A Forward Modeling Approach to AGN Variability–Method Description and Early Applications

Lia F. Sartori[1,2] , Benny Trakhtenbrot[3] , Kevin Schawinski[1,4] , Neven Caplar[5] , Ezequiel Treister[2] , and Ce Zhang[6]

[1] Institute for Particle Physics and Astrophysics, ETH Zürich, Wolfgang-Pauli-Str. 27, CH-8093 Zürich, Switzerland
[2] Instituto de Astrofísica, Facultad de Física, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile
[3] School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel
[4] Modulos AG, Technoparkstrasse 1, CH-8005 Zürich, Switzerland
[5] Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA; ncaplar@princeton.edu
[6] Systems Group, ETH Zürich, Universitätstrasse 6, CH-8006 Zürich, Switzerland

Abstract

We present a numerical framework for the variability of active galactic nuclei (AGNs), which links the variability of AGNs over a broad range of timescales and luminosities to the observed properties of the AGN population as a whole, and particularly the Eddington ratio distribution function. We implemented our framework on GPU architecture, relying on previously published time-series-generating algorithms. After extensive tests that characterize several intrinsic and numerical aspects of the simulations, we describe some applications used for current and future time-domain surveys and for the study of extremely variable sources (e.g., "changing-look" or flaring AGNs). Specifically, we define a simulation setup that reproduces the AGN variability observed in the (intermediate) Palomar Transient Factory survey and use it to forward model longer light curves of the kind that may be observed within the Large Synoptic Survey Telescope (LSST) main survey. Thanks to our efficient implementations, these simulations are able to cover, for example, over 1 Myr with a roughly weekly cadence. We envision that this framework will become highly valuable to prepare for, and best exploit, data from upcoming time-domain surveys, such as, for example, LSST.

1. Introduction

Variable emission is a ubiquitous property of active galactic nuclei (AGNs). It can be observed or inferred at essentially all timescales, from hours to billions of years (e.g., McHardy et al. 2004; Novak et al. 2011; MacLeod et al. 2012; Goyal et al. 2013; Schawinski et al. 2015; Sartori et al. 2018a; Smith et al. 2018), and across the entire electromagnetic spectrum (e.g., Paolillo et al. 2004, 2017; Uttley & Mchardy 2004; Caplar et al. 2017; Assef et al. 2018). Because the AGN luminosity originates from the accretion process, the study of AGN variability can provide crucial information about the physics of the supermassive black hole (SMBH) and central engine, including at spatial scales that are beyond the resolving power of most of the current and future facilities for most AGNs. In addition, it allows the close link between SMBHs and their hosts to be probed by taking into account the fact that the energy injection from the AGN varies with time (e.g., Hickox et al. 2014). AGN variability is also often employed as an efficient method to select AGNs from large multiepoch surveys (e.g., Trevese et al. 2008; Villforth et al. 2010; De Cicco et al. 2015; Sánchez-Sáez et al. 2019).

In the past decades, different campaigns with recurrent photometry and/or spectroscopy such as the Sloan Digital Sky Survey (SDSS) Stripe 82 (York et al. 2000; Ivezić et al. 2007; Sesar et al. 2007), the (intermediate) Palomar Transient Factory (PTF/iPTF; Bellm 2014), the Palomar Observatory Sky Survey (POSS; Minkowski & Abell 1963), and the Catalina Real-Time Transient Survey (CRTS; Drake et al. 2009; Djorgovski et al. 2011) have provided important insights into the variability behavior of the AGN population on timescales of days to decades (e.g., MacLeod et al. 2012; Morganson et al. 2014; Caplar et al. 2017; Graham et al. 2017). Indeed, these works show that the

amplitude of variability is increasing for increasing timescales, and investigate additional dependencies with physical parameters of AGNs. For a fixed timescale, the amplitude of variability is observed to anticorrelate with luminosity, rest-frame wavelength, and Eddington ratio, while little or no dependence is found as a function of redshift (e.g., Wilhite et al. 2008; Ai et al. 2010; MacLeod et al. 2010; Caplar et al. 2017; Rumbaugh et al. 2018). The dependence on black hole mass is still unclear because different studies found either positive, negative, or absent correlation (e.g., Wilhite et al. 2008; Kelly et al. 2009; MacLeod et al. 2010; Zuo et al. 2012; Caplar et al. 2017). Interestingly, the derived dependencies, as well the observed timescales at which variability is occurring in the optical and UV, cannot be fully explained with standard accretion disk theory (e.g., MacLeod et al. 2010; Zuo et al. 2012; Caplar et al. 2017).

In addition to these overall AGN variability properties, in the last years, new intriguing objects such as changing-look quasars (or more generally, changing-look AGNs, CL-AGNs) have been found in increasingly large numbers (e.g., LaMassa et al. 2015; McElroy et al. 2016; Ruan et al. 2016; Runnoe et al. 2016; Katebi et al. 2018; Mathur et al. 2018; Ross et al. 2018; Stern et al. 2018; Wang et al. 2018; Yang et al. 2018; Zetzl et al. 2018; MacLeod et al. 2019; Trakhtenbrot et al. 2019). CL-AGNs are characterized by the appearance or disappearance of broad Balmer emission line components over periods ranging from months to years,[7] often accompanied by changes in luminosity of over one order of magnitude in the

---

[7] Another group of so-called CL-AGNs are AGNs with X-ray spectra that show a switch from being Compton thick to Compton thin, or vice versa (e.g., Matt et al. 2003; Piconcelli et al. 2007; Marchese et al. 2012; Ricci et al. 2016). The relation between "optical CL-AGNs" and "X-ray CL-AGNs" is not clear yet.

same period, which again challenges our understanding of AGN accretion physics. Indeed, the observed large magnitude changes in short timescales are not consistent with the predictions from standard thin accretion disks (e.g., Krolik 1999; LaMassa et al. 2015; MacLeod et al. 2016). CL-AGNs also challenge the standard AGN unification model (Antonucci 1993; Urry & Padovani 1995), where the presence and absence of broad emission lines are explained by different line-of-sight orientations, as the angular position of the obscuring material is not expected to vary on such short timescales. One open question is if a similar level of variability should be expected for every AGN, or if different AGNs may belong to distinct classes with significantly different variability properties and physical mechanisms.

As mentioned above, the rapid variability observed in the optical/UV continuum, in particular in extreme cases such as CL-AGNs, is hard to reconcile with standard accretion disk theory. Indeed, timescales of months to years are much shorter compared to the viscous timescales at which variability due to overall accretion changes would appear, which is on the order of ~100–1000 yr (see Lawrence 2018 for a discussion of the "viscosity crisis"). The observed timescales are more consistent with thermal timescales, which are relevant for instabilities or local perturbations in the accretion disk. Such instabilities and perturbations could explain the moderate variability characteristic of the majority of AGNs, as well as produce structural changes in the inner part of the accretion disk, which could lead to the extreme variability observed in CL-AGNs (e.g., Ross et al. 2018; Stern et al. 2018). The possible causes of such instabilities and perturbations are still a matter of debate, but they include magnetorotational instabilities (MRI; Balbus & Hawley 1991; Reynolds & Miller 2009), local temperature fluctuations driven, e.g., by X-ray heating (Shappee et al. 2013), iron opacity (Jiang et al. 2016), as well as perturbations due to stellar-mass black holes, stellar remnants, and stars moving within the dense medium of the accretion disk (e.g., Syer et al. 1991; McKernan et al. 2014; Bartos et al. 2017). Other possible explanations are that accretion disks are magnetically elevated, which could lead to larger scale heights and shorter variability timescales compared to those expected for standard thin accretion disks (Dexter & Begelman 2019), or that the behavior of the accretion disk is dominated by nonlocal physics such as magnetic fields connecting different regions on timescales shorter than viscous timescales (Lawrence 2018).

The above refers to variability that can be directly traced through observations. In addition to this, indirect arguments based on the photoionization state of large-scale gas in (and outside) galaxies indicate that AGNs may dramatically change their luminosity on timescales that are significantly longer than what we can directly probe, i.e., $\gtrsim 10^4$ yr (e.g., Lintott et al. 2009; Keel et al. 2012a; Gagne et al. 2014; Schawinski et al. 2015; Sartori et al. 2016, 2018a). This type of long-timescale variability is addressed by simulations (e.g., Novak et al. 2011; Gabor & Bournaud 2013) and theoretical models (e.g., Martini & Schneider 2003; King & Nixon 2015), and may have implications for our attempts to understand the SMBH–host galaxy coevolution (Hickox et al. 2014; Volonteri et al. 2015).

With the advance of large, multiwavelength time-domain surveys such as the Large Synoptic Survey Telescope (LSST; Ivezic et al. 2008; LSST Science Collaboration et al. 2009), the Time-Domain Spectroscopic Survey in SDSS-IV (TDSS; Morganson et al. 2015), SDSS-V (Kollmeier et al. 2017), the Zwicky Transient Facility (ZTF; Bellm et al. 2019), the All-Sky Automated Survey for SuperNovae (ASAS-SN; Shappee et al. 2014), and the extended ROentgen Survey with an Imaging Telescope Array (eROSITA; Merloni et al. 2012), we are entering an exciting era of time-domain astronomy which will allow us to probe the variable universe with unprecedented cadence (~days), depth, sky area, and time span. These surveys will allow us to probe the variability properties of individual AGNs and of the AGN population as a whole, as well as new types of flares and other extreme variability phenomena associated with SMBH accretion such as CL-AGNs or tidal disruption events (TDEs; e.g., Rees 1988; van Velzen et al. 2011; Gezari et al. 2012; Arcavi et al. 2014; Chornock et al. 2014; Hung et al. 2017; Auchettl et al. 2018). This will provide crucial information to quantify the variability phenomena and to test various models for its origin. However, challenges in modeling and characterization of the observed light curves will be important. These are due to the finite time resolution and length of the obtained light curves (e.g., Uttley et al. 2002; Emmanoulopoulos et al. 2010), as well as sample selection and observational biases inherent to each survey. A careful study of AGN variability therefore requires sophisticated statistical analysis methods as well as computationally expensive modeling and simulations.

In Sartori et al. (2018b), we proposed that the AGN variability observed at optical/UV wavelengths may be modeled based on the distribution of the Eddington ratio ($L/L_{\rm Edd}$) among the (observed) AGN population. In fact, the emission at such wavelengths mainly arises from the accretion process (Shakura & Sunyaev 1973) and can therefore be modulated by changes in $L/L_{\rm Edd}$, although other processes such as reprocessing of high-energy photons from the hot corona may also affect the observed luminosity at short timescales (e.g., Uttley et al. 2003). Specifically, we suggest that AGN light curves can be fully simulated starting from two statistical functions: the Eddington ratio distribution function (ERDF, representing the $L/L_{\rm Edd}$ probability density function PDF; see Section 2.1.1) defining the possible $L/L_{\rm Edd}$ values in the simulation, and a power spectral density (PSD; see Section 2.1.2) describing the variability and therefore the time ordering of the $L/L_{\rm Edd}$ points. The obtained $L/L_{\rm Edd}$ time series can then be converted to (optical/UV) light curves in the observed band in question following commonly used conversion factors (i.e., reciprocal bolometric corrections). In this framework, every AGN light curve is therefore one realization of the underlying variability process described by the assumed ERDF+PSD set and represents the entire AGN population (following the ergodic hypothesis). A forward modeling approach allows the predictions of our models to be compared to observations to test if and how the observed variability features can be reproduced with this simple model and, if this is the case, constrain the underlying ERDF and PSD. In addition, given a model, it is possible to produce light curves that would be observed by specific facilities. We also stress that what Sartori et al. (2018b) allows to link and discuss, in the same framework, AGN variability observed at very different timescales.

In this paper, we present a new simulation setup to produce AGN light curves following the framework proposed in Sartori et al. (2018b), based on the algorithm presented in Emmanoulopoulos et al. (2013, E13 hereafter), and discuss possible applications. The paper is structured as follows. In Section 2, we describe the model and method. An extensive description of the simulations code and the tests to characterize both intrinsic

and numerical behaviors of the simulations is given in Section 3. Finally, in Section 4, we discuss some early application of our framework based on the observed variability in the PTF/iPTF survey and aimed at interpreting extremely variable SMBHs in the LSST era. An up-to-date version of the simulations code, along with examples and additional information about its usage, can be found at https://github.com/nevencaplar/AGN-Variability-Simulations.

2. Model and Method

The goal of our framework is to statistically model the (observed) AGN variability due to changes in accretion (thus, variability in Eddington ratio, $L/L_{\rm Edd}$), by simulating the $L/L_{\rm Edd}$ time series based on the distribution of $L/L_{\rm Edd}$ among the (active) galaxy population. Specifically, we simulate the $L/L_{\rm Edd}$ time series with PDF inspired by the ERDF and variability behavior described by a chosen input PSD.

In this section, we outline the methods that we use to simulate the $L/L_{\rm Edd}$ time series. Specifically, we discuss how we define the input PDF and PSD, and how the simulated $L/L_{\rm Edd}$ time series can be converted to (observed) AGN light curves. Details about the code implementation and testing are given in Section 4. Finally, some early applications aimed at illustrating the potential of our new framework are discussed in Section 3.

We note that, although the framework presented here and the algorithm are optimized for AGN and $L/L_{\rm Edd}$ time series, they can be adapted to investigate other variable processes such as, e.g., stellar variability (e.g., Catelan & Smith 2015 and references therein; Labadie-Bartz et al. 2017) or star formation histories (Caplar & Tacchella 2019).

2.1. Simulations of Eddington Ratio Time Series

To simulate $L/L_{\rm Edd}$ time series, we start from the time-series-generating algorithm proposed by E13. This algorithm, which is based on previous algorithms presented in Timmer & Koenig (1995, TK95 hereafter) and Schreiber & Schmitz (1996, S96 hereafter), produces time series with PDF and PSD consistent with the desired (input) ones. These input PDF and PSD can be taken from real observations (e.g., to reproduce multiple observations with the same variability behavior as an observed light curve), or from models that have to be tested against the data. The first step consists of creating a time series of the desired length (e.g., in years) and time resolution, whose periodogram scatters around the underlying input PSD (TK95 algorithm), therefore determining the variability in the data. A second time series with the desired number of steps and PDF corresponding to the input one is then obtained through a random draw. Finally, the two time series are combined following the iterative amplitude-adjusted Fourier transform algorithm described S96 to produce a time series with both PDF and PSD consistent with the input ones (see Figure 1 in E13). We stress that although this algorithm can be used to generate the time evolution of any given quantity (e.g., count rates, fluxes, or magnitudes), in our framework it is only employed to directly simulate the $L/L_{\rm Edd}$ time series, while the conversion from $L/L_{\rm Edd}$ to other observed quantities is performed in post-processing.

In our framework, we assume that the PDF of the $L/L_{\rm Edd}$ time series corresponds to the ERDF of the AGN population (as discussed in Section 2.1.1). We therefore implemented the

algorithm in such a way that it allows us to choose between different ERDF shapes proposed in the literature, such as a broken power-law or log-normal distribution (see, e.g., Caplar et al. 2015, 2018; Weigel et al. 2017). We note that our fundamental assumption that the PDF corresponds to the ERDF is physically acceptable only if the simulation is long enough for the AGN to span the whole ERDF. Although current knowledge about AGN variability does not allow this timescale to be constrained yet, we assume that it should be on the order of $\sim 10^6$ yr (see Section 4.1). On the other hand, in order to compare the predictions to observations, the simulations require a time resolution of years, and preferably even shorter timescales. As a consequence, we need to simulate light curves with $>10^7$ steps. This makes the simulations extremely challenging from a computational point of view. In order to maximize the performance and length of the simulated light curves, we thus implemented our entire algorithm to run on an architecture of graphical processing units (GPUs; Section 3.1).

2.1.1. The ERDF as the Variability PDF

As mentioned above, we propose using the ERDF as the PDF of the AGN (population) variability. By construction, the ERDF describes the distribution of $L/L_{\rm Edd}$ among the galaxy population, i.e., the fraction of galaxies in a given $L/L_{\rm Edd}$ range, for a given sample of galaxies at a given redshift range.[8] Assuming that the ERDF shape does not vary significantly during the considered time span, and that all galaxies in the considered sample have the same variability properties, the ERDF can also be interpreted as the distribution of $L/L_{\rm Edd}$ that a galaxy can have during this time span (up to a normalization factor): a galaxy moves across the ERDF, spending more time at $L/L_{\rm Edd}$ states corresponding to higher ERDF values.[9]

Many studies have attempted to infer the ERDF directly from observations (e.g., Kollmeier et al. 2006; Kauffmann & Heckman 2009; Schulze & Wisotzki 2010; Aird et al. 2012; Jones et al. 2016). However, directly measuring the ERDF is challenging, mainly due to selection effects and difficulties in measuring black hole masses $M_{\rm BH}$ for statistically significant samples (e.g., Trakhtenbrot & Netzer 2012; Shen 2013; Peterson 2014; Mejía-Restrepo et al. 2016). Caplar et al. (2015) and Weigel et al. (2017, W17 hereafter) proposed a different approach, where they assume a simple ERDF shape and apply forward modeling to deduce the parameterization of the ERDF by deconvolving the AGN luminosity function. Specifically, W17 determined the the ERDF for local AGNs, while Caplar et al. (2018) showed that the typical $L/L_{\rm Edd}$ of the (broken power-law) ERDF increases with redshift following $\lambda^*(z) \propto (1 + z)^{2.5}$ (at least to $z \sim 2$).

Different parameterizations have been proposed in the literature to describe the observed ERDF. The most common are the broken power-law (e.g., Caplar et al. 2015; Bongiorno et al. 2016; W17, Caplar et al. 2018) and the log-normal distributions (Kollmeier et al. 2006; Conroy & White 2013; Kauffmann & Heckman 2009). Because the aim of our study is to reproduce the observed AGN variability, and therefore the behavior of (mainly) radiatively efficient AGNs, we consider as the local ($z \sim 0$) PDF the (normalized) ERDF inferred by W17 [10] (see below). In later stages of the present work, we will use

---

[8] We note that in many empirical studies the ERDF is determined for the AGN population instead of the total galaxy population.

[9] In a monotonically declining ERDF, this translates to spending more time at lower $L/L_{\rm Edd}$ values than at higher ones.

