A photometric population study on nuclear transient events from the Zwicky Transient Facility

MSc thesis

Tim van der Vuurst

Universiteit Leiden
Observatory

A photometric population study on nuclear transient events from the Zwicky Transient Facility

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science
in
Astronomy

Author:	Tim van der Vuurst
Student ID:	s2557398
Supervisor:	Dr. Sjoert van Velzen
Second corrector:	Prof. Dr. Jelle Kaastra
Project duration:	October 2023 – July 2024

Cover image:	The Night and All Its Stars by Nate Rayfield
Template style:	Thesis style by Richelle F. van Capelleveen

Universiteit Leiden
Observatory

P.O. Box 9513, 2300 RA Leiden, The Netherlands

Abstract

Context: Wide-field (optical) surveys are presenting more insights on nuclear transient events than ever possible before, causing population studies on all types of transient to start emerging. However, little is known about the mechanism behind extreme flares in Active Galactic Nuclei (AGN). Multiple mechanisms exist that may explain where they come from, including Tidal Disruption Events (TDEs) in the galactic nucleus.
Aims: This work aims to investigate the possibility of TDEs being the mechanism behind extreme AGN flares, as well as investigating trends in the sub-populations nuclear transient events.
Methods: We use a set of 14258 nuclear transient alerts from the Zwicky Transient Facility (ZTF) survey and look at the photometric properties in the optical. We attempt to fit a Gaussian rise plus exponential decay model to the optical light curve and in so doing identify extreme flares. We further estimate the black hole mass and use observations from the NEOWISE survey for the purpose of calculating the strength of possible dust echoes.
Results: We find that of the total of 725 nuclear transients marked as extreme, 617 are as of yet unclassified. Through a two-sample Kolmorgorov-Smirnoff (KS) test, we find that the black hole masses of extreme AGN and non-extreme AGN with dust echo potential are likely drawn from different distributions (p-value of 0.00907). We further find that the certain model parameters for TDEs and AGN correlate, though none of the parameter correlations that were not expected due to definition appear for both TDE and AGN. We moreover present a list of 50 unclassified nuclear transients selected on having the strongest dust echoes. We find some to have been reported as a type of Supernova (SN) to the Transient Name Server (TNS), but suggest this is likely false for some of these transients based on infrared afterglow measurements.
Conclusions: The unique parameter correlations suggest that TDEs and AGN do not constitute the same population. A follow-up two-sample KS test on the black hole mass estimates enforces this further. While we cannot rule out the possibility of some of the extreme AGN flares being caused by TDEs, we do find that it does not hold for all of the extreme AGN flares investigated in this work.

Nomenclature

Acronyms

Notation	Description	Page List
AGN	active galactic nucleus	1–3, 9, 13–15, 17, 18, 21–23, 25, 26, 28, 29, 33–38, 47
BH	black hole	1, 2, 13, 20, 23, 24, 28, 29, 31
CDF	cumulative distribution function	18–25, 28, 34, 47
Dec	declination	4
DN	Digital Number	6
jd	Julian Date	9, 17
KS	Kolmorgorov-Smirnoff	28, 38, 47
PDF	probability density function	18–25

Notation	Description	Page List
PSF	Point Spread Function	2, 6
RA	right ascension	4
RMS	root mean square	5, 6, 14
SN	supernova	1, 3, 14, 17, 18, 21, 25, 31, 32, 36–38
TDE	Tidal Distruption Event	1–3, 14, 16–18, 20, 23, 25, 26, 29, 33–38, 47, 48
TNS	Transient Name Server	16, 22, 29, 31, 36–38
ZTF	Zwicky Transient Facility	2–5, 31, 36

1. Introduction

Astronomical transient events are generally relatively short-lived and/or variable events, with variability observed on timescales anywhere between a few milliseconds and many years. In some types of transients, such as Supernovae (SNe) and Tidal Disruption Events (TDEs), some object - often a star - is (at least partially) destroyed, violently releasing light. In this process such energy is released that the events are visible over cosmic scales - so much so that records of the first observed SN date back well over a thousand years. With the invention of the telescope and the eventual introduction of wide-field surveys, the observations of a wide variety of transient events have never been more bountiful, finding many instances of different transient types such as SNe, active galactic nuclei (AGN) and TDEs. The latter two are intrinsically related due to their more similar nature - both being induced by the gravity of a supermassive black hole (BH), often in the nucleus of a galaxy. Transient events that occur in galactic nuclei are usually referred to as nuclear transients and will serve as the main focus of this work.

AGN were first observed in the 20th century (e.g. Seyfert 1943) and were later theorized to be the result of a supermassive BH ($M_{BH} \approx 10^6 - 10^{10}\ M_{\odot}$) accreting mass (Salpeter, 1964; Zel'dovich & Shakura, 1964) - a theory that has worked phenomenally well to describe the basic properties of accretion discs. However, in recent years a number of extreme, relatively short-lived and smoothly evolving flares that are directly associated with AGN have been observed (e.g. Drake et al. 2011; Blanchard et al. 2017; Frederick et al. 2021; Graham et al. 2023). The light emitted by such flares significantly exceeds the level expected by the stochastic variability of AGN and thus need a different driving mechanism to explain their origin.

Only a few such mechanisms are known. Frederick et al. (2021) put forth that these mechanisms include: particularly UV-bright transient flare events associated with a change in accretion rate (Trakhtenbrot et al., 2019), transient events with a double peak in the line profile which have been linked to accretion disc emission (Halpern & Eracleous, 1994) and changing-look AGN - which are AGN whose spectroscopic classification changes drastically following a rise in continuum level, theorized to be connected to unstable changes in accretion state (see Ricci & Trakhtenbrot (2023) for a recent and extensive review). Another class of extreme AGN flares are the Bowen Fluorescence Flares (see e.g. Makrygianni et al. 2023). Moreover there is the intriguing possibility of extreme flares in AGN being caused by tidal disruption events (as in e.g. Frederick et al. 2021).

TDEs occur when a star passes within the Roche limit of a (massive) BH and gets torn apart by the tidal forces, releasing a huge amount of energy across the electromagnetic spectrum in the process (see the somewhat recent review by Gezari (2021) for a thorough overview). The concept of TDEs first turned up in theory in the 1970s and -80s (?Rees, 1988) and the first observational evidence came years later, first using soft X-ray bursts (Donley et al., 2002) and later with photometric studies in the optical (van Velzen et al., 2011).

At the time of writing, many wide-field optical surveys are producing significant numbers of detections of transient events, including the intermediate Palomar Transient Factory (iPTF) (e.g. Hung et al. 2017; Caplar et al. 2017), ASAS-SN (e.g. Holoien et al. 2014; Brown 2018; Necker et al. 2022), Pan-STARRS (e.g. Tonry 2013; Heinis et al. 2016; Nicholl et al. 2019b) and the Zwicky Transient Facility (ZTF) (e.g. Ward et al. 2021; van Velzen et al. 2021b; Hammerstein et al. 2023; Somalwar et al. 2024; Stein et al. 2024; Wiseman et al. 2024; van Velzen et al. 2024). This work will make use of data gathered from the ZTF mission.

Thanks to surveys such as the ZTF, sufficiently many transient events of all types have been observed and classified for population studies to start emerging (e.g. van Velzen et al. 2021b; Hammerstein et al. 2023). Wide-field optical surveys have been vital for discovering TDEs and accounted for almost two-thirds of all the reported TDEs in 2021 (Gezari, 2021). The optical properties of TDEs (e.g. van Velzen et al. 2011, 2020; Mummery et al. 2023) and AGN (e.g. Smith et al. 2018; Ward et al. 2021) have been studied extensively and play vital roles in expanding our understanding of the mechanisms behind them.

Photometric surveys such as the ZTF often make use of difference imaging to observe nuclear transient events. This is because it might be the case a transient event is sufficiently weak that it does not appear to be statistically discernible from light emitted by the host galaxy. By the nature of difference imaging, the observations all have a so-called reference image - which captures the host galaxy in the case of nuclear transients. This means the observed flux for a given source is the relative offset of the transient with respect to that of the host galaxy in the reference image (Bellm et al., 2018). The ZTF pipeline for processing a detection includes many steps, including Point Spread Function (PSF) fit photometry - the goodness of which is used to rescale errors as we will see later. For a full overview of the ZTF detection pipeline, see Masci et al. (2018).

Besides just the optical, transient events tend to be observed at various wavelengths. This is because they have multiple different properties that require a multi-wavelength approach. For instance, surveys such as the Near-Earth Object Widefield Infrared Survey Explorer (NEOWISE) provide many valuable insights into the properties of transient events in the infrared, such as the presence of dust reverberations (e.g. Dou et al. 2016a; van Velzen et al. 2024).

In short, a dust reverberation happens when a transient event is surrounded by dust. Following the event, the dust efficiently absorbs photons from the transient event which causes it to heat up. This yields isotropic emission that peaks at a few micron (van Velzen et al., 2021a). Due to the time delay between the initial burst and the IR emission the term "dust echo" is often used, as will be the case in this work.

Dust echoes are particularly useful in studies of transient events overall since they provide a possible way to discern SNe and types of AGN (van Velzen et al., 2024). Dust echoes have been well established for an extended period of time for both AGN (Edelson & Malkan, 1986; Robson et al., 1986; Barvainis, 1987) and SNe (Dwek, 1983; Bode & Evans, 1983) but they have not been observed in TDEs until fairly recently. TDE dust echoes were first theorized by Komossa et al. (2009) (see also the important theoretical paper by Lu et al. 2016) only to be observationally confirmed not long after by van Velzen et al. (2016); Jiang et al. (2016) and Dou et al. (2016b).

Observations of (nuclear) transient events are often further supplemented by spectroscopic analysis, see e.g. Koss et al. (2017); Laurenti et al. (2022) (AGN) and Charalampopoulos et al. (2022) (TDEs). Certain spectroscopic features can be used to classify (sub-)types of transients (e.g. Fremling et al. 2020). In this work however, we constrain ourselves to the optical lightcurves and the infrared afterglow of various transients - subdivided into SNe, AGN, TDEs and unclassified transients - and denote eventual spectroscopic investigation as outside the scope of this particular work.

The main focus of this research will be to further investigate extreme AGN associated flares. We will investigate whether the population of such extreme AGN flares is statistically separate from regular AGN activity and look into the possibility of the mechanism behind the extreme flares to be TDEs. We will do so by investigating a sample of the difference photometry measurements of more than 10000 nuclear flares from the ZTF survey, supplemented by photometric infrared observations from the NEOWISE survey for the purposes of investigating possible dust echoes.

The data retrieval and cleaning is discussed in more detail in chapter 2. The subsequent methodology of investigating trends in the populations is covered in detail in chapter 3 and was in very large part inspired by van Velzen et al. (2024). The results are presented in chapter 4 and are subsequently discussed in chapter 5. The final conclusions from this work are summarized in chapter 6.

2. Data

In this work we use data from two photometric surveys: the Zwicky Transient Facility (Bellm et al., 2018; Masci et al., 2018) and NEOWISE (Wright et al., 2010; Mainzer et al., 2011, 2014). Below, we describe these and explain how the data was retrieved and reduced.

2.1 Gathering data

The Zwicky Transient Facility survey scours the northern sky in search of transient events and has done so since 2017. The ZTF surveys work with alerts which each get a unique identifier, meaning some sources that alerted the system twice may have multiple ZTF IDs but constitute the same transient. An alert is to be interpreted as a 5σ detection in the difference imaging (Bellm et al., 2018). Not all transient alerts originate from a nuclear transient specifically, so these need to be searched for specifically (e.g. Wiseman et al. 2024).

In this work we use the results of a re-run which attempted to gather all results from searches for nuclear transients in the ZTF survey up until the summer of 2023. This resulted in a list of 14258 alerts. Therefore the list is not currently exhaustive, with more and more transients being discovered and classified using data from the ZTF survey constantly (e.g Somalwar et al. (2024)). Our data includes the ZTF identifier, right ascension (RA) and declination (Dec) of the transients in question.

2.1.1 Forced photometry data

Given the coordinates of an object, “forced” photometry at the specified location is possible. It is forced in the sense that the position of an object of interest is known from prior knowledge (e.g. detection in another bandpass), following which, photometric measurements at that location are taken. We attempt to gather the forced photometry data for every one of the 14258 transients from the ZTF. The ZTF surveys are done at the Palomar Observatory[1]. The used filters are named ZTF-g, ZTF-r and ZTF-i (which will henceforth be referred to as the g, r and i filters respectively), the relevant details of

[1]For more details, visit the Spanish Virtual Observatory filter profile service website here.

which are specified in table 2.1.

Of the 14258 transients, 13 had unspecified coordinates, thus making it impossible to gather more data from the ZTF. During the retrieval of the data, three more instances retrieved an empty table, reducing the as of yet viable set of data before cleaning to 14241 instances.

For the purposes of finding the dust echo strength and infrared luminosity of certain transients we look at the NEOWISE survey, specifically the W1 filter[2] centered at 3.35 μm (Wright et al., 2010), as shown in table 2.1. This information is available for 9603 of the transients in the dataset.

Filter	λ_{eff} (Å)	ν_{eff} (10^{14} Hz)
ZTF-g	4746.48	6.316101
ZTF-r	6366.38	4.708994
ZTF-i	7867.41	3.810561
WISE W1	$3.35{\cdot}10^4$	0.894

Table 2.1: *Effective wavelength of the ZTF-g, ZTF-r, ZTF-i and WISE W1 filters as specified by the Spanish Virtual Observatory filter profile service for the ZTF filters and Wright et al. (2010) for the WISE W1 filter. The corresponding effective frequencies are given as well.*

2.2 Data reduction

The retrieved data needs to be investigated to see if it is at all viable and reliable. This process starts with a preliminary quality control as described in section 6.1 of Masci et al. (2023). We furthermore remove epochs where the zero-point root mean square (RMS) is more than 0.05 dex from the median. The requirements for marking bad detections in the pre-processing are shown in table 2.2. After this pre-processing, the set is reduced to 14176 instances, as in some instances none of the data was deemed viable according to the criteria given in table 2.2.

2.2.1 Further quality control

After the preliminary quality control, some more cleaning and transforming of the data is done on a per filter basis. This means the steps described below are performed individually for all filters - given that there are observations in that filter. For each transient we find that there is at least 1 filter with observations.

First, it must be noted that usually the data from the ZTF is taken in different CCD fields. Measurements taken in different fields are not directly compatible since it is possible that each field combination will have used its own reference image for image-differencing

[2] See the NEOWISE explanatory supplement here for more information.

Feature	Description	Requirement
`procstatus`	Per-epoch processing status code	$= 0$
`scisigpix`	Robust sigma per pixel in sci image [DN]	< 20
`sciinpseeing`	Effective FWHM of science image [pixels]	< 4.0
`zpmaginpscirms`	RMS of zero-point baseline	< 0.05
`infobitssci`	Status of processing & instrumental calibration	< 33554432

Table 2.2: *Pre-processing requirements for marking bad data epochs in a set of data for a given transient used in this work. For a full, comprehensive description of the features see Masci et al. (2018) and Masci et al. (2023). FWHM refers to the full width at half maximum.*

(Masci et al., 2023). Therefore, the choice is made here to look only at the data taken in the primary field for each transient. We define the primary field of the observation of a given transient as the field in which most measurements exist *after* the preliminary quality control has filtered out some epochs.

Next, we re-scale the error on this primary field data as described in section 6.3 of Masci et al. (2023). In short, we scale the uncertainties on the flux by the square-root of the median reduced-χ^2 metric of the PSF fit (denoted as `forcediffimchisq`). We choose the median instead of some other average metric as it was found to be the more robust choice.

Finally, the flux is converted from Digital Number (DN) to μJy, after which the error on the flux in each epoch is clipped such that the relative error is at least 1%.

An example of the resulting lightcurve of this data reduction procedure is shown in figure 2.1. In this figure it becomes clear that data points with the largest error margins as well as outliers tend to be filtered out by this process relatively well. A complete overview of how the cleaning process influences the number of data points in the relevant filter bandpasses is shown in figure 2.2. It can be seen that before the cleaning process both the g and r filter have a very similar distribution in number of epochs, but the g filter tends to suffer from greater decreases in this amount. It is moreover clear how much data is generally removed, before cleaning there are a couple of instances with $\gtrsim 6000$ epochs, but after cleaning the highest number of epochs in a certain filter of a certain transient is 1311.

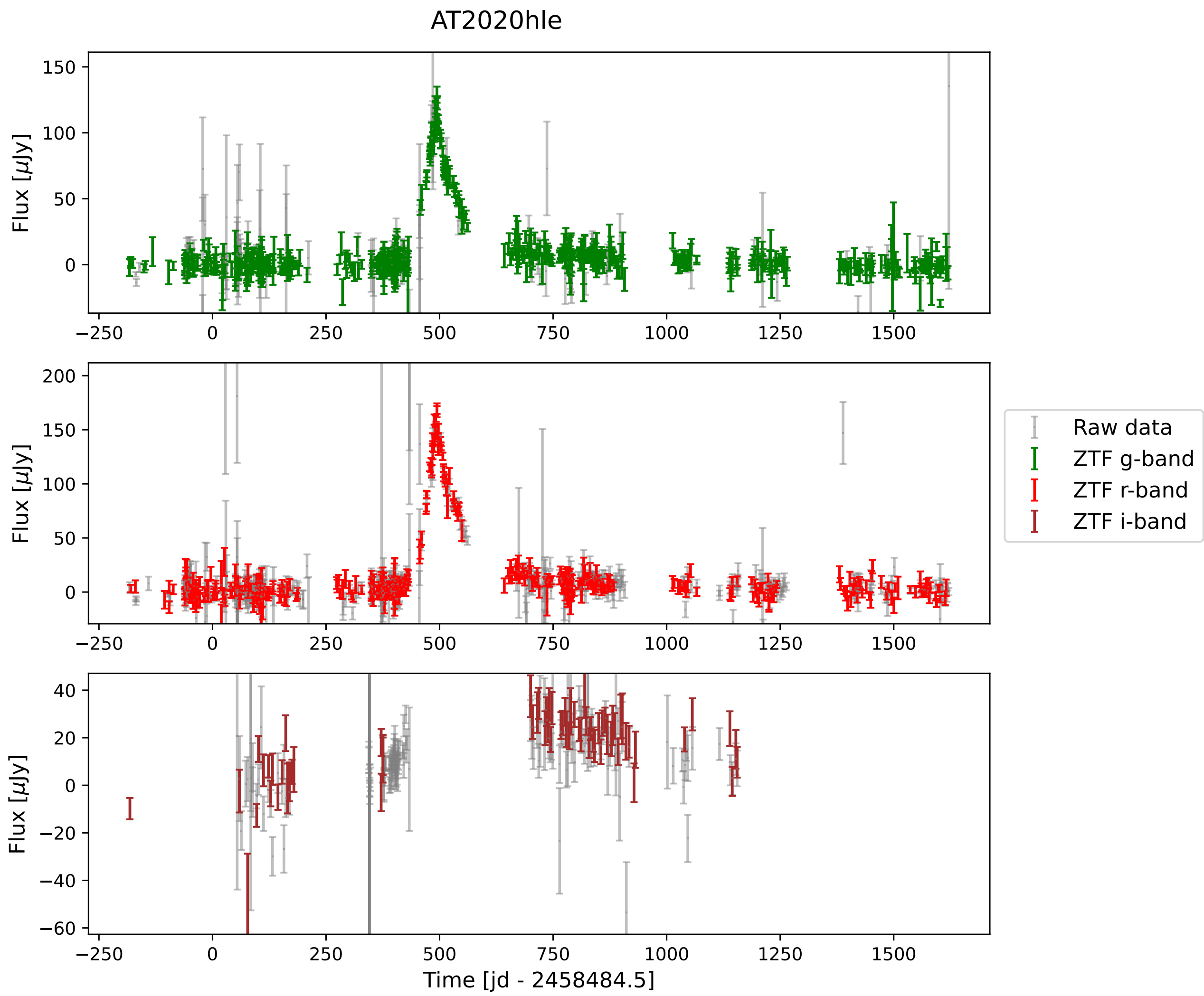

Figure 2.1: *The raw and clean data of AT2020hle (ZTF18abjjkeo) in all 3 bandpasses. The gray points in each plot show the total, unprocessed dataset in that filter. The colored points are over-plotted to show the remaining points after quality control as described in section* 2.2*. For ease, we set the time w.r.t. 01-01-2019 (jd 2458484.5).*

