Understanding AGN physics through variability

Elena López Navas

Supervisor: Dr. Patricia Arévalo Noordam (UV)
Internal reviewer: Dr. Lorena Hernández García (UV)
External reviewer: Dr. Paulina Lira (UCh)

Tesis para optar al grado de Doctora en Astrofísica
Instituto de Física y Astronomía
Facultad de Ciencias
Universidad de Valparaíso

Mayo 2023
Valparaíso, Chile.

To my friends, for their support and tenderness.
And to all the women who fight and inspire.

A mis amigas, por su apoyo y su ternura.
Y por todas las mujeres que luchan e inspiran.

This thesis is solely my own composition,
except where specifically indicated in the text.

Elena López Navas
Mayo 2023
Valparaíso, Chile.

Elena L.N.

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor, Dr. Patricia Arévalo, for her unwavering support, guidance, and encouragement throughout my PhD journey. Their insights, constructive criticism, and expertise have been invaluable in shaping my research and helping me overcome various challenges.

I am also immensely grateful to the members of my research group. Their discussions, feedback, and collaboration have been instrumental in advancing my research and broadening my perspective on the field. In particular, I would like to thank Dr. Matthew Graham and Dr. Paula Sánchez-Sáez for their hospitality and support during my PhD internships.

I also would like to thank financial support from Agencia Nacional de Investigación y Desarrollo / Programa de Becas/ Doctorado Nacional 21200718, the Max Planck Society and the Millenium Nucleus NCN19_058 (TITANs).

Last but not least, I would like to express my deep appreciation to my family and friends from all around the world, who have given me their unwavering love and encouragement, and to all who helped me make my artistic outreach project come true.

Thank you all from the bottom of my heart.

Abstract

In recent years, the discovery of the so-called changing-look (CL) active galactic nuclei (AGNs) has evidenced that several scenarios can lead to variations in the broad emission lines (BELs, with widths $>$1,000 km s^{-1}) and so to the AGN optical classifications, which has challenged the classical view of AGNs defined by the simplest unified models (UM).

This PhD thesis presents a comprehensive study of optical variability in AGNs with the aim of shedding light on the Type 1/Type 2 dichotomy and improving the search for the intriguing and rare CL AGNs. Our approach is based on the exploration of optical fluctuations of spectrally-classified Type 2 sources, that is, AGNs whose accretion disc and broad line region (BLR) should be obscured, and so their BELs and optical continuum variability. For that purpose, we make use of the Zwicky Transient Facility (ZTF), a state-of-the-art optical time-domain survey, and the Automatic Learning for the Rapid Classification of Events (ALeRCE), a Chilean-led broker using machine learning models, in preparation for the big data era with the arrival of the Legacy Survey of Space and Time (LSST).

By analysing systematically the ZTF light curves of a large ($>$15,000) Type 2 sample, we find that $\sim$ 11 per cent of sources show evidence for optical variations, which leads to the discovery of misclassified Type 1s with weak BELs and CL candidates. We then apply the same strategy (i.e., searching for optical variations in Type 2 sources) using the current light curve classifications given by ALeRCE, and find $\sim$ 60 new CL candidates. We took second epoch spectra of 36 candidates and confirmed 50 per cent of sources as turning-on CLs, resulting in one of the selection techniques with the highest success rate of CL confirmations up to date.

Overall, this thesis strengthens the importance of variability studies in understanding the physics of AGNs, and contributes to the use of machine learning algorithms together with all-sky variability surveys to search for intrinsically rare objects. The findings of this research will ultimately contribute to our understanding of the fundamental processes that drive the evolution of AGNs.

Contents

Chapter 1

Introduction

Active Galactic Nuclei (AGNs) are energetic astrophysical sources whose emission cannot be attributed to the nuclear fusion powering stars as in non-active galaxies. They are among the most energetic objects in the universe, emitting huge amounts of energy across the entire electromagnetic spectrum, from radio waves to gamma rays, up to a bolometric luminosity $L_{bol} \approx 10^{48}$ erg s^{-1}. Such significant luminosities make it possible to detect powerful AGNs at high redshifts (currently $z = 7.6$, Wang et al., 2021). Since their luminosity varies on very short time-scales —as short as minutes in the X-ray band— the size of the central region must be really compact (e.g. Ulrich et al., 1997), so AGNs are inferred to be powered by accretion of matter onto supermassive ($\gtrsim 10^6 M_\odot$) black holes (SMBHs). As matter falls toward the SMBH, it heats up and emits vast amounts of energy in the form of radiation and high-energy particles. The accretion of gas onto a SMBH is known for its exceptional efficiency: depending on the spin of the SMBH, approximately $5 - 42$ per cent of the energy available from the accretion process is converted into radiation (Kerr, 1963; Shapiro & Teukolsky, 1983).

AGNs present a diversity of observational properties, leading to a large zoo of AGN classes with distinct characteristics, such as their spectra, variability, and morphology. 'AGN unification' is the idea that all their observed properties can be explained by a small number of physical parameters, namely: orientation (e.g. Antonucci, 1993; Urry & Padovani, 1995; Netzer, 2015), accretion rate (Heckman & Best, 2014), the presence of strong two-sided collimated outflows called jets (e.g. Padovani, 2016), and possibly the host galaxy and the environment.

According to Heckman & Best (2014), the low-redshift population of AGNs can be divided into two main categories. The first category comprises sources whose energy is dominantly released in form of electromagnetic radiation produced by the conversion of the potential energy of the gas accreted by the SMBH. This includes Seyferts and QSOs and are called radiative-mode AGNs. The second category consists of jet-mode AGNs, historically known as low-excitation radio galaxies. These objects produce relatively little radiation and their prime energy output is bulk kinetic energy transported in jets. In this work, we focus on the study of radiative-mode AGNs, which are generally classified according to the standard Unified Model.

1.1 AGN components

Surrounding the central SMBH, AGNs are believed to include several components, which are illustrated in Fig. 1.1 together with their spectral energy distribution (SED):

- Accretion disc: A geometrically thin, optically thick accretion disc reaches into the radius of the innermost stable orbit around the central SMBH (e.g. Sun & Malkan, 1989; Laor & Netzer, 1989, and references therein). The accreted gas loses angular momentum through molecular, magnetic or other sources of viscosity as it spirals inwards towards the SMBH, and releases gravitational potential energy into heat forming a radial temperature gradient $T \propto R^{-3/4}$ (Shakura & Sunyaev, 1973). This results in a multi-color blackbody (power-law) continuum emission emerging from the extreme *UV* through to the optical portion of the electromagnetic spectrum.

- Corona: The accretion disc is surrounded by a hot electron corona that Compton-up-scatters the soft photons from the disc to X-ray energies. The intrinsic X-ray emission from the corona is tightly connected to the accretion-disc emission (e.g. Steffen et al., 2006; Lusso & Risaliti, 2016), demonstrating a causal relationship and showing the universality of X-ray emission in AGNs. Fluctuations in the X-ray flux on short time-scales (hours to days) suggest that the X-ray corona is small and compact, but its geometry is still under discussion.

- Broad line region (BLR): The BLR consists of large-column-density ($\sim 10^{23}$ cm^{-2}), high-density ($\sim 10^{10}$ cm^{-3}) dusty-free clouds at about 0.1–1 pc from the SMBH

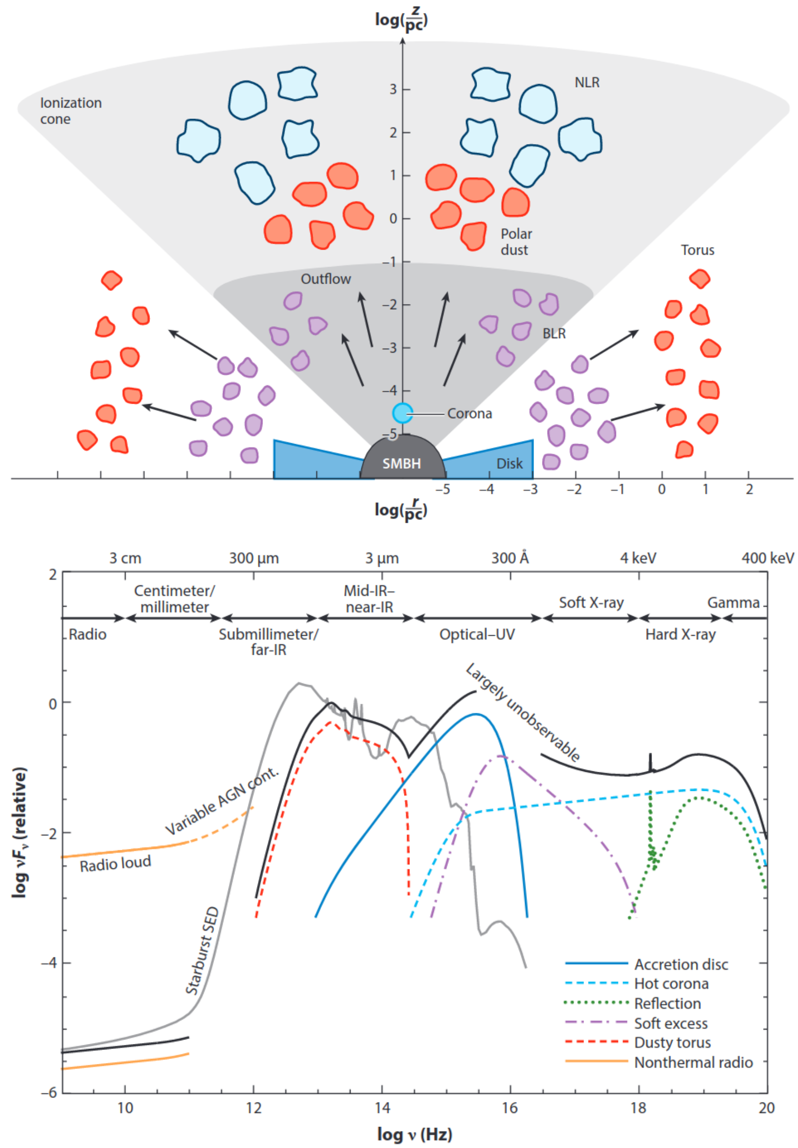


Figure 1.1: Schematic representation of the AGN model and its main components (top) and the SED (bottom) of an unobscured AGN (black curve) compared to the SED of a star-forming galaxy (gray curve). The emission from the disc, corona and torus are colored accordingly in both representations. Figures adapted by Hickox & Alexander (2018) from Ramos Almeida & Ricci (2017) (top) and Harrison (2014) (bottom).

(Czerny & Hryniewicz, 2011). The clouds of the BLR illuminated by the radiation coming from the accretion disc are highly ionized, and its size and structure are related to the luminosity of the AGN (e.g. Baldwin, 1997). This gas is the source of broad permitted and semi-forbidden emission lines (BELs) with typical Doppler widths between 1,000-20,000 kms^{-1}.

- Torus: Axisymmetric, clumpy and dusty structure on scales larger than that of the BLR (with a radius R <10 pc), with an inner boundary set by the sublimation temperature of the dust grains (Netzer & Laor, 1993). The existence of the torus is key to understand both the emission from hot dust in the mid-infrared (MIR), and obscuration of the accretion disc and corona emission depending on the line of sight (LOS) (Netzer, 2015). However, its geometric distribution and kinematics are still uncertain and can vary significantly between AGNs, leading to a broad range of covering factors (fc , the fraction of sky covered by the obscuring material, as seen from the SMBH) and observed column densities (e.g. Ramos Almeida et al., 2011; Burtscher et al., 2013).

- Narrow line region (NLR): The NLR consists of lower density (10^3–10^6 cm^{-3}), lower ionization gas that typically extends out to a few thousand parsecs from the central SMBH (Capetti et al., 1996), along the direction of the opening in the torus. It is characterized by its narrow emission lines, which have widths of 300 to 1,000 kms^{-1}. Because of its extent, the NLR emission is generally not obscured by the torus, although a fraction of the emission could be obscured by dust in the host galaxy.

- Jets: About 10 per cent of the radiative mode sources are radio-loud (Padovani, 2016), showing a highly collimated, relativistic radio jet and, occasionally, a γ-ray jet.

1.2 Unified Model

The Unified Model (UM) for AGNs is a theoretical framework that attempts to explain the observed differences in the appearance of AGNs through the IR-optical-UV-X-ray bands. The simplest model was proposed by Antonucci (1993) and later reviewed by Netzer (2015). It involves the existence of the torus, which presents a column density that can completely obscure the central source depending on its inclination to the

LOS: small inclinations allow a direct view of the central source, whereas high inclinations result in large obscuration. However, apart from orientation, the UM has other fundamental pillars: covering factor (fc) and AGN luminosity. On the one hand, the location of the inner wall of the torus is set by the distance from the nucleus at which dust reaches its sublimation temperature (e.g. Barvainis, 1987). Since in brighter objects this distance is greater, the critical angle that determines whether an observer can directly view the central engine increases (assuming a constant height of the torus). As a result, the fraction of unobscured AGNs becomes dependent on the luminosity. This is known as the 'receding torus' model and was first suggested by Lawrence (1991). On the other hand, fc has been found to depend primarily on mass-normalized accretion rate (Ricci et al., 2017), and it has been considered as an additional, independent variable (Elitzur, 2012).

In general, the UM and the existence of the torus have been quite successful in explaining the appearance of a wide variety of AGNs. However, many other considerations must be taken into account in order to explain all the observed AGN properties, such as the presence of other obscurers (Ramos Almeida & Ricci, 2017) and the identification of a growing number of sources that may not belong to this scheme (e.g. sources apparently obscured in the optical but unobscured in X-rays, Panessa & Bassani, 2002).

1.2.1 AGN Classification

According to the classical UM, the observed properties of an AGN depend on the orientation of the system relative to the observer. Specifically, there are two main classes of AGNs:

i) Type 1: These are AGNs that have a clear view of the central engine and the BLR, so their optical/UV spectra show both broad and narrow emission lines.

ii) Type 2: Type 2 AGNs are thought to be viewed from a direction that is close to the plane of the torus so it blocks the view of the central engine and BLR. As a result, their spectra only show narrow emission lines from the NLR.

In addition to Type 1 and Type 2 AGNs, there are also intermediate types that exhibit weak BELs, which could be due to a partial view of the BLR. Based on increasingly fainter BELs, the different classifications are Type 1.2, 1.5, 1.8 and 1.9.

In the X-rays, there is a good agreement between the Type 1/Type 2 classification and the X-ray absorption (Koss et al., 2017): the vast majority of Type 1s correspond to 'unobscured' AGNs, with column densities $N_H < 10^{22} cm^{-2}$, while Type 2s show $N_H \geq 10^{22} cm^{-2}$ and are commonly referred to as 'obscured'.

Apart from these optical and X-ray classification schemes, AGNs are also classified as high luminosity or 'quasars' ($L_X > 10^{44}$ erg s^{-1}), moderate luminosity or 'Seyfert galaxies' ($L_X = 10^{42} - 10^{44}$ erg s^{-1}), and low-luminosity AGNs (LLAGNs, with $L_X < 10^{42}$ erg s^{-1}, Ho, 2008).

1.3 AGN Variability

The emission from the different components of the accretion flow/outflow is known to change rapidly over a wide range of time-scales, as short as minutes in the lowest-mass AGNs. One of the main benefits of AGN variability studies is that they allow us to probe the structure and geometry of the accretion flow and the surrounding environment, as well as the physical conditions and mechanisms responsible for the observed emission.

1.3.1 X-ray

The X-ray flux variability is generally caused by changes in the normalization of the power-law continuum (Markowitz et al., 2003; Vaughan & Fabian, 2004, and references therein), which is the dominant emission component in the X-ray band. However, there is also evidence for variable absorption driving some of the observed X-ray variability, showing that this phenomenon is rather common in AGNs (Risaliti et al., 2002; Markowitz et al., 2014; Laha et al., 2020). In about a dozen of cases, variations of the column densities of the obscuring material were found to be very extreme, with the LOS obscuration going from Compton-thick ($N_H \geq 10^{24}$cm^{-2}) to Compton-thin ($N_H < 10^{24}$cm^{-2}) or viceversa (e.g. Guainazzi, 2002; Matt et al., 2003; Risaliti et al., 2005). Due to the strong changes in their spectral shapes, these objects are called changing-look (CL) AGN or changing-obscuration AGN (CO–AGN). Such events are typically observed in the X-rays, since the X-ray corona is considerably easier to obscure than more extended regions such as the BLR. Variability in N_H has provided clear indication that the material around the SMBH is clumpy, and could be associated with the BLR or with the putative torus. While the most widely accepted explanation for CL events is related to column-density variability due to clouds moving in and out the LOS, several additional explanations have been proposed over the years, such as outflows or changes in the intrinsic X-ray luminosity (see review by Ricci & Trakhtenbrot, 2022).

1.3.2 Optical/UV

Depending on the time-scale, AGN optical continuum variability is explained by a composite picture. On the one hand, the variations from the *r* band through to the *UV* are very well correlated, with observed short red lags by up to a day (e.g. Wanders et al., 1997; Sergeev et al., 2005). This can be explained if the optical variability is driven by heating from the intrinsically varying X-ray source, where the lags are interpreted as differential light-travel time between the emitting regions assuming a standard disc temperature profile.

On the other hand, on longer time-scales of months–years, larger amplitude variations are seen in the optical compared to X-rays (Uttley et al., 2003; Arévalo et al., 2008, 2009; Breedt et al., 2009). The natural interpretation is that a significant component of long term optical variability are driven by intrinsic variations in the accretion disc (Gaskell, 2008). This component is also correlated with long-term X-ray variations, which further suggests that instabilities in the accretion disc may ultimately drive the X-ray variability on those time-scales.

From the spectroscopic point of view, during the last decade it has become clear that several mechanisms can lead to changes in the optical/*UV* spectral properties of AGNs. In the most dramatic cases, the BELs completely appear/disappear in time-scales from months to years (e.g. LaMassa et al., 2015; Yang et al., 2018), defying the expectations from our basic understanding of AGN structure, and specifically the classical UM. Due to their change in appearance, they are called changing-look (CL) AGNs as in the X-ray field, but the physical processes driving such changes have been found to be generally different. As an example, we show in Fig. 1.2 the optical spectra of two optical CLs identified in this work, where striking differences in the continuum emission and BELs can be seen. For most of the cases, the optical CLs have been associated to rapid changes in the accretion-driven radiation field itself, which lead to the suppression, enhancement, or (dis-)appearance of the blue continuum and BELs typical of Type 1 AGNs, so they are called changing-state AGNs (CS AGNs, Graham et al., 2020). Depending on the origin of the BLR, a change in the accretion flow (or rate) could either dramatically transform the BLR or change its ionization properties, giving rise to the observed variable BELs. We summarize the possible BLR formation

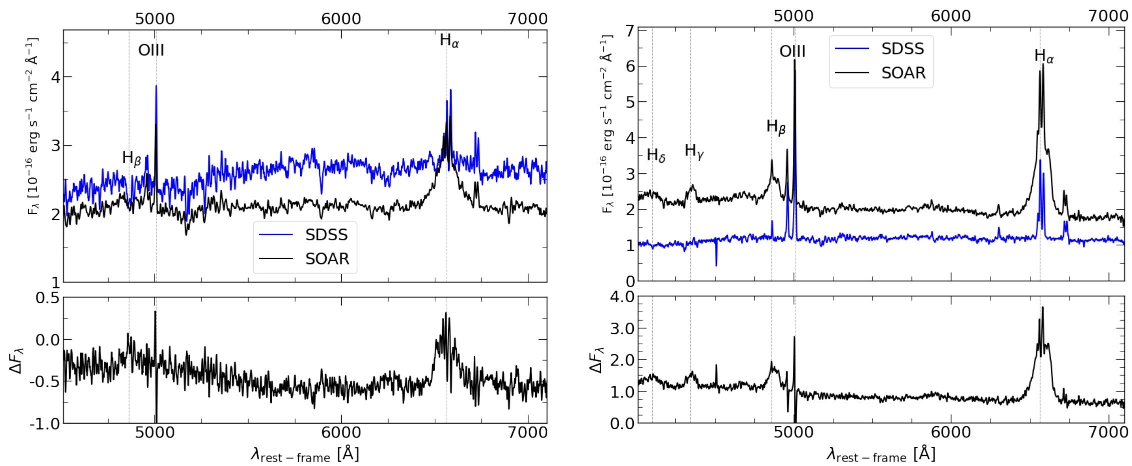


Figure 1.2: Archival (blue line) and new optical spectra (black line) for two CLs identified in this work (see López-Navas et al., 2022, or Chapter 3). The lower plots show the difference between the new and the old spectra, where a bluer continuum and BELs typical from Type 1 AGNs appear.

scenarios and its relation to the CS transitions in Sections 1.4 and 1.5. From now on, we will use the term CL to refer to sources suffering drastic changes in the optical BELs (and not to X-ray CLs), independently of the origin of such changes, except where specifically indicated in the text.

1.3.3 Infrared

The dust from the torus reprocesses the emission of the accretion disc into the infrared (IR), dominating the AGN SED from wavelengths longer than $\sim 1\mu$m up to a few tens of micron (e.g. Ramos Almeida et al., 2009). This emission is very prominent in the near-IR (NIR, 1–3 μm) and mid-IR (MIR, 3–50 μm) for both Type 1 and Type 2 AGNs, and its strength depends on the fc of the dust around the accretion disc. For longer wavelengths belonging to the far-IR (FIR, 50–500 μm), dust emission from star-formation activity in the host galaxy is more likely to dominate the SED (e.g. Hatziminaoglou et al., 2010).

As the AGN flux varies in the optical, a time lag in the fluctuations of NIR flux can be interpreted as the light-travel time from the accretion disc to the dusty torus. Thus, reverberation mapping studies have been used to estimate the radius of the innermost dust torus (e.g. Koshida et al., 2014; Lira et al., 2015).

In the MIR, a number of AGN selection criteria have been proposed over the years.

These criteria rely primarily on colours to separate AGNs from stars or non-active galaxies, as AGNs are expected to be significantly redder in the shorter wavelength 3 – 5 μm MIR bands (e.g. Stern et al., 2005). However, variability analyses on large-scale AGN population remain relatively unexplored at wavelengths of 3 – 5 μm, with the definitive study performed on 8.1 deg^2 (Kozłowski et al., 2010) using the four-epoch *Spitzer* Deep Wide Field Survey (SDWFS, Ashby et al., 2009). In the study, the authors found that the majority (76 per cent) of the ~5,100 MIR variable sources in the SDWFS field were AGNs.

1.4 BLR and accretion flow

Observations lead to a very complex picture about the BLR, indicating that the motion of the material is predominantly Keplerian, with traces of inflow and clear signatures of outflow, and that the distribution is flattened rather than spherical, but the covering factor is high (0.1 – 0.3) so part of the material is far from the equatorial plane (e.g. see review by Czerny, 2019).

However, the mechanism of the formation of the BLR and its relation with the accretion flow is still under debate. There are several scenarios about the origin of the BLR material, such as inflow models where the BLR material comes from the outer regions (Wang et al., 2017), formation *in situ* from fragmentation of the accretion disc (Collin & Zahn, 1999, 2008; Wang et al., 2011) or the accretion disc wind. The accretion disc wind is the most complex scenario, and several variants have been proposed over the years. In this model, the underlying stationary cold optically thick disc extends as far as the BLR. One of the advantages is that the presence of the winds from accretion discs in AGNs is expected theoretically (for a review, see Proga, 2007). Based on the accretion disc wind idea, several mechanisms could drive the outflow: magnetically (centrifugally) driven wind (Blandford & Payne, 1982), thermally-driven wind (Czerny & King, 1989; Witt et al., 1997; Blandford & Begelman, 1999), line-driven wind (Murray et al., 1995; Chiang & Murray, 1996; Murray & Chiang, 1997), or dust-driven wind (Czerny & Hryniewicz, 2011).

1.5 Changing-state transitions

Over the last years, >200 AGNs with emerging or disappearing optical BELs have been found. Most of the studies have focused on a systematic search of drastic changes in the broad H α and H β lines through repeated spectroscopy of SDSS programs and/or other spectroscopic surveys (MacLeod et al., 2016; Yang et al., 2018; Green et al., 2022). Such studies indicate these events are rare: in a blind search for optical CL AGNs, Yang et al. (2018) found that just 19 out of 330795 (the 0.006 per cent) galaxies with repeated spectroscopy in the SDSS and/or Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) are CL AGNs. Other studies have focused on a search for CS transitions based on their expected multiwavelength properties, such as optical and MIR flux and color changes associated to the variable BELs as the AGN transitions to a bright or dim state (e.g., Sheng et al., 2017; Yang et al., 2018; Lyu et al., 2022). As an example, Graham et al. (2020) found 111 CS quasars by applying two different criteria: strongly enhanced optical variability over some time-scales and a large absolute difference between the initial and final state in the MIR *WISE* light curve (i.e, $|\Delta W1| > 0.2$ or $|\Delta W2| > 0.2$, see definitions in Section 1.6.2). That work led to a CS sample at higher luminosity than previous CL AGNs in the literature.

This rapidly developing topic offers a new window to study a time-resolved picture of the BLR formation and its relation to the accretion flow. On the one hand, the line emission from a pre-existing BLR is expected to respond to any strong ionizing continuum variability, as unambiguously seen in reverberation mapping studies. CS events could thus be regarded as the most extreme cases of BLR response, where variable irradiation activates/deactivates the emission line production without any changes in the BLR structure. On the other hand, according to the accretion disc wind scenario the disappearance of BELs in CS AGNs would arise from a strong decrease in the accretion rate (Elitzur & Shlosman, 2006), which would lead to a decrease in the mass outflow rate of the disc wind. In particular, the disappearance of the BLR is predicted to happen at luminosities lower than $5 \cdot 10^{39} (M/10^7 M_\odot)^{2/3}$ erg s^{-1}, where M is the SMBH mass (Elitzur & Ho, 2009). According to this model, AGNs follow an intrinsic spectral evolution as the luminosity increases, transitioning from Type 2 to Type 1 through the intermediate (Type 1.9, 1.8, 1.5, 1.2) types (Elitzur et al., 2014).

1.6 Variability surveys

To find and characterise rare variability in AGNs such as the CS transitions, we need high-cadence monitoring of a huge amount of sources. In this work, we make use of multi-wavelength photometric surveys that allow to investigate the light curve variability of AGNs: the Zwicky Transient Facility (ZTF), the Wide-field Infrared Survey

Explorer (WISE) and the extended ROentgen Survey with an Imaging Telescope Array (eROSITA). In this section we provide a brief overview of such surveys and introduce the Legacy Survey of Space and Time (LSST) – an upcoming optical survey that is expected to revolutionize our understanding of the universe during the next years.

1.6.1 Zwicky Transient Facility

The Zwicky Transient Facility (ZTF, Bellm, 2014; Bellm et al., 2019b) is a state-of-the-art optical time-domain survey that is designed to discover and study transient and variable phenomena in the night sky. It is located at the Palomar Observatory in Southern California and began operating in 2017.

The ZTF uses a 48-inch telescope equipped with a camera that images the sky in three different bands (g, r, i). The camera is equipped with 16 6K x 6K CCD detectors, providing a high resolution and a large 47 deg^2 area of sky coverage. The ZTF can observe 3760 deg^2 per hour to a 5 σ detection limit of 20.5 mag in r with a 30 s exposure.

Due to its funding profile, 50 per cent of ZTF observing time is dedicated to a 2-night cadence public survey of the entire Northern sky in g and r bands (Bellm et al., 2019a). In addition, ZTF produce alerts of transient events and variable objects in real time, which are made publicly available through the ZTF public alert stream. Generally, ZTF generates alerts for sources that are at least 5 σ above the background noise in a single exposure. This allows for rapid follow-up observations by the global astronomical community, and enables scientists to study transient and variable phenomena in unprecedented detail.

ZTF is a precursor to the upcoming LSST, which will have a similar alert system but with a more advanced data processing pipeline and a lower threshold for generating alerts.

1.6.2 Wide-field Infrared Survey Explorer

The Wide-field Infrared Survey Explorer (WISE, Wright et al., 2010) is a NASA space telescope that was designed to conduct an all-sky survey in the MIR portion of the electromagnetic spectrum. The telescope had a 40-cm (16-inch) primary mirror and a wide-field infrared camera that enabled it to image large areas of the sky with high sensitivity and resolution.

During the first phase of the mission, which took place from January to August 2010, WISE conducted a full-sky survey at four infrared wavelengths (3.4, 4.6, 12, and

22 microns, called *W1, W2, W3* and *W4*). The survey covered the entire sky, with a sensitivity that was 100 to 1,000 times better than previous infrared surveys. The survey data were used to create a catalog of over 500 million astronomical objects, including stars, galaxies, and other infrared sources.

The second phase of the mission, called NEOWISE (Mainzer et al., 2011), began in 2013 and focused on the discovery and characterization of near-Earth objects (NEOs), such as asteroids and comets, as well as the study of other transient and variable objects. NEOWISE used the same telescope and camera as the first phase, but with different observing strategies and just the *W1* and *W2* filters optimized for detecting and characterizing NEOs.

Currently, a new MIR time-domain catalog is publicly available, which is based on WISE and NEOWISE observations spanning the period from 2010 to 2020 (the unTimely Catalog, Meisner et al., 2023). Detections are generated by stacking the ≥ 12 single exposures at each sky location every six months during which WISE has been operational, therefore obtaining 16 epochs in each of the *W1* and *W2* channels. This catalog will permit all-sky studies of long-time-scale variability in the MIR for a variety of sources.

1.6.3 Extended ROentgen Survey with an Imaging Telescope Array

The extended ROentgen Survey with an Imaging Telescope Array (eROSITA, Predehl et al., 2021) is a wide-field X-ray telescope on-board the Russian-German Spectrum-Roentgen-Gamma (SRG) observatory, and it was developed under the leadership of the Max-Planck Institute for extraterrestrial Physics (MPE) in Germany. SRG was successfully launched on July 13, 2019, and placed in a halo orbit around the L2 point.

The telescope has seven identical X-ray mirror modules, each with 54 nested mirror shells made of highly polished nickel. The mirrors are designed to focus X-rays onto seven corresponding X-ray cameras, which are sensitive to X-rays in the energy range of 0.2–10 keV.

The primary scientific goal of eROSITA is to conduct a deep and comprehensive survey of the X-ray sky: the eROSITA all-sky survey (eRASS). It started mid December 2019 and has an envisaged duration of 4 years, and consists of eight consecutive 6-months all-sky scans. In the soft band (0.5–2 keV), the deep All-sky survey will be 30 times more sensitive than the previous ROSAT All-sky survey.

The design-driving science of eROSITA is to detect a significant number of galaxy clusters, up to approximately 100,000 objects, at redshifts $z > 1$. This approach aims to explore and analyze the large-scale structure of the Universe and to test cosmological models, including Dark Energy. Thanks to the large area coverage, another anticipated outcome of eROSITA is the X-ray identification of millions of AGNs. Moreover, its 6-months all-sky scans will allow us to measure some X-ray variability, providing a unique perspective on the evolution of an unprecedentedly large amount of SMBHs.

