THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

APPLICATION OF MODERN STATISTICAL TECHNIQUES TO REVERBERATION MAPPING IN THE OZDES SAMPLE

Hugh McDougall
BSc, BE

Under the supervision of
Prof Tamara Davis & Dr Benjamin Pope

A THESIS SUBMITTED TO THE UNIVERSITY OF QUEENSLAND
IN PARTIAL FULFILMENT OF THE DEGREE OF BACHELOR OF SCIENCE WITH HONOURS
SCHOOL OF SCHOOL OF MATHS AND PHYSICS
JUNE 2023

Typeset in $\LaTeX\,2_\varepsilon$.

The work presented in this Thesis is, to the best of my knowledge and belief original, except as acknowledged in the text, and has not been submitted either in whole or in part, for a degree at this or any other university.

Hugh McDougall

Abstract

At each galaxy's core is a supermassive black hole, some of which are extremely luminous 'active galactic nuclei' (AGN) that are visible out to high redshift. There is an ongoing effort to measure the central black hole mass in such AGN, with the main technique for doing so being 'reverberation mapping' (RM), in which we measure the time delay between central broadband optical emission from the accretion disk and the resulting spectral line emission from orbiting gas clouds, allowing us to infer the geometry of the AGN.

Many-object RM campaigns are limited by the problem of 'aliasing', in which seasonal gaps in our observations yield high uncertainties and misleading results in lag recovery. Aliasing leads to problems both statistical and numerical, and severely impacts the measurements in RM. In this thesis, I investigate the numerical challenges associated with aliasing, focusing specifically on the failure modes of the leading RM software package `JAVELIN`. I demonstrate that `JAVELIN` is an unsuitable tool for the challenges posed by aliasing, and present a new alternative program that I have named `LITMUS`, which makes use of modern statistical techniques to overcome these issues. I apply this new program to simulated AGN signals and demonstrate that it provides significantly improved results as compared to `JAVELIN`, more accurately recovering underlying likelihoods by resolving significant numerical artefacts present in existing software.

I also investigate the utility of performing RM for multiple lags simultaneously for those sources in which more than one reverberating emission line is visible. I apply `LITMUS` to the 92 OzDES sources with multiple lines, contrasting their results for independent and simultaneous lag recovery, and compare to existing recoveries in the literature. I find that, though `LITMUS` outperforms `JAVELIN` in recovering the true likelihood distribution of lag parameters, the existing tools for identifying and removing false positives do not work when used in conjunction with our new program. Without appropriate measures for quality cuts, the improved accuracy of `LITMUS` does not overcome the inherent noise of the aliased signals. As a result, our sample is severely contaminated by aliases and we are unable to identify any new trends in our limited subset of the OzDES sample. Further work is needed to develop a robust method for false-positive identification before LITMUS can be applied at scale to high noise data sets. Where my results are able to be compared to existing recoveries, I them to be in good agreement. In application to simulated and real data, I find that fitting of multiple lines yields no significant improvement in results for the OzDES sample.

More work is needed to characterize the noise and false positive rate of `LITMUS` when applied to real data. However, this new program solves critical numerical issues present in `JAVELIN`, and has a demonstrably improved accuracy in recovering lag likelihood distributions. With further statistical developments, `LITMUS` presents the opportunity to significantly improve the number and accuracy of reverberation mapping sources in high-redshift surveys.

Acknowledgements

I would like to thank my supervisors, Tamara and Ben, for their guidance and support during this thesis, the OzDES team for their provision of the AGN data upon which this thesis relies, and Andrew Penton in particular for readily sharing his wealth of experience with the particulars of `JAVELIN`.

I would also like to thank my family and friends, in particular my sister Jade and oldest friend Aiden, for their support during difficult times. I thank Aaron Glanville and Christopher Chang for lending their experience, advice and perspective, and my fellow honours students Maddy and Owen for their moral support. I also thank my cat Persephone for providing motivation and distraction in equal measure during long nights of writing.

Finally, I would like to sincerely thank those who convinced me to return to physics. It never was quite right being away.

Contents

Abstract **v**

Acknowledgements **vii**

List of Figures **xi**

List of Tables **xiii**

1 Introduction **1**
- 1.1 OzDES: The Australian Dark Energy Survey . . . 3

2 Background **5**
- 2.1 Physical Basis for Reverberation Mapping . . . 5
- 2.2 The Radius-Luminosity Relationship . . . 8
- 2.3 An Overview of Lag Recovery Methods . . . 9
  - 2.3.1 Interpolated Cross Correlation Method . . . 9
  - 2.3.2 Gaussian Process Modelling (`JAVELIN`) & The Damped Random Walk 11
- 2.4 The Aliasing Problem & Its Impacts . . . 12
  - 2.4.1 A Conceptual Explanation of the Aliasing Problem . . . 15
  - 2.4.2 Methods of False Positive Rejection & Aliasing Removal . . . 16
- 2.5 Multi-Line Fitting . . . 18

3 Modelling & Methods of Lag Recovery via Gaussian Processes and MCMC **21**
- 3.1 A General Approach to Gaussian Process Modelling . . . 21
  - 3.1.1 The Damped Random Walk . . . 23
  - 3.1.2 Lag Recovery via Monte Carlo Markov Chains . . . 25
- 3.2 Geometry of The Posterior . . . 26
- 3.3 Behaviour of `JAVELIN` & Complications in Lag Recovery . . . 32
  - 3.3.1 The Affine Invariant Sampler & `emcee`, . . . 32
  - 3.3.2 Failure Modes of `JAVELIN` . . . 34

4 LITMUS: Lag Inference Through The Mixed Use of Samplers - An Alternative Program for Lag Recovery **37**
- 4.1 Alternative Samplers . . . 37
  - 4.1.1 Nested Sampling, `JAXNS` . . . 37
  - 4.1.2 Hamiltonian Monte Carlo / The No U-Turn Sampler . . . 39

4.2 Procedure Overview 40
4.3 Sampler Tuning 41
4.4 Software Packages 43
4.5 Validation of Results & Comparison with `JAVELIN` 44
4.5.1 Performance for Signals with No Underlying Lag 49

5 Results & Discussion 51
5.1 Application of `LITMUS` to the OzDES 2-Line Sample 51
5.2 Impact of Multi-Line Fitting in the OzDES 2-Line Sample 54

6 Future Work 57
6.1 New Methods for Rejection of False Positives 57
6.2 Model Extensions 58
6.3 Improved Sampling Algorithms for Better Numerical Performance 60
6.4 Hierarchical Modelling of Population 60

7 Conclusion 63

A Recovered Lag Summary for OzDES Two-Line Sources 65

B Quality Cuts of OzDES Sources & Impact of Simultaneous Lag Recovery 71

C Summary of Signal to Noise Ratios For OzDES Two-Line Sources 77

References 81

List of Figures

1.1 Standard Model of AGN geometry 2
1.2 Histogram of number of AGN sources in the OzDES sample against redshift 4

2.1 The structure of an AGN 6
2.2 Observable Ranges of Hβ, MgII and CIV 8
2.3 R-L Relationship for OzDES MgII Sources from Yu et al. [45] 9
2.4 Linearly interpolated continuum & emission response 10
2.5 Biases in lag summary statistics 11
2.6 Continuum & response measurements for OzDES source 2925858108 13
2.7 Lag recovery distributions from `JAVELIN` for simulated data at varying levels of sample quality. 14
2.8 Qualitative demonstration of the source of aliasing peaks 15
2.9 Example aliasing in a simulated AGN light-curve 16
2.10 Example of spectrum showing CIV line at edge of detection range 19

3.1 Example of an interpolated stochastic light curve 24
3.2 Illustration of MCMC sampling compared to grid integration 26
3.3 Demonstration of non-smooth likelihood function arising from measurement cadence 28
3.4 Example of Signal and Recovery for a low-SNR OzDES dataset 30
3.5 Corner plots demonstrating typical features of the posterior geometry for for 2-line lag recoveries 31
3.6 Demonstrating of the stretch-move proposal 32
3.7 A demonstration of the failure mechanism of the stretch-move sampler in multimodal distributions 33
3.8 Demonstration of `JAVELIN`'s failure mode in multi-modal distributions 35
3.9 Example of continuum property contours for a signal fit by `JAVELIN` 36

4.1 Shrinkage of contours in nested sampling over successive iterations 38
4.2 Example of an HMC path sampling from a Gaussian distribution 39
4.3 Demonstration of the way `LITMUS` combines nested sampling and Hamiltonian Monte Carlo (HMC) 42
4.4 Comparison of `LITMUS` and `JAVELIN` lag recoveries 45
4.5 Comparison of 2-line lag recoveries for `LITMUS` and `JAVELIN` 47
4.6 Comparison of `LITMUS` and `JAVELIN` behaviour in simulated data 48

4.7 Comparison of `JAVELIN` and `LITMUS` behaviour in the case of a false positive lag recovery 50

5.1 Example of lag recoveries accepted and rejected by quality cuts 52
5.2 R-L Domain plot of all 92 recovered lag distributions for independent fitting of MgII from the 92 OzDES 2-line sources 53
5.3 Examples of the impact of simultaneous fitting on lag recovery in the OzDES sample 56

6.1 Recovered continuum properties for OzDES source 69-B-2940901822, showing multiple behaviours at multiple timescales 59
6.2 Sketch of a Probabilistic Graphical Model for hierarchical model of lag recovery 61

List of Tables

4.1 Standard tuning parameters for `LITMUS`'s nested sampling phase 43
4.2 Comparison of `JAVELIN` and `LITMUS` performance for simulated data 46
4.3 Performance of Quality Cuts On Accepted Data 46
4.5 Performance of false positive removal for `LITMUS` and `JAVELIN` 49

5.2 Summary of the number of lag recoveries that pass quality cuts in the OzDES 2-line sample 52
5.3 Comparison of lag recoveries for MgII sources accepted by Yu et al. [45] . . . 54
5.5 Summary of impact of simultaneous fitting for sources that pass quality cuts 55

A.1 Recovered `LITMUS` lags for Hβ for independent & simultaneous 2 line fitting . 65
A.2 Recovered `LITMUS` lags for MgII for independent & simultaneous 2 line fitting 68
A.3 Recovered `LITMUS` lags for CIV for independent & simultaneous 2 line fitting 70

B.1 Quality cut passes for Hβ two-line OzDES sources. 71
B.2 Quality cut passes for MgII two-line OzDES sources. 74
B.3 Quality cut passes for CIV two-line OzDES sources. 76

C.1 Summary of measurements for all OzDES sources with two emission line signals 80

1

Introduction

It is a now well established fact that each galaxy is host to a supermassive black hole (SMBH) at is core, a dense galactic nucleus orders of magnitude more massive than the familiar stellar-mass black holes. A subset of these objects have absorb in-falling matter to create a super-luminous accretion disk, with these 'Active Galactic Nuclei (AGN)' acting as the origin of many seemingly disparate classes of in-sky objects. AGN are bright enough to be seen out to cosmological distances, providing a probe into galaxy evolution and the as-of-yet unresolved question of SMBH formation deep into the cosmological past. If viewed at the correct angle, AGN present as 'Seyfert 1 galaxies', in which we can view both the accretion disk 'engine' and the fast-orbiting dust cloud of the 'Broad Line Region' (BLR). The BLR is in a tightly bound, high speed orbit about the SMBH, where it absorbs the intense radiation of the engine and re-emits it as Doppler broadened emission lines. This broadening allows us to infer the kinematics of the BLR, which in turn provides an estimate of the black hole mass. However, such measurement requires that we also know the orbital radius of the BLR, which is far too small to resolve visually.

A solution exists in the form of 'reverberation mapping' (RM), a novel technique in which a we use the characteristic delay between the photometric observations of the continuum-light from the engine and the spectroscopic observations of the 'echo' it produces from the BLR to infer the geometric scale of the system. RM allows us to substitute temporal resolution in place of spatial resolution, and as such provides the primary means of constraining black hole masses for galaxies beyond redshifts of $z = 0.1$[3].

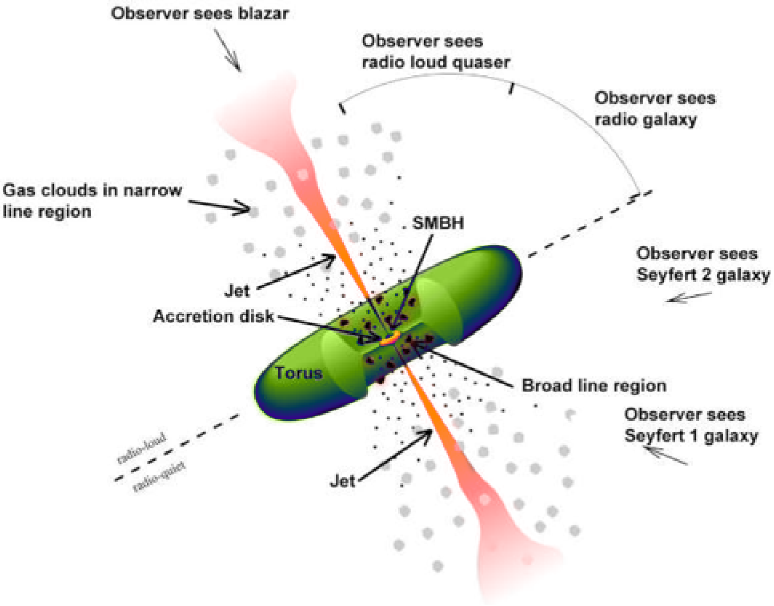

FIGURE 1.1: Standard model of AGN geometry, adapted from [40].

RM necessarily requires long observation windows of both photometric and spectroscopic measurements of AGN, and the last decade has seen efforts to extend the total number of SMBH mass estimates through multi-year, multi-object observation campaigns. In this thesis, I focus on the work of the Australian Dark Energy Survey (OzDES), which provides ground-based spectroscopy for 735 AGN measured photometrically by the Dark Energy Survey (DES).

Such multi-year surveys are characterized by half-yearly seasonal gaps in our measurements owing to the low in-sky brightness of AGN. These gaps lead to the problem of 'aliasing', a collection of statistical and numerical problems that gives rise to high uncertainty and erroneous results in reverberation mapping. There has been much past work investigating the problems of aliasing, seeking to properly characterise and mitigate its impacts, but little attention has been paid to suitability of the numerical tools in used RM to the problem of aliased signals.

The statistical techniques for performing RM are well established, with the most complete approach being to model the AGN light curve as a 'damped random walk', a highly parameterized stochastic model. This approach is employed by the program `JAVELIN`, which employs Markov-Chain Monte Carlo methods to recover lags via this stochastic modelling and has emerged as the standard tool for AGN RM. Despite its broad usage, `JAVELIN` is known to encounter significant issues when employed against AGN subject to aliasing [32].

In this thesis, I investigate the failure modes of `JAVELIN`, both their causes and impacts, and propose a new method leveraging modern numerical techniques that is robust against the obstacles posed by aliasing. To employ this new method, I present `LITMUS`: a new reverberation mapping program and alternative to `JAVELIN`. This new program makes use of the `JAX` framework for 'just-in-time' compilation to achieve speed on par with `JAVELIN`, and employs modern highly efficient MCMC techniques made possible by `JAX`'s 'autodiff' functionality.

I also employ `LITMUS` to test another novel technique in reverberation mapping: the simultaneous recovery of lags for multiple emission lines within a single source. The established technique is to fit only one lag at a time even in cases where multiple emission lines are visible, and such 'multi-line' fitting has not yet been employed on large scale surveys like OzDES. To test the impact of this multi-line fitting, I examine the 92 sources in the OzDES sample at redshifts which allow multiple reverberating emission to be seen. For these sources, I perform lag recovery for the 'simultaneous' fitting of lags and compare to the typical 'independent' fitting, and contrast the findings for both with previously established lags for this same sample.

1.1 OzDES: The Australian Dark Energy Survey

In this thesis, I aim my attention at AGN observed by the Australian Dark Energy Survey (OzDES), a multi-year, multi-object reverberation mapping campaign that provides spectroscopic measurements for these sources complimentary to the photometric measurements of the Dark Energy Survey (DES). OzDES tracks 793 sources in the redshift range $z \in [0.1, 3.8]$ over a 6-year period, with spectroscopy spanning the optical wavelength range of $\lambda \in [3750\AA, 8900\AA]$.

Photometric measurements of the AGN continuum are sourced from the Dark Energy Camera (DECam), a tool mounted to the 4 meter Victor M. Blanco telescope at the Cerro Tololo Inter-American Observatory in northern central Chile. These measurements provide light-curves for the 5 grizY photometric filters with roughly weekly cadence during observation seasons. The response line signals are measured by the 4 meter Anglo Australian Telescope (AAT) at the Siding Spring Observatory in New South Wales, Australia. These measurements are for visual-band wavelengths, and have a roughly monthly cadence.

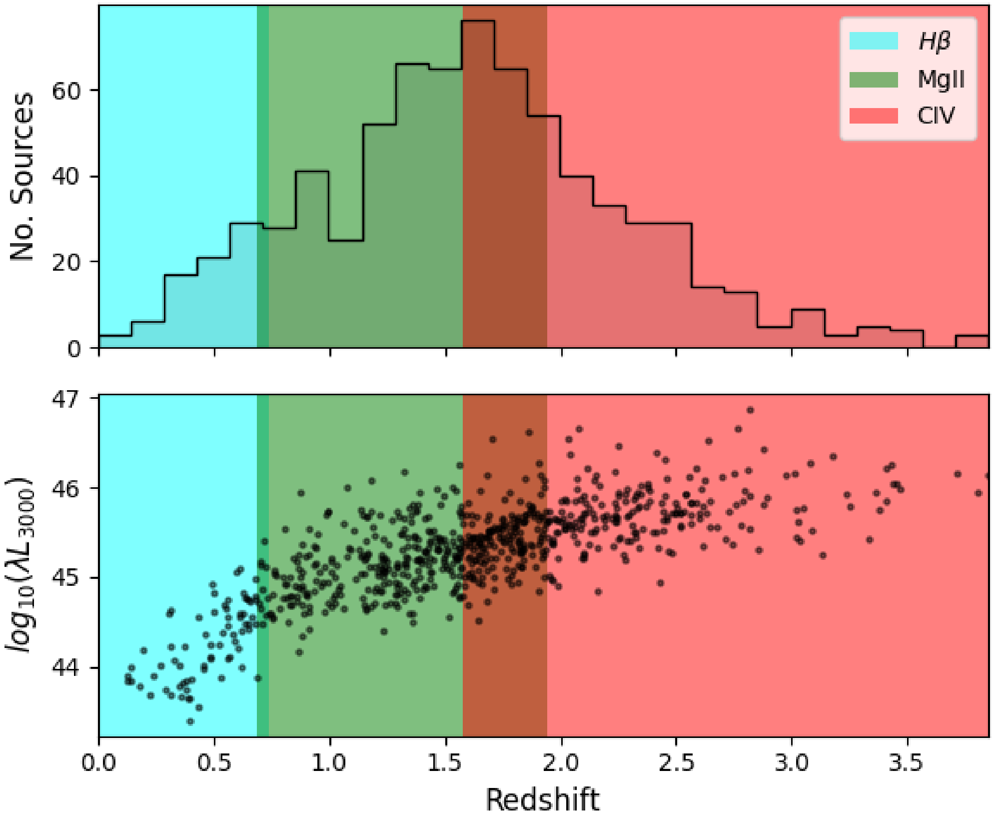

FIGURE 1.2: Histogram of AGN sources in the OzDES sample against redshift, with visible ranges for Hβ, MgII & CIV. 793 sources are available in total, with 92 sources existing at redshift ranges in which two lines are visible.

The regular cadence of the OzDES sample's measurements for any given AGN are separated into half-yearly observation windows, separated by similarly sized seasons in which the source is obscured by the background brightness of the sun. This windowing function can cause significant error in lag recoveries, and can lead to entirely spurious results that contaminate population level analysis if not identified and removed [30]. Aliasing is both difficult to detect and difficult to characterize, and so the OzDES team has done significant prior work in mitigating and accounting for its impact. We discuss some of these techniques in Section 2.4.2. Despite these obstacles, lags have been recovered from the OzDES sample for all three of the identified reverberating emission lines: Hβ[29], MgII[44] and CIV[22].

2

Background

AGN are highly luminous and highly variable sources, with different features of their geometry producing distinct signals that can be distinguished spectroscopically even when too distant to resolve spatially. The variations in the brightness of the AGN take time to 'reverberate' through its surroundings, and the timescale of this reverberation can be used to 'map' the AGN geometry ([5], [34]). This geometry can in turn can be used as a component in estimating the mass of the AGN's central black hole. In this section, I provide the reader with an introduction to the conceptual underpinnings of reverberation mapping: its applications, a broad-view look at the methods used to perform it, and the challenges it faces. A detailed review of the specific modelling and procedures used in AGN RM and this thesis in particular are available in Chapter 3.

2.1 Physical Basis for Reverberation Mapping

Though the statistical modelling of AGN RM are constantly evolving, the underlying physical mechanism have been well understood since they were introduced by Blandford and McKee [5]. The core principle is that fluctuations in brightness of the central accretion disk interact with the rest of the AGN geometry to produce similar variations, but that these variations take time to propagate through the AGN's geometry. This delay scales with distance that the continuum light must travel, and so this delay time encodes information about the system's physical size.

In reverberation mapping, we simplify the AGN to two components: the central 'engine' of the SMBH with its bright accretion disk which drivnig the system's variability, and the fast-orbiting cloud of the broad line region (BLR) which reprocesses the engine's continuum of radiation as heavily Doppler broadened spectral emission lines. High resolution RM can be used to infer more detail about the geometry and velocity of the BLR [33], or estimate the scale of the accretion disk itself [49] using the x-ray 'lamp post' model of Cackett et al.

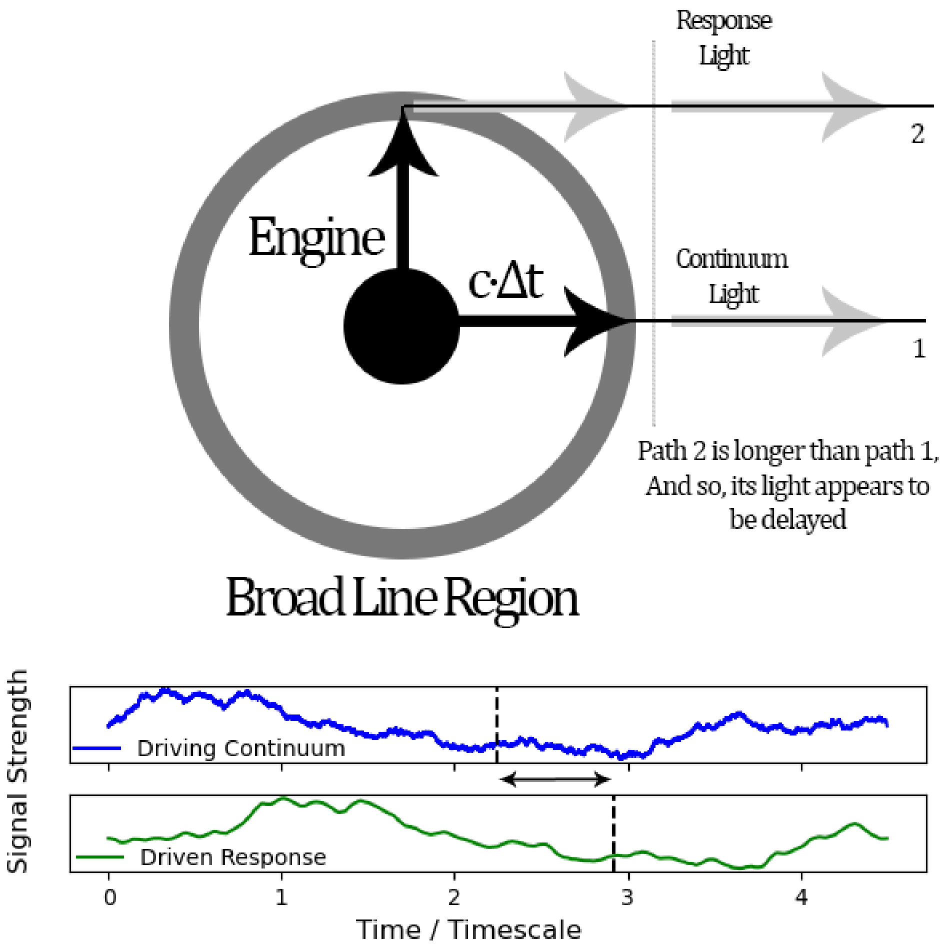

FIGURE 2.1: Simplified AGN structure, showing the differing path lengths for continuum & response light. Because light must first traverse intervening gap before exciting a response, signals from the BLR appear to lag behind the light directly from the engine, and this lag encodes the physical size of the gap.

[10], but the typical approach is to simplify the engine to a point-source and the BLR to a face-on disk of small but finite thickness (Figure 2.1).

Near in to the engine, the radiation of the accretion disk is intense enough that any dust is ionized. For the BLR to reprocess the continuum radiation, must be un-ionized, and so its innermost radius occurs where the flux from the engine drops below some critical threshold. Though the exact shape of the BLR is an open question, most RM modelling simplifies it to be a ring of small but finite width surrounding the continuum producing engine, which is simplified to a point source.

If the AGN is aligned such that BLR and engine are both visible, we perceive a 'Seyfert 1' galaxy, from which our photometric observations allows us to track variations on the accretion disk's continuum brightness, while our spectroscopic measurements allow us to track the variations in the BLR response. As the light must travel across the radial gap to excite the BLR, there is a characteristic delay between variations in these two signals:

$$R = c \cdot \Delta t \tag{2.1}$$

The high orbital speed of the BLR clouds Doppler-broaden the emission lines, giving the region its name. Measuring this broadening informs us of this speed at which the BLR orbits, while the lag encodes the radius at which it does so. Put together, we can infer the mass of the central black hole from basic kinematics:

$$\frac{GM_{\mathrm{BH}}}{R} = \langle f \rangle \left\langle V^2 \right\rangle \tag{2.2}$$

Here, $\langle f \rangle$ is a dimensionless 'virial factor' of order unity that accounts for the geometry and kinematics of the system. In this way, we can use lags derived from time-series data to infer both the geometry and mass of black holes well beyond the distances at which they can be resolved spatially. Provided we have consistent enough observations of a source to resolve the lag, we can then probe the masses of supermassive black holes out as far as our spectroscopic measurements allow. For this reason, RM is the dominant method of inferring black hole mass for sources beyond redshifts of z 0.1 [3].

To observe the 'echo' from the BLR, we need a clear emission line in its spectra with which to observe these variations. In the OzDES sample, three such lines are used, with different lines becoming visible at different redshifts as the wavelength of their corresponding emission drifts in and out of the visual-band range of our spectroscopy. These lines are 'hydrogen-beta' (Hβ) at 4861 Å, 'magnesium two' (MgII) at 2798 Å and 'carbon four' (CIV) at 1549 Å [26]. Note that there is an overlap in these redshift ranges, with two bands in which the MgII line is visible simultaneous along with Hβ at low redshift and CIV at high redshift (Figures 2.2 & 1.2).

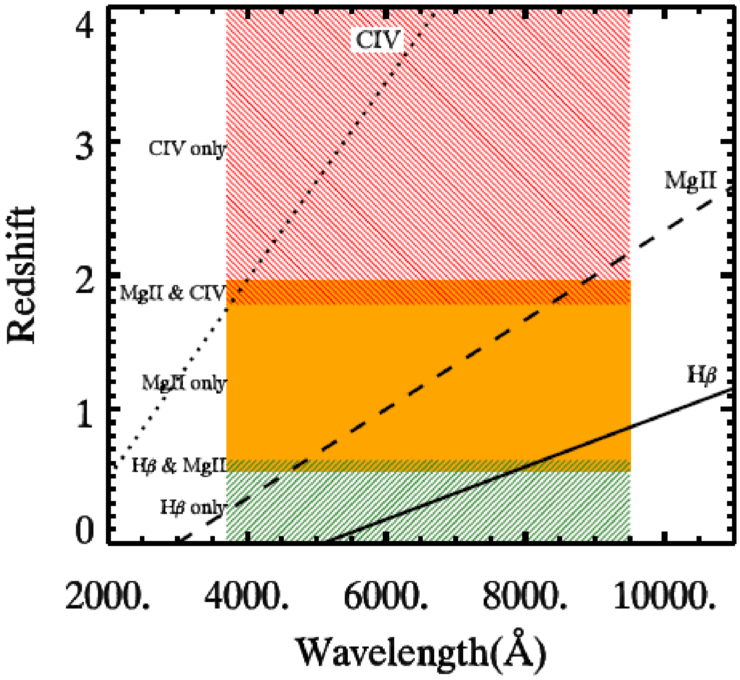

FIGURE 2.2: Visible wavelength vs redshift for the Hβ, MgII and CIV emission lines, with ranges of visibility highlighted [26]

.

