Optical Variability of Active Galactic Nuclei in Accretion Disks

A Major Qualifying Project Report

Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfilment of the requirements for the

Degree of Bachelor of Science

in Physics and Mathematical Sciences

By

Bhanuj Jain

Date: April 27, 2022

Advisors:

Prof. Rudra Prasad Kafle
Prof. Douglas Todd Petkie
Department of Physics

Prof. Francesca Bernardi
Department of Mathematical Sciences

Abstract

A damped random walk (DRW) is a stochastic process defined by its exponential covariance matrix which behaves as a normal random walk for short time scales and asymptotically achieves a finite variability amplitude at long time scales. The DRW is employed to provide a statistical description of observed Active Galactic Nuclei (AGN) variability in the optical wavelength range 620 to 750 nm. The best fitting DRW, posterior probability distributions for the fitted DRW parameters, and the binned power spectral density alongside its autocorrelation function are fitted for an AGN and provide insight into the optical variability for an AGN.

Acknowledgments

I would like to thank my adviser’s professor Bernardi, professor Kafle, and professor Petkie for their continuous support, guidance, and encouragement over the year. Without them, this project would not have been possible. I would also like to thank my parents who have always believed in my potential and have encouraged me to pursue my passions with astronomy, without any free or burdens. And to my friends, who never make life dull.

Contents

Table of Figures

1 Introduction

In 1974,[1] astronomers observed a compact radio source at the center of our galaxy for the very first time, which they later identified as the supermassive black hole Sagittarius A*. Supermassive black holes (SMBH) are some of the largest types of black holes we have discovered to date, with their mass ranging from millions to billions of solar masses ($M_{\odot}$).[2] Since their discovery, astronomers have theorized that SMBHs lie at the center of every galaxy in our universe. Black holes tend to be invisible and undetectable through our telescopes, making them notoriously difficult to observe directly. However, when these black holes are absorbing matter that falls into their event horizon, they tend to emit large quantities of light and radiation, which are brighter than the entirety of their galaxy combined.[3] These bright cores/centers are known as Active Galactic Nuclei (AGN) and serve as proof that SMBH exists.

The objective of this article is to measure the optical variability of AGN's through the study and analysis of ARP 151, an SMBH that is about 6.5-7 million $M_{\odot}$, located in the constellation Ursa Major.[4] The light curves obtained from these AGN's can be used to infer the physical parameters for a SMBH such as mass, luminosity, distance from Earth, and Eddington rates. The light curve of ARP 151 will be analyzed through the help of the power spectral density (PSD) and structure-function (SF) and modeled by the continuous autoregressive (CAR) models and the damped random walk (DRW) model.

2 Methods

2.1 Methodology Outline

Below is a simple diagram showcasing the mathematical analysis utilized in this paper:

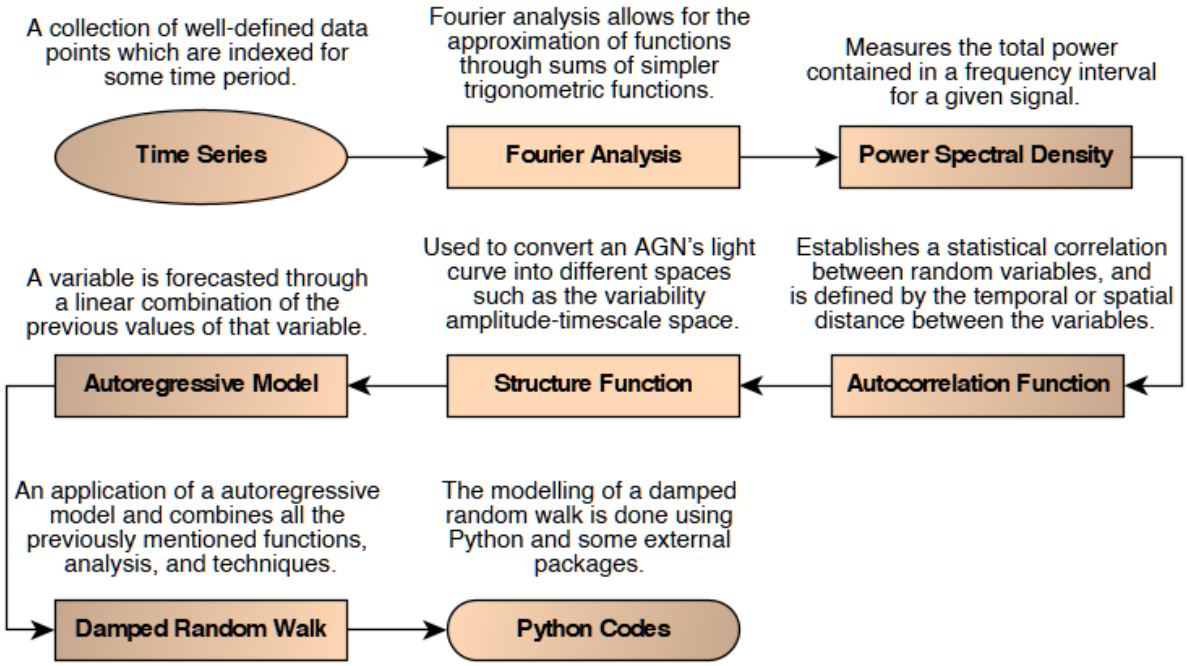

Figure 1: An outline of the methodology with a brief description for each analysis.

