

A structure function analysis of VST-COSMOS AGN^{★,★★}

D. De Cicco^{1,2,3}, F. E. Bauer^{1,2,4}, M. Paolillo^{3,5,6}, P. Sánchez-Sáez^{7,2,1}, W. N. Brandt^{8,9,10}, F. Vagnetti^{11,12},
G. Pignata^{13,2}, M. Radovich¹⁴, and M. Vaccari^{15,16,17}

¹ Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile
e-mail: demetra.decicco@unina.it, demetradecicco@gmail.com

² Millennium Institute of Astrophysics (MAS), Nuncio Monseñor Sotero Sanz 100, Of. 104, Providencia, Santiago, Chile

³ Dipartimento di Fisica, Università degli Studi di Napoli “Federico II”, Via Cinthia 9, 80126 Napoli, Italy

⁴ Space Science Institute, 4750 Walnut Street, Suite 2015, Boulder, CO 80301, USA

⁵ INFN – Sezione di Napoli, Via Cinthia 9, 80126 Napoli, Italy

⁶ INAF – Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy

⁷ European Southern Observatory, Karl-Schwarzschild-Strasse 2, 85748 Garching bei München, Germany

⁸ Department of Astronomy and Astrophysics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA

⁹ Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

¹⁰ Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA

¹¹ Dipartimento di Fisica, Università di Roma “Tor Vergata”, Via della Ricerca Scientifica 1, 00133 Roma, Italy

¹² INAF – Istituto di Astrofisica e Planetologia Spaziali, Via del Fosso del Cavaliere 100, 00133 Roma, Italy

¹³ Departamento de Ciencias Físicas, Universidad Andres Bello, Avda. Republica 252, Santiago, Chile

¹⁴ INAF – Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, 35122 Padova, Italy

¹⁵ Inter-university Institute for Data Intensive Astronomy, Department of Physics and Astronomy, University of the Western Cape, 7535 Bellville Cape Town, South Africa

¹⁶ Inter-university Institute for Data Intensive Astronomy, Department of Astronomy, University of Cape Town, 7701 Rondebosch, Cape Town, South Africa

¹⁷ INAF – Istituto di Radioastronomia, Via Gobetti 101, 40129 Bologna, Italy

Received 25 November 2021 / Accepted 23 May 2022

ABSTRACT

Context. We present our sixth work in a series dedicated to variability studies of active galactic nuclei (AGN), based on the survey of the COSMOS field by the VLT Survey Telescope (VST). Its 54 *r*-band visits over 3.3 yr and single-visit depth of 24.6 *r*-band mag make this dataset a valuable scaled-down version that can help forecast the performance of the Rubin Observatory Legacy Survey of Space and Time (LSST).

Aims. This work is centered on the analysis of the structure function (SF) of VST-COSMOS AGN, investigating possible differences in its shape and slope related to how the AGN were selected, and explores possible connections between the AGN ensemble variability and the black-hole mass, accretion rate, bolometric luminosity, redshift, and obscuration of the source. Given its features, our dataset opens up the exploration of samples ~ 2 mag fainter than most literature to date.

Methods. We identified several samples of AGN – 677 in total – obtained through a variety of selection techniques partly overlapping. Our analysis compares the results for the various samples. We split each sample in two based on the median of the physical property of interest, and analyzed the differences in the SF shape and slope, and their possible causes.

Results. While the SF shape does not change with depth, it is highly affected by the type of AGN (unobscured or obscured) included in the sample. Where a linear region can be identified, we find that the variability amplitude is anticorrelated to the accretion rate and bolometric luminosity, consistent with previous literature on the topic, while no dependence on black-hole mass emerges from this study. With its longer baseline and denser and more regular sampling, the LSST will allow for an improved characterization of the SF and its dependencies on the mentioned physical properties over much larger AGN samples.

Key words. galaxies: active – X-rays: galaxies – infrared: galaxies – surveys – methods: statistical

1. Introduction

Variability is a characteristic of active galactic nuclei (AGN) at all wavelengths, observed in the continuum emission as well as in the emission lines. Because of its universal char-

acter, it can be used as a selection criterion for AGN; in particular, the search for AGN in multivisit surveys via optical variability has been the subject of a large number of studies over the past decades (e.g., MacLeod et al. 2012; De Cicco et al. 2015, 2019, 2021; Falocco et al. 2015; Simm et al. 2016; Sánchez-Sáez et al. 2018; Sartori et al. 2019; Poulain et al. 2020, and references therein). This variability is stochastic and aperiodic, as suggested by their featureless power spectra (e.g., Kelly et al. 2009). Optical emission from AGN is generally thought to originate from the accretion disk

* Table 6 is only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/664/A117>

** Observations were provided by the ESO programs 088.D-4013, 092.D-0370, and 094.D-0417 (PI G. Pignata).

surrounding the central supermassive black hole (SMBH), and the corresponding variations are most commonly thought to originate from instabilities in the accretion disk and changes in the accretion rate (e.g., [Shakura & Sunyaev 1976](#)), even though this has been a somewhat controversial take in the last decades (e.g., [Aretxaga & Terlevich 1994](#); [Kawaguchi et al. 1998](#); [Hawkins 2002](#); [Kimura et al. 2020](#)). As a consequence, AGN monitoring campaigns are a powerful tool when conducting deeper investigations into the topic.

One way to quantitatively describe the time dependence of the variability of a sample of AGN is via the structure function (SF; e.g., [Simonetti et al. 1985](#); [di Clemente et al. 1996](#); [Kawaguchi et al. 1998](#); [Hawkins 2002](#); [de Vries et al. 2005](#); [Bauer et al. 2009](#); [Graham et al. 2014](#); [Kozłowski 2016](#), and references therein). This function is built from the light curves of the sample of sources and is basically a measure of the ensemble root mean square (rms) magnitude difference as a function of the time lag between different visits, where measurements are typically grouped together into bins. Such a statistical approach allows for the handling of large samples of objects and a characterization of their overall variability and average properties, and its strength lies in the mutual independence of the different bins over which we measure variability, which does not hold for individual sources (e.g., [de Vries et al. 2005](#)).

While the physical process driving the observed red-noise variability is still unknown, the damped random walk (DRW; e.g., [Kelly et al. 2009](#)) model has proven to be an effective characterization of AGN light curves. The DRW describes AGN light curves as a stochastic process via an exponential covariance matrix defined by two parameters: a variability amplitude and a characteristic damping timescale. For short timescales the random walk provides a good description of the process, while for longer timescales the variations are damped and tend to an asymptotic amplitude (e.g., [Ivezić et al. 2014](#); [Zu et al. 2013](#)). Some works from the literature (e.g., [MacLeod et al. 2010](#)) have found the DRW parameters to be correlated with properties such as wavelength, AGN luminosity, and black-hole (BH) mass; in addition, there are speculations about the damping timescale being related to the thermal timescale of the AGN accretion disk (e.g., [Kozłowski 2017](#); [Burke et al. 2021](#)). Nevertheless, as a model DRW is limited by its inability to constrain the underlying physical process, and the fact that the same variability can be described effectively also by other models (e.g., [Kozłowski 2016](#)). In addition, the use of DRW to model AGN light curves is affected by some limitations due to the data, as detailed in [Kozłowski \(2017\)](#) where the author generally proves that the sampled timescales should be at least ten times the damping timescale (1 yr, leading to >10 yr rest-frame length for the light curves) in order to allow for a proper modeling and recovery of accurate DRW parameters. These results have also been confirmed by [Suberlak et al. \(2021\)](#), who investigated the impact of the ratio of the timescale and the light curve baseline by means of simulated light curves, finding that, as this ratio decreases, it allows for the recovery of unbiased DRW parameters. Thus, it does not affect the analysis of the relations with physical properties of the AGN.

Here we present an analysis of the SF of AGN centered on the 1 sq. deg. area of the Cosmic Evolution Survey (COSMOS; [Scoville et al. 2007](#)) imaged by the VLT Survey Telescope (VST; [Capaccioli & Schipani 2011](#)). We note that the dataset was introduced and extensively described in [De Cicco et al. \(2015, 2019\)](#). Together with [De Cicco et al. \(2021\)](#), these works are part of a series of five dedicated to variability studies which are relevant in the context of performance forecasting

for the Legacy Survey of Space and Time (LSST; see, e.g., [LSST Science Collaboration 2009](#); [Ivezić et al. 2019](#)), which will be conducted with the Simonyi Survey Telescope at the Vera C. Rubin Observatory. The series illustrates the development of an efficient and automated methodology for the identification of optically variable AGN; in particular, [De Cicco et al. \(2021\)](#) selected a sample of AGN candidates via a random forest (RF; [Breiman 2001](#)) classifier built using optical variability, optical and near-infrared (NIR) colors, and a morphology indicator as features. A set of multiwavelength diagnostics was used for confirmation, examining the properties imparted by each selection technique.

In this work we aim to characterize the sample of known AGN in the VST-COSMOS area, which consists of the confirmed sources that we selected via optical variability plus the samples of multiwavelength-selected AGN that we used for the validation of our selection. We therefore want to assess whether the shape of the SF changes when using different AGN samples based on observed and intrinsic properties such as flux, redshift, and obscuration, with the aim of understanding their variability properties.

Via the SF we also aim to investigate possible connections between AGN variability – in terms of SF amplitude and slope – and bolometric luminosity L_{bol} , BH mass M_{BH} , and accretion rate – as quantified by the Eddington ratio λ_E – onto the SMBH. The existence of these connections has been discussed by several studies in the last decades, especially in the X-rays, which map the AGN inner regions (e.g., [McHardy et al. 2006](#); [González-Martín & Vaughan 2012](#); [Vagnetti et al. 2016](#); [Paolillo et al. 2017](#), and references therein). If we manage to determine the connection between AGN variability and L_{bol} , M_{BH} , and λ_E of the central SMBH, we can use the first to characterize the latter, especially in the context of the wide-field surveys to come, such as the LSST. This approach is model independent and helpful when the sampled timescales are too short to adopt, for instance, a DRW-based approach.

A large number of relatively recent studies have been dedicated to this inquiry, but the investigation of the potential relations between variability and M_{BH} remains controversial, as different studies arrive at conflicting conclusions. For example, [MacLeod et al. \(2010\)](#) modeled the optical/UV variability of a sample of ≈ 9000 quasars in the Stripe 82, adopting a damped random walk model and making use of light curves in the *ugriz* bands from the Sloan Digital Sky Survey (SDSS; [York et al. 2000](#)) spanning 10 yr. They investigated possible correlations between two temporal or variability-related parameters – a characteristic damping timescale and the asymptotic rms variability – as well as some physical parameters, finding that both are correlated with M_{BH} . [Wilhite et al. \(2008\)](#) obtained similar results from their analysis of the ensemble variability of ≈ 8000 quasars in the SDSS Equatorial Stripe. In this case they focused on the variability amplitude via the SF of their sources. Conversely, [Kelly et al. \(2013\)](#) analyzed correlations of optical and X-ray variability (via power spectral density, PSD) with luminosity, M_{BH} , and λ_E for a sample of 39 AGN, and found an anticorrelation with M_{BH} for both wavebands, with a larger scatter characterizing the relation with the optical variability amplitude. Meanwhile, [Simm et al. \(2016\)](#) searched for correlations for AGN in the XMM-COSMOS catalog from [Brusa et al. \(2010\)](#), with AGN optical variability quantified by the excess variance, finding no dependence on M_{BH} . [Caplar et al. \(2017\)](#) resorted to characterization using the SF and PSD to analyze optical variability in quasars from the (intermediate) Palomar Transient Factory ([i]PTF) and its dependence on the physical parameters of the

central BH, finding a weak correlation of the variability amplitude with M_{BH} .

Many of these same works point to λ_E as the main driver for AGN variability: [Simm et al. \(2016\)](#) indeed found an anticorrelation with this quantity; [Sánchez-Sáez et al. \(2018\)](#) used data from the QUEST-La Silla AGN variability survey to show that the variability amplitude anticorrelates with λ_E ; and [Wilhite et al. \(2008\)](#), [MacLeod et al. \(2010\)](#), and [Kelly et al. \(2013\)](#) additionally support the existence of this anticorrelation. [Wilhite et al. \(2008\)](#), [MacLeod et al. \(2010\)](#), [Kelly et al. \(2013\)](#) also found an anticorrelation between AGN variability and luminosity, as well as [Simm et al. \(2016\)](#), who resorted to excess variance (EV) and PSD. [Laurenti et al. \(2020\)](#) investigated the optical and X-ray variability properties in a sample of 795 AGN from the Multi-Epoch XMM Serendipitous AGN Sample 2 (MEXAS2; [Serafinelli et al. 2017](#)), making use of data from the Catalina Surveys Data Release 2 (CSDR2¹), and finding a strong anticorrelation with luminosity, a weaker anticorrelation with λ_E , and a modest correlation with M_{BH} .

