The Astrophysical Journal Supplement Series, 272:26 (32pp), 2024 June https://doi.org/10.3847/1538-4365/ad3936
© 2024. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

The Sloan Digital Sky Survey Reverberation Mapping Project: Key Results

Yue Shen[1,2], Catherine J. Grier[3], Keith Horne[4], Zachary Stone[1,5], Jennifer I. Li[6,7], Qian Yang[1,8], Yasaman Homayouni[9,10], Jonathan R. Trump[11], Scott F. Anderson[12], W. N. Brandt[10,13,14], Patrick B. Hall[15], Luis C. Ho[16,17], Linhua Jiang[16,17], Patrick Petitjean[18], Donald P. Schneider[10,13], Charling Tao[19,20], Fergus. R. Donnan[21], Yusra AlSayyad[22], Matthew A. Bershady[3,23,24], Michael R. Blanton[25], Dmitry Bizyaev[26,27], Kevin Bundy[28], Yuguang Chen[29], Megan C. Davis[11], Kyle Dawson[30], Xiaohui Fan[31], Jenny E. Greene[22], Hannes Gröller[32], Yucheng Guo[17,33], Héctor Ibarra-Medel[34,35], Yuanzhe Jiang[10], Ryan P. Keenan[31], Juna A. Kollmeier[36,37], Cassandra Lejoly[32], Zefeng Li[38], Axel de la Macorra[39], Maxwell Moe[31,40], Jundan Nie[41], Graziano Rossi[42], Paul S. Smith[31], Wei Leong Tee[31], Anne-Marie Weijmans[4], Jiachuan Xu[31], Minghao Yue[31], Xu Zhou[41], Zhimin Zhou[41], and Hu Zou[41]

[1] Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2] National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3] Department of Astronomy, University of Wisconsin-Madison, 475 North Charter Street, Madison, WI 53703, USA
[4] SUPA Physics and Astronomy, University of St Andrews, Fife, KY16 9SS, UK
[5] Center for AstroPhysical Surveys, National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[6] Michigan Institute for Data Science, University of Michigan, Ann Arbor, MI 48109, USA
[7] Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
[8] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[9] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[10] Department of Astronomy and Astrophysics, 525 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA
[11] Department of Physics, 196A Auditorium Road, Unit 3046, University of Connecticut, Storrs, CT 06269, USA
[12] Astronomy Department, University of Washington, Box 351580, Seattle, WA 98195, USA
[13] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[14] Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[15] Department of Physics & Astronomy, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada
[16] Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China
[17] Department of Astronomy, School of Physics, Peking University, Beijing 100871, People's Republic of China
[18] Institut d'Astrophysique de Paris & Sorbonne Université, 98bis Boulevard Arago 75014, Paris, France
[19] Center for Astrophysics, Department of Astronomy, Tsinghua University, Beijing, People's Republic of China
[20] Centre de Physique des Particules de Marseille, IN2P3, CNRS, France
[21] Department of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK
[22] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[23] South African Astronomical Observatory, P.O. Box 9, Observatory 7935, Cape Town, South Africa
[24] Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa
[25] Center for Cosmology and Particle Physics, Department of Physics, 726 Broadway, Room 1005, New York University, New York, NY 10003, USA
[26] Apache Point Observatory and New Mexico State University, P.O. Box 59, Sunspot, NM 88349-0059, USA
[27] Sternberg Astronomical Institute, Moscow State University, Moscow, Russia
[28] UCO/Lick Observatory, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[29] Department of Physics & Astronomy, University of California, Davis, CA 95616, USA
[30] Department of Physics and Astronomy, University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA
[31] Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[32] Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721, USA
[33] Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France
[34] Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional, U.P. Adolfo López Mateos, C.P. 07738, Ciudad de México, Mexico
[35] Universidad Nacional Autónoma de México, Instituto de Astronomía, AP 70-264, CDMX 04510, Mexico
[36] The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[37] Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, ON M5S-98H, Canada
[38] Research School of Astronomy & Astrophysics, Australian National University, Weston Creek, ACT 2611, Australia
[39] Instituto de Física, Universidad Nacional Autónoma de México, Apdo. Postal 20-364, Mexico
[40] Department of Physics & Astronomy, University of Wyoming, Laramie, WY 82071, USA
[41] Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, People's Republic of China
[42] Department of Astronomy and Space Science, Sejong University, 209, Neungdong-ro, Gwangjin-gu, Seoul, Republic of Korea
Received 2023 May 11; revised 2024 March 26; accepted 2024 March 28; published 2024 May 15

Abstract

We present the final data from the Sloan Digital Sky Survey (SDSS) Reverberation Mapping (RM) project, a precursor to the SDSS-V Black Hole Mapper RM program. This data set includes 11 yr photometric and 7 yr spectroscopic light curves for 849 broad-line quasars over a redshift range of $0.1 < z < 4.5$ and a luminosity range of $L_{\rm bol} = 10^{44-47.5}$ erg s^{-1}, along with spectral and variability measurements. We report 23, 81, 125, and 110 RM lags (relative to optical continuum variability) for broad Hα, Hβ, Mg II, and C IV using the SDSS-RM sample, spanning much of the luminosity and redshift ranges of the sample. Using 30 low-redshift RM active galactic

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

nuclei with dynamical-modeling black hole masses, we derive a new estimate of the average virial factor of $\langle \log f \rangle = 0.62 \pm 0.07$ for the line dispersion measured from the rms spectrum. The intrinsic scatter of individual virial factors is 0.31 ± 0.07 dex, indicating a factor of 2 systematic uncertainty in RM black hole masses. Our lag measurements reveal significant R–L relations for Hβ and Mg II at high redshift, consistent with the latest measurements based on heterogeneous samples. While we are unable to robustly constrain the slope of the R–L relation for C IV given the limited dynamic range in luminosity, we found substantially larger scatter in C IV lags at fixed L_{1350}. Using the SDSS-RM lag sample, we derive improved single-epoch (SE) mass recipes for Hβ, Mg II, and C IV, which are consistent with their respective RM masses as well as between the SE recipes from two different lines, over the luminosity range probed by our sample. The new Hβ and Mg II recipes are approximately unbiased estimators at given RM masses, but there are systematic biases in the C IV recipe. The intrinsic scatter of SE masses around RM masses is $\sim$0.45 dex for Hβ and Mg II, increasing to $\sim$0.58 dex for C IV.

Unified Astronomy Thesaurus concepts: Reverberation mapping (2019); Surveys (1671); Quasars (1319)

1. Introduction

Accurate measurements of supermassive black hole (hereafter SMBH or BH for short) masses are cornerstones for understanding the cosmic assembly of SMBHs, their potential coevolution with host galaxies, and fundamental accretion physics. Far beyond the nearby Universe, it is difficult to use spatially resolved gas or stellar kinematics to measure a dynamical mass of the accreting SMBH. Currently, the primary method to measure black hole masses in distant active galactic nuclei (AGNs) or quasars is reverberation mapping (RM; e.g., Blandford & McKee 1982; Peterson 1993; Cackett et al. 2021). RM measures a characteristic size of the broad-line region (BLR) from the time lag between continuum variability and the response in the broad-line flux. By combining the size of the BLR (assumed to be viralized) with the broad-line width (as the surrogate for the virial velocity), one can derive a virial black hole mass estimate for the AGN. RM results for local AGNs provide the foundation of the so-called "single-epoch" method (e.g., Greene & Ho 2005; Vestergaard & Peterson 2006; Shen 2013) that allows efficient estimation of quasar black hole masses using single-epoch (SE) spectroscopy (e.g., Shen et al. 2008, 2011; Wu & Shen 2022). This SE technique is built on the observed tight relation between the Hβ BLR size and the optical luminosity of the AGN for the local RM sample (e.g., Kaspi et al. 2000; Bentz et al. 2013), although some recent RM studies targeting a broad range of AGN properties have shown deviations and increased scatter in the R–L relation (e.g., Du et al. 2016; Fonseca Alvarez et al. 2020). The systematic uncertainties of extrapolating the local RM results to distant, luminous quasars, however, are not yet well quantified (e.g., see the review in Shen 2013).

In the past few years, there have been two major advancements in the RM field. On the one hand, high-quality monitoring data have become available for bright, low-redshift (typically $z \lesssim 0.3$) AGNs to measure broad-line RM lags with unprecedented precision, and to cover a broader range of accretion parameters (e.g., Barth et al. 2013; Du et al. 2016, 2018; Fausnaugh et al. 2016; Brotherton et al. 2020; Kara et al. 2021). These high-quality RM data further enabled more powerful constraints on the dynamics of the BLR with velocity-resolved RM (e.g., Grier et al. 2013; De Rosa et al. 2018), or dynamical modeling of the BLR response to continuum variations (e.g., Pancoast et al. 2014; Li et al. 2018; Williams et al. 2020). These new measurements of high-quality RM data are offering new insights on the structure and kinematics of the BLR gas, as well as high-fidelity BH masses in AGNs. In particular, the SEAMBH project (Du et al. 2016, 2018) targeting high-accretion-rate low-redshift AGNs is extending RM studies to systems with extreme accretion parameters. However, most of these high-quality RM measurements are for Hβ, with only two cases for C IV using intensive UV spectroscopy from the Hubble Space Telescope (De Rosa et al. 2015; Kara et al. 2021).

On the other hand, there have been considerable efforts to push RM to the high-redshift regime, and for additional broad lines in AGN spectra. For example, there have been monitoring programs to measure RM lags in high-redshift and high-luminosity quasars (e.g., Kaspi et al. 2007; Lira et al. 2018; Czerny et al. 2019a; Kaspi et al. 2021), with decade-long baselines to capture the anticipated multiyear lags. These most luminous quasars are rare and sparsely distributed on the sky, thus requiring individual monitoring campaigns. In most of these cases, the data quality is only sufficient to derive a mean lag. One notable exception is the gravitationally lensed quasar SDSS J2222+2745 at $z = 2.8$, for which Gemini optical monitoring spectroscopy combined with ground-based photometric light curves enabled velocity-resolved RM measurements (Williams et al. 2021). The success of RM measurements in SDSS J2222+2745 rests on anticipated variability features from prior photometric monitoring for this lensed quasar. For other monitoring programs of high-redshift quasars, however, the variability patterns are unknown ahead of time, resulting in a reduced success rate of lag measurements. Therefore, individual object monitoring becomes expensive and inefficient in building up the statistics of high-redshift RM measurements. One solution to circumvent these observational difficulties is to perform multiobject spectroscopic (MOS) monitoring that piggybacks on a survey program to greatly improve the efficiency of RM monitoring.

The Sloan Digital Sky Survey (SDSS) RM project (Shen et al. 2015a) and the OzDES-RM project (King et al. 2015) recently emerged as the first two MOS-RM programs to perform RM measurements for large statistical samples of quasars far beyond the low-redshift Universe. Both MOS-RM programs target all quasars within small areas of the sky that can be simultaneously monitored with wide-field optical MOS facilities. Because the target quasars are typically fainter than $m \sim 18$ in the optical, the spectral signal-to-noise ratio (S/N) is on average much lower than those achieved for low-redshift RM campaigns, a situation to be improved with future larger-aperture MOS facilities for a similar target brightness (e.g., McConnachie et al. 2016; Swann et al. 2019). Nevertheless, these MOS-RM programs are starting to produce large numbers of lag measurements at $z > 0.3$ that cover major rest-frame UV

Table 1
Lag Detection Summary

No. of Lags	Hα	Hβ	Mg II	C IV
All attempted	53	186	714	494
PyROA (>2σ det)	23	81	125	110
PyROA (>1σ det)	31	108	189	163

Note. We report all lag measurements in Table 2. But we recommend to only use lags with >2σ detection in statistical analyses.

to optical broad emission lines (e.g., Shen et al. 2016, 2019a; Grier et al. 2017b, 2019; Hoormann et al. 2019; Homayouni et al. 2020; Malik et al. 2023; Yu et al. 2023).

In this work, we present the final data set from the SDSS-RM project, which includes 11 yr of photometric monitoring (2010–2020) and 7 yr of spectroscopic monitoring (2014–2020). Using this data set, we measure RM lags for four major broad lines in quasar spectra (see Table 1): Hα, Hβ, Mg II, and C IV. We describe the data in Section 2, and the preparation of light curves in Section 3. We detail our methodology of lag measurements in Section 4, and present our RM results along with discussions in Section 5. We conclude in Section 6 with future prospects. In the appendices, we provide additional information about data access and lag detection efficiencies. Throughout this paper, we adopt a flat ΛCDM cosmology with $\Omega_\Lambda = 0.7$ and $H_0 = 70\,\mathrm{km\,s^{-1}Mpc^{-1}}$. By default, time lags refer to those in the observed frame unless otherwise specified. We adopt the convention that a positive lag means line variability lags behind continuum variability.

2. Data

2.1. The SDSS-RM Sample

The SDSS-RM sample contains 849 broad-line quasars at $0.1 < z < 4.5$ spectroscopically confirmed in a single 7 $\mathrm{deg^2}$ field (R.A. J2000 = 213.704 deg, decl. J2000 = +53.083 deg). These quasars were discovered in previous SDSS surveys and were selected by different algorithms (see details in Shen et al. 2019b). The sample is flux limited to $i_{\rm psf} = 21.7$, although the completeness near this flux limit is not well quantified and is particularly low at $z < 0.5$ (see Figure 1 in Shen et al. 2019b). No additional selection cuts, such as variability or emission-line strength, were imposed on the SDSS-RM sample. The SDSS-RM sample covers a broad and contiguous range in redshift–luminosity space, and is representative of luminous quasars with $L_{\rm bol} \gtrsim 10^{45}\,\mathrm{erg\,s^{-1}}$. The detailed sample characterization is provided in Shen et al. (2019b), where spectral variability properties are based on the 2014 SDSS-RM spectra alone.

2.2. Spectroscopy

Optical spectroscopic monitoring was obtained with the BOSS multiobject spectrographs (Smee et al. 2013) on the SDSS telescope (Gunn et al. 2006) during 2014–2020, as an ancillary program within SDSS-III (Eisenstein et al. 2011) and SDSS-IV (Blanton et al. 2017). The spectroscopy had an average cadence of ∼4 days in 2014, with a total of 32 epochs at a nominal exposure time of 2 hr each. In subsequent years, SDSS-RM obtained ∼12 epochs per year (two per month) with a nominal exposure time of 1 hr each during 2015–2017 and ∼6 epochs per year (monthly cadence) during 2018–2020. The final spectroscopic baseline covers 7 yr, with a total of 90 spectroscopic epochs. The wavelength coverage of BOSS spectroscopy is ∼3650–10400 Å, with a spectral resolution of $R \sim 2000$. The typical S/N per 69 $\mathrm{km\,s^{-1}}$ pixel averaged over the g band in a 2 hr exposure is ∼4.5 at $g_{\rm psf} = 21.2$, but could be lower for epochs observed with poor observing conditions. The nominal broadband spectrophotometry accuracy is ∼5% in gri bands, but can be worse than ∼10% below 4000 Å (Shen et al. 2015a).

2.3. Photometry

Optical monitoring photometry provides continuum light curves to facilitate RM lag measurements. SDSS-RM obtained photometry in the g and i bands with the Steward Observatory Bok 2.3 m telescope on Kitt Peak and the 3.6 m Canada–France–Hawaii Telescope (CFHT) on Maunakea. The Bok and CFHT imaging roughly covers the same monitoring period as the SDSS-RM spectroscopy, with the highest cadence in 2014 and reduced cadences in subsequent years. The details of these photometric observations and data processing are described in Kinemuchi et al. (2020). In addition to dedicated Bok and CFHT photometry, the SDSS-RM field coincides with the MD07 Medium-Deep Field in the PanSTARRS-1 (PS1) survey (Kaiser et al. 2010), with few-night-cadence, multiband ($grizY$) photometric light curves during 2010–2013. These PS1 data were published in Shen et al. (2019b), and provide early continuum light curves to extend the effective baseline for our RM measurements. Finally, we compile available photometric light curves for the SDSS-RM sample from the Zwicky Transient Facility (ZTF) public survey (Bellm et al. 2019) during 2018–2020. The combined photometric light curves span a baseline of 11 yr (2010–2020), and were merged to provide a uniform continuum light curve for each quasar (Section 3.2).

3. Light-curve Compilation

3.1. PrepSpec

PrepSpec is a spectral refining procedure performed on the flux-calibrated multiepoch SDSS spectra in order to reduce further the scatter in the flux calibration (Shen et al. 2016). PrepSpec rescales the flux levels of each individual epoch by optimizing model fits (in parameterized functional forms) to describe the continuum and broad-line variability patterns as functions of time and wavelength, using the fluxes of the narrow emission lines (in particular [O III] λλ 4959,5007) as an internal calibrator (e.g., van Groningen & Wanders 1992), which are assumed to remain constant over the relatively short monitoring period compared with the light-crossing time of the narrow-line region. This procedure improves the calibration of the relative spectrophotometry to ∼2% for low-redshift quasars with strong narrow emission lines. This refining procedure is particularly useful in calibrating a small number of spectroscopic fiber+epochs that suffered significant flux losses in the fiber due to unknown systematics (Shen et al. 2015a).

PrepSpec also generates model light curves (and uncertainties) for a given set of broad emission lines and continuum fluxes, as well as model mean and rms profiles of the broad lines. We use these broad-line light curves in our lag measurements, and the line widths in the calculation of RM

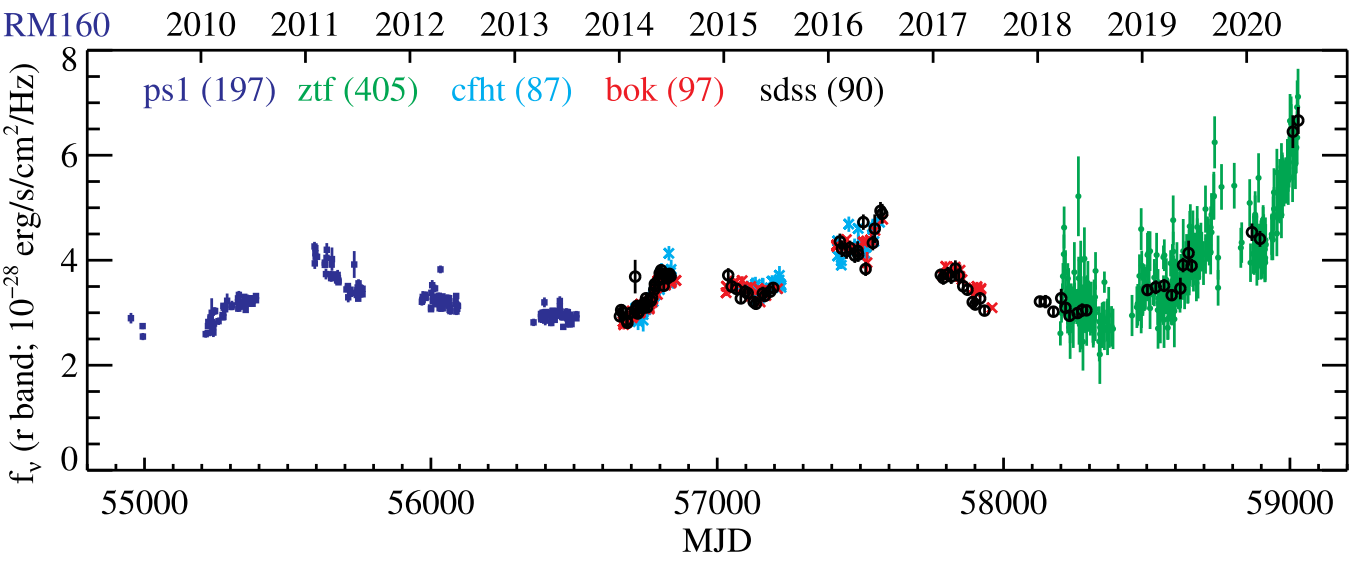

Figure 1. An example of merged 11 yr photometric light curves for the SDSS-RM sample (see Section 3.2 for details). Light curves from different surveys are calibrated to the SDSS *r* band. For each SDSS-RM quasar, there are typically hundreds of nightly averaged photometric epochs, to go with 90 spectroscopic epochs (black open circles) from SDSS.

