COVER LETTER
CNTAC Fast track 2024A

PI:
Patricia Arévalo, Universidad de Valparaíso, patricia.arevalo@uv.cl, faculty.

CO-Is:
Aldo Aliaga, Universidad de Valparaíso, aldo.aliaga@alumnos.uv.cl, undergraduate student
Paulina Lira, Universidad de Chile, plira@das.uchile.cl, faculty
Paula Sánchez-Sáez, ESO- Garching, Paula.SanchezSaez@eso.org, fellow

Title:
Towards finding intermediate mass black holes with large variability surveys

Abstract:
The origin of supermassive black holes (SMBHs) in the Universe is a key unanswered question. The analysis of the occupation fractions of intermediate mass black holes (IMBHs) and low-mass AGNs, can constrain the seeding mechanism involved in the formation and growth of SMBHs, but first we must be able to detect them robustly. There are only a few hundred known low-mass AGNs to date but many more are expected to exist in the local universe. One promising way to find them is to exploit their optical variability, analyzing large variability surveys (e.g. ZTF). However, known low-mass AGN are routinely missed in the ZTF alert stream and they are difficult to detect in the ZTF full light curves, owing to their rapid variability which is not well matched to the ZTF cadence. With this proposal we aim at obtaining higher cadence, higher S/N lightcurves of IMBH candidates in order to confirm them, establish their variability properties and further use them to refine the search algorithm.

Telescope: Speculoos

Time requested: 45 nights, minimum times: 3 nights
No preferred dates or moon phases.

SCIENTIFIC AIM AND RATIONALE

The origin of supermassive black holes (SMBHs) in the Universe is a key unanswered question. Common models invoke formation through "light seeds" ($M \approx 10^2\ M_\odot$), as remnants of the first Pop III stars, or "heavy seeds" ($M \approx 10^4 - 10^6 M_\odot$), from the direct collapse of giant pristine gas clouds (e.g., Volonteri 2010, Woods 2019, and references therein). While a fraction of these black hole (BH) seeds must grow rapidly to become $10^9\ M_\odot$ AGN at $z \sim 6$, many more may experience little or no growth, and be observed locally as intermediate mass black holes (IMBHs) with masses in the $10^2 - 10^5 M_\odot$, or low-mass AGNs ($M_{\rm BH} < 10^{6.3}\ M_\odot$). The analysis of the occupation fractions of IMBHs and low-mass AGNs, as well as the BH mass function can help us to constrain the seeding mechanism involved in the formation and growth of SMBHs (e.g., Greene 2019) but first we must be able to detect them robustly.

There are only a few hundred known low-mass AGNs to date. Around ~ 500 low-mass AGNs have been selected from Sloan Digital Sky Survey (SDSS; York 2000) spectra, by identifying the presence of broad emission lines in the optical spectrum (Greene & Ho 2004, 2007, Dong 2012b, Chilingarian 2018, Liu 2018), but these are only ∼0.6% of all the AGN spectroscopically confirmed by SDSS, so they are likely very underrepresented in current catalogs. As an alternative to finding these low mass AGN spectroscopically, another key trait of AGN can be exploited, i.e., their persistent flux variability. As an example, Martínez–Palomera (2020) selected low-mass AGN candidates by detecting intra-day variability in the nuclear region of an unbiased sample of local galaxies. On the other hand, Baldassare (2018, 2020) detected yearly variations in the central region of low-mass galaxies. Now that optical photometric monitoring of large fractions of the sky is becoming publicly available (e.g. ZTF, Bellm et al. 2014 or the upcoming Vera C. Rubin Observatory Legacy Survey of Space and Time, LSST; Ivezic 2019), it becomes possible to detect these low mass AGN in blind searches and for large numbers.