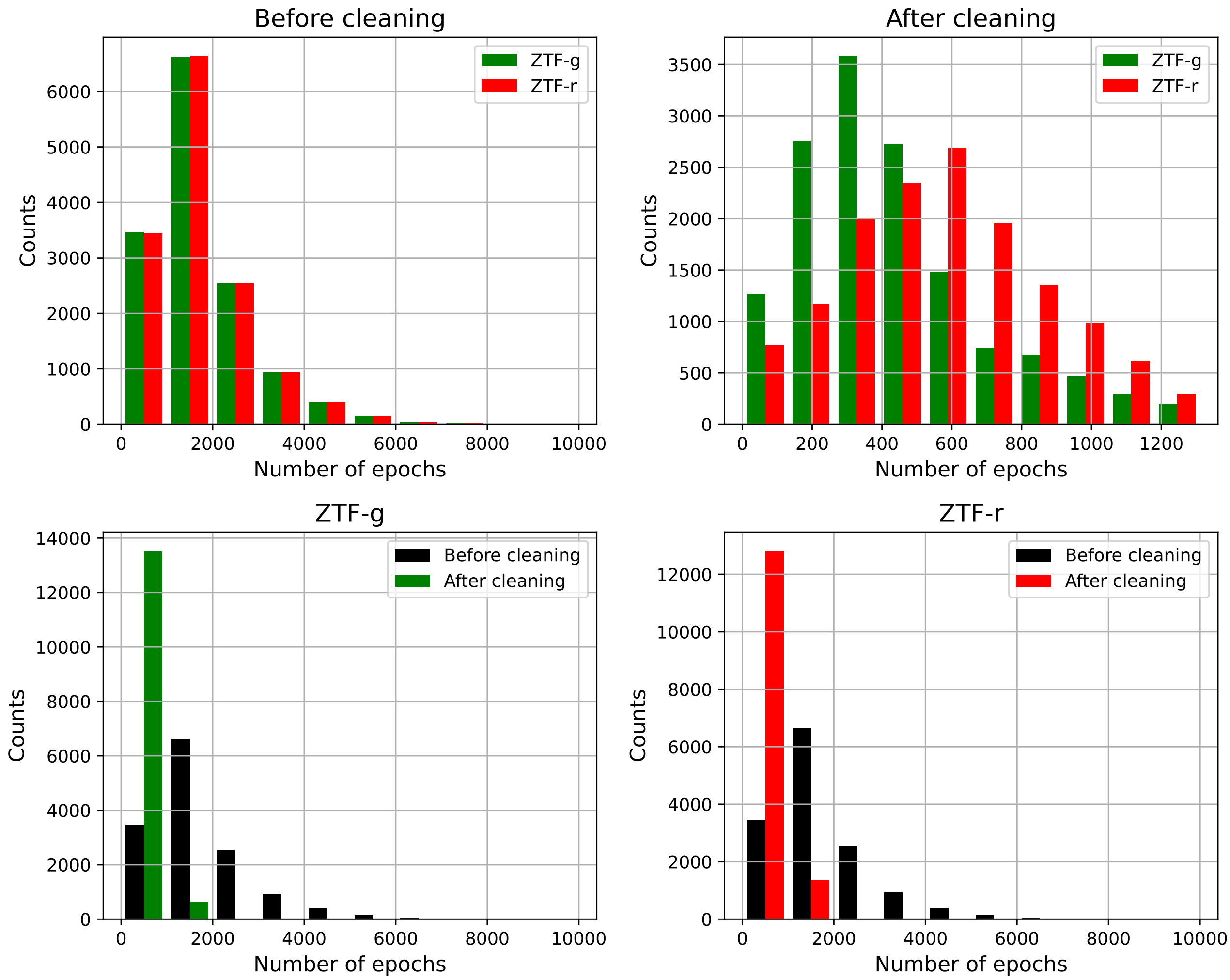

Figure 2.2: *Visualization of how much data is left after the cleaning process in both the ZTF-g and ZTF-r filters. The horizontal axes all show the number of epochs (that is to say how many data points) a ZTF forced photometry file contains and the vertical axes show how many files have that amount of epochs.*

3. Methods

Now that the data has been processed and reduced to a usable and reliable form, the remaining instances can be processed. For this, the assumption is made that each transient in the set has only a single significant burst. For most AGN this typically won't hold, since they tend to portray variability for their entire lifetime. Even so a single burst can often still be fitted to a subset of AGN transients as will be shown later. The processing of data consists of two parts. First, each transient will be fit to the expected model for a transient burst using a χ^2 minimization routine. After the parameters are found, studies can be done on the properties of the population to look for patterns and peculiarities.

The code used to implement the analysis below may be found on GitHub here.

3.1 Fitting

3.1.1 Lightcurve model

We fit the lightcurve of the alerts to a somewhat simplified model consisting of a Gaussian rise followed by an exponential decay in flux (as in Graham et al. (2023); van Velzen et al. (2021b)). This means it is a piece-wise function defined over the entire range of time which (usually) quickly rises and decays again. The choice of such a model was made since we want to find relatively short transients, since these constitute the extreme flares as we will see later. In that case a model like the one used here holds. The mathematical expression of the model implemented in this work is shown in equation (3.1).

$$F(t) = \mathcal{B}_r(T, \nu_0, \nu_1) \cdot \begin{cases} F_p \exp\left(\frac{(t-t_0)^2}{2\sigma_{\text{rise}}^2}\right) + F_0 & t \leq t_0 \\ F_p \exp\left(-\frac{t-t_0}{\tau_{\text{dec}}}\right) + F_0 & t > t_0 \end{cases} \tag{3.1}$$

Here $F(t)$ (in μJy) is the flux of the transient as a function of time t given in Julian Date (jd). F_p is the flux of the peak with respect to the baseline F_0, both of which are in μJy, t_0 is the time at which the burst peaks in jd and σ_{rise} and τ_{dec} are the e-folding time for the rise and decay (in days) of the transient flare respectively. Finally, $\mathcal{B}_r(T, \nu_0, \nu_1)$ is the ratio of a black body $B(T, \nu)$ with temperature T evaluated at frequency ν_1 to that same black body evaluated at frequency ν_0:

$$\mathcal{B}_r(T, \nu_0, \nu_1) = \frac{B(T, \nu_1)}{B(T, \nu_0)}. \tag{3.2}$$

Here, ν_1 is the effective frequency of the filter currently being looked at (as described in table 2.1 and ν_0 is defined as the weighted average of the effective frequencies of the g and r filter (again, as in table 2.1), with the weights being the the amount of viable epochs in each filter (N_g and N_r respectively):

$$\nu_0 = \frac{N_g \nu_g + N_r \nu_r}{N_g + N_r}. \tag{3.3}$$

This way, ν_0 approximates a "central" frequency for the combined g and r filter data. Finally, for completeness, the black body equation is as follows:

$$B(T, \nu) = \frac{2h\nu^3}{c^2} \frac{1}{\exp \frac{h\nu}{k_B T} - 1}, \tag{3.4}$$

with h the Planck constant, k_B the Boltzmann constant and c the speed of light.

The black body ratio term is added to the model so that finding the optimal parameters is viable for data on both the g and r filter at the same time. This means the fitting of the model in equation (3.1) will be done on the combined data of the g and r filter. During this process, each point is appropriately scaled according to the black body temperature depending on its central frequency, increasing the data available to use per fit, relieving the necessity for filter dependent parameter values, and allowing us to make an estimate of the temperature.

Due to the exponential nature of equation (3.1) and possible high values, some parameters are evaluated in log-space such that convergence on the parameters is quicker, see e.g. table 3.1.

3.1.2 Baseline correction

When analyzing a lightcurve of a transient event is important to correct for the baseline flux, which is generally non-zero. The origin of a non-zero baseline in the forced photometry data is due to systematics, specifically in the reference image, creating an offset from the expected zero baseline level (Masci et al., 2023). However, this holds mostly for “true” transients which have only a single burst - in this case the method described by Masci et al. (2023) is applicable. However, we are dealing with many sources which are variable in the baseline. This may cause negative flux in the difference image. Therefore, we opted for a separate parameter F_0 to act as the baseline in the model used in this work.

It was apparent in the data that the baseline offset was filter dependent. This is due to the fact that the reference images for both filters were not taken at the same time. As such the choice was made to fit the the model described in equation (3.1) separately for the g and r data first. Then, the two baseline parameters $F_{0,g}$ and $F_{0,r}$ are extracted from this fit and used to scale the combined g and r data set to get as close as possible to a zero baseline. Since these initial two fits are done on a per filter basis, no temperature correction through a black body ratio is necessary in these fits (e.g. $\mathcal{B}_r(T, \nu_0, \nu_1) = 1$). After this process, the model in equation (3.1) is fit to the scaled g and r data simultaneously, with F_0 set to 0 explicitly. This time, the temperature correction term is taken into account. In so doing, we attempt to infer an optimal value for T. The temperature correction refers to the scaling by a factor $\mathcal{B}_r$ on the parameter F_p based on the estimated temperature T and the central frequency of the bandpass ν (as in table 2.1) as described in equation (3.2).

3.1.3 Initial guesses and boundings

In order for the χ^2-minimization routine to converge to optimal parameters, it often needs good initial guesses for the various parameters. Furthermore, in order to restrict the values to both physically reasonable ones and to forgo divergence, we need to set uninformative uniform priors on the parameter values, which will henceforth also be referred to as bounding boxes or boundings. A complete overview is given in table 3.1.

For an initial guess of the time of the peak t_0, a cross-correlation technique is implemented here. This method is adopted over simply taking the time of the highest recorded flux as the initial guess for both t_0 and F_p, because extreme outliers might give a very wrong initial guess, stopping the model from properly converging.

The method works by calculating the cross-correlation between a Gaussian model $G(t)$ and the combined forced difference flux of the g and r filter $F(t)$. The Gaussian model is centered at points in time between the second earliest existing data-point and either the latest data point or 31-12-2022 23:59:59 (whichever is an earlier time) with a time step of 2 days such that it is centered at N different times. We choose this specific upper-bound on where the peak may be since we want there to be sufficient data post-peak such that we can reliably fit the decay of the model as well. See for example figure A.2 for an example of a second peak which is too recent to properly investigate the decay of the lightcurve.

The Gaussian model $G(t)$ has a standard deviation of 10 days and an amplitude of 1 μJy. The model, centered at time t_j, is then:

$$G(t_i, t_j) = \exp\left(\frac{(t_i - t_j)^2}{200}\right). \tag{3.5}$$

The data is linearly interpolated such that it may be evaluated at the same time steps as the Gaussian model. The cross-correlation of a Gaussian centered at t_j with the data (C_j) is then simply the sum of the product of the predicted flux by the Gaussian model and the linearly interpolated flux from the data at all time steps t_i:

$$C_j = \sum_{i=1}^{N} G(t_i, t_j) \cdot F(t_i). \tag{3.6}$$

A good guess for the time of the peak t_0 is then the central time at which C_j is maximal, denoted $t_{0,cc}$. Due to the simplifications in the model, such as the linear interpolation and the time step, $t_{0,cc}$ usually is not an accurate enough estimate in its own right, but it is found to be good enough to greatly increase the convergence of the χ^2 minimization routine.

The value for t_0 is bounded in the range $[t_{0,cc} - 365, t_{0,cc} + 365]$ during the fitting procedure, since the accuracy of the cross-correlation method is usually a good initial guess, but needs sufficient room for error.

The model itself is fit exclusively on the epochs in the range $[t_{0,cc} - 365, t_{0,cc} + 720]$. We expect this to be enough to encapsulate the rise and decay of a typical transient burst (see e.g. van Velzen et al. (2020),van Velzen et al. (2024) for results on TDE timescales).

For the baselines $F_{0,g}$ and $F_{0,r}$ the initial guess is set at the median value for the flux at epochs *outside* the range $[t_{0,cc} - 365, t_{0,cc} + 720]$. The values of these baselines are bound between the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles of the flux in the same epochs.

Parameter	Initial guess	Prior
$\log_{10}(\sigma_{\text{rise}})$	1	$[0, 4]$
$\log_{10}(\tau_{\text{dec}})$	2.5	$[0, 4]$
$\log_{10}(F_{\text{p}})$	$\log_{10}(\max(F_{\text{p}}))$	$[1, \log_{10}(\max(2F_{\text{p}}))]$
$\log_{10}(\text{T})$	4	$[3, 5]$
$F_{0,g}(*)$	$\text{med}(F_g)$	$[P_5(F_g), P_{95}(F_g)]$
$F_{0,r}(*)$	$\text{med}(F_r)$	$[P_5(F_r), P_{95}(F_r)]$
t_0	$t_{0,cc}$	$[t_{0,cc} - 365, t_{0,cc} + 365]$

Table 3.1: *Initial guesses and uniform priors set on the fittable parameters as described in equation* (3.1). *The* n^{th} *percentile of* x *is denoted as* $P_n(x)$. *The median of* x *is denoted* $\text{med}(x)$.
(*): F_g and F_r are evaluated only on the epochs that fall outside of the range $[t_{0,cc}-365, t_{0,cc}+720]$.

3.1.4 ZTF-i filter

The data in i filter is not used during the fitting process, since we are interested in the optical lightcurve. However, a separate partial fit was done to the i data (if any was available after cleaning) in order to visualize how the found model generalizes to lower frequencies. The peak $F_{p,i}$ and baseline $F_{0,i}$ are re-fit to the i-band data whilst keeping the other parameters from equation (3.1) the same as in the optical fit. This set of parameters is not used for anything other than visualization of the model (see appendix A), and as such $F_{p,i}$ and $F_{0,i}$ are strictly not included in any population study described in section 3.2. Any further analysis of the infrared in this work is strictly done with data from the NEOWISE survey.

3.1.5 Quality of fit

During the fitting process, it was found that for only 13456 instances the fitting procedure did not raise an error. For these remaining 709 the reason behind the error varies.

First, it may have occured that the maximum number of function calls for the χ^2 optimization process (set to 5000 here) was exceeded before convergence on the optimum parameters. This was usually caused by the data being too noisy or unlike the model described in equation (3.1).

Second, for several sources there was no viable data left in the g filter, r filter or both after the cleaning process. The choice was made to skip all such sources here, even if they were missing data in only one filter, since the fitting process focuses on fitting on the g and r filter simultaneously.

Third, several errors occured because there was no viable data outside of the range $[t_0 - 365, t_0 + 730]$ in at least one of the g and r filter. This means no baseline could be fit and as such these instances were skipped. It may also have occured that there was not enough data outside of this interval, causing the 5^{th} and 95^{th} percentiles that act as the bounding box for the baseline parameter to have the same exact value, thus making it impossible to find the optimum value.

Of the remaining fits that did converge, some inferred parameters may not be of good enough quality to use in population studies. As such, a set of rules was defined to classify bad fits: the error on the found optimum values of $\log_{10}(\sigma_{\text{rise}})$, $\log_{10}(\tau_{\text{dec}})$ and $\log_{10}(F_{\text{p}})$ must all be ≤ 0.2 dex for the fit on that transient to be considered "good".

3.2 Further properties and extreme flares

Given the result of the fits we can look at the distributions of the different parameters of the model. Beyond these, associated black hole mass estimates, infrared dust echoes and classifications are important to investigate. Finally, we must define what constitutes an extreme flare.

3.2.1 Black hole mass estimates

Estimates of the BH mass M_{BH} were found using the methods described by van Velzen et al. (2024, section 2.3). In short, a BH mass can be estimated if optical spectra are available. The BH mass is then estimated by one of two methods, either a scaling relation based reverberation mapping for type I AGN or a $M_{\text{BH}} - \sigma$ (mass-velocity dispersion) relation for the non-AGN transients. Since this does not require information from the fitting procedure, it is entirely possible to have a BH mass estimate for a flare which did not yield reliable fit results.

3.2.2 Infrared properties

Strong infrared dust echoes are expected for TDEs (Dou et al. (2016a),van Velzen et al. (2021a)), and as such it is of great value to quantify these and look at their distributions. The dust echo strength $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}}$ is defined here exactly as in van Velzen et al. (2024, section 2.2).

Besides just the dust echo strength, we also calculate the infrared luminosity L_{IR} associated with the post peak mean flux measurements of the NEOWISE mission. For this, redshift measurements - gathered from the same spectra as used for black hole mass estimates - are needed to convert from flux to lumonisity. The relation between L_{IR} and $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}}$ is useful for characterizing certain flares as can be seen in van Velzen et al. (2024) and we will see later in this work.

3.2.3 Classifications

A subset of the sources used in this work (1955 instances) have known classifications (van Velzen et al., 2024). These classifications are TDE, AGN and SN. The other, unclassified, flares we henceforth refer to as "unknown". The TDE and SN classifications originate from spectroscopic analysis, while the classifications of AGN originate from the type-I AGN sample by Liu et al. (2019).

3.2.4 Extreme flares

In order to define a flare as extreme and differentiate it from regular AGN variability, certain demands are set on its parameter values. To classify as an extreme flare we demand: $\sigma_{\mathrm{rise}} \leq 100$ days, $\tau_{\mathrm{dec}} \leq 500$ days and either $\Delta m_g \leq 1$ or $\Delta m_r \leq 1$, where Δm_x is the magnitude of the temperature corrected fitted peak of the burst F_p with respect to the reference magnitude in filter x - either g or r. Then we define the peak magnitude $m_{p,x}$ as this scaled flux converted to an AB magnitude in filter x. Then:

$$\Delta m_x = m_{p,x} - m_{\mathrm{ref},x}. \tag{3.7}$$

Notably then, $m_{p,x}$ is measured with respect to the fitted baseline $F_{0,x}$ and Δm is this quantity measured with respect to the reference magnitude in the relevant filter.

For 44 instances no reference magnitude m_{ref} was properly defined, 36 of which in the g filter and 8 in the r filter. As such, Δm can not be properly defined in the relevant filter for these instances. However, no instance is missing the reference magnitude in both the g and r filters, thus we set the demand that either $\Delta m_g \leq 1$ or $\Delta m_r \leq 1$ must hold for the purpose of classifying a flare as extreme.

Finally, we demand a strong dust echo: $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} \geq 3$. The latter two demands are different than the ones set in van Velzen et al. (2024); here they demand $\Delta m < -1$ and that the dust echo strength is larger than the significance of the baseline RMS variability: $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} > \frac{F_{\mathrm{rms}}}{\sigma_F}$. The different choices arise from the definition of Δm being different in this work and the simplicity of demanding a dust echo flux of at least 3 times the baseline variability.