[10] In the present work, we are considering values for radiatively efficient AGNs.
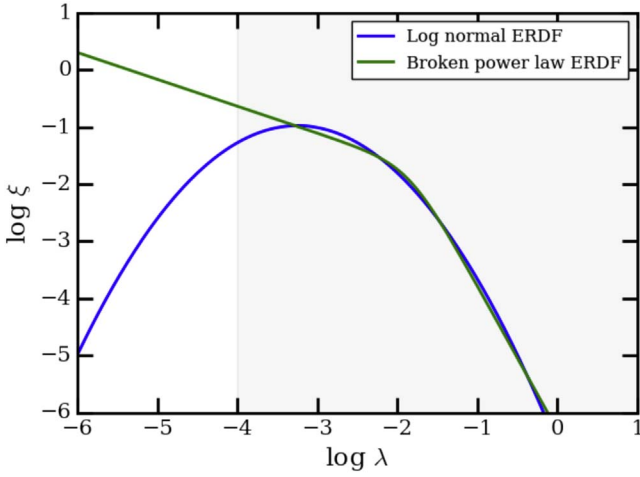
Figure 1. Eddington ratio distribution functions for local AGNs, from W17, described either as a log-normal (blue) or broken power-law (green) distribution. The two descriptions are essentially indistinguishable for the $L/L_{\rm Edd}$ (λ) range probed in that work, i.e., $\log \lambda \approx -4$ to 1 (lightly shaded region). We note that while here we present $\xi(\lambda) = dN/d \log \lambda$ (i.e., the space density per logarithmic unit in $L/L_{\rm Edd}$), as parameterized in Equations (1) and (2), our implementation uses a linear scaling, i.e., $dN/d\lambda$.

the analysis of Caplar et al. (2018) to evolve the ERDF to higher redshift. We stress again that these assumptions can be made only if the length of the simulated light curve is comparable to the time needed for the AGN to span the whole ERDF, and if the ERDF is not expected to vary significantly on these timescales.

Following W17, the broken power-law ERDF for local AGNs is parameterized as

$$\xi(\lambda) = \frac{dN}{d \log \lambda} = \xi^* \times \left[\left(\frac{\lambda}{\lambda^*} \right)^{\delta_1} + \left(\frac{\lambda}{\lambda^*} \right)^{\delta_2} \right]^{-1}, \qquad (1)$$

with parameter values $\log \lambda^* = -1.84$, $\delta_1 = 0.47$, $\delta_2 = 2.53$, and $\log \xi^* = -1.65$. On the other hand, the log-normal parameterization is

$$\xi(\lambda) = \frac{dN}{d \log \lambda} = \frac{\xi^*}{\tilde{\sigma}\sqrt{2\pi}} \times \exp\left(\frac{-(\log \lambda - \log \lambda^*)^2}{2\tilde{\sigma}^2} \right), \quad (2)$$

where $\log \lambda^* = -3.25$, $\tilde{\sigma} = 0.64$, $\log \xi^* = -0.77$.

It is important to note that the two parameterizations are almost equivalent for the $L/L_{\rm Edd}$ range probed in W17, i.e., $\log \lambda \approx (-4) - (+1)$. However, the currently available data do not allow the ERDF shapes to be constrained at lower ER values where the two parameterizations differ significantly, with the log-normal distribution leading to fewer low $L/L_{\rm Edd}$ AGNs (Figure 1). A drop at low $L/L_{\rm Edd}$ is, however, expected because the ERDF cannot increase indefinitely at the lower end. In order to allow future users of our implementation to use various ERDF shapes (as proposed in the literature), we implemented in our simulations code the possibility of using both the broken power-law and the log-normal PDF parameterizations (see Section 3.1). We note that in the simulations we will sample the PDF uniformly in linear space (as opposed to logarithmic space). Therefore, we will consider $dN/d\lambda$ instead of $dN/d \log \lambda$ (see Section 3.1 for more details).

2.1.2. Power Spectral Density

The second statistical function needed from our simulations is the PSD that describes the time variability of $L/L_{\rm Edd}$. The total PSD can be interpreted as a superposition of different physical processes driving the variability at different timescales, and therefore at different frequencies, and its overall shape is not yet fully understood. Multiple studies have performed detailed investigations of the PSD on timescales of hours to decades derived from light curves in the X-rays (in counts units; e.g., Markowitz et al. 2003; McHardy et al. 2004; González-Martín & Vaughan 2012; Paolillo et al. 2017) or optical/UV (in magnitude units; e.g., C17, Smith et al. 2018). Specifically, in the case of X-rays, the PSD is generally well described as a power law with slope approximately $\alpha = 2$, consistent with a random walk, at the highest probed frequencies (e.g., Green et al. 1993; Lawrence & Papadakis 1993), and one or two breaks leading to shallower low-frequency PSD (e.g., Uttley et al. 2002; McHardy et al. 2007). The break frequency appears to be correlated with BH mass and $L/L_{\rm Edd}$ (McHardy et al. 2004, 2006), but the physical reason for this correlation is still unclear (e.g., McHardy et al. 2006; González-Martín & Vaughan 2012). On the other hand, studies based on different optical/UV photometric surveys (e.g., Zu et al. 2013; Caplar et al. 2017) or on *Kepler* light curves (Mushotzky et al. 2011; Kasliwal et al. 2015; Smith et al. 2018) revealed steeper high-frequency PSD, strongly varying between different AGNs.

The studies mentioned above all refer to PSDs in counts or magnitude units, while to the best of our knowledge, no measurement or prediction for the PSD in $L/L_{\rm Edd}$ units is present in the literature. However, for the sake of simplicity, and inspired by the results above, we assume that the PSD of $L/L_{\rm Edd}$ also has a broken power-law shape. In this scenario, the bending may be associated with some specific physical process that suppresses the variability on progressively short frequencies, e.g., the response of the accretion disk to instabilities (e.g., Suberlak et al. 2017 and references therein).

Through the paper, we adopt the following parameterization of a broken power law:

$$\text{PSD}(f) = A \times \left[\left(\frac{f}{f_{\rm br}} \right)^{\alpha_{\rm low}} + \left(\frac{f}{f_{\rm br}} \right)^{\alpha_{\rm high}} \right]^{-1}, \qquad (3)$$

where $f_{\rm br}$ is the break frequency, and $\alpha_{\rm low}$ and $\alpha_{\rm high}$ are the slopes at lower and higher frequencies, respectively (longer and shorter timescales, respectively). This parameterization can, however, be modified to add additional breaks if needed. In addition, we do not consider any evolution of the variability behavior, and therefore of the PSD, with redshift, as supported by some observations (e.g., C17, Paolillo et al. 2017). We stress that throughout the paper, we will often refer to the periodogram, which is the statistical estimator of the PSD. The periodogram can be directly computed from a time series and scatters around the underlying PSD (see the Appendix).

2.2. Converting Eddington Ratio Time Series to Observables: Light Curves and Structure Function

As discussed in the Introduction and at the beginning of this section, the primary outputs of our simulations are time series in units of $L/L_{\rm Edd}$. In order to compare the simulations to observations, or to make predictions for future surveys, the

simulated $L/L_{\rm Edd}$ time series therefore have to be converted to observables. First, the simulated $L/L_{\rm Edd}$ time series can be converted to light curves in luminosity or magnitude units at the same wavelength range as the observations we want to compare to. The forward modeling approach allows different conversion recipes to be tested, e.g., by considering constant or luminosity-dependent bolometric corrections (e.g., Marconi et al. 2004 and references therein), as well as by introducing flux limits or other observational biases proper of every observation. The obtained light curves can then be treated as real observations and used to compute other quantities characterizing the AGN variability behavior. Specifically, in this work, we concentrate on the (ensemble) structure function (SF), which quantifies the characteristic amount of variability for light-curve (or in general time-series) measurements separated by a given time interval τ. A detailed definition and discussion of the SF is given in the Appendix (see in particular Equation (14)).

3. Simulations

As discussed in the previous sections, in this paper, we present a new approach for investigating AGN variability on multiple timescales and among different objects, based on the simulation of $L/L_{\rm Edd}$ time series. In this section, we provide details about the implementation of our simulations code, and we illustrate the extensive tests performed in order to characterize both the intrinsic and numerical behavior of the simulations.

3.1. Code Implementation

We implemented the algorithm proposed by E13 to be executed on GPUs. Specifically, we wrote our implementation in C++ within the CUDA programming framework,[11] with the libraries cuFFT,[12] cuRAND,[13] and Thrust.[14] This platform choice allows us to maximize the length (in steps) of the simulations and is therefore more suited to our science goal than other available implementations based on, e.g., MATLAB (E13) or Python (Connolly 2015). This is particularly due to the fact that the algorithm requires multiple (inverse) discrete Fourier transform (DFT) and sorting operations,[15] which are time consuming especially on arrays of the lengths considered in our work (up to $\sim 10^8$ points) and are being optimized on this platform.

We note that the current implementation is optimized for broken power-law and log-normal PDFs, and for single or broken power-law PSDs. Specifically, for the random draw step, we apply slice sampling (Neal 2003) from a normalized ERDF that is parameterized as in Equations (1) and (2). This choice of PDF and PSD shape is motivated by the observed ERDF and PSD, as discussed in Section 2. However, additional functional forms can be added if needed. As mentioned in Section 2.1.1, we are sampling the PDF uniformly in linear space.

---

[11] Version 9.0, https://developer.nvidia.com/cuda-downloads.
[12] https://docs.nvidia.com/cuda/cufft/index.html
[13] https://docs.nvidia.com/cuda/curand/index.html
[14] https://docs.nvidia.com/cuda/thrust/index.html
[15] As an example, one simulation with $N_{\rm steps}$ iterations requires $O(N)$ DFT, $O(N)$ IDFT, and $O(N)$ sorting operations. See also Section 3.2.3.

3.2. Testing and Characterization of the Simulations Code

We tested and optimized our code for the TITAN X (Pascal) NVIDIA GPU. In the following, we describe a series of tests that we performed to characterize both numerical effects and intrinsic behavior of the simulations.

3.2.1. Distortions in the Timmer & König Algorithm

The first step of the E13 algorithm consists of running the TK95 algorithm to create a time series of the desired length with a periodogram consistent with the input PSD. The PDF of the output time series should asymptotically approach a normal distribution for an increasing number of steps (e.g., E13). However, we found that this is not true for power-law PSDs with steep slopes, $\alpha \gtrsim 1.5$ (PSD $\propto \nu^{-\alpha}$). We illustrate these distortions for several choices of α in Figure 2. We conducted a number of simple tests in an attempt to understand the origin of these distortions. We found that the deviation from a normal PDF does not depend on the number of steps and is therefore not a consequence of the finite simulation length. We also found that the same behavior is present in simulations performed with other open codes (e.g., astroML, Ivezić et al. 2014; Vanderplas et al. 2012; DELCgen, Connolly 2015) and is therefore not an artifact of our specific implementation. Because our simulations only rely on the PSD of the TK95 output and not on its PDF, this discrepancy between claimed and observed PDF does not affect our results. However, the TK95 algorithm is extensively used in the literature, and this aspect should be further investigated. We therefore decided to report it here for future reference.

3.2.2. Iterations and Convergence

The algorithm that we use to produce the time series is iterative in nature (E13, see also Section 2.1), and so has to converge in order to be robust. As described in E13, while the values that can be assumed by the points in the simulated time series are determined at the beginning of the algorithm and do not change during the iterative process, these values are reordered at every iterative step until their variability converges to the variability described by the input PSD. Following E13 and S96, we consider that this convergence is reached when the order of the time series values no longer (significantly) changes between two consecutive steps. The exact number of iterations needed for the algorithm to converge to the final time series depends on the total number of steps, the assumed PDF and PSD, as well as the seed used for the random draw process. E13 proposed a convergence test where, at each step, the periodogram of the resulting time series is fitted with the same functional form as the input PSD, and the convergence of the fit parameters with respect to the input ones is considered. However, due to the significantly larger number of steps used in our simulations—more than three orders of magnitude larger than what was considered in previous applications of the algorithm—this convergence test would be too time consuming to implement in our case. We therefore decided to base our convergence criterion on the ratio between the time series obtained from two consecutive iterations. Specifically, at every iteration i, we consider the following distance measurement $\delta_{\rm i}$
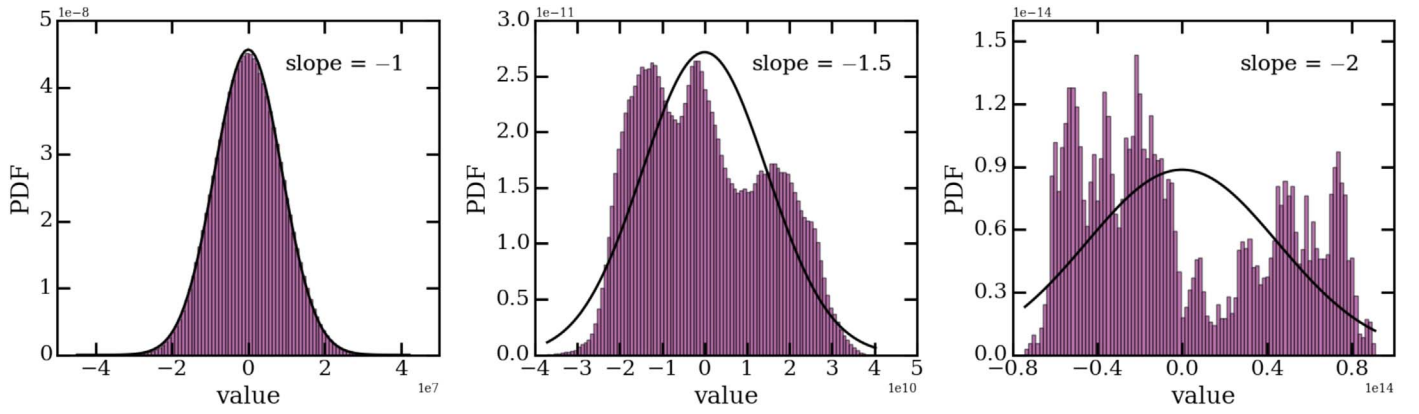
Figure 2. Example output PDFs of simulated time series based on the TK95 algorithm and on single power-law PSDs with various slopes (as indicated in each panel). In each panel, the black lines show normal distributions, which match the output PDFs in mean values and in standard deviations. For steep PSDs, the resulting PDF deviates from a normal distribution. We note that, differently from the text, for this image we assume $PSD \propto \nu^{\alpha}$ instead of $PSD \propto \nu^{-\alpha}$.

between the time series $x_i(t)$ and $x_{i-1}(t)$:

$$\delta_i = \mathrm{rms}\left[\log\left(\frac{x_i(t_n)}{x_{i-1}(t_n)}\right)\right], \; n = 1, \; ..., \mathrm{num.steps}. \quad (4)$$

The use of logarithm is motivated by the large dynamical range of the values in the time series that are simulated using our observationally motivated ERDF as PSD. The threshold corresponding to convergence strongly depends on the input PSD and PDF and should therefore be chosen by the user depending on their simulation setup and the goals of the performed analysis. We ran several tests with 200 up to 1000 iterations and found that, for the considered setups, the simulated time series remain essentially unchanged, except for some minor differences at the highest frequencies probed. We caution and stress that using our framework for simulating the most extreme cases of AGN variability (e.g., >3 dex difference in <1 yr) may require a simulation setup with an increased number of iterations to reach robust results. In what follows, however, we do not put particular focus on such extreme events and instead demonstrate how the framework can be used to describe the broad phenomena of AGN variability seen in large, wide-field time-domain surveys.

3.2.3. Performance

The maximum number of steps per simulation allowed in the current implementation is 2^{27} ($\sim 10^8$), which is set by memory limitations.[16] In the astrophysical context of the present study, this may correspond to, e.g., a $\sim$13 Gyr long simulation with $\sim$100 yr resolution, or a $\sim 10^6$ yr long simulation with $\sim$ day resolution.

The running time is currently limited by the random draw and sorting steps of the algorithm (steps ii and iv in Figure 1 of E13, respectively). On the other hand, using cuFFT ensures that the (inverse) Fourier transform steps are performed in the most efficient way. Because the running time is mostly determined by the random draw step, it can vary depending on the assumed PDF shape. However, typical running times for

one simulation with 200 iterations on TITAN X (Pascal) NVIDIA GPU machines do not exceed a few/several minutes.

### 3.2.4. Dependencies on Input Parameters, Combining and Dividing Light Curves

In this subsection, we discuss how the final L/L_{Edd} time series, PDF, periodogram, and SF shapes depend on the following input quantities:

1. t_{bin}: time separation between consecutive steps
2. N_{steps}: number of steps
3. T: simulation length in the same units of time as t_{bin}. This quantity combines t_{bin} and N_{steps}, i.e., $T = t_{\mathrm{bin}} \times N_{\mathrm{steps}}$.
4. μ: the mean of the simulated L/L_{Edd} time series. For large N_{steps} this corresponds to the mean of the input PDF.
5. σ: the standard deviation of the simulated L/L_{Edd} time series. For N_{steps} that is large enough, this corresponds to the standard deviation of the input PDF.

In addition, we look at the effect of combining or dividing light curves obtained from different simulations. Indeed, given the challenge in running long simulations (see Section 3.2.3), a naive approach would have been to consider "stitching" together separate shorter simulations. On the other hand, understanding the behavior of subsamples is critical for the interpretation of observed light curves, which actually span only a short time window of the whole AGN life (or, in the case of these simulations, of the time needed for the AGN to span the whole ERDF). We note that in this subsection we refer to quantities computed in L/L_{Edd} units, while a discussion of the actual observables, i.e., magnitudes, is given in Section 3.2.10.