The ALeRCE (Automatic Learning for the Rapid Classification of Events; Förster 2021) broker is currently processing the ZTF alert stream, providing classifications of different variable and transient objects, in preparation for the LSST era. In particular the light curve classifier (Sánchez-Sáez 2021) uses mainly flux variability features to classify astronomical objects into 15 variable classes including three types of AGN. **However, known low-mass AGN are routinely missed in the ZTF alert stream** because of their low luminosity and variability amplitude, which almost always leaves their variations below the significance threshold to be detected as ZTF alerts and therefore are simply not processed. Of the 500 low-mass AGN from SDSS only 12 generated ZTF alerts by 2020/06/09 and were therefore visible to ALeRCE. Of these, 11 were classified by ALeRCE as host-dominated AGN and 1 as a blazar.

To remedy the current shortcomings, Sánchez-Sáez et al. (2023) has developed a new classifier that uses the entire 4-year long ZTF light curves, not only the alerts, so now it includes light curves of much less variable objects. This new classifier has returned a promising 32 AGN identifications from a sample of the 298 known low mass AGN that have ZTF lightcurves, and that reside in low stellar mass galaxies ($M_{\rm stellar} < 10^{10}\ M_\odot$). Although the classifier does well in recovering the higher mass/higher luminosity end of the low mass AGN sample (see Fig. 1, left) it detects a very low fraction below $\sim 10^5 M_\odot$. **The goal of this proposal is to understand why the lower mass objects are harder to identify and whether there is hope to find IMBHs from planned optical variability surveys or not.**

The optical light curves of AGN appear to have a broken powerlaw power spectrum, where the break is the only observable characteristic timescale. This timescale τ has been shown to correlate very well with black hole mass, for masses ranging from a few times $10^5 M_\odot$ to $10^9 M_\odot$ (Burke et al. 2021). These authors find that τ is approximately 200 days for $10^8 M_\odot$ and drops as the square root of mass towards lower masses. Therefore for our objects of interest, τ is expected to be 20 days for low mass AGN and 6 days or less for IMBH. These short characteristic timescales approach the sampling rate of ZTF, making the variations harder to discern (see Fig. 2, and text in the figure caption).

In Fig. 1 we show the distribution of variability features of all the low mass AGN and IMBH identified spectroscopically that have ZTF lightcurves and inhabit low stellar mass galaxies (289 objects). In red we show the galaxies that are identified as AGN by the new light curve classifier and in blue the ones that were not. The two properties shown correspond to the characteristic timescale of variability (τ) in the y-axis and the amplitude of the variations (σ) in the x-axis. Both these parameters are estimated by modeling the ZTF lightcurves with a damped random walk model. The amplitude of variability σ does not appear to have a large effect on the success of the detection, but the characteristic timescale τ of variability appears to be determinant. We note however that the ZTF lightcurves are sampled only every 3 or 4 days, so if the characteristic timescale of variability is of this order or shorter, it cannot be reliably measured. Higher cadence lightcurves are needed to show whether the objects in blue indeed have short variability timescales or whether we are just recovering noise.

Therefore, there are several reasons why IMBHs might not be detected by the variability classifier:

1. they might be too dim and therefore the variability is hidden in the observational noise;

2. the variability might be too rapid for the cadence of ZTF and therefore the variability pattern is not discernible and confused with noise, or

3. the spectroscopic candidates are not really AGN and therefore there is no intrinsic variability.

IMBH can be disproportionately affected by these difficulties because they are effectively dimmer, their characteristic timescales of variability are expected to be shorter and they are more easily confused spectroscopically with starforming galaxies due to the small width of their broad lines. As a pilot project we aim to observe galaxies in the sample of spectroscopic IMBH and low mass AGN candidates that are missed by the classifier, with a 10-fold improvement in the cadence and better S/N when compared to the ZTF lightcurves, to establish which of these situations apply, if any.