2.2 Time Series Analysis

A time series is a collection of well-defined data points which are indexed for some time period. They are classified as evenly or unevenly spaced time series. Evenly sampled time series are characterized by their uniform and constant time intervals, whereas unevenly sampled time series have irregular time intervals. In astronomy, most time series are unevenly sampled, with a low signal-to-noise ratio, and contain heteroscedastic errors.[5] These occur due to the natural limitation of our current technology and observational abilities and can be attributed to varying weather conditions, observational time slots, and planetary configurations.

The time series considered in this paper will be contained within pair of random variables (t_N, y_N), where y_N is the observed parameter over a sampled time coordinate t_N. Time series analysis characterizes the temporal correlation between the values of y_N and forecasts the future values of y.

2.3 Fourier Analysis

Fourier analysis is crucial for analyzing time series. Fourier analysis allows for the approximation of functions through sums of simpler trigonometric functions. The Fourier transform of a function $h(t)$ is defined as:

$$H(f) = \int_{-\infty}^{+\infty} h(t) e^{-i2\pi ft} dt, \quad (1)$$

with the inverse transform:

$$h(t) = \int_{-\infty}^{+\infty} H(f) e^{i2\pi ft} df, \quad (2)$$

where t is time in seconds (s), and f is the frequency in hertz (Hz). $h(t)$ can have any input units alongside the units of Hz. The units for $H(f)$ are a product from the units of $h(t)$ and sec. $H(f)$ is a complex function for a real function $h(t)$. If $h(t)$ is an even function, then $H(f)$ is a real function. An application of these properties of Fourier pairs- $h(t)$ and $H(f)$, can be seen through the Fourier transform of the probability density function of a zero-mean Gaussian $N(\mu = 0, \sigma)$ in the time domain is a Gaussian, where μ is the mean and σ is the standard deviation for a probability density function. In the frequency domain, the probability density function is defined as:

$$H(f) = e^{-2\pi^2\sigma^2 f^2}. \quad (3)$$

For an arbitrary function $h(t)$, if the time axis is shifted by Δt, the Fourier transform of $h(t + \Delta t)$ is:

$$\int_{-\infty}^{+\infty} h(t + \Delta t) e^{-i2\pi ft}\, dt = H(f) e^{-i2\pi f \Delta t}, \quad (4)$$

substituting the results of Eq. 3, Eq. 4 can be represented as:

$$H_{Gauss}(f) = e^{-2\pi^2\sigma^2 f^2 + i2\pi f\mu}, \quad (5)$$

which is known as the Fourier transform of a Gaussian $N(\mu, \sigma)$. Since there is no dependency on frequency, $H_{Gauss}(f)$ is known as Johnson's noise or white noise.[6]

2.4 Power Spectral Density

Another important entity in time series analysis is the Power Spectral Density function (PSD). The PSD measures the total power contained in a frequency interval for a given signal.[7] For the interval $0 \leq f < \infty$, PSD is defined as:

$$PSD(f) = |H(f)|^2 + |H(-f)|^2. \quad (6)$$

PSD calculates the power contained within the interval f to $f + df$. The calculation for total power is similar for both frequency and time domain:

$$P_{tot} \equiv \int_0^{+\infty} PSD(f)df = \int_{-\infty}^{+\infty} |h(t)|^2 dt. \quad (7)$$

This is known as Parseval's Theorem.[8] To help analyze the PSD, Lomb–Scargle periodogram will be utilized.[9] The uncertainties in the PSD will be estimated using the bootstrap technique, which allows for the estimation of the PSD for a population by averaging estimates from multiple small data samples.[10] To estimate the 1σ uncertainty for the PSD, the 16th and 84th percentiles will be used. Then, the PSD will be binned for its median and errors in equal $\log_{10} f$ spacing. The binned PSD will be fitted with a broken power-law model[11] using the Levenberg–Marquardt method[12] for nonlinear least-squares minimization:

$$P \propto \frac{1}{\left(\frac{f}{f_{br}}\right)^{\alpha} + \left(\frac{f}{f_{br}}\right)^{\beta}}, \quad (8)$$

where P is the PSD amplitude, f_{br} is the break frequency, α and β are the slopes of the power law at high and low frequency ends respectively.