2.2 The Radius-Luminosity Relationship

Reverberation mapping is a powerful tool for constraining black hole masses, but its applicability is limited by its need for long campaigns of spectroscopic and photometric observations. After considering the high loss rate of source rejection that comes from aliasing (discussed in Section 2.4.2), the number of sources we can fit is limited even for high-object count surveys like OzDES. Fortunately, an alternative exists in the single-epoch lag estimation of the ‘radius-luminosity relationship’ (R-L Relationship), an empirically observed correlation between the source brightness of AGN and their size / rest-frame lag [25]. The physical mechanism of the R-L relationship is that more luminous AGN ionize a larger surrounding volume, forcing the BLR out to a larger radius. In a simplified model, we can say that BLR begins at some critical flux, f_{crit}, indicating that lags should scale with the square root of the AGN luminosity:

$$f_{\mathrm{crit}} = \frac{L}{R_{\mathrm{BLR}}^2} \Delta t \propto R \propto L^{0.5}$$

The actual mechanics of the BLR’s ionization are not so straight-forward, and so this relationship is generalized to a power law (Equation 2.3), which we constrain using lags recovered from RM. This power law has been empirically observed for local Hβ sources with a slope of $\alpha \approx 0.5$, and an ongoing effort of the recent generation of RM campaigns is to constrain new R-L relationships for MgII and CIV at higher redshifts. R-L relations have been constrained in the OzDES sample for all three emission lines[22, 29], and in this thesis I focus my attention on the R-L relation for MgII recovered by Yu et al. [45], which is based on 25 MgII sources in the OzDES sample and recovers a slope of $\alpha \approx 0.39$.

$$\log_{10}|\Delta t| = C + \alpha \cdot \log_{10}|L| \pm \sigma \tag{2.3}$$

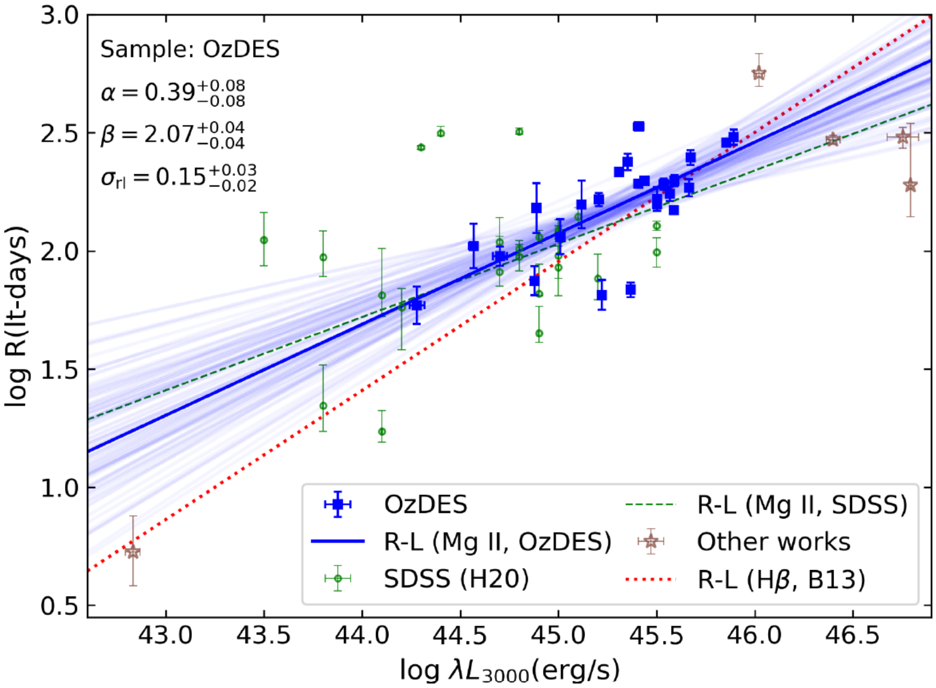

FIGURE 2.3: The R-L relation recovered for MgII by Yu et al. [45] using 25 sources in the OzDES sample.

A properly tuned R-L relation allows for lag recoveries and black hole mass estimates to be expanded out to a much larger sample, as it can be applied with only a single epoch of information. For this reason, one of the terminal goals of reverberation mapping is the recovery of lags specifically for the purpose of tuning of an R-L relationship.

2.3 An Overview of Lag Recovery Methods

Modern applications of reverberation mapping to AGN make use of two main techniques for recovering lags: the simplified but numerically inexpensive interpolated cross correlation function (ICCF), and the more precise but computationally costly Gaussian Process modelling used by programs like `JAVELIN`.

2.3.1 Interpolated Cross Correlation Method

Though `JAVELIN` has emerged as the primary tool for RM lag recovery, the simpler alternative of the ICCF provides a coarser but much faster alternative that is less susceptible to numerical artefacts. Though the ICCF is not used in this thesis, it is an important feature of RM campaigns and often used as a point of comparison to validate results (e.g. Section 2.4.2), and so is presented here as informative context.

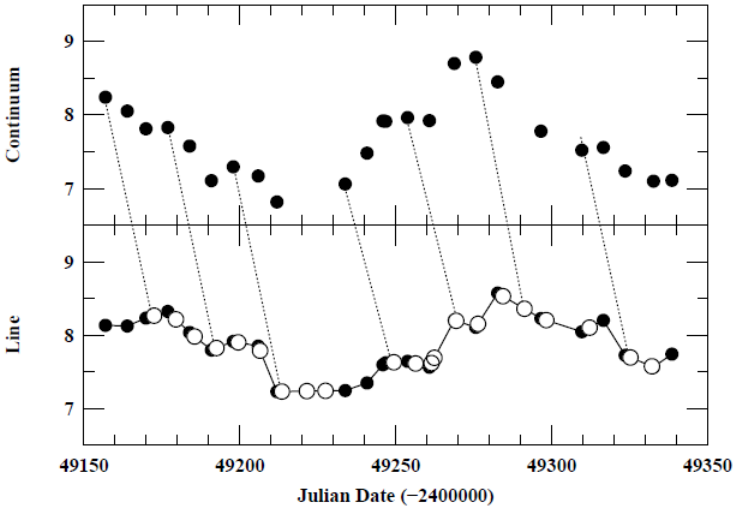

FIGURE 2.4: Demonstration of the process of the ICCF [35]. By linearly interpolating, we can align variations in the continuum and emission signals to identify the lag between them.

Introduced by Peterson [35], the ICCF method linearly interpolates between measurements to construct a continuous light-curve and calculates the correlation function as would be done for a continuous signal, i.e. numerically integrating to find the covariance. This covariance, a measure of how much the variations in the signals move in concert balanced against the amount they move against one another, is found by:

$$\text{covar}(y_1(t), y_2(t)) = \langle y_1, y_2 \rangle \approx \frac{1}{t_b - t_a} \int_{t_a}^{t_b} y_1(t) y_2(t) dt \tag{2.4}$$

where t_a and t_b are the times between which we have observations of both signals. This is often expressed instead in the dimensionless form of correlation, the covariance normalized by the variance of the two signals:

$$\phi_{12} = \frac{1}{\sigma_1 \sigma_2} \text{covar}(y_1, y_2), \quad \sigma_i^2 = \int_{\infty}^{\infty} y_1^2(t) dt \tag{2.5}$$

This correlation is a measure of how similar the signals are, while the 'correlation function' measures how this similarity changes as we slide one of the signals back and forth.

$$\phi_{12}(\Delta t) = \frac{1}{\sigma_1 \sigma_2} covar(y_1(t), y_2(t - \Delta t) \tag{2.6}$$

The correlation function then gives a 'goodness of fit' distribution for different lags, and shows a peak at the lags where the two signals align. There is some freedom in how this distribution is interpreted, and estimates of the lag can be taken from its peak, by taking the mean or median the entire distribution (e.g. Figure 2.5).

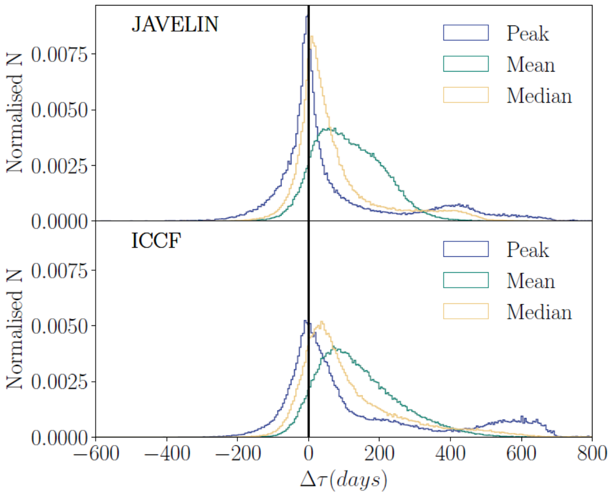

Figure 2.5: Biases introduced in different lag likelihood summary statistics for GP modelling via `JAVELIN` (top) and the ICCF (bottom) from, for OzDES-like data from [32]. All methods are biased towards over-reporting the lag, with mean being the least reliable

Uncertainties in the recovered lag are then estimated from ‘bootstrapping’, randomly varying data points within their measurement error and/or sub-sampling to remove a fraction of the measurements and seeing how the chosen measure (the correlation function’s mean median or mode) varies over many instances in a Monte Carlo fashion [16].

Despite its simplicity, the ICCF is still in use in AGN reverberation mapping and has been found to broadly agree with more sophisticated methods, though with less precision [30, 32, 42]. The ICCF’s low cost allows exhaustive algorithms that make it more robust against statistical artefacts, and its relatively low numerical cost makes it an attractive option for studies that require lag recovery on large sets of simulated data.

2.3.2 Gaussian Process Modelling (`JAVELIN`) & The Damped Random Walk

Though the ICCF is still in active use in AGN RM, more precise results can be found by leveraging our understanding of the statistical properties of the AGN’s underlying stochastic variability. First introduced by Rybicki and Press 37 and now forming the core of the methods used by `JAVELIN` [47] and `CREAM` [39], this method makes use of the fact that the variations in the continuum brightness of AGN appear to be instances of the ‘damped random walk’ [27, 28, 49], a stochastic Gaussian Process (GP; 36). Though these fluctuations are not entirely deterministic, they do exhibit well characterized statistical behaviour that allows us to make more precise inferences than the crude interpolation of the ICCF.

The central feature of GP modelling is that the light-curve’s variations are structured with some characteristic timescale, such that measurements at one time also provide information about the signal at nearby times. This structure allows us to parameterize the light curve even in the absence of a simple deterministic model. In addition to the lag, these stochastic models characterize the variations in AGN brightness by a number of other parameters, e.g. the timescale and amplitude of the fluctuations.

Programs like `JAVELIN` and `CREAM` recover lags by Markov Chain Monte Carlo (MCMC; 31) algorithms, randomly varying the model parameters (variance, timescale, lag, response amplitude etc.) to map out high-likelihood regions of parameter space. Marginalising over the other parameters, a likelihood curve for the delay is recovered, allowing this distribution to then be interpreted in a similar way to those produced by the Interpolated Cross-Correlation Function (ICCF; 35) method.

2.4 The Aliasing Problem & Its Impacts

As distant AGN are dim compared to the background brightness of the sky, ground-based observation campaigns like OzDES are necessarily broken into regular half-yearly ‘seasons’ separated by times in which the AGN is masked by the brightness of the sun. This windowing function leads to a problem in RM called ‘aliasing’, which creates spikes in the recovered lag likelihood distribution for lags that fall within these seasonal gaps [30].

Aliasing leads to two categories of problems: numerical and statistical. Firstly, the presence of multiple peaks can obscure the true underlying lag, meaning that sources heavily affected by aliasing must be discarded in a so called ‘false negative’. Conversely, the fact that the aliasing peaks are of finite width can give the false appearance of a well constrained lag recovery even when there is no underlying lag being detected in the signal, resulting in a so called ‘false positive’. Failure to remove false positives contaminates our population data and obscures important underlying trends (e.g. the R-L relationship). Removing false positives from a data-set is a non-trivial problem, and significant work has been done in finding effective ‘quality’ cuts to remove false positives (e.g. Penton et al. [32]).

Aliasing creates multiple modes in the marginal lag-likelihood distribution. This multimodality can lead some lag recovery algorithms to either misrepresent the lag recovery, exacerbating the problems false positives and negatives, or to fail to properly converge at all. These numerical issues are a central area of focus for this thesis, and are discussed in detail in Sections 3.3.2 and 3.2.

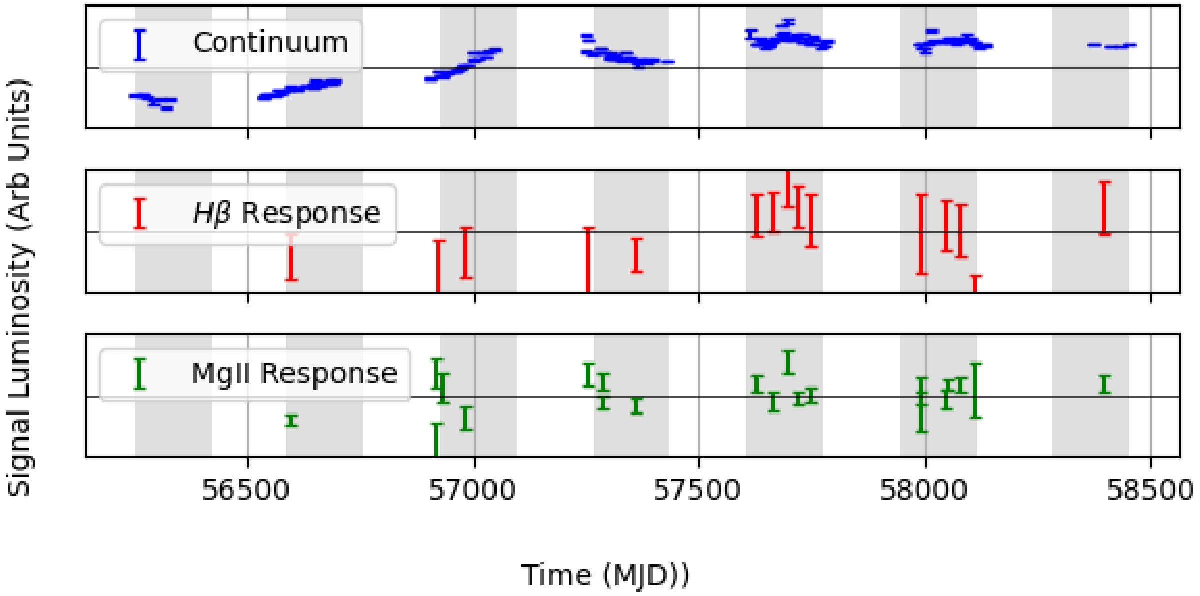

Figure 2.6: Continuum signal & response signals for Hβ and MgII for OzDES source 2925858108, showing the clear seasonal gaps.

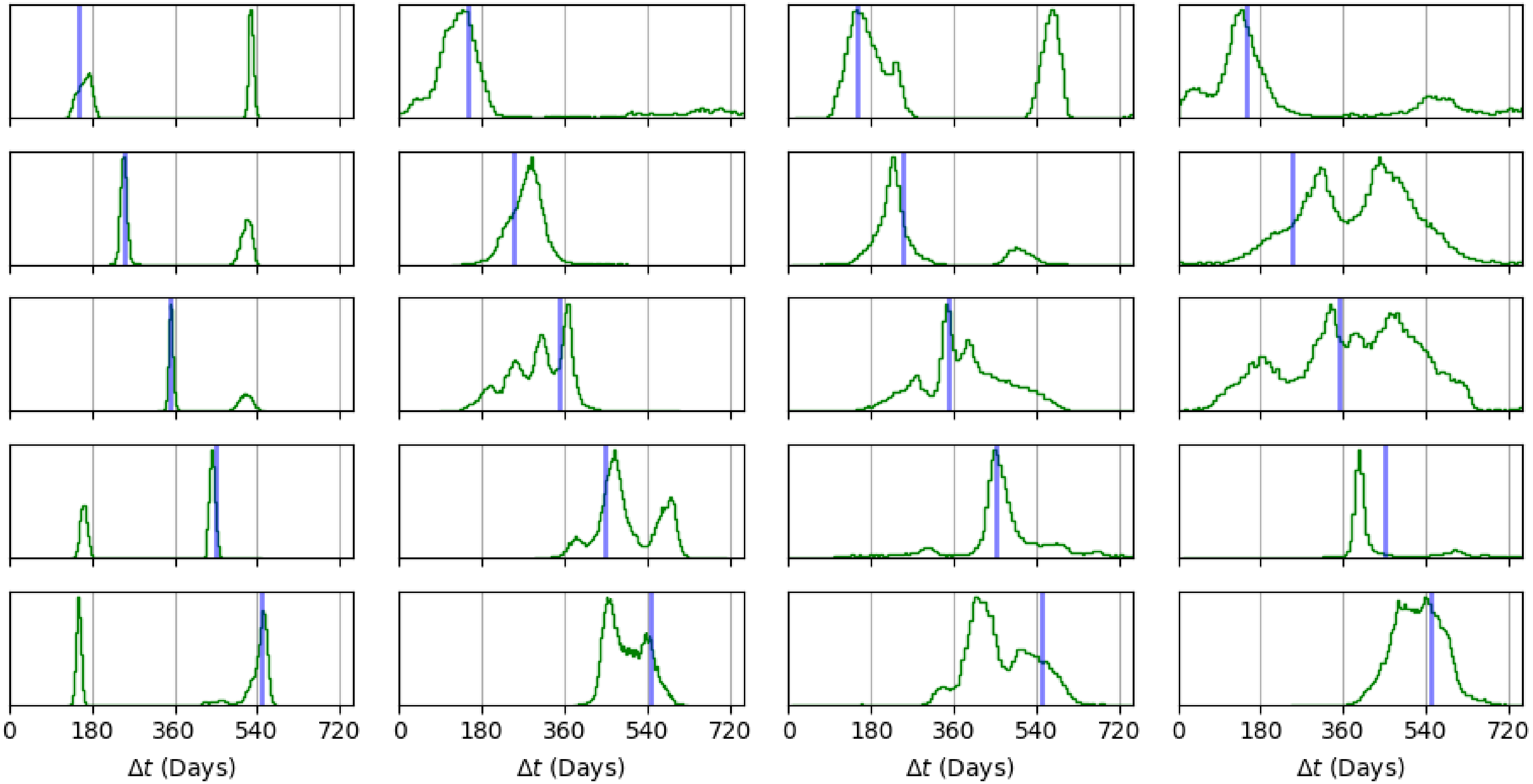

FIGURE 2.7: Lag recovery distributions from `JAVELIN` for simulated data at decreasing levels of sample quality from left to right. The true underlying lags are marked with a vertical line. These simulated sources emulate the cadence, seasonal gaps and relative measurement uncertainty of OzDES sources, but sample signals generated from time series integration of the stochastic damped random walk. Aliasing peaks can occur irrespective of the measurement quality, and become more prominent when the lag is near a $n + 1/2$ year mark, e.g. 180 d, 540 d etc [30].

2.4.1 A Conceptual Explanation of the Aliasing Problem

The aliasing effect of seasonal gaps is a central problem in reverberation mapping of the OzDES sample, and one of the main areas of focus for this report. This section presents a qualitative explanation of the source of these aliasing peaks to provide the unfamiliar reader with a conceptual sense of their behavior.

Fundamentally, all lag recovery relies on inferring and comparing the behavior of two light curves, identifying the degree to which they conflict as one of them is offset in time by some lag, and accepting or rejecting these lags based on some measure of the 'badness of fit'. In the ICCF method, the inference takes the form of interpolation and disagreement is measured from correlation function, while in GP modelling the disagreement measured in the form of a statistical tension between the measurements when fit with some set of model parameters (discussed in detail in Section 3.1).

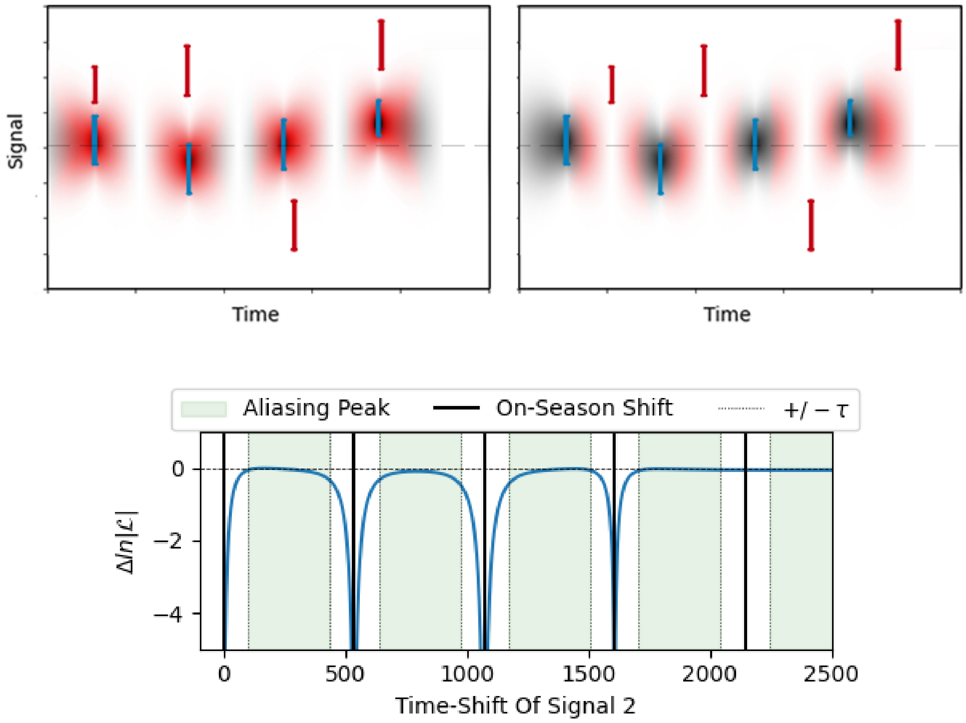

Figure 2.8: Qualitative demonstration of the source of aliasing peaks. When measurements overlap (figure top-left) disagreement between the signals is easy to observe. By shifting the two measurement series a half-step out of phase with each other (top right), the regions of disagreement (red) are moved to regions at which the signal inference (shading) is vague.
The bottom figure shows the log-likelihood of a range of lags, with lags where measurements overlap (black line) being strongly rejected, while the gaps between them (green shading) form regions of weakened rejection in the form of aliasing peaks.

An element common to all such methods is that their ability to measure the 'badness of fit' is limited by the degree to which the measurements overlap, and this overlap changes at different lags. When measurements have a large overlap, we can easily notice differences in the signals and strongly reject that lag, and when there is little overlap we have little constraining power (e.g. Figure 2.8).

This gives rise to the aliasing effect in cases where we have regularly sized and spaced gaps in the data. If the measurements are in-phase, the observations of each signal overlap closely and we can confidently reject lags that are a poor fit. Our confidence of rejection drops As we move between these on-phase lags, resulting in lags that fall within the gaps producing a local minimum in this constraining power as the measurements 'fit between' one another.

Figure 2.8 demonstrates a worst-case scenario in which correlation timescale, the span of time over which each measurement meaningfully constrains the light-curve, is much smaller than the data gaps. As such, the aliasing peaks represent a complete lack of data overlap, and so see a complete regression to the prior.

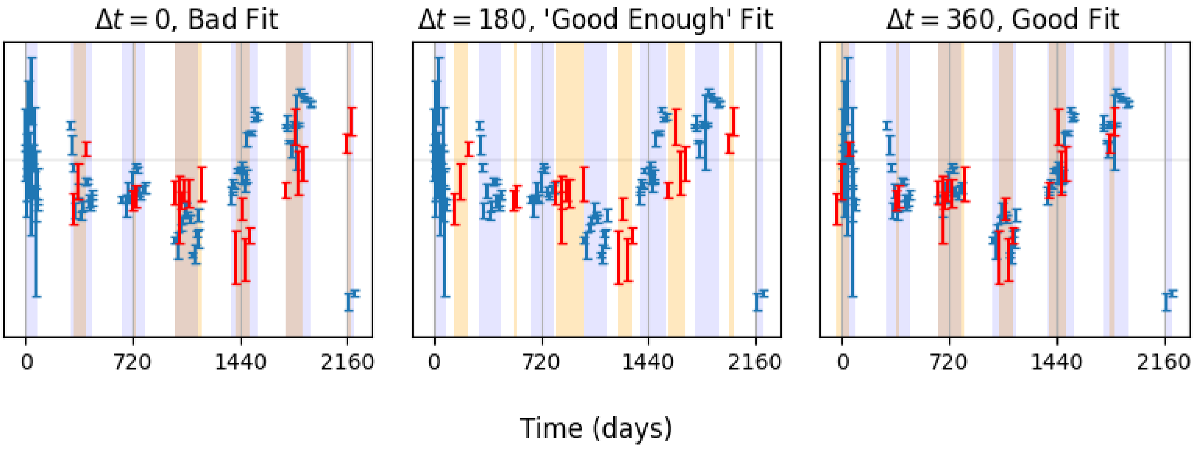

FIGURE 2.9: An example aliasing in a simulated AGN light-curve. The true lag is 360 d, and the measurement cadence and seasonal gaps emulate a source in the OzDES sample. At a lag of 180 d, the lack of overlap between measurements of the the continuum (blue) and line response (red) make difficult to determine if the two signals are in conflict or agreement with each other

The greatest danger of aliasing is that it can produce sharp, narrow peaks where no underlying lag exists. In such cases, the apparent constraint presented is entirely spurious, consisting entirely of the aliasing artefacts. Such 'false positives' are a major concern in AGN RM, as they are hard to identify and obscure true population trends [30, 32].

2.4.2 Methods of False Positive Rejection & Aliasing Removal

The low apparent brightness of AGN mean that multi-year RM campaigns are unavoidably restricted by seasonal gaps in their observations, and their results impacted by the effects of aliasing. There has been considerable prior investigation into ways of reducing these impacts, either by suppressing lags that are suspect of being associated with aliasing, or discarding entire sources of questionable reliability. Even after such quality cuts have been applied, it is still necessary to estimate the rate of false positives in the sample and the overall reliability of each recovery, and there are similarly varied methods for this as well.

A common technique of to suppress aliasing is to apply a weighting function to the posterior lag likelihood that down-weights lags associated with poor data overlap. This method, introduced by the Sloan Digital Sky Survey (SDSS; [19]) team, constructs a weighting function based on the amount of data points that overlap at each lag '$N(\Delta t)$', and then convolving this with an estimate of the signal's autocorrelation function, i.e.:

$$W(\Delta t) = \left(\frac{N(\Delta t)}{N(0)} \right)^2 * \mathrm{ACF}^+(\Delta \mathrm{t}) \tag{2.7}$$

where ACF^+ indicates setting the ACF to be 0 if it becomes negative. After applying this weighting, SDSS studies (e.g. Grier et al. [19], Homayouni et al. [21]) accept and reject lag recoveries based on whether they pass a set of reliability criteria. For a recovered lag Δt with uncertainty σ, a source is accepted if:

1. The reported lag is significantly larger than zero, $|\Delta t| > \sigma$, giving a general test for the significance of the recovery.
2. The weighting in Equation 2.7 removes no more than half the likelihood in the lag recovery, giving a measure how how heavily affected by aliasing the likelihood distribution is.
3. For lags in the region within $\Delta t \pm \sigma$ of the recovered lag, the Pearson correlation of the continuum and response signals (as measured by ICCF) must be > 0.5. This acts as a test that the two signals meaningfully co-vary at this lag, instead of it simply representing the 'least worst' fit.

SDSS also extent their their lag prior to negative lags, under the rationale that the amount of the lag likelihood distribution at these non-physical values gives a diagnostic of the overall reliability of the recovery. Rather than using this as a quality cut measure, they instead use the fraction of negative lags as an estimate of the False Positive Rate (FPR) in their sample, with the strictness of their quality cuts being tuned for FPR $< 10\%$.

A more rigorous approach is taken by Penton et al. [32] in previous work with the OzDES sample. A series of quality cuts are used to identify sources of high, low and moderate reliability. These cuts are:

1. That the main (highest) peak be within 100 d of the distribution median. This gives a measure of the degree to which the primary peak dominates the distribution.
2. That the recovered peak lag agree between `JAVELIN` and ICCF to within 30 d, protecting against method-specific artefacts.
3. That the uncertainty in this peak be < 80 d, giving a measure of the overall strength of the constraint.

The size of these cuts are chosen by simulating many AGN signals emulating of the OzDES sample's statistical properties, then determining cut sizes that best constrain the resulting R-L relationship while minimizing bias in `JAVELIN`'s lag recoveries (e.g. Figure 2.5).

By tracking the FPR of the simulated data before and after these cuts, we also gain an estimate of the FPR in our real sample without needing to retain negative lags as SDSS does. These quality cuts are extremely aggressive, with only 12% of sources expected pass all three cuts to produce an FPR of 12% in the highest 'gold standard' of reliability. Prior to such cuts, estimates of the FPR in the OzDES sample can be as high as 30% [32].

Studies often employ a mixture of these techniques. For example, in their analysis of MgII lags in the OzDES sample, Yu et al. [45] employs the weighting function of SDSS, but the quality cut procedures of Penton et al. [32]. Under these cuts, only only 25 of the 435 OzDES sources for MgII lags are retained, an acceptance ratio of only 5.7%. In this thesis, I apply my attention to the 92 MgII sources in the OzDES sample for which a second response signal is also available. Assuming a similar reliability, we would anticipate that only 6 such sources would recover reliable, meaningful lags.

2.5 Multi-Line Fitting

To date, most applications of reverberation mapping to AGN have focused on recovering the delay and/or tuning an R-L relationship for a single spectroscopic emission line, e.g, MgII by Yu et al. [45] or Hβ by Grier et al. [18]. There has been some investigation into a more complete use of the available light curves, e.g. looking for reverberations in the photometric measurements [50] or combining photometry from different band passes Grier et al. [19], but one as of yet unexplored avenue is the fitting of two lags simultaneously for sources with multiple visible emission lines. There is good motivation to do this, as leveraging more data at once should generally be expected to better constrain our lag recoveries. In particular, combining information from multiple light curves will reduce the number of lags in which seasonal gaps align, potentially mitigating the aliasing problem.