In order to quantify the strength of the burst in relation to the baseline variability in the optical, we define the optical peak to baseline variability ratio $\frac{\Delta F_g}{F_{\mathrm{rms}}}$ as a direct optical analogue to the dust echo strength. Here, ΔF_g is the flux difference between the peak F_p and the fitted baseline F_0 in the g band, and F_{rms} quantifies the variability in the baseline flux. For this we take the root mean square of the flux at all epochs $t < t_0 - 3\sigma_{\mathrm{rise}}$ and $t > t_0 + 3\tau_{\mathrm{dec}}$ given there are least 30 of such epochs, otherwise we say F_{rms} is not defined. For AGN we expect F_{rms} to be large due to large baseline flux variability expected in AGN even after some peak burst. This ratio is then expected to be smaller for AGN flares, which may result in this property being utilized to more accurately discern between AGN flares and other types of nuclear transients.

4. Results

Here, the result of the above methodology are presented, in the same order of operations as in chapter 3.

4.1 Fitting

Using the rules of a good fit construed previously, we find that 9577 sources have reliable values for the fit parameters. An example of such a fit is shown in figure 4.1. This particular flare, AT2020hle, is likely a TDE (Frederick et al., 2021) but has not been assuredly classified as such. The fit seems to agree reasonably well with the data upon visual inspection, further enforced by the reduced-χ^2 metric of 1.866. We refer to appendix A for more examples of lightcurves on different types of transients: TDE AT2019qiz (ZTF19abzrhgq, Siebert et al. 2019), AT2019aalc (ZTF19aaejtoy, e.g. van Velzen et al. 2024), supernova SN2019fck (ZTF18aaqkcso, Fremling et al. 2019) and ZTF18aahybgv. The latter is as of yet unreported and not present in the Transient Name Server (TNS).

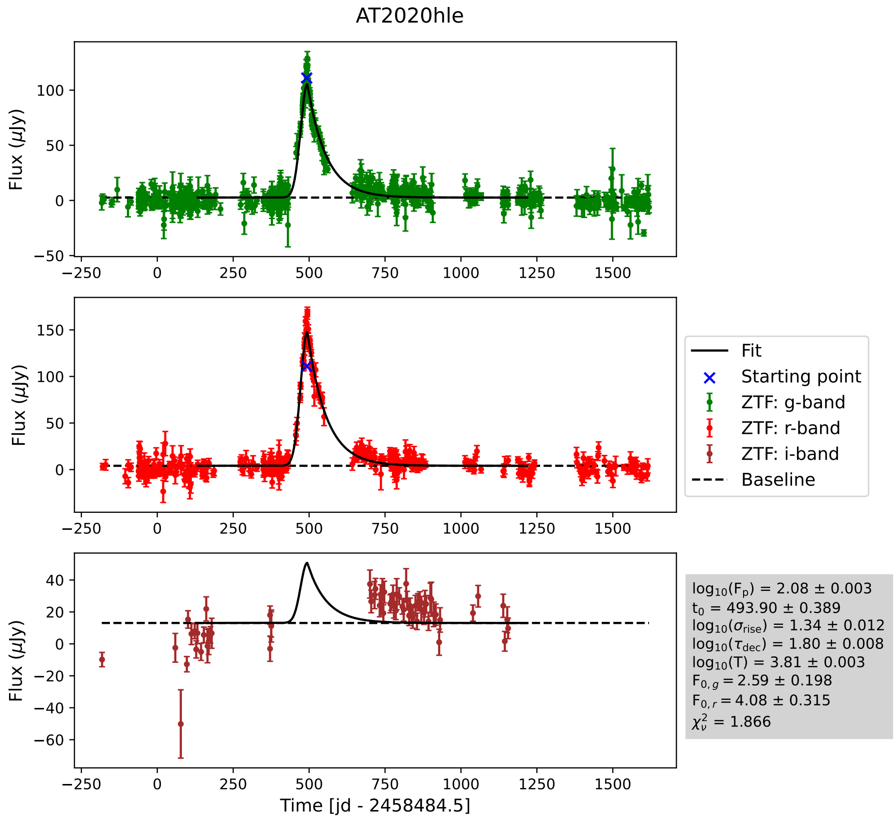

Figure 4.1: *The clean data of AT2020hle (ZTF18abjjkeo) in all 3 filters as well as the corresponding fit to this data. On the vertical axis the forced difference flux in μ Jy is shown, and on the horizontal axis the time in jd (w.r.t. 01-01-2019) is shown. The starting point corresponds to the initial guess of the peak according to the cross-correlation method. Shown are the found parameter values as well as the reduced-χ^2 metric (χ^2_ν) as a goodness of fit indicator.*

4.2 Population studies

From the sources with reliable fits, we find that 725 sources are classified as an extreme flare. Of these extreme flares, 617 (roughly 85.1%) are as of yet unclassified. The varying sample sizes for the various attributes of all different classifications of flares are shown in table 4.1.

It is of note that the TDEs in the set that were not classified as extreme were not classified as such due to the demand on the dust echo strength, with the exception of only a single TDE, AT2023cyz (ZTF20aahmtso). This TDE was not classified as an extreme flare here because of its inferred decay e-folding time $\tau_{\mathrm{dec}} = 506.44 \pm 4.87$ days - slightly outside of our set demands. The demand on the dust echo strength is clearly the biggest filter in regards to number of extreme SNe and TDEs. Conversely, the demands on the optical light curve parameters filter out the most AGN from being marked as extreme.

	AGN	SN	TDE	Unknown	All
NEOWISE detected	1558	140	48	7857	9603
Good ZTF fit	1391	132	52	8002	9577
Extreme optical	255	123	51	4114	4543
Extreme optical + dust echo	86	9	13	617	725
$\mathrm{M_{BH}}$ **estimated**	1829	46	27	960	2862
Total number	1831	169	55	11401	13456

Table 4.1: *The sizes of various sub-samples of the fitted sample of transient flares. The sub-samples shown in the rows are: the number of instances that have a detection in NEOWISE, that have reliable fit results from the optical ZTF data, that are classified as an extreme burst (for which a reliable fit result is a prerequisite, see section 3.2) and that have a BH mass estimate. The extreme bursts are further split into extreme as per the demands on the optical lightcurve parameters only, and with the added demand that that $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} \geq 3$ (denoted with "+ dust echo"). These are then subdivided into the different classifications (columns). The bottom row shows the totals of the different classifications and finally the total size of the dataset. Note that any single flare may very well have all or none of the qualities described in the rows of the table and as such the bottom row does not show the sum of the various columns.*

4.2.1 Investigating parameter space

Given these different sub-samples, we can visualize the parameter space of various combinations of two parameters. Attached to this parameter space plot are cumulative distribution functions (CDFs) of the parameter on the horizontal axis for the different classifications as well as probability density functions (PDFs) of the parameter on the vertical axis. The set demands for extreme flares are shown as dashed lines if applicable.

We might find some instances, usually of the unclassified or AGN type, at exactly the bounding box for certain parameters. This indicates that the fits of these instances got stuck at the boundaries, but were not marked as having a bad fit due to sufficiently small errors.

In figure 4.2 a cluster of SNe can clearly be seen around the point $(\log_{10}(\sigma_{\mathrm{rise}}), \log_{10}(\tau_{\mathrm{dec}})) \approx (0.9, 1.4)$. We even see the PDF of the unknown transients suddenly peak exactly where the PDF of the SNe does, indicating the presence of many SNe in that subset. There is a large variance in the timescales for AGN, making it hard to discern between them and TDEs based on this alone.

We further show that most of the strongest optical bursts ($\Delta m_g \leq 1$) - which tend to be TDEs and SNe - happen for events with $\mathrm{M_{BH}} \leq 10^8 M_\odot$ (see figure 4.4). In fact, this is exclusively true for all TDEs show in figure 4.4. This mass is the apparent threshold for strong dust echoes found by van Velzen et al. (2024) and appears to coincide with the Hills mass with the maximum mass for a visible disruption from a solar-type star (Hills 1975).

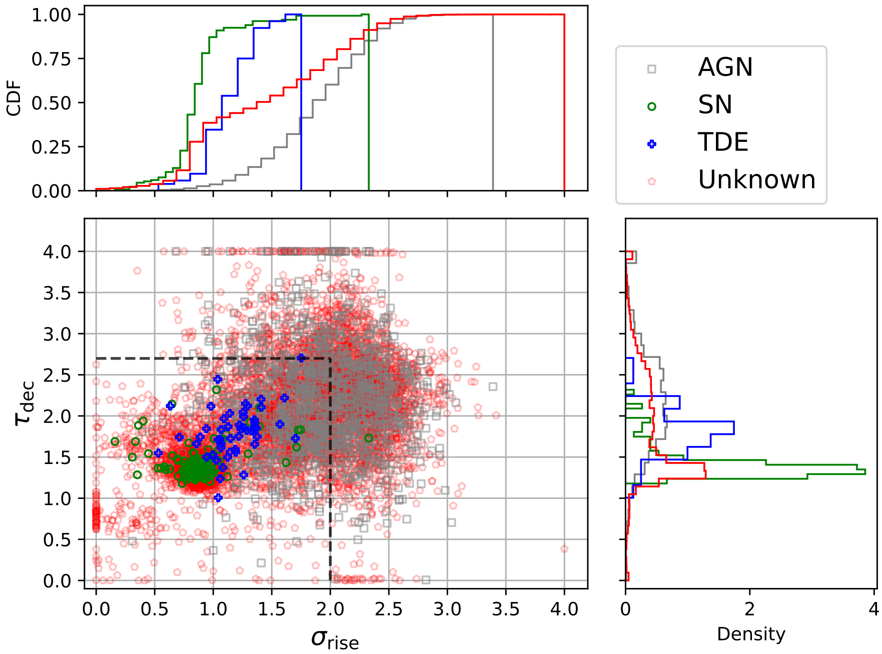

Figure 4.2: *This figure shows the e-folding rise time (horizontal axis) and e-folding decay time (vertical axis) inferred from a Gaussian rise plus exponential decay model. The dashed line shows the demands set for extreme flares on these parameters. Above, the* CDFs *of the e-folding rise time for the different classifications is shown, and to the right the* PDFs *of the e-folding decay time.*

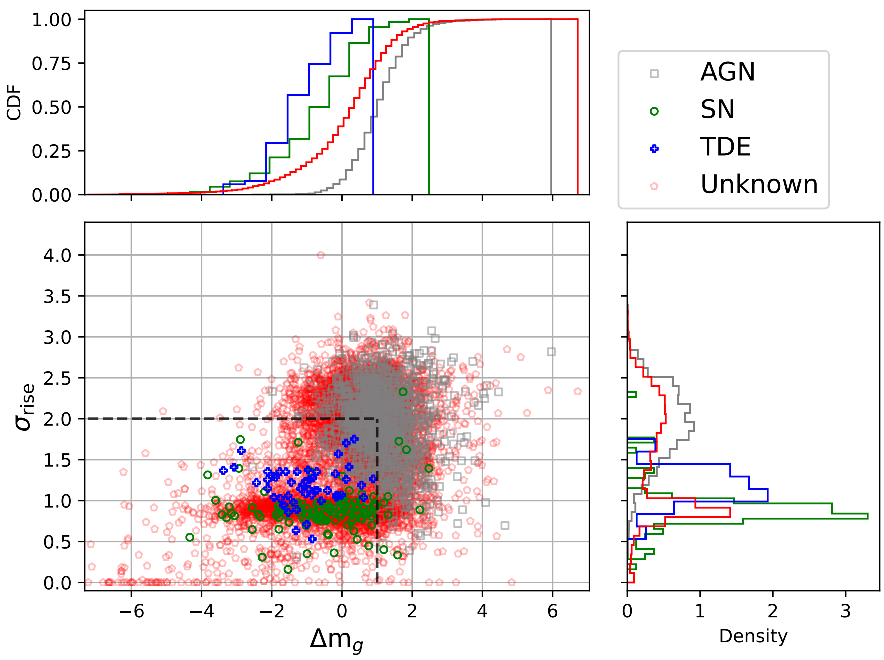

Figure 4.3: *This figure shows the peak magnitude in the g-band (horizontal axis) and the e-folding rise time (vertical axis) inferred from a Gaussian rise plus exponential decay model. The dashed line shows the demands set for extreme flares on these parameters. Above, the* CDFs *of the peak magnitudes for the different classifications is shown, and to the right the* PDFs *of the e-folding rise time.*

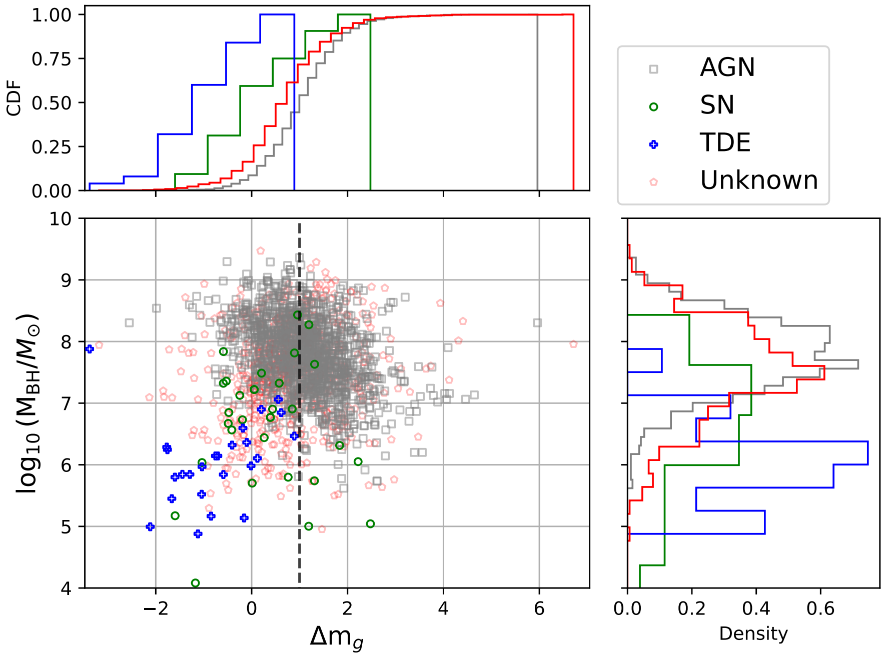

Figure 4.4: *This figure shows the peak magnitude (horizontal axis) against the black hole mass estimates (vertical axis) inferred from a Gaussian rise plus exponential decay model. The dashed line shows the demands set for extreme flares on these parameters. The CDFs of the peak magnitudes for the different classifications is shown on the top, and a PDF of the BH masses for the different classifications on the right.*

It must be noted that there is an extra TDE present outside the drawn box in figure 4.5. This is AT2020qhs (ZTF20abowque), a featureless TDE (Hammerstein et al., 2023) and the only TDE in the set with estimated $M_{BH} > 10^8 M_\odot$. This TDE is lacking observations from during the rise of its burst. As such, the fit yielded partially unreliable results for some of the parameters (see figure A.1). Therefore it is only present in figure 4.5 and not in the others.

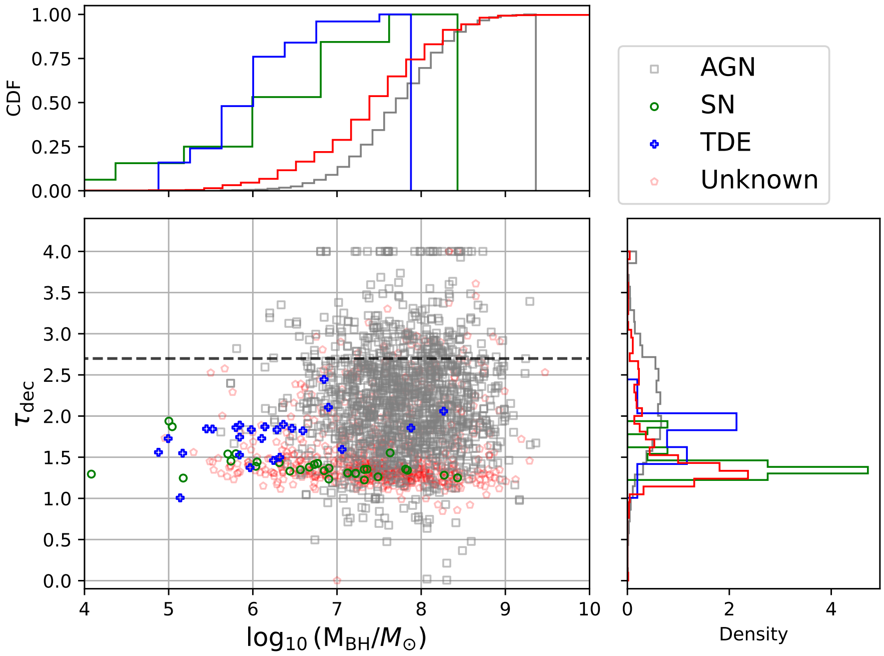

Figure 4.5: This figure shows the black hole mass estimates (horizontal axis) against the e-folding decay time (vertical axis). The CDFs of the black hole estimates for the different classifications is shown on the top, and a PDF of the decay times for the different classifications on the right.

Next, we look at how the dust echo strength holds in regard to some other parameters. First, we look at the rise-time and dust echo strength (figure 4.6). We note that instances with very high dust echo strength, i.e. $\frac{\Delta F_{IR}}{F_{rms}} \geq 10$, tend to have relatively low rise times, i.e. $\log_{10}(\sigma_{rise}) \lesssim 2$. We see the CDF of the unknown transients starts to exactly follow that of the AGN at $\frac{\Delta F_{IR}}{F_{rms}} \gtrsim 5$. Furthermore, just as in figures 4.2, 4.3 and 4.5, we see a spike in the PDF of both the unknown transients that coincides with the PDF of SNe. These spikes suggest a clear presence of SNe in the unknown sub-sample.

