Citation

Lusso, E.; Fotopoulou, S.; Selwood, M.; Allevato, V.; Calderone, G.; Mancini, C.; Mignoli, M.; Scodeggio, M.; Bisigello, L.; Feltre, A.; Ricci, F.; La Franca, F.; Vergani, D.; Gabarra, L.; Le Brun, V.; Maiorano, E.; Palazzi, E.; Moresco, M.; Zamorani, G.; Cresci, G.; Jahnke, K.; Humphrey, A.; Landt, H.; Mannucci, F.; Marconi, A.; Pozzetti, L.; Salucci, P.; Salvato, M.; Shankar, F.; Spinoglio, L.; Stern, D.; Serjeant, S.; Aghanim, N.; Altieri, B.; Amara, A.; Andreon, S.; Auphan, T.; Auricchio, N.; Baldi, M.; Bardelli, S.; Bender, R.; Bonino, D.; Branchini, E.; Brescia, M.; Brinchmann, J.; Camera, S.; Capobianco, V.; Carbone, C.; Carretero, J.; Casas, S.; Castellano, M.; Cavuoti, S.; Cimatti, A.; Congedo, G.; Conselice, C. J.; Conversi, L.; Copin, Y.; Corcione, L.; Courbin, F.; Courtois, H. M.; Dinis, J.; Dubath, F.; Duncan, C. A. J.; Dupac, X.; Dusini, S.; Farina, M.; Farrens, S.; Ferriol, S.; Fourmanoit, N.; Frailis, M.; Franceschi, E.; Franzetti, P.; Fumana, M.; Galeotta, S.; Garilli, B.; Gillard, W.; Gillis, B.; Giocoli, C.; Grazian, A.; Grupp, F.; Haugan, S. V. H.; Holmes, W.; Hook, I.; Hormuth, F.; Hornstrup, A.; Kümmel, M.; Keihänen, E.; Kermiche, S.; Kubik, B.; Kunz, M.; Kurki-Suonio, H.; Ligori, S.; Lilje, P. B.; Lindholm, V.; Lloro, I.; Mansutti, O.; Marggraf, O.; Markovic, K.; Martinet, N.; Marulli, F.; Massey, R.; Medinaceli, E.; Mei, S.; Mellier, Y.; Merlin, E.; Meylan, G.; Moscardini, L.; Munari, E.; Niemi, S.-M.; Padilla, C.; Paltani, S.; Pasian, F.; Pedersen, K.; Percival, W. J.; Pettorino, V.; Polenta, G.; Poncet, M.; Popa, L. A.; Raison, F.; Rebolo, R.; Renzi, A.; Rhodes, J.; Riccio, G.; Romelli, E.; Roncarelli, M.; Rossetti, E.; Saglia, R.; Sapone, D.; Sartoris, B.; Schneider, P.; Secroun, A.; Seidel, G.; Serrano, S.; Sirignano, C.; Sirri, G.; Stanco, L.; Surace, C.; Tallada-Crespí, P.; Taylor, A. N.; Teplitz, H. I.; Tereno, I.; Toledo-Moreo, R.; Torradeflot, F.; Tutusaus, I.; Valentijn, E. A.; Valenziano, L.; Vassallo, T.; Veropalumbo, A.; Vibert, D.; Wang, Y.; Weller, J.; Zoubian, J.; Zucca, E.; Biviano, A.; Bolzonella, M.; Bozzo, E.; Burigana, C.; Colodro-Conde, C.; Di Ferdinando, D.; Graciá-Carpio, J.; Mainetti, G.; Mauri, N.; Neissner, C.; Sakr, Z.; Scottez, V.; Tenti, M.; Viel, M.; Wiesmann, M.; Akrami, Y.; Anselmi, S.; Baccigalupi, C.; Ballardini, M.; Bethermin, M.; Borgani, S.; Borlaff, A. S.; Bruton, S.; Cabanac, R.; Calabro, A.; Cappi, A.; Carvalho, C. S.; Castignani, G.; Castro, T.; Cañas-Herrera, G.; Chambers, K. C.; Cooray, A. R.; Coupon, J.; Cucciati, O.; Davini, S.; De Lucia, G.; Desprez, G.; Di Domizio, S.; Dole, H.; Díaz-Sánchez, A.; Escartin Vigo, J. A.; Escoffier, S.; Ferrero, I.; Ganga, K.; García-Bellido, J.; Giacomini, F.; Gozaliasl, G.; Guinet, D.; Hall, A.; Hildebrandt, H.; Jiminez Muñoz, A.; Kajava, J. J. E.; Kansal, V.; Kirkpatrick, C. C.; Legrand, L.; Loureiro, A.; Macias-Perez, J.; Magliocchetti, M.; Maoli, R.; Martinelli, M.; Martins, C. J. A. P.; Matthew, S.; Maturi, M.; Maurin, L.; Metcalf, R. B.; Migliaccio, M.; Monaco, P.; Morgante, G.; Nadathur, S.; Patrizii, L.; Pezzotta, A.; Popa, V.; Porciani, C.; Potter, D.; Pöntinen, M.; Rocci, P.-F.; Sánchez, A. G.; Schneider, A.; Sefusatti, E.; Sereno, M.; Shulevski, A.; Simon, P.; Spurio Mancini, A.; Stadel, J.; Stanford, S. A.; Steinwagner, J.; Testera, G.; Teyssier, R.; Toft, S.; Tosi, S.; Troja, A.; Tucci, M.; Valieri, C.; Valiviita, J. and Zinchenko, I. A. (2024). Euclid preparation : XXXVIII. Spectroscopy of active galactic nuclei with NISP. Astronomy

& Astrophysics, 685, article no. A108.

Euclid preparation

XXXVIII. Spectroscopy of active galactic nuclei with NISP[⋆]

Euclid Collaboration: E. Lusso[1,2], S. Fotopoulou[3], M. Selwood[3], V. Allevato[4], G. Calderone[5], C. Mancini[6], M. Mignoli[7],
M. Scodeggio[6], L. Bisigello[8,9], A. Feltre[1], F. Ricci[10,11], F. La Franca[10], D. Vergani[7], L. Gabarra[8,12], V. Le Brun[13], E. Maiorano[7],
E. Palazzi[7], M. Moresco[14,7], G. Zamorani[7], G. Cresci[1], K. Jahnke[15], A. Humphrey[16,17], H. Landt[18], F. Mannucci[1], A. Marconi[2,1],
L. Pozzetti[7], P. Salucci[19,20], M. Salvato[21], F. Shankar[22], L. Spinoglio[23], D. Stern[24], S. Serjeant[25], N. Aghanim[26], B. Altieri[27],
A. Amara[28], S. Andreon[29], T. Auphan[30], N. Auricchio[7], M. Baldi[31,7,32], S. Bardelli[7], R. Bender[21,33], D. Bonino[34],
E. Branchini[35,36,29], M. Brescia[37,4,38], J. Brinchmann[16], S. Camera[39,40,34], V. Capobianco[34], C. Carbone[6], J. Carretero[41,42],
S. Casas[43], M. Castellano[11], S. Cavuoti[4,38], A. Cimatti[44], G. Congedo[45], C. J. Conselice[46], L. Conversi[47,27], Y. Copin[48],
L. Corcione[34], F. Courbin[49], H. M. Courtois[50], J. Dinis[51,52], F. Dubath[53], C. A. J. Duncan[46,54], X. Dupac[27], S. Dusini[12], M. Farina[23],
S. Farrens[55], S. Ferriol[48], N. Fourmanoit[30], M. Frailis[5], E. Franceschi[7], P. Franzetti[6], M. Fumana[6], S. Galeotta[5], B. Garilli[6],
W. Gillard[30], B. Gillis[45], C. Giocoli[7,56], A. Grazian[9], F. Grupp[21,57], S. V. H. Haugan[58], W. Holmes[24], I. Hook[59], F. Hormuth[60],
A. Hornstrup[61,62], M. Kümmel[33], E. Keihänen[63], S. Kermiche[30], B. Kubik[48], M. Kunz[64], H. Kurki-Suonio[65,66], S. Ligori[34],
P. B. Lilje[58], V. Lindholm[65,66], I. Lloro[67], O. Mansutti[5], O. Marggraf[68], K. Markovic[24], N. Martinet[13], F. Marulli[14,7,32],
R. Massey[69], E. Medinaceli[7], S. Mei[70], Y. Mellier[71,72], E. Merlin[11], G. Meylan[49], L. Moscardini[14,7,32], E. Munari[5], S.-M. Niemi[73],
C. Padilla[41], S. Paltani[53], F. Pasian[5], K. Pedersen[74], W. J. Percival[75,76,77], V. Pettorino[78], G. Polenta[79], M. Poncet[80], L. A. Popa[81],
F. Raison[21], R. Rebolo[82,83], A. Renzi[8,12], J. Rhodes[24], G. Riccio[4], E. Romelli[5], M. Roncarelli[7], E. Rossetti[31], R. Saglia[33,21],
D. Sapone[84], B. Sartoris[33,5], P. Schneider[68], A. Secroun[30], G. Seidel[15], S. Serrano[85,86,87], C. Sirignano[8,12], G. Sirri[32],
L. Stanco[12], C. Surace[13], P. Tallada-Crespí[88,42], A. N. Taylor[45], H. I. Teplitz[89], I. Tereno[52,90], R. Toledo-Moreo[91], F. Torradeflot[42,88],
I. Tutusaus[92], E. A. Valentijn[93], L. Valenziano[7,94], T. Vassallo[33,5], A. Veropalumbo[29,36], D. Vibert[13], Y. Wang[89], J. Weller[33,21],
J. Zoubian[30], E. Zucca[7], A. Biviano[5,95], M. Bolzonella[7], E. Bozzo[53], C. Burigana[96,94], C. Colodro-Conde[82], D. Di Ferdinando[32],
J. Graciá-Carpio[21], G. Mainetti[97], N. Mauri[44,32], C. Neissner[41,42], Z. Sakr[98,92,99], V. Scottez[71,100], M. Tenti[32], M. Viel[95,5,19,20,101],
M. Wiesmann[58], Y. Akrami[102,103], S. Anselmi[8,12,104], C. Baccigalupi[19,5,20,95], M. Ballardini[105,106,7], M. Bethermin[107,13],
S. Borgani[108,95,5,20], A. S. Borlaff[109,110,111], S. Bruton[112], R. Cabanac[92], A. Calabro[11], A. Cappi[7,113], C. S. Carvalho[90],
G. Castignani[14,7], T. Castro[5,20,95,101], G. Cañas-Herrera[73,114], K. C. Chambers[115], A. R. Cooray[116], J. Coupon[53], O. Cucciati[7],
S. Davini[36], G. De Lucia[5], G. Desprez[117], S. Di Domizio[35,36], H. Dole[26], A. Díaz-Sánchez[118], J. A. Escartin Vigo[21], S. Escoffier[30],
I. Ferrero[58], K. Ganga[70], J. García-Bellido[102], F. Giacomini[32], G. Gozaliasl[119,65], D. Guinet[48], A. Hall[45], H. Hildebrandt[120],
A. Jimenez Muñoz[121], J. J. E. Kajava[122,123], V. Kansal[124,125,126], C. C. Kirkpatrick[63], L. Legrand[64], A. Loureiro[127,128],
J. Macias-Perez[121], M. Magliocchetti[23], R. Maoli[129,11], M. Martinelli[11,130], C. J. A. P. Martins[131,16], S. Matthew[45], M. Maturi[98,132],
L. Maurin[26], R. B. Metcalf[14,7], M. Migliaccio[133,134], P. Monaco[108,5,20,95], G. Morgante[7], S. Nadathur[28], L. Patrizii[32], A. Pezzotta[21],
V. Popa[81], C. Porciani[68], D. Potter[135], M. Pöntinen[65], P.-F. Rocci[26], A. G. Sánchez[21], A. Schneider[135], E. Sefusatti[5,95,20],
M. Sereno[7,32], A. Shulevski[136,93,137], P. Simon[68], A. Spurio Mancini[138], J. Stadel[135], S. A. Stanford[139], J. Steinwagner[21], G. Testera[36],
R. Teyssier[140], S. Toft[62,141], S. Tosi[35,36,29], A. Troja[8,12], M. Tucci[53], C. Valieri[32], J. Valiviita[65,66], and I. A. Zinchenko[33]

(Affiliations can be found after the references)

ABSTRACT

The statistical distribution and evolution of key properties of active galactic nuclei (AGN), such as their accretion rate, mass, and spin, remains a subject of open debate in astrophysics. The ESA *Euclid* space mission, launched on July 1 2023, promises a breakthrough in this field. We create detailed mock catalogues of AGN spectra from the rest-frame near-infrared down to the ultraviolet – including emission lines – to simulate what *Euclid* will observe for both obscured (type 2) and unobscured (type 1) AGN. We concentrate on the red grisms of the NISP instrument, which will be used for the wide-field survey, opening a new window for spectroscopic AGN studies in the near-infrared. We quantify the efficiency in the redshift determination as well as in retrieving the emission line flux of the Hα+[N II] complex, as *Euclid* is mainly focused on this emission line, given that it is expected to be the brightest one in the probed redshift range. Spectroscopic redshifts are measured for 83% of the simulated AGN in the interval where the Hα is visible (i.e. $0.89 < z < 1.83$ at a line flux of $>2 \times 10^{-16}$ erg s^{-1} cm^{-2}, encompassing the peak of AGN activity at $z \simeq 1-1.5$) within the spectral coverage of the red grism. Outside this redshift range, the measurement efficiency decreases significantly. Overall, a spectroscopic redshift is correctly determined for about 90% of type 2 AGN down to an emission line flux of roughly 3×10^{-16} erg s^{-1} cm^{-2}, and for type 1 AGN down to 8.5×10^{-16} erg s^{-1} cm^{-2}. Recovered values for black hole mass show a small offset with respect to the input values by about 10%, but the agreement is good overall. With such a high spectroscopic coverage at $z < 2$, we will be able to measure AGN demography, scaling relations, and clustering from the epoch of the peak of AGN activity down to the present-day Universe for hundreds of thousands of AGN with homogeneous spectroscopic information.

Key words. galaxies: active – quasars: general – quasars: supermassive black holes

---

1. Introduction

Active galactic nuclei (AGN) are signposts of accretion of matter onto the supermassive black holes (SMBHs) found at the centre of the majority of galaxies. AGN activity appears to be one of the key mechanisms needed to quench star formation in massive galaxies (e.g. Gabor et al. 2010), to reproduce observed galaxy properties (e.g. SMBH and stellar mass function, Granato et al. 2004; Shankar et al. 2006; Croton 2006; Croton et al. 2006), and to circumvent overproduction of very massive galaxies in cosmological simulations (e.g. Ward et al. 2022). Nevertheless, the detailed conditions under which the feedback process initiates and is delivered to the host galaxy remain a topic of active research. Moreover, the determination of the black hole mass for a very large sample of AGN can provide further evidence to bolster or counter the idea that dark matter could also contribute to SMBH growth (De Laurentis & Salucci 2022). Among the many challenges associated with AGN studies is the rarity of available observations of these sources (e.g. Mullaney et al. 2012) and there is currently no universal selection criterion to identify them (Lacy et al. 2004; Stern et al. 2005, 2012; Donley et al. 2012; Kirkpatrick et al. 2012; Berta et al. 2013; Ji et al. 2022), which often leads to controversial results.

Euclid, the European Space Agency mission launched on July 1st 2023 (Laureijs et al. 2011), will observe ~15 000 deg² of the extragalactic sky in the optical with VIS (the visible imager; Cropper et al. 2016) and in the near-infrared (NIR) with NISP (the Near-Infrared Spectrometer and Photometer; Maciaszek et al. 2022). About ten million AGN are expected to be observable with the *Euclid* imager (Selwood et al., in prep.; Euclid Collaboration 2024a), including over 100 high-redshift quasars at $7.0 < z < 7.5$ (Euclid Collaboration 2019). From this sample, several hundred thousand AGN will be observed spectroscopically across a wide redshift range, which will represent the largest AGN sample with NIR spectroscopy to date. Here, we present detailed mock AGN spectra from the rest-frame NIR up to the UV, including emission lines, to simulate what *Euclid* will observe for both obscured (type 2) and unobscured (type 1) AGN. When properly analysed, such spectra can reveal key properties of SMBHs, and our aim is to show that the foreseen quality of the *Euclid* spectral data will be sufficient to extract such information.

In this paper, we forecast the performance of the NISP red grism regarding the observable parameters for the spectroscopically identified AGN with *Euclid*. The NISP photometric channel (Maciaszek et al. 2016) will cover the 0.95–2.02 μm range to a 5σ point-source median depth of 24.4 AB mag (Euclid Collaboration 2022a). A spectrum will be extracted for each 5σ detection position with the NISP slitless spectrometer through two 'red' grisms, namely the reflection grating spectrometers RGS000 and RGS180, which can be tilted by $-4°$ and $+4°$ respectively to obtain four RGS orientations in total. The RGS covers the wavelength range 1.21–1.89 μm with a spectral resolution of $R = 450$ for a $0\farcs5$ diameter source and an expected unresolved line flux at a 3.5σ sensitivity of 2×10^{-16} erg s^{-1} cm^{-2} (Euclid Collaboration 2022b, 2023b). NISP is also equipped with an additional 'blue' grism (0.93–1.37 μm) at a similar spectral resolution (see Costille et al. 2018, for technical details on the NISP spectrograph), which will not be employed for the wide survey but will be used for calibration purposes in the deep survey. We present the spectral properties (mainly focussing on the Hα emission line) of the expected AGN spectra that will be observed by the NISP red grisms, which will open a new window for spectroscopic studies in the NIR. We created mock AGN spectra from the rest-frame NIR to the UV that resemble those that are expected to be observed by *Euclid* for both type 2 and type 1 AGN.

For type 1 AGN, we generated average (empirical) spectra in intervals of redshift, rest-frame equivalent width (EW), and full width at half maximum (FWHM), starting from broad-line (FWHM > 2000 km s^{-1}) blue quasars, where we conservatively excluded radio-bright and broad absorption line (BAL) objects. As a result, the type 1 AGN sample is not fully representative of the entire AGN population, but rather of its blue tail. Nonetheless, BALs would have to have very large equivalent widths (e.g. Hall et al. 2002) to impact the final stack of type 1 AGN, and therefore most (classical) BALs should be accounted for in our templates. A very small population of broad iron absorption line quasars (FeLoBALs) is missed, but this AGN class represents an already minor (~10%) AGN subpopulation (e.g. McGraw et al. 2015; Choi et al. 2022), which tends to show strong reddening in the optical (e.g. Villforth et al. 2019). Likewise, the spectral properties of radio-bright AGN (e.g. continuum and emission lines) are similar to those of radio-quiet ones in the bands we consider here, with the main differences arising in the X-ray frequencies and at radio energies, both of which are not investigated in the present analysis.

Regarding the type 2 AGN, we build a set of (semi-empirical) templates following Bisigello et al. (2021). We started from different host galaxies (e.g. elliptical, star forming, and star-burst) and emission-line components where the number density of type 2 AGNs is derived from the observed galaxy luminosity function. Therefore, the final type 2 AGN sample that we considered to construct the stacks has a realistic number of objects at different redshifts and luminosities.

Whilst we defer the reader to the relevant sections for details, we note that these two approaches are quite different. This is motivated by the fact that, in this work, we are not aiming to recover the AGN variety and classification starting from the NISP spectra, but rather to understand what the NISP performances are in terms of retrieving the emission line flux of a variety of AGN, and estimating the black hole mass for the type 1 AGN.

The paper is structured as follows. We introduce the procedure used to build the empirical and semi-empirical AGN spectra for type 1 and type 2 AGN in Sects. 2 and 3, respectively. In Sect. 4 we present the simulation used to create the *Euclid* NISP spectra from the incident AGN samples, we show how we calculated the efficiency and purity of the redshift measurements for the simulated AGN spectra in Sect. 5. Section 6 details the spectral fitting procedure employed to compute the spectral properties. Finally, we discuss the expected observable parameters and resolution limitations for the spectroscopically identified AGN with *Euclid* in Sect. 7. All the type 1 and type 2 AGN templates built as part of the present study are available at the CDS. Whenever luminosity values are reported, we assume a standard flat ΛCDM cosmology with $\Omega_{\rm m} = 0.3$ and $H_0 = 70$ km s^{-1} Mpc^{-1}.

2. Empirical template spectra for type 1 AGN

In this section, we describe how the empirical spectra for type 1 AGN are built. We created a set of composite spectra over a wide range of redshifts ($0.3 < z < 6$) and emission line properties, which are then utilised to simulate *Euclid* spectra.

Table 1. Number, N, of galaxies for the different AGN samples considered to create the spectral stacks at low redshifts ($z < 0.4$).

BinID	Intervals	N
1	$15 \leq$ EW < 30 and $2000 \leq$ FWHM < 3000	47
2	$15 \leq$ EW < 30 and $3000 \leq$ FWHM < 5000	94
3	$15 \leq$ EW < 30 and FWHM ≥ 5000	176
4	$30 \leq$ EW < 60 and $2000 \leq$ FWHM < 3000	243
5	$30 \leq$ EW < 60 and $3000 \leq$ FWHM < 5000	467
6	$30 \leq$ EW < 60 and FWHM ≥ 5000	615
7	EW ≥ 60 and $2000 \leq$ FWHM < 3000	323
8	EW ≥ 60 and $3000 \leq$ FWHM < 5000	766
9	EW ≥ 60 and FWHM ≥ 5000	716

Notes. Rest-frame equivalent widths and FWHM values for the Hβ emission line are in units of Å and km s^{-1}, respectively.