Modelling AGN signals as Gaussian Processes, the approach common to both `JAVELIN` and the alternate method presented in this thesis, extends easily to the fitting of multiple lags (e.g. Equation 3.10). This extension comes at the cost of increasing the number of model parameters in MCMC fitting, and so provides also an opportunity for testing the viability of lag recovery algorithms in higher dimensions. For this reason, we focus on such two-line fitting in this thesis for both simulated data and the data of the OzDES sample.

We expect simultaneous multi-line fitting to have the greatest impact for sources in which either both lines have similarly high signal to noise ratios, or cases in which the continuum signal has poor signal to noise such that the contribution of the line response to the signal resolution becomes significant. To identify such sources, we follow the lead of Penton et al. [32] and use the 'fractional variability' as a measure of a time series signal to noise ratio, though with the minor adjustment of using the weighted estimate of the signal mean instead of unweighted.

$$\mathrm{SNR} = \frac{1}{\bar{y}}\sqrt{\frac{1}{N}\sum_i (y_i - \bar{y})^2 - E_i^2} \quad \bar{y} = \frac{\sum_i \left(\frac{1}{E_i^2} y_i^2\right)}{\sum_i \left(\frac{1}{E_i^2}\right)} \tag{2.8}$$

Of the 92 two-line sources in our sample, many are characterized by extremely high uncertainty on at least one of the signals, such that the summation in the square root of

Equation 2.8 becomes negative. For such signals, we treat the SNR as being zero. These SNR estimates are available for all sources in Appendix C.

Using this estimate, we identify 6 sources in which both response lines have similarly high signal to noise, and 10 sources in which at least one response line has an SNR 4 times that of the corresponding continuum signal. Two line sources are necessarily drawn from redshifts at the edge of those line's visible range, and so most of the response lines in our sample, particular those for CIV, have high relative measurement uncertainty and poor signal to noise ratios.

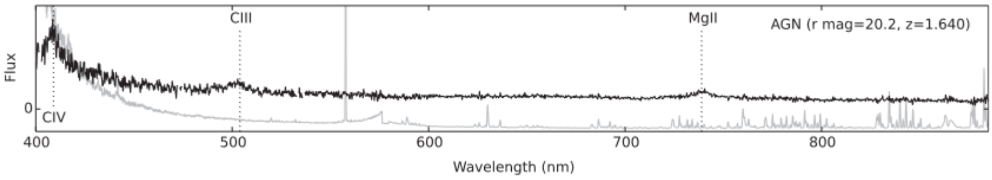

Figure 2.10: Spectrum for source near the edge of redshift range for CIV visibility, from Figure 2 of Yuan et al. [46]. Measurements are noisier at the edge of the wavelength domain, and so measurements of the line strength are similarly poorly constrained. In this example, proper measurement of the CIV line would be difficult.

3

Modelling & Methods of Lag Recovery via Gaussian Processes and MCMC

In this section, we provide an overview of the main techniques used by programs like `JAVELIN` to perform reverberation mapping through the modelling AGN light curves as stochastic processes. This chapter is intended to help the reader familiarise themselves with the general statistical concepts at play in existing methods, while specific details on the numerical techniques used in this thesis are available in Chapter 4.

3.1 A General Approach to Gaussian Process Modelling

In this section, we provide a general overview of the modelling of a time series signal as a GP, including the methods by which this modelling can be used for lag recovery and signal constraint / interpolation. Such methods have been applied in astrophysics for some time [37], and form the statistical core of programs like `JAVELIN` and `CREAM`. Details of this procedure as applied specifically to the AGN signals are in section 3.1.1.

The driving signals of an AGN are inherently stochastic, but not entirely random. Instead of producing meaningless noise, their signals are structured in that they take time to vary from high to low: measurements at nearby times are correlated (Figure 3.1). If we observe the signal at some time, we also gain information about what that signal was doing at nearby times. This is described by an auto-covariance function that relates how similar the signal is to itself at any two points in time:

$$\Phi_{i,j} = \Phi(t_i, t_j) = \langle y(t_i), y(t_j) \rangle \tag{3.1}$$

This covariance function is the defining feature of the GP and encodes all of its statistical behaviour. Different GPs will have different covariance functions of varying complexities, but for our purposes this always a symmetric [1], positive definite function. For a signal with well defined variability $\sigma^2 = \langle y(t), y(t) \rangle$, this can be in terms of the dimensionless 'correlation function':

$$\Phi_{i,j} = \Phi(t_i, t_j) = \sigma^2 \phi \left(\frac{|t_i - t_j|}{\tau} \right) \tag{3.2}$$

In this way, a covariance function requires at least two parameters: the amplitude of the signal's variability 'σ' and a correlation timescale 'τ_d' describing how quickly the variations occur.

If we have some time series of signal measurements $y_i = y(t_i)$, this covariance means that each measurement also encodes information about the signal at nearby times, including the other neighbouring measurements. A convenient way to think of this is to imagine some underlying signal '$u(t)$' to which our observations are linearly related:

$$y_i = \Phi_{i,1} u_1 + \Phi_{i,2} u_2 + \dots \tag{3.3}$$

Defining a 'covariance matrix' $C_{i,j} = \phi_{i,j}$, this linear system can be used to evaluate the likelihood, $\mathcal{L}$, a measure of how well the GP, for a given set of parameters θ, describes our observations:

$$P(\theta|\vec{y}) \propto \mathcal{L} \propto \frac{\exp(-\frac{1}{2}\vec{y}^T C^{-1} \vec{y})}{\sqrt{|C|}} \tag{3.4}$$

The above applies only to an ideal Gaussian process with zero mean and no measurement uncertainty. In practice, our observations are of a signal which is a combination of the structured Gaussian process '$s(t)$' and the uncorrelated measurement noise '$n(t)$', i.e. $y(t) = s(t) + n(t)$. This noise manifests as extra diagonal terms in the covariance matrix. For a set of 1σ measurement uncertainties $\{E_i\}$, the covariance matrix becomes:

$$\begin{gathered} C = S + N \\ N_{ii} = E_i, \quad S_{ij} = \Phi(t_i, t_j) \end{gathered} \tag{3.5}$$

To account for uncertainty in the mean of each signal, there are two approaches. The first is to simply subtract off the mean and introduce it as another model parameter:

$$P(\theta|\vec{y}) \propto \mathcal{L} \propto \frac{1}{\sqrt{|C|}} \exp \left(-\frac{1}{2} (\vec{y} - \bar{y})^T C_\perp^{-1} (\vec{y} - \bar{y}) \right) \tag{3.6}$$

---

[1] Many sources describe correlation functions in the form $\phi(\Delta t), \Delta t = |t_i - t_j|$. Here, this is not done to avoid possible confusion with the signal lags, also denoted Δt.

or all possible values of the mean may be marginalized over analytically as is done in `JAVELIN` and the work of Zu et al. [48]:

$$P(\theta|\vec{y}) \propto \mathcal{L} \propto \frac{\exp(-\frac{1}{2}\vec{y}^T C_\perp^{-1} \vec{y})}{\sqrt{|C| \times |L^T C^{-1} L|}} \tag{3.7}$$

where

$$\begin{aligned} C_\perp^{-1} &= C^{-1} - C^{-1} L C_q L^T C^{-1} \\ C_q &= (L^T C^- 1 L)^{-1} \\ L &= [1, 1, ...]^T \end{aligned}$$

These covariances can, for a set of fixed parameters, be used interpolate the statistical behaviour of the signal between measurements (e.g. Figure 3.1). For some set of 'data' measurements $\vec{y}_d$, and seeking to recover a signal curve $\vec{y}_c$, the mean and variance are:

$$\begin{aligned} \vec{s_c} &= S_{cd}^T C_{dd}^- 1 \vec{y} \\ \Delta \vec{s}_c^2 &= S_{cc} - S_{cd}^T C_{dd}^- 1 S_{cd}^T \end{aligned} \tag{3.8}$$

Where subscript '*cc*' indicates the square covariance matrix of the interpolated time series, '*dd*' is the same for the measurement times and '*cd*' is the rectangular covariance matrix connecting the two.

This procedure is for only a single signal, but we can extend it to cases of 'driving' and 'response' signal, such as the continuum and reverberating emission lines of an AGN. Provided the response is linear, i.e. that the driven signal '$y_a(t)$' is related the continuum signal '$y_c(t)$' via a transfer function '$\psi_a(t^{'})$':

$$y_a(t) = \int \psi_a(t - t') \cdot y_c(t') dt' \tag{3.9}$$

The covariance between any two measurements 'y_i' and 'y_j' on any two signals 'a' and 'b' that are responses to a signal with auto-covariance function $\Phi_c(\Delta)$ is:

$$\Phi_{i,j} = \int \int \Phi_c(t_i, t_j) \psi_a(t_i - t^{'}) \psi_b(t_j - t^{''}) dt' dt^{''} \tag{3.10}$$

This can be used to recover auto-covariances within a response signal by setting $a = b$, and covariances with measurements on the driving continuum by using $\psi_c(t^{'}) = \delta(t^{'})$.

3.1.1 The Damped Random Walk

The continuum signals from AGN have been found to match well to the statistical properties of the 'damped random walk' (DRW), a continuous first order auto-regressive ($CAR(1)$) process in which the signal 'wanders' up and down randomly (Figure 3.1), but with a restoring force returning it to a baseline with a characteristic timescale: [24, 28, 50]

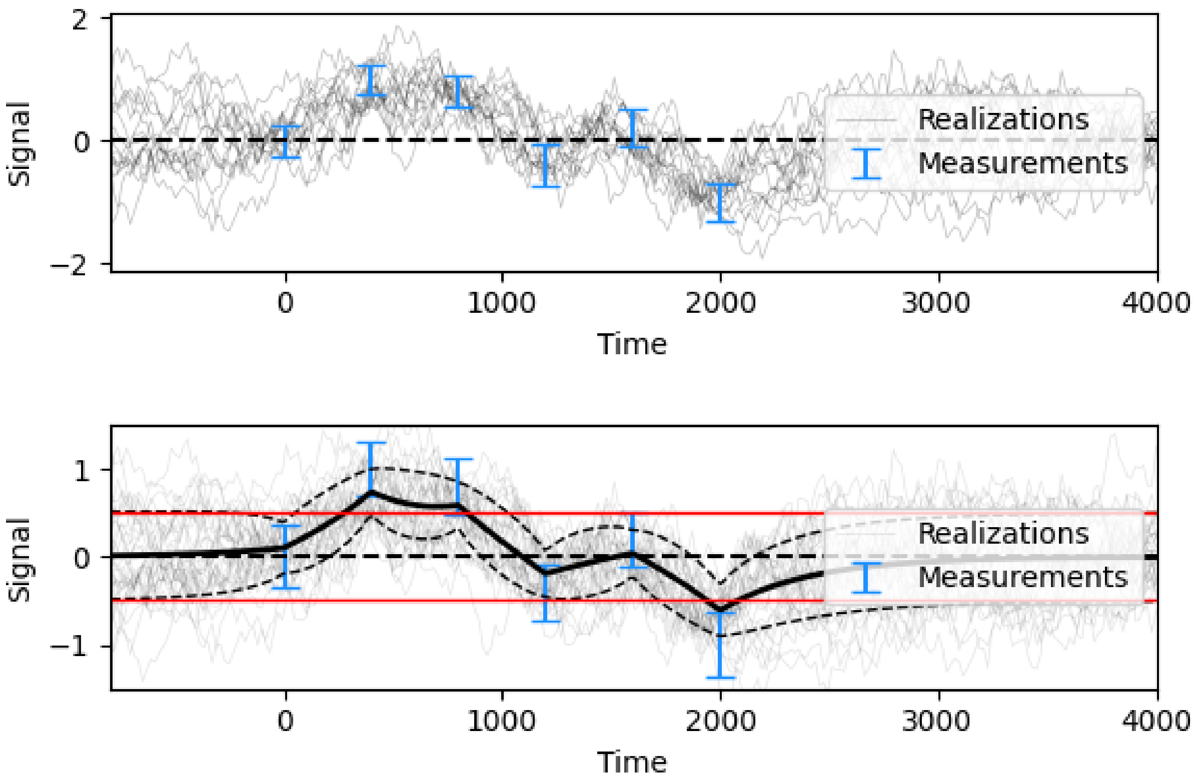

Figure 3.1: Example of an stochastic light curve interpolated per Equation 3.8. The top panel shows the way that measurements (blue) constrain the possible paths of the stochastic signal (thin black lines), while the bottom panel shows the analytical mean and variance of all such possible paths (black dashed lines). Far from measurements, theses constraints devolve to the mean and variance of the signal (red lines).

$$\frac{dy(t)}{dt} = \frac{-y(t)}{\tau_d} + \frac{\sigma_c}{2\tau_d} dW \tag{3.11}$$

where dW is the continuous stochastic element with variance 1, σ_c is the signal variability amplitude that determines the magnitude of the excursions about the baseline and τ_d is the timescale of the 'damping' force which in turn sets the timescale of the entire process (Figure 3.1).

The defining feature of the DRW is its exponential covariance function, for which the damping timescale also acts as the correlation timescale:

$$\Phi_{DRW}(\Delta t) \ = e^{\frac{-|t_i - t_j|}{\tau_d}} \tag{3.12}$$

To model the response of the emission lines, a rectangular 'top-hat' transfer of width w_a and centered at characteristic delay Δt_a is used. This has the effect of making the response signal follow the continuum, but 'shifted, scaled and smoothed' (e.g. Figure 2.1)

$$\psi(t') = \frac{A_a}{w_a} \begin{cases} 1, & \text{if } t' \in [\Delta t - \frac{w}{2}, \Delta t + \frac{w}{2}] \\ 0, & \text{otherwise} \end{cases} \tag{3.13}$$

Such modelling then has $3 + 3n$ parameters, where 'n' is the number of response lines / lags being fit, or $2 + 2n$ if marginalizing over the signal means as per Equation 3.7. As lag recoveries are found to be generally insensitive to the smoothing timescale [42], it is common practice to fix this parameter in order to simplify the system [19, 21, 32].

3.1.2 Lag Recovery via Monte Carlo Markov Chains

Provided the underlying signal follows the behaviour of a damped random walk, the process for finding the marginal lag likelihood distribution is well established. However, it requires the inversion of a matrix that scales with the total number of observations (Equations 3.7 & 3.6), which can be on the order of ≈ 100, meaning the each evaluation is numerically expensive. Such inversion is not prohibitively costly for a small number of evaluations, but the many parameters of our models make an exhaustive grid-integration intractable.

Instead, programs like `JAVELIN` make use of MCMC techniques, in which a 'chain' of evaluations navigate towards regions of good fit, and then sample this region of parameter space. MCMC techniques are numerous and varied, but typically involve a 'burn-in' or 'warm-up' phase, used to draw in towards the regions of highest probability, and then a main 'sampling phase' to explore this region. If the chains have properly converged, and are using an appropriate sampler, these samples will be distributed proportional to the posterior likelihood. The density of the samples then allows recovery of the location and uncertainty of the best fit parameters.

The defining feature of an MCMC algorithm is how the 'proposal' is generated and accepted at each step in the chain. The effects of aliasing produce a distribution with multiple highly separated regions of high probability, which can lead to a number of sampler-specific artefacts that affect lag recovery. We discuss these samplers and their suitability to RM in Section 3.3.2 and Section 4.1.

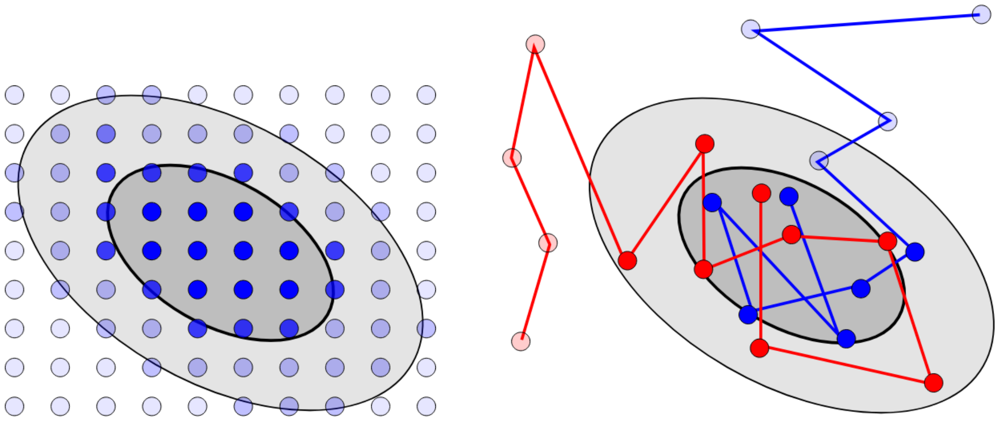

Figure 3.2: An illustration of the general procedure for MCMC integration (right), which evaluates only the high-likelihood regions of interest, compared to an exhaustive grid-search (left), which wastes many evaluations in low probability space.

3.2 Geometry of The Posterior

When using GP models to perform lag recovery, we describe the properties of the AGN's light curves with a number of other parameters in addition to the lag, leading to a high dimensional likelihood function. Much existing discussion of this distribution focuses solely on the broadest features in the marginalized lag-domain, doing so to the exclusion of finer details and behaviour in higher dimensions. These details are of practical interest, as they influence what statistical and numerical techniques are applicable to reverberation mapping.

Most model parameters are 'well behaved', with distributions that are smooth, uni-modal, uncorrelated with other parameters and are easily encapsulated by a sufficiently broad prior (e.g. Figure 3.5). Exceptions to this are strong correlations between the amplitude and timescale of the continuum signal, irregular and multimodal behaviour of the lag parameters, and the poor conditioning that arises from low signal to noise measurements. In this section, we examine each of these properties and the challenges they present.

Aliasing & Lag-Space Properties

Though the posterior distribution is smooth and uni-modal in all other parameters, the same is not true of the signal lags. The most well studied feature in the lag domain is the seasonal aliasing peaks: regular spikes in likelihood every 180 d (e.g. Figure 3.3). These peaks occur as a result of the observation windowing function, and so are always present, but their relative prevalence is dependent on the signal variations. Faster cadence and lower uncertainty in our observations narrow the width of these peaks, while the specific variations in the signal during our observations affect their relative prevalence. Though a low signal to noise leads to an entirely unconstrained lag, a single spurious aliasing peak can emerge even if there is no meaningful lag encoded. Such 'false positives' are difficult to distinguish from meaningful recoveries, and occur most often when the true lag is near an aliasing peak such that the lack of data overlap make it to less likely to emerge as the dominant feature of the distribution [30].

The aliasing peaks associated with seasonal gaps are well studied, but there is little discussion in past RM work of a second aliasing effect associated with measurement cadence. Aliasing arises from any regular gaps in measurement, and so OzDES-like data produces low likelihood peaks of roughly weekly spacing, corresponding to the cadence of the continuum observations. When fitting two lags simultaneously, a similar spacing occurs diagonally in lag-lag space at 30 d intervals. Aliasing peaks are heavily suppressed when the correlation timescale is larger than the time-series spacing of the data, and so these 'weekly' peaks are extremely shallow to the point of not being an easily visible feature when viewing likelihood distributions or MCMC results.

Cadence-scale aliasing is an inherently low likelihood effect, but can nonetheless produce significant numerical obstacles. When exploring parameter space, either by MCMC or optimization, any method to vulnerable to pinning at local optima is at risk of becoming trapped in these modes (e.g. Figure 3.3, top). Lags where measurements align create 'walls' of extremely low likelihood, which are both deep and have extremely sharp gradients due to their small width, meaning any walker that navigates with the likelihood function's local gradients will have difficulty moving through the 'flat' space between aliasing modes.

The multi-modal distributions that arise from aliasing and the roughness of the inter-modal space present obstacles for many sampling methods. Properly handling these issues is one of the main goals of this thesis, and I discuss solutions to these problems in Chapter 4.

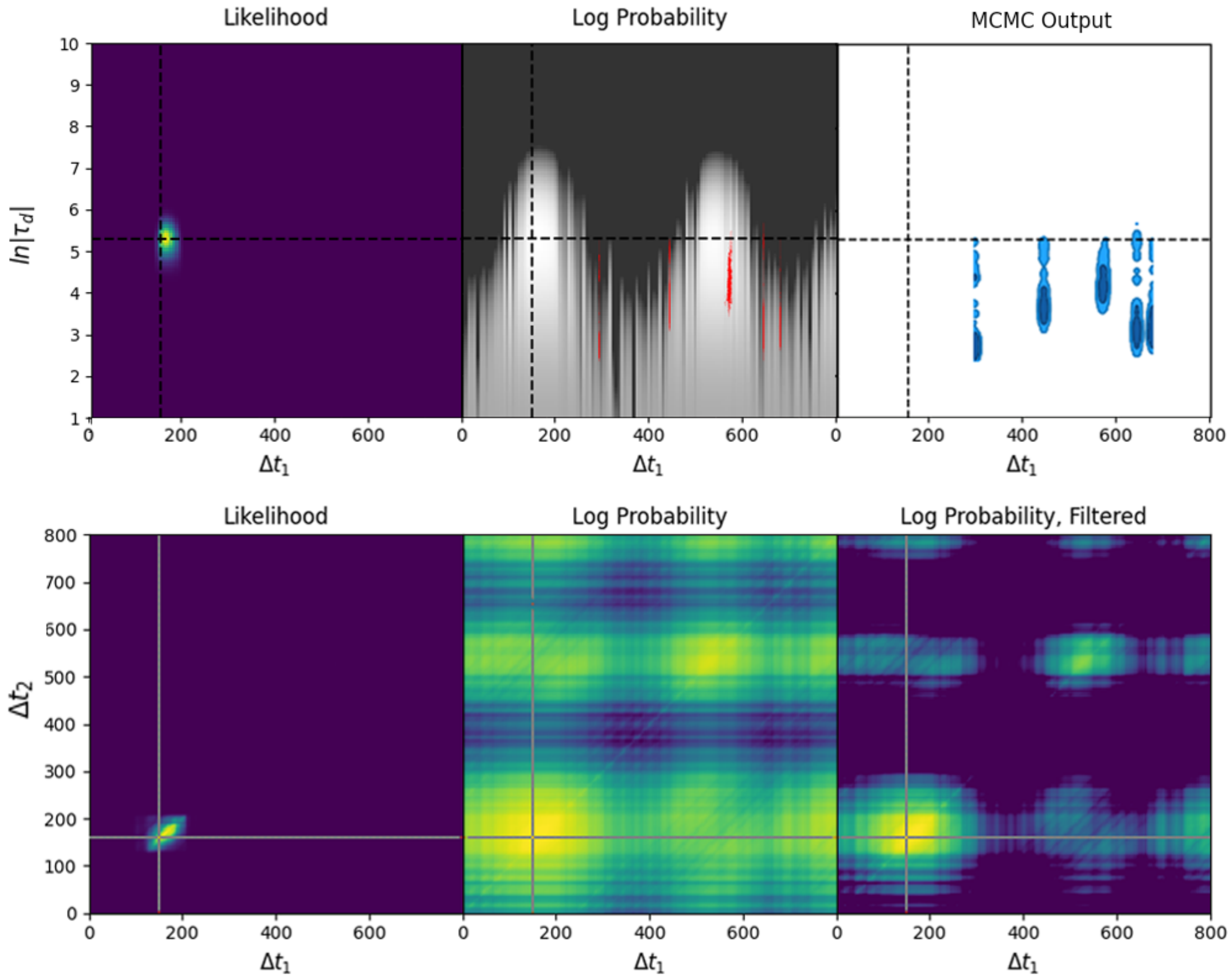

Figure 3.3: Demonstration of the roughness in the likelihood function arising from measurement cadence. All figures show the posterior for a simulated sources with true lags at 150 d. The top row shows the geometry in the dimensions of correlation timescale and one of the AGN lags, while the bottom shows the geometry in lag-lag space.
Top middle shows the failure of MCMC chains (red) to navigate through the rough the inter-modal parameter space, causing a complete failure of the MCMC output (top-right) to recover the true likelihood distribution (top left).
Bottom middle shows the high contrast log-likelihood function in lag-lag space, in which the regular striations arising from the weekly-scale cadence can be seen, as well as the diagonal striations from the monthly cadence of the spectroscopic measurements. Bottom right shows the same as bottom middle, but with the the contrast adjusted to more clearly show the affect this roughness has on the shape of seasonal aliasing modes.

Degeneracy of Continuum Signal Properties

In fitting the DRW model to an AGN, most parameters are uncorrelated, but there is an exception in the timescale and amplitude of the continuum signal. The DRW has decreasing power at higher frequencies, and so there is an ambiguity as to whether any variations we observe represent the full behaviour of the signal, or are a small fraction of a much larger, slower variation. This results in the timescale τ_d and amplitude of the signal σ_c being correlated, and a long tail in the posterior distribution for these continuum properties.

In the limit of $\tau_d \rightarrow \infty$, this correlation collapses to a complete degeneracy between the parameters, as can be seen in Equation 3.11. This limit of infinitely long damping timescale of corresponds to an un-damped random walk (UDRW), which has only a single parameter to describe its variability that scales with $\frac{\sigma}{\sqrt{\tau_d}}$. Because the correlation timescale can span several orders of magnitude, it is common practice to parameterize both it and σ_c logarithmically. In log-space, this high-τ_d degeneracy results in a diagonal band with a 1:2 slope of $ln|\tau_d|$ against $ln|\sigma_c|$ that extends to the edge of the prior space, (e.g. continuum properties of Figure 3.5, right) with a primary 'lobe' at low τ_d / σ_c corresponding the best fit for DRW-specific properties. The band is always present in the posterior distribution at low likelihood, but only becomes a significant feature when the observed signal variations fail to strongly demonstrate damping.

It is important to note that the high τ_d ends of the band are still entirely reasonable results for modelling the signal. The UDRW is still a valid Gaussian process for modelling AGN variability, one used by `JAVELIN` alternative `CREAM`, and lag recoveries are not particularly sensitive to the specific choices of GP[42]. However, it should be noted that it is impossible to choose a prior that fully encapsulates the band, as it extends indefinitely, and this can introduce artefacts and inefficient sample rejection rates in some samplers (e.g. Section 3.3.2).

Effect of Signal Quality

For post parameters, we can easily enclose high likelihood modes with a reasonably broad prior range. There is a common exception to this when fitting the amplitude of low signal to noise response lines, a case that commonly occurs in OzDES sample for the low redshift CIV lines that I examine in this thesis. In such cases (e.g. Figure 3.4) the signal variation is small compared to the measurement uncertainty, and the 'best fit' solution is a flat line of zero amplitude at the signal mean.

Much as the continuum band borders the upper range of the search boundary, such cases result in the amplitude parameter bordering the lower $A = 0$ boundary of our prior space. The physical basis for the response line requires that the amplitude be strictly positive: a positive or negative variation in the continuum should be expected to produce a variation of the same sign in the response line. As with the continuum, contours bounding the search space is not an inherent problem, but can lead to sampling artefacts and high rejection rates. Such cases can also lead to issues when using optimization, e.g. to find the peak likelihood of the non-lag parameters in a scan.

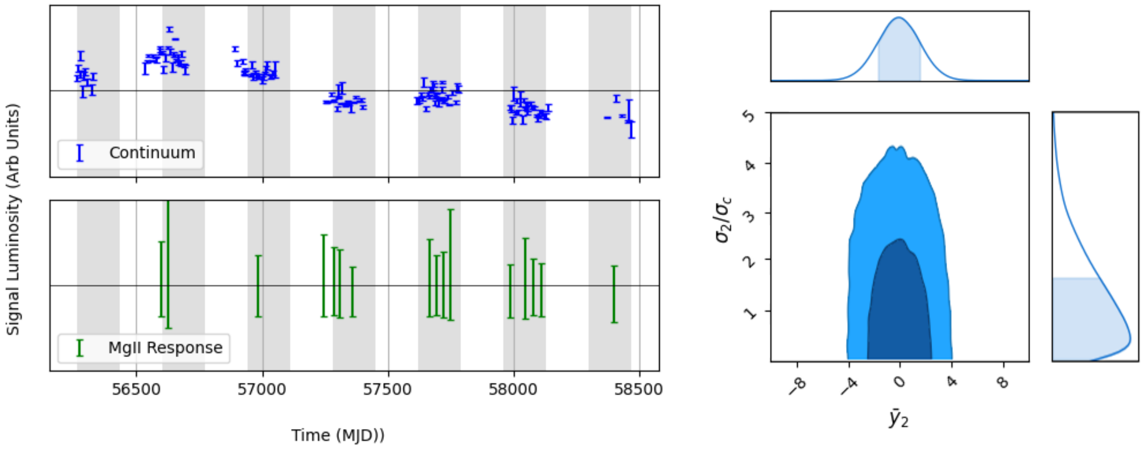

Figure 3.4: Example signal (left) and recovered distribution of the response line mean and amplitude (right) for a low SNR signal in the OzDES sample.