2.5 Stochastic Variability

The behavior of some physical systems over a period of time is considered stochastic if the system is indeterministic and affected by random variables, which introduce an element of randomness and variation.[13] Systems that are not predictable and high in variability are known as Stochastic or Gaussian processes. AGN's are considered highly variable and stochastic since they are aperiodic and have variable amplitudes for all observed wavelengths.[14]

Due to their erratic behavior, stochastic systems and AGN's have extremely complex underlying physics, and their future values and states cannot be predicted deterministically. However, stochastic processes can be quantified through models like autocorrelation, autoregressive, and random damped walks.[15]

2.6 Autocorrelation and Structure Function

A correlation function establishes a statistical correlation between random variables and is defined by the temporal or spatial distance between the variables.[16] An autocorrelation function is an application of correlation functions, where two functions $f(t)$ and $g(t)$ are scaled by their standard deviation and defined by their time lag Δt:

$$ACF(\Delta t) = \frac{\lim_{T\to\infty} \frac{1}{T} \int_{(T)} f(t)g(t+\Delta t)dt}{\sigma_f \sigma_g}, \tag{9}$$

where σ_f and σ_g are the standard deviations for $f(t)$ and $g(t)$, respectively. Due to this normalization, the correlation function for unity is $\Delta t = 0$. Without the normalization via standard deviations, the correlation function $CF(\Delta t)$ is equal to a covariance function. $f(t)$ and $g(t)$ are assumed to be statistically weak stationary functions, such that their autocorrelation function and average do not depend on time.

The correlation function relays information about the time delay between two functions. If the time series are dependent and produced from each other, only to be differed by shifting the time axis and introducing a time lag, their correlation function produces a peak at $\Delta t = t_{lag}$. With $f(t) = g(t) = y(t)$, the autocorrelation function of $y(t)$ for a time lag Δt is defined as:

$$ACF(\Delta t) = \frac{\lim_{T\to\infty} \frac{1}{T} \int_{(T)} y(t+\Delta t)dt}{\sigma_y^2}. \tag{10}$$

This autocorrelation function provides information about the variable timescales for a process. If y is uncorrelated (e.g., this could happen due to white noise without any signal) $ACF(\Delta t) = 0$ (except for $ACF(0) = 1$). For a characteristic time τ, if a process retains information of its state during this time, the autocorrelation function vanishes for $\Delta t \gg \tau$.

The PSD (Eq. 6) and the autocorrelation function (Eq. 10) for a function $y(t)$ are Fourier pairs in the time and frequency domains. This is known as Wiener-Khinchin Theorem and affects the stationary random process.[17] The autocorrelation function is an analysis method for the time domain, while the PSD is an analysis method for the frequency domain.

The structure-function (SF) analysis is a technique used to convert an AGN's light curve into different spaces such as the variability amplitude-timescale space.[15,18] The SF is related to the autocorrelation function and is defined as:

$$SF(\Delta t) = SF_{\infty}\sqrt{1 - ACF(\Delta t)}, \tag{11}$$

where SF_{∞} is the standard deviation of the time series over a large interval of time ($t \gg \tau$).

2.7 Autoregressive Model

In an autoregressive (AR) model, a variable is forecasted through a linear combination of the previous values of that variable. The values in an AR model are derived from a time series that is regressed on previous values from the same series. A random walk is an example of an AR model, wherein every new value of the series is obtained by adding noise to the preceding value:

$$y_i = y_{i-1} + e_i, \tag{12}$$

where e_i is the noise term and does not have to be Gaussian. If y_{i-1} is multiplied by a constant, then the random walk model is known as a geometric random walk model and is mostly used to model stock market data.[19]

The random walk can be generalized for a linear AR model with a dependency k on their previous values. An AR model of order k, $AR(k)$ is defined as:

$$y_i = \sum_{j=1}^{k} a_j y_{i-j} + e_i, \tag{13}$$

where a_j is a constant. The values for y are a linear combination of the previous k values for y, with an extra noise term. For a random walk, $k = 1$ and $a_1 = 1$, giving us Eq. 12. The ACF for an $AR(k)$ model is nonzero for all time lags but quickly decays over time.

The AR model defined in Eq. 13 is only applicable for an evenly sampled time series. The generalization of the AR model for any time series is called the continuous autoregressive [$CAR(k)$] model.[20] $CAR(1)$ process is a popular model that is widely used in modeling quasar and AGN variability.

Covariance matrices are also used for modeling data alongside AR models. For the $CAR(1)$ model, the covariance matrix is defined as:

$$S_{ij} = \sigma^2 e^{-\left|\frac{t_{ij}}{\tau}\right|}, \tag{14}$$

where σ and τ are model parameters. σ^2 controls the timescale covariance for $t_{ij} \ll \tau$, which decay exponentially for a timescale τ.