BH masses. The full technical details of PrepSpec are described in Shen et al. (2016). All PrepSpec outputs are released as part of the final data products of SDSS-RM.

Compared with earlier PrepSpec results used for intermediate SDSS-RM lag measurements (Shen et al. 2016, 2019b; Grier et al. 2017b, 2019; Homayouni et al. 2020), there are several improvements. First of all, we have run PrepSpec on the full 7 yr spectroscopic data, which provide better constraints on variability models than earlier shorter light curves. Second, PrepSpec measured more reliable broad-line widths in the mean spectrum by subtracting the emission of the Fe II complex. Third, PrepSpec adopted more accurate systemic redshifts provided in Shen et al. (2019b) and refined line windows for better line-width and flux measurements.

The broad-line rms spectrum constructed by PrepSpec is the true rms profile for the variable line component. This is also the standard approach used by most of the recent RM campaigns on low-z AGNs (e.g., the LAMP project, Barth et al. 2015), where the individual-epoch spectra are decomposed into separate emission-line and continuum components. Some earlier RM studies (e.g., Peterson et al. 2004; Grier et al. 2012) construct the rms spectrum from the total (emission line +continuum) epoch spectra, and subtract a local continuum to derive a line-only rms spectrum. There are systematic biases in the measured rms line widths from the latter approach when the monitoring period is not much longer than the lag (e.g., Barth et al. 2015; Wang et al. 2019). We will revisit this issue in Section 4.6.

3.2. Merged Continuum Light Curves

We combine the photometric light curves from different facilities/surveys to form a merged continuum light curve for lag measurements, as commonly done for RM studies. For the PS1 data, we use observations in *gri* filters to improve the sampling. The *zY* data are not used since the continuum variability there is small and more contaminated by host-galaxy emission. We use the PS1 data published in Shen et al. (2019b), and coadd the nightly observations weighted by the inverse variance of the individual flux measurements.

For the purpose of measuring broad-line lags, we ignore the interband continuum lags. Such continuum lags trace the structure of AGN accretion disks, which are typically much more compact than the BLR; hence, these continuum lags are on much shorter timescales (e.g., less than a few rest-frame days, but can be longer for more luminous quasar accretion disks; Fausnaugh et al. 2016). A complication is that a fraction (e.g., ∼10%–20%) of the optical continuum emission may come from the diffuse BLR emission (e.g., Guo et al. 2022), and cause interband continuum lags that are nonnegligible compared with broad-line lags. This is still an ongoing investigation, and without a robust method to correct for interband continuum lags for SDSS-RM quasars, we simply follow previous RM work and ignore this detail. We scale the merged light curves (in f_ν) to physical flux scales in *r* band. This band was chosen because it offers the best spectrophotometry in SDSS spectra (Shen et al. 2015a). We use the synthetic broadband flux computed from the SDSS spectroscopy as the baseline to calibrate and merge other photometric light-curve sets.

The continuum light-curve merging is performed with the public code `PyCali` (Li et al. 2014). `PyCali` is a Bayesian method to calibrate the light curves obtained with different facilities at different times to a reference set of light curves, during which the flux uncertainties of individual light-curve sets can also be adjusted to account for overestimated or underestimated flux uncertainties in the original data set. Before running `PyCali`, we coadd (weighted by inverse variance) intranight photometry points from the same facility and same band. However, for CFHT and Bok observations, some objects are covered on multiple detector chips, and we only coadd the nightly photometry from the same chip, using the light curves compiled in Kinemuchi et al. (2020).

We show an example of the merged 11 yr continuum light curves in Figure 1. The full set of `PyCali`-merged continuum

Table 2
Format of the Summary Table

Name (1)	Format (2)	Units (3)	Description (4)
RMID	LONG	⋯	RMID of the target
ZSYS	DOUBLE	⋯	Systemic redshift from Shen et al. (2019b)
RA	DOUBLE	deg	J2000 R.A.
DEC	DOUBLE	deg	J2000 decl.
DL_MPC	DOUBLE	Mpc	Luminosity distance
LOGL5100_YR1	DOUBLE	[erg s^{-1}]	Monochromic luminosity at rest-frame 5100 Å from 2014
LOGL3000_YR1	DOUBLE	[erg s^{-1}]	Monochromic luminosity at rest-frame 3000 Å from 2014
LOGL1350_YR1	DOUBLE	[erg s^{-1}]	Monochromic luminosity at rest-frame 5100 Å from 2014
F_H_5100_YR1	DOUBLE	⋯	Host-to-total fraction (within SDSS fiber) from 2014
L5100_MEAN	DOUBLE	erg s^{-1}	Mean rest-frame 5100 Å luminosity during 2014–2020
L5100_MEDIAN	DOUBLE	erg s^{-1}	Median rest-frame 5100 Å luminosity during 2014–2020
L3000_MEAN	DOUBLE	erg s^{-1}	Mean rest-frame 3000 Å luminosity during 2014–2020
L3000_MEDIAN	DOUBLE	erg s^{-1}	Median rest-frame 3000 Å luminosity during 2014–2020
L2500_MEAN	DOUBLE	erg s^{-1}	Mean rest-frame 2500 Å luminosity during 2014–2020
L2500_MEDIAN	DOUBLE	erg s^{-1}	Median rest-frame 2500 Å luminosity during 2014–2020
L1350_MEAN	DOUBLE	erg s^{-1}	Mean rest-frame 1350 Å luminosity during 2014–2020
L1350_MEDIAN	DOUBLE	erg s^{-1}	Median rest-frame 1350 Å luminosity during 2014–2020
HB_LOGMSE	DOUBLE	[$M_\odot$]	Single-epoch BH mass based on Hβ (Vestergaard & Peterson 2006)
MG2_LOGMSE	DOUBLE	[$M_\odot$]	Single-epoch BH mass based on Mg II (Shen et al. 2011)
C4_LOGMSE	DOUBLE	[$M_\odot$]	Single-epoch BH mass based on C IV (Vestergaard & Peterson 2006)
LAG_OBS_EXP	DOUBLE	days	Expected observed-frame lag for Hβ
$line$_W_MEAN	DOUBLE[2]	km s^{-1}	[FWHM, line dispersion] from the mean spectrum
$line$_W_RMS	DOUBLE[2]	km s^{-1}	[FWHM, line dispersion] from the rms spectrum
$line$_W_MEAN_ERR	DOUBLE[2]	km s^{-1}	Uncertainty in $line$_W_MEAN
$line$_W_RMS_ERR	DOUBLE[2]	km s^{-1}	Uncertainty in $line$_W_RMS
$line$_STAT_NAME	STRING[24]	⋯	Names of line light-curve statistics
$line$_STAT_VALUE	DOUBLE[24]	⋯	Values of line light-curve statistics
DRW_SIGMA	DOUBLE	⋯	Amplitude from continuum DRW fit
DRW_SIGMA_ERR	DOUBLE[2]	⋯	Uncertainty in DRW_SIGMA
DRW_TAU	DOUBLE	days	Observed-frame damping timescale from continuum DRW fit
DRW_TAU_ERR	DOUBLE[2]	days	Uncertainty in DRW_TAU
CONT_FRAC_RMS	DOUBLE	⋯	Intrinsic fractional rms variability from 11 yr continuum light curve
CONT_FRAC_RMS_ERR	DOUBLE	⋯	Uncertainty in CONT_FRAC_RMS
LAG_SEARCH	DOUBLE[2]	days	Lag search range (in observed frame)
$line$_LAG_DONE	INT	⋯	=1 if a lag measurement is attempted
$line$_LAG_DET	DOUBLE	days	Fiducial observed-frame lag measurement
$line$_LAG_DET_ERR	DOUBLE[2]	days	Uncertainty in $line$_LAG_DET [−, +]
$line$_LAG_FPR	DOUBLE	⋯	False-positive rate of the lag; −1 if the lag is not detected
$line$_MRM	DOUBLE	$M_\odot$	RM BH mass based on the fiducial lag
$line$_MRM_ERR	DOUBLE[2]	$M_\odot$	Uncertainty in $line$_MRM [−, +]
$line$_ICCF_*	DOUBLE	⋯	ICCF results
$line$_JAVELIN_*	DOUBLE	⋯	`Javelin` results
$line$_PYROA_*	DOUBLE	⋯	PyROA results
$line$_LAG_GRADE	LONG	⋯	Lag quality grade by visual inspection (1: poorest; 5: best)

Note. For each of the lag measuring methods (ICCF, `Javelin` and `PyROA`), we provide the best lag (*_LAG), lag uncertainties (*_LAG_ERR), bounds of the identified primary peak in the lag posterior (*_PEAK_BOUNDS), the fraction of rejected samples (*_FRAC_REJ), and the maximum Pearson correlation coefficient within $\pm 1\sigma$ of the best lag (*_LAG_RMAX). In a handful of cases, the reported `PyROA` lag may lie outside of the LAG_SEARCH range. These are cases where the original `PyROA` lag measurement reached 90% of the search boundary, and we have expanded the search range (for that specific line) to the maximum allowed range ([−2500, 4000] days). Finally, we caution that the false-positive rate estimate for individual lag detections ($line$_LAG_FPR) is often an overestimation, and should be combined with the lag measurement diagnostic plot to assess the fidelity of the reported lag.

light curves is released as part of the data products from this work.

3.3. Light-curve Properties

We quantify the variability characteristics of the light curves and compile the results in the summary of Table 2. We use a maximum-likelihood (ML) estimator (Shen et al. 2019b) to measure the intrinsic variability (rms) of the light curve, corrected for flux uncertainties. These rms variability measurements are performed for both the 11 yr photometric light curves and the 7 yr emission-line light curves from spectroscopy +PrepSpec analysis. In addition to the rms variability metric, we use another empirical metric to assess evidence for intrinsic

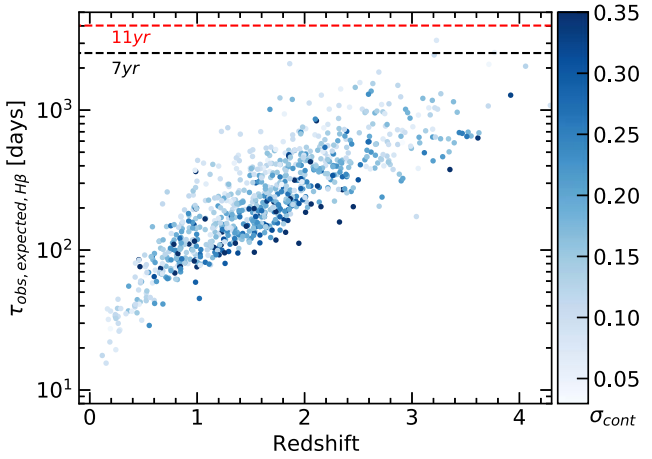

Figure 2. Expected Hβ lag in the observed frame for the SDSS-RM sample, estimated using the mean Hβ R–L relation from Bentz et al. (2013). The points are color coded by the fractional rms variability $\sigma_{\rm cont}$ from the 11 yr photometric light curve, corrected for flux measurement uncertainties. The black and red dashed lines are the 7 yr spectroscopic baseline and the 11 yr photometric baseline, respectively. There is a general tendency of increasing variability toward lower luminosities (shorter lags) at fixed redshift.

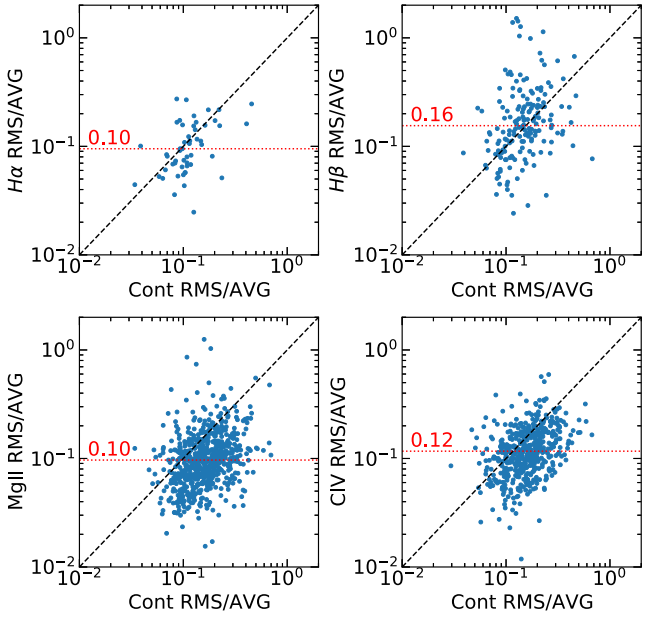

Figure 3. Fractional (intrinsic) rms variability for the four broad lines considered in this work vs. that of the continuum light curve. The dashed diagonal lines show the unity relation, and the horizontal red lines (with numbers) indicate the median fractional line variability. There are reasonably good correlations between the line and continuum rms variability, as expected from reverberation.

variability of the light curve (Shen et al. 2019b), defined as[43] SNR2 $= \sqrt{\chi^2 - {\rm dof}}$. Here, $\chi^2 \equiv \Sigma(y_i - \langle y_i \rangle)^2/\sigma_i^2$ is relative to the optimal average of the light curve $\langle y_i \rangle$, using flux measurement errors σ_i only; degrees of freedom (dof) equals $N-1$, where N is the number of light-curve points. Larger intrinsic variability (with respect to flux uncertainties) tends to have a larger SNR2.

Figure 2 displays the distribution of the SDSS-RM sample in the redshift versus expected (Hβ) lag space, color coded by the intrinsic fractional rms variability of the continuum. The continuum rms variability has a sample median of $\sim$15%, and broadly increases with decreasing luminosity at a given redshift (e.g., Vanden Berk et al. 2006; MacLeod et al. 2012). Figure 3 displays the comparisons between the continuum and broad-line (fractional) rms variabilities. Given these typical fractional variability amplitudes in the continuum and in the responding lines, measuring a lag requires good flux calibration and spectral S/N. Many targets in the SDSS-RM sample do not have the adequate S/N in the emission-line light curves to allow a robust lag detection. In addition, to measure a lag, the light-curve variability must contain distinctive inflection "features," as a monotonically changing light curve will produce broad correlations over a range of lags. Nevertheless, given the large sample size, there are many SDSS-RM quasars with well-detected variability patterns to measure the lags. A handful of the best-quality light curves are even good enough to measure velocity-resolved lags, which will be presented elsewhere (e.g., Fries et al. 2023).

Compared with the optical continuum variability, different broad lines exhibit different levels of fractional variability. Figure 3 shows that, while there are broad correlations between the continuum and line fractional rms variability, Mg II and Hα have slightly lower average variability compared with C IV and Hβ. The lower variability for Mg II, as seen in earlier studies (e.g., Sun et al. 2015; Yang et al. 2020), can be understood if the Mg II-emitting gas is farther out than the BLRs of other species, and/or the response of Mg II is in general weaker than Hβ, as shown in photoionization calculations (e.g., Goad et al. 2012; Guo et al. 2020). The high-ionization C IV line usually shows larger rms variability than Hβ (Peterson et al. 2004) in local low-luminosity AGNs. However, for high-luminosity quasars, it is argued that a portion of the C IV line does not reverberate to continuum variability, thus diluting the line fractional rms variability (e.g., Denney 2012; Wang et al. 2020b). The situation may be further complicated by the fact that the ionizing continuum (powering the lines) and the optical continuum have different rms variability, as the variable accretion flow onto the SMBH results in varying shapes of the spectral energy distribution (SED).

Although not covered in our lag measurements, we show the fractional rms variability for additional broad lines in Figure 4. The two strong UV broad lines, Si IV and C III], show similar rms variability as C IV, while the two high-ionization lines He II 1640 and He II 4686 show notably stronger rms variability than other lines. The behaviors of the He II broad lines are consistent with local RM observations (e.g., Peterson et al. 2004).

Finally, we show the comparisons between the FWHM measured from the mean spectrum and the line dispersion $\sigma_{\rm line,rms}$ measured from the rms spectrum in Figure 5, using PrepSpec outputs. A strong correlation between these two line widths for the same line is a necessary (but not sufficient) condition for the validity of SE virial mass estimators. The average ratios of FWHM/$\sigma_{\rm line}$ for all four broad lines are smaller than 2.35, implying the line profile is more Lorentzian (e.g., Collin et al. 2006; Villafaña et al. 2023).

[43] In rare cases within the SDSS-RM sample, if the light curve is consistent with noise or the flux uncertainties are overestimated, $\chi^2 - {\rm dof}$ may become negative. These cases will not yield a meaningful lag detection. In such cases, PrepSpec calculates SNR2 $= -\sqrt{|\chi^2 - {\rm dof}|}$, which is negative.

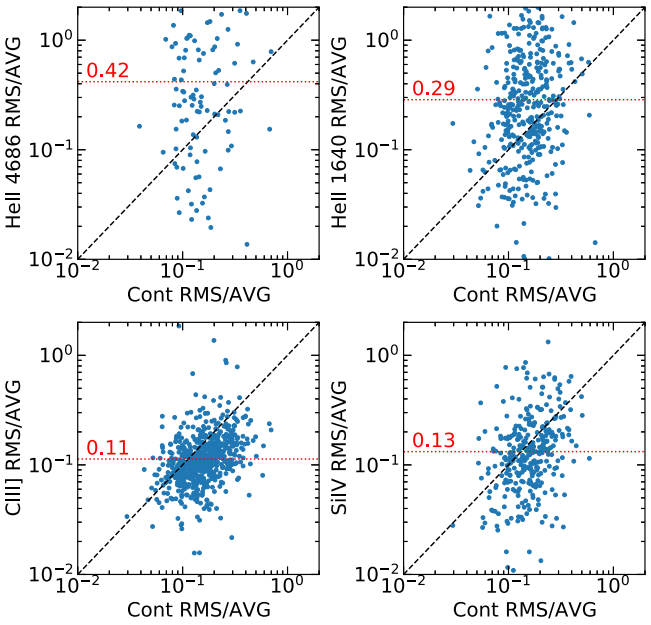

Figure 4. Fractional (intrinsic) rms variability for additional broad lines vs. that of the continuum light curve. We do not report lag measurements for these lines, but their variability properties are compiled in Table 2.

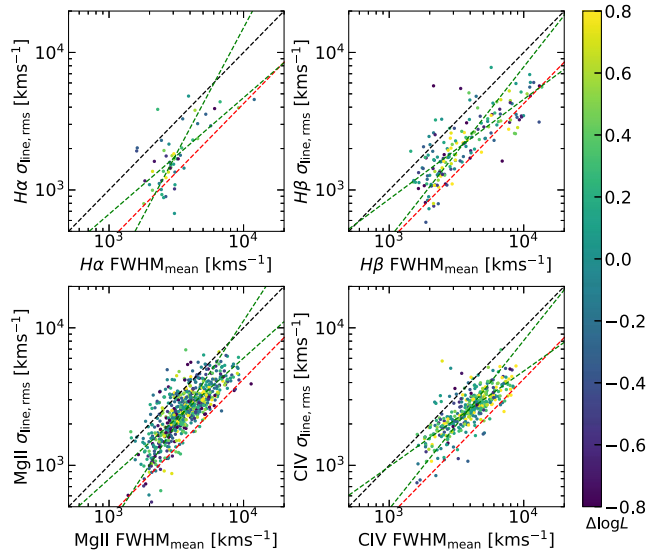

Figure 5. Comparisons between the FWHM from the mean spectrum (for SE estimators) and the $\sigma_{\rm line,rms}$ from the rms spectrum (for RM masses) for Hα, Hβ, Mg II, and C IV, color coded by the deviation from the sample median luminosity for the corresponding line (e.g., $L5100$, $L3000$, etc.). The Mg II $\sigma_{\rm line,rms}$ here has been corrected for the velocity split of the doublet, but not for the SDSS spectral resolution (negligible). In each panel, the black dashed line is the unity relation, and the red dashed line represents the FWHM–$\sigma_{\rm line}$ relation for a Gaussian profile. The best-fit linear regressions of $(Y|X)$ and $(X|Y)$ are also shown. There is no clear luminosity trend (measured with the deviations from the sample median luminosity, $\Delta \log L$) in the line-width correlation for all lines.

