meta-approaches combining different estimators, gradient optimisation, etc. The techniques applied are the following: AdaBoost (Drucker 1997), Random Forest (Breiman 1998, 2001), Support Vector Machines with linear and non-linear kernels (Smola & Schölkopf 2004; Rong-En 2008; Chang & Lin 2011), Ridge Regression (Hoerl & Kennard 1970; Rifkin & Lippert 2007), Stochastic Gradient Descent (Zhang 2004; Tsuruoka, Tsujii & Ananiadou 2009; Bottou 2012), Gradient Tree Boosting (Friedman 2001, 2002), Bayesian

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

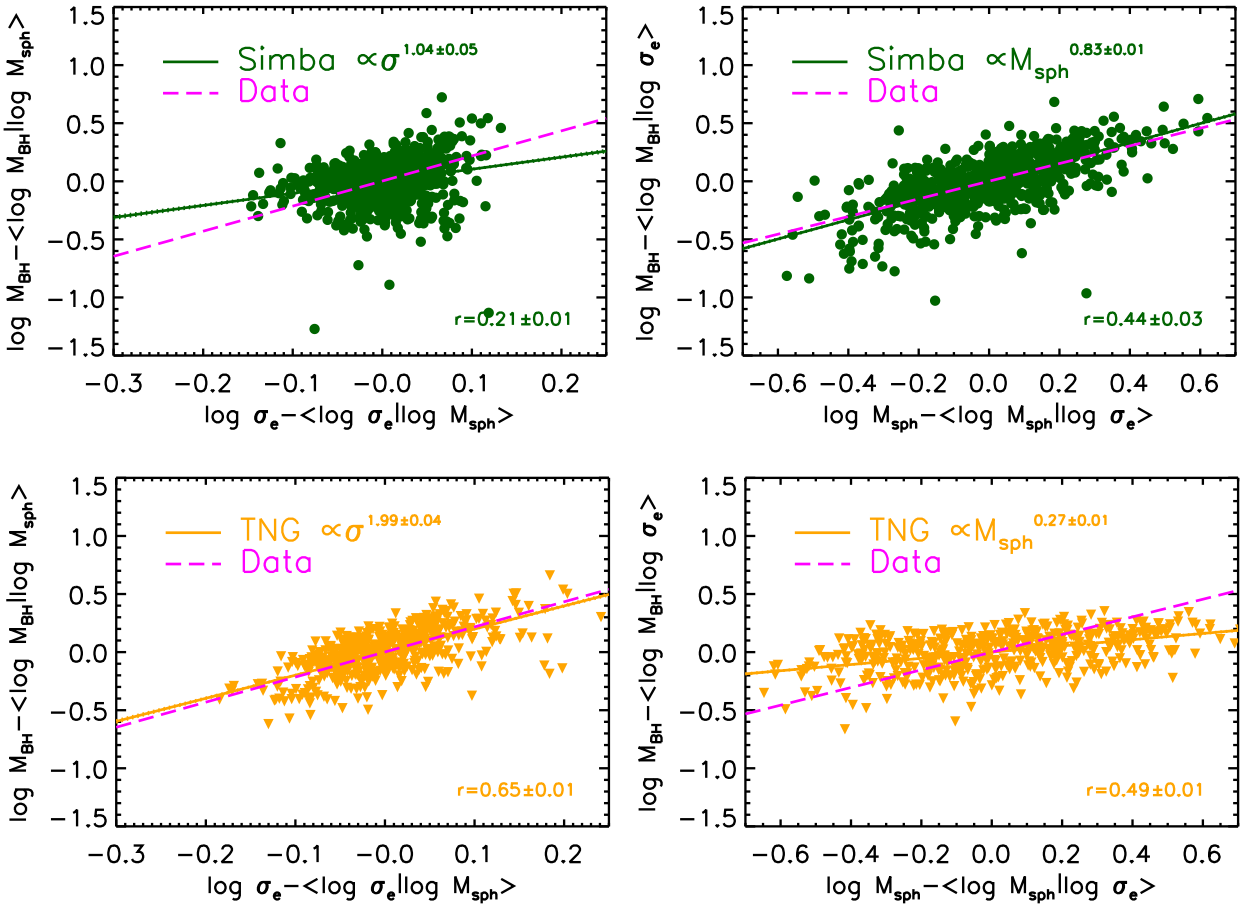

Figure 5. Same format as Fig. 4 but replacing total with bulge stellar mass. Simulations predict different results between them and with the data, with TNG being closer to what expected from a correlation between SMBH mass and host binding energy.

Ridge Regression (MacKay 1992; Tipping 2001), Decision Trees (Breiman et al. 1984), and K-nearest neighbours (Bentley 1975; Fix & Hodges 1989; Omohundro 1989). All these techniques have been implemented using the scikit-learn library (Pedregosa et al. 2011). The methodology applied is as follows. The data used in each regression are pre-processed by scaling them between zero and one to facilitate the regression. For each regressor, a grid search of its main hyperparameters is conducted in order to select those that obtain the best result evaluated by minimizing the mean squared error of the regression. The regression fit and its evaluation are carried out by means of cross-validation with fivefolds. This implies that each regression is carried out five times in each case, using different parts of the data to perform the fit and to evaluate the result. This ensures that the choice of which part of the data is used to perform the regression is not biasing the results, in line with the random selection of data we performed when computing the pairwise residuals. In each case, 80 per cent of the data is used for the regression and 20 per cent for its evaluation. The final result of the coefficient of determination and its error is calculated from the average and standard deviation of these five regressions.

The top, middle, and bottom panels of Fig. 7 report the results of the adopted ML algorithms, labelled on the y-axis, in terms of the R^2 parameter on the x-axis, applied, respectively, to the full data set of the local SMBH sample (top panels of Fig. 1), the (reference) Simba and TNG simulations (with reference values for the AGN feedback efficiencies). The first clear result is that all the ML predictors agree in finding a similarly significant correlation in the observational data between $M_{\rm bh}$ and σ and between $M_{\rm bh}$ and $M_{\rm sph}$ (top left and right panels), supportive of a fundamental plane relation, with values of the coefficient of determination on average $R^2 \sim 0.6$, but not an equally strong relation of $M_{\rm bh}$ with total stellar mass $M_{\rm gal}$ with an average $R^2 \sim 0.3$ (top, middle panel). These results are in line with what was concluded from the pairwise residuals in the top panels of Fig. 3, where the Pearson coefficient was higher for the correlation with σ and $M_{\rm sph}$, but showing a significantly weaker correlation with $M_{\rm gal}$ at fixed σ (Fig. 2). We also found that when restricting the analysis to only ETGs, the ML algorithms continue to point to a similar correlation with $M_{\rm sph}$, with $R^2 \sim 0.5$–0.6, and an even more marked correlation with σ, with $R^2 \sim 0.7$, in line with what was retrieved from pairwise residuals in the bottom panels of Fig. 3. We checked that the ML algorithms point to much weaker correlations of SMBH mass with both bulge Sérsic index and half-light radius, with an average $R^2 \sim 0.3$, a trend that, as suggested by Fig. B2, could be entirely ascribed to the underlying dependence of SMBH mass on $M_{\rm sph}$. Interestingly, the ML predictors do not identify any clear correlation between $M_{\rm bh}$ and σ or $M_{\rm gal}$ in the Simba simulation (middle panels), but a moderate one between $M_{\rm bh}$ and $M_{\rm sph}$ (middle right panel), a result which is consistent with the pairwise residuals analysis performed in the top panels of Fig. 5, which identified in Simba a noticeable correlation of SMBH mass with stellar bulge mass $M_{\rm sph}$ at fixed σ. On the other hand, the TNG simulation, according to the ML predictors, shows a strong, well-defined correlation between

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

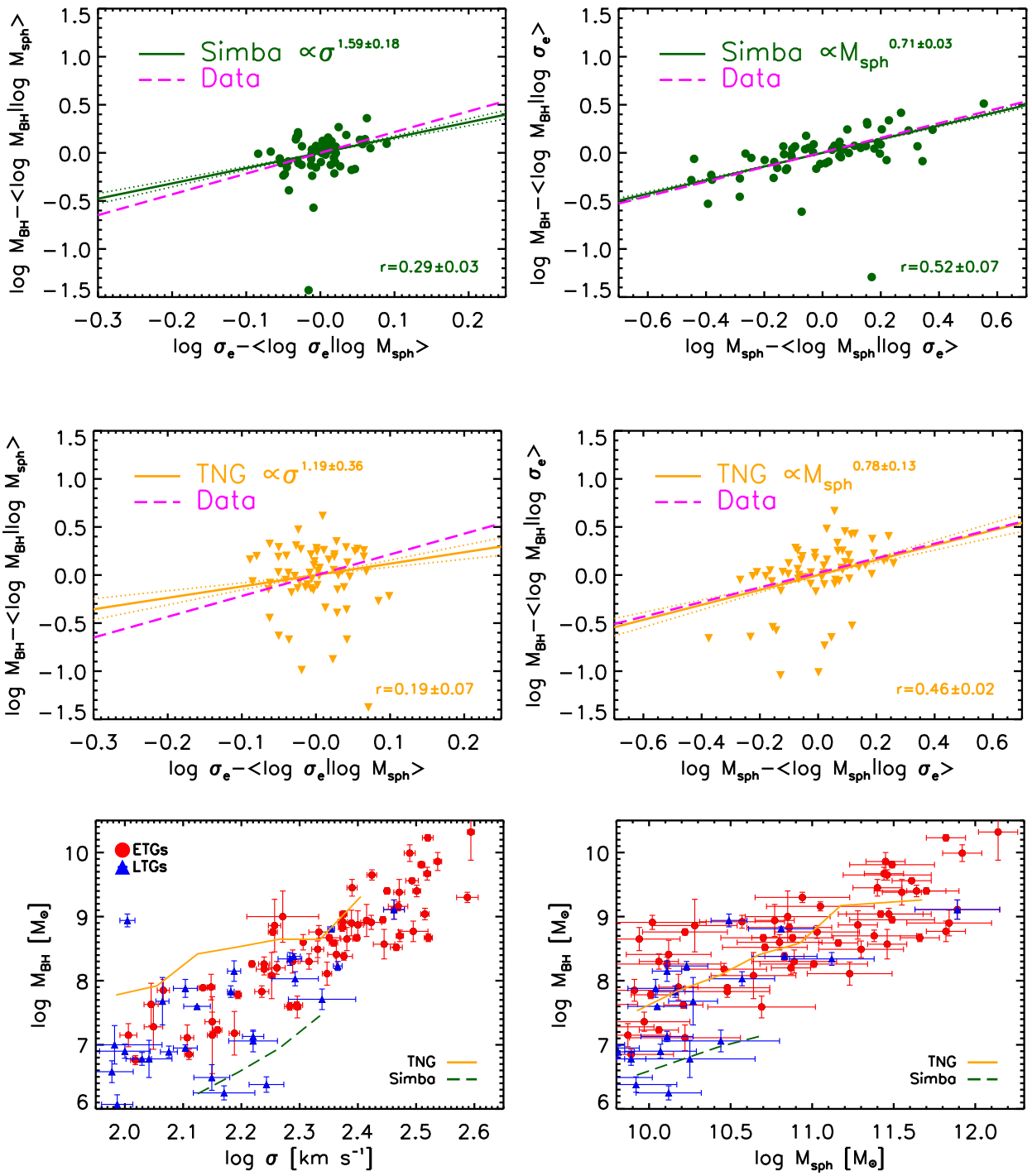

Figure 6. Same format as Fig. 4 but with increased AGN feedback efficiency (see the text for details). Bottom panels include the predicted SMBH–galaxy scaling relations in the new simulations as in the top panels of Fig. 1. An increased AGN feedback efficiency tends to increase the significance of the residuals, at least in the Simba simulation, at the cost of a significantly lower normalization in the scaling relations.

