OzDES Reverberation Mapping Program: Mg II Lags and R–L relation

Zhefu Yu,1 Paul Martin,1,2,3 A. Penton,4 T. M. Davis,4,5 C. S. Kochanek,1,2 G. F. Lewis,6 C. Lidman,5,7 U. Malik,8 R. Sharp,8 B. E. Tucker,5,8 M. Aguena,9 J. Annis,10 E. Bertin,11,12 S. Bocquet,13 D. Brooks,14 A. Carnero Rosell,9,15,16 D. Carollo,5,17 M. Carrasco Kind,18,19 J. Carretero,20 M. Costanzi,17,21,22 L. N. da Costa,9 M. E. S. Pereira,23 J. De Vicente,24 H. T. Diehl,10 P. Doel,14 S. Everett,25 I. Ferrero,26 J. García-Bellido,27 M. Gatti,28 D. W. Gerdes,29,30 D. Gruen,13 R. A. Gruendl,18,19 J. Gschwend,9,31 G. Gutierrez,10 S. R. Hinton,4 D. L. Hollowood,32 K. Honscheid,2,33 D. J. James,34 K. Kuehn,7,35 J. Mena-Fernández,24 F. Menanteau,18,19 R. Miquel,20,36 B. Nichol,37 F. Paz-Chinchón,19,38 A. Pieres,9,31 A. A. Plazas Malagón,39 M. Raveri,28 A. K. Romer,40 E. Sanchez,24 V. Scarpine,10 I. Sevilla-Noarbe,24 M. Smith,41 E. Suchyta,42 M. E. C. Swanson,14,34 G. Tarle,30 M. Vincenzi,41,43 A. R. Walker,44 N. Weaverdyck,45
(Affiliations listed at the end of the paper)

Accepted XXX. Received YYY; in original form ZZZ

Abstract

The correlation between the broad line region radius and continuum luminosity (R–L relation) of active galactic nuclei (AGN) is critical for single-epoch mass estimates of supermassive black holes (SMBHs). At z ∼ 1–2, where AGN activity peaks, the R–L relation is constrained by the reverberation mapping (RM) lags of the Mg II line. We present 25 Mg II lags from the Australian Dark Energy Survey (OzDES) RM project based on six years of monitoring. We define quantitative criteria to select good lag measurements and verify their reliability with simulations based on both the damped random walk stochastic model and the re-scaled, re-sampled versions of the observed lightcurves of local, well-measured AGN. Our sample significantly increases the number of Mg II lags and extends the R–L relation to higher redshifts and luminosities. The relative iron line strength \( R_{Fe} \) has little impact on the R–L relation. The best-fit Mg II R–L relation has a slope \( \alpha = 0.39 \pm 0.08 \) with an intrinsic scatter \( \sigma_\Delta = 0.15^{+0.03}_{-0.02} \). The slope is consistent with previous measurements and shallower than the H\( \beta \) R–L relation. The intrinsic scatter of the new R–L relation is substantially smaller than previous studies and comparable to the intrinsic scatter of the H\( \beta \) R–L relation. Our new R–L relation will enable more precise single-epoch mass estimates and SMBH demographic studies at cosmic noon.

Key words: galaxies: nuclei – quasars: general

1 INTRODUCTION

Accurate mass measurements of supermassive black holes (SMBHs) are critical for understanding their growth over cosmic time. In the local universe, studies have used spatially resolved kinematics of stars or gas at the centres of galaxies to determine the SMBH mass (e.g., Kormendy & Richstone 1995; Gebhardt & Thomas 2009; Barth et al. 2016). However, it is difficult to extend this method to higher redshifts due to the angular resolution limits of current facilities.

The reverberation mapping (RM) technique is a robust method to measure the SMBH mass in active galactic nuclei (AGN) outside the local universe (Blandford & McKee 1982; Peterson 1993). The AGN broad emission lines vary in response to the stochastic variation of the continuum emission after a time delay due to the light travel time from the central accretion disk to the broad line region (BLR). The time lag \( \tau \) is correlated with the SMBH mass through the virial theorem

\[
M_{\text{BH}} = \frac{f c \tau \Delta v^2}{G}
\]  

(1)

where \( f \) is a “virial factor” determined by the dynamics and structure of the BLR. \( R = c \tau \) gives the characteristic BLR size and \( \Delta v \) is the velocity width of the broad lines. To measure the time lag, RM campaigns generally monitor the target for months to years to obtain the photometric and spectroscopic lightcurves, which requires substantial observational resources.

An important result of RM studies is a correlation between the BLR radius \( R \) and the AGN continuum luminosity \( L \). The existence of an \( R – L \) correlation enables a measurement of the SMBH mass from just a single-epoch spectrum. Such single-epoch estimates can be applied to large sample of AGN, making them critical for SMBH...
demographic studies (e.g., Vestergaard et al. 2008; Kelly & Shen 2013). The $R - L$ relation has been well-constrained in nearby AGN using time lag measurements of the Hβ line (e.g., Kaspi et al. 2000; Bentz et al. 2009, 2013; Grier et al. 2017). However, the Hβ line is not present in optical spectra of higher redshift AGN. At higher redshifts we need $R - L$ relations for other lines, such as Mg II and C IV. Despite previous RM studies of Mg II (e.g., Metzroth et al. 2006; Shen et al. 2016; Lira et al. 2018; Czerny et al. 2019; Homayouni et al. 2020; Zajaček et al. 2020; Yu et al. 2021) and C IV (e.g., Kaspi et al. 2007; Lira et al. 2018; Hoormann et al. 2019; Grier et al. 2019), the $R - L$ relations of these lines remain poorly constrained.

Star formation and AGN activity peak at redshifts of $z \sim 1 - 2$ (e.g., Wolf et al. 2003; Ueda et al. 2014). The single-epoch mass estimates at this epoch of cosmic noon are mainly based on the $R - L$ relation of the Mg II line because it is the major broad line observable in optical spectra. Unfortunately, there were less than ten Mg II lags available until recently (Metzroth et al. 2006; Shen et al. 2016; Lira et al. 2018; Czerny et al. 2019). Therefore, early studies of SMBH demographics generally calibrated the Mg II $R - L$ relation to match the Hβ lags instead of using the direct Mg II lags (e.g., Vestergaard & Osmer 2009). This could lead to a bias because the collisionally-excited Mg II line may respond differently to the continuum variability than the photoionised Balmer lines (e.g., Guo et al. 2020). It is therefore critical to better constrain the Mg II $R - L$ relation directly with a larger sample of lag measurements.

A promising way to quickly increase the number of lag measurements is through RM campaigns with wide-field multi-fibre spectrographs and imaging facilities, such as the Sloan Digital Sky Survey (SDSS) RM project (e.g., Shen et al. 2015) and the Dark Energy Survey (DES) - Australian DES (OzDES) RM project (King et al. 2015; Hoormann et al. 2019; Yu et al. 2021). Both surveys have monitored $\sim 800$ AGN at redshifts up to $z \sim 4.5$ for $\sim 5 - 6$ years. The large sample size and long time duration potentially allow these projects to produce a large number of lag measurements. However, these RM campaigns also face significant challenges, most notably due to the complexity of flux calibrating fibre spectra, the relatively low cadence and signal-to-noise ratio (SNR) of the spectroscopy, and the gaps between observing seasons. As a result, the probability distribution of the lag can be complicated and it is non-trivial to properly select reliable lag measurements and define lag uncertainties.

Homayouni et al. (2020) presented Mg II lags for 57 quasars from the SDSS RM project, from which they identified 24 quasars as their “gold sample” with the most reliable lags. They derived a Mg II $R - L$ relation with a slope of 0.31 and an intrinsic scatter of 0.36 dex from this gold sample. Their $R - L$ relation has a shallower slope and larger scatter than the Hβ relations that typically have a slope of $\sim 0.5$ and an intrinsic scatter of $\sim 0.13 - 0.2$ dex (e.g., Bentz et al. 2013; Du et al. 2016). Yu et al. (2021, hereafter Y21) presented nine Mg II lags from the first five years of data on about half of the OzDES RM sample. Their results show much less scatter than Homayouni et al. (2020), but the sample had a small dynamic range in luminosity and therefore could not independently constrain the slope of the $R - L$ relation.