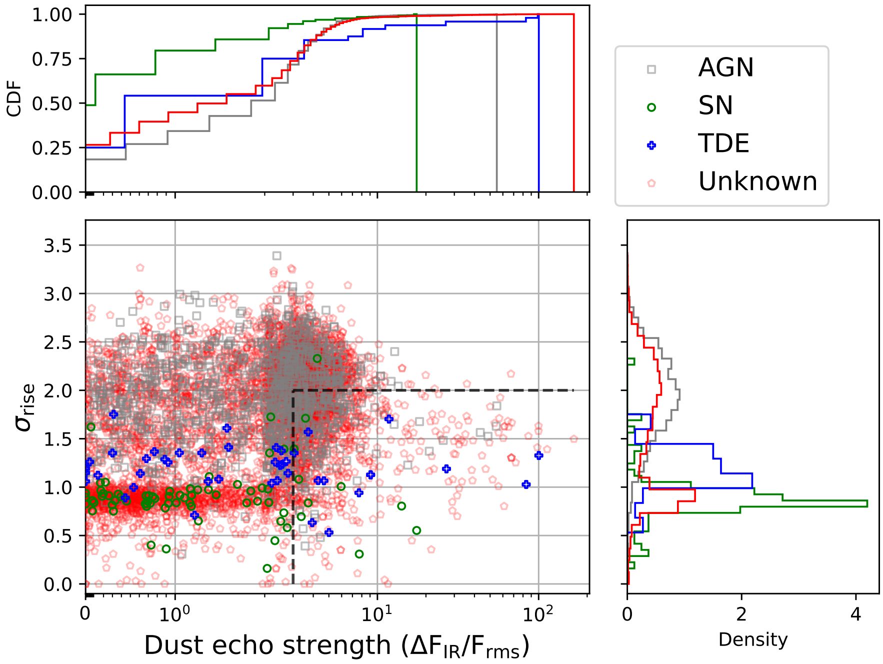

Figure 4.6: This figure shows the infrared dust echo strength (horizontal axis) against the rise e-folding time (vertical axis). The dust echo strength is plotted on a symmetrical log-scale, which allows for negative values in a log-scale and is linear around 0. The CDFs *of the dust echo strength for the different classifications is shown on the top, and the* PDFs *of the rise time on the right.*

We find here, analogous to van Velzen et al. (2024), that strong dust echoes ($\frac{\Delta F_{IR}}{F_{rms}} \geq 10$) happen almost exclusively at $M_{BH} \lesssim 10^8 M_\odot$ (see figures 4.7 and 4.8). Again, this appears to coincide with the Hills mass with the maximum mass for a visible disruption from a solar-type star (Hills 1975). The two sources in the set for which this does not hold are ZTF18accdoui and AT2019hks (ZTF19aarrrup), the first being classified as an AGN but has not yet been reported to TNS and the latter is marked as unclassified in our dataset. In figure 4.7 certain flares are marked with their IAU designation, including two candidate TDEs AT2020hle (Frederick et al., 2021) and AT2019fdr (Reusch et al., 2022). The latter, together with the flares AT2019dsg (TDE) and AT2019aalc (AGN) were found to have neutrino coincidences (van Velzen et al., 2024), these two are moreover marked. Finally some other transients with strong, intermediate or weak dust echoes are marked.

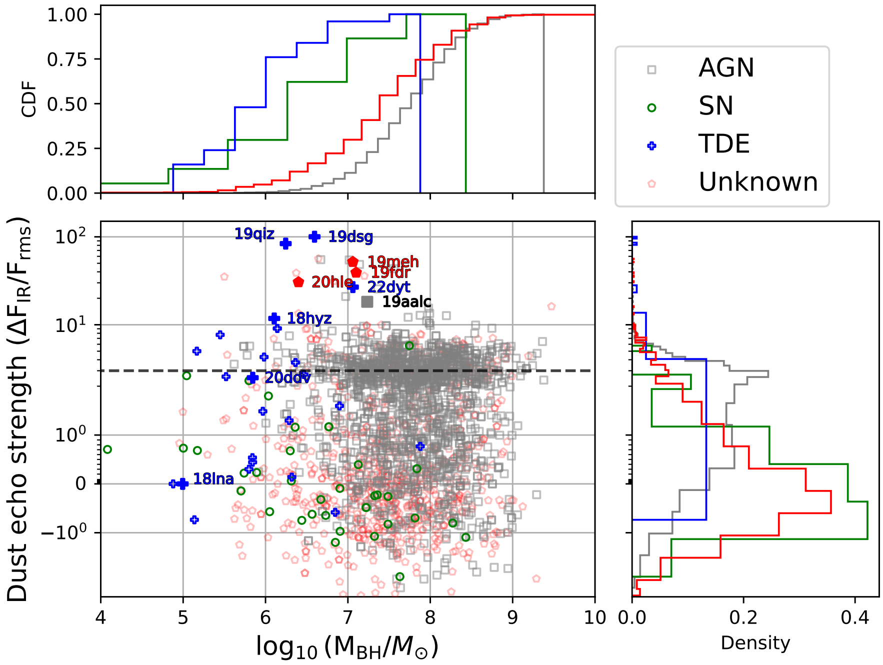

Figure 4.7: This figure shows the black hole mass estimates (horizontal axis) against the infrared dust echo strength (vertical axis). The dust echo strength is plotted on a symmetrical log-scale, which allows for negative values in a log-scale and is linear around 0. The CDFs of the BH masses for the different classifications is shown on the top, and a PDF of the dust echo strengths on the right. Certain interesting flares are marked explicitly. See figure 4.8 for a a figure zoomed in on the dashed box.

These same transients are marked in figure 4.9 which shows M_{BH} against the optical peak to baseline variability ratio $\frac{\Delta F_g}{F_{rms}}$. We find no classified AGN with $\frac{\Delta F_g}{F_{rms}} \geq 100$ in our sample, only TDEs and unclassified transients. Moreover, we find almost exclusively TDEs and unclassified transients with $\frac{\Delta F_g}{F_{rms}} \geq 10$ and $M_{BH} \leq 10^7 M_{\odot}$. However, the population in this region is in large part unclassified, so no strict conclusions may be drawn from this observation as of yet. This may motivate for further inspection and classification of such transients to possibly establish a pattern and subsequently increase the possibilities in discerning between AGN activity and other types of nuclear transients using $\frac{\Delta F_g}{F_{rms}}$.

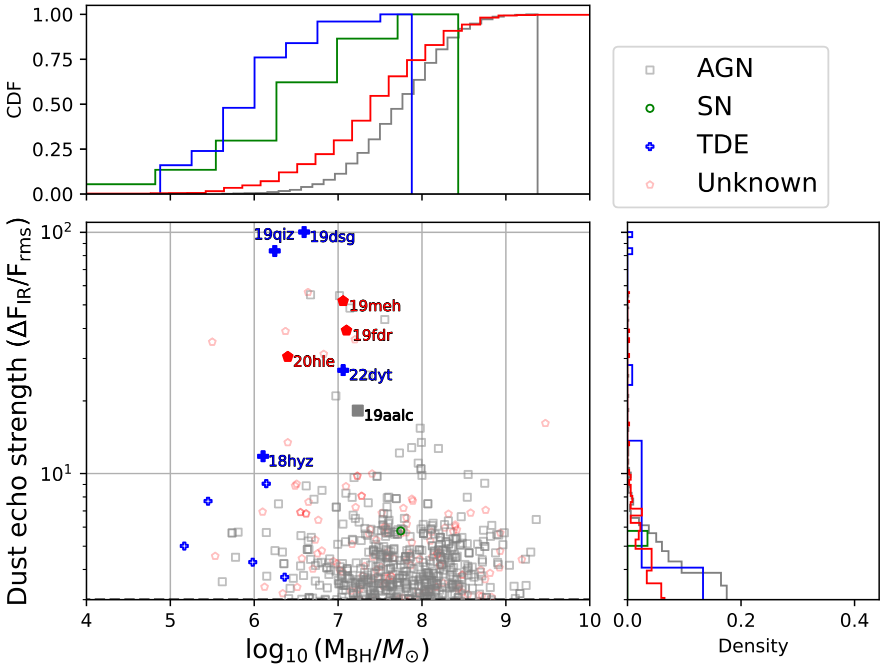

Figure 4.8: *Zoomed in version of figure 4.7, specifically on the region* $\frac{\Delta F_{IR}}{F_{rms}} \geq 3$ *and* $10^5 \leq M_{BH}/M_\odot \leq 10^{10}$.

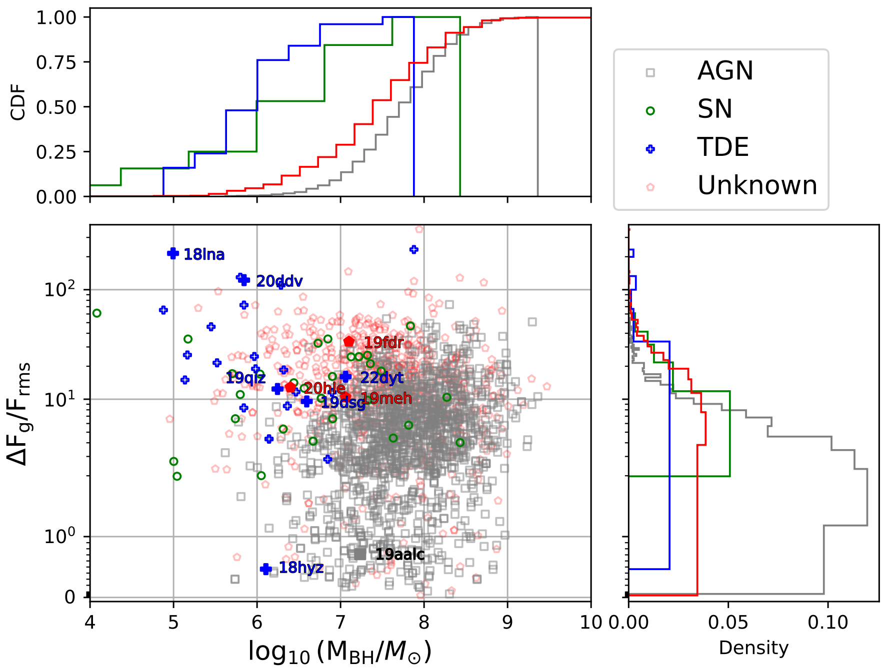

Figure 4.9: *This figure shows the black hole mass estimates (horizontal axis) against the optical peak to baseline variability ratio (vertical axis). The CDFs for the different classifications of the BH masses for the different classifications is shown on the top, and a PDF of the peak to baseline variability ratio is shown on the right. Certain interesting flares are marked explicitly.*

Analogous to figure 3 from van Velzen et al. (2024), figure 4.10 shows the infrared luminosity plotted against the dust echo strength. We see a clear clustering of AGN around $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} = 3$ and $L_{\mathrm{IR}} = 10^{43}$. Notably, SNe seem to have relatively low infrared luminosities and dust echo strengths compared to the other classifications.

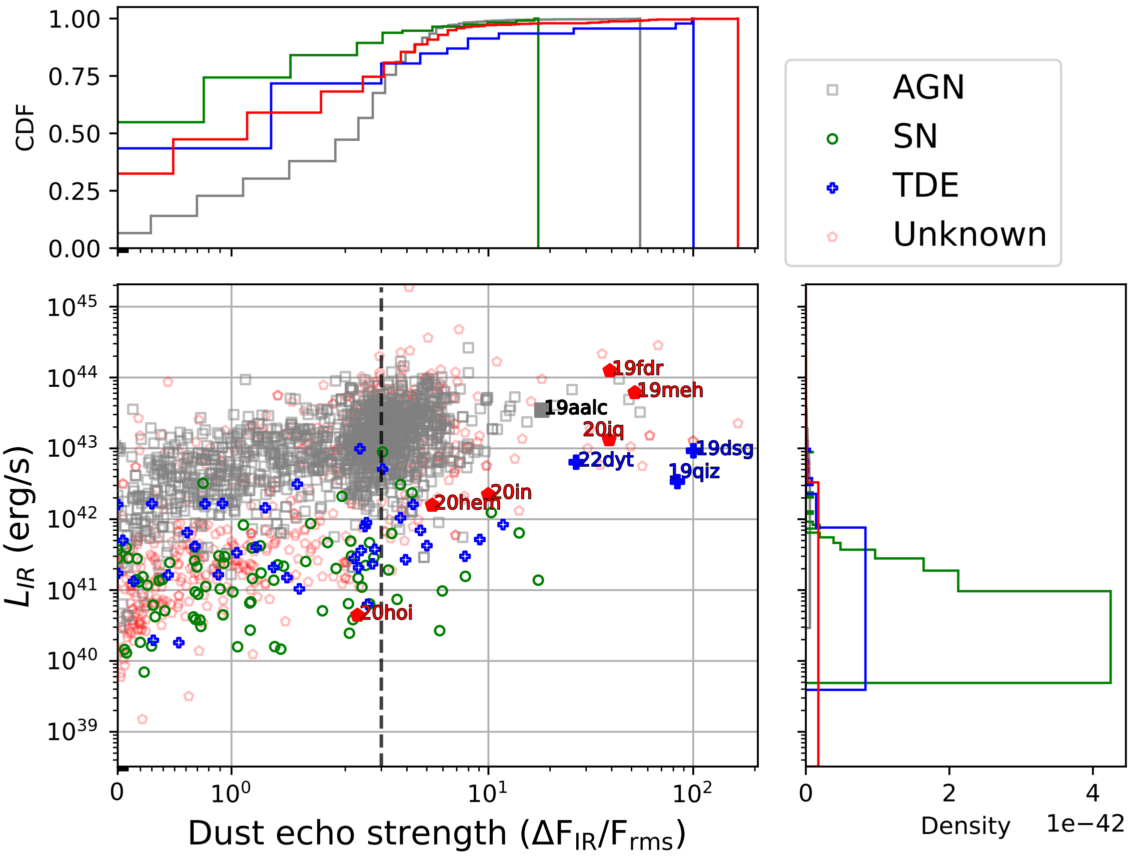

Figure 4.10: This figure shows the infrared dust echo strength (horizontal axis) against the measured infrared red luminosity after the peak in erg (vertical axis). Only sources with non-negative dust echo strength are included in this figure. The dashed line at $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} = 3$ *shows the threshold for extreme flares set in this work. The CDFs for the different classifications of the dust echo strengths for the different classifications is shown on the top, and a PDF of infrared luminosity is shown on the right. Certain interesting flares are marked explicitly.*

Given these various parameters, it is of interest to investigate if any correlations exist between them. For this reason we perform various Kendall τ tests (Kendall, 1938) on pairs of parameters separately for all TDEs and extreme AGN, the p-values of which may be seen in table 4.2 and 4.3 respectively. We use a confidence interval of $p < 0.05$ for the purpose of rejecting the null hypothesis that a pair of parameters is uncorrelated. With this, for the TDE sub-population, we report relevant correlations between σ_{rise} and τ_{dec}, σ_{rise} and M_{BH}, τ_{dec} and M_{BH}, Δm_g and M_{BH} and $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}}$ and $\frac{\Delta F_{\mathrm{g}}}{F_{\mathrm{rms}}}$. Interestingly, we find no such significant correlations for extreme AGN, instead reporting only a relevant correlation between M_{BH} and $\frac{\Delta F_{\mathrm{g}}}{F_{\mathrm{rms}}}$ and $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}}$ and Δm_r, which were not (significantly) present for TDEs. Other significant correlations present in both cases are entirely expected as a result of definition; Δm_g and Δm_r and Δm_g and $\frac{\Delta F_\mathrm{g}}{F_\mathrm{rms}}$ are found to correlate for both TDEs and AGN. Subsequently, since a very strong correlation is expected between Δm_g and Δm_r, it is unsurprising to find a correlation between Δm_r and $\frac{\Delta F_\mathrm{g}}{F_\mathrm{rms}}$. It may be noted that these expected correlations are weaker for AGN than for TDEs - most notably Δm_r and $\frac{\Delta F_\mathrm{g}}{F_\mathrm{rms}}$ - though we cannot discount them as insignificant.

Since we are investigating 20 different pairs of parameters we must take into account the look-elsewhere effect (with reasoning similar to Hammerstein et al. (2023)). Of the 20 unique pairs of parameters, we cannot claim Δm_g and Δm_r as well as Δm_g and $\frac{\Delta F_\mathrm{g}}{F_\mathrm{rms}}$ to be independent due to parameter definition. This leaves 18 independent pairs of parameters. We expect to find a p-value of 0.05 once every 20 tests or $\approx 1/18$ times. For the TDEs we find 6 significant independent correlations, the probability of this happening through random chance is $1.6 \cdot 10^{-4}$ (via a binomial distribution with a 0.05 percent chance at success). For AGN we find 3 independent correlations the probability of which is 0.04726 - significantly higher but still low enough to consider them unaffected by the look-elsewhere effect.

TDEs

	σ_{rise}	τ_{dec}	Δm_g	Δm_r	M_{BH}	$\Delta F_{\text{IR}}/F_{\text{rms}}$
σ_{rise}	-	$2.86{\cdot}10^{-3}$ (52)	$7.64{\cdot}10^{-1}$ (52)	$5.38{\cdot}10^{-1}$ (52)	$4.71{\cdot}10^{-2}$ (25)	$8.80{\cdot}10^{-1}$ (48)
τ_{dec}	$2.86{\cdot}10^{-3}$ (52)	-	$5.18{\cdot}10^{-1}$ (52)	$6.93{\cdot}10^{-1}$ (52)	$2.19{\cdot}10^{-2}$ (26)	$5.52{\cdot}10^{-1}$ (48)
Δm_g	$7.64{\cdot}10^{-1}$ (52)	$5.18{\cdot}10^{-1}$ (52)	-	$6.93{\cdot}10^{-13}$ (52)	$2.65{\cdot}10^{-2}$ (25)	$3.69{\cdot}10^{-1}$ (48)
Δm_r	$5.38{\cdot}10^{-1}$ (52)	$6.93{\cdot}10^{-1}$ (52)	$6.93{\cdot}10^{-13}$ (52)	-	$7.21{\cdot}10^{-2}$ (25)	$7.56{\cdot}10^{-1}$ (48)
M_{BH}	$4.71{\cdot}10^{-2}$ (25)	$2.19{\cdot}10^{-2}$ (26)	$2.65{\cdot}10^{-2}$ (25)	$7.21{\cdot}10^{-2}$ (25)	-	$1.02{\cdot}10^{-1}$ (25)
$\Delta F_{\text{IR}}/F_{\text{rms}}$	$8.80{\cdot}10^{-1}$ (48)	$5.52{\cdot}10^{-1}$ (48)	$3.69{\cdot}10^{-1}$ (48)	$7.56{\cdot}10^{-1}$ (48)	$1.02{\cdot}10^{-1}$ (25)	-
$\Delta F_g/F_{\text{rms}}$	$4.30{\cdot}10^{-1}$ (52)	$5.28{\cdot}10^{-1}$ (52)	$1.23{\cdot}10^{-12}$ (52)	$8.56{\cdot}10^{-6}$ (52)	$6.50{\cdot}10^{-2}$ (25)	$3.67{\cdot}10^{-2}$ (48)

Table 4.2: *This table shows the p-values from Kendall's τ test between all pairs of parameters for the TDE sub-population. It is of note that σ_{rise}, τ_{dec} and M_{BH} were evaluated in log-space for this test. The number of transients accounted in each test is denoted between brackets since this may vary between pairs of parameters (see table 4.1). A confidence interval of $\alpha = 0.05$ was chosen for the purpose of rejecting the null hypothesis that two parameters do not correlate.*

Extreme AGN

	σ_{rise}	τ_{dec}	Δm_g	Δm_r	M_{BH}	$\Delta F_{\text{IR}}/F_{\text{rms}}$
σ_{rise}	-	$7.80{\cdot}10^{-1}$ (86)	$2.43{\cdot}10^{-1}$ (86)	$8.61{\cdot}10^{-1}$ (86)	$1.90{\cdot}10^{-1}$ (84)	$5.58{\cdot}10^{-1}$ (86)
τ_{dec}	$7.80{\cdot}10^{-1}$ (86)	-	$7.29{\cdot}10^{-1}$ (86)	$7.34{\cdot}10^{-1}$ (86)	$3.48{\cdot}10^{-1}$ (85)	$1.90{\cdot}10^{-1}$ (86)
Δm_g	$2.43{\cdot}10^{-1}$ (86)	$7.29{\cdot}10^{-1}$ (86)	-	$5.91{\cdot}10^{-11}$ (86)	$7.66{\cdot}10^{-1}$ (84)	$1.22{\cdot}10^{-1}$ (86)
Δm_r	$8.61{\cdot}10^{-1}$ (86)	$7.34{\cdot}10^{-1}$ (86)	$5.91{\cdot}10^{-11}$ (86)	-	$6.68{\cdot}10^{-1}$ (84)	$6.06{\cdot}10^{-2}$ (86)
M_{BH}	$1.90{\cdot}10^{-1}$ (84)	$3.48{\cdot}10^{-1}$ (85)	$7.66{\cdot}10^{-1}$ (84)	$6.68{\cdot}10^{-1}$ (84)	-	$2.33{\cdot}10^{-1}$ (88)
$\Delta F_{\text{IR}}/F_{\text{rms}}$	$5.58{\cdot}10^{-1}$ (86)	$1.90{\cdot}10^{-1}$ (86)	$1.22{\cdot}10^{-1}$ (86)	$6.06{\cdot}10^{-2}$ (86)	$2.33{\cdot}10^{-1}$ (88)	-
$\Delta F_{\text{g}}/F_{\text{rms}}$	$8.37{\cdot}10^{-1}$ (86)	$7.18{\cdot}10^{-1}$ (86)	$4.16{\cdot}10^{-7}$ (86)	$2.70{\cdot}10^{-2}$ (86)	$4.02{\cdot}10^{-2}$ (84)	$1.35{\cdot}10^{-1}$ (86)

Table 4.3: *This table shows the p-values from Kendall's τ test between all pairs of parameters for the extreme AGN sub-population (see section 3.2 for how extreme is defined). It is of note that σ_{rise}, τ_{dec} and M_{BH} were evaluated in log-space for this test. The number of transients accounted in each test is denoted between brackets since this may vary between pairs of parameters (see table 4.1). A confidence interval of $\alpha = 0.05$ was chosen for the purpose of rejecting the null hypothesis that two parameters do not correlate.*

4.2.2 Extreme AGN vs. other AGN

With our definition of extreme flares we implicitly create two sub-populations of AGN flares; extreme and other. We perform a two-sample Kolmorgorov-Smirnoff (KS) test on the black hole mass estimates of the extreme and non-extreme sub-populations of AGN to see if these sub-population samples are likely to be drawn from the same underlying distribution. In order to make the comparison fair, we set the additional demand that the non-extreme AGN do need to have the potential for a dust echo, e.g. we demand for these $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} > 0$.