$M_{\rm bh}$ and σ with an average $R^2 \sim 0.5$, and, to a lesser extent with $M_{\rm sph}$, again in line with what was concluded from the pairwise residuals in the bottom panels of Fig. 5. All in all, the ML tests agree with the results of pairwise residuals, pointing to $M_{\rm sph}$, in the case of Simba, or σ and $M_{\rm sph}$, in the case of TNG, as the main galaxy properties correlated with the SMBH mass $M_{\rm bh}$. The ML algorithms adopted here do not, however, constrain the mathematical relation among the different variables and are thus unable to pin down any specific model of, e.g. AGN feedback.

5 DISCUSSION

One of the main results of this work is the evidence of a strong correlation between $M_{\rm bh}$ and σ. This empirical result represents a significant step forward in our understanding of the origin and evolution of SMBH scaling relations. The seminal work by Bernardi et al. (2007) had pointed out via residuals the key role played by the stellar velocity dispersion in the SMBH scaling relations, and the residual analysis carried out by Shankar et al. (2016), Shankar

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

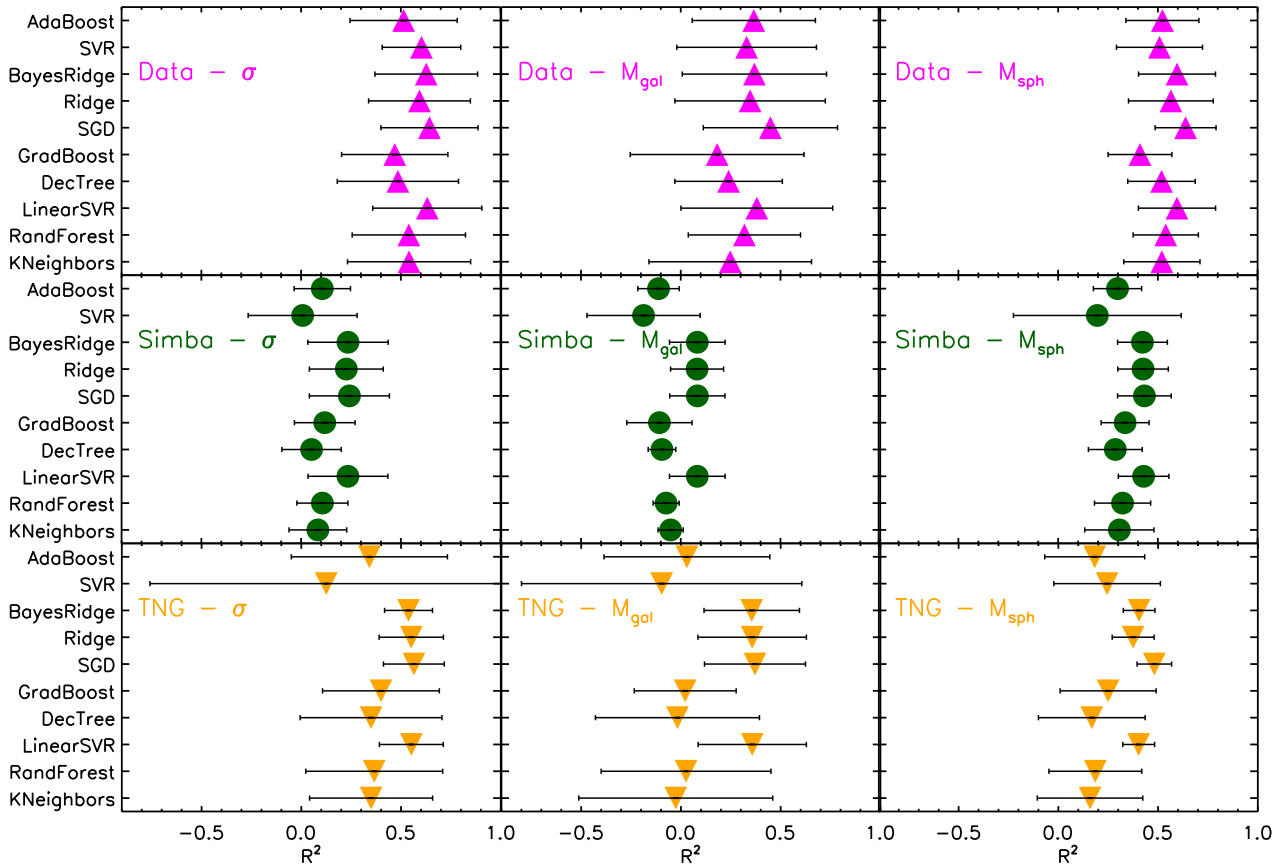

Figure 7. Calculation of the degree of correlation as labelled by the R^2 parameter, via a number of machine learning regression algorithms as listed on the left of the figure. Top, middle, and bottom panels refer to the results of the ML algorithms applied to the full data set of local galaxies with SMBH mass measurements, the predictions of the (reference) Simba and the TNG simulations, respectively. The left, middle, and right panels refer to the correlation with velocity dispersion σ, galaxy total stellar mass $M_{\rm gal}$, and galaxy bulge stellar mass $M_{\rm sph}$, respectively. The Simba simulation does not predict any strong correlation with any variable, except with $M_{\rm sph}$, while TNG predicts a correlation with σ comparable with what seen in the data and, to a lesser extent, also with $M_{\rm sph}$, in line with also derived from the pairwise residuals (see the text for details).

et al. (2017), and Marsden et al. (2020) confirmed this trend. The exquisite photometric homogeneity achieved by Sahu et al. (2019a) and Davis et al. (2019) on the local SMBH sample, now allows for a more accurate analysis of the residuals, providing further robust evidence for stellar velocity dispersion being more fundamental than total stellar mass or other photometric galactic properties, as shown in Appendix B.

We also found possible evidence for the existence of a fundamental plane between SMBH mass and stellar velocity dispersion and bulge stellar mass, of the type $M_{\rm bh} \propto \sigma^{2.2} M_{\rm sph}^{0.8}$, broadly aligned with what claimed by Hopkins et al. (2007b, see also Iannella et al. 2021). The 'dynamical' fundamental plane of SMBHs (not to be confused with the fundamental plane of SMBH activity by Merloni, Heinz & di Matteo 2003) becomes more tilted when only ETGs are considered (bottom panels of Fig. 3) and evidence for its existence is significantly weakened by the residual analysis performed on the SMBH sample by Saglia et al. (2016), which is also characterized by accurate stellar spheroidal measurements (Fig. B). Uniform and careful analyses on larger SMBH samples are required to unveil the actual existence of a dynamical SMBH fundamental plane, possibly also taking advantage of uniform measurements of host properties and SMBH masses of type 1, low/moderate luminosity AGN, which will soon become a reality with the Euclid (Euclid Collaboration: Y. Mellier et al. 2024) and Vera Rubin LSST (Ivezić et al. 2019) surveys.

The pairwise residuals might favour AGN feedback as a key driver shaping SMBH galaxy scaling relations, either directly via the $M_{\rm bh}$–σ relation (e.g. Silk & Rees 1998), or via a tilted correlation involving also spheroid stellar mass. A simultaneous correlation between SMBH mass with both stellar velocity dispersion and stellar spheroid mass, if confirmed, could favour a two-phase SMBH evolution (e.g. Cook, Lapi & Granato 2009; Oser et al. 2010; Boco et al. 2023), where SMBH's growth within the proto-spheroid, also possibly triggered by a gas-rich major merger and regulated by AGN feedback, could generate a tilted relation between SMBH mass and host dynamical mass (e.g. Hopkins et al. 2007a). The later phase of galaxy evolution, which may include the formation of a surrounding stellar disc (e.g. Cook et al. 2009; Hopkins et al. 2009; Cook et al. 2010), not necessarily linked with further growth onto the central SMBH, may loosen the correlation between SMBH mass and total galaxy stellar mass, as indeed suggested by current data.

On the other hand, detailed hydrodynamic simulations such as Simba, inclusive of cutting-edge implementations of AGN outflows, do not necessarily point to an AGN-driven origin of the SMBH scaling relations. Anglés-Alcázar et al. (2017, see also Anglés-Alcázar et al. 2015) showed that the shape of the $M_{\rm bh}$–$M_{\rm gal}$ relation is largely independent of the strength of AGN feedback. Similar normalizations and slopes are recovered in their simulations when moving from a model with no AGN feedback to one with a strong

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

outflow characterized by 20 times the fiducial value (see their fig. 4), with only a modest decrease in normalization by a factor of ~ 2 in the latter model. Similar conclusions were also reached by Menci et al. (2023) adopting a semi-analytic galaxy evolution model with a new treatment of AGN-driven winds. They showed that all SMBH scaling relations are largely preserved in both normalization and slope when including a significant AGN outflow component, which mostly controls the dispersion around the relations. Anglés-Alcázar et al. (2017) discussed that in their simulations SMBHs and galaxies grow in lockstep along the $M_{\rm bh}$–$M_{\rm gal}$ relation, the normalization of which is largely regulated by the parameters controlling the physics of accretion on to the central SMBHs. Dattathri et al. (2025, and references therein) have also discussed how the scaling with galaxy stellar mass arises from the tight relationship between SMBH accretion rate and galaxy star formation rate inducing, in some models like the ASTRID simulation (Bird et al. 2022), a nearly linear and redshift-independent $M_{\rm bh}$–$M_{\rm gal}$ relation.

The two state-of-the-art hydrodynamic simulations considered in this work, Simba and TNG, which implement different prescriptions for the SMBH gas accretion and AGN feedback (see Section 2.2), align with the previous theoretical studies discussed above, showing that a correlation with galaxy stellar mass, or a steep correlation with stellar velocity dispersion, are not necessarily a direct byproduct of an underlying AGN feedback-regulated process. As highlighted by several previous works (e.g. Barausse et al. 2017; Shankar et al. 2017; Menci et al. 2023), the apparent match between model predictions and data in the SMBH scaling relations may in fact sometimes simply arise as a result of a combination of different factors (see e.g. Cavaliere & Vittorini 2000 and discussion and appendices in Shankar et al. 2017). Inspired by the pairwise residuals' results, we could write the global relation between SMBH mass $M_{\rm bh}$ and galaxy properties as $M_{\rm bh} \propto \sigma^{\beta} M_{\rm sph}^{\alpha} \propto \sigma^{\beta+\alpha\,\gamma}$, where γ comes from the $M_{\rm sph} \propto \sigma^{\gamma}$ relation. The Sahu et al. (2023) SMBH sample yields a correlation of the type $M_{\rm sph} \propto \sigma^{3.5}$, which in turn would imply, when combined with the slopes from the top panels of Fig. 3, $M_{\rm bh} \propto \sigma^{2.2} M_{\rm sph}^{0.8}$, or $M_{\rm bh} \propto \sigma^{2.2+0.8\times3.5} \propto \sigma^{5}$. So the SMBH data may still yield a strong and steep correlation with stellar velocity dispersion, but mostly as a byproduct of the additional correlation with the spheroidal component. By repeating the same exercise for the Simba simulation, we would get $M_{\rm bh} \propto \sigma^{1+0.8\times3.8} \propto \sigma^{4}$ and for the TNG $M_{\rm bh} \propto \sigma^{2+0.3\times4.2} \propto \sigma^{3.3}$, which are close to the correlations seen in the bottom left panel of Fig. 1. In other words, both simulations could provide relatively steep $M_{\rm bh}$–σ correlations but for somewhat different reasons which are ultimately dictated by the different physics implemented in the models.

Pairwise residuals, more than the scaling relation themselves (e.g. Habouzit et al. 2021b), have the potential to distinguish among the most successful models and ultimately reveal the underlying physics regulating the co-evolution of SMBHs and their host galaxies. It is thus vital to compare theoretical predictions with the data not only in terms of absolute scaling relations but also in terms of their residuals to pin down the true performance of a given model. The statistical tests carried out in Fig. 7 via a variety of ML regression models, largely support the results from the pairwise residuals, but they are less informative, thus caution needs to be applied when interpreting causality among variables just based upon the outputs of ML-based regression algorithms (e.g. Bluck et al. 2020).