Guo et al. (2020) used photoionization models to show that the response of the Mg II line to continuum variability was weaker than the Balmer lines, which could potentially explain the scatter of the Mg II $R - L$ relation. However, their model does not quantitatively predict the scatter of the Mg II $R - L$ relation. RM studies of the Hβ line found that AGN with larger Eddington ratios generally had smaller lags at fixed luminosities (e.g., Du et al. 2016, 2018; Dalla Bontà et al. 2020), which may also explain the shallower slope and larger intrinsic scatter of the $R - L$ relation when this effect is not included. Martínez-Aldama et al. (2020) reduced the scatter of the Mg II $R - L$ relation to $\sim 0.1$ dex by including the Eddington ratio as an additional parameter, but their Eddington ratio estimates depended on the lag measurements. Khadka et al. (2022a) used the ratio $R_{\beta\gamma}$ of the iron line flux to the Mg II line flux as an independent indicator of the Eddington ratio and found that it had no significant impact on the scatter of the Mg II $R - L$ relation. A larger Mg II lag sample with a wider redshift and luminosity range is critical for better constraining the Mg II $R - L$ relation and understanding its intrinsic scatter.

In this paper we present 25 Mg II lags from the full OzDES RM sample with six years of data and derive a new Mg II $R - L$ relation. Our sample is homogeneously defined through lag quality criteria that are verified by multiple simulations and statistical tests. We describe our observations and spectroscopic analysis in Section 2. Our time series analysis and lag measurements are discussed in Section 3. Section 4 discusses the reliability assessments of the lag measurements. We present the black hole mass and $R - L$ relation in Section 5. Section 6 summarizes the paper. This paper adopts a ΛCDM cosmology with $H_0 = 70$ km/s/Mpc, $\Omega_m = 0.3$, $\Omega_{\Lambda} = 0.7$.

## 2 OBSERVATIONS AND SPECTROSCOPIC ANALYSIS

DES is a 6-yr wide-area photometric survey that began in 2013 (Abbott et al. 2018). The survey took images in the griY bands using the Dark Energy Camera (DECam, Flaugher et al. 2015) with a 2.2° diameter field of view on the 4-m Victor M. Blanco telescope at the Cerro Tololo Inter-American Observatory. In addition to the 5000 deg$^2$ wide-area survey, DES observed 10 supernova (SN) fields approximately every week for the first five years and every three weeks for Y6.

OzDES is a spectroscopic follow-up survey in the DES SN fields that covers roughly the same time baseline as DES (e.g., Yuan et al. 2015; Childress et al. 2017; Lidman et al. 2020). The spectra cover the $\sim 3700 - 8900$ Å wavelength range and were taken using the AAOmega spectrograph (Smith et al. 2004) with the Two Degree Field (2dF) multi-fibre positioner (Gray et al. 1993) on the 4-m Anglo-Australian Telescope (AAAT). The OzDES RM project is one of the key OzDES science projects. Over the six DES observing seasons, it monitored 735 quasars in the DES SN fields with about monthly cadence. An observing season spans about six months from July to January. Figure 1 shows the apparent magnitude and redshift distribution of the OzDES RM quasars. The sample spans an AB magnitude range of $g \sim 17 - 23$ mag and a redshift range of $z \sim 0.1 - 4.5$.

We use the pipeline from Hoormann et al. (2019) to calibrate the spectra. The pipeline first calculates the scaling factors from the instrumental flux derived by integrating the extracted and wavelength calibrated spectra within the DES filters to the DES photometric flux in the $grizY$ bands. The DES flux is derived from the linear interpolation of the two DES photometric epochs bracketing the spectroscopic epoch. It then fits the scaling factors with a second-order polynomial and generates flux calibrated spectra by multiplying the polynomial to the extracted and wavelength calibrated spectra.

The OzDES RM project has developed simulation frameworks for survey design and lag quality assessment (King et al. 2015; Penton et al. 2022; Malik et al. 2022) and published early measurements of continuum lags (Mudd et al. 2018; Yu et al. 2020a), Mg II lags
A challenge in analyzing the Mg II line is the strong contribution from nearby iron emission lines (e.g., Wills et al. 1980, 1985; Verner et al. 1999). The iron lines could also have reverberation signals that contaminate the Mg II lag signal (e.g., Barth et al. 2013). We fit and subtract the continuum + iron emission and derive the Mg II line flux using the pipeline from Y21. Our model

\[ f_{\text{model}}(\lambda) = f_c(\lambda) + f_{\text{Fe}}(\lambda) \]

consists of a power-law continuum

\[ f_c(\lambda) = A_c(\lambda/\lambda_0)^\alpha \]

with an index \( \alpha \) and a normalisation \( A_c \) at \( \lambda_0 = 2599 \, \text{Å} \), and an iron emission line component

\[ f_{\text{Fe}}(\lambda) = A_{\text{Fe}} f_1(\lambda) * G(w) \]

modeled by an iron template \( f_1(\lambda) \) of a normalisation \( A_1 \) convolved with a Gaussian kernel \( G(w) \) of a width \( w \) to account for the velocity broadening of the BLR. The four free parameters of the model are \( A_c \), \( \alpha \), \( A_1 \), and \( w \).

We adopt the empirical iron template from Vestergaard & Wilkes (2001) based on the Seyfert galaxy I Zwicky 1. While Vestergaard & Wilkes (2001) template did not model the iron emission under the Mg II line, Y21 tested other iron templates by Tsuzuki et al. (2006) and Salvianader et al. (2007) which modeled the iron emission under Mg II and found that the choice of iron template had little impact on the lag measurements. We fit the spectra over the rest-frame wavelength ranges 2260 - 2690 Å and 2910 - 3050 Å using a Markov chain Monte-Carlo (MCMC) sampler. We do not include the wavelength range near the Mg II line, since it is difficult to distinguish the iron emission from the strong Mg II line within this range. When fitting the single-epoch spectra, we cannot constrain the broadening width \( w \) very well due to the low SNR, so we fix it to the best-fit value found for the co-added spectra. While the width of the iron lines could vary in response to the continuum variability (e.g., Korista & Goad 2004; Guo et al. 2020; Wang et al. 2020), Y21 found such variability had little impact on the derived Mg II line flux.

We derive the Mg II flux as \( F_{\text{line}} = F_{\text{total}} - F_{\text{model}} \), where \( F_{\text{total}} \) is the integrated spectral flux over the rest-frame wavelength range 2700 - 2900 Å before continuum and iron subtraction and \( F_{\text{model}} \) is the integrated continuum + iron model flux. The line flux uncertainty is estimated as

\[ \sigma_{\text{line}}^2 = \sigma_{\text{total}}^2 + \sigma_{\text{model}}^2 \]

where

\[ \sigma_{\text{total}} = \sqrt{\sum_i \sigma_{\text{total},i}^2 \Delta \lambda} \]

is the uncertainty in the total flux \( F_{\text{total}} \). \( \sigma_{\text{total},i} \) is the uncertainty of the \( i \)-th pixel, \( \Delta \lambda \) is the wavelength pixel size and \( \sigma_{\text{model}} \) is the uncertainty in the model flux \( F_{\text{model}} \). We estimate \( \sigma_{\text{model}} \) as the scatter of the model fluxes in the MCMC chain. Figure 2 shows an example of the continuum + iron modeling. The model generally matches the observed spectra well. We visually inspect all spectra and exclude epochs from further analysis where the best-fit model fails to match the spectra or where the spectra are contaminated by instrumental artifacts, such as the bump at \( \sim 7100 \, \text{Å} \) created by an LED in the 2dF gripper gantry. Figures for all spectra of our final sample of quasars are available in the supplementary material.