For these tests, we considered multiple simulations with broken power-law PDFs inspired by the ERDF from W17 (see Equation (1)) and a broken power-law PSD as in Sartori et al. (2018b; BPL3), but with different choices of t_{bin} and N_{steps}. A summary of the input parameters for the different simulations is given in Table 1. Specifically, we consider L/L_{Edd} time series simulated assuming different combinations of t_{bin}, N_{steps}, and T (ER_sim_1 - ER_sim_6 in Table 1; we will refer to these simulations as "direct simulations"), as well as L/L_{Edd} time series obtained by dividing (ER_sim_2_cut) or stitching together (ER_sim_1_comp) direct simulations. We note that although some of the observed features depend on the specific

---

[16] The number of steps is defined in powers of 2 to maximize the simulation speed, in particular with regard to the discrete Fourier and inverse DFTs. All of the arrays involved in the simulations are allocated in shared memory for the CPU and GPU using cudaMallocManaged.
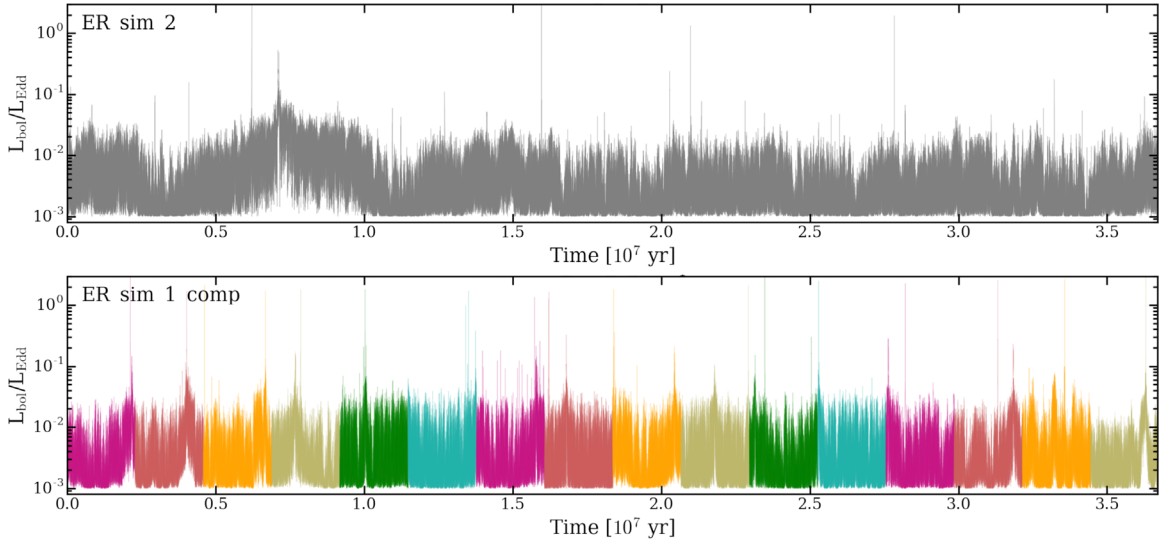
Figure 3. Testing the prospects of simulating increasingly long time series. Top: L/L_{Edd} time series from a single simulation spanning the entire PDF (i.e., the entire ERDF; `ER_sim_2`). Bottom: composite L/L_{Edd} time series obtained by "stitching" together multiple shorter simulations (different colors), each of them also spanning the entire PDF (`ER_sim_2_comp`). Evidently, although both example data sets span the same physical length (in yr) and L/L_{Edd} range, the direct (top) and composite (bottom) simulations present different features. Specifically, the composite simulations show multiple peaks (periods of elevated L/L_{Edd}), approximately one per short simulation, while the single simulation only has one such peak (around 0.75×10^7 yr). This can be naturally understood given that each of the short simulations is forced to cover the entire ERDF. As a consequence, stitching together light curves is not a feasible solution to obtain a realistically long L/L_{Edd} time series (with a given computational setup), because it will not preserve the L/L_{Edd} time-series shape.

Table 1
Summary of Simulations for the Tests in Section 3.2.4

	Name	t_{bin} (ks)	N_{steps}	T (10^6 yr)	Comment
Initial simulation	`ER_sim_1`	8640	$2^{23} \sim 8 \times 10^6$	~2.30	
Longer simulation but with same t_{bin}	`ER_sim_2`	8640	$2^{27} \sim 10^8$	~36.75	16× longer in yr
Append simulations	`ER_sim_1_comp`	8640	$2^{27} \sim 10^8$	~36.75	Append 16 realizations of `ER_sim_1`
Cut long simulation	`ER_sim_2_cut`	8640	$2^{23} \sim 8 \times 10^6$	~2.30	Cut `ER_sim_2` in 16× pieces with same length and binning as `ER_sim_1`
Change t_{bin} and N_{steps}, same T	`ER_sim_3`	4320	$2^{24} \sim 2 \times 10^7$	~2.30	$2 \times N_{steps}$, $0.5 \times t_{bin}$
	`ER_sim_4`	2160	$2^{25} \sim 3 \times 10^7$	~2.30	$4 \times N_{steps}$, $0.25 \times t_{bin}$
Change t_{bin} and T, same N_{steps}	`ER_sim_5`	2160	$2^{23} \sim 8 \times 10^6$	~0.60	$0.25 \times t_{bin}$ and T
	`ER_sim_6`	34560	$2^{23} \sim 8 \times 10^6$	~9.20	$4 \times t_{bin}$ and T

Note.
For reference, 8640 ks corresponds to 100 days.

adopted PSD and PDF, the main conclusions from these tests are also valid for different simulation setups.

Shape of the L/L_{Edd} time series: For the considered PSD and PDF set, the shape (global behavior) of the L/L_{Edd} time series obtained from direct simulations (`ER_sim_1` - `ER_sim_6` in Table 1) is similar and does not depend on t_{bin}, N_{steps}, and T. Specifically, in this particular case, all of the simulations show one main prolonged period of elevated L/L_{Edd} (a "burst" or "switch on"), and a few additional, shorter "spikes." This is a direct consequence of the fact that each of the simulations spans the entirety of the same PDF, and of the (bent) power-law shape of the periodogram, which gives more power to the lowest frequencies (longer time separations).

The composite simulation (`ER_sim_1_comp`) shows more bursts, and therefore a higher level of variability on short timescales, compared to a direct simulation of the same length (`ER_sim_2`; see Figure 3). In fact, in the composite simulation,

every subsample has characteristic features (e.g., burst and spikes) similar to the direct simulation, although at different timescales. This means that stitching together shorter L/L_{Edd} time series will not conserve the variability behavior proper of the input PSD (see also the discussion below). As a consequence, simulating and stitching together shorter light curves, although computationally cheaper (see Section 3.2.3), is not a feasible way to obtain longer L/L_{Edd} time series with both L/L_{Edd} values and variability properties consistent with the assumed PDF and PSD.

Cutting a long simulation into subsamples (`ER_sim_2_cut`) also returns L/L_{Edd} time series with different shapes compared to direct ones with the same final length (`ER_sim_3`, `ER_sim_4`). These are in general smoother as they are not covering the whole L/L_{Edd} range allowed by the PDF, only a (consecutive) part of it.

In summary, combining or cutting simulations returns L/L_{Edd} time series with different shapes and variability features compared to direct simulations with the same final length. This
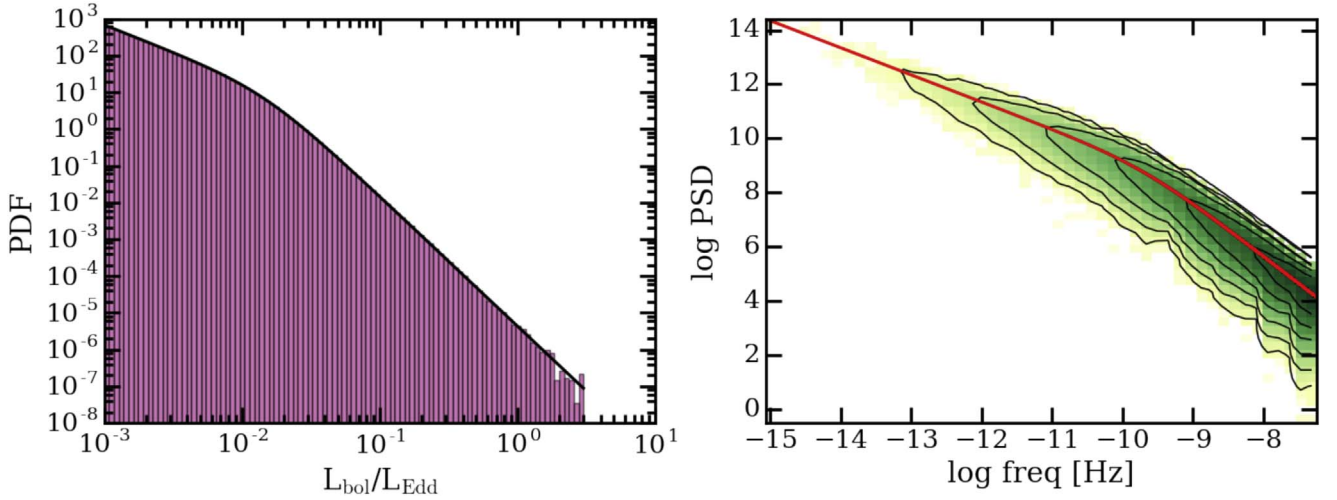
Figure 4. Comparison of input and output for a realistic choice of PDF (i.e., ERDF) and PSD (i.e., broken power laws). Left: the PDF of the `ER_sim_2` simulation (purple) compared with the input PDF (black line). By construction, the PDF (i.e., ERDF) of the simulated time series is statistically consistent with the input one. As expected, the sampling of the PDF becomes more accurate with increasing N_{steps}, as is particularly evident at the high L/L_{Edd} end, where the sampling probability decreases. Right: the periodogram of the same simulation. The red line illustrates the input PSD, normalized according to the μ, σ, t_{bin}, and N_{steps} of the simulation (see Equations (12) and (13)). Here, too, the output periodogram is statistically consistent with the input PSD.
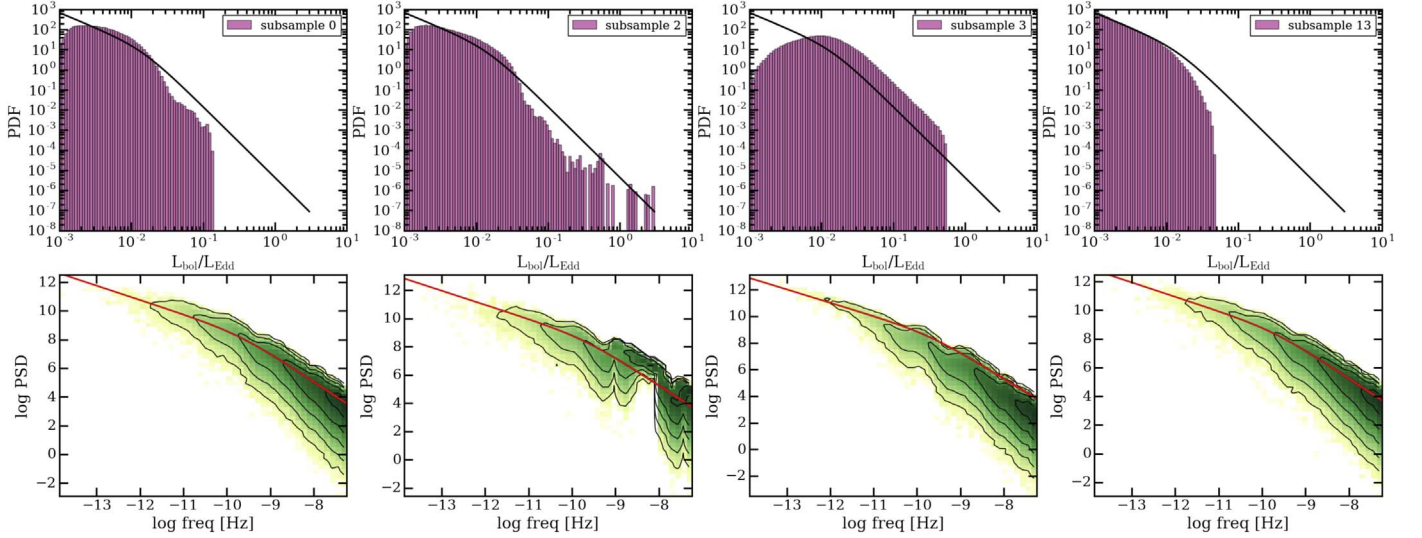
Figure 5. Same as Figure 4 but for four subsamples of `ER_sim_2` (from `ER_sim_2_cut`; that is, shorter simulations). The PDFs of the subsamples do not match the input PDF well, as the simulated time series are not sufficiently long to cover the entire PDF. Moreover, the periodogram shows distortions, which are more important for subsamples covering a broader L/L_{Edd} range. See text for discussion.

effect is visible also in the PDF and periodogram, as discussed below.

Probability density function: By construction, the PDF of every direct simulation (`ER_sim_1`-`ER_sim_6`), and of the composite simulation (`ER_sim_1_comp`), is statistically consistent with the input PDF (e.g., Figure 4, left). In addition, as expected, the sampling of the PDF is more accurate for high values of N_{steps}, especially at the increasingly high L/L_{Edd} end, which is sampled increasingly rarely. The lack of points with high L/L_{Edd} for low N_{steps} simulations has to be taken into account when looking at specific features in the L/L_{Edd} time series, which may ultimately be interpreted as extreme variability events (e.g., CL-AGNs).

On the other hand, the different subsamples (`ER_sim_2_cut`) also show PDFs that are (mostly) continuous in L/L_{Edd} space, but these do not overlap in any way/range with the input PDF (Figure 5, top). Specifically, most of the subsamples span a

narrower L/L_{Edd} range. This is due to the fact that the L/L_{Edd} time series values are not randomly distributed (in time) but follow an order defined by the PSD. As a consequence, the PDF of light curves corresponding to different times during the AGN's life can be significantly different.

Periodogram: By construction, all periodograms computed for the direct simulations (`ER_sim_1` - `ER_sim_6`) have shapes consistent with the input PSD (e.g., Figure 4, right). As described in the Appendix, the normalization of the output periodogram differs from the input one and depends on all other aforementioned parameters, as well as on the covered frequency range. We verified that the normalizations of the periodograms computed for direct simulations are consistent with the expected value.

The periodograms computed for the subsamples (`ER_sim_2_cut`) show some distortions compared to the input PSD, which are more important for subsamples with a larger dynamical range
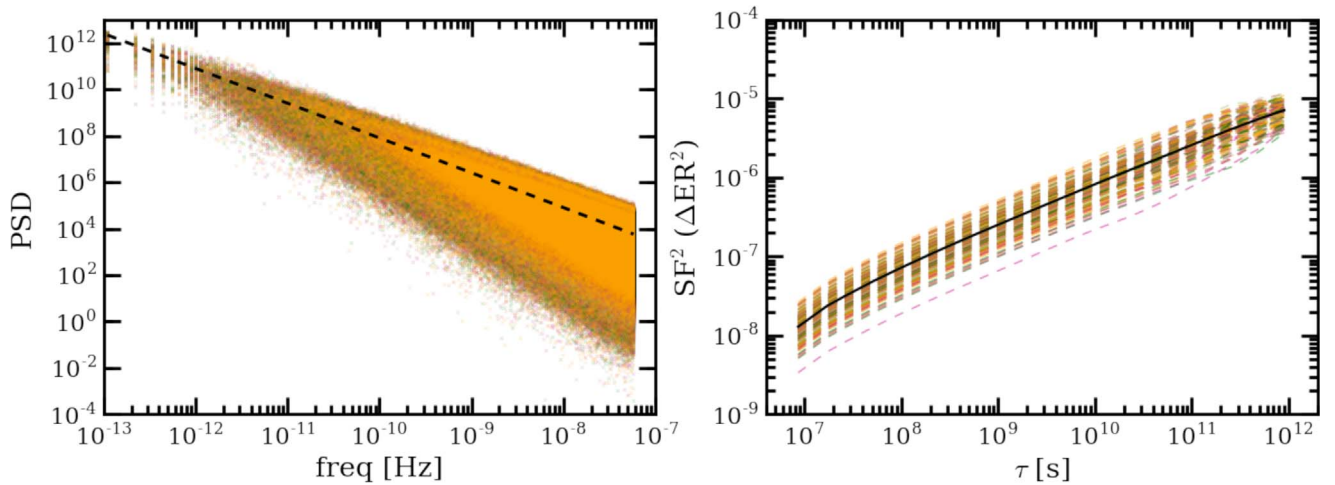
Figure 6. Scatter and robustness in repeated realizations of a time-series simulation. Left: periodograms for 200 realizations of the same simulation input with a simple power-law PSD (slope $\alpha = 1.5$; colored crosses). The black line traces the input PSD (renormalized; see the Appendix for a discussion about PSD normalization). As expected, the periodograms measured from the simulated time-series scatter around the intrinsic (input) PSD. Right: SF^2 for the same simulations (dashed color lines). The SF^2 measurements for different simulations have a similar shape, but different normalizations. The representative (mean) SF^2, defined following Equation (6), is shown in black. We note that, as discussed in Section 3.2.8, some distortions can be observed at the longest probed timescales (i.e., lowest frequencies).

in L/L_{Edd} (Figure 5). A detailed discussion of these distortions, including both well-known spectral distortions and effects related to the simulations, is given in Sections 3.2.7 and 3.2.9. On the other hand, the periodogram of the composite simulation (ER_sim_1_comp) also shows some distortion which in this case may be explained by a suppression of variability at long timescales (every subsample has same L/L_{Edd} values, therefore not allowing the formation of a single peak as observed in the direct simulation, Figure 3 top) and an increase of variability at short timescales (every subsample spans the whole PDF; Figure 3, bottom).

Structure function: We construct the SF following Equation (14) and Sartori et al. (2018b). All of the SFs computed for direct L/L_{Edd} time series (ER_sim_1-ER_sim_6) show a broken power-law shape, which seems to reflect the PSD shape. Indeed, the two slopes coincide asymptotically (for $\tau \to 0$ and $\tau \to \infty$) with the prediction of the Wiener–Khinchin theorem,[17] and the break is broadly consistent with $\tau_{br} \sim 1/\nu_{br}$.

In the case of composite and divided L/L_{Edd} time series (ER_sim_1_comp, ER_sim_2_cut), the SFs show a behavior similar to that described above for the periodogram, again confirming a relation between the two quantities. Specifically, distortions of the PSD at high frequencies are reflected to distortions of the SF at low τ.

The normalization of the SF differs between our various simulations. For a fixed PSD, it depends on the total length of the simulation (T), as well as on the standard deviation of the time series (σ). Specifically, for a pure power-law PSD. we empirically found that the normalization has the following proportionality:

$$\text{norm}_{SF} \propto \frac{\sigma}{(t_{bin} \times N_{steps})^\beta}, \qquad (5)$$

where $\beta = (\alpha - 1)/2$ is the same exponent as in the Wiener–Khinchin theorem, i.e., the exponent expected for the SF (in an ideal case). For a broken power-law PSD, the difference

between SFs obtained from different simulation setups (i.e., different T and σ) is less well defined, as the ratio between two SFs changes as a function of τ (i.e., the SF have slightly different shapes). Although we were not able to find an analytical description of the normalization in this case, we see similar dependencies: the overall SF is higher for increasing σ or decreasing T, although the dependency on T is more important for steeper PSDs. This can be understood as follows. By construction, the time series obtained with our simulations span the entire PDF (i.e., the ERDF). If the simulation length is shorter, this means that the PDF has to be covered faster, which translates to an overall higher level of variability and thus a higher SF. Similarly, a larger σ corresponds to a broader PDF, which, for fixed T, has to be again covered faster. This dependency of the SF normalization on T and σ has to be considered when comparing simulations to observations, as discussed in detail in Section 4.1. Specifically, the length of the simulation has to be chosen such that the normalization of the output SF (i.e., the SF of the simulated time series) matches the observed one. Moreover, we note that the normalization does not depend on t_{bin} or N_{steps} separately, but on the combination of the two (i.e., simulations with different t_{bin} and N_{steps} but the same T have the same normalization).