1.6.4 Next-generation survey: LSST

Vera C. Rubin Observatory (Ivezić et al., 2019) consists of an integrated system that combines an 8.4-meter primary mirror and the world's largest digital camera, with 3.2 gigapixels. It is currently under construction in Chile, and is expected to become operational in 2024.

The main goal of the Rubin Observatory is to conduct the 10-year Legacy Survey of Space and Time (LSST). The LSST will survey the entire southern sky over a period of 10 years, taking images of the same regions of sky repeatedly with a wide-field camera. This will result in a 500 petabyte set of images and data products, making it one of the largest astronomical datasets ever produced.

To achieve these goals, the LSST will use a state-of-the-art telescope and camera system. The 8.4-meter Simonyi Survey Telescope uses a special three-mirror design, which creates an exceptionally wide field of view, and has the ability to survey the entire sky in only three nights. The camera is a large-aperture, wide-field optical imager capable of viewing light from the near ultraviolet to near infrared (0.3–1 μm) wavelengths.

The work presented here using ZTF data can be replicated in the future with LSST to much higher depth, allowing us to study a larger sample and more distant AGNS.

1.7 Motivation

As mentioned above, the simple AGN-BLR unification schemes are capable of explaining many different properties of the Type 1/Type 2 dichotomy in AGNs. However, there are other subgroups of AGNs that are not contained in such schemes and need alternative explanations. In particular, a fraction of Type 2 objects appears to be X-ray unobscured, the so-called *True Type 2* (Panessa & Bassani, 2002). These systems often appear to be LLAGN whose emission could be diluted by the host-galaxy contribution (Shi et al., 2010; Barth et al., 2014), but some of them could be intrinsically unable to sustain the obscuring environment (the BLR and the torus) as proposed theoretically in the disc wind scenario (Nicastro, 2000; Elitzur & Ho, 2009).

With the discovery of CS AGNs, a new window to test the *True Type 2* population, the BLR formation mechanisms and its relation to the accretion flow has been opened. There are still numerous unresolved issues concerning CS transitions, such as the frequency and time-scales, the physical mechanisms driving such extreme changes, and their relation to main physical parameters such as the SMBH mass or accretion rate.

Future large samples of CS AGNs will allow us to address these open questions and improve further our overall understanding of AGN physics. This work is thus motivated by the advent of large sky-coverage photometric and spectroscopic surveys, with which we expect to identify and characterise hundreds of new CS transitions and other rare phenomena in AGNs.

The following Chapters 2 and 3 have already been published as manuscripts in *Monthly Notices of the Royal Astronomical Society* (López-Navas et al., 2022, 2023), and Chapter 4 has been submitted to the same journal.

Chapter 2

The Type 1 and Type 2 AGN dichotomy according to their ZTF optical variability

In recent years, the classical view of active galactic nuclei (AGNs) defined by the Unified Model (UM, Antonucci, 1993) has been challenged by the discovery of various anomalous objects. Within the framework of the UM, an anisotropic, geometrically and optically thick dusty torus can hide, partially or completely depending on the line-of-sight of the system, the direct emission from the accretion disc and the broad line region (BLR). Phenomenologically, this allows to classify AGNs depending on their optical/UV properties: sources with a blue continuum and broad emission lines (BELs) are called Type 1, whereas obscured, Type 2 AGNs are characterized by the lack of BELs in their optical/UV spectra.

The dusty torus obscuration corresponds to a typical extinction of 5 mag in the V band, and to an equivalent absorbing column density of $N_H > 10^{22}$ cm^{-2} in the X-ray band for typical dust-to-gas ratios (Predehl & Schmitt, 1995). In general, there is a good agreement between the X-ray absorption and the Type 1 and Type 2 dichotomy (94 per cent of agreement for the majority of Seyferts in the *Swift*–BAT AGN Spectroscopic Survey, Koss et al., 2017). However, there is a few percent of X-ray unobscured

AGNs whose optical spectra resemble Type 2 AGNs (Pappa et al., 2001; Xia et al., 2002; Wolter et al., 2005; Panessa et al., 2009; Brightman & Nandra, 2008; Bianchi et al., 2008). For most of these cases, the BELs are weak and are diluted by host galaxy contamination or by low signal to noise spectra (Shi et al., 2010; Barth et al., 2014), so these sources are misclassified Type 1 AGNs. Interestingly, there are still a few more exotic sources, called *naked* or *true Type 2* AGNs, that *could* intrinsically lack the denser gas that gives rise to the BLR, and/or the photoionizing continuum radiation that drives the broadline emission (Shi et al., 2010; Tran et al., 2011).

Independently of the origin of these unusual X-ray unobscured Type 2 AGNs, the discovery of the so-called changing-look (CL)/changing-state (CS) AGNs has evidenced that several scenarios can lead to variations in the BELs and so to the AGN optical classifications. In principle, the changes in the BELs could be associated with a large change of transient dust obscuration along the line-of-sight (Yang et al., 2019; Wang et al., 2019), in a similar way to the observed for the CL AGNs in X-ray astronomy (e.g. Rivers et al., 2015). Transient events such as tidal disruption events (TDEs) of a star by the supermassive black hole (SMBH) have also been claimed as possible drivers of CL phenomena (e.g. the case of 1ES 1927+654, which was previously classified as *true Type 2*, Trakhtenbrot et al., 2019). However, the studies of most other optical CL AGNs have ruled out these scenarios in favor of intrinsic changes to the accretion flow. In this case, variations in the accretion rate would lead to a disappearing BLR or to a dimming (brightening) of the AGN continuum and so to a reduced (increased) supply of ionizing photons available to excite the gas around the SMBH (LaMassa et al., 2015; Runnoe et al., 2016; MacLeod et al., 2016; Sheng et al., 2017; Hutsemékers et al., 2017; Noda & Done, 2018a; Hutsemékers et al., 2019; Graham et al., 2020). As mentioned above, this scenario also explains the existence of *true Type 2* AGNs. For instance, Guolo et al. (2021) found that NGC 2992 transitions recurrently from Type 2 to intermediate-type at an Eddington ratio (λEdd) of $\sim$1 per cent. This means that at lower values of λEdd, the AGN is still unobscured but intrinsically lacks BELs, which is by definition a *true Type 2* AGN.

Another hallmark in AGNs that is heavily affected by the obscuration of the system is the observed temporal variability. According to the UM, the continuum coming from the central source in obscured AGNs is blocked by the dusty torus, so the optical variability is highly suppressed in Type 2 sources (this prediction has been confirmed observationally, Yip et al., 2009; Sánchez et al., 2017). Based on these considerations, optical variability in spectroscopic Type 2 AGNs has led to the finding of sources that in principle could challenge the UM such as *true Type 2* candidates (Hawkins, 2004), misclassified Type 1 AGNs (Barth et al., 2014) and CL/CS AGNs (see Chapter 3).

In this chapter, we take advantage of the real time, deep, large sky-coverage monitoring survey Zwicky Transient Facility (ZTF, Bellm, 2014; Bellm et al., 2019b), to carry out one of the largest systematic investigations of Type 2 AGN optical variability to date, in comparison to a weak-Type 1 sample. The ZTF had first light in 2017 and employs an extremely wide ~ 47 deg^2 field-of-view camera mounted on the Samuel Oschin 48-inch Schmidt telescope. It is designed to scan the entire Northern sky every two days in the *gri* filters, which enables a wide variety of novel multiband time-domain studies. Here, we investigate different variability features that help us distinguish between Type 1 and Type 2 AGNs with the ZTF light curves, which leads to the discovery of a new sample of weak Type 1 AGNs and CL/CS candidates.

2.1 Sample and data

The parent sample consists of the 30520 galaxies classified as GALAXY AGN in the Sloan Digital Sky Survey (SDSS) Data Release 16 (DR16, Ahumada et al., 2020), which have detectable emission lines that are consistent with being a Seyfert or LINER according to the BPT-type (Baldwin et al., 1981) criteria employed by the SDSS pipeline ($\log_{10}$(OIII/H α)$<$ 0.7 – 1.2($\log_{10}$(NII/H α) + 0.4). These sources were not identified with a quasar (QSO) template by the SDSS pipeline, so they are generally weaker or obscured AGNs (we refer to Bolton et al., 2012, for details on how the pipeline classifies each spectrum). Light curves for the 29057 sources observed by the ZTF in the g and r filters were extracted from a forced photometry data set that has been produced based on all difference images available for ZTF DR5 (Mroz et al., in prep) and source positions from the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) Data Release 1 catalog (PS1 DR1 Chambers & et al., 2017). To clean the light curves for bad photometry data points we performed a 3-σ clipping twice and discarded the points whose error was greater than twice the mean value of the errors from each particular light curve. The variability analysis was performed on the 3-days binned light curves to reduce the errors on the light curves data points. After the cleaning and binning, we discarded the light curves with $<$10 data points, which reduces the sample to 20583. The final light curves have a mean of 100 data points in each band, spanning 900 days between April, 2018 and December, 2020.

In the sample, 2737 sources are sub-classified as BROADLINE by the SDSS pipeline, which means that lines with width > 200 km s^{-1} can be detected at least at the 5-σ level. Given the relatively low threshold for the line widths, these do not necessarily correspond to Type 1 AGNs. Additionally, broad lines detected at lower significance are not reported, so the pipeline cannot identify *weak* Type 1 AGNs. Therefore, more detailed analyses are needed to properly separate the different optical types of AGNs in the SDSS sample.

Here, we differentiate various subsamples according to previous studies and catalogues. In particular, we cross-matched the sample within a 1 arcsec radius with the Type 1s with weak BLRs from Oh et al. (2015) and Liu et al. (2019), selected via detailed modelling of SDSS DR7 spectra. We note that most of the Type 1 AGNs from Oh et al. (2015) were also in Liu et al. (2019)'s catalogue, so we merged both subsamples together (the *Weak Type 1* sample). These studies are based on the existence of a significant broad component and not to the width of the BEL, leading to the discovery of Type 1s at the low-mass and low-luminosity end. In particular, Liu et al. (2019) report H α luminosities in the range $10^{38.5}$–$10^{44.3}$erg s^{-1}, and line widths (FWHMs) of 500–34,000 km s^{-1}. In spite of the wide FWHMs and similar requirement on line significance, only 20 per cent of sources in the Weak Type 1 sample were classified as BROADLINE in the SDSS. On the other hand, 50 per cent of BROADLINE galaxies have not been identified as Type 1 in these catalogues despite they were presumably in their parent sample (galaxies and QSOs from the SDSS DR7, October 2008). Then, the difference in the selection of weak Type 1 and BROADLINE sources is not clear, and it might just arise from a more detailed spectroscopic method, with a better substraction of the galaxy contribution by Oh et al. (2015) and Liu et al. (2019).

We also separated the sources subclassified as BROADLINE in the SDSS and the blazars from The Roma-BZCAT Multi-Frequency Catalog of Blazars (*Blazar ROMA*, Massaro et al., 2015). In Table 2.1 we show the samples considered in this chapter, with the one called *Type 2 SDSS* being the parent sample without the sources that are in the other subsamples (BROADLINE, Blazar ROMA, Weak Type 1 and AGN lcc, the latter being explained in the next section). We note that, while the Weak Type 1 sample is quite pure, a level of contamination of very low-luminosity Type 1 objects is expected in the Type 2 sample. Since this chapter is focused on the comparison between the Type 1 and Type 2 objects, we selected the mean magnitudes and redshifts to be comparable between the samples: $g < 21$ and $r < 20$ mag and $z < 0.3$ (see Fig. 2.1). The final number of sources analysed after this selection are listed in Table 2.1. We note that, except for the Type 2 SDSS sources, some AGNs are included in more than

Table 2.1: Samples analysed in this chapter.

Name	N° of sources	g <21 / r <20 mag & z <0.3
Type 2 SDSS	16999	12743 / 13322
Weak Type 1	978	921 / 960
Blazar ROMA	26	20 / 21
BROADLINE	2737	2014 / 2118
AGN lcc	260	245 / 251

one subsample.

2.2 Light curve variability

We performed the variability analysis of the ZTF light curves on the SDSS DR16 AGN sample by extracting some of the variability features used by the Automatic Learning for the Rapid Classification of Events (ALeRCE, Förster et al., 2021) broker light curve classifier (LCC, Sánchez-Sáez et al., 2021b). The ALeRCE broker is currently processing the alert stream from the ZTF and has been selected as a Community Broker for the Vera C. Rubin Observatory and its Legacy Survey of Space and Time (LSST). The goal of the LCC is to provide a fast classification of transient and variable objects by applying a balanced random forest algorithm. For this purpose, a total of 174 features, including variability features and colours obtained from AllWISE and ZTF photometry, are computed for every object with at least 6 alerts in either g or r band. The complete set of features are described in `http://alerce.science/features/`, and a python library to extract variability features in astronomical light curves is publicly available at `https://github.com/alercebroker/lc_classifier`. In this chapter, we selected a set of variability features included in the LCC python library and applied them on the 3-days binned g and r ZTF light curves. For comparison, we also cross-matched our DR16 AGN sample with the sources classified as AGN or Blazar by the LCC, and treat them as another subsample in our variability analysis (*AGN lcc*). We note that this is the only subsample selected by flux variability, as opposed to spectral features, considered in this chapter.

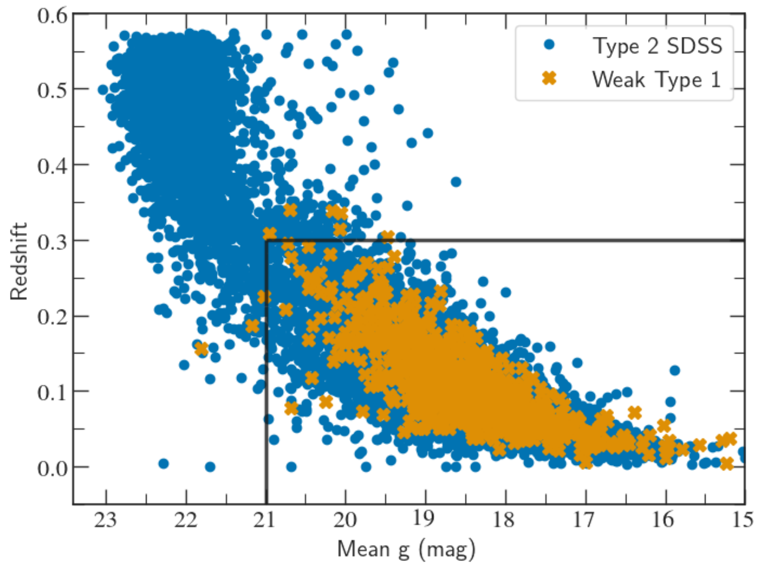


Figure 2.1: Redshift versus mean g magnitude for different samples considered in this chapter. The solid black lines indicate the limits used to compare different samples, z <0.3 and Mean g $<$ 21 mag.

2.2.1 Variability features

We characterize the optical variability of the ZTF light curves by extracting the following features included in the LCC python library:

- Excess Variance (ExcessVar): measure of the intrinsic variability amplitude. In the LCC library it is defined as $\mathrm{ExcessVar}_{LCC} = (\sigma_{\mathrm{LC}}^2 - \overline{\sigma}_m^2)/\overline{m}^2$, where σ_{LC} is the standard deviation of the light curve, $\overline{\sigma}_m$ is the average error, and $\overline{m}$ is the average magnitude. Here, we use ExcessVar$= (\sigma_{\mathrm{LC}}^2 - \overline{\sigma}_m^2)$, which is a magnitude version of excess variance analogous to the normalized excess variance that is typically used in linear flux units (Sánchez et al., 2017). Similar features have been broadly used to study X-ray variability (Nandra et al., 1997; Edelson et al., 2002; Vaughan et al., 2003, and references therein).

- Damped random walk (DRW) parameters: the DRW model is generally used to describe the AGN optical variability (see Kelly et al., 2009, and references therein). The model considers the light curves as continuous time stochastic processes, and provides an estimation of the characteristic time-scale τ and amplitude square σ^2 of the variations. Here, we will report the asymptotic value of the structure function on long time-scales (SF_∞) given by $\mathrm{SF}_\infty = \sqrt{2}\cdot\sigma$, which is broadly used in the literature (MacLeod et al., 2010). We also correct τ by the redshift of the sources as reported in the DR16 to obtain the intrinsic time-scales of the variations in the rest frame.

- Mexican hat power spectrum (MHPS, Arévalo et al., 2012): This method isolates variability on different time-scales in unevenly sampled light curves by applying a Mexican-hat type filter. It convolves each light curve with two Gaussian profiles of slightly different widths and takes the difference of the convolved light curves. This difference is dominated by fluctuations at time-scales of σ /0.225, where σ is the average width of the Gaussian filters, removing variations on shorter and longer time-scales. The variance of the resulting difference, as a function of frequency, is an estimation of the power spectrum. The time-scales by default are 10 and 100 days, which are convenient to separate AGNs to other stochastic sources such as long-period variable stars and young stellar objects. Since we will analyse just AGNs, we use the same feature than the LCC but we compute it at 300 (MH σ^2_{300}) and 150 days (MH σ^2_{150}).

2.3 Results

In order to characterize the optical variability of the SDSS DR16 AGN sample, we first computed the Excess Var for all the objects. This feature evaluates the variance after subtracting the contribution expected from measurement errors. As it can be seen in Fig. 2.2, the excess variance distributes approximately symmetrically around zero for magnitudes below $g \sim$21 mag, with a few sources showing clear positive values. This is expected when the intrinsic variance is essentially zero, so this behaviour is produced by an unbiased uncertainty in the noise estimate. In contrast, at dimmer magnitudes the Excess Var drops toward negative values, showing that the noise is systematically overestimated at larger values than the ZTF limiting magnitude (g $\sim$20.8 mag, AB, 5 σ in 30 sec, Masci et al., 2019) .

After the redshift and magnitude selection, we computed the variability features for each sample, whose results are presented in Table 2.2. As it can be seen, the samples are significantly distinct, with one order of magnitude difference between the variances of the less (Type 2 SDSS) and the most (AGN lcc) variable sources. Notably, the BROADLINE and the Type 2 SDSS samples present very similar features, with lower variances and DRW parameters than the Weak Type 1 objects. In parallel, the

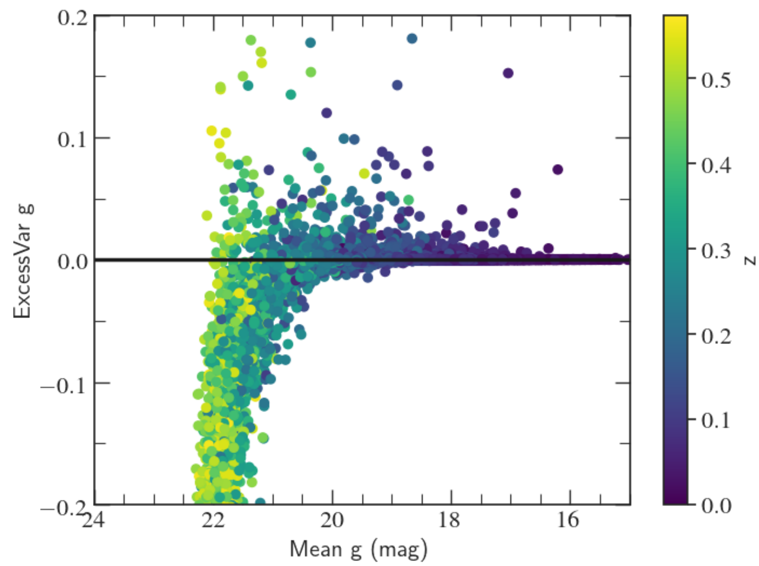


Figure 2.2: Excess variances (ExcessVar) versus mean magnitude in the g band for the SDSS DR16 AGN sample ZTF light curves. Above 21 magnitudes, most of the excess variances are negative, which indicates an overestimation of the average errors.

features from the Blazar ROMA sample are comparable to those of the Weak Type 1s.

Since the Weak Type 1 sample contains the only objects that have been classified via detailed spectroscopy, we next focus on the differences between this and the Type 2 SDSS sample to examine the ability of the selected features to separate both populations.

2.3.1 Type 2 versus Weak Type 1 variability features

Firstly, we investigated the Excess Var for the two samples, which is shown in Fig. 2.3a. From now on, we will present the results just in the g band, but the same conclusions are recovered when studying the r band. In this figure, it can be seen that for the Type 2 SDSS sample a small fraction of sources with large negative excess variances remains, but in general the data points distribute around zero as expected for non-variable sources. Although both samples have objects with positive and negative excess variances, the distribution of Weak Type 1s is significantly skewed towards positive values, indicating that as a sample they have stronger variability. A Kolmogorov-Smirnov (ks) test demonstrates the significance of this difference with ks=0.35 and p-value=$2 \cdot 10^{-93}$. We tested whether the difference in the samples could be due to the different distributions in magnitude, in particular to the higher fraction of Type 2

Table 2.2: Variability results for each sample considered in this chapter. The light curves were selected to have <21 mag in the g band, <20 mag in the r band, and z <0.3. The values reported correspond to the 16th, 50th and 84th percentiles.

Sample	Counts	ExcessVar g	ExcessVar r	MH σ^2_{g300}	MH σ^2_{r300}	$SF_{\infty,g}$	$SF_{\infty,r}$	τ_g	τ_r
	g/r	$\cdot 10^{-5}$	$\cdot 10^{-5}$	$\cdot 10^{-2}$	$\cdot 10^{-2}$	mag	mag	days	days
Type 2 SDSS	12743/13322	7^{60}_{-40}	$7^{24}_{-0.7}$	$0.17^{1.6}_{-0.5}$	$0.10^{0.5}_{-0.07}$	$0.018^{0.04}_{1.9e-6}$	$0.008^{0.018}_{4e-8}$	$6^{50}_{1.1}$	$5^{30}_{1.2}$
BROADLINE	2014/2118	10^{70}_{-25}	$8^{30}_{0.9}$	0.16^{3}_{-3}	$0.10^{1.0}_{-0.21}$	$0.018^{0.05}_{4e-6}$	$0.010^{0.021}_{9e-7}$	$7^{60}_{1.3}$	$5^{30}_{1.2}$
Weak Type 1	921/960	$60^{500}_{1.2}$	30^{140}_{5}	$1.3^{14}_{0.012}$	$0.5^{3}_{0.03}$	$0.04^{0.11}_{4e-5}$	$0.019^{0.05}_{7e-7}$	30^{100}_{3}	$21^{100}_{2.4}$
Blazar ROMA	$20/2170^{300}_{40}$	37^{300}_{15}	$1.5^{8}_{0.7}$	$0.7^{5}_{0.15}$	$0.06^{0.08}_{0.03}$	$0.03^{0.09}_{0.010}$	15^{50}_{7}	30^{80}_{7}	
AGN lcc	245/251	500^{2200}_{90}	170^{600}_{40}	$12^{70}_{1.2}$	$3^{20}_{0.7}$	$0.11^{0.24}_{0.04}$	$0.06^{0.12}_{0.024}$	70^{180}_{19}	60^{180}_{19}

sources in the range $20 < g < 21$ mag, where the Excess Var starts to drop to negative values. Cutting both samples at g <20 and repeating the ks test resulted in ks=0.34 and p-value=$3{\cdot}10^{-88}$, showing that the lower magnitudes in the Type 2 sample are not producing the difference in the Excess Var. Similar results are obtained by analyzing the MHPS variances computed at the intrinsic time-scales of 300 and 150 days (see Fig. 2.3b for the 300 d time-scale). For the 300 (150) d time-scale in the g band, the ks test results in ks=0.33 (0.32) and p-value=$4{\cdot}10^{-69}$ ($5{\cdot}10^{-65}$), which demonstrates again the significance of the distinction between the samples.

Finally, we examined the characteristic time-scale τ and the long term structure function (SF_∞) of the variations, given by the DRW model. In Fig. 2.3c we show the DRW parameters in the g band for the two samples. It can be seen that many points take non-physical values, with extremely small SF_∞ (for example all the cloud of points with $SF_\infty < 10^{-2.5}$ mag) or time-scales smaller than the actual data sampling (3 days bins) or larger than the maximum light curve length (1000 days). These results indicate that the DRW model is fitting noise primarily for most of the Type 2 SDSS sources, and not real variations. Moreover, the normalized histograms show a very similar behaviour for both samples except for the limit of highest SF_∞, which is significantly more populated in the Weak Type 1 sample.

To investigate whether any reliable features can be recovered from these results, we compared the DRW parameters of the Type 2 SDSS and Weak Type 1 samples within the broad but more physically meaningful limits of $0.01 < SF_{\infty,g} < 1$ mag and $1 < \tau_g < 10000$ d (black square in Fig. 2.3c). The outcome of this exercise is plotted in Fig. 2.3d. In spite of the overlap between the samples, the distribution of Weak Type 1s shifts to higher $SF_{\infty,g}$ and τ_g values than the Type 2 SDSS sample, which shows that AGNs with weak BELs in their spectra correspond to a more variable population in flux.

These results confirm the expectation that the weak BELs indeed correspond to a Type 1 activity where the variable continuum emission is also visible. Additionally, weak Type 1 AGNs can be identified through their optical light curves, allowing in principle the selection of weakly accreting black holes from existing and future large photometric variability surveys.

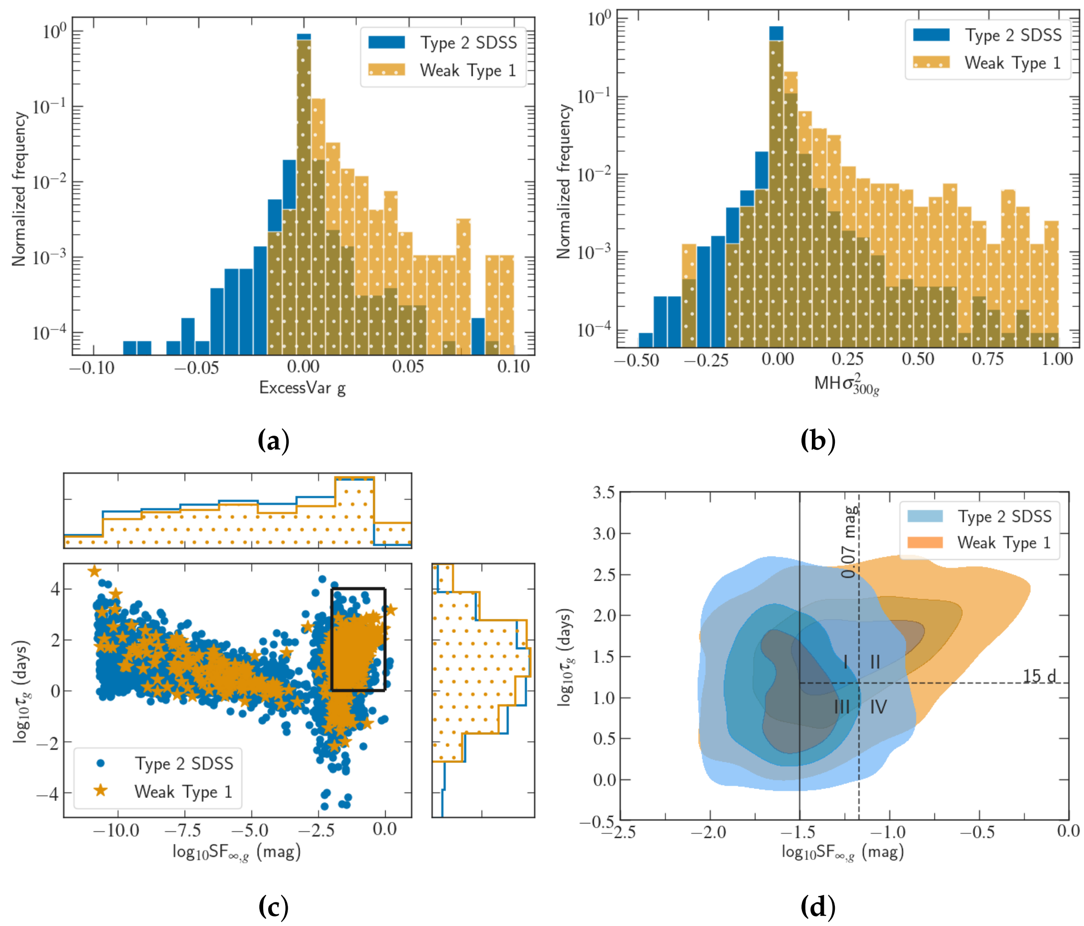


Figure 2.3: Comparison of the variability features in the g band for the Type 2 SDSS and the Weak Type 1 samples. (a) Excess variance (b) Variance at the 300 days time-scale. (c) Damped random walk (DRW) parameters: intrinsic time-scale (τ_g) of the variations versus the long term structure function ($\mathrm{SF}_{\infty,g}$). The black square indicates the region enlarged in sub-figure (d). Histograms at the sides of the plot show the normalized frequency of each parameter. (d) DRW parameters within the limits $0.01 < \mathrm{SF}_{\infty,g} < 1$ mag and $1 < \tau_g < 10^4$ d. The solid black line indicates the DRW region chosen to investigate the spectral properties of the Type 2 SDSS sample ($0.03 < \mathrm{SF}_{\infty,g} < 1$ mag), and the dashed lines show the different subdivisions used in the spectral analysis. The distinct behaviour of these variability features reinforces both samples belong to different populations.

2.3.2 Subsamples selected through variability

DRW parameters and spectral properties

The comparison between the Type 2 SDSS and the Weak Type 1 samples suggests that both classes are distributed differently in the DRW space. To evaluate this result, we can model the optical spectra of two different groups selected by their DRW parameters and determine whether their variability and spectral properties correlate.

To this end, we selected the Type 2 SDSS sources that overlap in the DRW space with the locus of the Weak Type 1 distribution at $\log_{10}\mathrm{SF}_{\infty,g} > -1.5$ (black solid line in Fig. 2.3d) with well determined g and r DRW parameters as follows:

- $g < 20$ mag, to select the brightest sources,
- $-1.5 < \log_{10}\,\mathrm{SF}_{\infty,g,r} < 0$ (most variable sources) and
- $0 < \log_{10}\,\tau_{g,r} < 3$ (significant time-scales).

We obtained 353 sources that met these requirements, which comprises 2 per cent of Type 2 SDSS sources (353/16999). In comparison, there are 28 per cent (272/978) of Weak Type 1s within this DRW space.

From the 353 Type 2s, we selected randomly 160 objects to perform the spectral analysis, 80 with $\tau_g < 15$ d and 80 more with $\tau_g > 15$ d, which corresponds to the value that separates most of the Weak Type 1 sources as shown in Fig. 2.3d. We downloaded the archival SDSS spectra available for each source and fitted the spectra using the Penalized Pixel-Fitting (pPXF) software (Cappellari, 2017). To model the spectra we used the E-MILES library (Vazdekis et al., 2016) to account for the stellar continuum component, a set of power law templates for the accretion disc contribution and two components for the emission lines, one with both permitted and forbidden lines to model the narrow emission and one just with the permitted lines to model the possible BELs. To compute the fitting errors, a total of 50 Monte Carlo simulations were performed for each spectrum using the residual of the best-fit to generate random noise. This noise was then added to the original spectrum and fitted with the same procedure.