A key novelty of this work is the depth of the analyzed sample, which probes up to $r \simeq 23.5$ mag, thus opening up the exploration of more frequently observed AGN, as compared to the typically 1.5–2 mag brighter samples in previous studies. A relevant feature is that the single-visit depth of VST-COSMOS r -band images is ≈ 24.6 mag for point sources, at a 5σ confidence level, hence roughly the same as the LSST typical single-visit depth of 24.7 mag for the same band.

This paper is organized as follows: Sect. 2 introduces the samples of AGN selected for our analysis on the basis of different properties, while Sect. 3 describes the definition of the SF adopted in this work and presents the SF obtained for the different samples of AGN; Sect. 4 investigates the connection between optical variability and M_{BH} , λ_E , L_{bol} , redshift, and absorption; we summarize our main findings in Sect. 5. Unless otherwise stated, throughout this paper, we refer to optical variability even when omitting the word “optical”.

2. The VST-COSMOS AGN sample

The VST-COSMOS dataset used in this work consists of 54 visits surveying a ≈ 1 sq. deg. area in the r band, spanning a 3.3 yr baseline and spread over three observing seasons, with two gaps. Table 1 reports information about the observed (i.e., not redshift-corrected) baseline and the number of visits for the three observing seasons, with basic statistics for the sample of sources analyzed. We refer to [De Cicco et al. \(2015, 2019, 2021\)](#) for further details. Magnitudes are in the AB system.

Since COSMOS is a widely studied area, a large number of sources are already known to be AGN based on a number of properties and diagnostics. Specifically, limited to the COSMOS area surveyed by the VST, we identified 677 sources confirmed as AGN (hereafter, main sample) on the basis of at least one of the following: optical spectroscopy ([Marchesi et al. 2016](#)), X-ray to optical flux ratio ([Maccacaro et al. 1988](#)), mid-infrared (MIR) properties ([Donley et al. 2012](#)). These diagnostics have been widely used in [De Cicco et al. \(2015, 2019, 2021\)](#) to validate the samples of AGN candidates selected via optical variability, which is the remaining selection criterion here adopted. These 677 AGN have counterparts in the sample of VST-COSMOS sources used in [De Cicco et al. \(2021\)](#), which includes sources with a magnitude $r \leq 23.5$ mag, and a redshift estimate, either spectroscopic (91% of sources) or photometric (9% of sources),

available from different COSMOS catalogs. Details about the catalogs and on how we assigned a redshift to each source can be found in Sect. 2.1 of [De Cicco et al. \(2021\)](#). By construction, the catalog contains sources with a minimum of 27 points in their light curves, the maximum being defined by the total number of visits (54); specifically 97% of the sources in our main sample contain at least 40 points in their light curves. Figure 1 shows a selection of our light curves, reporting for each one the length of the baseline, the number of visits, and the subsample(s) the corresponding source belongs to.

In the following, we analyze the SF of the main AGN sample and also of five subsamples, which partly overlap each other, with the aim of investigating the dependence on the selection criteria and the resulting variability properties². In particular, we focus on: X-ray selected AGN; Type I and Type II AGN subsamples, defined after the properties of their optical spectra reported in the X-ray catalogs by [Marchesi et al. \(2016\)](#), [Brusa et al. \(2010\)](#)³; MIR selected AGN; and AGN identified from their optical variability. Specifically, this last subsample is the one obtained from the analysis presented in [De Cicco et al. \(2021\)](#). The Venn diagram⁴ in Fig. 2 provides a visual assessment of the overlap among the five subsamples of AGN constituting our main sample. We stress that since the subsamples of spectroscopically confirmed Type I and Type II AGN are drawn from X-ray catalogs (see above), we have X-ray information for each of these sources and hence they are completely included in the X-ray subsample. We also note that the subsample of spectroscopic Type I sources is completely included in the subsample of optically variable AGN as well, since optical variability is highly biased towards this type of AGN (e.g., [Padovani et al. 2017](#), and references therein), and our selection method based on optical variability allowed us to retrieve all the sources in the Type I subsample ([De Cicco et al. 2021](#)).

In order to unearth possible correlations between variability and some physical SMBH properties in our sample of AGN, we selected those sources in the main AGN sample for which an estimate of M_{BH} and λ_E are known. We make use of estimates for these quantities drawn from several works, listed in the following according to the preference with which they were used: [Lusso et al. \(2012\)](#), where X-ray selected Type I and Type II AGN in COSMOS are investigated; M_{BH} estimates for Type I AGN are virial and were obtained from Mg II or H β lines (81 matches; 0.25–0.4 dex uncertainties, depending on the line);

² Hereafter we typically refer to these five subsamples including the prefix “sub” to stress that they are part of our main sample. We omit the prefix when talking about all six samples together or generically, as in that case we will also be including the main sample, which is not a subsample.

³ In [Marchesi et al. \(2016\)](#) sources are classified as BLAGN if their spectra exhibit at least one broad ($FWHM > 2000 \text{ km s}^{-1}$) emission line, while sources labeled as non-BLAGN could be NLAGN or star-forming galaxies: this can be due to low S/N spectra, or to the lack of disentangling diagnostics based on emission lines in the waveband in which the spectra are obtained. In this case we crossmatch the classification with the one provided in the XMM-COSMOS catalog by [Brusa et al. \(2010\)](#), where sources are classified as BLAGN, NLAGN, and inactive galaxies. Here, the criterion identifying BLAGN is the same as in [Marchesi et al. \(2016\)](#); the spectra of sources flagged as NLAGN are typically characterized by unresolved high-ionization emission lines with line ratios suggesting AGN activity, while inactive galaxy spectra are generally consistent with those of star-forming or normal galaxies and, when detected in the hard X-rays, they generally have rest-frame luminosity $L_X < 2 \times 10^{42} \text{ erg s}^{-1}$ in that band.

⁴ The Venn diagram in Fig. 2 was rendered via <http://bioinformatics.psb.ugent.be/webtools/Venn/>.

¹ <http://nesssi.cacr.caltech.edu/DataRelease/>

Table 1. Basic statistical information about the three observing seasons of data used in this work.

	Duration	Obs. baseline (d)	Mean obs. baseline (d)	Median obs. baseline (d)	tot. # of visits	Mean # of visits	Median # of visits	Mean sampling rate (d)	Median sampling rate (d)
Season 1	Dec. 2011–May 2012	151	150	151	26	24	25	6.04	3
Season 2	Dec. 2013–Apr. 2014	101	101	101	20	19	20	5.32	3
Season 3	Dec. 2014–Mar. 2015	106	103	106	8	7	8	15.14	15
Seasons 1+2+3	Dec. 2011–Mar. 2015	1187	1185	1187	54	51	52	22.40	4

Notes. The table reports (left to right): maximum observed baseline; mean and median value of the observed baseline for the 677 sources in the main sample; total number of visits; mean and median number of visits for the 677 sources in the main sample; mean and median sampling rate.

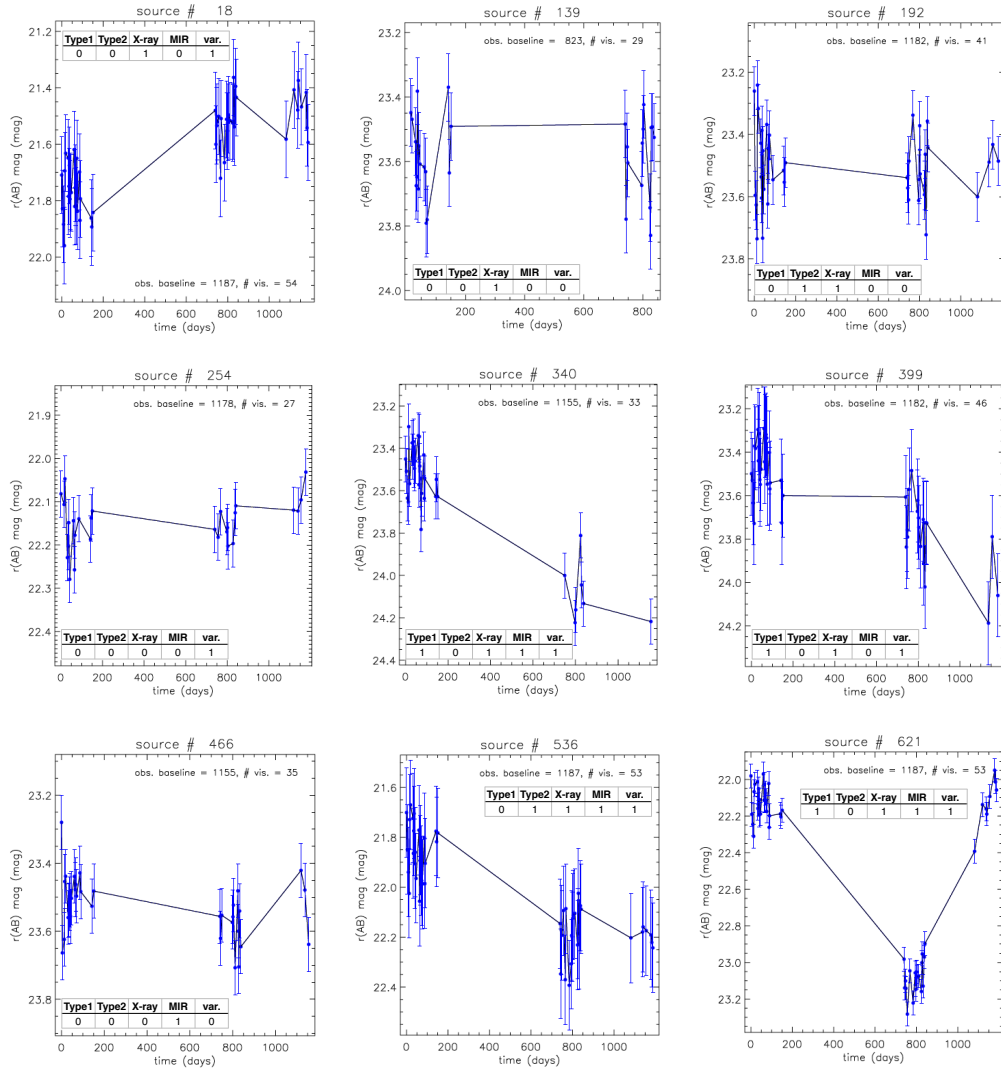


Fig. 1. Selection of light curves from the main sample (baselines are in days). The source numbered 18 is an example of the most commonly observed ones in this sample, where both the length of the observed baseline and the number of visits coincide with the corresponding maxima (see Table 1); source 139 is the one with the shortest baseline, while source 254 is one of the two detected in 27 visits (i.e., the minimum number of visits based on our selection threshold; see main text); the other instances correspond to other combinations of observed baseline and number of visits. We note that the light curve of source 621 is characterized by the largest magnitude variation over the analyzed baseline. The table in each panel indicates which subsample(s) the corresponding source belongs/does not belong to (1/0).

whereas for Type II AGN (106 matches; 0.5 dex average uncertainty) they were obtained via scaling relations and uncertainties on the estimates of stellar masses, bolometric luminosities as well as on the intrinsic scatter in the relation between M_{BH} and stellar mass were taken into account via Monte Carlo simulations; Schulze et al. (2018), where a sample of X-ray selected

Type I AGN is investigated via NIR spectroscopy and virial M_{BH} estimates (80 matches) are obtained from $H\alpha$, $H\beta$, and Mg II ; Rosario et al. (2013, priv. comm.), where the possible dependence of star formation rate on redshift, M_{BH} , nuclear L_{bol} is investigated in a sample of spectroscopically observed quasars in COSMOS, out to a redshift $z \sim 2$, and virial M_{BH} estimates

Table 2. Number of AGN included in the main sample and in the various subsamples selected following different criteria and used in this study.

