Astronomy & Astrophysics manuscript no. aanda
December 18, 2025

Doubling NIRSpec/IFS capability to calibrate the single epoch black hole mass relation at high redshift

Eleonora Parlanti[1,⋆], Bartolomeo Trefoloni[1,2], Stefano Carniani[1], Francesco D'Eugenio[4,5], Michele Perna[3], Giulia Tozzi[6], Hannah Übler[6], Giacomo Venturi[1], and Sandra Zamora[1]

[1] Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy
[2] INAF – Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, I-50125 Firenze, Italy
[3] Centro de Astrobiología (CAB), CSIC-INTA, Ctra. de Ajalvir km 4, Torrejón de Ardoz, E-28850, Madrid, Spain
[4] Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK
[5] Cavendish Laboratory - Astrophysics Group, University of Cambridge, 19 JJ Thomson Avenue, Cambridge, CB3 0HE, UK
[6] Max-Planck-Institut für extraterrestrische Physik (MPE), Gießenbachstraße 1, 85748 Garching, Germany

ABSTRACT

The recent discovery of a large population of overmassive black holes (BHs) in the early Universe allowed by *JWST* challenges the validity of the BH–host galaxy coevolution framework. However, the reliability of the estimated BH masses is being questioned, as these are typically derived using single-epoch (SE) mass relations calibrated locally. Calibrating SE relations directly in the early Universe would therefore enable more accurate BH mass estimates and allow us to identify potential biases.

In this work, we release a data-reduction technique for JWST/NIRSpec IFU observations that doubles the effective wavelength coverage, enabling detection of otherwise inaccessible emission. Therefore, whenever adjacent dispersers are required, observers should carefully evaluate the trade-off between integrating longer in the bluer configuration alone versus distributing the exposure time across two dispersers.

We apply this new pipeline to observations of a sample of five quasars at $z \sim 2$ with BH masses independently measured through reverberation mapping (RM) campaigns, enabling a joint analysis of both Hβ and Hα; the latter being located well beyond the NIRSpec nominal wavelength range.

We assess the reliability of the most widely adopted SE calibrations, finding that Hβ yields the closest agreement with the BH masses estimated from RM, whereas Hα-based estimators exhibit a substantially larger scatter (~0.5 dex). For the least massive BH in our sample ($M_{\rm BH,RM} \sim 10^{7.5} M_\odot$), which is accreting at a rate close to the Eddington limit ($\lambda_{\rm Edd}$ = 0.8), all SE calibrators return a BH mass that is an order of magnitude larger than the RM estimate. This discrepancy may indicate a systematic overestimation of BH masses for highly accreting BHs at high redshift.

Finally, using our new measurements together with GRAVITY dynamical masses, we provide the first high-redshift calibrations of SE BH mass estimators based on Hα and Hβ. Although a larger and broader BH-mass sample is needed to further reduce the parameter uncertainties, our calibration can already be applied to the newly discovered JWST BH population in the early Universe.

Key words. Galaxies: high-redshift, quasars: supermassive black holes, quasars: emission lines

1. Introduction

The majority of massive galaxies, both at high redshift and in the local Universe, are thought to host supermassive black holes (BHs, with masses exceeding $10^6 - 10^9 M_\odot$) at their centers (Magorrian et al. 1998). The mass of the BH ($M_{\rm BH}$) is found to correlate with several properties of the host galaxy (such as the stellar mass of the bulge or that of the entire galaxy, or the stellar velocity dispersion, Kormendy & Ho 2013), suggesting a common evolutionary scenario between the assembly of galaxies and the growth of their BHs. Despite this widely accepted coevolution framework between the BH and galaxy growth (Heckman & Best 2014), supported by the above tight correlations observed in the local Universe, new questions have recently started to arise. The advent of the *James Webb* Space Telescope (JWST, Gardner et al. 2023) has allowed the discovery of a large number of BHs that appear to deviate from the local scaling relations. In particular, the majority of the newly discovered BHs appear to be overmassive with respect to their host galaxies (Harikane et al. 2023; Matthee et al. 2023; Übler et al. 2023; Maiolino et al. 2024), with ratios of $M_\star/M_{\rm BH}$ approaching unity (Kokorev et al. 2023; Juodžbalis et al. 2024), while in the local Universe the typical ratio is ~0.001 (e.g., Kormendy & Ho 2013; Reines & Volonteri 2015).

To explain these new objects' peculiar properties, a large variety of different evolutionary scenarios and models have been proposed. If these objects are actually overmassive, heavy BH seeds (Natarajan et al. 2024; Cenci & Habouzit 2025), or super-Eddington accretion rates (Trinca et al. 2024; King 2025) are required. Despite the puzzling nature of these sources, which still calls for a comprehensive physical description, their overmassive nature should still be carefully tested. In particular, while the stellar masses of these objects suffer from large uncertainties, which depend on the fraction of light attributed to stellar emission (e.g., Wang et al. 2024; Akins et al. 2025; Leung et al. 2025), another major concern is that the BH mass could be overestimated (e.g., Lambrides et al. 2024; Naidu et al. 2025) due to the questionable applicability of typical calibrations in this new and elusive class of objects. In particular, the typical rela-

⋆ eleonora.parlanti@sns.it

tions used to infer M_{BH} are calibrated on local AGN, and then extrapolated to the high-z and high-luminosity regimes, where their applicability is still a matter of debate (e.g., Bertemes et al. 2025; Bosman et al. 2025) and featuring typical uncertainties of the order of $\sim 0.4-0.5$ dex (Shen 2013; Shen et al. 2024). The discovery of this larger population of overmassive BHs in the early Universe made by JWST has raised questions about the validity of these relations, highlighting the need for more accurate BH mass measurements at high redshifts. Such measurements are crucial not only to understand how galaxies and their BHs co-evolve but also to shed light on the formation of supermassive BHs at high redshift, providing key insights into the nature of BH seeds and their initial masses.

For local BHs, the mass can be robustly measured through kinematic methods that require resolving the sphere of influence of the BH, such as exploiting stellar kinematics (Gebhardt et al. 2000; Nguyen et al. 2025). Alternatively, we can use the detection of emission lines arising from the clouds in the broad line region (BLR), which are photoionized by the radiation from the accretion disk. Under the assumption of a virialized BLR, the BH mass can be estimated as:

$$M_{\rm BH} = \frac{fV^2R}{G} \quad (1)$$

where V is the rotational velocity of the BLR clouds, and R is the distance between the BLR and the ionizing source. Yet, both the geometry of the BLR as well as the inclination of the line of sight are, in most cases, unconstrained. In order to account for this, the f factor encapsulates the BLR geometry, kinematics, and inclination of the line of sight. Regarding the rotational velocity (V), this can be estimated by employing the width of some broad permitted line. In particular, different line width measurements have been explored in the literature, such as the full width half maximum (FWHM) or the line velocity dispersion (see e.g., Peterson et al. 2004), each of them having practical strengths and weaknesses.

The FWHM of the broad emission lines can be directly measured in observations, while the radius of the BLR can be estimated using reverberation mapping (RM) campaigns (see e.g., Dalla Bontà et al. 2020, and references therein). In particular, the measured time lag between variations in continuum flux and variations in broad-line luminosity can be modelled, assuming a given distribution of the BLR clouds, to derive the BLR radius. The RM approach requires accurate and frequent measurements of the luminosity of the BLR and the continuum in a span of months, which at high redshift can become decades in the observed frame due to cosmological time dilation. Hence, obtaining high-z measurements of the BLR sizes and then RM-derived BH masses can be achieved only through long-term monitoring campaigns (Shen et al. 2024; McDougall et al. 2025). Whilst RM campaigns are resource-intensive, they also provide us with a relation between the radius of the BLR (R) and the monochromatic luminosity at some wavelengths (e.g. 5100 Å), therefore supplying us with a robust proxy for the BLR radius ($R-L$ relation; Kaspi et al. 2000; Bentz et al. 2009). This discovery has enabled the so-called single epoch (SE) virial black hole mass calibrations, where a single spectrum can provide both the quantities needed to compute $M_{\rm BH}$, that is, the FWHM of the line and the continuum luminosity, respectively representative of the rotational velocity (V) and the BLR radius (R).

Other attempts to obtain independent measures of the BH mass have been made with the near-IR GRAVITY interferometer (GRAVITY Collaboration et al. 2017), which can spatially resolve the BLR and measure its kinematic properties. The modelling of the BLR kinematics returns a dynamical BH mass (e.g., Gravity Collaboration et al. 2018; GRAVITY Collaboration et al. 2020, 2024). Yet, this technique is currently limited to only very luminous Quasars (QSOs) with a nearby star to allow for AO correction (Gravity+ Collaboration et al. 2022).

All the techniques mentioned above to obtain a direct measure of the M_{BH}, in recent years, pushed their frontier up to the high-z Universe (Abuter et al. 2024; Shen et al. 2024; Liao et al. 2025; GRAVITY+ Collaboration et al. 2025), but the $z > 4$ Universe has so far still been inaccessible with these methods. Thus, to derive the BH masses at $z > 4$, we can only rely on SE virial BH mass estimators. With this approach, M_{BH} is derived using a single spectroscopic measurement, yielding the width of a broad line and the continuum (or the line itself) luminosity.

In this work, we aim to test the reliability of SE BH mass estimators in the high-z Universe by directly comparing them with robust RM-based measurements. We focus on a sample of five quasars at z~2 with RM BH masses (Shen et al. 2024), and exploit the unprecedented sensitivity and spatial resolution of JWST/NIRSpec IFU data to obtain high-quality measurements of both the Hβ and Hα broad emission lines. This allows us to perform a direct comparison between two different SE estimators and assess their consistency at high redshift. To achieve this goal, we implement a new data-reduction strategy for the medium-resolution gratings that effectively extends the usable wavelength coverage to longer wavelengths. With this innovative method, we recover the Hα emission line, using archival data whose spectroscopic coverage, based on standard reduction of JWST data, was limited to Hβ and [OIII]emission.

The paper is structured as follows. In Sect. 2 we present the methods to perform the new data reduction. In Sect. 3, we describe the data analysis of the sample of five QSOs. In Sects. 4 and 5, we estimate the BH masses using some of the most widely used SE estimators, and we derive our own relation calibrated at $z \sim 2$, respectively. We draw our conclusion in Sect. 6. Throughout this work, we adopt the cosmological parameters from Planck Collaboration et al. (2016): $H_0 = 67.7$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.307$, and $\Omega_\Lambda = 0.691$.

2. Changes to the standard data reduction

Data from NIRSpec/IFS observed with medium-resolution (G140M/F070LP, G140M/F100LP, G235M/F170LP, G395M/F290LP), with resolving power $R = \lambda/\Delta\lambda$ varying between ~ 700 and ~ 1300 and an average of R~1000, (Jakobsen et al. 2022) occupy just the first detector, namely 'NRS1', while the second detector is not used in the standard reduction process. However, we notice emission lines in the second detector and continuum emission until the red end of 'NRS2', as shown in Fig. 1, displaying the count-rate maps downloaded from the Mikulski Archive for Space Telescopes (MAST) of the G140M/F100LP observations from PID: 2057, which targeted a QSO at $z \sim 2.6$. The count-rate maps reveal the presence of the [OIII] and Hβ emission lines in the nominal wavelength range, but also the Hα emission in the second detector beyond the nominal wavelength range for G140M/F100LP observations.

In this work, we extend the wavelength range of medium-resolution gratings to recover the emission in the second detector. Extending to the second detector basically doubles the wavelength coverage, allowing us to effectively obtain the equivalent of two grating/filter observations in a single observation. The possibility of extending the spectra has already been demonstrated for the high-resolution (R~2700) grating

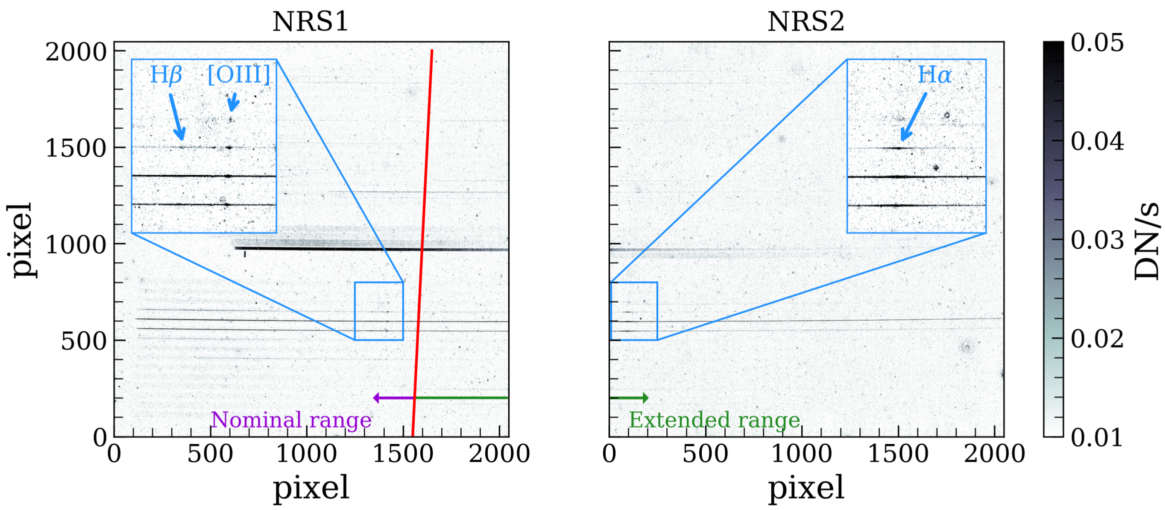

Fig. 1. Count-rate images for the G140M/F100LP observations of RM332, a QSO at $z \sim 2.6$ downloaded from MAST of the NRS1 and NRS2 detectors, on left and right, respectively. The vertical red line shows the maximum wavelength range of the nominal coverage of the observations. The inset panels show a zoomed-in view of the [Oiii]-Hβ complex on the left, and Hα emission on the right.