In Fig. 2.4 we show the distribution of the broad H α equivalent width (EW H α) for the Type 2 subsamples separating them at $\tau_g = 15$ d (region I+II vs III+IV), and at the region delimited by both $\tau_g = 15$ d and $\mathrm{SF}_{\infty,g} = 0.07$ mag (region II vs III+IV). As supported by the ks test shown in Table 2.3, the distribution of EW H α is consistently different between the two samples. In both cases, the spectroscopic analysis

Table 2.3: Kolmogorov-Smirnov (ks) test for different groups of the Type 2 SDSS sample selected according to their DRW parameters.

Regions	I+II III+IV	II+ IV I+III	II III+IV	II III
N° sources	80/80	97/63	50/80	50/33
ks	0.20	0.35	0.36	0.47
p-value	0.04	0.0001	0.0003	0.0001

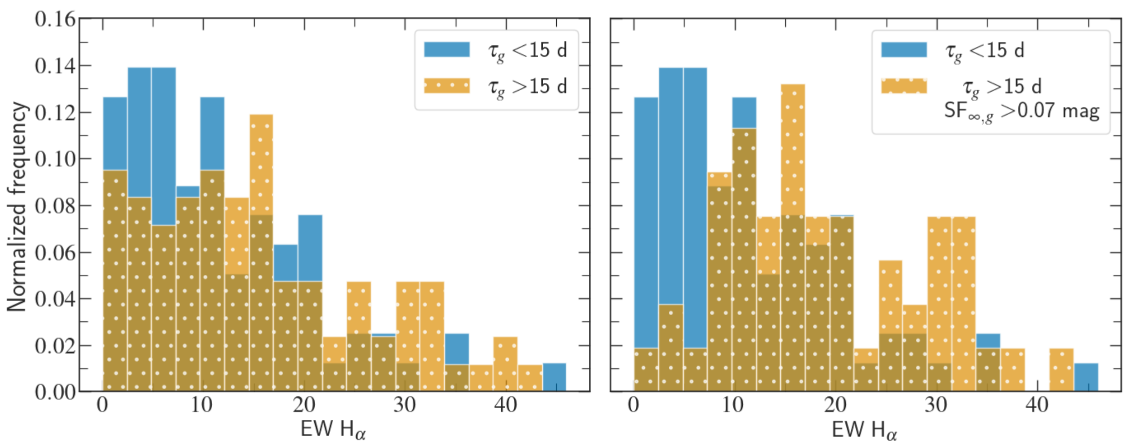


Figure 2.4: Normalized histograms for the H α equivalent width of two different subsamples of the Type 2 SDSS sample selected according to their DRW parameters. In both cases, we recover different distributions, with the EW H α of the sources outside the Weak Type 1 locus in the DRW space peaking at lower values than the sample that overlaps with the Type 1s ($\tau_g > 15$ d and $SF_{\infty,g} > 0.07$ mag).

leads to a EW H α distribution skewed towards larger values for the sources whose DRW parameters overlap those of the Weak Type 1s, with a higher significance when applying limits to both τ_g and $SF_{\infty,g}$ simultaneously (i.e., region II vs III+IV). We also evaluated the existence of a relationship between the individual EW H α and the DRW parameters themselves, but no clear correlation was found.

Most variable sources

The variability features used in this chapter have been shown to be a very powerful tool to separate different AGN populations. As a last step of the analysis, we selected the most variable sources of the entire sample to investigate their nature. We searched for current optical variability of sources that previously looked as Type 2 due to the

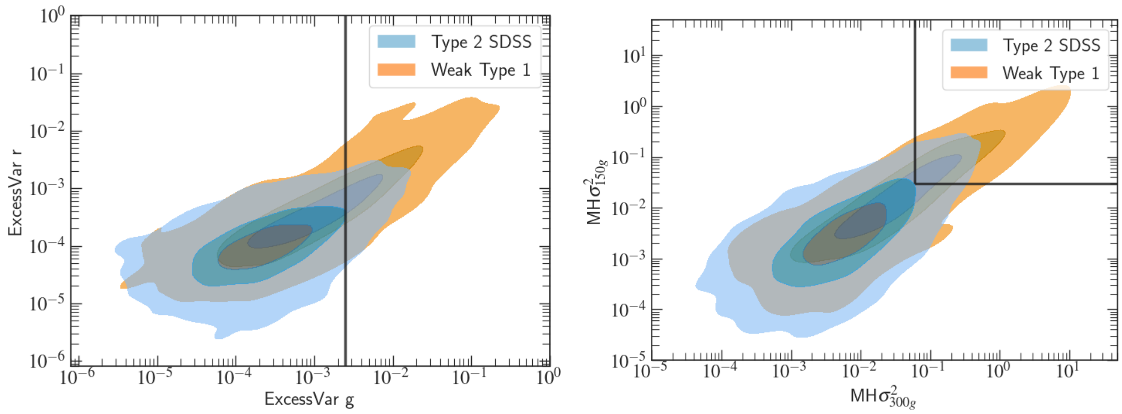


Figure 2.5: Comparison of the positive variances for two different samples of AGNs. Left: Excess variances in the r band against the g band. Right: g band variances at 150 d time-scale versus at 300 d time-scale. It can be seen that all the parameters distribute around higher values for the Weak Type 1 sample. The black solid lines indicate the limiting values of variances in the g band to select the most variable AGNs, which exclude the 67 per cent limit of Type 2 sources with positive variances.

absence of significant BELs in their spectrum, so they could be Type 1 AGNs that were misclassified, or CL/CS AGNs that have changed their type since their SDSS spectrum was taken. To determine the criteria to select the most variable sources, we first compared the positive variances for the Type 2 SDSS and Weak Type 1 samples, as illustrated in Fig. 2.5. Although we find an overlap between the distributions, we recover higher $\mathrm{ExcessVar}_{g,r}$ and MH $\sigma^2_{300,150g}$ for the Weak Type 1 sample. Here, we take as the limiting value for each variance the 2/3 level of the Type 2 SDSS sample, that is, which excludes the 67 per cent of Type 2 sources with positive variances. As shown in Fig. 2.5, the limiting values are found to be ExcessVar $g > 2.5 \cdot 10^{-3}$, MH $\sigma^2_{300g} > 6 \times 10^{-2}$ and MH $\sigma^2_{150g} > 3 \times 10^{-2}$. Then, we applied the following criteria simultaneously to select the most variable sources (in the g band):

- Mean $g < 21$ and $z < 0.3$
- ExcessVar $g > 2.5 \cdot 10^{-3}$
- $0.07 < \mathrm{SF}_{\infty,g} < 1$ mag and $15 < \tau_g < 1000$ d
- MH $\sigma^2_{300g} > 6 \cdot 10^{-2}$ and MH $\sigma^2_{150g} > 3 \cdot 10^{-2}$

We obtained 246 sources that met the requirements. From these, there are 133 that have been classified as Type 1 in Oh et al. (2015) and/or in Liu et al. (2019), 10 more

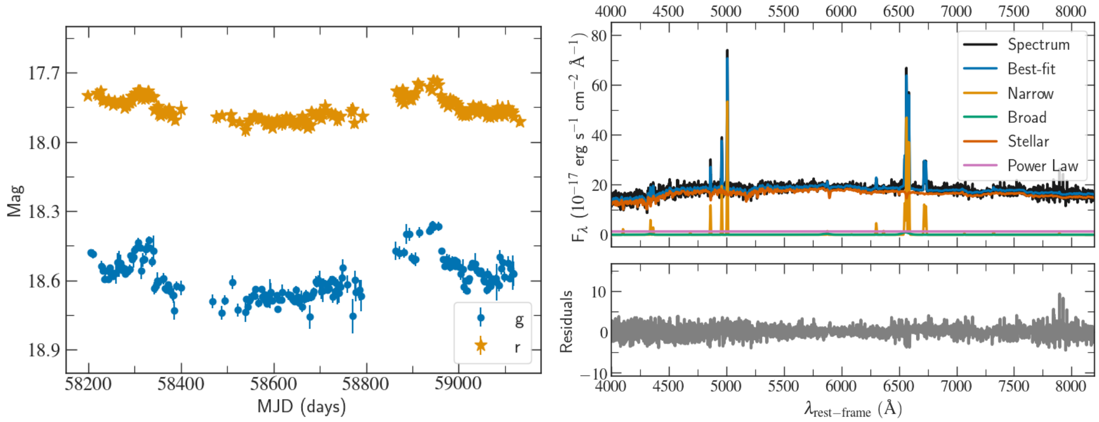


Figure 2.6: ZTF light curves (left) and fit to the archival SDSS spectrum (right, MJD=52443) of one of the CL candidates found in this chapter, J161219.56+462942.62. The light curves show significant aperiodic variability typical of Type 1 AGNs, while the spectrum shows very weak broad emission lines.

that have been subclassified as BROADLINE in the SDSS and 2 as blazar in Roma-BZCAT. From the remaining 98 sources, we discarded 12 that were classified as Sy1, QSO, LINER or BLLac in the SIMBAD database. We also checked that none of the 89 remaining sources were classified as QSO (Type 1 broad-line core-dominated) or AGN (Type 1 Seyferts/host-dominated) in The Million Quasars (MILLIQUAS) Catalog (Flesch, 2021), Version 7.5 (30 April 2022). We inspected visually the light curves of the 89 sources and we performed a spectral analysis of their archival SDSS spectra, in a similar way as described above (see Section 2.3.2). This analysis led to a list of 77 most likely misclassified Type 1 AGNs, with EW Hα >5Å (to be consistent with Liu et al. 2019's classification), and 12 CL candidates, with very weak or absent BELs (EW H α < 5 Å). As an example, in Fig. 2.6 we show the ZTF light curves and the fits to the SDSS spectrum for one of the CL candidates. The fitting results for all the sources and a brief discussion of some anomalous AGNs are presented in the Appendix A. We note, however, that some of these sources could have been already reclassified in other works that are not cited in this chapter.

2.4 Discussion

According to the simplest versions of unified models for AGNs, it is expected that optical variability in Type 2s should be suppressed due to the obscuration of the variable nuclear continuum by the dusty torus. In agreement with this picture, our results indicate that the Type 2 SDSS sample has much lower variances (given by the Excess Var and MH $\sigma^2_{300,150}$) and DRW parameters (the structure function on long time-scales SF_∞ and the time-scale τ of the variations) than the Weak Type 1 and Blazar ROMA samples that overlap with the SDSS AGN (not QSO) classification. These results are also broadly consistent with the conclusions from Yip et al. (2009), who found no evidence of continuum or emission-line variability in Type 2 AGNs on time-scales of months to a few years. Specifically, the Type 2 sources present excess variances that are fairly symmetrical around zero, which is expected for non-variable objects: $\mathrm{ExcessVar}_g = [-1000, -40, 7, 60, 700] \cdot 10^{-5}$ for the [1, 16, 50, 84, 99]-th percentiles. Similarly, for the MHPS variances at 300 d time-scale we get MH $\sigma^2_{g300} = [-12, -0.5, 0.17, 1.6, 21] \cdot 10^{-2}$. In addition, the DRW parameters take values within a very broad range of amplitudes and time-scales: $\mathrm{SF}_\infty, g = [6 \cdot 10^{-11}, 1.9 \cdot 10^{-6}, 0.018, 0.04, 0.15]$ mag and $\tau_g = [0.06, 1.1, 6, 50, 800]$d. These extremely small values up to the 50th percentile imply that the model is fitting noise primarily in these non-variable sources, and just a fraction <50 per cent of Type 2s have variations that can be characterized by physically meaningful values of amplitudes and time-scales.

Moreover, we find that the BROADLINE objects have a very similar variability behaviour to the Type 2 SDSS, which suggests the BELs that the SDSS pipeline detects are not necessarily coming from the BLR and might correspond to outflows. These results are consistent with the fact that $\sim$ 30 per cent of Type 2s at z <1 identified by Yuan et al. (2016) from the SDSS-III/Baryon Oscillation Spectroscopic Survey (BOSS) spectroscopic data base, selected on the basis of their emission-line properties, were subclassified as BROADLINE by the SDSS pipeline. We also considered to include Yuan et al. (2016)'s spectroscopically confirmed Type 2s as another subsample in the analysis, but these sources were too dim (23< g <21 mag) to get reliable variability results with the ZTF data.

On the other extreme of variability we have a subsample of objects that have been classified as AGN or Blazar according to the LCC (the AGN lcc sample). By construction, this sample is expected to be the most variable, since the ZTF produces alerts when a 5 σ variation in the template-subtracted images is detected. Accordingly, the AGN lcc sources show the highest variances and DRW parameters of all, with values up to 2 orders of magnitude above those of the Type 2 SDSS sample. For the LCC training set of the AGN class, Sánchez-Sáez et al. (2021b) considered the Type 1 Seyfert galaxies (i.e., AGNs whose emission is dominated by the host galaxy), selected from MILLIQUAS (broad type 'A'), and from Oh et al. (2015), and for the Blazar class they selected the BL Lac objects and Flat Spectrum Radio Quasars from The Roma-BZCAT Multi-Frequency Catalog of Blazars and MILLIQUAS. The reported g-band excess variances for both populations (Figure 20 in their paper) peak at $3 \cdot 10^{-5}$ (AGN) and $5 \cdot 10^{-4}$ (Blazar), which are in agreement with our results for the AGN lcc sample: $\mathrm{ExcessVar}_{LCC}, g = 1.6^{6}_{0.3} \cdot 10^{-5}$ (for the 16, 50 and 84-th percentiles).

One of the most cited works that include the time variability analysis of AGNs via the DRW is that from MacLeod et al. (2010), who model the optical variations of $\sim$9000 spectroscopically confirmed QSOs from the SDSS Stripe 82 (S82). To this end, the authors analyse SDSS *ugriz* photometric light curves with more than 60 epochs of observations over a decade. In Fig. 2.7 we show the comparison between the DRW parameters of the S82 QSOs, the Weak Type 1 and the AGN lcc samples, noting that this is an enlarged section of the entire DRW range. Since the S82 QSOs have a median redshift of 1.5, we selected the results for the infrared z band (9134 Å) to compare variations at similar emitted wavelengths to our $z < 0.3$ sources in the g band (4770 Å). As can be seen in the figure, the Weak Type 1s tend to have lower SF_∞ and τ values than the S82 QSOs. We can also see that most of the sources of the AGN lcc sample overlaps with the S82 QSOs, and just a few sources lie at lower DRW parameters. This comparison highlights the differences between variations of Type 1 objects, where AGNs with weaker BELs have lower amplitudes and characteristic time-scales than brighter QSOs. Specifically, just 16 per cent of Weak Type 1s varies strongly enough to generate alerts in the ZTF (so they overlap the AGN lcc objects) and are thus comparable to the S82 QSO sample. Incidentally, the S82 QSO sources reach larger characteristic time-scales than our objects. This is most likely due to a longer time span of the S82 light curves, which were taken in yearly seasons about 2–3 months long over the 2000-2010 decade. In fact, it has been demonstrated that the DRW τ determination is biased for light curve lengths shorter than 10 times the true τ value (Kozłowski, 2017; Sánchez et al., 2017), which is the case of our ZTF data.

Remarkably, all the variability features considered in this chapter distribute around different values for the Weak Type 1 (classified by a detailed spectroscopic analysis) and Type 2 SDSS samples. This implies the features can be used to statistically distinguish the obscured, Type 2 AGNs, from the variable, unobscured Type 1s even when the activity is weak. In particular, we tested the ability of the DRW parameters to separate the populations according to their variability. As a result, we confirmed that sources with DRW parameters that lie within the locus of Weak Type 1s in the DRW space show also a distribution of EW H α that skews towards larger

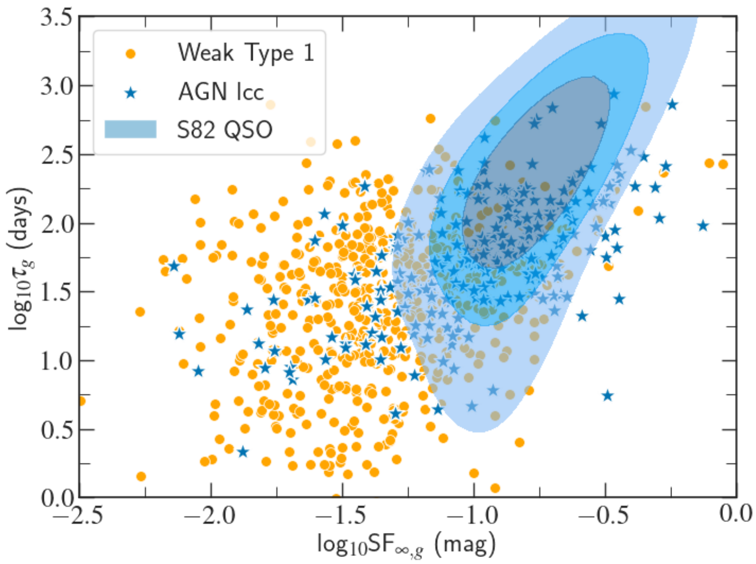


Figure 2.7: Comparison of the DRW parameters between different samples from this chapter and the SDSS Stripe 82 QSOs from MacLeod et al. (2010)

values than the sources with lower variability amplitude and shorter time-scales. We note that an overlap in the variability features is always present between these two samples, and could be occupied by intermediate type (1.8/1.9) AGNs that we have not considered in this work.

These encouraging results prove the ability of the optical variability to separate different populations. Following this idea, we can estimate the amount of variable objects in the Type 2 SDSS sample by applying limiting values in the variability features. Requiring all the variances to be positive (ExcessVar g>0 and MH $\sigma^2_{300,150g} > 0$) and the DRW parameters to be constrained within a certain range ($0.03 < \mathrm{SF}_\infty, g < 1$ mag, $1 < \tau_g < 1000$ d) leads to 1361 variable sources (∼11 per cent, 1361/12743). This value is comparable to the 10 per cent of variable sources found in a Type 2 sample in Barth et al. (2014) (17/173) and 10 per cent of *naked* AGNs reported by Hawkins (2004).

By applying more restricting limits in the variability features, we searched for the most variable, Type 1-like objects in the whole sample to investigate whether they could be misclassified Type 1s or CL/CS AGNs. We found that just the ∼1 per cent of sources (246/20583) met the requirements, of which more than half were already classified as Type 1 in different catalogues. From the remaining 89 sources, we found 77 whose SDSS spectrum was already consistent with a Type 1, suggesting that variability selection is an efficient method for identifying Type 1 contaminants in Type 2

AGN samples, as also concluded by Barth et al. (2014). We note that 32 of the most variable sources present very weak BELs (5 <EW H α <15 Å) and large variances, which could indicate a change in the accretion state of the sources, or in other words, they are CS candidates. Since several studies have reported extreme flux variability behaviour without significant spectroscopic changes (Graham et al., 2017, 2020), the CS nature of these sources needs to be verified through a new spectroscopic campaign.

Similarly, 12 of the most variable sources show no evidence of BELs in their SDSS spectra (EW Hα $<$ 5Å), which makes them potential CL/CS candidates. This selection method is very similar to the one applied in Chapter 3, where we confirm new CL AGNs found with ALeRCE. In fact, 20 of the reported CL candidates are also within our most variable sources. Here, we improve the completeness of the candidate list by characterising light curves that use the available epochs of ZTF instead of only using the ZTF alert stream. This includes more of the lower-variability AGNs, which comprise 90 per cent of known Type 1 sources. This variability-based selection method (that is, searching for current Type-1 variability in spectrally classified Type 2 AGNs) has been claimed to be one of the most successful methods in the search of CL AGNs, with a success rate (SR) of $\sim$ 60 per cent of CL confirmations (see Chapter 3). A spectroscopic confirmation is needed to determine the SR of the selection strategy presented in this chapter. In comparison, the highest SR to date is as high as 70 per cent, and has been reported by Hon et al. (2022). In the study, the authors searched for Type 1-like colours in a Type 2 sample coming from the spectroscopic Six-degree Field Galaxy Survey (6dFGS), which was observed $\sim$ 15 yr before.

2.5 Summary and conclusions

Although optical variability has been broadly used to characterize and separate different classes of objects, the variability in Type 2 AGNs has seldom been examined. In this work, we analyse systematically the ZTF light curves of $\sim$ 50 per cent (>15000) of AGNs from the SDSS DR16 to explore different variability features that allows to separate between obscured, Type 2 AGNs, from the variable, unobscured weak Type 1s. Our conclusions can be summarised as follows:

- Our results indicate Type 2 AGNs show negligible optical variations, which is consistent with the general expectations from the simplest unified models. In comparison, the variances in the g band are around 1 order of magnitude smaller than Type 1 objects with weak BELs.

- A small amount of variability features are able to separate distinct families of AGNs, including Type 1 and Type 2s AGNs.
- The characteristic time-scale τ and long term structure function $\mathrm{SF}_{\infty,g}$ of the variations, given by the DRW model, are a powerful tool to separate weak Type 1 from Type 2 AGNs. Here, we find significantly higher EW of broad H α for objects with $\tau_g > 15$ d and $\mathrm{SF}_{\infty,g} > 0.07$ mag.
- Around $\sim$11 per cent of Type 2 AGNs show evidence for optical variations, similarly to previous studies. This number of variable sources suggests a significant contamination of Type 1s or CL/CS AGNs in the DR16 AGN sample, that could be reduced by optical variability analysis.
- A spectroscopic analysis of the most variable Type 2 objects (<1 per cent, 89/12743) leads to the discovery of 77 weak Type 1 AGNs (EW Hα >5 Å) and 12 CL/CS candidates (EW Hα <5 Å). Follow-up spectroscopy would be needed to confirm the CL/CS nature of these sources and whether the weak Type 1s currently show larger BELs.

Future work with ZTF and next generation sky surveys such as the LSST, together with the use of machine learning algorithms, will allow to effectively improve the selection of pure samples of Type 2 AGNs in the optical range. This effort will be also crucial to better understand the CL/CS AGN population and the accretion physics at its most critical regimes, i.e., the advancing/receding accretion discs and the formation of the BLR clouds.

CHAPTER 3

Confirming new changing–look AGNs discovered through optical variability using a random-forest based light curve classifier

One remarkable property of active galactic nuclei (AGNs) is their variability, which is seen in a wide range of the electromagnetic spectrum. Of particular interest among AGNs with extreme variability are the changing-look (CL) AGNs, which display an appearance or disappearance of their optical broad emission lines (BELs) on time scales from months to years. Since the broad line region (BLR) is photoionized by the UV/optical continuum from the accretion disc (e.g. see the review by Netzer, 2015) dramatic variations in this waveband often occur simultaneously with the appearance/disappearance of the BELs, with a time-lag that is consistent with expectations from reverberation mapping studies (i.e light-days to light-weeks, Trakhtenbrot et al., 2019).

In the Unification Model (UM, e.g. Antonucci, 1993), AGNs can be classified as Type 1 or Type 2 depending on the orientation of the system. Type 1 refers to AGNs whose nuclei are visible, while in Type 2 AGNs a dusty structure (i.e., the torus) is expected to obscure a direct view to the accretion disc and the BLR, resulting in the absence of the disc's blue continuum and of the BELs in the spectrum. This configuration also hides the intrinsic variability of the disc emission, so the optical flux in classical Type 2 AGNs should be constant (Sánchez et al., 2017). However, the existence of CL AGNs challenges the UM, as the presence of BELs should be determined by the inclination of the source, which cannot change drastically in a few year's timescales.

The origin of the CL phenomenon in the optical range is still unclear, but nearly all observational tests have now disfavoured transient dust obscuration (as in CL AGNs in the X-ray range) or nuclear tidal disruption events (TDEs) as the source of most of the observed fading/brightening of the BELs and continuum emission (LaMassa et al., 2015; Runnoe et al., 2016; MacLeod et al., 2016, 2019; Hutsemékers et al., 2017, 2019). Interestingly, these sources could be suffering dramatic changes intrinsic to the accretion flow, in a way similar to what we observe when X-ray-binaries (XRB) go into outburst (e.g. Homan & Belloni, 2005). In this scenario, the CL AGN would make a spectral transition to/from an extreme UV bright accretion disc to a hot inner flow, which would produce a change in the BELs as a consequence of the changing shape of the ionizing spectrum. Recently, several studies have found evidence that supports the accretion state change as the physical origin of the CL phenomenon, although the observed time scale for this phenomenon is much shorter than expected if it was analogous to XRB outbursts. On the contrary, the events occur on thermal time scales, which has been associated with cooling/heating fronts propagating through the disc (Noda & Done, 2018a; Parker et al., 2019; Graham et al., 2020; Sniegowska et al., 2020). Thus, if Type 2 CL AGNs change their Type due to a change in the accretion state, they must belong to the unobscured, *true Type 2* AGN population (see Tran, 2001, and references therein).

In recent years, ~ 200 CL AGNs have been successfully identified with a variety of methods, starting with blind searches of BELs variability by comparing archival optical spectra from different epochs (Yang et al., 2018; Green et al., 2022). However, most of the CL AGNs have been confirmed via follow-up spectroscopy of candidates that had experienced great variations in any of their properties: in the optical flux (Yang et al., 2018; Frederick et al., 2019; MacLeod et al., 2019), in the colour (Hon et al., 2022), and/or in the optical and mid-IR photometric variability (Graham et al., 2020). These studies have shown the CL phenomenon is extremely rare. In a blind search for CL AGNs, Yang et al. (2018) found that just 19 out of 330795 (the 0.006 per cent) galaxies with repeated spectroscopy in the SDSS and/or LAMOST are CL AGNs. From an initial sample of 1.1 millions quasars, Graham et al. (2020) identified 47451 candidates that met specific optical and mid-IR photometric variability requirements, and found that 111 of them had significant spectral changes (the 15 per cent of their final sample with second epoch spectra). Since the origin of this variability is consistent with a change of state of activity and is not necessarily associated with a change in the optical Type but in the flux of the BELs, these sources are called ”changing-state” quasars. From this and previous works that include variability criteria to find new CL AGNs (Yang et al., 2018; Sánchez-Sáez et al., 2021a), it is clear that a diversity of phenomena can lead to extreme optical variations that are not associated with significant spectroscopic changes, and other observables are required to investigate further this behaviour.

Optimising the method to find new CL candidates will help to improve the statistics on the frequency and duration of this phenomenon. With the advent of real time, deep, large sky-coverage monitoring surveys as the Zwicky Transient Facility (ZTF, Bellm et al., 2019b) it is possible now, for the first time, to obtain a census of the rate of these changes (Sánchez-Sáez et al., 2021a). Here, we find new CL using the machine learning classifications provided by the Automatic Learning for the Rapid Classification of Events (ALeRCE, Förster et al., 2021) broker, which allows us to select a sample of 60 candidates that transitioned from Type 2 to Type 1 AGNs. To test our selection technique, we performed spectroscopic follow-up for six candidates, finding clear evidence of the appearance of BELs for at least four of them. Throughout this chapter, we assume a standard cosmological model with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

3.1 Selection of the sample

Our selection strategy is based on variability considerations: if an AGN shows rapid and strong optical variations we could have a direct view of its accretion disc and also of the BLR– if it exists. Therefore, current Type 1 variability seen in a spectrally-classified Type 2 AGN could mean that the AGN changed Type since its spectrum was taken.

In this chapter we used the current variability-based, publicly available classifications provided by the ALeRCE light curve classifier tool (LCC, Sánchez-Sáez et al., 2021b). The LCC uses a hierarchical imbalanced Random Forest classifier, fed with variability features computed from ZTF light curves and colours obtained from AllWISE and ZTF photometry, to classify each source with ZTF alerts into 15 subclasses, including variable stars, transient events, and three classes of AGNs (host-dominated or AGN, core-dominated or QSO, and Blazar). The LCC is able to separate QSO, Blazar and AGN from other stochastic sources with a 99 percent success rate. For the classification of AGN-like sources, the most relevant features are the ZTF/ALLWISE colours, features related to the amplitude of the variability at different time-scales, and features related to the time-scale of the variability.

In particular, the ZTF produces alerts of all sufficiently variable objects (5 σ variation in the template-subtracted images). Only 10 per cent of known Type 1 AGNs with $r < 21$ mag show alerts in the ZTF because the typical variability of these objects is too small to reach this threshold. Therefore, if former Type 2 AGNs with current Type 1 variability behave in the same way as the rest of the Type 1 population, we only expect to detect 10 per cent of the CLs in the sample with these data and method.

Our initial sample consists of 42027 AGNs that were classified as Type 2 in the Million Quasars Catalog (MILLIQUAS,Version 7.1, N and K Types, Flesch, 2021) or the Veron Catalog of Quasars & AGN (VERONCAT, 13th Edition, S2 and Q2 Types, Véron-Cetty & Véron, 2010). From these, we discarded sources that have been classified as Seyfert I, Low Ionization Nuclear Emission Region (LINER) or Blazar in any other study according to the SIMBAD Astronomical Database, and those without public "GALAXY" or "QSO" spectra in the SDSS database, nor subclassified as BROADLINE, which led to 22380 sources.

To find strong CL candidates, we performed a sky crossmatch within 1 arcsec between our Type 2 sample and the sources reported by the ZTF alerts that were classified primarily as AGN or QSO according to the ALeRCE LCC, which led to 60 sources. Of these, $\sim$30 possible misclassifications arise from a visual examination of the optical light curves and the SDSS spectra, so the further cleaning of the sample will be addressed in the next chapter.

This selection method results in a sample of 60 CL candidates that show a Type 1-like variability in the ZTF light curve, but that were previously classified as Type 2 by the absence of significant BELs in their spectra. We note that we are searching for AGNs that changed from a Type 2 to a Type 1 classification according to their optical spectral lines only, so the selection is not biased towards obscured/unobscured AGNs. In order to test and refine this selection method, the next step is to confirm the sample spectroscopically by quantifying the change of the BELs.

3.2 Spectroscopic follow-up

3.2.1 Observations and data reduction

With the aim of evaluating the selection method, we were allocated one night by the Chilean Time Allocation Committee (CNTAC) to observe a sample of five CL candidates with the Goodman High Throughput Spectrograph (GTHS) at the SOAR-4.1m telescope. The spectra were observed on November 2nd, 2021, using the red camera with the 400 lines/mm grating and the M2 filter (500–905 nm), a slit width of 1 arcsec, in normal readout and binning of 2x2. The total exposure time was 1h per source, divided in 3 different observations of 20 min each. On the other hand, we were also allocated observing time with FORS2-VLT at the UT1 Cassegrain focus to monitor a CL AGN event in real time. Since this AGN belongs to the same sample of candidates presented, we include here the first spectrum of the monitoring series, noting that the following spectra also contain significant BELs. The spectrum was taken on November 11th, 2021, with GRIS_300V+10 (GC435 filter), 1 arcsec slit, MIT red-optimized CCD, 100kHz 2x2 high readout mode. We took two different exposures of 12.5 min each. For the SOAR/GTHS sources, we used the master bias and flats produced by the SOAR pipeline[1], and the wavelength and flux calibration was obtained using the standard IRAF routines. In the case of the UT1/FORS2 object, we used the ESOReflex pipeline (Freudling et al., 2013). Both sets of observations were corrected by telluric bands using the spectrum of the standard star and the telluric task from IRAF. The spectra were corrected by redshift using the information from the SDSS database. The details of the observations are reported in Table 3.1.