	Main sample	X-ray	MIR	Opt. var.	Spec. Type I	Spec. Type II
Total number	677	605	225	301	225	122
$\bar{z}$	0.876	0.893	1.214	1.441	1.653	0.696
M_{BH} and λ_E estimate	264	260	117	183	152	85

Notes. The columns in the table indicate the number of AGN selected via (left to right): X/O ratio, MIR selection, optical variability selection, spectroscopic properties (returning Type I and Type II subsamples). For each sample we also report the median redshift $\bar{z}$ and the number of sources for which M_{BH} and λ_E estimates are available.

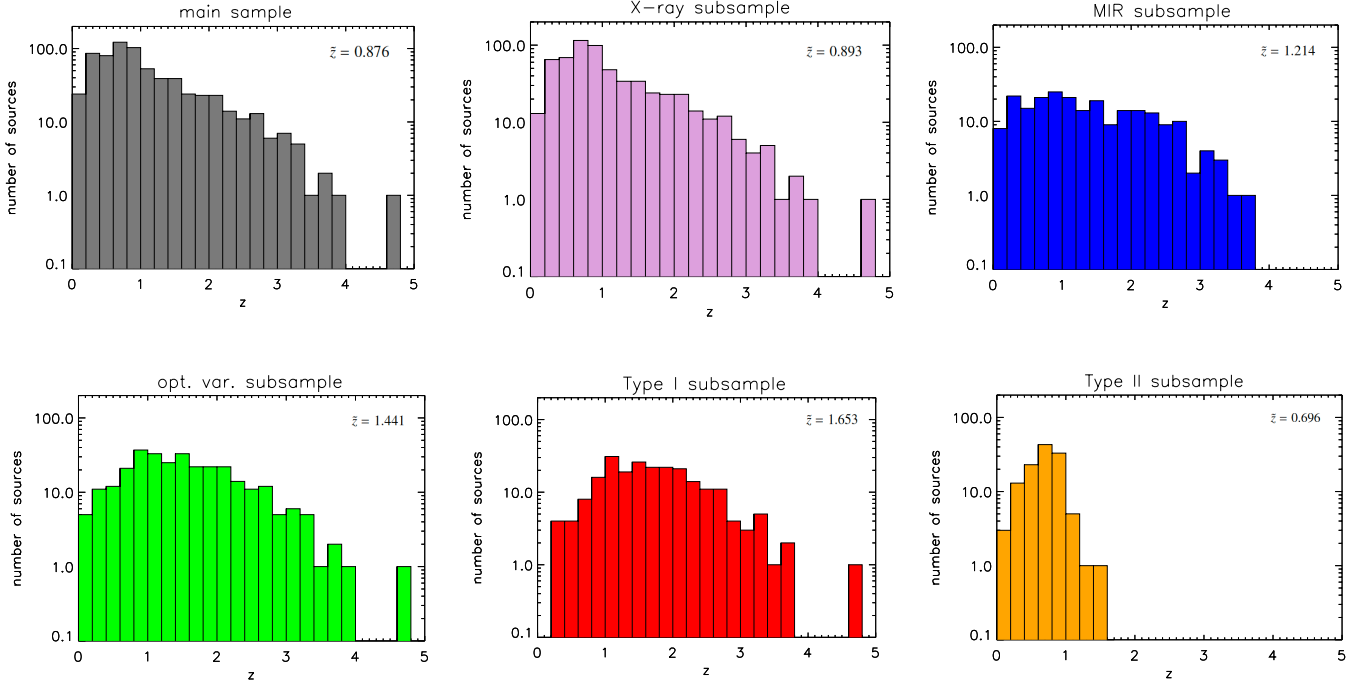


Fig. 3. Redshift distribution for the main sample of 677 AGN (*top left*), and for the five subsamples of AGN selected by means of different techniques: AGN selected via their X-ray properties (*top center*), AGN selected via their MIR properties (*top right*), AGN selected via optical variability (*bottom left*), and Type I and Type II AGN confirmed by spectroscopy (*bottom center* and *bottom right*, respectively). Each panel reports the median redshift value for the corresponding sample. Details about the selection criteria are reported in Sect. 2.

3.2. The SF of our six samples of AGN

In Fig. 5 we show the SFs obtained for each of our six samples of AGN selected by means of different properties and investigated in this study. Time differences are measured at rest frame in order to be able to study the dependence of the intrinsic variability on time. Each panel includes the SF obtained when adopting a 22.5 mag threshold, with the aim of assessing whether a different depth leads to significant differences in the SF shape.

The plot corresponding to the subsample of Type I AGN is the one representing the typical AGN structure function; the one corresponding to optically variable AGN is very similar, as that subsample is highly biased towards Type I AGN (75% of the sources). In these two plots three distinct regions can be identified. When time lags are too short to detect variability, we can see small amplitude variations characterized by larger error bars. Starting from time lags of ≈ 10 days, the SF can be represented via a power law and, thus, on a logarithmic scale it is approximately linear; this linear trend originates from the shape of the autocorrelation function describing the stochastic process leading to variability and, thus, it is related to the properties of the light curves. For timescales longer than ≈ 250 days ($\log(\text{time}) \approx 2.4$) we observe a drop in variability amplitude.

We may be tempted to believe that we reached a maximum in variability amplitude. Indeed, several works have shown how, as the time difference Δt increases, a turnover is observed at a characteristic timescale τ , and then the SF switches from a red-noise regime to a white-noise regime (e.g., Bauer et al. 2009; Kozłowski 2016). However this τ is estimated to be on the order of 10^2 rest-frame days at optical wavelengths for SDSS quasars (e.g., Collier & Peterson 2001; Kelly et al. 2009; MacLeod et al. 2010); this suggests that we need a more extended baseline in order to investigate this timescale without the current sampling limitations. Thus, we believe that the observed turnover is not real but an effect of the poor sampling at large time lags and it would not be observed if the sampling were even and if it covered longer timescales (e.g., Rengstorff et al. 2006; Bauer et al. 2009; Emmanoulopoulos et al. 2010). In particular, Bauer et al. (2009) investigate the windowing effects due to irregular data sampling by simulating quasar light curves with uniform cadence, and their results support the thesis that the turnover is not evidence of the existence of a maximum timescale in AGN variability but, rather, an effect of inadequate sampling.

We visually identify the linear region in correspondence of the logarithmic values 1.0 – 2.6 for the baseline, that is to say ≈ 10 –400 days. The measured slopes for the linear regions of the

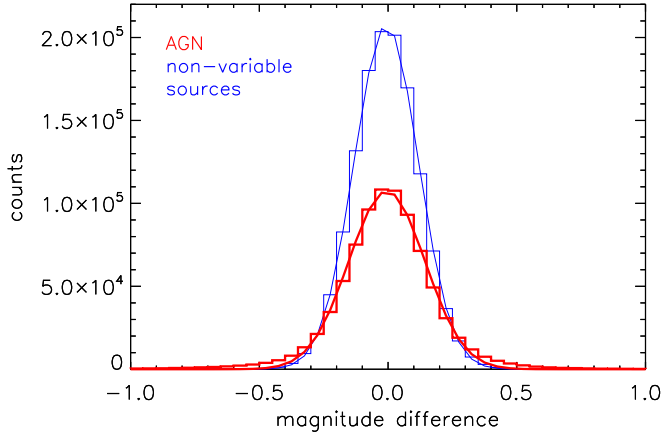


Fig. 4. Distribution of the magnitude difference for our main sample (thick red) and for the sample of non-variable sources selected to estimate the noise contribution (thin blue). Each solid line represents the Gaussian fit for the corresponding distribution, computed via non-linear least squares.

SFs of Type I AGN and optically variable AGN are obtained via a weighted least squares regression⁵ and are 0.39 ± 0.01 and 0.38 ± 0.01 , respectively, for the 23.5 mag threshold, and 0.33 ± 0.01 and 0.35 ± 0.01 , respectively, for the 22.5 mag threshold. These values are fairly consistent with some values reported in the literature, e.g., [Bauer et al. \(2009\)](#) and [Vanden Berk et al. \(2004\)](#), those being 0.3607 ± 0.0075 and 0.336 ± 0.033 , respectively.

If we focus on the SF corresponding to the subsample of Type II AGN, we see that the variations detected on short timescales are consistent with the ones detected for Type I AGN, but this does not hold for long timescales, where we clearly see that this subsample of AGN is characterized by smaller variations than Type I AGN. A line with a much shallower slope could in principle fit the whole set of points, but they are characterized in this case by much larger error bars than the previous subsamples. These “pure” subsamples of AGN, consisting only of Type I or Type II AGN, represent two extreme cases returning very different SFs. When we include AGN selected via different properties in a sample, we obtain a SF that reflects this heterogeneity: we can see this from the SFs corresponding to the remaining three subsamples of sources analyzed in this work, where the linear region is less defined, while the zone dominated by the measurement noise is larger. This is particularly evident for the largest samples analyzed, namely, the X-ray sample and the main sample, which largely overlap, as apparent from the Venn diagram in Fig. 2. We note that the fraction of known spectroscopic Type I AGN decreases as the sample includes more sources, being 60% in the MIR AGN sample, 37% in the X-ray AGN sample, and 33% in the main sample of AGN. Conversely, the corresponding fractions of known Type II AGN roughly increase, as they are 11%, 20%, and 18%, respectively. Assuming that this two trends hold for the fractions of sources lacking a spectroscopic classification in each of the three subsamples (29% for MIR AGN, 43% for X-ray AGN, and 49% for the main sample), all this would suggest that the inclusion of less variable AGN damps the average SF slope, as the inclusion of AGN seen along obscured lines of sight damps the intrinsic

variability properties on an individual basis. Therefore, all of the SF may be underestimating the intrinsic variability properties. Because of these characteristics, the subsample of Type II AGN turns out to be unsuitable for our study based on the optical variability of AGN. We note that we can still identify a clear linear region in the SF of the MIR subsample, its slope being 0.32 ± 0.02 for the 23.5 mag threshold and 0.31 ± 0.02 for the mag threshold. These values are slightly lower than the ones obtained for the subsamples of Type I AGN and optically variable AGN, supporting the thesis that the inclusion of Type II AGN in the MIR subsample is flattening the SF.

The noise contribution for each sample was estimated in two different ways. As mentioned above, the noise is what we measure on timescales too short to detect significant flux variations. Based on this, for each sample of sources we can obtain a noise estimate as the average of the first two points of the SF, corresponding to rest-frame timescales shorter than three days. The second way is more accurate and makes use of a sample of non-variable sources. For such objects, we expect the magnitude difference between two visits to be due only to stochastic fluctuations and show no dependence on timescale (i.e., white noise). For this estimate, we took advantage of a sample of 1000 sources used in [De Cicco et al. \(2021\)](#) as a labeled set of “inactive” galaxies. Based on all the COSMOS catalogs we consulted, these galaxies show no sign of nuclear activity and we can therefore assume them to be non-variable. In this case, consistent with the noise definition in the adopted SF expression (see Eq. (1)), we compute the square of the magnitude difference for this sample of sources and adopt its average value (hence corresponding to $\langle \sigma_{\text{noise}}^2 \rangle$) as our error estimate. Both methods return values that are consistent down to the third decimal digit.

3.3. The overlap issue

In Sect. 2 we mentioned that the various subsamples of AGN selected for this work are not disjointed; on the contrary, their overlap is in some cases even considerable, as is apparent from Table 2 and Fig. 2. This is a consequence of the multiple properties that AGN typically exhibit and of the different sensitivities of the different methods to AGN activity (as a function of various properties such as obscuration, orientation, radio-loudness, etc.), which allow for the selection to be carried out via more than one diagnostic or technique. If two or more subsamples of sources broadly overlap, we expect them to be characterized by similar SFs in terms of shape, amplitude, and steepness. Nonetheless, the partitions of objects that only belong to one subsample could in principle have peculiar properties – possibly explaining why they were not detected via other methods – or share the same properties of the rest of their parent subsample. It is therefore interesting to inspect these partitions; in this case 222 X-ray selected AGN, 50 MIR-selected AGN, and 19 optical variability-selected AGN.

