DUST REVERBERATION MAPPING IN NINE TYPE 1 ACTIVE GALACTIC NUCLEI

Joshua Weston

University of Southampton

Table of Contents

DUST REVERBERATION MAPPING IN NINE TYPE 1 ACTIVE GALACTIC NUCLEI

17th April 2020

Joshua Weston

University of Southampton

Abstract

Previous studies have indicated that the relation between the time delay of an active galactic nucleus' (AGN) accretion disk emission and dust region emission and its luminosity may be linearly correlated and fit for use as a standard candle. In this study we examine nine AGN of redshift $0.1 < z < 1.2$ via cross-correlation analysis in order to determine dust reverberation lags and their relation to the AGN luminosities. First we provide an overview of the unification model of AGN which details the structure assumed in our study. We present an overview of dust reverberation mapping and its pros and cons compared to other forms of AGN reverberation mapping, as well as studies conducted on the dust region structure that inform our results. We discuss the benefits that the use of AGN dust lags as standardizable candles would provide in determining new constraints on cosmological parameters. Carrying out the cross-correlation analysis on our AGN we determine the dust lags. Discarding outlying lags and objects we calculate a linear relation with slope x and offset y in keeping with previous results but in disagreement with theory, first through minimization of the relation and then refined through ensemble sampling. We discuss reasons for the discrepancy between experimental and theoretical findings, including the presentation of improvements that could be made to the methodology in future studies. We also present two different and unsuccessful attempts to remove the effects of the long-term variability in the AGN emission in order to better determine dust lags.

1.0 Introduction

Observations of supermassive black holes (SMBH) present in the centres of most galaxies indicate the existence of growth phases where the SMBH accretes interstellar gas and releases energy in response [1], [2], [3]. When observed in this growth stage the SMBH are classified as active galactic nuclei (AGN) with properties measured by examination of emitted electromagnetic radiation, gas outflows and ionized plasma. The accreted gas is responsible for a central source observable in the UV/optical frequencies which may be absorbed by surrounding components. Under the unification model of AGN one such component is introduced as a thick structure of dust and electrons capable of reemitting this radiation in the infrared.

It has been proposed that AGN may be used as standardizable candles through study of the time delays between emission variation in the broad line or dust regions and the accretion disk [4], [5], [6], [7], [8]. These time delays or lags are interpreted as the time

taken for light to travel from the accretion disk to the dust torus, assuming negligible time between absorption and reemission by the dust. Within a certain radius of the central engine of the AGN dust will sublimate due to high temperatures, allowing us to determine the radius of the dust torus with high accuracy (the "sublimation radius" R_{sub}) [9], [10], [11]. From the equations of radiative transfer at equilibrium, assuming that the AGN are standardizable candles, we should then expect to find a linear relation between AGN magnitudes and their respective dust lags. This method, known as reverberation mapping, would allow new constraints to be placed on cosmological parameters such as the Hubble constant or dark energy density. Reverberation mapping has often been successfully achieved through use of cross-correlation analysis [12], [13], [14], [15], with a range of potential lags induced in a data set in order to determine peaks in the cross-correlation function which correspond to the likeliest lag. Studies have varied between assuming the value of the slope in the lag-luminosity relation and finding the likeliest value for the data examined. Reverberation mapping has also been attempted through determination of the time delay in emission lines in tandem with interferometry [16], [17].

In this study we take nine AGN previously examined through reverberation mapping by Koshida et al. [15] and Lira et al. [18] in order to determine whether our experimental fitting of a relation between the dust lag value and the AGN luminosity matches the result obtained by theory. We first examine the general background of active galactic nuclei, the dust torus and the standard unification model of AGN, as well as an overview of the dust lag-luminosity relation expected from theory and supported by previous study. Stating the origin of our data we then outline our method, including relevant code in the appendices, involving lag induction into the data set and cross-correlation analysis before determining the likely value of the dust lags. We take the logarithms of our results and attempt to fit a linear model to the data through use of ensemble sampling. Finally we discuss our findings, comparing to theoretical and previously obtained results, before making our concluding arguments.

2.0 Theory

2.1 General Background of AGN

Many galaxies (as many as 10% of large galaxies in particular) contain highly luminous active galactic nuclei that strongly emit non-thermal radiation from the radio regime to the gamma-ray regime on the electromagnetic spectrum. This high luminosity makes them easy to detect at redshifts as high as >7, with predictable emission [2], [19], [20]. The emitting regions of AGN, found to have sizes of the order of parsecs, are powered by the accretion of matter by a supermassive black hole (SMBH) with a gravitational potential that acts as an energy source – the central engine. As the accreting matter (gas) has angular momentum this accretion occurs through a disk [21].

AGN are grouped into a variety of classes. Seyfert 1 types, characterized by broad emission lines and bright continua, have stronger non-stellar continua and greater X-ray emission than Seyfert 2 [21], [22], [23]; in addition, their permitted lines (typically hydrogen) are broader, with narrow forbidden lines in both. Differences in the full-width at half maximum (FWHM) between permitted and forbidden lines in Seyfert 1 galaxies imply different regions of production, unlike Seyfert 2 galaxies where the FWHM indicates the existence of only one emitting region. Another class, radio galaxies, again have two subclasses. Broad-line radio galaxies (BLRGs) and narrow-line

radio galaxies (NLRGs) are characterized by strong radio emission relative to regular spiral galaxies and strong emission lines [24], [25]. Radio maps of these galaxies show a double-lobed structure extending from the central engine with properties that can be used to further classify AGN [26], [27]. Despite differing morphologies radio galaxies can be considered to be radio-loud Seyfert galaxies. These classes are often referred to in literature as simply Type 1 and Type 2.

Although different classes of AGN have been observed to possess these distinct emission properties there is enough similarity to indicate that AGN have a standard structure that is observed to only possess these different properties when viewed at different angles [12], [27], [28]. Type 1 and Type 2 galaxies show a small but notable difference in IR emission wherein self-absorption of the obscuring structure leads to a suppression of dust emission between 3-5μm, implying the existence of a separate emission component. Observations in some Seyfert 2 objects show broad emission lines in polarized light, indicating reflection from a region of dust grains that obscure the broad-line region that is opaque to the emission. For narrow-line emission to remain unobscured this region exists within the narrow line region. In this unified model we then have a central engine that accretes gas, ejecting radio radiation normal to the disc, surrounded by a torus of dust and electrons. This model is widely accepted and is acceptable to assume for the purposes of this study, though there are still a number of caveats – e.g. certain AGN subclasses are still not clarified in the unification model [25].

The torus model must be constrained by several conditions; firstly, it must be dusty, with a high optical depth in the optical/UV frequencies [29], [30]. It is also necessary that the structure be geometrically thick and in the size scale of a parsec. As a result, infrared interferometry is only useful in resolving the torii of highly luminous local AGN. Evidence supports the structure of the dust region having a form that is not strictly toroidal [31], [32], [33], [34], [35]. Models that include clumpy structures provide accurate fits to the infrared (IR) spectral energy density distribution (SED) of many AGN supportive of dust reverberation measurements. IR observations, such as those carried out on the Type 2 galaxy NGC 1068 [36], [37], [38], support the existence of multiple dust emission sources within AGN. Mid-IR observations have been interpreted as the result of a resolved and unresolved emission source, the former elongated polar to the AGN that accounts for the point-like source observed in NGC 1068 and other similar galaxies [9]. The unresolved component is partially resolved as a thin disk in the closer, more luminous subpopulation of these galaxies [39], [40]. Models have thus been proposed that incorporate a thin disk on scales less than one parsec that experiences an inflow of dusty gas from scales of ~100pc [41]. The inner part of the disc distends under the IR pressure and unbinds gas from the central engine, launching wind from the AGN which feeds back into the galaxy. The mid-IR emission then originates from the disc component and the polar cone, with near-IR emission from the sublimation region.

2.2 Dust Reverberation Mapping

Previous studies of AGN have shown that the dust torus in more distant AGN can be examined via comparison between its infrared (IR) dust emission to the continuum emission [42], [43], [44]. AGN accretion disks emit in the optical/ultraviolet (UV) frequencies radiation which is absorbed and re-emitted in the infrared [45], [46], [47], [48], [49]. A time lag between the variation in optical flux and near-IR flux is then

interpreted as the time taken for light to travel from the accretion disk to the dust torus. The thermal processing and reemission of continuum UV and optical radiation by the dust torus reveals itself as a bump in the IR between 1μm and 100-1000μm with a peak at ~20μm, indicating a boundary between the dust region and a dust-free region between the torus and accretion disk [19]. This distance can be interpreted as a sublimation radius Rsub, inside which dust will sublimate due to the high temperature. Depending on the species of dust this sublimation temperature T_{sub} will range between 1200 – 1900K [50], [51], [52]. The lag in variation is then interpreted as the time taken for light to travel from the accretion disk to the radius of the innermost region of the dust torus.

Reverberation mapping in AGN was initially applied to the broad line region and has been useful in constraining the masses of the SMBH in distant objects, allowing extensive study of the size-luminosity relation [14], [53], [54]. By taking variations in continuum and the line responses, a time lag can be found between the similar emissions of the two. There are several points to consider when comparing reverberation mapping using the broad emission line region versus the dust continuum. Previous studies have shown that there exists a tighter relation between an AGN's lag and luminosity when using the dust continuum [15]. Conversely, broad emission line lags can be found at higher redshifts when compared to dust continuum lags [55], [56]. It has been reported that dust time lags can place an upper bound on the broad emission line lags which lends evidence to the AGN unification model [12].

Additional alternative approaches have been used to measure the reverberation lags in AGN. The Stochastic Process Estimation for AGN Reverberation (SPEAR) developed by several studies fits the light curves using a random walk model before aligning them to recover the time delay [57], [58], [59]. Another alternative method has been used involving the use of photoionization theory to estimate the BLR distance, which could apply to multiple AGN models [60]. The main advantage of the photoionization method is that it can be used for any AGN observed using low-resolution spectroscopy, significantly increasing the population of objects with known central masses and BLR sizes [61], [62]. This method is indirect and results in greater uncertainties when compared against the reverberation method, and so is often calibrated using the latter [63]. As a result the reverberation technique is selected as the best method for the purpose of this study, which examines low-redshift AGN with variable light curves.

The reverberation method is limited by its necessity for variability in the AGN and a well sampled light curve or spectra and thus is only useful for obtaining the time delay for certain objects, namely the low or moderate luminosity Type 1 galaxies [62].

Near-IR interferometry and reverberation mapping have shown an offset in smaller scales of the size-luminosity with respect to theoretical prediction [16], [64]. This can be explained via the assumption that the hot dust consists of large blackbody graphite grains, naturally supported by the fact that only the largest blackbody-like grains are able to resist sublimation at the 2.2μm emission [65], [66], [64]. Evidence of this has previously been discovered in that hot dust emission regions in many AGN are found to have similar characteristics; in particular the AGN used in this study.

2.3 Standard Candles

The local number densities of different AGN galaxy classes are found to be

between 10-4 and 10-9 Mpc-3; orders of magnitude lower than standard galaxy classes but still populous enough to be detected regularly in space [21]. This would make them excellent standard candles should their absolute magnitude be calibratable and equal for all objects in the class. A known luminosity would also prove useful in classifying AGN and lend further evidence to the unification theory.

It has been shown than the time delays between optical/UV and reprocessed emission can be used to determine constraints for certain cosmological parameters – most relevant to this study is the evidence that indicates the lag between the accretion disk emission and the dust reemission can be used as a standard candle due to its correlation with luminosity [67], [68], [7], [8]. Previous studies have also proposed the use of BLR lags or dust lags as standard rulers in order to set an absolute distant scale to AGN [6], [8]. A new standard ruler would be useful in cosmology, for example in constraining the value of the Hubble constant. Hönig et al. [19] have shown the viability of AGN dust lags for use as standard candles between redshifts $0.1 < z < 1.2$ via use of a mock sample. Previous studies have used near-infrared interferometry to measure distance to nearby AGN in order to calculate the Hubble constant as a means of calibrating the AGN hot-dust standard candle.

In order to prove the viability of AGN dust lags as standard candles we are required to prove a linear relation between the dust lag and the AGN luminosity.

2.4 The Dust Lag – Luminosity Relation

We are able to relate the dust lag to the AGN luminosity beginning from Wien's Law:

$$\lambda_{max} = \frac{2.93\times10^{-3}}{T}, \qquad [1]$$

where T is temperature and λ_{max} corresponds to a peak in the Planck spectrum. From the definition of T_{sub} this equates to a wavelength of order ~um at the sublimation radius. It is thus sensible to measure direct AGN emission in the V band and dust emission in the K band.

The Stefan-Boltzmann law relates blackbody flux F to temperature. Using the inverse-square law we can then relate the AGN luminosity L_{AGN} to the sublimation radius:

$$F = \sigma_{SB} T^4 \qquad [2]$$

$$L = 4\pi R^2 \sigma_{SB} T^4, \qquad [3]$$

where σ_{SB} is the Stefan-Boltzmann constant. In reality the correct relation contains additional factors [69]:

$$L_{AGN} = 16\pi R_{Sub}^2 f_{abs}^{-1} Q_{abs;P}(T_{sub}) \sigma_{SB} T_{sub}^4, \qquad [4]$$

where f_{abs} is the fraction of incident flux absorbed per dust grain and $Q_{abs;P}(T_{sub})$ is the absorption efficiency of the dust for a given sublimation temperature. Regardless we can bring the constant factors into a single parameter, and taking $R_{sub} / c = \tau_{sub}$ to trivially find that the luminosity should scale $L_{AGN} \propto \tau_{sub}^2$.