3.2.5. Mean Variability Behavior from Multiple Realizations

Because we are working with stochastic processes, every simulated ER curve will have a different periodogram and SF, i.e., they fluctuate around the input PSD and SF. By construction, at every frequency, the periodogram scatters around the intrinsic PSD as defined by the Timmer & König algorithm (see Section 3 in Timmer & Koenig 1995 or Appendix A2 in E13 for more details). On the other hand, the SFs of different realizations have similar shapes but different normalizations. Both effects are illustrated in Figure 6.

In order to compare the simulated SF to observations as proposed in Sartori et al. (2018b), we need to define a representative SF and properly take into account the scatter among the different realizations. To get a representative SF for

---

[17] The Wiener–Khinchin theorem states that a power-law PSD with slope $-\alpha$, $1 < \alpha < 3$ should correspond to a power-law SF with slope $\beta = (\alpha - 1)/2$ (under specific assumptions; see also Emmanoulopoulos et al. 2010).
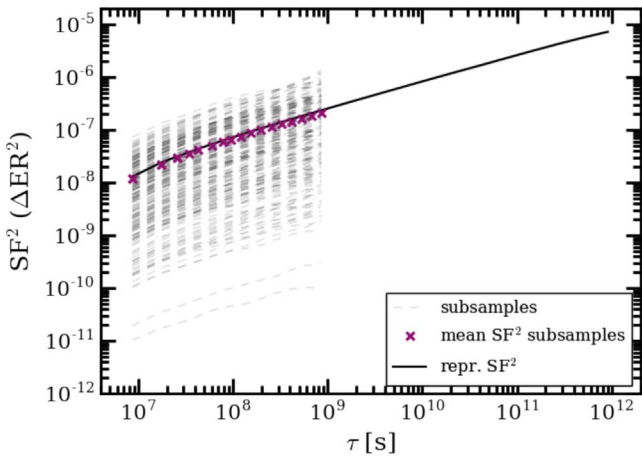
Figure 7. Robustness of the representative structure function. The SF^2 for the random subsamples of the simulations shown in Figure 6 (gray dashed lines; see text for more details). The representative SF^2 of the subsamples (purple crosses) overlaps with the representative SF^2 of the long simulations (black line).

a given input set, $SF(\tau)_{rep}$, we therefore repeat the simulation N_{sim} times ($N_{sim} > 50$; see discussion below), and compute the mean of the obtained $SF(\tau)^2$ (see also C17):

$$SF_{rep}(\tau)^2 = \langle SF(\tau)^2 \rangle = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} SF_i(\tau)^2. \quad (6)$$

The standard error on $SF(\tau)_{rep}$ at every τ can then be estimated using the scatter in the set of calculated $SF_i(\tau)^2$. Uttley et al. (2002) used a similar procedure to determine the intrinsic PSD from multiple light curves, and claimed that >50 realizations are needed in order to reliably estimate the standard errors on the simulated periodograms. The same seems to also hold for the SF, but more testing should be done for every specific input set.

We note that in general the time span probed by observations are much shorter compared to the simulated ones. In order to probe whether the observed SF can be directly compared to the representative one (i.e., there is no distortion or different normalization), we first simulated 200 time series with $\sim 10^6$ points each and computed the representative SF (Figure 6, right). From every time series, we then randomly selected a (consecutive) subsample of 1000 points, obtained the corresponding SF, and then computed the representative SF (Figure 7). Although the shorter SFs show a larger dispersion, the representative SF for the subsamples overlaps with the representative SF of the long simulations. This means that the representative SF can be directly compared to the SF computed by taking into account only subsamples (see also Section 3.2.6). We confirmed that this holds by considering various input PSD shapes and by computing the SF in magnitude units instead of L/L_{Edd} units.

3.2.6. Ensemble Analysis

Because the SF is a statistical measurement, it is possible to compute it for a single source only in the case of a well-sampled light curve (e.g., Suberlak et al. 2017 and references therein). However, for sparsely sampled light curves, the SF is usually computed in an ensemble way (e.g., Sesar et al. 2006; MacLeod et al. 2012; C17). In this case, it is assumed that all sources in the considered sample have the same variability

behavior, and that the SF at a given time lag τ is computed with Equation (14), by considering simultaneously all the pairs of data points separated by a corresponding time lag, in all sources under question.

In order to test if SFs computed in an ensemble way (i.e., considering subsamples of different light curves) are consistent with SFs computed directly from individual simulations, we performed a test similar to what is described at the end of Section 3.2.5. First, we ran 1000 simulations with $\sim 10^6$ points each and selected a (consecutive) subsample of 1000 data points from each simulation. For every simulation, we then masked the subsamples to mimic sparse measurements and computed the ensemble SF. As an example, Figure 8 shows the ensemble SF obtained by considering only 0.5%, 1%, and 10% of every subsample, respectively (randomly selected points). As expected, the ensemble SF always scatters around the representative SF, but the scatter significantly decreases with increasing number of considered points. This means that the mean shape and normalization of the SF are conserved with ensemble analysis, but the accuracy strongly depends on the number of measurements.

We stress that what is reported in the test above is only indicative of a trend. In fact, the scatter in the ensemble analysis strongly depends on the number of considered sources, and on the length and sampling of the light curves measured for each source. Therefore, in order to determine how well an ensemble SF computed for a given survey represents the intrinsic one (i.e., if we can compare the ensemble SF with our simulations), the test above should be repeated by assuming the same source numbers and observational cadence as in the considered survey.

### 3.2.7. Spectral Distortions: Red Noise Leak and Aliasing

The periodograms computed for finite, discrete light curves are subject to spectral distortions known as red noise leak and aliasing (see below for a definition, and Uttley et al. 2002 for a detailed discussion). In fact, measuring a light curve $l(t)$ with a time sampling given by the window function $w(t)$, where $w(t) = 1$ during the observations and 0 elsewhere, corresponds to convolving the Fourier transform of the underlying light curve $F(f)$ with the Fourier transform of the window function $W(f)$:

$$l_{obs}(t) = l(t) \times w(t) \quad \Rightarrow \quad F_{obs}(f) = F(f)*W(f). \quad (7)$$

Because the periodogram is computed from $F_{obs}(f)$ (see the Appendix), a distortion due to a window function is reflected as a distortion in the periodogram. Specifically, red noise leak (e.g., Deeter & Boynton 1982; Deeter 1984) appears when the length of the observed light curve is significantly shorter than the length of the intrinsic light curve, so that significant power is present at frequencies shorter than the ones probed with the observations. In this case, trends in the light curve due to frequencies below the observed limit (timescales longer than the light-curve total length) are not distinguishable from trends at higher frequencies, and power is transported from low to high frequencies. This effect is more pronounced for steep PSDs (see Uttley et al. 2002 and references therein). On the other hand, aliasing (e.g., van der Klis 1997) appears when the time sampling does not allow variability to be probed at high frequencies f, and in this case there is a foldback of power from $f_{Nyq} + \Delta f$ to $f_{Nyq} - \Delta f$, where f_{Nyq} is the Nyquist frequency
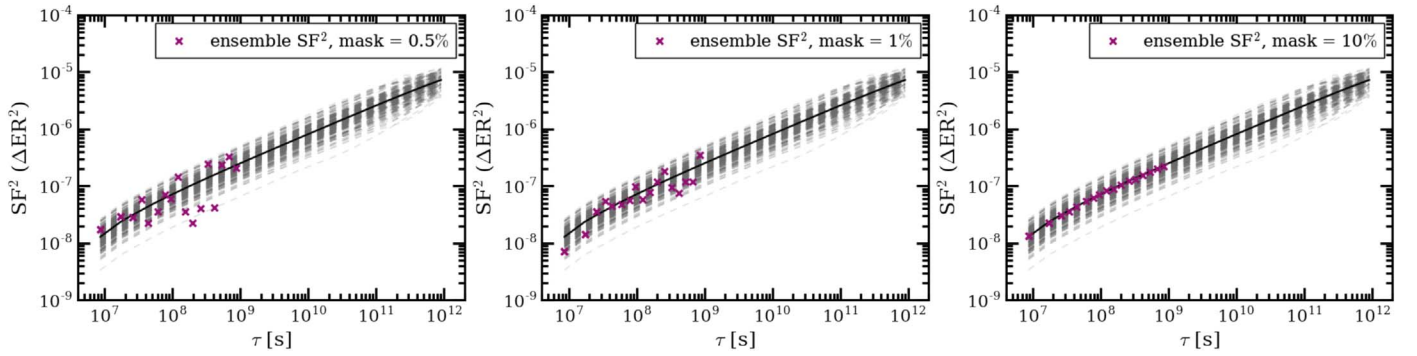
Figure 8. Ensemble structure functions for subsamples of the simulated time series—similar to Figure 6 (right) and Figure 7, but for ensemble analysis. Here, the purple crosses show the ensemble SF obtained by considering only 0.5%, 1%, and 10% (right) of the simulated data points in every subsample, (left, center, and right, respectively; randomly selected points). The ensemble SF always scatters around the representative SF, but the scatter decreases significantly with increasing number of considered points.
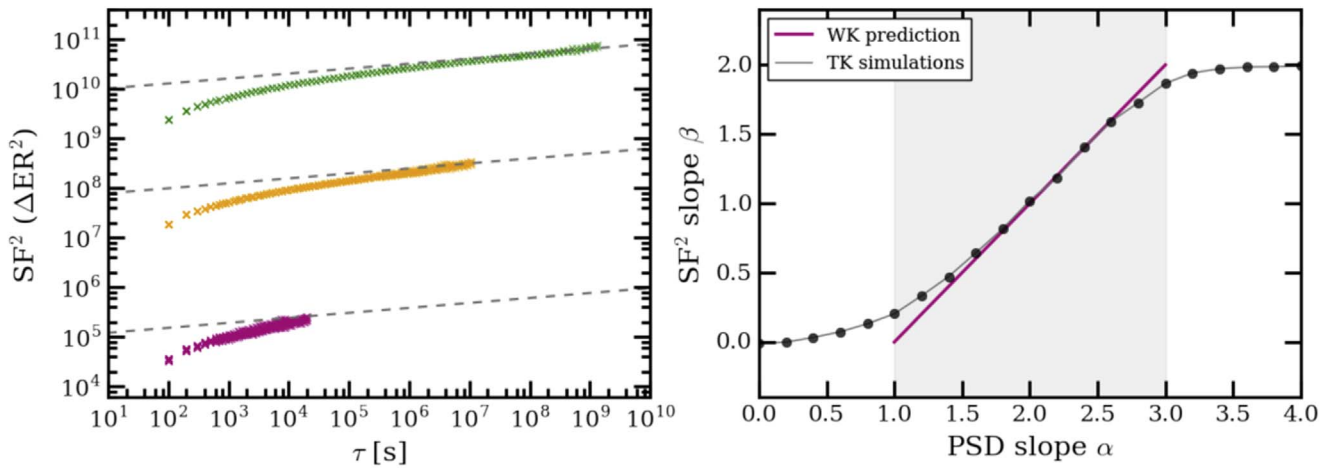
Figure 9. Left: Structure functions computed for TK95 simulation with PSD with slope $\alpha = 1.1$, a fixed time step, but different lengths. The gray dashed lines have slopes as expected from the Wiener–Khinchin theorem for this PSD (arbitrarily normalized). All of the simulated SFs show a downturn at the shortest probed timescales, which is independent of the simulation length. Right: resulting SF slopes as a function of input PSD slope (black points). Each point was obtained by fitting a line to 100 TK95 simulations (for every PSD slope). Here, SF $\propto \tau^{\beta}$ and PSD $\propto \nu^{-\alpha}$. The results deviate from expectation from the Wiener–Khinchin theorem (purple) as the PSD slopes deviate from $\alpha = 2$, with higher (lower) mean SF slopes for decreasing (increasing) PSD slopes.

(see the Appendix). This effect is more pronounced for shallow PSD.

Both red noise leak and aliasing cause a flattening of the measured periodogram, at a level that depends on the shape of the underlying PSD. Studies comparing simulated periodograms to observed ones should therefore include such distortions in the simulations. The standard way to do it is to simulate longer light curves (e.g., 100 times longer, for red noise leak) with higher sampling (e.g., 10% of the observed sampling, for aliasing), and then randomly select a subsample with length and sampling as in the observations (e.g., Uttley et al. 2002).

3.2.8. Distortions of the Structure Function

While the spectral distortions discussed in Section 3.2.7 can significantly modify the computed periodogram, the SF is commonly thought to be less affected by finite, irregularly sampled light curves. However, Emmanoulopoulos et al. (2010) showed that spurious breaks can appear in the SF of single light curves even in the case of a power-law PSD (i.e., no intrinsic characteristic timescales), with unpredictable behavior at timescales longer than the break. The position of the break increases with light-curve length and PSD steepness

(see Figure 5 in Emmanoulopoulos et al. 2010), so that the effect is less important in long light curves with steep PSDs. In addition, he showed that windowing of the light curve may also affect the precise SF shape (although the general behavior is conserved). As can be seen in, e.g., Figure 6, some distortions at the longest probed timescales are also present in our simulations, although they are not as extreme as reported in Emmanoulopoulos et al. (2010). However, these distortions smooth out when considering the representative or ensemble SF, so that these distortions do not significantly affect statistical studies (see also Guo et al. 2017; Kozłowski 2017a), as is the case in the applications presented in Section 4.

Although the distortions described in Emmanoulopoulos et al. (2010) do not affect ensemble studies, other distortions can be seen in our simulations and should be taken into account. According to the Wiener–Khinchin theorem, a power-law PSD with slope $-\alpha$, $1 < \alpha < 3$, should correspond to a power-law SF with slope $\beta = (\alpha - 1)/2$ (under specific assumptions, see also Emmanoulopoulos et al. 2010). However, as can be seen in Figure 9, for $\alpha \lesssim 1.5$ ($\gtrsim 2.5$), the obtained SF is steeper (shallower) than expected from the theorem. For a given time resolution, this effect is more
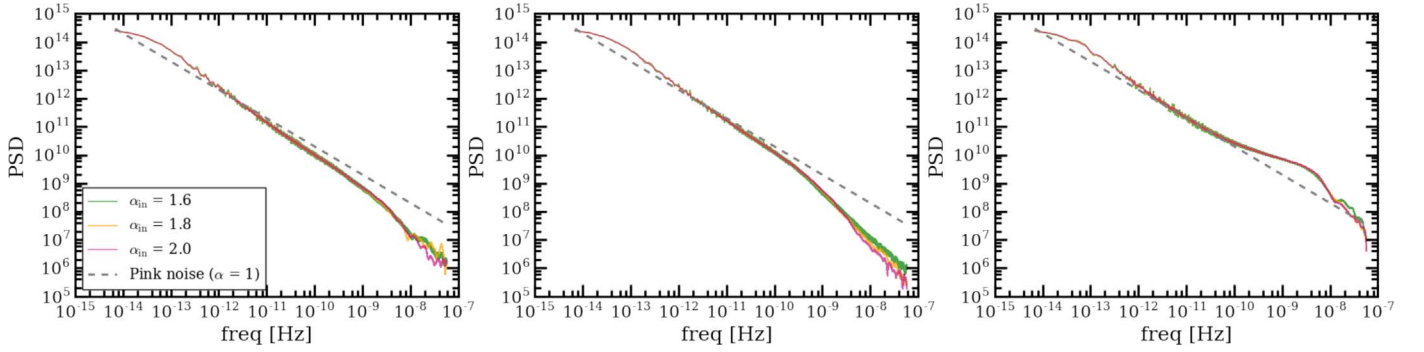
Figure 10. Example of the periodogram "collapse." We show three sets of simulations with broken power-law PDFs and power-law PSDs with slopes of $\alpha = 1.6$, 1.8, and 2.0 (green, orange, and purple, respectively). The three panels correspond to three initial random number generator seeds. The periodograms collapse to a line with negligible scatter and slope close to $\alpha = -1$ (pink noise), the specific shape of which mostly depends on the initial random seed.
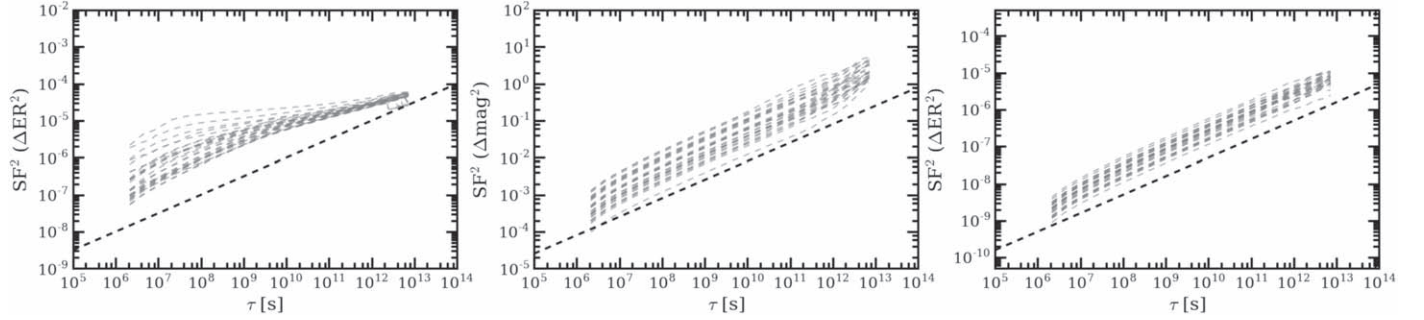
Figure 11. Distortions of the structure function due to the dynamic range of the PDF. Left: SF computed for 30 L/L_{Edd} time series simulated with the same input parameters (gray dashed lines), i.e., a power-law PSD with slope $\alpha = 1.5$ and a broken power-law PDF spanning 6 dex in L/L_{Edd}. The SFs are distorted compared to the predictions of the Wiener–Khinchin theorem (black dashed line, arbitrarily normalized). The distortions are clearly reduced if the SF is computed in magnitude instead of L/L_{Edd} (center), or if the PDF covers a narrower dynamical range (right; here 1 dex).

pronounced for short time series, as the difference between the predicted and observed SF slope is higher at the lowest probed tau τ, independently of the total number of steps (see Figure 9 for TK95 simulations). The main reason for this discrepancy is that the observed and simulated light curves are not infinite and continuous in time, as required by the theorem. In addition, for steep PSDs, we often see a distortion at the highest probed τ even in the representative SF. The main reason of these distortions is that the number of pairs for a given τ decreases with increasing τ, and therefore the SF, which is a statistical measure, cannot be reliably computed for large τ.