In order to do so, we divided the X-ray, MIR, and optically-variable subsamples in two complementary subsets each: the one overlapping one or more other AGN subsamples (hereafter, overlapping subset), and the one not overlapping any other AGN subsamples (hereafter, non-overlapping subset). We then computed the SF for each pair of subsets, showing our results in Fig. 6, where we also include the SF of the whole subsample at issue to ease comparisons. The figure shows that the two complementary subsets of AGN selected via optical variability have similar SFs in terms of shape; the linear region is less steep for the non-overlapping subset, which is also characterized by slightly lower amplitude values in that region. We include estimates for

⁵ Throughout this work, when referring to weighted least square regression, we consider as weights the error bars of each point included in the fit. As mentioned earlier in this section, each of these points is the result of a binning.

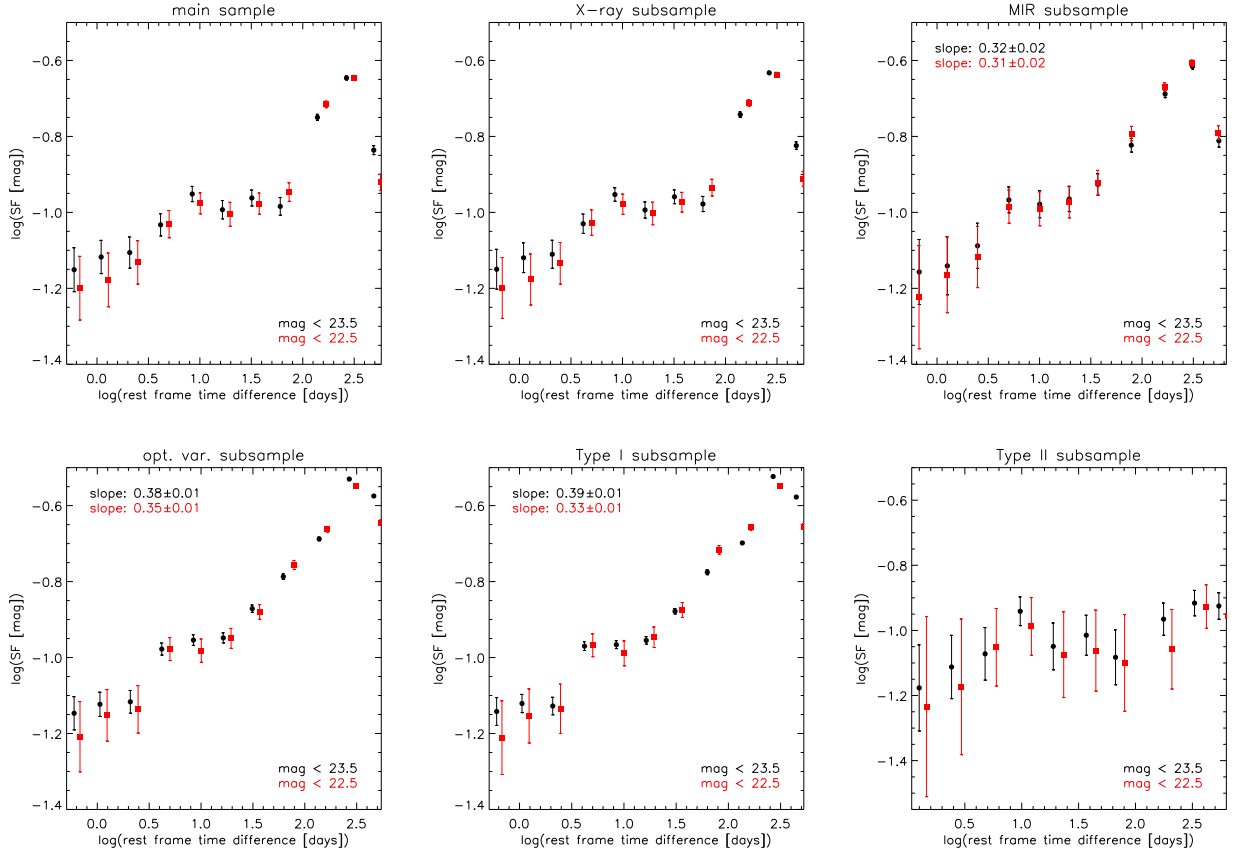


Fig. 5. SF of the six samples of AGN selected via different properties and investigated in this study: main sample (*top left*), X-ray AGN (*top center*), MIR AGN (*top right*), optically variable AGN (*bottom left*), Type I AGN (*bottom center*), and Type II AGN (*bottom right*). We tested two different magnitude thresholds: $r < 23.5$ mag (black dots) and $r < 22.5$ mag (red squares). The ranges on both axes are roughly the same to ease comparison among the various samples and thresholds. The three panels corresponding to MIR, optically variable, and Type I AGN, where a clear linear region in the SF can be identified in correspondence of the logarithmic range 1.0 – 2.6 for the baseline, include estimates for the corresponding slopes, for each of the two adopted thresholds. The slopes are the best-fit lines computed via weighted least squares regression.

the slope of each linear region. The amplitude of the global SF of this subsample is dominated by the overlapping subset, which constitutes the majority of the subsample (282 overlapping sources vs. 19 non overlapping). All this is not surprising, since the SF represents the variability of the investigated sources, and the sample at issue consists of AGN selected on the basis of their variability properties. These sources do not have a counterpart in the X-ray catalogs used in this work; and while they do have a MIR infrared counterpart, they are not classified as AGN based on that.

The results are very different for the other two subsamples of X-ray and MIR AGN: in both cases, the non-overlapping subset is characterized by a flat SF, while the SF of the overlapping subset is steeper and includes a linear region. As a consequence, the global SF of each subsample – dominated by the overlapping sources – is slightly less steep than the SF for the corresponding non-overlapping subset. Since optical variability is not a selection criterion for the X-ray and MIR AGN subsamples, these sources are not necessarily optically variable and, in particular, the two non-overlapping subsets can include Type II AGN, for which identification via optical variability tends to be difficult. A spectroscopic follow-up, at least for the brightest sources in these three non-overlapping subsets, would help shed some light on their nature; therefore, we plan to apply for observing time to pursue this goal. We note that, at present, only 16 of the sources in the three non-overlapping subsets (13 in the X-ray subsample

and three in the MIR subsample) have BH mass and Eddington ratio estimates available and used in this study.

4. Analysis of the main physical properties

One of the main goals of this work is the investigation of possible connections between the optical variability of our AGN samples and some physical properties, namely M_{BH} , λ_E , L_{bol} , redshift, and absorption.

4.1. Black hole mass, accretion rate, and bolometric luminosity dependence

We mentioned (see Sect. 2) that M_{BH} and λ_E estimates are available for 264 AGN in the main sample. In what follows, we do not take into account the subsample of Type II AGN as, based on Fig. 5 and on the arguments presented in Sect. 3.2, they do not vary significantly over the investigated timescale.

Here we analyze the possible dependence of the SF on the M_{BH} , λ_E , and L_{bol} of our AGN. For the luminosity, we limit the analysis to the sample of sources for which M_{BH} and λ_E estimates are available, in order to focus on the same sets of objects for each test. Figure 7 shows the distributions of M_{BH} , λ_E , and L_{bol} values, respectively, for each of the analyzed samples. We note that both M_{BH} and λ_E distributions cover larger ranges and,

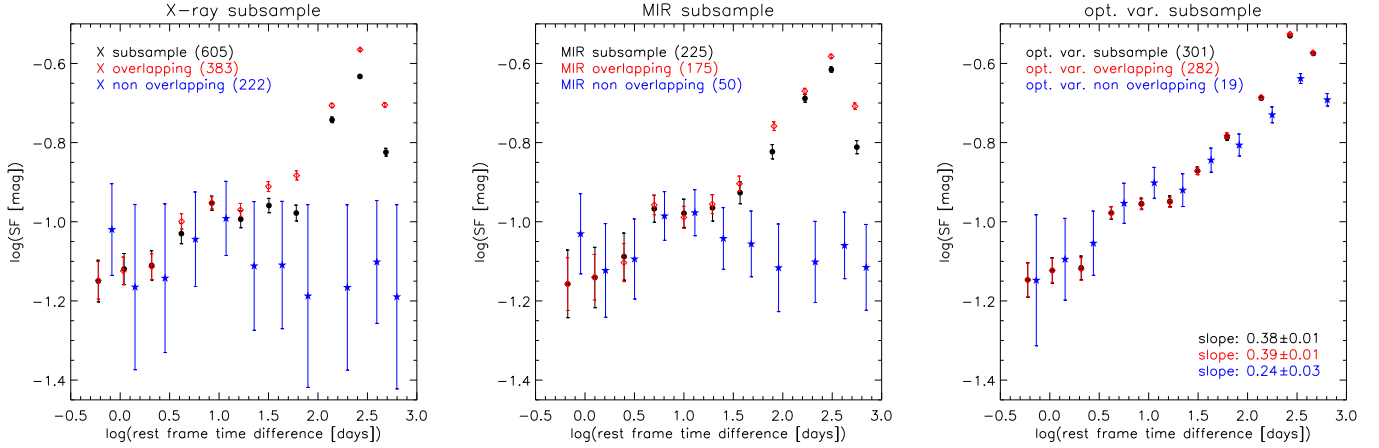


Fig. 6. SF of the three subsamples of AGN including a partition of sources not overlapping any other AGN subsamples: X-ray AGN (*left*), MIR AGN (*center*), optically variable AGN (*right*). Each subsample consists of two complementary subsets: one overlapping one or more AGN subsamples, and the other not overlapping any other AGN subsample. For each subsample we show the SF obtained from all of its sources (black dots), the SF of the overlapping subset (red diamonds), and the SF of the non-overlapping subset (blue stars). The panel corresponding to optically variable AGN includes estimates for the slope of the linear region of each SF, which can be identified in correspondence of the logarithmic range 1.0–2.6 for the baseline. The slopes are the best-fit lines computed via weighted least squares regression.

in particular, they include lower values than published in several past works such as, e.g., [Bauer et al. \(2009\)](#), [Simm et al. \(2016\)](#).

For each test, we divide each of the four analyzed samples of sources into two bins based on the median value (always indicated by a superimposed tilde) of the investigated physical property in each sample. Figure 8 presents the SFs obtained from the analysis of M_{BH} , λ_E , and L_{bol} dependence for each sample of sources (main sample, Type I AGN, MIR AGN, and optically variable AGN). For each pair of SFs reported in each panel, we also identify a region of linearity based on the SFs shown in Fig. 5. This roughly corresponds to the logarithmic values 1.0 – 2.6 for the baseline, that is to say, ≈ 10 –400 days; we show the zoomed-in linear regions in Fig. 9. For each linear region we estimated the best-fit line via weighted least squares regression. These estimates are gathered together in Table 3. The table also reports the median redshift value for each subset of sources.

Going forward, all the SFs computed will include the so-called V-correction, after [Vagnetti et al. \(2016\)](#). This allows us to take into account the dependence of variability on wavelength and, essentially, consists of a factor that is generally subtracted from the usual SF definition, producing an upwards shift of the SF, without affecting its slope. The V-correction depends on a parameter β , quantifying the variations of the spectral index in correspondence with monochromatic flux variations in the band of interest. Following [Laurenti et al. \(2020\)](#), we resort to their same estimate for this corrective factor, which is based on extrapolations from [Morganson et al. \(2014\)](#). Thus, we obtain the following corrective factor:

$$V - \text{corr} \equiv \delta \log \text{SF} \simeq -\log e \cdot \beta \log(1+z), \quad (4)$$

with $\beta \simeq 1$ (see [Vagnetti et al. 2016](#) and [Laurenti et al. 2020](#) for details).

The figures and the information in the section in Table 3 corresponding to the analysis of M_{BH} dependence show that the slopes in each pair of subsets (low vs. high M_{BH}) analyzed are consistent within the errors. In the beginning, before applying the V-correction by [Vagnetti et al. \(2016\)](#), we noticed that although the slope was roughly the same in most cases, in each panel the SF amplitude corresponding to higher M_{BH} values was systematically above the SF amplitude corresponding

to lower M_{BH} values. These results seemed to suggest that the sources with higher M_{BH} are characterized by a larger variability amplitude, not affecting the slope of the linear region of the SF. Nonetheless, from Table 3 we can see that the subsets with lower M_{BH} are systematically characterized by a lower value of the redshift. This suggested that we might be observing a correlation of the amplitude of variability with redshift, which, based on results from previous studies (e.g., [Vanden Berk et al. 2004](#); [Simm et al. 2016](#); [Sánchez et al. 2017](#)), seems to be the effect of the anticorrelation with rest-frame wavelength. The maximum difference between the variability amplitudes (i.e., measured on the y-axis, $\log \text{SF}$) for the various pairs of subsets was measured in correspondence with the maximum difference between the redshift values for a subset pair. All this led to the introduction of the above-mentioned V-correction to take into account the effect of wavelength. The panels related to M_{BH} dependence in Fig. 9 show almost no difference between the variability amplitude for each pair of subsets analyzed in each panel, so we can state that we do not detect any dependence of variability on M_{BH} .