A tight underlying connection between SMBH mass, stellar velocity dispersion and bulge stellar mass, could still point to a self-regulated SMBH growth, where AGN feedback is controlled by the potential well of its host (e.g. López et al. 2023, and references therein). In this context, one would then also expect a correlation between SMBH mass and host dark matter halo mass $M_{\rm halo}$. Several authors have in fact found evidence for a tight correlation between stellar velocity dispersion and the large-scale circular velocity as traced by either dynamical models or HI rotation curves measurements (e.g. Ferrarese 2002; Baes et al. 2003; Pizzella et al. 2005, see also Cirasuolo et al. 2005), which, if extrapolated, could indicate an underlying correlation with the host halo mass.

To investigate this intriguing possibility using the unique power of pairwise residuals, we have selected the subsample of 41 galaxies from the SMBH parent sample of Sahu et al. (2023) with dark matter halo mass measurements from Marasco et al. (2021). The latter have halo masses either derived from globular cluster dynamics coupled with assumptions on the host dark matter halo and stellar profiles from Posti & Fall (2021), or derived from the circular velocity of the flat part of the rotation curve converted into halo masses using the linear relation $M_{\rm halo}$–$v_{\rm flat}$ from Posti et al. (2019, see appendix A in Marasco et al. 2021 for full details). Fig. 8 shows the residuals of this subsample in terms of $M_{\rm halo}$ at fixed σ and $M_{\rm sph}$ (right panels) and in terms of σ and $M_{\rm sph}$ at fixed $M_{\rm halo}$ (left panels). It is evident that a significant correlation exists between SMBH mass and host halo mass, while the correlation with both σ and with $M_{\rm sph}$ are weaker at fixed halo mass which thus appears even more correlated with SMBH mass than either of the galactic variables, possibly because the latter are in turn controlled by the host halo mass (analogously to the case of the residual in R_e, the apparent anticorrelation with $M_{\rm sph}$ at fixed $M_{\rm halo}$ is a by-product of some random errors and not significant, as demonstrated by the nearly flat direct fit to the residuals). For completeness, we also checked the residual correlation between SMBH mass and potential well of the spheroidal component $W \propto \sigma^2 M_{\rm sph}$ at fixed halo mass, finding a very weak, slightly negative correlation with $r \sim -0.33$, while still a strong correlation with halo mass at fixed W, with $r \sim 0.66$. Interestingly, we do not detect any significant residual correlation with halo mass in Simba while a moderate correlation with halo mass is found in TNG. Caution has to be taken in overinterpreting the results in Fig. 8 as the halo mass measurements are still based on some ad-hoc assumptions and the sample is nearly half of the original one from Sahu et al. (2023), which is already quite small.

Nevertheless, the conclusions drawn from Fig. 8 would also align with the independent results from some AGN clustering measurements. Powell et al. (2022) have shown from forward modelling of the Swift/BAT AGN Spectroscopic Survey that AGN clustering measurements prefer models with a direct correlation between SMBH mass and host halo mass at fixed stellar mass, and their best-fitting $M_{\rm bh}$–$M_{\rm halo}$ relation is in line with what retrieved from other groups using independent clustering measurements and different AGN samples (e.g. Shankar et al. 2020a; Allevato et al. 2021). On the other hand, from pairwise residuals analysis we checked that the Simba and TNG simulations do not show any noticeable dependence between $M_{\rm bh}$ and $M_{\rm halo}$ at fixed σ ($r \lesssim 0.15$), further corroborated by the ML regression algorithms, which also point to a negligible correlation with $M_{\rm halo}$ ($R^2 \lesssim 0.25$).

Beyond host halo mass, other galactic properties have been put forward in the literature along the years as potentially providing a stronger correlation to SMBH mass than stellar velocity dispersion. For example, several groups have remarked that in cored ellipticals one of the strongest correlations is between the size of the central stellar core and SMBH mass, possibly as a result of the scouring of stars by mergers of binary SMBHs (e.g. Rusli et al. 2013, and references therein). We verified from pairwise residuals that, indeed, in the sample of ellipticals collected and studied by Rusli et al.

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

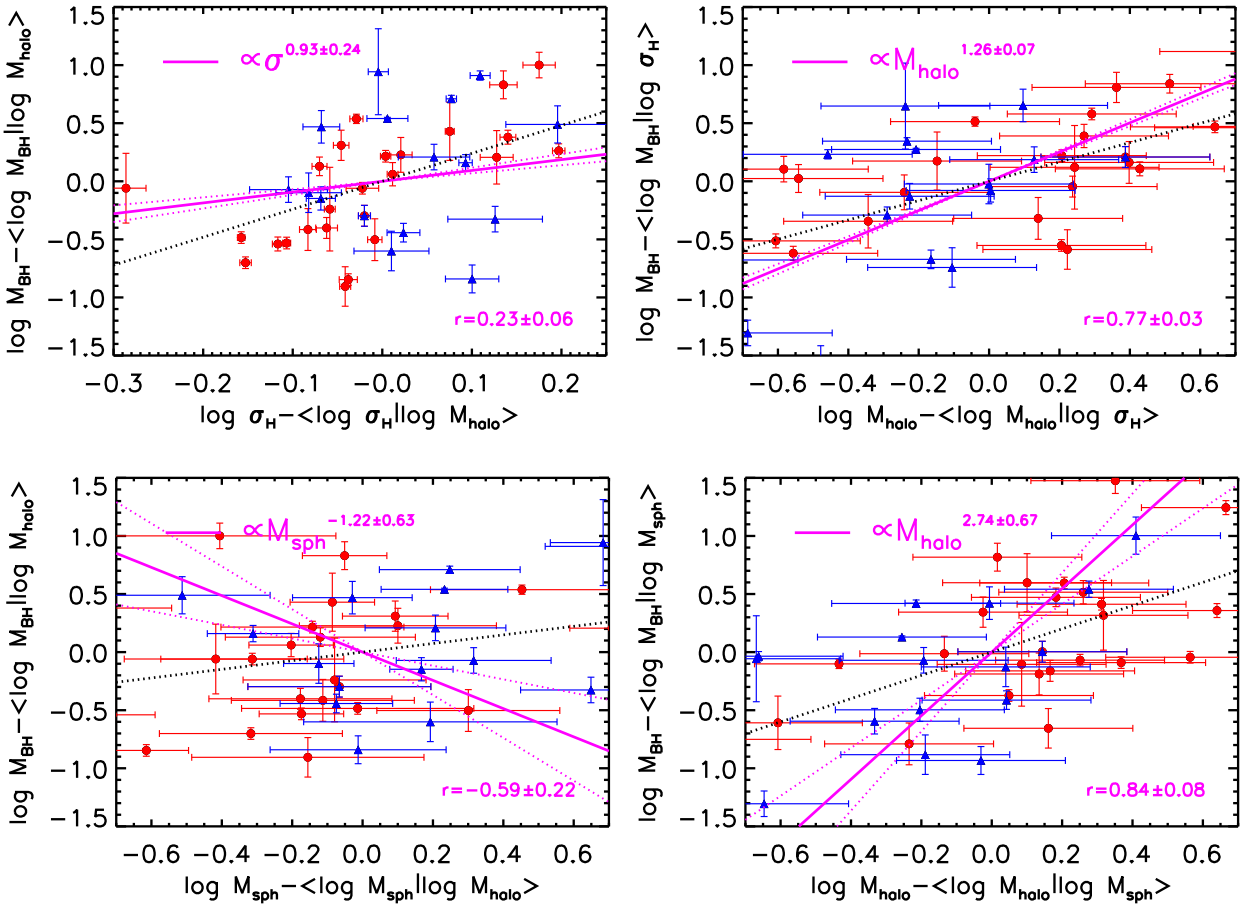

Figure 8. Residuals as a function of σ, $M_{\rm sph}$, and $M_{\rm halo}$ for the subsample of SMBHs from Sahu et al. (2023) with host dark matter halo mass measurement from Marasco et al. (2021). The residuals with host halo mass appear, respectively, stronger than those with σ and $M_{\rm sph}$ at fixed halo mass.

(2013), also inclusive of NGC1272 from the latest work by Saglia et al. (2024), the SMBH mass appears significantly more strongly correlated with core radius than with stellar velocity dispersion [taken from the Hyperleda data base, for consistency with the analysis carried above on the Sahu et al. (2023) SMBH sample]. However, this result mostly stems from the fact that the limited subsample of 24 cored galaxies with dynamical mass measurements of their central SMBHs (most of which are already included in the Sahu et al. (2023) SMBH sample) is characterized by an overall poor correlation between SMBH mass and host galaxy stellar velocity dispersion.

The possible overall dominance of stellar velocity dispersion (or possibly host potential well or host halo mass or even core radius) in driving the SMBH–galaxy scaling relations has profound consequences in the way observers and modellers approach the study of the co-evolution of these two systems. The slope and normalization of the $M_{\rm bh}$–$M_{\rm gal}$ relation must reflect the shape and evolution of the $M_{\rm bh}$–σ, $M_{\rm bh}$–$M_{\rm sph}$, or $M_{\rm bh}$–$M_{\rm halo}$ relations (e.g. Menci et al. 2023), thus the interpretation of SMBH evolution only through the lens of the $M_{\rm bh}$–$M_{\rm gal}$ plane may be incomplete, affected by large scatter and/or biases (see Appendix A), and possibly even misleading. Terrazas et al. (2020) for example studied the effects of AGN feedback on the $M_{\rm bh}$–$M_{\rm gal}$ relation coupled to specific star formation rate, but stellar velocity dispersion and, possibly, spheroid stellar mass could represent more robust variables to test recipes for AGN feedback both observationally and theoretically.

On a similar vein, observational evidence is accumulating for a significant evolution in the $M_{\rm bh}$–$M_{\rm gal}$ relation, at least when comparing similar types of galaxies at different epochs (e.g. Farrah et al. 2023), while new high-redshift measurements from local and high-redshift AGN from, e.g. JWST, are showing larger scatters and/or non-trivial evolution (e.g. Reines & Volonteri 2015a; Pacucci et al. 2023). As described by Maiolino et al. (2025), collectively the new JADES high-z SMBHs tend to better align with the local $M_{\rm bh}$–σ relation (Maiolino et al. 2024; Juodžbalis et al. 2024), suggesting that the latter is more fundamental, in line with the findings in this work, and has a relatively weak evolution (e.g. Shankar, Bernardi & Haiman 2009b; Shen et al. 2015). Indeed, many studies carried out on serendipitous AGN samples tend to favour a much weaker evolution in both scaling relations (e.g. Shen et al. 2015; Suh et al. 2020; Marsden et al. 2022; López et al. 2023; Zhuang & Ho 2023; Tanaka et al. 2025), as also suggested by arguments based on the time integrated emissivity of AGN converted to SMBH mass densities via a mean radiative efficiency (e.g. Shankar et al. 2009b). It is also interesting to remark that SMBH accretion models based on the time integration of empirical (mostly X-ray/IR based) Eddington ratio distributions $P(\lambda \propto L/M_{\rm gal}, z)$, continue to point to mean $M_{\rm bh}$–$M_{\rm gal}$ scaling relations significantly below those traced by local dynamically measured inactive SMBHs when adopting standard radiative efficiencies of $\gtrsim 10$ per cent, further suggesting some tensions in the demography of SMBHs in the $M_{\rm bh}$–$M_{\rm gal}$ relation (e.g. Reines & Volonteri 2015b; Yang et al. 2017; Shankar et al.

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

2019; Carraro et al. 2020; Suh et al. 2020; Shankar et al. 2020a, b; Carraro et al. 2022; Farrah et al. 2023; Terrazas, Aird & Coil 2024; Zou et al. 2024), which largely disappear when considering SMBH accretion models performed on the $M_{\rm bh}$–σ plane (e.g. Ricarte & Natarajan 2018a, b; Marsden et al. 2022).