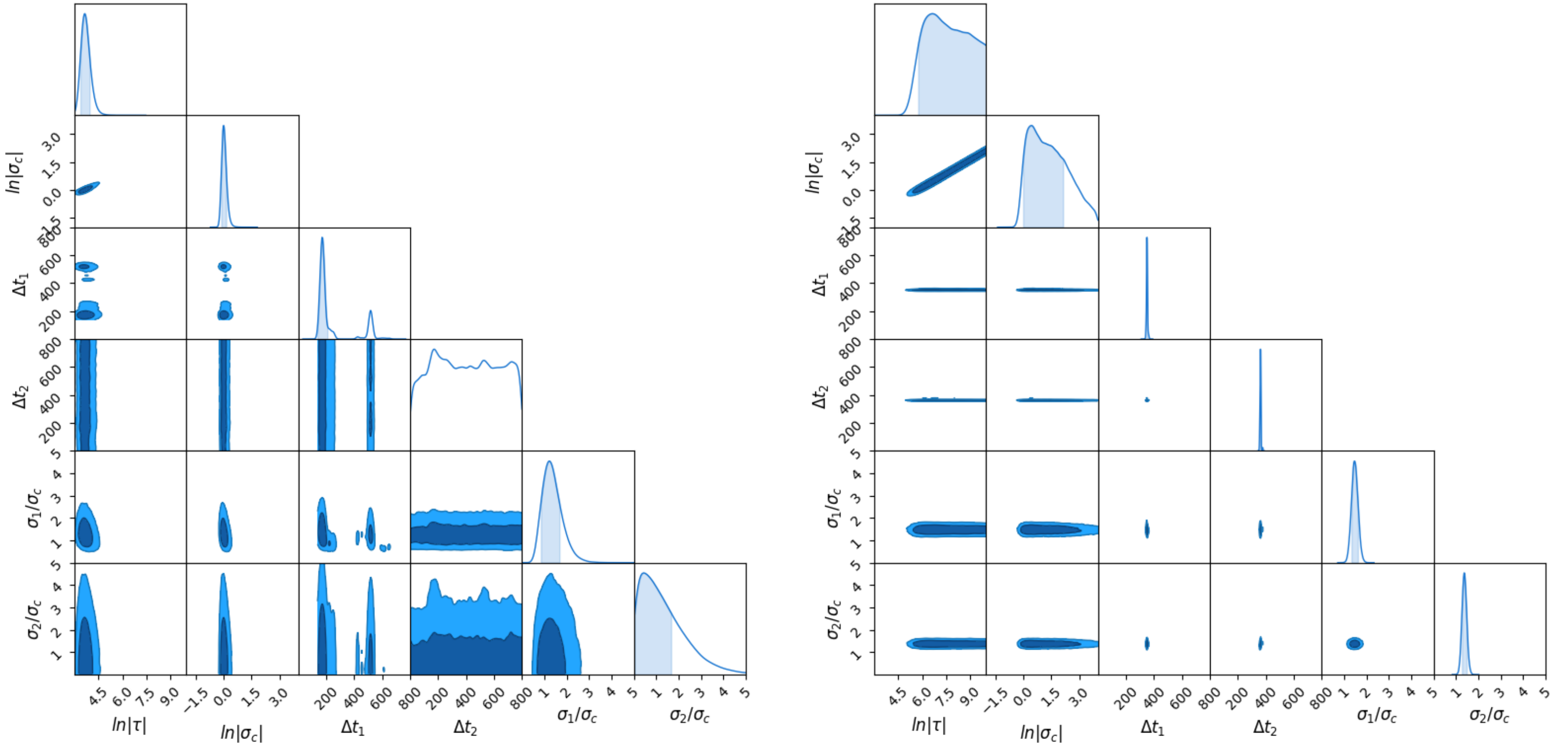

Figure 3.5: Typical corner plots of a 2-line fit, demonstrating the key features of the posterior geometry. Figure left is for OzDES source 2925373860, and figure right is for a simulated AGN with an underlying lag of $\Delta t_1 = \Delta t_2 = 360$ d.
For the low quality source, we can clearly see the multiple peaks from aliasing on signal 1, while signal 2 is completely unconstrained. For the high quality simulated source, we can see that all values are well constrained and mostly uncorrelated, with the only exception being the degeneracy in the continuum timescale τ_d and variability σ_c.

3.3 Behaviour of `JAVELIN` & Complications in Lag Recovery

To date, our most complete statistical model of AGN variability is that of the the damped random walk with smoothing in its response signals, described in detail in Section 3.1.1. The software package `JAVELIN`, which makes use of this model for its lag recoveries, has established itself as the current state of the art tool in AGN reverberation mapping. In this section, we introduce `JAVELIN` and the numerical methods it makes use of, and identify the failure modes that that arise when these methods are employed against AGN with seasonal gaps in their measurements.

3.3.1 The Affine Invariant Sampler & `emcee`,

For its MCMC runs, `JAVELIN` makes use of `emcee` from Foreman-Mackey et al. [15], an implementation of the 'Affine Invariant Ensemble Sampler' (AIES)[17]. The AEIS is a gradient-free sampler that operates by tracking an ensemble of live walkers across parameter space, using the distribution of this ensemble to generate proposals for each walker at each iteration. The central principle of `emcee` is the 'stretch-move': at each iteration i, a walker at position X_k^i selects a random 'complementary walker', X_j^i, using the line connecting the two as a proposal direction along which the proposal, Y_k, is made (e.g. Equation 3.6):

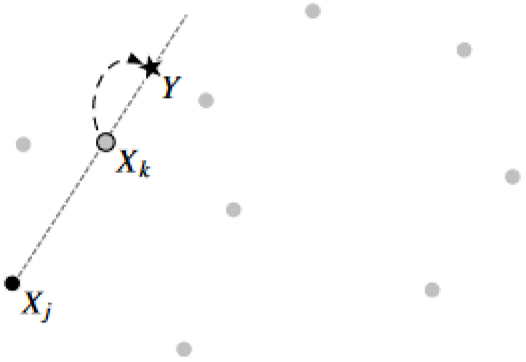

Figure 3.6: Demonstrating of the stretch-move proposal, from figure 2 of [17]

$$Y_k = z \cdot X_k + (1-z) \cdot X_j \tag{3.14}$$

Here, z determines the distance travelled in the proposal direction, with $z \in [0,1]$ moving towards the complementary walker, while $z > 1$ means moving away. z is drawn from the distribution in Equation 3.15, with the typical value of $a = 2$ meaning that a walker can, in any single step, move anywhere from doubling to halving the distance between it and the complementary walker.

$$g(z) = \begin{cases} \frac{1}{\sqrt{z}} & \text{if } x \in [\frac{1}{a}, a] \\ 0 & \text{Otherwise} \end{cases} \tag{3.15}$$

This proposal is accepted or rejected in the typical MCMC fashion but with a factor $z^{N_{Dim}-1}$ to account for the fact that the one dimensional proposal direction is acting as a proxy for shells in N_{Dim} dimensional parameter space:

$$P_{\text{Accept}} \propto min(1, z^{N_{\text{Dim}}-1} \cdot \frac{\mathcal{L}(Y_k)}{\mathcal{L}(x_k^i)} \tag{3.16}$$

Once properly converged, this sampling method is guaranteed to recover the mean and covariance of the posterior distribution [15].

Where `emcee` fails is in the case of multi-modal distributions, for which it cannot properly converge due to its proposal generation not easily allowing chains to migrate between modes. When updating a single walker, a complementary walker chosen from another mode will allow a proposal to be made, at furthest, halfway between the two. For well separated modes this proposal will be made in the low-likelihood inter-modal space, and so will be overwhelmingly rejected. If a complementary walker is chosen within the same mode, the majority of cases will still cause samples to be rejected by virtue of the proposal direction. A proposal can only be made in another mode if a complementary walker is selected from within the same mode, the proposal direction aligns with the adjacent mode, and the proposal is drawn at the very edge of its possible range. If the modes are narrow compared to their separation, even this mechanism becomes impossible. The result then is that walkers have no means of moving away from the first local optima they encounter, causing them to become 'pinned', oversampling low likelihood modes and failing to properly converge. In the proceeding section, we examine the impact this has on lag recovery.

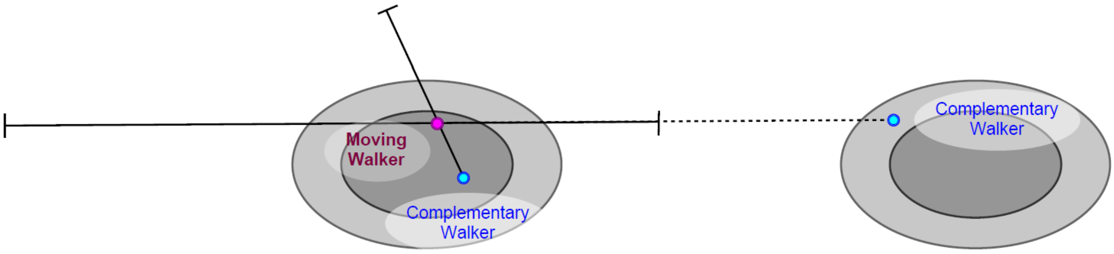

FIGURE 3.7: A demonstrating of the failure mechanism of the stretch-move sampler in multi-modal distributions. A given walker (purple) cannot migrate between modes unless a complementary walker (blue) is chosen such that it can 'stretch' across the inter modal gap.

The stretch-move AIES is also known to encounter issues in higher dimensional problems [23]. `emcee` can have poor convergence at as low as 10 parameters, and can fail to meaningfully converge at all in higher dimensions. This is not a major concern for lag recovery of one or two AGN signals, which involves 4-9 parameters at most, but does exclude `emcee` from being applied to higher dimensional models, such as population stacking (e.g. Brewer and Elliott [7]) or even to more complexly parameterized models for a single AGN.

3.3.2 Failure Modes of `JAVELIN`

To date, the Damped Random Walk (DRW; 28) represents our most complete model of AGN variability, and the program `JAVELIN` has established itself as the standard tool for lag recovery using this model. However, when dealing with the multi-modal distributions associated with seasonal gaps, `JAVELIN` is subject to numerical artefacts that can severely affect the quality of lag recovery. In this section, we examine the origins of these failure modes and explore their impact on reverberation mapping.

Pinning of Chains to Local Optima

As discussed in Section 3.3.1, `JAVELIN`'s MCMC sampler `emcee` encounters significant issues with multimodal distributions. The particular way that the stretch-move sampler generates proposals leaves little opportunity for MCMC chains to migrate between local optima, causing the them to become 'pinned' to the first mode they counter during burn-in.

If the modes are sufficiently compact, i.e. with sharp borders, the proposal generation in Equation 3.15 makes it extremely unlikely that a proposal will be drawn from within another mode. If the complementary walker for generating the proposal is selected from within the same mode as the moving walker, the the maximum possible distance the proposal can be made at is twice the width of the mode as measured from its leading edge. If the complementary walker is chosen from a separate mode, the maximum proposal will be drawn from roughly halfway between the two modes. In both cases, unless the modal width is of similar scale to the inter-modal separation, it becomes vanishingly unlikely that a proposal will be made that takes a walker from one mode to another in a single jump. Instead, a successful crossing between mode requires that at least one sample be accepted in the low-likelihood inter-modal 'flat space', which has a vanishingly small probability.

As a result of chains being unable to move between modes, they are pinned to the first local optima they encounter. As the ensemble is unable to properly converge, samples are instead drawn from theses modes at a rate disproportionate to their likelihood. In reverberation mapping of OzDES-like signals, this leads to `JAVELIN` over-sampling otherwise insignificant aliasing peaks, drastically inflating them relative to their true prevalence in the underlying posterior distribution (e.g. Figure 3.8). This obscures the 'true' peak in the distribution, leading to cases where a meaningful lag is unable to be recovered.

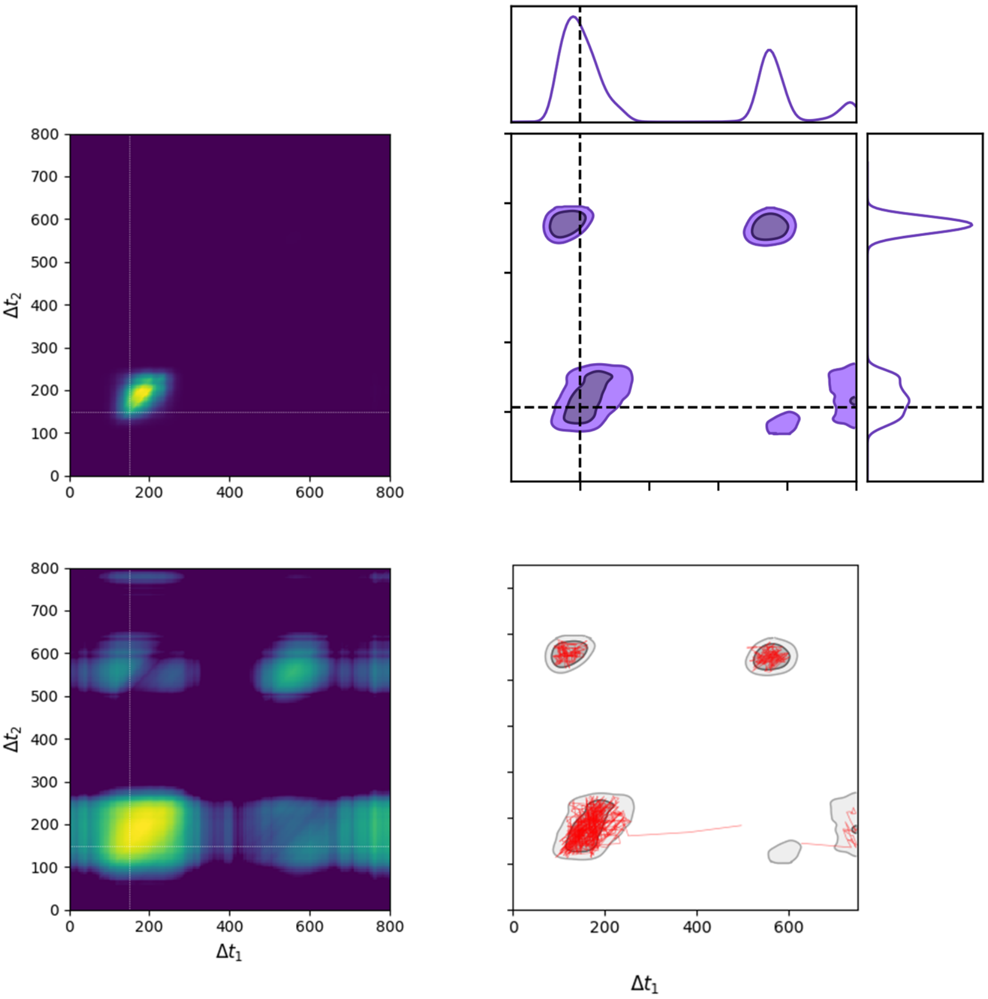

Figure 3.8: Demonstration of `JAVELIN`'s failure mode in multi-modal distributions. All plots are for a simulated AGN signal with a true lag at $\Delta t_1 = \Delta t_2 = 180\,\mathrm{d}$, mirroring the cadence and seasonal gaps of an OzDES source. Top left shows the true distribution as recovered from an exhaustive grid search, while top right shows the likelihood distribution as recovered by `JAVELIN`. Bottom left shows the log-likelihood of the true distribution, showing the locations of the low likelihood aliasing modes. Bottom right shows the post burn-in walker paths of `JAVELIN`, demonstrating the way the way in which they fail to migrate between the modes and over-sample the aliasing modes.

In addition to distorting the lag recovery, this issue also leads to a wasteful number of rejected samples being proposed. In the documentation for emcee, Foreman-Mackey et al. [15] notes that acceptance ratios below 25% indicate the `emcee` has failed to converge, likely being a sign of multimodality. A typical `JAVELIN` run on an aliased AGN signal can return acceptance ratios in the range 3% - 20%, demonstrating the program's poor suitability to signals of this nature.

Over-Sampling Near Prior Boundary

Though the `JAVELIN`'s failure when encountering multimodality is its most significant issue, there is a secondary concern with its sampler. `emcee`'s proposal method performs poorly when high likelihood modes extend to near the prior boundary, as proposals are frequently made beyond this boundary and then immediately rejected. When a proposal is rejected, the chain's previous position is retained as a sample. This leads to chains 'bunching up' and over-sampling near the edge of the prior space (e.g. Figure 3.9). For most parameters, this can be avoided with a sufficiently broad prior range, but, as we note in Section 3.4, there are cases in which high-likelihood contours extend to the edge of physically possible values, or extend indefinitely. This is not a major concern in lag recovery, as the affected parameters do not strongly impact lag recovery [43].

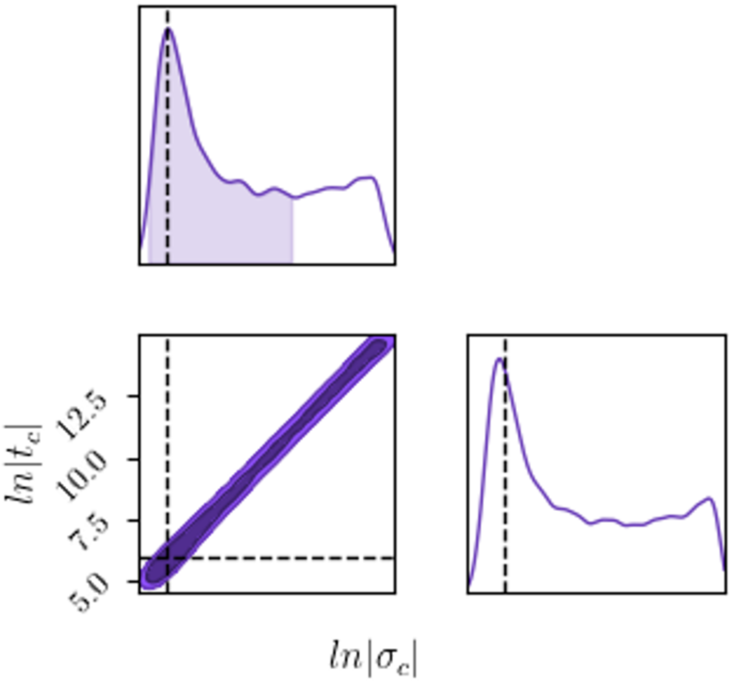

Figure 3.9: Example of continuum property contours for for a signal fit by `JAVELIN`. Results are for simulated data, with the truth values marked in dashed lines. The primary lobe in the bottom left represents a DRW solution, while the flat band extending to $\tau \to \infty$ is the degenerate UDRW solution. JAVELIN, showing the 'bunching up' near the boundary, causing a second peak at the upper edge of the domain due to `emcee` sampling artefacts.

In the case of the degenerate band in the continuum properties (e.g. Figures 3.9 and 3.5 (right)), `JAVELIN` has an optional ad hoc method for constraining this band in which it recovers continuum properties from a fit of the the continuum signal only, which has less confounding factors and so delivers well constrained contours, and then applies a broadened version of these constraints as a prior when performing lag recovery.

4

LITMUS: Lag Inference Through The Mixed Use of Samplers - An Alternative Program for Lag Recovery

Seasonal gaps are a characteristic feature of multi-year AGN surveys, and `JAVELIN` consistently fails to accurately recover the resulting multi-modal lag-likelihood distributions. This obfuscation of signal lags, as well as `emcee`'s wastefully high sample rejection rate in aliased signals, motivates us to employ different sampling algorithms that are better suited to the challenges of AGN lag recovery.

In this section, I present a new program for reverberation mapping, `LITMUS` (Lag Inference Through the Mixed Use of Samplers). I have designed this alternative method for AGN lag inference to be specifically suited to the challenges posed by seasonal observation gaps. This procedure combines two techniques, nested sampling and Hamiltonian Monte Carlo, leveraging the advantageous properties of each for lag recovery that is robust against the pitfalls that `JAVELIN` encounters, while remaining efficient in high dimensions.

4.1 Alternative Samplers

This section provides a conceptual introduction to the two sampler types used in `LITMUS`'s alternative lag recovery method. This is not a rigorous overview, and is instead intended only to equip the reader with a qualitative understanding of the methods and their behaviour.

4.1.1 Nested Sampling, `JAXNS`

`emcee`'s poor performance in multimodal posteriors motivates us to apply other sampling methods that are better suited to the problem at hand. One such option is Nested Sampling, a gradient free sampling method robust against multi-modality[38]. Nested sampling is not strictly formulated as a contour estimator, and is instead intended to be an integrator over all parameter space to find the 'evidence', the total probabilistic mass of the model:

$$Z = \int \mathcal{L}(\theta) d\theta \tag{4.1}$$

The core concept of nested sampling is to draw samples of strictly increasing likelihood to allow for Lebesgue integration of the posterior, with different implementations using a variety of methods to efficiently draw these samples. The general procedure common to all methods is, per Albert [1]:

---

1. Draw a random set of N_{Live} "Live points" from the prior.
2. By some method, advance each of these live points to a position in parameter space of greater likelihood such that the distribution of live points 'shrinks' in towards the modal peaks.
3. Use the lowest likelihood of all current samples as an estimate of the contour likelihood to check for convergence.
4. If not converged, go to 2.

The distribution of the points at each iteration gives an estimate of the size and shape of likelihood contours, which can be integrated over successive iterations to give the total model evidence. Even when we are not seeking to recover this integral, the shrinkage procedures for drawing samples are still an effective way of estimating the location, shape and prevalence of the different 'islands' of probability in a multimodal posterior. Though not an MCMC algorithm in the true sense, the evaluations made by nested sampling can be used in MCMC applications through 'weighting sampling', in which the likelihoods of these evaluation points are used to sub-sample them in a way that generates a sample representative of the underlying posterior distribution [1].

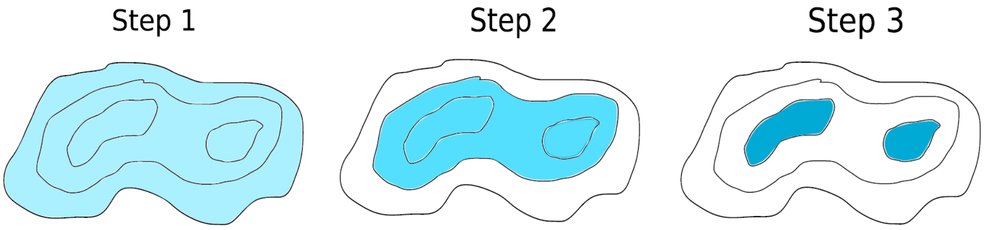

Figure 4.1: Shrinkage of contours in Nested Sampling over successive iterations (from figure 1 of Brewer et al. [8]).

In this project, I employ nested sampling via the `JAX` based implementation `JAXNS` from Albert [1], which readily integrates into `NumPyro`. There are two tuning parameters in this implementation: the total number of live points, and the maximum evaluations / number of samples per live point. The number of live points determines how reliably the modes can be identified: using too few can lead to modes being 'missed' or poorly estimated [2], while sufficient total evaluations must be made to ensure proper convergence. Albert [1] advises using a 'few' evaluations per live point, and a number of live points per Equation 4.2:

$$N_{\text{Live}} = 50 N_{\text{Modes}} (N_{\text{Dim}} + 1) \tag{4.2}$$

where N_{Modes} is the number of distinct modes in the posterior distribution and N_{Dim} is the number of dimensions / parameters being fit for. Too few live points can result in modes being entirely missed by the sampler, and so it is crucial not to under-tune this parameter.

4.1.2 Hamiltonian Monte Carlo / The No U-Turn Sampler

Gradient-free samplers like the AIES and Nested Sampler inherently require 'guess and check' type sampling, which can lead to low acceptance rates. An alternative exists in Hamiltonian Monte Carlo (HMC), which navigates through parameter space by leveraging gradients of the likelihood function. By analogy, the difference is similar to that between the secant method and Newton's method for root-finding: by having analytical gradients at our disposal we can achieve better performance and less wasted iterations. Such gradient based algorithms are normally human-time expensive to implement as they require readily evaluable expressions for these gradients, but this issue is sidestepped through power of `JAX`'s auto-differentiation[6].

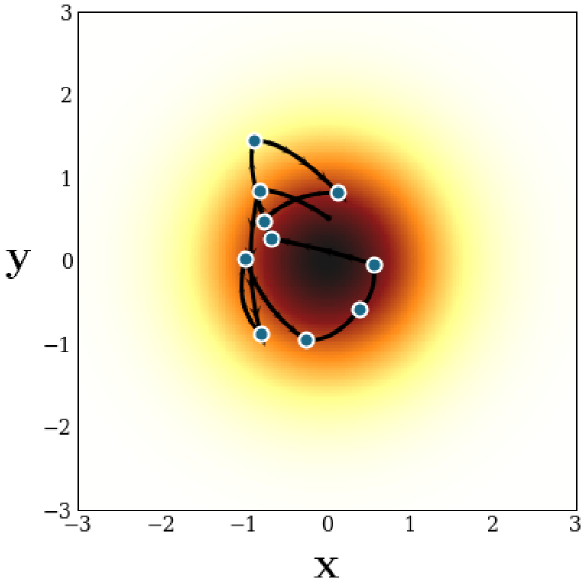

FIGURE 4.2: An example of an HMC chain navigating about a Gaussian distribution [11] with arbitrary parameters 'x' and 'y'. The walker moves about the distribution as though sliding around a 'bowl' that deepens at regions of high likelihood, with regular kicks randomizing the its direction and speed.

HMC operates by way of a kinematic analogy, constructing a landscape of potential energy that becomes deepest at regions of parameter space with high likelihood. Each walker is conceived of as a mass 'rolling' about this landscape with a set mechanical energy. The walker's path is calculated numerically using the typical approaches for time-series integration, such as the energy-conserving symplectic leapfrog method, much as we would for a physical system with conserved mechanical energy [4].

At intervals along its path, the particle is ‘kicked’ to randomize its energy and direction of motion. An ensemble of many such walkers can easily navigate in and out of shallow optima, but still have a draw towards sampling the most prevalent modes of the distribution.

HMC has three main advantages over gradient-free samplers like `emcee`. Firstly, the continuous kinematic paths of the walkers means that many meaningful samples can be drawn with an extremely high acceptance rate. Secondly, the sampling is free of any assumptions about the shape of the posterior: a single HMC walker can explore contours of any arbitrary shape without distortion. And thirdly, the sampling method scales well into arbitrarily high dimensions. These advantages come at the cost of poor behaviour in highly separated multimodal distributions. As the sampler follows a near-continuous path in parameter space, HMC is extremely susceptible to becoming trapped at local optima.

4.2 Procedure Overview

This section provides an introduction to the statistical and numerical procedures of `LITMUS` and contrasts with existing software. Like `JAVELIN`, `LITMUS` models the AGN light curve as a damped random walk with fixed timescale and amplitude. Unlike `JAVELIN`, `LITMUS` does not include the top-hat transfer function to smooth the response curves, instead treating all signals as being ‘scaled and shifted’ copies of the same signal so as to reduce the number of model parameters. The transfer function is known to have little impact on lag recovery for the low-cadence OzDES data [42], and is typically fixed at a nominal value when using `JAVELIN` [21, 32].

For compatibility with `TinyGP`, we do not marginalize over means as `JAVELIN` does (Equation 3.7), and instead include the mean of each signal as a model parameters (Equation 3.6). Our model then contains $3n+3$ parameters when fitting n lags: the amplitude and mean of each signal, the timescale of stochastic variability and the lag of each signal.

Prior to fitting the data, `LITMUS` normalizes the data using estimates of their mean and amplitude for uncorrelated data. i.e. for a set of measurements belonging to signal ‘a’, with measurements $\{y_i^a\}$ and measurement uncertainty $\{E_i^a\}$, the normalizing transformation is:

$$y_i^a \to \frac{y_i^a - \bar{y^a}}{\sigma^a}, \quad E_i^a \to \frac{E_i^a}{\sigma^a} \tag{4.3}$$

where the mean and amplitude estimates are defined:

$$\bar{y^a} = \frac{1}{\sum_i w_i} \sum_i w_i y_i, \quad \bar{\sigma^a} = \frac{1}{\sum_i w_i} \sum_i w_i (y_i - \bar{y_i})^2, \quad w_i = (E_i^a)^{-2} \tag{4.4}$$

This has no effect on the recovered lags, and is done only allows us to more easily set prior ranges on the non-time like parameters by allowing us to assume that all signals have means near zero and amplitudes of order unity. Following the lead of `JAVELIN`, the range of allowable lags are rest riced to positive values up to ≈ 0.3× the observation baseline, though this is an adjustable parameter in `LITMUS`. For consistency across all sources in the OzDES sample, all results in this thesis use a fixed range of $\Delta t \in [0, 800]\,\mathrm{d}$. Similar also to `JAVELIN`, the amplitude of the continuum signal and timescale of variation are modeled logarithmically using a log-uniform prior.

The model parameters and their chosen prior ranges are, for the continuum signal 'c':

- The damping timescale, fit in log-space, $ln|\tau_d| \in [2, 14]$.
- The amplitude of the continuum signal, $ln|\sigma_c| \in [-2, 4]$.
- The mean of the continuum signal, $\bar{y}_c \in [-20, 20]$.

And, for each fit response signal '*a*':

- The observer-frame signal lag, $\Delta t_a \in [0, 800]$.
- The amplitude relative to the continuum, $A_a = \sigma_a/\sigma_c \in [0, 10]$.
- The signal mean, $\bar{y}_a \in [-20, 20]$.

To properly explore parameter space in light of the concerns outlined in Section 3.2, `LITMUS` makes use of a 'mixed' sampling method. HMC, though a better exploratory sampler within a single mode, fails to navigate through the 'rough' lag-space likelihood distribution (e.g. Figure 3.3), and so `LITMUS` uses the more robust approach of nested sampling to first locate these 'islands' of likelihood. This is in place of the 'burn-in' phase employed by `JAVELIN`, but avoids the over-inflation of the aliasing modes, as nested sampling is robust against local optima issues. `LITMUS` then uses HMC for its 'main' sampling run, initiating a series of many walkers within each mode as recovered by nested sampling.