With a known bolometric luminosity LAGN and sublimation radius we should then expect to see a linear relation when these values are plotted logarithmically if AGN dust lags are indeed standard candles. This relation has been observationally supported by several papers [70], [71], [12], [15], [19], [42]. We should expect to see only a small scatter in the relation due to the aforementioned similar emission properties in the selected AGN. Previous studies have shown that the size of the BLR scales as the square root of the AGN luminosity in agreement with theory.

2.5 MAGNUM

The Multicolor Active Galactic Nuclei Monitoring (MAGNUM) project [72], [73] was conducted between 1995 and 2008 with preliminary observations beginning in 2001. Designed to carry out long-term observations of AGN in the optical and near-IR bands, the project is clearly useful for the purposes of dust reverberation mapping in our objects. The data was taken through a fully automated operation, involving the use of a single purpose 2m optical-infrared telescope and multicolour imaging photometer (MIP) attached to the former's Cassegrain focus, located at the Haleakala Observatory in Hawaii.

3.0 Experimental Method

3.1 Data Origin

For the purpose of the paper we study nine Type 1 AGN previously examined by Koshida et. al [15] and Lira et al. [18] with V-band magnitudes ranging from -15.75 to -22.23. The objects are relatively local with very low redshifts ranging from z~0.01 to z~0.025 and low luminosity. The timescales of the flux variations are thus relatively short producing light curves with numerous distinct features that allow for accurate cross-correlation analysis. It should be noted for posterity that recent study has indicated that Mrk 590 may not strictly be classed as a Type 1 with evidence indicating that it may in fact be a changing-look quasar with appearing and disappearing broad emission lines such that the AGN is able to change from a Type 1 to a Type 2 [74], [75]. At the time of observation, however, all nine galaxies were classed as Type 1, and this is a reasonable nomenclature to proceed with for the remainder of the paper. In addition to the Koshida et al. [15] dust reverberation study several of these objects have examined via BLR dust reverberation; NGC 4593 by Denney et. al. [76] and Onken et. al. [77], NGC 3516 by Onken et. al., NGC 3783 by Lira et al. [18] and the rest by Wandel et.al. [78] for the purpose of determining the central masses of the AGN.

The observations were initially taken using the Magnum MIP over the course of several years, with monitoring intervals for the objects configured to be in the range of several days [15], [72], [73]. With a field-of-view of 2.25 arcmin2 images were simultaneously obtained in the V and K bands by splitting the incident photon beam into two different detectors. The V band data contains the optical/UV continuum emission from the AGN accretion disk ignoring variable broad line emission contamination. The K band data contains the thermal near-infrared reemission from the innermost 'torus' region, where such emission dominates over the accretion disk continuum emission. These images are reduced via bias subtraction and flat fielding as standard. For NGC 3783 the J, H and K band data was taken over three years by the ANDICAM camera at CTIO, operated by the SMARTS consortium [18].

The relevant data then consists of K band flux taken over a long period of time for a given object, and V band flux in an overlapping epoch (see Appendix A). The V band flux consists of ten individual sets of simulated points – to obtain a better value for and understanding of τ_{peak} we will generate a CCF for each set. The flux initially contains a substantial contribution from stars in the host galaxies and thus is subtracted (see Table 4 in the appendices); these flux contributions have been previously calculated in studies by Sakata et al. [79], Suganuma et al. [12], and Koshida et al. [15] in the V band. These studies also calculated V band galaxy contributions for all AGN in the study, with the exception of AKN 120 which was provided by Tomita [80]. We can compare our results for the time delays

of these objects to those obtained by previous study in order to confirm that our findings agree with consensus and perhaps refine the accepted sublimation radii for the dust torus. We also use the V band magnitude of the objects, which ranges from M_V = 17.73 to -22.23, converted to luminosity using the relation:

$$M_V = -2.5 \log_{10} \frac{L_{AGN}}{L_0}, \quad [5]$$

where L_0 is the zero-point luminosity 3.0128×10^{28} W. We are assuming that the V band magnitude is equal to the bolometric magnitude of the AGN.

The first step is to induce a range of lags in the dataset; this is a simple transformation done by adding the lag value τ to all data points in a given band. We choose to induce the lag in the V band; this decision is trivial as should we induce the lag in the K band instead we will simply find a negative lag rather than a positive one. We sum over the number of data pairs N, e.g. the V and K fluxes at a given epoch. We thus must interpolate values of the flux such that we have flux values for both bands at any sampled time. In addition the cross-correlation function is only valid for epochs in which K and V overlap in time. We therefore require a function that will eliminate data outside this epoch of overlap and interpolate fluxes within it.

We choose a linear range of lags with a constant step of one Julian Day (JD) between -500 JD and 500 JD. A possible issue is that higher values for the AGN dust lags will be calculated using fewer points and thus is subject to greater errors; we will examine this and the object's deviation from our final fitted model in the discussion. Theory indicates that a negative time delay should not be possible to obtain; nonetheless we include the negative range for reasons that will become evident throughout the discussion. The primary reason is that inducing negative lags gives us a greater understanding of the overall shape of the cross-correlation function.

The custom function *matcharr*, which searches both data sets for matching timestamps and pairs them accordingly, attempts to sidestep the interpolation completely and requires no explicit search for an overlap period. However, an issue with this function is that data sets with a fixed offset, e.g. x=10,20,30 and y=11,21,31 will never be matched. We instead use *matcharr2*, which begins by truncating the data sets and uses linear interpolation to create fluxes for matching timestamps, which is the more appropriate choice. In our case we interpolate values in the K band, due to it being the band in which we have fewer data points.

3.2 Cross-correlation Analysis

In order to correctly determine the dust lag we use cross-correlation analysis for a range of different lags. The cross-correlation function gives a correlation coefficient for a given pair of time series – in our case this is the data taken in the V and K bands. Normalizing to obtain coefficients between 1.0 and -1.0 we use the function:

$$CCF(\tau) = \frac{1}{N} \sum_{i=1}^{N} \frac{[V(t_i-\tau)-\langle V \rangle][K(t_i)-\langle K \rangle]}{\sigma_V \sigma_K} \quad [6]$$

Where ⟨K⟩ is the average of the lag-induced K-band data K'(t-τ), ⟨V⟩ is the average of the V-band data V(t), N is the number of data 'pairs' which we divide over to achieve a normalized function and σ_K and σ_V are the respective standard deviations. The correct lag τsub = τpeak will be where CCF is highest CCF(τ_{peak}), which corresponds to the time taken for radiation to travel from the UV/optical source to the dust; Rsub = τsub.

Cross-correlation analysis has previously been used in a number of dust reverberation mapping studies, including both those involving BLRs and those involving the dust

continuum, to obtain values for sublimation radii consistent with theoretical results in a variety of AGN including quasars and Type 1 types; many of these studies have been referenced above and provide evidence that the technique will be appropriate moving forward [9], [10], [11], [12], [15], [18], [19].

3.3 Long-term Variability Subtraction

With the immediate method complete it is of interest to investigate factors that influence lag. The study of long-term variability in AGN is extensive and, due to the irrelevance of the theory for the purpose of this paper, shall not be covered in depth here.

Long-term variability in the flux across different observing periods could well alter the calculated lag if left unaccounted for. This long-term variability will make itself known in peaks across the cross-correlation function which may result in false values in sublimation radii – hence another reason to generate coefficients in the x-axis is to ease the detection of periodic and long-term variation in flux.

One possible method is to take an average flux value for each observing period and generate a spline from these average points. The spline can then be subtracted from each light curve to remove the long-term variability. Using the custom function *subspline* the data is searched for gaps in observations greater than a given length, in JD (we use a period of 90 JD). The data is then split by observing period (i.e. each split is made at an 'epoch break' of a set length of time; in our study we use an epoch break of JD) with an average flux value taken at each, with these averages generating a spline using the *scipy* interpolation function. The spline is then subtracted from the original data.

A different method would be to simply subtract the average flux in an observing period from every value in that period. The custom function *submu* works much the same as *subspline*, where the spline is disregarded and the averages are subtracted instead.

3.4 Finding a Peak

The question still remains of how to select the peak in the cross-correlation function. Even with subtraction of long-term variability false peaks would still occur due to coincidence in the V and K-band emission.

It would first be wise to consider the physically possible lags. Should the accretion disk emit in the V-band then it is obvious to assume that the K-band emission of the dust torus would be a translated V-band in the positive time direction, thus we examine only peaks for positive τ. As we translate the K-band over larger lags there will be fewer datum pairs to correlate – the aforementioned 'coincidence correlations' will be more prominent, and we should expect to see an increase in the number of peaks. We should assign greater value to peaks towards $\tau = 0$.

The custom function *exlag* detects the peak in a cross-correlation function, examining the function in either the positive or negative x-axis and determining viable candidates based on coefficient value and peak prominence. It then selects the top candidate by choosing the value with minimum x – this is the τpeak' for the corresponding simulated V-band. The benefit then of including negative lags in our cross-correlation function is that we maintain the full shape of the peaks towards $\tau = 0$ and the code is less likely to ignore viable candidates. We then average these values to get our τpeak for the AGN.

Examining the ten simulated peaks we find that the averages are liable to be skewed heavily by outlying results – take for example NGC 4593, where the omission of a single data point changes the average τ_{peak}

from 48.7 days to 35.2 days. It may be wise to discard the heavy outliers and proceed with the new average time delays. We will define an outlier as any delay greater than the 1.5 interquartile ranges below the first quartile or above the third. In the discussion we will compare results between the original set of time delays and the refined set with discarded values.

We take the luminosity and lags and find their logarithms to the base 10, and now attempt to fit a model to the data.

3.5 The Ensemble Sampler

The simplest way to fit a linear model to our data is through least squares analysis. A problem arises through this method – the uncertainties in the model parameters; the gradient m and intercept c, are rather small owing to the small uncertainties in our time lags – the least squares solution assumes that the error bars are correct. It would be wise to discard this assumption, due to an inability to discard long-term variability and propagation of the unknown errors arising from the interpolation. An alternate method would be to use the *polyfit* function included in the *numpy* package; once again however this does not always provide reasonable uncertainties. We instead use a combination of the scipy minimize function and an ensemble sampler to find the parameter values at maxima of the parameter likelihood function, namely the slope and offset of our linear fit and the fraction of the variance underestimation f [81], [82]. The ensemble sampler requires construction of a prior that encodes all known information about the parameters; we cover the construction in the discussion.

Several other studies have examined the slope of the lag – luminosity relation. Minezaki et al. [42] found the respective values of the slope and offset to be 0.424 ± 0.026 and -1.021 ± 0.023 through dust reverberation analysis. The GRAVITY collaboration found a similar slope value of 0.40 ± 0.04 [71]. From near-IR interferometry studies the value of the slope has been determined to have a value of 0.38.