SMBH mergers could also be playing a non-negligible role in shaping the SMBH-galaxy scaling relations (e.g. Hirschmann et al. 2010; Jahnke & Macciò 2011b; Shankar et al. 2012; Zou et al. 2024), but they must act in ways to preserve the $M_{\rm bh}$–σ relation first (e.g. Robertson et al. 2006), which is also the primary relation that should be considered when preparing galaxy-AGN mocks (e.g. Allevato et al. 2021). Our results also have implications for the interpretation of the stochastic gravitational wave background extracted from the 15-yr pulsar timing array (PTA) data set collected by the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) collaboration (Afzal et al. 2023). As discussed by Afzal et al. (2023), several SMBH binary populations are able to reproduce both the amplitude and shape of the observed low-frequency GW spectrum (Sesana et al. 2016; Agazie et al. 2023, and references therein). However, as pointed out by Lacy et al. (2024), the peaks of the posterior distributions in Agazie et al. (2023) point to a SMBH mass density up to an order of magnitude larger than many of the previous determinations of the local SMBH mass function (e.g. Marconi et al. 2004; Shankar et al. 2004; Tundo et al. 2007; Shankar, Weinberg & Miralda-Escudé 2009a; Shankar 2013; Shankar et al. 2020a, b; Sicilia et al. 2022).

Liepold & Ma (2024) have recently attempted a new measurement of the local SMBH function confirming the long-standing issue of a systematic discrepancy between the SMBH mass densities inferred from the $M_{\rm bh}$–$M_{\rm gal}$ and $M_{\rm bh}$–σ relations, with the latter usually providing a significantly lower value (Yu & Tremaine 2002; Shankar et al. 2004; Tundo et al. 2007; Shankar 2009). Shankar et al. (2016) showed that this discrepancy could be reconciled in light of a possible bias (see Appendix A) that preferentially affects the local $M_{\rm bh}$–$M_{\rm gal}$ relation. Burke et al. (2024) also recently estimated a local SMBH mass function consistent, at least at the high-mass end, with Liepold & Ma (2024) when adopting the $M_{\rm bh}$–$M_{\rm gal}$ relation, although significantly below it at lower SMBH masses. Liepold & Ma (2024) suggested that the (higher) value of the SMBH mass function implied by the $M_{\rm bh}$–$M_{\rm gal}$ relation should be preferred to the one arising from the $M_{\rm bh}$–σ relation, as the former is more consistent with the large masses of the SMBH binaries required by the PTA independent measurements. However, the pairwise residual results from this work would indicate that the $M_{\rm bh}$–σ relation is a more solid tracer of SMBH mass, and thus a more secure route towards a more robust census of SMBHs. The residuals would thus suggest that more accurate models of the SMBH mass function and SMBH merger rates should be anchored to the $M_{\rm bh}$–σ relation to provide a more insightful comparison with current and future GW measurements. Interestingly, at present, the large masses of the SMBH binaries derived from current PTA measurements are in tension with those extracted from the $M_{\rm bh}$–σ relation, as recently emphasized by Sato-Polito, Zaldarriaga & Quataert (2025), highlighting once more the well-known systematic discrepancy between the SMBH mass densities inferred from the two scaling relations with stellar velocity dispersion and stellar mass (e.g. Shankar et al. 2004; Tundo et al. 2007; Liepold & Ma 2024; Burke et al. 2025).

6 CONCLUSIONS

The apparently strong correlations between SMBH masses and their host galaxy properties carry the imprint of a possible co-evolution of these two systems. It is of vital importance to unveil the underlying nature and significance of these correlations if we want to truly advance in our knowledge of the growth of SMBHs and their putative impact on to the mass, structural, and dynamical evolution of their hosts. In this work, we have collected the latest sample of local galaxies from Sahu et al. (2023) with a uniform calibration of their photometric properties and with dynamically measured masses of their central SMBHs, analysed the scaling relations with pairwise residuals and ML regression algorithms, and compared the results with the outputs from two state-of-the-art hydrodynamic simulations, Simba and TNG. Our main results can be summarized as follows:

(i) The latest local SMBH data with uniform photometric calibrations and accurate bulge-to-total decompositions, point to some degree of correlation between SMBH mass $M_{\rm bh}$ and their host galaxy stellar velocity dispersion σ and stellar mass $M_{\rm gal}$. LTGs follow a steeper $M_{\rm bh}$–$M_{\rm gal}$ correlation than ETGs, while both ETGs and LTGs align better with spheroidal mass in the $M_{\rm bh}$–$M_{\rm sph}$ relation. The models broadly follow these trends (Fig. 1).

(ii) The pairwise residuals confirm that $M_{\rm bh}$ correlates more strongly with σ than with total stellar mass $M_{\rm gal}$, spheroidal effective radius, or Sérsic index (Figs 2 and B2).

(iii) When the spheroid stellar mass is considered, the local SMBH sample hints at a possible fundamental plane-type correlation of the type $M_{\rm bh} \propto \sigma^{2.2} M_{\rm sph}^{0.8}$ (Fig. 3), although this evidence is significantly weakened when switching to the Saglia et al. (2016) SMBH sample (Fig. B1).

(iv) The hydrodynamic simulations considered in this work, TNG and Simba, tend to align with the data in the residuals of $M_{\rm bh}$ versus σ at fixed $M_{\rm gal}$ and predict a negligible correlation with $M_{\rm gal}$ at fixed σ (Fig. 4).

(v) In addition, Simba favours $M_{\rm sph}$ as the galactic variable most closely linked to SMBH mass, while TNG points to σ (Fig. 5). Increasing the AGN kinetic output in these simulations does not change these trends (Fig. 6).

(vi) Another strong underlying correlation is found between $M_{\rm bh}$ and host halo mass $M_{\rm halo}$ (Fig. 8), albeit on a significantly smaller sample and with some assumptions on the derivation of the halo masses.

(vii) We show that the sample of local galaxies with dynamically measured SMBHs is biased high in σ with respect to the (significantly) larger sample of local galaxies as measured in the 3.6 μm or in the SDSS bands, especially towards lower luminosities (Fig. A1). We discuss this important point in detail in Appendix A.

Our results on pairwise residuals favour stellar velocity dispersion and host halo mass as more important global variables regulating the connection between SMBHs and their hosts. The comprehensive models explored here are able to broadly capture these trends, although they still require extreme fine-tuning in their physical prescriptions to simultaneously reproduce the observed SMBH scaling relations *and* their pairwise residuals, in particular with bulge stellar mass $M_{\rm sph}$ and possibly host halo mass $M_{\rm halo}$. Future models and data must continue to probe the strong link visible in the local Universe between $M_{\rm bh}$, σ and potentially even host halo mass $M_{\rm halo}$, both as a reference for SMBH mass measurements, as well as a more robust probe of the evolutionary link between SMBHs and their hosts and the implied stochastic gravitational wave background.

ACKNOWLEDGEMENTS

FS acknowledges partial support from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

Curie grant agreement no. 860744 'Big Data Applications for Black Hole Evolution Studies' (BiD4BESt; Coordinator F. Shankar). We acknowledge Nandini Sahu, Alister Graham, and Benjamin Davis for sharing some of their data and for useful clarifications on their data. FS also thanks Romeel Davé and Roberto Saglia for interesting discussions. HF acknowledges support from the Shanghai Super Post-doctoral Excellence Program grant no. 2024008.

DATA AVAILABILITY

All the data and simulations adopted in this paper are publicly available. The software underlying the statistical analysis can be made available upon reasonable request to the corresponding author.

REFERENCES

Afzal A. et al., 2023, ApJ, 951, L11
Agazie G. et al., 2023, ApJ, 952, L37
Allevato V., Shankar F., Marsden C., Rasulov U., Viitanen A., Georgakakis A., Ferrara A., Finoguenov A., 2021, ApJ, 916, 34
Anglés-Alcázar D., Özel F., Davé R., 2013, ApJ, 770, 5
Anglés-Alcázar D., Özel F., Davé R., Katz N., Kollmeier J. A., Oppenheimer B. D., 2015, ApJ, 800, 127
Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017, MNRAS, 464, 2840
Anglés-Alcázar D. et al., 2021, ApJ, 917, 53
Arjona-Galvez E., Di Cintio A., Grand R. J. J., 2024, A&A, 690, 286
Baes M., Buyle P., Hau G. K. T., Dejonghe H., 2003, MNRAS, 341, L44
Barausse E., Shankar F., Bernardi M., Dubois Y., Sheth R. K., 2017, MNRAS, 468, 4782
Batcheldor D., 2010, ApJ, 711, L108
Beifiori A., Courteau S., Corsini E. M., Zhu Y., 2012, MNRAS, 419, 2497
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289
Bentley J. L., 1975, Commun. ACM, 18, 509
Bernardi M., Sheth R. K., Tundo E., Hyde J. B., 2007, ApJ, 660, 267
Bernardi M., Roche N., Shankar F., Sheth R. K., 2011a, MNRAS, 412, L6
Bernardi M., Roche N., Shankar F., Sheth R. K., 2011b, MNRAS, 412, 684
Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697
Bernardi M., Meert A., Vikram V., Huertas-Company M., Mei S., Shankar F., Sheth R. K., 2014, MNRAS, 443, 874
Bernardi M., Meert A., Sheth R. K., Fischer J. L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017, MNRAS, 467, 2217
Bichang'a B., Kaviraj S., Lazar I., Jackson R. A., Das S., Smith D. J. B., Watkins A. E., Martin G., 2024, MNRAS, 532, 613
Bird S., Ni Y., Matteo T. D., Croft R., Feng Y., Chen N., 2022, MNRAS, 512,3703
Bluck A. F. L. et al., 2020, MNRAS, 499, 230
Boco L., Lapi A., Shankar F., Fu H., Gabrielli F., Sicilia A., 2023, ApJ, 954, 97
Bogdán Á. et al., 2024, Nat. Astron., 8, 126
Bottou L., 2012, in Neural Networks: Tricks of the Trade, 2nd edn. Springer, p. 421
Breiman L., 1998, Ann. stat., 26, 801
Breiman L., 2001, Mach. Learn., 45, 5
Breiman L., Friedman J., Stone C. J., Olshen R. A., 1984, Classification and regression trees. CRC press
Burke C. J., Natarajan P., Baldassare V. F., Geha M., 2025, ApJ, 978, 77
Burke C. J., Natarajan P., Baldassare V. F., Geha M., 2025, ApJ, 978, 77
Byrne Z. et al., 2023, MNRAS, 526, 1095
Cano-Díaz M., Maiolino R., Marconi A., Netzer H., Shemmer O., Cresci G., 2012, A&A, 537, L8
Cappellari M. et al., 2006, MNRAS, 366, 1126
Carniani S. et al., 2016, A&A, 591, A28
Carraro R. et al., 2020, A&A, 642, A65
Carraro R., Shankar F., Allevato V., Rodighiero G., Marsden C., Arévalo P., Delvecchio I., Lapi A., 2022, MNRAS, 512, 1185
Cattaneo A., Dekel A., Devriendt J., Guiderdoni B., Blaizot J., 2006, MNRAS, 370, 1651
Cavaliere A., Vittorini V., 2000, ApJ, 543, 599
Chang C.-C., Lin C.-J., 2011, ACM Transactions on Intelligent Systems and Technology (TIST), 2, 1
Cicone C. et al., 2014, A&A, 562, A21
Cirasuolo M., Shankar F., Granato G. L., De Zotti G., Danese L., 2005, ApJ, 629, 816
Cook M., Lapi A., Granato G. L., 2009, MNRAS, 397, 534
Cook M., Barausse E., Evoli C., Lapi A., Granato G. L., 2010, MNRAS, 402, 2113
Dattathri S., Natarajan P., Porras-Valverde A. J., Burke C. J., Chen N., Di Matteo T., Ni Y., 2025, ApJ, 984, 122
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Davis B. L., Graham A. W., Cameron E., 2019, ApJ, 873, 85
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Ding X. et al., 2023, Nature, 621, 51
Drucker H., 1997, in Icml. p. 107
Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, MNRAS, 463, 3948
Euclid Collaboration et al., 2024, A&A, 697, 94
Fabian A. C., 2012, ARA&A, 50, 455
Farrah D. et al., 2012, ApJ, 745, 178
Farrah D. et al., 2023, ApJ, 943, 133
Ferrarese L., 2002, ApJ, 578, 90
Ferrarese L., Ford H., 2005, Space Sci. Rev., 116, 523
Figueira M., Siudek M., Pollo A., Krywult J., Vergani D., Bolzonella M., Cucciati O., Iovino A., 2024, A&A, 687, A117
Fiore F. et al., 2017, A&A, 601, A143
Fix E., Hodges J. L., 1989, Int. Stat. Rev./Rev. Int. de Stat., 57, 238
Fontanot F., Monaco P., Shankar F., 2015, MNRAS, 453, 4112
Fontanot F. et al., 2020, MNRAS, 496, 3943
Forbes D. A., Sinpetru L., Savorgnan G., Romanowsky A. J., Usher C., Brodie J., 2017, MNRAS, 464, 4611
Friedman J. H., 2001, Ann. Stat., 29, 1189
Friedman J. H., 2002, Comput. Stat. Data Anal., 38, 367
Graham A. W., 2016, Galactic Bulges, 418, 263
Graham A. W., 2023a, MNRAS, 521, 1023
Graham A. W., 2023b, MNRAS, 522, 3588
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, ApJ, 600, 580
Greene J. E. et al., 2024, ApJ, 964, 39
Gültekin K., Tremaine S., Loeb A., Richstone D. O., 2011, ApJ, 738, 17
Habouzit M. et al., 2021a, MNRAS, 503, 1940
Habouzit M. et al., 2021b, MNRAS, 503, 1940
Habouzit M. et al., 2022a, MNRAS, 509, 3015
Habouzit M. et al., 2022b, MNRAS, 511, 3751
Hirschmann M., Khochfar S., Burkert A., Naab T., Genel S., Somerville R. S., 2010, MNRAS, 407, 1016
Ho L. C., Kim M., 2014, ApJ, 789, 17
Hoerl A. E., Kennard R. W., 1970, Technometrics, 12, 55
Hon D. S. H., Graham A. W., Davis B. L., Marconi A., 2022, MNRAS, 514, 3410
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., Quataert E., 2011, MNRAS, 415, 1027
Hopkins P. F., Hernquist L., Cox T. J., Robertson B., Krause E., 2007a, ApJ, 669, 45
Hopkins P. F., Hernquist L., Cox T. J., Robertson B., Krause E., 2007b, ApJ, 669, 67
Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, ApJ, 691, 1168
Hu H., Inayoshi K., Haiman Z., Li W., Quataert E., Kuiper R., 2022, ApJ, 935, 140
Iannella A. L., Greco L., Feoli A., 2021, Ap&SS, 366, 52
Inayoshi K., Ichikawa K., 2024, ApJ, 973, L49
Ivezić Ž. et al., 2019, ApJ, 873, 111