The issues discussed strongly illustrate why one should be cautious when considering the highest τ SF measurements of any simulation, as these are essentially unreliable. The actual threshold above which the SF becomes unreliable depends on light-curve sampling and length, as well as on the PSD shape. As a rule of thumb, in our analysis, we are considering SFs only up to $1/10$ of the total light-curve length (e.g., C17). We stress, however, that for the PSD slopes expected from observations (see Section 4), and for ensemble studies, we do not expect our results to be biased by these distortions.

3.2.9. PSD and SF Collapse

For some choices of PSD and PDF sets, the algorithm returns time series with periodograms that "collapse" toward utterly unrealistic PSD shapes. As an example, a PDF as described in W17 with $\lambda_{min} = 10^{-5}$, $\lambda_{max} = 10$ combined with a (broken) power-law PSD with low-frequency slope $\alpha \gtrsim 1.5$ leads to a collapse of the final periodogram to a line with negligible scatter

and a slope close to $\alpha = 1$ (pink noise). This effect is observed for both log-normal and broken power-law PDFs, and the final shape mostly depends on the initial random seed (and not on other parameters such as, e.g., the input slope). An example of this collapse for three simulations with a power-law PDF and three different slopes, $\alpha = 1.6$, 1.8 and 2.0, is given in Figure 10. We note that a distortion of the periodogram at high (low) frequencies is usually reflected in a distortion of the SF at short (long) τ.

As we elaborate below, our tests only showed PSD and SF collapses for simulations setups that considered an overall steep PSD ($\alpha \gtrsim 1.5$; see Figure 10) and/or large dynamical ranges for the values in the considered time series (e.g., $\gtrsim 6$ dex; see Figure 11). The observed collapse in the periodogram, as well as the distortions in the SF, disappears in the following circumstances (see example in Figure 11):

1. The overall slope is shallower. This will very likely always be the case in most of our analysis, because a steep low-frequency end would mean a steep long-timescale SF, and this does not seem to be allowed by the observations;

2. The dynamical range $[\lambda_{min}, \lambda_{max}]$ for the PDF is narrower (e.g., 2 dex instead of 6 dex);

3. We only consider a subsample of the light curve (LC), e.g., 10%. We note that in most cases this also corresponds to a smaller dynamical range (as the time series does not have enough time to traverse the entire PDF); and

4. We consider the SF in magnitude units instead of L/L_{Edd}. We note that, because Δmag $\propto \Delta \log L \propto \Delta \log L/L_{Edd}$ (see

below), working in magnitude instead of L/L_{Edd} results in a reduced dynamical range (i.e., a PDF covering 4 dex in L/L_{Edd} would be mapped to 10 magnitudes).

These conditions suggest that the observed distortions are due to a numerical issue (that is, the LC intrinsically have the right PSD and SF, but we are not able to compute it) instead of being intrinsic to the LC. Specifically, because the SF is computed by looking at the differences between two different points, it is numerically harder to compute it carefully when the dynamical range of the time series is larger (i.e., a roundoff error). In addition, it is possible that some choices of PSD+PDF sets are not allowed for mathematical reasons (as an extreme, a very sharp PDF would not allow almost any variability). Although we are not investigating this issue further, we acknowledge this effect and stress that this issue does not affect our results because the region of parameter space where this sort of collapse occurs is not considered in our analysis. Indeed, as discussed in Sections 2.1.2 and 4.1.2, the overall PSD is expected to break and flatten beyond a given frequency, such that the overall (mean) slope is shallow enough not to cause the collapse (see also Point 1 in the list above). In addition, the comparison to observed data is usually performed in magnitude (i.e., logarithmic) space instead of in linear L/L_{Edd} or luminosity space, which naturally reduces the dynamical range for the values in the considered time series, in a significant way.

3.2.10. Structure Function in Eddington Ratio versus Structure Function in Magnitude

As we mentioned in Section 2.2, the primary outputs of our simulations are L/L_{Edd} time series which, in order to be compared to observations, have to be converted to light curves. Specifically, most of the ensemble studies considered in this work compute SFs in magnitude units, SF_{mag} (e.g., de Vries et al. 2003; Sesar et al. 2006; MacLeod et al. 2010, 2012; C17). A description of how to convert from L/L_{Edd} to luminosity curves is given in Section 2.2. However, in the simplest case where the luminosity $L(t)$ in the considered band is proportional to the bolometric luminosity $L_{bol}(t)$,[18] the magnitude difference needed to compute the SF_{mag} (Equation (14)) can be directly derived from the L/L_{Edd} time series following

$$
\begin{aligned}
m(t_2) - m(t_1) &= -2.5 \log\left[\frac{L(t_2)}{L(t_1)}\right] = -2.5 \log\left[\frac{L_{bol}(t_2)}{L_{bol}(t_1)}\right] \\
&= -2.5 \log\left[\frac{\lambda_{Edd}(t_2)}{\lambda_{Edd}(t_1)}\right],
\end{aligned}
\tag{8}
$$

where $\lambda_{Edd}(t)$ is the Eddington ratio at time t (L/L_{Edd} throughout the text).

By inspecting multiple simulations we found that, if no distortions are present in $SF_{L/L_{Edd}}$ (see Section 3.2.9), then for the same simulation, the SF_{mag} computed following Equations (14) and (8) would have a similar shape, but different normalization, compared to $SF_{L/L_{Edd}}$. On the other hand, for cases where $SF_{L/L_{Edd}}$ shows distortions due to, e.g., a large dynamical range of the PDF (ERDF), these distortions disappear for SF_{mag}. The true underlying reason for this is not yet clear.

4. Applications

In the previous sections, we presented our framework to simulate AGN light curves starting from the distribution of L/L_{Edd} among the galaxy population. In the following, we illustrate some early science applications for the framework and describe how this framework (and, hopefully, similar studies pursued by the community) will become increasingly important to best exploit new observational data and surveys.

We note that the observational data available to date, and in particular the lack of observational constraints on superhuman timescales, do not allow our framework to be fully applied to constrain the underlying PDF+PSD driving AGN variability (or to probe its existence). However, new IFU observations of galaxies with extended AGN photoionized gas (e.g., the Voorwerpjes galaxies; Lintott et al. 2009; Schawinski et al. 2010; Keel et al. 2012a, 2012b; Gagne et al. 2014; Sartori et al. 2016; Keel et al. 2017; Sartori et al. 2018a) will allow us to reconstruct historical AGN light curves (e.g., Treister et al. 2018) and gain new information about AGN variability on $>10^4$ yr timescales. This will be crucial to constrain our models including data at timescales which are still poorly understood. In addition, new time-domain surveys, such as the upcoming LSST, will provide light curves for millions of sources (LSST Science Collaboration et al. 2009) which will allow us to better investigate the variable behavior of single AGNs, as well as compare it to the ensemble behavior of the general AGN population (e.g., MacLeod et al. 2010, 2012; C17).

4.1. Constraining Specific Models with Existing AGN Variability Data

The goal of this section is to provide an example of how existing AGN variability data, in this case the optical SF measurements from the PTF/iPTF survey (Bellm 2014) presented in C17, can be used to constrain the input model parameter space for the kinds of simulations enabled by our framework. Specifically, we want to find a simulation setup such that the derived 4000 Å (R-band) SF is consistent with the observed one, presented in C17.

As described in Section 3.2, the shape and normalization of the simulated SF depend on the input PSD and PDF, while the simulation's length in physical units (T; e.g., yr) also influences the overall normalization (in particular for steep PSDs). Because different normalizations correspond to different variability behaviors, it is crucial to choose physically motivated assumptions for these input parameters in order to obtain realistic SFs to use for our analysis. In addition, observational biases such as flux limits and sample selection can influence the shape and normalization of the observed SF, and therefore have to be included as additional steps in the simulation process.[19] In the following, we outline how to use available observations to define a reference simulation setup consistent with the PTF/iPTF SF and which will be used for other example applications in the next sections.

4.1.1. The PTF/iPTF Structure Function

C17 presented the ensemble 4000 Å SFs for a homogeneous sample of ∼28,000 quasars selected from SDSS-DR7 (Shen et al. 2011) divided into 64 $M_{BH} - z - L_{bol}$ bins (median $z \sim 1.3$,

---

[18] We assume that the BH mass is not increasing significantly during the time probed by the simulations.

[19] We stress that the simulations presented and discussed in the previous sections did not consider sample selection or flux limits.

median log M_{BH} = 9.1; see Figure 4 in C17). In order to obtain a representative SF for the whole PTF/iPTF sample, we divided the SF values C17 in six time bins and took the mean of all SF^2 measurements in that bin (see Figure 13, black points). In the following, we will refer to this representative SF as the PTF/iPTF structure function, $SF_{PTF,C17}$.

4.1.2. The Input PDF and PSD

Following Section 2.1.1, we defined the input PDF as a broken power law with slopes as in W17 (see Section 2.1.1) and $\lambda^* = 0.11$ as expected for $z \sim 1.3$ and the redshift evolution in Caplar et al. (2018). For these simulations, we fixed log $\lambda_{max} = 0.5$ in order to allow for short periods of super-Eddington accretion (and thus, super-Eddington sources). On the other hand, we chose log $\lambda_{min} = -3.75$ such that the active fraction, defined as the fraction of time the AGN spends above $\lambda = 0.01$, is ~10% (this is broadly consistent with observations; e.g., Aird et al. 2018). We stress that different definitions of the active fraction are present in the literature, and that its value is not fully constrained yet, as it may depend on multiple parameters such as black hole mass, luminosity, and redshift.

While the PDF can be motivated by observational constraints (e.g., ERDF, AGN fraction), the shape of the PSD is less well defined (see discussion in Section 2.1.2). However, inspired by the PSD obtained, e.g., in mag or flux units, we assumed a broken power-law shape with slopes and break to be determined through forward modeling, i.e., by comparing the SFs obtained through our suite of simulations to the observed one, $SF_{PTF,C17}$.

4.1.3. Creating Mock Ensemble Structure Functions

For every input PDF and PSD set (see Section 4.1.4 for the assumed inputs, including total simulation length), we ran 300 simulations (i.e., 300 realizations of the same input set) and created a mock ensemble SF to be compared to $SF_{PTF,C17}$ as follows. First, for every simulation, we looked for one subsample with similar length as the PTF survey (assumed here to simply be 2000 days in the observed frame), and observational biases and sample selection similar to the sample selection in C17, following the procedure below:

1. *SDSS selection*: select a random point in the L/L_{Edd} time series and convert L/L_{Edd} to the observed SDSS i-band magnitude assuming $z = 1.3$ and log M_{BH}= 9.1. Specifically, we convert L_{bol} to the observed magnitude by assuming the Vanden Berk et al. (2001) quasar composite UV-optical SED, shifted to $z = 1.3$, sampled at 3000 Å (rest frame), and applying a bolometric correction of $L_{bol}/\lambda L_\lambda(3000 \text{ Å}) = 3.25$ (e.g., Trakhtenbrot & Netzer 2012). Whenever the flux exceeds the SDSS main quasar sample flux limit, $i < 19.1$, we proceed to step (ii); otherwise, we select another random point and repeat the calculation. This step ensures that the mock AGNs would have been selected within the main SDSS quasar sample, and therefore, part of the initial sample used in C17.
2. *PTF sample*: take a 2000 day long subsample starting 5 yr after the SDSS point.[20]

---

[20] We chose 5 yr as representative of the gap between SDSS and PTF observations, but we tested that varying this gap by a couple of years does not significantly affect our results.

3. *PTF masking*: convert L/L_{Edd} to the observed PTF R-band magnitude (following the same conversions as in step (i) above) and apply a mask such that (1) all points with $R < 20$ are observed, (2) all points with $R > 21$ are not observed, and (3) the points with $20 < R < 21$ have a probability of being observed that a linear interpolation in mag between 20 and 21. This step aims to reproduce the observational limits in the PTF sample with a simple prescription.
4. *PTF selection*: if the median of the masked subsample is $R < 19.1$, accept the subsample as a legitimate mock PTF light curve, otherwise go back to step (i). This magnitude cut reproduces the specific sample selection applied in C17. We note that at the considered redshift ($z = 1.3$) and BH mass (log ($M_{BH}/M_\odot$) = 9.1), the $R = 19.1$ magnitude cut corresponds to an L/L_{Edd} cut of roughly log (L/L_{Edd}) = -1.2, which is consistent with that seen in the C17 sample. As an illustration, Figure 12 shows the regions above the PTF limit for 100 of the 300 simulations, as well as the position of a possible PTF region.

We then computed the ensemble observed SF, in magnitudes, from the 300 mock PTF light curves (i.e., observed R-band magnitudes), and converted it to intrinsic 4000 Å SF using Equation (4) in C17. For every choice of input parameter set (i.e., every PDF and PSD set), we repeated this procedure 500 times (i.e., we computed 500 ensemble SFs, each time using different subsamples), and computed the representative SF_{PTF} following Equation (6). We note that this is essentially equivalent to computing the ensemble SF for $300 \times 500 = 150,000$ AGNs.

4.1.4. The Reference Simulation

In principle, what was described in Section 4.1.3 should be repeated following an MCMC approach by stepping over a grid of input parameters (i.e., PSD break and slopes, and simulation length), in order to find the best possible model for the data. However, as pointed out above, the current data are not sufficient to fully constrain the PSD, especially due to the lack of measurements beyond human timescales. Therefore, for now we can only use a forward modeling approach to test if a given set of input parameters is consistent with the observed ensemble analysis. For the PDF described in Section 4.1.2, we found that a PSD with $\alpha_{low} = 1$, $\alpha_{high} = 2$, and $f_{br} = 1 \times 10^{-9}$ (~30 yr), and a simulation length of 1.4 Myr is consistent with the observations, as shown in Figure 13. The reference simulation setup is summarized in Table 2 and will be used in the next sections.

We stress that timescales on the order of the PSD break (~30 yr in intrinsic frame) or longer are not directly probed by the subsamples used for the ensemble analysis, as the iPTF/PTF data span only 5.5 yr. However, they are probed by the total light curves from which the subsamples are selected. Therefore, they indirectly also affect the timescales probed in the ensemble analysis. Specifically, in this case, we saw that the break position influences the SF normalization at the timescales that are probed by the observed data. In addition, we notice that a simulation length of ~1 Myr is long enough for the AGN to realistically span the whole ERDF, but advanced computational setup allowing simulations with $>10^8$ points (see Section 3.2.3) will allow even longer timescales to be probed.
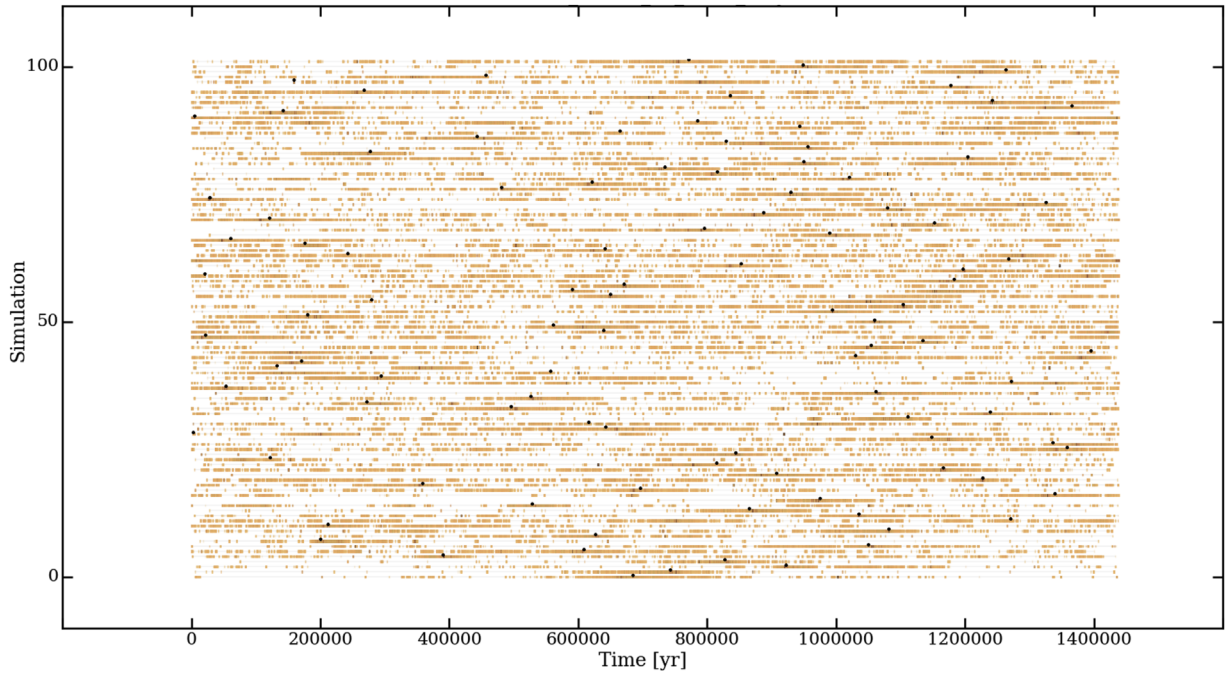
Figure 12. Identifying the PTF quasar sample in simulated light curves. This figure illustrates the regions above the PTF limit for 100 of the 300 reference simulations (orange; see Section 4.1.4), as well as the position of a possible PTF-observed region, selected following the procedure outlined in Section 4.1.3 (black points). Only ~1.7% of the points in every light curve meet the PTF selection (mainly the SDSS+PTF flux limits). This fraction may appear to be higher in this illustration due to the limited image resolution (see, e.g., Figure 15, top panel). Evidently, there is no preferential position within the ~1.4 Myr covered by the simulation where the the points that could be associated with the PTF quasars sample are clustered.

We also note that a break position of ~30 yr is longer compared to the breaks reported in the literature for both ensemble SF or PSD, which are on the order of several years (e.g., MacLeod et al. 2012; Simm et al. 2016 C17). However, as discussed in Kozłowski (2017b), break timescales directly computed from observed SFs may be considerably underestimated, due to the finite length of the AGN light curves, which is not sufficient to fully probe the variability behavior at timescales longer than the claimed break (that is, the surveys do not satisfy $T \gg \tau_{br}$). Moreover, PSD analyses almost always refer to individual AGNs and find break frequencies that vary from source to source (e.g., Simm et al. 2016). Such short timescale breaks could therefore be "washed out" when averaged over an ensemble of sources and could therefore not be robustly identified in the intrinsic PSD of the AGN population. Finally, we stress that the observed SF and PSD always refers to specific samples (e.g., with a magnitude limit), such that the observed variability features do not always correspond to the intrinsic ones. As an example, Figure 13 also shows a comparison between the simulated ensemble SF and the total SF obtained from the same simulations without applying any flux cut or sample selection (i.e., the intrinsic SF). The ensemble SF obtained in Section 4.1.3 has a lower normalization compared to the intrinsic one, suggesting that the sample selected in PTF/iPTF has an overall lower variability compared to the total sample of AGNs at the same redshift and BH mass. Because the sample selection mostly relies on the AGN luminosity, including only the most luminous sources (see Section 4.1.3), this effect is consistent with works suggesting the variability to be inversely proportional to luminosity and L/L_{Edd} (e.g., MacLeod et al. 2010; C17, Rumbaugh et al. 2018).