2.8 Damped Random Walk

A damped random walk (DRW) model, also known as the Ornstein-Uhlenbeck process,[21] is used to describe $CAR(1)$. The DRW stochastic model is known as a mean-reverting model, since over time, the model pushes a function $y(t)$ towards its mean.[22]

Analogous to a normal random walk which is known as a drunkard's walk[23], DRW is known as a married drunkard's walk, since they (the function) will always return home (to their mean), instead of drifting away. Following Eq. 14, the ACF for a DRW is given as:

$$ACF(t) = e^{\frac{-t}{\tau}}, \tag{15}$$

where τ is a characteristic/damping timescale. Given the ACF, it is apparent that the SF for a DRW is given by:

$$SF(t) = SF_\infty \sqrt{1 - e^{\frac{-t}{\tau}}}, \tag{16}$$

where $SF_\infty = \sigma\sqrt{2}$, and is the asymptotic variability amplitude for SF. When the SF is applied to the differences of the analyzed process, the PSD is defined as:

$$PSD(f) = \frac{SF_\infty^2 \tau^2}{1 + (2\pi f \tau)^2}. \tag{17}$$

For high enough frequencies, the DRW is a f^{-2} process, similar to an normal random walk.[24] For low frequencies $\left(f \ll \frac{2\pi}{\tau}\right)$ of PSD, the damped nature of random walks can be observed.

The DRW model is one of the simplest $CAR(1)$ models to implement for Gaussian processes.[25] A higher-order model $CAR(k)$ for any $k > 1$ may provide a better fit for the data, but it may also be more difficult to understand the multiple features and

characteristic timescales associated with those models. The DRW model will be the primary model employed in this paper.

2.9 Python Packages

The modeling of a DRW is done using Python and some external packages. A Python package consists of a repository of codes/modules which can be easily accessed by a programmer, and help minimize the complexity of a program.[26]

Since the DRW is a Gaussian process and is being used to fit and model the light curves of ARP 151, a package known as CELERITE will be employed.[27] An alternate package that could have been used is the CARMA_PACK,[28] however, the CELERITE is easier to implement and produces the same results as CARMA_PACK. As defined previously, a Gaussian process is described by its covariance/kernel function. And the covariance function for a DRW is given as:

$$k(t_{ij}) = 2\sigma_{DRW}^2 e^{\frac{-t_{ij}}{\tau_{DRW}}}, \tag{18}$$

where $t_{ij} = |t_j - t_i|$ is the time lag between measurements j and i, σ_{DRW} is the amplitude factor, and τ_{DRW} is the damping term. This equation is similar to the SF defined in Eq. 16. In addition to Eq. 18, an extra white noise term needs to be added to the kernel to simulate the white noise often found in light curves alongside any measurement errors[6]:

$$k(t_{ij}) = 2\sigma_{DRW}^2 e^{\frac{-t_{ij}}{\tau_{DRW}}} + \sigma_n^2 \delta_{ij}, \tag{19}$$

where σ_n is the white noise amplitude term, and δ_{ij} is the Kronecker δ function. σ_n, σ_{DRW}, and τ_{DRW} are the three parameters that need to be fitted for the DRW to model the optical variability/light curve for ARP 151.

3 Analysis

The data used for this paper is acquired from *"The AGN Black Hole Mass Database"*, a web-based interface that contains the compilation of all spectroscopic reverberation-mapped studies of active galaxies. Reverberation mapping (also known as echo mapping) is a technique used to measure the size of broad emission-line regions and central black holes masses for AGN's.[29] The repository containing data sets for ARP 151 is available at the *"The AGN Black Hole Mass Database"*.[30]

Fig. 2 plots the red-band light curve for the apparent magnitude of ARP 151 (MBH = $10^{6.67\pm0.05}\ M_{\odot}$) along with the best-fitted DRW model with a 1σ uncertainty, represented by the width of the orange shaded area. The greater the vertical width of the shaded region is, the more uncertainty the model has. A red-band filter is used to capture the visible light from an AGN for wavelengths in the red spectrum (from 620 to 750 nm). The apparent magnitude is a measurement of the brightness of a star/AGN as measured from Earth. The measurements for the apparent magnitude of ARP 151 are a multi-year light curve that contains several seasonal gaps. These gaps are produced simply due to the orbit of Earth around the Sun and our inability to observe a section of the sky for roughly half a year.