The starting position of these chains are drawn from the nested sampling evaluations using weighted sampling (a utility already provided by `JAXNS`). In this way, the starting locations of the HMC chains are already distributed across the modes proportional to each modes total 'likelihood mass'. Provided we draw enough start locations / run enough chains, we can be confident that our recovered MCMC outputs will reflect the prevalence of each mode, even in cases where their isolation from one another means the HMC chains are unable to migrate.

HMC chains propagate through parameter space through time-series integration, which requires a time-step size to be nominated. Rather than estimate the step-size a priori, we make use of a feature in `NUTS`, `NumPyro`'s implementation of HMC, to tune the step size during its burn-in phase. By running 200 pre-sampling steps, `NUTS` is able to locate a good fit for this step size.

4.3 Sampler Tuning

Though its main sampling run is performed by HMC, the inability of these chains to navigate between modes means that `LITMUS` is extremely sensitive to the accuracy of the initial nested sampling. Under-sampling at this stage can result in modes being under-represented, and

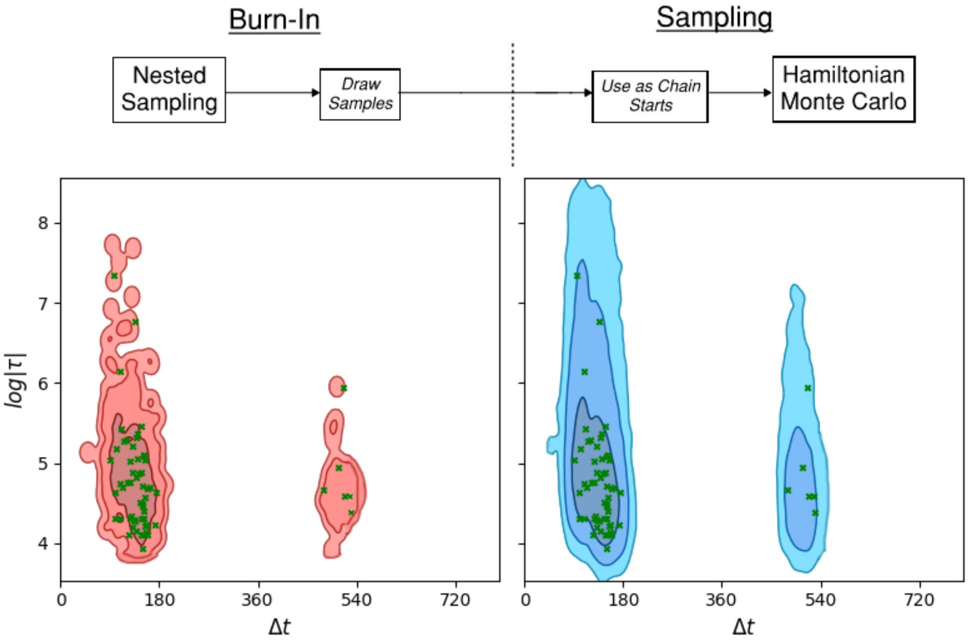

FIGURE 4.3: A representation of the way in which `LITMUS` combines nested sampling and HMC as demonstrated for the parameters of correlation timescale and lag. Nested sampling (red, left) is used to determine the location and prevalence of each mode. Samples (green crosses) are then drawn from this result and used as initial positions for HMC chains, which then sample the modes for a clear final result (blue, right).

may cause significant modes to be missed entirely. For this reason, it is important to not 'under-tune' either the number of live points or evaluations at this stage.

Albert [1] advises that number of live points scale with the number of parameters and expected modes samples per Equation 4.2. Under a worse-case scenario, the number of significant modes will be that of all possible aliasing peaks, occurring in a grid every 180 d, plus one 'true' peak:

$$N_{\text{modes}} = 1 + \left\lfloor \frac{\text{lag range}}{360} + \frac{1}{2} \right\rfloor^{N_{\text{signals}}-1} \tag{4.5}$$

Given how susceptible lag recovery is to under-tuning the live points, we double the number of live points nominated by Albert [1], and choose a similarly conservative 100 evaluations per live point to remove any question of convergence. Using also that the number of parameters is $3N_{\text{signals}}$, the number of evaluations used by `LITMUS` for the nested sampling is as given in Table 4.1.

Fit Type	Est. N_{Modes}	N_{Dim}	Num Live	Max Evals
Cont. Only	1	3	400	40,000
One Line	3	9	2,100	210,000
Two lines	5	15	500	500,000

TABLE 4.1: Standard tuning parameters for `LITMUS`'s nested sampling phase

When moving the main sampling phase, we have to consider the worst case scenario of zero mode migration for each chain. In this case, it is important that we draw a sufficient number of samples from the nested sampling results / initiate enough HMC chains that we representatively sample the modes. We can make some estimates of of the required number of samples by considering that the probability of any sample being drawn from a particular mode obeys a binary distribution. For a properly converged nested sampling run, the probability that a sample will be drawn from that modes is 'p', where p is the total fraction of the posterior likelihood associated with that mode. If we draw 'N_{samples}', the mean and variance of the total number of samples drawn from each mode is then:

$$\text{Mean}(N_i) = pN_{\text{samples}}, \quad \text{Var}(N_i) = p(1-p)N_{\text{samples}}. \tag{4.6}$$

For a representative sample, we require the variability to be small compared to the mean. Defining this relative error as $\Delta = \sqrt{\text{Var}(N_i)}/\text{Mean}(N_i)$, the required number of samples for a particular relative error is:

$$N_{\text{samples}} = \frac{1}{\Delta^2}\left(p^{-1} - 1\right). \tag{4.7}$$

.

For 'M' modes, the median probability is $\bar{p} = \frac{1}{M}$. For an estimate of an appropriate number of samples, we say that the accuracy of a small mode, say $r = \frac{p}{\bar{p}} = 20\%$, must be sampled accurately to within $\Delta = 20\%$. For 5 modes in a 2-line fit, per Equation 4.5, this gives a required number of samples:

$$N_{\text{samples}} = \frac{1}{\Delta^2}\left(\frac{M}{r} - 1\right) = 600. \tag{4.8}$$

All results presented in this thesis use these tuning parameters, with an HMC chain length of 600 samples per chain post burn-in.

4.4 Software Packages

A central computational problem of AGN RM is the numerical challenges posed by the high cost of evaluating the likelihood function for a GP. Though efficient sampling algorithms cab minimize the number of evaluations required, the computational efficiency of each individual evaluation is still a matter of practical concern. Though written predominantly in `Python`, `JAVELIN` achieves good performance by performing its matrix inversion using the extremely efficiency `FORTRAN`-based linear algebra package (`LAPACK`).

To achieve a similar end, `LITMUS` has been programmed in the "just-in-time" (JIT) compiled framework `JAX`. `JAX` is a python integrated language that allows code to be written in the easily maintainable syntax of `Python`, while achieving a order of magnitude speed-up by compiling certain sections of the code at run-time. In addition to this computational speed-up, `JAX` also has automatic differentiation (autodiff) features, which allows it to compute the exact gradients of any well-behaved function. Autodiff allows the use of gradient-based algorithms like HMC without the onerous 'by-hand' calculation of such derivatives.

To perform the Bayesian inference, `LITMUS` use the `JAX` based probabilistic programming language `NumPyro`, which provides a framework for both the statistical modelling and sampling. `NumPyro` provides an implementation of HMC in the form of `NUTS`, but does not have an in-built nested sampling algorithm. Instead, we make use of `JAXNS`, a `JAX`-based implementation of nested sampling which combines the slice sampling of PolyChord [13] and the Gaussian bubble estimation of MultiNest [20]. Gaussian Process modelling is performed using the `TinyGP` package, a 'GP accelerator' with a well maintained integration with `NumPyro`. `TinyGP`, also programmed in `JAX`, provides computationally efficient and end-user friendly utilities for constructing, interpolating and evaluating the likelihood of Gaussian processes.

Collectively, these tools provide a framework that allows for computationally efficient Bayesian modelling of the AGN signals. This framework can be easily extended to include variations on the physical modelling of the system or to implement new sampling algorithms, and so also gives a base from which to extend RM methods.

4.5 Validation of Results & Comparison with `JAVELIN`

To confirm that this new method properly recovers lags and compare to existing techniques, I test `LITMUS` against a sample of 20 simulated 2-line AGN and contrast with the lags recovered by applying `JAVELIN` to these same sources. Signals are generated by time series integration of Equation 3.11 at a uniform time-scale of $\tau_d = 400$ and using similar lags for both response signals at $\Delta t_1 \approx \Delta t_2$. The 20 AGN are simulated in groups of 4 at lags of 150 d through 550 d, with each group separated by 100 d increments. The measurement cadence, uncertainty and signal to noise ratio directly emulate sources in the OzDES sample, with simulated AGN for each lag ranging from high to low measurement quality.

Taking a broad view, we find that `LITMUS` and `JAVELIN` both produce similar contours for the 'true' mode, but that `JAVELIN` consistently over-reports spurious aliasing peaks. Some key examples for two-line fitting are shown in Figure 4.5, demonstrating `LITMUS`'s more reliable recovery of the underlying lag-likelihoods.

These distributions are summarized using median statistics, with the recovered lag being reported as the distribution median and the uncertainty as half the separation of the 16.8 and 78.2 percentiles. [1] Using these summaries, we find that `LITMUS` produces lags that are of a lower spread to `JAVELIN` about the true lags, and that have significantly better precision (e.g. Figures 4.4 and 4.6).

---

[1] Penton et al. [32] finds that the peak likelihood is the least biased measure of the recovered lag, but also finds that the median is of similar accuracy after applying quality cuts. We to not use the lag of peak likelihood to avoid potential inconsistencies arising from binning or smoothing of the posterior.

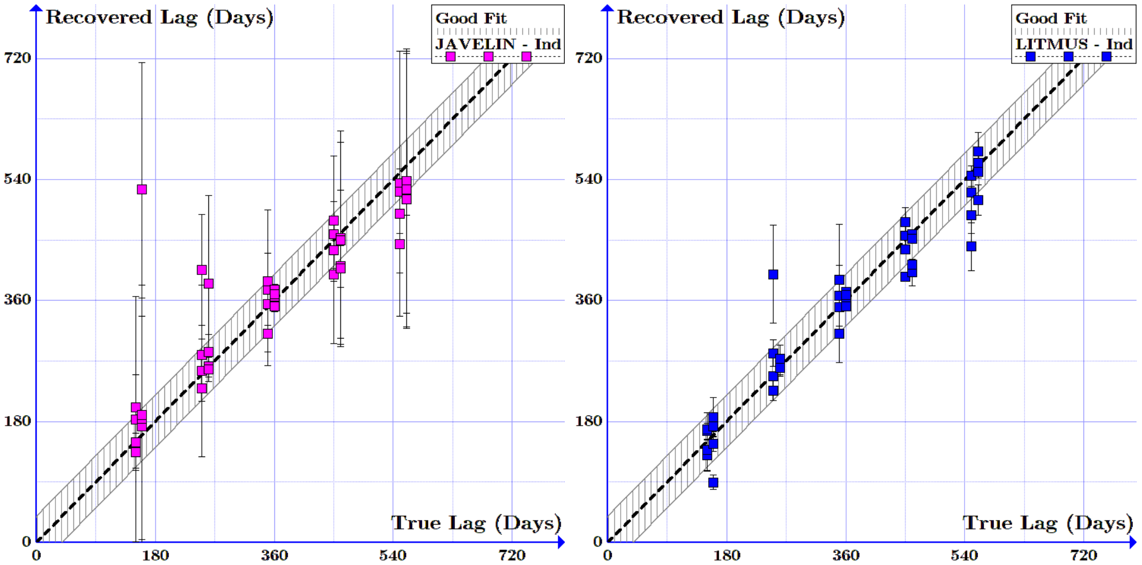

Figure 4.4: Comparison of `LITMUS` and `JAVELIN` lag recoveries when fitting each source's two lags independently of one another. We can see general agreement with truth for both, but with lower spread and measurement uncertainty in the `LITMUS` results. Shown also is a band 40 d either side of the truth value to indicate our criteria for a successful lag recovery.

A central challenge posed by aliasing is the identification and removal of false positives, more so than the precision of the accepted lags. To examine this matter, we test also the false positive rate of `LITMUS` and `JAVELIN` before and after suspect sources are removed, doing so for both independent and simultaneous lag recovery. I employ the quality cuts of Penton et al. [32], excluding the ICCF agreement cut for simplicity. I quantify the performance of these different methods through two measures: the average deviation from the true simulated lag and the fraction of 'correct' lag recoveries, where we consider a lag recovery to be correct if it is within 40 d of the underlying true value [2]. These measures are presented for all sources, the sources that pass quality cuts and the sources that return correct lags only, with the last of these representing the limit of a 100% accurate quality cut method.

---

[2] In their work, Penton et al. [32] define a correct lag recovery one that is within both 3 standard deviations and 80 d of the truth. Due to the dissimilar uncertainties returned by `JAVELIN` and `LITMUS` (e.g. Figure 4.4), I forego the uncertainty based measure in favour of a stricter fixed limit of 40 d.

	`JAVELIN`		`LITMUS`	
	Independent Fitting	Simultaneous Fitting	Independent Fitting	Simultaneous Fitting
False Positive Rate				
Pre-Cut	23%	38%	13%	18%
Post-Cut	20%	50%	19%	25%
RMS Deviation from Truth, $\sqrt{\langle(\Delta t - \Delta t_{\mathrm{true}})^2\rangle}$, in days				
Pre-Cut	75.0	136.8	35.2	31.9
Post-Cut	28.4	82.3	30.8	34.8
Correct Recoveries Only	24.6	22.1	16.8	20.1

Table 4.2: Comparison of `JAVELIN` and `LITMUS` performance for simulated data

`LITMUS` markedly outperforms `JAVELIN` when applied directly to the entire simulated sample, both in the number of correctly recovered lags and the average spread from these truth values (Table 4.2). However, while applying quality cuts has a significant improvement on the results of `JAVELIN`, it degrades the quality of the recoveries from `LITMUS`. Such quality cuts are tuned specifically to the characteristic behaviour of `JAVELIN` in independent line fitting, and appear to be poorly suited to the result of this new fitting method.

To examine the performance of quality cuts, we examine their ability to remove false positives and retain true positives in Table 4.3. The unsuitability of the existing source rejection methods are clear: used in combination with `LITMUS`, they consistently under-reject false positives while over-rejecting true positives as compared to their use with `JAVELIN`.

	`JAVELIN`		`LITMUS`	
	Independent Fitting	Simultaneous Fitting	Independent Fitting	Simultaneous Fitting
Fraction of Correct Lags Accepted	45%	16%	37%	27%
Fraction of Incorrect Lags Rejected	78%	33%	40%	0%
Overall Accuracy of Cuts	60%	65%	60%	60%

Table 4.3: Performance of Quality Cuts On Accepted Data

We also note that, for all fitting methods and quality cuts, the use of simultaneous fitting does not give any consistent significant improvement in the lag recovery by any of our measures, and causes quality cuts to perform markedly worse. It is clear then that our new fitting method produces a significant improvement in lag recovery in and of itself, but that new methods of detecting and removing false positive tuned specifically for `LITMUS` must be found in place of the existing methods tuned for `JAVELIN`.

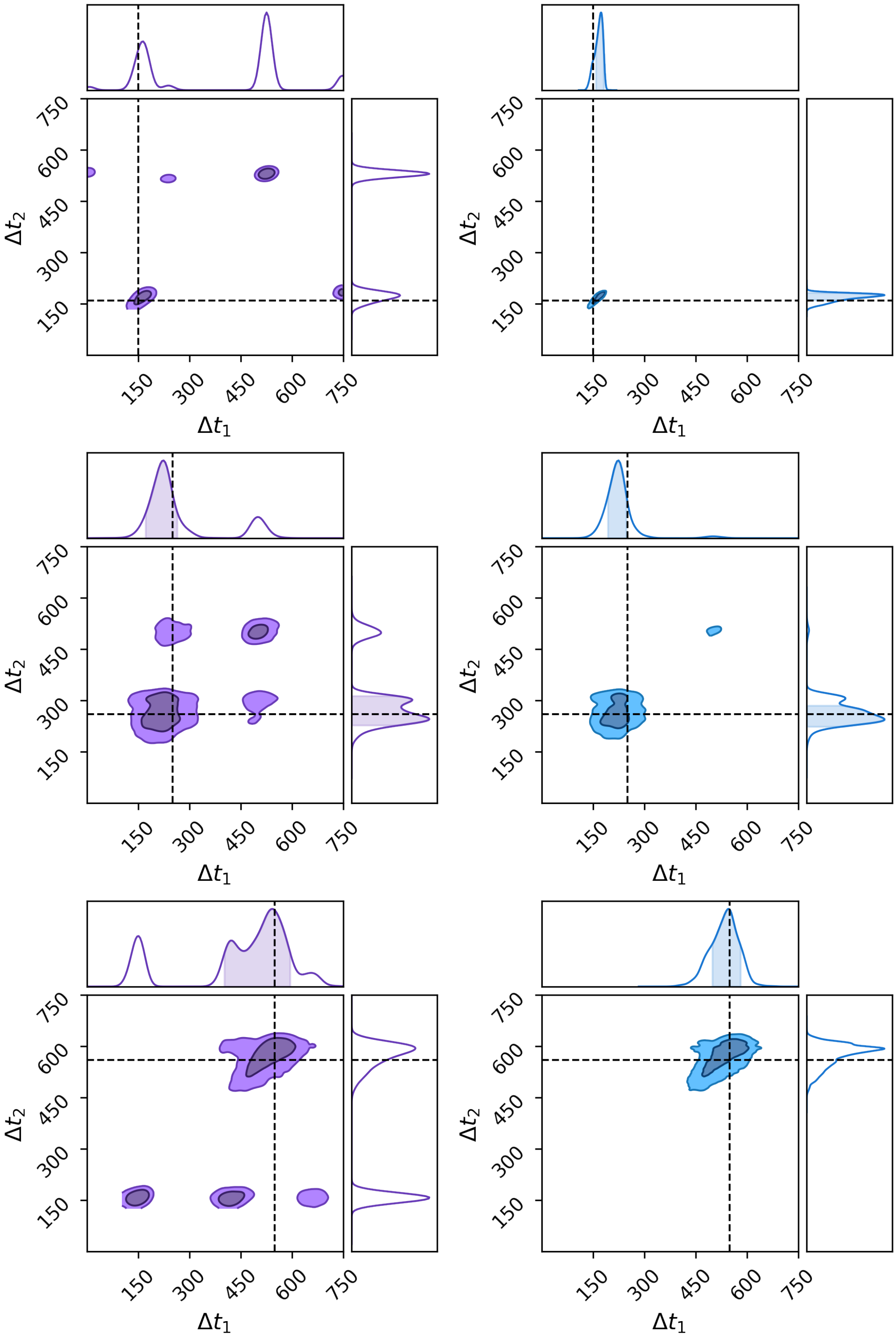

FIGURE 4.5: Comparison of marginalized 2-line lag recoveries for `JAVELIN` (purple, left) and `LITMUS` (blue, Right) for simulated AGN that reproduce the measurement characteristics of sources in the OzDES sample, with the true underlying lags marked with dashed lines. `LITMUS` consistently produces less prevalent aliasing peaks, allowing for a clear lag recovery in some cases where `JAVELIN` fails to produce a primary peak (e.g. top and bottom rows).

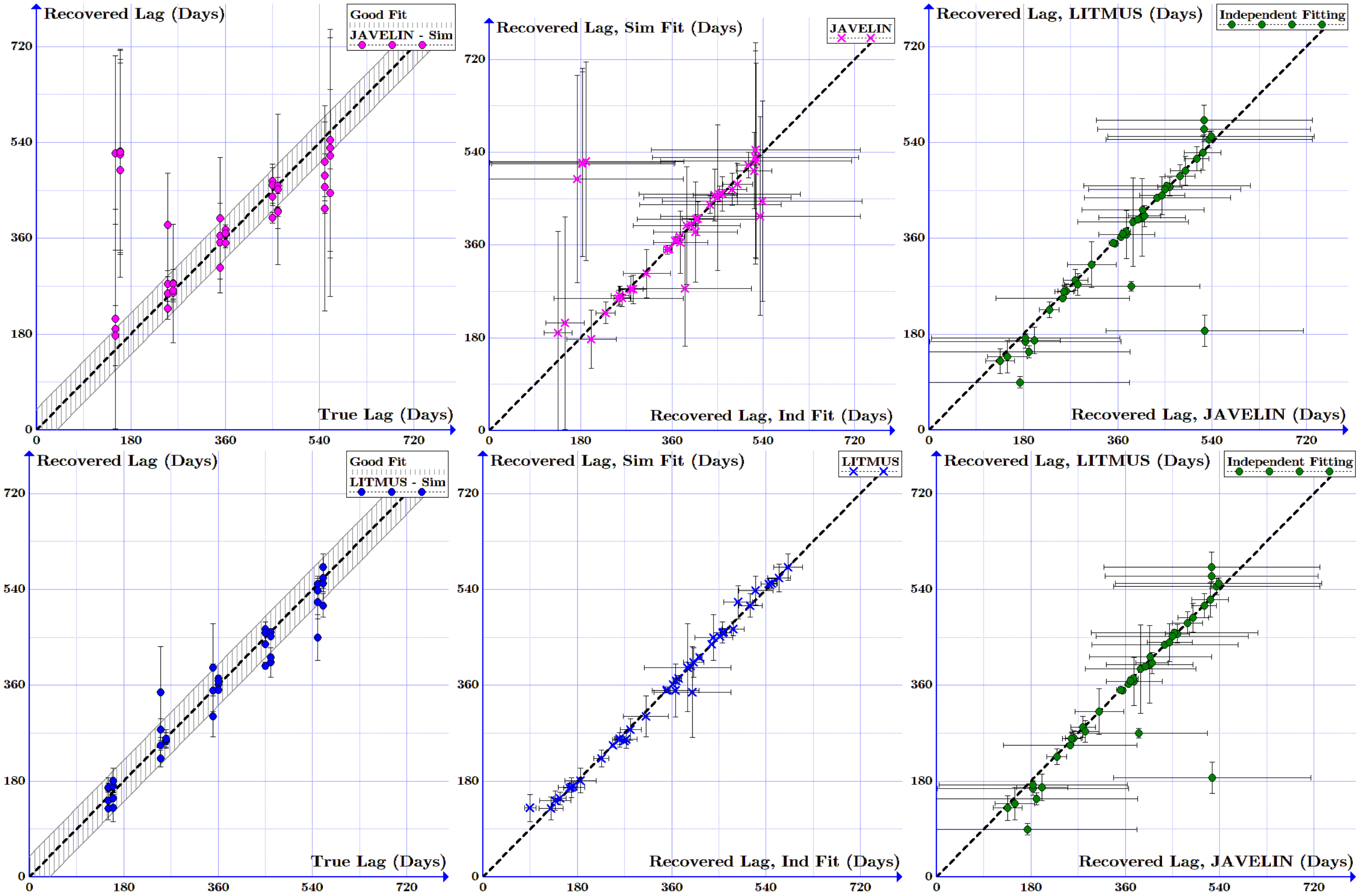

Figure 4.6: Comparison of `LITMUS` and `JAVELIN` behaviour in simulated data. Figures in the left column show the results of `JAVELIN` (top) and `LITMUS` (bottom) for simultaneous fitting of lags for each source, while the middle column compares the agreement between the independent and simultaneous fitting for each program. Figures in the right column show the agreement between `JAVELIN` and `LITMUS` for independent (top) and simultaneous (bottom) lag recovery. All plots are for the full 20 simulated sources with no data cuts.

4.5.1 Performance for Signals with No Underlying Lag

In Section 4.5, we considered only the behaviour of lag recovery programs as applied to signals that meaningfully encode a lag. The opposing case, in which there is no reverberation to observe, has little discussion in the literature. Such cases give a direct test of our ability to detect and remove false positives, as the 'true positive' rate is zero. I test this case here by creating 12 synthetic AGN through the random reassignment of the continuum and response signals of sources from the OzDES 2-line sample. In doing so, I synthesise a set of sources that directly emulate the statistical properties of the OzDES sample, but from which any lag recovery is guaranteed to be a false positive. I then test `LITMUS` and `JAVELIN` on these sources, both for independent and simultaneous lag recovery. Consistent with Section 4.5, we find that the existing methods of false positive removal fail as applied to the lag recoveries of our new method.

Success Rate of False Positive Removal	`JAVELIN`	`LITMUS`
Independent Fitting	52%	10%
Simultaneous Fitting	50%	10%

TABLE 4.5: Performance of false positive removal for `LITMUS` and `JAVELIN`

The underlying mechanism of this failure is that `JAVELIN` and `LITMUS` return dissimilar lag likelihood distributions in the case of a false positive. `JAVELIN`'s failure modes for aliased signals leads it to inflate the minor aliasing peaks, often causing multiple peaks of comparable prevalence to appear in case of poor lag recovery. `LITMUS`, by contrast, more accurately recovers the true prevalence of the modes, leading more often to a deceptively well-constrained primary peak. The quality cuts of Penton et al. [32] rely on the clear presence of multiple peaks to identify poor recoveries, and these criteria fail without these `JAVELIN`-specific hallmarks of aliasing (e.g. Figure 4.7). As a result, these quality cuts fail to identify false positives when applied to `LITMUS`, and fail to improve on the overall quality of our recoveries.

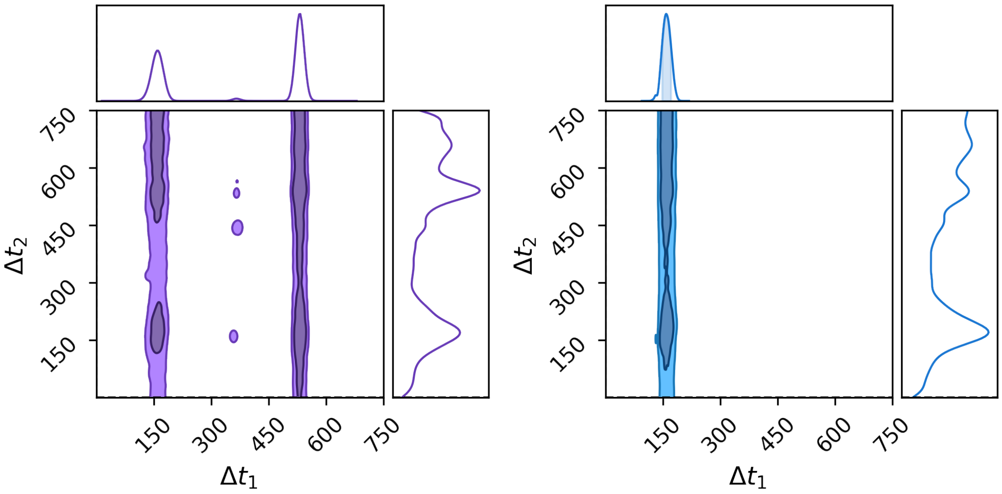

Figure 4.7: Lag Recoveries for `JAVELIN` (top) and `LITMUS` (bottom) for a signal with no true lag encoded. The `JAVELIN` output shows a clear inability to resolve a lag, while the `LITMUS` output may be accepted with unwarranted confidence.

5

Results & Discussion

In Section 4.5, we find that our new fitting procedure produces more accurate lag recoveries for simulated AGN than existing alternatives. We are then motivated to apply to the real data of the OzDES sample to see if any new information is revealed. In this section, I apply our my lag recovery program '`LITMUS`' to the 92 sources in the OzDES sample for which multiple response lines are observed, and compare to results for these same source from Yu et al. [45]. I interpret these recovered lags by examining two topics of interest: the ability of the 2-line MgII samples to constrain an R-L relationship, and the impact that simultaneous lag fitting has on lag recovery for these sources.

5.1 Application of `LITMUS` to the OzDES 2-Line Sample

In this section, I test `LITMUS` on the 92 sources in the OzDES sample for which multiple lines are available. I acquire lag likelihood distributions for both independent and simultaneous lag recovery, and interpret these distributions using median statistics as outlined in Section 4.5. Many sources fail to return a clear well constrained peak (e.g. Figure 5.1), and so we apply I first and third quality cut criteria of Penton et al. [32] for significant lag recoveries in an effort to remove such low quality results. Though roughly half of the MgII lags are retained after these cuts, too few recoveries remain for Hβ or CIV for a meaningful statistical analysis (Table 5.2), and so I examine the recovered lags for MgII only, comparing these to the results of Yu et al. [45]'s study of MgII lags in the entire OzDES sample. The full set of recovered lags are available in Appendix A, while their success or failure against criteria for significant detection are available in Appendix B.