(D'Eugenio et al. 2025b; Torralba et al. 2025), but in that case the gain in wavelength is much smaller, of the order of 0.1 μm. The extension is made possible by the fact that the photon conversion efficiency does not steeply drop as a function of wavelength, in particular for the G140M/F100LP and the G235M/F170LP grating/filter configurations (for details see Böker et al. 2022), which benefit most from the extension. For the G395M/F290LP, extending the wavelength range beyond the nominal limit of 5.27μm is limited by the detector sensitivity, which drops significantly, reaching 0 at $\lambda \sim 5.5\mu$m (Jakobsen et al. 2022). The G140M/F070LP could, in principle, be extended; however, with limited scientific gain and significantly more effort. For IFU observations, the wavelength range coverage of the G140M/F070LP is limited to the range between 0.90 and 1.26 μm, because the spectrum falls out of NRS1 at bluer wavelengths. Hence, the only effective gain in using the G140M/F070LP with respect to the G140M/F100LP is just to observe in the 0.90 and 0.97 μm interval. Additionally, by using the F070LP, the second and third-order spectra will contaminate the spectrum at shorter wavelengths, and we will have to deal with fourth and fifth-order contamination. For these reasons, we do not consider this setup in this work.

In this section, we describe the modification to the standard JWST data reduction pipeline and reference files, which allowed us to extend the wavelength range covered by NIRSpec/IFU R1000 observations, together with all the tests performed to ensure the goodness of the modified pipeline. Finally, we provide an estimate of the RMS and performance that these observations can achieve.

2.1. Data reduction

To extend the medium resolution gratings, we first create and modify the reference files to extract flux beyond the nominally calibrated wavelength range, and then we perform the flux calibration. For the initial tests and calibration steps, we exploited the observations of standard stars, observed in the calibration programs 6645, 1536, 1537, and 1538. The four programs observed three different stars with all the possible configurations of gratings and filters, with the goal of calibrating the flux and the flat fields for the IFU and other NIRSpec modes. We also used the observations of other targets (local galaxies, AGN, high-z galaxies) obtained with different gratings to verify the accuracy of the calibrations and calibrate our extension of the medium resolution gratings, which we then applied to other targets. The complementary sources that help verify and calibrate our extension were selected to respect the following criteria i) have observations with at least two consecutive medium resolution gratings (i.e., G140M/F100LP + G235M/F170LP or G235M/F170LP + G395M/F290LP) ii) being bright enough to have continuum detection with $S/N > 5$ across all the wavelength range covered.

For all our reductions, we retrieved the count rate maps (from both the first and second detectors) from the MAST archive. Then, we applied the standard data reduction steps `calwebb_spec2` and `calwebb_spec3` using pipeline version 1.18.1 and the context `jwst_1364.pmap`, performing a first reduction with the standard pipeline as a benchmark. Then, we modified the reference files to allow for the extension. In particular:

- we extrapolated each `S-flat` above the nominal wavelength range with ones (see Sec. 2.1.1);
- we modified the `F-flat` files such that each file is extended up to 5.5μm by concatenating the `F-flat` for every filter (see Sec. 2.1.2);
- we modified the wavelength range for each grating in the `cube_build` and in the `nirspec_wavelength_range` reference files.

In the following subsections, we provide a detailed description of the main changes made to the various configuration files. Moreover, the pipeline itself requires an additional modification: we removed the check of the presence of files from the second detector in the '`calwebb_spec2`' step of the pipeline, which otherwise causes an error and stops the data reduction. All the modified reference files, together with a guide on how to apply the changes to the official pipeline and the Python scripts to perform the flux calibration, are available on GitHub[1].

With all these modifications, the pipeline successfully creates the cubes up to the maximum wavelength range covered by the second detector. The extension naturally creates a 'gap' in

[1] https://github.com/eleonoraparlanti/nirspecIFU-extended

the final spectra similar to that visible in the data obtained with the high-resolution gratings, due to the physical gap between the two detectors. Table 1 reports the extended and nominal wavelength coverage, as well as the range of wavelengths that fall in the detector gap in the extended data.

Table 1. Nominal science range, extended science range, and new gap for each filter/grating combination analyzed in this work.

Grating/Filter	Nominal range [μm]	Extended range [μm]	Gap [μm]
G140M/F100LP	0.97 – 1.88	0.97 – 3.55	2.17 – 2.28
G235M/F170LP	1.70 – 3.15	1.66 – 5.27	3.66 – 3.82

Notes. Nominal range from Böker et al. (2022)

2.1.1. Spectroscopic Flat (S-flat)

The `S-flat` files consist of one file for each detector and each filter/grating configuration, and account for all the losses in the optical path from the aperture plane until the disperser. Each file is constituted by three image extensions and 2 Binary tables. The small corrections, slowly varying with wavelength, are stored in the "SCI" extension, while the Binary table extension labeled "FAST_VARIATION" encapsulates the rapidly wavelength-dependent variations with larger relative amplitude.

To extend the wavelength range of the reduced data, we modified the "SCI" and "DQ" extensions, which contain the information of the detector response at each pixel and the data quality flag for each pixel. Since in the IFU, for each grating and filter, the various wavelengths fall on the same pixel in the detector, the pixels containing wavelengths outside the nominal range have a value of "nan" in the SCI flat and are flagged as pixels not to be used in the DQ extension. Hence, for the S-flat in NRS1, we extended each trace until the end of the detector in the "SCI" and "DQ" images. Because the spectral traces exhibit a slight curvature and tilt on the detector, we fitted a linear polynomial to their edges and extended the resulting paths according to the best-fit model. The pixels in the extended "SCI" version have been fixed at 1, while in "DQ", we remove each flag, inserting the value 0. We show the original S-flat for NRS1 and the extended one in Fig. A.1. Moreover also the "FAST_VARIATION" table, which is only defined in the nominal wavelength range, must be modified. In particular, we extended it with a constant value after the last element defined until the maximum wavelength covered by the extended detectors. The original and extended fast variation values are shown in the right panel in Fig. A.1.

We retrieved the latest NRS2 S-flat files for each grating/filter combination from the CRDS website. Since NRS2 should not be used in the reduction of the medium resolution IFU products, its DQ extension is all flagged as "do not use", while the "SCI" extension is initialized with all ones. In this case, we just changed the "DQ" extension, removing the flag for each pixel, and we extended the "FAST_VARIATION" table as explained above.

2.1.2. Fore-optics Flat (F-flat)

The F-flat accounts for the losses from all the reflections in the telescope and in the NIRSpec fore optics. The F-flat does not depend on the detector, but is defined just for each filter/grating combination. In order to extend the wavelength coverage, we concatenated the Binary table extension describing the "FAST_VARIATION" of one filter/grating with the ones of the following one. For instance, the newly defined F-flat for the G140M/F100LP is now the concatenation with the ones of the G235M/F170LP until 3.15μm and the one of the G395M/F290LP until 5.2μm.

2.2. Flux calibration

With the aforementioned changes in the standard data reduction workflow, the output of the `calwebb_spec3` step is a datacube with an extended wavelength range (see Table 1), although only the spectra within the nominal range are flux calibrated. Our goal is therefore to determine an empirical correction curve that can be applied to the final datacube to calibrate the flux in the extended wavelength region.

First, we verified that the profile of the uncalibrated spectra in the extended wavelength region does not depend on the location of the targets within the field of view. Confirming this ensures that the empirical correction curve we derive needs to depend only on wavelength, with no spatial dependence. We thus utilized the data of the standard star P330-E, which was observed in cycles 1 and 3 across all grating/filter configurations in PID 1538 and 6645. In the PID program 6645, the star was also observed by centering it in different locations of the field of view. Figure 2 (B.1) shows the G140M/F100LP (G235M/F170LP) datacubes produced by `calwebb_spec3`, where the spectra are extracted from a circular aperture of 0.3″ centered on the star. Comparing the spectra, we find no significant differences for the same target observed two years apart and at different positions within the field of view. This result indicates that our empirical correction curve should depend only on wavelength and can be applied to any spectrum extracted from any location in the field of view.

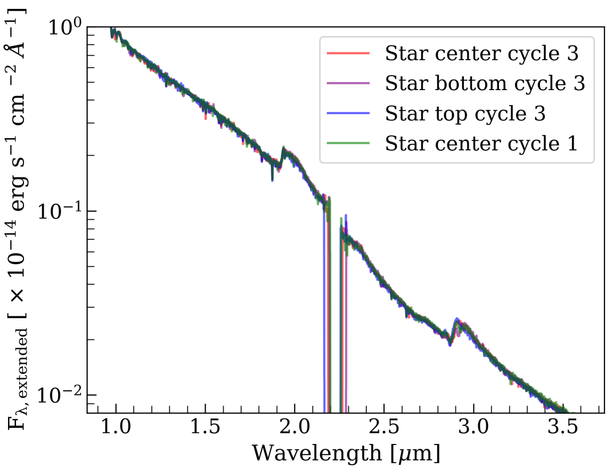

Fig. 2. Observed flux of the extended G140M/F100LP filter for the star P330-E observed as part of the programs 1538 and 6645. The different color curves show the same flux for the star observed in cycle 1 (green), in cycle 3 at the center of the FOV (black), and in the bottom and top parts of the FOV (blue and purple, respectively). The same for the G235M/F170LP is shown in Appendix B.

We then proceeded to estimate the empirical correction curve by using spectra of various standard stars observed with all filter–grating configurations from calibration programs PID 6645, 1536, 1537, and 1538, as well as bright sources (QSOs and local galaxies) observed in programs 2564 and 2186. The correction curve was derived by comparing the uncalibrated spectra in the extended wavelength region with the flux-calibrated spectra obtained from the filter–grating configurations covering longer wavelengths. To this end, we first extracted the flux-calibrated spectrum (hereafter referred to as the intrinsic spectrum) within each nominal wavelength range. However, we found that dividing the intrinsic spectrum by the uncalibrated spectrum is not sufficient to obtain an accurate correction curve. In particular, we must account for contamination from higher–order spectra (primarily second and third order), whose contribution becomes significant at wavelengths around $n, \lambda_{\rm min}$, where n is the spectral order and $\lambda_{\rm min}$ is the shortest wavelength of the filter.

Assuming the contributions only up to the third order, the output of `calwebb_spec3` can be expressed as:

$$\begin{cases} S_\lambda(\lambda) = f_\lambda(\lambda) & \lambda \leq \lambda_{\rm max}^{\rm nominal} \\ S_\lambda(\lambda) \approx k(\lambda) f_\lambda(\lambda) + \alpha(\lambda) g_\lambda(\lambda) + \beta(\lambda) h_\lambda(\lambda) & \lambda > \lambda_{\rm max}^{\rm nominal} \end{cases}$$

where $\lambda_{\rm max}^{\rm nominal}$ denotes the maximum wavelength of the NIRSpec nominal range, $f_\lambda(\lambda)$ represents the intrinsic source spectrum, and $g_\lambda(\lambda)$ and $h_\lambda(\lambda)$ correspond to the second- and third-order spectra, respectively. The functions $k(\lambda)$, $\alpha(\lambda)$, and $\beta(\lambda)$ describe the wavelength-dependent transmission efficiencies for each spectral order. The second- and third-order spectra can also be expressed as functions of the first-order spectrum, such that $g_\lambda(\lambda) \propto f_\lambda(\lambda/2)$ and $h_\lambda(\lambda) \propto f_\lambda(\lambda/3)$. Therefore, the uncalibrated spectrum beyond the nominal wavelength range can be written as

$$S_\lambda(\lambda) = k(\lambda) f_\lambda(\lambda) + \tilde{\alpha}(\lambda) f_\lambda(\lambda/2) + \tilde{\beta}(\lambda) f_\lambda(\lambda/3)$$

where $\tilde{\alpha}(\lambda)$ and $\tilde{\beta}(\lambda)$ are the effective transmission coefficients for the second- and third-order contributions as function of f_λ. Dividing both sides by the intrinsic spectrum $f_\lambda(\lambda)$ gives

$$\begin{aligned} R(\lambda) = \frac{S_\lambda(\lambda)}{f_\lambda(\lambda)} &= k(\lambda) + \tilde{\alpha}(\lambda) \frac{f_\lambda(\lambda/2)}{f_\lambda(\lambda)} + \tilde{\beta}(\lambda) \frac{f_\lambda(\lambda/3)}{f_\lambda(\lambda)} \\ &= k(\lambda) + \tilde{\alpha}(\lambda) r_2(\lambda) + \tilde{\beta}(\lambda) r_3(\lambda) \end{aligned}$$

where $r_2(\lambda)$ and $r_3(\lambda)$ are known functions determined by the intrinsic source spectrum $f_\lambda(\lambda)$. In conclusion, at a fixed wavelength $\bar{\lambda}$, there are three unknown variables: $k(\bar{\lambda})$, $\tilde{\alpha}(\bar{\lambda})$, and $\tilde{\beta}(\bar{\lambda})$. Therefore, if both $S_\lambda(\lambda)$ and $f_\lambda(\lambda)$ are known for at least three independent sources, a system of equations at each wavelength can be solved for $k(\lambda)$, $\tilde{\alpha}(\lambda)$, and $\tilde{\beta}(\lambda)$.