1 The constant \( \lambda_0 \) is chosen to reduce the error correlation between the power-law index \( \alpha \) and normalisation \( A_c \) by reducing the non-diagonal terms of the Fisher matrix.

---

**Figure 1.** *(Upper panel)* g-band apparent magnitude versus redshift of the OzDES RM quasars (grey points) and the sub-sample that has successful Mg II lag measurements (blue squares). The blue dashed lines show the redshift range of quasars that we analyzed in this paper. *(Lower panel)* Monochromatic luminosity at 3000 Å versus redshift. The blue squares, brown points, orange diamonds, pink pentagons and green hexagons represent the OzDES Mg II lag sample (this work), the OzDES C IV lag sample (Hoormann et al. 2019), the gold sample of the SDSS Mg II lags (Homayouni et al. 2020), other sources with Mg II lags from literature (Metzroth et al. 2006; Lira et al. 2018; Czerny et al. 2019; Zajacék et al. 2020, 2021) and other RM sources from literature (Peterson et al. 2004, 2005; Kaspi et al. 2007; Bentz et al. 2009; Denney et al. 2010; Barth et al. 2011b; Grier et al. 2011; Bentz et al. 2013; Barth et al. 2013; Bentz et al. 2014; Grier et al. 2014; Li et al. 2014; Trevese et al. 2014; Wang et al. 2014; Du et al. 2015, 2016; Bentz et al. 2016; Lu et al. 2016; Faussaugh et al. 2017; Grier et al. 2017; Du et al. 2018; Lira et al. 2018; Grier et al. 2019; Zhang et al. 2019; Hu et al. 2021; Li et al. 2021; Williams et al. 2021; U et al. 2022). The number of lags in each sample is given in the brackets. We converted the monochromatic luminosities at 5100 Å and 1350 Å to 3000 Å using the bolometric calibration could introduce correlated errors between the spectral pixels over a wide wavelength range. We estimate the calibration uncertainties based on the F-stars monitored by OzDES (Section 2.2).
2.2 Calibration Uncertainty

We estimate the calibration uncertainty and propagate it to the error budget of the Mg II line flux using the pipeline from Y21 based on the F-stars observed by OzDES in each epoch. For each F-star, we calculate the mean \( f_{s,j} \) and rms \( S_{s,j} \) of the calibrated spectra over all epochs, where \( i \) corresponds to the \( i \)-th pixel. Since F-stars do not have intrinsic variability, the residual variation \( S_{s,j} \) is due to a combination of the mostly uncorrelated photon noise and the correlated calibration uncertainty. We cannot directly derive the calibration uncertainty from \( S_{s,j} \) since it is dominated by the photon noise at the pixel level. We therefore bin the spectra by 300 \( \text{Å} \) to suppress the photon noise, while the correlated calibration uncertainty remains. The residual variation of the binned spectra is an estimate of the calibration uncertainty.

We propagate the calibration uncertainty to the line flux uncertainty using a Monte-Carlo method. We define a “warping function”

\[
W_{bj} = f_{s,bj}/\langle f_{s,b} \rangle
\]

where \( f_{s,bj} \) is the flux of the binned F-star spectra in the \( b^{\text{th}} \) bin of the \( j^{\text{th}} \) epoch and \( \langle f_{s,b} \rangle \) is the mean of the binned spectra over all epochs. We then derive a continuous warping function \( W_{f,j}(\lambda) \) by interpolating the discrete function \( W_{bj} \) with a 3rd order spline function. Each epoch of the F-star spectra provides a warping function \( W_{f,j}(\lambda) \), and each warping function is a realization of the spectral variability due to the calibration uncertainty. We create 2343 warping functions based on 161 F-stars. This provides a “library” of warping functions for each spectroscopic epoch. Multiplying the quasar spectrum with the warping functions of the corresponding epoch gives realizations of the quasar spectra warped by the calibration uncertainty. In each warped spectra, we calculate the Mg II flux using the method described in Section 2.1. We then multiply the fractional variation of the warped Mg II fluxes by the Mg II flux of the observed spectra to derive the line flux uncertainty contributed by the calibration procedure. This uncertainty is added in quadrature to the line flux uncertainty estimated in Section 2.1. The median calibration uncertainty is \( \sim 4\% \).

The calibration pipeline uses the linear interpolation of the two neighboring DES epochs to estimate the continuum flux at a spectroscopic epoch, which is a reasonable estimate since the cadence of the DES photometric lightcurve is much higher than the OzDES spectroscopy. While the photometric cadence in Y6 is lower than the first five years, the Y6 spectra are generally taken within a few days of the neighboring photometric epochs. The only cases where the linear interpolation can significantly increase the calibration uncertainty are two spectroscopic epochs taken at 2018-06-15 (MJD 58284) and 2018-06-24 (MJD 58293) for some quasars. They are in the seasonal gap between DES Y5 and Y6 and therefore have no nearby photometric epoch. This introduces additional calibration error due to the large uncertainty of the continuum flux. In these two cases we interpolate the photometric lightcurve with a damped random walk (DRW) stochastic process and calculate the fractional uncertainty of the DRW model at the time of the spectroscopic epoch. This characterizes the uncertainty in the overall normalisation of the calibration and we add it as an additional uncertainty to the error budget of the Mg II line flux. The additional uncertainty is \( \sim 10\% \) for these two cases.

3 TIME SERIES ANALYSIS

We use the \( g \)-band photometry from the DES database for the continuum lightcurves and the pipeline described above to create Mg II line lightcurves. Machine-readable lightcurves for our final sample are available in the supplementary material. We use JAVELIN (e.g., Zu et al. 2011, 2013) and the interpolated cross-correlation function (ICCF, e.g., Gaskell & Peterson 1987; Peterson et al. 1998, 2004) method to measure the time lags. JAVELIN uses a DRW stochastic process to interpolate the lightcurve and assumes that the line lightcurve is the continuum lightcurve convolved with a top-hat transfer function. JAVELIN fits the continuum and line lightcurves simultaneously with a MCMC sampler to derive the posterior probability distributions of the amplitude \( \sigma_{\text{drw}} \) and characteristic time scale \( \tau_{\text{drw}} \) of the DRW model, the scale \( \lambda_1 \) and width \( \lambda_2 \) of the transfer function, and the mean time lag \( \tau \). The algorithm sets the prior of the DRW parameters \( \sigma_{\text{drw}} \) and \( \tau_{\text{drw}} \) as their posterior distributions from fitting only the continuum lightcurve. We use a flat prior for the time lag \( \tau \) within \([\tau_{\text{min}}, \tau_{\text{max}}]\) days, while we allow the other parameters to vary freely.

We use PyCCF (Sun et al. 2018), a python implementation of the ICCF method. This method linearly interpolates the lightcurves and calculates the cross-correlation function (CCF) within a lag range \([\tau_{\text{min}}, \tau_{\text{max}}]\) days. It then uses the range of the CCF where the cross-correlation coefficient is at least 80% of the maximum and estimates the lag as its centroid or peak. For the lag uncertainty, it first creates lightcurve realizations by randomizing the single-epoch flux by its uncertainty and randomly sub-sampling the epochs with replacement. It then calculates the centroid and peak of the CCF for each realization to create the cross-correlation centre distribution (CCCD) and the cross-correlation peak distribution (CCPD). The scatter of these distributions gives the lag uncertainty estimate. We generate 8000 realizations and adopt the realizations with \( r_{\text{peak}} \) > 0.5 to create the CCCD and CCPD, where \( r_{\text{peak}} \) is the peak value of the CCF. We use CCCD as the fiducial lag probability distribution
of the ICCF method, since previous work found it provided better lag estimates than CCPD (e.g., Peterson et al. 2004).

The lag distributions from JAVELIN and ICCF usually have multiple peaks that are mostly due to the aliasing effects caused by the seasonal gaps, so it is non-trivial to identify successful lag measurements and estimate their uncertainties. We use two different methods to analyze the lag distributions, leading to two sets of criteria to select successful lag measurements.