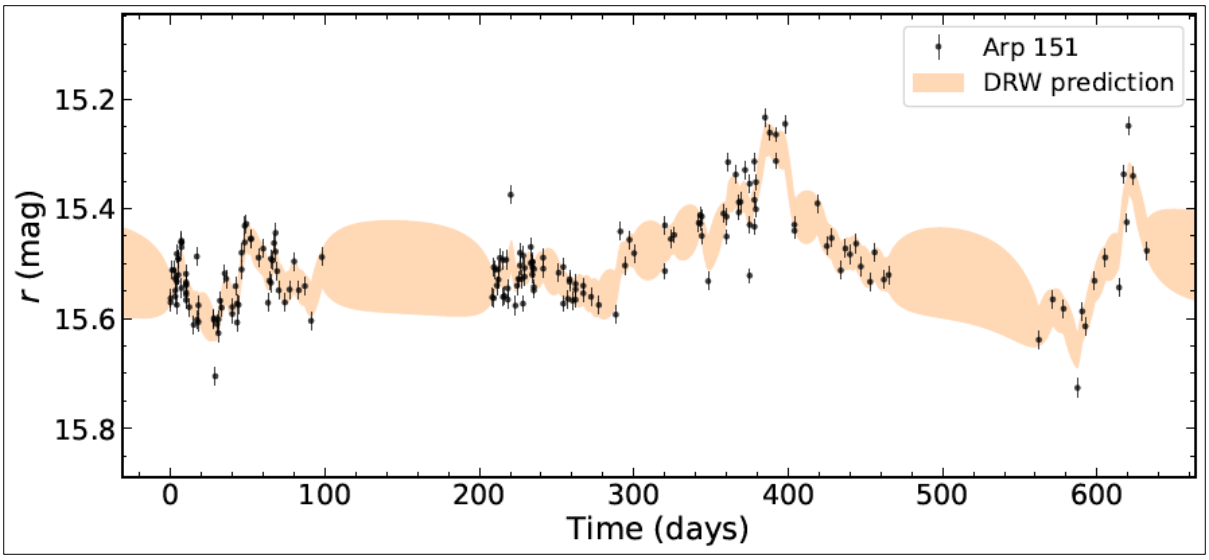

Figure 2: r-band light curve of ARP 151 and the best-fitting DRW model with $\mathbf{1\sigma}$ uncertainty (orange shaded area).

For the periods 0-120 and 200-280 days the magnitude values for densely packed, which helps the DRW model experience minimal uncertainty and predict the magnitude with a high degree of confidence since there are a lot of data points the

model can utilize and examine. However, for the periods 300-480 and 560-640 days the magnitude values are sparse and inconsistent, and while the DRW is still able to model this data, there is an increased level of uncertainty in the model due to this dispersion. While the DRW can successfully follow the general trends for the magnitude between the seasonal gaps, the uncertainty and variance of our model greatly increase during the gaps due to the lack of information.

It is worth noting that there are several data points present throughout Fig. 2 (such as 15.4 mag at 220th day) which are unmodelled for by the DRW. These outliers are not considered significant by the model since the model attaches a lower contribution weight for any outlying data through the parameters outlined in Eq. 18. It is possible that a higher-order Gaussian model would have also fitted these outliers well, but then we risk decreasing the interpretability of our data and increasing computational complexity. Outliers may also be the result of measurement/instrumentation errors presented within our telescopes and observational techniques.

There is a qualitative correlation between the uncertainty in the model and the number of data points available for a time period. This is a limitation for any Gaussian process and can only be improved upon by gathering better data and improving our observational techniques. However, there is a high enough correlation between the values for the apparent magnitude of ARP 151 and the DRW model that provides a sufficient approximation for the light curve of ARP 151.

Fig 3. shows a logarithmic plot for the posterior probability distributions of the fitted DRW parameters and their covariances. The normal distribution graphs plot the values for σ_n, σ_{DRW}, and τ_{DRW} over the iterations for the light curve as per Eq. 19. These fitted parameters carry no significance for analysis and only serve to help model the red-band light curve. The contours trace the $1, 2, 3\sigma$ levels superimposed on the sample density map, with black indicating higher density. The individual samples with the lowest-density regions are shown as black points. The red shaded regions correspond to time periods greater than 20% of the total light curve from Fig. 2.