2.1. Sample selection

The quasar sample we considered to construct the empirical incident AGN templates has been built following a similar approach as the one described in Lusso & Risaliti (2016). We started with the catalogue of quasar properties presented by Shen et al. (2011), which contains 105 783 spectroscopically confirmed broad-line quasars, optically selected from the Sloan Digital Sky Survey Data Release 7 (SDSS DR7). We conservatively removed from this catalogue all sources flagged as BAL quasars (i.e., sources with BAL_FLAG=0 are non-BALs) and radio-loud (i.e., $F_\nu(6\,\mathrm{cm})/F_\nu(2500\,\text{Å}) \geq 10$, this removed 8257 quasars, or 8% of the main SDSS sample). BAL quasars are often found in galaxies with red optical and UV colours, with broad absorption troughs (associated to the presence of winds or outflows) that could distort the continuum significantly. Radio bright objects seems to show steeper spectra than the ones for radio-quiet AGN (e.g. Zheng et al. 1997), whilst other studies have found that the observed continuum slopes are flatter (redder) compared to the composite spectrum for typical (radio faint) SDSS quasars (Kuźmicz et al. 2021).

Our selection resulted in 91 732 SDSS quasars. We further excluded 136 quasars classified as BALs by Gibson et al. (2009) and 17 quasars considered radio-loud in the catalogue published by Mingo et al. (2016), which is the largest available Mid-Infrared (WISE), X-ray (3XMM) and Radio (FIRST+NVSS) collection of AGN and star-forming galaxies. This precleaned SDSS quasar sample is thus composed of 91 579 sources.

To create the first composite at low redshifts (low-z stack hereafter), we selected all the AGN with $z < 0.4$ FWHM of the Hβ and Hα emission lines larger than 2000 km s^{-1}, leading to 3505 sources. We then considered three different intervals for the FWHM (only the broad component of the modelled line, see Shen et al. 2011 for details) and for the rest-frame EW as given in Table 1. Both EW and FWHM are measured using the Hβ emission line. We then created 9 low-z spectral stacks for each of the 9 different combinations of Table 1, with an additional low-z composite considering the whole sample. The number of available spectra within the 9 intervals as defined above are also summarised in Table 1.

To increase the coverage at bluer wavelengths and to keep the emission line properties homogeneous with respect to the low-z sample, we considered AGN with the EW and the FWHM of the Hβ emission line in the same intervals as listed in Table 1, with the additional requirement that the FWHM of the Mg II emission line be larger than 2000 km s^{-1} (mid-z stacks hereafter). To speed up the calculation, all the following spectral composites are built

Table 2. AGN samples considered to create the spectral stacks at high redshifts.

BinIDs	Intervals
1/4/7	$30 \leq$ EW ≤ 40, $3800 \leq$ FWHM ≤ 4200
2/5/8	$40 \leq$ EW ≤ 45, $4000 \leq$ FWHM ≤ 4400
3/6/9	$35 \leq$ EW ≤ 45, $5100 \leq$ FWHM ≤ 6000

Notes. Rest-frame equivalent widths and FWHM values for the broad component of the Mg II emission line are in units of Å and km s^{-1}, respectively.

Table 3. Summary of the redshift ranges spanned by the four different wavelength intervals (discussed in Sect. 2.1) considered to build the optical–UV stack.

Sample	Redshift range	$\langle z \rangle$
Low-z	$0.08 \leq z \leq 0.39$	0.30
Mid-z	$0.35 \leq z \leq 0.90$	0.64
High-z	$1.50 \leq z \leq 2.25$	1.80
Far-UV (Lyα)	$1.50 \leq z \leq 4.90$	2.20

by considering 100 random spectra in each single FWHM–EW bin and 1000 spectra for the total stack in the entire FWHM–EW selected region.

To further extend the wavelength coverage (high-z stack hereafter) to include the C IV emission line, we considered the Mg II line (only the broad component) as the reference since the Hβ line is no longer available. We then considered three different intervals for the following combinations of FWHM (again, only the broad component of the modelled line) and EW, shown in Table 2.

Again this has an additional requirement that the FWHM of the C IV emission line be larger than 2000 km s^{-1}. The three FWHM–EW ranges listed in Table 2 are chosen to closely match the values of the FWHM of the Mg II line in the mid-z sample. Specifically, the first FWHM–EW interval in Table 2 roughly matches the Mg II FWHM and EW values of the BinIDs 1, 4, and 7. The same considerations exist for the other two intervals.

The Lyα region and the Lyα forest are then included in the stack by selecting the quasars with $40 \leq$ EW(Å) ≤ 55 and $5200 \leq$ FWHM(kms^{-1}) ≤ 6000 for the C IV emission line. The FWHM–EW interval above roughly matches the values of the FWHM of the C IV line in the high-z sample. A summary of the redshift ranges spanned by the four different wavelength intervals considered to build the optical-UV stack is given in Table 3.

2.2. Average spectrum construction

We follow a procedure similar to that in Lusso et al. (2015, L15 hereafter) to construct the average spectrum in each interval. Specifically, our steps are as follows.

1. We correct the quasar flux density[1] (f_λ) for Galactic reddening by adopting the $E(B - V)$ estimates from Schlegel et al. (1998, SFD) and the Galactic extinction curve from Fitzpatrick (1999) with $R_V = 3.1$. We do not correct the spectra for intrinsic dust absorption.
2. We generate a rest-frame wavelength array with fixed dispersion $\Delta\lambda = 0.3$ Å. The dispersion value was set to be large

---

[1] In the following we will use the word 'flux' to mean the flux density (i.e., flux per unit wavelength).
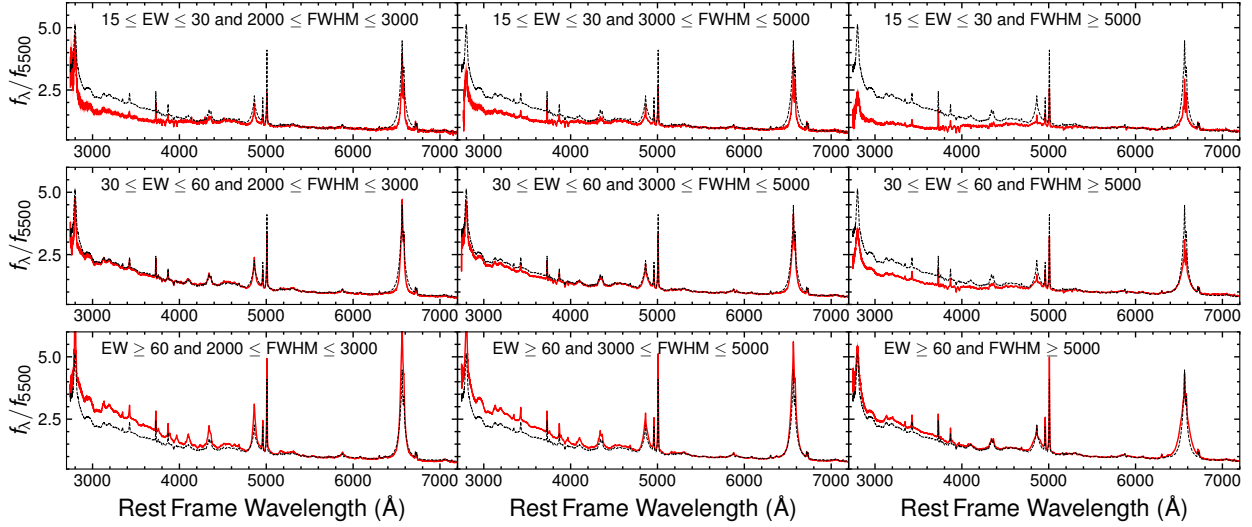
Fig. 1. Composite spectra of the low-z sample from ultraviolet to optical rest frame wavelengths. Black line: Stack obtained from the entire sample, shown as a reference. The different composites (shown with the red lines) obtained in each bin are marked as in Table 1 and are shown in separate panels. EW and FWHM are relative to the Hβ emission line and are in units of Å and km s^{-1}, respectively.

enough to include at least one entire pixel from the SDSS spectra around Lyα, where we assumed an average spectral resolution $R = 2000$ across the entire wavelength range (i.e., $\Delta\lambda \simeq 1250\,\text{Å}/2R$). This assumption slightly oversamples the spectral stack at low-z ($\Delta\lambda \simeq 1.3\,\text{Å}$ at 5100 Å).

3. Each quasar spectrum is shifted to the rest-frame and linearly interpolated over the rest-frame wavelength array with the fixed dispersion $\Delta\lambda$ defined above.

4. We normalise single spectra by their flux at rest $\lambda = 5500\,\text{Å}$ for the low-z, whilst the mid-z and high-z composites are both normalised at 2000 Å. The spectral stack at Lyα is normalised at $\lambda = 1550\,\text{Å}$.

5. All the flux values are then averaged (mean) to produce the stacked spectrum normalised to unity at the reference wavelengths above.

Uncertainties on each stack are computed as $\sigma/\sqrt{N-1}$ where σ and N are the standard deviation and the number of points in each spectral channel. All the spectral stacks are finally combined together by taking the low-z AGN composite (shown in Fig. 1) as the reference and by matching the flux in the overlapping wavelength range between two nearby spectra. The final optical/ultraviolet composite obtained from the entire sample, as well as in each bin, covers the rest-frame wavelength interval 800–7000 Å (see Fig. 2).

2.3. IGM transmission correction

Blueward of Lyα emission in the quasar rest frame, absorption from intergalactic H I attenuates the quasar flux, both in the Lyman series (creating the so-called H I forest), and in the Lyman continuum at rest $\lambda < 912\,\text{Å}$ (e.g. Moller & Jakobsen 1990). Therefore, to correct the observed stacks (Fig. 2) for the intergalactic medium (IGM) absorption by neutral hydrogen along the line-of-sight, we follow a similar (but very much simplified) approach as the one employed by L15 (see also Sects. 4 and 5 in Lusso et al. 2018).

To recover the IGM-corrected quasar emission, we constructed the IGM transmission functions (T_λ) making use of the publicly available `pyigm` package for the analysis of the

intergalactic medium[2] discussed in Prochaska et al. (2014, P14 and references therein). These functions have been computed through a cubic Hermite spline model which describes the H I absorber distribution function which, in turn, depends on both redshift and column density, i.e., $f(N_{HI}, z) = \partial^2 n/(\partial N_{HI}\partial z)$. For simplicity, we considered the default $f(N_{HI}, z)$ which is currently the P14 formulation (see their Sect. 2)[3]. Briefly, this modelling assumes that the H I forest is composed of discrete 'lines' with Doppler parameter $b = 24\,\text{km s}^{-1}$ and that the normalisation of $f(N_{HI}, z)$ evolves as $(1 + z)^2$ (Prochaska et al. 2009). This redshift evolution is consistent, within the uncertainties, with the one obtained in high-redshift surveys (Prochaska et al. 2014). Opacity due to metal line transitions was ignored since they contribute negligibly to the total absorption in the Lyman continuum.

Here we consider 8 realisations of $f(N_{HI}, z)$, and calculated T_λ in the wavelength range 500–1250 Å for the redshifts 0.3, 0.7, 1.5, 2.5, 3.5, 4.5, 7.5. We also checked how many sources are contributing to the bluer part of the total stack and what is their average redshift. There are about 450 quasars contributing more than 25 spectral channels at $\lambda < 1200\,\text{Å}$. Their average redshift is 2.9, so an additional IGM transmission function at $z = 3$ was added to the final T_λ library. Examples of the T_λ function at $z = 2.5$, $z = 3.5$, and at $z = 4.5$ are shown in Fig. 3.

The observed spectral flux ($f_{\lambda,\text{obs}}$) in the stack is finally divided by the IGM transmission curve to obtain the intrinsic corrected emission ($f_{\lambda,\text{corr}}$), i.e., $f_{\lambda,\text{corr}} = f_{\lambda,\text{obs}}/T_\lambda$.

2.4. Extending the stack to the NIR

As NISP will cover the rest-frame NIR for local AGN, we extended the composite to wavelengths longer than the SDSS spectral coverage by including 9 hard X-ray selected low-redshift type 1 AGN at $0.015 < z < 0.114$ ($K_s < 14$) from the *Swift* Burst Alert Telescope (BAT) published by Ricci et al. (2022). The AGN are observed with the folded-port infrared

---

[2] https://github.com/pyigm/pyigm

[3] The interested reader should refer to Sect. 1.7.3 at https://pyigm.readthedocs.io/_/downloads/en/latest/pdf/
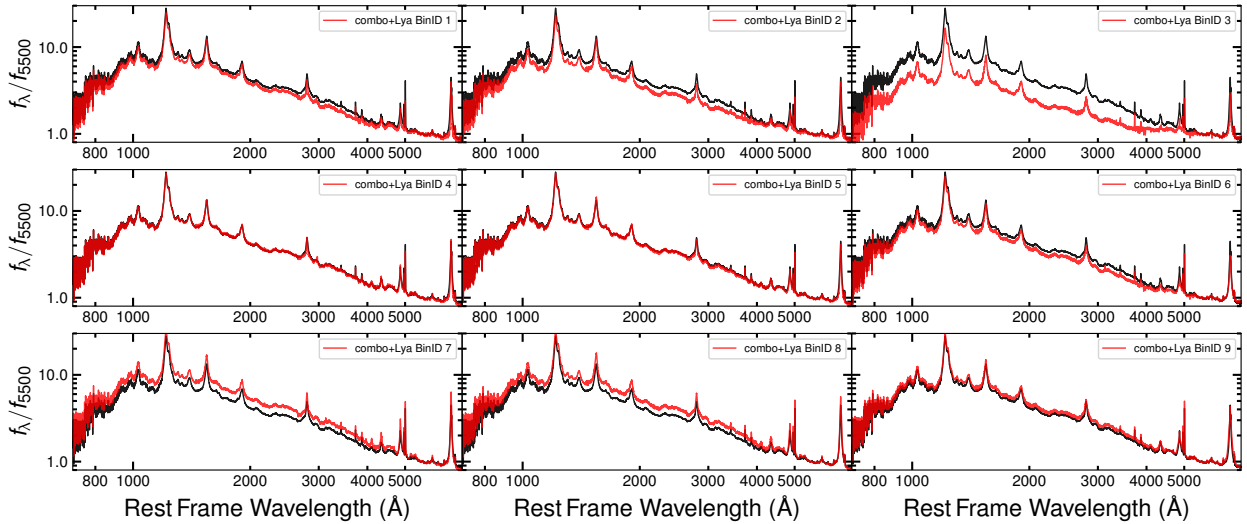
Fig. 2. Composite spectra from ultraviolet to optical rest frame wavelengths. The black line represents the stack obtained from the entire sample, shown as a reference. The red lines in the different panels represent the final stack in each bin from the combined composites from blueward of Lyα to Hα. The combined spectra covers the rest-frame wavelength interval 800–7000 Å. The spectral flattening at wavelengths longer than 4000 Å is likely due to the contribution of the host galaxy emission.
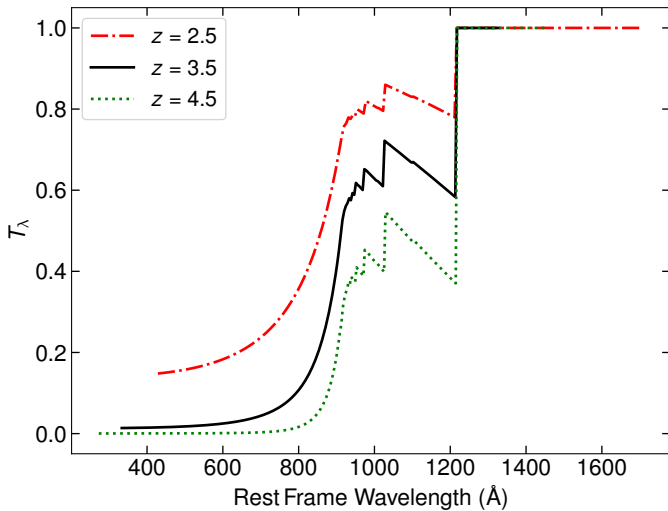
Fig. 3. Example of our adopted Lyman series opacity (T_λ) at different redshifts. The lines represent the effective optical depth due to intervening absorbers and includes the contribution of absorption in the hydrogen Lyman series and the Lyman continuum at $\lambda < 912$ Å.

echellette (FIRE, Simcoe et al. 2008) instrument in the high-resolution echelle mode (see, e.g. Ricci et al. 2022). FIRE covers the whole NIR bandpass in a single exposure, with nominal wavelength resolution of $R \approx 6000$ for a $0\farcs6$-slit width, i.e., $\Delta v \simeq 50$ km s^{-1}. In particular, from the starting sample of 13 broad-line AGN described in Ricci et al. (2022) we selected all sources classified as Seyfert 1 (Sy 1) to 1.5 (i.e. we excluded BAT ID 577) and excluded, after visual inspection, those targets whose emission lines where particularly affected by low atmospheric transmission (BAT ID 1045) and (or) showing a red spectral shape ($\alpha_\nu > 0$, where $f_\nu \propto \nu^{\alpha_\nu}$, i.e. BAT IDs 411 and 657) which is unusual for Sy1s in the NIR (Glikman et al. 2006). The list of sources is provided in Table D.1 These NIR spectra were then smoothed using a Savitzky-Golay filter, which preserves the average resolving power.

To increase the sample statistics and extend both the redshift and the luminosity range, we complemented the data above with the sample of 23 infrared spectra of type 1 AGN published by Glikman et al. (2006, see their Table 2). From the starting sample published by Glikman et al. (i.e. 27 quasars, we excluded SDSS J013418.1+001536.6 0 as the spectrum presents a strong absorption feature at 17 300 Å close to the Paα. We also removed SDSSJ130756.5+010709.6 and SDSSJ160507.9+483422.0 because both spectra do not have a common wavelength interval that can be used to normalise them to the FIRE data. This data set covers the redshift range $0.118 < z < 0.418$ with $K_s < 14.5$ and $M_i < -23$ ($\log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}}) \simeq 42.2$–46.3), and it is complementary both in terms of redshift range and luminosity with the low-z SDSS sample ($\log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}}) \simeq 44.1$–46.6).

All the spectra are averaged following the procedure described in Sect. 2.2 to produce the final stack down to Paα (see the Appendix D for further tests on additional NIR quasar datasets available in the literature). The spectral stack is then combined by taking the low-z AGN composite (shown in Fig. 1) as the reference and by matching the flux in the overlapping wavelength range between two nearby spectra. The final IGM corrected quasar composites, from the forest to the NIR obtained by stacking the entire quasar sample is shown in Fig. 4 (top panel). All the empirical spectra we constructed for the type 1 AGN (from UV to NIR) are available online as supporting material.

2.5. Creation of the mock type 1 AGN sample

Starting from the nine templates described in the previous sections, we constructed a grid of redshifts, $E(B - V)$, and bolometric luminosities to probe the expected observed parameter space of type 1 AGN: $0.1 < z < 7.0$ with a step width $\delta z = 0.1$, $0 < E(B - V) < 1.25$ with $\delta E(B - V) = 0.25$, and $42 \le \log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}}) \le 47$, with $\delta \log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}}) = 0.2$.

For each point in the grid, each of the nine templates is scaled to the bolometric luminosity using a bolometric correction value of 5.8 derived for the B-band (Duras et al.
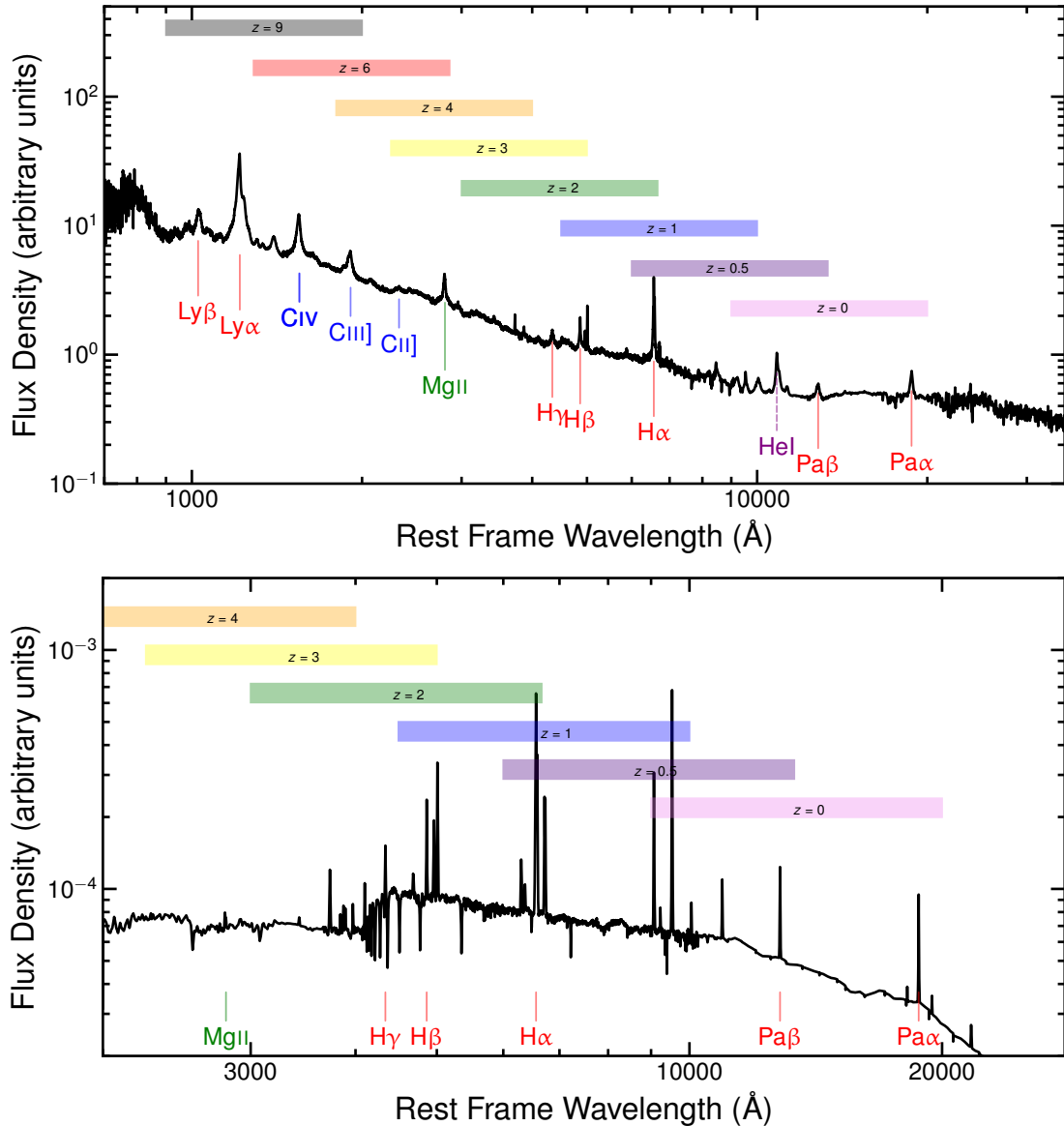
Fig. 4. Incident composite spectra from ultraviolet to near-infrared rest frame wavelengths. Final empirical stack (flux density values are reported in units of erg s^{-1} cm^{-2} Å^{-1}) for type 1 AGN covering the rest-frame wavelength range 700–36 000 Å. The main emission lines are marked. Shaded coloured regions mark the portion of the spectrum that will be covered by the low-resolution *Euclid* spectrograph up to redshift $z = 9$. All the incident spectra for the type 1 and type 2 AGN are available at the CDS. Bottom panel: final empirical stack (flux density values are reported in arbitrary units) for type 2 AGN (SB-AGN, SF-AGN, and AGN2) covering the rest-frame wavelength range 2000–28 000 Å.