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

Jahnke K., Macciò A. V., 2011a, ApJ, 734, 92
Jahnke K., Macciò A. V., 2011b, ApJ, 734, 92
Juodžbalis I. et al., 2024, Nature, 636, 594
King A., 2003, ApJ, 596, L27
Kokorev V. et al., 2023, ApJ, 957, L7
Kormendy J., 2020, in Valluri M., Sellwood J. A., eds, Galactic Dynamics in the Era of Large Surveys, Vol. 353. p. 186, preprint (arXiv:1909.10821)
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Lacy M., Engholm A., Farrah D., Ejercito K., 2024, ApJ, 961, L33
Läsker R., Ferrarese L., van de Ven G., Shankar F., 2014, ApJ, 780, 70
Li Y. et al., 2020, ApJ, 895, 102
Li J. et al., 2025, ApJ, 981, 19
Liepold E. R., Ma C.-P., 2024, ApJ, 971, L29
López I. E. et al., 2023, A&A, 672, A137
MacKay D. J., 1992, Neur. Comput., 4, 415
Maiolino R. et al., 2024, A&A, 691, 29
Maiolino R. et al., 2025, MNRAS, 538, 1921
Marasco A., Cresci G., Posti L., Fraternali F., Mannucci F., Marconi A., Belfiore F., Fall S. M., 2021, MNRAS, 507, 4274
Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, MNRAS, 351, 169
Marsden C., Shankar F., Ginolfi M., Zubovas K., 2020, Front. Phys., 8, 61
Marsden C., Shankar F., Bernardi M., Sheth R. K., Fu H., Lapi A., 2022, MNRAS, 510, 5639
Matthee J. et al., 2024, ApJ, 963, 129
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
Menci N., Gatti M., Fiore F., Lamastra A., 2014, A&A, 569, A37
Menci N., Fiore F., Shankar F., Zanisi L., Feruglio C., 2023, A&A, 674, A181
Merloni A., Heinz S., di Matteo T., 2003, MNRAS, 345, 1057
Mezcua M., Domínguez Sánchez H., 2024, MNRAS, 528, 5252
Morabito L. K., Dai X., 2012, ApJ, 757, 172
Musiimenta B. et al., 2023, A&A, 679, A84
Nanni L. et al., 2023, MNRAS, 522, 5479
Nelson D. et al., 2019, MNRAS, 490, 3234
Omohundro S. M., 1989, Five balltree construction algorithms. International Computer Science Institute Berkeley
Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312
Pacucci F., Nguyen B., Carniani S., Maiolino R., Fan X., 2023, ApJ, 957, L3
Paturel G., Petit C., Prugniel P., Theureau G., Rousseau J., Brouty M., Dubois P., Cambrésy L., 2003, A&A, 412, 45
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Penny S. J. et al., 2018, MNRAS, 476, 979
Pillepich A. et al., 2018, MNRAS, 473, 4077
Pizzella A., Corsini E. M., Dalla Bontà E., Sarzi M., Coccato L., Bertola F., 2005, ApJ, 631, 785
Posti L., Fall S. M., 2021, A&A, 649, A119
Posti L., Marasco A., Fraternali F., Famaey B., 2019, A&A, 629, A59
Powell M. C. et al., 2022, ApJ, 938, 77
Reines A. E., Volonteri M., 2015a, ApJ, 813, 82
Reines A. E., Volonteri M., 2015b, ApJ, 813, 82
Ricarte A., Natarajan P., 2018a, MNRAS, 474, 1995
Ricarte A., Natarajan P., 2018b, MNRAS, 481, 3278
Rifkin R. M., Lippert R. A., 2007
Robertson B., Hernquist L., Cox T. J., Di Matteo T., Hopkins P. F., Martini P., Springel V., 2006, ApJ, 641, 90
Rong-En F., 2008, J. Mach. Learn. Res., 9, 1871
Rusli S. P., Erwin P., Saglia R. P., Thomas J., Fabricius M., Bender R., Nowak N., 2013, AJ, 146, 160
Saglia R. P. et al., 2016, ApJ, 818, 47
Saglia R. et al., 2024, A&A, 692, A124
Sahu N., Graham A. W., Davis B. L., 2019a, ApJ, 876, 155
Sahu N., Graham A. W., Davis B. L., 2019b, ApJ, 876, 155
Sahu N., Graham A. W., Davis B. L., 2019c, ApJ, 887, 10
Sahu N., Graham A. W., Davis B. L., 2020, ApJ, 903, 97
Sahu N., Graham A. W., Hon D. S. H., 2023, MNRAS, 518, 1352
Sani E., Marconi A., Hunt L. K., Risaliti G., 2011, MNRAS, 413, 1479
Sato-Polito G., Zaldarriaga M., Quataert E., 2025, preprint (arXiv:2501.09786)
Savorgnan G. A. D., Graham A. W., 2016, ApJS, 222, 10
Savorgnan G. A. D., Graham A. W., Marconi A., Sani E., 2016, ApJ, 817, 21
Schaye J. et al., 2015, MNRAS, 446, 521
Sersic J. L., 1968, Atlas de Galaxias Australes
Sesana A., Shankar F., Bernardi M., Sheth R. K., 2016, MNRAS, 463, L6
Shankar F., 2009, New A Rev., 53, 57
Shankar F., 2013, Class. Quantum Gravity, 30, 244001
Shankar F., Salucci P., Granato G. L., De Zotti G., Danese L., 2004, MNRAS, 354, 1020
Shankar F., Lapi A., Salucci P., De Zotti G., Danese L., 2006, ApJ, 643, 14
Shankar F., Weinberg D. H., Miralda-Escudé J., 2009a, ApJ, 690, 20
Shankar F., Bernardi M., Haiman Z., 2009b, ApJ, 694, 867
Shankar F., Marulli F., Mathur S., Bernardi M., Bournaud F., 2012, A&A, 540, A23
Shankar F. et al., 2016, MNRAS, 460, 3119
Shankar F., Bernardi M., Sheth R. K., 2017, MNRAS, 466, 4029
Shankar F. et al., 2019, MNRAS, 485, 1278
Shankar F. et al., 2020a, Nat. Astron., 4, 282
Shankar F. et al., 2020b, MNRAS, 493, 1500
Shen Y. et al., 2015, ApJ, 805, 96
Sheth R. K., Bernardi M., 2012, MNRAS, 422, 1825
Sicilia A. et al., 2022, ApJ, 934, 66
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Silk J., Rees M. J., 1998, A&A, 331, L1
Smola A. J., Schölkopf B., 2004, Stat. Comput., 14, 199
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Springel V., 2010, MNRAS, 401, 791
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2018, MNRAS, 475, 676
Suh H., Civano F., Trakhtenbrot B., Shankar F., Hasinger G., Sanders D. B., Allevato V., 2020, ApJ, 889, 32
Tanaka T. S. et al., 2025, ApJ, 979, 215
Terrazas B. A. et al., 2020, MNRAS, 493, 1888
Terrazas B. A., Aird J., Coil A. L., 2024, preprint (arXiv:2411.08838)
Thomas N., Davé R., Anglés-Alcázar D., Jarvis M., 2019, MNRAS, 487, 5764
Tipping M. E., 2001, J. Mach. Learn. Res., 1, 211
Tsuruoka Y., Tsujii J., Ananiadou S., 2009, in Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP. p. 477
Tundo E., Bernardi M., Hyde J. B., Sheth R. K., Pizzella A., 2007, ApJ, 663, 53
van den Bosch R. C. E., Gebhardt K., Gültekin K., Yıldırım A., Walsh J. L., 2015, ApJS, 218, 10
Villaescusa-Navarro F. et al., 2021, ApJ, 915, 71
Villaescusa-Navarro F. et al., 2023, ApJS, 265, 54
Yang G. et al., 2017, ApJ, 842, 72
Yu Q., Tremaine S., 2002, MNRAS, 335, 965
Zhang T., 2004, in Proceedings of the Twenty-first International Conference on Machine Learning. ICML, p. 116
Zhuang M.-Y., Ho L. C., 2023, Nat. Astron., 7, 1376
Zou F., Brandt W. N., Gallo E., Luo B., Ni Q., Xue Y., Yu Z., 2024, ApJ, 976, 6

APPENDIX A: INVESTIGATING THE PRESENCE OF A BIAS IN THE LOCAL SCALING RELATIONS OF SMBHS AND THEIR HOST GALAXIES

In this Appendix, we revisit the long-standing issue of possible selections biases affecting the SMBH-galaxy scaling relations in the local Universe (e.g. Batcheldor 2010). Uncovering such biases