The functions $k(\lambda)$, $\alpha(\tilde{\lambda})$, and $\beta(\tilde{\lambda})$ were recovered for the G140M/F100LP and G235M/F170LP by exploiting all the datasets for which we have both the extended spectrum $S_\lambda(\lambda)$ from our reduction and the intrinsic spectrum $f_\lambda(\lambda)$ obtained by reducing the three gratings within their nominal wavelength range. For each combination of three datasets, we computed the three functions. To account for the presence of some bad pixels and spikes in these calibration functions, we averaged and smoothed them over a window of 40 channels along the wavelength axis. The three functions are reported in Fig. 3. We note that the $\tilde{\alpha}$ and $\tilde{\beta}$ functions, which quantify the level of the contamination of the second and third order spectra to the observed one, respectively, are of the order of 1–5% at their maximum value, that is the range between 2–2.3 μm for the G140M and 3.2 – 4 μm for the G235M.

Finally, knowing the calibration functions, we obtained the intrinsic spectrum of each source in the extended region as:

$$f_\lambda(\lambda) = \frac{[S_\lambda(\lambda) - \tilde{\alpha}(\lambda) f_\lambda(\lambda/2) - \tilde{\beta}(\lambda) f_\lambda(\lambda/3)]}{k(\lambda)} \quad (2)$$

where $\tilde{\alpha}(\lambda)$ and $\tilde{\beta}(\lambda)$ are defined equal to zero at wavelength smaller than $2\lambda_{\rm min}$ and $3\lambda_{\rm min}$, respectively; and $k(\lambda)$ is defined as one at $\lambda \leq \lambda_{\rm max}^{\rm nominal}$.

To assess the accuracy of Equation 2, and therefore the reliability of the flux calibration in the extended wavelength range, we show in Fig. 4 the ratio between the spectra obtained in the nominal ranges and the extended flux-calibrated spectra for both G140M/F100LP and G235M/F170LP. To perform this test, we used the data from the programs: 6645, 1536, and 1537, which observed a type G, a type-A star, and a white dwarf, respectively. For the G140M/F100LP, we also used the observations of two AGN, SDSSJ0749 and SDSSJ0841, which we dubbed AGN 1 and AGN 2 from the program 2654. For the G235M/F170LP, we utilized observations from three local ULIRGS observed as part of the program 2186. Figure 4 highlights that the calibration described above achieves an accuracy of ∼ 5% and ∼ 10%, respectively, for the G235M/F170LP and the G140M/F100LP configurations. These values are well within the current NIRSpec flux calibration uncertainties and discrepancies between different gratings (of the order of ∼ 20 – 25%, D'Eugenio et al. 2025a; Scholtz et al. 2025).

2.3. Properties of the extended spectra

After successfully creating the cubes and computing the functions necessary to account for the second- and third-order spectra, and correctly calibrating the flux, we obtained the data cube containing only the target's spectrum in the extended wavelength range. Figure 5 shows two examples of the extension of both G140M/F100LP and G235M/F170LP in the top and bottom panels, respectively. The two panels showcase some examples of how, after the calibration, we achieved a good agreement with the nominal spectrum, consistently removing the effect of the second and third order spectra and correcting for the unknown flat field level in the extended region of the first detector and the entire second detector.

From these spectra, it is, however, visible that the RMS of the extended region is higher than that of the nominal range. We note that in the case of the AGN SDSSJ0841, both filters have been observed with the same exposure time, while for the target UGC-5101 reported in the bottom panel, the exposure time for the observations with the G235M is double that of the G395M. The increase in the RMS must be attributed both to the decrease in the disperser transmission outside its nominal wavelength range (see JWST/NIRSpec documentation[2]) and also to the increase in spectral resolution. In particular, for the G140M, the transmission drops from 0.9 at ∼ 1μm, down to ∼ 0.2 at 3.5μm, the maximum wavelength covered in our extension. Similarly, the transmission curve of the G235M disperser reaches its peak of 0.9 at ∼ 2.1μm and drops to 0.3 at 5μm.

We quantified the increase in the RMS with wavelength using the datasets having observations in all the gratings/filters and renormalizing the measured RMS to account for the difference in exposure time in the different gratings/filters observations. Hence, we computed the RMS over 10 wavelength channels for both the nominal and the extended spectrum. Figure 6 shows, as a solid black line, the median ratio between the measured RMS

[2] https://jwst-docs.stsci.edu/jwst-near-infrared-spectrograph/nirspec-instrumentation/nirspec-dispersers-and-filters

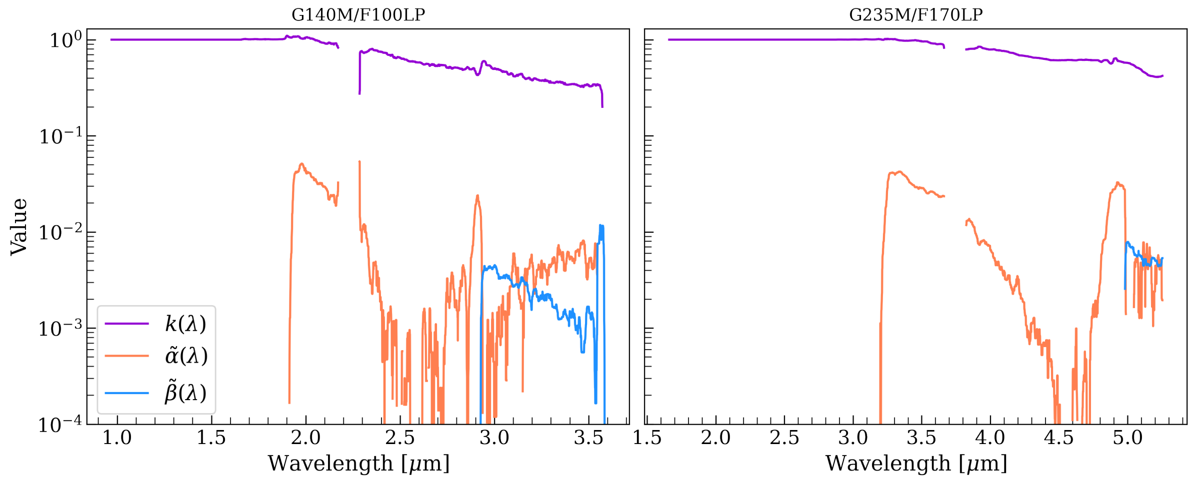

Fig. 3. Values of $\alpha\tilde{(\lambda)}$ (orange), $\beta\tilde{(\lambda)}$ (blue), $k(\lambda)$ (violet), as function of wavelength for the G140M/F100LP and the G235M/F170LP on the left and right panels, respectively.

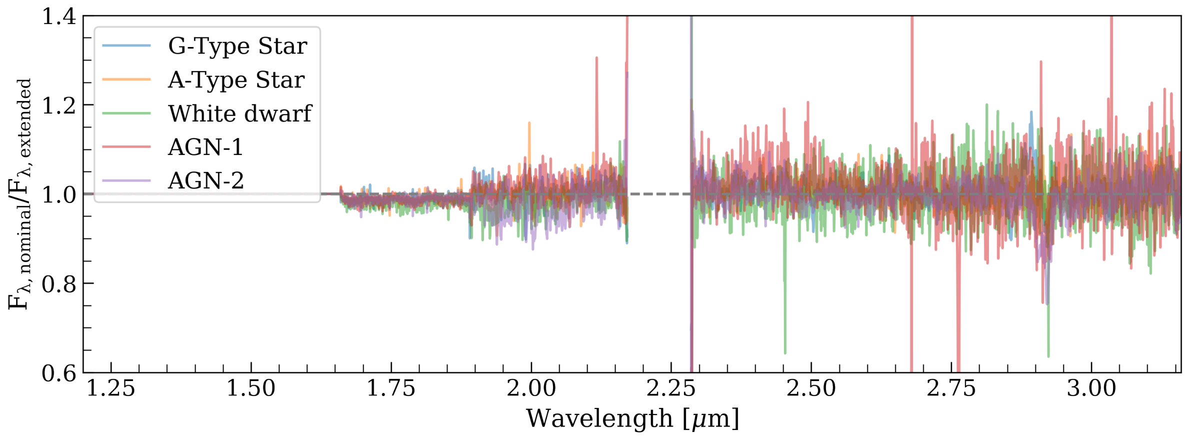

Fig. 4. Ratio between the observed flux in the nominal wavelength range of the filters and the flux observed with the extended G140M/F100LP filter. The same figure for the G235M/F170LP is shown in Figure B.2.

of the extended cubes and that measured in the nominal range computed for each cube (shown in gray) as a function of wavelength, for both the G140M/F100LP and the G235M/F170LP datasets. The value obtained by each dataset fluctuates, but the median steadily increases from one in the overlapping region between the two consecutive gratings up to 3 at the redder wavelengths of our extension.

As expected from the grating equation, the spectral resolution continues to increase with increasing wavelength, following the trend of Jakobsen et al. (2022). We confirm this trend also beyond the nominal wavelength range. Figure 7 shows the same Hα line from the galaxy DC-417567 at $z = 5.67$ observed in the program 3045 (Faisst et al. 2025). We show both the extended G235M/F170LP observations and the nominal range of G395M/F290LP, highlighting the different resolutions and RMS.

To measure the spectral resolution beyond the nominal range, we employed observations of emission lines detected both in a nominal filter (whose resolution is known and tabulated) and in the extended region, and we fitted them. From the FWHM of the line measured in the nominal range ($\mathrm{FHWM_{obs,nom}}$), we can derive the intrinsic FWHM by correcting for the instrumental broadening as $\mathrm{FWHM_{int}} = \sqrt{\mathrm{FHWM^2_{obs,nom}} - \mathrm{FWHM^2_{instr,nom}}}$. As a fiducial value for the resolution in the nominal range of each filter $\mathrm{FWHM_{instr,nom}}$, we employ the calibrations by Shajib et al. (2025), which highlight that the in-flight spectral resolution is better by more than 10% than the one estimated before the launch and provided in the JWST documentation (Jakobsen et al. 2022). Once the intrinsic line width is obtained, we can derive the instrumental FWHM in the extended wavelength range as $\mathrm{FWHM_{instr,ext}} = \sqrt{\mathrm{FHWM^2_{obs,ext}} - \mathrm{FWHM^2_{int}}}$, where $\mathrm{FHWM_{obs,ext}}$ is the measured FWHM of the line in the extended spectrum.