2020). To account for intrinsically reddened type 1 quasars, we then apply an intrinsic extinction by assuming the SMC law (Prevot et al. 1984, as appropriate for unobscured AGN; see also Hopkins et al. 2004; Salvato et al. 2009). The spectra are then redshifted and the appropriate IGM curve of Prochaska et al. (2014) is applied. Finally, we measure the expected observed *Euclid* magnitudes in I_E, Y_E, J_E, and H_E. Each corresponding spectrum is saved and the resulting *Euclid* magnitudes are stored in a catalogue. We then keep a random subsample of 1248 spectra, across the grid, that satisfies the *Euclid* depth of $I_E < 24.5$.

3. Semi-empirical template spectra for type 2 AGN

The number density, continuum and physical properties of type 2 AGN are derived starting from the spectro-photometric

realisations of infrared-selected targets at all-z (SPRITZ, Bisigello et al. 2021). The simulation starts from the observed infrared luminosity function, complemented with the K-band luminosity function of elliptical galaxies and the galaxy stellar mass function of dwarf irregulars, and it is in agreement with a wide set of observables, including luminosity functions at several wavelengths, number counts (from radio to X-rays), the total galaxy stellar mass function, and the relation between star-formation rate (SFR) and stellar mass (see Bisigello et al. 2021, for further details).

To build the templates for the type 2 AGN, we consider three AGN populations. Specifically, we consider a first population of star-forming galaxies hosting intrinsically faint AGN (SF-AGN), and two populations hosting a powerful AGN, namely a mildly obscured AGN (classical type 2 AGN, AGN2) and a heavily obscured AGN (hosted by a starburst galaxy,

SB-AGN)[4]. The number density of both SF-AGN and SB-AGN were derived from their observed IR luminosity functions (8−1000 μm; Gruppioni et al. 2013). The number density of AGN2 were similarly derived from the same IR luminosity function, but obtained by performing a simultaneous fit of the *Herschel* (Gruppioni et al. 2013) and UV observed luminosity functions (Croom et al. 2009; McGreer et al. 2013; Ross et al. 2013; Akiyama et al. 2018; Schindler et al. 2019), as derived by Bisigello et al. (2021).

The spectral energy distributions (SEDs) of the three AGN populations were described with a set of 12 empirical templates (Polletta et al. 2007; Gruppioni et al. 2010), chosen as they describe the multi-wavelength observations of the *Herschel* AGN used to derived the starting luminosity functions. These SED templates, with the addition of emission lines (see Sect. 3.1), were used to classify the type 2 AGN population and to derive the photometric fluxes in different *Euclid* ancillary filters. The host galaxy continuum is derived from SED fitting using the MAGPHYS code (da Cunha et al. 2008), as discussed in (Bisigello et al. 2021, see their Sect. 3.2). The physical properties, such as AGN accretion luminosity and optical depth in the *V*-band (and the relative contribution of the diffuse ISM and the birth clouds to dust attenuation, as described in Charlot & Fall 2000), were derived by fitting the mentioned empirical templates with the code SED3FIT (Berta et al. 2013), which assumes two sets of equally possible AGN libraries with smooth (Fritz et al. 2006; Feltre et al. 2012) and clumpy tori (Nenkova et al. 2008a,b). Finally, the total SFR was derived by summing the UV and IR component, derived from the 1600 Å and the IR stellar continuum, assuming the conversions by Kennicutt (1998a,b).

3.1. Inclusion of emission lines

We used the AGN accretion luminosity and the SFR as normalisation factors to incorporate the emission line fluxes due to both AGN activity and star-formation. We briefly describe the main aspects related to the emission features of interest below (see Sect. 2.3 of Bisigello et al. 2021, for further details). The nebular emission was incorporated into the SPRITZ galaxy templates as follows.

- First, we considered the line emission due to star-formation. The Hα flux was derived from the SFR (Kennicutt 1998a,b). All the other hydrogen lines were, in turn, derived by assuming the case-B hydrogen recombination coefficient for an electron temperature of $T_e = 10^4$ K and an electron density of $N_e = 100$ cm^{-3}. Other metal optical lines, namely [O II]λ3727, [Ne III]λ3869, [N II]$\lambda\lambda$6548, 6584, [O III]λ5007, [S II]$\lambda\lambda$6717, 6731 and [S III]λ9069 were included using a set of empirical relations (i.e., Kennicutt 1998b; Pettini & Pagel 2004; Jones et al. 2015; Kewley et al. 2013; Dopita et al. 2016; Kashino et al. 2019; Proxauf et al. 2014; Mingozzi et al. 2020).
- Next, we accounted for the line emission due to AGN activity. We considered theoretical predictions from photoionisation models developed by Feltre et al. (2016) using the code CLOUDY (version c13.3; Ferland et al. 2013). These models reproduce the emission from the gas in the narrow-line emitting regions of AGN. Among the adjustable parameters of the entire model grid (see Sect. 2.1 in Feltre et al. 2016,
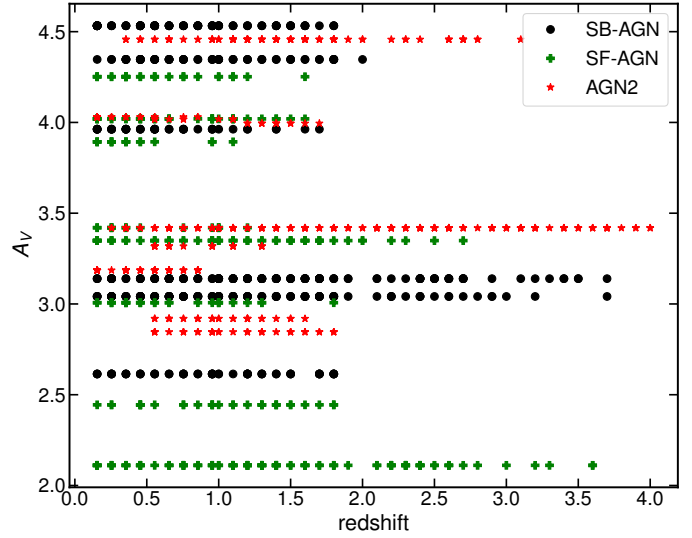

Fig. 5. Total (ISM+birth clouds, see Sect. 3) dust attenuation in the *V*-band as a function of the (input) redshift considered to build the type 2 semi-empirical library of incident spectra. The three different AGN classes are marked with different symbols as in the legend.

for details), we considered an ionisation parameter at the Strömgren radius between $\log_{10}(U_S) = -1.5$ and -3.5, metallicity from ≈ 0.5 to 2 times solar ($Z = 0.008, 0.017, 0.03$), a dust-to-metal ratio (which rules the dust depletion onto dust grains) of 0.3, a UV spectral index of the ionising radiation $\alpha = -1.4$, and an internal micro-turbulence velocity $v = 100$ km s^{-1}. The amplitude of the line fluxes were scaled to the AGN accretion luminosity derived as described above.

- Next, we considered the scattering of the line fluxes. For each line, we either include the scatter associated to each considered empirical relation referenced above (when quoted in the corresponding papers) or a generic scatter of 0.1 dex.
- We considered dust attenuation next. The total line fluxes (AGN+host) were attenuated for the presence of dust by considering the two-components (one describing the diffuse ISM and another for the dust in the birth-clouds, see Bisigello et al. 2021 for details) using the model of Charlot & Fall (2000). We thus assumed that the broad-line component is completely extincted, whilst the narrow AGN lines have the same dust attenuation as the stellar population. In this step, we also assumed that the AGN continuum has been totally extincted, and therefore the only continuum visible is the one from the host galaxy, which has been determined from the SED3FIT[5] code combining simple stellar populations from Bruzual & Charlot (2003) with the approach presented in MAGPHYS code (da Cunha et al. 2008). The dust attenuation is in the range $A_V \simeq 2.1$−4.5, with a distribution as a function of redshift shown in Fig. 5 for the three different AGN classes.
- Finally we considered the line widths. For the line emission due to star-formation, we derived the gas velocity dispersion from the stellar mass by taking into account the broadening of the nebular emission lines (Bezanson et al. 2018). For the AGN nebular emission, we assumed a FWHM in the typical range of type 2 AGN, between 500 and 800 km s^{-1}. The total

---

[4] The interested reader on details regarding the SED classification in the three different classes should refer to Sect. 2 in Gruppioni et al. (2013) (see their Fig. 1).

[5] http://steatreb.altervista.org/alterpages/sed3fit.html
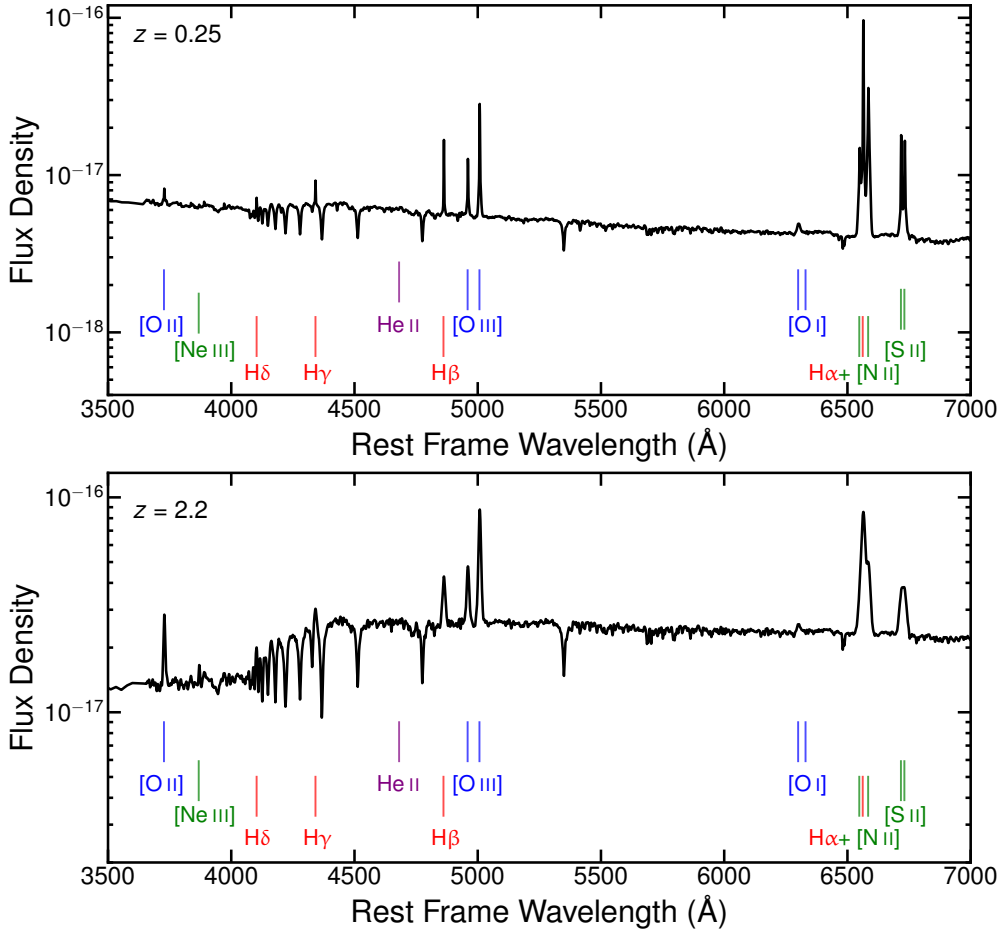
Fig. 6. Example obscured type 2 AGN incident spectra (flux density values are in units of erg s^{-1} cm^{-2} Å^{-1}) at optical wavelengths from 3500 Å to 7000 Å. Top panel: A $z = 0.25$ type 2 AGN with a bolometric luminosity $\log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}})$ of 43.9, hosted in a galaxy with a stellar mass $M_\star$ of $10^{9.2}\,M_\odot$ and SFR of $10^{0.9}\,M_\odot$ yr^{-1}. Bottom panel: Massive ($M_\star = 10^{11.8}\,M_\odot$) galaxy at $z = 2.2$ with intense star formation (SFR $= 10^{2.6}\,M_\odot$ yr^{-1}) hosting an AGN with $\log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}}) = 45.2$. The main emission lines are marked.

line width thus consider both star-formation and AGN contributions, which have been added to the continuum at rest. We finally assumed a Gaussian profile for the emission lines when creating the simulated spectra (Bisigello et al. 2021). A few examples of the resulting type 2 templates are shown in Fig. 6, zooming-in at optical wavelengths. The first is a $z = 0.25$ type 2 AGN with a bolometric luminosity $\log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}})$ of 43.9, hosted in a galaxy with a stellar mass $M_\star$ of $10^{9.2}\,M_\odot$ and SFR of $10^{0.9}\,M_\odot$ yr^{-1}. The second is a massive ($M_\star = 10^{11.8}\,M_\odot$) galaxy at $z = 2.2$ with intense star-formation (SFR $= 10^{2.6}\,M_\odot$ yr^{-1}) hosting an AGN with $\log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}}) = 45.2$.

3.2. Subsample selection

For the selection of the type 2 AGN templates, we proceed as follows.
– We considered all the galaxies, among the AGN2, SF-AGN, and SB-AGN SPRITZ populations, with AGN bolometric luminosity above a given threshold, $\log_{10}(L_{\rm bol}/{\rm erg\,s^{-1}}) > 42$ (see Sect. 3).
– We then selected all the templates with a line flux above $\log_{10}(F_{\rm line}/{\rm erg\,s^{-1}\,cm^{-2}}) > -16.2$. For the purposes of this selection, we considered specific emission lines depending on the redshift (see Table 4).

Table 4. Emission line per redshift bin chosen for the selection of AGN2 templates.

Redshift interval	Emission line
$0.15 < z \le 0.44$	Paβ
$0.44 < z \le 0.90$	[S III]λ9069
$0.90 < z \le 1.80$	Hα
$1.80 < z \le 2.80$	Hβ
$2.80 < z \le 4.00$	[O II]λ3727

– We applied a number density cut for a survey of 15 000 deg^2 resulting in 2561 templates from the SPRITZ simulations. Finally, we randomly selected (i.e. random uniform) 1248 spectra from the above sample. This random sample corresponds to the number of spectra fitted in the NISP simulation, which is sufficient for the paper's objectives.

The 1248 spectra cover 23 classes representative of different AGN and host galaxies combinations as discussed above (see Sect. 2.1.1 in Bisigello et al. 2021, for further details). Figure 4 (bottom panel) shows the spectral composite, from the Mg II to the NIR obtained by stacking the entire sample of AGN2, SF-AGN, and SB-AGN, whilst the 23 templates separated by the three classifications are presented in Fig. 7.

4. Spectra simulation

To generate a sample of simulated AGN spectra, we randomly picked 1248 sources from those selected with the previous steps described in Sect. 2.1 for type 1 AGN and the 1248 type 2 AGN described in Sect. 3.2. We then simulated two *Euclid* pointings with the NISP spectrometer, one pointing contains type 1 AGN while the other pointing contains type 2 AGN. We positioned the sources on a grid with sufficient spacing to ensure clear separation of individual spectra, while also avoiding the edges of the NISP detectors. This choice was made with the intention of maximising the number of AGN per pointing, in order to examine the NISP's ability to detect AGN based on their physical properties. It is important to note that the primary focus of this study does not involve the decontamination of spectra. However, the decontamination process is expected to be effectively addressed by the *Euclid* reference observation sequence (ROS), which incorporates various grism orientations and high-resolution direct images (further details on spectra decontamination in slitless spectroscopy are in Ryan et al. 2018).

4.1. Configuration of the NISP-S simulator: TIPS

The NISP spectroscopic channel (NISP-S) simulator (TIPS; Zoubian et al. 2014; see also Euclid Collaboration 2023a for the pilot run of the simulation campaign) is part of OU-SIM (Organisation Unit for the simulations of all instruments; Euclid Collaboration 2024b), which also includes the NISP photometric (NISP-P) simulator. The NISP-S channel simulator considers additional features with respect to the photometric channel. The detector noise, point spread function, and spectral dispersion were measured during ground tests (Waczynski et al. 2016; Barbier et al. 2018; Costille et al. 2019; Maciaszek et al. 2022, Kubik et al. in prep. for the detector on-ground characterisation) and were parameterised into TIPS. The detector noise, which is dominated by the dark current, is quite homogeneous for each of the 16 NISP detectors with a spatial mean level on the order of 0.011 ADU per frame per pixel. The maps of detector noise at the pixel-level were incorporated into TIPS. The spectral dispersion characterised at nine positions of the focal plane during the NISP ground test (Gillard et al. in prep for the spectrometer on-ground calibration) has been averaged at 13.4 $\text{Å}\,\text{px}^{-1}$ in the simulations. The astrophysical background simulated by TIPS includes the zodiacal and out-of-field stray light (see Euclid Collaboration 2022b, for details on background estimation for the *Euclid* Wide Survey) and are considered constant across the field of view.

To perform the simulation, we chose the coordinates $(\text{RA}, \text{Dec}) = (228\,394, 6.590)$ degrees where the background is dominated by zodiacal light. At these coordinates, the zodiacal light is estimated to be $2.2\,\text{photon}\,\text{s}^{-1}\,\text{px}^{-1}$. The zodiacal light in the *Euclid* Wide Survey is expected to vary between 1.1 and $3.0\,\text{photon}\,\text{s}^{-1}\,\text{px}^{-1}$ with a median at $1.6\,\text{photon}\,\text{s}^{-1}\,\text{px}^{-1}$. The out-of-field stray light is estimated at $0.3\,\text{photon}\,\text{s}^{-1}\,\text{px}^{-1}$ in our coordinates and the median estimated value for the *Euclid* Wide Survey is $0.4\,\text{photon}\,\text{s}^{-1}\,\text{px}^{-1}$.

In terms of data acquisition, we replicated the ROS for the *Euclid* Wide spectroscopic survey (Euclid Collaboration 2022b). This strategy involved a total integration time of 2212 seconds divided into four exposures. Each exposure was dithered, with an approximate offset of three pixels between them. Additionally, we simulated the four RGS grism orientations specified in the *Euclid* ROS, namely $+0°$, $-4°$, $+180°$, and $+184°$, in order to minimise the effects of cosmic ray impacts and

faulty pixels. This approach will also serve the purpose of decontamination.

4.2. Extraction of the 1D spectra

The extraction of the 1D spectra from the simulated images was performed using the *Euclid* Spectroscopic Image Reduction pipeline (SIR) which has been combined with the simulation pipeline by SIR_SPECTROSIM_RUNNER (Paganin 2022). SIR is the official *Euclid* pipeline for image reduction, wavelength and flux calibration, and extraction of the 1D spectra. The four dithered frames are processed separately by the SIR pipeline, extracting the first-order spectra using an aperture in pixels which is set at five pixels for point-like sources, as is the case for the AGN simulated in this work. The four resulting 1D spectra are then combined using inverse-variance weighting. Examples of simulated *Euclid*-like spectra for both type 1 and type 2 AGN are shown in Appendix A.

5. Redshift measurements

To estimate the efficiency and purity of the redshift measurements for the simulated AGN spectra described in Sect. 4, we performed a cross-correlation between simulated and optimised templates created from the input AGN spectra. This choice allows us to better understand the effects on the redshift measurements due to the limited spectral coverage provided by the red grism, as well as by its spectral resolution, which introduces a severe blending between adjacent spectral lines. The set of templates is composed by 9 type 1 composites (see Fig. 1) along with 23 type 2 semi-empirical spectra (see Fig. 7) that cover all the different combinations of AGN and host galaxy presented in Sect. 3.

To automatically measure the redshift, we use the IRAF task xcorr, which performs a cross-correlation between the Fourier-transformed simulated spectrum and the set of template spectra (Tonry & Davis 1979). The routine was adapted to uniformly cover the full redshift range $0 \le z \le 6$ when searching for the correlation peak. We considered the R_{xcorr} parameter, which is related to the shape of the correlation peak provided by xcorr, to assess the goodness of the correlation between the input and template spectra. This quantity is expected to be inversely proportional to the redshift measurement uncertainties (see Eq. (23) in Tonry & Davis 1979). We note that xcoor strongly depends upon the choice of the templates utilised to fit the data but, in our experiment, we have considered the same AGN templates used to build the simulations. The template library we employed is thus representative of the data we fit. The redshift distribution determined from xcoor is thus obtained in the most favourable condition. Moreover, since our main aim does not require a complete distribution of AGN, the redshift determination we present here only provides a test bed for the future OU-SPE pipeline, without having the ambitious goal of being complete. To quantify the quality of the redshift measurement, we considered the following definitions.