4. Lag Measurements

4.1. Methodology

Most RM studies measure a reverberation-weighted "average" lag across the BLR between the continuum and broad-line light curves. This is mainly limited by the quality of the data (i.e., cadence, baseline, and S/N), relative to the intrinsic variability of the quasar. For high-quality RM light curves, it is possible to measure the 1D transfer function $\Psi(\tau)$ between the continuum and integrated line fluxes (e.g., Horne et al. 2004). If the spectroscopic data are of sufficient quality to reveal the velocity-resolved BLR responses, it becomes feasible to measure the 2D transfer function ($\Psi(\tau, v)$; or velocity-delay map) and constrain the dynamical structure of the BLR. The ultimate goal of RM is to measure high-fidelity velocity-delay maps and to constrain the geometry and kinematics of the BLR, along with a dynamical measurement of the black hole mass.

SDSS-RM is a pioneering program to measure BLR lags in a large sample of distant and luminous quasars. As such, the quality of the monitoring data is generally insufficient to robustly measure the 1D or 2D transfer functions in individual objects, except for a small number of objects with very large variability amplitudes. Rather, the main science goal of SDSS-RM is to measure an average lag for quasars in a luminosity–redshift regime uncharted by earlier RM programs. While an average lag does not inform us about the detailed geometry and kinematics of the BLR, it allows us to measure an approximate mass of the black hole, assuming the BLR is virialized. The uncertainty of the inferred RM mass in individual objects is dominated by the assumed geometric factor (the virial factor, see Section 4.5) and the uncertainties in the measured average lag and line width.

To measure this average BLR lag, RM studies traditionally deploy nonparametric (or with minimal parameterization) cross-correlation analysis methods, such as the interpolated cross-correlation function (ICCF; Gaskell & Peterson 1987), discrete correlation function (Edelson & Krolik 1988), or the z-transform discrete correlation function (Alexander 2013). Other nonparametric methods are also available to measure the average time delay between two light curves (see, e.g., Czerny et al. 2019a, for some examples). Alternatively, recent approaches such as `Javelin` (Zu et al. 2011) and `CREAM` (Starkey et al. 2016) use a stochastic variability model to interpolate the light curves and measure the lag using Markov Chain Monte Carlo (MCMC) to sample the posterior lag distribution. `PyROA` (Donnan et al. 2021) takes a slightly different approach: it performs a Bayesian analysis on the light curves with a running optimal average, from which the mean lag and the shape and width of the transfer function can be simultaneously constrained. The later lag measuring techniques are more reliable in recovering the lag and its uncertainties when the light-curve quality is insufficient for the other nonparametric methods to measure a significant lag (e.g., Li et al. 2019; Yu et al. 2020). Nevertheless, earlier correlation analysis methods are less model dependent, and provide important confirmation of the mean lag measurements. In particular, we found that the empirical ICCF correlation coefficient provides an efficient and intuitive metric on the successful detection of a lag. For this reason, information in the original ICCF will be used in determining the success of a lag measurement, even if the best lag is taken from more sophisticated approaches such as `Javelin` or `PyROA`.

Each pair of continuum and broad-line light curves is passed through a lag detection pipeline (wrapper) that collects publicly available packages (Zu et al. 2011; Sun et al. 2018; Donnan et al. 2021) to perform ICCF, `Javelin (v0.33)`, and `PyROA (v3.1.0)` analyses. The pipeline first flags and removes outliers 5σ away from a damped random walk (DRW)

model fit to the entire continuum light curve, using the package *celerite* (Foreman-Mackey et al. 2017). This outlier rejection is only performed for the continuum, given the numerous epochs combined from different surveys. The best-fit (taken as the median of the posterior distribution) DRW parameters are adopted (and fixed) in the subsequent `Javelin` analysis for self-consistency. We report these best-fit continuum DRW parameters in Table 2. The best-fit damping timescales $\tau_{\rm DRW}$ for the SDSS-RM sample have a median value of ∼270 days in the quasar rest frame, which is consistent with the findings from past studies using ∼10 yr long light curves (e.g., MacLeod et al. 2012; Suberlak et al. 2021; Stone et al. 2022). However, there is evidence that the damping timescale may be even longer if measured from light curves with much longer baselines (e.g., Stone et al. 2022).

For the `Javelin` run, we use MCMC parameters $n_{\rm chain} = n_{\rm walkers} = n_{\rm burn} = 200$ to provide enough sampling of the posterior over the lag search range. For `PyROA`, we use a similar amount of MCMC samples, and turn on the extra variance and delay distribution (adopting a Gaussian kernel) options in `PyROA`. The pipeline computes the relevant lag measurement properties in each method and determines the best-fit lag and its uncertainty following the rules specified in Section 4.2. The pipeline (wrapper) is publicly available at https://pypetal.readthedocs.io/en/latest/.

Our long-term baseline allows the exploration of lags up to as much as ∼11 yr in the observed frame (but the success rate will drop dramatically when approaching the maximum baseline), given the leading photometric light curves from PS1. However, quasars in the SDSS-RM sample have a wide range of observed-frame lags, from a few days to a few years. Searching the entire possible lag range enabled by the baseline is expensive (given the required MCMC samples) and unnecessary for objects with much shorter lags, and greatly increases the chance of aliases that become difficult to mitigate with our methodology described in Section 4.2. We therefore limit the lag search range on an object-by-object basis.

We first estimate the expected observed-frame lag, $\tau_{\rm obs,exp}$, based on the median continuum luminosity over the spectroscopic baseline, using the Hβ *R–L* relation from Bentz et al. (2013). The C IV lag is typically shorter than the Hβ lag, and the Hα and Mg II lags are slightly longer than that of Hβ based on past RM results (e.g., Peterson et al. 2004; Homayouni et al. 2020). We then limit the lag search range within $\pm 5 \times \tau_{\rm obs,exp}$, where negative lags refer to the line variability leading the continuum variability. However, we also require a minimum search range of ±250 days and a maximum search range of ±2500 days for each object. The latter maximum lag search range is defined to be roughly the same as the spectroscopic baseline. Importantly, as we are imposing a symmetric lag search range, we do not expect a systematic bias toward aliased (positive) lag values. In a small number of cases, the lag approaches 90% of the search range, and we rerun the lag detection with expanded search ranges to avoid introducing an artificial boundary in the measured lags.

There is a small fraction (∼4%) of objects for which the expected $\tau_{\rm obs,exp}$ exceeds 1000 days (Figure 2). For these objects, we relax the symmetry requirement on the lag search range, and utilize the full range of detectable lags given the leading years of photometry. The lag search range for this subset of quasars is [−2500, 4000] days in the observed frame.

4.2. Alias Mitigation and Detection of Lags

The standard ICCF, `Javelin` and `PyROA` analyses, do not take into account the number of LC pairs contributing to the correlation calculation at each lag. For typical SDSS-RM quality light curves, there are often a number of aliasing peaks in the lag posterior distribution that impact the measurement of the true lag. We mitigate the impact from aliasing peaks in the lag posterior following the scheme introduced in earlier work (Grier et al. 2017b, 2019). In short, we calculate a weight $P(\tau) = [N(\tau)/N(0)]^2$, where $N(\tau)$ is the number of shifted emission-line data points that overlap in the date ranges with the continuum, and $N(0)$ is the number of overlapping points at $\tau = 0$. Shifted emission-line data points falling into seasonal gaps are not counted. Since our continuum light curve encompasses the emission-line light curve by design, $N(0)$ is the maximum number of spectroscopic epochs. We then derive a final weight $P_f(\tau) = P(\tau) * {\rm ACF}(\tau)$, where ACF($\tau$) is the auto correlation function of the continuum LC, and "$*$" is the convolution operator. The weight $P(\tau)$ accounts for the reduction in the statistical constraint on the lag due to the fewer data points. While the exponent of 2 in the $P(\tau)$ weight is somewhat arbitrary, we find it results in optimal rejection of aliasing peaks (e.g., Grier et al. 2019). The convolution with the continuum autocorrelation function (ACF) accounts for the effect of seasonal gaps on the detection of certain lags. If the ACF declines rapidly, the annual seasonal gaps will have a significant effect on our detection because we are less likely to account correctly for the light-curve behavior during the gaps. On the other hand, if the ACF declines slowly away from zero lag, it implies it is straightforward to interpolate across the seasonal gaps as the variability is slow; in this case, the seasonal gaps are less likely to have an effect on our lag measurements. In the latter case, it is possible to measure a lag even if most of the shifted emission-line data points fall into seasonal gaps in the continuum light curves (see examples in, e.g., Shen et al. 2019a).

Once we have the posterior lag distribution (for `Javelin` or `PyROA`, this is the posterior lag distribution; for ICCF, this is the cross-correlation centroid distribution; CCCD), we weight the distribution with $P_f(\tau)$. The weighted lag distribution is then smoothed by a Gaussian kernel with a width of 15 days, and we identify the tallest peak within this smoothed distribution as the "primary" peak. We identify local minima in the distribution to either side of the peak and adopt these minima as the minimum and maximum lags to be included in our final lag calculation. We then return to the original *unweighted* posterior, reject all lag samples that lie outside the determined range, and use the remaining samples to calculate the final lag (median of the truncated distribution) and its 1σ uncertainties (16th and 84th percentiles of the distribution). Figure 6 demonstrates this procedure with one example of our targets (RM078). We did not detect an Hβ lag for this object using the 2014 data (Grier et al. 2017b), but the final SDSS-RM data, with multiyear coverage, are able to firmly detect a lag. In Appendix D, we show several additional examples that sample different lag detection qualities, and all lag diagnostic figures are available from our public server (see Appendix A).

Figure 7 shows the distribution of ICCF lag measurements in the lag versus $r_{\rm max}$ plane (e.g., Shen et al. 2016; Grier et al. 2017b, 2019; Homayouni et al. 2020), where $r_{\rm max}$ is the maximum Pearson correlation coefficient within ±1σ from the reported ICCF lag, which can be positive or negative. For all

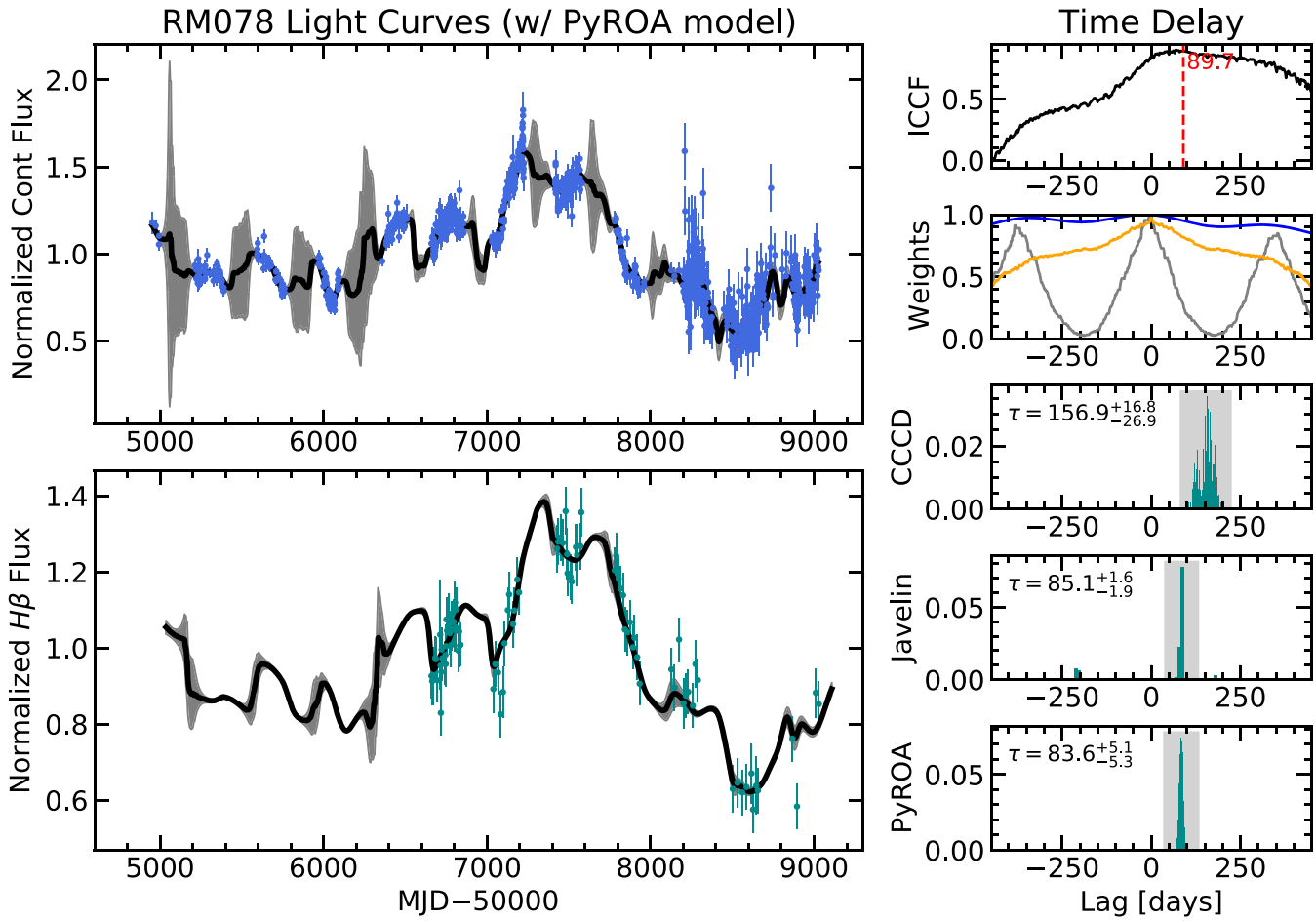

Figure 6. An illustration of lag measurements using ICCF, `Javelin`, and `PyROA`, as well as our alias mitigation scheme using lag posterior weights. The left two panels show the light-curve data in points, and the `PyROA` model light curves in solid lines and shaded area (best model and 1σ uncertainty). The right panels show the lag measurements. The "ICCF" panel shows the original ICCF, with the red vertical line (and number) indicating the expected (observed-frame) lag for Hβ. The "Weights" panel shows the ACF (orange), periodic overlapping light-curve data points $(N(\tau)/N_{\rm max})^2$ (gray), and the final weights (blue). The next three panels show the lag results for ICCF, `Javelin`, and `PyROA`, with the lag posterior displayed in the normalized histogram. The gray shaded band shows the boundaries of the identified primary peak, from which we measure the lag (median) and 1σ uncertainties (16th and 84th percentiles). The final lag and uncertainties are reported in each panel. For this particular object, the ICCF is relatively broad, resulting in a different lag using the cross-correlation centroid distribution (CCCD). However, the lag inferred from the cross-correlation peak distribution (CCPD; also see the top-right panel of this figure) is in excellent agreement with the `Javelin` and `PyROA` lags.

four lines, there is a clear asymmetry toward positive lags (i.e., line lagging continuum), especially when $r_{\rm max}$ is high. This figure demonstrates the statistical detection of RM lags for all four lines, given the symmetrical lag search ranges. When the light-curve quality is poor (or variability insignificant), the significance of correlation $r_{\rm max}$ is low, and there are more random incidents of positive and negative lags. At $r_{\rm max} > 0.4$, the overall negative-to-positive lag ratio within the plotted lag bounds is ∼0.04–0.17, indicating a false-positive rate (FPR) of ∼4%–17% in the reported (positive) lags. At $r_{\rm max} < 0.4$, although there are still more positive lags than negative ones, the FPR is substantially higher, and the measured positive lags there are of lower significance. Similar lag asymmetries are observed when we replace $r_{\rm max}$ with the line variability SNR2, which is expected as successful lag recoveries rely on well-detected line variability. However, there is no rigid boundary on SNR2 that would guarantee the detection or nondetection of a lag.

Figure 7 also dictates that using a positive lag search range in ICCF will not introduce a large confirmation bias if $r_{\rm max} > 0.4$. This may justify the usage of a positive lag search range for MOS-RM samples to reduce computational time. However, it is important to use the $r_{\rm max}$ value derived from the original ICCF to prefilter low-significance lag detections.

During our lag measurements with ICCF, `Javelin`, and `PyROA`, we discovered that many `Javelin` fits have lag posteriors that peak in the seasonal gaps. Because the ACF for SDSS-RM quasars is generally broad enough to allow the DRW model to interpolate into seasonal gaps, our weighting scheme is often unable to suppress these gap peaks. The situation is different from our previous analyses using shorter SDSS-RM light curves (e.g., Grier et al. 2017b, 2019; Homayouni et al. 2020). Now with our full multiyear baseline, it becomes more common for the `Javelin` posterior to be trapped in these seasonal gaps, resulting in spurious lag measurements.

In these failed `Javelin` fits, the favored model is where the line light curve is shifted into an annual gap, with minimal overlap with the continuum light curve in order to produce a good match between the two light curves. One particular reason for this gap alignment may be underestimated systematic uncertainties in the flux measurements that cause tension in matching up the continuum and line light curves. To remedy the situation, we deploy `PyROA`, which includes an excess

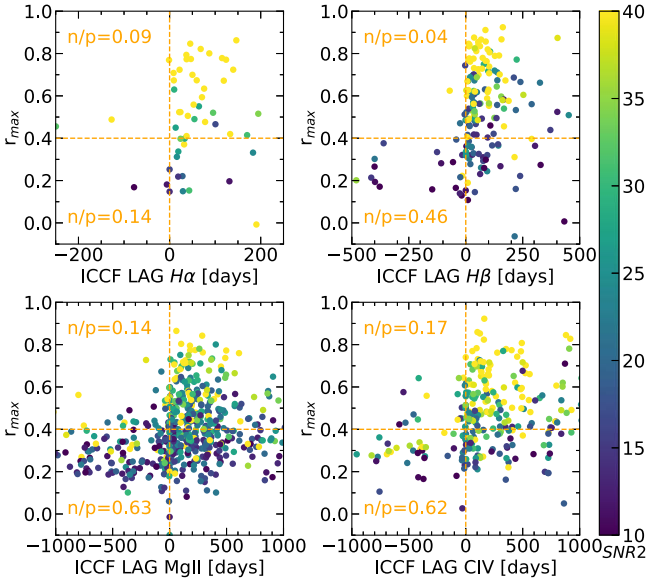

Figure 7. Statistical detection of lags for the SDSS-RM sample using ICCF. As the $r_{\rm max}$ value decreases, there are increased incidents of negative lags. While overall the asymmetry toward positive lags is clear, below $r_{\rm max} = 0.4$, the negative to positive lag ratio is substantially higher than that at $r_{\rm max} > 0.4$, indicating a high false-positive rate for the positive lags measured with low $r_{\rm max}$ values. The points are color coded by the line variability SNR2 (see Section 3.3). Objects with higher SNR2 values tend to cluster in the upper right quadrant, where a positive lag detection is more likely to happen, along with a high $r_{\rm max}$ value.

variance term to account for unknown systematic uncertainties in flux measurements. Like ICCF, PyROA is an empirical method that does not rely on a physical DRW model to interpolate the light curves. However, PyROA implements modules to account for the smearing of the line light curve due to the transfer function and for excess variance in flux measurements, with MCMC sampling to produce the lag posterior and rigorous lag uncertainties. Therefore, PyROA preserves the advantages of being model independent, while being less prone to underestimated flux uncertainties than Javelin.

We visually inspect the ICCF, Javelin, and PyROA results for each object, and conclude that indeed PyROA returns the most reasonable lag posterior and model light curves among all three methods. We therefore adopt the PyROA lag results as our fiducial results in the following analyses.