Expected results: We aim to first verify the feasibility of obtaining high cadence lightcurves of sufficient S/N of these targets with SPECULOOS. Use these to establish if the higher cadence reveals a correlated variability pattern or not. If it does then the scattered pattern seen in ZTF light curves will be due to the undersampling of intrinsic variations. Therefore, using the ZTF light curves to train a new classifier with a specific low mass AGN/IMBH class can potentially result in the correct identification of this type of objects. For this purpose, we need to select as many reliable IMBH candidates as possible, to approach a balance with the other classes that have at least 1000 objects. On the contrary, if our higher S/N lightcurves reveal no significant variability even on short timescales we can discard the galaxies as AGN candidates and be satisfied that the lightcurve classifier does not find them. If the higher cadence still reveals uncorrelated variations but the higher S/N shows that these variations are well above the observational noise then we can confirm the objects as variable but not as active galaxies with black hole masses in the range $10^5 - 10^6 M_{\odot}$, since their variability timescales would be too short. These objects would merit further study, beyond the scope of this proposal.

References: • Arévalo 2008 MNRAS, 389, 1479 • Arévalo 2009, MNRAS, 397, 2004 • Baldassare 2018, ApJ, 868, 152 • Baldassare 2020, ApJ, 896, 10 • Bellm 2014, The Third Hot-wiring the Transient Universe Workshop, 27-33 & R. Seaman, 27–33 • Blanton 2011, AJ, 142, 31 • Burke 2021, Science, 373, 789 • Chilingarian 2018, ApJ, 863, 1 • Dong 2012, ApJ, 755, 167 • Förster 2020, arXiv:2008.03303 • Greene & Ho 2004, ApJ, 610, 722 • Greene & Ho 2007, ApJ, 667,131 • Greene 2019, arXiv:1911.09678 • Ivezic et al. 2019, ApJ, 873,111 • Liu 2018, ApJS, 235, 40 • Martínez-Palomera 2020, ApJ, 889, 113 • Sánchez-Sáez 2018, ApJ, 864, 87 • Sánchez-Sáez 2021, AJ, 161, 141 • Sánchez-Sáez 2023, A&A, 675, 195 • Volonteri 2010, A&ARv, 18, 279 • Woods 2019, PASA, 36, e027 • York 2000, AJ, 120, 1579

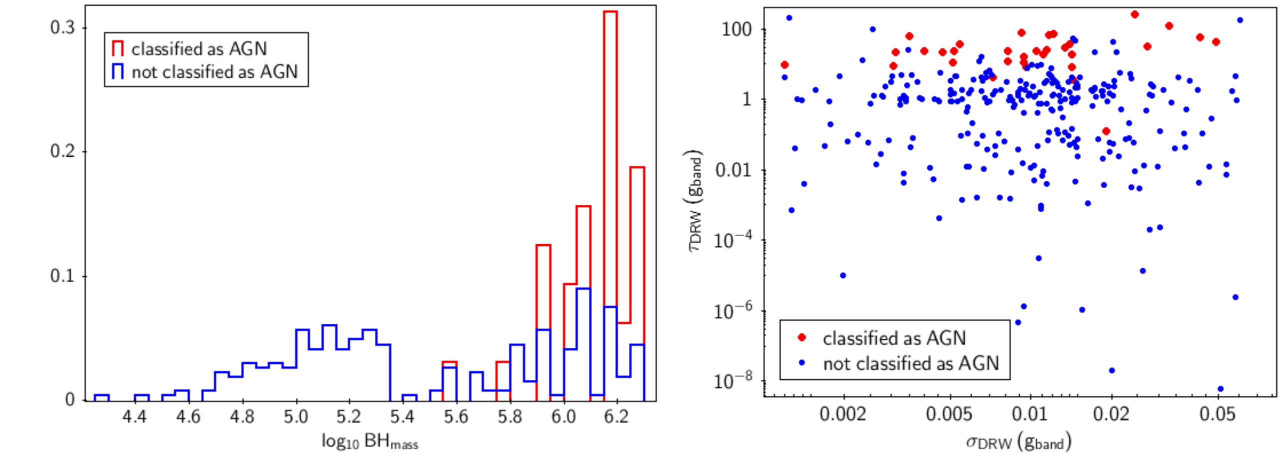

Figure 1: Left: histogram of the masses estimated spectroscopically for the sample of IMBH and low-mass AGN. The different colors show the distribution of masses of those objectes selected as AGN by the new light curve classifier (red) and those which are not selected (blue). Right:variability features of the spectroscopically identified low mass AGN and IMBH candidates. The axes denote the characteristic variability timescale τ in the vertical axis and the amplitude of variability in the horizontal axis. The galaxies identified by the new light curve classifier (red) occupy the longer characteristic timescale ranges, while almost all objects with short characteristic timescales are not identified as AGN (blue).