3.1 Symmetric Prior

Method 1 resembles the lag analysis in Y21 and the SDSS RM project (e.g., Grier et al. 2019; Homayouni et al. 2020). We set a lag prior range \([t_{\text{min}} = -1000 \text{ days}, t_{\text{max}} = 1000 \text{ days}]\) for both JAVELIN and ICCF and define a weighting function to suppress aliasing. The weighting function is the convolution of two components. The first component is defined as

\[
P(\tau) = \left[ N(\tau)/N(0) \right]^2
\]

where \(N(\tau)\) is the number of overlapping points between the continuum lightcurve shifted by the time lag \(\tau\) and the line lightcurve. This penalizes lags in the seasonal gaps where the shifted continuum has little overlap with the line lightcurve. The second component is the auto-correlation function (ACF) of the continuum lightcurve, which characterizes how fast the continuum varies. We set ACF = 0 when it is below zero.

We multiply the lag distribution by the weighting function and convolve the weighted distribution with a Gaussian kernel that has a width of 12 days, the same width used in Y21 and Homayouni et al. (2020). We define the major peak as the highest peak in the weighted and smoothed lag distribution. When there are multiple connected peaks, we define one as a separate peak if its prominence exceeds 10% of the prominence of the neighboring peaks and it is separated from the neighboring peaks by at least 10 days. We use the package scipy.signal.peak_prominences (Jones et al. 2001) to calculate the peak prominence defined as the vertical distance between the major peak and the higher minimum at its two sides. We then define the lag as the median of the unweighted lag distribution within the major peak and define the lag uncertainty based on its 16th and 84th percentiles.

We define the first set of selection criteria based on Method 1:

(a) \(f_{\text{peak}} > 0.6\) for the JAVELIN lag distribution, where \(f_{\text{peak}}\) is the probability within the major peak;
(b) The lag uncertainty from JAVELIN is less than 110 days; and
(c) The lags from JAVELIN and ICCF agree within 2\(\sigma\).

The first criterion ensures that the major peak has enough significance to be distinguished from the aliasing peaks. The second criterion requires that the lag distribution has enough constraining power and excludes lag distributions that are flat over a wide range of lags. The agreement between JAVELIN and ICCF required by the third criterion helps to increase the lag reliability. We obtain 62 lags that pass the Method 1 criteria, 48 of which are positive.

3.2 Positive Prior

Method 2 is defined by Penton et al. (2022) and is based on simulated lightcurves that have the same cadence and SNR as the OzDES observations. We use a lag prior range \([t_{\text{min}} = 0 \text{ day}, t_{\text{max}} = 1000 \text{ days}]\) to derive the lag distribution from JAVELIN and ICCF. We analyze the full lag distribution binned by 3 days without any weighting or major peak identification. We define the second set of selection criteria based on Method 2:

(a) The mean absolute deviation of the JAVELIN lag distribution is less than 110 days;
(b) The separation between the median and peak of the JAVELIN lag distribution is less than 110 days; and
(c) The separation between the peaks of the JAVELIN and ICCF lag distribution is less than 100 days.

The threshold values in each criterion are the same as Penton et al. (2022). The first and second criteria effectively require the lag distribution to have a strong major peak compared to the aliasing peaks, since the aliasing peaks would increase the mean absolute deviation and result in a difference between the median and peak. There are 56 quasars that pass the Method 2 criteria.

3.3 Final Measurements

We define a successful lag measurement as one which passes both the Method 1 and 2 criteria. This yields 25 lag measurements which we refer to as our final sample. We adopt the lag and uncertainty estimates in Section 3.1 based on the JAVELIN lag distributions, since previous studies showed that JAVELIN provides better lag and uncertainty estimates than the ICCF method (e.g., Li et al. 2019; Yu et al. 2020b). Table 1 gives the lag measurements of the final sample. The final sample spans a redshift range of \(0.84 \leq z \leq 1.86\) and an observed lag range of \(130 \text{ days} \leq \tau \leq 880 \text{ days}\).

Figure 3 shows examples of the lightcurves and lag distributions of our final sample. The lag distributions exhibit clear major peaks. The major peaks are significantly stronger than the secondary peaks for the JAVELIN lag distributions. The median JAVELIN major peak fraction \(f_{\text{peak}}\) of the final sample is 0.78. The lags from the symmetric and positive prior ranges all agree within \(1\sigma\). The JAVELIN and ICCF lags are all consistent within 100 days and within \(2\sigma\), as required by the selection criteria.

The lightcurves in Figure 3 exhibit clear variability features. The black points in the lower left panels show the continuum lightcurve shifted by the best-fit JAVELIN lag with the symmetric prior range. For the sources shown in the top two rows, the shifted continuum lightcurve matches the line lightcurve well, which supports the reliability of the lag measurement. The lag of the third source is close to the 1.5-yr seasonal gap, but its shifted continuum lightcurve is a reasonable “interpolation” between the line lightcurves in different seasons.

A special feature of DES J003206.50−425325.22 (bottom panel of Figure 3) is the two outliers at MJST 58073 and 58399. The excess flux is due to a significant broadening of the Mg II line. This can be intrinsic to the Mg II line or due to a drastic change in the iron emission that is not characterized by the model. The JAVELIN lag distribution would be dominated by a sharp aliasing peak at \(\sim 540 \text{ days}\) if we included these two epochs, which differs from the \(\sim 470\)-day ICCF lag. For lags of a few hundred days, we expect the signal to come from lightcurve features over several years rather than dramatic short time scale variability, so we exclude these two epochs for this particular source. Excluding these two epochs also gives a cleaner JAVELIN lag distribution compared to the Y5 results where we kept the MJST 58073 epoch, and the best-fit lag changes by \(\sim 40\) days relative to Y5.

Seven of the nine quasars from Y21 pass our final sample criteria in this paper after adding the Y6 data. The black dashed lines in the top and bottom panels of Figure 3 show the lag distributions presented in Y21 based on the first 5-yr lightcurves. The new and old lags agree within \(1\sigma\) except for the case of DES J003206.50−425325.22 discussed above. Adding the Y6 data...
Figure 3. Lightcurves and lag measurements. Each of the three main rows is for an AGN with its name given in the upper left corner. (Left column) The upper and lower panels show the $g$-band lightcurve and the Mg II line lightcurve, respectively. The black points in the lower panel show the $g$-band lightcurve shifted by the best-fit lag. (Middle column) The upper and lower panels show the lag distributions from JAVELIN and ICCF with a symmetric lag prior. The blue solid and red dashed lines represent the unweighted and weighted lag distributions, respectively. For sources presented in Y21, the black dash-dotted lines show the lag distributions from the 5-yr lightcurves. The yellow shaded area marks the major peak region. (Right column) Lag distributions with a positive lag prior. The green shaded area represents the $1\sigma$ range centered at the peak, where $\sigma$ is defined as the mean absolute deviation of the lag distribution.
suppresses some aliasing peaks and makes the lag signal cleaner. Two quasars from Y21 (DES J021612.83–044634.10 and DES J03553.51–27504.70) fail to pass the lag selection. While both pass the Method 1 criteria, they are excluded by the Method 2 criteria due to large mean absolute deviations and differences between the JAVELIN and ICCF lags caused by aliasing peaks.

4 LAG RELIABILITY

We assess the reliability of lag measurements using simulations based on both the DRW model and the observed lightcurves of intensively monitored local AGN. We then discuss the lag reliability of our final sample and compare with the samples selected by only one of the criteria.

4.1 Simulations with DRW

We adopt the simulation tool developed by Penton et al. (2022) to create simulated lightcurves following the same procedure as Y21. For each of the 25 final sample sources that pass both sets of criteria for Method 1 and 2, we create 1000 realizations of DRW lightcurves that have the same variability as the observed continuum lightcurve. We use the DRW lightcurve as the simulated continuum and convolve it with a top-hat transfer function to create the simulated line lightcurves. The input lag \( \tau_{\text{in}} \) of the transfer function is randomly drawn from a uniform distribution between 10 days and 1000 days. We then re-sample and add noise to the simulated lightcurves so that their cadence and SNR match the observed lightcurves.