We now quantify the significance of individual lag measurements, and define a set of criteria to declare successful detection of a positive lag. These cases have $line$_LAG_DET >0 in Table 2. To do so, we rely on a hybrid set of parameters from the ICCF and PyROA results:

1. The PyROA lag must be positive at $\geqslant 2\sigma$ significance. Although a lower-significance detection might still be valid, the large uncertainties will render their measurements less useful, and could complicate our comparative analyses.
2. The original ICCF must reach Pearson correlation coefficient $r_{\rm max} > 0.4$ within $\pm 1\sigma$ of the reported PyROA lag.
3. Less than half of the lag posterior samples can be removed by our alias-removal procedure. If this procedure eliminates more than half of the samples, it indicates we lack a solid measurement of the lag as most of the samples are rejected. For PyROA lag posterior, this criterion is usually always satisfied.

As a final check, we visually inspect all detected lags (see Section 4.3) and confirm there are no peculiarities in the light curves (e.g., predominantly corrupted data) or in the measured lag posterior. The choice of a threshold of $r_{\rm max} > 0.4$ is somewhat arbitrary, but provides an efficient means to remove a large number of false positives. If we lower this threshold, we would recover more lags, but the contamination rate will also rise quickly. On the other hand, we caution that imposing more stringent selection criteria may introduce additional selection biases in the resulting lag sample, and artificially reduce the observed (intrinsic) scatter or bias the slope in the *R–L* relation.

In Appendix C, we show the results with more stringent quality cuts on detected lags, which are generally consistent with our fiducial results using the full sample. However, there are noticeable differences when we limit to the manually inspected subset with the highest grades. These highest-grade lags, while being the most reliable, are limited to well-detected lags much shorter than the baseline. Thus, using this subset of lags imposes severe selection biases in the *R–L* plane.

In Figure 8, we compare the detected lags for the four lines using the three methods. As expected, overall, the PyROA lags agree with ICCF lags, with the latter producing larger lag uncertainties. On the other hand, Javelin produces too many artificial lags within the seasonal gaps, although there is agreement between Javelin and PyROA for many objects. The underperformance of Javelin on large MOS-RM samples with multiyear light curves is unexpected.[44] Considering the flexibility of adding extra variance in the light curves with PyROA, and the excellent agreement between PyROA results and the empirical, model-independent ICCF results, we recommend PyROA as an efficient method to measure lags for MOS-RM data.

4.3. Manual Inspection of Lags

For the manual inspection procedure, we follow the convention of our earlier work (Grier et al. 2017b, 2019; Homayouni et al. 2020) to assign an integer grade (1–5; higher grades correspond to better quality) to each detected lag (see below). This manual inspection process is of course subjective, but nevertheless provides a useful means to select the most reliable lags. Subsequently, we consider the subset with lag grade = 4 or 5 the "Gold" sample. These lag grades are included in Table 2.

For a detailed manual inspection, the grading criteria are as follows:

1. *Spectral variability.* We use the rms spectrum to ensure that the emission line of interest is variable over 7 yr of observation. Therefore, we inspect the 7 yr rms spectrum to ensure that the line variability signal is present, and is not overlapping with the edges of the spectrum.
2. *Light-curve variability.* The continuum and line light curves demonstrate sufficient variability, and features that deviate from a smooth trend, suggesting underlying dynamic processes.

[44] We tested limited cases where we inflate the flux uncertainties in the Javelin run, and find that the Javelin lags are still incorrectly placed in seasonal gaps. We plan to further investigate this issue in future work.

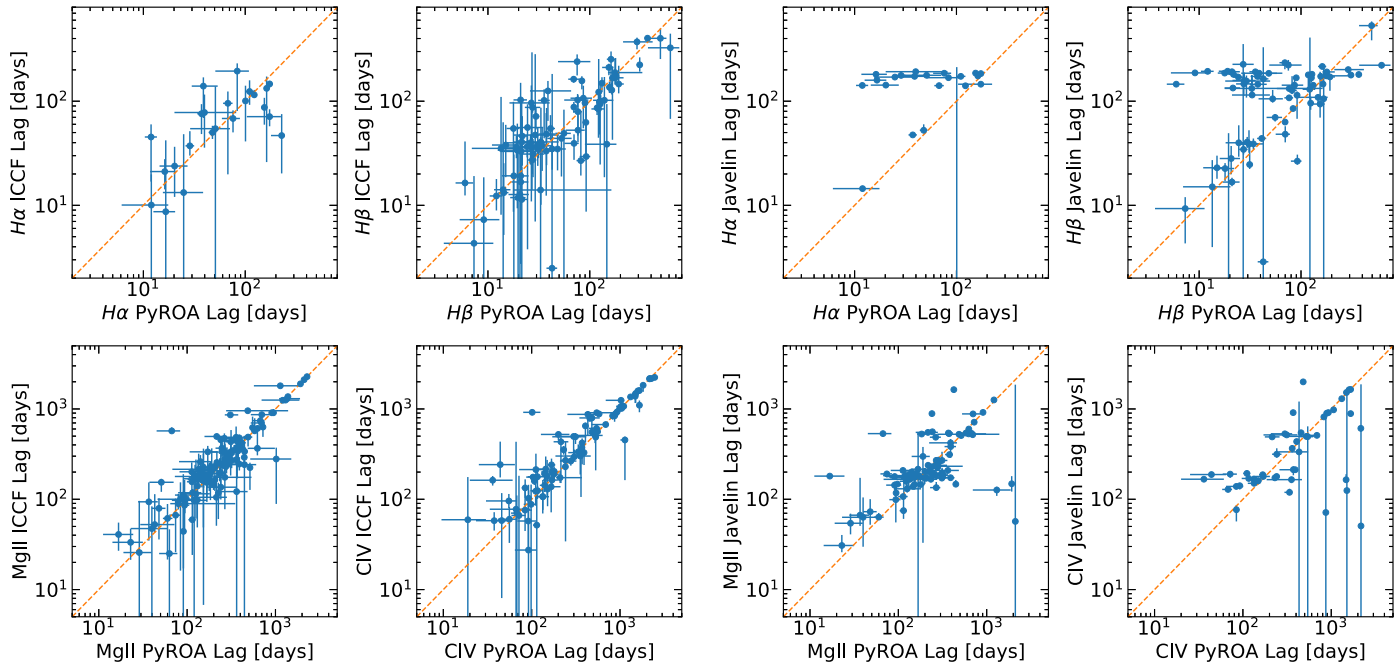

Figure 8. Comparisons of the detected lags between two different methods (ICCF vs. `PyROA` at left and `Javelin` vs. `PyROA` at right). In most cases, our fiducial `PyROA` lags are consistent with the ICCF lags. However, for our multiyear light-curve data, the `Javelin` method often traps the lag in the seasonal gaps, leading to artificial clustering of lags there.

3. *Lag posterior peak.* As an additional check on the quality of the measured lags, we require the PyROA lag probability density function (PDF) has a well-defined primary peak away from the lag search range boundaries.
4. *Lag validation.* We inspect the superposition of the shifted line light curve onto the continuum light curve to validate the measured lag against the continuum and line light curves. This is particularly important due to the seasonal gaps in SDSS-RM observations. This could coincide with lags of approximately 180, 540, etc. days, which could introduce false-positive lag detections in our 7 yr measurements. This validation process consists of two steps:
   (a) Good alignment between the continuum and line light curves suggests that the measured lag represents a reliable estimate of the underlying light travel time delay, characteristic of physical RM lag.
   (b) If the measured lag coincides with the seasonal gap, indicating fewer overlapping data points between the continuum and the shifted line light curve, then we examine to see if the overall trends between the two light curves and correlated variations match a reverberation picture.

   In either case, if the superimposed light curves show excessive overlap due to high data variance, we consider the light curves to be an inadequate match for reliable lag determination.

The visual inspection procedure introduces complex selection functions on the resulting best-quality lags. As a consequence, we caution on the danger of removing certain lags using visual inspection from the sample in measuring the *R–L* relation. We do not recommend on using the "Gold" lag sample alone or in combination with other heterogeneous lag samples to measure the *R–L* relation. On the other hand, studies that are not affected by lag selection effects, such as the investigation of host galaxy properties, can utilize this "Gold" sample for best-measured RM masses.

4.4. False-positive Rate for Individual Lags

Using `Javelin` or `PyROA` to estimate the FPR for individual lags would be prohibitive, given the large number of lag detections in the SDSS-RM sample and the computational demands of running `Javelin` or `PyROA` for the full light-curve sets. However, Figure 7 demonstrates that using the much faster ICCF method to estimate the FPR is a reasonable approach. Here, we pair the observed line light curve with randomly generated continuum light curves for mock lag measurements. To capture the variability characteristics of individual objects, we use the best-fit DRW parameters for each object to generate random continuum light curves.

For each detected positive lag, we generate 100 mock continuum light curves, and perform the ICCF analysis on each realization. If a mock lag detection from the CCCD has $r_{\rm max} > 0.4$ (within 1σ range of the CCCD median) and recovers a positive lag at $\geqslant 2\sigma$, we count it as a false positive. The fraction of such incidents among the 100 mock realizations is recorded as the FPR for the observed lag.

For the detected lags (i.e., $line$_LAG_DET >0 in Table 2), the median FPR for each line ranges from 6% to 14%, which is in line with the expected rates using our $r_{\rm max}$ cut (Section 4.2). There is a small fraction of detections (0%, 5%, 5%, and 15% for Hα, Hβ, Mg II, and C IV) for which the estimated FPR > 0.3. However, visually inspecting these cases, the light curves are often significantly variable, and the lags are robustly detected. Some examples are RM078 (Hβ lag), RM101 (Hα and Hβ lags), RM578 (Mg II), RM256 (C IV), etc. We argue that the FPR has been significantly overestimated in these cases for the following reasons:

(i) The dominant variability feature in the light curve is a broad bump or dip. Given the typical DRW damping

timescale, the mock continuum light curve often contains multiple peaks/dips over the 11 yr baseline, leading to aliasing lags in the cross-correlation, some of which will be identified as positive detections. In that sense, even for some of the best-measured lags in the local RM AGN sample, this FPR estimation approach would produce a nonnegligible fraction of false positives, if we use simulated continuum light curves that are much longer than the measured lag.

(ii) Our lag search range is generous (e.g., ±5 times the expected lag). When combined with the long light-curve baseline, we will often recover an alias lag with the mock continuum light curve, boosting the FPR. In other words, the prior information on the anticipated lag is not used in this FPR estimation.

(iii) Using ICCF in the FPR estimation is efficient, but not as robust as using `PyROA` in measuring the lag, which may also tend to overestimate the FPR.

Given these reasons, and our visual inspection of detected lags with high FPR estimates, we conclude that the individual FPR estimates should not be used as the sole metric to assess the robustness of individual lag measurements. Some high-fidelity lag detections (e.g., Figure 6) may have a high FPR. We do not recommend using the reported FPR to select a refined lag sample, unless the lag measurement quality (e.g., $r_{\rm max}$ and lag errors) and diagnostic plots are checked simultaneously.

Nevertheless, the most important utility of our FPR test is to ensure consistency. The statistical behaviors of our FPR estimates confirm that our reported lags have a low overall FPR of ∼10%, in concordance with our estimates in Section 4.2 based on the positive/negative lag asymmetry (e.g., Figure 7).

4.5. The Virial Coefficient f

To scale the measured virial product, $\mathrm{VP} \equiv c\tau_{\rm rest}\Delta V^2/G$, to the RM BH mass, we need to multiply by the virial coefficient f. The exact value of f depends on the choice of velocity indicator ΔV, which can be the FWHM or line dispersion ($\sigma_{\rm line}$) measured either from the mean spectrum or from the rms spectrum (e.g., Collin et al. 2006; Wang et al. 2020b). Throughout this paper, we focus on the line dispersion measured from the rms spectrum, $\sigma_{\rm line,rms}$, as commonly adopted in RM studies. Given a specified line width, this scale factor depends on the geometry and kinematics of the BLR gas, and can differ significantly from object to object. For example, if the BLR gas is primarily distributed in a flattened structure as suggested in earlier observations of broad Balmer lines (e.g., Wills & Browne 1986; Shen & Ho 2014; Mejía-Restrepo et al. 2018), then more face-on BLRs require larger f-factors than more edge-on BLRs. The situations may be more complicated for realistic kinematics of the BLR gas that could involve nonvirial motions, e.g., for the high-ionization C IV line (e.g., Proga et al. 2000; Waters et al. 2016).

For traditional RM work that utilizes the average lag, the virial coefficient f is usually determined by comparing the virial products to the expected BH masses from the local $M_{\rm BH}$–σ_* relation (e.g., Onken et al. 2004; Graham et al. 2011; Park et al. 2012; Woo et al. 2013). These calibrations can only provide an average viral coefficient $\langle f \rangle$ for the sample (because the $M_{\rm BH}$–σ_* relation itself has significant intrinsic scatter), varying from $\langle f \rangle = 2.8$ (Graham et al. 2011) to $\langle f \rangle = 5.9$ (Woo et al. 2013) if using $\sigma_{\rm line,rms}$ as the velocity indicator. Because the best-fit local $M_{\rm BH}$–σ_* relation has been updated over the years (see discussions in, e.g.,. Kormendy & Ho 2013), the calibration of the virial coefficient has also evolved. Moreover, because the $M_{\rm BH}$–σ_* relation depends on galaxy bulge properties, in principle, one should use different virial coefficients based on the bulge classification of the host galaxy (e.g., Ho & Kim 2014)—such bulge classifications, however, are usually difficult for broad-line AGNs, and their accuracy largely depends on the expertise of the classifier.

Recent progress on dynamical modeling of the BLR with RM data has made it possible to derive the individual virial coefficient for AGNs (e.g., Pancoast et al. 2014; Grier et al. 2017a; Li et al. 2018, 2022; Williams et al. 2018, 2020; Bentz et al. 2021, 2022; Villafaña et al. 2022). By dynamically modeling the BLR and constraining the model parameters with high-quality RM data, this approach can derive the underlying BH mass along with the geometrical and dynamical properties of the BLR. Comparing the dynamical BH mass with the average-lag-based virial product, one can derive the required virial coefficient for individual systems, rather than an average virial coefficient from the $M_{\rm BH}$–σ_* relation. This approach, while still dependent on the details of the dynamical modeling, eliminates the systematics and limitations in the estimated average f-factor with BH–host relations.

Based on 16 local RM AGNs with available dynamical-modeling BH masses, Williams et al. (2018) found an average $\langle \log f \rangle = 0.57$ dex with a dispersion of 0.19 dex, which is consistent (within 1σ) with the average f-factor derived from the $M_{\rm BH}$–σ_* relation. Several studies (e.g., Pancoast et al. 2014; Williams et al. 2018) also studied the correlations between the individual f-factor and AGN properties and found no significant trend (but see Linzer et al. 2022; Villafaña et al. 2023, for the latest results on potential correlations). However, the individual f-factor strongly depends on the inclination of the BLR derived from dynamical modeling, where the BLR gas is often distributed in a thick-disk geometry.

Here, we compile the latest sample of RM AGNs with dynamical-modeling BH masses from the literature. This sample contains 30 objects (NGC 5548 has two independent measurements), and the details are summarized in Table 3. We only compile the virial product computed using the line dispersion measured from the rms spectrum, since this is the most commonly adopted line width in RM work. Figure 9 (top) shows the distribution of the individual f-factors ($\log f(\sigma_{\rm line,rms})$) in the sample, which is peaked around $\log f \approx 0.6$.

To quantify the mean and intrinsic dispersion in $\log f(\sigma_{\rm line,rms})$, we follow the approach in, e.g., Pancoast et al. (2014). We assume the intrinsic $\log f$ distribution is a Gaussian distribution with mean value $\mu_{\rm logf}$ and dispersion $\sigma_{\rm logf}$. The posterior probability distributions of parameters $\mu_{\rm logf}$ and $\sigma_{\rm logf}$, given the sample of measured $\log f$, are calculated following Appendix B in Pancoast et al. (2014) that takes into account the posterior PDF of individual $\log f$ measurements. Because these dynamical-modeling papers typically do not publish the full PDF of $\log f$, we simply adopt symmetric Gaussian errors for these $\log f$ measurements. The constrained marginalized PDFs for $\mu_{\rm logf}$ and $\sigma_{\rm logf}$ are shown as black and red lines in Figure 9 (bottom), where the peak and 16th/84th percentile of the distribution are marked. The constrained mean $\log f$ is $\mu_{\rm logf} = 0.62 \pm 0.08$, and the constrained intrinsic dispersion $\sigma_{\rm logf} = 0.32^{+0.08}_{-0.06}$, where uncertainties are 1σ.

Table 3
Compiled VPs and Dynamical-modeling BH Masses from Previous RM Studies

Name (1)	logVP (2)	$\sigma_{\log VP}$ (3)	VP References (4)	$\log M_{\rm dyn}$ (5)	$\sigma_{\log M_{\rm dyn}}$ (6)	$M_{\rm dyn}$ References (7)	$\log f$ (8)	$\sigma_{\log f}$ (9)
NGC4151	6.70	0.04	DeRosa18	7.22	0.11	Bentz22	0.52	0.11
PG2209+184	6.81	0.09	U22	7.53	0.20	Villafana22	0.72	0.21
RBS1917	6.50	0.17	U22	7.04	0.29	Villafana22	0.54	0.34
MCG+04-22-042	6.53	0.07	U22	7.59	0.35	Villafana22	1.06	0.36
NPM1G+27.0587	6.71	0.27	U22	7.64	0.38	Villafana22	0.93	0.47
Mrk1392	7.15	0.06	U22	8.16	0.12	Villafana22	1.01	0.13
RBS1303	6.75	0.09	U22	6.79	0.15	Villafana22	0.04	0.17
Mrk1048	6.69	0.55	U22	7.79	0.46	Villafana22	1.10	0.72
RXJ2044.0+2833	6.43	0.07	U22	7.09	0.17	Villafana22	0.66	0.18
Mrk841	7.02	0.19	U22	7.62	0.40	Villafana22	0.60	0.44
Mrk335	6.66	0.05	Grier12	7.25	0.10	Grier17	0.59	0.11
Mrk1501	7.52	0.06	Grier12	7.86	0.18	Grier17	0.34	0.20
3C120	7.09	0.04	Grier12	7.84	0.17	Grier17	0.75	0.17
PG2130+099	6.92	0.04	Grier12	6.92	0.23	Grier17	0.00	0.24
Mrk50	6.78	0.07	Barth11	7.50	0.21	Williams18	0.72	0.23
Mrk141	6.76	0.38	Williams18; Barth15	7.46	0.18	Williams18	0.70	0.42
Mrk279	6.80	0.12	Peterson04	7.58	0.08	Williams18	0.78	0.14
Mrk1511	6.48	0.08	Barth13; Barth15	7.11	0.18	Williams18	0.63	0.20
NGC4593/Mrk1330	6.24	0.14	Barth13; Barth15	6.65	0.21	Williams18	0.41	0.25
Zw229–015	6.28	0.09	Barth11b	6.94	0.14	Williams18	0.66	0.17
PG1310–108[b]	6.74	0.30	Busch14	6.48	0.20	Williams18	−0.26	0.36
Mrk142	5.49	0.17	Bentz09b	6.23	0.35	Li18	0.74	0.39
NGC5548	7.17	0.14	Pei17	7.54	0.29	Williams20	0.37	0.32
Arp151/Mrk40[a]	⋯	⋯	Pancoast14	6.62	0.12	Pancoast14	0.51	0.12
Mrk1310[a]	⋯	⋯	Pancoast14	7.42	0.27	Pancoast14	1.63	0.27
NGC5548[a]	⋯	⋯	Pancoast14	7.51	0.18	Pancoast14	0.42	0.18
NGC6814[a]	⋯	⋯	Pancoast14	6.42	0.21	Pancoast14	−0.14	0.21
SBS1116+583A[a]	⋯	⋯	Pancoast14	6.99	0.28	Pancoast14	0.96	0.28
NGC3783	6.69	0.08	Bentz21	7.51	0.20	Bentz21	0.82	0.21
3C273	7.54	0.04	Zhang19	9.06	0.24	Li22	1.52	0.24

Notes. For simplicity, we have symmetrized all uncertainties in order to use these uncertainties in our PDF calculation (Section 4.5). Reference keys are: Barth11b (Barth et al. 2011), Barth13 (Barth et al. 2013), Barth15 (Barth et al. 2015), Bentz09b (Bentz et al. 2009), Bentz21 (Bentz et al. 2021), Bentz22 (Bentz et al. 2022), Busch14 (Busch et al. 2014), DeRosa18 (De Rosa et al. 2018), Grier12 (Grier et al. 2012), Grier17 (Grier et al. 2017a), Li18 (Li et al. 2018), Li22 (Li et al. 2022), Pancoast14 (Pancoast et al. 2014), Pei17 (Pei et al. 2017), Peterson04 (Peterson et al. 2004), U22 (U et al. 2022), Villafana22 (Villafaña et al. 2022), Williams18 (Williams et al. 2018), Williams20 (Williams et al. 2020), and Zhang19 (Zhang et al. 2019).