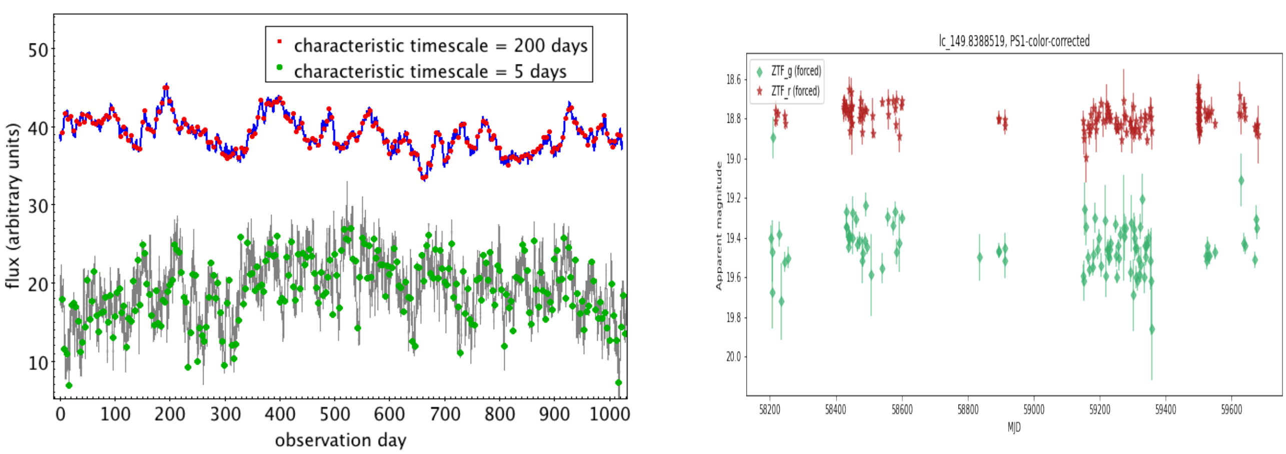

Figure 2: Left: Simulated lightcurves shown to highlight the effect of different characteristic timescales τ. For $\tau = 200$ days, expected for a $10^8 M_\odot$ black hole, the fluctuations are smooth and well resolved, even when sampled every 4 days as the ZTF lightcurves (red dots). For a shorter $\tau = 5$ days, expected for IMBHs with $10^5 M_\odot$ black holes, the rapid fluctuations in grey are not well recovered by the 4-day sampled light curve (green dots). Therefore, if IMBHs have these short characteristic timescales, their 4-day sampled ZTF light curves will look similar to white noise, i.e., only random scatter. We note that no observational noise has been added to these simulations. This **apparent** uncorrelated noise, expected for ZTF light curves of low mass black holes, might prevent the light curve classifier from identifying these objects as AGN. Right: ZTF lightcurves in g and r band of an IMBH candidate. The scattered variability, especially evident in the more variable g band can be confused with observational noise but it might correspond to intrinsic rapid variability as seen in the simulated light curves on the left.

TECHNICAL DESCRIPTION

We aim to obtain high-cadence light curves of our objects using the robotic 1-m telescopes SPECULOOS. Given the sparse distribution of our targets we can only accommodate one per pointing so we aim at observing them repeatedly through the night, for several consecutive nights.