4.2. Simulating Future Time-domain Surveys

With the advance of large, multiwavelength time-domain surveys such as LSST (Ivezic et al. 2008; LSST Science Collaboration et al. 2009), TDSS (Morganson et al. 2015), ZTF (Bellm 2014), and *eROSITA* (Merloni et al. 2012), we are entering an era of time-domain astronomy that will allow us to probe the variable universe with an unprecedented time resolution and time span for very large statistical samples (~millions of objects). Thanks to the high observing cadence, these surveys will allow us to probe for the first time the variability properties of individual AGNs in a large sample and to compare them to what was found in ensemble analyses, therefore gaining new insights into the diversity of variability behaviors among the galaxy population.

Based on current observed light curves, which show a variety of features among different AGNs (e.g., Smith et al. 2018), one would expect the light curves from the new surveys to be very diverse. In our framework, this diversity arises from the fact that (1) every AGN corresponds to a different realization of the underlying statistical process, and (2) that the observed time span (~decades) is significantly shorter than the AGN life, and therefore only probes a small part of it. Indeed, we assume that the AGN variability behavior is (statistically) the same for every AGN, only "integrated" over the time needed for the AGN to span the entire ERDF, which for our reference simulation is on the order of $\sim10^6$ yr. As a consequence, peculiar variability features lasting only for short times (i.e., rare events) will be observed only in a few galaxies, although they may be present in all of them.

To illustrate this point, Figure 14 shows some example AGN light curves that may be expected to arise from LSST assuming the reference simulation setup defined in Section 4.1. First, we simulated a $\sim10^6$ yr long (intrinsic frame) L/L_{Edd} time series
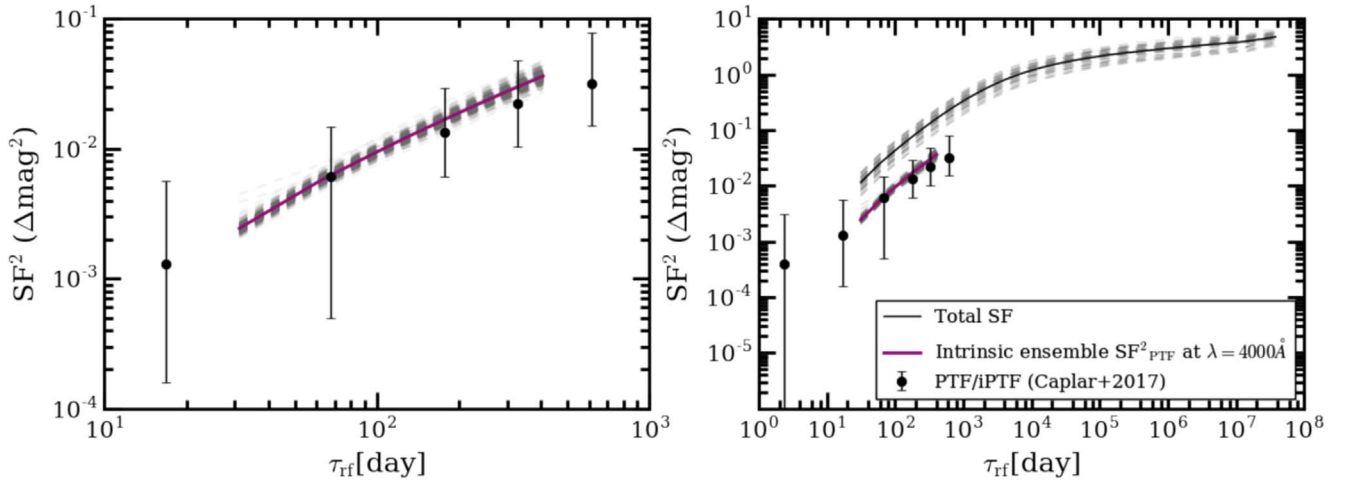
Figure 13. Matching the observed PTF/iPTF structure function with our reference simulation. Left: black points denote the ensemble PTF/iPTF SF, taken from C17; the gray lines trace 500 ensemble SFs from light curves obtained with the reference simulation setup. The indigo line traces the final, representative ensemble SF for the best model, computed from the 500 intrinsic ones. Right: same as left panel, but showing also the total SF obtained directly from the best model simulations without applying any cut or sample selection (black line). We note that the ensemble SF obtained as discussed in Section 4.1.3 has a lower normalization compared to the intrinsic one.

Table 2
Summary of Reference Simulation Setup for a Sample with $z \sim 1.3$,
log $M_{BH} = 9.1$, Consistent with the PTF/iPTF Sample in C17

PDF	Broken power law (Equation (1))
	$\lambda^* = 0.11$
	$\delta_1 = 0.47$
	$\delta_2 = 2.53$
	log $\lambda_{min} = -3.75$
	log $\lambda_{max} = 0.5$
PSD	Broken power law (Equation (3))
	$\alpha_{low} = 1$
	$\alpha_{high} = 2$
	$f_{br} = 10^{-9}$ (~ 30 yr)
Simulation length	$\sim 1.4 \times 10^6$ yr
Eddington ratio cut	log $\lambda_{cut} = -1.2$

based on the reference PSD and PDF (from Section 4.1) with a nine-day resolution in observed frame (3.9 days in intrinsic frame at $z \sim 1.3$), which is comparable to the $\sim$weekly LSST resolution. We then converted it to bolometric luminosity and LSST i-band magnitude in a way similar to that discussed in Section 4.1 and extracted 10 yr long "snapshots" to mimic LSST light curves. As is clearly visible in Figure 14, every short light curve, i.e., every mock observed time series, shows a different variability behavior. Specifically, the example shown in Figure 14 illustrates well how the AGN can experience a variety of variability features, such as "switch on," "switch off," and a sharp, year-long "flare," as well as periods of much more subtle variability (see the different panels in Figure 14). Naturally, not all of these features will be observed in every AGN because of the limited observation length, as well as magnitude limits and host galaxy contamination (as will be discussed below). This framework therefore allows the occurrence rate of specific variability features to be investigated by considering both intrinsic AGN behavior as well as observational limitations (e.g., the fact that we are observing every AGN only for a short time). In addition, the framework proposed in this paper will be highly valuable to constrain the true (intrinsic) AGN variability from artifacts and distortions,

due to the fact that we are not continuously probing the whole AGN light curve (see Section 3.2), as well as other observational biases.

One additional advantage of the new and upcoming surveys is the deeper flux limit, which will allow the variability properties of fainter AGNs to be detected and investigated. As an example, LSST will reach a magnitude limit $i = 24$,[21] which is almost 5 magnitudes deeper than SDSS and PTF. Given these deep flux limits, it is, however, important to consider the possible effect of host galaxy contamination on the observed AGN light curves. To illustrate this, we again simulated a $\sim 10^6$ yr long (intrinsic frame) L/L_{Edd} time series as described above (Figure 15 top, gray light curve). We then computed the i-band magnitude for a galaxy at $z \sim 1.3$ hosting a log $M_{BH} = 9.1$ BH (as in the PTF sample) by assuming the Bruzual & Charlot templates at 1 Gyr and solar metallicity (Bruzual & Charlot 2003), initial mass function (IMF) from Chabrier (2003), and the M_{BH}–M_{bulge} relation from Kormendy & Ho (2013). This gives a host galaxy magnitude $i = 22.2$. The green light curve in Figure 15 (top), which considers both AGN and galaxy contributions, corresponds to the total light curves. The magnitude change due to host galaxy contribution is ~ 0.06 mag for sources at the PTF limit and ~ 2 mag at the LSST limit. Adding the galaxy contribution therefore causes an increase in total flux of $\sim 6\%$ at the PTF limit and of a factor $\sim 6\times$ at the LSST limit. Although the effect is almost negligible for PTF, it should be considered when analyzing faint LSST observations.

4.3. Extremely Variable Quasars and Changing-look AGNs

In the last years, quasars exhibiting extreme variability behavior (i.e., large magnitude changes over months–years timescales) have been found in increasingly high numbers, both serendipitously (e.g., LaMassa et al. 2015; Husemann et al. 2016; McElroy et al. 2016; Runnoe et al. 2016; Katebi et al. 2018; Ross et al. 2018; Stern et al. 2018; Trakhtenbrot et al. 2019) or through systematic searches in surveys such as

---

[21] This corresponds to the 5σ detection limit for a point source in a single visit (Kahn 2018).
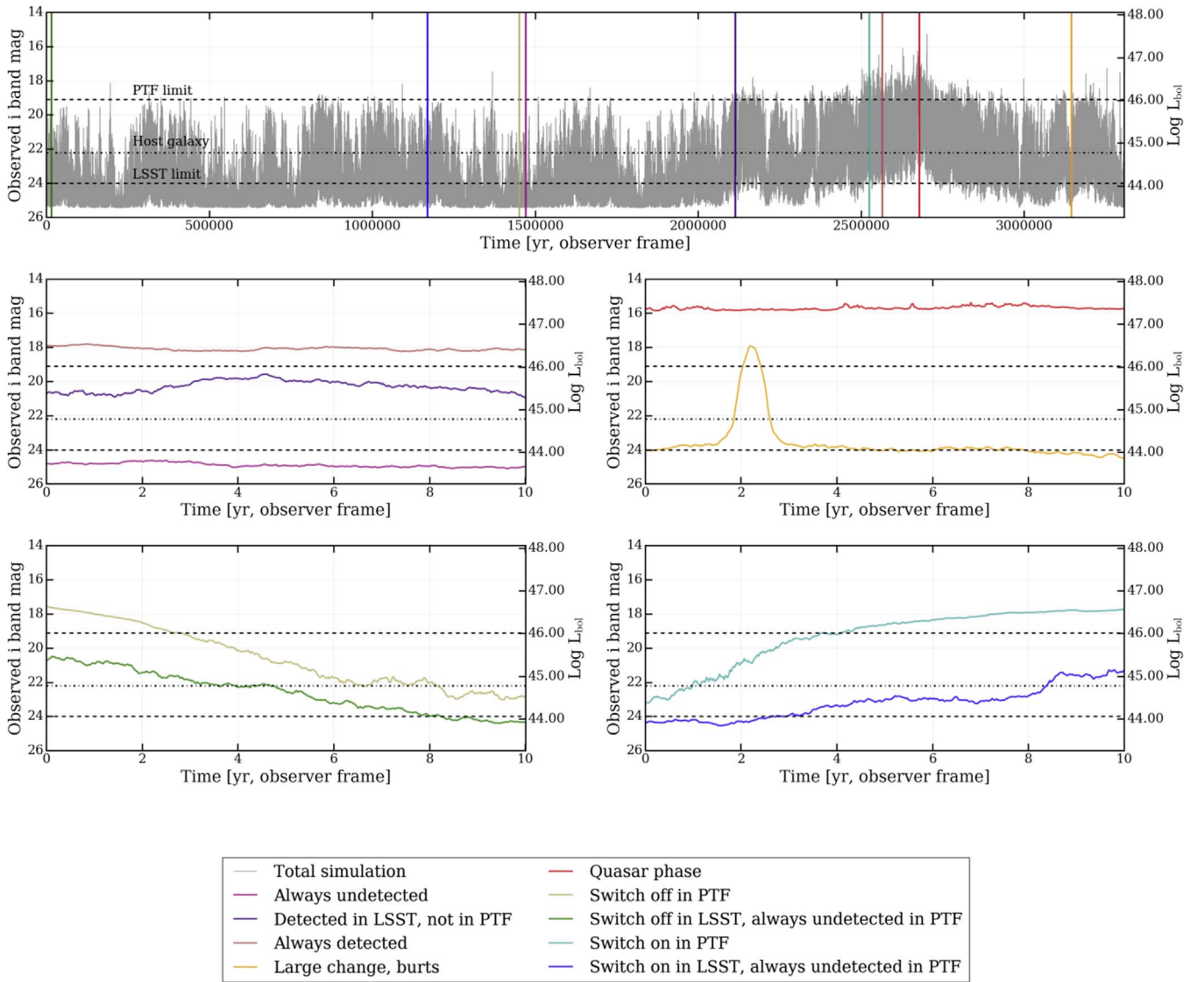
Figure 14. Simulating AGN variability as observed within the LSST. The top panel shows an example AGN light curve, simulated using the reference setup (described in Section 4.1), and a resolution of nine days in the observed frame—comparable to the LSST cadence. The horizontal lines mark the PTF and LSST flux limits (dashed lines) and the fiducial host galaxy emission (dashed–dotted line). In the panels below, we show several 10 yr long "snapshots" (in the observed frame) which mimic AGN light curves that may possibly be observed within the LSST main survey. The positions within the total simulated light curve (upper panel) at which the different "snapshots" are extracted from are marked with vertical lines of the same color. This figure illustrates how observing AGNs at different moments during their life can show significantly different variability features, namely, "switch on," "switch off," and "flaring" events, in addition to much more subtle AGN variability.

the SDSS (MacLeod et al. 2016, 2019; Ruan et al. 2016; Rumbaugh et al. 2018), PanSTARRS-1 (Lawrence et al. 2016), or CRTS (Graham et al. 2017). One example of extreme variability is observed in CL-AGNs which, in addition to changes in emission line structure, show changes in luminosity of an order of magnitude or more on years–decades timescales. Future time-domain surveys will allow these extreme variable objects to be found and followed up in a systematic way.

Although multiple explanations for extreme variability have been put forward (e.g., Graham et al. 2017; Lawrence 2018; Ross et al. 2018; Rumbaugh et al. 2018; Stern et al. 2018), it is not yet clear if these sources are driven by physical processes that are distinct from what drives "normal" AGN variability, or if their extreme variable behavior is simply the tail of a continuous distribution characterizing the overall AGN variability phenomenon (see, e.g., Sartori et al. 2018b; MacLeod et al. 2019). Our

framework provides an alternative way to address these questions by comparing the observed numbers (or rates) of variable sources with the prediction from our simulations. Indeed, our framework allows the probability of observing a given magnitude change to be quantified given a model for the underlying AGN variability (i.e., the input of the simulations) and specifics of the survey (such as flux limit, cadence, and length).

As an illustration, Figure 16 shows the probability P of an AGN undergoing a change in magnitude $|\Delta m| > \Delta m_{ref}$ between two epochs separated by τ, within our reference simulation (as specified in Section 4.1). Specifically, for a given simulated light curve, $P(|\Delta m_{obs}| > \Delta m_{ref}$ is defined as the ratio of the number of pairs separated by τ with $|\Delta m_{obs}| > \Delta m_{ref}$ to the total number of pairs separated by τ. As expected, the probability of undergoing a large magnitude change drops significantly with increasing Δm_{ref}, but increases
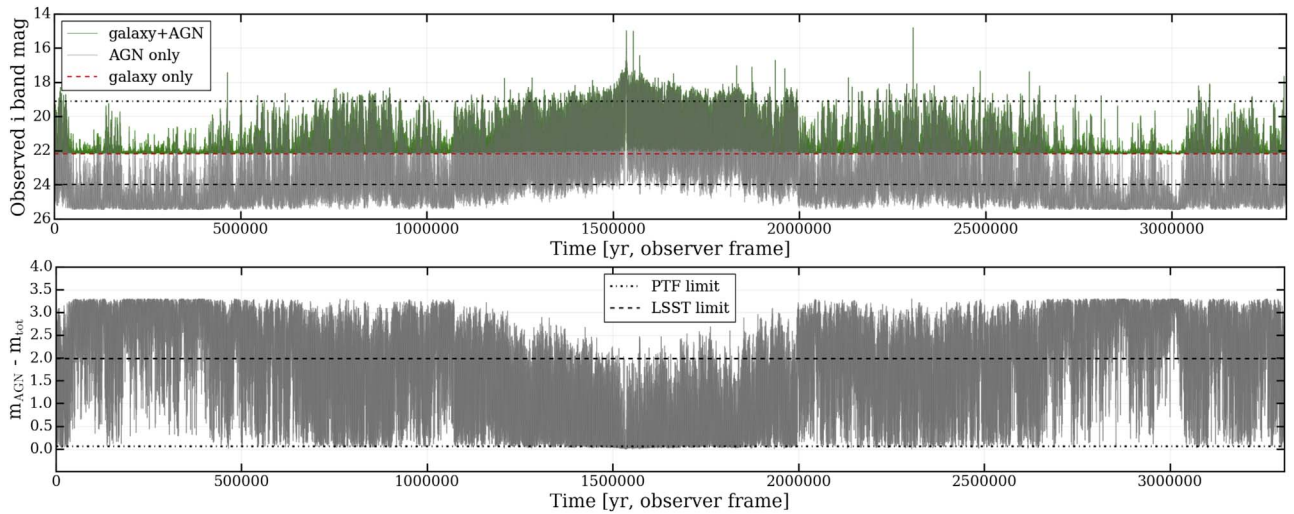
Figure 15. The effect of host galaxy contamination on realistic AGN light-curve simulations. Top: in gray we show the same total AGN light curve as in the top panel of Figure 14, but for a different realization of the simulation (i.e., identical input parameters). This light curve corresponds solely to the AGN emission. In green we show the light curve derived when adding the contribution of the host galaxy (at $i \simeq 22$; see text for details). Because the host is fainter than the PTF flux limit, the magnitude change due to host light is only ~0.06 mag for sources above the PTF flux limit (i.e., $i \lesssim 19.1$). However, it grows to ~2 mag for the LSST limit ($i < 24$). Bottom: the magnitude difference due to the addition of the galaxy contribution.

with increasing τ.[22] This is a direct consequence of the assumed general PSD shape, which gives more power to low frequencies (long τ).