The figures and the information in Table 3 corresponding to the analysis of λ_E dependence show that, for each pair of subsets (low vs. high λ_E), the slopes of the linear region of the SF are consistent within the errors except for the MIR subsample, where the line corresponding to lower λ_E values is steeper. The variability amplitude is systematically larger for each lower λ_E subset, supporting the thesis of an anticorrelation of the variability amplitude with λ_E , which is consistent with several findings from previous works (see Sect. 1).

The figures and the information in Table 3 corresponding to the analysis of L_{bol} dependence show conflicting results when we compare the slopes of the linear regions in each pair, while they seem to point towards an anticorrelation of the variability amplitude with L_{bol} , except for the sources in the main sample. We point out that this is the most heterogeneous sample of AGN used in this study, and thus includes a significant fraction of Type II AGN with respect to the other samples of AGN analyzed. In particular, 31% of the lower L_{bol} subset consists of Type II AGN, while the percentages of Type II AGN in each of the other subsets in each pair is one order of magnitude lower. In Sect. 3.2 we discussed how, based on Fig. 5, Type II AGN seem to be responsible for a flattening of the SF; what we observe for the

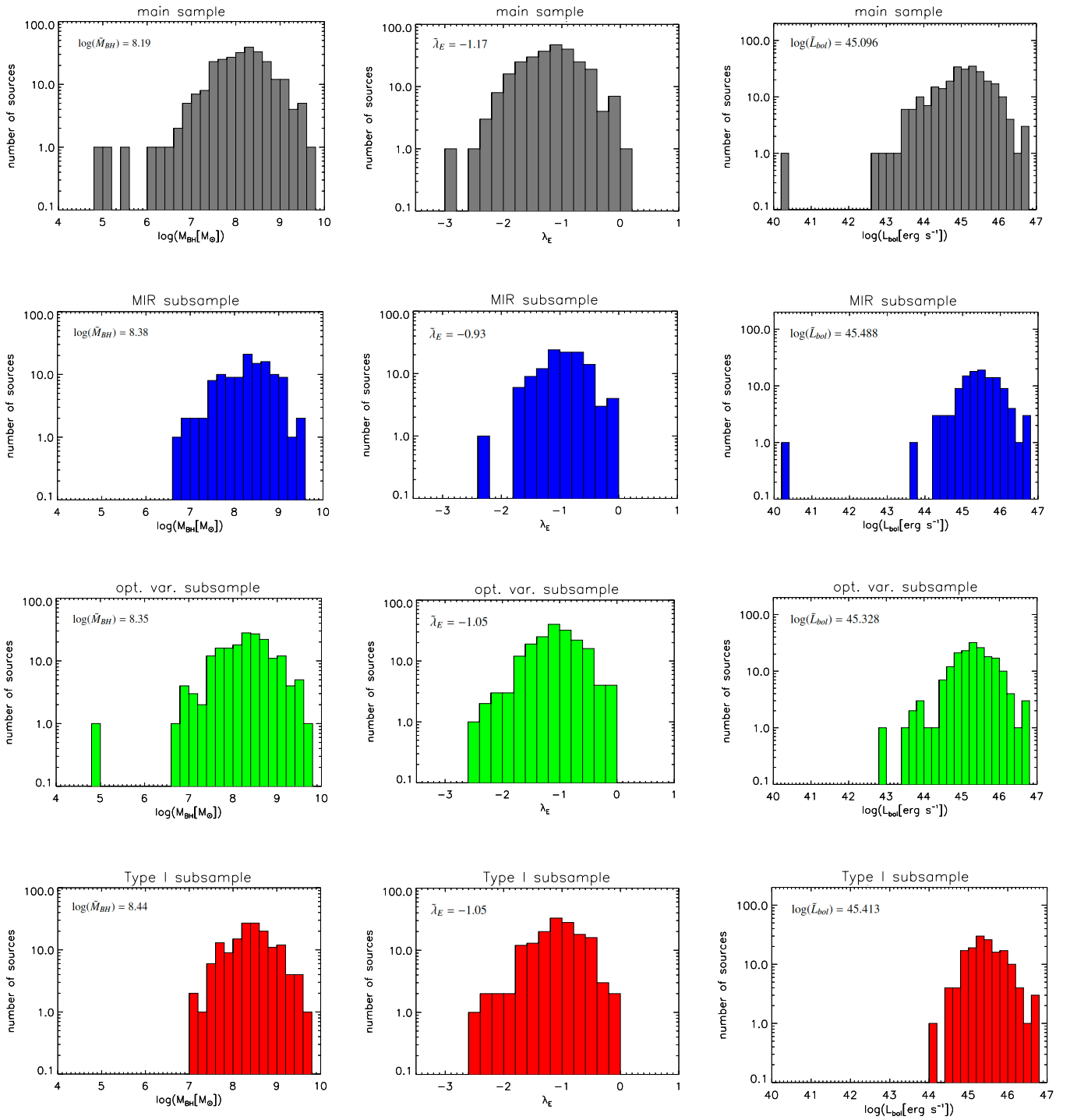


Fig. 7. Black-hole mass (left column), λ_E (middle column), and L_{bol} (right column) distributions for the main sample of 677 AGN (top line), and for the three subsamples of AGN selected to investigate dependence on these physical properties: AGN selected via their MIR properties, AGN selected via optical variability, and Type I AGN confirmed by spectroscopy (second-to-bottom lines). Each panel reports the (logarithm of the) median value of the analyzed physical quantity for the corresponding sample; M_{BH} are in solar mass units and bolometric luminosities are in erg s^{-1} . Details on the selection criteria are reported in Sect. 2.

lower L_{bol} subset corresponding to the main sample is therefore consistent with this.

We point out that, in the case of the sources with M_{BH} , λ_E , and L_{bol} estimates, the overlap among the various subsamples is almost total, leaving a few sources not overlapping with any other samples. For this reason it is not possible to analyze them as separate sets as we did in Sect. 3.3.

4.2. Redshift and obscuration dependence

In order to investigate the dependence of the SF on redshift, we consider our six original (i.e., independently of the availability of M_{BH} and λ_E estimates) samples of AGN and split each of them in two based on their redshift values. We then analyzed the slopes and amplitudes of their SFs in the linear regions of each,

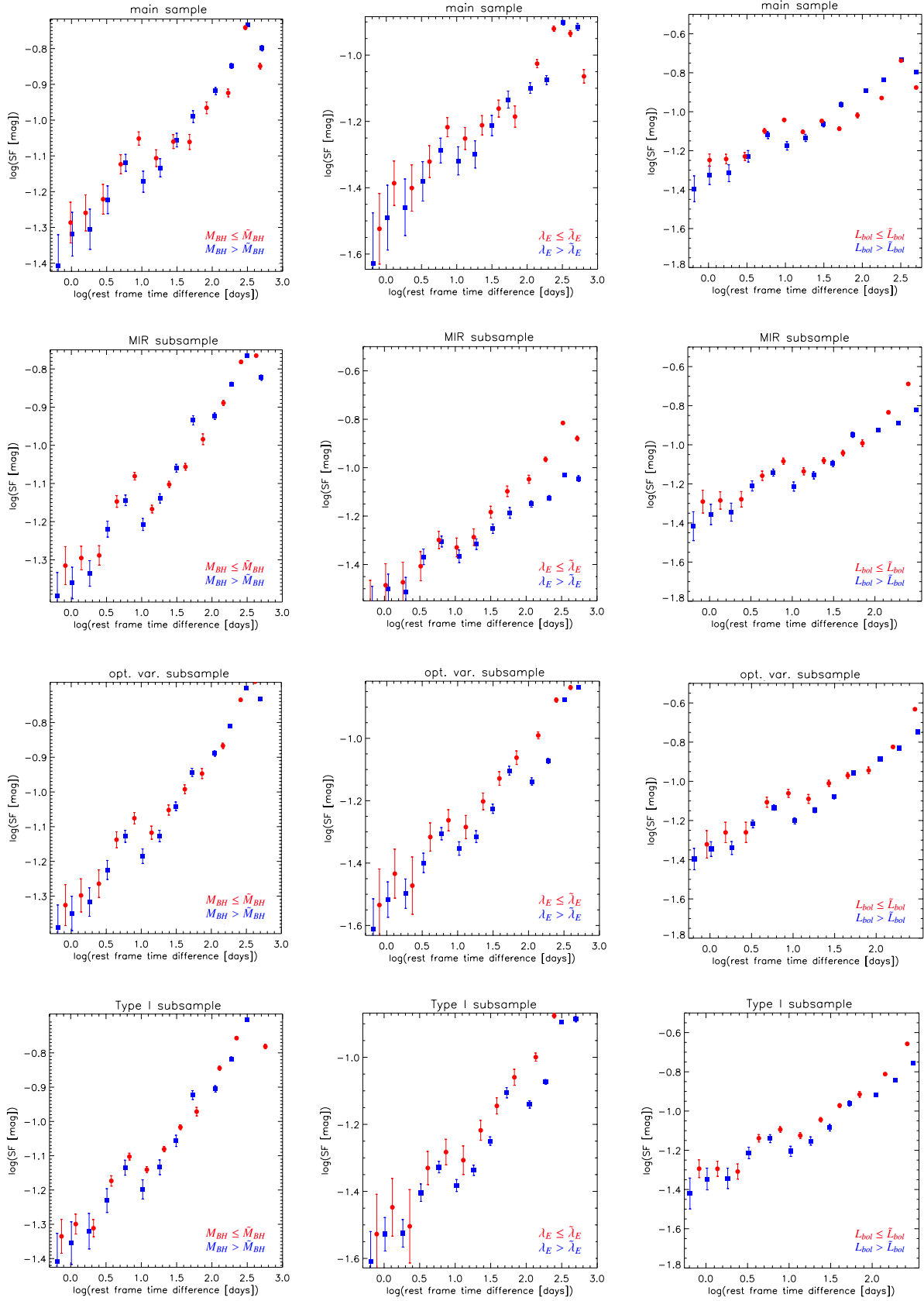


Fig. 8. SF for the four samples of AGN for which we investigate possible dependencies on M_{BH} (left column), λ_E (middle column), and L_{bol} (right column). The various panels show results for the main sample, MIR AGN, optically variable AGN, and Type I AGN (top to bottom lines). In each panel, the sample was divided into two subsets based on the (logarithm of the) median value of the physical property of interest: red dots indicate the samples of sources with lower $M_{BH}/\lambda_E/L_{bol}$ values and blue squares represent the sources with higher values.