4.0 Discussion

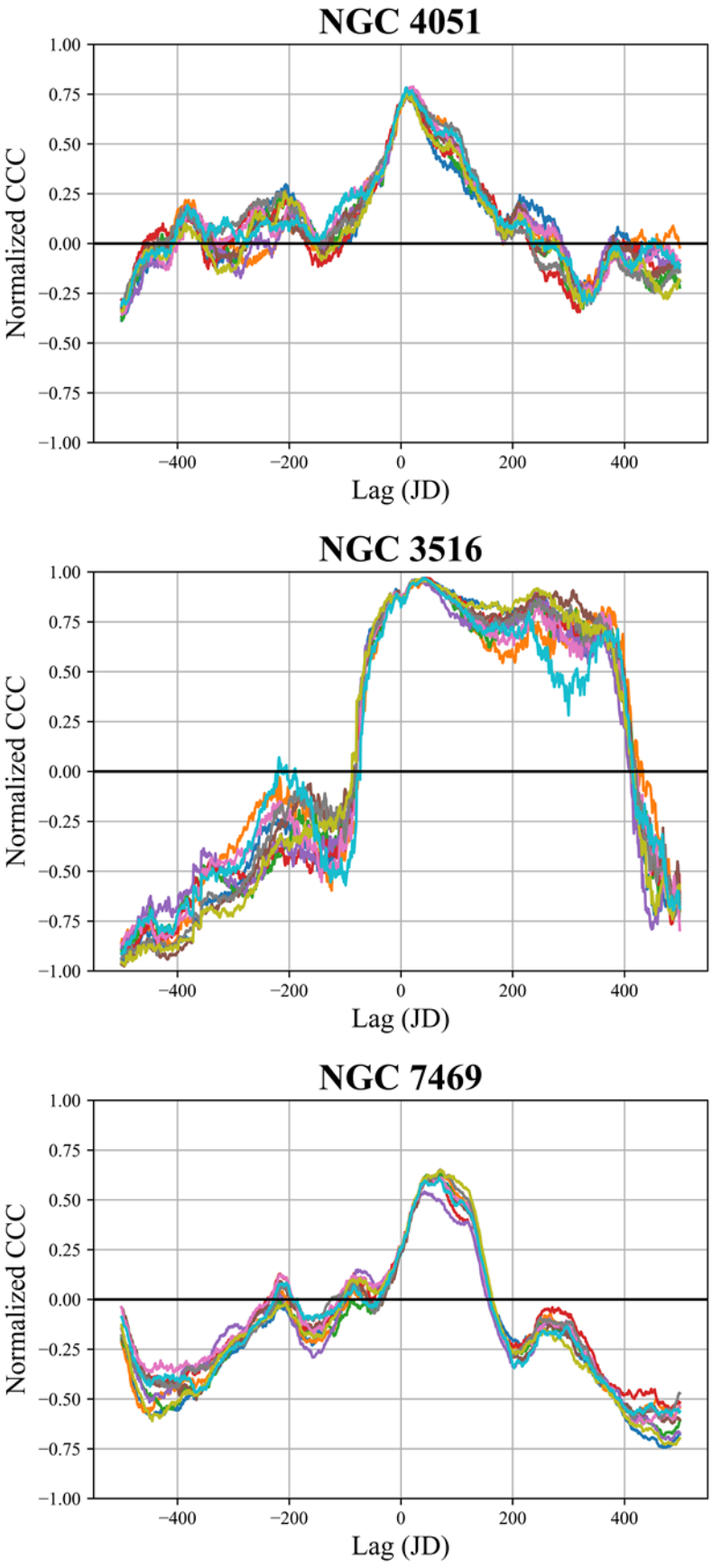
4.1 Cross Correlation Analysis

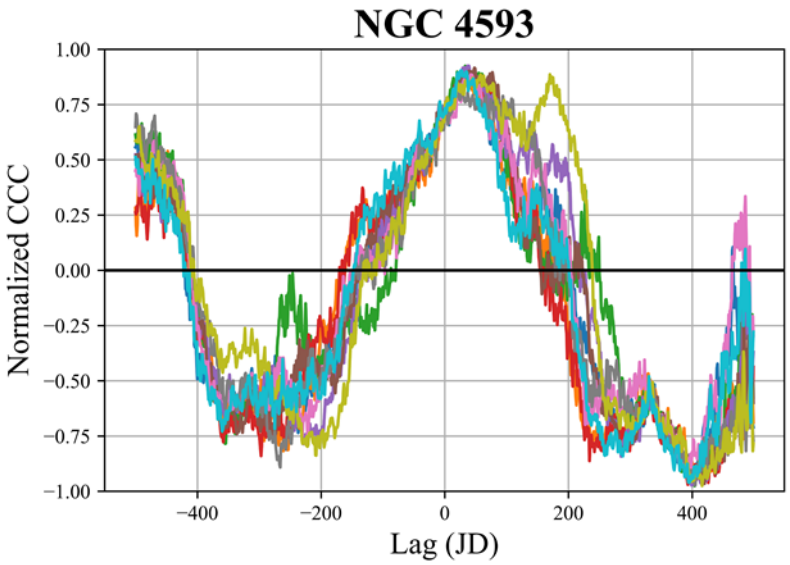
The peaks in the cross-correlation functions obtained for each of the ten respective simulated V band light curves of the AGN is displayed in Table 5 (see appendices). We find the lags to initially be in the range of 13.5 JD to 90.5 JD, with a moderate average fractional standard deviation (0.343), which extends as high as 0.840 in NGC 4593. The average cross-correlation coefficients can be said to be fairly strong, with the lowest of 0.524 for Mrk 817, and the next lowest at 0.618 for NGC 7469, but a mean over all AGN of 0.788 (and a corresponding fractional standard deviation of 0.0395).

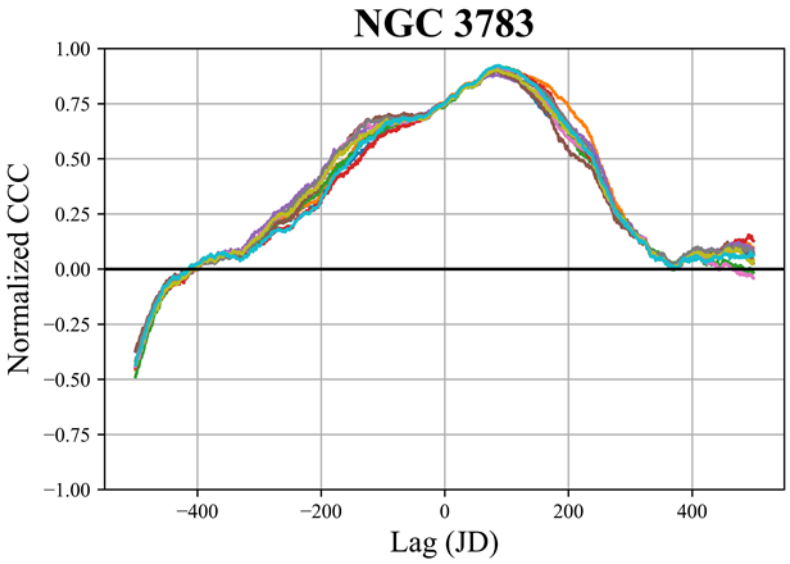
Removing the outliers using the method outlined above presents a more accurate picture of the results (Table 1 and Figure A). A total of five outliers were removed; one each from Mrk 817, NGC 3516 and NGC 4593 and two from Mrk 590, which demonstrably lead to small changes in the mean dust lag value but much smaller standard deviations, decreasing the average across all AGN to 0.209.

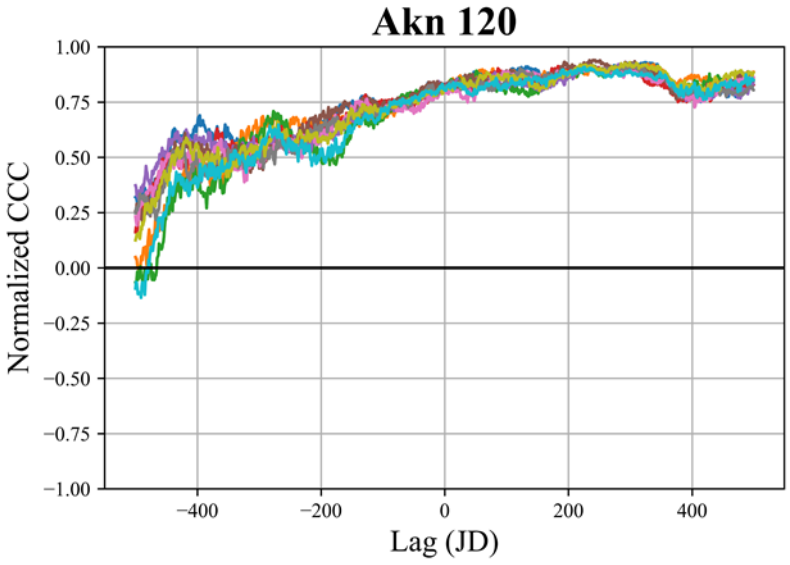
For the purpose of an in-depth discussion we examine four AGN with distinct cross-correlation functions given in Figure A in greater detail – Akn 120, Mrk 509, Mrk 590 and Mrk 817. The foremost has a cross-correlation function marked by a distinct lack of prominent peaks or the erratic variation we would expect to see at either end of the x-axis. The cause of this is not readily apparent upon inspection of the light curves; similar variations in both can be observed that should naturally translate to a typical CCF.

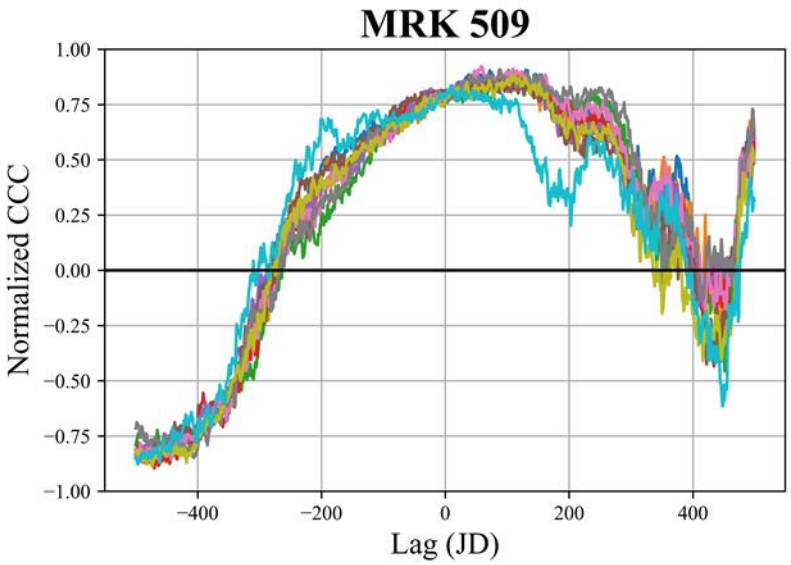
Figure A.
(pg.11,12,13) Cross correlation functions for the AGN sample. The multiple peaks in Mrk 817 are indicative of regular periodic variation. Removal of long-term variation in flux would perhaps aid in reducing the peak width of Mrk 509.

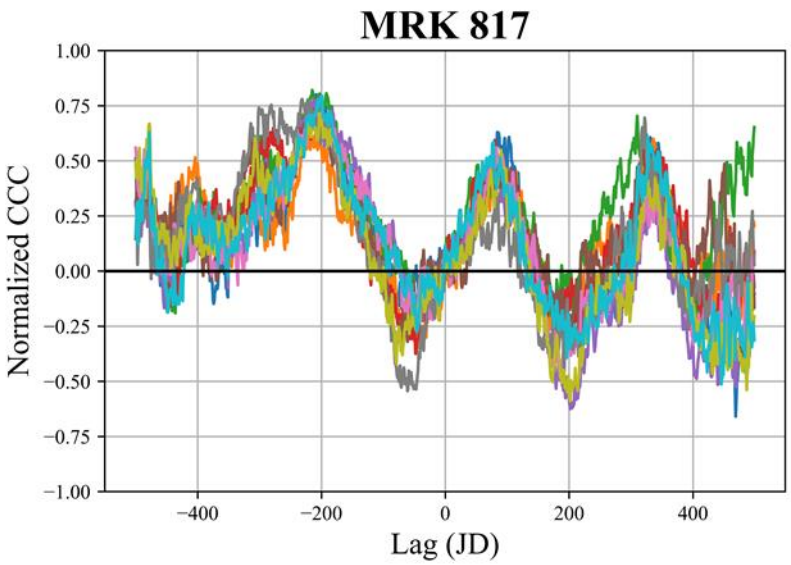


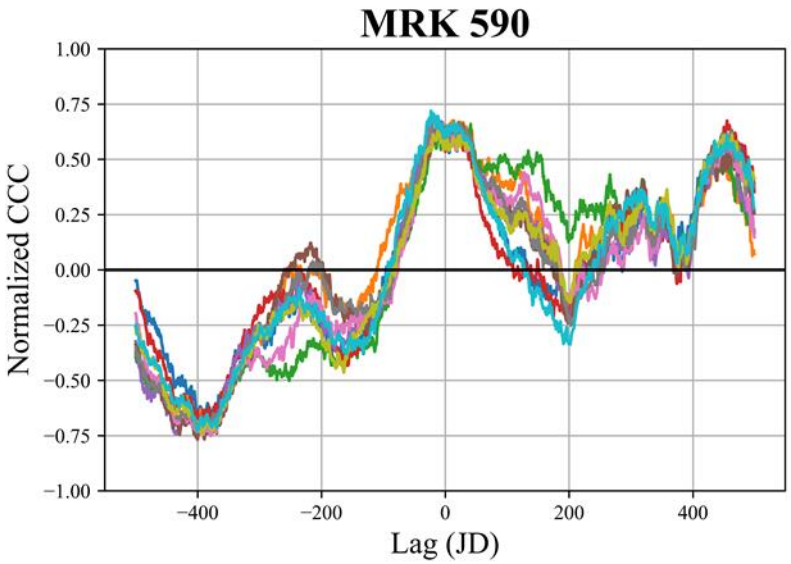

NGC 4593
Normalized CCC
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−400
−200
0
200
400
Lag (JD)


NGC 3783
Normalized CCC
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−400
−200
0
200
400
Lag (JD)


Akn 120
Normalized CCC
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−400
−200
0
200
400
Lag (JD)


MRK 509
Normalized CCC
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−400
−200
0
200
400
Lag (JD)


MRK 817
Normalized CCC
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−400
−200
0
200
400
Lag (JD)


MRK 590
Normalized CCC
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−400
−200
0
200
400
Lag (JD)

OBJECT	PEAKS										MEAN	STDEV
	1	2	3	4	5	6	7	8	9	10		
NGC 4593												
CCF	0.884	0.903	0.924	0.915	0.925	0.916	0.885	0.837	0.886	0.904	0.899	0.0261
Dust Lag (JD)	31	39	33	34	37	48	42	27	170*	26	35.2	6.7
Mrk 590												
CCF	0.633	0.705	0.68	0.619	0.675	0.645	0.639	0.683	0.615	0.669	0.656	0.0287
Dust Lag (JD)	29	28	28	1*	25	29	27	27	1*	28	22.3	10.71
NGC 3516												
CCF	0.97	0.96	0.965	0.968	0.958	0.963	0.966	0.966	0.964	0.969	0.966	0.00294
Dust Lag (JD)	44	45	42	46	21*	44	37	36	36	40	41.1	3.75
NGC 4051												
CCF	0.759	0.773	0.771	0.76	0.761	0.749	0.788	0.758	0.749	0.783	0.765	0.0126
Dust Lag (JD)	11	18	20	9	12	9	22	14	10	10	13.5	4.57
NGC 7469												
CCF	0.611	0.648	0.63	0.611	0.542	0.609	0.614	0.651	0.65	0.61	0.618	0.0303
Dust Lag (JD)	72	70	72	44	42	71	52	70	72	68	63.3	11.63
Mrk 817												
CCF	0.63	0.468	0.551	0.593	0.435	0.518	0.56	0.35	0.549	0.585	0.534	0.0781
Dust Lag (JD)	84	87	76	80	65*	85	83	85	90	78	83.1	4.18
Mrk 509												
CCF	0.907	0.896	0.906	0.892	0.871	0.856	0.923	0.908	0.882	0.836	0.888	0.0254
Dust Lag (JD)	130	124	102	103	112	65	58	89	113	9	90.5	35.1
NGC 3783												
CCF	0.909	0.915	0.92	0.912	0.882	0.898	0.915	0.903	0.903	0.924	0.908	0.0116
Dust Lag (JD)	83	81	84	85	81	83	85	86	86	86	84	1.84
Akn 120												
CCF	0.862	0.884	0.866	0.83	0.868	0.858	0.892	0.839	0.873	0.882	0.866	0.0185
Dust Lag (JD)	67	61	51	19	81	45	133	54	126	180	81.7	46.8