Figure 8 shows the resolution reported by the JWST documentation, the one reported by Shajib et al. (2025), and its extrapolation beyond the nominal wavelength range, and, as colored points, our measurement of the resolution in the extended wavelength range. For the resolution beyond the nominal range of the G235M, we used observations from the program 3045, which targeted the same galaxies with both the G235M/F170LP and G395M/F290LP filters/gratings. We fitted the bright [O III] and Hα lines falling at 3 – 4.5 μm at $z \sim 5.1 - 5.67$. For the G140M/F100LP, we just used the data from the program 3435, which targets the galaxy M51, where we detect the Brβ emission line at 2.63 μm, even if with lower S/N in the extended region. To our knowledge, there are no other suitable data in the archive with narrow emission lines observed in both the G140M/F100LP

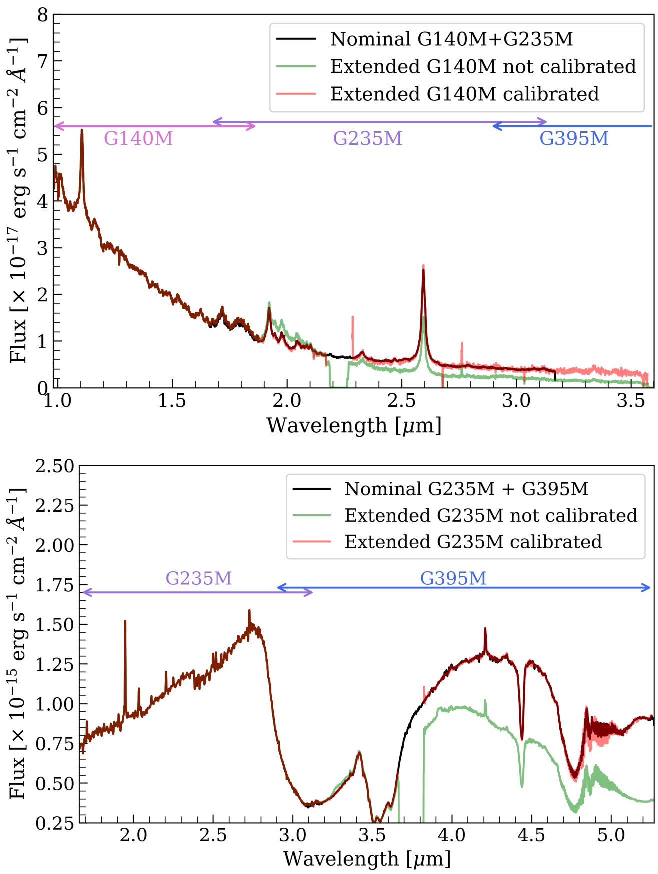

Fig. 5. Examples showing the spectrum in the nominal range (black), and the extended spectra before and after the calibration reported in Eq. 2 (green and red, respectively). In the top panel, we show the spectrum extracted from a circular aperture of radius 0.5″ centered on the brightest pixel of the AGN SDSSJ0841 (PID: 2654), the nominal G140M/F100LP+G235M/F170LP spectrum, and the extensions of the G140M/F100LP. In the bottom panel, we show the spectrum extracted from a circular aperture of radius 0.5″ centered on the brightest pixel of the ULRIG UGC-5101 (PID: 2186) from the nominal G235M/F170LP and G395M/F290LP (black) and the extended G235M/F170LP spectrum. The nominal wavelength range encompassed by each filter is shown as the pink, purple, and blue lines for the G140M/F100LP, G235M/F170LP, and G395M/F290LP, respectively

and the G235M/F170LP that we can exploit for our purposes. The resolution measured in the extended region follows the same trend of the curves by Shajib et al. 2025 within the uncertainties, which we hence assume in the rest of this paper. We also remark that, as a result of this new reduction, observations in the extended region have resolutions as high as twice the nominal ones. While this is advantageous as the spectral resolution can easily reach $R \sim 2000$, and Figure 8 shows that at the reddest wavelength our extension of the G140M has a resolution comparable with that of the high resolution G395H data, it also implies lower S/N and contributes to the increasing RMS at longer wavelengths.

Finally, we also tested the wavelength calibration beyond the nominal range. In Fig. 7, we show that the centroid of the emission line falls at the same wavelengths. We systematically checked for any discrepancy in the wavelength calibration by fitting emission lines detected in the nominal and extended gratings. Figure C.1 shows the difference between the centroid position of the same line measured in extended and nominal range, showcasing no systematic discrepancy between the two measurements and highlighting that the wavelength calibration in the extended range is accurate down to 1/4–1/10 of the spectral resolution.

2.4. Different applications

Due to the drawback of an increased RMS, while reaching a higher resolution, our wavelength extension showcases its full potential when deep observations are needed in a blue filter, with bright lines easily detectable in the red filter. When one aims to detect faint emission lines/features in the extended region, one should prefer the observation in the nominal filter, rather than relying on the extended wavelength coverage. In general, some examples of its application might be when there is a need for a search for a faint line in a blue filter, hence the brighter lines in the red filters are observed for free. This is especially relevant for dust-obscured sources, where the emission lines at shorter wavelengths may be heavily attenuated. On the other hand, this method is also well suited for the observation of bright sources, such as QSOs or nearby galaxies, where the S/N is not a limiting factor; hence, the increase in RMS does not heavily affect the final results.

3. Data analysis

In this work, we applied the procedure described above to the targets of the proposal PID: 2057, which observed 10 QSOs at $z \sim 2$ previously targeted by reverberation mapping (RM) campaigns (Shen et al. 2024). Among these, we focused on the five sources with available RM-based black hole mass measurements (Shen et al. 2024). The RM-based mass was estimated from the observation of the C IV (RM032, RM312, and RM539) and the Mg II (RM332 and RM401) emission lines.

The targets were observed with the G140M/F100LP for ~2400s with a "4-POINT-DITHER" dithering strategy. We retrieved the count-rate images from the MAST archive and applied all pipeline steps described above to obtain the final data cube with a "drizzle" weighting method and a spaxel size of 0.05″. We then corrected the flux by applying Eq. 2. For all targets, the [O III] and Hβ emission lines fall in the nominal range covered by the G140M/F100LP observations, while the Hα emission line can only be obtained thanks to our extension of the G140M data. For each cube, we extracted the spectrum from a circular aperture of 1 arcsec radius, centered on the brightest pixel to cover the extent of the PSF (D'Eugenio et al. 2024) and remove the effect of the "wiggles" visible in smaller apertures (Perna et al. 2023). Each integrated spectrum was then fitted using a custom-made Python code, based on the IDL MPFIT package (Markwardt 2009) that finds the best-fit parameters by minimizing the χ^2.

To reproduce the observed spectra, we modeled them as the sum between a power-law continuum arising from the accretion disk, broad lines (FWHM ≥ 1000 km s^{-1}) arising from the BLR, narrow lines (FWHM < 1000 km s^{-1}) arising from the narrow-line region (NLR) or the host galaxy emission, and a Fe II pseudo-continuum. When required, we also included broad components to reproduce NLR outflows. The Hβ–[O III] and Hα–[N II] line complexes were modeled independently in the rest-frame wavelengths between 4200–5500Å, and between 6200–6900Å for Hβ and Hα, respectively.

We employed different spectral shapes to reproduce as faithfully as possible both the broad lines and the narrow ones. In

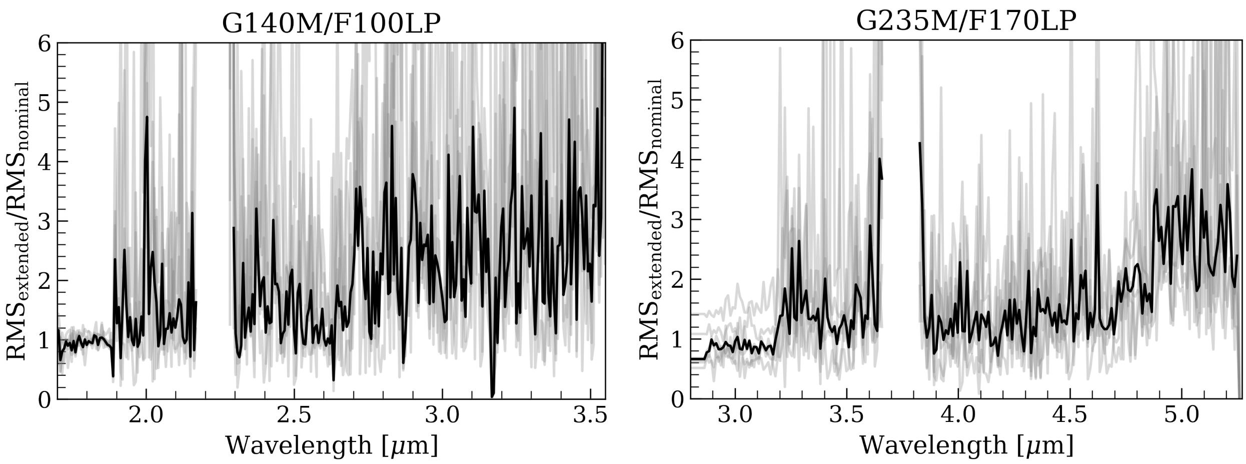

Fig. 6. Increase in the RMS computed in 15 channels for the extended and nominal ranges of the G140M/F100LP and the G235M/F170LP in the left and right panels, respectively. The solid line shows the average value of the ratio between the RMS computed in the extended version and the one in the nominal range. Gray lines show the RMS for each dataset for which we have observations of all the filters/grating combinations with the same exposure time.

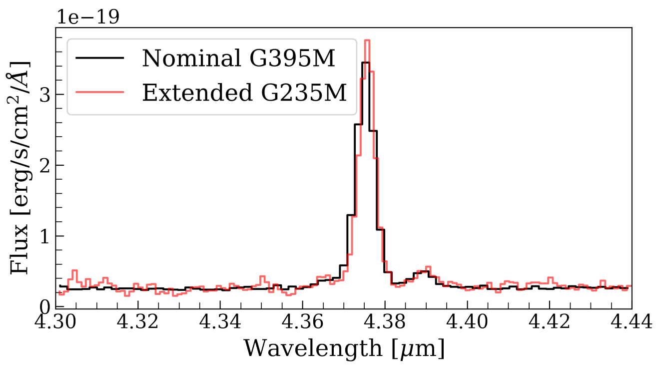

Fig. 7. Example showing the spectrum in the nominal range (black), and the extended spectra after the calibration in red extracted from a circular aperture of diameter 0.3″ centered on the brightest pixel of the SF-galaxy DC-417567 at $z = 5.67$ (see Faisst et al. 2025, for details). The spectra show the Hα-[NII] complex and highlight the increase in spectral resolution and RMS of the extended version.

particular, we used both Gaussian and Lorentzian profiles for the broad lines, while the former profile was generally preferred for the narrow lines. Additionally, we constrained the kinematics of the narrow and outflow components of the Hβ and [OIII]to be the same and set a 3:1 ratio between the [OIII]λ5007 and the [OIII]λ4959 components (Osterbrock & Ferland 2006). Regarding the broad FeII pseudo-continuum, instead of relying on an empirical template, we adopted a linear combination of theoretical templates obtained employing the CLOUDY software (Ferland et al. 2017). More details about the implementation can be found in (Trefoloni et al. 2025). Since the goal of this analysis is to constrain the broad line properties, in order to incorporate as much information as possible about the kinematics traced by the lines, we adopted a two-step process similar to that described in Trefoloni et al. in prep. We first performed the fit, aimed at reproducing as best as possible the global observed spectrum. Then, we subtracted the best-fit components with the exception of the broad line of interest, respectively Hβ and Hα. We used this isolated broad line profile to estimate the relevant information (e.g., flux, FWHM, line dispersion). The result of this procedure is shown in black in Figure 9, where we report the best-fit model

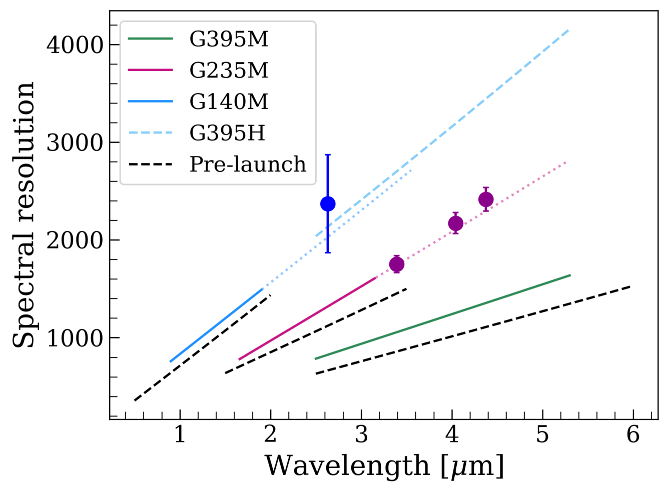

Fig. 8. Spectral resolution as function of wavelength. Black dashed lines show the JWST pre-launch expected resolution. Solid lines show the curves determined after the launch by Shajib et al. 2025, for the G140M, G235M, and G395M in blue, purple, and green, respectively. The dotted colored lines show the extrapolation beyond the nominal wavelength range of the curves determined after the launch. Purple and blue points show the measured resolution in the extended G235M and G140M data, respectively. The light blue dashed line shows the spectral resolution curve derived from Shajib et al. 2025 for the G395H grating.

for one of the QSOs we applied the IFU extension procedure. The others, together with the best fit values, can be found in Appendix D.