With these definitions, there are 84 extreme AGN and 1127 other AGN with a $\mathrm{M_{BH}}$ estimate present in the sample. We find a p-value of 0.00907 on the hypothesis that the BH masses of extreme AGN and other AGN are drawn from the same distribution. As such, we can reject the null hypothesis at a level of confidence of $\alpha = 0.010$. We report that the extreme AGN in the dataset used for this work generally have higher $\mathrm{M_{BH}}$ than other AGN and that it is likely the BH masses are drawn from different underlying distributions.

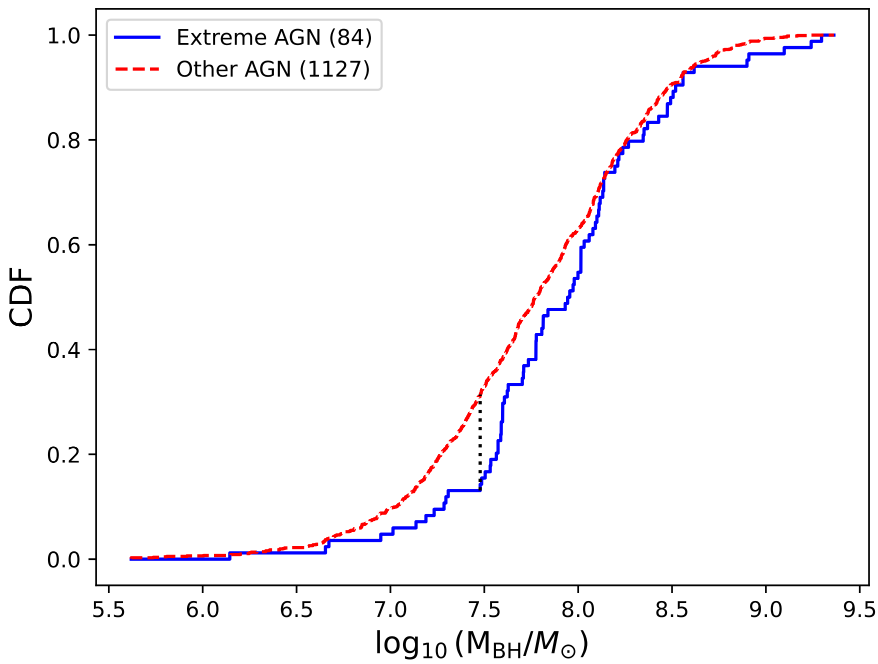

Figure 4.11: *CDFs of the BH mass estimates of AGN with extreme flares and regular variability given that they have non-negative dust echo strength (denoted as other AGN). The black line shows the maximal distance as per a KS test which is significant enough to reject the null hypothesis that these two sub-samples drawn from the same population with a p-value of* 0.00907. *The size of each sub-sample is shown in the legend.*

A noteworthy caveat to this result however is the difference in redshift distributions of the two different sub-populations (see figure 4.12). We see that they tend to follow different distributions, with the probability of finding a non-extreme AGN at relatively higher redshift dropping off strongly. This is not the case for the extreme AGN. This introduces a redshift related bias to our analysis which must be taken into account when

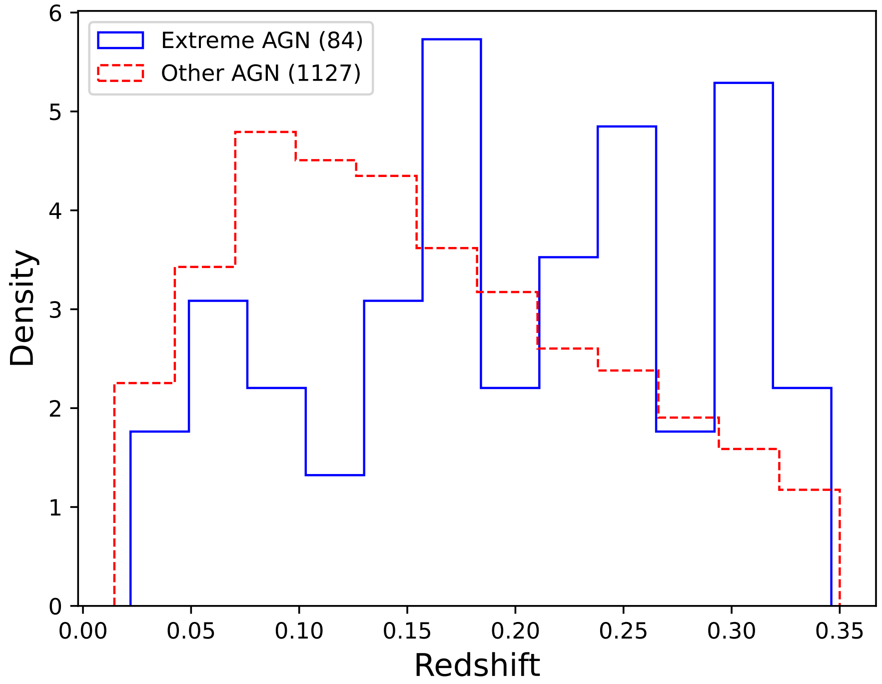

Figure 4.12: *Redshift distributions of AGN with extreme flares and other AGN, with the other AGN strictly having* $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} > 0$.

interpreting the results.

4.2.3 Unclassified transients with strong dust echoes

In table 4.4 the name, associated BH mass estimate, dust echo strength and optical peak to baseline variability ratio of 50 unclassified transients with the strongest dust echo are shown. These also comply with the demands that they have a known M_{BH} estimate and have been marked as an extreme flare as per the definition set in this work. Given that these flares have the most substantial dust echoes of the unknowns the dataset, it is of great interest to perform more research into classifying them since there we expect a greater chance finding a TDE or an extreme AGN flare in this subset than drawing at random from the greater total set of unknown flares. About half of the shown flares have not been reported to the TNS as discovered at the time of writing. Another 6 have been reported as both discovered and classified. Other instances have been put forward as candidates of a certain type of transient and have all been marked as such in the table.

ID	M_{BH} ($\log_{10} M_\odot$)	$\Delta F_{IR}/F_{rms}$	$\Delta F_g/F_{rms}$	Classification	Ref
AT2019meh	7.060	51.86	10.41	SLSN(*)	1
AT2019fdr	7.100	39.18	33.58	TDE?	2,3,4
AT2020iq	6.373	38.88	18.59	SN IIn (*)	5,6
AT2020hle	6.400	30.50	12.84	TDE?	7
AT2019pev	6.400	13.45	17.93	AGN?	8
SN2020in	7.406	9.982	10.75	SN IIn	6,9,10
ZTF20abbbyhe	6.506	9.055	20.74	-	-
AT2020afab	6.487	8.895	63.49	-	-
AT2020afap	6.647	7.618	13.87	TDE?	11
ZTF18aahybgv	6.551	6.905	85.99	-	-
AT2021tmz	7.521	6.857	14.10	-	-
ZTF20aagyaug	8.316	6.802	15.40	-	-
AT2021ytz (ZTF20aaxwxgq)	8.396	6.134	25.76	-	-
ZTF19abaezlx	7.548	6.024	15.50	-	-
AT2021ytz (ZTF21abwoxbv)	8.396	6.020	24.43	-	-
AT2019lod	8.416	5.854	6.149	-	-
SN2020hem	6.729	5.314	8.349	SN IIn	6,12
ZTF18acdwazb	7.166	5.092	14.38	-	-
AT2021qah	6.806	4.787	10.78	-	-
ZTF18abjyjik	8.931	4.768	10.26	-	-
AT2019cjn	8.199	4.740	4.278	-	-
AT2020afex	8.183	4.670	74.81	-	-
ZTF19aadfymo	7.855	4.640	6.363	-	-
AT2018lnh	8.183	4.628	8.937	-	-
AT2019nne	5.535	4.545	96.59	-	-
AT2019cuk	7.969	4.420	2.473	AGN?(**)	13,14
ZTF20abfxyvk	6.943	4.327	20.39	-	-
ZTF19aatbrtp	6.984	4.240	12.53	-	-
ZTF19aalznaj	8.215	4.160	4.301	-	-
ZTF18aavvdjt	8.215	4.160	4.288	-	-
ZTF20aarqhfi	7.869	4.159	91.85		
ZTF18abtrgsu	7.514	4.099	18.22	-	-
AT2019hfb (ZTF19adcgabk)	7.634	3.756	5.876	-	-
AT2019hfb (ZTF18aahxdfs)	7.634	3.756	6.028	-	-
AT2019acd	6.123	3.387	25.37	-	-
AT2020xtd	8.441	3.382	4.76	-	-

Continued on next page

ID	M_{BH} ($\log_{10} M_\odot$)	$\Delta F_{IR}/F_{rms}$	$\Delta F_g/F_{rms}$	Classification	Ref
ZTF19aarsuvf	6.512	3.379	16.54	-	-
ZTF19aaruusi	8.429	3.328	128.4	-	-
ZTF18acvghyv	8.079	3.256	8.592	-	-
ZTF20achosql	8.549	3.157	36.80	-	-
AT2023cl	8.681	3.125	6.054	-	-
ZTF18aalmtum	8.403	2.940	14.19	-	-
ZTF19abirtdv	8.511	2.883	23.00	-	-
AT2018igc	8.703	2.338	5.309	-	-
AT2021lfg	5.998	2.336	12.99	-	-
SN2020hoi	5.816	2.293	6.637	SN Ib	15
ZTF18aawwpbh	6.838	2.289	15.87	-	-
AT2023yzr	8.464	2.229	7.108	-	-
ZTF19aagroms	7.474	2.157	16.38	-	-
AT2019ywa	5.827	2.091	24.85	-	-

Table 4.4: *The 50 unclassified transient flares with the highest dust echo strength from the subset of flares that have been marked as extreme events and have a known black hole mass estimate. The first column shows the name of the transient. The next three columns give the black hole mass estimate, the dust echo strength and optical peak to baseline variability ratio respectively. Some instances are in fact classified despite them being marked as unclassified in the data used for this work. The classifications are given in the final columns (incl. references). Any inconclusive classification of a transient is marked with a question mark. It may be noted that AT2020ytz and AT2019hde are present twice, as they had two alerts in the ZTF survey, but constitute one source each. The inferred parameters from the Gaussian rise plus exponential decay lightcurve model may differ due to slight differences in ZTF data between alerts.*
References: 1. Nicholl et al. (2019a); 2.Jiang et al. (2023); 3. van Velzen et al. (2024); 4. Reusch et al. (2022); 5. Dahiwale & Fremling (2020a); 6. Strotjohann et al. (2021) 7. Frederick et al. (2021); 8. Yu et al. (2022); 9. Graham et al. (2020a); 10. Graham et al. (2020b) ;11.Gomez et al. (2023); 12. Dahiwale & Fremling (2020b); 13. Ward et al. (2024); 14 Hoshi et al. (2024); 15. Perley et al. (2020).
(*): This source has been reported to TNS as a type of SN, however we show in this work that this is likely untrue (see figure 4.10 and section 5.2).
(**): Binary supermassive BH merger candidate (Jiang et al., 2022).

5. Discussion

We discuss here the methodology and possible improvements thereof, as well as the implications of our results and the conclusions we can draw from them. We further propose some ideas for research that may follow up with the used data, methods and results presented in this work.

5.1 Reflection on methodology

We have shown various population trends resulting from a Gaussian rise plus exponential decay model fit on a large number of transient flares. A large subset of the transients that had reliable clean data was not able to be fitted or otherwise yielded unreliable fit results (e.g. with huge error margins) and were later filtered out. This left only 9577 out of an original 14258. This is in large part due to instances that were too noisy to properly fit the model to or had variability on timescales that the model was unable to properly capture.

The latter is due to the fact that the fitted model implicitly assumes a single burst to transpire over the duration of the measurement, making it a bad fit to instances with secondary bursts. An example of a transient with a second burst is AT2019aalc (ZTF19aaejtoy), the first flare of which has been found to coincide with a PeV-scale neutrino (van Velzen et al. (2024)). This transient recently experienced another, stronger, burst (see figure A.2). This motivates for a method that can detect, from the lightcurve, where a second burst starts. A specialized implementation of Bayesian Blocks (Scargle et al., 2013) might prove useful in this regard. Some attempts at this have been made in the past (see e.g. Worpel & Schwope 2015), but are not directly applicable in the context of this work. A more adaptable model, granted that it is sophisticated enough, might also be able to account for secondary peaks inherently (see e.g. the SN lightcurve model described in Müller-Bravo et al. (2021), which does not explicitly account for secondary peaks but has cases where a secondary peak is successfully fit).

With an algorithm such as Bayesian Blocks, a single burst Gaussian rise plus exponential decay model could potentially be fit to only the epochs that do not constitute a different burst. A simple test of the `astropy`[1] implementation of the Bayesian Blocks algorithm

[1] https://docs.astropy.org/en/latest/api/astropy.stats.bayesian_blocks.html

yielded no fruitful results upon testing on AT2019aalc - the algorithm finding 79 blocks whereas by eye we would expect only 2 - hence the reason why it is not implemented to differentiate between flares in this work.

Another possibility of discerning secondary bursts would be a pseudo-manual method of fitting any well-defined single burst model to the entire range of epochs available instead of only in $[t_{0,cc} - 365, t_{0,cc} + 730]$ and looking at the goodness of fit as quantified by χ^2_ν. If χ^2_ν increases significantly (a threshold for which would need to be carefully defined) at least one secondary burst is expected to exist. The instances for which this is true could then be manually investigated to confirm the existence of such secondary flares. Lots of AGN are variable on timescales shorted than the 3 year window in which we currently fit (see e.g. the lightcurve of AGN AT2019aalc in figure A.2) which shows slight variability on top of the fitted model, which is reflected in its high χ^2_ν value.

In order to improve upon which epochs are considered in the fitting process, future work may benefit from the introduction of a dynamic time window. As it stands, the window is based on the initial guess of the peak time $t_{0,cc}$ and is not dynamic with the best fit value for t_0, which might give slight inaccuracies and cause the model to not see some important epochs at the tails of the burst depending on how inaccurate $t_{0,cc}$ is. On average, we find an absolute offset of 32.06 ± 0.544 days and a (standard deviation of 53.25 days) between the fitted value t_0 and the initial guess from the cross correlation method $t_{0,cc}$, using all instances deemed as good fits. Therefore the offset from the inital guess and the subsequent epochs which are considered for the fit are usually relatively properly defined but this may not be the case for some instances with a high offset.

The relatively low average offset, however, solidifies the cross-correlation method as being reliable enough for an initial guess of the peak. Some improvements upon the method might even make it usable as a tool in and of itself to automatically deduce the most likely time of a burst peak. Possible improvements include: a decrease in the 2 day time-step used in this work, using non-linear interpolation for the data and/or using a Gaussian rise plus exponential decay model instead of a fully Gaussian one for the purpose of calculating the cross-correlation. More investigation is needed to confirm if these suggestions would actually improve the results.

The Gaussian-rise plus exponential decay model is the one used for fitting the optical light curve to in this work. One benefit of this model is the simplicity. Moreover, we use it since we are interested specifically in short-lived transients for which the model tends to hold well. However, TDEs with observations longer than 100 days after the peak of the burst show deviations from an exponential decay model (van Velzen et al., 2021b). We see that the minority of the TDE sub-population used in this work have a decay e-folding time of more than 100 days as can be seen in e.g. figure 4.2, with most TDE having an inferred decay e-folding time of $\log_{10}(\tau_{\mathrm{dec}}) \leq 2$. This suggests the model works relatively well for most of the TDEs in the dataset. Even so, accuracy could be improved by taking the known TDE in the dataset (re-)fit them to a Gaussian rise plus power-law decay model as described in equation (3) of van Velzen et al. (2021b). This is computationally feasible due to the small sub-set of transients that are known TDE. If this choice is made, however, it is important to consider the difference in model parameters (e.g. characteristic decay time) when comparing between different classifications of transients.

5.2 Implications of results

5.2.1 TDEs as a mechanism for extreme AGN flares

Given the results of our anaylsis on the populations of TDEs and AGN in our dataset, we can comment on the possibility of the mechanism of tidal disruption events being the driver behind extreme AGN flares.

Firstly, through performing a KS test we reject the null hypothesis that the two sub-populations of AGN - extreme and non-extreme with non-negative dust echo strength - are drawn from the same overarching distribution with a p-value of 0.00907. Our results show that extreme AGN are more likely to have a higher mass than non-extreme AGN (see figure 4.11) - which does not agree with the analysis of van Velzen et al. (2024), who find non-extreme flares to occur at higher mass. It might be due to the relatively small sample size of extreme AGN flares due to our definition of what marks a flare as extreme, but it was found that loosening this definition (e.g. by excluding the demand that $\frac{\Delta F_{\mathrm{IR}}}{F_{\mathrm{rms}}} \geq 3$ increasing the sub-sample to 255 extreme flares as show in table 4.1), the result only became more conclusive with a p-value of $2.30 \cdot 10^{-10}$ and extreme AGN still tending to higher M_{BH}. Whichever way, the result stands - in agreement with e.g. van Velzen et al. (2024) and Frederick et al. (2021) - that extreme AGN flares are likely not caused solely by regular AGN activity.

Importantly, however, it must be noted that the redshift distribution of these two sub-populations of AGN are not the same as can be seen in figure 4.12. Future follow-up on this result may improve upon our findings by using for instance (rejection) sampling methods to solve this inconsistency between the two redshift distributions so that the comparison is more fair and the result possibly more conclusive.