We then use the method described in Section 3 to measure lags from the simulated lightcurve pairs. We define a lag as a “false positive” if the measured lag \( \tau_{\text{m}} \) differs from the input lag \( \tau_{\text{in}} \) by 3\( \sigma \). For a final sample source with an observed lag \( \tau_{\text{obs}} \), we define the false positive rate (FPR) as

\[
FPR = \frac{N_{\text{bad, m}}(\tau_{\text{obs}})}{N_{\text{p, m}}(\tau_{\text{obs}})}
\]

where \( N_{\text{p, m}}(\tau_{\text{obs}}) \) is the number of realizations where the measured lag \( \tau_{\text{m}} \) passes the final sample selection and is within 1\( \sigma \) of \( \tau_{\text{obs}} \), and \( N_{\text{bad, m}}(\tau_{\text{obs}}) \) is the number of false positives among the \( N_{\text{p, m}}(\tau_{\text{obs}}) \) realizations. The FPR is an estimate of the probability that the observed lag \( \tau_{\text{obs}} \) is different from the true lag. Table 1 gives the FPR of each AGN in the final sample. The filled histograms in Figure 4 show the FPR distribution of the final sample. The median FPR is \( \sim 3\% \). The sum of the FPR for all 25 AGN is \( \sim 1.3 \), which

### Table 1. Characteristics of the 25 AGN in our final sample with successful Mg II lag measurements. Columns (1) and (2) give the DES name and redshift of the source. Columns (3) and (4) give the lags and uncertainties in the observer frame from JAVELIN and the ICCF method, respectively. Column (5) gives the probability within the major peak of the JAVELIN lag distribution (see Section 3.1). Column (6) gives the false positive rate from simulations (see Section 4.1). Column (7) gives the 3000 Å monochromatic luminosity and its uncertainty. Column (8) gives the ratio \( R_{\text{Fe}} \) of the ion flux to the Mg II flux (see Section 5.1). Columns (9) and (10) give the line dispersion and black hole mass. The black hole mass uncertainty is about 0.4 dex.

| Source Name           | DES J024340.09+001749.40 | DES J025254.18+001119.70 | DES J024831.08+005025.60 | DES J024723.54+002536.50 | DES J024494.09+003317.50 | DES J024455.45+011500.40 | DES J025225.52+003405.90 | DES J022716.52–050008.30 | DES J022751.50–044252.70 | DES J02208.15–065550.50 | DES J033836.19–295113.50 | DES J033903.66–293326.50 | DES J033328.93–275641.21 | DES J022436.64–063255.90 | DES J032311.42–284323.99 | DES J032136.34–283620.99 | DES J033710.86–444138.89 | DES J03922.97–430230.41 | DES J020923.85–435240.69 | DES J03207.44–433049.00 | DES J030015.00–430333.52 | DES J030052.76–430301.08 | DES J032326.61–433302.99 | DES J03234.33–431937.81 | DES J03206.50–425325.22 |
|-----------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| z                     | 1.44                     | 1.64                     | 0.89                     | 1.86                     | 1.48                     | 1.53                     | 1.62                     | 1.64                     | 1.79                     | 1.66                     | 1.15                     | 1.68                     | 0.84                     | 1.42                     | 1.24                     | 1.49                     | 1.07                     | 1.37                     | 1.00                     | 1.53                     | 1.65                     | 1.43                     | 1.49                     | 1.64                     |
| JAVELIN Lag (days)    | 818*39                   | 415*27                   | 286*56                   | 867*64                   | 412*31                   | 165*31                   | 520*26                   | 524*14                   | 538*28                   | 439*25                   | 225*46                   | 308*40                   | 175*16                   | 181*26                   | 132*21                   | 171*11                   | 382*30                   | 564*21                   | 571*13                   | 576*10                   | 508*30                   | 537*11                   | 656*33                   |
| ICCF Lag (days)       | 797*34                   | 420*24                   | 341*03                   | 880*44                   | 402*66                   | 210*39                   | 505*43                   | 495*32                   | 558*33                   | 423*37                   | 280*49                   | 284*36                   | 182*35                   | 147*44                   | 137*49                   | 192*31                   | 430*29                   | 637*17                   | 540*32                   | 362*13                   | 467*28                   | 388*35                   | 535*43                   |
| \( f_{\text{peak}} \) | 0.76                     | 0.75                     | 0.91                     | 0.84                     | 0.88                     | 0.77                     | 0.69                     | 0.68                     | 0.70                     | 0.96                     | 0.96                     | 0.67                     | 0.82                     | 0.83                     | 1.00                     | 0.85                     | 0.93                     | 0.76                     | 0.73                     | 0.65                     | 0.73                     | 0.78                     | 0.98                     |
| FPR (%)               | 7.9                      | 3.7                      | 0.9                      | 0.8                      | 1.0                      | 8.0                      | 0.6                      | 2.1                      | 0.0                      | 2.8                      | 5.7                      | 10.0                     | 2.3                      | 0.0                      | 1.4                    | 3.6                      | 1.9                      | 1.5                      | 1.9                      | 2.5                      | 1.3                      | 0.7                      |
| log(\( \Delta L/\text{Sun} \) erg/s) | 45.41 ± 0.02 | 45.12 ± 0.01 | 44.89 ± 0.02 | 44.89 ± 0.01 | 45.50 ± 0.01 | 45.22 ± 0.02 | 45.59 ± 0.01 | 44.44 ± 0.01 | 44.72 ± 0.02 | 44.31 ± 0.01 | 44.47 ± 0.02 | 44.01 ± 0.01 | 44.70 ± 0.04 | 44.88 ± 0.02 | 44.28 ± 0.04 | 44.37 ± 0.01 | 44.67 ± 0.01 | 44.35 ± 0.01 | 44.35 ± 0.01 | 45.86 ± 0.00 | 45.59 ± 0.00 | 45.54 ± 0.01 | 45.37 ± 0.00 | 45.95 ± 0.01 | 45.57 ± 0.01 |
variability could deviate from the DRW model on timescales below ~ month (e.g., Mushotzky et al. 2011; Kasliwal et al. 2015; Smith et al. 2018). AGN could also have “BLR holidays” when the broad line variability is not correlated with the continuum variability (e.g., Goad et al. 2016; Horne et al. 2021), and this could hamper the lag recovery (Yu et al. 2020b).

We perform additional simulations using the observed lightcurves of several intensively monitored local AGN as “reference lightcurves”. We first scale the time axis of the reference lightcurve by a factor of \( X_t \) so that its baseline length matches the 6-yr baseline of the OzDES continuum and line lightcurves. While scaling the baseline increases the variability time scale, Stone et al. (2022) found that the DRW time scale \( \tau_{\text{drw}} \) of AGN could span ~ 1.5 orders of magnitude, so the scaling does not make the variability time scale unrealistic for AGN. We then shift the scaled reference lightcurve of the emission line by \( \Delta t = \tau_i - X_t \tau_{\text{ref}} \) so that the lag between the shifted, scaled reference lightcurves of the continuum and emission line equal the desired input lag \( \tau_i \), where \( \tau_{\text{ref}} \) is the lag between original reference lightcurves from literature. The second row of Figure 5 shows an example of the scaled reference lightcurves.