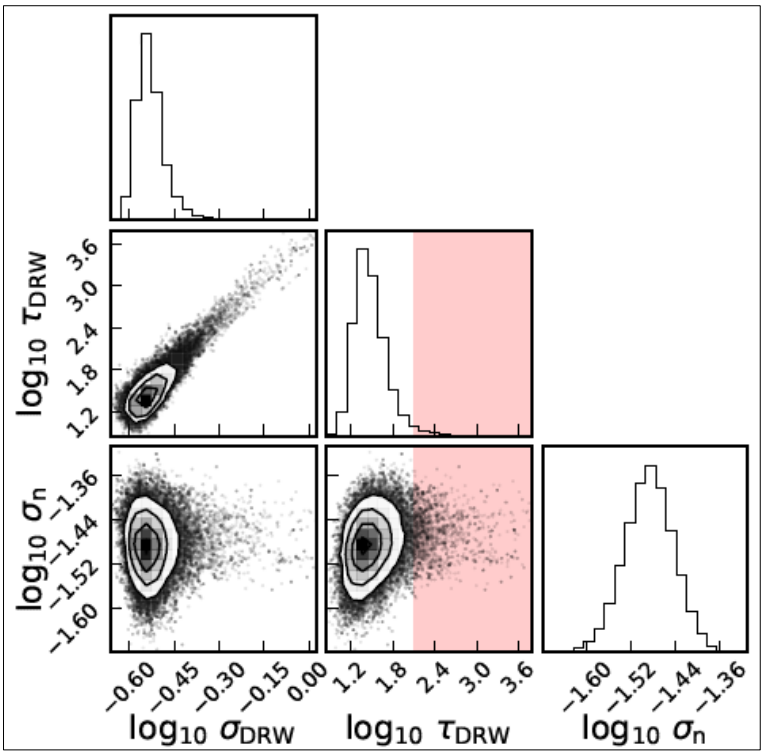

Figure 3: The posterior probability distributions for the fitted DRW parameters and their covariance contours.

The impact of σ_{DRW} and τ_{DRW} are small in this case: the normal distribution plot for $\log_{10}\sigma_{DRW}$ does not significantly differ from the $\log_{10}\tau_{DRW}$. For ARP 151, extending its light curve moves it toward the left (well-constrained) portion of the diagram, since for a fixed τ_{DRW}, increasing σ_n decreases σ_{DRW}, as seen by the relationship of these parameters in Eq. 19 and the contour plots for the fitted parameters, which highlight the negative covariance between these parameters. If the probability distribution function (PDF) peaks in the red shaded region, the light curve is not long enough due to the increased length of τ. For these increased timescales, the results become unreliable as the DRW model will overfit the light curve due to the increased randomness and unpredictability for an AGN, and produce an inaccurate model.[31] For characteristic timescales shorter than τ, best fit τ is underestimated and becomes biased to $\frac{t_{ij}}{2}$.[32] The contours show the $1, 2, 3\sigma$ levels (enclosing 68.3%, 95.5%, and 99.7% of the data), with the dense black center contour indicating the 1σ level. These contours are well defined and follow the anisotropic trends seen from the PDF function. The posterior probability distributions and covariance counters are similar to previous findings and literature available for ARP 151.[33]

Fig. 4 shows the normalized power spectral density (PSD) and binned PSD with 1σ uncertainties. The best-fitting broken power-law (PL) model is shown as a red line. The 1σ range of the DRW PSD from the posterior prediction is the orange shaded area. The corresponding break frequency f_{br} (from the broken power-law fit) and $f_\tau = (2\pi\tau_{DRW})^{-1}$ (from the DRW fitting) are shown as corresponding-colored arrows with the line segment below indicating the 1σ uncertainty. The red shaded region on the left corresponds to a period less than the mean cadence where the PSD is not well sampled, and the red shaded region on the right corresponds to a period greater than 20% of the light curve length from Fig. 2.

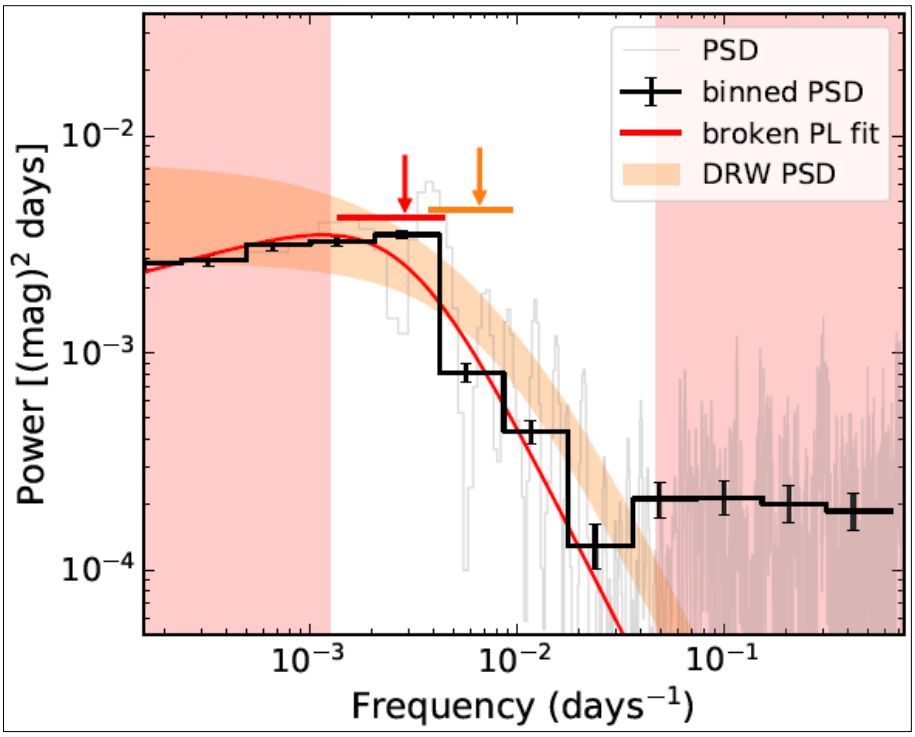

Figure 4: The normalized PSD and binned PSD with 1σ uncertainties. The best-fitting broken power-law model is shown as a red line.