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

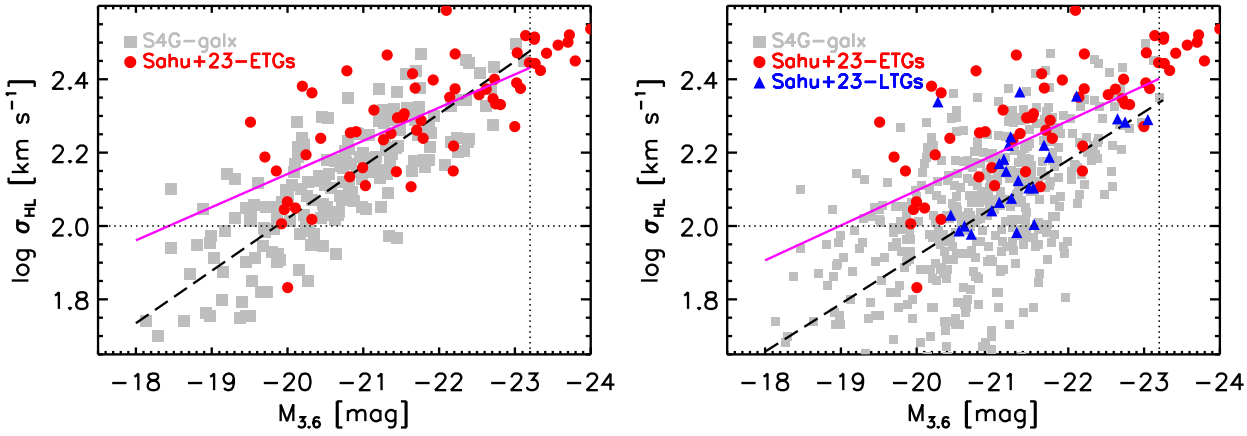

Figure A1. Scaling relation between the galaxy stellar velocity dispersion and galaxy magnitude at 3.6 μm for the local sample of ETGs from the S^{4}G sample with σ measurements from the Hyperleda data base (filled black circles and long-dashed lines), compared with the SMBH sample collected by Sahu et al. (2023) at 3.6 μm sample with *Spitzer* photometry. The left panel only includes ETG galaxies from Sahu et al. (2023) and ETGs from the S^{4}G sample with T-type < 1, while the right panel includes also the LTGs from Sahu et al. (2023) and T-type < 6.5 from the S^{4}G sample. The SMBH sample has a tendency to be characterized, on average, by larger stellar velocity dispersions at fixed luminosity.

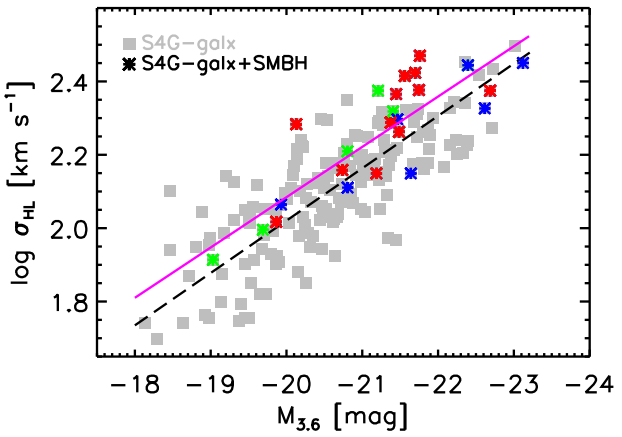

Figure A2. Scaling relation between the galaxy stellar velocity dispersion and galaxy magnitude at 3.6 μm for the local sample of ETGs from the S^{4}G sample with σ measurements from the Hyperleda data base (filled black circles and long-dashed lines). The coloured symbols mark the galaxies from the original samples adopted by Shankar et al. (2016) which are in common to both Hyperleda and the ETG S^{4}G samples, namely from Beifiori et al. (2012, green stars) and Savorgnan et al. (2016, red stars), and some additional galaxies from Sahu et al. (2023, blue stars). It is evident that the original sample adopted by Shankar et al. (2016) shows a sharp offset in stellar velocity dispersion at fixed magnitude, which is slightly reduced when including the new galaxies from Sahu et al. (2023, blue stars).

is vital to pin down the true shape and time evolution of SMBH–galaxy scaling relations (e.g. Shankar et al. 2016; Farrah et al. 2023). Shankar et al. (2016) also showed via extended Monte Carlo tests that selection effects are not expected to significantly impact the residual analysis, which therefore represents a powerful tool to extract the intrinsic correlations (in terms of slopes) between SMBH mass and galaxy properties.

Following the seminal papers by Bernardi et al. (2007), Gültekin et al. (2011), and Morabito & Dai (2012), Shankar et al. (2016) showed that several local samples of galaxies with dynamically measured masses of their central SMBHs tend to show, on average, higher stellar velocity dispersions at fixed galaxy stellar mass than what predicted by the mean distribution of galaxies in the SDSS galaxy sample. This result was derived by comparing the local sample of SDSS galaxies with four independent SMBH galaxy samples with distinct galactic photometries from Savorgnan et al. (2016), Beifiori et al. (2012), Läsker et al. (2014), and McConnell & Ma (2013). SDSS galaxy stellar masses were derived from Sérsic (Sersic 1968) plus exponential fits by Bernardi et al. (2014) and colour-dependent mass-to-light ratios from Bell et al. (2003). To infer host galaxy luminosities, Savorgnan et al. (2016) adopted 3.6 μm *Spitzer* images with Sérsic profiles plus, wherever relevant, additional components such as bars and rings. Shankar et al. (2016) also included in their analysis galaxies from the original McConnell & Ma (2013) sample with 3.6 μm luminosities derived from the Sérsic plus exponential fits by Sani et al. (2011). Beifiori et al. (2012) instead extracted homogeneous host galaxy luminosities from bulge-to-disc decompositions of SDSS i-band images, from which Shankar et al. (2016) derived stellar masses using the same colour-dependent mass-to-light ratios from Bell et al. (2003) self-consistently adopted for the comparison SDSS galaxy sample. Läsker et al. (2014) extracted galaxy K-band photometries from deep, high spatial resolution images obtained from the wide-field WIRCam imager at the Canada–France–Hawaii–Telescope, and luminosities were then converted to stellar masses by Shankar et al. (2016) using an average standard mass-to-light ratio.

Sahu et al. (2023) criticized the Shankar et al. (2016) result of a systematic bias between SMBH and galaxy samples reducing it to a simple byproduct of systematic discrepancies in the stellar mass scales between SDSS galaxies and the SMBH galaxy samples, rather than induced by offsets in stellar velocity dispersion at fixed host galaxy stellar mass. To support their claim, Sahu et al. (2023) focused specifically on the local sample of SMBHs by Savorgnan et al. (2016) at 3.6 μm, further enriched by Sahu et al. (2019b), i.e. the one also adopted in this work, and ignored the other local samples of SMBHs with different photometries included in Shankar et al. (2016) and summarized above. To estimate the mean stellar mass correction between the stellar masses derived from 3.6 μm photometry and SDSS imaging, Sahu et al. (2023) followed two steps. First, they took a sample of 43 galaxies in common between SDSS and the S^{4}G local sample of galaxies with *Spitzer* photometry, and estimated a mean discrepancy of ~ 0.2 dex between the two stellar mass systems. The SDSS galaxy subsample they chose for this step had total galaxy

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

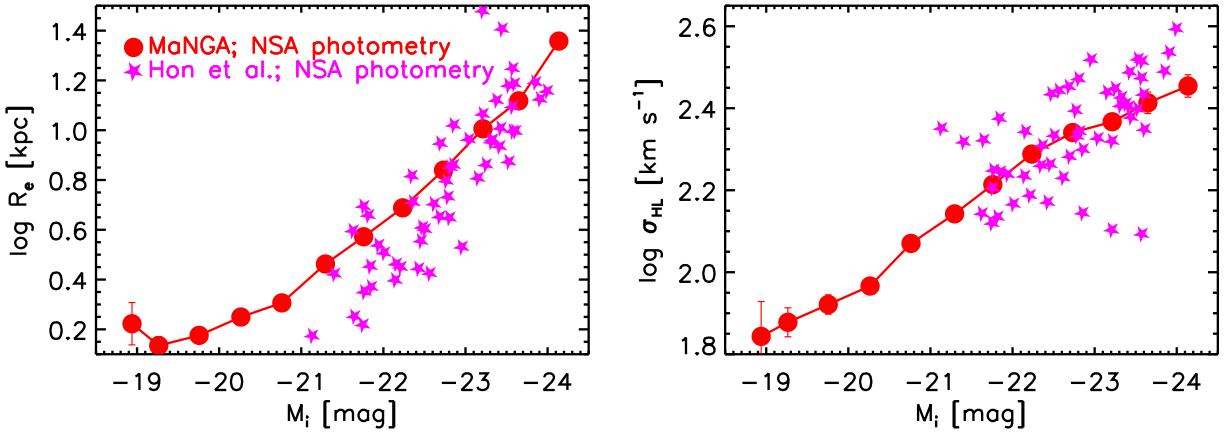

Figure A3. *Left panel*: Mean scaling between effective radius and i-band magnitude extracted from the local MaNGA sample with NSA photometry (filled, red circles) against the Hon et al. (2022) data sets of local compact galaxies with the same NSA photometry (magenta stars). *Right panel*: Same format as the left panel but in the stellar velocity dispersion (rescaled to the Hyperleda aperture) versus i-band magnitude plane. The Hon et al. (2022) galaxies tend to be more compact than the average MaNGA galaxies and with slightly larger stellar velocity dispersion, in particular at bright luminosities.

luminosities derived by Hon et al. (2022) from the SDSS i band. Sahu et al. (2023) then checked that the 37 galaxies in common between the S^{4}G sample and the Sahu et al. (2019b) SMBH samples have very similar photometry, and thus the same stellar mass correction between SDSS and S^{4}G could be applied between SDSS and the Sahu et al. (2019b) SMBH sample, an offset which is sufficient to remove any apparent systematic offset ('bias') between the two σ–$M_{\rm gal}$ relations followed by the SMBH galaxy and SDSS samples.

The methodology followed by Sahu et al. (2023) makes the fundamental and untested assumption that the S^{4}G and local SMBH samples follow the same identical stellar velocity dispersion-3.6 μm luminosity relation, and thus the same stellar mass correction can be applied to both the S^{4}G and SMBH samples when comparing with SDSS galaxies. However, we have verified, that this is not the case. We have cross-correlated the S^{4}G sample with the Hyperleda data base and the resulting sample is reported in Fig. A1 with grey, filled squares. The left panel only includes S^{4}G ETGs with T-type ≤ 1 to include S0/a and Ellipticals, as in the reference ETG sample from Sahu et al. (2023). It is clear that the ETG SMBH sample does *not* strictly follow the distribution of the S^{4}G sample. The vast majority of ETGs below $M_i \lesssim -22$ (approximately $M_{\rm gal} \sim 10^{11}\,{\rm M}_\odot$, using the Sahu et al. (2023) mass-to-light ratio at 3.6 μm of ${\rm M}_\odot/{\rm L}_\odot = 0.6$) lie above the mean relation traced by the S^{4}G sample. To remark this point, we show two linear fits to the Sahu et al. (2023) and S^{4}G samples (solid magenta and long-dashed black line, respectively), where we have assumed the same statistical errors on magnitude and $\log\sigma$ of 0.2 mag and 0.1 dex, respectively. The ETG SMBH sample tends to be skewed towards larger stellar velocity dispersions, in particular there are virtually no galaxies with SMBH masses below 100 km s^{-1} (horizontal dotted lines).

The discrepancy between the S^{4}G and the SMBH samples becomes even more pronounced in the right panel of Fig. A1 where we include all the LTGs of Sahu et al. (2023) and all the corresponding S^{4}G galaxies matched in morphology with T-type < 6.5, e.g. down to Sc galaxies. The two samples reach an average offset of up to ~ 0.2 dex in $\log\sigma$ at fixed luminosity. In addition, the S^{4}G sample with velocity dispersion measurements in the Hyperleda database is around 1/4 of the original sample, and thus it could suffer from incompleteness. We stress that the photometric systems and definitions of total host galaxy luminosity adopted in the S^{4}G and the Sahu et al. (2023) samples are not strictly identical and could further bias the comparison. A dedicated study tailored at applying the same measurement techniques and assumptions in a (complete) reference galaxy sample and the local SMBH sample is required to truly advance in our understanding of systematic biases in the SMBH scaling relations traced by the local sample of dynamically measured SMBHs.