Given our finding that it is likely that extreme AGN flares are caused by other mechanisms - which is corroborated by many prior investigations (see e.g. Frederick et al. (2021) and references therein) - it is of interest to find what is in fact the underlying mechanism. We look here at the possibility of TDEs being that driving mechanism. In the populations of known extreme AGN flares and TDEs, we find that unique correlations exist between parameters for both both of them - that is to say no two parameters, except for those expected by definition, correlate significantly for both TDEs and AGN (see tables 4.2 and 4.3). This finding implies that TDEs are not the obvious cause of all the extreme AGN flares in our sample. Indeed, upon performing a KS test in an analogous manner as done for the two sub-populations of AGN but now for the extreme AGN flares and TDE we confidently reject the null hypothesis that these are drawn from the same distribution with a p-value of $1.637 \cdot 10^{-17}$ (see figure A.8 in appendix A for the CDFs). It must be noted, however, that the sample sizes are quite small here and a similar caveat as per the redshift distributions hold (see figure A.9). Still, this does not rule out the possibility that some TDEs may be the cause of extreme AGN flares. For example Merloni et al. (2015) presented the possibility of what was first classified as a result of a changing-look AGN to be chalked up to a TDE (see moreover van Velzen et al. (2021b) and references given there). Our results do not rule out the possibility of the presence of TDEs being the mechanism behind some extreme AGN flares, though we do conclude that it does not hold for the investigated population of extreme AGN flares as a whole.

Investigations of TDEs causing unexplained extreme AGN flares is an ongoing field of research. Hydrodynamical simulations of TDEs in the accretion disc are important supplements in understanding the effects and resulting emitted light of such event (see e.g. Chan et al. 2019, 2020). Such supplementary investigations and implications are important and can prove useful in future follow-up research which may hope to pick out extreme AGN flares that were in fact caused by TDEs.

Besides just finding the mechanism behind extreme AGN flares however, the results of this work may be useful the other way around as well. It is of great import to be able to distinguish AGN activity from TDEs for the purpose of creating TDE candidate samples. We refer to Zabludoff et al. (2021) for an in-depth oversight into the process of filtering out other transients from such a candidate sample.

5.2.2 Trends in the populations

Besides just the implications and relation between extreme AGN flares and TDEs, our results hold more information on trends and parameter correlations in (sub-)populations as a whole as can be seen in tables 4.2 and 4.3.

For instance, we report a strong correlation between the rise timescale and decay timescale for the TDEs in our set, in direct conflict with the findings of van Velzen et al. (2021b) but in agreement with those of Hammerstein et al. (2023). This suggests that such a correlation may very well exist and might be an important result for TDEs, but given the discrepancy in literature and low sample size a strict conclusion about the existence of such a relation cannot be drawn yet.

We further report, for TDEs, a significant correlation to exist between the peak magnitude and black hole mass in the g band. However, such a correlation does not appear in the r band (with a confidence interval of $p < 0.05$). This is an unexpected result due to the high level of correlation between Δm_g and Δm_r. This suggests that the associated BH mass has a more significant effect on emitted light in the g band than in the r band for TDEs or that the black body temperature correction as described in equation (3.2) might not be sufficiently accurate in accounting for photometric differences for TDEs. However, this result might very well be an artefact of low sample size or skewed redshift distribution. Hammerstein et al. (2023) report a correlation to exist between the peak luminosity and galaxy mass (which is used as a proxy for black hole mass) which is similar to our finding but explicitly account for the redshift since they use a luminosity.

Future work that expands on this one may benefit from adding such an analysis - e.g. looking at peak luminosity as well as magnitude - to draw further conclusions on possible correlations.

5.2.3 Dust echoes

We report that strong dust echoes in data used in this work appear almost exclusively for transients with $\log_{10}\left(\frac{\mathrm{M_{BH}}}{M_\odot}\right) \leq 8$, and do so completely exclusively for TDEs with a strong dust echo, in direct accordance with previous work (van Velzen et al. 2024; Wevers et al. 2017; Mockler et al. 2019; Wevers et al. 2020). However, it is of note that the black hole mass estimate for the (featureless) TDE AT2020qhs (ZTF20abowque) (Hammerstein et al., 2023) exceeds this boundary with an estimated $\log_{10}(\frac{mbh}{M_\odot}) = 8.267$. This instance has no measured dust echo hence why it is not in visible in figure 4.7. This shows that $\log_{10}\left(\frac{\mathrm{M_{BH}}}{M_\odot}\right) \leq 8$ is less of a rule and more of a trend which need not stricly be true. However, it is of interest for future work to investigate if there are more TDEs with higher $\mathrm{M_{BH}}$ and to investigate if they are featureless TDEs as well.

5.2.4 Unclassified flares

With the definition of an extreme flare set in this work, we find 725 flares to be extreme of which 617 (roughly 85.1%) are as of yet unclassified. Since these extreme flares as as of yet ill-understood, it is vital that nuclear transient events are properly classified. The more classified transients are available, the more certain we can be about population trends and possible origins of extreme flares in AGN.

In table 4.4 we present 50 flares that were marked as unclassified in our dataset, sorted and selected on the by dust echo strength. It is of note that some of these flares have in fact been previously classified or have been put forward as candidates of a certain type. Those that were marked as classified in the TNS were classified as certain types of SNe; most notably AT2019meh as a SLSN (Nicholl et al., 2019a) and AT2020iq as a SN IIn (Dahiwale & Fremling, 2020a). Under certain circumstances, nuclear SNe can look spectroscopically very similar too rapid AGN flares in the optical (Frederick et al., 2021; Moriya et al., 2018). We see in figure 4.10 that supernovae tend to have low infrared luminosities and dust echo strength compared to other classifications. AT2019meh and AT2020iq, among others, have been explicitly marked in figure 4.10. We report it is much more likely that AT2019meh and AT2020iq are either some type of AGN or TDE. The results shown in figure 4.10 are very much similar to those of (van Velzen et al., 2024, figure 3) which further enforces the implications of the results. The other transients marked as a type of SN in table 4.4 are not pulled into question following the same logic, given their placement in figure 4.10.

Misclassification of transient bursts as SNe is not unheard of. For example real-time classification by the ZTF alert process marked some transients as SNe while later analysis by Frederick et al. (2021) revealed this to be untrue. We suggest more investigation on the particular flares mentioned here to verify their TNS classification, using e.g. detailed spectra besides only the optical lightcurve.

Using machine learning for classification

The classification of nuclear transients is a vital part of any population study. As can be seen in table 4.1, a large majority of transients investigated in this work are as of yet unclassified. The classification of transients is an important issue, since it is a necessity in investigating population trends of certain types of transients. Usually, the classification of transients is tedious and time-consuming work. This can, however, be made easier by certain sophisticated algorithms.

For instance, real-time classification can play a vital role in prioritizing measurements of certain flares while they are ongoing (Mahabal et al., 2008). The classification of a transient is a process that needs investigating into multiple different aspects of the transient - not just the optical lightcurve properties, or the infrared dust echo but the spectroscopic properties need attention as well (see e.g. Hammerstein et al. 2023). Real-time classification, however, is still prone to making mistakes, as mentioned in this work for certain flares marked as SNe in the TNS and shown in Hammerstein et al. (2023).

Machine learning can offer unique opportunities to classify transients both in real time and after detection; see Gomez et al. (2023),van Roestel et al. (2021),Muthukrishna et al. (2019),Mahabal et al. (2008) and most recently Stein et al. (2024). An abundant set of features is used in these works and it is put forth here that it might be of great interest to use the properties of the thousands of transients - both previously classified and not - discovered in this work together with e.g. the `tdescore` framework (Stein et al., 2024) in order to create a reliable TDE sample. We have shown here various empirical correlations between properties that differ between TDEs and extreme AGN (tables 4.2 and 4.3) and how extreme AGN and non-extreme AGN differ from each other (figure 4.11). Moreover, we can see areas where certain types of TDEs seem to cluster, such as the SNe in figures 4.2 and 4.5, or the AGN in figure 4.9. Such patterns are vital information for a ML framework such as `tdescore` or `FLEET` (Gomez et al., 2023).

In this work, we have not made use of the temperature estimate for anything other than convenience and generalization. However, the temperature estimate can in fact help in classifying SNe. Therefore future work that hopes to classify unclassified transients may benefit greatly from the temperature estimates presented here.

We moreover report the relation between the optical peak to baseline variability ratio $\frac{\Delta F_g}{F_{rms}}$ and M_{BH} as a possible benchmark in discerning between regular AGN and TDEs (see figure 4.9). More investigation into this property is needed to fully explore its usefulness, though the initial results are presented here show promise, especially at relatively low black hole mass ($\lesssim 10^7 M_\odot$). Such a feature might prove useful for machine learning algorithms to adopt. We therefore suggest for future work to expand on this one by using the information of the thousands of transients found here and investigate the usefulness of properties such as $\frac{\Delta F_g}{F_{rms}}$.

6. Conclusions

In this research, we investigated the trends in populations of nuclear transient events, specifically the possibility of TDEs being the main mechanism behind as of yet ill-explained extreme flares in active AGN. Using an initial list of 14258 transient alerts from the Zwicky Transient Facility, we tried to fit a Gaussian rise plus exponential decay model to the optical light curve. This method was succesful in inferring the parameter values for 9577 of these instances. We moreover find black hole mass estimates and classifications as per the methodology in van Velzen et al. (2024), as well as dust echo strength measurements found through NEOWISE observations.

With a two-sample Kolmolgorov-Smirnoff test yielding a p-value of 0.0907, we conclude that the black hole mass estimates for extreme and non-extreme AGN are likely drawn from different underlying distributions at a level of confidence $\alpha = 0.010$. By investigating correlation between parameters for both AGN and TDEs we show the existence of various parameter correlations for Tidal Distruption Events that do not exist for extreme AGN flares and vice versa. This suggests that TDEs cannot be said to be the mechanism behind all extreme AGN flares in our set, which is further confirmed by another KS-test on the black hole mass estimates yielding a p-value of $1.637 \cdot 10^{-17}$. Though the possibility of TDEs being the cause behind some of the extreme AGN in our set cannot be ruled out, we do find that it is highly unlikely that TDEs are the sole mechanism responsible for them.

We further present a list of 50 unclassified transients that have strong dust echoes that are of interest for classification and further inspection. Some have already been shown to be candidates of other types, while others have reported classifications in the TNS as some type of SN, which we show here to be likely false for the instances AT2019meh and AT2020iq based on infrared afterglow measurements.

7. Acknowledgements

I would like to greatly thank Sjoert van Velzen for the continued support, mentorship and ideas that made this thesis possible.
Special thanks goes out to Simeon Reusch for generating the initial list of ZTF alerts this work makes use of.
This work makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a joint project of the Jet Propulsion Laboratory/California Institute of Technology and the University of Arizona. NEOWISE is funded by the National Aeronautics and Space Administration.
Software: `numpy` (Harris et al., 2020), `scipy` (Virtanen et al., 2020), `matplotlib` (Hunter, 2007), `astropy` (Robitaille et al., 2013), `pandas` (McKinney, 2010) and the `numpy` encoder for `json` files by Hunter M. Allen (available at https://pypi.org/project/numpyencoder/).

A. Supplementary figures

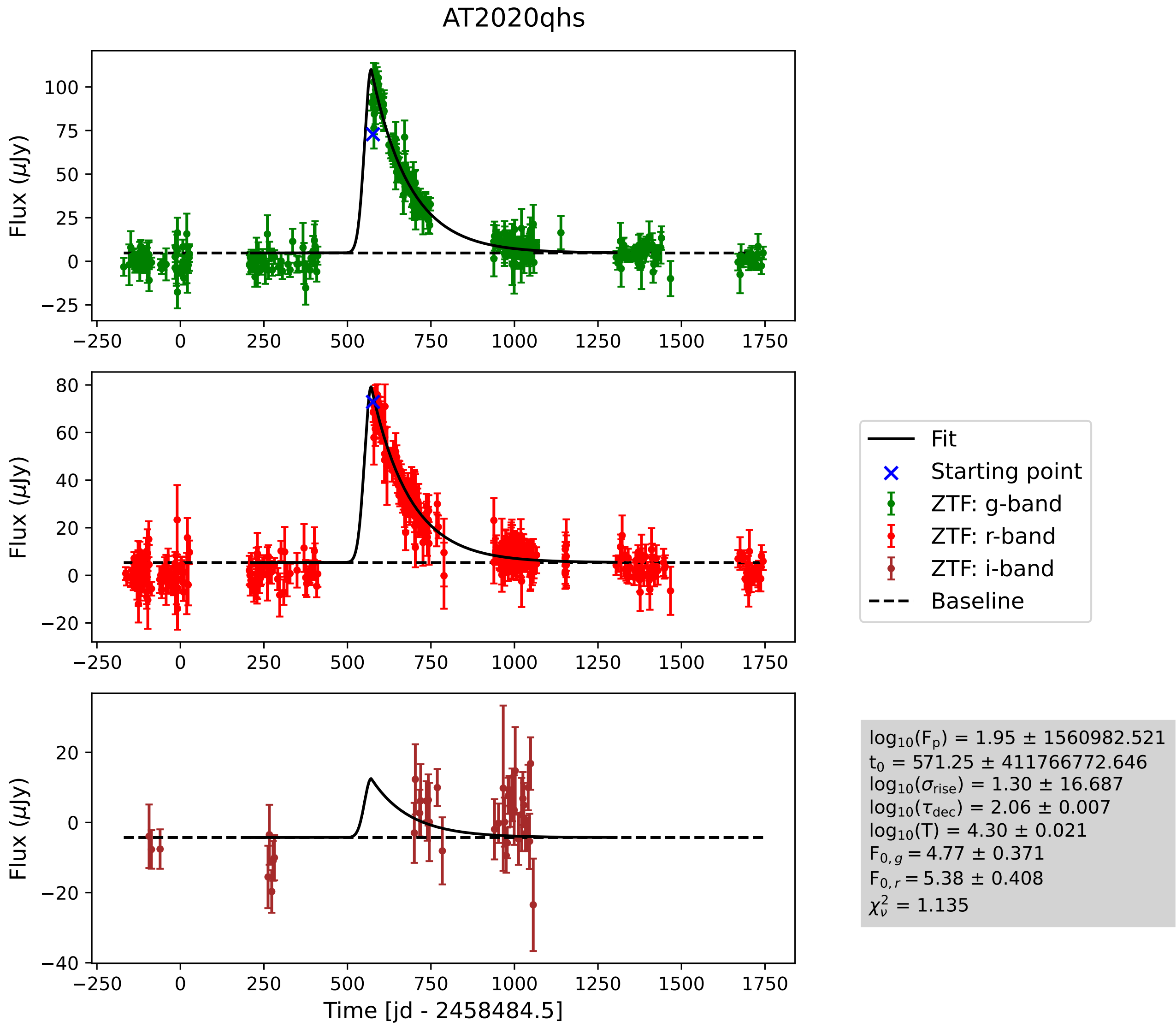

Figure A.1: *Lightcurve of TDE AT2020qhs (ZTF20abowque), with a Gaussian rise plus exponential decay model fit to it. On the vertical axis is shown the forced difference flux in μJy, and on the horizontal axis the time in mjd (w.r.t. 01-01-2019). The starting point corresponds to the initial guess of the peak according to the cross-correlation method. Shown are the found parameter values as well as the reduced-χ^2 metric (χ^2_ν) measured on the combined g and r band data as a goodness of fit indicator. Reliable data was missing entirely during the rise of its burst, and as such some parameters could not reliably be fit (as apparent by the errorbars).*

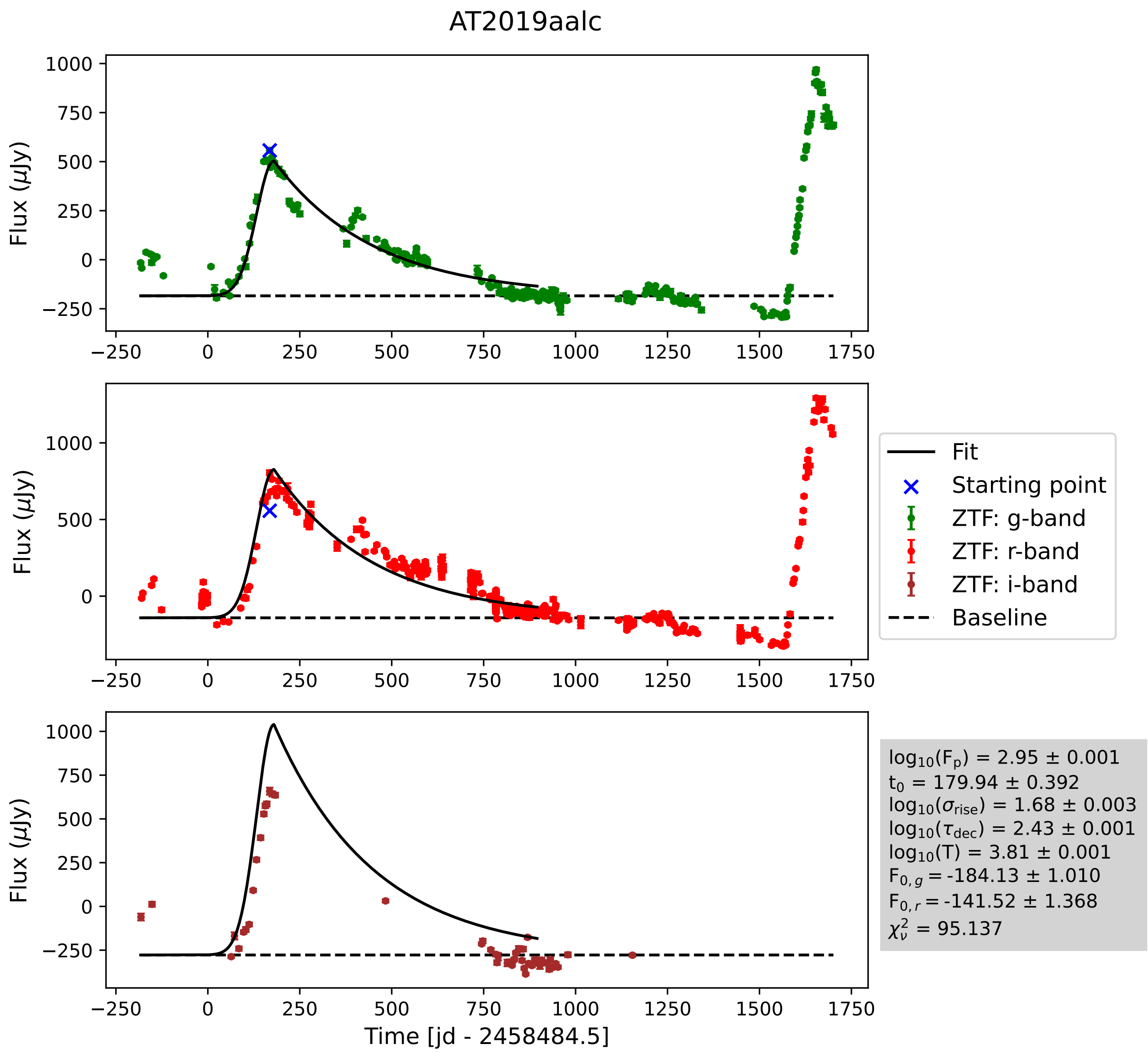

Figure A.2: *The clean data of AGN AT2019aalc (ZTF19aaejtoy) in all 3 filters as well as the corresponding Gaussian rise plus exponential decay model fit to this data. On the vertical axis is shown the forced difference flux in μJy, and on the horizontal axis the time in mjd (w.r.t. 01-01-2019). The starting point corresponds to the initial guess of the peak according to the cross-correlation method. Shown are the found parameter values as well as the reduced-χ^2 metric (χ^2_ν) measured on the combined g and r band data as a goodness of fit indicator.*