[a] $\log f$ values directly taken from Pancoast14, which did not include errors in logVP.

[b] VP in Busch14 was computed using $\sigma_{\rm line}$ from the mean spectrum and an estimated lag from the mean R–L relation (Bentz et al. 2013). Note that NGC 5548 has two independent $\log f$ measurements, both of which are included in our analysis.

Alternatively, given the set of $\log f$ measurements and their uncertainties, we can calculate the mean and intrinsic dispersion in $\log f$ using an ML estimator described in Shen et al. (2019b). This ML approach also assumes that the intrinsic $\log f$ values are Gaussian distributed around mean value $\mu_{\rm ML}$ with a dispersion of $\sigma_{\rm ML}$. The results are $\mu_{\rm ML} = 0.62 \pm 0.07$, and $\sigma_{\rm ML} = 0.31 \pm 0.07$, which are nearly identical to the results using the PDF approach above.

Thus, we conclude, based on a sample of 30 independent measurements of $\log f$ ($\sigma_{\rm line,rms}$) with both VPs and dynamical-modeling BH masses, that the mean logarithmic virial factor is $\langle \log f \rangle = 0.62 \pm 0.07$. This value is consistent with most of the recent determinations of the virial factors either using BH–host scaling relations or using dynamical masses from RM (Onken et al. 2004; Graham et al. 2011; Park et al. 2012; Woo et al. 2013; Ho & Kim 2014; Pancoast et al. 2014; Grier et al. 2017a; Williams et al. 2018). Importantly, this literature sample compiled in Table 3 allows us to robustly constrain the intrinsic dispersion of the individual virial factor to be ∼0.3 dex (or a factor of 2 in f). This dispersion in individual f factors sets a typical systematic uncertainty of RM-based BH masses derived using a single, average f factor. This intrinsic scatter in $\log f$ is notably larger than (but consistent within $\sim 2\sigma$) the value of $\sigma_{\log f} = 0.14 \pm 0.10$ reported in Williams et al. (2018) based on 16 objects with RM data and dynamical masses. The additional objects in our sample require a larger $\sigma_{\log f}$ to account for their dispersion.

In our analysis, we have only focused on the f-factor based on $\sigma_{\rm line,rms}$, and ignored potential dependencies of the f-factor on AGN properties. Villafaña et al. (2023) used a similar AGN sample with dynamical BH masses to investigate correlations between the virial factors based on different line-width definitions and AGN properties, for Hβ only. Such correlations, if confirmed with future larger samples with dynamical BH masses, will further shed light on the structure of AGN BLRs and improve BH mass estimation using average RM lags.

In the following discussion, we adopt the same average f-factor for all four broad lines. As shown in Section 5.1, the overall agreement between the RM BH masses from two different lines indicates adopting the same f-factor for all lines is a reasonable approach. However, we acknowledge that the SDSS-RM sample is different from the local RM AGN sample

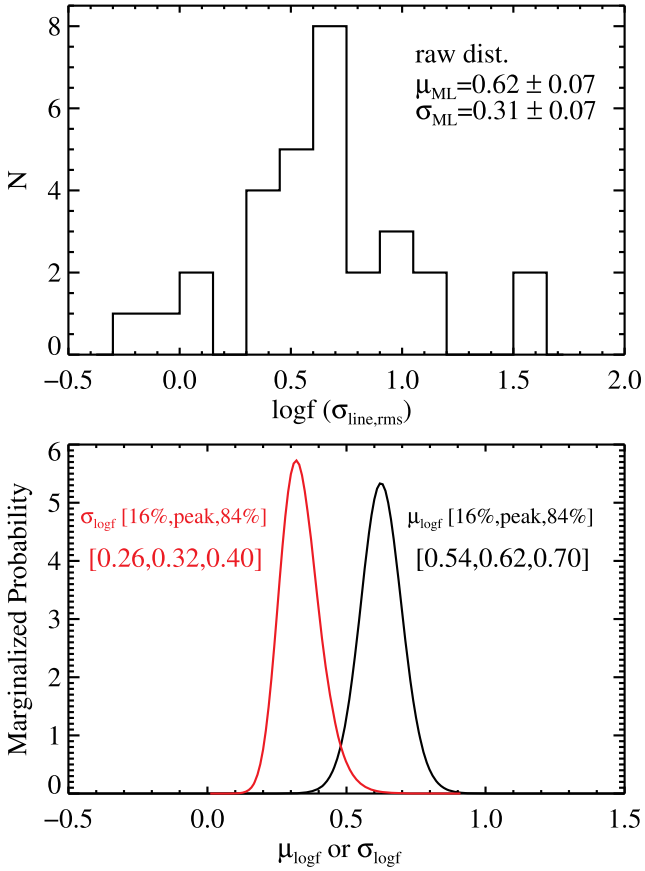

Figure 9. Constraints on the mean and dispersion of the distribution of individual f factors based on the line dispersion measured from the rms spectrum, using a sample of 30 objects with dynamical-modeling and RM BH masses as compiled in Table 3. Top: raw distribution of individual f factors. The ML estimates of the sample mean and intrinsic 1σ dispersion are marked. Bottom: Marginalized probability distributions of the mean and intrinsic 1σ dispersion of f factors. The 16%, peak, and 84% locations of the distributions are marked.

in terms of luminosity and BH mass. Whether or not the average f-factors differ among different samples is a topic for future studies.

4.6. Lags and RM Masses

Our fiducial RM masses are computed from the measured broad-line lags combined with the line dispersion measured from the rms spectrum of the monitoring data. In mathematical form,

$$M_{\rm RM} \equiv \frac{\langle f \rangle c\tau_{\rm rest}\sigma^2_{\rm line,rms}}{G}, \quad (1)$$

where $\langle f \rangle$ is the geometric mean f factor when using $\sigma_{\rm line,rms}$ for the line width.

The PrepSpec $\sigma_{\rm line,rms}$ measurements are based on the continuum-subtracted rms spectrum, which corresponds to the true line-only variability. Some earlier approaches (e.g., Peterson et al. 2004) generate the rms spectrum from the (total) emission-line+continuum spectra and measure the line dispersion by removing a local continuum from the rms spectrum. The latter approach can produce line dispersions that are on average biased low by ∼20%–30% if the monitoring period is not much longer than the lag (e.g., by a factor of $\lesssim$3, Barth et al. 2015; Wang et al. 2019). Fortunately, this

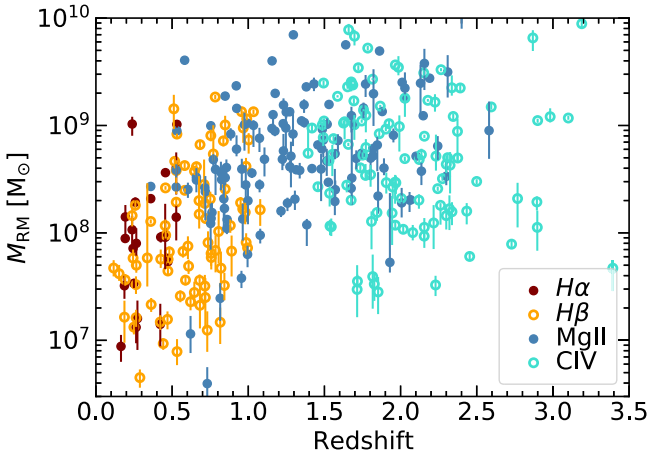

Figure 10. Redshift distribution of RM masses derived from Hα, Hβ, Mg II, and C IV lags for the SDSS-RM sample. Error bars are measurement uncertainties only and do not include the ∼0.3 dex systematic uncertainty in $M_{\rm RM}$.

monitoring duration criterion is satisfied in most of the recent low-z RM campaigns listed in Table 3, and many of these campaigns (such as the LAMP project) measured $\sigma_{\rm rms}$ from the properly constructed line-only rms spectrum. Therefore, we can safely apply the mean f factor derived in Section 4.5 to the SDSS-RM sample.

We compile our lags and RM masses in Table 2. In Figure 10, we show the distribution of RM masses and redshift for the SDSS-RM sample. With ∼300 unique objects, this sample represents the largest sample of quasars with direct RM-based BH masses up to $z \sim 3.5$.

5. Discussion

5.1. RM Results

For an MOS-RM program like ours that targets a flux-limited quasar sample across broad redshift and luminosity ranges, the common baseline of the program may impose significant selection biases in the measured lags, missing long lags that cannot be measured, given the length of the light curves. Fortunately for the SDSS-RM program, the final light curves span an effective baseline of 11 yr, sufficient to recover the vast majority of expected lags for the SDSS-RM sample (see Figure 2).

Figure 11 shows our measured lags for the four lines in their respective R–L planes, where the luminosity is taken as the usual local continuum luminosity corresponding to each line and is the median luminosity over the spectroscopic baseline. For $L5100$, we have corrected for host contamination, using the estimated host fraction (within the SDSS fiber) from the 2014 spectroscopy (Shen et al. 2015b). When fitting an R–L relation to these data, we only use the SDSS-RM sample, rather than combining the SDSS-RM sample with other literature samples. This is because by number the SDSS-RM sample over-represents the population over the probed luminosity range, and would likely dominate the global fit. A joint fit with other lag samples probing different luminosity ranges will be the objective of future work, after selection effects are properly taken into account.

For Hβ and Mg II, the luminosity range in the SDSS-RM lag sample spans ∼2 dex, and an R–L relation is clearly present. Overall, the SDSS-RM lag sample follows the reported global

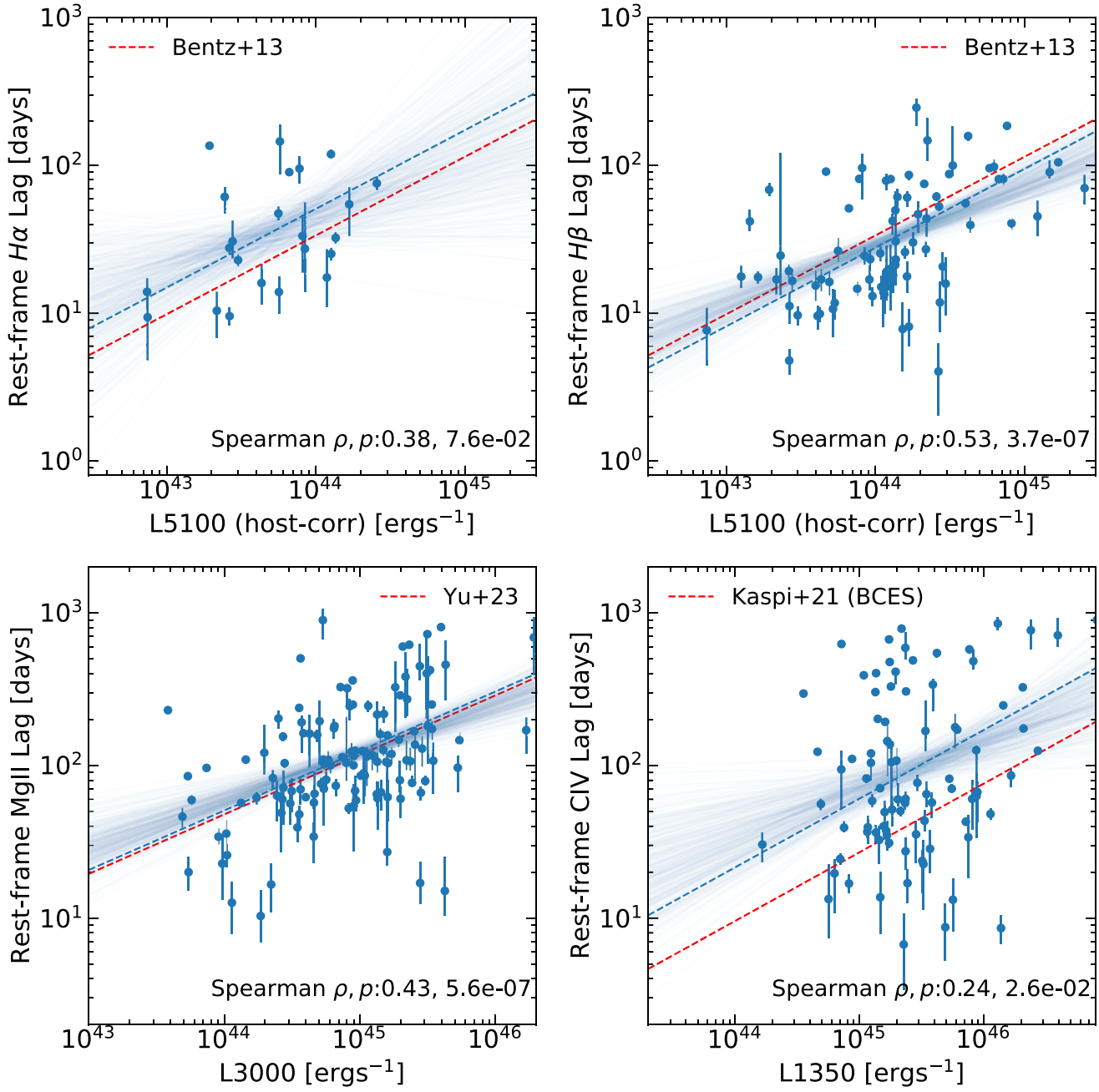

Figure 11. *R–L* relations for the four broad lines in this work. In each panel, the red dashed line shows the most recent literature relation (Bentz et al. 2013; Kaspi et al. 2021; Yu et al. 2023). The light blue lines are random draws from the linear regression of predicting $\log\tau$ with $\log L$ using the Bayesian approach in Kelly (2007). The blue dashed line is a forced regression fit with slope fixed to that measured in earlier work. Overall, measured SDSS-RM lags follow these earlier relations, but there are notable deviations (in particular for C IV). Our lag measurements suggest that the Hβ and Mg II *R–L* relations are reasonably tight (with an intrinsic scatter of ∼0.3 dex in lag), while the C IV *R–L* relation has substantially larger scatter (∼0.5 dex) than the other two lines. The Spearman rank-order test results on the SDSS-RM sample (correlation coefficient ρ and p-value p) are marked in each panel. These results are discussed in detail in Section 5.1 and Table 4.

R–L relation in the latest studies that derived such relations using heterogeneous samples (Bentz et al. 2013; Yu et al. 2023). However, the Hβ lags for SDSS-RM quasars on average fall slightly below the local relation by ∼0.1 dex, if we fix the slope to that in the local relation in the regression. A similar, albeit slightly larger, average offset in Hβ lags was reported in earlier SDSS-RM results based on shorter baselines (Grier et al. 2017b; Fonseca Alvarez et al. 2020). Because the baseline is limited in earlier SDSS-RM lag measurements, this average offset from the local *R–L* relation might have been overestimated in our earlier papers. Note that the host correction in Shen et al. (2015b) is on average smaller by ∼30%–40% than that estimated from image decomposition (Li et al. 2021, 2023). If we further reduce the 5100 Å luminosities by ∼0.1 dex, the lag offset in the *R–L* relation would become negligible. The intrinsic scatter in our Hβ lags around the mean *R–L* relation is ∼0.32 dex.

The dispersion around the canonical Bentz et al. (2013) *R–L* relation for Hβ has gained substantial interest in the past few years (e.g., Du et al. 2016; Czerny et al. 2019b; Du & Wang 2019; Martínez-Aldama et al. 2019; Dalla Bontà et al. 2020). It is suggested (e.g., Czerny et al. 2019b; Du & Wang 2019; Fonseca Alvarez et al. 2020) that this dispersion is due to different SEDs between the local RM AGN sample in

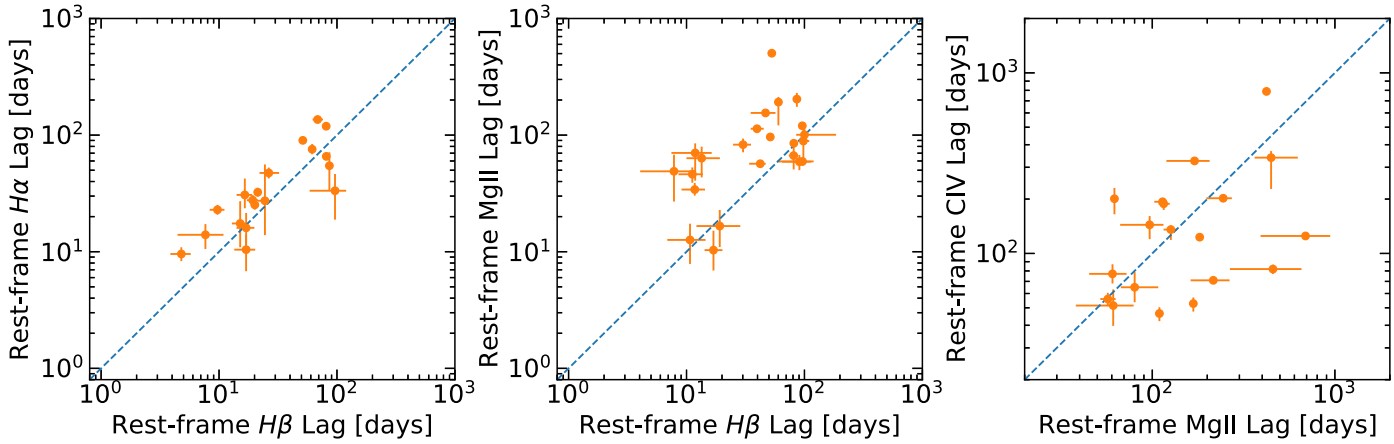

Figure 12. Comparisons between lags from two different lines for the same object. Hα and Mg II lags are typically longer than Hβ lags, while C IV lags are somewhat shorter (but the sample statistics are limited, and the scatter is large) than Mg II lags.