3.2.2 Spectral analysis

We analysed the new SOAR/GTHS and UT1/FORS2 observations and the archival SDSS spectra to compare the strength of the BELs in different epochs. In Fig. 3.1 we present the plots of the average flux and difference spectra of the six CL candidates observed. It can be seen that ZTF19abixawb, ZTF20abshfkf, ZTF19aalxuyo and ZTF18accdhxv show a clear change both in the continuum and in the H α and H β BELs. To quantify the changes of the emission lines, we fitted the spectra using the Penalized Pixel-Fitting (pPXF) software (Cappellari, 2017). The SDSS spectra were fitted using the MILES library (Vazdekis et al., 2010) for the stellar pseudocontinuum

[1] https://github.com/soar-telescope/goodman_pipeline

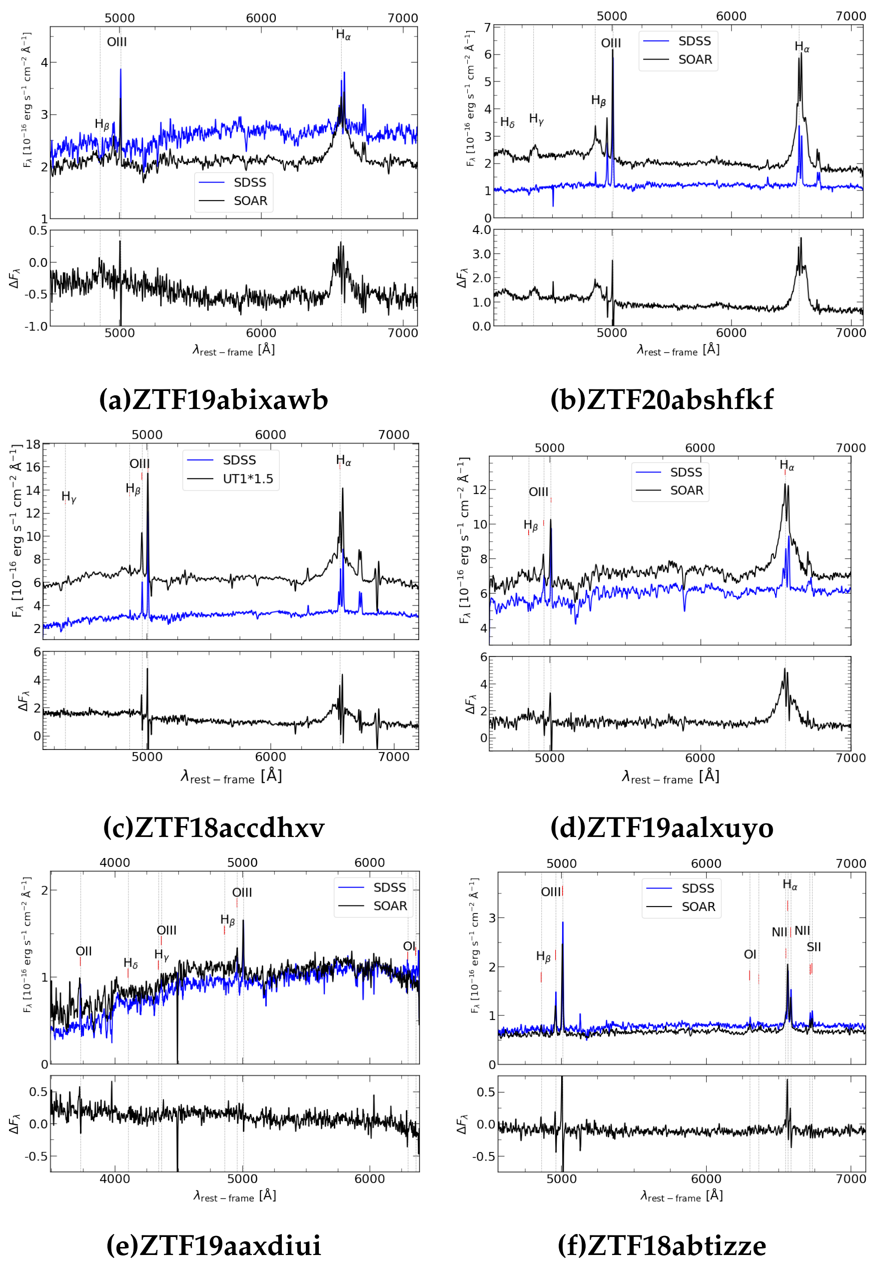
Table 3.1: Sample of the CL candidates observed with [a]: SOAR/GTHS and [b]:UT1/FORS2, and results of the spectral fitting. z: redshift from the SDSS database. EW: equivalent width of the emission lines. Δ_t: time difference between the SDSS and the SOAR/GTHS or UT1/FORS2 spectra. [c]: derived properties for new (top) and archival (bottom) spectra. Slash (/): the broad component is not needed in the fit. **Bold**: sources with $> 3\sigma$ change in the EW of H α and H β. Errors on derived quantities reflect the statistical uncertainties only, the associated systematic error is 0.5 dex on the mass and 0.3 dex on the bolometric luminosity used to calculate λ_{Edd}.

ZTF ID	SDSS name	z	epoch	SNR_{5100}[c]	EW H α[c]	EW H β[c]	H α/OIII[c]	Δ_t	log $\mathrm{M_{BH}}$	λ_{Edd}
					Å	Å		yrs	$\mathrm{M}_\odot$	
ZTF19abixawb[a]	J001014.86+000820.7	0.1022	new	33	$47.0^{+10}_{-0.3}$	$5.8^{+5}_{-0.6}$	10.1±0.5	21.17	8.0 ±0.2	0.013±0.007
			old	14	8^{+3}_{-1}	/	2.3±0.5			
ZTF20abshfkf[a]	J011311.82+013542.4	0.2375	new	24	133±2	29 ±1	5.8±0.4	6.13	8.0±0.1	0.03±0.01
				old	18	7^{+20}_{-4}	0^{+2}	0.2 ±0.2		
ZTF18accdhxv[b]	J075544.35+192336.3	0.1083	new	91	90^{+3}_{-20}	15^{+1}_{-10}	5.1±0.4	17.02	8.5±0.1	0.005±0.001
			old	20	21^{+9}_{-1}	1±1	0.97±0.08			
ZTF19aalxuyo[a]	J081240.76+071528.5	0.0849	new	36	54^{+1}_{-3}	4±1	9.8±0.3	17.64	7.76±0.04	0.027±0.005
			old	27	$2.2^{+5}_{-0.4}$	/	0.3±0.2			
ZTF19aaxdiui[a]	J214046.03+091631.7	0.4030	new	17	–	4±1	–	11.08	–	–
			old	17	–	/	–			
ZTF18abtizze[a]	J215055.73-010654.1	0.0879	new	24	13^{+1}_{-4}	/	0.42±0.05	17.39	7.4±0.2	0.005±0.003
			old	11	3^{+21}_{-1}	/	0.1±0.1			

component, a set of power law models for the accretion disc contribution and two components for the emission lines, one with both permitted and forbidden lines to model the narrow emission and one just with the permitted lines to model the BELs. The second epoch spectra were fitted with the same components, using the stellar populations models from the best-fit to the SDSS spectra. To account for the fitting errors a total of 50 Monte Carlo simulations were performed for each spectrum using the residual of the best-fit to generate random noise. This noise was then added to the original spectrum and fitted with the same procedure. The errors reported for the EW in Table 3.1 correspond to the 10 and 90 percentiles of the simulations results.
For the sources with broad H α emission, we estimated the black hole masses (M_{BH}) and continuum luminosity at 5100 Å (L_{5100}) following the approach outlined in Reines et al. (2013), using the FWHM and luminosity of broad H α obtained from the new spectra. With these values, we computed the current Eddington ratios $\lambda_{\rm Edd} = L_{\rm bol}/L_{\rm Edd}$, where $L_{\rm Edd} = 1.5 \times 10^{38}(M_{\rm BH}/M_{\odot})$ erg s^{-1} is the Eddington luminosity and $L_{\rm bol}$ is the bolometric luminosity defined as $L_{\rm bol} = 40 \cdot (L_{5100}/10^{42})^{-0.2}$ according to Netzer (2019). The results of the spectral fitting and the derived physical quantities are shown in Table 3.1.

3.3 Discussion

The selection of our CL candidates sample is based on current Type 1 optical variability reported by the ZTF and previous Type 2 spectral classification. Thus, this method leads to sources that showed very weak or absent BELs and now behave as normal Type 1 AGNs, with obvious BELs in their optical spectra. Our sample is similar to the CL AGNs identified by comparing repeated spectroscopy as in Yang et al. (2018), and differs from the changing-state AGNs in that the latter usually refer to changes in the flux of existent BELs. Our method results in 60 CL candidates, from which half appear to be misclassified based on visual inspection of their spectra and light curves, leading to a list of $\sim$ 30 promising CL candidates. Here we confirm a significant spectral change (a $> 3\sigma$ change in the EW of H α and H β) in four out of the six sources that we re-observed, which were chosen from our list of promising candidates based on observability considerations (i.e. that were observable on the same night we had been allocated). Despite the small size of this sub-sample, our first results suggest a success rate (SR) that is comparable to the 70 per cent SR reported in Hon et al. (2022), which is the most successful search method for CL AGNs to date in comparison to previous



(a)ZTF19abixawb **(b)ZTF20abshfkf**

(c)ZTF18accdhxv **(d)ZTF19aalxuyo**

(e)ZTF19aaxdiui **(f)ZTF18abtizze**

Figure 3.1: Archival (blue line) and new optical spectra (black line) for the subsample of CL candidates identified in this chapter, scaled to the flux of [OIII] λ=5007 Å in the earliest spectra. The lower plots show the difference between the new and the old spectra. All spectra are smoothed with a 5 Å box filter.

works with variability selection criteria (15-35 per cent, e.g. Yang et al., 2018; MacLeod et al., 2019).

With the results of this chapter, we can also make a lower limit estimate of the frequency of Type 2 to Type 1 transitions. There are 29943 Type 2 sources (BROAD_TYPE N or K in MILLIQUAS) that can be detected by the ZTF alert stream (dec>-28 deg and $r < 21$ mag). From these, 158 are classified as QSO or AGN by the LCC. As pointed out in Section 3.1, we need to take into consideration misclasifications in the parent sample, which could reach 30-50 per cent. This high probability comes from the fact that we are searching for oddities, so the frequency of misclassifications, which is overall small, matches the frequency of real candidates, which is also very small. Considering 50 per cent misclassifications and a minimum success rate of 2/3 (as found in this chapter) for spectra separated by an average of 15 years, we estimate a fraction of 0.18 per cent of Type 2 AGN turning into Type 1 per 15 years, or 0.01 per cent per year, that could be detected with this method. This number must be corrected by the percentage of sources that actually are variable enough to generate ZTF alerts (10 per cent for known Type 1 AGNs), leading to a lower estimate of 0.12 per cent Type 2 sources changing into Type 1 per year (1.8 per cent in 15 years). This lower limit is similar to the value reported by Hon et al. (2022), who found a minimum turn-on CL AGN rate of 3 per cent over $\sim$15 yr.

Our results indicate that the sources in this chapter are currently accreting at a few per cent L_{Edd}, which is close to the value for the state transitions observed in XRB, and suggests an accretion state change as the origin of the CL phenomena in these sources (Noda & Done, 2018a; Ross et al., 2018; Hutsemékers et al., 2019; Graham et al., 2020; Guolo et al., 2021). In these transitions, the soft/hard accretion states are caused by pronounced changes in the accretion disc contribution to the total radiative output. According to the latest AGN accretion-oriented diagrams, the soft state in AGNs would include broad-line Seyferts, showing highly excited gas and radio-quiet cores consistent with disc-dominated nuclei, while most true Seyfert 2 nuclei and the bright LINERs would show low excitation at high accretion luminosities and could be identified with the bright hard and intermediate states (Fernández-Ontiveros & Muñoz-Darias, 2021). Here, our preliminary fits to the optical spectra suggest a significant increase of 40-70 per cent in the AGN component for ZTF19abixawb and ZTF20abshfkf, in agreement with the scenario where our sources transitioned from true Type 2 AGNs, with some or negligible contribution from a cold accretion disc, to Type 1 disc dominated sources. Incidentally, the light curves of these sources show persistent stochastic variations as typical Type 1 AGNs, whereas the other two confirmed CLs, ZTF19aalxuyo and ZTF18accdhxv, show an increasing optical flux consistent with a transition to a disc dominated state. Independently of the AGN continuum contribution (that can be very uncertain), the Hα/OIII ratios of the four CL AGNs are now between 5 and 30 times larger than in the archival spectra (see Table 3.1), which most likely reveals the change in the ionising flux. On the contrary, the two candidates that were not confirmed as CL show a strong declining trend over the last 2-3 years before the new spectra were taken. This suggests that we could have missed the transitions from true Type 2 to Type 1 and again to Type 2. In fact, preliminary fits to the light curve of ZTF18abtizze point to a TDE as the origin of the optical variations, but further analysis is required to better understand the nature of this source. We note that for the other not-confirmed CL AGN, ZTF19aaxdiui, H α falls out of the observed wavelength range, but the EW H β is the same as for the CL ZTF19aalxuyo (which displays very significant changes in H α). Therefore, we cannot rule out this source is actually a CL AGN.

3.4 Summary and conclusions

We have selected a sample of 60 CL AGNs candidates that were classified spectroscopically as Type 2, but currently show significant photometric variations according to the ALeRCE LCC. To test our selection criteria, we re-observed six of these sources with SOAR and VLT telescopes. By comparing these and archival SDSS observations we find that at least four of the sources have experienced a significant ($> 3\sigma$) increase in the emission of their Balmer BELs, which implies a promising success rate of ≥ 66 per cent. Our spectral fits suggest the CL AGNs are currently accreting at a few per cent L_{Edd}, which is consistent with an accretion state change as the origin of the CL phenomena in these sources. To the best of our knowledge, this is the first time the ALeRCE broker LCC has been used to select new potential CL AGNs candidates, and we re-observed the full promising CL sample to refine the statistics on the frequency of Type 2 to Type 1 transitions, which will be presented in the next chapter.

CHAPTER 4

Improving the selection of changing-look AGNs through multi-wavelength photometric variability

For the past few years, a growing (>200) population of active galactic nuclei (AGNs) with emerging or disappearing optical broad emission lines (BELs) has been found, arousing great interest among the astrophysics community (see review by Ricci & Trakhtenbrot, 2022). Most studies favour an accretion rate change as the origin of such dramatic changes in unobscured AGNs, so these sources are often called changing-state (CS) AGNs. Other mechanisms such as variable absorption and tidal disruption events (TDEs) are also expected to produce variations in the BELs, so the term changing-look (CL) is generally used to refer to all AGNs that show such spectral transitions, regardless of the physical mechanisms driving these changes. This term is borrowed from the X-ray community, where a CL event is led by extreme variable X-ray absorption, causing a switch between Compton-thin ($N_{\rm H} < 10^{24}$ cm^{-2}) and Compton-thick ($N_{\rm H} \gtrsim 10^{24}$ cm^{-2}) states in AGNs (e.g. Matt et al., 2003).

The CL phenomenon is characterised by drastic changes in the optical BELs. The

BELs consist of permitted and semi-forbidden emission lines with typical line widths FWHM $\geq$ 1000 km s^{-1}, formed by high density gas clouds called the broad line region (BLR) located close to the central engine (e.g. Netzer, 2015). Therefore, most of the effort to find CL AGNs has focused on systematic searches of broad Balmer line variations (generally $>3\sigma$ flux change in broad H β) in sources with multi-epoch spectroscopy (although other lines such as Mg II are also possible, see MacLeod et al., 2016; Ross et al., 2018; Guo et al., 2019). In particular, some sizable samples have been found comparing repeated spectra from different surveys such as the Sloan Digital Sky Survey (SDSS) and the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) (Yang et al., 2018; Green et al., 2022).

CL events are often accompanied by large photometric changes in the optical and ultraviolet (UV) bands, and this has been used as a selection criteria to find new CL AGNs (i.e., $|\Delta g| > 1$ mag, $|\Delta r| > 0.5$ mag in MacLeod et al., 2016, 2019). However, the link between extreme spectroscopic and photometric changes is uncertain, since just 10–15 per cent of photometrically variable AGNs have been found to display CL behaviour (MacLeod et al., 2019). This uncertainty can be affected by the time-scales involved in CL events, which have been constrained just in a few sources (Trakhtenbrot et al., 2019), and the fact that the CL behaviour has been found to occur repeatedly in some sources (e.g. depending on the Eddington ratio, Guolo et al., 2021).

More recently, some studies have concentrated on the search for CL events based on the physical expectations for an accretion-state change. In this scenario, it is expected to see gradual changes in the optical and mid-infrared (MIR) flux and colours associated with monotonically varying BEL strengths and/or continuum changes, as the AGN goes to bright (AGN dominated) or dim (host dominated) states (e.g., Sheng et al., 2017; Yang et al., 2018; Lyu et al., 2022). In the optical, CS AGN candidates have been selected by searching for anomalous variability (Sánchez-Sáez et al., 2021a) and bluer optical colours in turning-on AGNs (Hon et al., 2022), where the latter method shows a higher success rate for confirmed CL compared to other selection techniques. In the MIR band, individual CL AGNs have been found by identifying highly MIR-variable quasars in the *Wide-field Infrared Survey Explorer* (*WISE*) and Near-Earth Object WISE Reactivation (NEOWISE) data stream (Stern et al., 2018; Assef et al., 2018). In Graham et al. (2020), 111 CS quasars were found by applying two different criteria: strongly enhanced optical variability over some time-scales and a large absolute difference between the initial and final state in the *WISE* light curve (i.e, $|\Delta W1| > 0.2$ or $|\Delta W2| > 0.2$). That work led to a CS sample at higher luminosity than previous CL

AGNs in the literature.

On the other hand, individual CL events have been associated with changes in soft–X-ray/UV emission, responsible for photoionizing the BLR gas, as in the case of Mrk 1018 (Cohen et al., 1986; McElroy et al., 2016; Noda & Done, 2018b). More extreme X-ray spectral and flux variability was found in the CL source 1ES 1927+654, which has been suggested to be caused by a TDE in an AGN (Trakhtenbrot et al., 2019; Ricci et al., 2020, 2021) or a magnetic flux inversion event (Laha et al., 2022). Recently, in a CL search of sources with multi-epoch optical spectroscopy within the *Swift*-BAT AGN Spectroscopic Survey (BASS), it was reported that five out of nine events with *Swift*-BAT data available could be associated with significant flux changes in the 14–195 keV hard X-ray band (Temple et al., 2023).

With the advent of deep, large sky-coverage monitoring surveys such as the Zwicky Transient Facility (ZTF, Bellm et al., 2019b) and the upcoming Legacy Survey of Space and Time (LSST, Ivezić et al., 2019), the identification of CL AGNs will be possible using machine-learning techniques. In Chapter 3, we present a method specifically looking for turn-on events using a balanced random forest algorithm with the ZTF alert stream (Sánchez-Sáez et al., 2021b), confirming CL behavior in four out of six sources that we re-observed based on follow-up spectroscopy. Extending this work further, we obtained a second epoch spectra of 30 additional CL candidates, confirming $\sim$50 per cent as CLs. In this chapter, we present the new observations and perform a multi-wavelength (optical, MIR and X-ray) variability analysis of the CL sources. This effort enables us to improve the selection technique and reinforces the common features of these CL events. Throughout this chapter, we assume a standard cosmological model with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

4.1 Selection of the sample

According to the classic view of AGNs, the optical variability in Type 2 sources is highly suppressed due to obscuration of the continuum coming from the central source by the dusty torus (Antonucci, 1993). Based on this consideration, looking for Type 1-like optical flux variability (coming from the accretion disc) in spectrally-classified Type 2 AGNs (whose accretion disc *should* be hidden) has led to the finding of turning-on CL sources. We note that in these cases, the previous Type 2 classification is due to the absence of significant BELs in their spectra, and not due to the identification of the viewing angle of the system. This is the selection strategy we followed in Chapter 3 to find potential CL candidates. Here, we update the candidate list reported in Chapter 3 and clean it further.

Our initial sample consists of all the spectrally classified Type 2 AGNs included in the Million Quasars Catalog (MILLIQUAS, Version 7.7, N and K types, Predehl et al., 2021). We removed sources classified as Seyfert 1, Low Ionization Nuclear Emission Region (LINER) or blazar in any other study according to the SIMBAD Astronomical Database, and those included in the Type 1 AGN catalogues from Oh et al. (2015) and Liu et al. (2019). We also required the sources to have a public 'GALAXY AGN' or 'QSO AGN' spectrum in SDSS DR17, and discarded objects with a BROADLINE classification. This led to 20834 Type 2 AGNs. We also checked that none of these sources were included in the Roma-BZCAT Multifrequency Catalogue of Blazars (Massaro et al., 2015).

To select potential CL candidates, we searched for current Type 1 optical flux variability in our Type 2 sample. In particular, we used the host- and core-dominated AGN classifications given by the Automatic Learning for the Rapid Classification of Events (ALeRCE, Förster et al., 2021) broker light curve classifier (LCC, Sánchez-Sáez et al., 2021b). The ALeRCE broker is currently processing the alert stream from the Zwicky Transient Facility (ZTF, Bellm, 2014; Bellm et al., 2019b) to provide a fast classification of variable objects by applying a balanced random forest algorithm. In particular, the LCC computes a total of 174 features, including colours obtained from AllWISE and ZTF photometry and variability features, for all objects that had generated at least 6 alerts in either g or r band in the ZTF alert stream. Since each alert is produced when a 5 σ variation in the template-subtracted image occurs, only sufficiently variable objects are detected and classified by the LCC. In the case of AGNs, only 10 per cent of known Type 1s with $r < 20.5$ mag exhibit variations reaching this threshold and produce alerts in ZTF. Therefore, using this method we expect to select the 10 per cent most variable CLs in the parent sample.

We performed a sky crossmatch within 1 arcsec between our Type 2 parent sample and the sources classified primarily as AGN or QSO by the LCC (updated on 2022 November 30), which led to 71 matches. Ten sources had two different identification names in ZTF, resulting in 61 CL candidates. Of these, $\sim$30 objects were identified as bad candidates from a visual examination of the optical light curves and the SDSS spectra. In the first place, some sources show BELs in their optical spectra and were apparently misclassified in the catalogues. This generally occurs in the lower-luminosity and lower-black hole-mass regimes, where the BELs fail to meet the

FWHM ≥ 1000 km s^{-1} criterion (Liu et al., 2019), and also in intermediate Type 1.8/1.9 AGNs, which show weak broad Balmer lines (e.g. Hernández-García et al., 2017). Secondly, other sources appear misclassified by the ALeRCE LCC due to a small number of data points or transient events in the reference image used to construct the difference images (mostly supernovae, Sánchez-Sáez et al., 2023).

We note that the ZTF alert light curves only contain alerts generated since the ZTF started its operations in 2018, but the actual length of the alert light curves depend on how variable the object is. We stress that the first ZTF alert does not necessarily mark the time of the change of state, considering that most known Type 1 AGNs have not triggered alerts or have taken years after the start of ZTF to show their first alert. This method does not give information about time the CL transition occurred other than that it happened some time between the SDSS spectrum was taken 10–20 years ago and the ALeRCE LCC classification as Type 1 AGN. In order to confirm the selected candidates as CLs, we need to perform optical follow-up spectroscopy that quantifies the changes in the BELs with respect to the SDSS spectrum.

In this work, we further investigate the variability and spectral properties of the most promising CL candidates to improve the selection method and shed light on the origin of these phenomena.

4.2 Spectroscopic follow-up

4.2.1 Data

We obtained second epoch spectra for 36 of our 61 CL candidates, allowing us to confirm CL objects by quantifying BEL changes with respect to archival SDSS spectra. These sources were selected via visual examination of the optical light curves and archival spectra of the CL candidates. Six out of the 36 sources were reported in Chapter 3. In this chapter, we present spectral analysis for the remaining 30 objects. The new optical spectra were taken during February and April 2022 using either the Double Spectrograph (DBSP) on the Palomar 200-inch Hale Telescope (P200) or the Low Resolution Imaging Spectrometer (LRIS) spectrograph on the Keck I telescope at the W. M. Keck Observatory, as specified in the Appendix, Table B.1. The spectra were obtained using a blue and a red arm with the 600 and 316 lines/mm grating respectively, 1.5″ slit widths and 1x1 binning, and processed using standard procedures. All the observed sources fall within the redshift range $0.04 \leq z \leq 0.22$, therefore covering the H α for all sources. In some of the cases, the H β and [O III] emission lines were in the border between the blue (2500 – 5700 Å) and red (4800 – 10700 Å) useful regions, leading to uncertain fits. Thus, we compared the broad H α and the H α/[S II] ratio between epochs, instead of the broad H β or the H α/[O III] ratio as performed in other studies (e.g. in Graham et al., 2020).

4.2.2 Spectroscopic fitting

We fit the archival DR17 SDSS spectra and the second epoch spectra from the Keck I and P200 telescopes using the Penalized Pixel-Fitting (pPXF) software (Cappellari, 2017). To account for the stellar continuum component we used the E-MILES library (Vazdekis et al., 2010), and to model the AGN emission we added the following components:

- A power low template for the accretion disc contribution of the form $(\frac{\lambda}{\lambda_N})^\alpha$ where λ is the wavelength, $\lambda_N = 5000$ Å is a normalization factor and α goes from -3 to 0 in steps of 0.1.
- One kinematic component with both permitted and forbidden emission lines, with free normalizations, to model the narrow lines.
- One kinematic component with permitted emission lines, with free normalizations, to model the possible BELs.
- One component for possible outflows with velocity dispersion values from 400 to 1000 km s^{-1}.

To fit the second epoch spectra we used the same stellar population templates obtained from the SDSS fit and we left the normalization free during the fitting process. We obtained errors for each spectrum by performing Monte Carlo simulations using the best-fit model and simulating random noise generated from the standard deviation of the best-fit residuals. Then we fit the simulated spectra using the same procedure described for the SDSS and second epoch spectra, providing an error on the model parameters.

Table 4.1 shows the results of the spectral fitting for broad H α and the 1-σ error from the simulations. We identify 13 confirmed CL AGNs with $>3\sigma$ change in the EW of broad H α *and* $>3\sigma$ change in the H α/[S II] ratio. We also identify as CL two sources with a $>3\sigma$ change in the H α/[S II] ratio, whose change can be confirmed via visual inspection. In total, we find 15 CL sources (highlighted in bold, see Table 4.1). Their optical spectra are shown in the Appendix, Fig. B.1. Some of the sources show bluer continuum emission and/or asymmetric and complex BEL profiles (e.g., ZTF18acbzrll, ZTF19aavyjdn and ZTF20aaxwxgq) during the high state, as found in previous CL works (Oknyansky et al., 2021).

For the 36 observed sources from our CL candidates, we distinguish the 19 confirmed CL AGNs as the CL sample (including the four CLs reported in Chapter 3) and the 17 not confirmed CLs as the NOT CL sample (including two such sources reported in Chapter 3).

4.3 Results

4.3.1 Improvement of the selection method: ALeRCE features for the alert light curves

All the CL candidates considered in this chapter have generated at least 6 ZTF alerts in either g or r band and have been classified by the AleRCE LCC. An alert is generated when a 5σ variation in the template-subtracted image occurs. In this section, we analyse the variability of the sample to determine whether the CL phenomenon is related to any physical parameter or if we can make other improvements to the selection method and thereby to the CL candidate list.

The LCC uses a total of 174 features, most of them computed solely with the public ZTF g and r data. The complete set of features is described in the ALeRCE website [1], and can be requested using the ALeRCE PYTHON client. In this work, we separate the features that dominate the classifier (both the 'Top' level and the 'Stochastic' level of the LCC) as reported in Sánchez-Sáez et al. (2021b) and the secondary, not-ranked features. For comparison, we obtained the features for the known AGNs that were used to train the LCC, which includes the Weak Type 1 sources from Oh et al. (2015) and the host-dominated AGNs (class 'A') from MILLIQUAS, totalling 4612 sources.

Top-ranked variability and colour features

Most of the features that dominate the LCC consist of ZTF and AllWISE colours and variability features related to the amplitude and time-scale of the variability and to a

[1] `http://alerce.science/features/`

Table 4.1: Sample of the CL candidates observed in this chapter and results of the spectral fitting. z denotes redshift from the DR17 SDSS data base. MJD, fiberid and plate denote the date and observational information from the SDSS spectra. EW denotes the equivalent width of the BELs. Sources identified as CL are shown in bold. Asterisks denote CL sources showing a $>3\sigma$ change in the H α/[S II] ratio, but with no significant changes in EW H α.