- Measurements are termed good if $|z_{\text{out}} - z_{\text{in}}| < 0.01$, where z_{in} and z_{out} are the redshifts of the input and simulated spectrum and $R_{\text{xcorr}} \ge 5$, respectively.
- Measurements are termed false if $|z_{\text{out}} - z_{\text{in}}| > 0.01$ even if the parameter $R_{\text{xcorr}} \ge 5$. The fraction of redshift estimates flagged as false gives us information about the purity of the AGN spectroscopic mock sample.
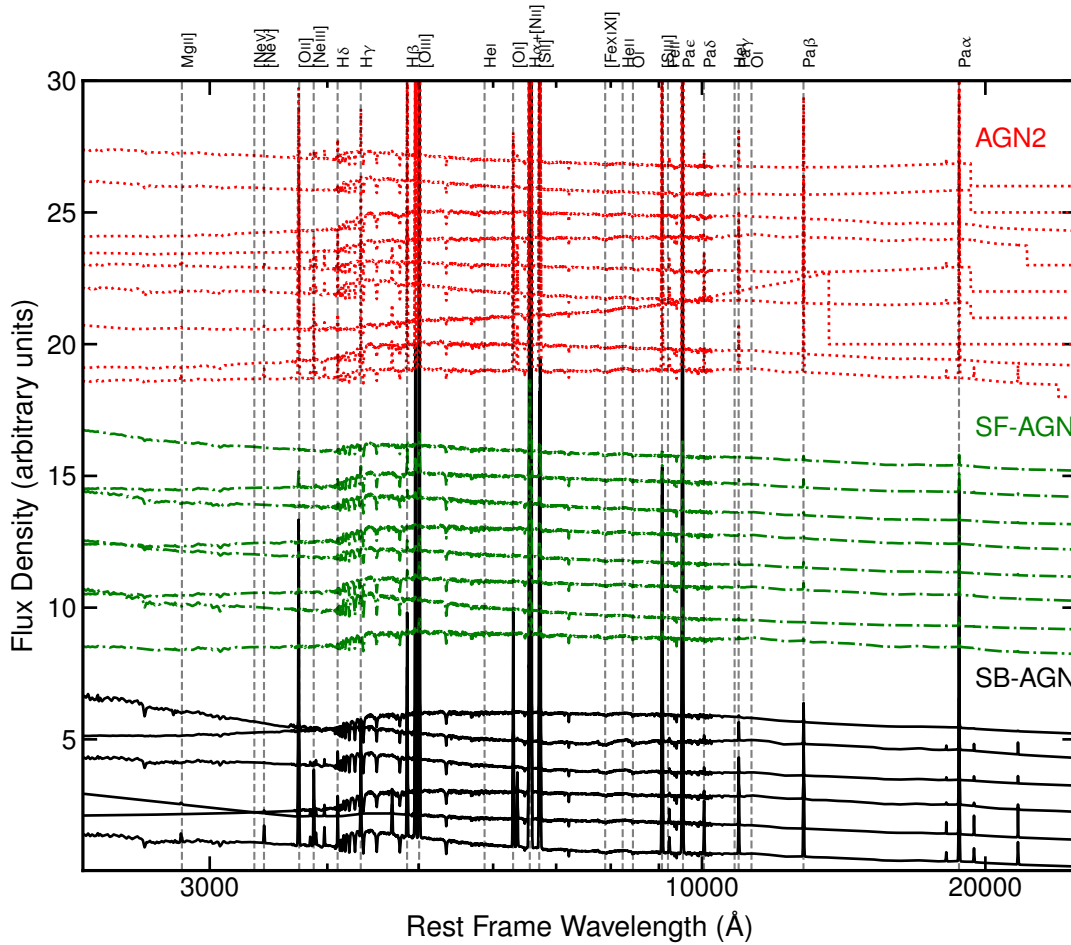
Fig. 7. Type 2 incident templates employed for the cross-correlation analysis in the redshift determination (i.e. `xcoor`, see Sect. 5). Black solid line: Six templates representing SB-AGN. Green dot-dashed line: Eight SF-AGN templates. Red dotted line: Nine AGN2 templates.

– Measurements are termed lost if $R_{xcorr} < 5$. This regime also includes AGN with a low signal-to-noise ratio (S/N) spectrum and (or) absence of strong emission lines. This is especially true at low redshift and high redshift where Hα and Hβ+[O III] are outside the wavelength coverage of the grism. As a first test, we adopted no prior assumption on either the redshift or the AGN classification of the simulated source. By running the cross-correlation completely blind, we measure the correct redshifts for almost 60% of the AGN spectra (1487/2496), but with a relatively high contamination (7%), i.e., 176 objects in which the redshift is measured with good confidence but with an incorrect redshift estimate (flagged as 'false', see Fig. 8, top panel). We do not find a significant improvement in the redshift measurements when adopting a type 1 or 2 template depending on the classification of the analysed AGN spectrum.

We then included a simulated set of photometric redshifts, with an uncertainty assumed to be $\delta z/(1 + z) = 0.2$, as a prior in the cross-correlation, by considering the photometric redshift value as the centre of a redshift window with a width of 0.5. The search for the correlation peak is carried out only within this redshift range. With this prior of the photometric redshift, we significantly improved the efficiency in measuring correct redshifts[6]. The objects with a redshift measurement flagged *good*

($|z_{out} - z_{in}| < 0.01$ and $R_{xcorr} \geq 5$) constitute 65% of the total (1622/2496), with an increase of the efficiency that is due to the significant decrease of false measurements (49/2496): only 2% of objects have a confident but incorrect redshift estimate (i.e., $R_{xcorr} \geq 5$).

In terms of redshift intervals, we have identified three ranges. A low-redshift ($z < 0.89$) interval where the Hα emission line is not yet redshifted into the NISP spectral range. A second range where the Hα emission line is within the NISP spectral coverage of the red grism ($0.89 < z < 1.83$) and a third range at high redshift ($z > 1.83$) where the Hα is, again, redshifted outside the observed NISP spectral range.

Our results are very encouraging in the redshift interval where Hα is visible. In this range, considering the results obtained including the photo-z prior, our simulation comprises 844 AGN (307 type 1 and 537 type 2 AGN), and 689 of them (82%) have good and highly ranked redshifts, with further 144 spectra (17%) with redshifts determined from `xcorr` but they are not reliable based on our adopted criterion ($R_{xcorr} < 5$). The purity in this redshift range is very high (98%), with only 11 objects classified as false. We caution that the fraction of AGN with good redshifts will be lower in the real data, since we are missing faint AGN in the simulated sample and the success rate is a function of flux (as it will be shown in Fig. 13).

Outside this redshift range, the measurement efficiency decreases significantly. This is expected at high redshift, where we have fainter objects and a wider redshift interval, and the

---

[6] We caution that the effect of catastrophic errors is not included in our analysis.
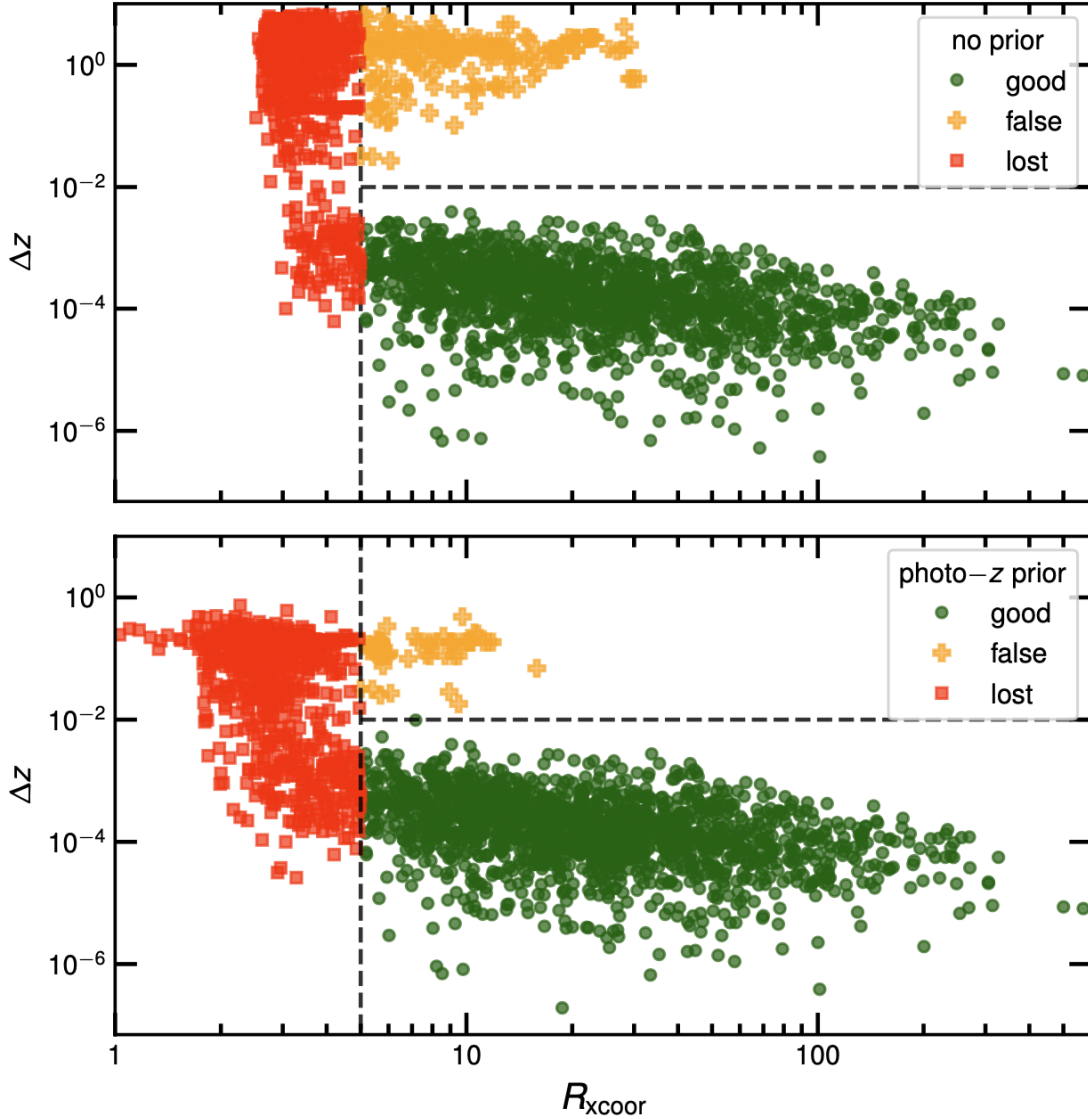
Fig. 8. Redshift uncertainty Δz (defined as $|z_{\text{out}} - z_{\text{in}}|$) for the total sample of 2496 AGN as a function of the parameter R_{xcorr} (which is related to the shape of the correlation peak, and is inversely proportional to the redshift measurement uncertainty) when using no prior (upper panel) and a photometric redshift prior (lower panel). Different colours refer to different classifications of the redshift estimates (see Sect. 5 for more details).

emission lines are weak if not absent. At low redshift, the low efficiency is related to the lack of intense emission lines.

We also analysed the redshift measurements obtained from the latest version of the OU-SPE (organisation Unit for the spectroscopic redshifts) pipeline. The OU-SPE pipeline estimates the correct redshift for 30% considering the full sample. This fraction becomes 48% when we limit the AGN sample to the redshift interval where Hα is visible. The principal cause for this inconsistency lies in the set of templates implemented in the pipeline, which are tailored for non-active galaxies (e.g. broad lines are not considered). As the pipeline is currently under development (Euclid Collaboration: Le Brun et al., in prep.), we prefer to restrict the analysis to the redshift values as discussed above.

6. Spectral analysis

As the calibrated *Euclid* spectra will be available to the scientific community online, we detail here a 'user-like' spectral analysis of the simulated data mainly focussing on the emission line fluxes. Emission line fluxes for both incident and simulated *Euclid*-like mock AGN spectra are measured using a modified version of the Quasar Spectral Fitting library (QSFIT, Calderone et al. 2017). The updated QSFIT is a Julia language[7] based automatic spectral fitting tool (Selwood et al. 2023) for both type 1 and type 2 AGN spectra from the optical to the ultraviolet, using the redshift values estimated in Sect. 5. Properties of the AGN continuum, Balmer continuum and pseudo-continuum, iron complexes, host galaxy contributions, and emission lines are derived in the rest-frame wavelength range 1215–7300 Å (see also e.g. Shen et al. 2011; Rakshit et al. 2020). Scattered light and nebular continuum are not included in the spectral fit of type 2 AGN. These two contributions are nonetheless negligible with respect to the continuum and star formation (e.g. Vidal-García et al. 2024).

QSFIT utilises a flexible fitting recipe which includes several physically motivated AGN spectral components to model the data. Whilst we refer the reader to the relative publication for details, we summarise below the main components considered to fit the data.

---

[7] https://julialang.org/

Table 5. Optical and UV emission lines considered in QSFIT analysis of type 1 and type 2 incident spectra and their wavelengths (rest-frame, vacuum).

Line	Wavelength (Å)	Type 1	Type 2
C III]	1908.53	B+N	N
C II]	2327.64	B	–
Mg II	2796.352	B+N	N
[Ne V]	3426.5	N	N
[O II]	3727	N	N
[Ne III]	3870.16	N	N
Hδ	4102.892	B	–
Hγ	4341.684	B	N
[O III]	4364.436	N	N
He II	4687.02	B	–
Hβ	4862.683	B+N	N
[O III]	4960.295	N	N
[O III]	5008.240	N	N
[O III] (blue wing)	5008.240	N	N
He I	5877.25	B	–
[O I]	6302.046	N	N
[O I]	6365.536	N	N
[N II]	6549.85	N	N
Hα	6564.61	B+N+VB	N
[N II]	6585.28	N	N
[S II]	6718.29	N	N
[S II]	6732.67	N	N

Notes. B refers to broad lines, N to narrow lines, and VB to very broad lines.

The first major component is the AGN continuum. In the general case for both type 1 and type 2 AGN, the AGN nuclear continuum is modelled as a single power-law model of the form

$$L_\lambda = A \left(\frac{\lambda}{\lambda_0} \right)^{\alpha_\lambda}, \tag{1}$$

across the available (rest-frame) wavelength range, where λ_0 is a reference wavelength, A is the luminosity density at $\lambda = \lambda_0$ and α_λ is the spectral slope at $\lambda \ll \lambda_0$. The parameter λ_0 is fixed at the median wavelength in the available spectral range, A is constrained to be positive and α_λ is constrained to be in the range $[-5, 5]$.

The next major component is the Balmer continuum and pseudo-continuum. The Balmer continuum for type 1 AGN is modelled in QSFIT according to Grandi (1982) and Dietrich et al. (2002):

$$L_\lambda = A \, B_\lambda(T_e) \left\{ 1 - \exp\left[\tau_{BE} \left(\frac{\lambda}{\lambda_{BE}} \right)^3 \right] \right\}, \tag{2}$$

where A is the luminosity density at 3000 Å, $B_\lambda(T_e)$ is the blackbody curve at the electron temperature T_e (fixed at $T_e = 15\,000$ K), τ_{BE} is the optical depth at the Balmer edge (fixed at $\tau_{BE} = 1$), and λ_{BE} is the Balmer edge wavelength, fixed at 3645 Å. Higher order Balmer components ($7 \leq n \leq 50$, i.e., the Balmer pseudo-continuum), are accounted for in the QSFIT model using the line ratios from Storey & Hummer (1995) with a fixed electron density of 10^9 cm^{-3} and a temperature of 15 000 K. A free parameter in the fit determines the ratio of the higher-order blended line complex to the Balmer continuum at the Balmer edge. The whole component is finally broadened by a

Table 6. Default known emission line type properties available to be implemented with QSFIT.

Line type	FWHM (km s^{-1})
Narrow	100–2000
Broad	900–15 000
VeryBroad	10 000–30 000

convolution with a Gaussian profile with a FWHM equal to 5000 km s^{-1}.

Another significant component is the host galaxy template, especially for low-luminosity AGN. QSFIT models the contribution of the host galaxy emission to the AGN spectrum using a host galaxy template. A library of host galaxy SEDs is incorporated in QSFIT, consisting of the templates from SWIRE (*Spitzer* wide-area infrared extragalactic survey) collated by Polletta et al. (2007) supplemented with twelve starburst galaxy templates generated to model cosmological evolution survey (COSMOS) AGN in Ilbert et al. (2009). This provides a selection of SEDs ranging from quiescent elliptical galaxies through to starbursts. By default, the template of an elliptical galaxy with an age of 5 Gyr drawn from the SWIRE library is used when the host galaxy is enabled in a fit.

We then considered the UV and optical iron templates. For type 1 AGN, broad and narrow (permitted and forbidden) iron lines are modelled using the UV and the optical iron complex templates of Vestergaard & Wilkes (2001) and Véron-Cetty et al. (2004), respectively.

Known emission lines represent a subset of spectral transitions that are considered in a given QSFIT fitting recipe. Known emission lines can be defined as narrow, broad or very broad lines (referred to in Table 5 as N, B and VB respectively), which carry different limits on their allowed parameter ranges as summarised in Table 6. Known emission lines can be modelled with a Gaussian, a Lorentzian, or a Voigt profile.

The integrated line luminosity ratios are fixed to a ratio 1:3 only for the oxygen doublet, [O III]λ4959 and [O III]λ5007, and (separately) for the nitrogen doublet, [N II]λ6549 and [N II]λ6583. Also, velocity offsets with respect to the redshift are fixed for the above mentioned lines. The only constraint for the narrow line widths is to be $<10^3$ km s^{-1}. No other constraint involving [O III]$\lambda\lambda$4959, 5007 and [N II]$\lambda\lambda$6549, 6583 is applied since they would be simultaneously available in the same spectrum only in a limited redshift range.

Finally, Gaussian-profiled 'unknown' lines are placed at areas of highest residuals in the latter stages of a spectral fit, with the aim of modelling additional emission lines features that are present in the spectrum but have not been listed in the current spectral lines library. Unknown lines are able to be placed at any wavelength contained in the spectrum, but are excluded from being placed within the Hα+[N II] and Hβ+[O III] complexes. The number of considered unknown lines is a parameter that should be adjusted based on the redshift and the spectral coverage of the spectra being analysed, to strike a balance between under- and over-utilisation. We constrain unknown line FWHM to the range 500–10 000 km s^{-1} for type 1 AGN and 10–2000 km s^{-1} for type 2 AGN.

6.1. Spectral analysis of the incident spectra

We limit our analysis of incident spectra to the NISP-S red grism observed-frame wavelength range of 12 500–18 500 Å. Lines

Table 7. Summary of model components considered in QSFIT analysis of type 1 and type 2 incident spectra.

Model component	Type 1	Type 2
AGN continuum	Single PL	Single PL
Balmer (pseudo-)continuum	Yes	–
Host galaxy template	–	Yes
UV and optical iron templates	Yes	–
Known emission lines	Yes	Yes
Known emission line profiles	Gaussian	Voigt
Max number of unknown lines	5	4

Notes. PL means power law.

Fig. 9. Example showing a *Euclid*-like simulated type 2 AGN spectrum where the maximum number of unknown lines are used.

outside of this range will not be present in the simulated spectra which are generated for the red grism observed-frame wavelength range only. Due to this observed-frame wavelength range cut, any samples that subsequently fall outside QSFIT's rest-frame wavelength analysis range (1215–7300 Å) are not considered in our analysis[8]. Table 5 provides an overview of the known optical and UV emission lines considered in our analysis of type 1 and type 2 AGN incident spectra.

Prior to the QSFIT analysis, we resampled the spectra to have evenly spaced wavelength samples in logarithmic space, thus ensuring a constant resolution of $100 \, \mathrm{km \, s^{-1}}$. We found five unknown lines to be an appropriate maximum amount for modelling unconsidered elements in type 1 incident spectra in the NISP-S wavelength range. Since type 2 spectra exhibit less complex features with respect type 1 AGN, we find that a maximum number of four unknown lines are appropriate to model type 2 incident spectra in the NISP-S wavelength range. The maximum number of unknown lines is utilised in 54% of type 1 incident spectra fits and 96% of type 2 incident spectra fits with QSFIT. To model known emission lines in the type 2 incident spectra, we considered a Gaussian spectral profile (a test considering different spectral profiles is discussed in Appendix B), to be consistent with the simulated data (see the next section). Table 7 provides a summary of the individual recipes used to fit type 1 and type 2 incident mock spectra.

6.2. Spectral analysis of the simulated spectra

Identical to the treatment of the incident spectra, we resampled the linearly spaced simulated spectra to be logarithmically spaced in wavelength. We deconvolved the simulated spectra using a Gaussian NISP-S instrumental resolution of $667 \, \mathrm{km \, s^{-1}}$ (corresponding to a spectral resolution of $R = 450$). The maximum number of unknown lines allowed to perform the spectral fitting for type 1 and type 2 simulated spectra is reduced with respect to the incident spectra models due to the decreased wavelength range covered by the simulated spectra, which are generated for the NISP-S red grism only. A total of two unknown lines is found to be sufficient for the modelling of unconsidered elements in type 1 and type 2 simulated spectra. We recall here that the unknown lines are 'nuisance lines' useful to deal with minor spectral features in the data which, if neglected by our model, would produce an overall bad fit. An example is shown in Fig. 9

---

[8] We do not include emission lines shorter than C III] in the spectral fit, as we are only analysing the spectra in the red grism wavelength range (for which the simulated spectra are created). Since the highest redshift AGN in the catalogue are at $z \simeq 7$ and the red grism coverage starts at $12\,500$ Å observed-frame, the shortest rest-frame wavelength we can possibly observe is at 1563 Å.
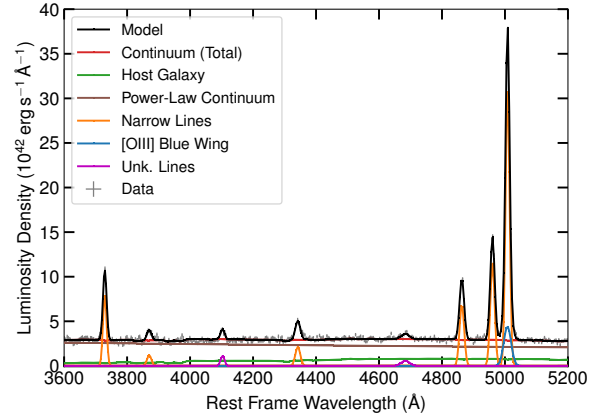
for a simulated type 2 spectrum. We remind that these lines are never used to model spectral features such as the Hα+[N II] complex, for instance. The maximum number of unknown lines is utilised in 77% of type 1 simulated spectra fits and 95% of type 2 simulated spectra fits. The fraction in type 2 AGN is the same as for the incident data, whilst it is higher than in type 1 AGN, but we must consider that there are only two unknown lines included in the spectral fit recipe and also that many of the spectra have a low S/N which sometimes results in an unknown line (or two) being used for unphysical noise features.