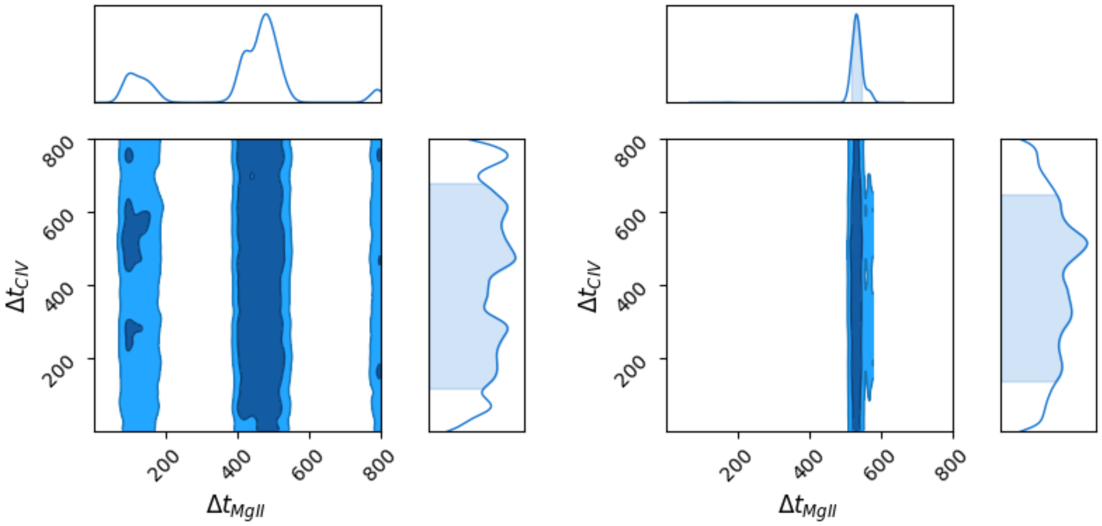

FIGURE 5.1: Example of lag recoveries accepted and rejected by quality cuts. The left figure shows the lag recovery contours for OzDES Source 2925693212, which fails to produce a primary peak for either lag and so is rejected by quality cuts. The right figure right shows lag recoveries for source 2938756405, which has a single well constrained peak for MgII, and so passes these cuts. In both cases, the CIV lag is poorly constrained, and so does not pass the cuts.

		Independent Fitting		Simultaneous Fitting	
Line Type	Total Sources	Passes Cut 1	Passes Cut 2	Passes Cut 1	Passes Cut 2
Hβ	6	6	1	6	3
MgII	92	86	46	84	45
CIV	86	28	0	24	0

TABLE 5.2: Summary of the number of lag recoveries that pass quality cuts in the OzDES 2-line sample

Even after removing sources that fail to meet our significance criteria, we can still see that the remaining sources fail to demonstrate a any consistent R-L relation (Figure 5.2). Instead, we observe a clear bi-modality, with lag recoveries forming into two distinct groups corresponding to the 180 d and 720 d aliasing peaks.

I compare these results to the MgII R-L relation fit by Yu et al. [43] for the entire OzDES sample. In these past results, which applied similar quality cuts to lags as recovered by `JAVELIN`, only 25 sources sources were accepted across the entire 435 objects in the OzDES MgII sample. By contrast, we see double this number just in our limited sub-sample, with a higher acceptance rate by an order of magnitude. This unreasonably high acceptance rate, along with the concentration of lags at the aliasing modes, suggests that our sample is overwhelmingly contaminated by false-positives. This failure of existing quality cuts to sufficiently reduce the false-positive rate is consistent with the findings for simulated data in Section 4.5.1.

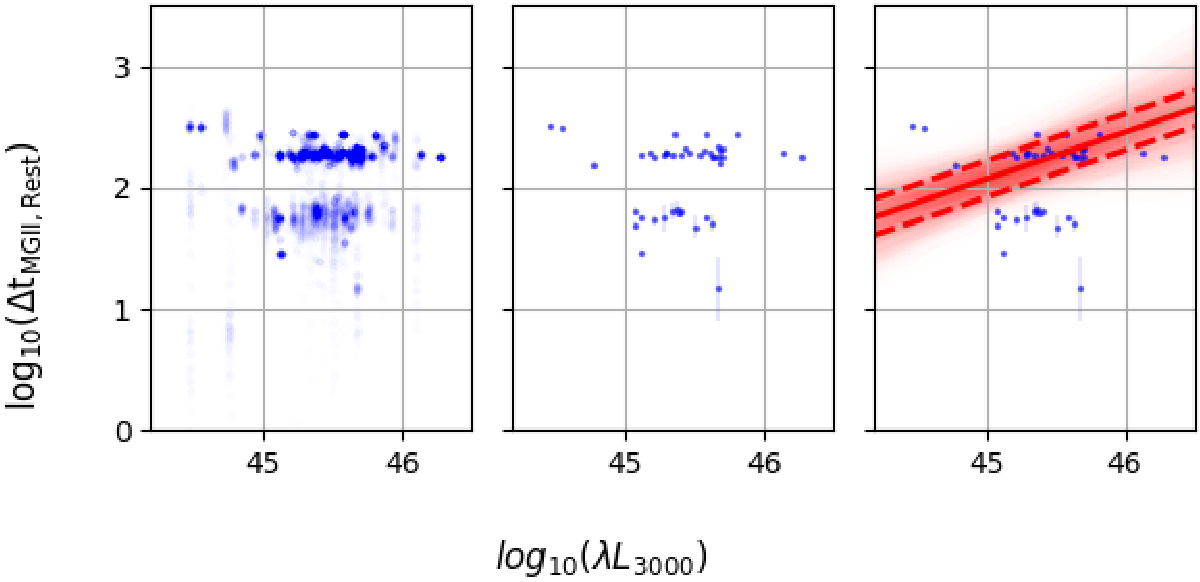

FIGURE 5.2: R-L Domain plot of all 92 recovered lag distributions for independent fitting of MgII from the 92 OzDES 2-line sources (left) and the summarized lag recoveries for the 48 sources that pass the quality cuts of Penton et al. [32] (middle). The right figure The right figure shows the same as the middle, but with an overlay of the MgII R-L relationship as recovered by Yu et al. [45] for the entire OzDES sample.

It is worth noting that the bulk of our lag recoveries are near the 540 d mark. Though a suspect result for reasons of aliasing, this is still somewhat consistent with the prediction of Yu et al. [45]'s R-L relationship. Sources with a true lag near such aliasing values are associated with poor reliability of lag recovery [30], and so the high degree of noise in this sample is not unexpected. By limiting ourselves to only the OzDES 2-line sources, we also unavoidably restrict the luminosity range of our sample, making general trends in R-L space difficult to resolve, and allowing the aliasing noise to dominate.

Yu et al. [45] identify 3 sources with reliable lag recoveries in our 92 source sample (Table 5.3), giving us a small set to check for consistency with our recoveries from `LITMUS`. We see good agreement for source 2, and can attribute the difference for source 1 to our lag prior extending only to 800 d. In their treatment of source 3, Yu et al. [45] note that the MgII light curve has two outlier epochs, which they remove to achieve their recorded lag. They note that not removing these causes the source to return a strong recovery at 540 d, which is consistent with our results from `LITMUS`.

Source Properties				Past Recovery [45] (days)		`LITMUS` Recovery (days)	
	OzDES ID	Redshift	$\log_{10}\lvert\lambda L_{3000}\rvert$	Lag	Error	Lag	Error
1	2925515125	1.86	45.89	867	64	782.1	33.3
2	2938756405	1.79	45.51	538	28	532.8	20.5
3	2971214955	1.75	45.57	479	40	545.8	13.7

TABLE 5.3: Comparison of lag recoveries for MgII sources accepted by Yu et al. [45]

As with previous sections, we see that `LITMUS` has good agreement with existing studies in cases where we can confirm a true positive, but that our ability to identify these reliable sources is completely absent with current techniques. Significance criteria tuned for `JAVELIN` and ICCF consistently under-filter poor recoveries when used in concert with `LITMUS`, and such false positives then contaminate our data to the point of pure noise. It is clear that, if we are to apply `LITMUS` at scale to OzDES like samples, we must first find a more suitable means of quantifying the reliability of recovered lags.

5.2 Impact of Multi-Line Fitting in the OzDES 2-Line Sample

In this section, I examine the impact that the simultaneous fitting of two AGN response signals has on the lag recoveries for the OzDES sample. I examine this impact through two lenses: firstly, the change in the number of recoveries that are well constrained by the measure of Penton et al. [32]'s lag recovery quality cuts, and, secondly, a broader view of how much influence simultaneous fitting has on the lag likelihood distribution in general.

I limit my analysis to only the sources that pass the quality cuts for at simultaneous or independent fitting, as these represent the recoveries with some meaningful constraint on the lags. As outlined in Sections 4.5 and 4.5.1, these cuts are of questionable applicability to the fitting method used in this thesis, but still at very least narrow the sample of interest to the sources that return one clear peak in their lag distribution.

To quantify the impact of multi-line fitting, I classify each source by the degree to which it changes the recovered lag-likelihood distribution. A 'high impact' is any source for which the recovered lag changes by more than 50 d, while 'low' and 'moderate' impacts are classified based on the total shift in the recovered lag distribution. This shift is measured using the correlation between the lag-likelihood distributions with and without multi-line fitting:

$$R = \int_0^{\Delta t_{\max}} \mathcal{L}'(\Delta t_{\text{Ind}}) \cdot \mathcal{L}'(\Delta t_{\text{Sim}}) d\Delta t \tag{5.1}$$

where $\mathcal{L}'(\Delta t_{\text{Ind}})$ and $\mathcal{L}'(\Delta t_{\text{Sim}})$ are the normalized marginal lag-likelihood distributions for independent and simultaneous fitting, and $[0, \Delta t_{\max}]$ is the range of our lag prior. A 'moderate' impact is any in which this correlation is below 95%. The number and degree of impact for every source in the 2-line sample is listed in Table 5.5, and demonstrative examples are shown in Figure 5.3.

Significant Detections	Total Sources	Low Impact	Moderate Impact	High Impact
Gain	9	3	2	3
Loss	8	7	0	1
No Change	34	33	1	0
Total	51	44	3	4

TABLE 5.5: Summary of impact of simultaneous fitting for sources that pass quality cuts

In general, multi-line fitting has very little impact on our results. The high-cadence and low-error signal of the continuum already provides the overwhelming majority of constraining power for all sources, and so the most sources see a low impact. In particular, most MgII recoveries see no meaningful effect from the simultaneous fitting of a CIV line, as the measurement error of most of the CIV signals is so low as to give little to no meaningful constraint (e.g. the CIV lags in Figure 5.1). We find no consistent improvement in the uncertainty of lag recoveries or the number of sources that pass our quality cuts, with a comparable number of sources being gained as 'significant' detections as are being lost (Table 5.5).

In combination with our finding that multi-line fitting has no improvement on lag recovery or the detection of false positives in Section 4.5, fitting lags simultaneously does not appear to be a promising avenue of further research for reverberation mapping OzDES-like samples. However, there is still the possibility that these results may be a function of our high false positive rate, or a characteristic arising from the typical low signal to noise ratio of the CIV lines. Two-line sources are necessarily located at redshifts in which at least one line is at the limit of the observable wavelength range, leading them to more often have a poor SNR, meaning at least one signal will provide little constraining power. Future samples with different seasonal gaps or clearer measurements of the complementary response line signal may see better results from the application of multi-line fitting.

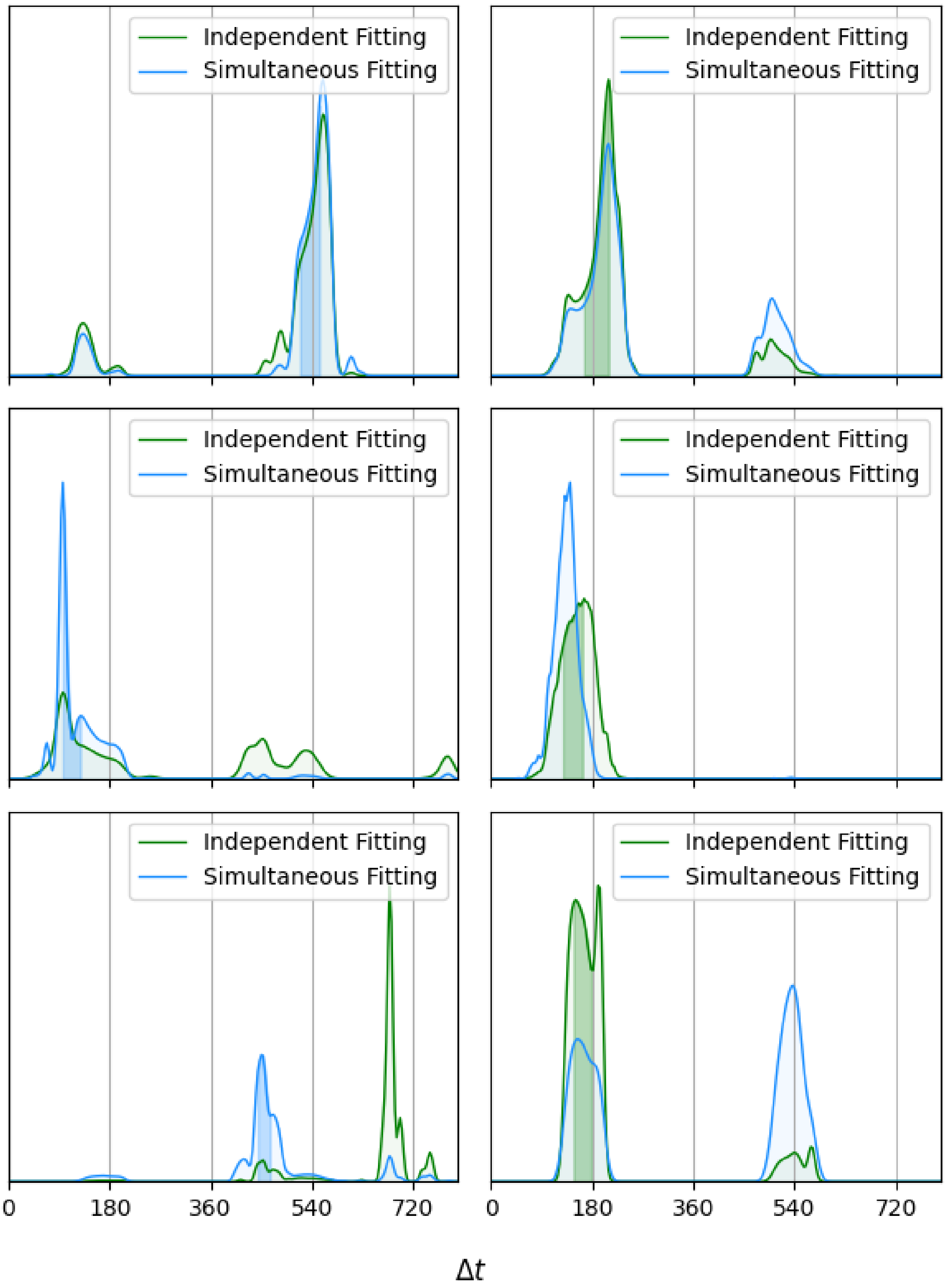

FIGURE 5.3: Examples of the impact of simultaneous multi-line fitting on lag recovery in the OzDES sample. From top to bottom, shown are recoveries with a low, moderate and high impact from multi line fitting.

6

Future Work

In this thesis, I have demonstrated a new method for AGN reverberation mapping that outperforms the existing standard, but there are number of opportunities to further extend this work. In this section, I outline the most immediate areas of interest in improving `LITMUS`, in terms either of its applicability to real data and its general efficiency.

6.1 New Methods for Rejection of False Positives

In its application to both simulated and real data, we find that `LITMUS` outperforms `JAVELIN` at a base level in lag recovery, but that the existing criteria for removing false positives from our sample fail almost entirely when used in conjunction with this new fitting method. The first and highest priority for future work is then to find new methods of estimating the significance of lag recoveries, such that we can apply our new advanced tools to real data with full confidence in their correctness.

The quality cuts applied in Chapters 4 and 5 determine the reliability of a recovery only by the constraint of the recovered lag, not by the actual quality of this 'best fit'. Even where lag recoveries are well constrained to a single primary peak, acceptance of the results are still conditional on the assumption that a meaningful lag actually exists. Rather, we should test against the null hypothesis: that the continuum and lines are completely independent Gaussian Processes with no correlation between them. We can measure this rigorously by comparing the total evidence for a model with lag as a parameter to the evidence for one without:

$$r = \frac{Z_{Delay}}{Z_{Independent}}, \quad Z = \int_{Prior Space} \mathcal{L}(\theta) d\{\theta\}$$

Such evidence estimation through numerical integration of the likelihood is the primary use-case of nested sampling, and so we have a readily available tool-set for making this extension. The threshold for significance in this measure will need to be tuned to achieve the best false positive rate. To do so, we can follow the prior work of Penton et al. [32], in which many AGN simulations are used to optimize quality cuts for a particular AGN sample.

Most current procedures of identifying false positives focus on the hallmarks of aliasing that appear in the marginal lag-likelihood distribution, or on enforcing agreement between different lag recovery methods. Such methods make use of only a limited amount of the available data space, and so are liable to discard meaningful lag recoveries. Using evidence ratios, which leverage all of our available information, may also lead to improvements in the 'false negative' rate as compared to existing methods. For sources like MgII, which only have a total number of RM recoveries on the order of dozens, even a marginal improvement in the source acceptance rate will be a significant advancement.

Having established that `LITMUS` eliminates the significant artefacts that `JAVELIN` produces in aliased sources in section Chapter 4, proper interpretation of these improved results through optimized quality cuts present the opportunity to markedly increase the number and quality of results in AGN reverberation mapping. Equipped with such a means to quantify source reliability, we can then extend this new lag recovery method to the entire 735 AGN of the OzDES sample. This thesis focuses only only the 92 two-line sources, a sub-sample that is necessarily poorly behaved due to its low signal to noise, poor conditioning and limited luminosity range, and so we would expect a better acceptance ratio when broadening our scope.

6.2 Model Extensions

In this thesis, we apply only the simplest model of AGN variability as a Gaussian Process, in which continuum and response are all drawn from the same underlying DRW with a constant baseline brightness. There are two immediate extensions to this model we may consider, one for modelling nearby AGN and another for future long-baseline surveys.

Programs like `JAVELIN` include a smoothing effect between the continuum and response signals, motivated by the finite thickness of the broad line region. This smoothing, encoded in the model in the form of transfer functions (e.g. Equation 3.13), has a negligible effect on lag recovery when the observation cadence is coarser than the timescale of the smoothing [42], and so is often not included as a parameter in RM of OzDES-like data (e.g. Grier et al. [19], Homayouni et al. [21], Penton et al. [32]). This smoothing timescale is an important element of high-cadence AGN RM, and may be a necessary feature for modelling high redshift AGN in which the smoothing timescale is time-dilated to longer values. As GP modelling naturally adapts to such transfer functions (Equation 3.10), it is a natural next step to include this in `LITMUS` to broaden generalize its use case.

Another extension we may consider is the inclusion of a non-constant baseline in one or more of the observed signals. At present, we treat the mean all light curves as varying about a fixed mean, but some sources in the OzDES sample clearly demonstrate variability at different timescales (e.g. Figure 6.1). The techniques for describing time-varying baselines with a linear function basis are well established [37], and we can naturally implement them in `NumPyro` at the cost of a few extra model parameters. Alternately, we may consider modelling signals as a mixture of multiple super-imposed Gaussian Processes. Such a model is similarly easy to implement, as mixtures of GP's are described by a simple linear combination of their respective covariance matrices / functions (e.g. the DRW and 'noise' processes in Equation 3.5).

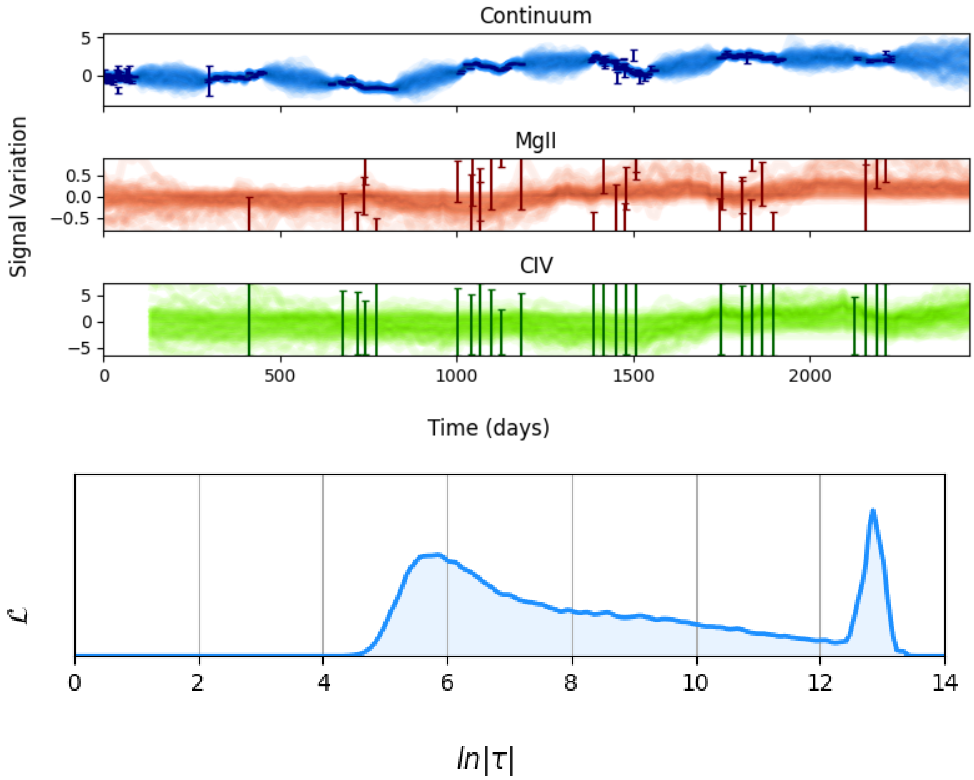

Figure 6.1: Recovered continuum properties for OzDES source 69-B-2940901822, showing multiple behaviours at multiple timescales. The bottom figure shows the marginalized likelihood distribution for the timescale of stochastic variation, while the top figure shows the confidence intervals for the interpolated light-curves marginalized over all parameters. We can clearly see long-term variability in the continuum that is not reflected in the MgII signal, resulting in an inability to meaningfully correlate the signals and a total failure to interpolate the behaviour of the response curves.

6.3 Improved Sampling Algorithms for Better Numerical Performance

In its current state, `LITMUS` has areas in which it can be significantly improved in computational efficiency, and we here identify three immediate avenues of interest. The current tuning of the Nested Sampling parameters (Section 4.3) defers to a conservatively high value for the number of evaluations to ensure good convergence. There exist numerical estimates of convergence are well established for nested sampling [38], and these may be implemented as a halting condition for the burn-in phase. This will provide both a speed-up in cases where convergence is fast, and ensure convergence in poorly behaved cases.

Secondly, the transition from the nested sampling to the main HMC sampling phase is presently somewhat wasteful, as it requires us to initiate on the order of 100's of chains to ensure we draw a proportionate number of samples from each mode (Equation 4.7). Instead, we may consider using some clustering algorithm to identify the modes at the end of the nested sampling phase (e.g. Wegmann et al. [41]), and initiate only a few chains per mode. After mapping each mode individually, we can then combine their samples in a representative way using a weighting via importance sampling [12]. In doing so, we can remove the computational overhead of running 100's of chains, providing a valuable speed-up in `LITMUS`'s lag recovery run-time. `JAX` already offers efficient optimisation tools, and so this too is a natural extension.

Finally, we note that the likelihood function in lag recovery, though high-dimensional, is broadly uncorrelated in its various parameters (Section 3.2). Given nested sampling becomes expensive in high dimensions (Equation 4.2), a faster alternative may existing in using nested sampling over only the complex behaviour of the lag domain, with all other parameters at fixed values of maximum likelihood.

6.4 Hierarchical Modelling of Population

The standard approach of reverberation mapping for a population of AGN is to examine each source independently, assuming no prior or collective information about their lags. We understand that AGN of similar luminosity should exhibit similar rest frame lags, with such similarity being the physical basis for R-L relationships. Along this line of reasoning, it has been demonstrated that ‘stacking’, the combining of lag-likelihood distributions of AGN of similar luminosities as if they were measurements of a single source, is an effective means of extracting predictive power from poorly constrained sources [14]. This approach has also been extended to a more general hierarchical model, in which we allow for an inherent spread in the lags, with a similar effectiveness [7]. We are then motivated to extend an entire population of AGN spanning a range of luminosities in the form of a full hierarchical population model, using recovered lags to constrain an R-L relationship, while simultaneously using this relation as a prior to better constrain the lags of each source.

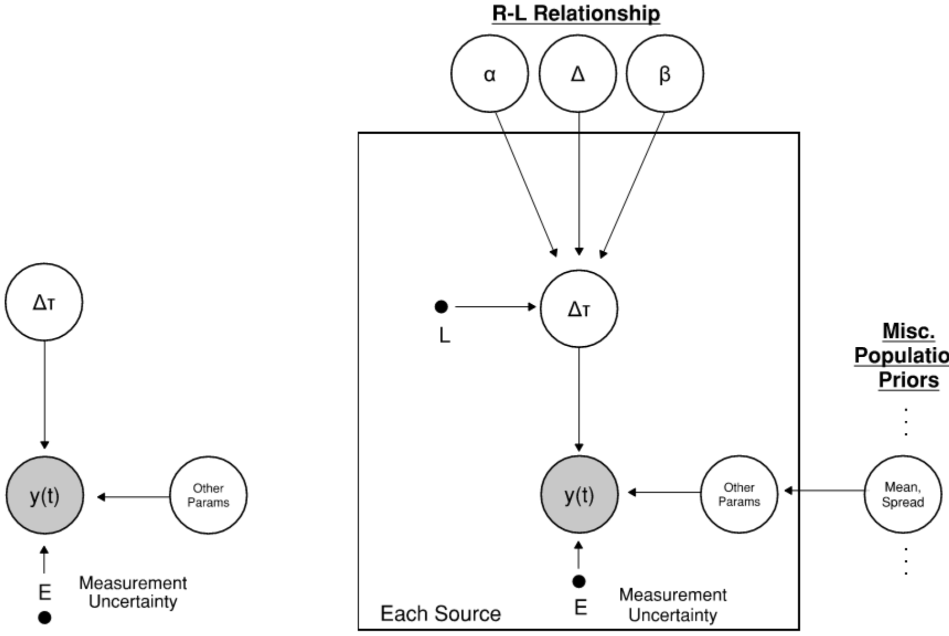

FIGURE 6.2: Sketch of a Probabilistic Graphical Model for single-source model (left) and hierarchical population model (right) of lag recovery for one response line. In the single-source model, lags are modeled for each source individually by a Bayesian model that encompasses only one source at a time. In a hierarchical model, lags for a ‘plate’ of sources are fit simultaneously, with their lags related by population levels priors like the R-L relationship. Properties of these priors are tuned as model parameters, with the total Bayesian model encompassing the entire set of AGN at once.

Along a similar line of reasoning, there is good physical reason to believe different emission lines from a single AGN should produce similar lags, as they are both emitted from the same geometric features. It has been demonstrated that the lags for Hβ and MgII are of similar scale for a single source, while CIV lags are $\sim 3-4$ times smaller [21]. Though we have found in this thesis that simultaneous fitting of multiple lags in a single source does not markedly improve results by itself, such a relationship between the different lag parameters does provide an opportunity to extract some constraining power from low signal to noise source. In many 2-line sources in the OzDES sample, one of the lags fails to be constrained at all (e.g. Figure 5.1). By employing a ‘lag similarity’ prior, constraining $\Delta t_1 \approx \Delta t_2$, we can use a good lag recovery for one line to infer the lag for another. The properties of such a prior, much like the parameters of the R-L relationships, can also be constrained in a hierarchical way.

Such priors will have the greatest effect on poorly constrained sources. Such sources tend to occur at the edge of our observation range, (e.g. the low signal to noise of low redshift CIV sources), and so any predictive power we can derive from such sources will improve not only the number of data points we have, but also their range. We note in Section 5.1 that the 2-line MgII sources, by the R-L relationship recovered by Yu et al. [45], are expected to have lags near 540 d. This range is near a seasonal gap, which is associated with poor lag recovery [30], and is an example of a data range that would benefit from additional constraining power.

7

Conclusion

The recent generation of multi-object AGN reverberation campaigns presents an unprecedented opportunity for measurements of supermassive black hole masses deep into the cosmic past. As these surveys draw to a close, we have found that the seasonal observation windows of such surveys have lead to significant measurement noise, and the existing methods for counter-acting this noise drastically limit the number of objects we can reliably extract information from. In this thesis, I demonstrate that the existing tools for reverberation mapping are poorly suited to these seasonal observations, and that there is a clear need for modern alternatives that are more appropriate for these unique challenges.

In this thesis, I have demonstrated that `JAVELIN`, a program that stands as the existing state of the art in AGN reverberation mapping, has failure modes that are significant and common when applied to AGN with seasonal gaps. I have identified the mechanisms and impact of these failure modes, and have presented a new alternative program that makes use of new sampling techniques that does not share these issues. I demonstrate that this new program, `LITMUS`, reliably outperforms `JAVELIN` in recovering the true underlying likelihood distributions in AGN reverberation mapping, and that this improved performance returns results with higher precision and less deviation from the underlying truth values.

However, `LITMUS`'s practical utility is limited by it's lack of suitable and effective methods for rejecting false-positives, as the the existing tools are either ineffective or maleficial. I find that established methods of false positive rejection fail when applied to `LITMUS` in both simulated and real data, and identify a fundamental incompatibility between the mechanisms of these quality cuts and the lag likelihood distributions recovered by `LITMUS`. Without effective alternatives to these methods, `LITMUS` is of limited applicability to samples with a high false-positive rate like OzDES. For sources with previously published lags, I find that my new method produces results that are in good agreement.