Table 1. *(above) Values of the peak cross-correlation coefficient for each simulated V band and the corresponding dust lag value. The mean peak values and their standard deviations are given on the rightmost columns. Outliers that have not been included in the calculation are marked with *. For mean values including the outliers, see Table 5 in the appendices.*

A possible reason then is that these variations are not great enough for the cross-correlation function to be able to detect correlation. With a mean dust lag of 81.7 JD and a corresponding standard deviation of 46.8 JD the results obtained for Akn 120 seem unlikely to contribute to any useful conclusion. If the long-term variability subtraction is unable to yield more successful results we will have to discard the object. Conversely the cross-correlation function of Mrk 509 appears to consist of one very

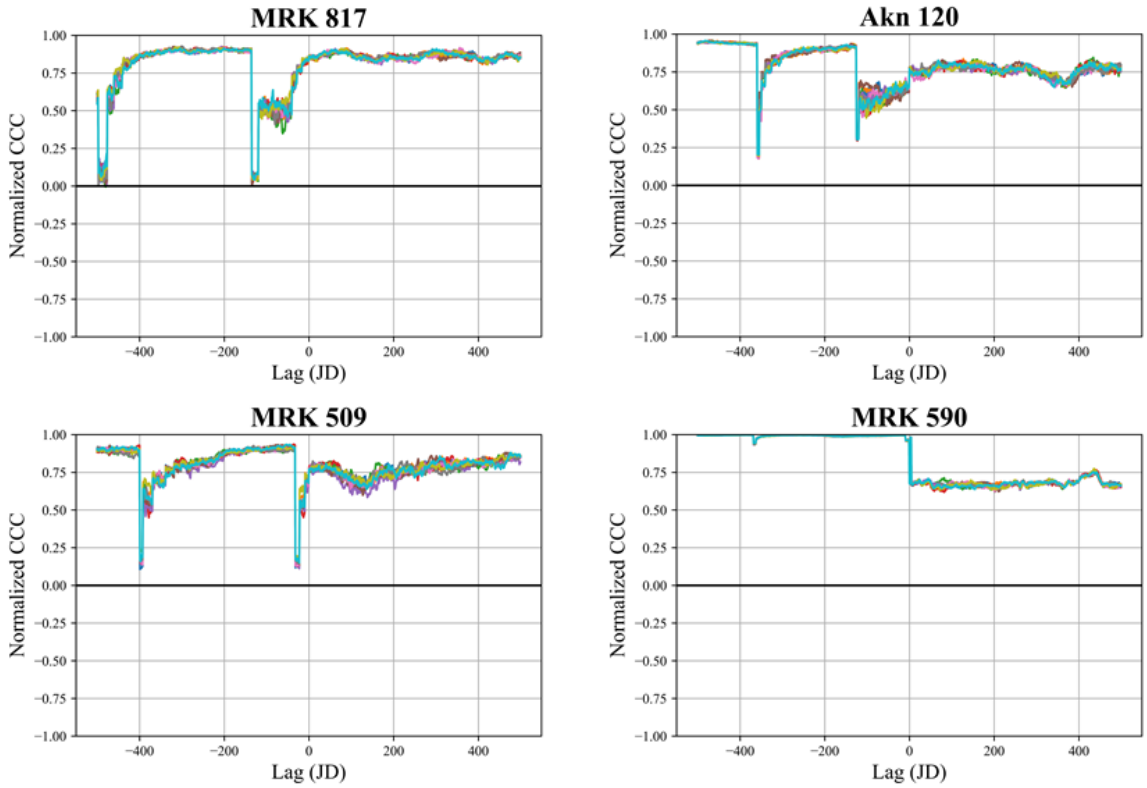


Figure B. *(above) Cross-correlation functions for the mean-subtracted light curves for selected objects.*

broad peak, with strong anticorrelation at one x-axis boundary and a spike in correlation at the other. Once again we have a strong standard deviation in the results; 35.1 JD for a mean of 90.5 JD – this is likely due to the broadness of the peak. Mrk 590 shows a narrow peak that plateaus at the top, where all the dust lags obtained lie. Discarding the outliers pushes the dust lag to a higher value away from the origin and decreasing the standard deviation to a much smaller value. Finally, Mrk 817 shows three peaks of similar full width at half maximum, evenly spaced, indicating a periodic variation The peak fitting method, which prioritises the lower-lag peaks, makes selecting the correct peak trivial and we obtain a mean value of 83.1 JD with a low standard deviation (4.18 JD).

4.2 Long Term Variability Subtraction

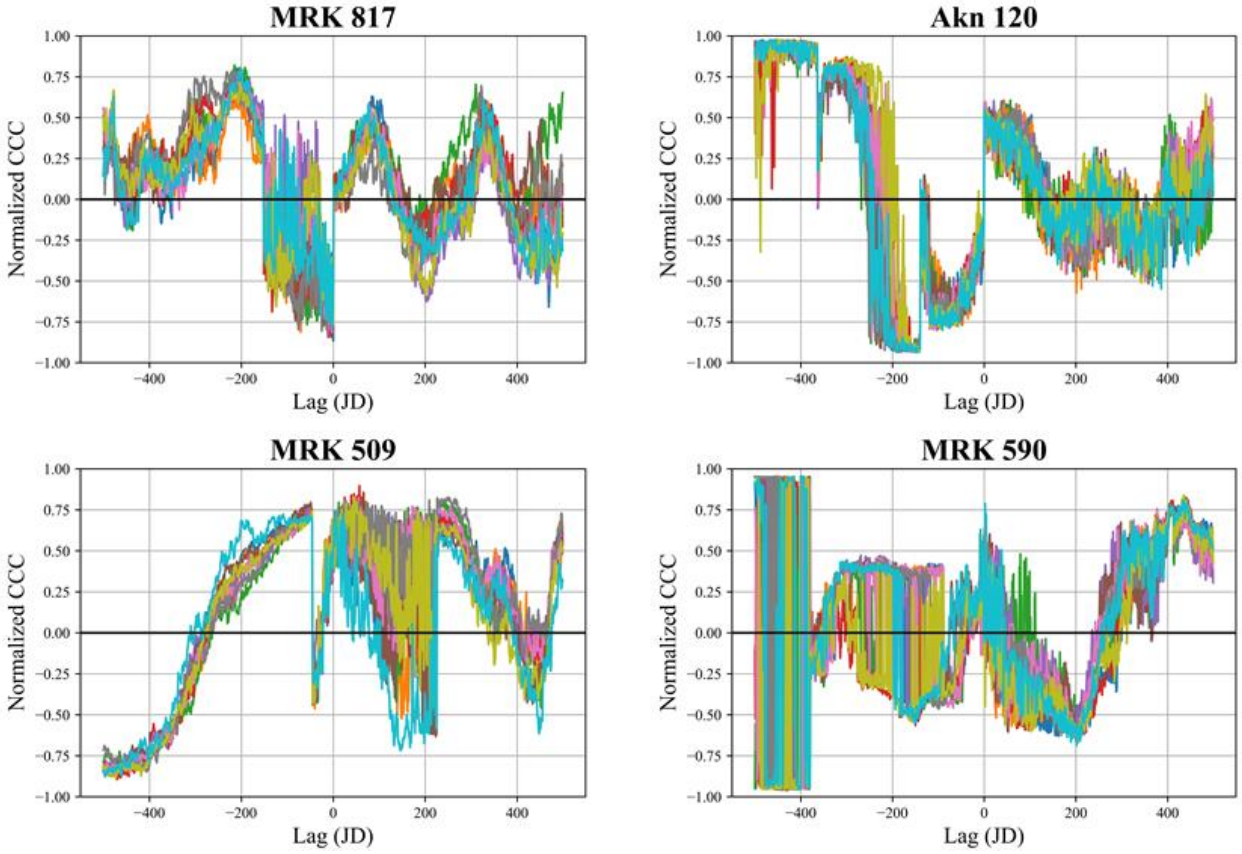
Performing the same analysis on the mean-subtracted light curves leads to a greater number of discarded results – two each in NGC 4593, NGC 3516, NGC 4051 and Mrk 817, with three discarded from NGC 3783 for eleven in total (see Table 6 in the appendices). We note immediately from Figure B that the functions have lost many of their expected c characteristics; prominent peaks and sharp variation at the boundaries have been replaced by constant semi-regular behaviour. The most notable feature that all functions possess are sharp, often short drops in the cross-correlation coefficient. These are likely due to short observing periods where an accurate value of the average flux is difficult to obtain. It is then interesting to note that the results obtained through this method are more consistent than those obtained previously, with an average fractional standard deviation of 0.158. Comparing the mean results from both methods results in remarkable consistency between the dust lags for NGC 3516, NGC 7469, NGC 3783 and Mrk 817. Once again the mean result for Akn 120

carries a large fractional standard deviation, as does Mrk 509.

Analysis of the spline-subtracted light curves (Figure C) begins with examination of the erratic behaviour of the respective cross-correlation functions. Due to the spline generation requiring a certain number of detected observing periods several AGN have not undergone complete spline subtraction; notably NGC 3516 and NGC 4593 have not undergone any spline subtraction whatsoever. Spline subtraction was only able to be fully carried out on Mrk 590, NGC 4051, NGC 7469, NGC 3783 and Akn 120, and the mean dust lags for these AGN carry the highest fractional standard deviations. Despite this, only eight results are discarded as outliers due to the large interquartile ranges, and the dust lags for NGC 7469 are consistent with our original results (see Table 7 in the appendices).

It is worth comparing our dust lags to those obtained by Koshida et al. [15] and Lira et al. [18]. Through cross-correlation analysis the Lira et al. [18] study found the dust lag for NGC 3783 using the B and K bands to be 76.3 JD; a ~6% percentage difference from our own result (it is worth noting that NGC 3783 was marked by high cross-correlation coefficients and low dust lag standard deviations). For the rest of the AGN we compare our results to Koshida et al. [15], who determined dust lags through both cross-correlation analysis and the JAVELIN software developed by Zu et al. [58] Our dust lags are close to consistent with those obtained through the cross-correlation method as expected, with the exception of NGC 4593 and NGC 7469 which show more agreement with the JAVELIN method.

Figure C. *(below) Cross-correlation functions for the spline-subtracted light curves for selected objects.*



Mrk 509 additionally shows little agreement with any results from the study. Our spline-subtracted dust lags show even less agreement, with only half the AGN (Mrk 590, NGC 3516, NGC 4593 and Mrk 817) having any consistency; two of which have not undergone any spline subtraction with only one undergoing complete spline subtraction. For our average-subtracted dust lags the results are inconsistent with any results obtained in previous studies, save for possibly NGC 7469 and Akn 120. The latter shows remarkable consistency considering a previous inability to match our results.

Due to the inability to obtain a complete and cohesive set of results for spline-subtracted light curves we are forced to admit that the method is in need of improvement and discard the results. An alternate methodology should aim to improve the method of peak-fitting to account for the sharp variations in the cross-correlation function and develop a method for more precise observation epoch detection. While the mean-subtracted light curves result in cross-correlation functions that initially seem to have few discernible features we are nonetheless able to obtain consistent dust lags that in some cases agree with our initial results.

With an inability to reduce the error in the dust lag for Akn 120 in any case we should consider discarding the object from the study. Mrk 509 may also be considered a candidate for discard. None of the results obtained agree with those determined in previous studies [15] and contain large errors. The primary difference between our results for Mrk 509 and Akn 120 is that the initial cross-correlation curve contains a peak, albeit a broad one. For now we will retain both objects and examine its contribution to our final result.

OBJECT	MAGNI-TUDE	L_{AGN} (e35 Ws^{-1})	DUST LAG (J.D.)
NGC 4593	−17.93 ± 0.04	4.477	35.2 ± 6.70
Mrk 590	−18.69 ± 0.08	9.015	27.6 ± 1.22
NGC 3516	−18.14 ± 0.09	5.432	41.1 ± 3.75
NGC 4051	−15.75 ± 0.09	0.601	13.5 ± 4.57
NGC 7469	−19.79 ± 0.06	24.83	63.3 ± 11.63
Mrk 817	−20.86 ± 0.02	66.52	83.1 ± 4.18
Mrk 509	−22.13 ± 0.03	214	90.5 ± 35.1
NGC 3783	-19.09	13	84.0 ± 1.84
Akn 120	−22.23 ± 0.02	234.9	81.7 ± 46.8

Table 2. *(above) Magnitudes, luminosities and dust lags of the AGN sample.*

4.3 Ensemble Sampling

We find the logarithm to the base 10 of the dust lags and the AGN luminosity (see table 2), assuming symmetric errors for the dust lags. Observing the plotted results in Figure G for our spline-subtracted and mean-subtracted light curves it is clear that discarding the methodology was the correct choice as any relation between dust lag and AGN luminosity appears to have been removed. Using the numpy polyfit function to obtain starting parameters we find a slope m of 0.296 and an offset b of -9.02. We set the range of our prior using this and theoretical result to be $-2.0 < m < 0.0$ and $-15.0 < b < 0.0$, taking into account extremities and observation of the plotted data in Figure E.