Due to the low spectral resolution of NISP-S, we find that the Hα complex (Hα+[N II]) cannot be reasonably deblended in simulated spectra. The blending of the lines results in our standard three-component fitting of the Hα+[N II] complex becoming highly degenerate. The individual components included for the complex therefore no longer follow their intended use and will blend together to fill the residual of the feature. Figure 10 presents an example showing the difference between the Hα complex in a type 1 incident spectrum and its *Euclid* simulation. We compared two alternative methods of modelling the blended Hα complex in type 1 simulated spectra: (1) a single broad Gaussian component and (2) both a narrow and broad Gaussian component. We drew a comparison between the total Hα+[N II] complex flux measured for the incident spectra against the flux measured using our two models for simulated spectra. A single Gaussian profile provided a median percentage difference compared to incident measurements of 26%, with a median flux measurement difference (incident flux minus simulated flux) of $-6.5 \times 10^{-18} \, \mathrm{erg \, s^{-1} \, cm^{-2}}$. By employing a narrow and broad component, we found a median percentage difference compared to incident measurements of 24%, with a median flux measurement difference of $-3.8 \times 10^{-16} \, \mathrm{erg \, s^{-1} \, cm^{-2}}$. Hence, we find a roughly even percentage difference compared to our incident spectra flux measurements using both models, but a lower median flux measurement difference with the single Gaussian model. We therefore opt to use the more simplistic single Gaussian model for type 1 simulated spectra Hα complexes as there are no significant gains in measurement performance using a more complex model.

The same blending effect of the Hα complex is observed in type 2 AGN simulated spectra (see Fig. 10). We find the blended Hα complex for type 2 AGN to be well modelled with a broad Gaussian component for which the FWHM is constrained in the range 900–3500 $\mathrm{km \, s^{-1}}$. We found that Gaussian emission line profiles most effectively capture the flux that is present in type 2
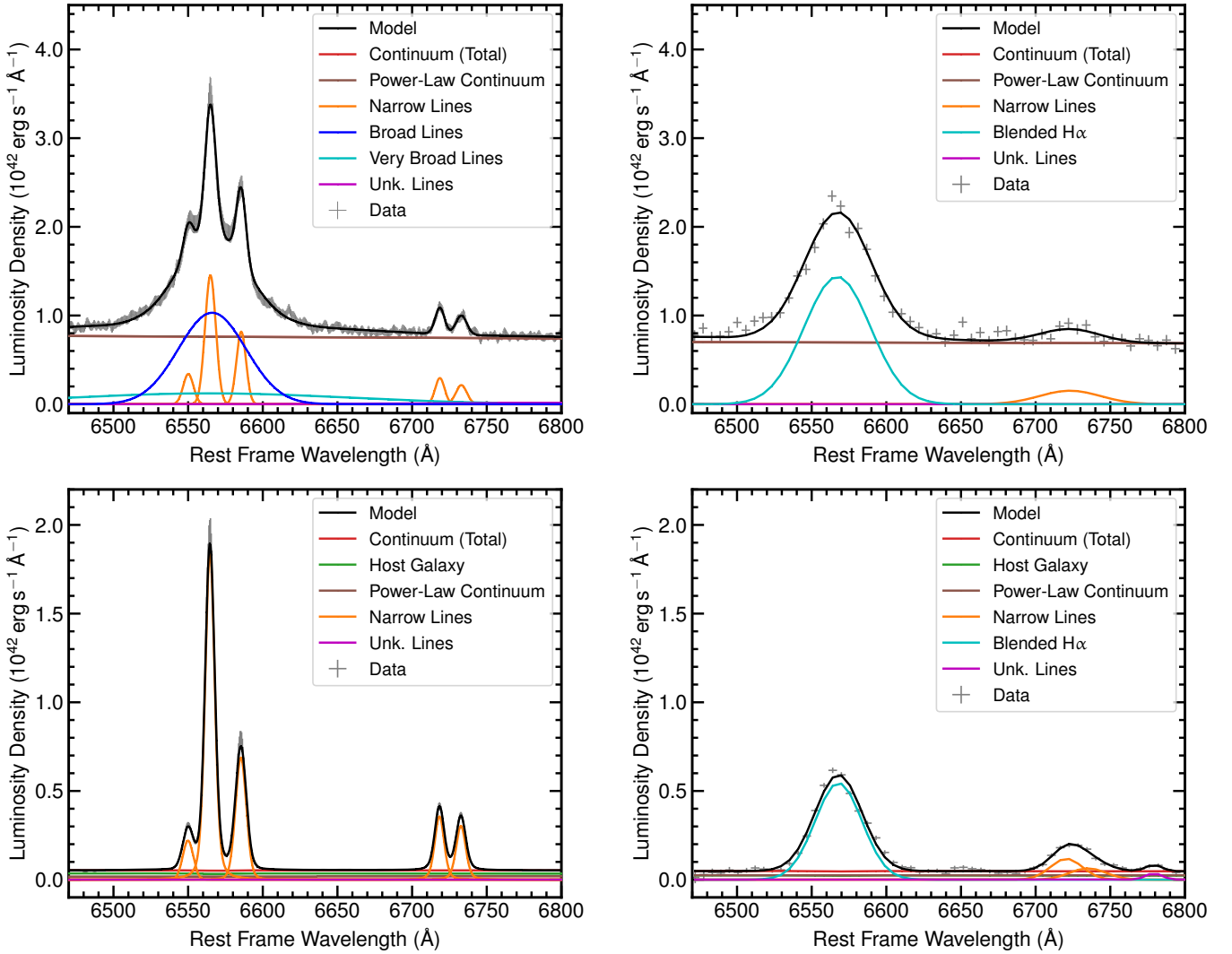
Fig. 10. Example showing type 1 (top row) and type 2 (bottom row) AGN Hα complex of incident (left) and *Euclid*-like simulated spectra (right). The sources shown in this plot have redshifts $z = 1.7$ (type 1) and $z = 1.2$ (type 2).

simulated spectra (90% of that measured in the incident spectra, see Sect. 7.1). We therefore used Gaussian emission line profiles to model known emission lines in these spectra.

Other than the changes discussed above, all the other components considered in the spectral fit of the type 1 and type 2 AGN simulated spectra models are identical to those used to analyse the type 1 and type 2 incident spectra[9]. Table 8 provides a summary of the individual recipes used to fit type 1 and type 2 simulated *Euclid*-like spectra.

7. Results and discussion

7.1. Comparison between incident and simulated data

In this section, we compare emission line fluxes measured with QSFIT in both the incident and simulated spectra. During the *Euclid* mission, all the spectral measurements will be computed by an automatic pipeline that is currently under development by the Euclid collaboration. At the time of writing, the algorithm performance is still under evaluation, and it is beyond the scope

of this work to provide an in-depth spectral analysis of the simulated data with this pipeline (details will be provided in Le Brun et al. in prep.). Nonetheless, a preliminary comparison of the emission line flux measurements for the Hα complex shows a broad agreement between the values obtained from our user-like analysis and those obtained with the SPE pipeline in its latest available version (see Appendix C).

Since *Euclid* is mainly focussed on Hα emission as it is expected to be the brighter emission line in the probed redshift range, we concentrated on the comparison of flux measurements for this emission line only. This exercise serves to qualify how close the spectral measurements obtained with *Euclid*-like spectra are to the truth and to simulate a 'user-like' analysis with a spectral fitting pipeline that is not the *Euclid* official one. Due to the discrete wavelength values used in the construction of mock spectra, the Hα complex is only observed in simulated spectra (red grism only) in the redshift range $0.9 < z < 1.8$ for type 1 sources (281 objects) and $0.855 < z < 1.8$ for the type 2 sources (582 objects). As the Hα complex is blended in simulated spectra (see Sect. 6.2), we compare the flux measured for the single broad component used to model the complex in simulated spectra to the summation of the individual components used to model the Hα complex for incident spectra (i.e., broad and

Table 8. Summary of model components considered in QSFIT analysis of type 1 and type 2 simulated spectra. PL = power law.

Model Component	Type 1	Type 2
AGN continuum	Single PL	Single PL
Balmer (pseudo-)continuum	Yes	–
Host galaxy template	–	5 Gyr elliptical
UV and optical iron templates	Yes	–
Known emission lines	Yes (modified Hα)	Yes (modified Hα)
Known emission line profiles	Gaussian	Gaussian
Max number of unknown lines	2	2

narrow component for the Hα, [N II]$\lambda\lambda6549$, 6583, see Sect. 6.1). To perform the fit and to compute the luminosity values, we considered the input (true) redshift. This is just a (zero-order) working assumption to evaluate the performance of the fitting procedure. The Hα flux was determined from the simulated spectra of the type 1(2) AGN for 263(353) sources, with 172(344) having a spectroscopic redshift flagged as 'good'.

Figure 11 (top row) presents the comparison of the Hα+[N II] flux measurement between the simulated and the incident spectra for the type 1 (left panel) and type 2 (right panel) AGN samples. We highlighted all objects with a spectroscopic redshift measured from the simulated spectra flagged as 'good' (see Sect. 5) with blue symbols. We quantified the relative flux difference between the simulated and incident data as $\Delta F/F = (F_{\mathrm{H\alpha,sim}} - F_{\mathrm{H\alpha,inc}})/F_{\mathrm{H\alpha,inc}}$, as shown in the bottom panels of each plot. The Hα+[N II] flux values obtained with the simulated and the incident data are in overall good agreement, with the narrower dispersion obtained for the subsample for which we have a good spectroscopic redshift, as expected. The median $\Delta F/F$ are about -28% and -10% for the type 1 and type 2 AGN, respectively. We inspected the best fits of the simulated spectra for both the type 1 and type 2 AGN, finding that the profiles are overall well modelled for both AGN types (see Fig. 10). Nonetheless, a fraction of the flux is lost, especially in the case of type 1 AGN as shown in Fig. 11 (top left panel).

To investigate the reason for this flux loss, we fit all the incident data again with the same prescription employed for the simulated ones: one single broad component to model the whole Hα+[N II] complex. Results are shown in the bottom row of Fig. 11 for type 1 (left) and type 2 (right) AGN. The median $\Delta F/F$ for type 1 is now reduced to about -8%, whilst the flux comparison for the type 2 significantly worsened. The latter case is explained by the fact that one single Gaussian component is clearly not sufficient to retrieve the Hα+[N II] emission in the incident data. In the case of type 1 AGN, instead, the comparison significantly improved, implying that 1 single component is sufficient to recover most of the Hα+[N II] emission in the simulated data.

We also checked whether the remaining fraction of the missing flux could be attributed to the slitless spectra extraction procedure. By comparing the median flux values of the incident and simulated spectra, we find that flux loss could be roughly 11% for type 1 AGN and 2–8% for type 2 AGN. This implies that flux loss could partly explain the remaining fraction of flux missing in the comparison between simulated and incident flux values (see Appendix F for details). We conclude that the low resolution of the NISP instrument will not allow the deblending of the Hα+[N II] complex in AGN of any type, with a single Gaussian profile able to recover $> 85\%$ of the Hα+[N II] emission in the simulated data.

### 7.2. Expected fraction of AGN with spectroscopic redshifts

As discussed in Sect. 6.2, the low resolution of the NISP instrument will not deblend the Hα+[N II] complex in AGN of any type. This implies that the number of type 1 AGN for which the redshift will be determined from the observed *Euclid* spectra is lower than for type 2 AGN, especially for AGN close to the line flux limit. The emission line S/N of the Hα+[N II] complex[10] observed in the simulated spectra for type 1 and type 2 AGN, for which the spectroscopic redshift is measured, is shown in Fig. 12 where the emission line S/N is plotted as a function of the integrated observed emission line flux. The extrapolations to the flux limit of the relations fall below the required S/N (3σ), which is more pronounced for type 1 AGN, as expected, given that the noise associated to a broad line is greater than the one associated to the narrow one (i.e. the number of pixels is higher for a broad line and thus the related noise). We also computed the ratio between the number of sources with a 'good' redshift measurement and the total number of objects as a function of the integrated Hα+[N II] flux in the simulated spectra for both type 1 and type 2 AGN (Fig. 13). A spectroscopic redshift is determined for $\sim$90% of the simulated type 2 AGN down to an emission line flux of roughly 3×10^{-16} erg s^{-1} cm^{-2}, whilst the emission line flux value is more than a factor of 2 higher for type 1 AGN at the same percentage, i.e., 8.5×10^{-16} erg s^{-1} cm^{-2}.

### 7.3. Estimating black hole mass using Hα

The measurement of the black hole mass (M_{BH}) is challenging, even when reverberation mapping campaigns are available (e.g. Bentz & Manne-Nicholas 2018). Yet, M_{BH} is one of the key parameter for studies focussed on determining the scaling relations between black hole and host galaxy properties. These scaling relations are pivotal to test black hole feedback mechanisms (e.g. Steinborn et al. 2015; Weinberger et al. 2017) and cosmological hydrodynamical simulations of structure formation that aim at investigating galaxy-black hole growth (e.g. DeGraf et al. 2015). *Euclid* will observe hundreds of thousand AGN spectroscopically, thus providing a prime dataset for such studies.

To estimate M_{BH} for both the incident and simulated datasets, we considered the recalibrated single-epoch virial relationships recently published by Kynoch et al. (2023, see their Table 4) that make use of the integrated line flux and FWHM components of the Hα emission line. We thus have to investigate first how well the FWHM of the broad Hα can be

---

[10] The emission line S/N is defined as the ratio between the total Hα+[N II] flux (obtained from the fit of a single Gaussian profile to the line) and the (statistical) uncertainty on this measurement.
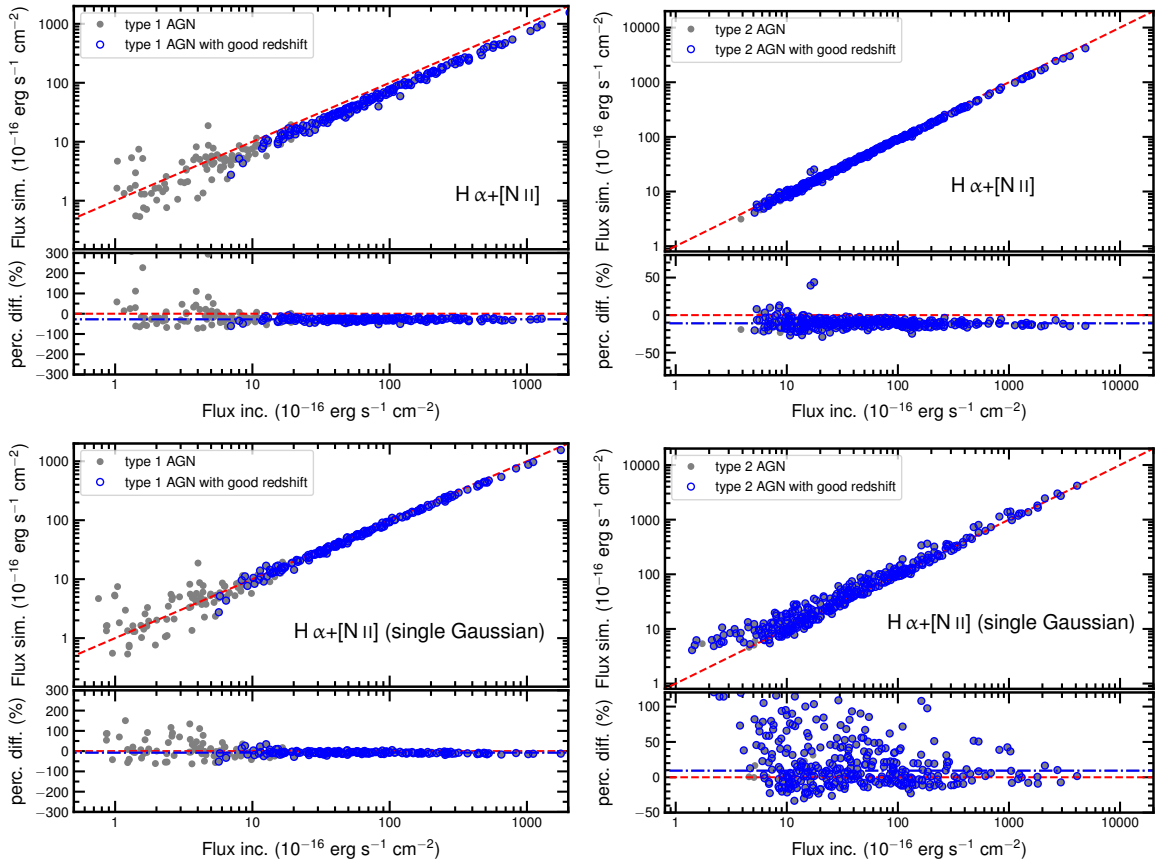
Fig. 11. Comparison of the Hα+[N II] flux measurement between the simulated and the incident spectra for the type 1 (left) and type 2 (right) AGN samples. Sources with a spectroscopic redshift measured from the simulated spectra and flagged as 'good' (see Sect. 5 for details) are marked with blue symbols. The bottom panel in each plot shows the percentage luminosity difference defined as $\Delta F/F = (F_{H\alpha,sim} - F_{H\alpha,inc})/F_{H\alpha,inc}$. The one-to-one flux relation is shown with a red dashed line, which is set to zero in the bottom panel. The dot-dashed blue line in the bottom panel represents the 50th percentile of the $\Delta F/F$ distribution. Top row: Simulated data: best fit flux obtained from the integration of a single broad component used to model the Hα+[N II] complex. Incident: Best fit flux obtained from the summation of the individual components used to model the Hα+[N II] complex for incident spectra (i.e., broad and narrow component for the Hα, [N II]$\lambda\lambda$6549, 6583). Bottom row: The best fit flux of the Hα+[N II] complex is obtained from the integration of a single broad component for both simulated and incident data.
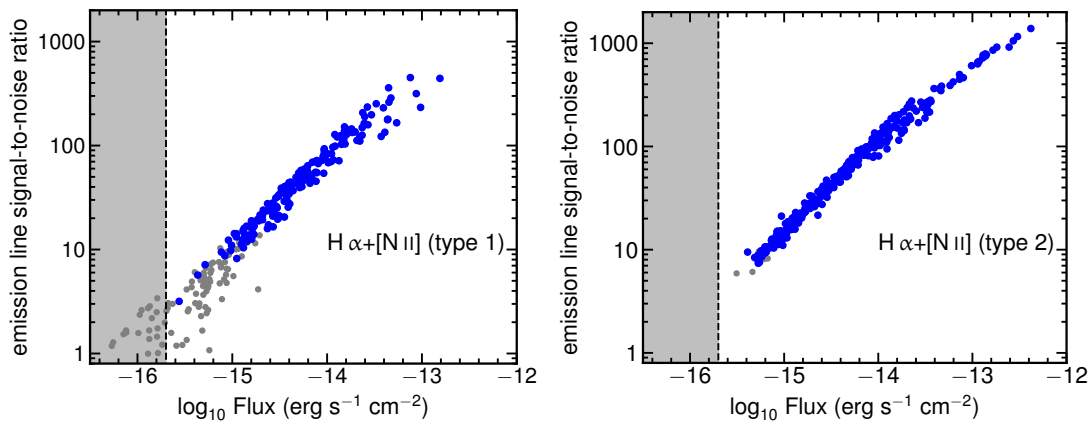
Fig. 12. Emission line S/N as a function of the total integrated flux for Hα+[N II] observed in the simulated spectra in type 1 (left) and type 2 (right) AGN. The dashed line marks the Hα flux limit of the *Euclid* survey of 2×10^{-16} erg s^{-1} cm^{-2}. Symbols are as in Fig. 11.

recovered, despite the strong blending expected for the *Euclid* NISP spectra.