Sample selection: Our parent sample is that of spectroscopic candidates of low mass AGN and IMBH compiled from the black hole mass estimates of Greene & Ho 2004, 2007, Dong 2012, Chilingarian 2018 and Liu 2018. To improve the purity of the sample, which contained some dubious high redshift galaxies, we cross matched this sample with a sample of low stellar mass galaxies, with a threshold of $M_{\mathrm{stellar}} < 10^{10}\,M_{\odot}$, selected from the NASA-Sloan Atlas (NSA; Blanton 2011). This selection returned about 400 targets, of which 298 have ZTF lightcurves of enough quality to calculate variability features (other are outside the ZTF footprint, too dim or too sparsely sampled). Of these, 32 were classified as AGN by the new light curve classifier, our targets belong to the rest of the 298 candidates. From these, we selected all targets with dec<= 25 and with values of the characteristic timescale $\tau < 10$ days and $\tau > 0.5$ which returned 104 potential targets, of which 70 are visible throughout the night for at least a fraction of A semesters. We will select 15 of these objects at a time, depending on the observing nights scheduled so that they can be observed repeatedly during each night.

Cadence: As discussed above, the characteristic timescales of variability of low mass AGN are about 20 days and for IMBHs are about 6 days or less, while the ZTF sampling is about 3–4 days. We propose to use the 3 consecutive night blocks offered by SPECULOOS to observe a a target several times per night, in order to produce about 30 data points per target in between ZTF observations. This rate nicely extends the timescale coverage of the existing lightcurves and can precisely be used to distinguish between scattered noise and smooth trends between ZTF data points. This would allow us to find trends rather than pure scatter in the light curves (see Fig. 2) confirming or rejecting the low mass black hole scenario in each galaxy.

Photometric requirements: We will used the g-band filter, which is the bluest available in SPECULOOS and matches the one used by ZTF public light curves. This is preferred over the r band filter, also used by ZTF, because the stellar light contamination is normally smaller in the g band and the intrinsic variability amplitude of known AGN is also larger at shorter wavelengths (Sánchez-Sáez et al. 2018). The typical size of fluctuations in the ZTF lightcurves of our targets is 0.1 to 0.2 magnitudes, similar to 10-20% variations in flux. To be able to track these we aim at a photometric S/N=100. We note that we will perform relative photometry comparing with other objects in the FOV so this precision only needs to be achieved relative to brighter point sources and not as absolute calibration errors (e.g. as in Arévalo et al. 2008, 2009). We used the Las Cumbres Observatory ETC for the 1m telescopes, since they are similar in size to SPECULOOS, to estimate the exposure required to reach a S/N=100 in an object with g band magnitude of 19, obtaining 10 minutes exposures. Previous experience with SPEcULOOS shos that it is more feasible to make repeated observations of a single object than cycle through many targets, and tis setup will also allow use to bin observations together to control the point to point scatter.

Here we request 45 nights to target the maximum number of objetcs possible, but any multiple of 3-night blocks is useful and we will continue requesting time until all targets are observed.