We interpolate the scaled reference continuum and line lightcurves using the predicted lightcurves from JAVELIN (red solid line in Figure 5) and re-sample them to the same cadence as the OzDES lightcurves (top row of Figure 5) to create the simulated lightcurves. We then assign an uncertainty \( \sigma_{\text{sim},j} = K\sigma_{\text{oz},j} \) to each simulated epoch, where \( j \) corresponds to the \( j^{\text{th}} \) epoch, \( K \) is a constant coefficient and \( \sigma_{\text{oz},j} \) is the uncertainty of the corresponding OzDES epoch. The constant \( K \) is derived such that the variability

\[
\chi^2 = \sum_j \frac{(f_j - \langle f \rangle_j)^2}{\sigma_j^2}
\]

of the simulated lightcurve equals that of the OzDES lightcurve, where \( f_j \) and \( \sigma_j \) are the flux and uncertainty of the \( j^{\text{th}} \) epoch and \( \langle f \rangle \) is the mean flux. If the assigned uncertainty \( \sigma_{\text{sim},j} \) is larger than the original uncertainty \( \sigma_{\text{oz},j} \) of the reference lightcurve, we add additional Gaussian noise with a variance of \( \sigma_{\text{oz},j}^2 - \sigma_{\text{sim},j}^2 \) to the simulated epoch. While matching the overall variability of the lightcurves does not guarantee the short time scale variability is identical, we expect the major lag signal to be from lightcurve features over multiple years, so this approximation would not significantly affect our simulation results. The third row of Figure 5 shows an example of a final simulated lightcurve.

Each reference AGN effectively gives a realization of the simulated lightcurve. We use 12 reference AGN summarized in Table 2 from Grier et al. (2012), Du et al. (2014), Wang et al. (2014), De Rosa et al. (2015) and De Rosa et al. (2018) where the cadences of the scaled reference lightcurves are higher than 15 days and 20 days for the continuum and emission line, respectively. All but Mrk 382 have velocity resolved RM results (Grier et al. 2013a; De Rosa et al. 2018; Li et al. 2018; Horne et al. 2021). NGC 5548 has lightcurves for two different seasons, which we use as separate reference lightcurves.

### 4.2 Simulations with Reference Lightcurves

The simulations in Section 4.1 assume that the AGN variability is a DRW stochastic process and that the line lightcurve is related to the continuum lightcurve by a top-hat transfer function. These assumptions may not hold in real AGN. For example, the AGN variability could deviate from the DRW model on timescales below ~ month (e.g., Mushotzky et al. 2011; Kasliwal et al. 2015; Smith et al. 2018). AGN could also have “BLR holidays” when the broad line variability is not correlated with the continuum variability (e.g., Goad et al. 2016; Horne et al. 2021), and this could hamper the lag recovery (Yu et al. 2020b).

We perform additional simulations using the observed lightcurves of several intensively monitored local AGN as “reference lightcurves”. We first scale the time axis of the reference lightcurve by a factor of \( X_t \) so that its baseline length matches the 6-yr baseline of the OzDES continuum and line lightcurves. While scaling the baseline increases the variability time scale, Stone et al. (2022) found that the DRW time scale \( \tau_{\text{drw}} \) of AGN could span ~ 1.5 orders of magnitude, so the scaling does not make the variability time scale unrealistic for AGN. We then shift the scaled reference lightcurve of the emission line by \( \Delta t = \tau_i - X_t \tau_{\text{ref}} \) so that the lag between the shifted, scaled reference lightcurves of the continuum and emission line equal the desired input lag \( \tau_i \), where \( \tau_{\text{ref}} \) is the lag between original reference lightcurves from literature. The second row of Figure 5 shows an example of the scaled reference lightcurves.

We interpolate the scaled reference continuum and line lightcurves using the predicted lightcurves from JAVELIN (red solid line in Figure 5) and re-sample them to the same cadence as the OzDES lightcurves (top row of Figure 5) to create the simulated lightcurves. We then assign an uncertainty \( \sigma_{\text{sim},j} = K\sigma_{\text{oz},j} \) to each simulated epoch, where \( j \) corresponds to the \( j^{\text{th}} \) epoch, \( K \) is a constant coefficient and \( \sigma_{\text{oz},j} \) is the uncertainty of the corresponding OzDES epoch. The constant \( K \) is derived such that the variability

\[
\chi^2 = \sum_j \frac{(f_j - \langle f \rangle_j)^2}{\sigma_j^2}
\]

of the simulated lightcurve equals that of the OzDES lightcurve, where \( f_j \) and \( \sigma_j \) are the flux and uncertainty of the \( j^{\text{th}} \) epoch and \( \langle f \rangle \) is the mean flux. If the assigned uncertainty \( \sigma_{\text{sim},j} \) is larger than the original uncertainty \( \sigma_{\text{oz},j} \) of the reference lightcurve, we add additional Gaussian noise with a variance of \( \sigma_{\text{oz},j}^2 - \sigma_{\text{sim},j}^2 \) to the simulated epoch. While matching the overall variability of the lightcurves does not guarantee the short time scale variability is identical, we expect the major lag signal to be from lightcurve features over multiple years, so this approximation would not significantly affect our simulation results. The third row of Figure 5 shows an example of a final simulated lightcurve.

Each reference AGN effectively gives a realization of the simulated lightcurve. We use 12 reference AGN summarized in Table 2 from Grier et al. (2012), Du et al. (2014), Wang et al. (2014), De Rosa et al. (2015) and De Rosa et al. (2018) where the cadences of the scaled reference lightcurves are higher than 15 days and 20 days for the continuum and emission line, respectively. All but Mrk 382 have velocity resolved RM results (Grier et al. 2013a; De Rosa et al. 2018; Li et al. 2018; Horne et al. 2021). NGC 5548 has lightcurves for two different seasons, which we use as separate reference lightcurves.

### Table 2. The AGN used for the simulations in Section 4.2. Column (1) gives the object name. These objects generally have lightcurves for multiple continuum bands and emission lines. Columns (2) and (3) give the continuum band and emission line we adopt for the reference lightcurves. Column (4) gives the source of the reference lightcurves. NGC 5548 has lightcurves from different seasons which we use as separate reference lightcurves.

| Name   | Continuum Band | Line   | Source                  |
|--------|----------------|--------|-------------------------|
| NGC 5548 | 1367 Å        | Ly α   | De Rosa et al. (2015)   |
| Mrk 335 | V band        | Hβ     | Grier et al. (2012)     |
| Mrk 1501 | V band        | Hβ     | Grier et al. (2012)     |
| 3C 120  | V band        | Hβ     | Grier et al. (2012)     |
| Mrk 6   | V band        | Hβ     | Grier et al. (2012)     |
| PG 2130+099 | V band    | Hβ     | Grier et al. (2012)     |
| Mrk 704 | 5100 Å        | Hβ     | De Rosa et al. (2018)   |
| NGC 3227 | 5100 Å        | Hβ     | De Rosa et al. (2018)   |
| NGC 3516 | 5100 Å        | Hβ     | De Rosa et al. (2018)   |
| NGC 4151 | 5100 Å        | Hβ     | De Rosa et al. (2018)   |
| NGC 5548 | 5100 Å        | Hβ     | De Rosa et al. (2018)   |
| Mrk 142 | 5100 Å        | Hβ     | Du et al. (2014)        |
| Mrk 382 | V band        | Hβ     | Wang et al. (2014)      |

indicates that there could be one incorrect lag in the sample. The median and average of the FPR are similar to the Y21 sample where the median was ~ 4% and the sum was 0.44 for 9 AGN.
Figure 5. Examples of the simulations based on the observed lightcurves of local AGN. (Upper 3×2 panels) Procedure for creating simulated lightcurves based on the reference lightcurves. The lightcurves are in arbitrary units and the absolute flux scales have no effect on our simulations. The left and right columns show the continuum and line lightcurves, respectively. The top row shows the OzDES continuum and line lightcurves of DES J003207.44−433049.00. The middle row shows the scaled reference lightcurves based on 3C 120. The input lag between the scaled reference lightcurves is 370.1 days, which equals the observed lag between the OzDES lightcurves of DES J003207.44−433049.00. The red lines represent the best-fit lightcurves from JAVELIN. The bottom row shows the simulated lightcurves, which have the same cadence and variability amplitude as the OzDES lightcurves in the top row. (Lower 2×2 panels) Lag distributions from the simulated lightcurves. The left and right columns show results from the symmetric and positive lag priors, respectively. The black solid line is drawn at the input lag. The black dashed lines show the 1σ uncertainty of the input lag, which is the uncertainty of the lag between the original reference lightcurves multiplied by the baseline scaling factor X. Other symbols have the same meanings as Figure 3.
2x2 panels of Figure 5 shows an example of the lag distributions from the simulated lightcurves, and the recovered lags are consistent with the input lag. The same figures for other reference AGN and input lags are available in the supplementary material. There are three realizations for $\tau_{\text{r}} = 370.1$ days and six realizations for $\tau_{\text{r}} = 540.8$ days that pass the final sample criteria. None of them are “false positives” as defined in Section 4.1. While this simulation still does not perfectly mimic the behavior of the Mg II line due to the difference between $H\beta$ and Mg II and the simplifications used in the method, it is a qualitative verification of our lag selection criteria based on real AGN lightcurves rather than idealized models.