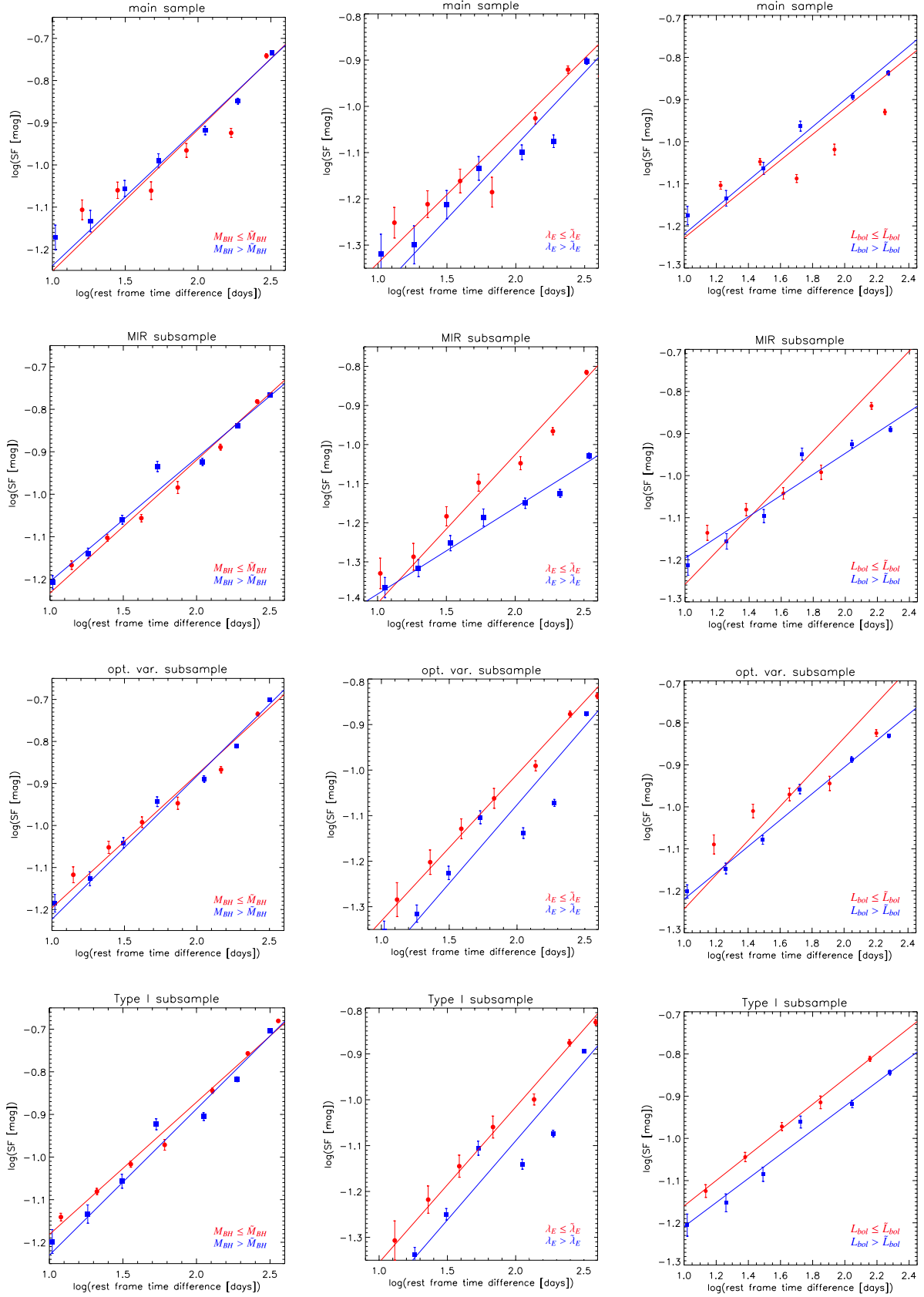


Fig. 9. Zoom on the linear region of the SF for the four samples of AGN selected to investigate possible dependencies on M_{BH} (left column), λ_E (middle column), and L_{bol} (right column). The various panels show results for the main sample, MIR AGN, optically variable AGN, and Type I AGN (top to bottom lines). As in Fig. 8, in each panel we show the two subsets of sources delimited by the (logarithm of the) median value of the physical property of interest: red dots indicate the samples of sources with lower $M_{\text{BH}}/\lambda_E/L_{\text{bol}}$ values and blue squares indicate the sources with higher values. The represented rest-frame baseline corresponds to $\approx 10 - 400$ days.

Table 3. Results from the investigation of possible dependencies of the SF on M_{BH} (top section), λ_E (middle section), and L_{bol} (bottom section) for our four samples of AGN (listed in the top row) selected via different properties and used for this part of the analysis.

	Main sample	MIR	Opt. var.	Spec. Type I
M_{BH} dependence				
$\log(M_{\text{BH}})$ range ($M_{\text{BH}}[M_{\odot}]$)	4.95 to 9.65	6.68 to 9.47	4.95 to 9.65	7.02 to 9.65
1st quartile	7.73	7.89	7.89	8.07
Median ($\log(\tilde{M}_{\text{BH}})$)	8.19	8.38	8.35	8.44
3rd quartile	8.54	8.70	8.69	8.72
$M_{\text{BH}} < \tilde{M}_{\text{BH}}$ slope	0.34 ± 0.01	0.31 ± 0.01	0.32 ± 0.01	0.31 ± 0.01
$\bar{z}$ of this subset	0.801	1.061	1.122	1.298
$M_{\text{BH}} > \tilde{M}_{\text{BH}}$ slope	0.33 ± 0.01	0.29 ± 0.01	0.34 ± 0.01	0.34 ± 0.01
$\bar{z}$ of this subset	1.450	1.560	1.531	1.597
λ_E dependence				
λ_E range	-2.85 to 0.09	-2.21 to -0.09	-2.40 to -0.14	-2.40 to -0.14
1st quartile	-1.55	-1.18	-1.35	-1.32
Median ($\tilde{\lambda}_E$)	-1.17	-0.93	-1.05	-1.05
3rd quartile	-0.85	-0.70	-0.81	-0.79
$\lambda_E < \tilde{\lambda}_E$ slope	0.30 ± 0.02	0.38 ± 0.01	0.32 ± 0.02	0.34 ± 0.02
$\bar{z}$ of this subset	0.938	1.195	1.306	1.345
$\lambda_E > \tilde{\lambda}_E$ slope	0.32 ± 0.02	0.22 ± 0.01	0.34 ± 0.01	0.34 ± 0.01
$\bar{z}$ of this subset	1.258	1.531	1.444	1.550
L_{bol} dependence				
$\log(\tilde{L}_{\text{bol}})$ range L_{bol} in erg s^{-1}	40.276 to 46.676	45.299 to 46.676	42.865 to 46.676	44.023 to 46.676
1st quartile	44.634	45.149	44.995	45.149
Median ($\log(\tilde{L}_{\text{bol}})$)	45.096	45.488	45.328	45.413
3rd quartile	45.513	45.815	45.665	45.76s
$L_{\text{bol}} < \tilde{L}_{\text{bol}}$ slope	0.31 ± 0.01	0.40 ± 0.01	0.41 ± 0.01	0.31 ± 0.01
$\bar{z}$ of this subset	0.762	0.970	1.033	1.279
$L_{\text{bol}} > \tilde{L}_{\text{bol}}$ slope	0.32 ± 0.01	0.25 ± 0.01	0.31 ± 0.01	0.29 ± 0.01
$\bar{z}$ of this subset	1.527	1.729	1.646	1.681

Notes. For each physical quantity we report the (logarithm of the) corresponding range boundaries, as well as the 1st quartile, median, and 3rd quartile values. Each sample is divided in two subsets on the basis of the (logarithm of the) median value of the physical property of interest in that sample. Each section reports, for each subset, the median redshift value and the estimated slope value for the best-fit line approximating the corresponding set of points in the linear region of the SF, with its error. We show the linear plots for each of the investigated samples of sources in Fig. 9, comparing results for each of the performed tests.

following our steps listed in Sect. 4.1. We report the slopes of the linear regions in each pair, delimited as usual, in the top section of Table 4, together with the redshift values (which are the same reported in Table 2), and we show the SFs obtained for each pair of subsets in Fig. 10, limited to their linear regions. We notice that the slopes in a pair are roughly consistent within the errors in the case of Type I AGN and optically variable AGN. Conversely, for the main sample and the X-ray subsample, which largely overlap, we obtain the largest difference between the corresponding slopes. Mid-infrared AGN are in between, with a moderate difference between the two slopes, while Type II AGN, as usual, behave differently than all the other samples, and are characterized by almost flat SFs both for lower and higher redshift subsets, and also by larger errors on these estimates. We also notice that, apart from Type II AGN, the slopes corresponding to higher redshifts are all roughly consistent within the errors, while this does not hold for the slopes corresponding to lower redshifts. Since we know that Type II AGN are generally responsible for the flattening of the SF, we investigated the fraction of known

Type I and Type II AGN in each subset, for each pair. We find that, in general, Type I AGN dominate the various subsets corresponding to higher redshifts. They are also dominant in the subsets at lower redshifts corresponding to optically variable AGN (and of course, by construction they constitute the total subset for the subsample of Type I AGN). Conversely, in the main sample and in the X-ray subsample Type II AGN are the ones dominating the lower redshift subsets, namely, the ones characterized by a flattening with respect to the corresponding higher redshift subsets. The situation is intermediate for the MIR AGN subsets, which include a significant fraction of Type II AGN but are still dominated by Type I AGN. All this suggests that redshift itself does not affect the slope of the SF, but this is rather a consequence of the fraction of unobscured/obscured AGN in each sample. In addition, we should take into account that the basic classification capability likely changes with redshift; for instance, because of the different rest-frame lines available in optical spectra, the different rest-frame X-ray bands sampled, and so on.

Table 4. Results from the investigation of possible dependencies of the SF on redshift (top section) and obscuration (bottom section) for the main sample as well as the other five subsamples of AGN selected via different properties and analyzed in this study (listed in the top row).

	Main sample	X-ray	MIR	Opt. var.	Spec. Type I	Spec. Type II
Redshift dependence						
z range	0.121 to 3.715	0.121 to 3.715	0.350 to 3.715	0.346 to 3.715	0.346 to 3.715	0.121 to 1.447
1st quartile	0.530	0.620	0.704	0.961	1.171	0.520
Median ($\bar{z}$)	0.876	0.893	1.214	1.441	1.653	0.696
3rd quartile	1.402	1.450	2.021	2.032	2.152	0.890
$z < \bar{z}$ slope	0.02 ± 0.03	0.03 ± 0.02	0.20 ± 0.03	0.34 ± 0.01	0.38 ± 0.01	0.01 ± 0.04
$z > \bar{z}$ slope	0.34 ± 0.03	0.35 ± 0.02	0.33 ± 0.02	0.37 ± 0.01	0.37 ± 0.02	0.10 ± 0.03
Absorption dependence						
N_H range ($\times 10^{22} \text{ cm}^{-2}$)	0.0 to 25.9	0.0 to 25.9	0.0 to 25.9	0.0 to 25.9	0.0 to 25.9	0.0 to 13.0
$N_H < 10^{22} \text{ cm}^{-2}$ slope	0.38 ± 0.02	0.38 ± 0.02	0.34 ± 0.01	0.39 ± 0.01	0.40 ± 0.01	0.10 ± 0.03
$N_H > 10^{22} \text{ cm}^{-2}$ slope	0.15 ± 0.02	0.15 ± 0.02	0.25 ± 0.01	0.29 ± 0.01	0.28 ± 0.01	0.04 ± 0.07

Notes. Absorption is quantified by the hydrogen column density N_H . For each quantity we report the corresponding range boundaries. In the redshift-related section of the table (top) we report the 1st quartile, median, and 3rd quartile values of redshift. Each sample is divided in two subsets on the basis of the median redshift of the sample (top section), or on the basis of the value $N_H = 10^{22} \text{ cm}^{-2}$ (bottom section). For each subset we report the estimated slope value for the best-fit line, obtained as usual via a weighted least squares regression, approximating the corresponding set of points in the linear region of the SF, with its error.

The analysis performed so far has shown how the AGN variability properties are different for Type I and Type II AGN. To test further the dependence on AGN obscuration, we investigated how our SFs change with absorption – which in the X-rays is typically quantified by the hydrogen column density N_H – for the six samples of AGN. An estimate of this quantity is available for 71% of the AGN in our main sample from the *Chandra* COSMOS Legacy Survey Multiwavelength Catalog (Marchesi et al. 2016). We chose the value of 10^{22} cm^{-2} to split each sample in two as this is usually assumed as the limit between unabsorbed ($N_H < 10^{22} \text{ cm}^{-2}$) and absorbed ($N_H > 10^{22} \text{ cm}^{-2}$) sources. Once again, we estimated the slopes of the linear regions of the SFs of unabsorbed and absorbed sources in each sample. We report these estimates in the bottom section of Table 4, and we show the SFs obtained for each pair of subsets in Fig. 10, limited to their linear regions. For each sample, the slope of the linear region corresponding to lower N_H is always steeper than the one obtained for higher N_H . As usual, optical Type II AGN exhibit flatter SFs for both subsets than the rest of the analyzed samples. If we compare the lower N_H slopes of the various samples with each other, we notice that they all are roughly consistent within the errors. When we do the same for higher N_H slopes, we notice that the values obtained for Type I AGN, optically variable AGN, and MIR AGN are roughly consistent within the errors, and are higher than the values obtained for the main sample and the X-ray subsample, which are in turn consistent with each other. From the plots in the figure we can see how the angle formed by the two lines in each pair decreases as we move from the main sample towards Type I AGN (from top left to bottom central panels), that is, as the contribution of Type II AGN becomes less relevant. All these details suggest that absorption is responsible for the flattening observed in the SF.