Figure 14 compare the Hα FWHM between incident and simulated data when the best fit of incident spectra considers one single Gaussian component of the Hα emission line pro-

file. Despite the scatter, the values obtained for the type 1 AGN with a robust redshift measurement follow a one-to-one relation, with a median percentage error on the Hα FWHM of roughly −9%. We then plotted the comparison of the Hα flux and FWHM between incident and simulated data only for the broad
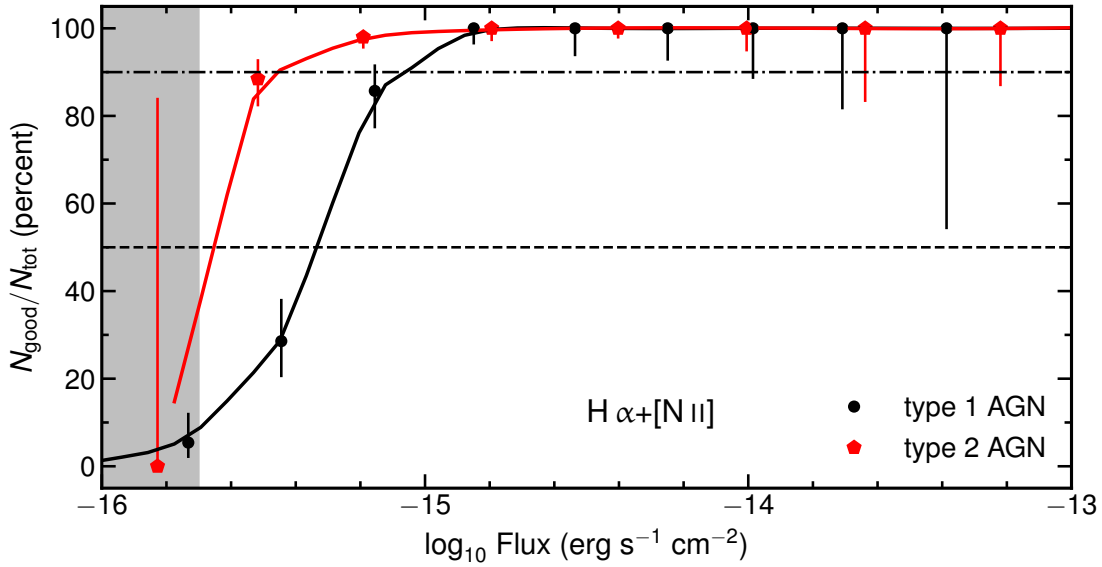
Fig. 13. Ratio between the number of sources with a 'good' redshift measurement (see Sect. 5 for details) and the total number of objects (in percent) as a function of the integrated flux of the observed Hα+[N II] emission complex in the simulated spectra for type 1 (black filled symbols) and type 2 (red pentagons) AGN. The dashed and dot-dashed lines mark the 50th and 90th percentiles, respectively.
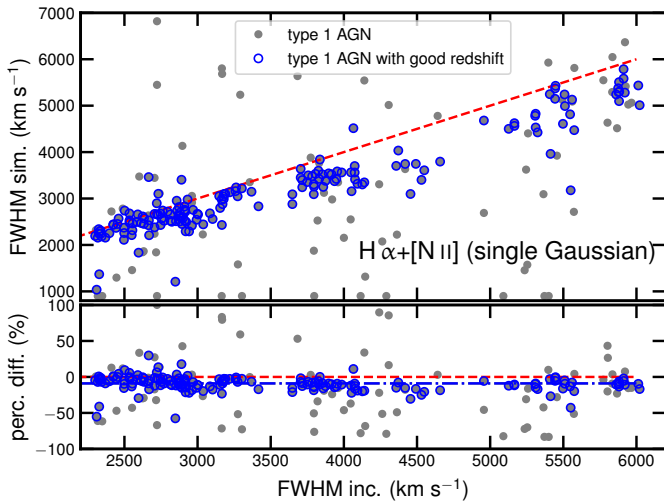
Fig. 14. Comparison of the Hα FWHM between incident and simulated data. The best fit of incident spectra considers one single Gaussian component of the Hα emission line profile (see Sect. 7.1). Symbols are as in Fig. 11.

component of the Hα emission line in Fig. 15. The median percentage error for Hα flux and FWHM are about 55% and -14%, respectively, with a rather scattered distribution of FWHM values.

Finally, Fig. 16 presents the comparison of M_{BH} estimates between the incident and the simulated data. We quantified the relative difference between M_{BH} for simulated and incident data as $\Delta M_{BH}/M_{BH} = (M_{BH,sim} - M_{BH,inc})/M_{BH,inc}$, which is shown in the bottom panel. The left panel displays the M_{BH} values obtained from the incident data where only the broad component of the Hα emission line profile is considered (see Sect. 6.1). The agreement is good overall, with a median $\Delta M_{BH}/M_{BH}$ of about -10%. The scatter in the M_{BH} distribution reduces when we consider one single Gaussian component of the Hα emission

line profile (right panel of Fig. 16) to derive M_{BH} for incident spectra. Yet, the average shift in the latter case is about -20%.

Summarising, although the estimate of both Hα flux and FWHM are individually offset with respect to the reference values, the effect of these two measures on the determination of the M_{BH} is opposite. The combination of these different shifts in Hα flux and FWHM produces the different offsets in M_{BH} discussed above.

8. Conclusions

In this work, we present estimations of the performance of the spectroscopic characterisation of both unobscured (type 1) and obscured (type 2) AGN with the NISP spectrograph, which is mounted in the ESA *Euclid* space telescope. NISP operates in the spectral range 0.9–2 μm, opening a new window for spectroscopic AGN studies in the NIR. Several hundred thousand AGN will be observed spectroscopically across a wide redshift range, which will represent the largest AGN sample with NIR spectroscopy so far. We built AGN template libraries to provide realistic spectra for type 1 and type 2 AGN that resemble those that will be detected by the NISP spectrograph. Starting from these incident spectral models, we then constructed a dedicated set of simulations to create mock AGN spectra from the rest-frame NIR to the UV, and determined the emission line flux of the Hα+[N II] complex in both the incident and simulated data. Our main findings can be summarised as follows.

– Redshift measurements are robust in the interval where Hα is visible within the spectral coverage of the red grism (i.e. $0.89 < z < 1.83$, probing the peak of AGN activity at $z \simeq 1$–1.5). In this range, our simulation comprises 889 AGN, with 83% having good redshifts (and an additional 6% have good redshift but are not reliable; i.e. $R_{xcorr} < 5$). We determined an upper limit on the purity in this redshift range to be 98% (with only 11 objects classified as 'false'). Outside this redshift range, the measurement efficiency decreases significantly. At high redshift, where we have fainter objects and a wider redshift interval, the emission lines are weak or absent. At low redshift ($z < 0.83$), the high inefficiency in the redshift determination is related to the lack of intense emission lines.
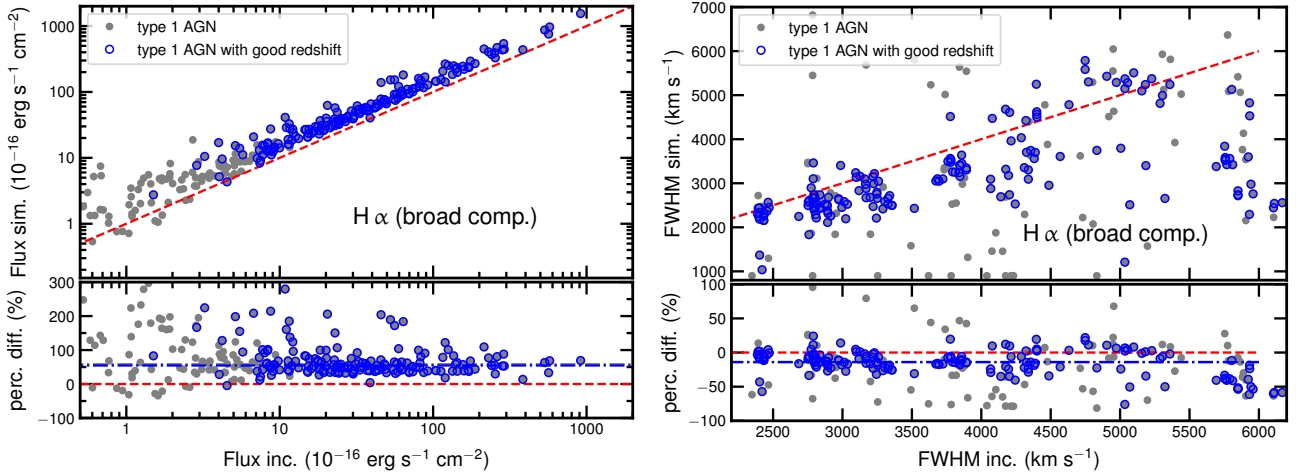
Fig. 15. Comparison of the Hα flux (left panel) and FWHM (right panel) between incident and simulated data for type 1 AGN. The flux and FWHM values derived for incident spectra considers only the broad component of the Hα emission line profile (see Sect. 6.1). Symbols are as in Fig. 11.
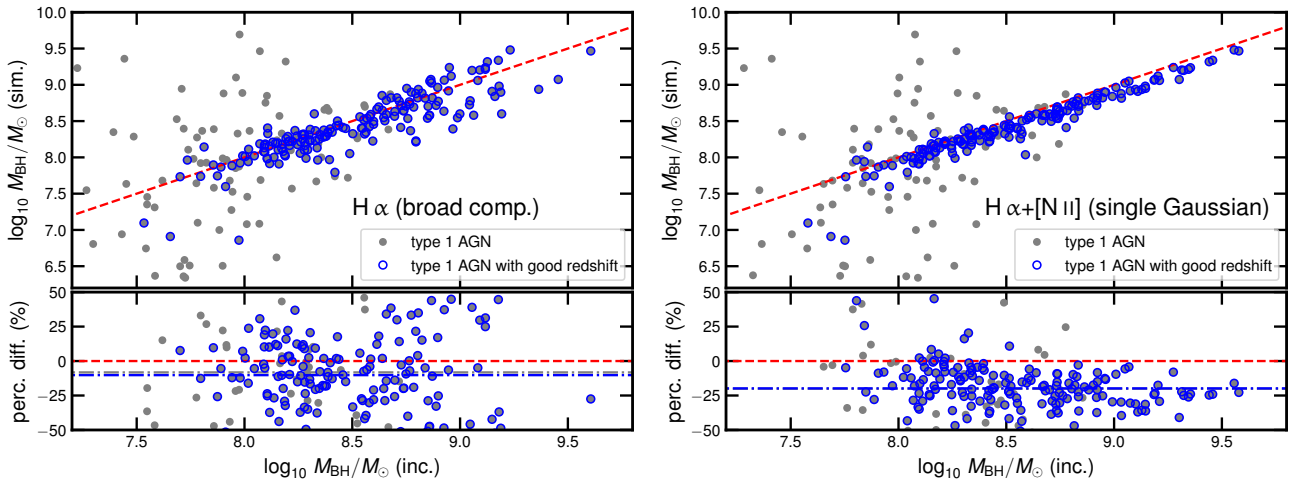
Fig. 16. Comparison of M_{BH} estimates between incident and simulated data derived from the Hα (see Sect. 7.3 for details). Left panel: M_{BH} derived from the best fit of the incident data that considers only the broad component of the Hα emission line profile (see Sect. 6.1). Right panel: M_{BH} derived from the best fit of the incident data that considers one single Gaussian component of the Hα emission line profile (see Sect. 7.1). Symbols are as in Fig. 11.

- The low spectral resolution of the NISP instrument does not allow the deblending of the Hα+[N II] complex in AGN of any type, with a single Gaussian profile able to recover > 85% of the Hα+[N II] emission in the simulated data (see Sect. 6.2 for details). A small fraction of the observed flux could also be lost at the slitless spectra extraction level, which is expected to be mitigated by the optimal extraction algorithm that is implemented in the upcoming version of the OU-SIR pipeline.

- We determined a spectroscopic redshift for about 90% of the simulated type 2 AGN down to an emission line flux of roughly 3×10^{-16} erg s^{-1} cm^{-2}. The emission line flux value is more than a factor of 2 higher for type 1 AGN at the same percentage; that is 8.5×10^{-16} erg s^{-1} cm^{-2}.

- We investigated how well the luminosity and FWHM of the broad component of Hα can be recovered, despite the strong blending expected for the *Euclid* NISP spectra. The median percentage errors for Hα luminosity and FWHM are about 55% and −14%, respectively, with a rather scattered distribution of FWHM values. Regarding our estimate of M_{BH}, we

find a good agreement overall, with a median $\Delta M_{BH}/M_{BH}$ of about −10%.

We caution that the simulation used in the present analysis assumed (1) no contamination from neighbouring spectra, (2) a single value for both sky background and out-of-field stray light, and (3) that luminosity values are extracted at the true redshift. Regarding the latter, our main results are presented only for the fraction of objects with a good redshift measurement, and therefore even if the purity is different in the real *Euclid* dataset, this approximation is suitable for a sensitivity analysis regarding redshift, Hα luminosity measurement, and black hole mass estimation. Extracted NISP-S spectra should have low contamination from neighbouring objects due to the fact that four orientations of spectra are combined. Therefore, the simulations we employed, although simplified at the level of creation, should reflect reasonable NISP-S spectra (see also Sect. 6 in Gabarra & Consortium 2023). Our simplification regarding stray light is justified if we can assume that stray light from very bright stars outside the current field of view does not have a strong gradient along a NISP-S spectrum, and so should simply represent

a raised background. Moreover, stray light from nearby stars should also be somewhat different for the four spectrum orientations and should therefore be washed out to some degree when combined. We also mention that our results for the type 1 AGN are biassed to the bright blue AGN population, and given that we excluded BALs and radio-bright AGN from the main AGN sample, our analysis could be extended to include these interesting AGN types, especially the most extreme classes in these two categories (e.g. FeLoBALs, LoBALs). We could also perhaps fold in the AGN luminosity function to quantify the capability of NISP in retrieving the AGN classification from the NISP data.

The presented analysis further extends the campaign to assess the performance of the NISP spectrograph in the context of active galaxies. Finally, the expected high spectroscopic coverage of AGN at $z < 2$ will be of prime importance for studies of AGN demography, scaling relation, and AGN clustering from the epoch of the peak of AGN activity down to the present-day Universe.

References

Akiyama, M., He, W., Ikeda, H., et al. 2018, PASJ, 70, S34
Alexandroff, R., Strauss, M. A., Greene, J. E., et al. 2013, MNRAS, 435, 3306
Barbier, R., Buton, C., Clemens, J. C., et al. 2018, SPIE Conf. Ser., 10709, 107090S
Bentz, M. C., & Manne-Nicholas, E. 2018, ApJ, 864, 146
Berta, S., Lutz, D., Santini, P., et al. 2013, A&A, 551, A100
Bezanson, R., van der Wel, A., Straatman, C., et al. 2018, ApJ, 868, L36
Bisigello, L., Gruppioni, C., Feltre, A., et al. 2021, A&A, 651, A52
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calderone, G., Nicastro, L., Ghisellini, G., et al. 2017, MNRAS, 472, 4051
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Choi, H., Leighly, K. M., Terndrup, D. M., et al. 2022, ApJ, 937, 74
Costille, A., Caillat, A., Rossin, C., et al. 2018, SPIE Conf. Ser., 10698, 106982B
Costille, A., Carle, M., Prieto, E., et al. 2019, SPIE, 11180, 111807L
Croom, S. M., Richards, G. T., Shanks, T., et al. 2009, MNRAS, 399, 1755
Cropper, M., Pottinger, S., Niemi, S., et al. 2016, SPIE Conf. Ser., 9904, 99040Q
Croton, D. J. 2006, MNRAS, 369, 1808
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
DeGraf, C., Di Matteo, T., Treu, T., et al. 2015, MNRAS, 454, 913
De Laurentis, M., & Salucci, P. 2022, ApJ, 929, 17
Dietrich, M., Appenzeller, I., Vestergaard, M., & Wagner, S. J. 2002, ApJ, 564, 581
Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142
Dopita, M. A., Kewley, L. J., Sutherland, R. S., & Nicholls, D. C. 2016, Ap&SS, 361, 61
Duras, F., Bongiorno, A., Ricci, F., et al. 2020, A&A, 636, A73
Euclid Collaboration (Barnett, R., et al.) 2019, A&A, 631, A85
Euclid Collaboration (Schirmer, M., et al.) 2022a, A&A, 662, A92
Euclid Collaboration (Scaramella, R., et al.) 2022b, A&A, 662, A112
Euclid Collaboration (Gabarra, L., et al.) 2023a, A&A, 676, A34
Euclid Collaboration (Paterson, K., et al.) 2023b, A&A, 674, A172
Euclid Collaboration (Bisigello, L., et al.) 2024a, A&A, submitted
Euclid Collaboration (Serrano, S., et al.) 2024b, A&A, submitted [arXiv:2401.01452]
Feltre, A., Hatziminaoglou, E., Fritz, J., & Franceschini, A. 2012, MNRAS, 426, 120
Feltre, A., Charlot, S., & Gutkin, J. 2016, MNRAS, 456, 3354
Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mex. Astron. Astrofis., 49, 137
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, MNRAS, 366, 767
Gabarra, L., & Consortium, E. 2023, in 41st International Conference on High Energy Physics, 1220
Gabor, J. M., Davé, R., Finlator, K., & Oppenheimer, B. D. 2010, MNRAS, 407, 749
Gibson, R. R., Jiang, L., Brandt, W. N., et al. 2009, ApJ, 692, 758
Glikman, E., Helfand, D. J., & White, R. L. 2006, ApJ, 640, 579
Granato, G. L., De Zotti, G., Silva, L., Bressan, A., & Danese, L. 2004, ApJ, 600, 580
Grandi, S. A. 1982, ApJ, 255, 25
Gruppioni, C., Pozzi, F., Andreani, P., et al. 2010, A&A, 518, L27
Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, MNRAS, 432, 23
Hall, P. B., Gunn, J. E., Knapp, G. R., et al. 2002, in Mass Outflow in Active Galactic Nuclei: New Perspectives, eds. D. M. Crenshaw, S. B. Kraemer, & I. M. George, ASP Conf. Ser., 255, 161
Hao, L., Strauss, M. A., Tremonti, C. A., et al. 2005, AJ, 129, 1783
Hopkins, P. F., Strauss, M. A., Hall, P. B., et al. 2004, AJ, 128, 1112
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Ji, Z., Giavalisco, M., Kirkpatrick, A., et al. 2022, ApJ, 925, 74
Jones, T., Martin, C., & Cooper, M. C. 2015, ApJ, 813, 126
Kashino, D., Silverman, J. D., Sanders, D., et al. 2019, ApJS, 241, 10
Kennicutt, R. C., Jr 1998a, ARA&A, 36, 189
Kennicutt, R. C., Jr 1998b, ApJ, 498, 541
Kewley, L. J., Dopita, M. A., Leitherer, C., et al. 2013, ApJ, 774, 100
Kirkpatrick, A., Pope, A., Alexander, D. M., et al. 2012, ApJ, 759, 139
Kuźmicz, A., Sethi, S., & Jamrozy, M. 2021, ApJ, 922, 52
Kynoch, D., Mitchell, J. A. J., Ward, M. J., et al. 2023, MNRAS, 520, 2781
Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166
Landt, H., Bentz, M. C., Ward, M. J., et al. 2008, ApJS, 174, 282
Landt, H., Elvis, M., Ward, M. J., et al. 2011, MNRAS, 414, 218
Landt, H., Ward, M. J., Peterson, B. M., et al. 2013, MNRAS, 432, 113
Landt, H., Ward, M. J., Elvis, M., & Karovska, M. 2014, MNRAS, 439, 1051
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints [arXiv:1110.3193]
Lusso, E., & Risaliti, G. 2016, ApJ, 819, 154
Lusso, E., Worseck, G., Hennawi, J. F., et al. 2015, MNRAS, 449, 4204
Lusso, E., Fumagalli, M., Rafelski, M., et al. 2018, ApJ, 860, 41
Maciaszek, T., Ealet, A., Jahnke, K., et al. 2016, SPIE Conf. Ser., 9904, 99040T
Maciaszek, T., Ealet, A., Gillard, W., et al. 2022, SPIE, 12180, 121801K
McGraw, S. M., Shields, J. C., Hamann, F. W., et al. 2015, MNRAS, 453, 1379
McGreer, I. D., Jiang, L., Fan, X., et al. 2013, ApJ, 768, 105
Mignoli, M., Vignali, C., Gilli, R., et al. 2013, A&A, 556, A29
Mingo, B., Watson, M. G., Rosen, S. R., et al. 2016, MNRAS, 462, 2631
Mingozzi, M., Belfiore, F., Cresci, G., et al. 2020, A&A, 636, A42
Moller, P., & Jakobsen, P. 1990, A&A, 228, 299
Mullaney, J. R., Pannella, M., Daddi, E., et al. 2012, MNRAS, 419, 95
Nenkova, M., Sirocky, M. M., Ivezić, Ž., & Elitzur, M. 2008a, ApJ, 685, 147
Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008b, ApJ, 685, 160
Paganin, L. 2022, Ph.D. Thesis, Genoa University, https://inspirehep.net/files/a2b879a4912972803e3b019cc51394d9, Italy
Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59
Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81
Prevot, M. L., Lequeux, J., Maurice, E., Prevot, L., & Rocca-Volmerange, B. 1984, A&A, 132, 389
Prochaska, J. X., Worseck, G., & O'Meara, J. M. 2009, ApJ, 705, L113
Prochaska, J. X., Madau, P., O'Meara, J. M., & Fumagalli, M. 2014, MNRAS, 438, 476
Proxauf, B., Öttl, S., & Kimeswenger, S. 2014, A&A, 561, A10
Rakshit, S., Stalin, C. S., & Kotilainen, J. 2020, ApJS, 249, 17
Reyes, R., Zakamska, N. L., Strauss, M. A., et al. 2008, AJ, 136, 2373
Ricci, F., Treister, E., Bauer, F. E., et al. 2022, ApJS, 261, 8
Ross, N. P., McGreer, I. D., White, M., et al. 2013, ApJ, 773, 14
Ryan, R. E., Casertano, S., & Pirzkal, N. 2018, PASP, 130, 34501
Saccheo, I., Bongiorno, A., Piconcelli, E., et al. 2023, A&A, 671, A34
Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, ApJ, 690, 1250

Schindler, J.-T., Fan, X., McGreer, I. D., et al. 2019, ApJ, 871, 258
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Selwood, M., Calderone, G., Fotopoulou, S., & Bremer, M. N. 2023, MNRAS, 518, 130
Shankar, F., Lapi, A., Salucci, P., De Zotti, G., & Danese, L. 2006, ApJ, 643, 14
Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45
Simcoe, R. A., Burgasser, A. J., Bernstein, R. A., et al. 2008, SPIE Conf. Ser., 7014, 70140U
Steinborn, L. K., Dolag, K., Hirschmann, M., Prieto, M. A., & Remus, R.-S. 2015, MNRAS, 448, 1504
Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163
Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30
Storey, P. J., & Hummer, D. G. 1995, MNRAS, 272, 41
Tonry, J., & Davis, M. 1979, AJ, 84, 1511
Véron-Cetty, M. P., Joly, M., & Véron, P. 2004, A&A, 417, 515
Vestergaard, M., & Wilkes, B. J. 2001, ApJS, 134, 1
Vidal-García, A., Charlot, S., Bruzual, G., & Hubeny, I. 2017, MNRAS, 470, 3532
Vidal-García, A., Plat, A., Curtis-Lake, E., et al. 2024, MNRAS, 527, 7217
Villforth, C., Herbst, H., Hamann, F., et al. 2019, MNRAS, 483, 2441
Waczynski, A., Barbier, R., Cagiano, S., et al. 2016, SPIE Conf. Ser., 9915
Ward, S. R., Harrison, C. M., Costa, T., & Mainieri, V. 2022, MNRAS, 514, 2936
Weinberger, R., Springel, V., Hernquist, L., et al. 2017, MNRAS, 465, 3291
Yuan, S., Strauss, M. A., & Zakamska, N. L. 2016, MNRAS, 462, 1603
Zakamska, N. L., Strauss, M. A., Krolik, J. H., et al. 2003, AJ, 126, 2125
Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, ApJ, 475, 469
Zoubian, J., Kümmel, M., Kermiche, S., et al. 2014, in Astronomical Data Analysis Software and Systems XXIII, eds. N. Manset, & P. Forshay, ASP Conf. Ser., 485, 509