Because the PDF assumed in this example spans the range $\log (L/L_{\mathrm{Edd}}) \in [-3.75, 0.5]$, the maximum possible $|\Delta m|$ that can be found in the simulated light curves is $|\Delta m| \sim 10.6$ mag, which corresponds to a change in luminosity of more than four orders of magnitude (4.24 dex). As can be clearly seen in Figure 16, such extreme variability can be found at all timescales probed by our simulations, although the probability for such changes is extremely low. As an example, the probability for the simulated AGN to undergo a magnitude change $|\Delta m| > 10$ mag within $\tau = 1$ yr is $P \sim 10^{-8}$. Because such large $|\Delta m|$ over such short timescales are very unlikely to be physically explainable (see, e.g., Katebi et al. 2018 and references therein for a discussion of possible timescales linked to AGN variability), this feature could arise from the fact that the assumed model (both for the PSD and for the L/L_{Edd} to magnitude conversion) is too simplistic, as the SED shape could change drastically (Trakhtenbrot et al. 2019). However, the probability of these features is low enough that it does not affect the main results presented in this paper, and we only acknowledge this issue to be investigated in future studies.

It is important to stress that Figure 16 shows the probability of a simulated AGN undergoing a certain change in emission, which does not directly correspond to the probability of observing or detecting such a change. In fact, given the flux limit of any reasonable considered survey, combined with the contamination from the host galaxy, low L/L_{Edd} sources are (usually) not observed (or not identified as AGNs), which limits the range of $|\Delta m|$ that may be observable.

To address this limitation, in Figure 17 we show a similar analysis to that shown in Figure 16, but now assuming the different magnitude limits discussed in the previous sections:

the PTF magnitude limit $r = 19.1$, the LSST magnitude limit $i = 24$, as well as the host galaxy contamination, at $i = 22.2$. Specifically, for a given survey depth (magnitude limit) m_{cut}, we first substitute all of the points with $m > m_{\mathrm{cut}}$ with an upper magnitude limit $m_{\mathrm{ul}} = m_{\mathrm{cut}}$, and then compute the probabilities as in the previous test, although now considering only the pairs of data points in which at least one point is above the detection limit. As expected, the probability of detecting large magnitude changes over a given timescale decreases for higher flux limits (shallower surveys). As an example, for the considered simulation, the probability of detecting a magnitude change of $|\Delta m| = 2.5$ (1 dex change in luminosity) over 10 yr is $P \sim 10^{-2}$ assuming the LSST depth, and $P \sim 10^{-4}$ for the PTF. We stress that the probabilities showed in Figures 16 and 17 are not directly comparable because they are computed using different assumptions for the total number of considered pairs (total number of pairs versus number of pairs where at least one point is detected), and because Figure 17 also includes upper limits.

If our hypothesis of an underlying common PSD+PDF set is correct and the reference simulation defined in Section 4.1 realistically describes the AGN population, then, for sources with redshifts and BH masses comparable to the PTF quasar sample, we would expect to detect a level of variability in LSST consistent with the predictions from our simulations (after accounting for sample selection biases in a similar way to what is discussed in Section 4.1). Finding different values would, however, also be informative. For example, detecting a lower than expected number of highly variable quasars may indicate that L/L_{Edd} changes do not directly translate to luminosity changes, but the variability is suppressed due, e.g., to a luminosity-dependent bolometric correction (that is, a luminosity-dependent SED shape; see, e.g., Vignali et al. 2003; Marconi et al. 2004; Just et al. 2007), and to the fact that the AGN may become radiatively inefficient at low L/L_{Edd} (Narayan et al. 1998; Panessa et al. 2006; Ho 2009). On the other hand, a higher detection rate may be expected if additional external processes, such as microlensing (e.g., Lawrence et al. 2016), dust obscuration (e.g., Maiolino et al. 2010; Ricci et al. 2016), or

---

[22] We caution that the trend observed at the highest probed Δm_{ref}, where the probability appears to decrease for $\tau > 30$ yr, may be due to the fact that the number of corresponding pairs within the simulated light curves is not large enough to reliably sample this low-probability regime.
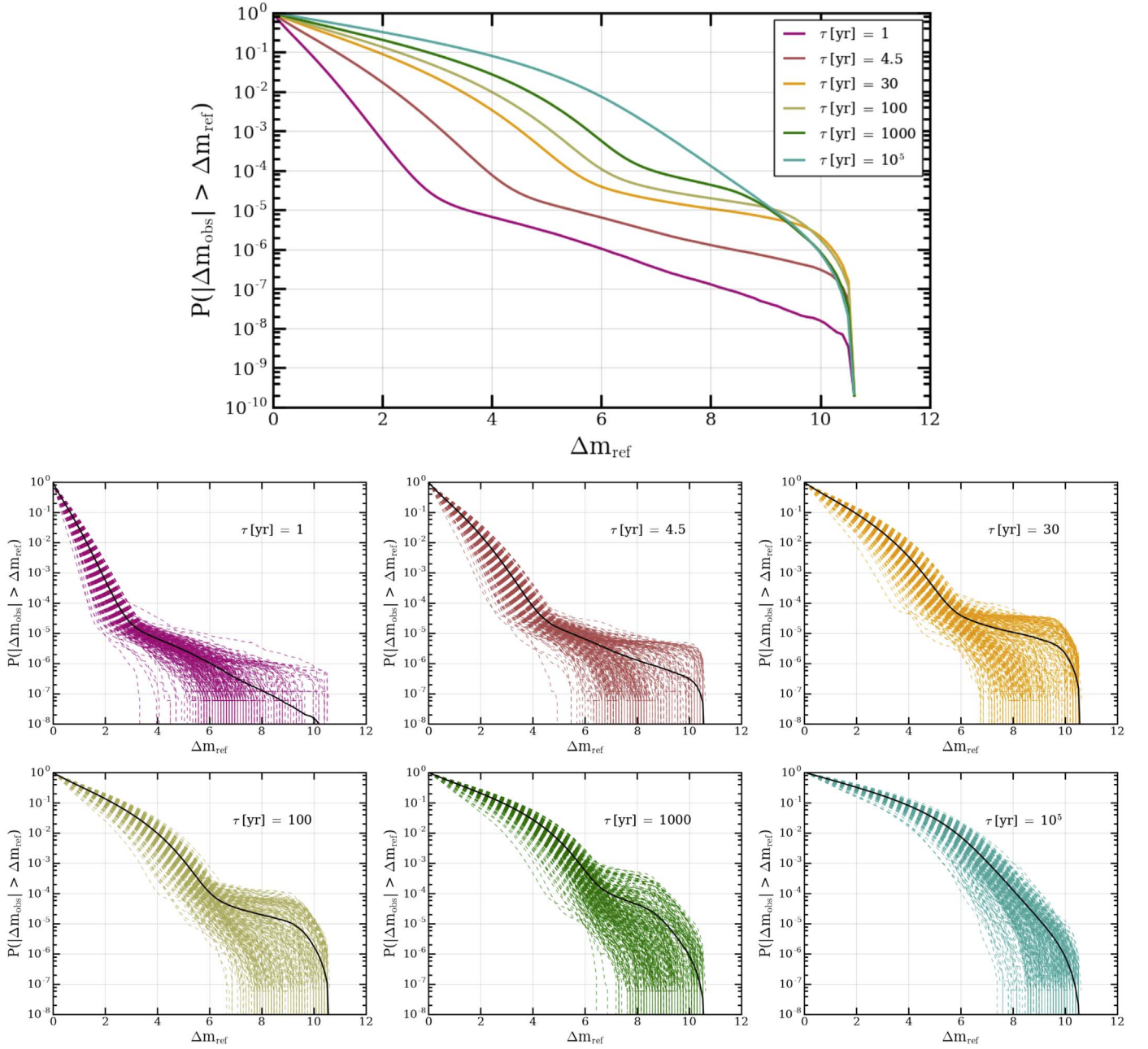
Figure 16. Expected distributions of AGN flux variability. We show the calculated probability of a simulated quasar, drawn from our reference simulation, to undergo a change in magnitude $|\Delta m| > \Delta m_{ref}$ between two epochs separated by τ (in the intrinsic/rest frame; see Sections 4.1 and 4.2). The small panels below show the analysis for the different τ separately, where the colored lines are the different realizations and the black lines show the mean trends. The large panel summarizes the mean trends for all considered τ. We note that at the redshift assumed for the reference simulation, $z = 1.3$, an intrinsic $\tau_{int} = 4.5$ yr corresponds to an observed $\tau_{obs} \sim 10$ yr, i.e., the foreseen length of the LSST main survey. Because in general the probability of the AGN undergoing a given $|\Delta m|$ decreases with decreasing τ, for every possible τ probed by the LSST, the probability of the AGN undergoing a given $|\Delta m|$ is lower than that found for $\tau_{int} = 4.5$ yr.

supernovae also contribute to the observed luminosity changes. Other possible explanations of high variability proposed in the literature are stellar-mass binary black hole mergers within the dense gas of the AGN accretion disk (e.g., McKernan et al. 2014; Bartos et al. 2017) or TDEs (e.g., Gezari et al. 2012; Arcavi et al. 2014; Hung et al. 2017; Auchettl et al. 2018). Although these events may contribute to the instantaneous emission changes seen in the concerned galaxy, their appearance is not expected to have the same probability in every galaxy (see e.g., Arcavi et al. 2014; French et al. 2016; Law-Smith et al. 2017 for the host galaxy dependence of TDEs). Therefore, such

bursts of "external" origin are not expected to be present in our simulations, which aim to reproduce the intrinsic stochastic AGN behavior only.

Systematic searches for CL-AGNs require large multiepoch spectroscopic surveys and elaborate spectral fitting analysis to determine the changes in emission line structure (e.g., MacLeod et al. 2016). Therefore, current studies aimed at systematically searching for highly variable objects are concentrating on changes in the overall luminosity. As an example, starting from a sample of $\sim$900,000 quasars in CRTS, Graham et al. (2017) found 51 sources with $|\Delta m| > 1$ (V-band) over $\sim$900 days,
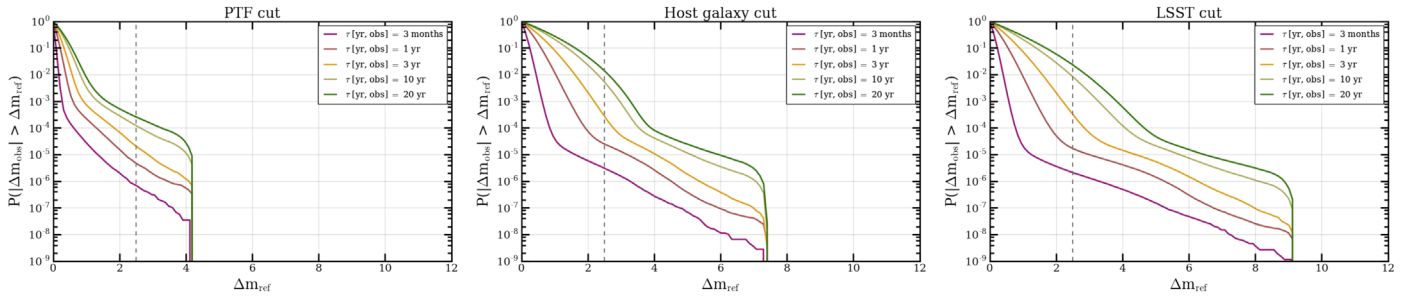
Figure 17. Similar to Figure 16, but here τ is given in the observer frame, and taking into account different magnitude cuts (left to right: the PTF flux limit, the host galaxy contribution, and the LSST limit). The gray dashed lines mark a change of $\Delta m_{ref} = 2.5$ (i.e., a $10\times$ change in flux or luminosity).

which corresponds to a fraction of $\sim 5 \times 10^{-5} = 0.005\%$. On the other hand, Rumbaugh et al. (2018) detected a variability of $|\Delta m| > 1$ (g band) over a 15 yr period in $\sim 10\%$ of sources among a quasar sample drawn from SDSS and DES (Flaugher et al. 2015), and further estimate that this fraction would increase to $\sim 30\%$–50% accounting for selection effects. Furthermore, Lawrence et al. (2016) performed a joint analysis of SDSS and PanSTARRS-1 and estimated the fraction of AGNs showing $|\Delta m| > 1.5$ in the g band to lie between 10^{-5}–10^{-3}. A direct comparison between the results of these studies is impossible, due to the different sample selection criteria and the analysis methodology (see Rumbaugh et al. 2018 for details). Similarly, a comparison between these observational results and the predictions from simulations would require a whole analysis aimed at reproducing the sample selection and observational biases in a similar way to that showed in Section 4.1 for PTF, which is beyond the scope of this paper. We stress that the proposed framework will be very valuable to test if the results from different studies can be reproduced starting from the same simple model, i.e., if the observed differences are only attributable to different sample selections and analysis methods, or if different works are finding intrinsically different classes of objects. Indeed, this framework allows us to start from the intrinsic AGN variability predicted by simple models, and forward model all observational effects which may be affecting the observed variability.

It is also important to note that, although CL-AGNs usually show a change in the overall luminosity $|\Delta m| > 1$, not all highly variable sources also present changes in the line structure. For example, MacLeod et al. (2019) report that only 10%–50% of the quasars with $|\Delta g| > 1$ in the SDSS and PanSTARRS-1 surveys also show CL-AGN behavior in the optical spectrum. By combining upcoming photometric and spectroscopic time-domain surveys such as LSST and SDSS-V, it will be possible to better constrain this fraction. Together with the predictions from the simulations enabled by our framework, this will allow the relation between changes in accretion rate and changes in the broad line structure, and in general of the AGN structure, to be investigated, therefore gaining new insights into the interplay between the accretion disk and broad line regions (e.g., Marin 2017).

5. Summary and Conclusions

We presented a framework designed to model AGN variability over a broad range of timescales and in different objects, based on the observed Eddington ratio distribution among the AGN population. The framework was initially proposed in Sartori et al. (2018b). In this paper, we discussed in detail the fundamental assumptions of the framework, its

implementation using GPU architecture, and various simulations that test and characterize both the intrinsic and numerical behavior of our implementation. We then demonstrate several possible applications of the framework, based on the observed variability in the PTF/iPTF survey, with the aim to interpret the light curves expected to be observed with the LSST and the prospects of discovering extremely variable AGNs in upcoming time-domain surveys. An up-to-date version of the simulations code, along with examples and additional information about its usage, can be found at https://github.com/nevencaplar/AGN-Variability-Simulations.

The central points of our framework can be summarized as follows (see Section 2 for details):

1. We propose that AGN light curves can be modeled starting from the distribution of L/L_{Edd} among the galaxy population. Specifically, we assume that the distribution of L/L_{Edd} covered by an AGN during its lifetime is consistent with the ERDF of the galaxy population, and that the variability behavior is described by a PSD that is similar for all AGNs.
2. This implies that the variety of AGN variability features observed in different AGNs arises from the fact that every AGN is a different realization of the same underlying process (here described by the PSD and the ERDF), and that the observed period is much shorter than the AGN lifetime.[23]
3. The primary outputs of our simulations are L/L_{Edd} time series, which can be converted to light curves and other observables in post-processing. This requires taking into account both the AGN physics (e.g., different bolometric corrections) as well as properties of the observations (e.g., flux limits and sample selection).

In order to perform the simulations, we implemented the iteration algorithm from E13 to run on a GPU architecture (Section 3):

1. Thanks to the CUDA environment, our GPU implementation allows the performance both in terms of length of the simulation (up to $\sim 10^8$ points) and speed ($\sim$minutes for the longest simulations) to be maximized.
2. The code is optimized for broken power-law or log-normal PDFs and broken power-law PSDs; however, different functional forms can be added.
3. By construction, the PDF and the periodogram of the final light curves are consistent with the input PDF and PSD. On the other hand, the normalization of the

---

[23] This can be thought of as following the ergodic hypothesis for test particles.

corresponding SF depends on the physical length (e.g., in yr) of the simulated light curve (Equation (5)).

As an example application, we looked for a model (i.e., input set) consistent with the variability behavior observed in the PTF/iPTF sample (C17) and used it to simulate possible light curves, of the kind that are expected to be observed with LSST (Section 4):

1. After taking into account sample selection and flux limit, we found that the PTF/iPTF ensemble SF is consistent with simulations performed assuming the ERDF proposed in W17 and C17, and a broken power-law PSD characterized by a random walk at low frequencies and a break at ∼30 yr. In this case, the time needed for the AGN to span the entire ERDF is ∼1.4 Myr (Table 2).
2. We showed that the ensemble SF computed by including only light-curve points above the PTF flux limit has a lower normalization (and therefore a lower overall variability) compared to the SF computed by considering the entire simulated light curve (Figure 13). This is consistent with observations showing that the amplitude of variability is anticorrelated with the AGN luminosity and/or Eddington ratio (e.g., MacLeod et al. 2010; C17, Rumbaugh et al. 2018).
3. The simulated LSST light curves show that the same AGN can experience a variety of rather extreme variability features such as "switch on," "switch off," and year-long "flares," in addition to prolonged periods of much more subtle variability (Figure 14). Naturally, because the foreseen observations span much shorter periods that the AGN lifetime, not all of these features are observed for each AGN.
4. We expect that the framework presented here would become highly valuable to understand whether any observed variability feature is consistent with the general behavior of the AGN population (i.e., if it can be reproduced with our general model), and to constrain the intrinsic variability from artifacts due to, e.g., sample selections and observational cadence.
5. Our framework provides a forward modeling approach that can be used to compare results of different time-domain surveys, for "normal" AGN variability, extreme variability events, and, with suitable adaptations, for the variability of other astrophysical sources.

Appendix
PSD and Structure Function

A time series corresponds to a single realization of an underlying PSD and probability density function (PDF). Given a time series $x(t_k)$, $k = 1, 2, ..., N$, the statistical estimator of the underlying PSD is called a periodogram $P(f_j)$ and is computed based on Fourier analysis.

Let us assume a time series $x(t_k)$, $k = 1, 2, ..., N$, with N_{steps} equidistant observations with sampling period $t_k - t_{k-1} = t_{bin}$. Depending on the study, the quantity x can correspond e.g., to luminosity, magnitude, counts rate or Eddington ratio. The DFT of $x(t_k)$ is given by

$$\text{DFT}(f_j) = \sum_{k=1}^{N} x(t_k) e^{2\pi i(k-1)j/N}, \tag{9}$$

where f_j, $j = 0, 1, ..., N-1$, are the Fourier frequencies which depend on the parity of N_{steps}:

N_{steps} even

1. $f_j^+ = j/(Nt_{bin})$ for $j = 1, ..., N/2 - 1$
2. $f_{N/2} = f_{Nyq} = 1/(2t_{bin})$ for $j = N/2$
3. $f_j^- = -(N - j)/(Nt_{bin})$ for $j = N/2 + 1, ..., N - 1$

N_{steps} odd

1. $f_j^+ = j/(Nt_{bin})$ for $j = 1, ..., (N - 1)/2$
2. $f_j^- = -(N - j)/(Nt_{bin})$ for $j = (N + 1)/2, ..., N - 1$

The periodogram $P(f_j)$ is then computed from the DFT (up to a normalization constant depending on the convention used):

$$P(f_j) = \frac{2}{N^2}\{Re[DFT(f_j)]^2 + Im[DFT(f_j)]^2\}, \tag{10}$$

where $j = 0, 1, l ..., N/2$ for even N_{steps} and $j = 0, 1, ..., (N - 1)/2$ for odd N_{steps}. The units of the PSD, and consequently of its estimator $P(f_j)$, are (power/Hz), so that the integral of the PSD is proportional to the total power in the process:

$$\text{total power} \propto \int_0^\infty \text{PSD}(f)df. \tag{11}$$

For a given frequency f_i, $\text{PSD}(f_i)$ is therefore proportional to the power of variations with frequency f_i.