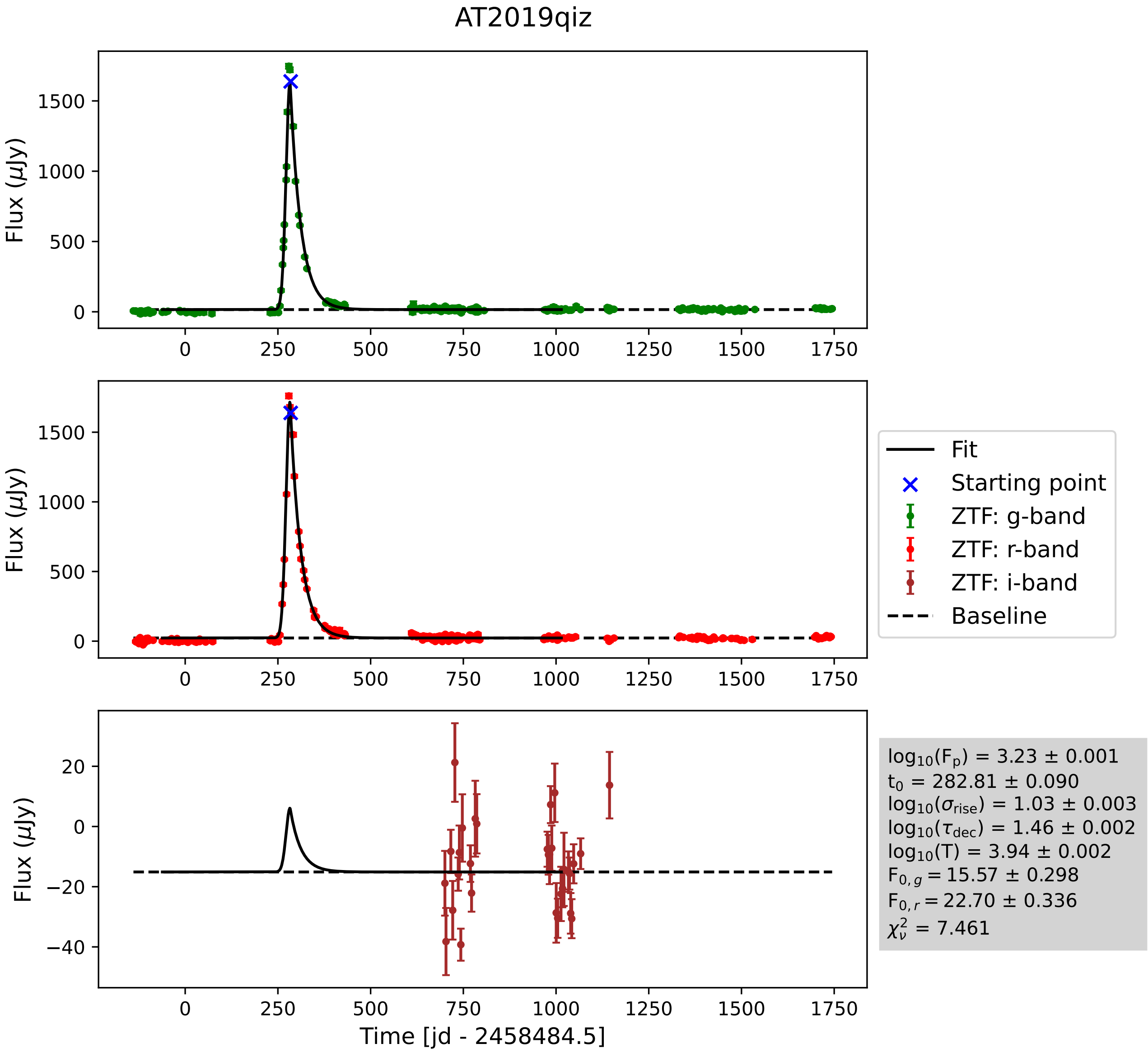

Figure A.3: *The clean data of the TDE AT2019qiz (ZTF19abzrhgq) in all 3 filters as well as the corresponding Gaussian rise plus exponential decay model fit to this data. On the vertical axis is shown the forced difference flux in μJy, and on the horizontal axis the time in mjd (w.r.t. 01-01-2019). The starting point corresponds to the initial guess of the peak according to the cross-correlation method. Shown are the found parameter values as well as the reduced-χ^2 metric (χ^2_ν) measured on the combined g and r band data as a goodness of fit indicator.*

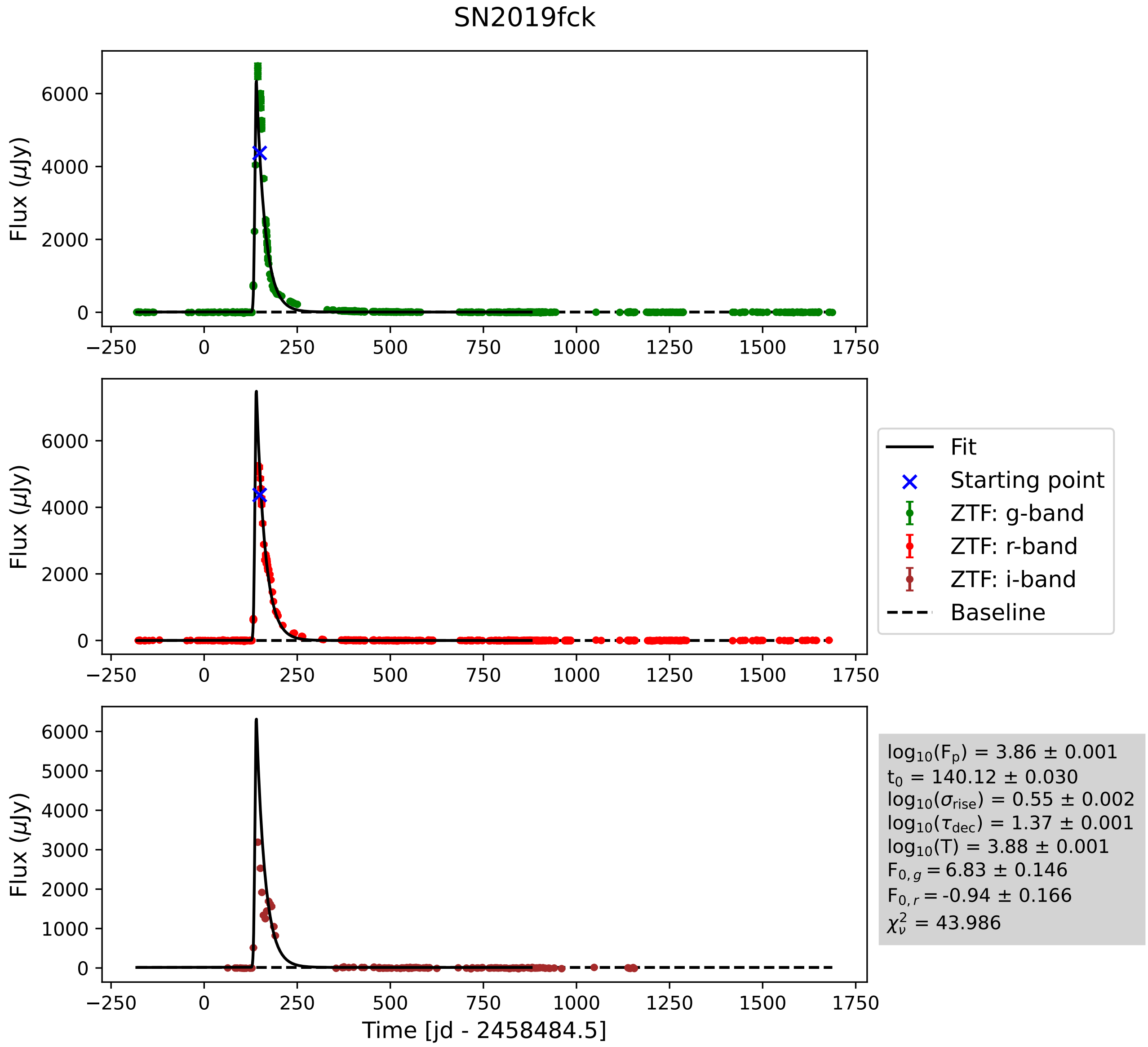

Figure A.4: *The clean data of SN2019fck (ZTF18aaqkcso) in all 3 filters as well as the corresponding Gaussian rise plus exponential decay model fit to this data. On the vertical axis is shown the forced difference flux in μJy, and on the horizontal axis the time in mjd (w.r.t. 01-01-2019). The starting point corresponds to the initial guess of the peak according to the cross-correlation method. Shown are the found parameter values as well as the reduced-χ^2 metric (χ^2_ν) measured on the combined g and r band data as a goodness of fit indicator.*

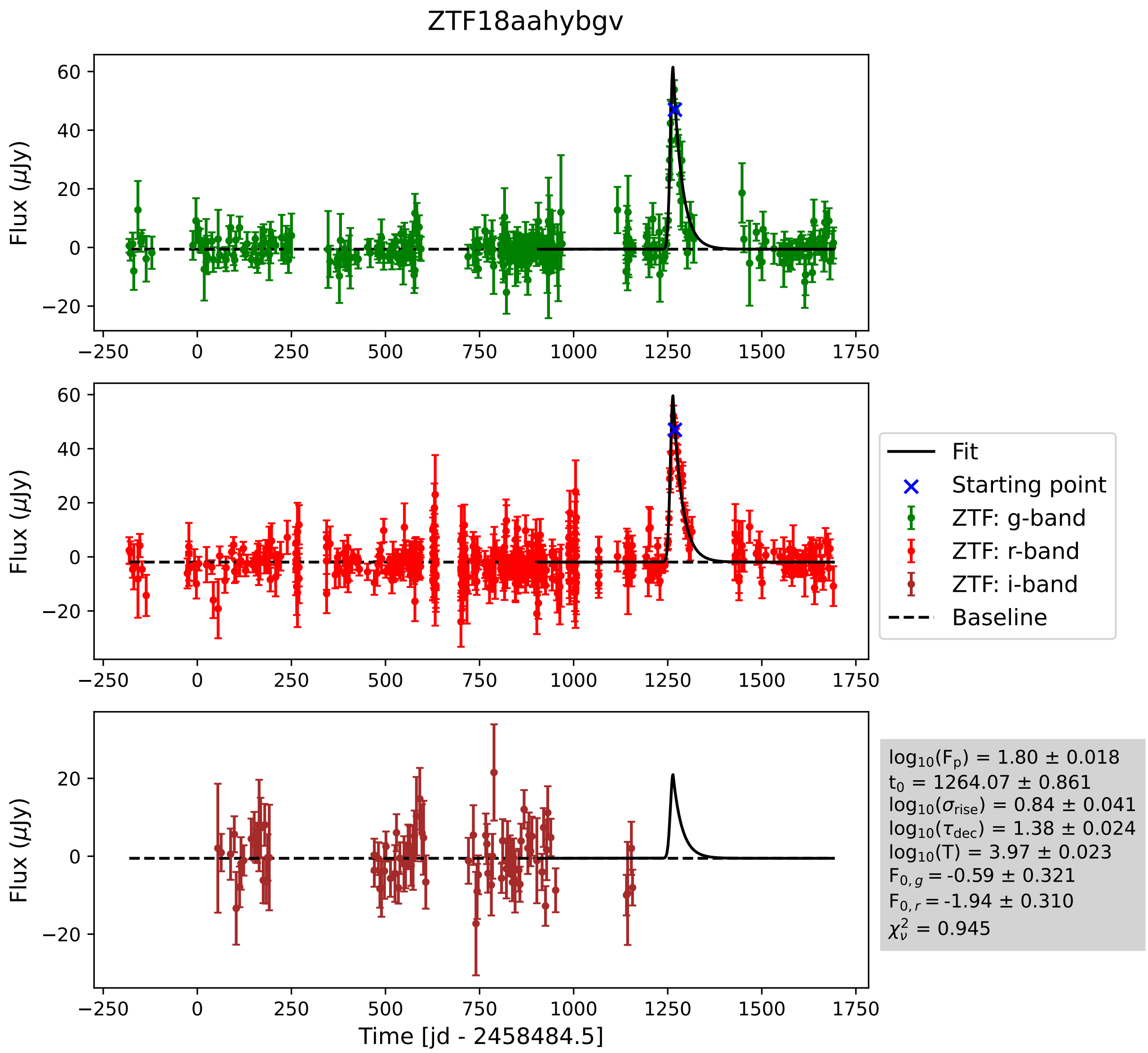

Figure A.5: *The clean data of ZTF18aahybgv in all 3 filters as well as the corresponding Gaussian rise plus exponential decay model fit to this data. On the vertical axis is shown the forced difference flux in μJy, and on the horizontal axis the time in mjd (w.r.t. 01-01-2019). The starting point corresponds to the initial guess of the peak according to the cross-correlation method. Shown are the found parameter values as well as the reduced-χ^2 metric (χ^2_ν) measured on the combined g and r band data as a goodness of fit indicator.*

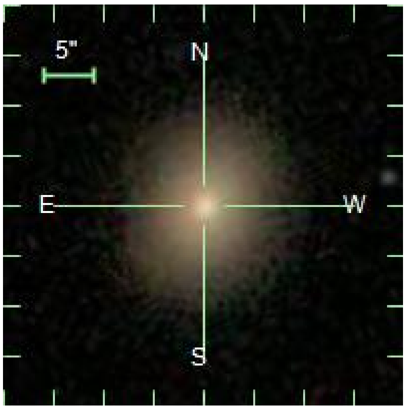

Figure A.6: *SDSS image of object SDSS J143429.19+463459.1 where transient ZTF18aahybgv took place based on coordinates. Retrieved from https://cas.sdss.org/dr7/en/tools/explore/obj.asp?id=588017711514910899.*

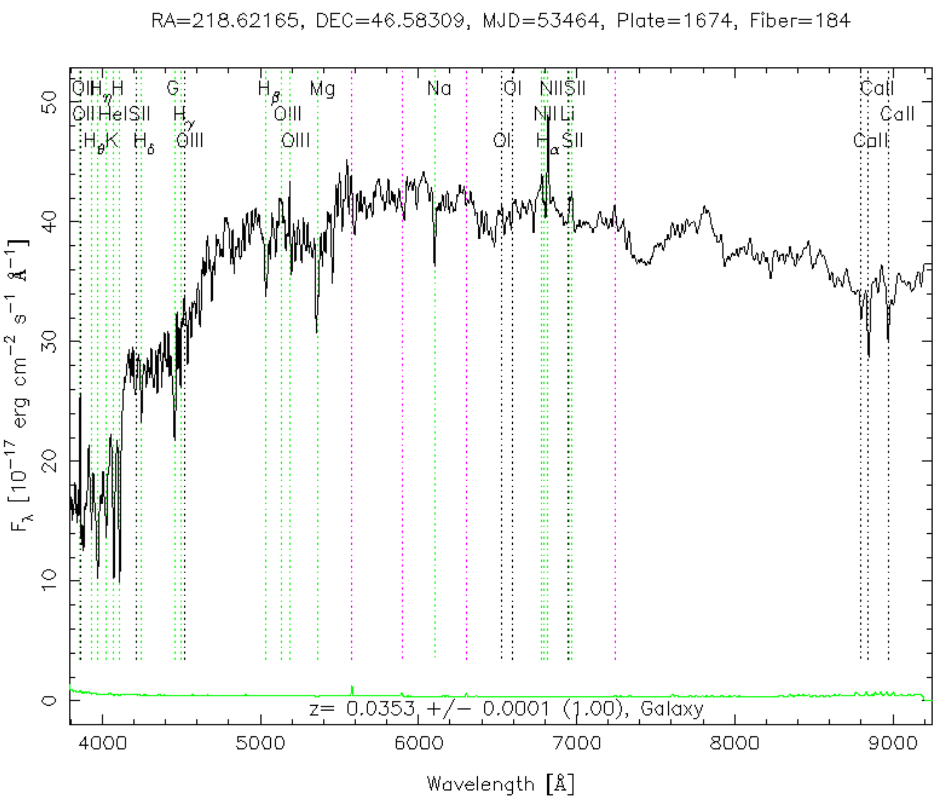

Figure A.7: *SDSS spectrum of object SDSS J143429.19+463459.1 where transient ZTF18aahybgv took place based on coordinates. Retrieved from https://cas.sdss.org/dr7/en/tools/explore/obj.asp?id=588017711514910899.*

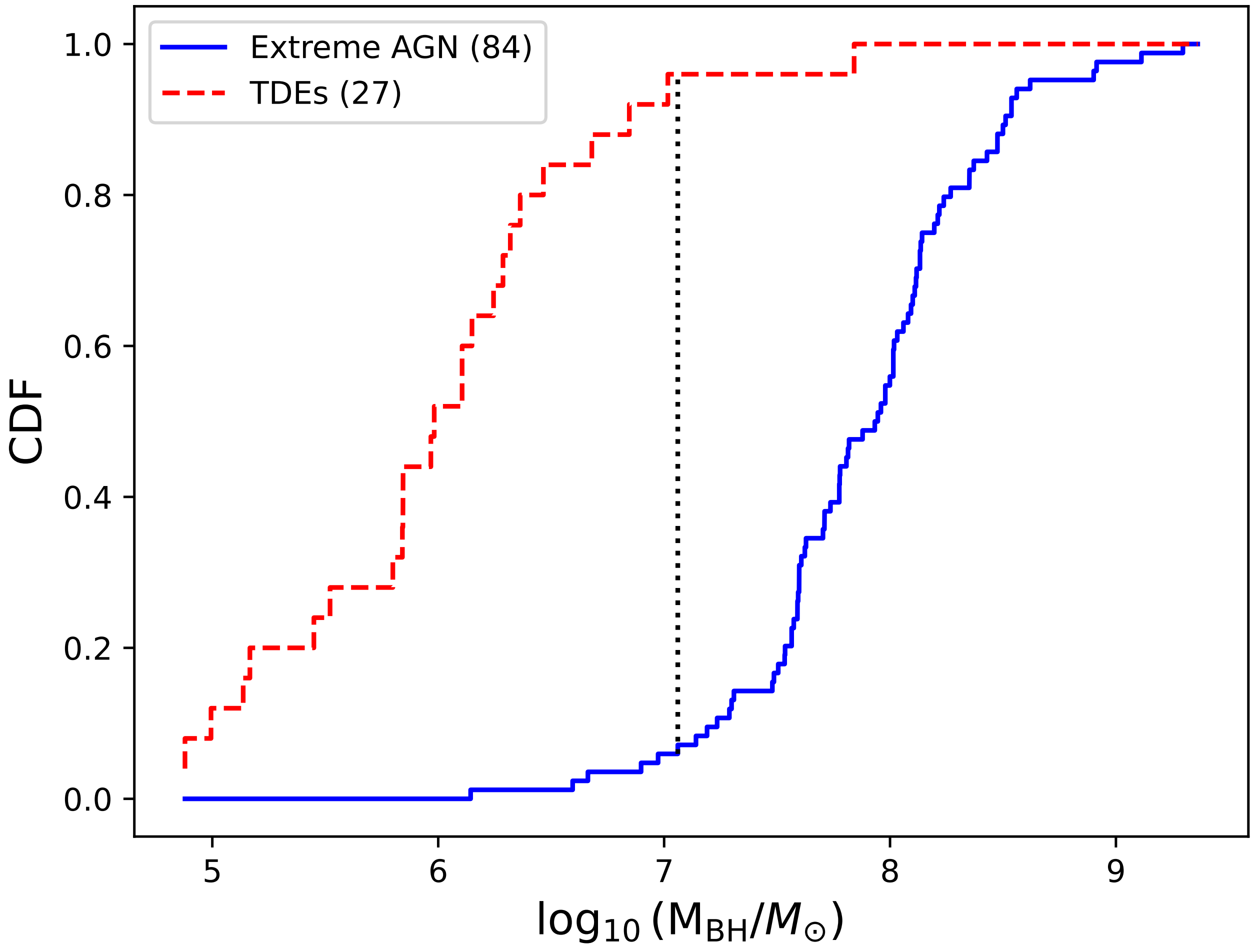

Figure A.8: CDFs of black hole mass estimates of AGN with extreme flares and TDEs. The black line shows the maximal distance as per a KS test which is significant enough to reject the null hypothesis that these two sub-samples drawn from the same population. The size of each sub-sample is shown in the legend.

