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

E. López-Navas, M.L. Martínez-Aldama, S. Bernal, P. Sánchez-Sáez, P. Arévalo, Matthew J. Graham, L. Hernández-García, P. Lira, and P.A. Rojas Lobos.

1 Instituto de Física y Astronomía, Facultad de Ciencias, Universidad de Valparaíso, Gran Bretaña 1111, Valparaíso, Chile
2 Departamento Astronomía, Universidad de Chile, Casilla 36D, Santiago, Chile
3 European Southern Observatory, Karl-Schwarzschild-Strasse 2, 85748 Garching bei München, Germany
4 Millennium Institute of Astrophysics (MAS), Nuncio Monseñor Sítero Sanz 100, Providencia, Santiago, Chile
5 California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA

ABSTRACT

Determining the frequency and duration of changing–look (CL) active galactic nuclei (AGNs) phenomena, where the optical broad emission lines appear or disappear, is crucial to understand the evolution of the accretion flow around supermassive black holes. We present a strategy to select new CL candidates starting from a spectroscopic type 2 AGNs sample and searching for current type 1 photometric variability. We use the publically available Zwicky Transient Facility (ZTF) alert stream and the Automatic Learning for the Rapid Classification of Events (ALeRCE) light curve classifier to produce a list of CL candidates with a highly automated algorithm, resulting in 60 candidates. Visual inspection reduced the sample to 30. We performed new spectroscopic observations of six candidates of our clean sample, without further refinement, finding the appearance of clear broad Balmer lines in four of them and tentative evidence of type changes in the remaining two, which suggests a promising success rate of ≥66 per cent for this CL selection method.

Key words: galaxies: active – accretion, accretion discs – (galaxies:) quasars: emission lines

1 INTRODUCTION

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 2018; 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 color (Hon et al. 2021), 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 11 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. 2021b), 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. 2019) it is possible now, for the first time, to obtain a census of the rate of these changes (Sánchez-Sáez et al. 2021b). 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 work, we assume a standard cosmological model with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3$, and $\Omega_{\Lambda} = 0.7$.

2 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 work we used the current variability-based, publicly available classifications provided by the ALeRCE light curve classifier tool (LCC, Sánchez-Sáez et al. 2021a). The LCC uses a hierarchical imbalanced Random Forest classifier, fed with variability features computed from ZTF light curves and colors 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 colors, features related to the amplitude of the variability at different timescales, and features related to the timescale 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 (see Table 1). Of these, ~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 a forthcoming paper.

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 SPECTROSCOPIC FOLLOW-UP

3.1 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 ESOReflx 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 2.

### 3.2 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 Figure 1 we present the plots of the average flux and difference spectra of the six CL candidates observed. It can be seen that ZTF19abixawb, ZTF20abshkf, ZTF19aaxuyo and ZTF18accdxv show a clear change both in the continuum and in the H\(\alpha\) and H\(\beta\) 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 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 2 correspond to the 10 and 90 percentiles of the simulations results.

For the sources with broad H\(\alpha\) emission, we estimated the black hole masses (M\(\text{BH}\)) 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\(\alpha\) obtained from the new spectra. With these values, we computed the current Eddington ratios \(\lambda_{\text{Edd}} = L_{\text{bol}}/L_{\text{Edd}}\), where \(L_{\text{Edd}} = 1.5 \times 10^{38} (M_{\text{BH}}/M_\odot)\) erg s\(^{-1}\) is the Eddington luminosity and \(L_{\text{bol}}\) is the bolometric luminosity defined as \(L_{\text{bol}} = 40 \cdot (L_{5100}/10^{42})^{-0.2}\) according to Netzer (2019). The results of the spectral fitting and the derived physical quantities are shown in Table 2.

### 4 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 ~ 30 promising CL candidates. Here we confirm a significant spectral change (a > 3\(\sigma\) change in the EW of H\(\alpha\) and H\(\beta\)) 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. (2021), which is the most successful search method for CL AGNs to date in comparison to previous works with variability selection criteria (15-35 per cent, e.g. Yang et al. 2018; MacLeod et al. 2019).

With the results of this work, 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 2, we need to take into consideration misclassifications 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 work) 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. (2021), who found a minimum turn-on CL AGN rate of 3 per cent over ~15 yr.

Our results indicate that the sources in this work are currently accreting at a few per cent L\(_{\text{Edd}}\), 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 2018; Ross et al. 2018; Huisemakers 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 radioquiet 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 ZTF20abshkf, 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, ZTF19aaxuyo and ZTF18accdxv, 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 H\(\alpha\)/OIII ratios of the four CL AGNs are now between 5 and 30 times larger than in the archival spectra (see Table 2), 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
Table 2. 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. Δt: time difference between the SDSS and the SOAR/GTHS or UT1/FORS2 spectra. Derived quantities only, the associated systematic error is 0.5 dex on the mass and 0.3 dex on the bolometric luminosity used to calculate \( L_{\text{edd}} \).