With no estimation of f we run the ensemble sampler with a prior range of -20.0

Parameter	Polyfit	Minimized	Sampled
Slope	0.296	0.327	$0.324^{+0.0775}_{-0.0741}$
Offset	-9.02	-10.1	$-10.0^{+2.68}_{-2.67}$
$\log_{10} f$	N/A	-2.57	$-2.43^{+0.329}_{-0.287}$

Table 3. *(above) Parameter values determined through numpy.polyfit, scipy.minimize and the emcee ensemble sampler.*

Figure D. *(below) Corner plot of the sampled parameters for the dust lag-luminosity relation. Starting parameters obtained from the scipy.minimize function are indicated in blue.*

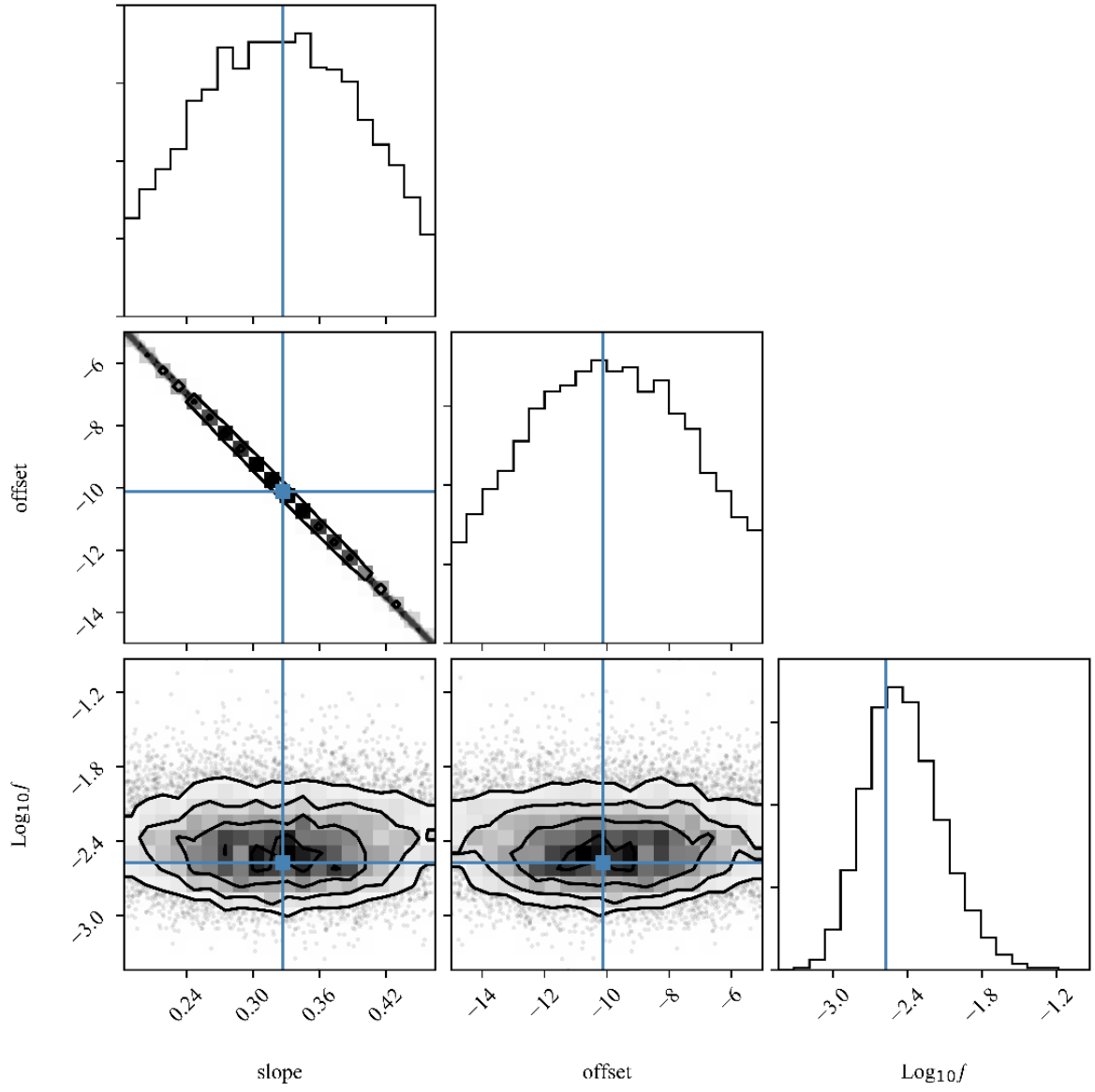
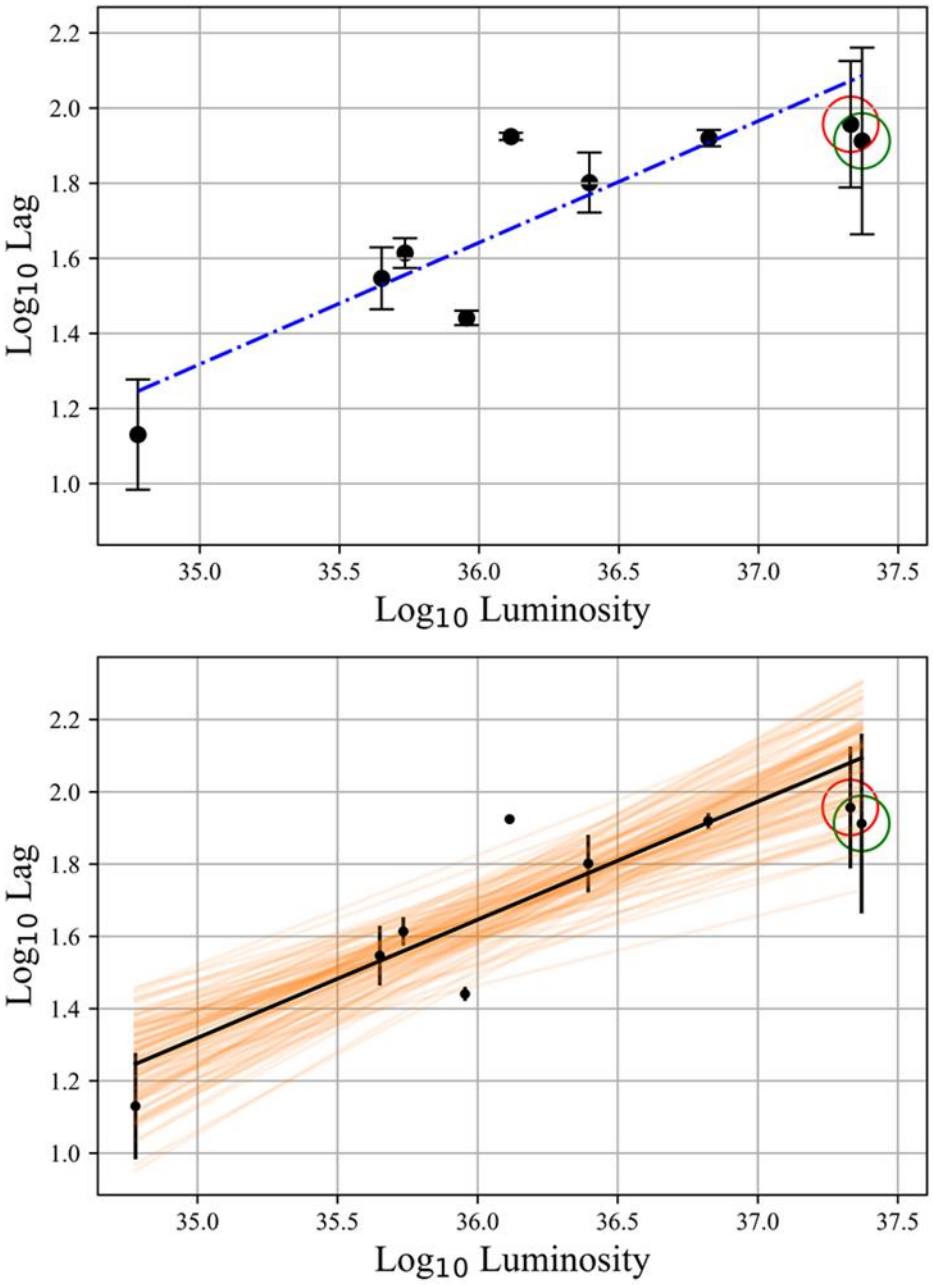


Figure E. *(below) Logarithmic plot of the dust lag-luminosity relation (in Watts and JD respectively). Linearly fitted model using best-fit parameters obtained by the ensemble sampler is shown top, in blue. Additional samples are plotted bottom. Outlier candidates Mrk 509 and Akn 120 are marked red and green respectively.*



$< \log_{10} f < 20.0$ initially and refine the range based on our result. For our first run of the sampler we then find a $\log_{10} f$ value of -11.1, and so refine the range to $-20.0 < \log_{10} f < 0.0$. Now running the sampler with the appropriate prior we find for our original set of dust lags a slope of 0.324 and an offset -10.0 with $\log_{10} f$ -2.43 (see table 3). A corner plot of the samples can be found in Figure D.

Running the ensemble again with the same prior but with the data for Mrk 509 and Akn 120 omitted results in very little change to any of our parameters or their respective errors. As we see in Figure E, with additional samples projected onto the plot, the data for

Mrk 509 and Akn 120 lie in an acceptable range for the model. We therefore decide against discarding the AGN and will include it in analysis of our fitted model.

Quoting the errors in Table 3 based on the 16th, 50th and 84th percentiles we find the parameter results to be reasonably accurate; average errors of ± 0.0758, ± 2.68 and ± 0.308 for the slope, offset and $\log_{10} f$ respectively. When compared to previous study [15], [42], [71], our values are consistent; all previous values for the slope lie within standard error ranges and we can state with confidence that our methodology is as accurate as those previously used. Compared to the expected theoretical result however we find our slope to be roughly 40% lower. While fault may lie with the methodology there are a number of additional theoretical and experimental factors to consider. A primary issue is the number of AGN used and their common properties; these are low luminosity Type 1 galaxies, two of which we have discussed as potential outliers, and hence only represent a small sample of the complete AGN class. Conducting a study using a larger sample of AGN may go some way in confirming the cause of the discrepancies, as well as one with larger observing periods so that we can sample multiple peaks for each object and determine dust lags with greater accuracy. It has been shown that anisotropic illumination of the V band region may result in differences between the equatorial and polar sublimation radii [35], [83]. In addition, studies of the SED luminosity dependence of AGN have shown that more luminous AGN may show a bluer optical continuum compared to lower-luminosity AGN, which show a far-UV continuum [84]. Studies may need to be conducted that account for this dependence when attempting cross-correlation analysis to further understand the behaviour.

It has been found that while upon a drop in luminosity the inner dust torus will contract this contraction does not occur immediately [85], [86], [87]. This can result in observations of dimmed accretion disks with larger than expected dust torii. For our low-luminosity AGN in which flux variations occur on smaller timescales this could explain the shallower slope, and reinforces the need for study with more luminous AGN.

With regards to the methodology an inability to remove the long-term variability may result in contributions to the cross-correlation function that affected the obtained dust lags. We have also failed to account for uncertainties in the subtracted host galaxy fluxes and the AGN absolute magnitude which may propagate to reduce the value of the slope.

5.0 Conclusions

In this study we have examined the dust lag-luminosity relation for nine active galactic nuclei previously studied by Koshida et al. [15] and Lira et al. [18] We have first collated and laid out the theoretical framework for our model, initially through explanation of the basic structure of AGN, classification and the unification model accepted by the majority of the astronomical community, in addition to justifying our proposal of using AGN dust lags as standardizable candles. The methodology and history of dust reverberation mapping has been discussed, as well as advantages and disadvantages both on its own and in comparison to alternative methods of AGN reverberation mapping. We finish examination of the theoretical aspect of the study through discussion of the expected lag-luminosity relation, both through physical theory and previous studies conducted regarding this topic.

We then carried out cross-correlation analysis via the inducing of a range of lags in ten simulated V band data sets and

corresponding K band data, which originate as emission from the AGN accretion disk and the corresponding dust torus reemission regions respectively, using custom python functions. The use of cross-correlation analysis is justified through its use in numerous previous studies of AGN reverberation mapping, many of which have been referred to throughout the paper. Two possible methods of subtracting long-term AGN variability have been examined and attempted on our data to obtain better values for our sublimation radii, with little success. Extracting the dust lags from the peaks in our cross-correlation functions we proceeded to remove outliers and found magnitudes for our objects. On inspection of our final data we found it acceptable to retain Akn 120 for the final step of the study as the average dust lag showed good agreement with the model, as well as Mrk 590.