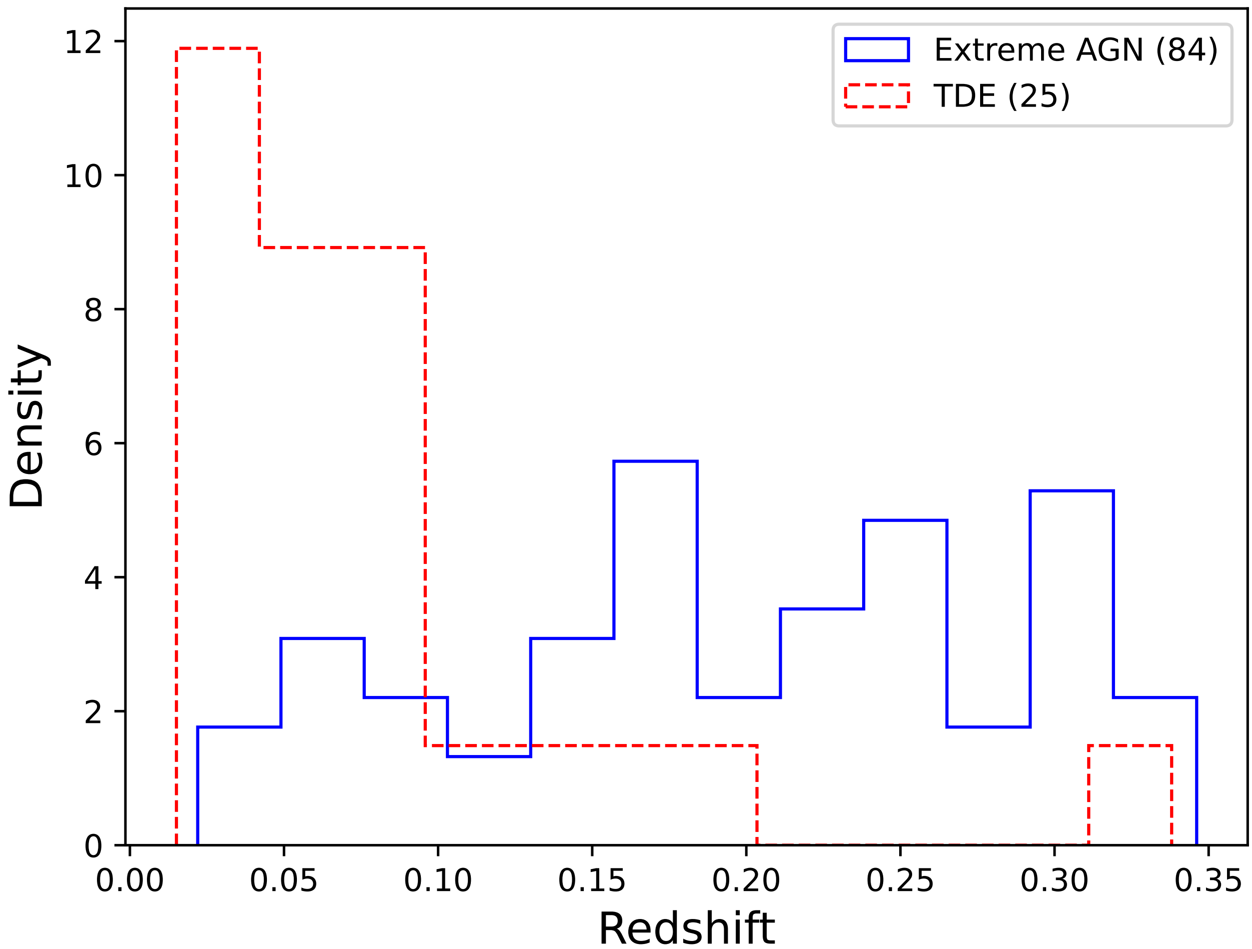

Figure A.9: *Redshift distributions of AGN with extreme flares and* TDEs *with a known black hole mass.*

Bibliography

Barvainis, R. 1987, The Astrophysical Journal, 320, 537, doi: 10.1086/165571

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2018, Publications of the Astronomical Society of the Pacific, 131, 018002, doi: 10.1088/1538-3873/aaecbe

Blanchard, P., Nicholl, M., Berger, E., et al. 2017, The Astrophysical Journal, 843, 106

Bode, M. F., & Evans, A. 1983, NASA/ADS. https://ui.adsabs.harvard.edu/abs/1983QJRAS..24...83B/abstract

Brown, J. S. 2018, OhioLINK ETD: Brown, Jonathan S, Jr. http://rave.ohiolink.edu/etdc/view?acc_num=osu1531743543481036

Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2017, The Astrophysical Journal, 834, 111, doi: 10.3847/1538-4357/834/2/111

Chan, C.-H., Piran, T., & Krolik, J. H. 2020, The Astrophysical Journal, 903, 17

Chan, C.-H., Piran, T., Krolik, J. H., & Saban, D. 2019, The Astrophysical Journal, 881, 113

Charalampopoulos, P., Leloudas, G., Malesani, D. B., et al. 2022, Astronomy amp; Astrophysics, 659, doi: 10.1051/0004-6361/202142122

Dahiwale, A., & Fremling, C. 2020a, Transient Name Server Classification Report. https://ui.adsabs.harvard.edu/abs/2020TNSCR1667....1D

—. 2020b, NASA/ADS. https://ui.adsabs.harvard.edu/abs/2020TNSCR1176....1D/abstract

Donley, J. L., Brandt, W. N., Eracleous, M., & Boller, T. 2002, The Astronomical Journal, 124, 1308–1321, doi: 10.1086/342280

Dou, L., Wang, T.-g., Jiang, N., et al. 2016a, The Astrophysical Journal, 832, 188, doi: 10.3847/0004-637x/832/2/188

—. 2016b, The Astrophysical Journal, 832, 188, doi: 10.3847/0004-637x/832/2/188

Drake, A., Djorgovski, S., Mahabal, A., et al. 2011, The Astrophysical Journal, 735, 106

Dwek, E. 1983, The Astrophysical Journal, 274, 175, doi: 10.1086/161435

Edelson, R. A., & Malkan, M. A. 1986, The Astrophysical Journal, 308, 59, doi: 10.1086/164479

Frederick, S., Gezari, S., Graham, M. J., et al. 2021, The Astrophysical Journal, 920, 56, doi: 10.3847/1538-4357/ac110f

Fremling, C., Dugas, A., & Sharma, Y. 2019, ZTF transient classification report for 2019-05-15. https://ui.adsabs.harvard.edu/abs/2019TNSCR.790....1F/abstract

Fremling, C., Miller, A., Sharma, Y., et al. 2020, The Astrophysical Journal, 895, 32

Gezari, S. 2021, Annual Review of Astronomy and Astrophysics, 59, 21, doi: https://doi.org/10.1146/annurev-astro-111720-030029

Gomez, S., Villar, V. A., Berger, E., et al. 2023, The Astrophysical Journal, 949, 113, doi: 10.3847/1538-4357/acc535

Graham, M. J., McKernan, B., Ford, K. E., et al. 2023, The Astrophysical Journal, 942, 99, doi: 10.3847/1538-4357/aca480

Graham, M. L., Dahiwale, A., & Fremling, C. 2020a, NASA/ADS. https://ui.adsabs.harvard.edu/abs/2020TNSCR.205....1G/abstract

Graham, M. L., Sipocz, B., Bellm, E. C., et al. 2020b, ZTF bright transient survey classifications with Apache Point Observatory. https://ui.adsabs.harvard.edu/abs/2020ATel13418....1G/abstract

Halpern, J. P., & Eracleous, M. 1994, Astrophysical Journal, Part 2-Letters (ISSN 0004-637X), vol. 433, no. 1, p. L17-L20, 433, L17

Hammerstein, E., van Velzen, S., Gezari, S., et al. 2023, The Astrophysical Journal, 942, 9, doi: 10.3847/1538-4357/aca283

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357–362, doi: 10.1038/s41586-020-2649-2

Heinis, S., Gezari, S., Kumar, S., et al. 2016, The Astrophysical Journal, 826, 62, doi: 10.3847/0004-637x/826/1/62

Hills, J. G. 1975, Nature, 254, 295–298, doi: 10.1038/254295a0

Holoien, T. W.-S., Prieto, J. L., Bersier, D., et al. 2014, Monthly Notices of the Royal Astronomical Society, 445, 3263–3277, doi: 10.1093/mnras/stu1922

Hoshi, A., Yamada, T., & Ohta, K. 2024, Publications of the Astronomical Society of Japan, 76, 103, doi: 10.1093/pasj/psad083

Hung, T., Gezari, S., Blagorodnova, N., et al. 2017, The Astrophysical Journal, 842, 29, doi: 10.3847/1538-4357/aa7337

Hunter, J. D. 2007, Computing in Science Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

Jiang, N., Dou, L., Wang, T., et al. 2016, The Astrophysical Journal Letters, 828, doi: 10.3847/2041-8205/828/1/l14

Jiang, N., Zhou, Z., Zhu, J., Wang, Y., & Wang, T. 2023, The Astrophysical Journal Letters, 953, doi: 10.3847/2041-8213/acebe3

Jiang, N., Yang, H., Wang, T., et al. 2022, arXiv.org. https://arxiv.org/abs/2201.11633

Kendall, M. G. 1938, Biometrika, 30, 81–93, doi: 10.1093/biomet/30.1-2.81

Komossa, S., Zhou, H., Rau, A., et al. 2009, The Astrophysical Journal, 701, 105–121, doi: 10.1088/0004-637x/701/1/105

Koss, M., Trakhtenbrot, B., Ricci, C., et al. 2017, The Astrophysical Journal, 850, 74, doi: 10.3847/1538-4357/aa8ec9

Laurenti, M., Piconcelli, E., Zappacosta, L., et al. 2022, Astronomy amp; Astrophysics, 657, doi: 10.1051/0004-6361/202141829

Liu, H.-Y., Liu, W.-J., Dong, X.-B., et al. 2019, The Astrophysical Journal Supplement Series, 243, 21

Lu, W., Kumar, P., & Evans, N. J. 2016, Monthly Notices of the Royal Astronomical Society, 458, 575–581, doi: 10.1093/mnras/stw307

Mahabal, A., Djorgovski, S. G., Williams, R., et al. 2008, AIP Conference Proceedings, doi: 10.1063/1.3059064

Mainzer, A., Bauer, J., Grav, T., et al. 2011, The Astrophysical Journal, 731, 53, doi: 10.1088/0004-637x/731/1/53

Mainzer, A., Bauer, J., Cutri, R. M., et al. 2014, The Astrophysical Journal, 792, 30, doi: 10.1088/0004-637x/792/1/30

Makrygianni, L., Trakhtenbrot, B., Arcavi, I., et al. 2023, The Astrophysical Journal, 953, 32

Masci, F. J., Laher, R. R., Rusholme, B., et al. 2018, Publications of the Astronomical Society of the Pacific, 131, 018003, doi: 10.1088/1538-3873/aae8ac

—. 2023, arXiv.org. https://doi.org/10.48550/arXiv.2305.16279

McKinney, W. 2010, Proceedings of the Python in Science Conference, doi: 10.25080/majora-92bf1922-00a

Merloni, A., Dwelly, T., Salvato, M., et al. 2015, Monthly Notices of the Royal Astronomical Society, 452, 69

Mockler, B., Guillochon, J., & Ramirez-Ruiz, E. 2019, The Astrophysical Journal, 872, 151, doi: 10.3847/1538-4357/ab010f

Moriya, T. J., Sorokina, E. I., & Chevalier, R. A. 2018, Space Science Reviews, 214, 1

Mummery, A., van Velzen, S., Nathan, E., et al. 2023, Monthly Notices of the Royal Astronomical Society, 527, 2452, doi: 10.1093/mnras/stad3001

Muthukrishna, D., Narayan, G., Mandel, K. S., Biswas, R., & Hložek, R. 2019, Publications of the Astronomical Society of the Pacific, 131, 118002, doi: 10.1088/1538-3873/ab1609

Müller-Bravo, T. E., Sullivan, M., Smith, M., et al. 2021, Monthly Notices of the Royal Astronomical Society, 512, 3266–3283, doi: 10.1093/mnras/stab3065

Necker, J., de Jaeger, T., Stein, R., et al. 2022, Monthly Notices of the Royal Astronomical Society, 516, 2455–2469, doi: 10.1093/mnras/stac2261

Nicholl, M., Short, P., Lawrence, A., Ross, N., & Smartt, S. 2019a, NASA/ADS. https://ui.adsabs.harvard.edu/abs/2019TNSCR1586....1N/abstract

Nicholl, M., Blanchard, P. K., Berger, E., et al. 2019b, Monthly Notices of the Royal Astronomical Society, 488, 1878–1893, doi: 10.1093/mnras/stz1837

Perley, D. A., Taggart, K., Dahiwale, A., & Fremling, C. 2020, NASA/ADS. https://ui.adsabs.harvard.edu/abs/2020TNSCR1155....1P/abstract

Rees, M. J. 1988, Nature, 333, 523–528, doi: 10.1038/333523a0

Reusch, S., Stein, R., Kowalski, M., et al. 2022, Phys. Rev. Lett., 128, 221101, doi: 10.1103/PhysRevLett.128.221101

Ricci, C., & Trakhtenbrot, B. 2023, Nature Astronomy, 7, 1282

Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, Astronomy amp; Astrophysics, 558, doi: 10.1051/0004-6361/201322068

Robson, E. I., Gear, W. K., Brown, L. M., et al. 1986, Nature, 323, 134–136, doi: 10.1038/323134a0

Salpeter, E. E. 1964, The Astrophysical Journal, 140, 796, doi: 10.1086/147973

Scargle, J. D., Norris, J. P., Jackson, B., & Chiang, J. 2013, The Astrophysical Journal, 764, 167, doi: 10.1088/0004-637x/764/2/167

Seyfert, C. K. 1943, The Astrophysical Journal, 97, 28, doi: 10.1086/144488

Siebert, M. R., Strasburger, E., Rojas-Bravo, C., & Foley, R. J. 2019, Transient classification report for 2019-09-26. https://ui.adsabs.harvard.edu/abs/2019TNSCR1921....1S/abstract

Smith, K. L., Mushotzky, R. F., Boyd, P. T., et al. 2018, The Astrophysical Journal, 857, 141, doi: 10.3847/1538-4357/aab88d

Somalwar, J., Chornock, R., Ravi, V., et al. 2024, ZTF classification of TDE candidates. https://ui.adsabs.harvard.edu/abs/2024TNSAN.150....1S/abstract

Stein, R., Mahabal, A., Reusch, S., et al. 2024, The Astrophysical Journal Letters, 965, doi: 10.3847/2041-8213/ad3337

Strotjohann, N. L., Ofek, E. O., Gal-Yam, A., et al. 2021, The Astrophysical Journal, 907, 99, doi: 10.3847/1538-4357/abd032

Tonry, J. L. 2013, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371, 20120269, doi: 10.1098/rsta.2012.0269

Trakhtenbrot, B., Arcavi, I., Ricci, C., et al. 2019, Nature Astronomy, 3, 242

van Roestel, J., Duev, D. A., Mahabal, A. A., et al. 2021, The Astronomical Journal, 161, 267, doi: 10.3847/1538-3881/abe853

van Velzen, S., Holoien, T. W.-S., Onori, F., Hung, T., & Arcavi, I. 2020, Space Science Reviews, 216, doi: 10.1007/s11214-020-00753-z

van Velzen, S., Mendez, A. J., Krolik, J. H., & Gorjian, V. 2016, The Astrophysical Journal, 829, 19, doi: 10.3847/0004-637x/829/1/19

van Velzen, S., Pasham, D. R., Komossa, S., Yan, L., & Kara, E. A. 2021a, Space Science Reviews, 217, doi: 10.1007/s11214-021-00835-6

van Velzen, S., Farrar, G. R., Gezari, S., et al. 2011, The Astrophysical Journal, 741, 73, doi: 10.1088/0004-637x/741/2/73

van Velzen, S., Gezari, S., Hammerstein, E., et al. 2021b, The Astrophysical Journal, 908, 4, doi: 10.3847/1538-4357/abc258

van Velzen, S., Stein, R., Gilfanov, M., et al. 2024, Monthly Notices of the Royal Astronomical Society, 529, 2559, doi: 10.1093/mnras/stae610

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261–272, doi: 10.1038/s41592-019-0686-2

Ward, C., Gezari, S., Frederick, S., et al. 2021, The Astrophysical Journal, 913, 102, doi: 10.3847/1538-4357/abf246

Ward, C., Gezari, S., Nugent, P., et al. 2024, The Astrophysical Journal, 961, 172, doi: 10.3847/1538-4357/ad147d

Wevers, T., van Velzen, S., Jonker, P. G., et al. 2017, Monthly Notices of the Royal Astronomical Society, 471, 1694–1708, doi: 10.1093/mnras/stx1703

Wevers, T., Stone, N. C., van Velzen, S., et al. 2020, Monthly Notices of the Royal Astronomical Society, 493, 1498–1499, doi: 10.1093/mnras/staa461

Wiseman, P., Williams, R., Arcavi, I., et al. 2024, arXiv preprint arXiv:2406.11552

Worpel, H., & Schwope, A. D. 2015, Astronomy amp; Astrophysics, 578, doi: 10.1051/0004-6361/201525946

Wright, E. L., Eisenhardt, P. R., Mainzer, A. K., et al. 2010, The Astronomical Journal, 140, 1868–1881, doi: 10.1088/0004-6256/140/6/1868

Yu, Z., Kochanek, C. S., Mathur, S., et al. 2022, Monthly Notices of the Royal Astronomical Society, 515, 5198–5210, doi: 10.1093/mnras/stac2073

Zabludoff, A., Arcavi, I., La Massa, S., et al. 2021, Space Science Reviews, 217, doi: 10.1007/s11214-021-00829-4

Zel'dovich, Y. B., & Shakura, N. I. 1964, NASA/ADS. https://ui.adsabs.harvard.edu/abs/1969SvA....13..175Z/abstract