ZTF ID	RA deg	DEC deg	z	MJD	fiberid	plate	EW H α SDSS (Å)	EW H α new (Å)	H α/[S II] SDSS	H α/[S II] new
ZTF18aaiescp	207.21292	57.646792	0.13	52668	198	1158	30±2	14±1	1.75±0.12	1.14±0.07
ZTF18aaiwdzt	199.48361	49.258651	0.09	52759	390	1282	22±2	24±2	0.7±0.07	0.69±0.04
ZTF18aajywbu	205.45327	37.013091	0.20	53858	513	2101	23±4	94±2	2.4±0.4	9.01±0.23
ZTF18aaqftos	180.95505	60.888181	0.07	52405	420	954	23±3	116±3	1.81±0.21	25.09±0.74
ZTF18aaqjyon	180.42264	38.47264	0.06	53473	358	2108	6±2	0±0	0.56±0.18	0.0±0.25
ZTF21aaqlazo	170.03619	34.312731	0.04	53713	346	2100	15±1	43±1	1.79±0.11	6.43±0.2
ZTF18aavxbec*	243.08151	46.495172	0.13	52370	408	814	17±2	16±0	1.47±0.18	3.14±0.1
ZTF18aawoghx	156.10498	37.650863	0.10	52998	254	1428	9±2	34±1	0.91±0.2	3.95±0.17
ZTF18aawwcaa	128.10120	35.859979	0.14	52668	459	1197	26±2	73±2	5.49±0.77	16.74±1.06
ZTF19aabyvtv	197.87773	31.866893	0.07	53819	632	2029	17±1	6±3	17.57±3.38	8.09±9.03
ZTF18acbzrll	124.82294	30.32660	0.10	52619	94	931	7±2	90±3	1.29±0.43	20.13±0.88
ZTF18acgvmzb	148.96896	35.965616	0.04	52998	251	1596	8±1	8±1	0.52±0.03	0.73±0.04
ZTF18achdyst	157.39925	24.777606	0.11	53734	545	2349	26±1	15±1	5.72±0.02	1.37±0.1
ZTF18acusqpt	177.91898	12.036714	0.07	53142	116	1609	18±1	17±2	3.27±0.32	6.07±0.65
ZTF19aafcyzr	125.71008	15.673859	0.12	53713	517	2272	11±1	30±1	1.2±0.17	3.98±0.24
ZTF19aaixgoj	146.80538	12.205624	0.12	53053	575	1742	14±1	13±1	2.65±0.29	2.61±0.24
ZTF19aaoyjoh	180.18956	14.967685	0.11	53463	158	1763	10±5	59±1	1.25±0.37	10.44±0.31
ZTF19aapehvs	199.74519	57.501847	0.10	52759	338	1320	32±2	18±1	3.02±0.27	2.34±0.17
ZTF19aavqrjg	181.24583	15.58718	0.22	53467	447	1764	51±3	67±9	35.65±2.0	10.61±10.55
ZTF19aavyjdn	202.39941	-1.509453	0.08	52377	524	910	14±4	77±18	2.94±0.63	31.1±6.41
ZTF20aaeutuz	164.81581	12.483378	0.15	55956	726	5357	30±1	34±1	2.83±0.11	5.31±0.2
ZTF20aagwxlk	153.42829	55.432205	0.15	52407	179	946	31±2	27±8	3.69±0.29	1.14±0.64
ZTF20aagyaug	172.41282	36.883602	0.20	55673	426	4648	41±3	96±2	1.35±0.09	4.17±0.11
ZTF20aakreaa	181.52333	42.169888	0.10	53120	149	1448	17±1	20±10	4.85±0.92	1.03±2.34
ZTF20aaorxzv	132.39052	3.68048	0.08	52224	606	564	18±1	16±16	1.63±0.16	1.04±0.31
ZTF21abwoxbv	234.63610	46.126392	0.20	52781	559	1332	11±3	68±10	2.47±0.74	56.73±9.03
ZTF20abcvgpb*	238.24977	21.046358	0.17	53557	220	2171	49±4	44±1	2.98±2.3	11.77±0.69
ZTF20abgnlgv	232.52715	7.172269	0.13	54208	322	1820	41±2	35±1	3.98±0.38	4.8±0.17
ZTF21aafkiyq	174.93478	-1.727439	0.11	52294	614	327	18±2	54±1	1.79±0.23	4.42±0.17
ZTF21abcsvbr	184.64829	18.771713	0.22	54477	299	2611	19±4	44±1	1.37±0.25	5.35±0.29

Table 4.2: Features that show different distributions for the CL and the NOT CL samples. Features recovered for both the alert light curves and the forced-photometry light curves are highlighted in bold.

Name	Filter	Description
		Top ranked variability and colour features used in the LCC
SF_ML_amplitude	*g*	rms magnitude difference of the structure function computed over a one year time-scale
MHPS_low	*g*	Variance associated with the 100 days time-scale ('low' frequency)
SPM_tau_rise	*g*	Initial rise time-scale from the supernova parametric model
GP_DRW_sigma	g	Amplitude of the variability at short time-scales, from the damp random walk (DRW) model
GP_DRW_tau	g	Relaxation time from the DRW model
IAR_phi	g	Level of autocorrelation from an irregular autoregressive (IAR) model
positive_fraction	*g*,*r*	Number of detections in the difference images that are brighter than the template
delta_mag_fid	*g*	Difference between the maximum and minimum observed magnitudes in a given band
r–*W3*		colour computed using the ZTF mean *r* magnitude and the AllWISE *W3* filter
g_r_mean_corr		ZTF *g*-*r* colour using the mean magnitudes of each band
g_r_max		ZTF *g*-*r* colour using the brightest magnitudes of each band
		Other features used in the LCC
n_pos	*g*,*r*	Number of positive detections in the alert light curve
n_neg	*g*,*r*	Number of negative detections in the alert light curve
iqr	*g*	Difference between the 3rd and the 1st quartile of the light curve
MHPS_ratio	*g*	Ratio between the variances at 100 days and 10 days time-scales for a given band, applying a Mexican hat filter
		Variability features for the forced-photometry light curves
SPM_A	*r*	Amplitude from the supernova parametric model
LinearTrend	*g*,*r*	Slope of a linear fit to the light curve
ExcessVar	*g*	Intrinsic variability amplitude
Meanvariance	*g*	Ratio of the standard deviation to the mean magnitude
Std	*g*	Standard deviation of the light curve
Amplitude	*g*	Half of the difference between the median of the maximum 5 per cent and of the minimum 5 per cent magnitudes
SPM_tau_rise	*g*, *r*	See above
MHPS_low	*r*	See above
MHPS_ratio	*r*	See above
GP_DRW_sigma	*g*,*r*	See above
SF_ML_amplitude	*g*,*r*	See above

decrease/increase of the luminosity. In order to evaluate the difference in distribution of these features between the CL and NOT CL samples we applied the Kolmogorov-Smirnov (KS) test to all their ranked-features. In Table 4.2 we present the features that dominate the LCC and have a p-value <0.05, that is, where we can reject the null hypothesis that the two distributions (from the CL and the NOT CL samples) are identical.

We note that the DRW parameters determination is generally biased for light curve lengths shorter than 10 times the true τ value (Kozłowski, 2017; Sánchez et al., 2017), which is the case of our ZTF data. Therefore, the DRW parameters obtained in this chapter should be considered just as variability features and not as physically correct estimations.

In Fig. 4.1 we show the distribution of some of the variability features computed in the g band that present different distributions between the CL and NOT CL sources. In particular, the rise time-scale from the supernova (SN) parametric model (SPM_tau_rise) seems to separate both samples, with the NOT CL objects placed at

SPM_tau_rise < 30 d. These lower values are similar to characteristic SN time-scales, which suggests the variability of some NOT CL could be due to transient events such as SN or TDEs. Along the same lines, the DRW relaxation time for the NOT CL objects peaks at the minimum data sampling (~ 1 d), and spreads up to $>$1000 d, larger than the maximum light curve length. This indicates a DRW model is unable to properly model the optical variability for some NOT CLs, and thus it is unlike Type-1 AGN. For the CLs however, the DRW relaxation time peaks at 10–100 days as expected for Type 1 AGNs. In terms of the amplitude of the variability, from the GP_DRW_sigma distributions we see that some NOT CL objects reach much smaller values (log10(GP_DRW_sigma) < -6), which again indicates they have most likely been misclassified as Type 1 AGN by the LCC. Interestingly, the amplitude of the variations for all our objects peaks at a smaller value than the distribution for the AGN training set, suggesting their variability could be diluted by the host galaxy contribution. On the other hand, the autocorrelation of the light curves, given by the IAR_phi parameter, reaches smaller values for the NOT CL sample than for the CL sample. This feature, together with the SPM_tau_rise parameter, could be used to further clean the CL candidate list.

Apart from the variability features, the classifier is also dominated by ZTF and AllWISE colours and the morphological properties of the images. In general, the optical colours for both samples look similar to each other but show a redder tendency than the AGN training set distribution, as shown in Fig. 4.2a.

In Fig. 4.2b we present the 2010-2011 AllWISE *W1*–*W2* versus the *W2*–*W3* colours for the CL and NOT CL objects, in comparison to the Type 2 parent sample and the AGN training set. Most of our objects have fairly similar AllWISE colours, but the distribution of the *W1*–*W2* colour peaks at a lower value than the AGN training set and closer to the Type 2 distribution, which implies that the MIR colours are generally dominated by the stellar populations. We note that the AllWISE observations were taken between 2010 and 2011, so these features are not indicative of their *current* MIR colour. We further investigate the behaviour of the MIR colours in Section 4.3.2, including contemporaneous *WISE* observations.

Not-ranked features

The top-ranked features are the most important features that the LCC uses for the classification of variable sources. However, there are secondary features that could potentially allow us to evaluate whether a source is a good CL candidate or not. In

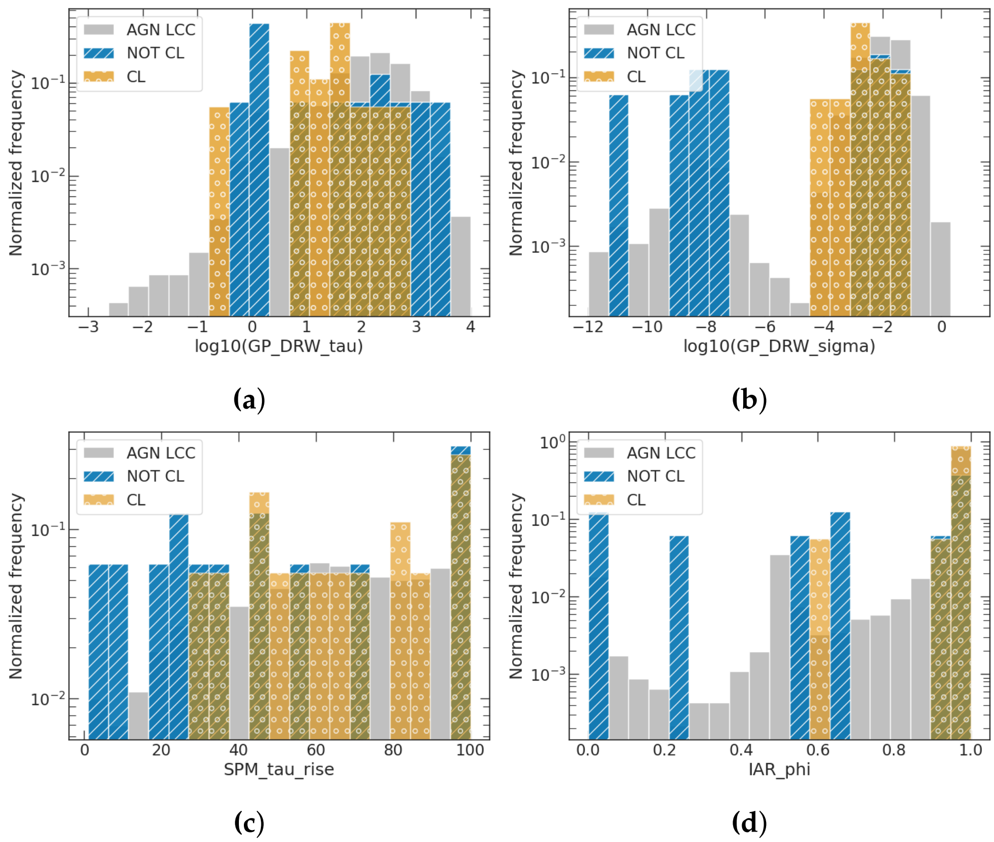
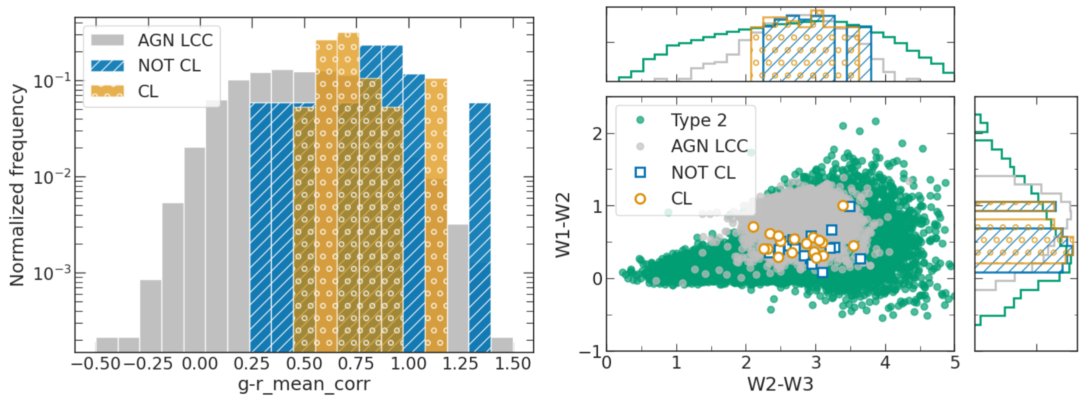


Figure 4.1: Normalised distributions of the alert light curves top-ranked variability features for the CL and NOT CL samples and the AGN LCC training set: (a) relaxation time and (b) amplitude of the variability on short time-scales, obtained from the damp random walk (DRW) model; (c) initial rise time-scale from the supernova parametric model and (d) level of auto-correlation from an irregular autoregressive (IAR) model. These features show distinct distributions between the CL and NOT CL samples and could be used to select the most promising CL candidates.



(a) **(b)**

Figure 4.2: Alert light curves top-ranked colour features for the CL and NOT CL samples and the AGN LCC training set: (a) g-r colour obtained with ZTF data and (b) AllWISE MIR colours in comparison with the Type 2 parent sample. Both the CL and NOT CL samples show optical and MIR colours more host-galaxy dominated than the AGN LCC training set.

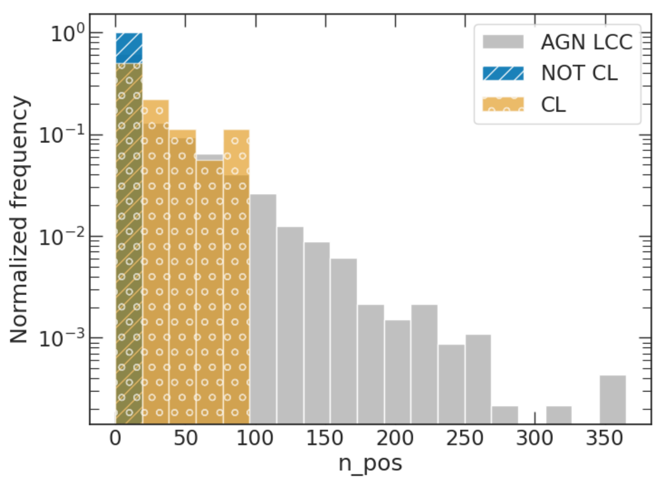


Figure 4.3: Number of positive detections in the alert light curves for the CL and NOT CL samples and the AGN LCC training set, where the CL sample tends to increase their optical flux.

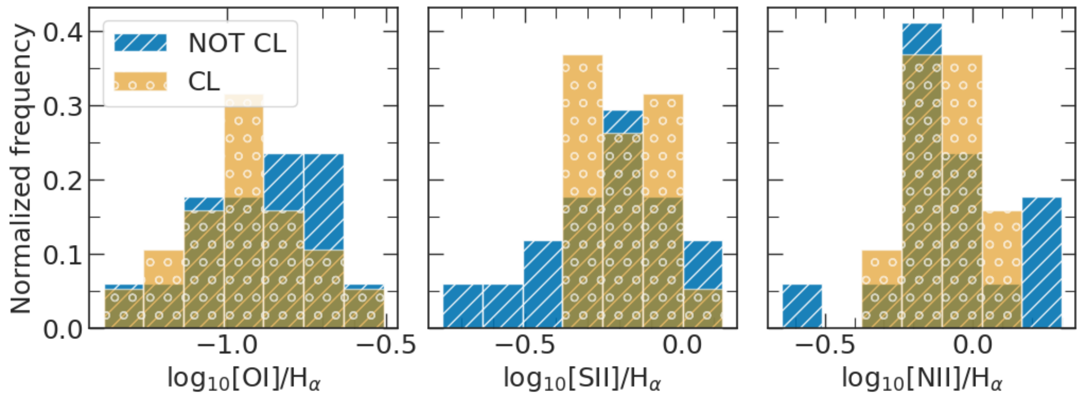


Figure 4.4: BPT line ratio distributions for the CL and NOT CL samples obtained from their archival SDSS spectra. The comparison indicates these samples are indistinguishable in terms of the three BPT diagnostic criteria.

Table 4.2 we present other secondary features showing distinct distributions between the CL and NOT CL samples. A particular example is the number of positive detections in the alert light curves (n_pos, see Fig. 4.3), which reaches higher values for the CL sample, meaning that CL objects tend to increase more their flux with respect to the template image (as expected for turning-on events).

4.3.2 Characteristics of the CL vs NOT CL AGNs

BPT diagnostics

To investigate the emission-line properties of the samples, we calculated the BPT (Baldwin et al., 1981) diagnostics from the archival SDSS spectra. We used the classification system defined by Kewley et al. (2006) for different ionisation mechanisms utilizing the three BPT diagnostic criteria ([N II], [S II], and [O I]) and found all sources are consistent with a Seyfert classification. The comparison of the line ratios for the CL and NOT CL samples is plotted in Fig. 4.4. The KS-test leads to large p-values >0.4 for the three cases, showing that these samples are indistinguishable in terms of the emission lines properties from their old (pre-CL) optical spectra. Similar results were found by analysing second epoch spectra.

Eddington ratio estimates

We estimated the black hole masses ($M_{\rm BH}$) and continuum luminosity at 5100 Å (L_{5100}) using the full width at half maximum (FWHM) and luminosity of broad H α as out-

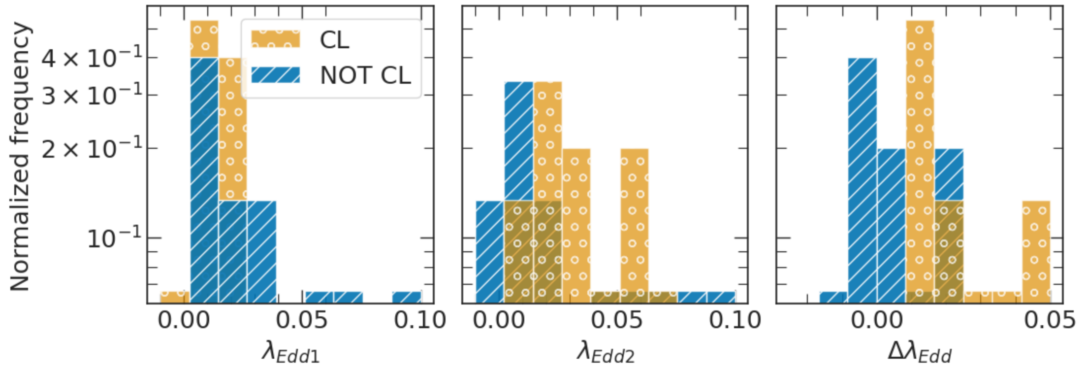


Figure 4.5: Eddington ratios for the CL and NOT CL samples obtained from the archival SDSS spectrum (left panel) and current spectra (middle panel), and their difference ($\Delta\lambda_{\rm Edd}$ = $\lambda_{\rm Edd2}$ - $\lambda_{\rm Edd1}$, right panel). CL objects have increased their Eddington ratios and are now accreting at 1–5 per cent $L_{\rm Edd}$.

lined in Reines et al. (2013), using the values obtained from the new spectra. Then, we computed the Eddington ratios for the old and new spectra $\lambda_{\rm Edd} = L_{\rm bol}/L_{\rm Edd}$, where $L_{\rm Edd} = 1.5 \times 10^{38}(M_{\rm BH}/M_{\odot})$ erg s^{-1} is the Eddington luminosity and $L_{\rm bol}$ is the bolometric luminosity defined as $L_{\rm bol} = 40 \cdot (L_{5100}/10^{42})^{-0.2}$ erg s^{-1} according to Netzer (2019). In Fig. 4.5 we present $\lambda_{\rm Edd}$ for both samples computed for the old ($\lambda_{\rm Edd1}$) and new spectra ($\lambda_{\rm Edd2}$), and the difference of accretion rate ($\Delta\lambda_{\rm Edd}$). We find that the old accretion rate is similar for both samples, but in the second epoch spectra the distribution shifts towards higher values for the CLs. These are the expected results, since both the method to compute $\lambda_{\rm Edd}$ and the criteria to confirm the sources as CL use the properties of broad H α.

ALeRCE features for the forced-photometry light curves

To compare with the alert light curves, we also analysed variability in the full forced-photometry light curves, which are produced based on all ZTF difference images available. We requested the most updated forced photometry (up to 2022 September 26) of the entire sample from the ZTF forced-photometry service and generated the cleaned light curves according to the recommendations outlined in Masci et al. (2019). A PYTHON library to extract variability features in astronomical light curves is publicly available[2]. The forced-photometry light curves have a mean of 376 (453) data points in the g (r) filter, in comparison with the 30 (27) detections from the alert

[2] https://github.com/alercebroker/lc_classifier

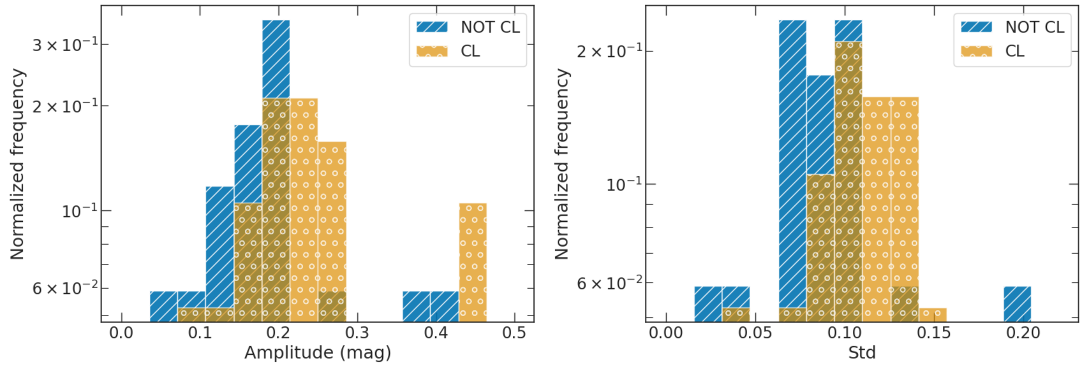


Figure 4.6: Amplitude (left) and standard deviation (right) in the g filter for the CL and NOT CL forced-photometry light curves. CL objects present higher values for both features, indicating a stronger optical variability.

light curves. As a result, the variability features can be better constrained and we find more top-ranked features that show different distributions according to the KS-test between the CL and NOT CL samples, as shown in Table 4.2. We recover many variability features related to the amplitude of the variations and the deviations from the mean (e.g. ExcessVar, Meanvariance or Std), which are missing in the comparison of features from the alert light curves. Fig. 4.6 shows the amplitude and the standard deviation in the g filter for the forced-photometry light curves for both samples. The CL objects present generally higher values for both features, indicating that their variability is more similar to the expected Type 1 behaviour than the NOT CL sample.

Mid-infrared variability

Some of the features that the LCC uses to classify variable objects are computed with AllWISE data, which are indicative of the state of the sources ten years ago. To investigate whether the CL and NOT CL samples show distinct MIR behaviour we downloaded all the AllWISE multi-epoch and NEOWISE-R single exposure (L1b) photometric data spanning from 2010 to 2021 and averaged every six months. For each source, we obtained the following variability features for the *W1* and *W2* bands: the maximum magnitude and colour variations (Δ*W1*, Δ*W2*, Δ*W1*–*W2*), the colour from the last epoch (*W1*–$W2_f$), the intrinsic variability (σ_{m1} and σ_{m2}) computed as in Lyu et al. (2022) and the slopes of a linear trend fit to the MIR magnitude light curves (a_1 and a_2) and to the *W1*–*W2* colour (a_{12}). Table 4.3 shows the comparison between the median values with the 1-σ errors for the CL and NOT CL samples. All the features except for the maximum colour variation Δ*W1*–*W2* show distinct distributions according to the

Table 4.3: Mid-infrared variability features. Asterisks indicate the features that show distinct distributions between the samples according to the KS-test (p-value<0.05). The errors correspond to the 1-sigma deviation from the median.

Feature	CL	NOT CL
$\langle \Delta W1 \rangle$ (mag)*	$0.5^{+0.3}_{-0.1}$	0.4 ± 0.2
$\langle \Delta W2 \rangle$ (mag)*	0.7 ± 0.3	$0.5^{+0.1}_{-0.3}$
$\langle \Delta W1–W2 \rangle$(mag)	$0.3^{+0.2}_{-0.1}$	$0.3^{+0.2}_{-0.1}$
$\langle W1–W2_f \rangle$(mag)*	$0.5^{+0.2}_{-0.1}$	$0.4^{+0.1}_{-0.2}$
$\langle \sigma_{m1} \rangle$ *	0.15 ± 0.05	0.09 ± 0.07
$\langle \sigma_{m2} \rangle$*	$0.18^{+0.09}_{-0.05}$	$0.14^{+0.04}_{-0.09}$
$\langle a_1 \rangle \cdot 10^{-5}$ (mag^{-1})*	-4^{+4}_{-9}	0.5^{+4}_{-5}
$\langle a_2 \rangle \cdot 10^{-5}$ (mag^{-1})*	-6^{+9}_{-13}	2^{+4}_{-8}
$\langle a_{12} \rangle \cdot 10^{-5}$(mag^{-1})*	2^{+3}_{-4}	0^{+3}_{-2}

KS-test, with CL sources having a stronger variability. Moreover, the results from the linear fits indicate that the CLs have become brighter in both bands and have higher *W1*–*W2* values (see Fig. 4.7 and B.1), whereas for the NOT CL the distributions peak closer to zero resulting in no net increase or decrease in brightness or colour.

X-ray variability

In order to obtain the X-ray fluxes of our sources, we have used the individual eROSITA (extended ROentgen Survey with an Imaging Telescope Array, Predehl et al., 2021) All-Sky Surveys (eRASS1 to eRASS5). The data were processed with the eROSITA Standard Analysis Software System (eSASS, Brunner et al. 2022). We used the newest available pipeline processing version c020 which is an updated version of the software used for the first eROSITA Data Release (Merloni et al. 2023, in prep).

The counterparts are determined using the same procedure adopted in the eROSITA/eFEDS field (Salvato et al., 2022), but applied to Legacy Survey DR10[3] and Gaia DR3 separately. After the identification of the CL candidates with the counterparts, we obtained the 0.2–2.3 keV flux from the corresponding eROSITA catalog (see Brunner et al. 2022 for a description of the eROSITA catalog processing).

From our list of 61 CL candidates, there are 28 sources within the eROSITA-DE footprint (Galactic longitude $179.9442 < l < 359.9442$ deg): 11 CLs, seven NOT

[3] `https://www.legacysurvey.org/dr10/`

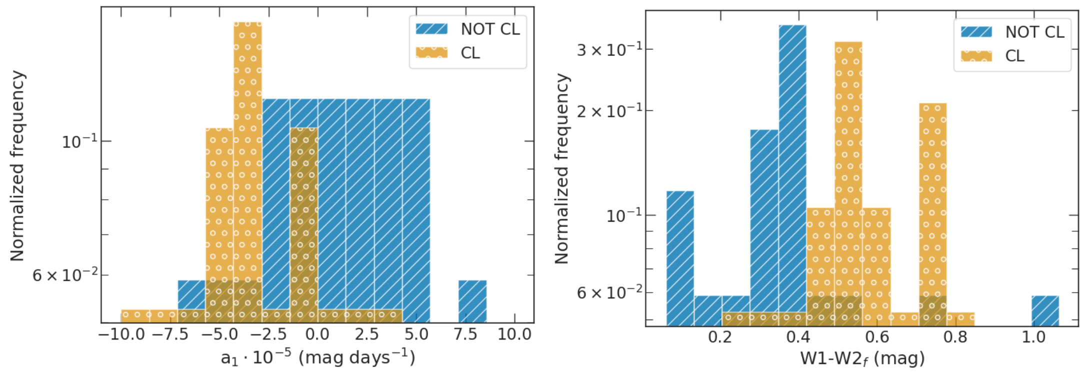


Figure 4.7: Distributions of the linear slope of the *W1* 10-years-long light curve (a1, left) and the last epoch colour (*W1*–*W2*$_f$, right) for the CL and NOT CL samples. The lower a1 values for the CL sample indicates the sources are getting brighter in the MIR waveband, while the slopes for the NOT CL distribute around zero, indicating that as a sample they are neither brightening nor dimming. The higher last epoch colour (*W1*–*W2*$_f$) for the CLs is expected for AGN-dominated galaxies (*W1*–*W2* >0.5).

CLs and ten CL candidates without a second epoch optical spectrum. From the CLs, ten sources have at least one detection within the five different eRASS, and one (ZTF18aawoghx) has only upper limits. The upper limits are calculated based on X-ray photometry on the eROSITA standard pipeline data products (science image, background image, and exposure time) following the Bayesian approach described by Kraft et al. (1991). For details about eROSITA upper limits, see Tubín-Arenas et al. (2023, in prep). We consider a circular aperture with a radius given by a PSF encircled energy fraction of EEF = 0.75 ($\sim 30''$) and a single-sided 3σ confidence level. From the NOT CLs, four sources have been detected in at least one eRASS. The remaining three (ZTF18acgvmzb/ZTF18aclfugf, ZTF19aaixgoj, and ZTF20aaorxzv) have only upper limits. Interestingly, six CL sources show an X-ray flux increase between eROSITA scans by factors of 2 to 15 times. For two sources (ZTF19aavyjdn and ZTF21abcsvbr) the difference between the maximum and the minimum values is similar to the error of the minimum value. For the remaining four sources (ZTF18accdhxv, ZTF19aalxuyo, ZTF21aafkiyq and ZTF21aaqlazo) the difference is at least eight times the error (see Fig. 4.8).

We also checked archival X-ray fluxes from other missions to compare to the eROSITA fluxes. All the 11 CLs and 7 NOT CLs in the eROSITA-DE footprint have at least one X-ray upper limit from either the *XMM-Newton* Slew (Saxton et al., 2008) or *ROSAT* Survey (Boller et al., 2016). However, due to the low sensitivity of the data,

ZTF ID	*ROSAT* flux $\cdot\, 10^{-13}$ erg s^{-1} cm^{-2}	eROSITA flux $\cdot\, 10^{-13}$ erg s^{-1} cm^{-2}
ZTF18accdhxv	$<$2.853	5.0 ± 0.8
ZTF19aalxuyo	$<$3.279	4.6 ± 0.8
ZTF21aafkiyq	$<$1.049	7.7 ± 0.9
ZTF21aaqlazo	0.51 ± 0.06	11.1 ± 0.1

Table 4.4: 0.2–2.3 keV X-ray fluxes for four CL sources that show an increase between the archival 1990–1993 *ROSAT* 1σ upper limits or fluxes and the 2021 eROSITA data.

most of the archival upper limits fall above the current eROSITA measurements. This hinders us from finding the possible changes, with the notable exception of four of the CLs that show a significant ($\geq 2\sigma$) increase in the eROSITA 2021 flux with respect to the archival 1990–1993 *ROSAT* 1σ upper limits or fluxes, which are shown in Table 4.4. In the table, we converted the observed *ROSAT* fluxes to the 0.2–2.3 keV band for a direct comparison, using an absorbed power law model with photon index $\Gamma = 2$ and column density $N_{\mathrm{H}} = 3 \cdot 10^{20}$ cm^{-2}. These sources are also the CLs that experience a significant X-ray increase during the eROSITA monitoring as shown in Fig. 4.8.