4. Comparing different single epoch calibrations

After measuring the broad lines (luminosity and FWHM of Hα and Hβ) and continuum (luminosity at 5100Å) properties from the extended spectra (reported in Table D.1), we use them to derive the single epoch BH-masses. In this section, we compare the BH masses inferred from RM, derived by Shen et al. 2024, with those inferred from different single epoch calibrations, all calibrated from local BHs from the measurement of the Hα and the Hβ emission lines. While individual study of indi-

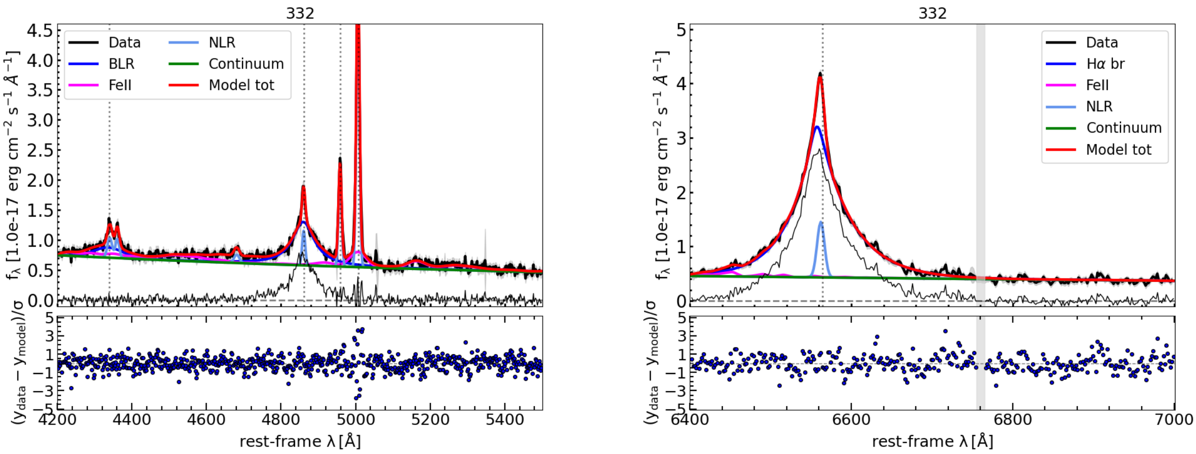

Fig. 9. Rest-frame spectrum and best-fit model for RM332 for the Hβ–[OIII] line complex and the Hα (from the extended region), on the left and right panels, respectively. The data are shown in black, while the model is shown are the red solid line. The total model is constructed as the sum of the BLR, NLR, FeII pseudocontinuum, and continuum emission, which are shown in blue, light blue, magenta, and green, respectively (see legend). The thin black line represents the isolated broad line profile resulting from the best-fit model subtraction. In the bottom panels, we show the residuals, computed as (data-model)/error.

vidual objects, thanks to GRAVITY, already revealed some discrepancies (Abuter et al. 2024; GRAVITY+ Collaboration et al. 2025) between the BH mass derived from modeling of the BLR and from SE scaling relations, the now available RM measurements at high redshift provide a valuable sample benchmark to assess which SE relations yield the most reliable estimates in the early Universe.

To perform this test, we exploited the calibrators of Vestergaard & Peterson (2006, hereafter: VP06, both the one exploiting the monochromatic luminosity at 5100Å and the luminosity of the Hβ emission line) and Dalla Bontà et al. (2020, hereafter: DB20), which rely on the measure of the Hβ; and Reines et al. (2013, hereafter: R13), Dalla Bontà et al. (2025, hereafter: DB25), and Greene & Ho (2005, hereafter: GH05), which exploit the Hα properties (FWHM and luminosity). While the typical uncertainty associated with reverberation mapping M_{BH} is of the order of 0.3 dex (see e.g., Shen et al. 2024), that of SE M_{BH} is of the order of 0.2–0.5 dex (Shen 2013; Dalla Bontà et al. 2025).

The BH masses that we estimate from our 5 targets from the fitted broad emission line by applying the aforementioned SE calibrators are reported in Table 2, together with the literature value of the RM-based BH mass. Figure 10 also highlights the comparison between the different BH mass estimates. It is evident that for some targets, such as RM312, RM401, and RM539, the majority of the calibrators agree with the RM-estimated BH mass within their uncertainties, with scatters smaller than 0.5 dex. In the case of RM332, there is a larger scatter between the various calibrators, employing different emission lines, similar to what has been found in other high-z targets (Abuter et al. 2024; Bertemes et al. 2025; Marshall et al. 2025). On the other hand, RM032, the least massive BH according to the RM measurement, shows a striking difference between all the different calibrators and the RM-estimated mass: all SE estimates are higher compared to the RM measurement. In particular, this analysis highlights that by applying some SE calibrators, the derived BH mass can be one order of magnitude higher than the one measured with the RM technique. We note that the RM mass for RM032 is estimated by employing the FWHM of the CIV emission line, which can be affected by uncertainties in its modeling since it can trace non-virial motions, such as outflows (Denney 2012). However, also RM312 and RM539 employ the same line measure, showing good agreement with the RM and SE measures, and we do not see any evidence for the presence of strong outflows in the [OIII] emission line or RM032 (see Fig. D.1).

With the exception of RM032, we find that the calibrators using the Hβ emission line show a better agreement in estimating the RM-BH mass for our sample, with the smallest deviation achieved by using the VP06 calibrations. The Hα-based calibrators, on the other hand, in most cases show a larger scatter of the derived BH mass by employing different calibrators; however, they are usually always within 0.5 dex of the RM-based M_{BH}.

To investigate the origin of the deviations in BH mass estimates, we looked at possible correlation with the Eddington ratio. This is motivated by multiple studies (e.g., Du & Wang 2019; King 2024; Lupi et al. 2024) that highlighted that when close to or above the Eddington limit, SE virial estimators of the BH mass might not be applicable anymore. For instance, King (2024) highlighted that approaching the Eddington limit, the BLR could be far from virialized and be instead dominated by outflowing motions. Such behaviour has also been recently observed in the BLR of a super-Eddington QSO ($\lambda_{Edd} \sim 6-19$; GRAVITY+ Collaboration et al. 2025), suggesting that standard calibrations, relying on the virialization of the BLR, could fall short in this case. On the other hand, Lupi et al. (2024) pointed out to the fact that in the case of super-Eddington accretion, the accretion is thought to happen through a slim disk, which would screen some of the ionizing photons, preventing them from reaching the BLR. In this case, assuming the standard $R-L$ relation would overestimate the actual BLR size, as also reported in various observations (Du & Wang 2019; Abuter et al. 2024). Both these different mechanisms would result in an overestimation of the estimated virial BH mass.

The Eddington luminosity was estimated as:

$$L_{EDD} = \frac{4\pi G M_{BH} m_p c}{\sigma_T} \simeq 1.26 \times 10^{38} \left(\frac{M_{BH}}{M_\odot}\right) \mathrm{erg\,s^{-1}} \quad (3)$$

where M_{BH} is the BH mass estimated from reverberation mapping, G is the gravitational constant, m_p is the proton mass, c is the speed of light, and σ_T is the Thomson scattering cross section.

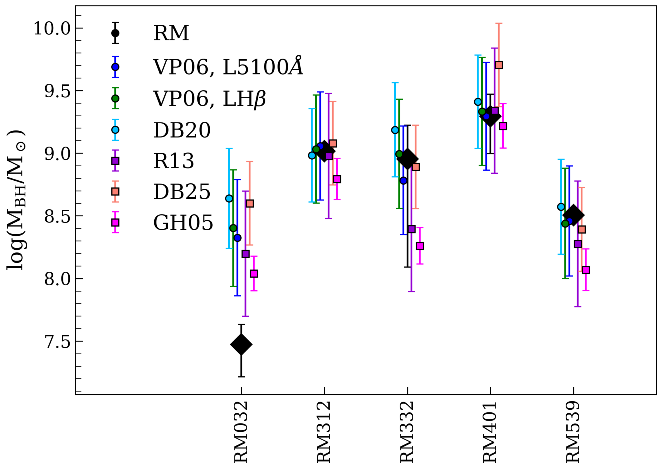

Fig. 10. Comparison between the RM-estimated BH mass (large black diamond), and the single epoch calibrators employing the Hβ emission line VP06-L5100, VP06-LHβ, DB20, which are shown are blue, green and light blue circles, respectively, and the SE calibrators employing the Hα, R13, DB25, GH05, shown as the purple, orange and pink squares, respectively.

The Eddington ratio ($\lambda_{\rm EDD}$) is then computed as the ratio between the observed bolometric luminosity reported in Shen et al. 2019, that is derived from the continuum luminosity employing the bolometric corrections described in Richards et al. 2006, and the Eddington luminosity. While all the other targets have Eddington ratios between 0.03 and 0.12, RM032 has a $\lambda_{\rm EDD} = 0.8 \pm 0.3$.

In the case in which the RM measurements of RM032 are not affected by the presence of non-virial motion traced by the C IV line, the results presented in this work would suggest that, at high redshift, the majority of the SE calibrators overestimate the masses of low-mass rapidly accreting BHs. If confirmed, this could reconcile new results of JWST-discovered broad line AGN, where BH masses are found to be comparable to the stellar masses of the host (Kokorev et al. 2023; Juodžbalis et al. 2024). However, our current sample includes only one target with $\log(M_{\rm BH}/M_\odot) < 8$, hence larger statistics are needed in order to confirm this trend.

5. Calibrating the single epoch relation at z~2

In the previous section, we tested the various SE calibrators to estimate the BH masses for high-z BHs. Since all the SE calibrators so far have been derived based on local AGN, we set to derive a first SE calibration at high-z, exploiting the 5 targets above (plus the ones with dynamical estimation of the BH mass by GRAVITY at $z \sim 2$ and $z \sim 4$ from Abuter et al. 2024 and GRAVITY+ Collaboration et al. 2025, and another target with BH mass inferred from RM and Hβ detection at $z \sim 4$ (Saturni et al. 2018). In particular, we fitted the relation

$$\log(M_{\rm BH}/M_\odot) = a + b[\log(L) - 42] + c[\log({\rm FWHM}) - 3] \quad (4)$$

where $M_{\rm BH}$ is the mass from RM campaigns (or from the modeling of the BLR for the GRAVITY targets), L is the luminosity (of the Hα, Hβ emission line or monochromatic continuum emission at 5100Å in units of erg s^{-1}) and FWHM is the FWHM of the Hα or Hβ broad emission lines in units of km s^{-1} that we obtain from our fitting. We simultaneously fitted the parameters a, b, and c, together with the intrinsic scatter $\sigma_{\rm int}$. We performed the fit using a Markov Chain Monte Carlo (MCMC) approach, which allowed us to explore the posterior distributions of the parameters and account for the measurement uncertainties. We adopted uniform priors on a, b, c, and $\sigma_{\rm int}$, ran 50 chains with 100.000 steps to ensure convergence, and discarded the initial 10.000 as the initial burn-in phase. The posterior distributions of the parameters are shown in Figures E.1, E.2, and E.3 while in Table 3 we report the best value for each calibrator and its uncertainties estimated as the 16$^{\rm th}$ and 84$^{\rm th}$ percentiles of the posterior distribution.

Figure 12 shows the comparison between the BH mass from RM mapping and kinematics measure and the SE mass estimated from our calibration for the Hα, and Hβ emission line measurements. Our calibrations give $M_{\rm BH}$ values consistent with the other calibrators for our sources, within the large uncertainties due to the small number of points fitted and the small range of BH masses of our sample, spanning only 2–2.5 dex. We note that we can reproduce the mass of RM032 within 1σ and reproduce well the mass from the GRAVITY targets, both the one at $z \sim 2$ by using Hα and especially the one at $z \sim 4$ by using Hβ, for which all the other calibrators give a BH mass an order of magnitude higher (GRAVITY+ Collaboration et al. 2025).

We also apply our calibration to the newly discovered *JWST* high-z AGN population, using the Hα fluxes and FWHM values reported in Harikane et al. (2023) and Maiolino et al. (2024). We note that these sources have redshifts between 4 and 7, beyond the range in which we calibrate our relation. The revised BH-mass estimates are consistent within the uncertainties with the values reported in the literature, which were derived using the calibrations of Greene & Ho (2005) and Reines et al. (2013). Even with our new calibration, the *JWST* high-z AGN population lies above the local relation (Fig. 11), indicating that the detection of "overmassive" BHs does not depend on the specific BH-mass estimator adopted. However, we emphasize that our relation is calibrated only for $M_{\rm BH} \gtrsim 10^{7.5} M_\odot$, and extrapolating it to lower masses may be unreliable and could lead to biased BH-mass estimates.

Finally, we tested our relations on a larger sample of local QSOs from Wu & Shen (2022). In appendix F we show the differences between the BH mass inferred from our relations and that inferred by using the SE relations by VP06 and R13. We show that, using our relation based on Hβ, $M_{\rm BH}$ is typically overestimated, or consistent with standard calibrators, at the low-luminosity end, while it is on average underestimated by up to 0.4 dex at the highest luminosities when compared to VP06 calibrators. This trend is significantly less pronounced when using Hα, although it does not disappear entirely. Larger, unbiased samples spanning a broad range of luminosities and BH masses are required to draw more robust conclusions and address the reasons for these discrepancies.