The broken power-law fit described in Eq. 8 is a poor fit for this data set, particularly around the estimated break frequency, as seen by the red and orange arrows. The binned PSD shows how the power for ARP 151 drops off rapidly after the break frequency, to a point where the broken power-law and DRW fits are unable to model these values. The PSD analysis shows a flattening of the PSD towards the low-frequency end, although the location of the break frequency/timescale cannot be accurately determined and only approximated, especially for this multi-year light curve with seasonal gaps. For the low-frequency end, there exists a high degree of autocorrelation between adjacent and near-adjacent observations for the PSD. For the

high-frequency end, there is no autocorrelation between the observations for the PSD and indicates a higher degree of randomness for the PSD, which cannot be modeled accurately.[34]

It is also noteworthy that after the break frequency (red arrow), there are large systematic deviations for the binned PSD, due to the oscillatory nature of the PSD. A better model for the PSD could have been developed through the creation of a function where multiple broken power-law could be connected to have the power-law at lower frequencies bend smoothly and become a steeper power-law at higher frequencies. This function should be able to model the sharp change in power for the higher frequency as seen in Fig. 4. The fitted PSD slope could deviate from the DRW model due to well-known effects such as red noise leakage, sampling, and windowing effects.[35]

4 Conclusions

Over the last decade, there has been significant progress in both data availability and the modeling of stochastic AGN variability. The damped random walk model provides a satisfactory statistical description for ARP 151. The model has a sufficient correlation between the values for the apparent magnitude, and it provides an approximate fit for the AGNs light curve. The parameters of the DRW fitted for ARP 151 are well-defined and show that the length of the light curve is sufficient to be modellable for smaller characteristic timescales. For larger timescales, overfitting may occur, and the results would be unreliable. This conclusion is further supported by the analysis of the power spectral density function. For low-frequency end and shorter timescales, there exists a high degree of correlation between the observed values for the PSD, while for the high-frequency end and larger timescales, there exists a high degree of randomness and no correlation between the observed values for PSD.

The scope of this paper was limited to modelling the optical variability for ARP 151. However, the procedure described in chapter 2 can be used to model other AGNs, provided they have a sufficiently large enough light curve. Through the analysis of light curves, PSD, and characteristic timescales for different AGNs, it may be possible to develop a linear regressive model to help determine their physical parameters such

as their mass, luminosity, distance from Earth, and Eddington rates, which cannot be spatially resolved due to the limitations of our current instrumentation.

The aperiodic and variable nature of AGNs make it extremely difficult to model and predict their light curves. The uncertainty constrained within the DRW model is dependent on the quality of data available. Currently, our observational techniques and astronomical instruments are subpar and not extremely effective at gathering data from AGNs and quasars. However, observatories like the Sloan Digital Sky Survey (SDSS) are continuously innovating and developing new methods to gather more information about AGNs and quasars. Over time, the quality of data available would be sufficient to help model these AGNs precisely, maybe even quantify their variability through more deterministic methods.