5. Discussion and conclusions

In this work we presented an analysis of the SF of 677 AGN in the VST-COSMOS area, defining a sample of AGN as a result of different diagnostics based on: spectroscopy, X-ray properties, MIR properties, and selection via optical variability. We analyzed the SF of the main sample of sources as well as the

SF of the various subsamples defined by the various diagnostics. We investigated the properties of these samples at a depth that is approximately two orders of magnitude deeper than most of the similar analyses from the literature. This study therefore represents, to our knowledge, one of the few investigating the ensemble variability of fainter AGN than most datasets to date, thus exploring new ranges of luminosity and, therefore, of M_{BH} and λ_E . On the other hand, this work suffers from two main limitations: the very irregular sampling and the small size of the sample of AGN with available estimates of the physical quantities of interest for this analysis. Indeed, we also examined the behavior of the SF in relation to two major properties of our AGN, namely, the mass of the central SMBH and the accretion rate, the latter being quantified by λ_E (both with estimates available for 264 sources), and also in relation to the AGN L_{bol} , redshift, and obscuration.

Our catalog of 677 AGN is available in the electronic version of this paper as Table 6 (available at the CDS). In what follows we summarize our main findings.

The shape of the SF is affected by the sample of AGN used to build it: in Fig. 5 we report the SFs obtained for the six samples of AGN used in this work, and it is apparent that when the sample is dominated by Type I AGN, the shape of the SF shows a “clean” region with a fairly clear linear slope. This is consistent with the fact that Type I AGN are typically easier to select via optical variability. Conversely, the presence of a significant fraction of Type II AGN in the sample heavily affects the shape of the SF. Based on the comparison of the SFs shown in Fig. 5 for the various samples of AGN used in this study, this effect is more relevant at timescales in the range ≈ 10 –65 days, where the SF starts to rise for Type I AGN-dominated samples, while it is still rather flat otherwise. It is well known that Type II AGN are harder to detect via optical variability; clearly there will be a reprocessed component on larger physical scales, and hence it will also “reprocess” or damp the timescales. If we assume that the origin of optical variability is the same for Type I and Type II AGN, then the main difference between the two types lies in the fact that for Type I AGN we detect the disk variability directly; whereas, in the case of Type II, we do not do so since it is absorbed, but we do detect a scattered component of

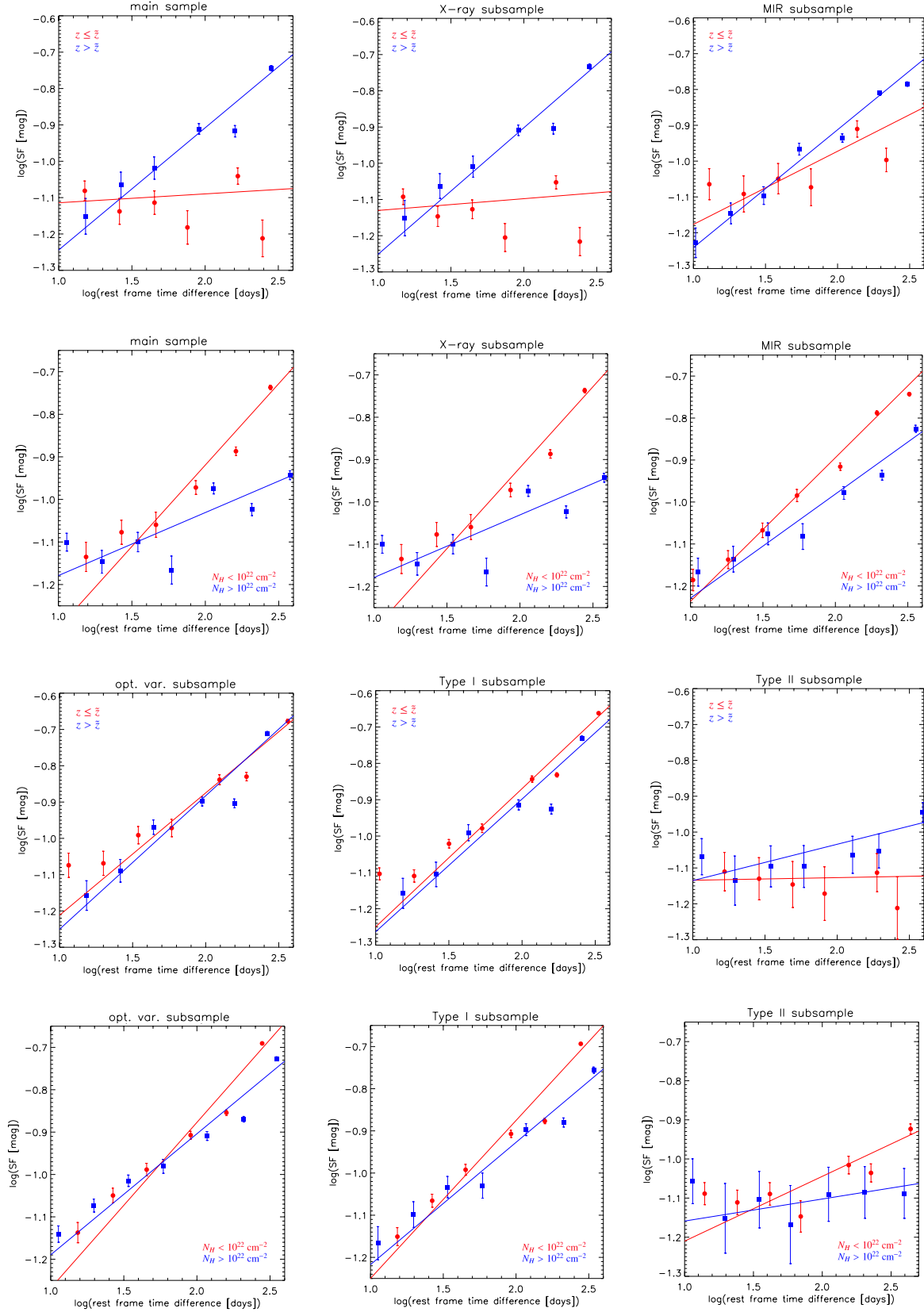


Fig. 10. Analysis of the dependence of variability on redshift and absorption (quantified by the hydrogen column density N_H), focusing on the linear regions of the SF for the six samples of AGN selected for this study: the two top lines show panels for the main sample (*left*), X-ray AGN (*center*), and MIR AGN (*right*); the two bottom lines show panels for optically variable AGN (*left*), Type I AGN (*center*), and Type II AGN (*right*). Redshift (odd lines) and absorption (even lines) panels are coupled for each sample of sources, in order to ease the comparison of the results. Red dots indicate the samples of sources with lower redshift or absorption, depending on what the panel at issue represents; blue squares indicate the sources with higher redshift or absorption values. As usual, the represented rest-frame baseline corresponds to $\approx 10 - 400$ days.

Table 5. Comparison of our results with some findings from the literature (see Sect. 1), where several methods to analyze the dependence of AGN variability on M_{BH} , λ_E , and L_{bol} were used.

Publication	Method	Amplitude			Timescale		
		M_{BH}	λ_E	L_{bol}	M_{BH}	λ_E	L_{bol}
This work	SF	N	A	A	–	–	–
Wilhite et al. (2008)	SF	C	A	A	–	–	–
MacLeod et al. (2010)	DRW	C	AA	AA	C	–	N
Kelly et al. (2013)	PSD	A	A	A	–	–	–
Simm et al. (2016)	EV, PSD	N	A	A	N	N	N
Caplar et al. (2017)	SF	C	–	AA	–	–	C
Sánchez-Sáez et al. (2018)	SF	N	A	–	–	–	–
Laurenti et al. (2020)	SF	C	A	AA	–	–	–

Notes. The table reports (left to right): reference paper; method applied for the analysis; dependence of amplitude/timescale on M_{BH} , λ_E , and L_{bol} . C = correlation (CC = strong); A = anticorrelation (AA = strong); N = no relation found; – = dependence not investigated.

this variability coming from the narrow line region. Since this is supposed to be located at larger distances from the central BH (typically 10–1000 pc, vs. the ≈ 1 pc radius of the accretion disk), this suggests that we need longer timescales in order to be able to detect larger variations for Type II AGN. Indeed, our previous studies (De Cicco et al. 2015, 2019) already showed how the fraction of Type II AGN detected via optical variability increases from 6% to 18% when the baseline is extended from five months to 3.3 yr; hence, we expect some improvement with a, for instance, 10 yr baseline (i.e., the full baseline for the LSST main survey, but also the baseline that we will be able to cover with the new observing season of VST-COSMOS observations; for more details, see below). While the shape of the SF changes with the type of sources, it does not seem to change with the depth of the sample, as can be seen from Fig. 5, where two distinct magnitude thresholds ($r \leq 23.5$ and $r \leq 22.5$) are tested.

In cases where a region of linearity can be identified, the slope of this region is fairly consistent with results from several past studies dedicated to the AGN SF. It is worth mentioning that the slope of the SF is highly affected by the way the noise contribution in the SF definition is estimated (e.g., Kozłowski 2016).

The study of possible connections between the shape of the SF and the mass of the central SMBHs in our sample of AGN shows no significant relations for what concerns both the amplitude and the slope of the SF (see Figs. 8 and 9, and Table 3). The analysis of connections with λ_E , on the other side, suggest an anticorrelation with variability amplitude, based on Figs. 8 and 9, and Table 3. This is consistent with most results from past investigations on the topic and it is important to stress once again that, with our study, we are extending this result to fainter sources. The investigation of possible relations with the L_{bol} also points towards an anticorrelation of the variability amplitude with luminosity, which is consistent once again with previous findings. Following Suberlak et al. (2021), in Table 5 we summarize the main results of the works mentioned in Sect. 1 analyzing the variability dependence on M_{BH} , λ_E , and L_{bol} , and we compare them with the results obtained in the course of this work. Based on this comparison, in what concerns the variability amplitude, the most controversial results are obtained for the dependence on M_{BH} , while all the works investigating a dependence on λ_E or on L_{bol} are consistent in showing an anticorrelation, which in some cases can be very strong. The dependence of the variability timescale on these same quantities was investigated in only

a few works, and so we cannot state anything significant about that.

We also analyzed the dependence of the SF on redshift and found that it is affected by the type of AGN included in the various samples: we observe a strong correlation with redshift for our main sample of sources as well as the X-ray subsample, but in both cases the lower redshift subsets turn out to be dominated by Type II AGN, which seem to be responsible for the corresponding flattening in the SF. Indeed, this correlation is weaker for MIR AGN, where Type I AGN start to outnumber Type II AGN in both lower and higher redshift subsets, while it is not observed at all in the subsample of Type I AGN; nor is it observed in the subsample of optically variable AGN, where Type I AGN constitute 75% of the total. In addition we investigated the dependence on absorption, quantified by the hydrogen column density N_{H} , and our tests show that the slope of the linear region of the SF is shallower for absorbed sources.

As the other works in our series making use of VST data, this study can provide some forecasts for the LSST. The survey will include ultra-deep investigation of regions known as deep-drilling fields (DDFs; they include the COSMOS area), widely surveyed in the past decades and therefore with valuable multi-wavelength and spectroscopic coverage available, which makes them ideal for AGN science as well as training sets in the context of the main survey. The surveyed area will cover 9.6 sq. deg. per DDF. Based on the results from De Cicco et al. (2021), we expect to be able to identify at least ≈ 3500 AGN per DDF via optical variability combined with the colors that will be available for the LSST. We point out that this estimate depends on the characteristics of our VST-COSMOS survey; hence we should take into account that, while the single-visit depth is roughly the same for the two surveys, for the LSST we will have the chance to stack close visits together, as we expect observations with about a two-night cadence. This will return deeper images and, therefore, will allow us to identify fainter AGN and increase their number per square degree. We should also consider that the number of confirmed AGN per sq. deg. in the DDFs is expected to roughly double as the multiwavelength coverage returns other samples of AGN selected via other diagnostics. To give a general idea of what we can expect: the estimated sky density limited to BLAGN and obtained from the XMM-SERVS survey of two other DDFs, namely W-CDF-S and ELAIS-S1, is ≈ 200 per sq. deg., which translates into ≈ 1900 BLAGN per DDF (Ni et al. 2021).