6. Summary and conclusions

In this work, we present and release a modification of the JWST Science Calibration Pipeline for NIRSpec IFU observations that more than doubles the wavelength coverage achievable with the medium-resolution (R1000) G140M/F100LP and G235M/F170LP configurations (see Table 1). We also provide the corresponding reference and calibration products as machine-readable files, enabling the community to run the data-reduction workflow directly and obtain final flux-calibrated datacubes over the extended wavelength range. The extended coverage maximizes the scientific return from archival datasets and will offer new opportunities for the upcoming observing pro-

Table 2. Comparison between different BH masses estimates for the 5 targets of this paper. The first and second columns report the measure and the calibrators used to derive it, respectively. In the columns from 2 to 7, we report all the BH masses estimated for each target.

	Calibrator	RM032	RM312	RM332	RM401	RM539
z	-	1.714	1.92	2.58	1.822	2.27
$\log(M_{\rm BH,RM}/M_\odot)$	-	$7.47^{+0.22}_{-0.26}$	$9.01^{+0.03}_{-0.05}$	$8.95^{+0.27}_{-0.26}$	$9.29^{+0.23}_{-0.30}$	$8.50^{+0.05}_{-0.02}$
$\log(M_{\rm BH,SE}/M_\odot)$	VP06, FWHM(Hβ), L5100	8.33 ± 0.46	9.06 ± 0.43	8.78 ± 0.43	9.29 ± 0.46	8.45 ± 0.44
$\log(M_{\rm BH,SE}/M_\odot)$	VP06, FWHM(Hβ), LHβ	8.40 ± 0.47	9.03 ± 0.43	8.99 ± 0.43	9.33 ± 0.43	8.44 ± 0.44
$\log(M_{\rm BH,SE}/M_\odot)$	DB20, FWHM(Hβ), LHβ	8.64 ± 0.39	8.98 ± 0.37	9.18 ± 0.37	9.40 ± 0.37	8.57 ± 0.38
$\log(M_{\rm BH,SE}/M_\odot)$	R13, FWHM(Hα), LHα	8.19 ± 0.5	8.97 ± 0.5	8.39 ± 0.5	9.34 ± 0.5	8.28 ± 0.5
$\log(M_{\rm BH,SE}/M_\odot)$	DB25, FWHM(Hα), LHα	8.60 ± 0.33	9.03± 0.33	8.89 ± 0.33	9.70 ± 0.3	8.39 ± 0.3
$\log(M_{\rm BH,SE}/M_\odot)$	GH05, FWHM(Hα), LHα	8.04 ± 0.14	8.79 ± 0.17	8.26 ± 0.14	9.2 ± 0.17	8.06 ± 0.15

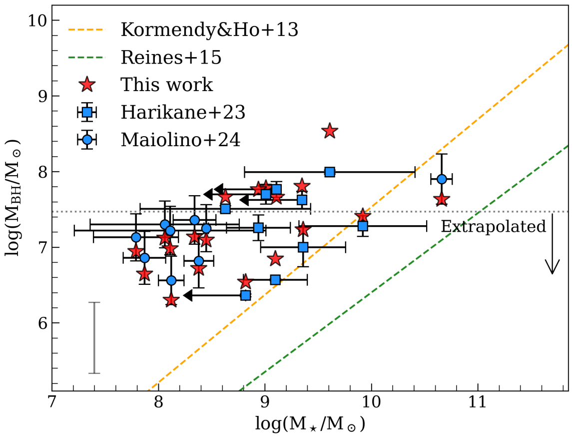

Fig. 11. $M_{\rm BH}$–stellar mass relation for the high-z BHs studied by Harikane et al. (2023) and Maiolino et al. (2024). Blue points show the values of $M_{\rm BH}$ derived by Harikane et al. (2023) and Maiolino et al. (2024) using the SE relations from GH05 and R13, respectively. The local scaling relations Red stars show the values of $M_{\rm BH}$ that are obtained by using the SE Hα-based equation derived in this work. The typical systematic uncertainties on the SE-$M_{\rm BH}$ are shown as the gray line in the bottom left corner. The yellow and green dashed lines show the local $M_{\rm BH}$-$M_\star$ relation from Kormendy & Ho 2013 and Reines & Volonteri 2015, respectively. The gray dashed line marks the regime below which our calibration is extrapolated.

Table 3. Best-fit parameters of the the three different relation employing the Hα line, the Hβ FWHM and luminosity and the Hβ FWHM and the continuum luminosity at 5100Å.

Calibrators	a	b	c	σ_{int}
$L_{\rm H\alpha}$, $\rm FWHM_{H\alpha}$	$6.48^{+0.71}_{-0.76}$	$0.34^{+0.26}_{-0.20}$	$2.54^{+0.91}_{-0.93}$	0.46
$L_{\rm H\beta}$, $\rm FWHM_{H\beta}$	$6.12^{+0.77}_{-0.66}$	$0.23^{+0.15}_{-0.12}$	$3.20^{+0.90}_{-1.05}$	0.40
$L_{5100Å}$, $\rm FWHM_{H\beta}$	$5.79^{+0.70}_{-0.52}$	$0.23^{+0.12}_{-0.11}$	$3.01^{+0.78}_{-0.92}$	0.35

grams with JWST. We note that spectra in the extended wavelength range exhibit a modest increase in the RMS level (up to a factor of 2—3 compared to observations obtained in the nominal filter). As a result, the method is best suited for bright targets or for detecting longer-wavelength spectral features (e.g., rest-frame optical/near-IR lines) as a secondary objective in observing programs with long exposure times that primarily focus on the detection of fainter, shorter-wavelength emission as the rest-frame UV lines. The extended spectra reach a resolution up to R ∼ 2500, approximately twice that of spectra in the nominal range and comparable with that of the high-resolution gratings (R ∼ 2700).

We exploit our new data reduction to recover the Hα emission line of a sample of five QSO at $z \sim 2$, for which only Hβ and [OIII] were included in the nominal wavelength range. We selected those targets because they have an independent measure of the BH mass through reverberation mapping campaigns (Shen et al. 2024) derived from the MgII or CIV emission line. Our extension of the G140M/F100LP data also allows us to obtain the Hα flux and FWHM, to directly compare the BH mass from reverberation mapping to the most commonly used SE calibrators at high redshift, which typically employ the Hα and Hβ emission lines. All the SE relations that we employ are calibrated locally, but are routinely used to infer the BH masses up to $z \sim 9$ (e.g., Tripodi et al. 2024; Taylor et al. 2025).

We find that the majority of the calibrators provide a BH mass consistent within the uncertainties with the RM one for the high-mass end. However, the QSO with the lowest RM-based mass ($\log M_{\rm BH}/M_\odot \sim 7.5$) is not reproduced by any of the calibrations. In this case, several SE estimators predict a BH mass higher by more than an order of magnitude. Notably, this BH also has the highest Eddington ratio of our sample, reaching close to the super Eddington limit ($\lambda_{\rm EDD} = 0.8 \pm 0.3$). Although based on our limited sample, these results underlined the possibility of a bias in the estimation of the BH masses for the low-mass, highly accreting BHs, highlighting the need for kinematic and RM measurements at lower masses and higher redshifts to validate the SE scaling relations across the range in BH masses and cosmic times.

Finally, we exploited our new measures to provide the first SE mass estimators based on Hα and Hβ calibrated at high redshift, by combining our measurements with the dynamical BH masses obtained at high redshift from GRAVITY. Although our limited sample of a few targets with $\log(M_{\rm BH}/M_\odot) \sim 7.5$–10 shows no evidence for evolution in the SE BH mass scaling relation at $z \sim 2$ within the large uncertainties, a larger sample with masses spanning a broader range is needed to draw a more statistically robust conclusion.

Acknowledgements. BT, SC, GV, SZ acknowledge support by European Union's HE ERC Starting Grant No. 101040227 - WINGS. MP acknowledges support through the grants PID2021-127718NB-I00, PID2024-159902NA-I00, and RYC2023-044853-I, funded by the Spain Ministry of Science and Innovation/State Agency of Research MCIN/AEI/10.13039/501100011033 and El

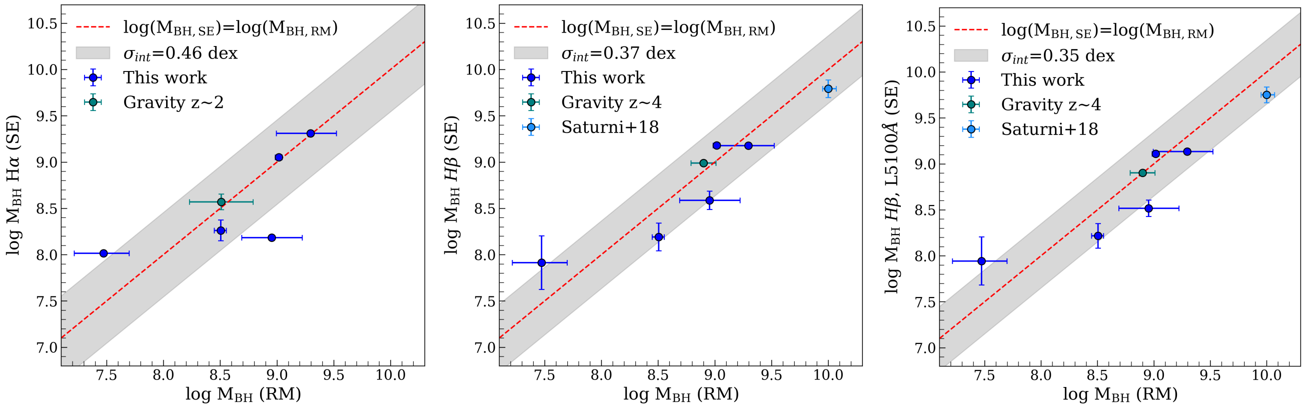

Fig. 12. Comparison between the RM (or kinematic) BH mass and the SE-BH mass derived in this work. From left to right, we report the SE M_{BH} estimated from the $L_{H\alpha}$ and $FWHM_{H\alpha}$, $L_{H\beta}$ and $FWHM_{H\beta}$, and $L_{5100Å}$ and $FWHM_{H\beta}$, respectively. The error bars shown represent only the statistical uncertainties on the measured FWHM and luminosities and do not include the intrinsic scatter and the errors on the parameters *a*, *b* and *c*. Blue points show the targets from this work, the teal points refer to the QSO with kinematic BH-mass measurements from GRAVITY, (Abuter et al. 2024; GRAVITY+ Collaboration et al. 2025), and the light blue point refers to the target reported in Saturni et al. (2018). The dashed red line shows the 1:1 relation, while the gray-shaded area shows the intrinsic scatter derived from the relation.