5 References

1. Carlson, E. K. How We Discovered the Black Hole at the Center of Our Galaxy | Astronomy.com. https://astronomy.com/news/2018/07/how-we-discovered-the-black-hole-at-the-center-of-our-galaxy (2018).
2. Supermassive Black Hole | COSMOS. https://astronomy.swin.edu.au/cosmos/s/supermassive+black+hole.
3. Active galactic nuclei. Choice Rev. Online **34**, 34-2141-34–2141 (1996).
4. Arp 151 Fact Sheet - StarDate's Black Hole Encyclopedia. https://blackholes.stardate.org/objects/factsheet-Arp-151.html.
5. Leong, Y. P. & Stringer, M. Unevenly Spaced Data. datascopeanalytics.com https://www.ideo.com/datascope/unevenly-spaced-data (2016).
6. Jones, D. Thermal noise. Nature vol. 393 631 (1998).
7. LMS. What is a Power Spectral Density (PSD)? 1–16 (2020).
8. Mathuranathan. Parseval's theorem - derivation - GaussianWaves. https://www.gaussianwaves.com/2020/09/parsevals-theorem-derivation/.
9. VanderPlas, J. T. Understanding the Lomb–Scargle Periodogram. Astrophys. J. Suppl. Ser. **236**, 16 (2018).
10. Johnson, R. W. An Introduction to the Bootstrap. Teach. Stat. **23**, 49–54 (2001).
11. Freeman, W. J. & Zhai, J. Simulated power spectral density (PSD) of background electrocorticogram (ECoG). Cogn. Neurodyn. **3**, 97–103 (2009).
12. Weisstein, E. W. Levenberg-Marquardt Method. Wolfram MathWorld https://mathworld.wolfram.com/Levenberg-MarquardtMethod.html.
13. Lury, D. A., Kendall, M. G. & Buckland, W. R. A Dictionary of Statistical Terms. Stat. **20**, 76 (1971).
14. Peterson, B. M. Variability of Active Galactic Nuclei. in Advanced Lectures on the Starburst-AGN Connection 3–68 (2001). doi:10.1142/9789812811318_0002.
15. Kozłowski, S. Optical Variability of Active Galactic Nuclei. Front. Astron. Sp. Sci. **4**, 14 (2017).
16. Braun, S. Correlation Functions. in Encyclopedia of Vibration 294–302 (Elsevier, 2001). doi:10.1006/rwvb.2001.0170.
17. Wiener-Khinchin Theorem -- from Wolfram MathWorld. https://mathworld.wolfram.com/Wiener-KhinchinTheorem.html.
18. Miller, D. A. B. An introduction to functional analysis for science and engineering. (2019) doi:10.48550/arxiv.1904.02539.
19. 5.4 Geometric Random Walks. http://sfb649.wiwi.hu-berlin.de/fedc_homepage/xplore/tutorials/sfehtmlnode21.html.

20. Kelly, B. C., Bechtold, J. & Siemiginowska, A. Are the variations in quasar optical flux driven by thermal fluctuations? Astrophys. J. **698**, 895–910 (2009).

21. Mishura, Y. & Zili, M. Fractional and Sub-fractional Brownian Motions. Stoch. Anal. Mix. Fract. Gaussian Process. 31–73 (2018) doi:10.1016/B978-1-78548-245-8.50002-1.

22. Zu, Y., Kochanek, C. S., Kozłowski, S. & Udalski, A. Is Quasar Optical Variability a Damped Random Walk? Astrophys. J. **765**, (2012).

23. Random Walks. in Encyclopedia of Social Network Analysis and Mining 1488–1488 (2014). doi:10.1007/978-1-4614-6170-8_100100.

24. Takacs, L. Random Walk Processes and their Applications in Order Statistics. Ann. Appl. Probab. **2**, (2007).

25. Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A. & Uttley, P. Flexible and scalable methods for quantifying stochastic variability in the era of massive time-domain astronomical data sets. Astrophys. J. **788**, 33 (2014).

26. Udacity Team. What is a Python script? | Curious Efficiency. https://www.curiousefficiency.org/posts/2011/03/what-is-python-script.html.

27. Rasmussen, C. E. & Williams, C. K. I. Gaussian Processes for Machine Learning. Gaussian Processes for Machine Learning (The MIT Press, 2018). doi:10.7551/mitpress/3206.001.0001.

28. Kelly, B. C. MCMC Sampler for Performing Bayesian Inference on Continuous Time Autoregressive Models. GitHub https://github.com/brandonckelly/carma_pack (2018).

29. Cackett, E. M., Bentz, M. C. & Kara, E. Reverberation mapping of active galactic nuclei: from X-ray corona to dusty torus. iScience vol. 24 (2021).

30. Bentz, M. C. & Katz, S. The AGN Black Hole Mass Database. Publ. Astron. Soc. Pacific **127**, 67–73 (2015).

31. Kozlowski, S. Limitations on the recovery of the true AGN variability parameters using damped random walk modeling. Astron. Astrophys. **597**, (2017).

32. Suberlak, K. L., Ivezić, Ž. & MacLeod, C. Improving Damped Random Walk Parameters for SDSS Stripe 82 Quasars with Pan-STARRS1. Astrophys. J. **907**, 96 (2021).

33. Raimundo, S. I., Pancoast, A., Vestergaard, M., Goad, M. R. & Barth, A. J. Modelling the AGN broad line region using single-epoch spectra - I. The test case of Arp 151. Mon. Not. R. Astron. Soc. **489**, 1899–1918 (2019).

34. Suny Polytechnic Institute. 1.3.3.1. Autocorrelation Plot. Eng. Stastics Handb. 4–7 (2012).

35. Uttley, P., McHardy, I. M. & Papadakis, I. E. Measuring the broad-band power spectra of active galactic nuclei with RXTE. Mon. Not. R. Astron. Soc. **332**, 231–250 (2002).