I also test the effectiveness of fitting multiple lags simultaneously for AGN in which we can observe more than one reverberating emission lines in place of the common approach of recovering them independently. I find that, in both real and simulated data, this technique does not offer any consistent improvement in results for the measurement quality of the OzDES sample. This multi-line fitting does not present any promise in application to current data, but may show better results in future surveys in which the response signals have a higher signal to noise.

`LITMUS`'s improved accuracy has the potential to decrease the number of discarded false negatives in real data, and, if applied properly, presents an opportunity to increase the number of meaningful lag recoveries from multi-object campaigns like OzDES. Only a few dozen MgII lags have been recovered at high redshift in total, and so even a marginal increase in the number of these recoveries will present a significant advancement. Without effective tests of the reliability of its measurements, `LITMUS` presents only half a solution to the existing problems of aliasing in AGN RM, and so development of such tests is an important avenue of future research. As the first generation of surveys like OzDES draw to a close, there is strong motivation to extract as much information as possible from their measurements. `LITMUS`, if properly complemented by surrounding tools and procedures as `JAVELIN` is, presents the potential to measurably improve the number and quality of results recovered from such surveys.

A

Recovered Lag Summary for OzDES Two-Line Sources

These tables list the recovered lags for all 92 OzDES sources with multiple lines visible as recovered by `LITMUS`. Lags are measured from the median of the recovered distribution, and the recovery uncertainty is given as the difference between the lags of the 18.2 and 62.8 percentiles. Results are presented for fitting lags independently and simultaneously.

		Independent Fitting		Simultaneous Fitting	
Index	OzDES ID	Lag	Error	Lag	Error
7	2925552152	107.0	93.5	111.9	95.3
19	2925858108	549.8	439.4	551.6	471.0
79	2970604169	149.8	76.5	160.4	88.7
88	2971028700	140.0	117.4	152.5	111.3
90	2971086054	134.0	41.1	155.1	59.3
91	2971134055	138.8	51.2	143.1	357.0

TABLE A.1: Recovered `LITMUS` lags for Hβ for independent & simultaneous 2 line fitting

		Independent Fitting		Simultaneous Fitting	
Index	OzDES ID	Lag	Error	Lag	Error
0	2925372393	210.7	330.4	204.8	72.1
1	2925373860	191.4	427.5	161.2	390.0
2	2925375402	178.6	346.5	511.8	351.1

3	2925420688	42.4	48.1	42.2	19.5
4	2925515125	782.1	33.3	781.4	34.8
5	2925523772	548.8	444.7	509.8	408.7
6	2925551147	605.9	318.1	606.5	319.2
7	2925552152	562.8	183.9	540.9	166.7
8	2925606181	518.1	19.3	518.0	19.9
9	2925609647	153.2	78.6	155.9	77.1
10	2925637387	529.9	37.9	530.0	38.5
11	2925674035	507.0	462.3	510.2	467.5
12	2925685619	170.2	428.2	171.1	410.6
13	2925693212	462.6	364.4	456.2	377.2
14	2925707000	455.0	71.7	678.4	120.3
15	2925718880	518.2	154.4	221.2	389.6
16	2925786831	227.5	458.8	185.9	442.9
17	2925835393	135.5	23.9	135.6	23.8
18	2925857917	195.3	59.9	194.9	59.2
19	2925858108	12.5	86.1	13.0	185.1
20	2937741147	597.1	28.6	597.1	28.4
21	2937810288	568.2	579.4	569.3	575.0
22	2937856748	229.5	610.0	207.7	608.6
23	2937895789	534.2	63.3	535.9	70.1
24	2937961955	524.8	19.1	525.0	19.4
25	2938049569	657.1	240.8	662.6	231.6
26	2938055829	117.1	118.6	120.0	119.6
27	2938228331	210.6	369.3	511.8	365.6
28	2938254204	150.8	27.2	150.6	27.2
29	2938258860	156.3	22.0	156.3	21.6
30	2938268664	179.2	28.6	193.4	372.4
31	2938398472	577.4	423.5	579.9	33.0
32	2938497808	517.8	26.4	517.4	27.1
33	2938498296	530.9	526.5	538.6	531.5
34	2938638637	189.5	88.4	193.3	324.7
35	2938666749	782.5	54.5	782.4	57.6
36	2938756405	532.8	20.5	532.7	20.5
37	2938870373	515.0	401.7	523.1	45.3
38	2938878570	526.3	166.9	531.7	181.8
39	2938969598	507.2	95.9	507.7	96.8
40	2938985411	494.0	104.3	492.0	113.5
41	2939225137	211.5	394.9	195.0	239.9
42	2939317867	80.6	1.5	80.7	75.0
43	2939348426	523.5	348.3	189.1	418.9
44	2939394160	236.2	385.6	552.9	351.7

45	2939489514	528.3	377.6	519.2	368.6
46	2939520972	506.9	391.9	168.2	51.2
47	2939531854	523.6	408.9	542.4	270.1
48	2939622630	633.1	26.9	632.7	26.4
49	2939627477	613.5	402.9	628.7	34.1
51	2939649603	751.4	91.3	751.9	89.3
52	2939652807	157.3	112.0	156.0	108.4
53	2939694336	492.0	389.8	524.9	42.3
54	2939771369	528.1	436.0	288.4	440.6
55	2939776918	556.9	53.0	556.1	56.2
56	2939791816	525.0	63.2	525.6	64.2
57	2940060892	103.4	79.9	186.9	429.6
58	2940088831	202.8	334.0	200.5	24.8
59	2940097386	530.6	60.0	521.8	89.3
60	2940140085	547.9	54.8	542.6	90.4
61	2940147277	162.5	13.2	162.5	13.2
62	2940260013	538.9	323.4	524.7	331.3
63	2940327071	208.0	266.0	195.6	57.2
64	2940351862	180.6	61.1	181.6	61.7
65	2940510474	525.2	29.3	527.1	29.6
66	2940557913	525.8	385.6	509.2	386.5
67	2940670625	176.1	37.4	179.4	34.5
68	2940876978	558.1	68.9	552.6	69.7
69	2940901822	217.9	379.4	216.0	375.6
70	2940908514	186.5	392.7	161.7	382.9
71	2940922093	637.7	606.8	188.9	623.3
72	2940963619	745.1	297.9	750.3	296.8
73	2970356187	479.7	93.3	478.2	95.7
74	2970386160	512.4	37.1	512.7	39.6
75	2970400768	543.1	44.8	546.0	41.5
76	2970414144	523.1	360.7	510.8	377.5
77	2970460831	140.4	45.9	140.7	44.6
78	2970569124	559.4	26.0	560.2	25.9
79	2970604169	262.0	43.4	259.4	44.3
80	2970611652	521.8	19.0	522.0	18.9
81	2970744291	500.7	49.5	178.3	349.6
82	2970786038	174.8	28.3	175.2	28.1
83	2970791376	784.8	2.6	784.7	2.7
84	2970798013	134.0	53.8	133.3	57.7
85	2970807282	582.6	42.8	585.5	38.9
86	2970850376	510.6	23.3	510.5	22.9
87	2970951335	557.9	27.9	557.3	33.3

88	2971028700	97.2	160.4	12.1	126.2
89	2971049565	524.3	378.7	526.6	386.1
90	2971086054	530.9	29.2	525.8	372.3
91	2971134055	550.9	42.0	537.9	44.8
92	2971214955	545.8	13.7	545.9	13.6

TABLE A.2: Recovered `LITMUS` lags for MgII for independent & simultaneous 2 line fitting

		Independent Fitting		Simultaneous Fitting	
Index	OzDES ID	Lag	Error	Lag	Error
0	2925372393	409.7	540.8	418.5	534.2
1	2925373860	361.1	549.1	360.0	564.5
2	2925375402	402.2	530.4	405.8	538.2
3	2925420688	417.4	571.6	423.9	568.8
4	2925515125	506.6	530.7	505.4	533.3
5	2925523772	108.9	432.0	135.3	160.8
6	2925551147	485.2	527.1	496.1	525.3
8	2925606181	432.0	534.8	390.7	563.1
9	2925609647	331.3	528.7	341.4	529.1
10	2925637387	370.4	512.2	340.4	510.2
11	2925674035	395.6	560.2	397.2	554.4
12	2925685619	496.6	552.8	496.5	547.9
13	2925693212	433.0	512.2	432.5	497.6
14	2925707000	387.7	544.6	389.7	541.0
15	2925718880	279.9	528.6	232.0	442.5
16	2925786831	370.8	543.0	365.8	537.8
17	2925835393	367.7	553.5	441.1	553.7
18	2925857917	396.0	547.9	394.2	540.0
20	2937741147	337.9	564.4	348.9	570.0
21	2937810288	465.2	518.7	465.8	515.6
22	2937856748	391.0	530.6	401.6	523.5
23	2937895789	576.4	505.5	595.3	509.7
24	2937961955	412.4	508.6	406.4	526.5
25	2938049569	503.3	562.5	506.0	564.7
26	2938055829	452.9	553.0	444.5	565.4
27	2938228331	405.6	548.8	402.5	553.1
28	2938254204	346.9	559.1	330.7	566.9
29	2938258860	409.7	589.1	408.2	614.9
30	2938268664	367.1	511.9	383.4	554.8
31	2938398472	393.0	559.3	392.9	558.5

32	2938497808	265.4	500.4	390.6	585.4
33	2938498296	417.5	342.1	412.8	345.6
34	2938638637	351.5	525.5	367.4	550.3
35	2938666749	429.6	552.5	418.2	542.8
36	2938756405	402.6	473.5	373.7	488.1
37	2938870373	456.1	587.1	464.1	596.6
38	2938878570	457.1	519.8	435.9	533.8
39	2938969598	510.6	617.6	441.9	631.8
40	2938985411	430.1	571.2	395.2	575.8
41	2939225137	413.6	570.4	412.5	563.1
42	2939317867	389.2	461.4	378.4	483.3
43	2939348426	381.9	537.8	382.5	533.2
44	2939394160	441.1	556.5	449.6	568.9
45	2939489514	394.6	554.4	395.5	555.8
46	2939520972	413.0	559.9	415.6	553.7
47	2939531854	442.8	647.9	453.8	652.3
48	2939622630	331.3	412.3	310.6	360.3
49	2939627477	384.2	534.0	383.0	532.1
51	2939649603	354.4	551.2	350.3	542.6
52	2939652807	258.5	547.7	246.1	533.0
53	2939694336	405.6	560.5	408.9	559.4
54	2939771369	432.2	533.5	430.9	538.9
55	2939776918	385.1	558.1	381.2	569.5
56	2939791816	528.5	581.7	515.9	594.3
57	2940060892	381.1	565.2	387.0	552.7
58	2940088831	447.4	462.0	472.8	461.6
59	2940097386	354.0	552.2	367.1	548.6
60	2940140085	396.2	538.3	394.0	542.4
61	2940147277	295.2	511.5	287.2	530.4
62	2940260013	388.3	540.9	385.6	538.7
63	2940327071	426.8	547.2	442.0	535.9
64	2940351862	357.5	518.4	396.5	541.4
65	2940510474	138.9	375.4	119.0	158.0
66	2940557913	405.9	521.9	389.6	552.7
67	2940670625	389.5	528.7	388.2	532.8
68	2940876978	471.7	481.1	454.2	487.9
69	2940901822	410.5	539.5	415.8	534.0
70	2940908514	366.2	527.5	352.7	525.4
71	2940922093	375.2	555.1	375.6	551.8
72	2940963619	401.2	507.4	402.3	508.7
73	2970356187	404.8	517.9	400.9	525.8
74	2970386160	616.5	535.8	619.2	518.2

75	2970400768	337.8	525.0	286.0	509.2
76	2970414144	409.4	555.2	393.3	566.5
77	2970460831	211.2	450.7	262.2	475.4
78	2970569124	462.7	484.2	419.9	511.5
80	2970611652	488.5	555.3	467.3	538.7
81	2970744291	386.7	498.3	386.6	495.4
82	2970786038	358.7	488.2	359.0	527.0
83	2970791376	527.1	490.5	525.6	506.0
84	2970798013	356.3	523.4	353.2	522.2
85	2970807282	370.2	534.3	348.9	540.1
86	2970850376	337.5	498.1	314.8	483.6
87	2970951335	394.1	536.0	402.5	543.7
89	2971049565	380.7	487.8	314.5	455.3
92	2971214955	486.3	532.7	417.0	568.9

TABLE A.3: Recovered `LITMUS` lags for CIV for independent & simultaneous 2 line fitting

B

Quality Cuts of OzDES Sources & Impact of Simultaneous Lag Recovery

These tables list the impact of multi-line fitting on the lags recovered for all 92 sources in the MgII sample. Each table lists whether a particular lag recovery passes or fails the quality cuts of Penton et al. [32] for both independent and simultaneous lag recovery, as well as the correlation parameter 'R', which describes how similar the lag likelihood distribution is between these two fit cases.

	Independent Fitting		Simultaneous Fitting		
OzDES ID	Passes Cut 1	Passes Cut 2	Passes Cut 1	Passes Cut 2	R
2925373860	TRUE	FALSE	TRUE	FALSE	0.995
2925707000	TRUE	FALSE	TRUE	FALSE	0.999
2970400768	TRUE	FALSE	TRUE	TRUE	0.97
2970807282	TRUE	FALSE	TRUE	FALSE	0.988
2970951335	TRUE	TRUE	TRUE	TRUE	0.816
2971049565	TRUE	FALSE	TRUE	TRUE	0.975

TABLE B.1: Quality cut passes for Hβ two-line OzDES sources.

	Independent Fitting		Simultaneous Fitting		
OzDES ID	Passes Cut 1	Passes Cut 2	Passes Cut 1	Passes Cut 2	R
2925552152	TRUE	TRUE	TRUE	FALSE	0.98
2925858108	TRUE	FALSE	TRUE	FALSE	0.949
2970604169	TRUE	FALSE	TRUE	FALSE	0.578

2971028700	TRUE	TRUE	TRUE	TRUE	1
2971086054	TRUE	TRUE	TRUE	TRUE	1
2971134055	TRUE	FALSE	TRUE	FALSE	0.906
2925372393	TRUE	FALSE	TRUE	FALSE	0.999
2925373860	TRUE	FALSE	TRUE	FALSE	0.979
2925375402	TRUE	TRUE	TRUE	TRUE	1
2925420688	TRUE	TRUE	TRUE	TRUE	0.997
2925515125	TRUE	TRUE	TRUE	TRUE	1
2925523772	TRUE	FALSE	TRUE	FALSE	0.997
2925551147	TRUE	FALSE	TRUE	FALSE	1
2925606181	TRUE	FALSE	TRUE	FALSE	0.994
2925609647	TRUE	FALSE	TRUE	TRUE	0.251
2925637387	TRUE	FALSE	TRUE	FALSE	0.444
2925674035	FALSE	FALSE	FALSE	FALSE	0.974
2925685619	TRUE	TRUE	TRUE	TRUE	1
2925693212	TRUE	TRUE	TRUE	TRUE	1
2925707000	TRUE	FALSE	TRUE	FALSE	1
2925718880	TRUE	TRUE	TRUE	TRUE	1
2925786831	FALSE	FALSE	FALSE	FALSE	0.999
2925835393	TRUE	FALSE	FALSE	FALSE	0.976
2925857917	TRUE	TRUE	TRUE	TRUE	0.999
2937741147	TRUE	TRUE	TRUE	TRUE	1
2937810288	FALSE	FALSE	FALSE	FALSE	0.999
2937856748	TRUE	FALSE	TRUE	FALSE	1
2937895789	TRUE	FALSE	TRUE	FALSE	0.842
2937961955	TRUE	TRUE	TRUE	TRUE	1
2938049569	TRUE	TRUE	TRUE	TRUE	1
2938055829	TRUE	FALSE	TRUE	TRUE	0.823
2938228331	TRUE	TRUE	TRUE	FALSE	0.989
2938254204	TRUE	TRUE	TRUE	TRUE	1
2938258860	FALSE	FALSE	FALSE	FALSE	0.994
2938268664	TRUE	FALSE	TRUE	FALSE	0.993
2938398472	TRUE	TRUE	TRUE	TRUE	1
2938497808	TRUE	TRUE	TRUE	TRUE	1
2938498296	TRUE	TRUE	TRUE	FALSE	0.966
2938638637	TRUE	FALSE	TRUE	FALSE	0.967
2938666749	TRUE	FALSE	TRUE	FALSE	1
2938756405	TRUE	FALSE	TRUE	FALSE	0.998
2938870373	TRUE	FALSE	TRUE	FALSE	0.954
2938878570	TRUE	TRUE	TRUE	TRUE	0.996
2938969598	TRUE	FALSE	TRUE	FALSE	0.675
2938985411	TRUE	FALSE	TRUE	FALSE	0.876

2939225137	TRUE	FALSE	TRUE	FALSE	0.946
2939317867	TRUE	TRUE	TRUE	FALSE	0.705
2939348426	TRUE	FALSE	TRUE	FALSE	0.805
2939394160	TRUE	TRUE	TRUE	TRUE	1
2939489514	TRUE	TRUE	TRUE	FALSE	0.95
2939520972	TRUE	FALSE	TRUE	FALSE	1
2939531854	TRUE	FALSE	TRUE	FALSE	0.999
2939622630	TRUE	TRUE	FALSE	FALSE	0.717
2939627477	FALSE	FALSE	TRUE	FALSE	0.94
2939649603	TRUE	TRUE	TRUE	TRUE	0.999
2939652807	TRUE	TRUE	TRUE	TRUE	0.999
2939694336	TRUE	FALSE	TRUE	TRUE	0.841
2939771369	TRUE	TRUE	TRUE	FALSE	0.983
2939776918	TRUE	FALSE	TRUE	TRUE	0.912
2939791816	TRUE	FALSE	TRUE	TRUE	0.993
2940060892	TRUE	TRUE	TRUE	TRUE	1
2940088831	TRUE	FALSE	TRUE	FALSE	0.912
2940097386	TRUE	TRUE	TRUE	FALSE	0.985
2940140085	TRUE	TRUE	TRUE	TRUE	1
2940147277	TRUE	TRUE	TRUE	TRUE	0.995
2940260013	TRUE	FALSE	TRUE	FALSE	0.968
2940327071	TRUE	TRUE	TRUE	TRUE	0.99
2940351862	TRUE	TRUE	TRUE	TRUE	0.987
2940510474	FALSE	FALSE	FALSE	FALSE	0.995
2940557913	TRUE	FALSE	TRUE	FALSE	0.936
2940670625	TRUE	FALSE	FALSE	FALSE	0.741
2940876978	TRUE	FALSE	TRUE	FALSE	0.999
2940901822	TRUE	FALSE	TRUE	FALSE	0.999
2940908514	TRUE	TRUE	TRUE	TRUE	0.999
2940922093	TRUE	TRUE	TRUE	TRUE	0.997
2940963619	TRUE	FALSE	TRUE	FALSE	0.87
2970356187	TRUE	TRUE	TRUE	TRUE	0.999
2970386160	TRUE	TRUE	TRUE	TRUE	1
2970400768	TRUE	TRUE	TRUE	TRUE	0.994
2970414144	TRUE	TRUE	TRUE	TRUE	1
2970460831	TRUE	FALSE	TRUE	TRUE	0.356
2970569124	TRUE	TRUE	TRUE	TRUE	1
2970611652	TRUE	TRUE	TRUE	TRUE	1
2970744291	TRUE	TRUE	TRUE	TRUE	0.998
2970786038	TRUE	TRUE	TRUE	TRUE	0.987
2970791376	TRUE	TRUE	TRUE	TRUE	1
2970798013	TRUE	TRUE	TRUE	TRUE	0.994

2970807282	TRUE	FALSE	TRUE	FALSE	0.955
2970850376	TRUE	FALSE	TRUE	FALSE	0.986
2970951335	TRUE	FALSE	TRUE	TRUE	0.961
2971049565	TRUE	TRUE	TRUE	TRUE	0.983
2971214955	TRUE	TRUE	TRUE	TRUE	1

TABLE B.2: Quality cut passes for MgII two-line OzDES sources.

	Independent Fitting		Simultaneous Fitting		
OzDES ID	Passes Cut 1	Passes Cut 2	Passes Cut 1	Passes Cut 2	R
2925552152	FALSE	FALSE	FALSE	FALSE	0.999
2925858108	FALSE	FALSE	FALSE	FALSE	0.995
2970604169	FALSE	FALSE	FALSE	FALSE	0.996
2971028700	FALSE	FALSE	FALSE	FALSE	0.999
2971086054	FALSE	FALSE	FALSE	FALSE	0.997
2971134055	TRUE	FALSE	TRUE	FALSE	0.867
2925372393	FALSE	FALSE	FALSE	FALSE	0.999
2925375402	FALSE	FALSE	TRUE	FALSE	0.977
2925420688	TRUE	FALSE	TRUE	FALSE	0.993
2925515125	FALSE	FALSE	FALSE	FALSE	0.994
2925523772	TRUE	FALSE	FALSE	FALSE	0.999
2925551147	FALSE	FALSE	FALSE	FALSE	0.999
2925606181	TRUE	FALSE	TRUE	FALSE	0.998
2925609647	FALSE	FALSE	TRUE	FALSE	0.999
2925637387	TRUE	FALSE	FALSE	FALSE	0.928
2925674035	TRUE	FALSE	TRUE	FALSE	0.999
2925685619	FALSE	FALSE	FALSE	FALSE	0.968
2925693212	FALSE	FALSE	FALSE	FALSE	0.999
2925718880	FALSE	FALSE	FALSE	FALSE	0.998
2925786831	FALSE	FALSE	FALSE	FALSE	1
2925835393	TRUE	FALSE	TRUE	FALSE	0.993
2925857917	FALSE	FALSE	FALSE	FALSE	0.989
2937741147	TRUE	FALSE	FALSE	FALSE	0.991
2937810288	FALSE	FALSE	FALSE	FALSE	0.998
2937856748	FALSE	FALSE	FALSE	FALSE	0.998
2937895789	FALSE	FALSE	FALSE	FALSE	0.999
2937961955	FALSE	FALSE	FALSE	FALSE	0.986
2938049569	FALSE	FALSE	FALSE	FALSE	0.973
2938055829	TRUE	FALSE	FALSE	FALSE	0.964
2938228331	TRUE	FALSE	TRUE	FALSE	0.998

2938254204	FALSE	FALSE	FALSE	FALSE	0.865
2938258860	TRUE	FALSE	TRUE	FALSE	0.998
2938268664	TRUE	FALSE	FALSE	FALSE	0.99
2938398472	FALSE	FALSE	FALSE	FALSE	0.998
2938497808	TRUE	FALSE	FALSE	FALSE	0.985
2938498296	FALSE	FALSE	TRUE	FALSE	0.99
2938638637	TRUE	FALSE	TRUE	FALSE	0.986
2938666749	FALSE	FALSE	FALSE	FALSE	0.977
2938756405	FALSE	FALSE	TRUE	FALSE	0.994
2938870373	FALSE	FALSE	FALSE	FALSE	0.998
2938878570	FALSE	FALSE	TRUE	FALSE	0.989
2938969598	TRUE	FALSE	FALSE	FALSE	0.999
2938985411	FALSE	FALSE	FALSE	FALSE	0.996
2939225137	FALSE	FALSE	FALSE	FALSE	0.999
2939317867	FALSE	FALSE	FALSE	FALSE	0.998
2939348426	FALSE	FALSE	FALSE	FALSE	0.994
2939394160	TRUE	FALSE	TRUE	FALSE	0.939
2939489514	TRUE	FALSE	FALSE	FALSE	0.999
2939520972	FALSE	FALSE	FALSE	FALSE	0.999
2939531854	FALSE	FALSE	FALSE	FALSE	0.997
2939622630	FALSE	FALSE	FALSE	FALSE	0.999
2939627477	TRUE	FALSE	TRUE	FALSE	0.999
2939649603	FALSE	FALSE	FALSE	FALSE	0.998
2939652807	FALSE	FALSE	FALSE	FALSE	0.994
2939694336	TRUE	FALSE	FALSE	FALSE	0.999
2939771369	FALSE	FALSE	FALSE	FALSE	0.952
2939776918	FALSE	FALSE	FALSE	FALSE	0.995
2939791816	TRUE	FALSE	FALSE	FALSE	0.998
2940060892	FALSE	FALSE	FALSE	FALSE	0.979
2940088831	FALSE	FALSE	TRUE	FALSE	0.999
2940097386	FALSE	FALSE	FALSE	FALSE	0.998
2940140085	FALSE	FALSE	FALSE	FALSE	0.98
2940147277	TRUE	FALSE	TRUE	FALSE	0.897
2940260013	FALSE	FALSE	FALSE	FALSE	0.987
2940327071	FALSE	FALSE	FALSE	FALSE	0.998
2940351862	FALSE	FALSE	FALSE	FALSE	0.996
2940510474	TRUE	FALSE	FALSE	FALSE	0.996
2940557913	FALSE	FALSE	FALSE	FALSE	0.997
2940670625	FALSE	FALSE	FALSE	FALSE	0.999
2940876978	TRUE	FALSE	TRUE	FALSE	1
2940901822	TRUE	FALSE	TRUE	FALSE	0.999
2940908514	TRUE	FALSE	TRUE	FALSE	0.996

2940922093	TRUE	FALSE	FALSE	FALSE	0.95
2940963619	FALSE	FALSE	FALSE	FALSE	0.995
2970356187	FALSE	FALSE	TRUE	FALSE	0.969
2970386160	FALSE	FALSE	TRUE	FALSE	0.978
2970414144	FALSE	FALSE	FALSE	FALSE	0.991
2970460831	FALSE	FALSE	FALSE	FALSE	0.998
2970569124	TRUE	FALSE	FALSE	FALSE	0.98
2970611652	FALSE	FALSE	TRUE	FALSE	0.981
2970744291	FALSE	FALSE	FALSE	FALSE	0.999
2970786038	FALSE	FALSE	FALSE	FALSE	0.991
2970791376	FALSE	FALSE	FALSE	FALSE	0.976
2970798013	TRUE	FALSE	FALSE	FALSE	0.998
2970850376	FALSE	FALSE	FALSE	FALSE	0.976
2971214955	FALSE	FALSE	TRUE	FALSE	0.956

Table B.3: Quality cut passes for CIV two-line OzDES sources.

C

Summary of Signal to Noise Ratios For OzDES Two-Line Sources

The following table lists a summary of the measurement quality for all sources in the OzDES two-line sample. For each source, the number of measurements and signal to noise ratio is provided for the continuum and each of the two response line signals.