Fondo Social Europeo Plus FSE+. GT, acknowledges funding by the European Union (ERC Advanced Grant GALPHYS, 101055023). HÜ acknowledges funding by the European Union (ERC APEX, 101164796). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Abuter, R., Allouche, F., Amorim, A., et al. 2024, Nature, 627, 281
Akins, H. B., Casey, C. M., Lambrides, E., et al. 2025, ApJ, 991, 37
Amorim, A., Bourdarot, G., Brandner, W., et al. 2024, Astronomy & Astrophysics, 684, A167
Bentz, M. C., Peterson, B. M., Netzer, H., Pogge, R. W., & Vestergaard, M. 2009, ApJ, 697, 160
Bertemes, C., Wylezalek, D., Rupke, D. S. N., et al. 2025, A&A, 693, A176
Böker, T., Arribas, S., Lützgendorf, N., et al. 2022, A&A, 661, A82
Bosman, S. E. I., Álvarez-Márquez, J., Davies, F. B., et al. 2025, arXiv e-prints, arXiv:2511.02902
Cenci, E. & Habouzit, M. 2025, MNRAS, 542, 2597
Dalla Bontà, E., Peterson, B. M., Bentz, M. C., et al. 2020, ApJ, 903, 112
Dalla Bontà, E., Peterson, B. M., Grier, C. J., et al. 2025, A&A, 696, A48
Denney, K. D. 2012, ApJ, 759, 44
D'Eugenio, F., Cameron, A. J., Scholtz, J., et al. 2025a, ApJS, 277, 4
D'Eugenio, F., Maiolino, R., Perna, M., et al. 2025b, arXiv e-prints, arXiv:2503.11752
D'Eugenio, F., Pérez-González, P. G., Maiolino, R., et al. 2024, Nature Astronomy, 8, 1443
Du, P. & Wang, J.-M. 2019, ApJ, 886, 42
Faisst, A. L., Fujimoto, S., Tsujita, A., et al. 2025, arXiv e-prints, arXiv:2510.16111
Ferland, G., Chatzikos, M., Guzmán, F., et al. 2017, Revista mexicana de astronomía y astrofísica, 53
Gardner, J. P., Mather, J. C., Abbott, R., et al. 2023, PASP, 135, 068001
Gebhardt, K., Kormendy, J., Ho, L. C., et al. 2000, ApJ, 543, L5
GRAVITY Collaboration, Abuter, R., Accardo, M., et al. 2017, A&A, 602, A94
Gravity+ Collaboration, Abuter, R., Alarcon, P., et al. 2022, The Messenger, 189, 17
GRAVITY Collaboration, Amorim, A., Bauböck, M., et al. 2020, A&A, 643, A154
GRAVITY Collaboration, Amorim, A., Bourdarot, G., et al. 2024, A&A, 684, A167
GRAVITY+ Collaboration, El Dayem, K. A., Aimar, N., et al. 2025, arXiv e-prints, arXiv:2509.13911
Gravity Collaboration, Sturm, E., Dexter, J., et al. 2018, Nature, 563, 657
Greene, J. E. & Ho, L. C. 2005, ApJ, 630, 122
Harikane, Y., Zhang, Y., Nakajima, K., et al. 2023, ApJ, 959, 39
Heckman, T. M. & Best, P. N. 2014, Annual Review of Astronomy and Astrophysics, 52, 589
Jakobsen, P., Ferruit, P., Alves de Oliveira, C., et al. 2022, A&A, 661, A80
Juodžbalis, I., Maiolino, R., Baker, W. M., et al. 2024, Nature, 636, 594
Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, ApJ, 533, 631
King, A. 2024, MNRAS, 531, 550
King, A. 2025, MNRAS, 536, L1
Kokorev, V., Fujimoto, S., Labbe, I., et al. 2023, ApJ, 957, L7
Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv preprint arXiv:1711.03234
Kormendy, J. & Ho, L. C. 2013, ARA&A, 51, 511
Lambrides, E., Garofali, K., Larson, R., et al. 2024, arXiv e-prints, arXiv:2409.13047
Leung, G. C. K., Finkelstein, S. L., Pérez-González, P. G., et al. 2025, ApJ, 992, 26
Liao, M., Assef, R. J., Tsai, C.-W., et al. 2025, arXiv e-prints, arXiv:2504.13409
Lupi, A., Trinca, A., Volonteri, M., Dotti, M., & Mazzucchelli, C. 2024, A&A, 689, A128
Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285
Maiolino, R., Scholtz, J., Curtis-Lake, E., et al. 2024, A&A, 691, A145
Markwardt, C. B. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 411, Astronomical Data Analysis Software and Systems XVIII, ed. D. A. Bohlender, D. Durand, & P. Dowler, 251
Marshall, M. A., Windhorst, R. A., Ferrami, G., et al. 2025, arXiv e-prints, arXiv:2502.20550
Matthee, J., Naidu, R. P., Brammer, G., et al. 2023, arXiv e-prints, arXiv:2306.05448
McDougall, H., Davis, T. M., Yu, Z., et al. 2025, arXiv e-prints, arXiv:2512.01261
Naidu, R. P., Matthee, J., Katz, H., et al. 2025, arXiv e-prints, arXiv:2503.16596
Natarajan, P., Pacucci, F., Ricarte, A., et al. 2024, ApJ, 960, L1
Nguyen, D. D., Ngo, H. N., Cappellari, M., et al. 2025, arXiv e-prints, arXiv:2509.20519
Osterbrock, D. E. & Ferland, G. J. 2006, Astrophysics Of Gas Nebulae and Active Galactic Nuclei (University science books)
Perna, M., Arribas, S., Marshall, M., et al. 2023, A&A, 679, A89
Peterson, B. M., Ferrarese, L., Gilbert, K., et al. 2004, ApJ, 613, 682
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Reines, A. E., Greene, J. E., & Geha, M. 2013, ApJ, 775, 116
Reines, A. E. & Volonteri, M. 2015, ApJ, 813, 82
Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, ApJS, 166, 470
Saturni, F. G., Bischetti, M., Piconcelli, E., et al. 2018, A&A, 617, A118
Scholtz, J., Carniani, S., Parlanti, E., et al. 2025, arXiv e-prints, arXiv:2510.01034
Shajib, A. J., Treu, T., Melo, A., et al. 2025, arXiv e-prints, arXiv:2507.03746
Shen, Y. 2013, arXiv preprint arXiv:1302.2643
Shen, Y., Grier, C. J., Horne, K., et al. 2024, ApJS, 272, 26
Shen, Y., Hall, P. B., Horne, K., et al. 2019, ApJS, 241, 34
Taylor, A. J., Kokorev, V., Kocevski, D. D., et al. 2025, ApJ, 989, L7
Torralba, A., Matthee, J., Pezzulli, G., et al. 2025, arXiv e-prints, arXiv:2510.00103
Trefoloni, B., Ji, X., Maiolino, R., et al. 2025, Astronomy & Astrophysics, 700, A203
Trinca, A., Valiante, R., Schneider, R., et al. 2024, arXiv e-prints, arXiv:2412.14248
Tripodi, R., Martis, N., Markov, V., et al. 2024, arXiv e-prints, arXiv:2412.04983
Übler, H., Maiolino, R., Curtis-Lake, E., et al. 2023, A&A, 677, A145
Vestergaard, M. & Peterson, B. M. 2006, ApJ, 641, 689
Wang, B., Leja, J., de Graaff, A., et al. 2024, ApJ, 969, L13
Wu, Q. & Shen, Y. 2022, ApJS, 263, 42

Appendix A: Extended S-flat

In Figure A.1, we show the comparison between a nominal and an extended S-flat file. All the modified S-flat files provided as machine-readable versions can be found at https://github.com/eleonoraparlanti/nirspecIFU-extended.

Appendix B: Flux calibration G235M/F170LP

Following the same approach as for the G140M/F100LP, we extended the G235M/F170LP up to 5.3 μm, that is, the wavelength range encompassed by the G235M/F170LP and the G395M/F290LP. Since the flats are extended with ones the reduced spectrum is contaminated by the presence of the second and third order spectra, we require observations of the same targets with the G235M/F170LP and the G395M/F290LP to compute the correction factors $\tilde{\alpha}(\lambda)$, $\tilde{\beta}(\lambda)$ adn $k(\lambda)$ to calibrate the flux. Before computing them, we verified any possible dependency on the position of the source on the detector and the observation time. Figure B.1 shows the same as Figure 2, but for the extended G235M. Even in this case, by exploiting various observations of the same standard star, we noticed no major dependence on the cycle in which it was observed, nor on the position where the spectrum fell on the detector.

Figure B.2 shows the same as Figure 4, but for the extended G235M, highlighting that flux calibration remains below 20%.

Appendix C: Wavelength calibration

We also verify the wavelength calibration beyond the nominal wavelength range by measuring the centroid of various emission lines that are detected both in the extended spectra and in the nominal wavelength range. Fig. C.1 reports the difference between the two measured centroids, highlighting that the observed variations are well within the spectral resolution typically compatible with zero.

Appendix D: Spectral fitting best-fit results

In this section, we show the spectrum and the best-fit modelling of the other four targeted in this work. Figure D.1 shows the Hβ, [OIII] complex (left) and Hα-[NII] complex on the right panels for each target. The best-fit values of luminosities and FWHM for each target are reported in Table D.1.

Appendix E: Single Epoch relation calibrations

In this section, we show the posterior distributions of the free parameters used to derive the single-epoch scaling relation at high redshift. Figures E.1, E.2, and E.3 display the corner plots posterior distributions for the Hα, Hβ FWHM and Hβ luminosity, and Hβ FWHM and 5100Å luminosity, respectively.

Appendix F: Verification of our SE-relations

In this section, we compare the M_{BH} obtained with our SE-relations with those from VP06 and R13 for a large sample of local QSOs to investigate any sign of evolution and difference. Figure F.1 and Figure F.2 show the comparison for our Hα- and Hβ- based calibrators as a function of the continuum luminosity measured at 5100Å, as a proxy for the bolometric luminosity. In general, we find that the R13 and the VP06 relations, respectively for Hα and Hβ, tend to underestimate M_{BH} at low luminosities ($\log(L_{5100Å}/(\mathrm{erg\,s^{-1}})) \lesssim 45.25$) and overestimate it at high luminosities, with respect to our calibration. This effect is minor for the Hα, with the difference being on average $\lesssim$0.1 dex, whereas it is more apparent for the Hβ, where the average difference reaches ~0.3 dex. Obviously, a larger sample is desirable to investigate the robustness of our calibration, overcome possible biases due to the limited representativity of the QSO population, and shrink the statistical uncertainties. While on-going and future high-redshift RM campaigns will provide larger data-sets to anchor SE calibrations (e.g. the ‘Black hole mapper’ program, Kollmeier et al. 2017), we speculate that this behavior might be due to a physical effect, such as a flattening of the R–L relation at the high-luminosity end (e.g. Amorim et al. 2024).

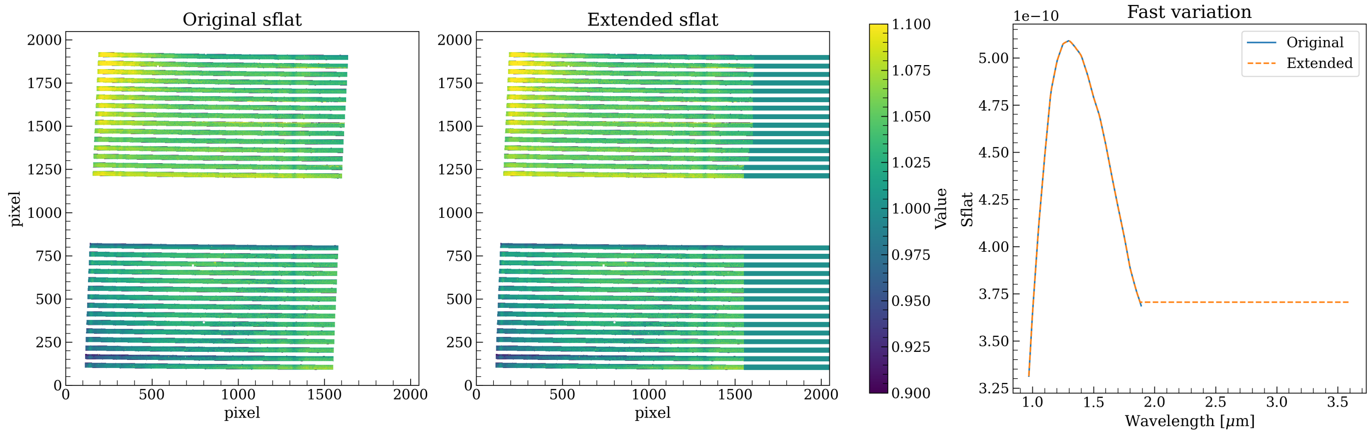

Fig. A.1. Detector response of S-flat for NRS1 before (left) and after (center) the extension. In the right panel, we show the original "FAST_VARIATION" table as a function of wavelength in blue and the extended one in orange.

Table D.1. Best fit results from the integrated spectra.

Target	$\log\left(\frac{\lambda L(5100\text{Å})}{\text{ergs}^{-1}}\right)$	$\log\left(\frac{L(H\beta)}{\text{erg s}^{-1}}\right)$	$\log\left(\frac{L(H\alpha)}{\text{erg s}^{-1}}\right)$	FWHM Hα [km s^{-1}]	FWHM Hβ [km s^{-1}]
RM032	44.94 ± 0.02	43.25 ± 0.01	43.77 ± 0.01	2343 ± 4	2975 ± 595
RM312	44.80 ± 0.02	42.98 ± 0.07	43.46 ± 0.03	6611 ± 50	7452 ± 121
RM332	45.16 ± 0.02	43.64 ± 0.03	44.08 ± 0.01	2486 ± 4	4418 ± 301
RM401	45.46 ± 0.01	43.61 ± 0.01	44.22 ± 0.01	6615 ± 24	6686 ± 75
RM549	44.79 ± 0.06	42.98 ± 0.09	43.16 ± 0.03	3530 ± 120	3769 ± 387

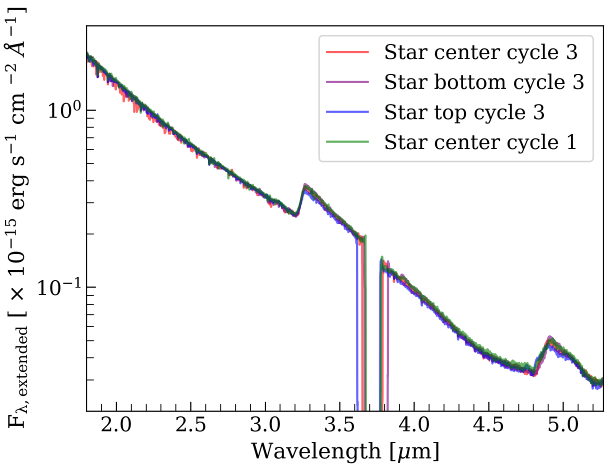

Fig. B.1. Observed flux in the nominal wavelength range of the extended G235M/F170LP filter for the star P330-E observed as part of the proposals 1538 and 6645. The curves of different colors show the same flux for the star observed in cycle 1 (green), in cycle 3 at the center of the FOV (red), and in the bottom and top parts of the FOV (blue and purple, respectively).