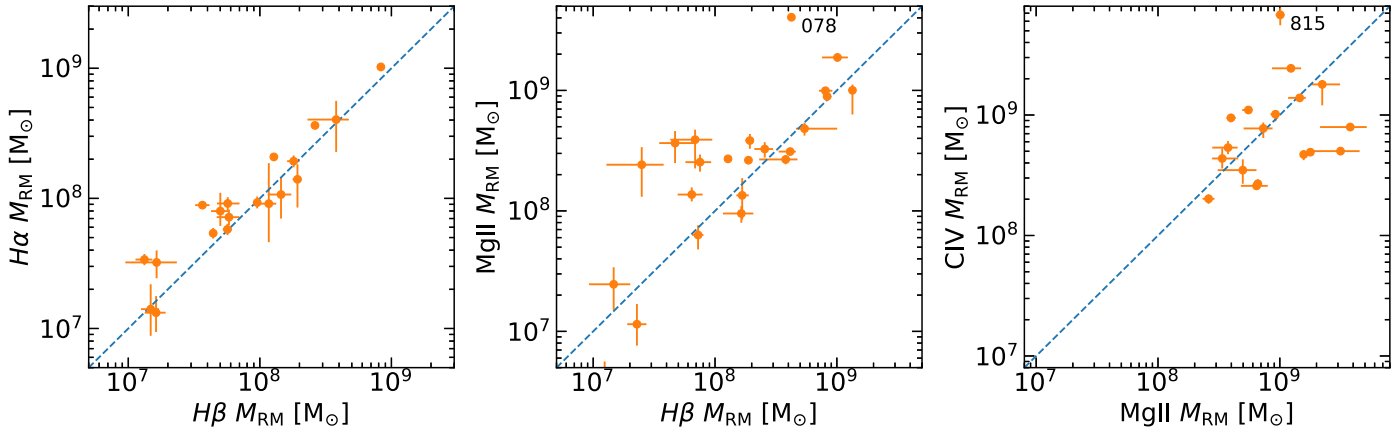

Figure 13. Comparisons between $M_{\rm RM}$ from two lines for the same object. For Mg II, we have corrected the rms line dispersion for the velocity split $V \approx 750$ km s^{-1} of the doublet ($\sigma_{\rm true} = \sqrt{\sigma^2_{\rm measure} - (V/2)^2}$). There is general agreement between the RM masses measured from two different lines, validating the RM technique of measuring BH masses. The two most significant outliers in the line comparisons are marked with their RMIDs.

Bentz et al. (2013) and other samples. In particular, the SEAMBH collaboration has targeted local AGNs with high-accretion rates, and found that the lag deviation from the canonical Bentz et al. (2013) *R–L* relation can reach as much as a factor of a few (e.g., Du et al. 2016; Du & Wang 2019). Different accretion rates of the AGN lead to different ratios between the ionizing continuum (directly responsible for the lag) and the local continuum near the line, which in turn leads to an offset from the mean *R–L* relation. From a theoretical point of view, it is expected that there is a dispersion in lags at fixed optical luminosity caused by the diversity in BH mass and accretion rate and subsequently the underlying SED (Q. Wu et al. 2024, in preparation).

For Mg II, we found that there is a strong *R–L* relation over the luminosity range probed by our lag sample (Figure 11). The slope is somewhat shallower than 0.5. Combining the OzDES-RM results on Mg II lags and literature Mg II lag measurements, Yu et al. (2023) measure a global R–L_{3000} relation for Mg II, which is consistent with our measurements. The intrinsic scatter in the Mg II *R–L* relation is $\sim$0.32 dex.

The luminosity range for the SDSS-RM C IV lag sample is limited to $\sim$1 dex. Thus, similar to Grier et al. (2019), we cannot reliably constrain an *R–L* relation using the SDSS-RM sample alone. Kaspi et al. (2021) recently compiled literature C IV lag results that span $\sim$8 orders of magnitude in luminosity and derived a global *R–L* relation for C IV. The SDSS-RM lags generally fall on this relation, with a slight offset to longer lags. It is possible that we were able to recover these longer lags given the baseline of our light curves. But more importantly, within the luminosity range probed by the SDSS-RM C IV lag sample, the scatter around a mean *R–L* relation is $\sim$0.5 dex, substantially larger than those seen for Hβ and Mg II. This large scatter argues against using the SE C IV BH mass estimator, which assumes there is a tight *R–L* relation for C IV. Since C IV is a high-ionization line, the diversity in the SED due to accretion parameters could be the main driver for the large scatter of lags around the mean *R–L* relation, where we are using the local continuum L_{1350} near C IV as the luminosity indicator.

Given the large sample size of SDSS-RM, there are subsets of quasars for which we successfully measured lags for two lines. Figure 12 compares the lags measured from two different lines in the same quasar. Compared with Hβ lags, Hα and Mg II lags are on average longer, while C IV lags are somewhat shorter than Mg II lags overall (but the sample size is small). These results are consistent with earlier RM results for local AGNs (Peterson et al. 2004), where BLR stratification and radiative transfer effects are invoked to explain these lag differences (e.g., Goad et al. 2012; Guo et al. 2020).

We also compare the RM masses from two different lines in the same quasar in Figure 13. The RM masses between two

lines are generally consistent with each other, suggesting that the line dispersion measurements are adequate. In addition, this consistency implies using the same virial factor (derived from Hβ) for all lines does not introduce significant mass offsets in the average sense. There are, however, individual objects for which the RM masses from two lines differ significantly. We examined the lag and line-width measurements for these objects, and found overall these measurements are reasonable.[45] It is possible that these objects have different BLR kinematics for different lines, and their respective virial factors differ (e.g., Runnoe et al. 2013; Fries et al. 2023). Either way, our results highlight the importance of deriving velocity-resolved RM results for Mg II and C IV to fully understand the BLR structure and kinematics for these lines.

5.2. Comparison with Earlier SDSS-RM Results

Our final lag results are generally consistent with earlier SDSS-RM lag measurements based on shorter baselines (Shen et al. 2016; Grier et al. 2017b, 2019; Homayouni et al. 2020). However, over the multiyear baseline, some SDSS-RM quasars reported in earlier work have undergone significant luminosity changes, resulting in changes in lags as well. One notable improvement of our final light curves is that we are able to recover many additional lags that were missed in earlier studies (e.g., Figure 6). Those lags require either longer baselines or the capture of significant variability features for their successful measurements. In terms of RM masses, some quasars have significantly revised values compared with their earlier SDSS-RM results, due to changes in the rms line-width measurements (i.e., PrepSpec has been run on different light-curve baselines), as well as changes in the lag. While there is no overall concern on these final RM masses upon our random inspections, we will perform a more thorough investigation on individual cases and update their RM masses if necessary in the near future.

On the other hand, the much longer final baselines also led to more aliases in the lag measurements for some objects. One such example is the C IV lag in RM231. In previous work with 4 yr SDSS-RM data, Grier et al. (2019) was able to measure an observed-frame C IV lag around 200 days for this object. Examining the ICCF for the C IV lag measurement in RM231 with the final SDSS-RM data, the peak around 200 days is still there and strong. But there are additional ICCF peaks, and the highest one happened to be in the negative lag regime, leading to a nondetection for this object. Nevertheless, the number of such missed lags is negligible compared with the additional lags recovered with the final SDSS-RM light curves.

5.3. Comparisons with Single-epoch Estimators

As mentioned earlier, SE mass estimators are extensively used to estimate quasar black hole masses near and far (see the review in, e.g., Shen 2013). There are a large number of SE calibrations for each line from various groups (e.g., Greene & Ho 2005; Vestergaard & Peterson 2006; Wang et al. 2009; Shen & Liu 2012; Ho & Kim 2014; Coatman et al. 2017; Park et al. 2017; Dalla Bontà et al. 2020; Dix et al. 2023). The differences in the luminosity and line-width slopes, the adopted virial f factor, and the methodologies of measuring line widths, combine to produce systematic differences among these SE mass recipes even for the same data. There have not been detailed comparisons between SE and RM masses for the same line and for a large statistical sample (e.g., $N > 10$), except for Hβ using the local RM AGN sample (e.g., Vestergaard & Peterson 2006).

The validity of SE mass estimators relies on two conditions: (1) a tight R–L relation; (2) a tight correlation between line widths measured from the mean (or SE) spectrum and from the rms spectrum. Condition (2) is generally satisfied for the four major broad lines considered here (see Figure 5), although anomalous line "breathing" behaviors would lead to luminosity-dependent biases in the derived SE masses (e.g., Wang et al. 2020b).

However, Condition (1) has only been tested for Hβ, but not for Mg II and C IV, despite recent efforts in deriving a global R–L relation for Mg II and C IV using heterogeneous samples, for which selection functions are not well quantified. As discussed in Section 5.1, there is evidence for a global R–L relation for Hβ and possibly for Mg II as well, although it seems additional physical parameters are involved to introduce scatter in these two R–L relations. For C IV, an R–L relation is only apparent when the dynamic range in luminosity is sufficiently large (Kaspi et al. 2021). Compared with Hβ and Mg II, there is substantially larger scatter in C IV lags around a mean R–L relation, as probed by the SDSS-RM sample (Figure 11). As a result, we expect the correlation between RM mass and SE mass would be the worst for C IV.

Figure 14 compares RM and SE masses for Hβ, Mg II, and C IV. For this comparison, we use fiducial SE estimators from Vestergaard & Peterson (2006; Hβ and C IV) and from Shen et al. (2011; Mg II).[46] As for the luminosity and line width in the SE mass calculation, we adopt the median luminosity over the 90 spectroscopic epochs, and the FWHM measured from the mean spectrum generated from the 90 epochs. By doing so, we are ignoring additional scatter in SE masses from anomalous line "breathing," where, for the same quasar, varying luminosity does not lead to anticipated changes in the SE FWHM value (Wang et al. 2020b).

Figure 14 demonstrates that these SE mass recipes can produce average values that are within $\sim$0.2 dex of the RM masses, with typical scatter of $\sim$0.4–0.5 dex. This systematic offset can be explained by the differences in the adopted R–L relation, the average f factor, or systematic differences in the line-width measurements. As expected, the general correlation between SE and RM masses is the best for Hβ, and the worst for C IV. In fact, using Mg II and C IV SE masses could artificially narrow the mass dynamic range. Namely, for a flux-limited quasar sample like the SDSS-RM sample, the dynamic range in mean FWHM is smaller than that in $\sigma_{\rm line,rms}$, and the dynamic range in L_{3000} or L_{1350} is smaller than that in BLR lags. Thus, the dynamic range in SE masses would be smaller than the dynamic range in RM (true) masses. These caveats of SE masses for flux-limited samples were discussed in detail in

[45] We found two exceptions where the lag measurements may be problematic. For RM078 (the most significant outlier in the Hβ–Mg II comparison), the Mg II lag is much longer than Hβ, which is potentially due to different BLR structures for the two lines. For RM815 (the most significant outlier in the Mg II-C IV comparison), the variability in Mg II and C IV is weak, and both lag measurements may be spurious.

[46] The reason that we adopt the Mg II SE recipe in Shen et al. (2011) instead of that in, e.g., Vestergaard & Osmer (2009) is that the former yields consistent results compared with both Hβ and C IV SE recipes in Vestergaard & Peterson (2006) for luminous SDSS quasars. This apparent agreement is likely a coincidence from the methodologies of measuring the mean FWHMs from the spectra. There is no consensus on what specific SE mass recipes are the best, since the calibration sample and spectral fitting methodologies vary from study to study.

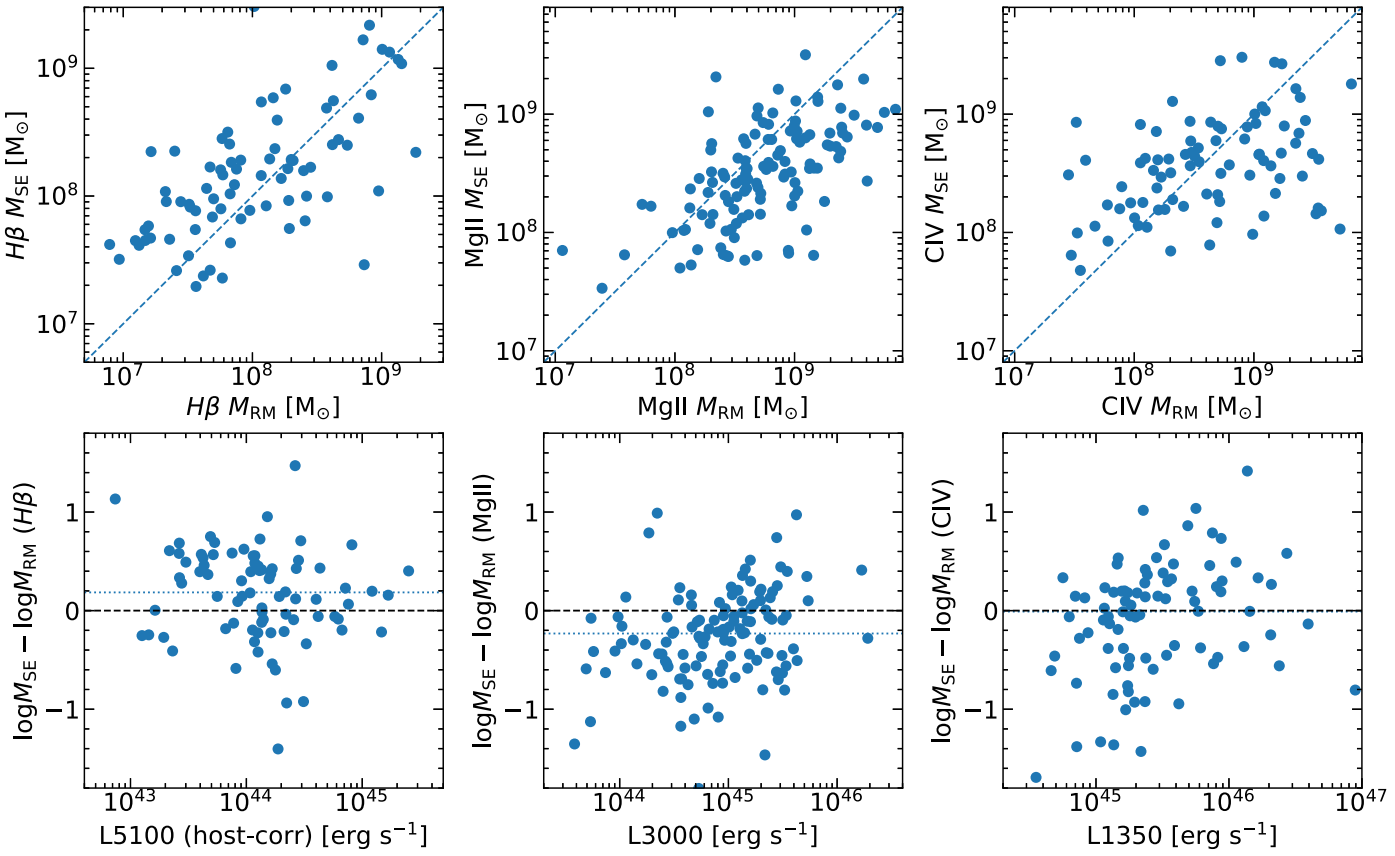

Figure 14. Comparisons between M_{RM} and M_{SE} (using earlier SE recipes) for the same line. Hβ has the best correlation while C IV has the worst correlation. However, for all three lines, there are good correlations between $\sigma_{line,rms}$ and $FWHM_{mean}$, so the poor mass correlation for C IV is mainly driven by the poor R–L relation. Because the single-epoch mass is solely determined by continuum luminosity and line width, its dynamic range can be artificially suppressed compared with RM-based masses (Shen 2013). In the bottom panels, the blue dotted line is a simple median of the residuals, rather than the average offset weighted by measurement uncertainties.

Shen (2013), and procedures to account for the resulting statistical biases in the BH mass distribution are outlined in, e.g., Shen & Kelly (2012) and Kelly & Shen (2013).

5.4. Improved SE Mass Recipes

Next, we consider potential refinements of SE mass recipes using our new measurements of RM masses for Hβ, Mg II, and C IV. The SE estimator is defined as

$$\log M_{SE} \equiv \alpha \log L + \beta \log FWHM_{mean} + C, \quad (2)$$

where L is the corresponding local continuum luminosity for each line. By default, luminosity is in units of 10^{44} erg s^{-1} for L_{5100} and 10^{45} erg s^{-1} for L_{3000} and L_{1350}, roughly the midpoint of the luminosity distribution for each line sample. Here, both continuum luminosity and FWHM are measured with negligible measurement uncertainties, since they are the average values during the monitoring period. We only consider $FWHM_{mean}$ for the SE mass recipes, given its greater reproducibility compared with the more challenging $\sigma_{line,mean}$ measurements (see discussions in, e.g., Wang et al. 2019). As shown in Figure 5, there are good correlations between $FWHM_{mean}$ and $\sigma_{line,rms}$ used for RM mass calculations, justifying the use of FWHM in SE recipes. Nevertheless, there are physical motivations to use σ_{line} instead of FWHM in SE recipes, as reasoned in, e.g., Dalla Bontà et al. (2020).

Because there are correlations between $FWHM_{mean}$ and $\sigma_{line,rms}$, and correlations between continuum luminosity and BLR size for each line, we expect M_{SE} as defined in Equation (2) will be correlated with M_{RM} in general. With the right choices of slopes α and β, the calibration of SE mass recipes becomes a one-parameter fitting problem to constrain the normalization C in Equation (2) using a sample of quasars with M_{RM} measurements (e.g., Vestergaard & Peterson 2006). One option is to adopt slopes that are similar to those in the measured R–L relation and $FWHM_{mean} - \sigma_{line,rms}$ correlations. In the work of Vestergaard & Peterson (2006), they adopt luminosity slopes of $\alpha \approx 0.5$, and $\beta = 2$. The adopted luminosity slopes are consistent with the expected canonical value, and the adopted line-width slopes assume there is a linear relation between $FWHM_{mean}$ and $\sigma_{line,rms}$.