Index	OzDES ID	Signal Type	Continuum Measurements	Response 1 Measurements	Response 2 Measurements	Continuum SNR	Response 1 SNR	Response 2 SNR
0	2925372393	MgII / CIV	131	23	17	0.12	0.38	0.00
1	2925373860	MgII / CIV	137	23	16	0.17	0.40	0.00
2	2925375402	MgII / CIV	132	24	19	0.10	0.17	0.00
3	2925420688	MgII / CIV	139	23	19	0.16	0.17	0.00
4	2925515125	MgII / CIV	136	20	15	0.23	0.22	0.00
5	2925523772	MgII / CIV	134	23	16	0.10	0.21	0.00
6	2925551147	MgII / CIV	129	20	16	0.28	0.08	0.00

7	2925552152	Hβ / MgII	133	16	21	0.38	0.09	0.11
8	2925606181	MgII / CIV	130	21	15	0.31	0.21	0.00
9	2925609647	MgII / CIV	132	20	8	0.07	0.30	0.00
10	2925637387	MgII / CIV	135	20	16	0.21	0.27	0.00
11	2925674035	MgII / CIV	132	21	15	0.09	0.52	0.00
12	2925685619	MgII / CIV	130	21	15	0.26	0.99	0.00
13	2925693212	MgII / CIV	124	19	16	0.15	0.17	0.00
14	2925707000	MgII / CIV	131	19	16	0.05	0.37	0.00
15	2925718880	MgII / CIV	129	21	16	0.12	0.35	0.00
16	2925786831	MgII / CIV	135	6	6	0.05	0.26	0.00
17	2925835393	MgII / CIV	134	19	15	0.17	0.45	0.00
18	2925857917	MgII / CIV	130	20	16	0.13	0.17	0.00
19	2925858108	Hβ / MgII	133	15	21	0.23	0.09	0.11
20	2937741147	MgII / CIV	138	25	20	0.11	0.26	0.00
21	2937810288	MgII / CIV	137	22	20	0.20	0.15	0.00
22	2937856748	MgII / CIV	135	24	16	0.10	0.15	0.00
23	2937895789	MgII / CIV	136	18	15	0.15	0.04	0.00
24	2937961955	MgII / CIV	133	29	21	0.14	0.08	0.00
25	2938049569	MgII / CIV	143	21	18	0.18	0.08	0.00
26	2938055829	MgII / CIV	142	21	19	0.19	0.12	0.00
27	2938228331	MgII / CIV	140	22	19	0.09	0.13	0.00
28	2938254204	MgII / CIV	131	22	16	0.33	0.44	0.00
29	2938258860	MgII / CIV	142	21	5	0.15	0.34	0.00
30	2938268664	MgII / CIV	94	21	19	0.26	0.35	0.00
31	2938398472	MgII / CIV	136	32	23	0.12	0.23	0.00
32	2938497808	MgII / CIV	137	18	16	0.22	0.21	0.00
33	2938498296	MgII / CIV	132	13	12	0.14	0.00	0.47
34	2938638637	MgII / CIV	135	22	19	0.31	0.15	0.00
35	2938666749	MgII / CIV	136	23	20	0.32	0.19	0.00
36	2938756405	MgII / CIV	134	33	25	0.20	0.23	0.00

37	2938870373	MgII / CIV	132	22	18	0.10	0.36	0.00
38	2938878570	MgII / CIV	133	30	25	0.27	0.22	0.00
39	2938969598	MgII / CIV	136	33	24	0.18	0.15	0.00
40	2938985411	MgII / CIV	135	32	24	0.11	0.10	0.00
41	2939225137	MgII / CIV	170	29	21	0.16	0.13	0.00
42	2939317867	MgII / CIV	156	28	14	0.10	0.70	0.00
43	2939348426	MgII / CIV	149	40	24	0.05	0.32	0.00
44	2939394160	MgII / CIV	139	28	22	0.17	0.28	0.00
45	2939489514	MgII / CIV	169	32	20	0.18	0.25	0.00
46	2939520972	MgII / CIV	166	32	21	0.11	0.18	0.00
47	2939531854	MgII / CIV	154	29	19	0.12	0.13	0.00
48	2939622630	MgII / CIV	205	34	22	0.17	0.17	0.00
49	2939627477	MgII / CIV	172	33	21	0.09	0.21	0.00
51	2939649603	MgII / CIV	144	25	21	0.47	0.26	0.00
52	2939652807	MgII / CIV	106	23	19	0.21	0.25	0.00
53	2939694336	MgII / CIV	154	28	22	0.10	0.14	0.00
54	2939771369	MgII / CIV	85	24	19	0.10	0.24	0.00
55	2939776918	MgII / CIV	147	25	22	0.12	0.26	0.00
56	2939791816	MgII / CIV	225	55	26	0.12	0.14	0.00
57	2940060892	MgII / CIV	139	24	20	0.06	0.13	0.00
58	2940088831	MgII / CIV	131	23	19	0.14	0.30	0.00
59	2940097386	MgII / CIV	139	23	19	0.09	0.18	0.00
60	2940140085	MgII / CIV	162	40	27	0.00	0.28	0.00
61	2940147277	MgII / CIV	176	39	26	0.19	0.31	0.00
62	2940260013	MgII / CIV	149	27	22	0.07	0.13	0.00
63	2940327071	MgII / CIV	142	22	19	0.11	0.16	0.00
64	2940351862	MgII / CIV	134	24	20	0.11	0.75	0.00
65	2940510474	MgII / CIV	180	42	27	0.11	0.32	0.00
66	2940557913	MgII / CIV	119	36	23	0.17	0.37	0.00
67	2940670625	MgII / CIV	152	16	12	0.09	0.55	0.00

68	2940876978	MgII / CIV	186	35	25	0.14	0.17	0.00
69	2940901822	MgII / CIV	190	38	25	0.18	0.15	0.00
70	2940908514	MgII / CIV	182	40	27	0.13	0.12	0.00
71	2940922093	MgII / CIV	197	40	26	0.25	0.13	0.00
72	2940963619	MgII / CIV	147	34	23	0.16	0.19	0.00
73	2970356187	MgII / CIV	115	31	17	0.20	0.17	0.00
74	2970386160	MgII / CIV	115	39	22	0.08	0.10	0.00
75	2970400768	MgII / CIV	142	38	22	0.10	0.48	0.00
76	2970414144	MgII / CIV	145	42	24	0.20	0.18	0.00
77	2970460831	MgII / CIV	178	63	23	0.19	0.14	0.00
78	2970569124	MgII / CIV	151	38	24	0.14	0.24	0.00
79	2970604169	Hβ / MgII	143	21	35	0.13	0.21	0.16
80	2970611652	MgII / CIV	141	25	21	0.17	0.31	0.00
81	2970744291	MgII / CIV	141	39	21	0.10	0.25	0.00
82	2970786038	MgII / CIV	146	38	22	0.10	0.30	0.00
83	2970791376	MgII / CIV	145	39	5	0.20	0.33	0.00
84	2970798013	MgII / CIV	147	37	24	0.13	0.12	0.00
85	2970807282	MgII / CIV	145	41	24	0.13	0.20	0.00
86	2970850376	MgII / CIV	147	41	24	0.11	0.23	0.00
87	2970951335	MgII / CIV	117	43	23	0.08	0.11	0.00
88	2971028700	Hβ / MgII	117	22	38	0.08	0.12	0.22
89	2971049565	MgII / CIV	115	34	18	0.09	0.24	0.00
90	2971086054	Hβ / MgII	114	23	35	0.15	0.11	0.27
91	2971134055	Hβ / MgII	115	22	39	0.11	0.09	0.17
92	2971214955	MgII / CIV	116	39	23	0.19	0.21	0.00

TABLE C.1: Summary of measurements for all OzDES sources with two emission line signals

References

[1] Joshua G. Albert. Jaxns: a high-performance nested sampling package based on jax, 2020.

[2] Greg Ashton, Noam Bernstein, Johannes Buchner, Xi Chen, Gábor Csányi, Andrew Fowlie, Farhan Feroz, Matthew Griffiths, Will Handley, Michael Habeck, Edward Higson, Michael Hobson, Anthony Lasenby, David Parkinson, Livia B. Pártay, Matthew Pitkin, Doris Schneider, Joshua S. Speagle, Leah South, John Veitch, Philipp Wacker, David J. Wales, and David Yallup. Nested sampling for physical scientists. *Nature Reviews Methods Primers*, 2(1), may 2022. doi: 10.1038/s43586-022-00121-x. URL `https://doi.org/10.1038%2Fs43586-022-00121-x`.

[3] Misty C. Bentz. AGN reverberation mapping. In *Astronomy at High Angular Resolution*, pages 249–266. Springer International Publishing, 2016. doi: 10.1007/978-3-319-39739-9_13. URL `https://doi.org/10.1007%2F978-3-319-39739-9_13`.

[4] Michael Betancourt. A conceptual introduction to hamiltonian monte carlo, 2018.

[5] R. D. Blandford and C. F. McKee. Reverberation mapping of the emission line regions of Seyfert galaxies and quasars. , 255:419–439, April 1982. doi: 10.1086/159843.

[6] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018. URL `http://github.com/google/jax`.

[7] Brendon J. Brewer and Tom M. Elliott. Hierarchical reverberation mapping. *Monthly Notices of the Royal Astronomical Society: Letters*, 439(1):L31–L35, jan 2014. doi: 10.1093/mnrasl/slt174. URL `https://doi.org/10.1093%2Fmnrasl%2Fslt174`.

[8] Brendon J. Brewer, Livia B. Pártay, and Gábor Csányi. Diffusive nested sampling, 2010.

[9] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng, editors. *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC, may 2011. doi: 10.1201/b10905. URL `https://doi.org/10.1201%2Fb10905`.

[10] Edward M. Cackett, Keith Horne, and Hartmut Winkler. Testing thermal reprocessing in active galactic nuclei accretion discs. *Monthly Notices of the Royal Astronomical*

Society, 380(2):669–682, aug 2007. doi: 10.1111/j.1365-2966.2007.12098.x. URL `https://doi.org/10.1111%2Fj.1365-2966.2007.12098.x`.

[11] Colin Carroll. Hamiltonian monte carlo from scratch, 2019. URL `https://colindcarroll.com/2019/04/11/hamiltonian-monte-carlo-from-scratch/`.

[12] Víctor Elvira and Luca Martino. Advances in importance sampling, 2022.

[13] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: an efficient and robust bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, 398(4):1601–1614, oct 2009. doi: 10.1111/j.1365-2966.2009.14548.x. URL `https://doi.org/10.1111%2Fj.1365-2966.2009.14548.x`.

[14] S. Fine, T. Shanks, P. Green, B. C. Kelly, S. M. Croom, R. L. Webster, E. Berger, R. Chornock, W. S. Burgett, K. C. Chambers, N. Kaiser, and P. A. Price. Stacked reverberation mapping. *Monthly Notices of the Royal Astronomical Society: Letters*, 434(1):L16–L20, jun 2013. doi: 10.1093/mnrasl/slt069. URL `https://doi.org/10.1093%2Fmnrasl%2Fslt069`.

[15] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. ttemcee/tt: The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306–312, mar 2013. doi: 10.1086/670067. URL `https://doi.org/10.1086%2F670067`.

[16] C. Martin Gaskell and Bradley M. Peterson. The Accuracy of Cross-Correlation Estimates of Quasar Emission-Line Region Sizes. , 65:1, September 1987. doi: 10.1086/191216.

[17] Jonathan Goodman and Jonathan Weare. Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80, January 2010. doi: 10.2140/camcos.2010.5.65.

[18] C. J. Grier, B. M. Peterson, R. W. Pogge, K. D. Denney, M. C. Bentz, Paul Martini, S. G. Sergeev, S. Kaspi, T. Minezaki, Y. Zu, C. S. Kochanek, R. Siverd, B. Shappee, K. Z. Stanek, C. Araya Salvo, T. G. Beatty, J. C. Bird, D. J. Bord, G. A. Borman, X. Che, C. Chen, S. A. Cohen, M. Dietrich, V. T. Doroshenko, T. Drake, Yu. S. Efimov, N. Free, I. Ginsburg, C. B. Henderson, A. L. King, S. Koshida, K. Mogren, M. Molina, A. M. Mosquera, S. V. Nazarov, D. N. Okhmat, O. Pejcha, S. Rafter, J. C. Shields, J. Skowron, D. M. Szczygiel, M. Valluri, and J. L. van Saders. REVERBERATION MAPPING RESULTS FOR FIVE SEYFERT 1 GALAXIES. *The Astrophysical Journal*, 755(1):60, jul 2012. doi: 10.1088/0004-637x/755/1/60. URL `https://doi.org/10.1088%2F0004-637x%2F755%2F1%2F60`.

[19] C. J. Grier, Yue Shen, Keith Horne, W. N. Brandt, J. R. Trump, P. B. Hall, K. Kinemuchi, David Starkey, D. P. Schneider, Luis C. Ho, Y. Homayouni, Jennifer I-Hsiu Li, Ian D. McGreer, B. M. Peterson, Dmitry Bizyaev, Yuguang Chen, K. S. Dawson, Sarah Eftekharzadeh, Yucheng Guo, Siyao Jia, Linhua Jiang, Jean-Paul Kneib, Feng Li, Zefeng

Li, Jundan Nie, Audrey Oravetz, Daniel Oravetz, Kaike Pan, Patrick Petitjean, Kara A. Ponder, Jesse Rogerson, M. Vivek, Tianmeng Zhang, and Hu Zou. The sloan digital sky survey reverberation mapping project: Initial c iv lag results from four years of data. *The Astrophysical Journal*, 887(1):38, dec 2019. doi: 10.3847/1538-4357/ab4ea5. URL `https://doi.org/10.3847%2F1538-4357%2Fab4ea5`.

[20] W. J. Handley, M. P. Hobson, and A. N. Lasenby. polychord: next-generation nested sampling. *Monthly Notices of the Royal Astronomical Society*, 453(4):4385–4399, sep 2015. doi: 10.1093/mnras/stv1911. URL `https://doi.org/10.1093%2Fmnras%2Fstv1911`.

[21] Y. Homayouni, Jonathan R. Trump, C. J. Grier, Keith Horne, Yue Shen, W. N. Brandt, Kyle S. Dawson, Gloria Fonseca Alvarez, Paul J. Green, P. B. Hall, Juan V. Herná ndez Santisteban, Luis C. Ho, Karen Kinemuchi, C. S. Kochanek, Jennifer I-Hsiu Li, B. M. Peterson, D. P. Schneider, D. A. Starkey, Dmitry Bizyaev, Kaike Pan, Daniel Oravetz, and Audrey Simmons. The sloan digital sky survey reverberation mapping project: Mg ii lag results from four years of monitoring. *The Astrophysical Journal*, 901 (1):55, sep 2020. doi: 10.3847/1538-4357/ababa9. URL `https://doi.org/10.3847%2F1538-4357%2Fababa9`.

[22] J K Hoormann, P Martini, T M Davis, A King, C Lidman, D Mudd, R Sharp, N E Sommer, B E Tucker, Z Yu, S Allam, J Asorey, S Avila, M Banerji, D Brooks, E Buckley-Geer, D L Burke, J Calcino, A Carnero Rosell, D Carollo, M Carrasco Kind, J Carretero, F J Castander, M Childress, J De Vicente, S Desai, H T Diehl, P Doel, B Flaugher, P Fosalba, J Frieman, J García-Bellido, D W Gerdes, D Gruen, G Gutierrez, W G Hartley, S R Hinton, D L Hollowood, K Honscheid, B Hoyle, D J James, E Krause, K Kuehn, N Kuropatkin, G F Lewis, M Lima, E Macaulay, M A G Maia, F Menanteau, C J Miller, R Miquel, A Möller, A A Plazas, A K Romer, A Roodman, E Sanchez, V Scarpine, M Schubnell, S Serrano, I Sevilla-Noarbe, M Smith, R C Smith, M Soares-Santos, F Sobreira, E Suchyta, E Swann, M E C Swanson, G Tarle, S A Uddin, and (DES Collaboration). Civ black hole mass measurements with the Australian Dark Energy Survey (OzDES). *Monthly Notices of the Royal Astronomical Society*, 487(3): 3650–3663, 06 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz1539. URL `https://doi.org/10.1093/mnras/stz1539`.

[23] David Huijser, Jesse Goodman, and Brendon J. Brewer. Properties of the affine invariant ensemble sampler in high dimensions, 2015. URL `https://arxiv.org/abs/1509.02230`.

[24] Ž eljko Ivezić and Chelsea MacLeod. Optical variability of quasars: a damped random walk. *Proceedings of the International Astronomical Union*, 9(S304):395–398, oct 2013. doi: 10.1017/s1743921314004396. URL `https://doi.org/10.1017%2Fs1743921314004396`.

[25] Shai Kaspi, Paul S. Smith, Hagai Netzer, Dan Maoz, Buell T. Jannuzi, and Uriel Giveon. Reverberation measurements for 17 quasars and the size-mass-luminosity relations in

active galactic nuclei. *The Astrophysical Journal*, 533(2):631–649, apr 2000. doi: 10.1086/308704. URL `https://doi.org/10.1086%2F308704`.

[26] Anthea L. King, Paul Martini, Tamara M. Davis, K. D. Denney, C. S. Kochanek, Bradley M. Peterson, Andreas Skielboe, Marianne Vestergaard, Eric Huff, Darach Watson, Manda Banerji, Richard McMahon, Rob Sharp, and C. Lidman. Simulations of the OzDES AGN reverberation mapping project. *Monthly Notices of the Royal Astronomical Society*, 453(2):1701–1726, aug 2015. doi: 10.1093/mnras/stv1718. URL `https://doi.org/10.1093%2Fmnras%2Fstv1718`.

[27] Szymon Kozłowski. REVISITING STOCHASTIC VARIABILITY OF AGNs WITH STRUCTURE FUNCTIONS. *The Astrophysical Journal*, 826(2):118, jul 2016. doi: 10.3847/0004-637x/826/2/118. URL `https://doi.org/10.3847%2F0004-637x%2F826%2F2%2F118`.

[28] C. L. MacLeod, Ž . Ivezić, C. S. Kochanek, S. Kozłowski, B. Kelly, E. Bullock, A. Kimball, B. Sesar, D. Westman, K. Brooks, R. Gibson, A. C. Becker, and W. H. de Vries. MODELING THE TIME VARIABILITY OF SDSS STRIPE 82 QUASARS AS a DAMPED RANDOM WALK. *The Astrophysical Journal*, 721(2):1014–1033, sep 2010. doi: 10.1088/0004-637x/721/2/1014. URL `https://doi.org/10.1088%2F0004-637x%2F721%2F2%2F1014`.

[29] U Malik, R Sharp, A Penton, Z Yu, P Martini, C Lidman, B E Tucker, T M Davis, G F Lewis, M Aguena, S Allam, O Alves, F Andrade-Oliveira, J Asorey, D Bacon, E Bertin, S Bocquet, D Brooks, D L Burke, A Carnero Rosell, D Carollo, M Carrasco Kind, J Carretero, M Costanzi, L N da Costa, M E S Pereira, J De Vicente, S Desai, H T Diehl, P Doel, S Everett, I Ferrero, J Frieman, J Garcí a-Bellido, D W Gerdes, D Gruen, R A Gruendl, J Gschwend, S R Hinton, D L Hollowood, K Honscheid, D J James, K Kuehn, J L Marshall, J Mena-Fernández, F Menanteau, R Miquel, R L C Ogando, A Palmese, F Paz-Chinchón, A Pieres, A A Plazas Malagón, M Raveri, M Rodriguez-Monroy, A K Romer, E Sanchez, V Scarpine, I Sevilla-Noarbe, M Smith, M Soares-Santos, E Suchyta, M E C Swanson, G Tarle, G Taylor, D L Tucker, N Weaverdyck, and R D Wilkinson. OzDES reverberation mapping program: H lags from the 6-yr survey. *Monthly Notices of the Royal Astronomical Society*, 520(2):2009–2023, jan 2023. doi: 10.1093/mnras/stad145. URL `https://doi.org/10.1093%2Fmnras%2Fstad145`.

[30] Umang Malik, Rob Sharp, Paul Martini, Tamara M Davis, Brad E Tucker, Zhefu Yu, Andrew Penton, Geraint F Lewis, and Josh Calcino. Observational window effects on multi-object reverberation mapping. *Monthly Notices of the Royal Astronomical Society*, 516(3):3238–3253, aug 2022. doi: 10.1093/mnras/stac2263. URL `https://doi.org/10.1093%2Fmnras%2Fstac2263`.

[31] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953. doi: 10.1063/1.1699114. URL `https://doi.org/10.1063/1.1699114`.

[32] A Penton, U Malik, T M Davis, P Martini, Z Yu, R Sharp, C Lidman, B E Tucker, J K Hoormann, M Aguena, S Allam, J Annis, J Asorey, D Bacon, E Bertin, S Bhargava, D Brooks, J Calcino, A Carnero Rosell, D Carollo, M Carrasco Kind, J Carretero, M Costanzi, L N da Costa, M E S Pereira, J De Vicente, H T Diehl, T F Eifler, S Everett, I Ferrero, P Fosalba, J Frieman, J Garcí a-Bellido, E Gaztanaga, D W Gerdes, D Gruen, R A Gruendl, J Gschwend, G Gutierrez, S R Hinton, D L Hollowood, K Honscheid, D J James, A G Kim, K Kuehn, N Kuropatkin, M A G Maia, J L Marshall, F Menanteau, R Miquel, R Morgan, A Möller, A Palmese, F Paz-Chinchón, A A Plazas, A K Romer, E Sanchez, V Scarpine, D Scolnic, S Serrano, M Smith, E Suchyta, M E C Swanson, G Tarle, C To, S A Uddin, T N Varga, W Wester, R D Wilkinson, and G Lewis and. OzDES reverberation mapping program: Lag recovery reliability for 6-yr c scpiv/scp analysis. *Monthly Notices of the Royal Astronomical Society*, 509(3): 4008–4023, oct 2021. doi: 10.1093/mnras/stab3027. URL `https://doi.org/10.1093%2Fmnras%2Fstab3027`.

[33] B. M. Peterson and K. Horne. Echo mapping of active galactic nuclei. *Astronomische Nachrichten*, 325(3):248–251, mar 2004. doi: 10.1002/asna.200310207. URL `https://doi.org/10.1002%2Fasna.200310207`.

[34] Bradley M. Peterson. Reverberation Mapping of Active Galactic Nuclei. *Publications of the Astronomical Society of the Pacific*, 105:247, March 1993. doi: 10.1086/133140/. URL `https://iopscience.iop.org/article/10.1086/133140/pdf`.

[35] Bradley M. Peterson. VARIABILITY OF ACTIVE GALACTIC NUCLEI. In *Advanced Lectures on the Starburst-AGN Connection*, pages 3–68. WORLD SCIENTIFIC, apr 2001. doi: 10.1142/9789812811318_0002. URL `https://doi.org/10.1142%2F9789812811318_0002`.

[36] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning.* Adaptive computation and machine learning. MIT Press, 2006. ISBN 026218253X.

[37] George B. Rybicki and William H. Press. Interpolation, Realization, and Reconstruction of Noisy, Irregularly Sampled Data. , 398:169, October 1992. doi: 10.1086/171845.

[38] John Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833 – 859, 2006. doi: 10.1214/06-BA127. URL `https://doi.org/10.1214/06-BA127`.

[39] D. A. Starkey, Keith Horne, and C. Villforth. Accretion disc time lag distributions: applying CREAM to simulated AGN light curves. *Monthly Notices of the Royal Astronomical Society*, 456(2):1960–1973, dec 2015. doi: 10.1093/mnras/stv2744. URL `https://doi.org/10.1093%2Fmnras%2Fstv2744`.

[40] C. Megan Urry and Paolo Padovani. Unified schemes for radio-loud active galactic nuclei. *Publications of the Astronomical Society of the Pacific*, 107:803, sep 1995. doi: 10.1086/133630. URL `https://doi.org/10.1086%2F133630`.

[41] Marc Wegmann, Domenique Zipperling, Jonas Hillenbrand, and Jürgen Fleischer. A review of systematic selection of clustering algorithms and their evaluation, 2021.

[42] Z Yu, C S Kochanek, B M Peterson, Y Zu, W N Brandt, E M Cackett, M M Fausnaugh, and I M McHardy. On reverberation mapping lag uncertainties. *Monthly Notices of the Royal Astronomical Society*, 491(4):6045–6064, dec 2019. doi: 10.1093/mnras/stz3464. URL `https://doi.org/10.1093%2Fmnras%2Fstz3464`.

[43] Zhefu Yu, Paul Martini, T. M. Davis, R. A. Gruendl, J. K. Hoormann, C. S. Kochanek, C. Lidman, D. Mudd, B. M. Peterson, W. Wester, S. Allam, J. Annis, J. Asorey, S. Avila, M. Banerji, E. Bertin, D. Brooks, E. Buckley-Geer, J. Calcino, A. Carnero Rosell, D. Carollo, M. Carrasco Kind, J. Carretero, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, J. De Vicente, S. Desai, H. T. Diehl, P. Doel, T. F. Eifler, B. Flaugher, P. Fosalba, J. Frieman, J. Garcí a-Bellido, E. Gaztanaga, K. Glazebrook, D. Gruen, J. Gschwend, G. Gutierrez, W. G. Hartley, S. R. Hinton, D. L. Hollowood, K. Honscheid, B. Hoyle, D. J. James, A. G. Kim, E. Krause, K. Kuehn, N. Kuropatkin, G. F. Lewis, M. Lima, E. Macaulay, M. A. G. Maia, J. L. Marshall, F. Menanteau, R. Miquel, A. Möller, A. A. Plazas, A. K. Romer, E. Sanchez, V. Scarpine, M. Schubnell, S. Serrano, M. Smith, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, E. Swann, M. E. C. Swanson, G. Tarle, B. E. Tucker, D. L. Tucker, and V. Vikram. Quasar accretion disk sizes from continuum reverberation mapping in the DES standard-star fields. *The Astrophysical Journal Supplement Series*, 246(1):16, jan 2020. doi: 10.3847/1538-4365/ab5e7a. URL `https://doi.org/10.3847%2F1538-4365%2Fab5e7a`.

[44] Zhefu Yu, Paul Martini, A Penton, T M Davis, U Malik, C Lidman, B E Tucker, R Sharp, C S Kochanek, B M Peterson, M Aguena, S Allam, F Andrade-Oliveira, J Annis, J Asorey, E Bertin, D Brooks, D L Burke, J Calcino, A Carnero Rosell, D Carollo, M Carrasco Kind, M Costanzi, L N da Costa, M E S da Silva Pereira, H T Diehl, S Everett, I Ferrero, B Flaugher, J Frieman, J Garcí a-Bellido, E Gaztanaga, D W Gerdes, D Gruen, R A Gruendl, J Gschwend, G Gutierrez, S R Hinton, D L Hollowood, D J James, A G Kim, R Kron, K Kuehn, N Kuropatkin, G F Lewis, M A G Maia, M March, J L Marshall, F Menanteau, R Miquel, R Morgan, A Möller, A Palmese, F Paz-Chinchón, A A Plazas, E Sanchez, V Scarpine, S Serrano, I Sevilla-Noarbe, M Smith, M Soares-Santos, E Suchyta, G Tarle, D Thomas, C To, and D L Tucker. OzDES reverberation mapping programme: the first mg scpii/scp lags from 5 yr of monitoring. *Monthly Notices of the Royal Astronomical Society*, 507(3):3771–3788, aug 2021. doi: 10.1093/mnras/stab2244. URL `https://doi.org/10.1093%2Fmnras%2Fstab2244`.

[45] Zhefu Yu, Paul Martini, A Penton, T M Davis, C S Kochanek, G F Lewis, C Lidman, U Malik, R Sharp, B E Tucker, M Aguena, J Annis, E Bertin, S Bocquet, D Brooks, A Carnero Rosell, D Carollo, M Carrasco Kind, J Carretero, M Costanzi, L N da Costa, M E S Pereira, J De Vicente, H T Diehl, P Doel, S Everett, I Ferrero, J Garcí a-Bellido, M Gatti, D W Gerdes, D Gruen, R A Gruendl, J Gschwend, G Gutierrez, S R Hinton, D L Hollowood, K Honscheid, D J James, K Kuehn, J Mena-Fernández, F Menanteau, R Miquel, B Nichol, F Paz-Chinchón, A Pieres, A A Plazas Malagón,

M Raveri, A K Romer, E Sanchez, V Scarpine, I Sevilla-Noarbe, M Smith, E Suchyta, M E C Swanson, G Tarle, M Vincenzi, A R Walker, and N Weaverdyck. OzDES reverberation mapping programme: Mg scpii/scp lags and ir/i-il/i relation. *Monthly Notices of the Royal Astronomical Society*, 522(3):4132–4147, apr 2023. doi: 10.1093/mnras/stad1224. URL `https://doi.org/10.1093%2Fmnras%2Fstad1224`.

[46] Fang Yuan, C. Lidman, T. M. Davis, M. Childress, F. B. Abdalla, M. Banerji, E. Buckley-Geer, A. Carnero Rosell, D. Carollo, F. J. Castander, C. B. D'Andrea, H. T. Diehl, C. E Cunha, R. J. Foley, J. Frieman, K. Glazebrook, J. Gschwend, S. Hinton, S. Jouvel, R. Kessler, A. G. Kim, A. L. King, K. Kuehn, S. Kuhlmann, G. F. Lewis, H. Lin, P. Martini, R. G. McMahon, J. Mould, R. C. Nichol, R. P. Norris, C. R. O'Neill, F. Ostrovski, A. Papadopoulos, D. Parkinson, S. Reed, A. K. Romer, P. J. Rooney, E. Rozo, E. S. Rykoff, M. Sako, R. Scalzo, B. P. Schmidt, D. Scolnic, N. Seymour, R. Sharp, F. Sobreira, M. Sullivan, R. C. Thomas, D. Tucker, S. A. Uddin, R. H. Wechsler, W. Wester, H. Wilcox, B. Zhang, T. Abbott, S. Allam, A. H. Bauer, A. Benoit-Lévy, E. Bertin, D. Brooks, D. L. Burke, M. Carrasco Kind, R. Covarrubias, M. Crocce, L. N. da Costa, D. L. DePoy, S. Desai, P. Doel, T. F. Eifler, A. E. Evrard, A. Fausti Neto, B. Flaugher, P. Fosalba, E. Gaztanaga, D. Gerdes, D. Gruen, R. A. Gruendl, K. Honscheid, D. James, N. Kuropatkin, O. Lahav, T. S. Li, M. A. G. Maia, M. Makler, J. Marshall, C. J. Miller, R. Miquel, R. Ogando, A. A. Plazas, A. Roodman, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, E. Suchyta, M. E. C. Swanson, G. Tarle, J. Thaler, and A. R. Walker. OzDES multifibre spectroscopy for the Dark Energy Survey: first-year operation and results. *Monthly Notices of the Royal Astronomical Society*, 452(3):3047–3063, 07 2015. ISSN 0035-8711. doi: 10.1093/mnras/stv1507. URL `https://doi.org/10.1093/mnras/stv1507`.

[47] Ying Zu, C. S. Kochanek, and Bradley M. Peterson. JAVELIN: Just Another Vehicle for Estimating Lags In Nuclei. Astrophysics Source Code Library, record ascl:1010.007, October 2010.

[48] Ying Zu, C. S. Kochanek, and Bradley M. Peterson. An Alternative Approach to Measuring Reverberation Lags in Active Galactic Nuclei. , 735(2):80, July 2011. doi: 10.1088/0004-637X/735/2/80.

[49] Ying Zu, C. S. Kochanek, Szymon Kozłowski, and Andrzej Udalski. IS QUASAR OPTICAL VARIABILITY a DAMPED RANDOM WALK? *The Astrophysical Journal*, 765(2):106, feb 2013. doi: 10.1088/0004-637x/765/2/106. URL `https://doi.org/10.1088%2F0004-637x%2F765%2F2%2F106`.

[50] Ying Zu, C. S. Kochanek, Szymon Kozłowski, and B. M. Peterson. APPLICATION OF STOCHASTIC MODELING TO ANALYSIS OF PHOTOMETRIC REVERBERATION MAPPING DATA. *The Astrophysical Journal*, 819(2):122, mar 2016. doi: 10.3847/0004-637x/819/2/122. URL `https://doi.org/10.3847%2F0004-637x%2F819%2F2%2F122`.