We also note that the subsample of SDSS galaxies chosen by Sahu et al. (2023) to calibrate the galaxy stellar mass offset with the S^{4}G local sample is not ideal, being itself biased with respect to the full galaxy population as described by the MaNGA survey, being more compact at fixed host galaxy luminosity, and with a tendency to have larger stellar velocity dispersion, especially at $M_i \lesssim -22.5$, as seen in Fig. A3.

It is true that with an increased size in the SMBH sample and more accurate photometry the offset in stellar velocity dispersion between the SMBH sample compared to the galaxy sample is somewhat reduced, as also noted by Sahu et al. (2023). None the less, the offset is still noticeable at lower luminosities, approaching $\Delta\log\sigma \sim 0.1$ at $M_i \lesssim -22$ when comparing the two full samples, and even larger at fainter luminosities. Note that an offset of $\gtrsim 0.2$ dex in $\log\sigma$ at fixed luminosity would correspond to a noticeable average offset in SMBH mass of a factor of ~ 5, if $M_{\rm bh} \propto \sigma^{3.6}$, as inferred from the residual analysis presented in the top left panel of Fig. 2. For completeness, Fig. A2 compares the S^{4}G ETG-Hyperleda sample with the fraction of inactive ETGs with SMBHs in common to both S^{4}G and Hyperleda from the original SMBH sample collected by Shankar et al. (2016, red stars), plus some galaxies from the SMBH sample of Beifiori et al. (2012, green stars) and some additional galaxies with SMBH from Sahu et al. (2023, blue stars). It is clear that, overall, the original sample adopted by Shankar et al. (2016), from Savorgnan et al. (2016), is clearly biased high in stellar velocity dispersion at fixed luminosity. The addition of new galaxies from Sahu et al. (2023) tends to lower the mean offset compared to the original sample adopted by Shankar et al. (2016).

We conclude that a bias between the local SMBH sample and the larger sample of galaxies without dynamical SMBH mass measurements, persists in the form of an offset in stellar velocity dispersion at fixed host galaxy luminosity, which is thus not a byproduct of different stellar mass-to-light ratios between the two samples. Such an offset in stellar velocity dispersion at fixed luminosity was also identified by Bernardi et al. (2007), van den Bosch et al. (2015),

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

and, more recently, by Kormendy (2020), who reported in their fig. 2 a clear offset in σ at fixed L_V for both core and coreless galaxies. In addition, many local AGN appear to sit significantly below the scaling relations of dormant SMBHs, as discussed by, e.g. Reines & Volonteri (2015a) and Shankar et al. (2019), when adopting reasonable assumptions for the virial factors, and it is evident even among AGN samples calibrated with the same photometry as in the comparison galaxy sample. This result appears common to all local AGN independently of their host galaxy morphology, although some dependence of the offset on the Eddington ratio may be present (e.g. Farrah et al. 2023).

More recently, Byrne et al. (2023) adopted an alternative approach to probe the existence of a selection bias in the SMBH mass-host galaxy stellar mass relation by computing the SMBH masses in a mass-complete sample of 18 ETGs from the Virgo cluster. They were able to extract SMBH masses in 11 out of the 18 galaxies, and thus, on the assumption that the remaining galaxies with undetected SMBHs do not contain a SMBH, they were able to place a conservative lower limit to the mean SMBH mass in their sample by dividing the sum of all their detected SMBH masses by 18 (instead of 11). They claimed a mean lower limit of $M_{\rm bh} = 3.7 \times 10^7\,{\rm M}_\odot$ for host galaxies with mean total stellar mass $M_{\rm gal} = (1.8 \pm 1.1) \times 10^{10}\,{\rm M}_\odot$. At face value, this value of the mean SMBH mass would be significantly higher than what predicted by, e.g. Model I in Shankar et al. (2016, their equation 6). However, there are several caveats to be considered here before drawing any definitive conclusion. First off, the results by Shankar et al. (2016) were based on Monte Carlo simulations applied to thousands of galaxies from the SDSS survey, and thus one would need to check the consistency between the (small) galaxy sample from Byrne et al. (2023) with the much larger SDSS galaxy sample, in particular on the σ–$M_{\rm gal}$ plane. We verified that, when cross-correlating the 18 galaxies in the Byrne et al. (2023) sample with the Hyperleda database, the latter yields an average $\log \sigma_{\rm HL} \sim 2.02$, which, at face value, would be tentatively ~ 0.1 dex higher than the mean $\log \sigma_{\rm HL}$ calculated by Shankar et al. (2016) at an average stellar mass of $M_{\rm gal} = 1.8 \times 10^{10}\,{\rm M}_\odot$. However, possible systematic differences between the ATLAS dynamical mass measurements and the SDSS-based stellar masses from Shankar et al. (2016) prevent a robust comparison between two samples. Secondly, several other indicators and/or independent measurements of SMBH masses tend to point to lower masses, as also highlighted by Byrne et al. (2023) in their fig. 5, and as also stressed by Shankar et al. (2019). Last but not least, the Monte Carlo method put forward by Shankar et al. (2016) did not include the larger SMBH sample by Sahu et al. (2023) which, as we discussed above, tends to show a lower degree of bias, comparable to $\Delta \log \sigma \sim 0.1$ dex at the mass scale probed by Byrne et al. (2023), as suggested by the left panel of Fig. A1.

APPENDIX B: ADDITIONAL RELEVANT PAIRWISE RESIDUAL CORRELATIONS

Here, we present additional residuals derived for the Saglia et al. (2016) SMBH sample, with accurate measurements of the stellar velocity dispersion within the effective radius and the stellar bulge component (Fig. B1), along with additional residuals extracted from the Sahu et al. (2023) SMBH sample as a function of half-light effective radius and Sérsic index.

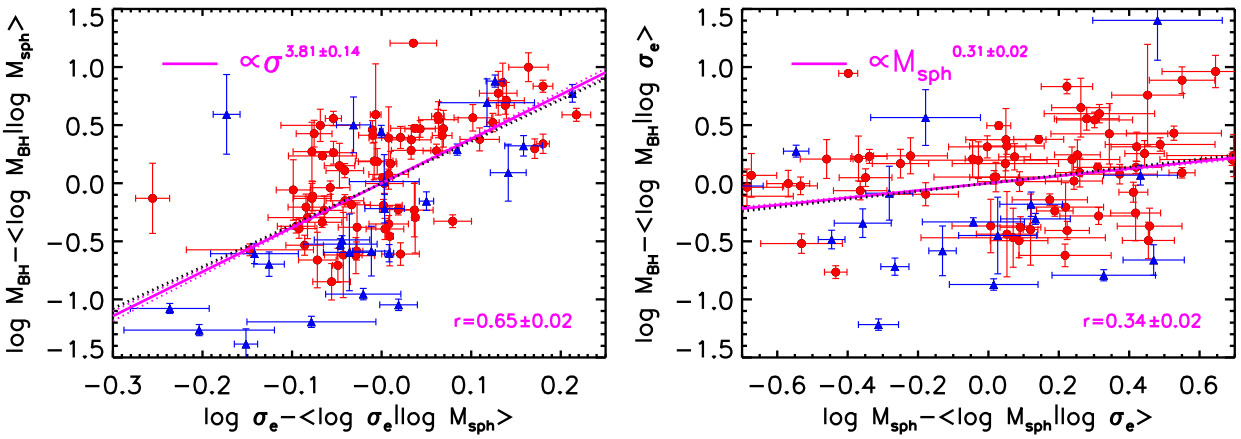

Figure B1. Pairwise residuals as a function of stellar velocity dispersion at fixed spheroidal mass (left) and vice versa (right) for the Saglia et al. (2016) SMBH sample which only considers bulge stellar mass and central velocity dispersions measured at the half-light radius. Even in this sample, we still find a stronger dependence on stellar velocity dispersion confirming the results retrieved from the Sahu et al. (2023) sample when considering only the bulge component (Fig. 3).

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

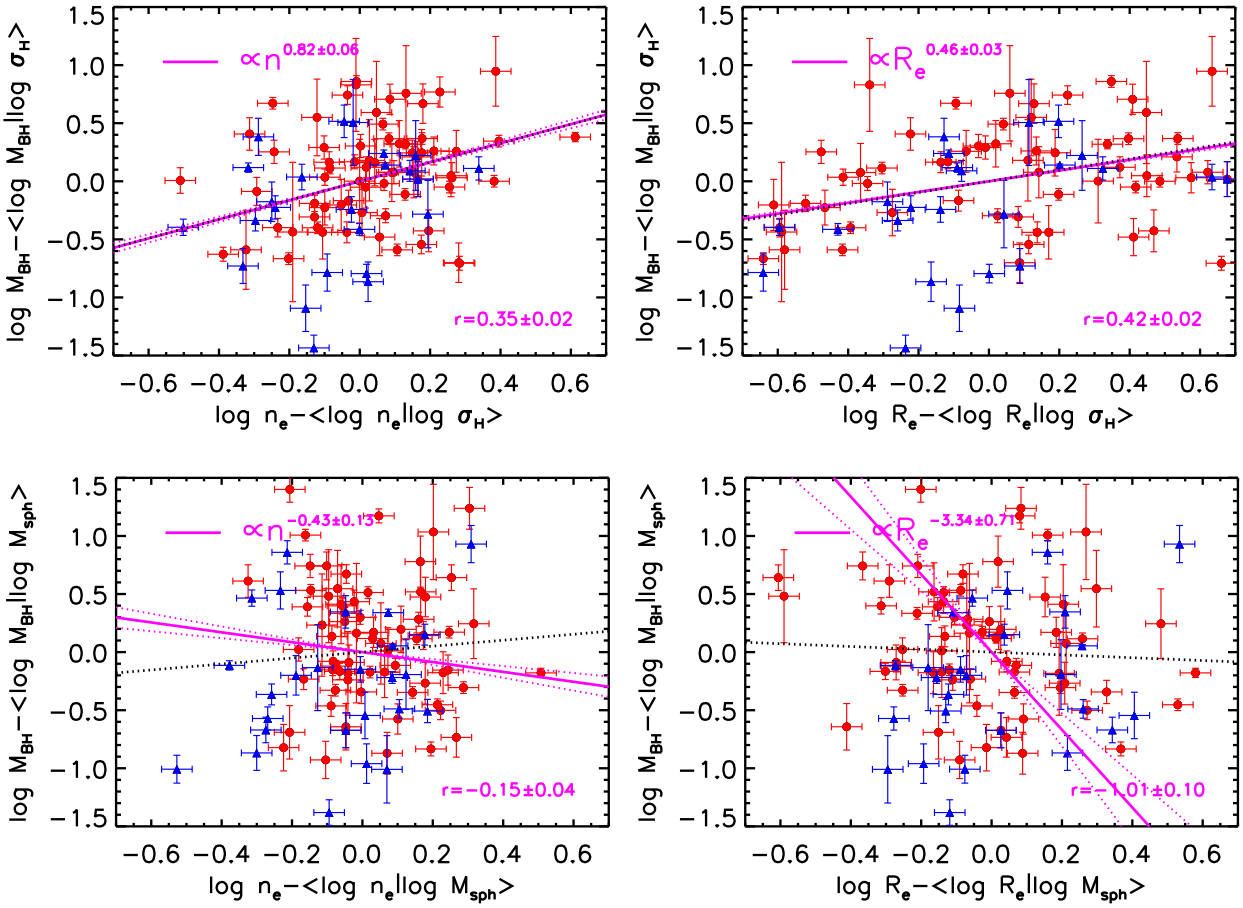

Figure B2. Pairwise residuals as a function of Sérsic index n (left) and bulge effective radius R_e (right) at fixed stellar velocity dispersion (top) and spheroidal mass (bottom) for the Sahu et al. (2023) SMBH sample. We do not find any significant residual dependence of SMBH mass on any of these quantities when the underlying dependence on stellar velocity dispersion and/or spheroidal (bulge) mass is subtracted.