One subtlety in the above calibration procedure is that the regression is performed as matching the observed $\log M_{RM}$ (Y) against the model $\log M_{SE}$ (X), such that the distribution of ($Y|X$) is unbiased (i.e., the expectation value $\langle \log M_{RM} \rangle = \log M_{SE}$ at fixed $\log M_{SE}$). However, to design an unbiased SE mass estimator, we want $\langle \log M_{SE} \rangle = \log M_{RM}$ at fixed true mass $\log M_{RM}$ instead. If the scatter in the R–L relation and in the $FWHM_{mean} - \sigma_{line,rms}$ relation is small, it would be adequate to directly use the slope in the R–L relation (predicting R with L) and the slope in the $FWHM_{mean} - \sigma_{line,rms}$ (predicting $\sigma_{line,rms}$ with $FWHM_{mean}$). In reality, there is substantial scatter in these correlations, and the slope of predicting Y at X could differ significantly from that of predicting X at Y. These differences in

Table 4
Regression Results

$Y = a + bX +$ Scatter						
$Y\|X$	X	X Range	a	b	$\sigma_{\rm int}$	$N_{\rm fit}$
$\log \tau_{\rm rest,H\beta}$	$\log L_{5100,{\rm host-corr}}$	[42.87, 45.40]	$1.458^{+0.038}_{-0.038}$	$0.41^{+0.07}_{-0.07}$	$0.32^{+0.03}_{-0.03}$	80
$\log \tau_{\rm rest,H\beta}$	$\log L_{5100,{\rm host-corr}}$	[42.87, 45.40]	$1.445^{+0.037}_{-0.038}$	**0.533**	$0.32^{+0.03}_{-0.03}$	80
$\log \tau_{\rm rest,MgII}$	$\log L_{3000}$	[43.58, 46.28]	$2.086^{+0.030}_{-0.031}$	$0.31^{+0.06}_{-0.06}$	$0.32^{+0.02}_{-0.02}$	124
$\log \tau_{\rm rest,MgII}$	$\log L_{3000}$	[43.58, 46.28]	$2.095^{+0.030}_{-0.029}$	**0.39**	$0.32^{+0.02}_{-0.02}$	124
$\log \tau_{\rm rest,CIV}$	$\log L_{1350}$	[44.22, 46.95]	$1.840^{+0.075}_{-0.073}$	$0.32^{+0.11}_{-0.11}$	$0.51^{+0.04}_{-0.04}$	89
$\log \tau_{\rm rest,CIV}$	$\log L_{1350}$	[44.22, 46.95]	$1.783^{+0.055}_{-0.054}$	**0.45**	$0.51^{+0.04}_{-0.04}$	89
$\log \sigma_{\rm line,rms,H\beta}$	${\rm logFWHM}_{\rm mean,H\beta}$	[3.16, 4.13]	$0.76^{+0.17}_{-0.16}$	$0.73^{+0.05}_{-0.05}$	$0.152^{+0.009}_{-0.008}$	183
$\log \sigma_{\rm line,rms,MgII}$	${\rm logFWHM}_{\rm mean,MgII}$	[3.14, 4.06]	$0.26^{+0.10}_{-0.10}$	$0.88^{+0.03}_{-0.03}$	$0.119^{+0.003}_{-0.003}$	702
$\log \sigma_{\rm line,rms,CIV}$	${\rm logFWHM}_{\rm mean,CIV}$	[3.14, 4.05]	$0.91^{+0.10}_{-0.10}$	$0.69^{+0.03}_{-0.03}$	$0.099^{+0.003}_{-0.003}$	487
$\log L_{5100,{\rm host-corr}}$	$\log \tau_{\rm rest,H\beta}$	[0.61, 2.39]	$-1.06^{+0.22}_{-0.22}$	$0.78^{+0.14}_{-0.14}$	$0.43^{+0.04}_{-0.03}$	80
$\log L_{3000}$	$\log \tau_{\rm rest,MgII}$	[1.01, 2.95]	$-1.50^{+0.26}_{-0.26}$	$0.68^{+0.13}_{-0.13}$	$0.47^{+0.03}_{-0.03}$	124
$\log L_{1350}$	$\log \tau_{\rm rest,CIV}$	[0.83, 3.05]	$-0.15^{+0.22}_{-0.20}$	$0.30^{+0.10}_{-0.10}$	$0.49^{+0.04}_{-0.04}$	89
${\rm logFWHM}_{\rm mean,H\beta}$	$\log \sigma_{\rm line,rms,H\beta}$	[2.76, 3.89]	$0.88^{+0.17}_{-0.17}$	$0.80^{+0.05}_{-0.05}$	$0.160^{+0.009}_{-0.008}$	183
${\rm logFWHM}_{\rm mean,MgII}$	$\log \sigma_{\rm line,rms,MgII}$	[2.73, 3.84]	$1.30^{+0.07}_{-0.07}$	$0.67^{+0.02}_{-0.02}$	$0.103^{+0.003}_{-0.003}$	702
${\rm logFWHM}_{\rm mean,CIV}$	$\log \sigma_{\rm line,rms,CIV}$	[2.87, 3.84]	$0.81^{+0.11}_{-0.11}$	$0.82^{+0.03}_{-0.03}$	$0.108^{+0.004}_{-0.003}$	487

$Y = C + \alpha X1 + \beta X2 +$ scatter							
$Y\|X1,X2$	$X1$	$X2$	C	α	β	$\sigma_{\rm int}$	$N_{\rm fit}$
$\log M_{\rm SE,H\beta}$	$\log L_{5100,{\rm host-corr}}$	${\rm logFWHM}_{\rm mean,H\beta}$	0.85 $(0.75^{+0.05}_{-0.05})$	**0.5**	**2.0**	$0.45^{+0.04}_{-0.04}$	80
$\log M_{\rm SE,MgII}$	$\log L_{3000}$	${\rm logFWHM}_{\rm mean,MgII}$	−2.05 $(-1.97^{+0.04}_{-0.04})$	**0.6**	**3.0**	$0.47^{+0.03}_{-0.03}$	124
$\log M_{\rm SE,CIV}$	$\log L_{1350}$	${\rm logFWHM}_{\rm mean,CIV}$	1.40 $(1.31^{+0.06}_{-0.06})$	**0.5**	**2.0**	$0.58^{+0.05}_{-0.04}$	88

Note. Default units for $L_{5100,{\rm host-corr}}$, L_{3000}, L_{1350}, line widths, time delays, and BH masses are 10^{44} erg s^{-1}, 10^{45} erg s^{-1}, 10^{45} erg s^{-1}, kilometers per second, days, and $M_\odot$. Linear regressions for the R–L relations and width correlations were performed using the Bayesian approach in Kelly (2007). Regressions with fixed slopes were performed using an MCMC model with `emcee`. Line dispersion for Mg II has been corrected for the velocity split of the doublet. Bold values (without error bars) mean these parameters are fixed.

correlation slopes for scattered samples are apparent in Figure 5 and Figure 11, as well as the compiled regression results in Table 4.

So how do we choose slopes of α and β for each line? The SDSS-RM lag sample alone does not constrain the R–L relations particularly well to allow fitting α and β as free parameters. Instead, we follow Vestergaard & Peterson (2006) to fix these slopes, and use the lag sample to determine the normalization C in Equation (2).

For Hβ and C IV, we adopt $\alpha = 0.5$, the canonical slope in the R–L relation, and $\beta = 2$. These choices are similar to the choices in Vestergaard & Peterson (2006). Adopting $\beta = 2$ is reasonably well justified as the bisector relation between ${\rm FWHM_{mean}}$ and $\sigma_{\rm line,rms}$ does appear to be approximately linear for Hβ and C IV (Figure 5)—recall, for an unbiased SE estimator, we require the regression relation of ${\rm FWHM_{mean}}|\sigma_{\rm line,rms}$. Our measured R–L relation for Hβ is slightly shallower than but consistent with $\alpha = 0.5$ (e.g., Bentz et al. 2013). We are unable to robustly constrain the slope of the R–L relation for C IV, but $\alpha = 0.5$ is roughly consistent with the findings in Kaspi et al. (2021). The comparison between the uncalibrated $M_{\rm SE}$ and the measured $M_{\rm RM}$ for Hβ and C IV is shown in the top row of Figure 15. For Hβ, the correlation is nearly diagonal, indicating our choices of α and β are reasonable. For C IV, however, the correlation is poor, mainly resulting from the large scatter in the C IV R–L relation.

For Mg II, we adopt somewhat different slopes of α and β. The measurement of the Mg II R–L relation is still in an early stage, and a slope different from 0.5 is possible. In addition, the ${\rm FWHM_{mean}} - \sigma_{\rm line,rms}$ is more nonlinear than those for Hβ and C IV (Figure 5). This nonlinearity implies β should not be 2, since the fiducial RM masses are computed as $\propto \sigma^2_{\rm line,rms}$. We found that, in order to produce the most consistent results, we require $\beta > 2$ and $\alpha > 0.5$ for Mg II. After experimenting, we adopt $\alpha = 0.6$ and $\beta = 3.0$ for Mg II. This luminosity slope is not too different from the canonical value, and this ${\rm FWHM_{mean}}$ slope is consistent with the measured slope in predicting ${\rm FWHM_{mean}}$ with $\sigma_{\rm line,rms}$ (see Table 4). Compared with the canonical case of $\alpha = 0.5$ and $\beta = 2$, our fiducial Mg II recipe produces (1) a less biased estimator at fixed $M_{\rm RM}$; (2) more consistent results with Hβ and C IV SE masses for the common objects, over the luminosity range probed by SDSS quasars.

Given the chosen slopes of α and β, we show the uncalibrated $\log M_{\rm SE}$ against $\log M_{\rm RM}$ in the top row of Figure 15. There are good correlations between the SE mass and the RM mass for Hβ and Mg II, and a weak correlation for C IV. We then constrain the normalization C using an MCMC model that incorporates a term for intrinsic scatter, in addition to the measurement uncertainties in $M_{\rm RM}$. The constrained ranges of C are listed in Table 4. We adopt fiducial C values that are slightly different (but consistent within 2σ) from the MCMC results. These tweaked normalizations produce the best agreement between the SE masses from two lines for the same quasars.

The middle row in Figure 15 displays the residuals in log mass as a function of continuum luminosity, which show no significant luminosity trend. The bottom row displays the

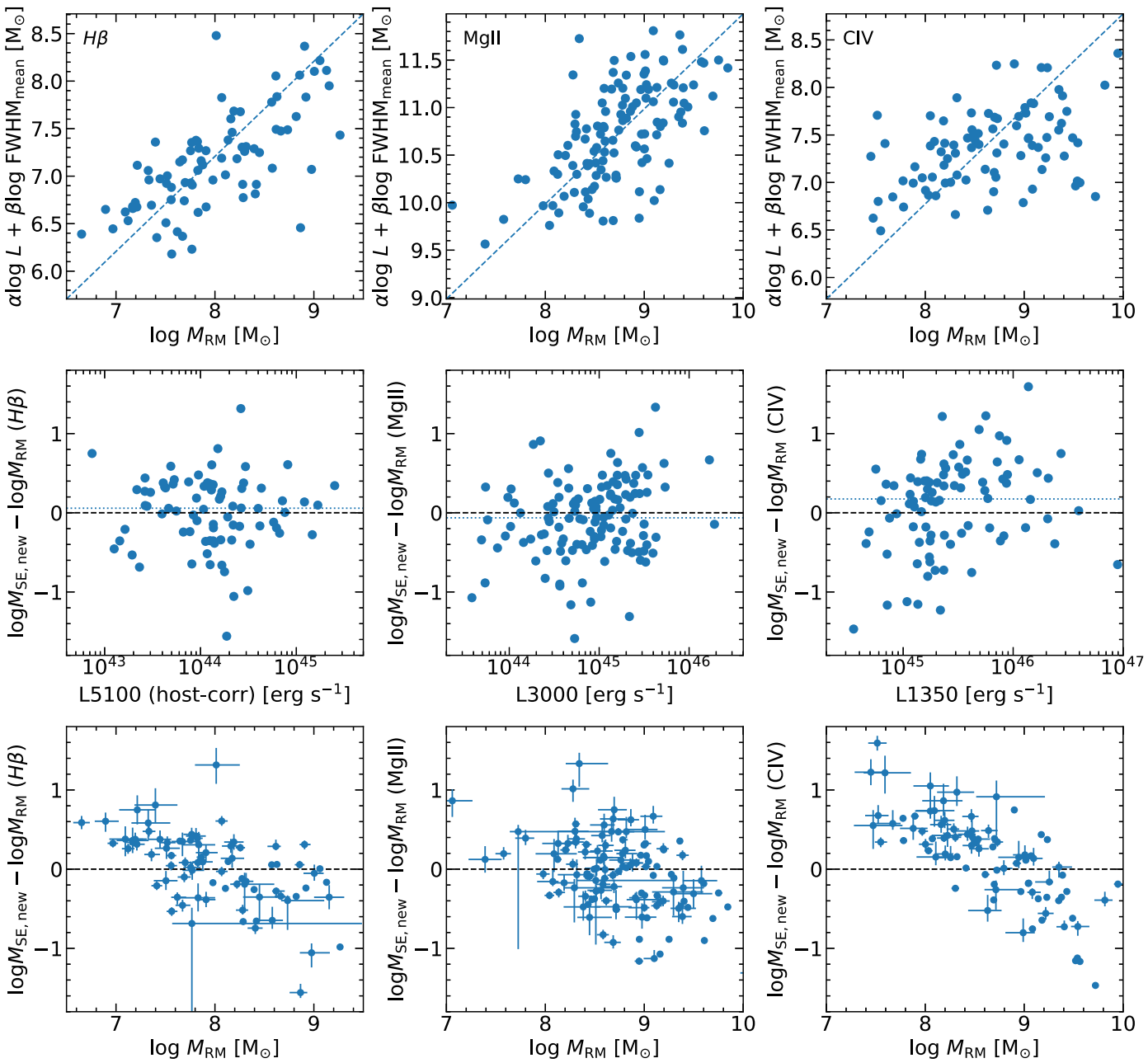

Figure 15. Same as Figure 14, but for the comparisons between $M_{\rm RM}$ and updated $M_{\rm SE}$ estimates (Table 4) for the same line. The new Mg II SE recipe is now an approximately unbiased estimator (i.e., at a given RM mass, they return an unbiased expectation value), while the Hβ and C IV SE recipes are still biased estimators (as we adopted similar slopes on L and mean FWHM as the earlier recipe in Vestergaard & Peterson 2006). However, the normalizations have been recalculated using our RM samples.

residuals as a function of $\log M_{\rm RM}$. By design, the Mg II SE estimator is approximately unbiased, with no significant correlation between the mass residual and $M_{\rm RM}$. For Hβ and C IV, however, the adopted slopes are appropriate for producing unbiased $\log M_{\rm RM}$ at given $\log M_{\rm SE}$. As a consequence, there are systematic biases in $\log M_{\rm SE}$ as a function of $\log M_{\rm RM}$ for Hβ and C IV, as also seen in Vestergaard & Peterson (2006). Nevertheless, this SE mass bias is much less severe for Hβ than for C IV. It is possible to design less biased SE estimators for C IV and for Hβ, which we plan to investigate with more RM data from upcoming MOS-RM programs (e.g., Kollmeier et al. 2017).

Our results thus signify the severe limitations and caveats of SE BH mass recipes, especially for C IV. It is our opinion that individual quasar studies that require robust BH mass estimates should not rely too heavily on C IV SE masses. On the other hand, Hβ and Mg II SE masses seem to reproduce the RM masses well, with an additional ∼0.45 dex scatter.

5.5. Caveats

The dynamical time of the BLR is typically a few years to a few decades for luminous SDSS quasars. While rms quasar variability is on average ∼10% over multiyear timescales, a small fraction of quasars display significantly larger variability amplitude, and sometimes even monotonically increasing or decreasing light curves over time. It is possible that the structure of the BLR in some hypervariable quasars undergoes significant changes over multiyear timescales.

In the SDSS-RM sample, several quasars have notably increased their luminosity over the past few years by more than a factor of a few (e.g., RM017, RM160, etc., Dexter et al. 2019; Fries et al. 2023). While we have attempted to measure an average lag over the full SDSS-RM baseline, we found that the measured lags are substantially longer than those measured from their earlier light curves when they were at a lower-luminosity state (Grier et al. 2017b). It would be interesting to study how the BLR sizes have evolved when the luminosity changes significantly over a timescale comparable to the dynamical time of the BLR, and see if there is evidence for structural changes in the BLR, in addition to the anticipated breathing effect of the BLR.

A related issue is whether or not one should remove a long-term (typically assumed linear) trend in the light curves before measuring the lag. Linear trends in the light curve do not contribute to the short-term cross-correlated variations of interest but may tilt the resulting ICCF, and therefore, there is an argument for detrending before measuring the lag. On the other hand, the removal of a linear trend from the data is somewhat arbitrary: the long-term trend may not be linear, but using higher-order detrending cannot be easily justified; the rms spectra are constructed without detrending; and the removed trend may be part of a real long-term signal in the BLR response. For these reasons, we opt to measure lags without detrending, as the multiyear baseline is typically much longer than the damping timescale of the light curve for our targets. Nevertheless, we perform a test of measuring ICCF lags with (linearly) detrended light curves for our fiducial lag sample. We did not find significant differences for the bulk of lags with and without detrending.

Another important caveat is that the overall detection rate for SDSS-RM quasars is low (see Table 1), as necessarily limited by the light-curve quality of the project. Selection biases from incompleteness in lag detection may have consequences on the inferred *R–L* relations. In Appendix B, we investigate the detection incompleteness for Hβ, Mg II, and C IV. We found that the success of a lag detection primarily depends on how well we can measure the line variability, e.g., via the SNR2 metric discussed in Section 3.3. Most of the lags are undetected simply because their line variability is not well measured (e.g., SNR2 < 20).

For Hβ, we found that the lag detection rate is very high ($\gtrsim$70%) once SNR2 reaches $\sim$35, and there is no obvious difference in quasar properties (e.g., expected lag, quasar luminosity, line equivalent width, line width, etc.) between detected and undetected cases matched in SNR2. These lag nondetections are caused by random fluctuations of the light curves that led to weak cross-correlations in the light curves. Indeed, our visual inspection of nondetection cases with SNR2 > 40 suggests that most of the cases still have some hint of a lag, but the lag is usually small and below the 2σ threshold for our lag detection. There are a handful of cases where the Hβ line is not responding to continuum variations in the expected way, which will be investigated in future work. Thus, we conclude there is limited impact on the Hβ *R–L* relation from detection incompleteness.

For Mg II and C IV, we found similar trends for Hβ, and no obvious differences in quasar properties between the detections and nondetections. However, their lag detection fractions are notably lower than that for Hβ at fixed SNR2. In Appendix B, we elaborate on the potential overestimation of SNR2 for Mg II and C IV due to the insignificant improvement on the spectrophotometry with PrepSpec for these high-redshift quasars. In addition, it is possible that Mg II and C IV in high-redshift and high-luminosity quasars are more often to show abnormal responses to continuum variations, making it more difficult to detect a lag.

Finally, there is sill room to improve the lag measurement methodology, spectrophotometry refinement with PrepSpec, and scrutiny on individual light curves. While these additional efforts are beyond the scope of the current work, we plan to investigate these issues as we continue monitoring the SDSS-RM field in SDSS-V with the Black Hole Mapper program (Kollmeier et al. 2017). In particular, we plan to improve the spectrophotometry at the blue end of SDSS spectra, which would improve the lag detection for Mg II and C IV, as well as other UV broad lines, in high-redshift quasars.

5.6. Applications of SDSS-RM Data

The rich data from SDSS-RM enable a broad array of applications. The large number of RM lags and reliable lag-based RM masses have proven valuable to study the evolution of the BH mass–host galaxy correlations toward high redshift (e.g., Shen et al. 2015b; Matsuoka et al. 2015; Li et al. 2021, 2023). By extending RM measurements to high redshift and measuring large samples of Mg II and C IV lags, we are starting to explore the diversity in the BLR structure for these UV broad lines, and to critically evaluate their potential for designing efficient SE BH mass recipes.

In the meantime, the multiyear photometric and spectroscopic light curves for the SDSS-RM sample allow investigations of the general variability of quasars (Sun et al. 2015; Dyer et al. 2019), and of the variability of particular subsets of quasars, such as broad absorption line quasars (Grier et al. 2015; Hemler et al. 2019) and extreme variability quasars (Dexter et al. 2019; Fries et al. 2023). In addition to the unprecedented spectroscopic monitoring data that allow detailed spectral variability analyses, the SDSS-RM sample provides high-quality multiwavelength information (Shen et al. 2019b; Liu et al. 2020) for a well-defined quasar sample. These data have enabled additional RM science, such as constraining the accretion disk structure using continuum lags (Homayouni et al. 2019) and modeling the propagation of variability in the accretion disk (e.g., Stone & Shen 2023).

In this paper, we focused on the main science goal of SDSS-RM, i.e., the measurements of lags for major broad emission lines in quasar spectra. We will explore other topics with the final SDSS-RM data in future work. We also invite the community to exploit the rich SDSS-RM data for RM science or general AGN science.

6. Summary and Outlook

6.1. Main Findings

This work marks the official conclusion of the SDSS-RM program, although we will continue to analyze this data set. We present the final SDSS-RM data, including 7 yr spectroscopy (2014–2020) and 11 yr photometry (2010–2020). Using this data set, we measure velocity-integrated average RM lags for Hα, Hβ, Mg II, and C IV, over broad ranges of redshifts ($0.1 < z < 4.5$) and quasar luminosities ($L_{\rm bol} = 10^{44\text{–}47.5}$ erg s^{-1}). We report 23, 81, 125, and 110 lags for each of the four lines, respectively. The main findings are summarized as follows.