Proceeding with minimization of our finalised data we find an initial slope of x and offset y for a linearly fitted model. Constructing a prior function from the theoretical framework we then use ensemble sampling to refine these parameters, concluding the following:

i) The linearly fitted model is best parametrised by a slope of 0.324 and an offset of -10.0, with the logarithmic underestimation of the data $\log_{10} f$ -2.43. The small errors ± 0.0758, ± 2.68 and ± 0.308 respectively and suggest an accurate model with respect to the dataset.

ii) Our slope lies just outside a value that would be wholly consistent with the theoretical relation L -> t1/2. The value is however consistent with those obtained by previous studies [15] [42] [71], suggesting that there may be additional factors to consider in theory some of which have been outlined in our discussion; namely torus contraction delay, SED luminosity dependency, and anisotropic illumination of the V band emission region.

This study and future similar studies may be improved by an increased AGN sample size that includes a greater percentage of high-luminosity objects. Additional steps in the methodology may be attempted, such as a successful method to subtract long-term variability or accounting for redshift of the more distant of the sampled AGN. Despite these potential shortcomings the evidence given in this paper provides a clear path moving forward in the theoretical and experimental study of AGN dust lags, and indicates that the dust lags may be used as standardizable candles.

ACKNOWLEDGEMENTS

We would like to thank Sebastian F. Hoenig, Samuel Williamson and Michal Wypych for their assistance in the study. We would also like to thank Jessica and Spencer Walls for their assistance during the writing of the report.

References

[1] Fabian A. C., 2012, Annual Review of Astronomy and Astrophysics, 50, 455-489

[2] Almeyda T.; Robinson A.; Richmond M.; Nikutta R.; McDonough B., 2020, Astrophysical Journal, 891

[3] Ferrarese L.; Ford H., 2005, Space Science Reviews, 116, 523-624

[4] Watson D.; Denney K. D.; Verstergaard M.; Davis T. M.; 2011, Astrophysical Journal, 740

[5] Hönig S. F., Astrophysical Journal, 2014, 784

[6] Hönig S. F.; Watson D.; Kishimoto M.; Hjorth J., 2014, Nature, 515, 528

[7] Elvis M.; Karovska M.; 2002, Astrophysical Journal, 581

[8] Horne K.; Korista K. T.; Goad M. G. 2003, MNRAS, 339, 367-368

[9] Hönig S. F., 2019, Astrophysical Journal, 884

[10] Shablovinskaya E.S.; Afanasiev V.L.; Popvic. L.C., 2020, Astrophysical Journal, 892

[11] Netzer H., 2015, Annual Review of Astronomy and Astrophysics, 53, 365-408

[12] Suganuma M. et al., 2006, Astrophysical Journal, 639, 46-63

[13] Shen Y. et al., 2016, Astrophysical Journal, 818, 30

[14] Lyu J.; Rieke G.H.; Smith P.S., 2019, Astrophysical Journal, 886

[15] Koshida S. et al., 2014, Astrophysical Journal, 788

[16] Kishimoto M.; Hönig S. F.; Antonucci, R.; Millour F.; Tristram K.R.W.; Weigelt G., 2009, Astronomy and Astrophysics, 507

[17] Kishimoto M.; Hönig S. F.; Antonucci, R.; Millour F.; Tristram K.R.W.; Weigelt G., 2011, Astronomy and Astrophysics, 536

[18] Lira P.; Arevalo P.; Uttley P.; McHardy I.; Breedt E., 2011, MNRAS, 415, 1290-1303

[19] Hönig S.F. et al., 2016, MNRAS, 464, 1693-1703

[20] Urry M. 2003. In Active Galactic Nuclei: From Central Engine to Host Galaxy, ed. S Collin, F Combes, I Shlosman. ASP Conf. Ser. 290:3. San Francisco: ASP

[21] Mo H.; van den Bosch F.; White S.; Galaxy Formation and Evolution (Cambridge University Press, New York 2010), 14, 618-639

[22] Seyfert C. K., 1943, Astrophysical Journal, 97

[23] Peterson B. M., An Introduction to Active Galactic Nuclei (Cambridge University Press, New York 1997)

[24] Urry C. M., 1995, Publications of the Astronomical Society of the Pacific, 107, 803-845

[25] Fanaroff B. L.; Riley J. M., 1974, MNRAS, 167

[26] Wilson A. S.; Pooley G. G.; Willis, A. G.; Clements E. D; 1980, Astrophysical Journal, 237

[27] Ulvestad J.S., 2015, Astrophysical Journal, 278, 544-557

[28] Antonucci R., 1993, Annual Review of Astronomy and Astrophysics, 31, 473-521

[29] Hoenig S. F., 2013, ArXiv e-prints, arXiv:1301.1349

[30] Almeyda T.; Robinson A.; Richmond M.; Vazquez B.; Nikutta R., 2017, Astrophysical Journal, 843

[31] Krolik J. H.; Begelman M. C., 1988, Astrophysical Journal, 329, 702

[32] Burtscher L.; Jaffe W.; Raban, D.; Meisenheimer K.; Tristram K. R. W.; Rottgering H., 2009, Astrophysical Journal, 705

[33] Markowitz A. G.; Krumpe M.; Nikutta R., 2014, MNRAS, 439, 1403

[34] Stalevski M.; Fritz J.; Baes M.; Nakos T.; Popovic L. C., 2012, MNRAS, 420, 2756, 2756

[35] Kawaguchi T.; Mori M., 2011, Astrophysical Journal, 737, 105

[36] Gallimore J. F. et al., 2016, Astrophysical Journal, 829

[37] Combes F. et al., 2019, Astronomy and Astrophysics, 623

[38] GRAVITY Collaboration et al., 2019, Astronomy and Astrophysics, 634

[39] Tristram K. R. W. et al., 2014, Astronomy and Astrophysics, 563

[40] Lopez-Gonzaga N.; Jaffe W.; Burtscher L.; Tristram K. R. W.; Meisenheimer K., 2014, Astronomy and Astrophysics, 565

[41] Vollmer B. et al., 2018, Astronomy and Astrophysics, 615

[42] Minezaki T. et al., 2019, Astrophysical Journal, 886

[43] Penston M. V.; Penston, M. J.; Neugebauer G.; Tritton K. P., Becklin E. E., Visvanathan N., 1971, MNRAS, 153, 29-40

[44] Glass I. S., 2004, MNRAS, 350, 1049

[45] Rees M. J.; Silk J. I.; Werner M. W.; Wickramasinghe N. C., 1969, Nature, 223, 788

[46] Kobayashi Y.; Sato S.; Yamashita T.; Shiba H.; Takami H., 1993, Astrophysical Journal, 404, 94

[47] Landt H. et al., 2011, MNRAS, 413

[48] Deo R. P., 2011, Astrophysical Journal, 729

[49] Mor R.; Netzer H.; Elitzur M., 2009, Astrophysical Journal, 705

[50] Barvainis R., 1987, Astrophysical Journal, 320, 537

[51] Phinney E. S.; Duschl W.; Meyer F.; Frank J.; Meyer-Hofmeister E., 1989, Theory of Accretion Disks, 42, 457

[52] Kishimoto M.; Hönig S. F.; Beckert T.; Weigelt G., 2007, Astronomy and Astrophysics, 476, 713

[53] Blandford R. D.; McKee C. F., 1982, Astrophysical Journal, 255, 419

[54] Peterson B. M. et al., 2004, Astrophysical Journal, 613, 682-699

[55] King A. L.; Davis T. M.; Denney K. D.; Vfestergaard M.; Watson D., 2014, MNRAS, 441, 3454

[56] King A. L. et al., 2015, MNRAS, 453, 1701-1726

[57] Kovacevic A.; Popovic L. C.; Shapovalova A. I.; Ilic D.; Burenkov A. N.; Chavushyan V. H., 2014, Advances in Space Research, 54

[58] Zu Y.; Kochanek C. S.; Peterson B. M, 2010, Astrophysical Journal, 735, 80

[59] Zhang X. G., 2013, MNRAS, 431

[60] Netzer H.; Peterson B.M.; Maoz D.; Sternberg A.; Leibowitz E. M., 1997, Astronomical Time Series, 85

[61] Rokaki E.; Boisson C.; Coullin-Suffrin S., 1992, Astronomy and Astrophysics, 253, 57

[62] Wandel A.; Peterson B. M.; Malkan M. A., 1999, Astrophysical Journal, 526

[63] Wandel A. et al., 1998, Structue and Kinematics of Quasar Broad Line Regions (23-26 March 1998, Lincoln, Nebraska)

[64] Salpeter E. E., 1977, Annual Review of Astronomy and Astrophysics, 15, 267-293

[65] Mor R.; Netzer H.; Elitzur M.; 2009, Astrophysical Journal, 705, 298

[66] Mor R.; Netzer H., 2012, MNRAS, 420, 526

[67] Yoshii Y.; Kobayashi Y.; Minezaki T.; Koshida S.; Peterson B. A., 2014, Astrophysical Journal, 784

[68] Oknyansky V. L.; Horne K., 2001, ASPC, 224, 149

[69] Hönig S. F.; Kishimoto M., 2011, Astronomy and Astrophysics, 534, 121

[70] Weigelt, G. et al., 2012, Astronomy and Astrophysics, 541

[71] GRAVITY Collaboration et al., 2020, Astronomy and Astrophysics, 635

[72] Kobayashi Y.; Yoshii Y.; Minezaki T., 2004, Astronomische Nachrichten, 325

[73] Kobayashi Y. et al., 1998, Astronomical Telescopes and Instrumentation, Kona Hawaii, 3352

[74] Mathur S. et al., 2018, Astrophysical Journal, 866

[75] Denney K. D. et al., 2014, Astrophysical Journal, 796

[76] Denney K. D. et al., 2006, Astrophysical Journal, 653, 152

[77] Onken C. A.; Peterson B. M.; Dietrich M.; Robinson A.; Salamanca I. M., 2003, Astrophysical Journal, 585, 121

[78] Wandel A.; Peterson B. M.; Malkan M. A., 1999, Astrophysical Journal, 526, 579

[79] Sakata, Y. et al., 2010, Astrophysical Journal, 711, 461

[80] Tomita, H. 2005, Doctoral Dissertation, The University of Tokyo

[81] Foreman-Mackey D. et al., 2012, The Ensemble Sampler - emcee 3.0.2 documentation, 29th February 2019, < https://emcee.readthedocs.io/en/stable/user/sampler/ >

[82] Foreman-Mackey D. et al., 2012, Fitting a model to data - emcee 3.0.2 documentation, 29th February 2019, < https://emcee.readthedocs.io/en/stable/tutorials/line/>

[83] Kawaguchi T.; Mori M., 2010, Astrophysical Journal, 724, 2

[84] Krawczyk C. M. et al., 2013, Astrophysical Journal, 206, 1

[85] Kokubo M.; Minezaki T. 2019, arXiv e-prints, arXiv:1904.08946

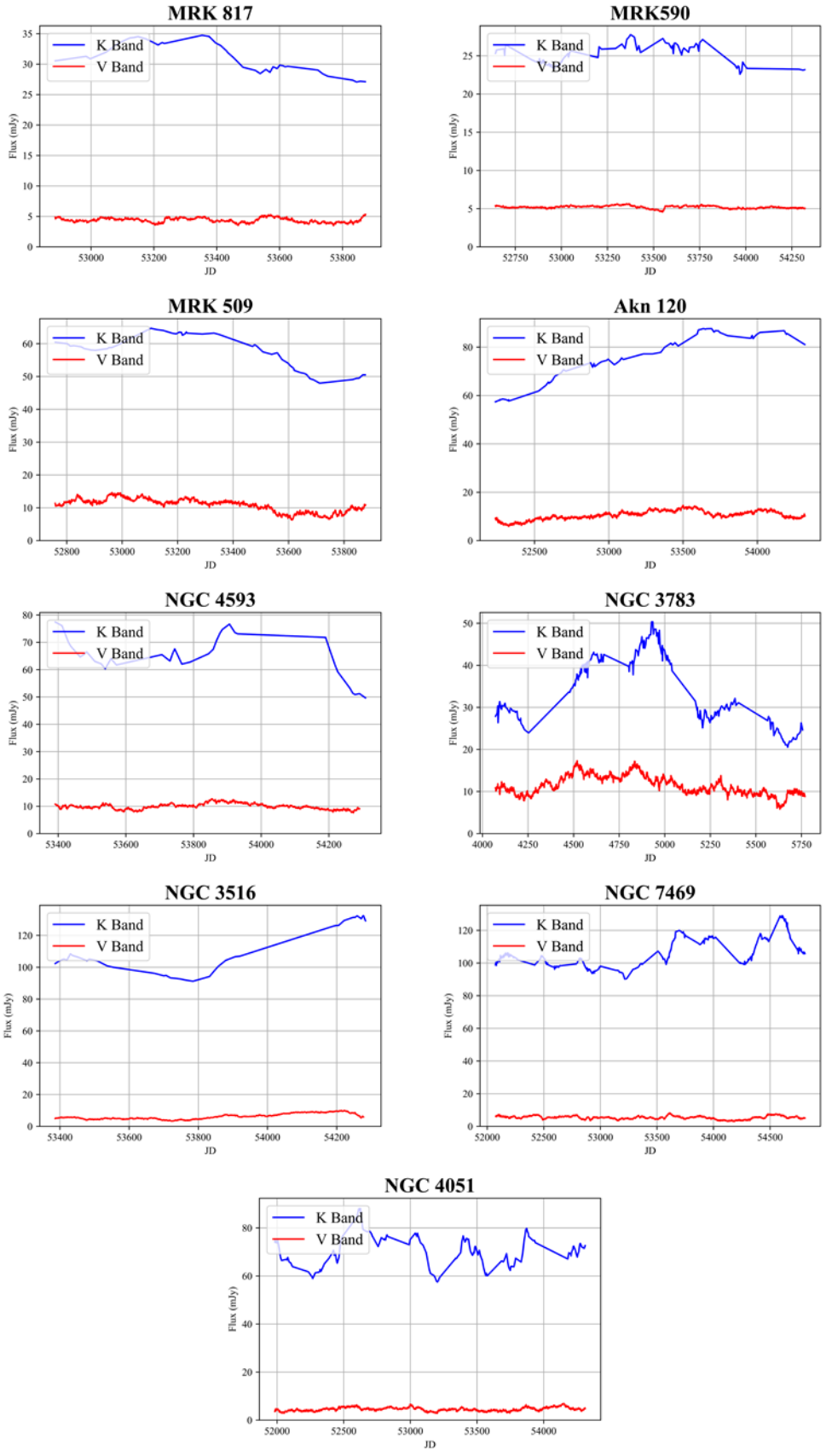
[86] Koshida S. et al., 2009, Astrophysical Journal, 700, 2