4.3 Comparison of Lag Selection Criteria

In addition to the final sample, there are AGNs that only pass the criteria for Method 1 or 2. Figure 6 shows examples are ambiguous and dominated by aliasing signals at the seasonal gaps. This qualitatively shows how combining the two criteria can help exclude ambiguous measurements. The symmetric lag prior range in Method 1 reduces the strength of spurious signals caused by only allowing positive lags, while Method 2 is more sensitive to spurious lag distributions with multiple aliasing peaks at significantly different lags.

To quantitatively compare the reliability of different samples, we repeat the FPR calculation in Section 4.1 but use only one set of criteria to select successful lag measurements. Figure 4 compares the FPR from a single set of criteria to that from the final sample criteria. Using just one set of criteria results in a larger overall FPR than using both. We then analyze the simulated lightcurves from Section 4.2 using only one set of selection criteria. For an input lag $\tau_{\text{r}} = 370.1$ days, we get threecorrect lag recoveries when only using the Method 1 criteria. For $\tau_{\text{r}} = 540.8$ days, we get incorrect lag recovery only when using the Method 2 criteria. This indicates that only using one set of criteria is less robust than the combined criteria.

Another commonly used method of assessing the sample reliability is comparing the number of positive and negative lag measurements (e.g., Grier et al. 2019; Homayouni et al. 2020; Yu et al. 2021). The negative lags are from artifacts, so the comparison between the number of positive and negative lags gives an overall estimate of the contamination from spurious detection. For Method 1, we get 14 negative lags and 48 positive lags. The comparison is not directly applicable to Method 2 since it only allows positive lags when running JAVELIN and ICCF. We do a one-sided search of $[-1000,0]$ days as symmetric to the positive prior $[0,1000]$ days while keeping other criteria of Method 2 the same. This gives 29 negative lags compared to 56 positive lags. When combining the two sets of criteria, we get the 25 positive lags in the final sample and no negative lags. These comparisons indicate that the final sample is less contaminated by artifacts than the samples obtained using only a single set of criteria. There are likely to be physical lag measurements in the single-method samples. However, since they are more likely to be contaminated by spurious lags than our final sample and could increase the risk of biasing the $R - L$ relation, we do not include them in our $R - L$ relation. The lightcurves of the single-method sample are available in the supplementary material.

5 BLACK HOLE MASS AND R – L RELATION

We calculate the black hole mass with Equation 1, and parameterize the line width using the line dispersion defined by

$$\sigma_{\text{line}}^2 = \int \lambda^2 P(\lambda) d\lambda \int P(\lambda) d\lambda - P_0(\lambda)^2$$

(11)

where $P(\lambda)$ is the line profile and $P_0(\lambda)$ is the first moment of $P(\lambda)$. Using the line dispersion as the line width estimator generally provides better black hole mass estimates than using the full-width at half maximum (e.g., Peterson et al. 2004; Dalla Bontà et al. 2020). Given the low SNR of our single-epoch spectra, we measure the line dispersion from the mean spectra instead of the rms spectra commonly used in RM studies. Previous studies show that the mean spectra line dispersion reasonably approximates the rms spectra line dispersion (e.g., Wang et al. 2020; Dalla Bontà et al. 2020). We adopt a virial factor of $f = 4.31 \pm 1.05$ from Grier et al. (2013b). Table 1 gives the line dispersion and black hole mass estimates for our final sample. The black hole masses have an uncertainty of $\sim 0.4$ dex, which is mainly from the intrinsic scatter in the calibration of the virial factor (e.g., Peterson 2014). We do not include detailed estimates of other measurement uncertainties since they are small relative to the uncertainty in the virial factor.

Figure 7 shows the Mg II $R - L$ relation for our final sample. Before this work, the largest Mg II lag sample was the 24 sources in the gold sample of Homayouni et al. (2020). Our sample nearly doubles the total number of high-quality Mg II lags. We significantly extend the Mg II lag measurements toward higher redshifts and bridge the gap in luminosity between the Homayouni et al. (2020) sample and the high luminosity sources studied by Lira et al. (2018), Czerny et al. (2019), Zajaček et al. (2020) and Zajaček et al. (2021), as illustrated in Figures 1 and 7. Compared to the OzDES Y3 results from Y21, the full OzDES Mg II lag sample covers about a factor of five wider range in $\log(L/L_\odot)$, which allows us to better constrain the $R - L$ relation.

We parameterize the $R - L$ relation as

$$\log(R/\text{lt-days}) = \alpha \log(L/(10^{45} \text{ erg} \cdot \text{s}^{-1})) + \beta$$

(12)

with an intrinsic scatter $\sigma_{\text{d}}$. We fit our final sample with the MCMC sampler emcee (Foreman-Mackey et al. 2013) and get best-fit parameters $\alpha = 0.39 \pm 0.08$, $\beta = 2.07 \pm 0.04$ and $\sigma_{\text{d}} = 0.15^{+0.03}_{-0.02}$. Figure 7 shows the best-fit $R - L$ relation and uncertainties. The slope is shallower than the $\sim 0.5$ slope of the $H\beta$ $R - L$ relations (e.g., Bentz et al. 2013), although they are marginally consistent at 1.5$\sigma$. The slope is broadly consistent with the Homayouni et al. (2020) gold sample given their uncertainties, but our intrinsic scatter is significantly smaller than the $\sim 0.36$-dex intrinsic scatter in their sample. It is marginally larger than the intrinsic scatter $\sigma_{\text{d}} \approx 0.13$ dex of the $H\beta$ $R - L$ relation from Bentz et al. (2013) but within uncertainties, while it is smaller than $\sigma_{\text{d}} \approx 0.19$ dex of the $H\beta$ $R - L$ relation from Du et al. (2016) after they accounted for the effect of Eddington ratio. This indicates that the Mg II line may still have a shallower slope and larger intrinsic scatter, although the intrinsic
scatter is still smaller than that from Homayouni et al. (2020) and is close to the \(0.26\)-dex intrinsic scatter from Du et al. (2016) without considering the Eddington ratio. The lags from different works are selected using different criteria and have different definitions of uncertainties, which could bias the \(R - L\) relation, so we adopt our fit to only the OzDES sample as the most robust Mg II \(R - L\) relation.

### 5.1 Effect of Iron Line Strength

Previous studies measured smaller H\(\beta\) lags from AGN with higher Eddington ratios at fixed luminosities (e.g., Du et al. 2016, 2018; Dalla Bontà et al. 2020). If also true of Mg II, it could explain the shallower slope of our Mg II \(R - L\) relation. However, the Eddington ratios in previous studies depend on the virial black hole mass estimates and therefore depend on the lag measurements themselves. One independent indicator of the Eddington ratio is the strength of the iron emission (e.g., Boroson 2002; Negrete et al. 2018; Panda et al. 2019). Du & Wang (2019) reduced the intrinsic strength of the iron emission (e.g., Boroson 2002; Negrete et al. 2018; Panda et al. 2019). Du & Wang (2019) reduced the intrinsic strength of the iron emission (e.g., Boroson 2002; Negrete et al. 2018; Panda et al. 2019). Du & Wang (2019) reduced the intrinsic strength of the iron emission (e.g., Boroson 2002; Negrete et al. 2018; Panda et al. 2019).