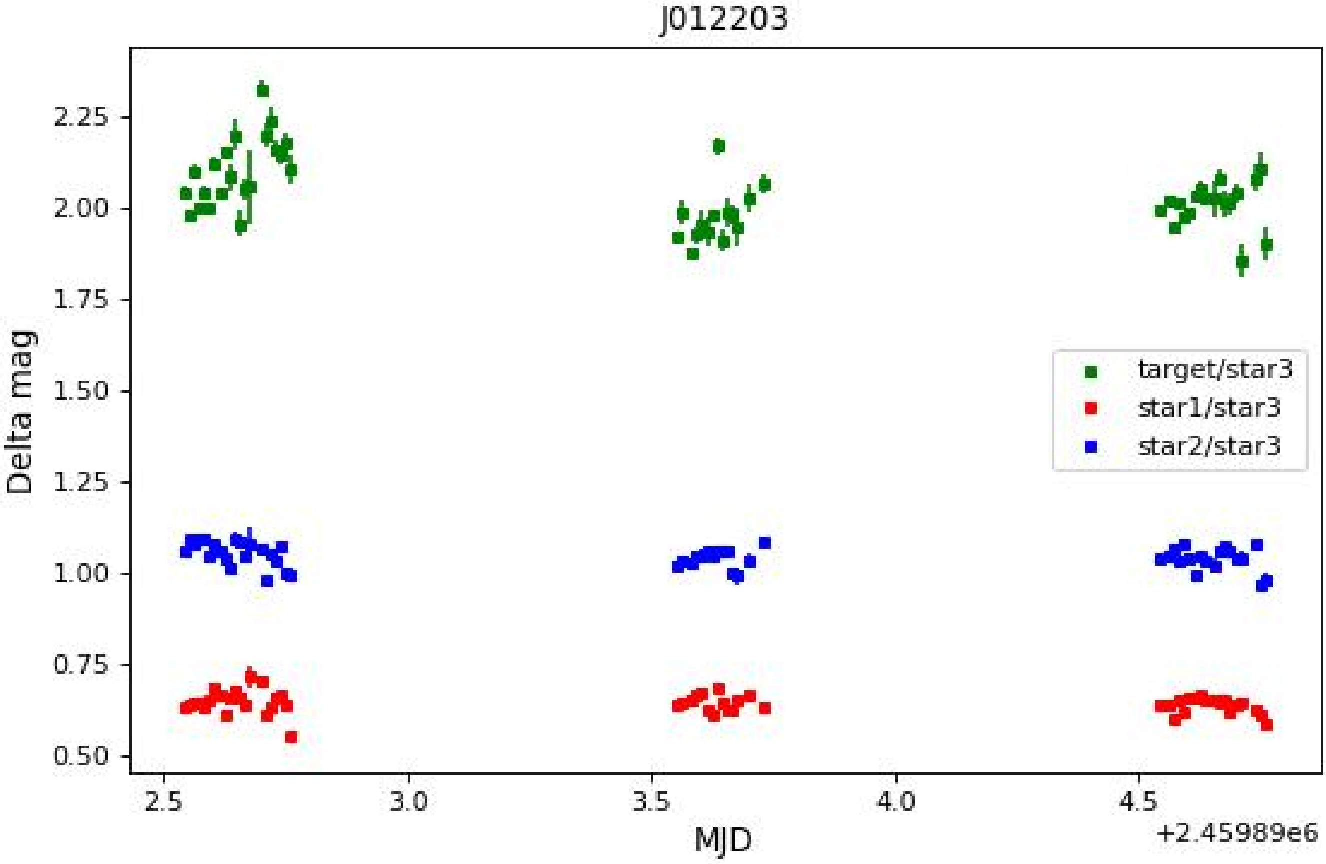

Figure 3: Difference photometry light curves tajen with SPECULOOS on of target in this sample.

`CURRENT STATUS OF THE PROJECT`
The ZTF variability sample was obtained by using a machine learning procedure presented in Sanchez-Saez et al. 2023, which uses complete light curves from ZTF, not just the alert stream, so it includes more photometric points. The use of the ZTF data release light curves allowed us to detect low amplitude variations, that are not easily detected in the ZTF alerts (due to its 5-σ variability threshold). Efforts are being made to use the new classifier to select robustly low mass AGN and IMBH and this proposal is key to select a reliable training set of ZTF light curves of this type of objects.

SPECULOOS light curves of 6 objects in this sample were taken in 2022B. In Fig. 3 we show the light curve of one of the objects, together with the light curves of two comparison stars. The lightcurves are computed through difference photometry to other stars in the field. Although there is scatter, we do detect higher variance in the target than in the stars, particularly from one night to the others. These fluctuations are on the order of 0.1 magnitude, as expected for AGN of this luminosity.

`STUDENT THESIS` The data will be used for the undergraduate thesis of an identified bachelor student. He has so far develop the Python scripts to reduce the existing SPECULOOS data, homogenize the PSFs and build difference photometry light curves. With data from this proposal he will use his pipeline to produce lightcurves on the fly and measure their variability in time for his defence at the end of the first semester of 2024.