[87] Oknyansky V. L.; Metlova N. V.; Taranova O. G.; Shenavrin V. I.; Artamonov B. P.; Gaskell C. M., 2014, Astronomy Letters, 40, 527-536

APPENDICES

Appendix A: Light Curves

Figure F. *(right) K band light curves and one of the simulated V band light curves for out nine AGN. [15] [18]*



Appendix B: Tables

OBJECT	V-BAND FLUX CONTRIBUTION (mJy)
NGC 4593	6.89 ± 0.08
Mrk 590	4.39 ± 0.06
NGC 3516	14.34 ± 0.25
NGC 4051	8.03 ± 0.37
NGC 7469	9.3 ± 0.31
Mrk 817	1.53 ± 0.05
Mrk 509	3.78 ± 0.21
NGC 3783	2.2
Akn 120	3.27 ± 0.07

Table 4. *(above) AGN host galaxy V band flux contributions [15] [18].*

OBJECT	PEAKS										MEAN	STDEV
	1	2	3	4	5	6	7	8	9	10		
NGC 4593												
CCF	0.884	0.903	0.924	0.915	0.925	0.916	0.885	0.837	0.886	0.904	0.898	0.0251
Dust Lag (JD)	31	39	33	34	37	48	42	27	170*	26	48.7	40.9
Mrk 590												
CCF	0.633	0.705	0.68	0.619	0.675	0.645	0.639	0.683	0.615	0.669	0.656	0.0287
Dust Lag (JD)	29	28	28	1*	25	29	27	27	1*	28	22.3	10.71
NGC 3516												
CCF	0.97	0.96	0.965	0.968	0.958	0.963	0.966	0.966	0.964	0.969	0.965	0.00362
Dust Lag (JD)	44	45	42	46	21*	44	37	36	36	40	39.1	7.01
NGC 4051												
CCF	0.759	0.773	0.771	0.76	0.761	0.749	0.788	0.758	0.749	0.783	0.765	0.0126
Dust Lag (JD)	11	18	20	9	12	9	22	14	10	10	13.5	4.57
NGC 7469												
CCF	0.611	0.648	0.63	0.611	0.542	0.609	0.614	0.651	0.65	0.61	0.618	0.0303
Dust Lag (JD)	72	70	72	44	42	71	52	70	72	68	63.3	11.63
Mrk 817												
CCF	0.63	0.468	0.551	0.593	0.435	0.518	0.56	0.35	0.549	0.585	0.524	0.0798
Dust Lag (JD)	84	87	76	80	65*	85	83	85	90	78	81.3	6.72
Mrk 509												
CCF	0.907	0.896	0.906	0.892	0.871	0.856	0.923	0.908	0.882	0.836	0.888	0.0254
Dust Lag (JD)	130	124	102	103	112	65	58	89	113	9	90.5	35.1
NGC 3783												
CCF	0.909	0.915	0.92	0.912	0.882	0.898	0.915	0.903	0.903	0.924	0.908	0.0116
Dust Lag (JD)	83	81	84	85	81	83	85	86	86	86	84	1.84
Akn 120												
CCF	0.862	0.884	0.866	0.83	0.868	0.858	0.892	0.839	0.873	0.882	0.866	0.0185
Dust Lag (JD)	67	61	51	19	81	45	133	54	126	180	81.7	46.8

Table 5. *(above) Values of the peak cross-correlation coefficient for each simulated V band and the corresponding dust lag value. The mean peak values and their standard deviations are given on the rightmost columns. Potential outliers included in the calculation are marked with *.*

OBJECT	PEAKS										MEAN	STDEV
	1	2	3	4	5	6	7	8	9	10		
NGC 4593												
CCF	0.915	0.906	0.905	0.907	0.903	0.91	0.904	0.911	0.907	0.909	0.907	0.00262
Dust Lag (JD)	76*	72	49*	72	73	72	68	70	71	71	71.13	1.45
Mrk 590												
CCF	0.982	0.982	0.982	0.98	0.981	0.981	0.98	0.983	0.981	0.982	0.981	0.00099
Dust Lag (JD)	2*	2*	3	3	3	3	3	3	3	4*	3	0
NGC 3516												
CCF	0.962	0.958	0.962	0.961	0.959	0.96	0.96	0.96	0.96	0.96	0.96	0.0012
Dust Lag (JD)	39	43	42	49*	21	42	37	40	32*	40	39.4	3.31
NGC 4051												
CCF	0.957	0.957	0.955	0.96	0.953	0.963	0.954	0.965	0.963	0.958	0.959	0.00428
Dust Lag (JD)	10*	10*	71	100	89	73	71	95	103	71	84.1	13.2
NGC 7469												
CCF	0.976	0.972	0.976	0.974	0.968	0.974	0.975	0.974	0.975	0.976	0.974	0.00232
Dust Lag (JD)	70	70	70	70	66	63	69	66	66	67	67.7	2.33
Mrk 817												
CCF	0.911	0.914	0.904	0.906	0.911	0.908	0.907	0.9	0.912	0.905	0.908	0.00426
Dust Lag (JD)	83	88	80	84	62*	76	79	81	85	64*	82	3.54
Mrk 509												
CCF	0.819	0.819	0.823	0.825	0.795	0.807	0.814	0.824	0.811	0.805	0.814	0.00923
Dust Lag (JD)	10	50	51	59	42	7	10	9	49	41	32.8	20
NGC 3783												
CCF	0.95	0.939	0.948	0.951	0.94	0.949	0.945	0.943	0.948	0.95	0.947	0.00384
Dust Lag (JD)	97	106*	95	99	94	80*	97	90	79*	97	95.6	2.72
Akn 120												
CCF	0.814	0.837	0.833	0.824	0.812	0.828	0.806	0.827	0.828	0.825	0.823	0.00925
Dust Lag (JD)	161	127	160	122	89	79	159	68	154	68	118.7	37.5

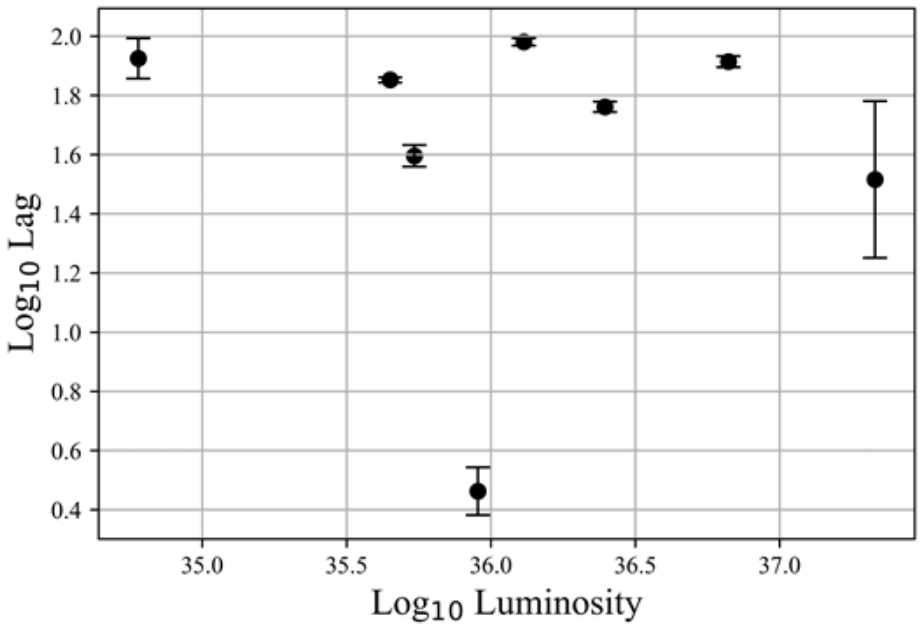
Table 6. *(above) Values of the peak cross-correlation coefficient for each mean flux-subtracted simulated V band and the corresponding dust lag value. The mean peak values and their standard deviations are given on the rightmost columns. Outliers that have not been included in the calculation are marked with *.*

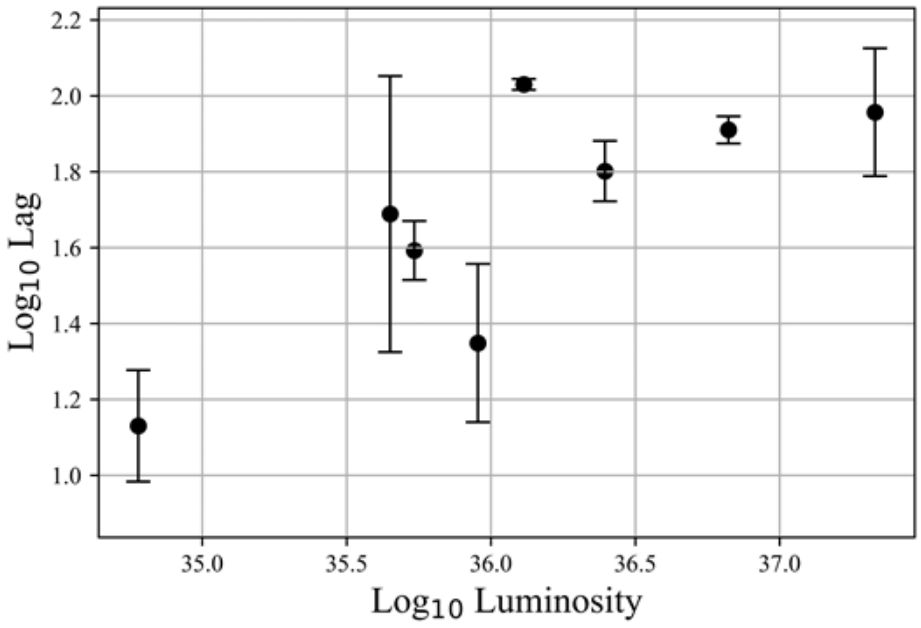
OBJECT	PEAKS										MEAN	STDEV
	1	2	3	4	5	6	7	8	9	10		
NGC 4593												
CCF	0.884	0.903	0.924	0.915	0.925	0.916	0.885	0.837	0.886	0.904	0.899	0.0261
Dust Lag (JD)	31	39	33	34	37	48	42	27	170*	26	35.2	6.7
Mrk 590												
CCF	0.622	0.699	0.481	0.664	0.742	0.602	0.493	0.672	0.409	0.788	0.632	0.112
Dust Lag (JD)	2	4	79*	3	2	12	2	3	18	2	5.33	5.4
NGC 3516												
CCF	0.97	0.96	0.965	0.968	0.958	0.963	0.966	0.966	0.964	0.969	0.966	0.00294
Dust Lag (JD)	44	45	42	46	21*	44	37	36	36	40	41.1	3.75
NGC 4051												
CCF	0.964	0.968	0.974	0.975	0.973	0.973	0.954	0.974	0.973	0.974	0.97	0.00629
Dust Lag (JD)	89	32	24	102	34	14	24	106	95	12	53.2	37.4
NGC 7469												
CCF	0.919	0.914	0.927	0.898	0.862	0.917	0.92	0.926	0.918	0.92	0.912	0.0183
Dust Lag (JD)	66	55	55	51	70	69	61	50	59	66	60.2	6.97
Mrk 817												
CCF	0.63	0.468	0.551	0.593	0.435	0.518	0.56	0.35	0.549	0.585	0.534	0.0781
Dust Lag (JD)	84	87	76	80	65*	85	83	85	90	78	83.1	4.18
Mrk 509												
CCF	0.786	0.829	0.822	0.896	0.764	0.826	0.823	0.806	0.826	0.746	0.888	0.0254
Dust Lag (JD)	60	62	62	57	50	25	55	48	45	7	90.5	35.1
NGC 3783												
CCF	0.981	0.983	0.984	0.982	0.976	0.975	0.98	0.98	0.981	0.984	0.98	0.0032
Dust Lag (JD)	110	100	110	120*	105	111	86*	105	106	110	107.1	3.55
Akn 120												
CCF	0.603	0.604	0.605	0.557	0.605	0.572	0.597	0.598	0.554	0.556	0.585	0.0213
Dust Lag (JD)	40	44	57	1	21	1	1	37	5	1	20.8	20.73

Table 7. *(above) Values of the peak cross-correlation coefficient for each flux spline-subtracted simulated V band and the corresponding dust lag value. The mean peak values and their standard deviations are given on the rightmost columns. Outliers that have not been included in the calculation are marked with *.*