For each spectroscopic epoch, we calculate the iron ratio as \(R_{Fe} = F_{Fe}/F_{line}\), where \(F_{Fe}\) is the integrated iron flux over rest-frame 2250 - 2650 Å and \(F_{line}\) is the Mg II line flux. We derive the mean \((\langle R_{Fe} \rangle)\) and the standard deviation \(S_{R_{Fe}}\) over \(N_e\) epochs. The uncertainty of the mean is \(\sigma_{R_{Fe}} = S_{R_{Fe}}/\sqrt{N_e}\). Table 1 gives the iron ratio \((\langle R_{Fe} \rangle)\) and its uncertainty \(\sigma_{R_{Fe}}\) for the AGN in our final sample. Figure 8 shows the deviation of the observed lags from our Mg II \(R - L\) relation versus the iron ratio \(R_{Fe}\). It also includes the AGN from previous Mg II RM studies, with the exception of Metzroth et al. (2006) and Lira et al. (2018) which did not report \(R_{Fe}\) measurements. There is no significant correlation between the deviation from the \(R - L\) relation and the iron ratio \(R_{Fe}\).

To quantify the potential improvement from the 2-parameter \(R - L\) relation to a 3-parameter relation with \(R_{Fe}\), we consider the two parameterizations studied in Khadka et al. (2022a):

\[
\log(R/\text{lt-days}) = \alpha \log[L/(10^{45} \text{ erg} \cdot \text{s}^{-1})] + \beta + \gamma R_{Fe},
\]

\[
\log(R/\text{lt-days}) = \alpha \log[L/(10^{45} \text{ erg} \cdot \text{s}^{-1})] + \beta + \gamma \log(R_{Fe}).
\]

For the OzDES sample, we obtain an intrinsic scatter of \(\sigma_{R_{Fe}} = 0.15 \pm 0.02\) 3 for both parameterizations. When including the literature lags, both parameterizations give \(\sigma_{R_{Fe}} = 0.23 \pm 0.03\). The intrinsic scatter of these 3-parameter correlations are nearly the same as the 2-parameter \(R - L\) relation. These results indicate that the iron ratio \(R_{Fe}\) has little impact on the Mg II \(R - L\) relation, in agreement with the results of Khadka et al. (2022a).

6 SUMMARY

We use six years of photometry and spectroscopy from the OzDES RM project to measure Mg II lags. We calibrate the spectra using the pipeline developed by Hoormann et al. (2019) based on the DES photometry and estimate the calibration uncertainties using the Monte-Carlo approach from Y21 based on the F-stars observed in the OzDES fields. We use the algorithm from Y21 to model and subtract the continuum + iron emission around the Mg II line. We define quantitative criteria to select reliable measurements and verify the lag reliability using simulations based on both the DRW stochastic process and the observed lightcurves of AGN. Our major results are:
Figure 7. Mg II $R - L$ relation from the OzDES RM project and previous works. The blue squares are our high-quality measurements from the full six years of data. The green circles are the gold sample of Homayouni et al. (2020). The brown stars are measurements from Metzroth et al. (2006), Lira et al. (2018), Czerny et al. (2019), Zajaček et al. (2020) and Zajaček et al. (2021). The blue solid line shows the best-fit Mg II $R - L$ relation constrained using our final sample with the slope $\alpha$ and intrinsic scatter $\sigma_{rl}$ shown in the upper left corner. The lighter blue lines are 100 realizations randomly drawn from the MCMC chain, which illustrates the uncertainty of our $R - L$ relation. The green dashed line shows the Mg II $R - L$ relation from Homayouni et al. (2020). The red dotted line shows the H$\beta$ $R - L$ relation from Bentz et al. (2013) with a slope of $\sim 0.5$ and an intrinsic scatter of $\sim 0.13$ dex. We use the bolometric correction from Richards et al. (2006) to convert the 3000 Å monochromatic luminosity to 5100 Å for the H$\beta$ $R - L$ relation.

(i) We obtain high quality Mg II lag measurements for 25 quasars. Seven quasars were presented in Y21, and their lags from the full 6-yr lightcurves are consistent with those from Y21 based on the first five years of data. Our sample substantially increases the number of Mg II lags and extends the $R - L$ relation to higher redshifts and luminosities.

(ii) Our sample provides a new constraint of the Mg II $R - L$ relation with a slope $\alpha = 0.39 \pm 0.08$, an intercept $\beta = 2.07 \pm 0.04$ and an intrinsic scatter $\sigma_{rl} = 0.15^{+0.03}_{-0.02}$. The slope is consistent with the Mg II $R - L$ relation from Homayouni et al. (2020), while it is shallower than the H$\beta$ $R - L$ relation based on local AGN. The intrinsic scatter is significantly smaller than that from Homayouni et al. (2020) and is close to that of the H$\beta$ $R - L$ relations, which makes it promising to apply our $R - L$ relation to large samples of single-epoch mass estimates and SMBH demographic studies at cosmic noon.

(iii) The residual from the $R - L$ relation has no significant correlation with the relative iron strength $R_{Fe}$ as an indicator of the Eddington ratio. Adding $R_{Fe}$ as a third parameter does not reduce the intrinsic scatter of the $R - L$ relation.

Future RM campaigns based on wide-field photometric and spectroscopic surveys, such as the Rubin Observatory Legacy Survey of Space and Time (LSST, e.g., LSST Science Collaboration et al. 2009), the Black Hole Mapper in SDSS-V (e.g., Kollmeier et al. 2017) and the Time-Domain Extragalactic Survey (TiDES) with 4MOST (e.g., Swann et al. 2019), should provide larger sample of lag measurements and better constraints of the $R - L$ relation. In addition to deeper observations with future facilities, the optimization of survey strategy is critical for successful RM campaigns. For example, Malik et al. (2022) used simulated lightcurves to study the effect of observational windows on the lag recovery in future...
ACKNOWLEDGEMENTS

ZY and CSK are supported by Chandra grant GO9-20084X. PM is grateful for support from the Radcliffe Institute for Advanced Study at Harvard University. PM also acknowledges support from the United States Department of Energy, Office of High Energy Physics under Award Number DE-SC-0011726. CSK is supported by NSF grants AST-1908570 and AST-1814440. Parts of this research were supported by the Australian Government through the Australian Research Council Laureate Fellowship FL180100168.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Department of Energy OneWorld Research Initiative, the Grainger Foundation, and the Keck Foundation. The DES Collaboration includes the University of Chicago, the University of Pennsylvania, Stanford University, the University of Sussex, Texas A&M University, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l’Espai (IEEC/CSIC), the Institut de Física d’Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF’s NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory at NSF’s NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Program (FP7/2007–2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

DATA AVAILABILITY

We present supplementary materials in the online journal:

(i) Figures of the single-epoch spectra in the same format as Figure 2 for the final sample sources. These figures also include the co-added and rms spectra for each source.

(ii) Machine readable lightcurves for the final sample sources and sources that only pass one set of criteria.

(iii) Figures of the lightcurves and lag measurements in the same format as Figure 3 for the final sample sources.

(iv) Figures of the simulated lightcurves based on the observed lightcurves of local AGN (Section 4.2). They are in the same format as Figure 5 but for other reference AGN and input lag.

The underlying DES and OzDES data are available in Abbott et al. (2021) and Lidman et al. (2020).

REFERENCES

Abbott T. M. C., Abdalla F. B., Allam S., et al., 2018, ApJS, 239, 18
Abbott T. M. C., et al., 2021, ApJS, 255, 20
Barth A. J., et al., 2011a, ApJ, 732, 121
Barth A. J., et al., 2011b, ApJ, 743, L4