To compare the eROSITA fluxes between the CL and NOT CL sources, we selected the last X-ray detection within the five eRASS, or the last eROSITA upper limit for the sources without detections. We also computed the ratio between the X-ray flux and the optical flux in the g band, obtained from the ZTF forced photometry light curves. To avoid spurious results coming from variability, we chose pairs of contemporaneous fluxes, i.e., that were taken within the same days or week in the X-ray and optical bands. The results are plotted in Fig. 4.9, which shows the CL sources are generally brighter in the X-ray band than the NOT CL sources, both in absolute terms and relative to their g band fluxes. The KS-test indicates the X-ray flux distribution is significantly distinct between the CL and NOT CL samples (p-value < 0.05), both considering just detections (p-value $= 0.01$) and considering detections and upper limits (p-value $= 0.01$). However, the X-ray to optical ratio distributions are not significantly distinct according to the KS-test, either considering just detections (p-value $= 0.08$) or considering detections and upper limits (p-value $= 0.10$). Therefore, although the X-ray to optical ratios tend to be higher for the CLs than for the NOT CLs, the difference is not statistically significant for the sources considered in this work, and more extended samples are needed to improve the statistics in terms of the X-ray behaviour for CLs.

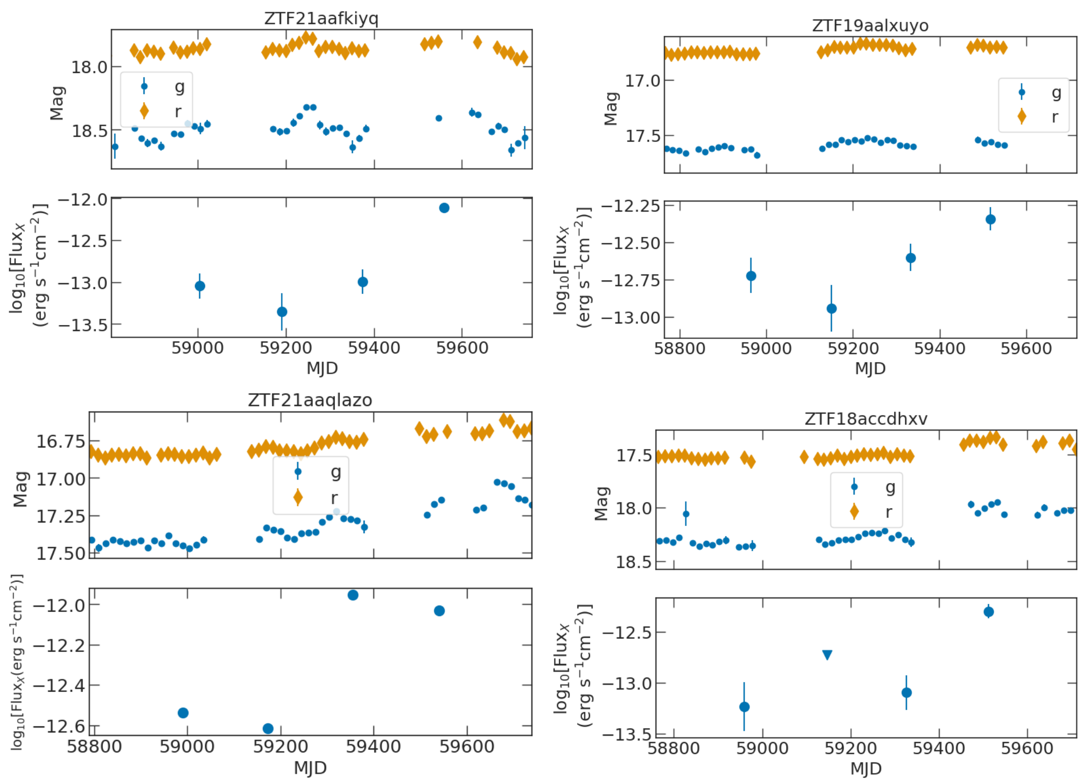


Figure 4.8: ZTF forced-photometry light curves and contemporaneous eROSITA fluxes in the 0.2–2.3 keV band. The triangle in the last plot indicates an upper limit. These sources experience an increase in their X-ray flux during the eROSITA monitoring.

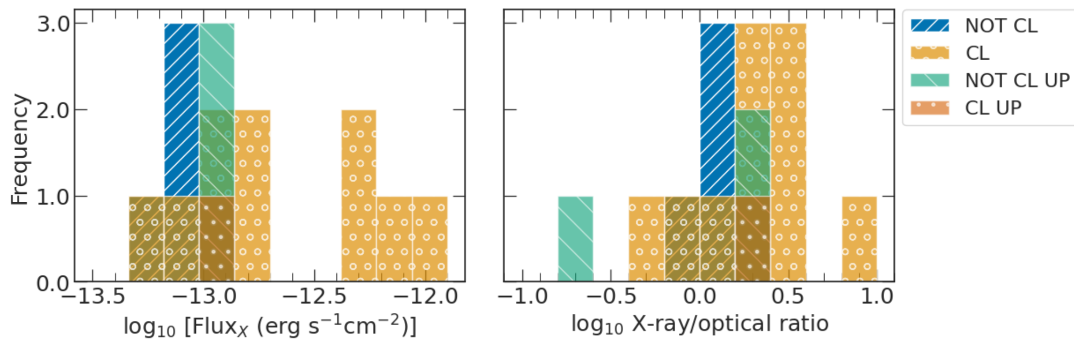


Figure 4.9: Distributions of the 0.2–2.3 keV eROSITA fluxes (left) and ratios between the 0.2–2.3 keV flux and the g-band flux from the ZTF forced photometry light curves, taken within the same day or week (right). *UP* indicates the eROSITA upper limits.

As a final step we also checked the harder, 2.3–5 keV eROSITA fluxes. Most of the sources have just upper limits in this band, thus we cannot draw further conclusions about the X-ray spectral shape.

4.4 The origin of the NOT CL sources

The previous section showed that the CL and NOT CL samples are significantly different in terms of their optical, MIR and X-ray flux variability properties, with the CLs showing stronger optical and MIR variability with a tendency to increase their MIR flux and colour. In order to understand the nature of the NOT CL sources we visually inspected the ZTF forced photometry and alert light curves and the reference images used by ZTF to compute the difference images. In Table 4.5 we present the characteristics of the variability in individual NOT CL objects, indicating the most probable cause of their variations. As a result, we find two sources that have been most likely misclassified by the LCC due to a small number of alerts (ZTF20aakreaa and ZTF20aaorxzv) and another two sources are possibly false detections due to a bad template image subtraction (ZTF18acgvmzb and ZTF18acusqpt). Three sources apparently show a transient event (that is, a flare-type variation in an otherwise flat curve) in the optical light curves: two of them resemble SN events (ZTF18aaiwdzt and ZTF18aaqjyon) and one shows a long decay that we speculate could be due to a TDE (e.g. a slowly-decaying TDE as in Lin et al., 2017). The occurrence of TDEs in turning-on AGNs is theorized to be more likely than in other galaxies, due to the possibility of 'Starfall' (McKernan et al., 2022). This TDE candidate in a Type 2 AGN (shown in

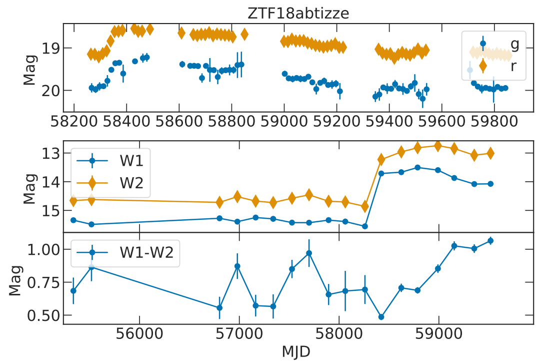


Figure 4.10: Optical ZTF forced-photometry light curves and evolution of the *WISE* MIR fluxes and *W1*–*W2* colour for a TDE candidate belonging to the NOT CL sample. Note the different time-scales of the ZTF and *WISE* data: the optical monitoring starts at the end of the MIR light curves.

Fig. 4.10) could potentially be happening in a turning-on AGN whose BELs are still too weak to be detected, and merits further study which is beyond the scope of this thesis. This discovery highlights the possibility of finding TDE candidates in AGNs, in order to compare their rate of occurrence to TDEs in quiescent galaxies.

Notably, the remaining ten sources show small-amplitude, stochastic optical variations, characteristic of Type 1 AGNs. This is also consistent with their optical spectra, which show weak broad H α emission lines, indicative of weak Type 1 AGNs. Interestingly, eight of these ten objects show a decrease in their optical flux along with a decrease in their MIR flux and colour, which suggests they are now transitioning to a dimmer state. Fig. 4.11 shows the optical and MIR light curves of two clear examples of this behaviour, ZTF19aapehvs and ZTF20aagwxlk. Generally, the MIR colours from these NOT CL weak Type 1s are galaxy dominated (i.e., *W1*–*W2* < 0.5), which suggests their weaker variability and BELs are not due to an orientation effect, but to an intrinsically lower AGN luminosity diluted by the emission of the host galaxy. The remaining two sources, ZTF19aavqrjg and ZTF20abgnlgv, also show Type-1 like variability and stronger broad H α emission (EW H α SDSS > 40 Å), indicative of Type 1 AGNs.

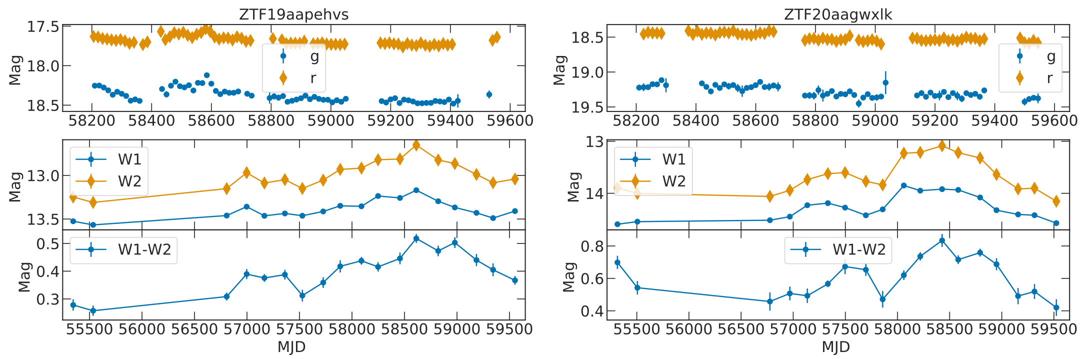


Figure 4.11: Optical ZTF forced photometry light curves and evolution of the *WISE* MIR fluxes and *W1*–*W2* colour for two NOT CLs sources. Note the different time-scales of the ZTF and *WISE* data: the optical monitoring starts at the end of the MIR light curves, where there is a dimming in the MIR emission. This evolution suggests the sources are now transitioning back to a dim state.

Table 4.5: Characterization of the NOT CL sources analysed in this chapter. *decr.* denotes a decreasing trend in the optical forced-photometry or MIR light curve. *var.* denotes a variable trend. Dashes denote the light curve is fairly flat. Slashes denote no available eROSITA-DE data. Asteriks denote the two NOT CLs reported in Chapter 3.

ZTF ID	Optical flux variability	MIR flux trend	MIR colour *W1*–*W2*	0.2–2.3 keV flux ($\cdot 10^{-14}$ erg s^{-1} cms^{-2})	Most likely cause
ZTF18aaiescp	yes, decr.	decr.	0.5–0.3	/	Weak Type 1
ZTF18aaiwdzt	transient	decr.	1–0.7	/	Transient event
ZTF18aaqjyon	transient	–	<0.3	/	Transient event
ZTF19aabyvtv	yes, decr.	decr.	decr. <0.4	/	Weak Type 1
ZTF18acgvmzb	no	decr.	decr. <0.5	< 9.9	Bad image subtraction
ZTF18achdyst	yes, decr.	decr.	decr. <0.5	8 ± 3	Weak Type 1
ZTF18acusqpt	no	var.	<0.3	8 ± 3	Bad image subtraction
ZTF19aaixgoj	yes, decr.	decr.	decr. <0.5	<11	Weak Type 1
ZTF19aapehvs	yes, decr.	decr.	decr. <0.5	/	Weak Type 1
ZTF19aavqrjg	yes	–	∼0.45	6 ± 3	Type 1 (BELs)
ZTF20aaeutuz	yes, decr.	decr.	decr. <0.5	13 ± 5	Weak Type 1
ZTF20aagwxlk	yes, decr.	decr.	decr. 0.8–0.4	/	Weak Type 1
ZTF20aakreaa	no	var.	0.2–0.6	/	Misclassified, small # of alerts
ZTF20aaorxzv	no	–	∼ 0.2	<9.5	Misclassified, small # of alerts
ZTF20abgnlgv	yes	var.	0.4–0.6	/	Type 1 (BELs)
ZTF18abtizze *	yes	var.	0.4–1	/	Possible TDE
ZTF19aaxdiui *	yes, decr.	decr.	decr. 0.6–0.2	/	Weak Type 1

4.5 Discussion

4.5.1 Improvement of the CL selection method through ALeRCE

One of the main aspects of the selection of CL candidates through ALeRCE relies on the correct classification of their alert light curve variability by the LCC. As mentioned in Section 4.4, there are seven out of 17 NOT CLs that have been misclassified by the LCC due to a bad subtraction of the ZTF images used to compute the alerts, a small number of data points and/or transient events in the alert light curves. On the other hand, we also found that ten ($\sim$60 per cent) NOT CLs have been correctly classified by the LCC as Type 1 AGN and are spectroscopically consistent with being weak Type 1 AGNs whose optical and MIR properties are dominated by the host galaxy contribution.

Thus, the most straightforward way to select the most promising turning-on CL candidates consists of adding constraints to the Type 1 AGNs provided by ALeRCE. These constraints can be drawn from the ALeRCE variability features that show the largest differences between the CL and NOT CL samples (see Section 4.3.1). The limiting values for cleaning the sample can be chosen according to different criteria. Here, we will take the 16th percentile of the variability features from the CL sample as a limiting lower value, as shown for the g band in Table 4.6. In the table we also show the percentage of NOT CL sources with the variability features having equal or lower values than the actual limiting value, which could be cleaned from the sample. Most of the features by themselves are able to discard around 50–60 per cent of NOT CLs (with the exception of the colour features), at the expense of losing 16–20 per cent of CLs.

It is possible to increase the completeness of the CL candidate list as well as the purity by applying two additional criteria on the ALeRCE features simultaneously. The best combinations in the g band, with a loss of 0–5 per cent of CLs and removal of 50 per cent of NOT CLs, were found to be all the possible pairs between GP_DRW_sigma, GP_DRW_tau and delta_mag_fid, and the combinations of iqr with MHPS_low or delta_mag_fid. We note that our samples are too small to judge the universality of this additional cleaning with bootstrap methods. However, similar cleaning can be devised in the future for ZTF forced photometry and/or the classifications of the ZTF data release light curves (Sánchez-Sáez et al., 2023), which would help to

Table 4.6: Variability and colour features in the g band that can be used to clean the CL candidate list. The limiting lower values correspond to the 16th percentile obtained for the CL sample. The third column indicates the percentage of NOT CL sources that can be discarded by limiting each feature to its limiting value.

Feature	**Limiting value**	**% cleaning**
positive_fraction	0	56
SF_ML_amplitude	-0.2	56
SPM_tau_rise	44	50
GP_DRW_sigma	$4{\cdot}10^{-5}$	56
GP_DRW_tau	9	56
IAR_phi	0.97	44
MHPS_low	0.012	63
delta_mag_fid	0.17	56
r–W3	6.7	6
g_r_mean_corr	0.60	18
g_r_max	-0.10	12
n_pos	0	56
iqr	19.7	63

improve the selection of CL candidates.

4.5.2 Characterizing the turning-on CL properties

Apart from the ALeRCE LCC features, we have analysed other multi-wavelength properties of the sample to investigate what can increase the likelihood of finding a turning-on transition.

BPT diagrams

While most of the reported CLs in the literature are Seyferts, some of them have been found to lie on the borderline between a LINER and Seyfert classification, with extreme order-of-magnitude changes in continuum and emission-line flux compared to less dramatic CLs occurring in Seyferts (Frederick et al., 2019).

Here, we tested whether the CL transitions could be related to the emission-line properties of the sources by computing the line-ratio diagnostic diagrams involving the line ratios [O III]/H β, [N II]/H α, [O I]/H α, and [S II]/H α (Baldwin et al., 1981; Kewley et al., 2006). By selection, all our CL candidates are consistent with a Seyfert classification, and the BPT criteria indicate that the line ratios from both their archival and new optical spectra distribute similarly, regardless of whether they experience a CL transition or not.

Forced photometry variability

Due to a higher number of data points, analyzing the complete light curves (instead of the alert curves) gives us a much more reliable determination of the variability of the sources. In general, we find a stronger variability typical of luminous Type 1 AGNs in the CL sample (during or post transition), which indicates that most of the candidates with non-variable BELs (i.e., the NOT CL sample) belong to a different population. Recently, ZTF light-curve variability analysis has been found to be a powerful tool to seek new CL candidates. In Chapter 2, by analysing ZTF forced photometry light curves of ~ 15000 Type 2 AGNs, we were able to distinguish between weak Type 1 and Type 2 sources and select CL candidates that have also been found through ALeRCE in the work presented here. Sánchez-Sáez et al. (2021a) used light curves from ZTF data release 5 (ZTF DR5) instead, and applied deep learning techniques to find anomalous behaviour in AGNs, leading to the identification of 75 promising CS candidates.

MIR variability

The study of the photometric variability in the MIR band can help us understand the physical mechanisms that may be involved in CL events. This emission is believed to come from dust thermally heated by the AGN, which emits relatively unimpeded by dust extinction. Indeed, simple *W1*–*W2* (i.e., [3.4]-[4.6] μm) colour cuts with *WISE* data are able to reliably differentiate AGNs from stars and galaxies (e.g. *W1*–*W2* $\geq$ 0.8, Stern et al., 2012). Along these lines, an important feature in our analysis is a_{12}, which indicates the linear trend of the *W1*–*W2* colour. Our CL sample shows a general upward trend during the last $\sim$11 years, indicating a redder colour in the MIR band. This behavior is consistent with a turning-on transition, where the sources change from galaxy-like (i.e., *W1*–*W2* $<$0.5) to AGN-like (i.e., *W1*–*W2* $>$ 0.8) MIR colour, as it has been found in previous CL studies (Sheng et al., 2017; Yang et al., 2018; Sheng et al., 2020; Lyu et al., 2022). This phenomenon can be explained by an increase of the brightness of the AGN, which illuminates the torus and produces a stronger contribution of hot dust in the *W1* and *W2* bands, with a larger effect on the latter. This scenario is supported by the observed *W1* and *W2* fluxes in our CL sample, which also show an upward trend given by the slopes a_1 and a_2.

To estimate the variability of the CLs in comparison with the NOT CL sample, we computed the parameters $\langle\sigma_{m1}\rangle$ and $\langle\sigma_{m2}\rangle$ (indicative of the intrinsic variability) and the maximum variations of *W1* and *W2* ($\langle\Delta W1\rangle$ and $\langle\Delta W2\rangle$) as in Lyu et al. (2022). Those authors present a study of the variability of the *W1* and *W2* light curves of a population of CLs in comparison with low-luminosity AGN and QSO samples. They also use the AllWISE multi-epoch photometry and NEOWISE single exposure (L1b) source table data, though we average the data every six months. Our results, presented in Table 4.3, are consistent with their conclusions. On the one hand, we find that the CL population has a higher $\langle\sigma_{m1}\rangle$ and $\langle\sigma_{m2}\rangle$ variability than the NOT CL population. On the other hand, the CL sample exhibits maximum variations of *W1* and *W2* ($\langle\Delta W1\rangle$ and $\langle\Delta W2\rangle$) greater than 0.3 mag. In the scenario of variable obscuration, Yang et al. (2018) estimate that a variation in the *W1* band due to dust extinction would yield a factor of $\sim$21 change in the g band magnitude, assuming the extinction curve in the optical and MIR and considering micrometer-sized grains (Wang et al., 2015). A variation of 0.3 mag in the *W1* band would lead to $\sim$6.3 mag difference in the g band, which is not consistent with the observed properties of CLs. Therefore, the variability and general increase in the MIR flux and colours of the CL sample, over the last 11 years, are most likely due to an intrinsic brightening of the AGN and cannot be solely explained by the motion of absorbing clouds within our line of sight.

X-rays

The CL sample shows higher X-ray fluxes and higher X-ray to g-band flux ratios than the NOT CL sample in the recent eROSITA scans. We note that the AGN continuum in the optical spectra of these sources is sub-dominant or undetectable under the starlight contribution, so a significant fraction of the g-band flux can be associated with the host galaxy. For a given host-galaxy flux, the CL sources are X-ray brighter, as expected if their AGNs are more dominant than in the NOT CL sample.

There is a relatively small number of CL sources with contemporaneous X-ray data in the literature, so it is still unclear whether these events are always accompanied by a clear change in the X-ray band. We note that we are talking about optical CLs, not X-ray CLs where the changes are due to obscuration. In the well studied case of the CL Mrk 1018, a strong soft X-ray excess was found when the source was in the bright state, which dropped by much more than the hard X-rays when the source transitioned to the dimmest state (McElroy et al., 2016; Husemann et al., 2016). In this case, the changes in the BELs were associated with an increase/decrease of ionizing photons from the soft X-ray excess (Noda & Done, 2018b). Additionally, Temple et al. (2023) found that five out of nine CL events with *Swift*-BAT data available could be associated with significant X-ray changes in the hard 14–195 keV band. The hard band is less affected by obscuration and thus rules out a varying column density as the origin of the flux changes. In this chapter, we find further evidence that at least some of the CL transitions could be associated with significant X-ray flux changes. In particular, four out of the 11 CLs with eROSITA data have experienced an X-ray flux rise with respect to archival *ROSAT* data taken 30 years earlier. These sources also show variable flux within the eROSITA 2019–2022 light curves, with the X-ray flux being four, eight and up to 15 times higher than the minimum value (see Fig. 4.8). However, due to the small number of data points and the intrinsic X-ray variability in unobscured AGNs, we are unable to discard the possibility that these changes are purely stochastic and not related to the CL event.

The CL AGN that has been most extensively observed in X-ray monitoring is 1ES 1927+654. This source was found to show dramatic changes in its X-ray spectral shape, showing the disappearance and re-appearance of the X-ray corona after the CL event (Trakhtenbrot et al., 2019; Ricci et al., 2020, 2021). The authors suggest that this particular phenomenon was caused by a TDE in an AGN. This strengthens the importance of multi-wavelength campaigns studying CLs, which enable constraints on the evolution of the AGN components during the transitions.

4.5.3 Promising CL candidates selected through multi-wavelength flux variability

The multi-wavelength photometric analysis performed in this work provides a powerful tool to identify turning-on CL events, allowing us to select the most promising CL events from the candidate list we have not re-observed. In Table 4.7 we show the sources from our CL candidates that are currently varying in the optical according to the refined ZTF feature criteria (all the possible pairs of best combinations mentioned in Section 4.5.1), and that are experiencing a MIR flux and colour increase. For the sources with eROSITA-DE data, we also checked whether the sources are brightening in the soft X-ray band than to their previous archival data. In Fig. B.2 we present the optical and MIR light curves for these sources. Some of them were not selected to be re-observed because they already showed obvious BELs in their archival optical spectra, such as ZTF18aaodaie and ZTF21abmjobi. However, according to their increase

Table 4.7: Potential CL sources from our list of CL candidates identified through their current optical variability and the criteria mentioned in the comments.

Name	Ra	Dec	z	MJD	fiberid	plate	Comments
ZTF18aalmtum	136.365522	17.928647	0.40	56246	330	5769	Increasing MIR fluxes and *W1*–*W2*.
ZTF18aaodaie	191.572165	28.342741	0.10	54205	442	2238	Increasing MIR fluxes and *W1*–*W2*.
ZTF18aaoudgg	221.975974	28.556697	0.16	53764	134	2141	Increasing MIR fluxes and *W1*–*W2*.
ZTF18acakoya	43.709715	-2.797898	0.12	56978	181	7823	Increasing MIR fluxes and *W1*–*W2*.
ZTF19adcfhxp	143.823381	34.320597	0.45	56336	727	5805	Increasing MIR fluxes and *W1*–*W2*.
ZTF20acpcfgt	339.995984	0.860685	0.38	52201	620	674	Increasing MIR and optical fluxes and *W1*–*W2*.
ZTF21abjprnu	230.292534	20.544009	0.14	54328	421	2159	Increasing MIR and optical fluxes. Decreasing *W1*–*W2*, but still *W1*–*W2* > 0.7
ZTF21abdvcpz	335.843667	0.771973	0.22	52140	454	375	Increasing MIR fluxes and *W1*–*W2*.
ZTF21abmjobi	358.363382	0.123463	0.17	52523	520	684	Increasing MIR and optical fluxes and *W1*–*W2*.

in the MIR fluxes and colour, it is likely that these sources now present stronger BELs and are therefore worth re-observing. We strongly encourage follow-up spectroscopy to confirm the CL behaviour in the promising candidates presented here, which will improve the statistics and refine further the selection criteria of CL candidates.

4.5.4 Key questions about the CL phenomena

Frequency and time-scales

CL events have been found to occur on time-scales as short as one or two months (Trakhtenbrot et al., 2019; Zeltyn et al., 2022). However, most research to date uses archival spectral and/or photometric data to look for CL transitions, making it difficult to obtain a proper estimate of the typical CL time-scales and thus understand the drivers of such extreme variations (Stern et al., 2018). The time difference between the first and last spectral epochs for our CL sample lies between 10 and 20 years, which must be taken as an upper limit on the time-scale for these transitions. Given their low incidence rate (discussed below) it is difficult to find a large amount of CL events as they take place. However, much larger and/or intensively sampled spectroscopic campaigns, such as SDSS-V (Kollmeier et al., 2017) and 4MOST (de Jong et al., 2019) will provide better constraints on the typical CL time-scales, potentially finding shorter (weeks and even days) CL events.

Another key question we can address with our results is the occurrence rate of turning-on transitions, following the discussion in Chapter 3. There are 30,333 AGNs classified as Type 2 (N and K types in MILLIQUAS, cleaned for weak Type 1s and LINERS) that can be detected by ZTF (Dec$< -28^{\circ}$ and r<21 mag). From them, 178 have alerts and are classified as AGN or QSO by the LCC. The 50 per cent confirmation rate found in this chapter implies that 0.3 per cent of transitions can be detected by this method over an average timespan of 15 years. Taking into account that just 10 per cent of Type 1 AGNs are variable enough to generate alerts in ZTF, we estimate a lower limit of 3 per cent of turning-on events every 15 years (i.e., 0.2 per cent per year). These values are consistent with the results from Hon et al. (2022), who finds a minimum turn-on CL AGN rate of 3 per cent every 15 years, and from Temple et al. (2023), who reports a CL rate of 0.7–6.2 per cent on 10–25 year time-scales (including both turning-on and turning-off events).

Physical origin

In principle, several mechanisms could lead to variations in the BELs, including variable obscuration and changes in the accretion rate. Here, we find that the MIR emission, which is dominated by the dust response (i.e., from the torus) to the UV–optical variations of the central engine, tends to increase in the CL sources. Since the MIR emitting region is too large to be obscured/unobscured on time-scales of $<$20 years, and the emission is much less affected by dust obscuration, we infer that the changes in the MIR waveband and by extension the CL transitions are due to changes in the accretion rate.

To explain the changes in the accretion flow at these time-scales, the proposed scenarios include instabilities in the accretion disc and major disc perturbations such as those caused by TDEs. In general, the light curves of our CL sources do not follow the power law decay $t^{-5/3}$ from the peak brightness as traditionally expected for TDEs (Rees, 1990). On the other hand, the values of λ_{Edd} obtained for the CL sample in their bright state are similar to the results for CL AGNs and QSOs reported in recent works ($-2 \lessapprox \log\lambda_{\mathrm{Edd}} \lessapprox -1$, MacLeod et al., 2019; Frederick et al., 2019; Graham et al., 2020; Temple et al., 2023), and is consistent with attributing the changes in accretion rate occurring preferentially in lower activity systems. Interestingly, our sources are currently (post-CL) accreting at λ_{Edd} 1 and 5 per cent L_{Edd}, which is in agreement with the expectations from a hydrogen ionization disc instability (Noda & Done, 2018b; Ruan et al., 2019; Śniegowska et al., 2023) In this scenario, by analogy to the spectral transitions in black hole X-ray binaries, the changes in the structure of the inner accretion disc occur around a critical value of $\lambda_{\mathrm{Edd}} \sim 0.02$, which is in agreement with the most recent CL studies (Graham et al., 2020; Guolo et al., 2021; Temple et al., 2023). Furthermore, our CL sources show a redder-when-brighter tendency in the MIR, with AGN-like MIR colours (i.e. *W1*–*W2* > 0.5) when they enter the bright state, which also has been found in other CL sources (Yang et al., 2018; Lyu et al., 2022) and supports the accretion state transition.

In this scenario, either the BELs appear due to the increase of ionizing photons that excite the BLR (LaMassa et al., 2015), or the BLR itself re-appears according to the expectations in disc-wind BLR models (Nicastro, 2000; Elitzur & Ho, 2009). These models predict an evolutionary sequence for the BLR depending on the accretion rate, leading to different intermediate type spectral transitions and the BLR disappearance at very low luminosities (Lbol $\lessapprox 5 \cdot 10^{39} M_7^{2/3}$ erg s^{-1}, where $M_7 = M/10^7 M_\odot$, Elitzur & Ho, 2009; Elitzur et al., 2014). According to the Eddington ratio estimates, none of our sources fell close to these limits, which suggests the BLR already existed in these sources but was too weakly ionized to produce detectable broad lines. On the other hand, the residual spectra in Fig. B.1 show that the continuum of the CLs currently looks either flat or blue. This 'bluer when brighter' effect, although clearly real in quasars, could at least partially be due to differences in the relative contribution of the host galaxy as these lower-luminosity AGNs change their intrinsic flux.

4.6 Summary and conclusions

This chapter is the continuation of work introduced in Chapter 3, where we present a method to search for turning-on CL candidates. The selection method consists of searching for current Type-1 AGN variability in a sample of spectrally classified Type 2 AGNs, using classifications given by a random-forest based light curve classifier, the ALeRCE LCC (Sánchez-Sáez et al., 2021b). In order to refine the selection method, we obtained second epoch spectra for 36 of our 61 CL candidates, six of which were reported in Chapter 3, which allows the confirmation of the CL objects by quantifying the change of the BELs in comparison with the archival SDSS spectra taken 15–20 years earlier. As a result, we find 19 ($\sim$ 50 per cent) turning-on CL confirmations (the CL sample), and 17 sources without significant changes in their BELs (the NOT CL sample).

We have analysed the multi-wavelength properties of the CL sample in comparison with the NOT CL sample to investigate what would increase the likelihood of finding a turning-on transition and understand its origin. Firstly, we performed a variability analysis of the alert light curves from ZTF, finding several variability features that are distinct between the samples and that can be applied to select the most promising CL sources. We also find that the turning-on transitions are characterized by an increase in the *WISE*-MIR brightness and the MIR (*W1*–*W2*) colour, where the stronger H α emission corresponds to an AGN-dominated MIR colour (*W1*–*W2* > 0.5 mag). The current Eddington ratio estimations for the CLs are lower than the overall Type 1 population, falling between one and five per cent L_{Edd}. In the X-ray band, we find that the CLs are brighter than the candidates that have not transitioned, and for six CLs we observe a flux increase between 2–15 times during the 2019–2022 X-ray monitoring. These results are in agreement with previous CL/CS works, and with the expectations from an accretion-state transition as the origin of these phenomena.