Such larger samples of AGN will allow us to improve our analysis. Indeed, our main sample of 677 AGN reduced to 264 sources with available M_{BH} and λ_E estimates. This has affected the whole analysis of the correlations of variability with AGN physical properties, as further selections in the space of the parameters under investigation led to scant samples of a few tens of sources, unsuitable for a statistically significant analysis. The LSST will therefore provide us with a suitable dataset for the exploration of multidimensional parameter dependence for the SF of AGN.

We mentioned that our light curves are characterized by two large gaps of one year and seven months plus eight months, as shown in Table 1. From the table, it is apparent that mean and median observed baseline values, computed for the sources in the main sample for each season, are very close to the maximum observed baseline (when not exactly coincident with it) for the corresponding season. Similarly, the mean and median number of visits for each season are very close to the total number of visits (when not exactly coincident with it) for the corresponding season. This shows how, for individual seasons, our dataset can take advantage of a dense sampling, which plays a key role in the context of AGN detection efficiency, as shown in De Cicco et al. (2019). Nevertheless, the two gaps affect the shape of our SF with inadequate sampling in correspondence with some timescales. Sparse and/or irregular sampling is a very common issue in SF analysis (e.g., de Vries et al. 2003; Peters et al. 2015; Simm et al. 2016; Sartori et al. 2019). Indeed, there are works from the literature where the cadence is low but the sampling is regular: as an example, Hawkins (2002) uses quasar light curves from a long-term monitoring program with 24 yearly observations per source to investigate the origin of the emission mechanism in AGN. Nonetheless, one of the advantages in the use of the SF is its relative insensitivity to irregular sampling when sources are considered as an ensemble rather than individually (e.g., Hawkins 2007; Kozłowski 2016; Sartori et al. 2019). As mentioned in Sect. 3.2, Bauer et al. (2009) analyze the effect of irregular sampling by means of simulations, and conclude that the turnover that is observed in the light curve is an effect of sparse sampling at longer timescales, and not a real feature in the SF of AGN. Emmanoulopoulos et al. (2010) also resort to simulations in order to assess whether and to what extent the SF is robust against the presence of gaps in the light curves. They simulated a single light curve and then investigated the effect of three different gaps in the data, representing three different situations: almost periodic data gaps, dense and sparse sampling, and purely sparsely sampled data, corresponding to 57%, 83%, and 92% of the data being removed from a single simulated light curve that is 2000 time units long, respectively. They then used bootstrapping to extract 1000 light curves from each of the obtained light curves with gaps, then compared the results obtained with and without gaps. While they find that the presence of gaps is responsible for the presence of wiggles and bends in the SF, from the right panels of their Fig. 12 we can infer that these wiggles and bends do not alter the slope of the SF obtained from the light curve with no gaps. In this work we are not investigating the turnover as our baseline is not long enough (Sect. 3.2); our analysis is instead focused on the linear region of the SF (and the possible dependence on physical quantities of interest), where the irregularity of the sampling does not constitute a major issue.

The gaps in our baseline are one of the reasons why we need long observing seasons and high-cadence observations for this kind of studies. For the LSST light curves we expect more regular sampling and, hence, we expect the sampling issue to be

a minor one. Indeed, for the DDFs, a 2-day 2-filter cadence is planned (Brandt et al. 2018). In addition, the 10 yr duration of the survey will allow us to probe longer timescales than our 3.3 yr (and to extend the redshift coverage), which has been done in some works from the literature, but not at these depths. In this way we can investigate the above-mentioned characteristic timescale τ where a turnover in the SF is generally observed, and explore its possible connections with M_{BH} (see, e.g., Burke et al. 2021) and other AGN physical properties.

Some of the above challenges can be overcome even without the Vera C. Rubin Observatory being operational. Indeed, a fourth season consisting of 13 observations of the COSMOS field with the VST was recently completed (ESO P108, Dec. 2021 – Mar. 2022) and the data are currently being reduced. This can be combined with ~ 10 yr of archival DECam monitoring of the central DDFs (which probably lacks the high cadence, but grants a long baseline and reasonable depth). This means that, while we wait for the *Vera C. Rubin* Observatory to be operational, we can take advantage of the data from this new observing season to extend our baseline from 3.3 yr to ≈ 10 yr, that is, roughly the same baseline that we will have when the LSST will be completed. Including data from this new observing season means that the VST-COSMOS baseline will suffer from an additional, longer gap between its observing seasons; therefore, we plan to investigate the effect of the presence of these gaps thoroughly in a forthcoming paper in this series. Nevertheless, the VST-COSMOS dataset still remains one of the few that, to date, can benefit from a high observing cadence in each of the observing seasons, along with a considerable depth. This will allow us to lead additional AGN studies focused on their optical variability, to pursue the above-mentioned goals in the context of SF analysis and – last but not least – it will favor the identification of fainter AGN, for which longer baselines are needed, as shown in, for instance, De Cicco et al. (2019).

Acknowledgements. We thank D. J. Rosario and B. Trakhtenbrot for providing their catalog of black hole masses. We acknowledge support from: ANID grants FONDECYT Postdoctorado NO 3200222 (D.D.) and 3200250 (P.S.S.); PON R&I 2021, CUP E65F21002880003 (D.D.); ANID – Millennium Science Initiative Program ICN12_009 (F.E.B.); CATA-Basal AFB-170002 and FB210003 (F.E.B.); FONDECYT Regular 1190818 and 1200495 (F.E.B.); NSF grant AST-2106990 and V.M. Willaman Endowment (W.N.B.); South Africa's DSI-NRF Grant Number 113121 & 121291 (M.V.).

References

- Aretxaga, I., & Terlevich, R. 1994, *MNRAS*, **269**, 462
- Bauer, A., Baltay, C., Coppi, P., et al. 2009, *ApJ*, **696**, 1241
- Brandt, W. N., Ni, Q., Yang, G., et al. 2018, ArXiv e-prints [arXiv:1811.06542]
- Breiman, L. 2001, *Mach. Learn.*, **45**, 5
- Brusa, M., Civano, F., Comastri, A., et al. 2010, *ApJ*, **716**, 348
- Burke, C. J., Shen, Y., Blaes, O., et al. 2021, *Science*, **373**, 789
- Capaccioli, M., & Schipani, P. 2011, *The Messenger*, **146**, 2
- Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2017, *ApJ*, **834**, 111
- Collier, S., & Peterson, B. M. 2001, *ApJ*, **555**, 775
- De Cicco, D., Paolillo, M., Covone, G., et al. 2015, *A&A*, **574**, A112
- De Cicco, D., Paolillo, M., Falocco, S., et al. 2019, *A&A*, **627**, A33
- De Cicco, D., Bauer, F. E., Paolillo, M., et al. 2021, *A&A*, **645**, A103
- de Vries, W. H., Becker, R. H., & White, R. L. 2003, *AJ*, **126**, 1217
- de Vries, W. H., Becker, R. H., White, R. L., & Loomis, C. 2005, *AJ*, **129**, 615
- di Clemente, A., Giallongo, E., Natali, G., Trevese, D., & Vagnetti, F. 1996, *ApJ*, **463**, 466
- Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, *ApJ*, **748**, 142
- Emmanoulopoulos, D., McHardy, I. M., & Uttley, P. 2010, *MNRAS*, **404**, 931
- Falocco, S., Paolillo, M., Covone, G., et al. 2015, *A&A*, **579**, A115
- González-Martín, O., & Vaughan, S. 2012, *A&A*, **544**, A80
- Graham, M. J., Djorgovski, S. G., Drake, A. J., et al. 2014, *MNRAS*, **439**, 703
- Hawkins, M. R. S. 2002, *MNRAS*, **329**, 76
- Hawkins, M. R. S. 2007, *A&A*, **462**, 581

- Hook, I. M., McMahon, R. G., Boyle, B. J., & Irwin, M. J. 1994, *MNRAS*, **268**, 305
- Ivezić, Ž., & MacLeod, C. 2014, in *Multiwavelength AGN Surveys and Studies*, eds. A. M. Mickaelian, & D. B. Sanders, 304, 395
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Kawaguchi, T., Mineshige, S., Umemura, M., & Turner, E. L. 1998, *ApJ*, **504**, 671
- Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, *ApJ*, **698**, 895
- Kelly, B. C., Treu, T., Malkan, M., Pancoast, A., & Woo, J.-H. 2013, *ApJ*, **779**, 187
- Kimura, Y., Yamada, T., Kokubo, M., et al. 2020, *ApJ*, **894**, 24
- Kozłowski, S. 2016, *ApJ*, **826**, 118
- Kozłowski, S. 2017, *A&A*, **597**, A128
- Laurenti, M., Vagnetti, F., Middei, R., & Paolillo, M. 2020, *MNRAS*, **499**, 6053
- LSST Science Collaboration (Abell, P. A., et al.) 2009, ArXiv e-prints [arXiv:0912.0201]
- Lusso, E., Comastri, A., Simmons, B. D., et al. 2012, *MNRAS*, **425**, 623
- Maccacaro, T., Gioia, I. M., Wolter, A., Zamorani, G., & Stocke, J. T. 1988, *ApJ*, **326**, 680
- MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., et al. 2010, *ApJ*, **721**, 1014
- MacLeod, C. L., Ivezić, Ž., Sesar, B., et al. 2012, *ApJ*, **753**, 106
- Marchesi, S., Civano, F., Elvis, M., et al. 2016, *ApJ*, **817**, 34
- McHardy, I. M., Koerding, E., Knigge, C., Uttley, P., & Fender, R. P. 2006, *Nature*, **444**, 730
- Morganson, E., Burgett, W. S., Chambers, K. C., et al. 2014, *ApJ*, **784**, 92
- Ni, Q., Brandt, W. N., Chen, C.-T., et al. 2021, *ApJS*, **256**, 21
- Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, *A&ARv*, **25**, 2
- Paolillo, M., Papadakis, I., Brandt, W. N., et al. 2017, *MNRAS*, **471**, 4398
- Peters, C. M., Richards, G. T., Myers, A. D., et al. 2015, *ApJ*, **811**, 95
- Poulain, M., Paolillo, M., De Cicco, D., et al. 2020, *A&A*, **634**, A50
- Rakshit, S., Stalin, C. S., & Kotilainen, J. 2020, *ApJS*, **249**, 17
- Rengstorf, A. W., Brunner, R. J., & Wilhite, B. C. 2006, *AJ*, **131**, 1923
- Rosario, D. J., Trakhtenbrot, B., Lutz, D., et al. 2013, *A&A*, **560**, A72
- Sánchez, P., Lira, P., Cartier, R., et al. 2017, *ApJ*, **849**, 110
- Sánchez-Sáez, P., Lira, P., Mejía-Restrepo, J., et al. 2018, *ApJ*, **864**, 87
- Sartori, L. F., Trakhtenbrot, B., Schawinski, K., et al. 2019, *ApJ*, **883**, 139
- Schulze, A., Silverman, J. D., Kashino, D., et al. 2018, *ApJS*, **239**, 22
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, **172**, 1
- Serafinelli, R., Vagnetti, F., Chiaraluce, E., & Middei, R. 2017, *Front. Astron. Space Sci.*, **4**, 21
- Shakura, N. I., & Sunyaev, R. A. 1976, *MNRAS*, **175**, 613
- Simm, T., Salvato, M., Saglia, R., et al. 2016, *A&A*, **585**, A129
- Simonetti, J. H., Cordes, J. M., & Heeschen, D. S. 1985, *ApJ*, **296**, 46
- Suberlak, K. L., Ivezić, Ž., & MacLeod, C. 2021, *ApJ*, **907**, 96
- Vagnetti, F., Middei, R., Antonucci, M., Paolillo, M., & Serafinelli, R. 2016, *A&A*, **593**, A55
- Vanden Berk, D. E., Wilhite, B. C., Kron, R. G., et al. 2004, *ApJ*, **601**, 692
- Wilhite, B. C., Brunner, R. J., Grier, C. J., Schneider, D. P., & vanden Berk, D. E. 2008, *MNRAS*, **383**, 1232
- York, D. G., Adelman, J., & Anderson, Jr., J. E. 2000, *AJ*, **120**, 1579
- Zu, Y., Kochanek, C. S., Kozłowski, S., & Udalski, A. 2013, *ApJ*, **765**, 106