1 INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy
2 Dipartimento di Fisica e Astronomia, Università di Firenze, Via G. Sansone 1, 50019 Sesto Fiorentino, Firenze, Italy
e-mail: elisabeta.lusso@unifi.it
3 School of Physics, HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK
4 INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
5 INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
6 INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
7 INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
8 Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy
9 INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
10 Department of Mathematics and Physics, Roma Tre University, Via della Vasca Navale 84, 00146 Rome, Italy
11 INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
12 INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
13 Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
14 Dipartimento di Fisica e Astronomia "Augusto Righi" – Alma Mater Studiorum Università di Bologna, Via Piero Gobetti 93/2, 40129 Bologna, Italy
15 Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
16 Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal
17 DTx – Digital Transformation CoLAB, Building 1, Azurém Campus, University of Minho, 4800-058 Guimarães, Portugal
18 Department of Physics, Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK
19 SISSA-International School for Advanced Studies, Via Bonomea 265, 34136 Trieste, TS, Italy
20 INFN-Sezione di Trieste, Via Valerio 2, 34127 Trieste, TS, Italy
21 Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
22 Department of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK
23 INAF-Istituto di Astrofisica e Planetologia Spaziali, Via del Fosso del Cavaliere, 100, 00100 Roma, Italy
24 Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
25 School of Physical Sciences, The Open University, Milton Keynes MK7 6AA, UK
26 Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405 Orsay, France
27 ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada Madrid, Spain
28 Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
29 INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
30 Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
31 Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
32 INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
33 Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany
34 INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese, (TO), Italy
35 Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146 Genova, Italy
36 INFN-Sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy
37 Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126 Napoli, Italy
38 INFN section of Naples, Via Cinthia 6, 80126 Napoli, Italy
39 Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
40 INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
41 Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra, (Barcelona), Spain
42 Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra, (Barcelona), Spain
43 Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany
44 Dipartimento di Fisica e Astronomia "Augusto Righi" – Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
45 Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
46 Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
47 European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
48 University of Lyon, Univ. Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, 69622 Villeurbanne, France
49 Institute of Physics, Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
50 UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
51 Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande 1749-016, Lisboa, Portugal
52 Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande 1749-016, Lisboa, Portugal
53 Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
54 Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK

55 Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191 Gif-sur-Yvette, France

56 Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Via Irnerio 46, 40126 Bologna, Italy

57 University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany

58 Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029, Blindern 0315, Oslo, Norway

59 Department of Physics, Lancaster University, Lancaster LA1 4YB, UK

60 von Hoerner & Sulger GmbH, SchloßPlatz 8, 68723 Schwetzingen, Germany

61 Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark

62 Cosmic Dawn Center (DAWN), Copenhagen, Denmark

63 Department of Physics and Helsinki Institute of Physics, Gustaf Hällströmin katu 2, 00014 University of Helsinki, Finland

64 Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, 1211 Genève 4, Switzerland

65 Department of Physics, PO Box 64, 00014 University of Helsinki, Finland

66 Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland

67 NOVA optical infrared instrumentation group at ASTRON, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

68 Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany

69 Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK

70 Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France

71 Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014 Paris, France

72 Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France

73 European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands

74 Department of Physics and Astronomy, University of Aarhus, Ny Munkegade 120, 8000 Aarhus C, Denmark

75 Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

76 Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

77 Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada

78 Université Paris-Saclay, Université Paris Cité, CEA, CNRS, Astrophysique, Instrumentation et Modélisation Paris-Saclay, 91191 Gif-sur-Yvette, France

79 Space Science Data Center, Italian Space Agency, Via del Politecnico snc, 00133 Roma, Italy

80 Centre National d'Études Spatiales – Centre spatial de Toulouse, 18 avenue Édouard Belin, 31401 Toulouse Cedex 9, France

81 Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov 077125, Romania

82 Instituto de Astrofísica de Canarias, Calle Vía Láctea s/n, 38204 San Cristóbal de La Laguna, Tenerife, Spain

83 Departamento de Astrofísica, Universidad de La Laguna, 38206 La Laguna, Tenerife, Spain

84 Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile

85 Institut d'Estudis Espacials de Catalunya (IEEC), Carrer Gran Capitá 2-4, 08034 Barcelona, Spain

86 Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain

87 Satlantis, University Science Park, Sede Bld 48940, Leioa-Bilbao, Spain

88 Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain

89 Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA

90 Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda 1349-018, Lisboa, Portugal

91 Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain

92 Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Édouard Belin, 31400 Toulouse, France

93 Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700, AV Groningen, The Netherlands

94 INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy

95 IFPU-Institute for Fundamental Physics of the Universe, Via Beirut 2, 34151 Trieste, Italy

96 INAF-Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy

97 Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin, 69627 Villeurbanne Cedex, France

98 Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany

99 Université St Joseph; Faculty of Sciences, Beirut, Lebanon

100 Junia, EPA department, 41 Bd Vauban, 59800 Lille, France

101 ICSC-Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy

102 Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain

103 CERCA/ISO, Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA

104 Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, 92190 Meudon, France

105 Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy

106 Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy

107 Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France

108 Dipartimento di Fisica – Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34131 Trieste, Italy

109 NASA Ames Research Center, Moffett Field, CA 94035, USA

110 Kavli Institute for Particle Astrophysics & Cosmology (KIPAC), Stanford University, Stanford, CA 94305, USA

111 Bay Area Environmental Research Institute, Moffett Field, CA 94035, USA

112 Minnesota Institute for Astrophysics, University of Minnesota, 116 Church St SE, Minneapolis, MN 55455, USA

113 Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France

114 Institute Lorentz, Leiden University, PO Box 9506, Leiden 2300 RA, The Netherlands

115 Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

116 Department of Physics & Astronomy, University of California, Irvine, CA 92697, USA

117 Department of Astronomy & Physics and Institute for Computational Astrophysics, Saint Mary's University, 923 Robie Street, Halifax, Nova Scotia B3H 3C3, Canada

118 Departamento Física Aplicada, Universidad Politécnica de Cartagena, Campus Muralla del Mar, 30202 Cartagena, Murcia, Spain

119 Department of Computer Science, Aalto University, PO Box 15400, Espoo 00 076, Finland

120 Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing (GCCL), 44780 Bochum, Germany

121 Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, Avenue des Martyrs, 38000 Grenoble, France

122 Department of Physics and Astronomy, Vesilinnantie 5, 20014 University of Turku, Finland

[123] Serco for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada 28692, Madrid, Spain

[124] ARC Centre of Excellence for Dark Matter Particle Physics, Melbourne, Australia

[125] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Victoria 3122, Australia

[126] W.M. Keck Observatory, 65-1120 Mamalahoa Hwy, Kamuela, HI, USA

[127] Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, Stockholm 106 91, Sweden

[128] Astrophysics Group, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK

[129] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy

[130] INFN-Sezione di Roma, Piazzale Aldo Moro, 2 – c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy

[131] Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal

[132] Zentrum für Astronomie, Universität Heidelberg, Philosophenweg 12, 69120 Heidelberg, Germany

[133] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, Roma, Italy

[134] INFN-Sezione di Roma 2, Via della Ricerca Scientifica 1, Roma, Italy

[135] Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

[136] ASTRON, the Netherlands Institute for Radio Astronomy, Postbus 2, 7990 AA Dwingeloo, The Netherlands

[137] Anton Pannekoek Institute for Astronomy, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands

[138] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

[139] Department of Physics and Astronomy, University of California, Davis, CA 95616, USA

[140] Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA

[141] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
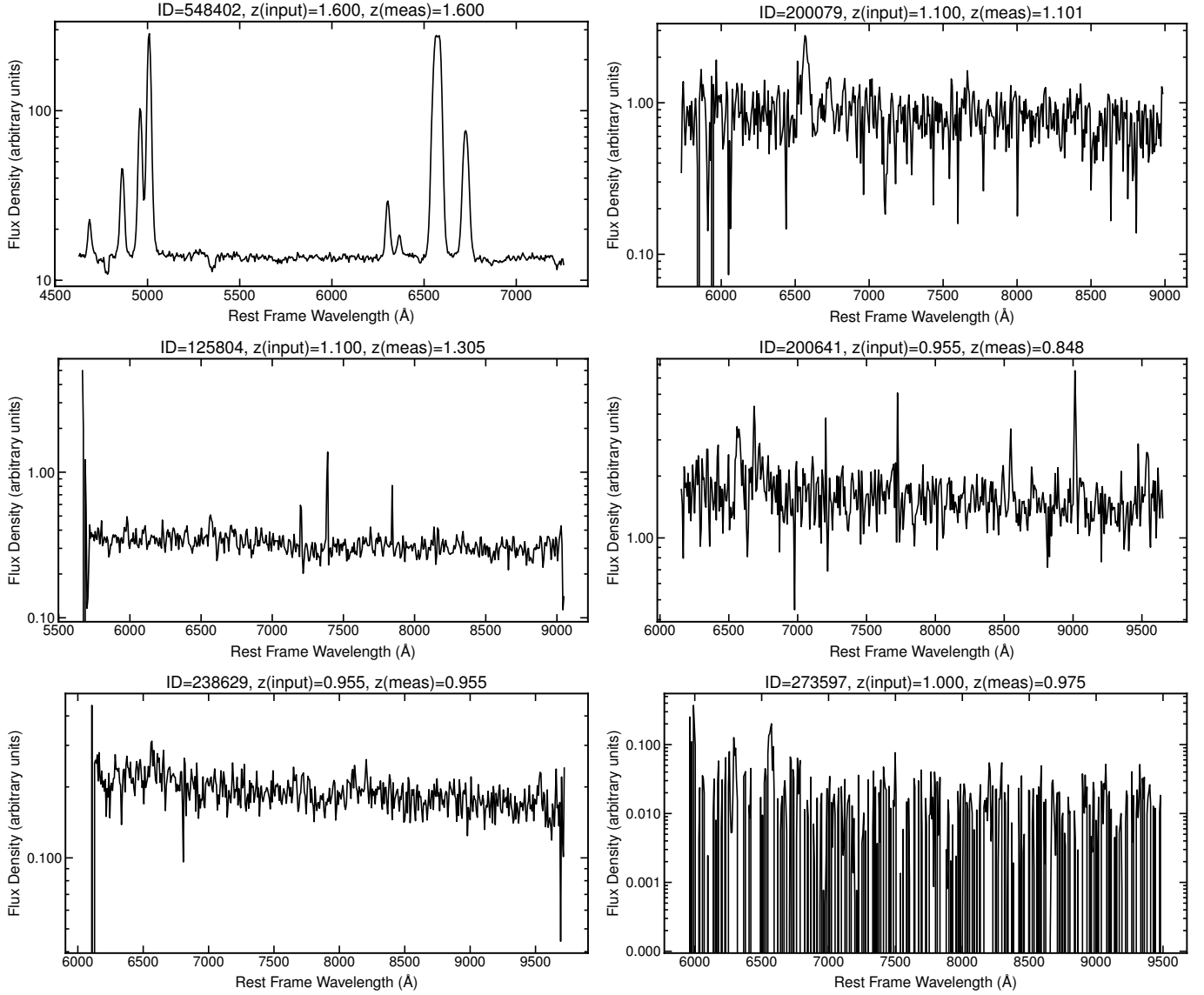
Appendix A: Example of *Euclid*-like simulated spectra

We present in Figure A.1 and in Figure A.2 a few examples of *Euclid*-like simulated spectra for type 1 and type 2 AGN, respec-

tively, in the redshift interval $0.89 \leq z \leq 1.83$, thus where the Hα+[N II] complex is within the spectral coverage of the red grism. Spectra are shown at different S/N and noise levels for the *good*, *false* and *lost*, redshift classes (see Sect. 5).
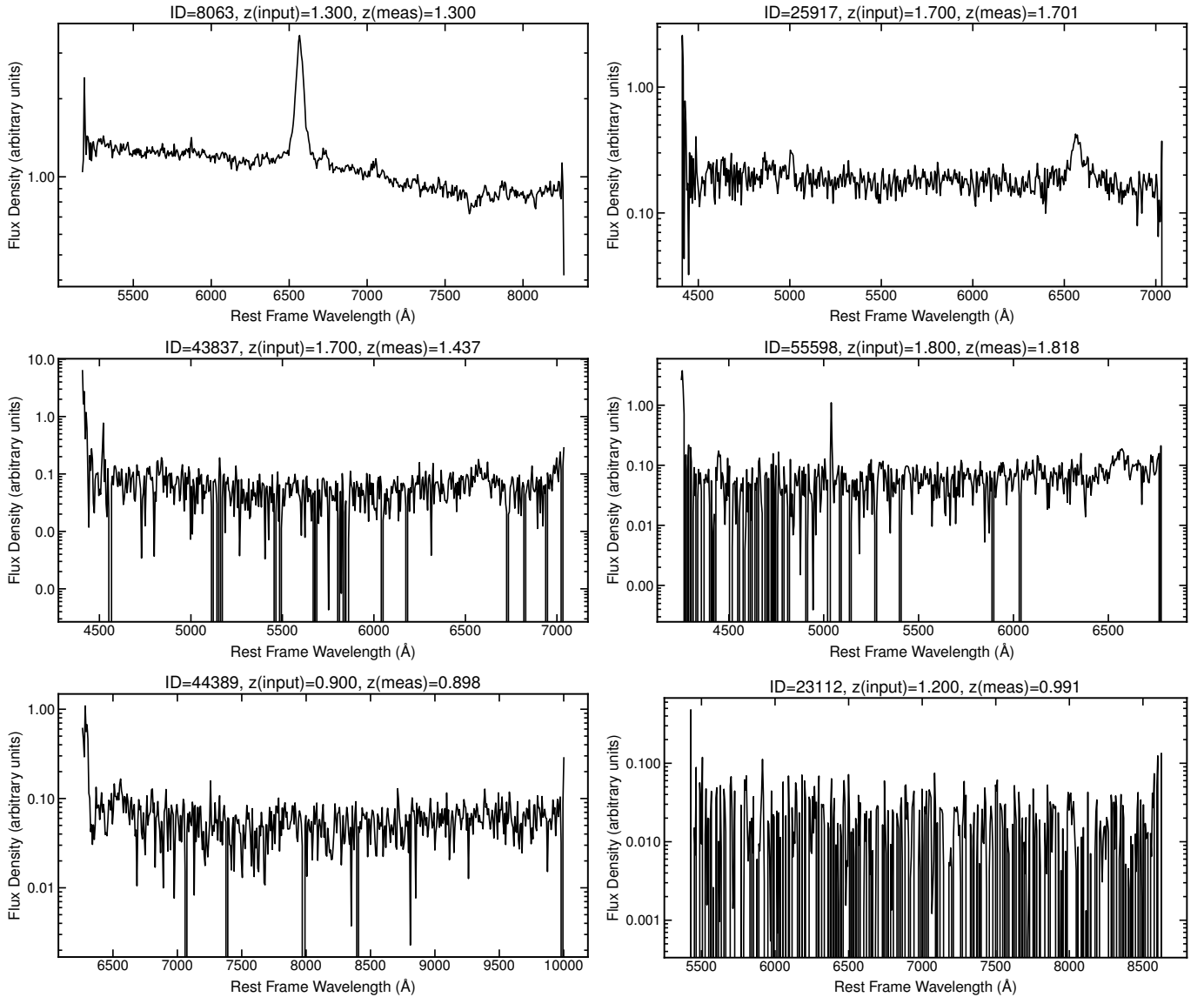
Fig. A.1. Example of *Euclid*-like simulated spectra for type 1 AGN in the redshift interval $0.89 \leq z \leq 1.83$, thus where the Hα+[N II] complex is within the spectral coverage of the red grism. Spectra are shown at different S/N and noise levels (medium/high on the left and low on the right) for the *good* (top row), *false* (mid row), and *lost* (bottom row) redshift classes (see Sect. 5).

Fig. A.2. Example of *Euclid*-like simulated spectra for type 2 AGN in the redshift interval $0.89 \leq z \leq 1.83$, thus where the Hα complex is within the spectral coverage of the red grism. Spectra are shown at different S/N and noise levels (medium/high on the left and low on the right) for the *good* (top row), *false* (mid row), and *lost* (bottom row) redshift classes (see Sect. 5).

Appendix B: Line profile modelling

To better understand whether we are capturing the entire line flux in the incident data, we test two spectral profiles (Lorentzian, and Voigt), in addition to the Gaussian one, to model known emission lines. Figure B.1 shows an example for the Hα complex fit in the type 2 incident spectra with each of the profile options. Gaussian profiles underestimate the peak of the emission lines, while effectively fitting the base and core of the lines. Lorentzian profiles overestimate the tips of the emission lines and underestimate the cores of the lines. Voigt profiles, which are a convolution of Gaussian and Lorentzian profiles, provide a middle ground where the majority of the line flux in the data is captured

by the model. Yet, the tips of the emission lines are still not fully modelled with a Voigt profile. We postulate that this is due to the type 2 incident spectra emission lines being constructed with an AGN and starburst component (i.e., two Gaussians) and so cannot be perfectly modelled using a single line profile. We find that Voigt profiles capture the most flux presented in the data and are hence the best trade-off to be used for type 2 incident spectra emission line modelling. To be consistent with the simulated spectra that are insensitive to the profile employed, we restrict the comparison to the results obtained with the Gaussian one only. Yet, we caveat that a (small) fraction of the flux may be already missed at the level of the spectral modelling of the incident data depending upon the profile assumed.
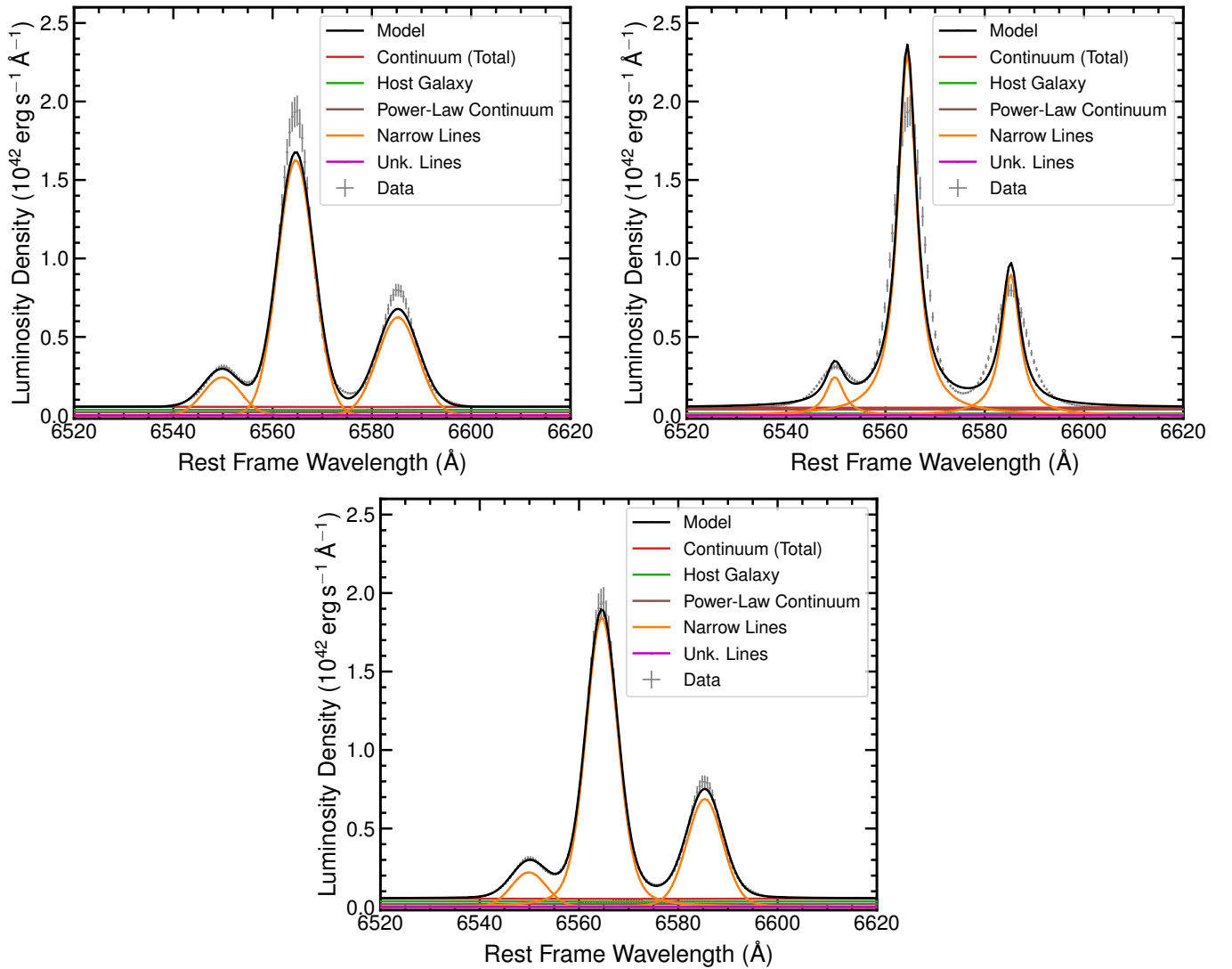

Fig. B.1. Example of a type 2 AGN incident spectrum Hα complex fitted using Gaussian (*left*), Lorentzian (*centre*) and Voigt (*right*) emission line profiles.