We also analyse the nature of the NOT CL sources according to their optical and MIR variability. For seven out of the 17 objects, the Type 2 sources were misclassified by the LCC due to a bad subtraction of the images, a small number of data points and/or transient events in the light curves such as SNe and one case of a likely TDE. Interestingly, we also find that ten sources are consistent with a Type 1 classification, where the optical and MIR emission is dominated by the host galaxy. This translates into lower amplitude variations in the optical and MIR wavebands, weaker BELs and galaxy-dominated MIR colours (i.e., *W1*–*W2* $<$0.5). Incidentally, we also find that seven of the NOT CL sources are currently decreasing their optical and MIR fluxes, suggesting they are currently transitioning to a dimmer state. The multi-wavelength differences between the CL and NOT CL sources allow us to select the most promising CL candidates from our list without spectroscopic follow-up, leading to nine sources that are worth re-observing.

The use of machine learning algorithms on complete optical light curves from the ZTF or the upcoming LSST can be combined with MIR data to unequivocally identify CLs, improve the statistics and ultimately understand the underlying physics of these phenomena.

Chapter 5

Conclusions and future prospects

AGN variability studies have been instrumental in advancing our understanding of the physical processes that govern the behavior of these enigmatic objects and have provided valuable insights into the evolution of galaxies. In this work, we have made use of state-of-the-art variability surveys in order to address the existence of AGNs that in principle challenges the simple UM, such as CS sources. Our main conclusions can be summarized as follows:

- Systematic optical variability analyses in spectrally-classified Type 2 objects are needed in order to obtain a pure Type 2 sample. In Chapter 2, by analysing the light curve variability of > 15.000 objects classified as Type 2 by the absence of BELs, we found evidence for a significant (11 per cent) fraction of optically-varying sources that could belong to the weak Type 1 or the CS population. Furthermore, the distinction between the weak Type 1 and Type 2 populations of AGNs can be differentiated by a small amount of variability features, such as the DRW parameters.
- As found in Chapter 3, the ALeRCE broker and other machine learning algorithms applied to large samples of light curves are a powerful tool to find rare phenomena like CS events and TDEs.
- Many CS events are characterized by a common multi-wavelength variability behaviour, which can be used to select new potential candidates. In Chapter 4

we found that most of the CS AGNs identified in this work have experienced an increase of MIR flux, MIR colour and –for those sources with available X-ray data– X-ray flux associated to a stronger broad H α line. Incidentally, we found that many of our CS candidates without changes in their BELs are consistent with a Type 1 classification, where the emission is dominated by the host-galaxy contribution, and therefore had been originally misclassified.

These analyses can be extended with data from next-generation photometric surveys such as the LSST, which will allow us to improve further the statistics on the CS typical timescales and frequency, but also to understand the physical processes driving such transitions and the formation of the BLR clouds. Furthermore, together with CS transitions we have found a sample of Type 1 AGNs whose emission is dominated by their host-galaxy. Since the properties of these sources could be associated to an intrinsic lower AGN luminosity, studying these systems will give us valuable insights on the factors contributing to the Type 1/Type 2 dichotomy, such as the relation between the torus covering factor and the mass accretion rate.

Appendix A

Most variable sources: Anomalous AGNs

Within the most variable sources of Chapter 2, we found three cases where two broad components where needed to fit the permitted emission lines in the SDSS spectrum: J124617.34+282033.92 (J1246+2820), J073149.29+361353.03 (J0731+3613) and J154246.71+334602.62 (J1542+3346). In these cases, the EW Hα is computed by adding the flux of the two components. For J0731+3613 and J1542+3346, the addition of a second semi-broad component with a similar line-of-sight velocity to the broad one was sufficient to fit the spectral shape of the Hα emission. On the contrary, for J1246+2820 the addition of a second broad component leads to some residuals around the Hα region (within the standard deviation), as a result of a more complex profile. In Fig. A.1 we present the fit to the SDSS spectrum of J1246+2820, where there seems to be a second peak of Hα redshifted with respect to the narrow emission lines. Interestingly, J1246+2820 (full) ZTF light curve is similar to that from the nearby Seyfert galaxy SDSS J143016.05+230344.4 (J1430+2303, see the bottom panel in Fig. A.1 for the most updated light curves retrieved from the ZTF Forced Photometry service), which has been recently reported as a candidate to host a supermassive black hole binary (SMBHB, Jiang et al., 2022; Dou et al., 2022). Its most recent optical spectrum, obtained on January 2022, also shows a complex velocity structure, which can be fitted with three Gaussians, including a significantly redshifted and a blueshifted component (Jiang

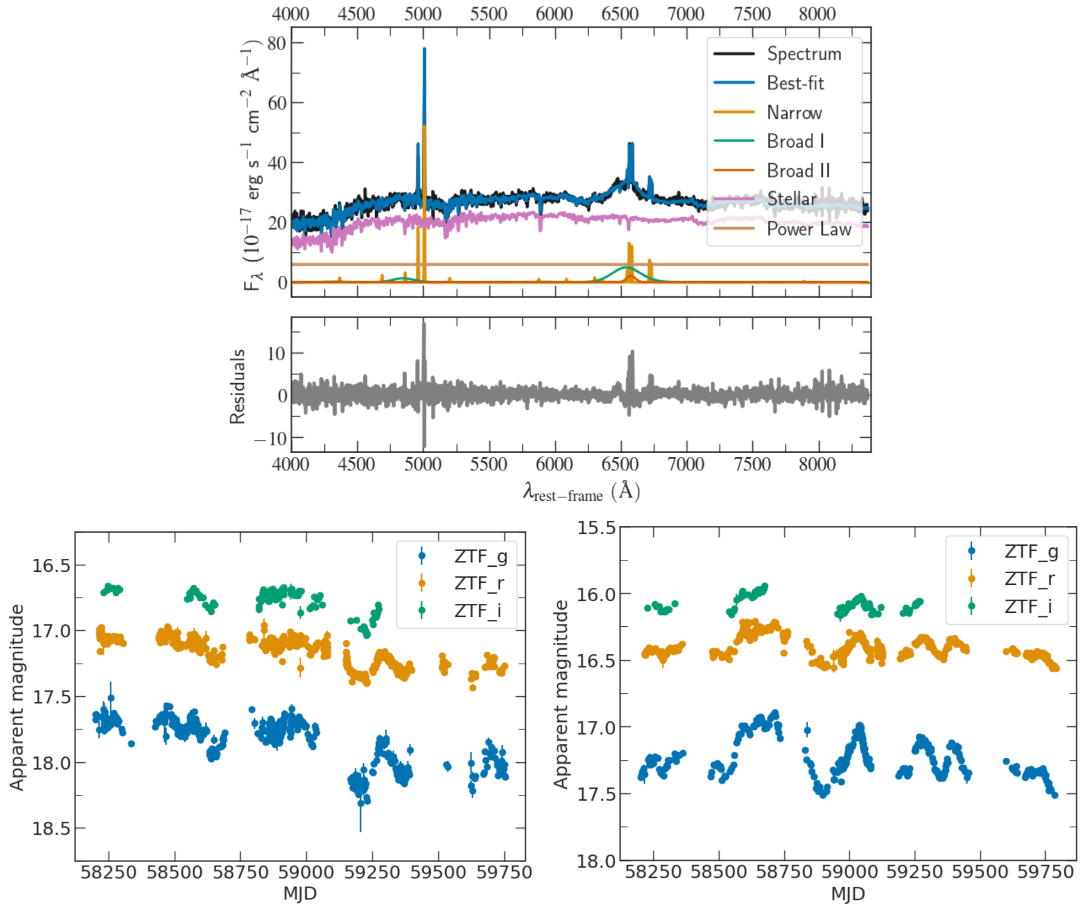


Figure A.1: Top: Fit to the SDSS spectrum of one of the most variable AGNs in our Type 2 SDSS sample, J1246+2820. Some residuals remain around the Hα line as a result of a more complex profile– notably a second peak of Hα redshifted with respect to the narrow lines. Bottom: ZTF light curves of J1246+2820 (left) and the SMBHB candidate host J1430+2303 (right).

et al., 2022). This suggests that similar mechanisms could be producing the observed optical properties, but future multi-wavelength observations are needed to better understand the nature of this source.

Table A.1: SDSS spectral information and fits for the most variable sources selected in Chapter 2. For the equivalent widths (EW) of Hα and Hβ we present the $^{90}_{10}$-th percentiles. Sources with EW Hα >5 Å are considered weak Type 1 AGNs, while sources EW Hα <5 Å are CL candidates. Names in bold refer to sources fitted with two broad components.

Name	Ra deg	Dec deg	redshift	MJD	fiberid	plate	EW Hα Å	EW Hβ Å
J000722.15+153811.56	1.842292	15.636544	0.12	52251	503	751	14^{15}_{13}	0^{1}_{0}
J002744.34+131300.22	6.93475	13.216728	0.09	56210	880	6190	21^{23}_{20}	0^{1}_{0}
J004247.82+231442.60	10.69925	23.245167	0.14	56248	300	6285	8^{13}_{5}	0^{0}_{0}
J005505.43+175256.18	13.772625	17.882272	0.1	56901	367	7623	21^{23}_{20}	1^{1}_{0}
J011142.76+262156.17	17.928167	26.365603	0.08	57367	852	7679	15^{16}_{14}	0^{1}_{0}
J011319.27+233449.13	18.330292	23.580314	0.11	55953	162	5699	5^{5}_{4}	0^{0}_{0}
J020858.95+274633.02	32.245625	27.775839	0.23	56325	320	6272	1^{22}_{7}	0^{0}_{0}
J073149.29+361353.03	112.955375	36.231397	0.14	55182	447	3662	53^{55}_{41}	0^{1}_{0}
J081917.51+301935.76	124.822958	30.3266	0.10	52619	94	931	3^{13}_{3}	0^{1}_{0}
J082250.42+154025.89	125.710083	15.673858	0.12	53713	517	2272	26^{28}_{23}	0^{1}_{0}
J082517.65+125855.07	126.323542	12.981964	0.17	54096	593	2422	29^{31}_{26}	0^{2}_{0}
J083224.29+355135.92	128.101208	35.859978	0.14	52668	459	1197	14^{18}_{12}	0^{1}_{0}
J083310.47+041036.73	128.293625	4.176869	0.10	52646	190	1186	20^{20}_{13}	0^{1}_{0}
J083934.03+104925.07	129.891792	10.823631	0.18	54061	437	2573	3^{5}_{2}	0^{1}_{0}
J091214.33+065722.28	138.059708	6.956189	0.13	52703	479	1194	11^{22}_{9}	0^{2}_{0}
J091357.17+250813.95	138.488208	25.137208	0.18	53415	217	2087	21^{23}_{19}	0^{2}_{0}
J093612.22+253226.10	144.050917	25.540583	0.13	54524	592	2294	0^{2}_{0}	0^{0}_{0}
J095137.27+341612.30	147.905292	34.270083	0.13	53388	351	1948	12^{26}_{10}	0^{2}_{0}
J095504.39+072606.93	148.768292	7.435258	0.21	52734	140	1235	15^{17}_{13}	0^{1}_{0}
J095526.28+305057.86	148.8595	30.849406	0.09	53436	342	1950	8^{11}_{8}	0^{1}_{0}
J102935.82+244639.38	157.39925	24.777606	0.11	53734	545	2349	14^{166}_{0}	0^{2}_{0}
J103450.93+462915.31	158.712208	46.487586	0.18	52620	294	962	8^{9}_{3}	$0^{1.6}_{0}$
J103812.23+244004.77	159.550958	24.667992	0.09	53770	150	2352	6^{7}_{5}	0^{1}_{0}
J110407.10+125005.21	166.029583	12.834781	0.13	53119	534	1603	3^{5}_{2}	0^{1}_{0}
J110529.70+051649.23	166.37375	5.280342	0.09	52356	400	581	12^{18}_{11}	0^{0}_{0}
J110920.32+531426.20	167.334667	53.240611	0.18	57135	363	8171	28^{31}_{27}	6^{7}_{5}
J111625.35+220049.37	169.105625	22.013714	0.14	54178	585	2492	39^{52}_{39}	0^{8}_{0}

J112229.25+102126.76	170.621875	10.357433	0.21	55976	584	5371	13_{13}^{21}	0_{0}^{2}
J112634.19+395539.72	171.642458	39.9277	0.19	53436	495	1996	38_{35}^{51}	1_{0}^{5}
J112939.08+365300.97	172.412833	36.883603	0.20	55673	426	4648	44_{40}^{84}	2_{0}^{10}
J114931.03+163743.16	177.379292	16.628656	0.29	56035	35	5892	31_{27}^{34}	4_{3}^{6}
J120045.49+145803.67	180.189542	14.967686	0.11	53463	158	1763	10_{7}^{21}	0_{0}^{1}
J120054.50+145850.97	180.227083	14.980825	0.08	53463	595	1763	4_{4}^{11}	0_{0}^{0}
J120349.21+605317.45	180.955042	60.888181	0.07	52405	420	954	24_{23}^{26}	0_{0}^{1}
J120402.12+335247.47	181.008833	33.879853	0.14	53469	38	2099	26_{24}^{49}	0_{0}^{6}
J120459.00+153513.85	181.245833	15.587181	0.22	53467	447	1764	25_{21}^{27}	0_{0}^{2}
J122011.98+153029.97	185.049917	15.508325	0.22	53436	382	1767	28_{24}^{34}	2_{0}^{5}
J122026.76+363327.88	185.1115	36.557744	0.10	53442	423	2003	20_{19}^{21}	0_{0}^{1}
J122053.47+283239.80	185.222792	28.544389	0.09	53816	569	2231	24_{22}^{25}	3_{2}^{4}
J122415.44+272506.48	186.064333	27.418467	0.09	56356	786	5976	22_{21}^{23}	4_{4}^{5}
J122439.67+180630.93	186.165292	18.108592	0.13	56034	566	5852	21_{19}^{22}	1_{0}^{2}
J122737.79+084406.84	186.907458	8.735233	0.09	53472	566	1626	12_{12}^{19}	0_{0}^{0}
J122801.79+223200.99	187.007458	22.533608	0.16	54495	355	2647	29_{27}^{31}	0_{0}^{1}
J124214.48+141146.99	190.560333	14.196386	0.16	53502	59	1769	15_{13}^{17}	0_{0}^{1}
J124445.48+282958.63	191.1895	28.499619	0.20	54205	369	2238	50_{47}^{57}	13_{12}^{14}
J124617.34+282033.92	191.57225	28.342756	0.10	54205	442	2238	57_{54}^{59}	10_{8}^{11}
J124629.36+465932.76	191.622333	46.992433	0.23	53089	118	1455	27_{26}^{29}	0_{0}^{2}
J131130.66+315200.81	197.87775	31.866892	0.07	53819	632	2029	8_{8}^{9}	0_{0}^{0}
J131508.78+365334.61	198.786583	36.892947	0.27	57519	836	8871	5_{3}^{6}	0_{0}^{1}
J131858.85+573006.65	199.745208	57.501847	0.10	52759	338	1320	13_{12}^{22}	0_{0}^{1}
J132558.71+151257.99	201.494625	15.216108	0.20	53759	449	1774	17_{16}^{29}	0_{0}^{1}
J132939.91+323144.28	202.416292	32.528967	0.24	58526	146	10257	36_{30}^{40}	9_{5}^{10}
J133052.10+202600.99	202.717083	20.433608	0.22	54230	212	2653	19_{18}^{22}	6_{4}^{8}
J134045.31+405333.45	205.188792	40.892625	0.09	57511	440	8391	24_{22}^{25}	3_{2}^{3}
J134148.78+370047.13	205.45325	37.013092	0.20	53858	513	2101	12_{10}^{33}	0_{0}^{4}
J134803.73+453728.47	207.015542	45.624575	0.16	53082	31	1465	13_{13}^{24}	0_{0}^{3}
J135007.70+124657.41	207.532083	12.782614	0.14	53857	205	1777	22_{21}^{28}	2_{2}^{5}
J135425.48+334254.98	208.606167	33.715272	0.25	55274	746	3861	8_{5}^{38}	0_{0}^{3}
J141821.98+612731.79	214.591583	61.458831	0.25	52365	129	606	11_{10}^{13}	0_{0}^{1}
J142052.22+472625.72	215.217583	47.440478	0.13	53462	303	1673	4_{3}^{5}	0_{0}^{1}
J142352.09+245417.14	215.967042	24.904761	0.07	53493	417	2132	20_{19}^{33}	0_{0}^{2}
J142736.37+265700.50	216.901542	26.950139	0.17	53876	208	2134	12_{9}^{20}	1_{0}^{3}

J143519.06+511739.78	218.829417	51.294383	0.08	52781	245	1327	5_{4}^{7}	0_{0}^{1}
J144021.49+141125.74	220.089542	14.190483	0.12	54234	147	2748	13_{11}^{15}	0_{0}^{1}
J144933.48+082355.69	222.3895	8.398803	0.12	54555	418	1814	8_{7}^{10}	0_{0}^{1}
J145005.11+154348.27	222.521292	15.730075	0.27	54535	308	2764	24_{21}^{30}	2_{0}^{4}
J150111.92+040422.87	225.299667	4.073019	0.16	52055	478	589	8_{5}^{29}	0_{0}^{1}
J153006.52+071020.17	232.527167	7.172269	0.13	54208	322	1820	25_{23}^{42}	0_{0}^{3}
J153832.66+460735.01	234.636083	46.126392	0.2	52781	559	1332	13_{11}^{16}	0_{0}^{1}
J154246.71+334602.62	235.694625	33.767394	0.28	55747	816	4971	95_{93}^{102}	17_{15}^{19}
J154907.53+372900.59	237.281375	37.483497	0.20	53172	270	1681	14_{16}^{26}	0_{0}^{1}
J155258.30+273728.41	238.242917	27.624558	0.09	53498	603	1654	25_{24}^{27}	0_{0}^{1}
J155259.94+210246.89	238.24975	21.046358	0.17	53557	220	2171	28_{25}^{63}	0_{0}^{6}
J155640.32+451338.41	239.168	45.227336	0.18	52753	471	1169	14_{10}^{52}	0_{0}^{5}
J161123.42+424139.79	242.847583	42.694386	0.25	57896	953	8528	0_{0}^{29}	0_{0}^{4}
J161219.56+462942.62	243.0815	46.495172	0.13	52443	414	814	4_{4}^{22}	0_{0}^{2}
J163344.96+112611.59	248.437333	11.436553	0.19	54585	212	2533	20_{18}^{21}	3_{2}^{4}
J163639.58+194201.73	249.164917	19.700481	0.15	53224	51	1659	8_{7}^{29}	0_{0}^{1}
J164360.00+321009.80	251.0	32.169389	0.14	52786	159	1341	22_{20}^{24}	0_{0}^{1}
J164432.94+213306.45	251.13725	21.551792	0.3	53149	80	1570	25_{23}^{30}	1_{0}^{2}
J165119.25+242011.41	252.830208	24.336503	0.15	52912	455	1424	2_{3}^{23}	0_{0}^{3}
J220138.03+121456.52	330.408458	12.249033	0.19	52224	520	734	5_{0}^{36}	0_{0}^{3}
J221044.76+245958.05	332.6865	24.999458	0.12	56213	258	5958	42_{33}^{46}	0_{0}^{1}
J222559.67+201944.75	336.498625	20.329097	0.16	55854	712	5024	20_{19}^{21}	1_{0}^{2}
J223249.33+035829.19	338.205542	3.974775	0.14	55525	402	4291	27_{26}^{29}	1_{0}^{1}
J231009.82+074928.51	347.540917	7.824586	0.16	56187	642	6168	52_{51}^{53}	0_{0}^{2}
J231138.89+274504.30	347.912042	27.751194	0.12	56559	358	6289	6_{5}^{7}	0_{0}^{0}
J231720.14+143855.97	349.333917	14.648881	0.15	52258	315	745	9_{9}^{13}	0_{0}^{1}
J233628.86+231956.29	354.12025	23.332303	0.18	56566	856	6519	8_{7}^{15}	0_{0}^{0}

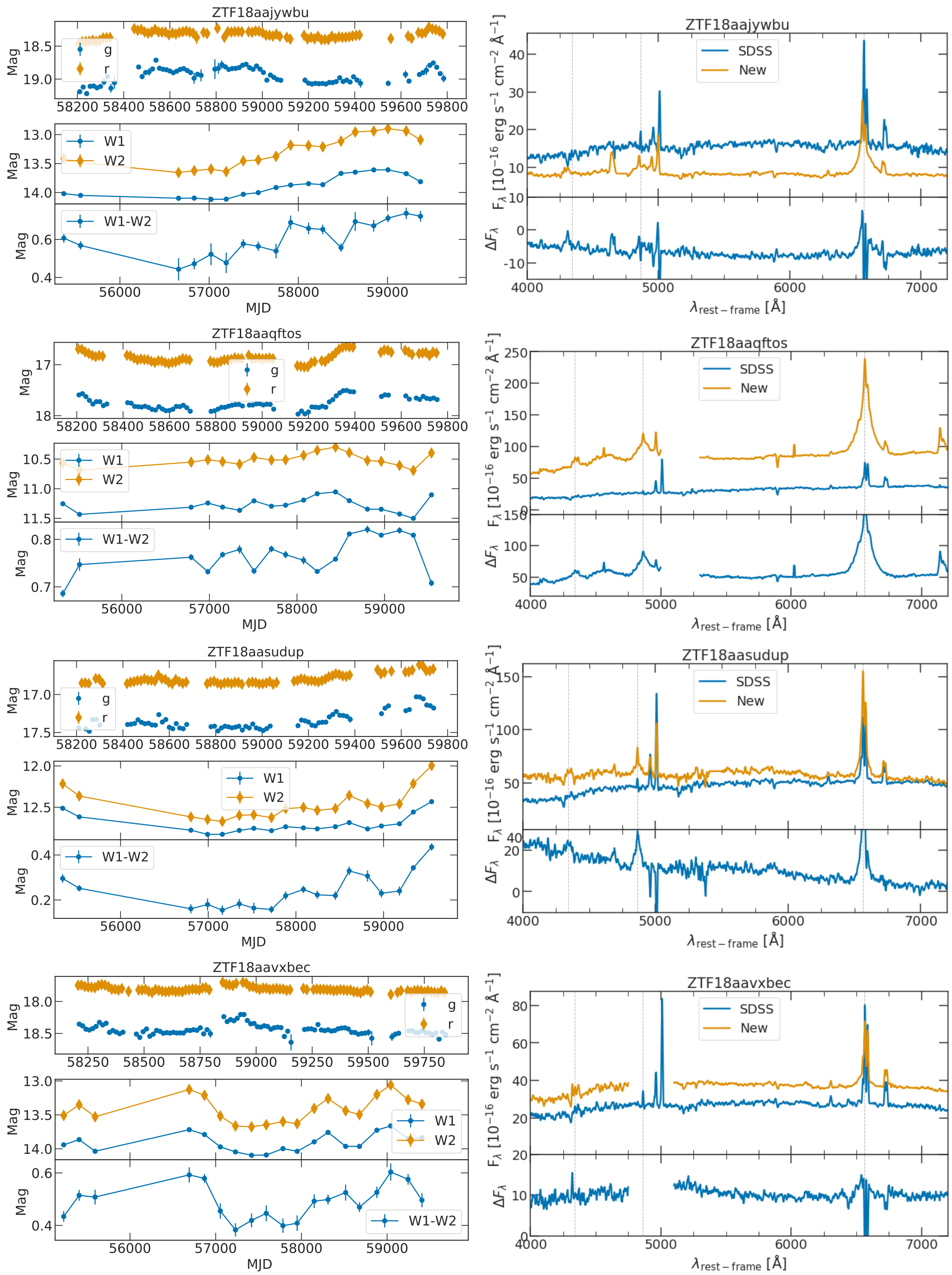
Appendix B

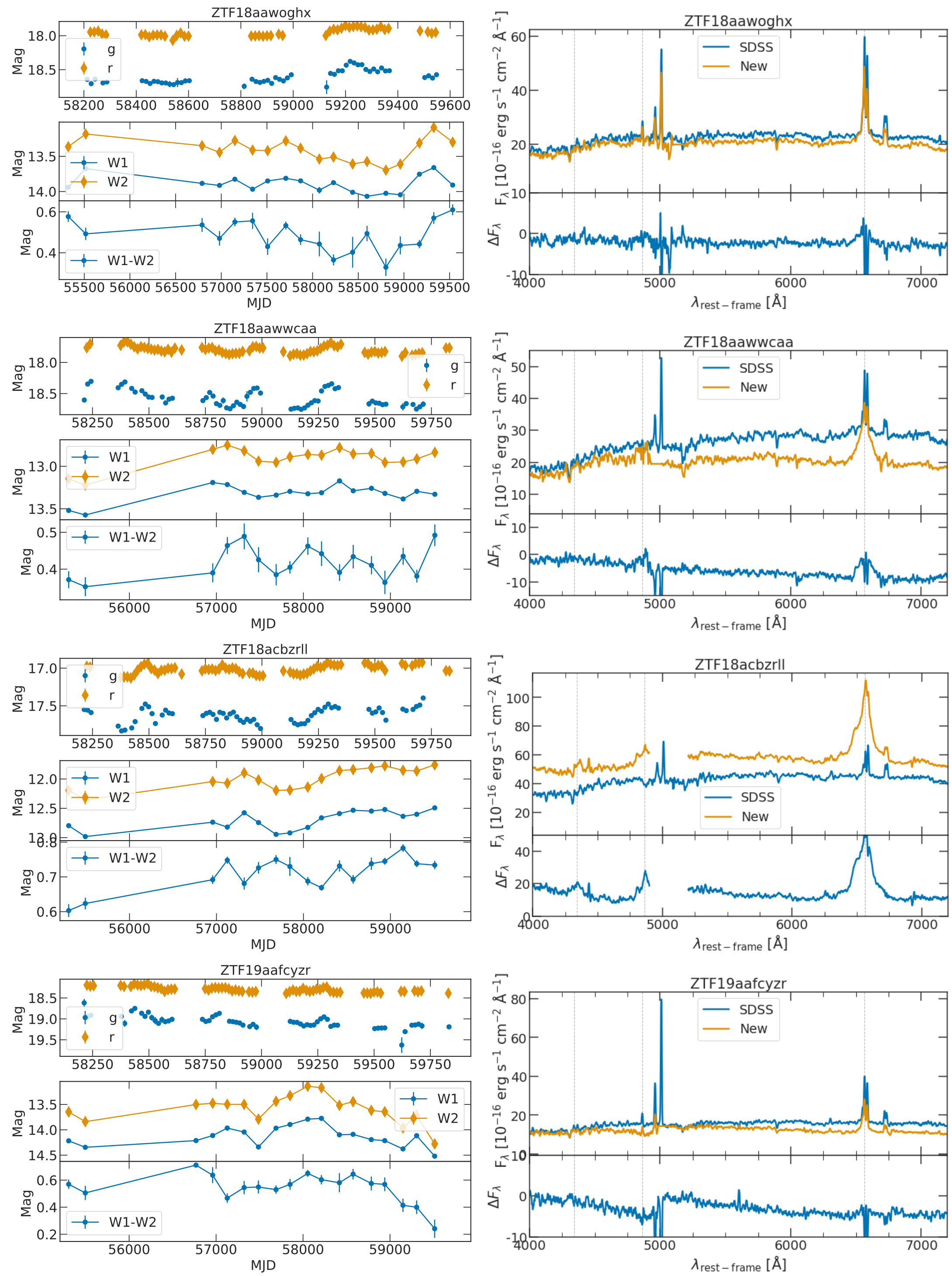
Light curves and spectra from the CL sample

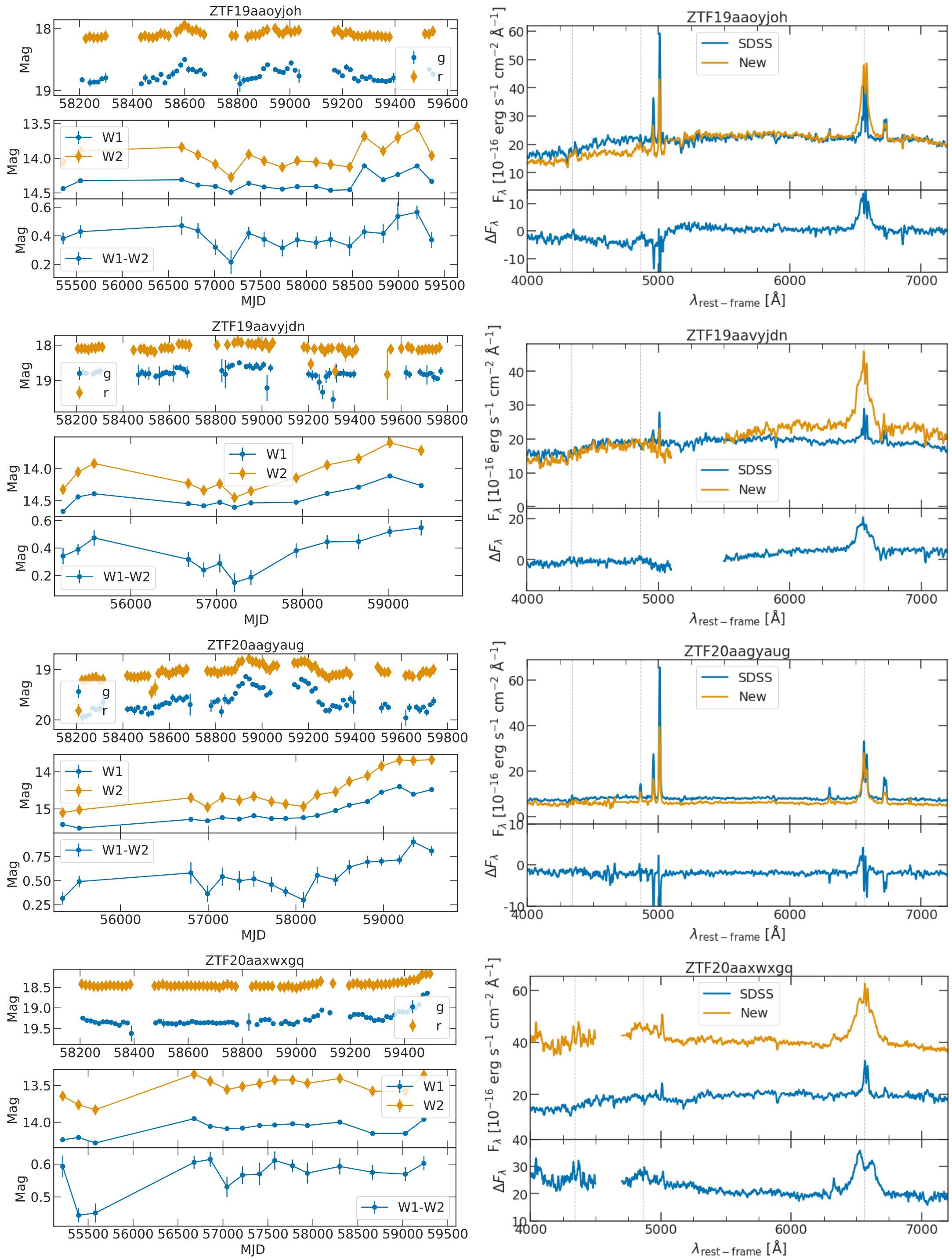
Table B.1 indicates the dates and instruments used for the second epoch spectra of all the CL candidates observed in Chapter 4. Fig. B.1 shows the ZTF optical and *WISE*-MIR light curves (left) and optical spectra (right) for the 15 CL sources identified in Chapter 4. The optical spectra are scaled to the flux of [S II] in the earliest spectra and smoothed with a 10 Å box filter. The lower plots in the right side show the difference between the new and the old spectra. In some cases, the new spectra were taken with a blue and a red arms, leading to very noisy intermediate regions that have been deleted. The other four CL sources considered in this work are reported in Chapter 2. Fig. B.2 shows the ZTF optical and *WISE*-MIR light curves for the most promising CL candidates selected according to their optical and MIR photometric variability.