One of the most widely used PSD normalizations is the fractional rms normalization defined by van der Klis (1997; see also Miyamoto et al. 1991; Vaughan et al. 2003):

$$P_{rms}(f_j) = \frac{2t_{bin}}{\mu^2 N}\{Re[DFT(f_j)]^2 + Im[DFT(f_j)]^2\}, \tag{12}$$

where μ is the mean value of the time series. With this normalization, the integral of the underlying PSD_{rms} between two frequencies f_i and f_j, $i < j$ corresponds to the contribution of the frequency window $[f_i, f_j]$ to the total rms squared variability σ^2/μ^2. The total rms squared variability can be obtained by integrating between f_1 and f_{max}, where $f_{max} = f_{N/2} = f_{Nyq}$ for even N_{steps} and $f_{max} = f_{(N - 1)/2}$ for odd N_{steps}:

$$\frac{\sigma^2}{\mu^2} = \int_{f_1}^{f_{max}} \text{PSD}_{rms}(f)df. \tag{13}$$

An alternative way to characterize the variability in a given time series is to construct the SF, which quantifies the amount of variability in the data for a given timescale. We stress that different definitions are present in the literature (e.g., MacLeod et al. 2012; Kozłowski 2016), so that caution has to be used when comparing the SFs obtained in different studies. In this work, for a time series $x(t_k)$ as defined above and a given time

lag τ, we adopt the following SF formulation:

$$\mathrm{SF}(\tau)^2 = \frac{1}{P} \sum_{i,j>i} [x(t_i) - x(t_j)]^2 = \langle [x(t) - x(t+\tau)]^2 \rangle,$$

(14)

where $t_j - t_i = \tau$ and P is the number of $\{x(t_i), x(t_j)\}$ pairs, mostly in L/L_{Edd} or mag units ($x(t_k)$ are the observed or simulated L/L_{Edd} and light curves).

In the case of a zero mean stationary time series with f varying between 0 and ∞, a power-law PSD with slope $-\alpha$, $1 < \alpha < 3$, corresponds to a power-law SF with slope $\beta = (\alpha - 1)/2$, where $\tau = 1/f$ (e.g., Emmanoulopoulos et al. 2010). This is a direct consequence of the Wiener–Khinchin theorem. However, the analytical relationship between PSD and SF for the general case is not clear yet.

ORCID iDs

Lia F. Sartori ⓘ https://orcid.org/0000-0001-8020-3884
Benny Trakhtenbrot ⓘ https://orcid.org/0000-0002-3683-7297
Kevin Schawinski ⓘ https://orcid.org/0000-0001-5464-0888
Neven Caplar ⓘ https://orcid.org/0000-0003-3287-5250
Ezequiel Treister ⓘ https://orcid.org/0000-0001-7568-6412

References

Ai, Y. L., Yuan, W., Zhou, H. Y., et al. 2010, ApJL, 716, L31
Aird, J., Coil, A. L., & Georgakakis, A. 2018, MNRAS, 474, 1225
Aird, J., Coil, A. L., Moustakas, J., et al. 2012, ApJ, 746, 90
Antonucci, R. 1993, ARA&A, 31, 473
Arcavi, I., Gal-Yam, A., Sullivan, M., et al. 2014, ApJ, 793, 38
Assef, R. J., Prieto, J. L., Stern, D., et al. 2018, ApJ, 866, 26
Auchettl, K., Ramirez-Ruiz, E., & Guillochon, J. 2018, ApJ, 852, 37
Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214
Bartos, I., Kocsis, B., Haiman, Z., & Márka, S. 2017, ApJ, 835, 165
Bellm, E. 2014, in Proc. Third Hot-wiring the Transient Universe Workshop,
    ed. P. R. Wozniak (Los Alamos, NM: Los Alamos National Laboratory), 27
Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002
Bongiorno, A., Schulze, A., Merloni, A., et al. 2016, A&A, 588, A78
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2015, ApJ, 811, 148
Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2017, ApJ, 834, 111
Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2018, ApJ, 867, 148
Caplar, N., & Tacchella, S. 2019, MNRAS, 487, 3845
Catelan, M., & Smith, H. A. 2015, Pulsating Stars (New York et al: Wiley)
Chabrier, G. 2003, PASP, 115, 763
Chornock, R., Berger, E., Gezari, S., et al. 2014, ApJ, 780, 44
Connolly, S. 2015, arXiv:1503.06676
Conroy, C., & White, M. 2013, ApJ, 762, 70
De Cicco, D., Paolillo, M., Covone, G., et al. 2015, A&A, 574, A112
Deeter, J. E. 1984, ApJ, 281, 482
Deeter, J. E., & Boynton, P. E. 1982, ApJ, 261, 337
de Vries, W. H., Becker, R. H., & White, R. L. 2003, AJ, 126, 1217
Dexter, J., & Begelman, M. C. 2019, MNRAS, 483, L17
Djorgovski, S. G., Drake, A. J., Mahabal, A. A., et al. 2011, arXiv:1102.5004
Drake, A. J., Djorgovski, S. G., Mahabal, A., et al. 2009, ApJ, 696, 870
Emmanoulopoulos, D., McHardy, I. M., & Papadakis, I. E. 2013, MNRAS,
    433, 907
Emmanoulopoulos, D., McHardy, I. M., & Uttley, P. 2010, MNRAS, 404, 931
Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, AJ, 150, 150
French, K. D., Arcavi, I., & Zabludoff, A. 2016, ApJL, 818, L21
Gabor, J. M., & Bournaud, F. 2013, MNRAS, 434, 606
Gagne, J. P., Crenshaw, D. M., Kraemer, S. B., et al. 2014, ApJ, 792, 72
Gezari, S., Chornock, R., Rest, A., et al. 2012, Natur, 485, 217
González-Martín, O., & Vaughan, S. 2012, A&A, 544, A80
Goyal, A., Gopal-Krishna, Paul, J., Stalin, W., & Sagar, R, C. S. 2013,
    MNRAS, 435, 1300
Graham, M. J., Djorgovski, S. G., Drake, A. J., et al. 2017, MNRAS, 470, 4112
Green, A. R., McHardy, I. M., & Lehto, H. J. 1993, MNRAS, 265, 664
Guo, H., Wang, J., Cai, Z., & Sun, M. 2017, ApJ, 847, 132

Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 782, 9
Ho, L. C. 2009, ApJ, 699, 626
Hung, T., Gezari, S., Blagorodnova, N., et al. 2017, ApJ, 842, 29
Husemann, B., Urrutia, T., Tremblay, G. R., et al. 2016, A&A, 593, L9
Ivezic, Z., Axelrod, T., Brandt, W. N., et al. 2008, SerAJ, 176, 1
Ivezić, Ž., Connolly, A., Vanderplas, J., & Gray, A. 2014, Statistics, Data Mining
    and Machine Learning in Astronomy (Princeton, NJ: Princeton Univ. Press)
Ivezić, Ž., Smith, J. A., Miknaitis, G., et al. 2007, AJ, 134, 973
Jiang, Y.-F., Davis, S. W., & Stone, J. M. 2016, ApJ, 827, 10
Jones, M. L., Hickox, R. C., Black, C. S., et al. 2016, ApJ, 826, 12
Just, D. W., Brandt, W. N., Shemmer, O., et al. 2007, ApJ, 665, 1004
Kahn, S. 2018, in COSPAR Meeting 42, XLII COSPAR Scientific Assembly,
    E1.16-5-18
Kasliwal, V. P., Vogeley, M. S., & Richards, G. T. 2015, MNRAS, 451, 4328
Katebi, R., Chornock, R., Berger, E., et al. 2018, arXiv:1811.03694
Kauffmann, G., & Heckman, T. M. 2009, MNRAS, 397, 135
Keel, W. C., Chojnowski, S. D., Bennert, V. N., et al. 2012a, MNRAS,
    420, 878
Keel, W. C., Lintott, C. J., Maksym, W. P., et al. 2017, ApJ, 835, 256
Keel, W. C., Lintott, C. J., Schawinski, K., et al. 2012b, AJ, 144, 66
Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, ApJ, 698, 895
King, A., & Nixon, C. 2015, MNRAS, 453, L46
Kollmeier, J. A., Onken, C. A., Kochanek, C. S., et al. 2006, ApJ, 648, 128
Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv:1711.03234
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Kozłowski, S. 2016, ApJ, 826, 118
Kozłowski, S. 2017a, ApJ, 835, 250
Kozłowski, S. 2017b, A&A, 597, A128
Krolik, J. H. 1999, Natur, 398, 678
Labadie-Bartz, J., Pepper, J., McSwain, M. V., et al. 2017, AJ, 153, 252
LaMassa, S. M., Cales, S., Moran, E. C., et al. 2015, ApJ, 800, 144
Lawrence, A. 2018, NatAs, 2, 102
Lawrence, A., Bruce, A. G., MacLeod, C., et al. 2016, MNRAS, 463, 296
Lawrence, A., & Papadakis, I. 1993, ApJL, 414, L85
Law-Smith, J., Ramirez-Ruiz, E., Ellison, S. L., & Foley, R. J. 2017, ApJ,
    850, 22
Lintott, C. J., Schawinski, K., Keel, W., et al. 2009, MNRAS, 399, 129
LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009,
    arXiv:0912.0201
MacLeod, C. L., Green, P. J., Anderson, S. F., et al. 2019, ApJ, 874, 8
MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., et al. 2010, ApJ, 721, 1014
MacLeod, C. L., Ross, N. P., Lawrence, A., et al. 2016, MNRAS, 457, 389
MacLeod, C. L., Ivezić, Ž., Sesar, B., et al. 2012, ApJ, 753, 106
Maiolino, R., Risaliti, G., Salvati, M., et al. 2010, A&A, 517, A47
Marchese, E., Braito, V., Della Ceca, R., Caccianiga, A., & Severgnini, P.
    2012, MNRAS, 421, 1803
Marconi, A., Risaliti, G., Gilli, R., et al. 2004, MNRAS, 351, 169
Marin, F. 2017, A&A, 607, A40
Markowitz, A., Edelson, R., & Vaughan, S. 2003, ApJ, 598, 935
Martini, P., & Schneider, D. P. 2003, ApJL, 597, L109
Mathur, S., Denney, K. D., Gupta, A., et al. 2018, ApJ, 866, 123
Matt, G., Guainazzi, M., & Maiolino, R. 2003, MNRAS, 342, 422
McElroy, R. E., Husemann, B., Croom, S. M., et al. 2016, A&A, 593, L8
McHardy, I. M., Arévalo, P., Uttley, P., et al. 2007, MNRAS, 382, 985
McHardy, I. M., Koerding, E., Knigge, C., Uttley, P., & Fender, R. P. 2006,
    Natur, 444, 730
McHardy, I. M., Papadakis, I. E., Uttley, P., Page, M. J., & Mason, K. O. 2004,
    MNRAS, 348, 783
McKernan, B., Ford, K. E. S., Kocsis, B., Lyra, W., & Winter, L. M. 2014,
    MNRAS, 441, 900
Mejía-Restrepo, J. E., Trakhtenbrot, B., Lira, P., Netzer, H., &
    Capellupo, D. M. 2016, MNRAS, 460, 187
Merloni, A., Predehl, P., Becker, W., et al. 2012, arXiv:1209.3114
Minkowski, R. L., & Abell, G. O. 1963, in The National Geographic Society-
    Palomar Observatory Sky Survey, ed. K. A. Strand (Chicago, IL: Univ.
    Chicago Press), 481
Miyamoto, S., Kimura, K., Kitamoto, S., Dotani, T., & Ebisawa, K. 1991, ApJ,
    383, 784
Morganson, E., Burgett, W. S., Chambers, K. C., et al. 2014, ApJ, 784, 92
Morganson, E., Green, P. J., Anderson, S. F., et al. 2015, ApJ, 806, 244
Mushotzky, R. F., Edelson, R., Baumgartner, W., & Gandhi, P. 2011, ApJL,
    743, L12
Narayan, R., Mahadevan, R., & Quataert, E. 1998, in Theory of black hole
    Accretion Disks, ed. M. A. Abramowicz, G. Björnsson, & J. E. Pringle
    (Cambridge: Cambridge Univ. Press), 148
Neal, R. M. 2003, Ann. Statist., 31, 705

Novak, G. S., Ostriker, J. P., & Ciotti, L. 2011, ApJ, 737, 26
Panessa, F., Bassani, L., Cappi, M., et al. 2006, A&A, 455, 173
Paolillo, M., Papadakis, I., Brandt, W. N., et al. 2017, MNRAS, 471, 4398
Paolillo, M., Schreier, E. J., Giacconi, R., Koekemoer, A. M., & Grogin, N. A. 2004, ApJ, 611, 93
Peterson, B. M. 2014, SSRv, 183, 253
Piconcelli, E., Fiore, F., Nicastro, F., et al. 2007, A&A, 473, 85
Rees, M. J. 1988, Natur, 333, 523
Reynolds, C. S., & Miller, M. C. 2009, ApJ, 692, 869
Ricci, C., Bauer, F. E., Arevalo, P., et al. 2016, ApJ, 820, 5
Ross, N. P., Ford, K. E. S., Graham, M., et al. 2018, MNRAS, 480, 4468
Ruan, J. J., Anderson, S. F., Cales, S. L., et al. 2016, ApJ, 826, 188
Rumbaugh, N., Shen, Y., Morganson, E., et al. 2018, ApJ, 854, 160
Runnoe, J. C., Cales, S., Ruan, J. J., et al. 2016, MNRAS, 455, 1691
Sánchez-Sáez, P., Lira, P., Cartier, R., et al. 2019, ApJS, 242, 10
Sartori, L. F., Schawinski, K., Koss, M., et al. 2016, MNRAS, 457, 3629
Sartori, L. F., Schawinski, K., Koss, M. J., et al. 2018a, MNRAS, 474, 2444
Sartori, L. F., Schawinski, K., Trakhtenbrot, B., et al. 2018b, MNRAS, 476, L34
Schawinski, K., Evans, D. A., Virani, S., et al. 2010, ApJL, 724, L30
Schawinski, K., Koss, M., Berney, S., & Sartori, L. F. 2015, MNRAS, 451, 2517
Schreiber, T., & Schmitz, A. 1996, PhRvL, 77, 635
Schulze, A., & Wisotzki, L. 2010, A&A, 516, A87
Sesar, B., Ivezić, Ž., Lupton, R. H., et al. 2007, AJ, 134, 2236
Sesar, B., Svilković, D., Ivezić, Ž., et al. 2006, AJ, 131, 2801
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Shappee, B. J., Grupe, D., Mathur, S., et al. 2013, ATel, 5059, 1
Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, ApJ, 788, 48
Shen, Y. 2013, BASI, 41, 61
Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45
Simm, T., Salvato, M., Saglia, R., et al. 2016, A&A, 585, A129
Smith, K. L., Mushotzky, R. F., Boyd, P. T., et al. 2018, ApJ, 857, 141
Stern, D., McKernan, B., Graham, M. J., et al. 2018, arXiv:1805.06920
Suberlak, K., Ivezić, Ž., MacLeod, C. L., Graham, M., & Sesar, B. 2017, MNRAS, 472, 4870
Syer, D., Clarke, C. J., & Rees, M. J. 1991, MNRAS, 250, 505
Timmer, J., & Koenig, M. 1995, A&A, 300, 707
Trakhtenbrot, B., Arcavi, I., MacLeod, C. L., et al. 2019, arXiv:1903.11084
Trakhtenbrot, B., & Netzer, H. 2012, MNRAS, 427, 3081
Treister, E., Privon, G. C., Sartori, L. F., et al. 2018, arXiv:1801.06190
Trevese, D., Boutsia, K., Vagnetti, F., Cappellaro, E., & Puccetti, S. 2008, A&A, 488, 73
Urry, C. M., & Padovani, P. 1995, pasp, 107, 803
Uttley, P., Edelson, R., McHardy, I. M., Peterson, B. M., & Markowitz, A. 2003, ApJL, 584, L53
Uttley, P., & Mchardy, I. M. 2004, PThPS, 155, 170
Uttley, P., McHardy, I. M., & Papadakis, I. E. 2002, MNRAS, 332, 231
van der Klis, M. 1997, in Statistical Challenges in Modern Astronomy II, ed. G. J. Babu & E. D. Feigelson (Berlin: Springer), 321
van Velzen, S., Farrar, G. R., Gezari, S., et al. 2011, ApJ, 741, 73
Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549
Vanderplas, J., Connolly, A., Ivezić, Ž., & Gray, A. 2012, in Proc. Conf. Intelligent Data Understanding (CIDU), Introduction to astroML: Machine Learning for Astrophysics, 47
Vaughan, S., Edelson, R., Warwick, R. S., & Uttley, P. 2003, MNRAS, 345, 1271
Vignali, C., Brandt, W. N., & Schneider, D. P. 2003, AJ, 125, 433
Villforth, C., Koekemoer, A. M., & Grogin, N. A. 2010, ApJ, 723, 737
Volonteri, M., Capelo, P. R., Netzer, H., et al. 2015, MNRAS, 452, L6
Wang, J., Xu, D. W., & Wei, J. Y. 2018, ApJ, 858, 49
Weigel, A. K., Schawinski, K., Caplar, N., et al. 2017, ApJ, 845, 134
Wilhite, B. C., Brunner, R. J., Grier, C. J., Schneider, D. P., & vanden Berk, D. E. 2008, MNRAS, 383, 1232
Yang, Q., Wu, X.-B., Fan, X., et al. 2018, ApJ, 862, 109
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579
Zetzl, M., Kollatschny, W., Ochmann, M. W., et al. 2018, A&A, 618, A83
Zu, Y., Kochanek, C. S., Kozłowski, S., & Udalski, A. 2013, ApJ, 765, 106
Zuo, W., Wu, X.-B., Liu, Y.-Q., & Jiao, C.-L. 2012, ApJ, 758, 104