| ZTF ID   | SDSS name            | z     | epoch | SNR_{5100} | \( \text{EW}_H \)\(^{\alpha} \) Å | \( \text{EW}_H \)\(^{\beta} \) Å | Hα/OIII\(^{\gamma} \) | Δt | log M_{\text{BH}} | \( L_{\text{edd}} \) |
|----------|----------------------|-------|-------|------------|----------------|----------------|----------------|-----|----------------|----------------|
| ZTF19abixawb\(^{a}\) | SDSS J001014.86+000820.7 | 0.1022 | new   | 33         | 47.0±10      | 8.3±0.3        | 5.8±0.6        | 10.1±0.5       | 21.17 | 8.0±0.2      | 0.013±0.007 |
|         | old                  | 14    |       |            |                |                |                |                 |                |              |              |
| ZTF20abshfkf\(^{b}\) | SDSS J011311.82+013542.4 | 0.2375 | new   | 24         | 133±2        | 29±1           | 7±20           | 5.8±0.4        | 6.13    | 8.0±0.1      | 0.03±0.01   |
|         | old                  | 20    |       |            |                |                |                |                 |                |              |              |
| ZTF18aceddxv\(^{b}\) | SDSS J075544.35+192336.3 | 0.1083 | new   | 91         | 90.3±1       | 21.3±0.4       | 15±10          | 5.1±0.4        | 17.02   | 8.5±0.1      | 0.005±0.001 |
|         | old                  | 20    |       |            |                |                |                |                 |                |              |              |
| ZTF19aalxyuo\(^{a}\) | SDSS J081240.76+071528.5 | 0.0849 | new   | 36         | 54±1         | 4±1            | 9.8±0.3        | 17.64          | 7.76±0.04 | 0.027±0.005 |
|         | old                  | 27    |       |            |                |                |                |                 |                |              |              |
| ZTF19aaxdiiri\(^{a}\) | SDSS J214046.03+091631.7 | 0.4030 | new   | 17         | –             | 4±1            | –              | 11.08          | –       | –            | –            |
|         | old                  | 17    |       |            |                |                |                |                 |                |              |              |
| ZTF18abtízze\(^{a}\) | SDSS J215055.73-010654.1 | 0.0879 | new   | 24         | 13±2         | 3±20           | 0.42±0.05      | 17.39          | 7.4±0.2   | 0.005±0.003 |
|         | old                  | 11    |       |            |                |                | 0.1±0.1        |                 |                |              |              |

transitions from true type 2 to type 1 and again to type 2. In fact, preliminary fits to the light curve of ZTF18abtízze 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, ZTF19aaxdiiri, H\(_\alpha\) falls out of the observed wavelength range, but the \( \text{EW}_H \) is the same as for the CL ZTF19aalxyuo (which displays very significant changes in H\(_\alpha\)). Therefore, we cannot rule out this source is actually a CL AGN.

5 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 \( L_{\text{edd}} \), 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 are currently re-observing the full promising CL sample to refine the statistics on the frequency of type 2 to type 1 transitions, which will be presented in a forthcoming paper. In future work, we will improve the completeness of the candidate list by characterising light curves that use all available epochs of ZTF instead of only using the ZTF alert stream. This will include more of the lower-variability AGNs, which comprise 90 per cent of known type 1 sources.

ACKNOWLEDGEMENTS

We acknowledge financial support from ANID Becas 21200718 and 21212344 (ELN, SB), Millenium Nucleus NCI19_058 TITANS (PA, MLMA), ICN12_009 MAS (LHG), FONDECYT Nº 3200250 and 1201748 (PSS, PL). ELN acknowledges Caltech for its hospitality. Based on observations collected at the European Southern Observatory under ESO programme 108.22BA.001, at the Southern Astrophysical Research telescope under CNTAC programme CN2021B-16, and with the Samuel Oschin Telescope 48-inch and the 60-inch Telescope at the Palomar Observatory as part of the ZTF project.

DATA AVAILABILITY

The SDSS data underlying this article were accessed from SDSS DR16 (http://skyserver.sdss.org/dr16). The UT1/FORS2 and SOAR/GTHS data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Antonucci R., 1993, ARA&A, 31, 473
Bellm E. C., et al., 2019, PASP, 131, 018002
Cappellari M., 2017, MNRAS, 466, 798
Fernández-Ontiveros J. A., Muñoz-Darias T., 2021, MNRAS, 504, 5726
Flesch E. W., 2021, arXiv e-prints, p. arXiv:2105.12985
Fürster 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., Moshier S., Neese M. J., 2013, A&A, 559, A96
Graham M. J., et al., 2020, MNRAS, 491, 4925
Green P. I., et al., 2022, arXiv e-prints, p. arXiv:2201.09123
Guolo M., Ruschel-Dutra D., Grupe D., Peterson B. M., Storchi-Bergmann T., Schimoia J., Nemmen R., Robinson A., 2021, MNRAS, 508, 144
Homan J., Belloni T., 2005, Ap&SS, 300, 107
Hon W. J., Wolf C., Onken C. A., Webster R., Auchettl K., 2021, MNRAS, 511, 54
Hutsemékers D., Agís González B., Sluse D., Ramos Almeida C., Acosta Pulido J. A., 2017, A&A, 604, L3
Hutsemékers D., Agís González B., Marin F., Sluse D., Ramos Almeida C., Acosta Pulido J. A., 2019, A&A, 625, A54
LaMassa S. M., et al., 2015, ApJ, 800, 144
MacLeod C. L., et al., 2016, MNRAS, 457, 389
MacLeod C. L., et al., 2019, ApJ, 874, 8
Netzer H., 2015, ARA&A, 53, 365
Netzer H., 2019, MNRAS, 488, 5185
Noda H., Done C., 2018, MNRAS, 480, 3898
Figure 1. Archival (blue line) and new optical spectra (black line) for the subsample of CL candidates identified in this work, scaled to the flux of [OIII] $\lambda$=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.