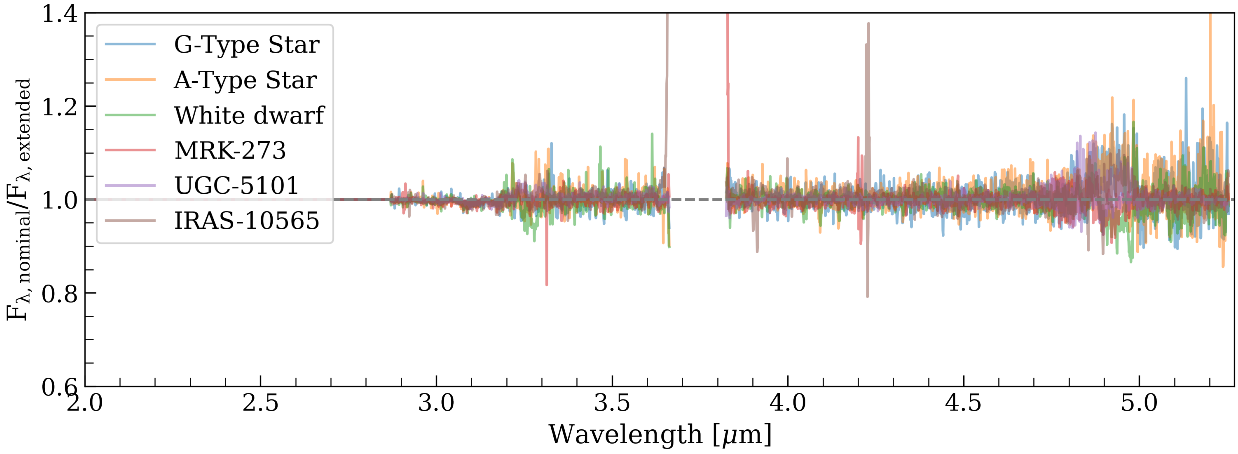

Fig. B.2. Ratio between the observed flux in the nominal wavelength range of the G235M/F170LP and G395M/F290LP and the flux observed with the extended G235M/F170LP.

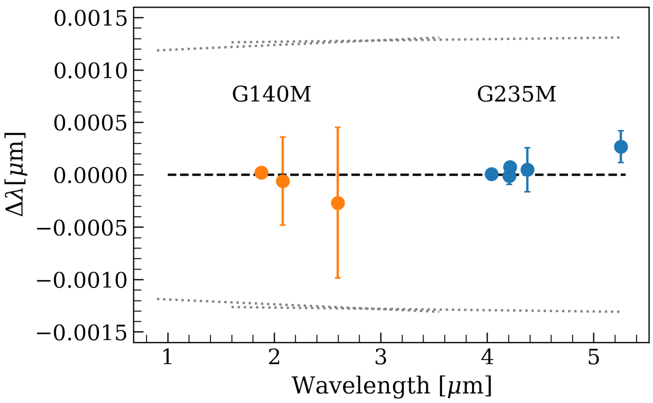

Fig. C.1. Comparison of the line centroids measured in the extended data and grating in the nominal range. In orange and blue, we report the measurements for the extended G140M/F100LP, and G235M/F170LP, respectively. The dotted lines represent the spectral resolution from Shajib et al. (2025).

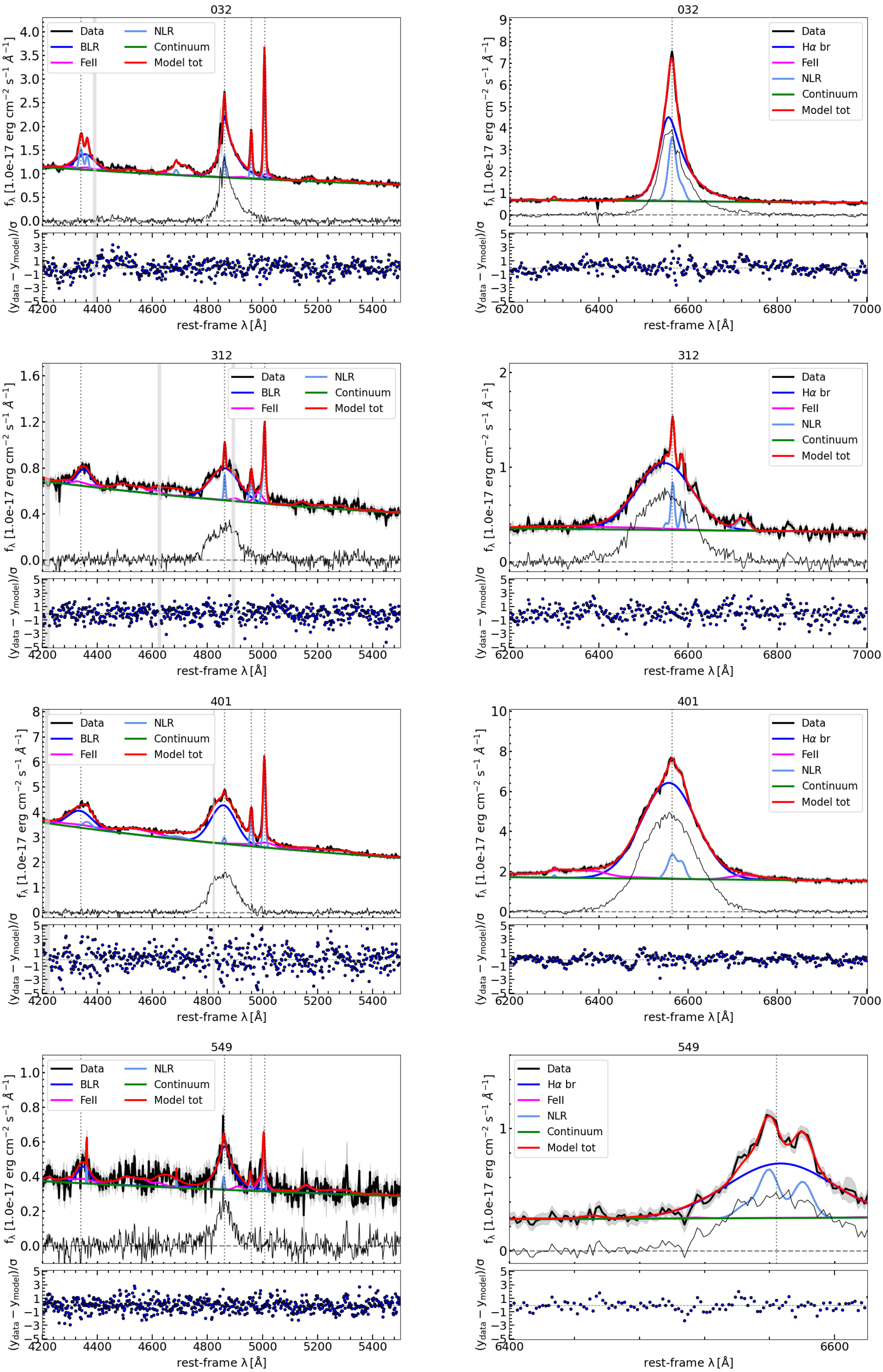

Fig. D.1. Same as Fig. 9 for the targets RM032, RM312, RM401 and RM539.

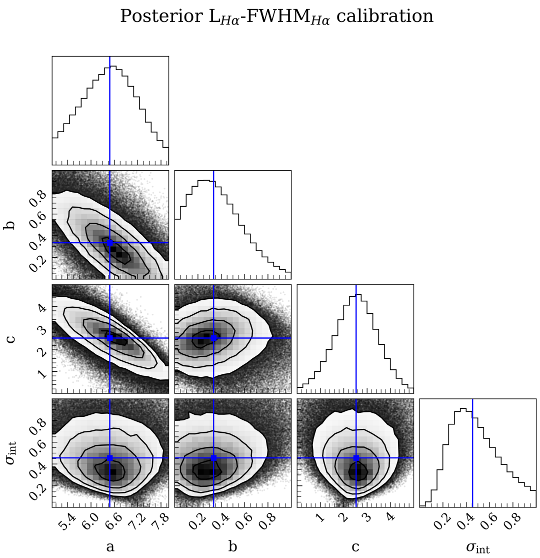

Fig. E.1. Posterior distributions of the parameters a, b, c, and σ_{int} for the single epoch BH-mass estimation derived in this work by exploiting the $L_{H\alpha}$ and $FWHM_{H\alpha}$.

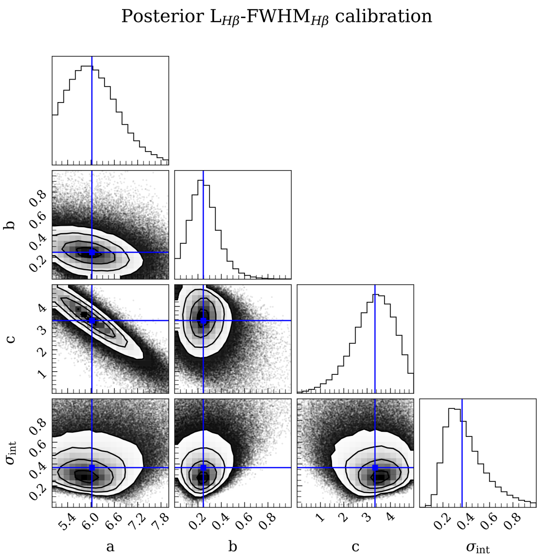

Fig. E.2. Same as Fig. E.1, but for Hβ.

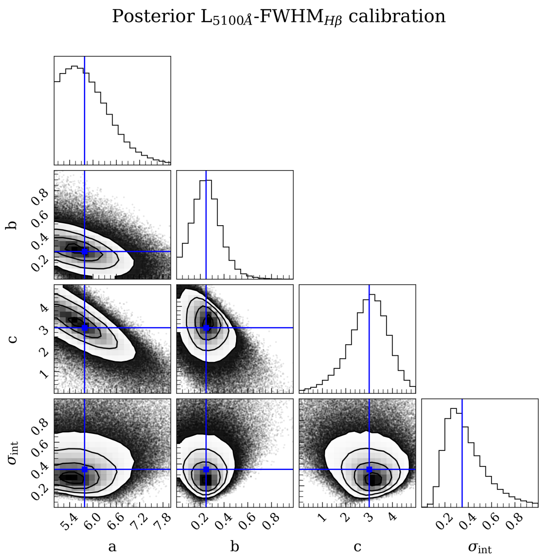

Fig. E.3. Same as Fig. E.1, but for Hβ FWHM and 5100Å luminosity.

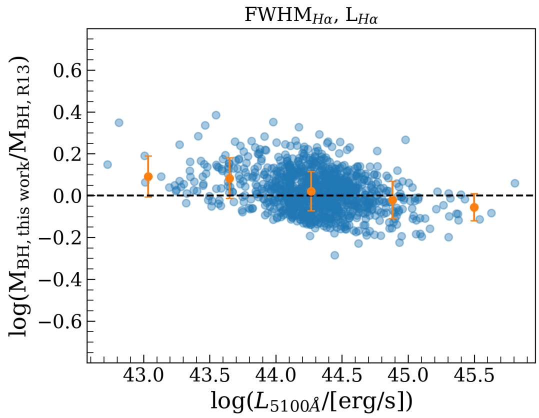

Fig. F.1. Difference between M_{BH} inferred from our Hα-based relation and the one from R13 as function of the continuum luminosity. Blue points show the value for each QSO, while orange points show the mean and standard deviation for each luminosity bin.

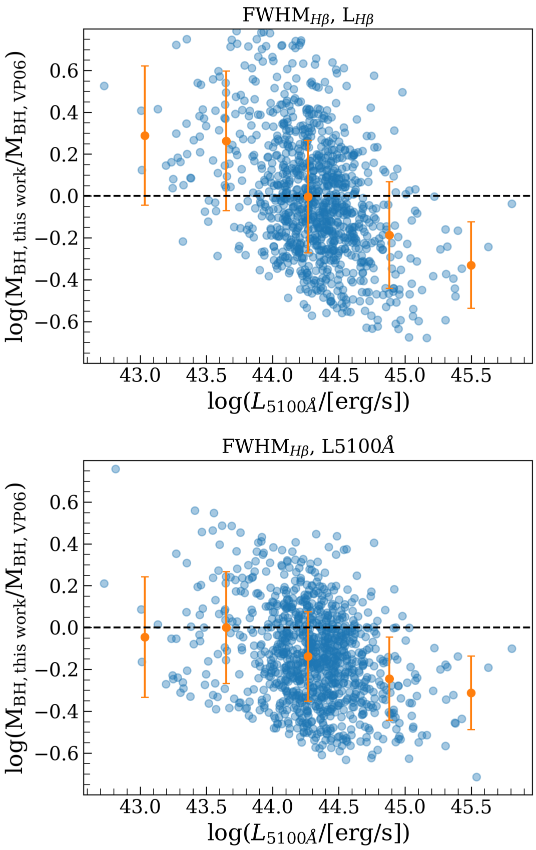

Fig. F.2. Same as Fig.F.1 but for Hβ-based relations compared to VP06.