Appendix C: Additional Figures

Figure G. *(below) Logarithmic plot of the dust lags and luminosities for the AGN sample, for both the mean-subtracted light curves (top) and spline-subtracted light curves (bottom).*





Appendix D: Code

Cross-Correlation Analysis

```
array = []

with open('NGC4051_K.txt', 'r') as f:
   for line in f:
      array.append((line.split()))
NGC4051_K = np.asarray(array)
NGC4051_K = NGC4051_K[:,0:2].astype(float)
#NGC4051_K = NGC4051_K[:,3:-1].astype(float)


array = []

with open('NGC4051_V_Sim.txt', 'r') as f:
   for line in f:
      array.append((line.split()))
NGC4051_VS = np.asarray(array)
NGC4051_VS = NGC4051_VS.astype(float)

def crosscorrelate(a1,a2,lrange,step,fc):
    """introduces lag to a data set and cross-correlates that lagged set with another. could
    maybe lag in opposite directions to save time?"""
    lags = np.arange(-lrange,+lrange,step) #create range of lags
    ccc = np.zeros(len(lags)) #create input for ccc
    a1old = a1
    for i in range(0,len(lags)): #loop for each
        a3 = a2 #could just replace a3 in line below with a2 and save the hassle of creating a whole new array
        a3[:,0] = a2[:,0] + lags[i] #lag array. for some reason after this a2 = a2+lags? must be a syntax thing
        a1,a3 = matcharr2(a1,a3,fc,0)
        #a1,a3 = subspline(a1),subspline(a3)
        a1,a3 = matcharr2(a1,a3,1,1)
        #a1,a3=submu(a1),submu(a3)
        ccc[i] = correlatecof(a1[:,1],a3[:,1]) #calculates ccc at this lag value
        a2[:,0] = a2[:,0] - lags[i] #set a2 back to original value
        a1 = a1old #set a1 back to original value
    return ccc,lags

def correlatecof(y1,y2):
    """calculates cross-correlation coefficient for two sets of data. see readme for details"""
    mu1 = np.mean(y1)
    sd1 = np.std(y1)
    mu2 = np.mean(y2)
    sd2 = np.std(y2)
    y1 = y1 - mu1
    y2 = y2 - mu2
    dn = sd1*sd2
    N = float(len(y1))
    dn = 1.0 / (dn*N)
    ccc = np.dot(y1,y2)*dn
    return ccc


def matcharr(y1,y2,fc):
    """matches two data sets by JD and discards non-matching data."""
    y3 = np.zeros((len(y1[:,0]),2)) #creates empty array the length of y1
    y2 = np.column_stack((y2[:,0],y2[:,fc])) #get JD of y2 and relevant flux column.
    y1d = []    #which rows to delete from y1
    y3d = []    #which rows to delete from y3
    for j in range(0,len(y1[:,0])):          #loop through y1
        match = 0                              #no match yet
        for k in range(0,len(y2[:,0])):     #loop through y2
            if int(y1[j,0]) == int(y2[k,0]):  #do the times match?
                y3[j,:] = y2[k,:]               #add to y3
                match = 1                       #match found
        if match == 0:                        #no match found
            y1d.append(j)                     #note rows to be
            y3d.append(j)                     #deleted
    y1 = np.delete(y1,y1d,0)             #and delete at
    y3 = np.delete(y3,y3d,0)             #the end
    """an issue with this function is that data sets with a fixed offset, e.g. x=10,20,30 and y=11,21,31 will never be
    matched. it may be useful to introduce a parameter that looks for matches within a certain range and assign the
    most suitable match."""
    return y1,y3
```

```
def matcharr2(y1,y2,fc,test):
    '''matches two data sets by JD using linear interpolation.'''
    y2 = np.column_stack((y2[:,0],y2[:,fc])) #get JD of y2 and relevant flux column
    '''we need to truncate y1 to fit within the range of JD in y2'''
    '''we should find the first jd of y1 in the range'''
    a = np.where(y1[:,0]<=y2[0,0])
    y1 = np.delete(y1,a,0)
    '''and the last'''
    a = np.where(y1[:,0]>=y2[-1,0])
    y1 = np.delete(y1,a,0)
    y3 = np.zeros((len(y1[:,0]),2))
    for j in range(0,len(y1[:,0])): #loop through JD of y1
        idx = int(np.argmin(np.abs(y2[:,0]-y1[j,0]))) #find the jd in y2 closest to the jd in y1
        sign = int(np.sign(y2[idx,0]-y1[j,0])) #sign = -1 if jd2 is lower, +1 if higher
        '''if sign = 0 both values are at the same JD and we don't need to interpolate'''
        if sign == 0.0:
            y3[j,:] = y2[j,:]
        else: #linearly interpolate
            xl = min(y2[idx-sign,0],y2[idx,0]) #lower x value
            xh = max(y2[idx-sign,0],y2[idx,0]) #higher x value
            yl = y2[np.argmin(np.abs(y2[:,0]-xl)),1]
            yh = y2[np.argmin(np.abs(y2[:,0]-xh)),1] #corresponding y values
            grad = (yh-yl)/(xh-xl)                                          #XH-XL ARE ZERO - WHY????
            yint = yl - (grad*xl)
            yj = grad*y1[j,0] + yint #linearly interpolated flux  #YINT AND GRAD ARE PROBLEMS
            y3[j,0] = y1[j,0]
            y3[j,1] = yj
    return y1,y3

def subspline(x,epc=90):
    '''generates a spline from the averages at different epochs.
       then subtracts the spline from the dataset.'''
    '''x values are in JD, y are corresponding flux'''
    '''need to divide data into epochs - say 90JD break is enough'''
    h = [0] #the epoch preceding a break in observation;the last epoch in an observing period
    xmean = []
    ymean = []
    for i in range(0,len(x[:,0])-1):
        #print(' ')
        #print('current date is ' +str(x[i,0]))
        #print('next date is ' +str(x[i+1,0]))
        #print('difference is ' +str(x[i+1,0]-x[i,0]))
        if (x[i+1,0]-x[i,0]) >= epc:
            #print('split here!')
            h.append(i+1)
    for j in range(0,len(h)-1):
        xtemp = x[h[j]:h[j+1],:]
        ymean.append(np.mean(xtemp[:,1])) #mean flux
        xmean.append(xtemp[np.argmin(np.abs(xtemp[:,1]-ymean[-1])),0])
    if len(h)<4:
        print("spline gen error!")
        return x
    #yspline = sp.interpolate.spline(xmean,ymean,x[:,0],order=3) #spline is deprecated

    yspline = sp.interpolate.CubicSpline(xmean,ymean)
    nuy = yspline(x[:,0])

    x[:,1] = np.subtract(x[:,1],nuy)

    return x

def submu(x,epc=90):
    '''subtracts the mean value of the flux from each value in an epoch.'''
    h = [0] #the epoch preceding a break in observation;the last epoch in an observing period
    for i in range(0,len(x[:,0])-1):
        if (x[i+1,0]-x[i,0]) >= epc:
            h.append(i+1)
    for j in range(0,len(h)-1):
        ymean = np.mean(x[h[j]:h[j+1],1])
        x[h[j]:h[j+1],1] = np.subtract(x[h[j]:h[j+1],1],ymean)
    return x
```

```
def exlag(x,y,lim=180):
    '''extracts the lag from a 1d dataset x.'''
    for i in np.arange(0,1,0.05):
        pks = find_peaks(x,prominence=i)
        pks = pks[0]    #indices of peaks
        pkv = y[pks]
        pkv = pkv[ (pkv >= 0) & (pkv <= lim) ]
        if len(pkv) == 1:
            peak = pkv[0]
            pc = pks[0]
            return pc,peak
    print('unable to retrieve peak!')
    pkv = 0
    peak = 0
    return peak, pkv

def exlag2(x,y,lim=180):
    '''extracts the lag from a 1d dataset x.'''

    posx = np.asarray(x[(y > 0) & (y <= lim)]) #values of ccc for positive lag
    posy = np.asarray(y[(y > 0) & (y <= lim)]) #corresponding lag values

    pkxm = posx[np.argmax(posx)] #maximum ccc
    #print(pkxm)
    pkym = posy[np.argmax(posx)] #corresponding x

    for i in np.arange(0,1,0.05):
        pks = find_peaks(x,prominence=i)
        pks = pks[0]    #indices of peaks

        pky = y[pks] #peak values for lag
        pkx = x[pks]

        pk = pks[(pky > 0) & (pky <= lim)] #look for peaks in the given range

        if len(pk) == 1: #best peak selected
            if y[pk[0]] <= pkym:
                return y[pk[0]], x[pk[0]]
            else:
                return pkym, pkxm

    return pkym, pkxm

rangel = 500


pks = []
pkc = []


for i in range(1,11):
    ccc,lags = crosscorrelate(NGC4051_K,NGC4051_VS,rangel,1,i)
    pylab.plot(lags,ccc)
    a,b = exlag2(ccc,lags) #a is lag, b is the ccc
    pks.append(b)
    pkc.append(a)
```

The Ensemble Sampler

```
def log_likelihood(theta, x, y, yerr):
    m, b, log_f = theta
    model = m * x + b
    sigma2 = yerr ** 2 + model ** 2 * np.exp(2 * log_f)
    return -0.5 * np.sum((y - model) ** 2 / sigma2 + np.log(sigma2))

def log_prior(theta):
    m, b, log_f = theta
    if 0.0 < m < 1.5 and -15.0 < b < -5.0 and -20.0 < log_f < 0.0:
        return 0.0
    return -np.inf

def log_probability(theta, x, y, yerr):
    lp = log_prior(theta)
    if not np.isfinite(lp):
        return -np.inf
    return lp + log_likelihood(theta, x, y, yerr)

```

```
29
30 def refine_parameters(x,y,yerr,f=0.454,minimize = '1',plotcorner='no',project='no'):
31     '''
32     uses emcee and returns a 4x3 array. column 1 = m, column 2 = c, column 3 = log_f.
33     row 1 = first estimate by numpy polyfit, row 2 = scipy minimized. Can also use scipy minimize to further refine.
34     '''
35
36     V = np.polyfit(x,y,1) # get initial parameters
37     m_initial = V[0]
38     c_initial = V[1]
39     f_initial = f
40
41     log_f_initial = log(f_initial,10)
42
43     soln = m_initial, c_initial, log_f_initial
44
45     parameters = np.zeros((5,3))
46     parameters[0,0] = m_initial
47     parameters[0,1] = c_initial
48     parameters[0,2] = log_f_initial
49
50
51     if minimize == 2:
52         from scipy.optimize import minimize
53
54         n11 = lambda *args: -log_likelihood(*args)
55         initial = np.array([m_initial,c_initial,log_f_initial])
56         soln = minimize(n11,initial,args=(x,y,yerr))
57         soln = soln.x
58
59         parameters[1,0] = soln[0]
60         parameters[1,1] = soln[1]
61         parameters[1,2] = soln[2]
62
63     if minimize ==1:
64         np.delete(parameters,1,0) #this isn't working right now?
65
66     m_m1, c_m1, log_f_m1 = soln
67
68     pos = soln + 1e-4 * np.random.randn(32, 3) #5000
69     print(pos.shape)
70     nwalkers, ndim = pos.shape
71
72     sampler = emcee.EnsembleSampler(nwalkers, ndim, log_probability, args=(x, y, yerr))
73     sampler.run_mcmc(pos, 5000, progress=True) #10000
74
75     flat_samples = sampler.get_chain(discard=100, thin=15, flat=True)
76
77     labels = ["slope", "offset", 'Log'+ r'$_1$'+'$_0$'+u'\u0192']
78
79     if plotcorner == 'yes':
80         fig = corner.corner(flat_samples, labels=labels, truths=[m_m1, c_m1, log_f_m1])
81         pylab.savefig('corner.png',dpi=1000)
82         pylab.show()
83     if project == 'yes':
84         inds = np.random.randint(len(flat_samples), size=100)
85
86         for ind in inds:
87             x0 = np.linspace(min(x),max(x),2)
88             sample = flat_samples[ind] #put this outside if statement
89             pylab.plot(x0, np.dot(np.vander(x0, 2), sample[:2]), "C1", alpha=0.1)
90             pylab.scatter(x[-2],y[-2],s=500,facecolors='none',edgecolors='red') #mrk590
91             #pylab.scatter(x[-1],y[-1],s=500,facecolors='none',edgecolors='green') #akn120
92             pylab.errorbar(x, y, yerr=yerr, fmt=".k", capsize=0)
93             pylab.plot(x0, m_m1 * x0 + c_m1, "k", label="truth")
94             #pylab.legend(fontsize=14)
95             pylab.xlabel('Log'+r'$_1$'+'$_0$'+' Luminosity',size=15)
96             pylab.ylabel('Log'+r'$_1$'+'$_0$'+' Lag',size=15)
97 #            pylab.ylim(min(y)-yerr[np.argmin(y)]*2)
98             pylab.grid(True)
99         #print('showing')
100        pylab.savefig('project.png',dpi=1000)
101        pylab.show()
102
103
104    for i in range(ndim):
105        mcmc = np.percentile(flat_samples[:,i],[16,50,84])
106        q = np.diff(mcmc)
107        parameters[2,i] = mcmc[1]
108        parameters[3,i] = -q[0]
109        parameters[4,i] = q[1]
110
111    return parameters
154 p = refine_parameters(x,y,yerr,f=0.1,minimize=2,plotcorner='no',project='yes') #change this to loop through different f?
```