Appendix C: Emission line flux from the OU-SPE automated spectral fitting

The Euclid collaboration is currently developing an automated pipeline with the main goal to compute spectral line measurements (e.g. emission lines, absorption lines, continuum breaks) in *Euclid* spectra. Briefly, this automatic routine computes line spectral parameters through Gaussian fitting with a non-linear least-squares minimisation. The Hα complex profile is modelled with three Gaussians, one for each feature (i.e., Hα and [N II]$\lambda\lambda$6548, 6584), on top of a polynomial which accounts for the continuum. For this model, the [N II]λ6584 to [N II]λ6548 ratio is fixed to 3:1 and the FWHMs of the lines are assumed to be the same. In the case of a broad profile (i.e., FWHM> 2000 km s^{-1}), the spectral fit is performed with two Gaussians only, one for the broad component of the Hα and another for

the narrow (i.e., no spectral components are considered for the [N II] doublet). At the time of writing, the pipeline provides tested results for the integrated flux of the Hα emission line complex (flux uncertainties are currently not implemented). Spectral parameters and relative uncertainties for other emission lines are currently under analysis and will be delivered in the updated version of the pipeline.

Figure C.1 presents the comparison of the Hα complex line flux between the QSFIT and the OU-SPE measurements. Fluxes are in very good agreement, with a slightly more scattered dispersion around the one-to-one for the type 1 AGN, as expected. The relative flux difference between the QSFIT and OU-SPE value is defined as $\Delta F/F = (F_{\text{H}\alpha,\text{QSFIT}} - F_{\text{H}\alpha,\text{OU}-\text{SPE}})/F_{\text{H}\alpha,\text{OU}-\text{SPE}}$. The median $\Delta F/F$ are about -9% and 3% for the type 1 and type 2 AGN, respectively.
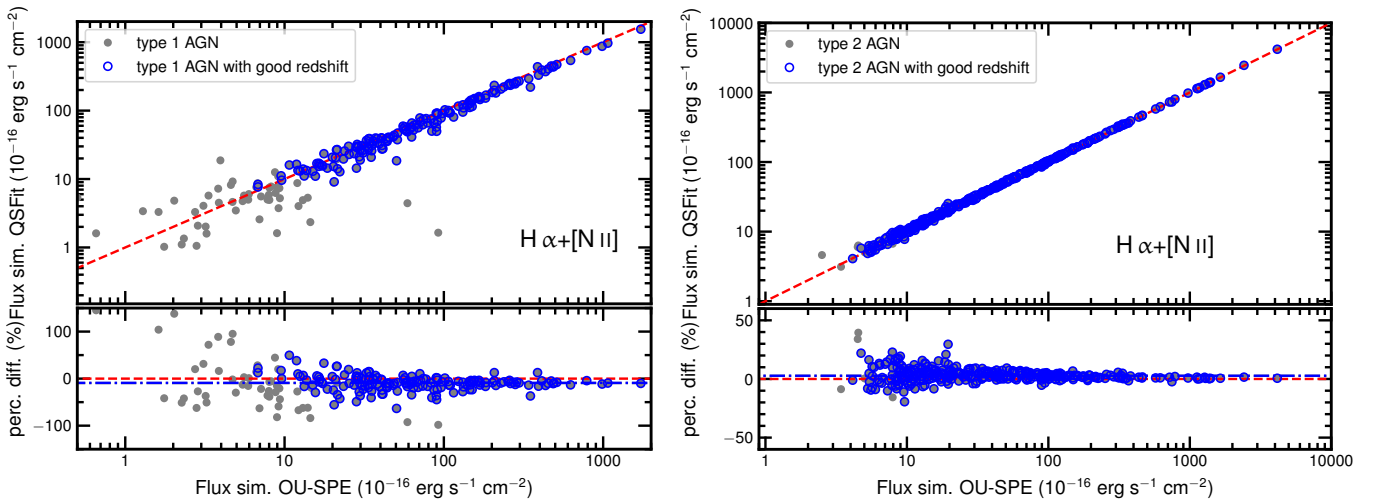
**Fig. C.1.** Comparison of the Hα+[N II] flux measurements of the simulated spectra between QSFIT and OU-SPE. The results for the type 1 and type 2 AGN are shown in the *left* and *right* panels, respectively. Sources with a spectroscopic redshift measured from the simulated spectra and flagged as *good* (see Sect. 5 for details) are marked with blue symbols. The bottom panel shows the percentage flux difference defined as $\Delta F/F = (F_{\text{H}\alpha,\text{QSFIT}} - F_{\text{H}\alpha,\text{OU}-\text{SPE}})/F_{\text{H}\alpha,\text{OU}-\text{SPE}}$. The one-to-one flux relation is shown with a red dashed line, which is set to zero in the bottom panel. The dot-dashed blue line in the bottom panel represents the 50th percentile of the $\Delta F/F$ distribution.

Appendix D: Comparison of the near-infrared spectral region of the type 1 composite with the literature

The list of the sources used to build the NIR composite is given in Table D.1. We further investigated whether the NIR composite we included in the type 1 AGN main stack (Sect. 2) can be improved by including additional NIR spectra of quasars in the literature. For this purpose we collected all the cross-dispersed NIR spectra published by Landt et al. (2008, 2011, 2013). This data set consists on 29 well-known local type 1 AGN observed at the NASA infrared telescope facility (IRTF), a 3 m telescope, and the 8 m Gemini North Observatory, both located on Mauna Kea, Hawai'i (see Table 1 in Landt et al. 2014 for a summary of the sample properties). We note that the a few sources in their sample that display a continuum dominated by the host galaxy emission were not considered here. The data were corrected for Galactic absorption, shifted to the rest-frame and stacked in the same way as done for the SDSS spectra (see Sect. 2.2). The resulting composite is shown in Fig. D.1 with the red line (available at the CDS). Our composite is flatter (redder), likely due to the contribution of the host galaxy emission, and it is more noisy around the Paα emission line, but there is good agreement overall. We defer a more in-depth analysis of the rest-frame NIR region to future work. A version of the composite shown in Fig. 4 corrected for host galaxy emission in the optical, for the IGM absorption in the UV, and that extends in the near-infrared through the spectral composite we build by considering the Landt et al. AGN sample, is also available online. The host galaxy correction at optical wavelengths ($3500 < \lambda/\text{Å} < 7400$) has been performed by assuming the intrinsic continuum of the type 1 template published by Saccheo et al. (2023).

Table D.1. List of AGN used to construct the NIR composite. The SDSS AGN in the table are drawn from Glikman et al. (2006), whilst the remaining are BAT AGN from Ricci et al. (2022).

Name	Redshift
SDSS J000943.1-090839.2	0.210
SDSS J001327.3+005231.9	0.362
SDSS J005709.9+144610.1	0.172
SDSS J005812.8+160201.3	0.211
SDSS J010226.3-003904.6	0.295
SDSS J011110.0-101631.8	0.179
SDSS J015530.0-085704.0	0.165
SDSS J015910.0+010514.5	0.217
SDSS J015950.2+002340.8	0.163
SDSS J021707.8-084743.4	0.292
SDSS J024250.8-075914.2	0.378
SDSS J031209.2-081013.8	0.265
SDSS J032213.8+005513.4	0.185
SDSS J034042.93-073125.5	0.217
SDSS J073623.12+392617.8	0.118
SDSS J125519.6+014412.2	0.343
SDSS J150610.5+021649.9	0.135
SDSS J170441.3+604430.5	0.372
SDSS J172711.8+632241.8	0.218
SDSS J211843.2-063618.0	0.328
SDSS J212619.6-065408.9	0.418
SDSS J232235.6-094438.1	0.372
SDSS J234932.7-003645.8	0.279
SDSS J235156.1-010913.3	0.174
419	0.034
432	0.060
441	0.039
447	0.057
488	0.057
689	0.084
1015	0.056
1079	0.144
1151	0.060
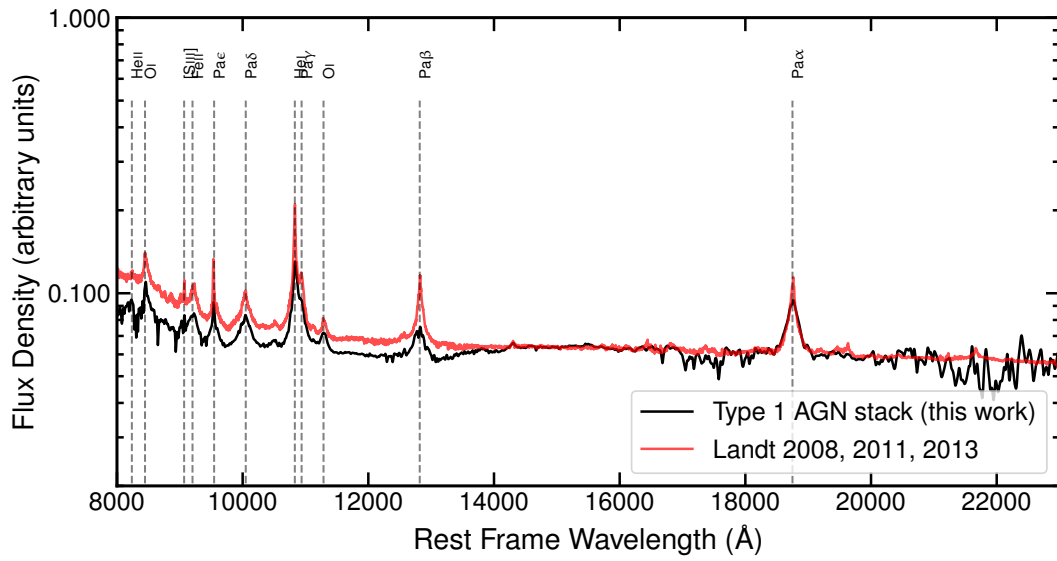
Fig. D.1. Comparison of the empirical composite incident spectrum of type 1 AGN utilised in this work (shown with the black solid line) with the one we constructed from the datasets published by Landt et al. (2008; 2011; 2013, in red). The Landt et al. composite is normalised at the 16 000 Å flux of our stack. Main emission lines are marked.
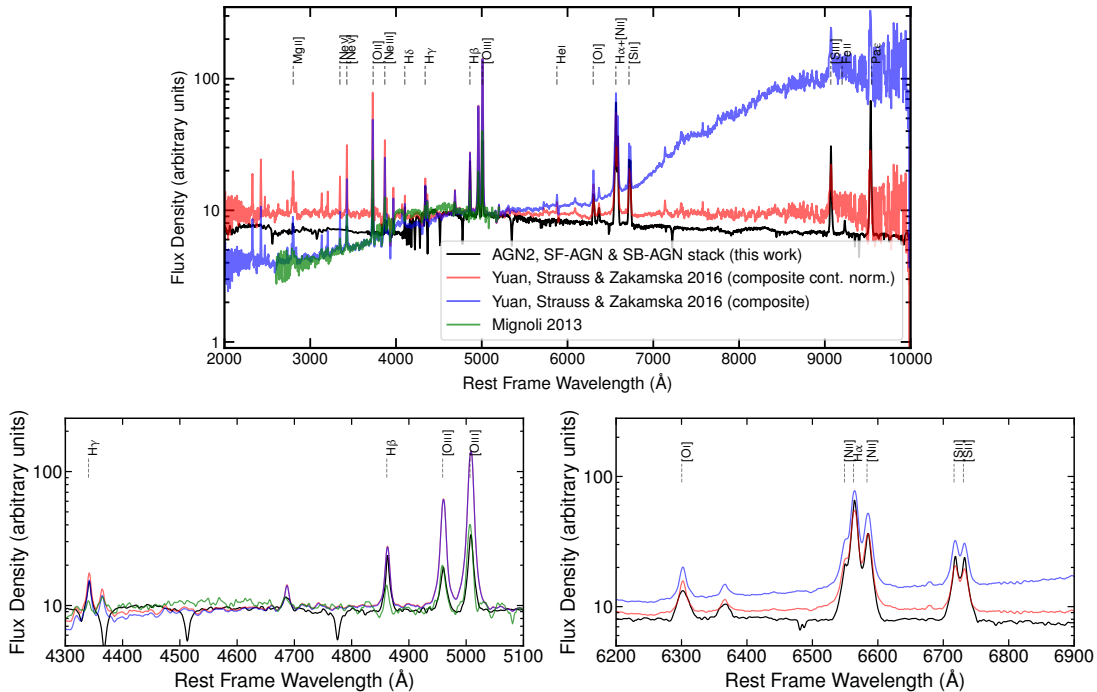
Appendix E: Comparison of the semi-empirical composite spectrum of type 2 AGN with the literature

The composite incident spectra utilised for the type 1 and the type 2 AGN have been computed in a very different way. The former have been derived from observed type 1 spectra of optically selected galaxies from the SDSS, while the latter from simulations (i.e. SPRITZ) with empirical composites of different kind of galaxies (e.g. classical obscured AGN, star-bursts and star forming galaxies with different AGN fractions, see Sect. 3 for details) considered as an input. The choice of using SPRITZ is justified by the fact that it is more difficult to define a clean sample of type 2 AGN. Indeed, even if type 2 AGN display strong narrow high-ionisation emission lines (with FWHM < 2000 km s^{-1}) and weak continua (Zakamska et al. 2003) in the rest-frame optical spectra, the literature works often refer to type 2 AGN as 'candidates', as their selection strongly depends on the diagnostic (and thus the available emission lines) considered (e.g. Hao et al. 2005; Reyes et al. 2008; Alexandroff et al. 2013; Mignoli et al. 2013).

We thus want to test whether our type 2 composite derived from a semi-empirical library is consistent with what is currently available in the literature. Figure E.1 shows the comparison of the semi-empirical template spectra of type 2 AGN utilised in this work (shown with the black solid line) with the empirical one published by Yuan et al. (2016). The blue line represents the composite spectrum of all type 2 quasars in the baryon oscillation spectroscopy survey (BOSS) analysed by Yuan et al. (2016).

The excess flux at $\lambda > 6500$ Å is due to the host galaxy emission, which also seems to be dominant at bluer wavelengths. Therefore, the difference between the composite by Yuan et al. (2016) and ours is mainly due to the different host galaxy. The red line is derived by normalising the composite spectrum by Yuan et al. (2016) to a continuum obtained by spline-interpolating between relatively line-free regions. This stack preserves the equivalent widths of the emission lines, although the intrinsic shape of the continuum is lost (see their Fig. 15 and their Sect. 3.7). Overall, there is good consistency between the two composites regarding the Hα+[N II] complex and the spectral region around the Hβ and [O III]. There are instead some differences in the strength of the emission lines in the UV at $\lambda < 4000$ Å, such as [O II], [Ne III] and Hδ, with a few spectral features missing in our template (e.g. [Ne V]).

To this comparison we also included the composite published by Mignoli et al. (2013, green line in Fig. E.1) obtained from a sample of 94 narrow-line AGN in the redshift range $0.65 < z < 1.20$. Their sample is selected from the 20k-Bright zCOSMOS galaxy sample by the detection of the high-ionisation [Ne V]λ3426 line. This composite shows strong, high-ionisation narrow lines, such as [Ne V], that is not present in our stack, and a very faint Mg II emission line consistent with the emission seen in our composite. We stress that these differences do not affect any result in our analysis, as our comparison is mainly focused on the Hα spectral region rather than the UV and the overall shape of the continuum. In the future, we will implement a more data-driven composite for the type 2 AGN to improve the incident template library for type 2 AGN in the UV.


Fig. E.1. Comparison of the semi-empirical template incident spectra of type 2 AGN utilised in this work (shown with the black solid line) with the empirical one published by Yuan et al. (2016). The blue line represents the composite spectrum of all BOSS type 2 quasars in the sample analysed by Yuan et al. (2016), whilst the red line is derived by normalising the composite spectrum to a continuum obtained by spline-interpolating between relatively line-free regions. The green line represents the spectral stack of 94 narrow-line AGN published by Mignoli et al. (2013). For sake of comparison, the literature composites are normalised at the 5100 Å flux of our stack. Main emission lines are marked. The bottom panels are two zoom-ins around the Hβ (*left*) and the Hα (*right*) the spectral regions. Stellar absorption features are also included and visible in our type 2 AGN spectrum between Hγ and Hβ (see also Fig. A1 by Vidal-García et al. 2017 and Figs. 9 e 10 by Bruzual & Charlot 2003), which will be blended to the continuum in any moderate (low) resolution spectrum.

Appendix F: Analysis of the flux loss

Here we present some additional test we performed to investigate whether a fraction of the missing flux could be lost at the slitless spectra extraction level (see Sect. 5.1.3 in Euclid Collaboration 2023a, for details). We first compared the distribution of the ratios r of the median observed flux values of the incident spectra F_{inc} and the simulated spectra F_{sim} (i.e., $r = \langle F_{sim} \rangle / \langle F_{inc} \rangle$, where the angle brackets $\langle \rangle$ represent medians); see Fig. F.1. The median of this distribution is about 12% for type 1 AGN and 2% for type 2 AGN, with a percentage error on the distributions of spectrum flux median values (defined as $1.253\,\sigma/\sqrt{N}$, where σ and N represent the standard dispersion of the distribution and the number of objects, respectively) at the 2.5% level for type 1 and 1% for type 2. As an example, we present a comparison between the incident (black) and simulated (red) spectra for a representative type 1 (left) and a type 2 (right) AGN at the top row of Fig. F.1. The inset plot is a zoom-in of a continuum (emission line free) region. The overall flux of the simulated spectra are systematically lower for both the type 1 and 2 AGN. For completeness, we also carried out the same exercise of comparing median luminosity density values for the best-fit models of the incident and simulated spectra to their respective data [i.e. ratio = median(data) / median(model)], finding that the QSFIT best-fit model is close to the data, on average, with the biggest discrepancy for type 1 simulated spectra at ≃2% level. This analysis implies that flux loss could partly explain the remaining fraction of flux missing in the comparison between simulated and incident flux values, although most of the difference on the Hα+[N II] emission between the simulated and incident data is due to the assumptions made in the spectral modelling of the emission lines. Nonetheless, the OU-SIR team is working on the spectral extraction pipeline and flux losses will be alleviated by the optimal extraction routine that is implemented in the upcoming version of the OU-SIR pipeline.
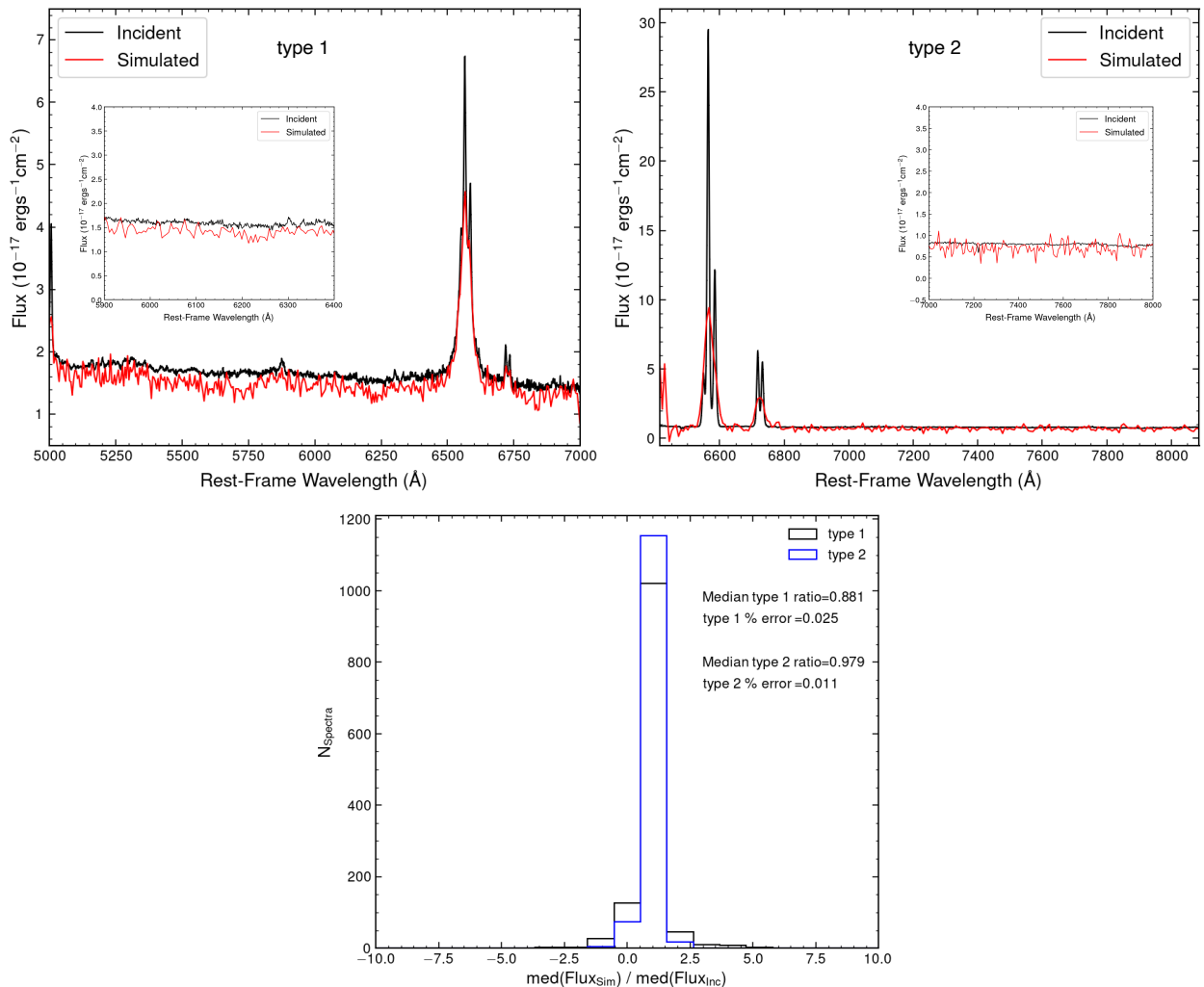
Fig. F.1. *Top row*: Example of a comparison between the incident (black) and simulated (red) spectra for a type 1 (*left*) and a type 2 (*right*) AGN around the Hα region. The inset plot is a zoom-in of a continuum (emission line free) region. *Bottom row*: distribution of the ratios of the median flux value (i.e., the median of the observed flux of the total spectrum) of the incident and simulated data for both type 1 (black) and type 2 (blue) AGN. The median values for the two distributions are also reported. The percentage error on the distributions of the spectrum flux median values are computed as $1.253\,\sigma/\sqrt{N}$, where σ and N represent the standard dispersion of the distribution and the number of objects, respectively.