Table B.1: Dates and instruments of the second epoch spectra for the CL candidates observed in this work. Sources identified as CL are shown in bold.

ZTF ID	RA deg	DEC deg	MJD	Instrument
ZTF18aaiescp	207.21292	57.646792	59603	DBSP(P200)
ZTF18aaiwdzt	199.48361	49.258651	59603	DBSP(P200)
ZTF18aajywbu	205.45327	37.013091	59604	LRIS(Keck I)
ZTF18aaqftos	180.95505	60.888181	59696	DBSP(P200)
ZTF18aaqjyon	180.42264	38.47264	59696	DBSP(P200)
ZTF18aasudup	170.03619	34.312731	59603	DBSP(P200)
ZTF18aavxbec	243.08151	46.495172	59696	DBSP(P200)
ZTF18aawoghx	156.10498	37.650863	59603	DBSP(P200)
ZTF18aawwcaa	128.10120	35.859979	59603	DBSP(P200)
ZTF19aabyvtv	197.87773	31.866893	59603	DBSP(P200)
ZTF18acbzrll	124.82294	30.32660	59696	DBSP(P200)
ZTF18acgvmzb	148.96896	35.965616	59616	DBSP(P200)
ZTF18achdyst	157.39925	24.777606	59616	DBSP(P200)
ZTF18acusqpt	177.91898	12.036714	59603	DBSP(P200)
ZTF19aafcyzr	125.71008	15.673859	59616	DBSP(P200)
ZTF19aaixgoj	146.80538	12.205624	59603	DBSP(P200)
ZTF19aaoyjoh	180.18956	14.967685	59603	DBSP(P200)
ZTF19aapehvs	199.74519	57.501847	59696	DBSP(P200)
ZTF19aavqrjg	181.24583	15.58718	59603	DBSP(P200)
ZTF19aavyjdn	202.39941	-1.509453	59603	DBSP(P200)
ZTF20aaeutuz	164.81581	12.483378	59696	DBSP(P200)
ZTF20aagwxlk	153.42829	55.432205	59696	DBSP(P200)
ZTF20aagyaug	172.41282	36.883602	59603	DBSP(P200)
ZTF20aakreaa	181.52333	42.169888	59603	DBSP(P200)
ZTF20aaorxzv	132.39052	3.68048	59604	LRIS(Keck I)
ZTF20aaxwxgq	234.63610	46.126392	59696	DBSP(P200)
ZTF20abcvgpb	238.24977	21.046358	59696	DBSP(P200)
ZTF20abgnlgv	232.52715	7.172269	59696	DBSP(P200)
ZTF21aafkiyq	174.93478	-1.727439	59603	DBSP(P200)
ZTF21abcsvbr	184.64829	18.771713	59603	DBSP(P200)


ZTF18aajywbu
Mag
g
r
W1
W2
W1-W2
MJD
SDSS
New
F_λ [10^{-16} erg s^{-1} cm^{-2} Å^{-1}]
ΔF_λ
$\lambda_{\rm rest-frame}$ [Å]
ZTF18aaqftos
ZTF18aasudup
ZTF18aavxbec


ZTF18aawoghx
g
r
W1
W2
W1-W2
Mag
MJD
SDSS
New
F_λ [10^{-16} erg s^{-1} cm^{-2} Å^{-1}]
ΔF_λ
$\lambda_{\mathrm{rest-frame}}$ [Å]
ZTF18aawwcaa
ZTF18acbzrll
ZTF19aafcyzr


ZTF19aaoyjoh
ZTF19aavyjdn
ZTF20aagyaug
ZTF20aaxwxgq
g
r
W1
W2
W1-W2
Mag
MJD
SDSS
New
F_λ [10^{-16} erg s^{-1} cm^{-2} Å^{-1}]
ΔF_λ
$\lambda_{\rm rest-frame}$ [Å]

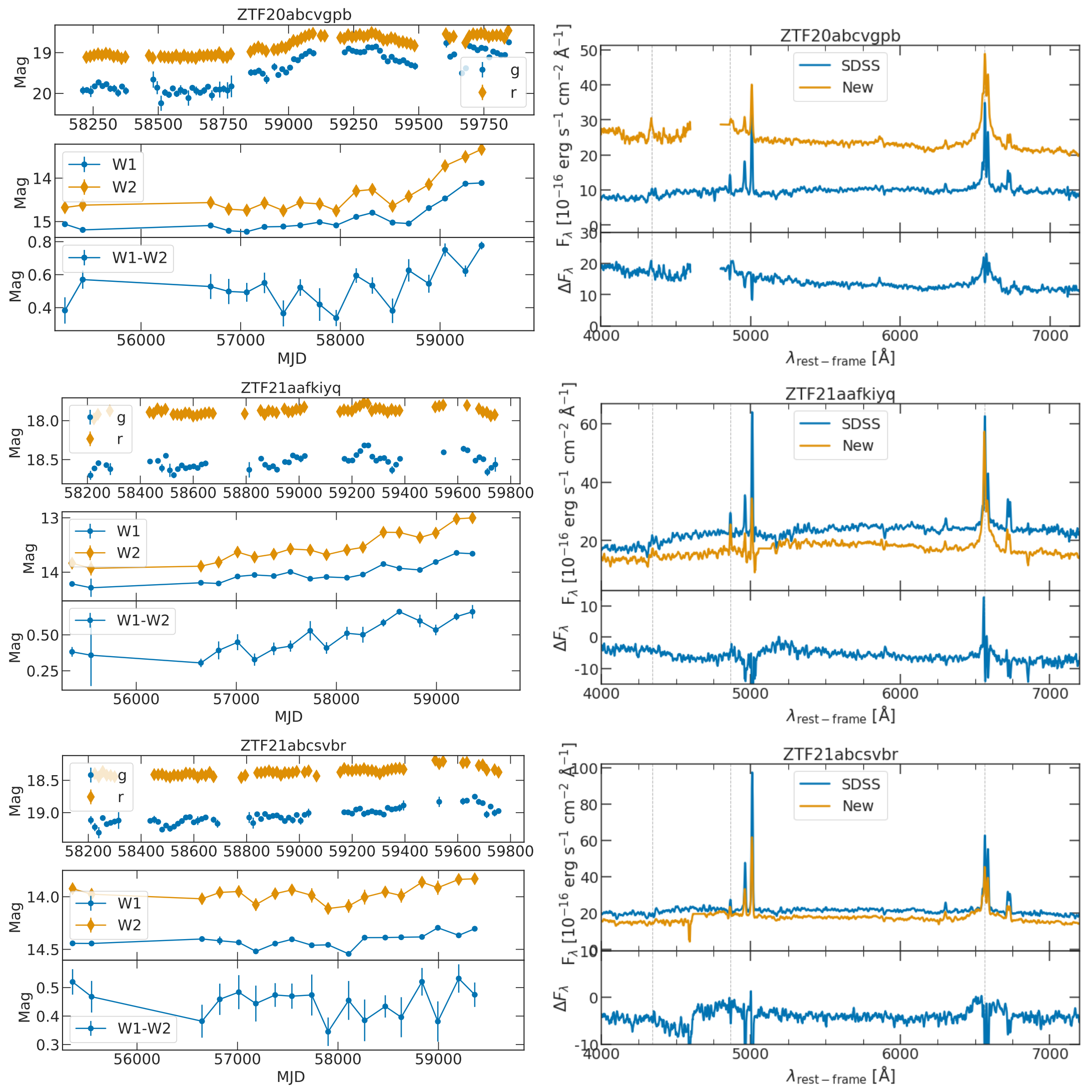
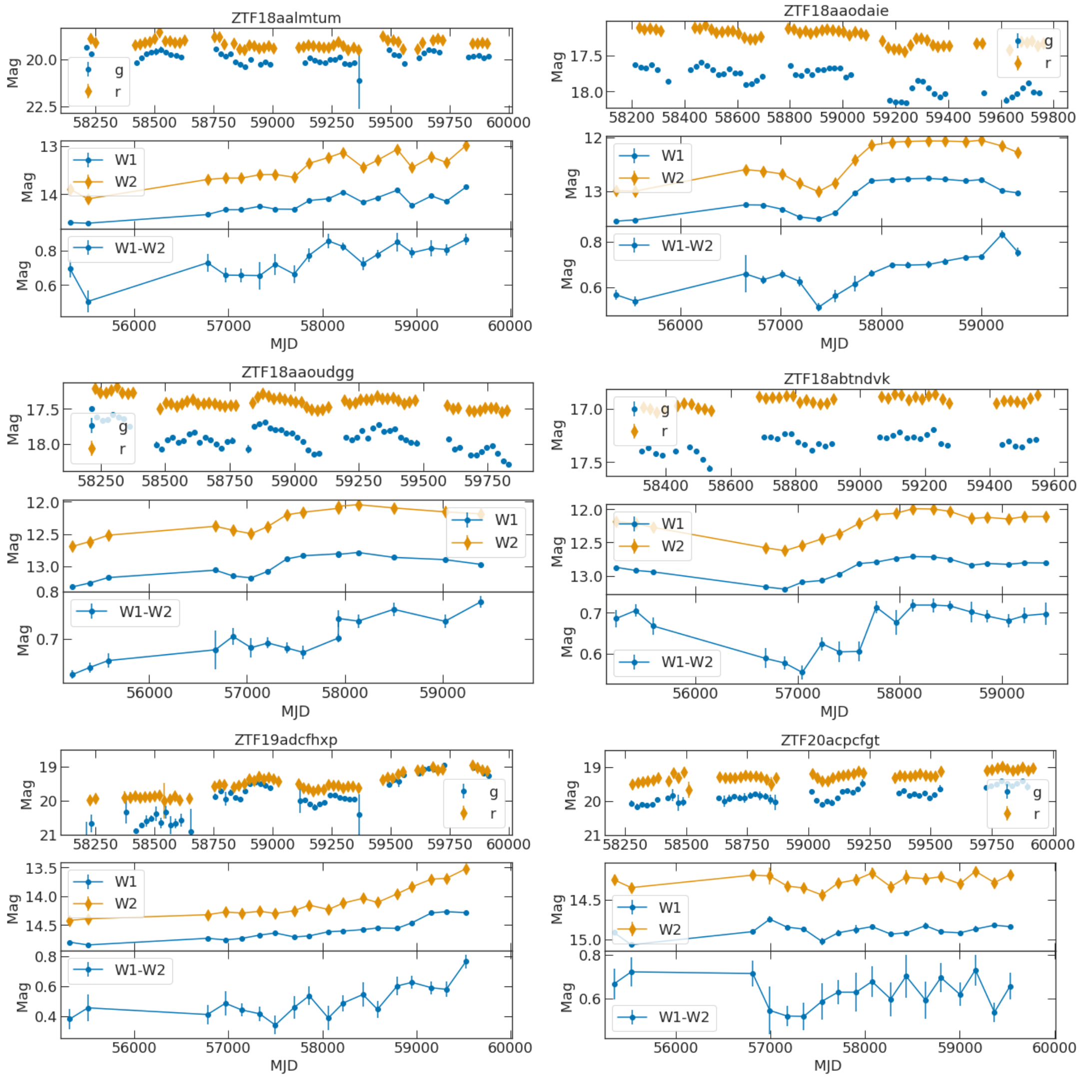


Figure B.1: ZTF optical and *WISE*-MIR light curves (left) and optical spectra (right) for the CL sources.


ZTF18aalmtum
ZTF18aaodaie
ZTF18aaoudgg
ZTF18abtndvk
ZTF19adcfhxp
ZTF20acpcfgt
Mag
MJD
g
r
W1
W2
W1-W2

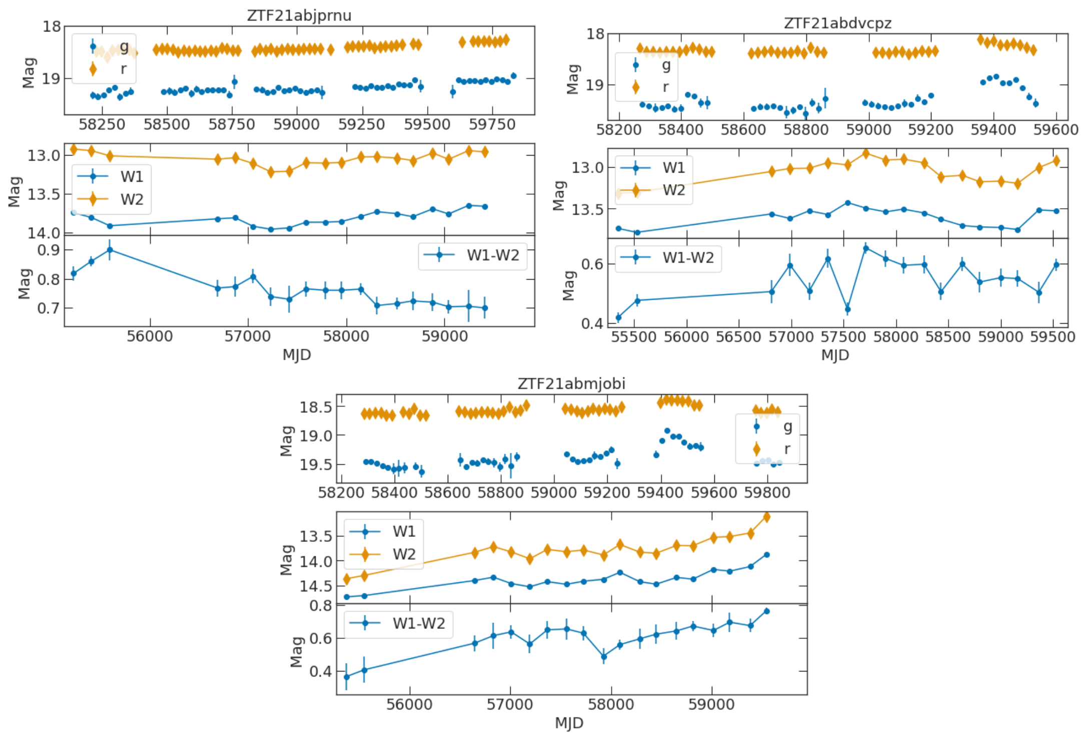


Figure B.2: ZTF optical and *WISE*-MIR light curves from the most promising CL candidates.

Bibliography

Ahumada R. et al., 2020, ApJS, 249, 3

Antonucci R., 1993, ARA&A, 31, 473

Arévalo P., Churazov E., Zhuravleva I., Hernández-Monteagudo C., Revnivtsev M., 2012, MNRAS, 426, 1793

Arévalo P., Uttley P., Kaspi S., Breedt E., Lira P., McHardy I. M., 2008, MNRAS, 389, 1479

Arévalo P., Uttley P., Lira P., Breedt E., McHardy I. M., Churazov E., 2009, MNRAS, 397, 2004

Ashby M. L. N. et al., 2009, ApJ, 701, 428

Assef R. J., Stern D., Noirot G., Jun H. D., Cutri R. M., Eisenhardt P. R. M., 2018, ApJS, 234, 23

Baldwin J. A., 1997, in Astronomical Society of the Pacific Conference Series, Vol. 113, Peterson B. M., Cheng F.-Z., Wilson A. S., eds, IAU Colloq. 159: Emission Lines in Active Galaxies: New Methods and Techniques, p. 80

Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5

Barth A. J., Voevodkin A., Carson D. J., Woźniak P., 2014, AJ, 147, 12

Barvainis R., 1987, ApJ, 320, 537

Bellm E., 2014, in Wozniak P. R., Graham M. J., Mahabal A. A., Seaman R., eds, The Third Hot-wiring the Transient Universe Workshop, p. 27

Bellm E. C. et al., 2019a, PASP, 131, 068003

Bellm E. C. et al., 2019b, PASP, 131, 018002

Bianchi S., Corral A., Panessa F., Barcons X., Matt G., Bassani L., Carrera F. J., Jiménez-Bailón E., 2008, MNRAS, 385, 195

Blandford R. D., Begelman M. C., 1999, MNRAS, 303, L1

Blandford R. D., Payne D. G., 1982, MNRAS, 199, 883

Boller T., Freyberg M. J., Trümper J., Haberl F., Voges W., Nandra K., 2016, A&A, 588, A103

Bolton A. S. et al., 2012, AJ, 144, 144

Breedt E. et al., 2009, MNRAS, 394, 427

Brightman M., Nandra K., 2008, MNRAS, 390, 1241

Brunner H. et al., 2022, A&A, 661, A1

Burtscher L. et al., 2013, A&A, 558, A149

Capetti A., Axon D. J., Macchetto F., Sparks W. B., Boksenberg A., 1996, ApJ, 469, 554

Cappellari M., 2017, MNRAS, 466, 798

Chambers K. C., et al. , 2017, VizieR Online Data Catalog, II/349

Chiang J., Murray N., 1996, ApJ, 466, 704

Cohen R. D., Rudy R. J., Puetter R. C., Ake T. B., Foltz C. B., 1986, ApJ, 311, 135

Collin S., Zahn J.-P., 1999, A&A, 344, 433

Collin S., Zahn J. P., 2008, A&A, 477, 419

Czerny B., 2019, Open Astronomy, 28, 200

Czerny B., Hryniewicz K., 2011, A&A, 525, L8

Czerny M., King A. R., 1989, MNRAS, 241, 839

de Jong R. S. et al., 2019, The Messenger, 175, 3

Dou L. et al., 2022, arXiv e-prints, arXiv:2208.11968

Edelson R., Turner T. J., Pounds K., Vaughan S., Markowitz A., Marshall H., Dobbie P., Warwick R., 2002, ApJ, 568, 610

Elitzur M., 2012, ApJ, 747, L33

Elitzur M., Ho L. C., 2009, ApJ, 701, L91

Elitzur M., Ho L. C., Trump J. R., 2014, MNRAS, 438, 3340

Elitzur M., Shlosman I., 2006, ApJ, 648, L101

Fernández-Ontiveros J. A., Muñoz-Darias T., 2021, MNRAS, 504, 5726

Flesch E. W., 2021, arXiv e-prints, arXiv:2105.12985

Förster F. et al., 2021, AJ, 161, 242

Frederick S. et al., 2019, ApJ, 883, 31

Freudling W., Romaniello M., Bramich D. M., Ballester P., Forchi V., García-Dabló C. E., Moehler S., Neeser M. J., 2013, A&A, 559, A96

Gaskell C. M., 2008, in Revista Mexicana de Astronomia y Astrofisica Conference Series, Vol. 32, Revista Mexicana de Astronomia y Astrofisica Conference Series, p. 1

Graham M. J., Djorgovski S. G., Drake A. J., Stern D., Mahabal A. A., Glikman E., Larson S., Christensen E., 2017, MNRAS, 470, 4112

Graham M. J. et al., 2020, MNRAS, 491, 4925

Green P. J. et al., 2022, ApJ, 933, 180

Guainazzi M., 2002, MNRAS, 329, L13

Guo H., Sun M., Liu X., Wang T., Kong M., Wang S., Sheng Z., He Z., 2019, ApJ, 883, L44

Guolo M., Ruschel-Dutra D., Grupe D., Peterson B. M., Storchi-Bergmann T., Schimoia J., Nemmen R., Robinson A., 2021, MNRAS, 508, 144

Harrison C. M. C., 2014, Ph.D. thesis, Durham University, UK

Hatziminaoglou E. et al., 2010, A&A, 518, L33

Hawkins M. R. S., 2004, A&A, 424, 519

Heckman T. M., Best P. N., 2014, ARA&A, 52, 589

Hernández-García L., Masegosa J., González-Martín O., Márquez I., Guainazzi M., Panessa F., 2017, A&A, 602, A65

Hickox R. C., Alexander D. M., 2018, ARA&A, 56, 625

Ho L. C., 2008, ARA&A, 46, 475

Homan J., Belloni T., 2005, Ap&SS, 300, 107

Hon W. J., Wolf C., Onken C. A., Webster R., Auchettl K., 2022, MNRAS, 511, 54

Husemann B. et al., 2016, A&A, 593, L9

Hutsemékers D., Agís González B., Marin F., Sluse D., Ramos Almeida C., Acosta Pulido J. A., 2019, A&A, 625, A54

Hutsemékers D., Agís González B., Sluse D., Ramos Almeida C., Acosta Pulido J. A., 2017, A&A, 604, L3

Ivezić Ž. et al., 2019, ApJ, 873, 111

Jiang N. et al., 2022, arXiv e-prints, arXiv:2201.11633

Kelly B. C., Bechtold J., Siemiginowska A., 2009, ApJ, 698, 895

Kerr R. P., 1963, Phys. Rev. Lett., 11, 237

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Kollmeier J. A. et al., 2017, arXiv e-prints, arXiv:1711.03234

Koshida S. et al., 2014, ApJ, 788, 159

Koss M. et al., 2017, ApJ, 850, 74

Kozłowski S., 2017, A&A, 597, A128

Kozłowski S. et al., 2010, ApJ, 716, 530

Kraft R. P., Burrows D. N., Nousek J. A., 1991, ApJ, 374, 344

Laha S., Markowitz A. G., Krumpe M., Nikutta R., Rothschild R., Saha T., 2020, ApJ, 897, 66

Laha S. et al., 2022, ApJ, 931, 5

LaMassa S. M. et al., 2015, ApJ, 800, 144

Laor A., Netzer H., 1989, MNRAS, 238, 897

Lawrence A., 1991, MNRAS, 252, 586

Lin D. et al., 2017, Nature Astronomy, 1, 0033

Lira P., Arévalo P., Uttley P., McHardy I. M. M., Videla L., 2015, MNRAS, 454, 368

Liu H.-Y., Liu W.-J., Dong X.-B., Zhou H., Wang T., Lu H., Yuan W., 2019, ApJS, 243, 21

López-Navas E., Arévalo P., Bernal S., Graham M. J., Hernández-García L., Lira P., Sánchez-Sáez P., 2023, MNRAS, 518, 1531

López-Navas E. et al., 2022, MNRAS, 513, L57

Lusso E., Risaliti G., 2016, ApJ, 819, 154

Lyu B., Wu Q., Yan Z., Yu W., Liu H., 2022, ApJ, 927, 227

MacLeod C. L. et al., 2019, ApJ, 874, 8

MacLeod C. L. et al., 2010, ApJ, 721, 1014

MacLeod C. L. et al., 2016, MNRAS, 457, 389

Mainzer A. et al., 2011, ApJ, 731, 53

Markowitz A., Edelson R., Vaughan S., 2003, ApJ, 598, 935

Markowitz A. G., Krumpe M., Nikutta R., 2014, MNRAS, 439, 1403

Masci F. J. et al., 2019, PASP, 131, 018003

Massaro E., Maselli A., Leto C., Marchegiani P., Perri M., Giommi P., Piranomonte S., 2015, Ap&SS, 357, 75

Matt G., Guainazzi M., Maiolino R., 2003, MNRAS, 342, 422

McElroy R. E. et al., 2016, A&A, 593, L8

McKernan B., Ford K. E. S., Cantiello M., Graham M., Jermyn A. S., Leigh N. W. C., Ryu T., Stern D., 2022, MNRAS, 514, 4102

Meisner A. M., Caselden D., Schlafly E. F., Kiwy F., 2023, AJ, 165, 36

Murray N., Chiang J., 1997, ApJ, 474, 91

Murray N., Chiang J., Grossman S. A., Voit G. M., 1995, ApJ, 451, 498

Nandra K., George I. M., Mushotzky R. F., Turner T. J., Yaqoob T., 1997, ApJ, 476, 70

Netzer H., 2015, ARA&A, 53, 365

Netzer H., 2019, MNRAS, 488, 5185

Netzer H., Laor A., 1993, ApJ, 404, L51

Nicastro F., 2000, ApJ, 530, L65

Noda H., Done C., 2018a, MNRAS, 480, 3898

Noda H., Done C., 2018b, MNRAS, 480, 3898

Oh K., Yi S. K., Schawinski K., Koss M., Trakhtenbrot B., Soto K., 2015, ApJS, 219, 1

Oknyansky V. L. et al., 2021, MNRAS, 505, 1029

Padovani P., 2016, A&A Rev., 24, 13

Panessa F., Bassani L., 2002, A&A, 394, 435

Panessa F. et al., 2009, MNRAS, 398, 1951

Pappa A., Georgantopoulos I., Stewart G. C., Zezas A. L., 2001, MNRAS, 326, 995

Parker M. L. et al., 2019, MNRAS, 483, L88

Predehl P. et al., 2021, A&A, 647, A1

Predehl P., Schmitt J. H. M. M., 1995, A&A, 293, 889

Proga D., 2007, in Astronomical Society of the Pacific Conference Series, Vol. 373, Ho L. C., Wang J. W., eds, The Central Engine of Active Galactic Nuclei, p. 267

Ramos Almeida C. et al., 2011, ApJ, 731, 92

Ramos Almeida C. et al., 2009, ApJ, 702, 1127

Ramos Almeida C., Ricci C., 2017, Nature Astronomy, 1, 679

Rees M. J., 1990, Science, 247, 817

Reines A. E., Greene J. E., Geha M., 2013, ApJ, 775, 116

Ricci C. et al., 2020, ApJ, 898, L1

Ricci C. et al., 2021, ApJS, 255, 7

Ricci C., Trakhtenbrot B., 2022, arXiv e-prints, arXiv:2211.05132

Ricci C. et al., 2017, Nature, 549, 488

Risaliti G., Elvis M., Fabbiano G., Baldi A., Zezas A., 2005, ApJ, 623, L93

Risaliti G., Elvis M., Nicastro F., 2002, ApJ, 571, 234

Rivers E. et al., 2015, ApJ, 815, 55

Ross N. P. et al., 2018, MNRAS, 480, 4468

Ruan J. J., Anderson S. F., Eracleous M., Green P. J., Haggard D., MacLeod C. L., Runnoe J. C., Sobolewska M. A., 2019, ApJ, 883, 76

Runnoe J. C. et al., 2016, MNRAS, 455, 1691

Salvato M. et al., 2022, A&A, 661, A3

Sánchez P. et al., 2017, ApJ, 849, 110

Sánchez-Sáez P. et al., 2023, arXiv e-prints, arXiv:2304.08519

Sánchez-Sáez P. et al., 2021a, AJ, 162, 206

Sánchez-Sáez P. et al., 2021b, AJ, 161, 141

Saxton R. D., Read A. M., Esquej P., Freyberg M. J., Altieri B., Bermejo D., 2008, A&A, 480, 611

Sergeev S. G., Doroshenko V. T., Golubinskiy Y. V., Merkulova N. I., Sergeeva E. A., 2005, ApJ, 622, 129

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Shapiro S. L., Teukolsky S. A., 1983, Black holes, white dwarfs, and neutron stars : the physics of compact objects

Sheng Z. et al., 2020, ApJ, 889, 46

Sheng Z., Wang T., Jiang N., Yang C., Yan L., Dou L., Peng B., 2017, ApJ, 846, L7

Shi Y., Rieke G. H., Smith P., Rigby J., Hines D., Donley J., Schmidt G., Diamond-Stanic A. M., 2010, ApJ, 714, 115

Sniegowska M., Czerny B., Bon E., Bon N., 2020, A&A, 641, A167

Śniegowska M., Grzędzielski M., Czerny B., Janiuk A., 2023, A&A, 672, A19

Steffen A. T., Strateva I., Brandt W. N., Alexander D. M., Koekemoer A. M., Lehmer B. D., Schneider D. P., Vignali C., 2006, AJ, 131, 2826

Stern D. et al., 2012, ApJ, 753, 30

Stern D. et al., 2005, ApJ, 631, 163

Stern D. et al., 2018, ApJ, 864, 27

Sun W.-H., Malkan M. A., 1989, ApJ, 346, 68

Temple M. J. et al., 2023, MNRAS, 518, 2938

Trakhtenbrot B. et al., 2019, ApJ, 883, 94

Tran H. D., 2001, ApJ, 554, L19

Tran H. D., Lyke J. E., Mader J. A., 2011, ApJ, 726, L21

Ulrich M.-H., Maraschi L., Urry C. M., 1997, ARA&A, 35, 445

Urry C. M., Padovani P., 1995, PASP, 107, 803

Uttley P., Edelson R., McHardy I. M., Peterson B. M., Markowitz A., 2003, ApJ, 584, L53

Vaughan S., Edelson R., Warwick R. S., Uttley P., 2003, MNRAS, 345, 1271

Vaughan S., Fabian A. C., 2004, MNRAS, 348, 1415

Vazdekis A., Koleva M., Ricciardelli E., Röck B., Falcón-Barroso J., 2016, MNRAS, 463, 3409

Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, MNRAS, 404, 1639

Véron-Cetty M. P., Véron P., 2010, A&A, 518, A10

Wanders I. et al., 1997, ApJS, 113, 69

Wang F. et al., 2021, ApJ, 907, L1

Wang J., Xu D. W., Wang Y., Zhang J. B., Zheng J., Wei J. Y., 2019, ApJ, 887, 15

Wang J.-M., Du P., Brotherton M. S., Hu C., Songsheng Y.-Y., Li Y.-R., Shi Y., Zhang Z.-X., 2017, Nature Astronomy, 1, 775

Wang J.-M. et al., 2011, ApJ, 739, 3

Wang S., Li A., Jiang B. W., 2015, MNRAS, 454, 569

Witt H. J., Czerny B., Zycki P. T., 1997, MNRAS, 286, 848

Wolter A., Gioia I. M., Henry J. P., Mullis C. R., 2005, A&A, 444, 165

Wright E. L. et al., 2010, AJ, 140, 1868

Xia X. Y., Xue S. J., Mao S., Boller T., Deng Z. G., Wu H., 2002, ApJ, 564, 196

Yang Q., Shen Y., Liu X., Wu X.-B., Jiang L., Shangguan J., Graham M. J., Yao S., 2019, ApJ, 885, 110

Yang Q. et al., 2018, ApJ, 862, 109

Yip C. W. et al., 2009, AJ, 137, 5120

Yuan S., Strauss M. A., Zakamska N. L., 2016, MNRAS, 462, 1603

Zeltyn G. et al., 2022, ApJ, 939, L16