APPENDIX C: DATA FOR THE SMBH SAMPLE ADOPTED IN THIS WORK

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

Table C1. Data for the 69 ETGs. σ and $\delta\sigma$ have units of km s^{-1}. All magnitudes are AB at 3.6 μm. Both effective radius and Sérsic index are the 'equivalent' measures from Sahu et al. (2020).

Name	$\log\left(\frac{M_{\rm BH}}{M_\odot}\right)$	$\delta\log\left(\frac{M_{\rm BH}}{M_\odot}\right)$	$\log\left(\frac{M_{\rm sph}}{M_\odot}\right)$	σ	$\delta\sigma$	$R_e('')$	n	Distance (Mpc)	Mag
IC4296	9.04	0.09	11.47	327.3	5.4	41.1	3.82	40.7	−23.3
IC1459	9.38	0.20	11.55	295.8	6.4	57.3	7.00	28.4	−23.4
NGC0404	4.85	0.13	7.96	34.6	3.1	3.89	0.90	3.1	−17.3
NGC0524	8.92	0.10	10.57	236.6	4.5	8.35	2.16	23.3	−22.2
NGC0821	7.59	0.17	10.69	197.7	2.8	18.9	6.10	23.4	−21.5
NGC1023	7.62	0.05	10.21	197.2	4.6	7.4	2.00	11.1	−21.5
NGC1275	8.90	0.20	11.84	244.9	12.8	53.6	4.31	72.9	−24.2
NGC1332	9.16	0.07	11.05	294.4	11.3	18.0	3.70	22.3	−22.2
NGC1374	8.76	0.05	10.22	179.5	3.2	11.74	1.65	19.2	−20.8
NGC1399	8.67	0.06	11.66	331.9	5.3	338.1	10.00	19.4	−23.7
NGC1407	9.65	0.08	11.46	265.5	5.1	47.29	3.89	28	−23.3
NGC1600	10.23	0.05	11.82	331.1	7.0	49.58	5.08	64	−24.1
NGC2549	7.15	0.60	9.59	141.3	2.7	3.1	1.50	12.3	−19.9
NGC2778	7.18	0.34	9.41	154.2	3.2	2.2	1.20	22.3	−19.7
NGC2787	7.60	0.06	9.13	191.9	3.9	2.88	1.27	7.3	−19.5
NGC3091	9.56	0.04	11.61	311.2	7.7	51.2	6.60	51.2	−23.6
NGC3115	8.94	0.25	10.77	260.0	3.0	34.4	5.10	9.4	−21.6
NGC3245	8.30	0.12	10.06	207.0	7.3	2.4	1.70	20.3	−21.1
NGC3377	7.89	0.04	10.48	136.1	2.3	91.7	9.20	10.9	−20.8
NGC3379	8.60	0.12	10.8	202.3	1.8	50.9	5.30	10.3	−21.5
NGC3384	7.23	0.05	10.06	144.2	2.5	5.6	1.80	11.3	−21
NGC3414	8.38	0.06	10.83	237.7	8.1	25.5	4.50	24.5	−21.7
NGC3489	6.76	0.07	9.54	104.2	2.0	1.7	1.30	11.7	−20.3
NGC3585	8.49	0.13	11.3	214.3	5.1	86.3	6.30	19.5	−22.8
NGC3607	8.11	0.18	11.23	222.3	4.1	65.5	5.60	22.2	−22.7
NGC3608	8.30	0.18	10.89	194.1	4.2	43.4	5.70	22.3	−21.8
NGC3665	8.76	0.10	11.03	215.3	8.5	12.78	2.74	34.7	−22.7
NGC3842	9.99	0.13	11.92	308.3	6.7	73.6	8.20	98.4	−24.4
NGC3923	9.45	0.13	11.4	245.5	4.9	78.78	4.77	20.9	−23
NGC3998	8.91	0.11	10.02	264.9	11.0	4.8	1.30	13.7	−20.8
NGC4026	8.26	0.11	10.11	173.4	3.8	2.35	3.98	13.2	−20.4
NGC4261	8.70	0.09	11.38	296.5	4.3	47.3	4.30	30.8	−23
NGC4291	8.52	0.05	10.71	292.4	6.9	15.4	5.90	25.5	−21.3
NGC4339	7.63	0.33	9.67	110.9	3.1	6.42	1.40	16.0	−20
NGC4342	8.65	0.18	9.94	240.4	5.7	4.69	3.99	23.0	−20.2
NGC4350	8.86	0.41	10.28	180.7	4.4	19.45	3.97	16.8	−20.9
NGC4371	6.85	0.08	9.89	128.8	2.2	8.9	3.19	16.9	−21
NGC4374	8.95	0.05	11.49	277.3	2.4	129.8	7.90	17.9	−23.3
NGC4429	8.18	0.09	10.46	173.4	5.4	11.29	2.31	16.5	−21.8
NGC4434	7.85	0.17	9.91	116.4	2.8	5.31	2.93	22.4	−20
NGC4459	7.83	0.09	10.48	171.8	4.8	13.0	2.60	15.7	−21.3
NGC4472	9.40	0.05	11.7	281.8	2.9	135.3	5.40	17.1	−23.8
NGC4473	8.08	0.36	10.64	178.6	2.5	36.9	2.90	15.3	−21.4
NGC4486	9.81	0.05	11.49	322.8	4.3	87.1	5.90	16.8	−23.3
NGC4526	8.67	0.05	10.7	224.4	9.4	14.88	2.96	16.9	−22.1
NGC4552	8.67	0.05	10.88	250.0	2.9	71.5	5.36	14.9	−21.9
NGC4564	7.78	0.06	10.01	156.3	2.2	6.0	3.00	14.6	−20.2
NGC4578	7.28	0.35	9.77	111.9	4.1	6.32	1.99	16.3	−20.1
NGC4596	7.90	0.20	10.18	140.6	2.2	9.0	3.00	17.0	−21.4
NGC4621	8.59	0.05	11.16	228.0	3.8	90.9	8.80	17.8	−22.5
NGC4649	9.67	0.10	11.44	330.4	4.6	80.59	5.21	16.4	−23.1
NGC4697	8.26	0.05	11.01	165.2	1.6	226.4	6.70	11.4	−22.2
NGC4742	7.15	0.18	9.87	101.4	3.4	3.41	3.20	15.5	−19.9
NGC4762	7.36	0.15	9.97	141.3	4.1	2.24	1.85	22.6	−22.2
NGC4889	10.32	0.44	12.14	392.6	5.3	60.8	6.80	103.2	−24.9
NGC5077	8.87	0.22	11.28	251.2	5.5	23.0	5.70	41.2	−22.7
NGC5252	9.00	0.40	10.85	186.6	26.5	1.47	2.95	96.8	−23
NGC5419	9.86	0.14	11.45	344.3	5.4	16.83	2.62	56.2	−24
NGC5576	8.20	0.10	10.87	182.4	6.0	49.3	3.70	24.8	−21.7
NGC5813	8.83	0.06	10.86	236.0	3.4	14.16	3.65	31.3	−22.6
NGC5845	8.41	0.22	10.12	230.7	7.9	5.29	3.27	25.2	−20.3
NGC5846	9.04	0.05	11.42	237.1	3.5	83.4	5.70	24.2	−23.1

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

Table C1 – *continued*

Name	$\log\left(\frac{M_{\rm BH}}{\rm M_\odot}\right)$	$\delta\log\left(\frac{M_{\rm BH}}{\rm M_\odot}\right)$	$\log\left(\frac{M_{\rm sph}}{\rm M_\odot}\right)$	σ	$\delta\sigma$	$R_e('')$	n	Distance (Mpc)	Mag
NGC6251	8.77	0.16	11.82	312.6	18.2	30.1	5.60	104.6	−24.1
NGC6861	9.30	0.08	10.94	387.3	16.5	20.13	3.52	27.3	−22.1
NGC7052	8.57	0.23	11.46	278.6	11.8	20.04	3.46	66.4	−23.2
NGC7332	7.11	0.20	10.22	127.9	3.3	2.43	2.15	24.9	−21.6
NGC7457	7.00	0.30	9.40	67.9	3.5	6.51	2.84	14	−20
NGC7619	9.40	0.09	11.64	317.0	4.9	58.0	5.20	51.5	−23.7
NGC7768	9.11	0.15	11.89	289.7	11.9	42.1	6.70	112.8	−24.2

Table C2. Data for the 26 LTGs. σ and $\delta\sigma$ have units of $\rm km\,s^{-1}$. All magnitudes are AB at 3.6 μm. Both effective radius and Sérsic index are the 'equivalent' measures from Sahu et al. (2020).

Name	$\log\left(\frac{\rm M_{BH}}{\rm M_\odot}\right)$	$\delta\log\left(\frac{\rm M_{BH}}{\rm M_\odot}\right)$	$\log\left(\frac{\rm M_{sph}}{\rm M_\odot}\right)$	σ	$\delta\sigma$	$R_e('')$	n	Distance (Mpc)	Mag
Circinus	6.25	0.11	10.12	148.0	18.0	23.13	1.80	4.2	−21.1
IC2560	6.49	0.20	9.63	141.0	10.0	3.92	1.63	31.0	−21.2
NGC0224	8.15	0.16	10.11	154.0	4.0	173.6	1.30	0.8	−21.8
NGC0253	7.00	0.30	9.76	96.0	18.0	27.89	2.33	3.5	−21.3
NGC1097	8.38	0.04	10.83	195.0	5.0	11.39	1.52	24.9	−23.1
NGC1300	7.71	0.16	9.42	218.0	29.0	7.39	2.83	14.5	−20.3
NGC1320	6.78	0.29	10.25	110.0	10.0	2.23	2.87	37.7	−21
NGC1398	8.03	0.11	10.57	196.0	18.0	10.38	3.00	24.8	−22.6
NGC2960	7.06	0.17	10.44	166.0	16.0	2.19	2.86	71.1	−21.7
NGC2974	8.23	0.07	10.23	232.0	4.0	6.53	1.17	21.5	−21.4
NGC3031	7.83	0.09	10.16	152.0	2.0	42.98	3.46	3.5	−21.1
NGC3079	6.38	0.12	9.92	175.0	12.0	4.35	0.58	16.5	−21.2
NGC3227	7.88	0.14	10.04	127.0	6.0	8.34	1.90	21.1	−21.5
NGC3368	6.89	0.09	9.81	119.0	4.0	4.83	1.00	10.7	−21.2
NGC3627	6.95	0.05	9.74	127.0	6.0	3.92	2.10	10.6	−21.5
NGC4151	7.68	0.37	10.27	116.0	3.0	6.0	1.85	19.0	−21.1
NGC4258	7.60	0.01	10.05	133.0	7.0	26.4	2.60	7.6	−21.3
NGC4303	6.58	0.17	9.42	95.0	8.0	2.16	0.90	12.3	−20.7
NGC4388	6.90	0.11	10.07	100.0	10.0	14.3	1.15	17.8	−20.6
NGC4501	7.13	0.08	10.11	166.0	7.0	20.35	2.83	11.2	−21.2
NGC4594	8.81	0.03	10.81	226.0	3.0	41.36	4.24	9.6	−22.1
NGC4699	8.34	0.10	11.12	192.0	9.0	29.75	6.77	23.7	−22.8
NGC4736	6.78	0.10	9.89	107.0	4.0	9.65	1.03	4.4	−20.4
NGC4826	6.07	0.15	9.55	97.0	6.0	11.93	0.76	5.6	−20.6
NGC5055	8.94	0.10	10.49	101.0	3.0	43.52	1.76	8.9	−21.6
NGC7582	7.67	0.09	10.15	147.0	19.0	4.55	2.21	19.9	−21.5

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/541/2/2070/8126743 by guest on 28 July 2025

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.