1. The measured lags from different lines in the same quasars reveal BLR stratification, consistent with earlier findings based on low-z RM work. Specifically, Hα and Mg II have on average longer lags than Hβ, while C IV lags are somewhat shorter than Mg II lags overall (but the sample size is small). This result suggests increasing distances of the (observable) emitting clouds from the BH in the order of C IV $<$ Hβ $<$ Hα $\lesssim$ Mg II. (See Section 5.1.)
2. The measured RM BH masses are consistent between two different lines, albeit with different lags and line widths for both lines (Section 5.1). This validates the basic principle of measuring BH mass with the RM technique. In calculating the RM masses, we have derived a new estimate of the average (geometric mean) virial factor of $\langle \log f \rangle = 0.62 \pm 0.07$ (when using $\sigma_{\rm line,rms}$ as line width), based on measurements of dynamical-modeling BH masses in 30 low-z RM AGNs. The intrinsic scatter in individual virial factors is 0.31 ± 0.07 dex, indicating a factor of 2 systematic uncertainties in RM masses calculated using the average virial factor. (See Section 4.5.)
3. There is a significant global R–L relation for Hβ and Mg II in SDSS-RM quasars, using $R = c\tau_{\rm rest}$. The best-fit slopes and normalizations are consistent with latest measurements using heterogeneous samples (Bentz et al. 2013; Yu et al. 2023). The slopes of the Hβ and Mg II R–L relations are also consistent with the canonical value of 0.5 expected from photoionization. The intrinsic scatter in the R–L relations (in the R direction) is $\sim$0.3 dex. Because the luminosity dynamic range for the SDSS-RM sample is still somewhat limited, it is possible that the constrained slopes are biased from the global slope across many orders of magnitude in AGN luminosity. (See Section 5.1.)
4. Given the limited dynamic range in luminosity for C IV, we are unable to well constrain the slope of an assumed R–L relation. However, our slope measurements are consistent with the latest results (Kaspi et al. 2021) based on a heterogeneous sample that spans a much larger luminosity range than the SDSS-RM sample. More importantly, the intrinsic scatter in the said R–L relation for C IV is substantially larger than that for Hβ or Mg II, likely caused by the larger dispersion in the ratio between C IV ionizing flux and 1350 Å continuum flux in quasars. (See Section 5.1.)
5. We present recalibrated SE BH mass recipes for Hβ, Mg II, and C IV, using the SDSS-RM lag sample. The new Hβ and Mg II SE recipes are approximately unbiased estimators of RM mass (i.e., expectation value $\langle \log M_{\rm SE} \rangle = \log M_{\rm RM}$), with an intrinsic scatter of $\sim$0.45 dex around RM masses. The C IV SE recipe, on the other hand, shows substantially larger scatter ($\sim$0.58 dex) and the weakest correlation with RM masses. Nevertheless, these new SE recipes show overall consistency across two lines over the luminosity range probed by SDSS-RM quasars. Considering the systematic uncertainty of $\sim$0.3 dex in RM masses (Section 4.5), one can argue that the absolute uncertainties of SE masses are $\sim$0.35 dex for Hβ and Mg II, and $\sim$0.5 dex for C IV. Our findings continue to support the usage of Hβ and Mg II SE masses, but caution on the usage of C IV SE masses for high-redshift quasars. The three SE mass recipes are summarized below. (See Section 5.4.)

$$\log\left(\frac{M_{\rm SE,H\beta}}{M_\odot}\right) = \log\left[\left(\frac{L_{5100,\rm AGN}}{10^{44}\ {\rm erg\ s^{-1}}}\right)^{0.5}\left(\frac{\rm FWHM}{\rm km\ s^{-1}}\right)^2\right] + 0.85. \quad (3)$$

$$\log\left(\frac{M_{\rm SE,MgII}}{M_\odot}\right) = \log\left[\left(\frac{L_{3000}}{10^{45}\ {\rm erg\ s^{-1}}}\right)^{0.6}\left(\frac{\rm FWHM}{\rm km\ s^{-1}}\right)^3\right] - 2.05. \quad (4)$$

$$\log\left(\frac{M_{\rm SE,CIV}}{M_\odot}\right) = \log\left[\left(\frac{L_{1350}}{10^{45}\ {\rm erg\ s^{-1}}}\right)^{0.5}\left(\frac{\rm FWHM}{\rm km\ s^{-1}}\right)^2\right] + 1.40. \quad (5)$$

6.2. Outlook

The results from the SDSS-RM project unambiguously confirm the feasibility and potential of performing MOS-RM in the high-redshift regime. Hβ continues to be the most reliable line for RM purposes. We have shown that it is possible to measure Mg II lags with optical spectroscopic RM, but the increased Mg II lag and diluted RM response make it more difficult to robustly measure the lag. Indeed, the Mg II lag detection rate is the lowest among the four lines studied in this work (see Table 1). We also demonstrated that measuring C IV lags in high-redshift, high-luminosity quasars is possible, provided that the monitoring baseline is sufficient to constrain the long time delays.

The final lag yields from SDSS-RM are broadly consistent with our predictions at the beginning of the project (Shen et al. 2015a). The forecast was made using simulated light curves of quasars with physical properties and observing cadence/baseline/depth similar to those in SDSS-RM, but the lag detection methodology and criteria in the forecast were somewhat different. Nevertheless, this overall consistency demonstrates the important value of planning an MOS-RM program with tailored simulations. In Appendix B, we provide additional lessons learned from SDSS-RM.

Given the success of SDSS-RM and other MOS-RM programs, there are several ways to improve RM studies in the high-z regime. First, the typical quality of lag measurements from current MOS-RM programs is still low to moderate. This is not a limitation of the principles of the MOS-RM approach, but a limitation from insufficient observing resources. Future MOS-RM programs with larger aperture telescopes (e.g., McConnachie et al. 2016; Swann et al. 2019) will provide higher S/N spectroscopy to better measure the reverberation responses in the broad-line flux. These future MOS-RM programs will also be able to target fainter quasars (thus extending the sample to lower luminosities), and measure velocity-resolved lags to better constrain dynamical models of the BLR. In anticipation of these future MOS-RM programs, there are also continued efforts to expand the sample size of high-redshift lag measurements with existing MOS facilities, such as the SDSS-V Black Hole Mapper RM program during 2020–2027 (Kollmeier et al. 2017). In particular, the SDSS-RM field continues to be monitored in SDSS-V, providing extended light-curve baselines for a significant fraction of SDSS-RM quasars.

Besides the optical MOS-RM approach, there are also important complementary approaches to advance RM in the high-z regime. To extend the R–L relations to the most luminous quasars (e.g., $L_{\rm bol} \gtrsim 10^{47}$ erg s^{-1}), it is necessary to perform decade-long monitoring campaigns for individual objects (e.g., Lira et al. 2018; Czerny et al. 2019a; Zajaček et al. 2020; Kaspi et al. 2021), since these most luminous quasars are rare and sparsely distributed on the sky. The high success rate of measuring Hβ lags at high-z and the overall best reliability of Hβ for RM mass measurements suggest the possibility of IR RM for Hβ at $z > 1$. Such IR RM programs would require large-aperture telescopes to measure the Hβ response with IR spectroscopy.

Finally, the advent of high-spatial-resolution spectroastrometric measurements with interferometry, e.g., with the GRAVITY instrument on the Very Large Telescopes, has enabled a new avenue toward constraining the inner structures of broad-line AGNs and quasars (e.g., Gravity Collaboration et al. 2018; GRAVITY Collaboration et al. 2020, 2021). With the upgraded GRAVITY+ instrument (GRAVITY+ Collaboration 2022), it will become possible to target relatively faint (e.g., $K \sim 13$–15) quasars at $z > 1$. Combining the spectro-astrometric measurements (e.g., Bailey 1998; Bosco et al. 2021) and RM measurements for the same quasar will provide powerful joint constraints on the underlying BLR structure and kinematics (e.g., Li et al. 2022), an approach that can also be used to measure geometric distances toward distant AGNs (e.g., Elvis & Karovska 2002; Hönig et al. 2014; Wang et al. 2020a; GRAVITY Collaboration et al. 2021; Songsheng et al. 2021). Far down the road, it may eventually become feasible to measure astrometric jitter signals due to reverberation (e.g., Shen 2012; Li & Wang 2023). These various techniques motivate continued monitoring programs to constrain AGN inner structures, measure BH masses, and improve our understanding of the growth of SMBHs and the expansion history of the Universe.

Acknowledgments

We thank the referee for comments that improved the manuscript, and Pu Du, Yan-Rong Li, and Aaron Barth for useful discussions. Y.S. acknowledges support from NSF grants AST-1715579 and AST-2009947. C.J.G. acknowledges support from NSF grants AST-2009949 and AST-2108667. J.I. L. is supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures program. Y.H. was supported as an Eberly Research Fellow by the Eberly College of Science at the Pennsylvania State University. J.R. T. acknowledges support from NSF grants CAREER-1945546, AST-2009539, and AST-2108668. W.N. B. acknowledges support from NSF grant AST-2106990 and the Eberly Endowment at Penn State. L.C.H. was supported by the National Science Foundation of China (11721303, 11991052, 12011540375, 12233001) and the China Manned Space Project (CMS-CSST-2021-A04, CMS-CSST-2021-A06). C.T. acknowledges Tsinghua University for the support to her work.

This work is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada–France–Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. The authors recognize the cultural importance of the summit of Maunakea to a broad cross section of the Native Hawaiian community. The astronomical community is most fortunate to have the opportunity to conduct observations from this mountain.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

The PS1 has been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of

Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under grant No. AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE).

ZTF. Based on observations obtained with the Samuel Oschin 48-inch Telescope at the Palomar Observatory as part of the Zwicky Transient Facility project. ZTF is supported by the National Science Foundation under grant No. AST-1440341 and a collaboration including Caltech, IPAC, the Weizmann Institute for Science, the Oskar Klein Center at Stockholm University, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron and Humboldt University, Los Alamos National Laboratories, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, and Lawrence Berkeley National Laboratories. Operations are conducted by COO, IPAC, and UW.

Software: astropy (Astropy Collaboration et al. 2013), celerite (Foreman-Mackey et al. 2017), emcee (Foreman-Mackey et al. 2013), `PyCCF` (Sun et al. 2018), `Javelin` (v0.33; Zu et al. 2011), `PyROA` (v3.1.0; Donnan et al. 2021).

Appendix A
Format of Data Products

The final SDSS-RM data set and related data products are provided via anonymous ftp and https services[47] (or contact Y. Shen for the latest archival info). We organize the data products by the object ID (RMID). In each object directory (e.g., *./rm000/*), we provide the original spectra, continuum and emission-line light curves, PrepSpec outputs, and lag measurements. Additional information about the SDSS-RM sample is provided in the sample characterization paper Shen et al. (2019b). Table 5 provides a brief description of the data products available. Future updates of these products will also be distributed via this site. Below, we describe the main data products released with this paper.

SDSS spectra. Customarily reprocessed (e.g., Shen et al. 2015a) optical spectroscopy from SDSS is provided in the *./spec/* directory in ascii format. A total of 90 epochs are available for each object. These are the original spectra before the PrepSpec run. However, these spectra were customarily reprocessed as described in Shen et al. (2015a), thus are slightly different from the spectra included in the SDSS-III/IV public data releases.

Continuum and emission-line light curves. The merged continuum light curve and emission-line light curves (after processed by PrepSpec) are provided in plain ascii files with the *_lc.txt* affix in each object directory (e.g., *./rm000/*). These light curves are used in the lag measurements. The emission-line light-curve files only include Hα, Hβ, Mg II, and C IV, with prefixes *ha, hb, mg2*, and *c4*, respectively. Additional emission-line light curves, if covered by spectroscopy, will be included in the *./prepspec/* directory in the future. We are aware that, in rare occasions, the merging of the continuum light curves failed, with discrepant fluxes from different facilities. These corrupted light curves usually do not produce successful lags, and we did not attempt to resolve these flux discrepancies in this work.

PrepSpec outputs. Model light curves, a flux calibration scaling factor (*_p0_t.dat*), mean and rms spectra, and line-width measurements from PrepSpec run are provided in the *./prepspec/* directory. These outputs include all available lines covered in SDSS spectra. The flux calibration scaling factor *p0_t* is used to scale the SDSS spectra to improve spectrophotometry. A subset of the PrepSpec outputs are also compiled in the summary Table 2.

[47] ftp://quasar.astro.illinois.edu/public/sdssrm/final_result/; https://ariel.astro.illinois.edu/sdssrm/final_result/

Table 5
Content of Data Products

File Path	Description
./	Top directory
./spec/	Directory for SDSS spectra (before PrepSpec)
./prepspec/	Directory for PrepSpec outputs
./rm/*	Object directory for each SDSS-RM quasar
./scripts/	Batch scripts and additional software tools
./summary.fits	Summary table (Table 2)
	Content of *./prepspec/*
File Path	Description
./rm/*	Object directory
⋯*./rm*_p0_t.dat*	Scaling factor of flux calibration (tabulated in $\ln p_0(t)$)
⋯*./rm*_avg_w.dat*	Average spectrum
⋯*./rm*_rms_w.dat*	rms spectrum; column 2 is the "RMSx" estimate (Shen et al. 2016)
⋯*./rm*_$cont$.dat*	Continuum light curve (e.g., "c3000," "c5100")
⋯*./rm*_$cont $_t_stats.dat*	Statistics of the continuum light curve
⋯*./rm*_vblr.dat*	Line-width measurements from the average spectrum
⋯*./rm*_vvblr.dat*	Line-width measurements from the rms spectrum
⋯*./rm*_$line$_t.dat*	Line light curve
⋯*./rm*_$line$_w.dat*	rms line profile
⋯*./rm*_$line $_t_stats.dat*	Statistics of the line light curve
	Content of *./rm*/*
File Path	Description
./lag_summary.pdf	Summary plot of lag measurements
./cont_lc.txt	Merged continuum light curve
./$line$_lc.txt	Original line light curve (first three columns in PrepSpec output *rm*_$line$_t.dat*)
./pypetal/	Lag measurement directory
⋯*./light_curves/*	Copied original light curves with flags for outliers added
⋯*./processed_lcs/*	Outlier-rejected light curves and light curves formatted for `Javelin` and `PyROA`
⋯*./cont/drw_rej/*	Continuum DRW fitting result and outlier mask
⋯*./$line$/*	Lag measurements for the line
⋯*./pyccf/*	ICCF results
⋯*./javelin/*	`Javelin` results
⋯*./pyroa/*	`PyROA` results
⋯*./weights/*	Weights for lag posteriors

Note. There are multiple files inside each directory. The content and format of these files are either self-explanatory or explicitly specified in the header. Occasionally, the "lag_summary.pdf" plot needs to be regenerated for better display ranges using the scripts provided in the *./scripts/* directory. Bold font is to emphasize the most important file.

Lag measurements. Outputs from our lag analyses are provided in the *./pypetal/* directory in each object directory. These outputs include the DRW fits to the continuum light curves and outlier masks, ICCFs, `Javelin`/`PyROA` results, and weights applied to the lag calculations. The essential subset of these measurements are compiled in the summary Table 2. For convenience, we provide a lag summary plot in each object directory (e.g., *./rm000/lag_summary.pdf*) to display the lag measurements for all available lines.

Software. Scripts of running lag detection and postprocessing are provided in the *./scripts/* directory.

Appendix B
Informing Future MOS-RM Programs

The lag yields from the final SDSS-RM data are roughly in line with the simulation forecast in Shen et al. (2015a), despite the slightly different total baseline and lag detection criteria assumed in the forecast, and the fact that the predictions were based on simulated light curves. In Figure 16, we show the final lag detection fraction as a function of the line variability SNR2 parameter described in Section 3.3 (also see Shen et al. 2019b). For a future MOS-RM program, it is necessary to have a sufficient number of epochs to sample variability patterns over a sufficiently long (e.g., multiyears) spectroscopic baseline. However, in order to detect the line variability properly, each spectroscopic epoch must also achieve a significant S/N for the line flux measurement. A value of SNR2 = 20 with the benchmark SDSS-RM cadence/baseline (with $N_{\rm epoch} \approx 100$) approximately corresponds to line RMS/$\langle\sigma_i\rangle \approx 2$, where line RMS is the rms line flux variability (absolute not fractional), and $\langle\sigma_i\rangle$ is the average per-epoch flux uncertainty. Thus, for a 20% fractional line variability, i.e., RMS/AVG = 0.2, we require a fractional line flux measurement uncertainty of $\langle\sigma_i\rangle$/AVG = 10% in order to reach SNR2 = 20. Deeper epoch spectra, e.g., with 5% flux measurements, will increase SNR2 to ∼40 and subsequently boost the probability of lag detection (Figure 16).

For MOS spectroscopy, there is also a systematic floor of flux uncertainties limited by the spectrophotometry accuracy of the MOS program. For example, in SDSS-RM, we achieved a typical systematic flux uncertainty floor of ∼5%. Using PrepSpec and for quasars with strong narrow lines (such as [O III] 5007) at low redshift, we were able to further improve the spectrophotometry to ∼2%–3%. This additional improvement is likely partially responsible for the higher lag detection rate for Hβ at fixed SNR2 (computed using statistical flux uncertainties only), compared with Mg II and C IV for which the PrepSpec improvement on spectrophotometry is marginal. Moreover, the systematic uncertainty of spectrophotometry increases toward the blue end of SDSS spectra (e.g., Shen et al. 2015a), resulting in a larger impact on lag detection for Mg II and C IV in high-redshift quasars than for Hβ in the low-redshift subset.

The outcome from the SDSS-RM final data set thus provides useful guidance for future MOS-RM programs with similar baseline and cadence. Given the typical ∼10% rms variability of broad lines, a low systematic floor of spectrophotometry will greatly improve the lag detection rate. At the same time, we recommend to reach a statistical (per epoch) line flux uncertainty of better than ∼10% at the flux limit of the sample. For SDSS-RM, the latter goal is achieved with 2 hr integration per epoch at a limiting magnitude of $i_{\rm psf} < 21.7$, for the general quasar population.

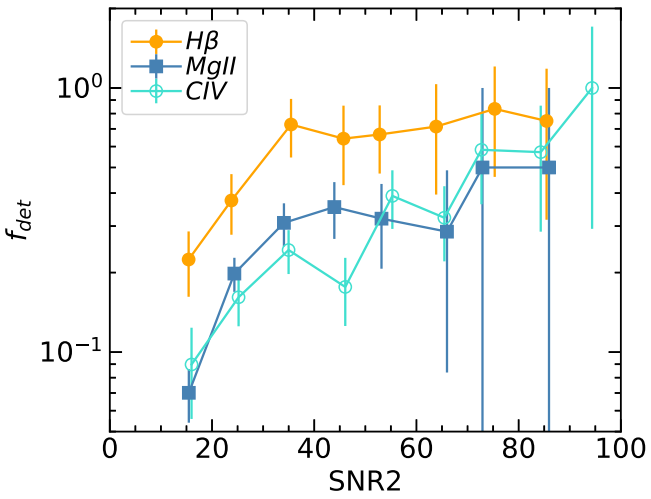

Figure 16. Lag detection fraction as a function of line variability parameter SNR2, for Hβ, Mg II, and C IV, respectively. Error bars are Poisson counting uncertainties.

Appendix C
Additional Quality Cuts on Lags

As discussed in Section 4.2, we impose a set of quantitative cuts for lag detection. Our choice of the $r_{\rm max} > 0.4$ cut is more lenient than those used in other high-quality RM studies. However, while imposing this less stringent cut inevitably introduces more false positives in the lag sample, it also ensures that we are not removing lags near the detection boundary that could potentially bias the *R–L* relation measurements.

We test the effects of imposing more stringent quality cuts on lag detection. Figure 17 shows the comparison in the *R–L* plane with our fiducial lag results for (1) a higher $r_{\rm max} > 0.6$ cut; (2) an SNR2 > 35 cut; (3) high-quality lags from manual inspection as detailed in Section 4.3. These additional quality cuts ($r_{\rm max}$, SNR2, grade) roughly reduce the fiducial lag sample by half for each line.

As shown in Figure 17, for the quantitative cuts with $r_{\rm max}$ and SNR2, we generally found consistent results on the best-fit *R–L* relation in terms of the slope, normalization, and intrinsic scatter. For the visual inspection grade cut, however, the Mg II and C IV *R–L* relations are notably more flattened and shifted to lower normalizations. This is because long lags with limited light-curve overlap are generally less likely to be perceived as a good-quality detection. As a result, a manual inspection preferentially removes long lags and will bias the measurement of the *R–L* relation. This highlights the danger of cherrypicking the best-quality lags in the *R–L* relation measurements from a systematic RM program such as ours.

For full disclosure, we provide the best-fit *R–L* relation for these higher-quality lag subsets in Table 6, but we strongly caution that these relations are not intended for any practical uses for the reasons discussed above.

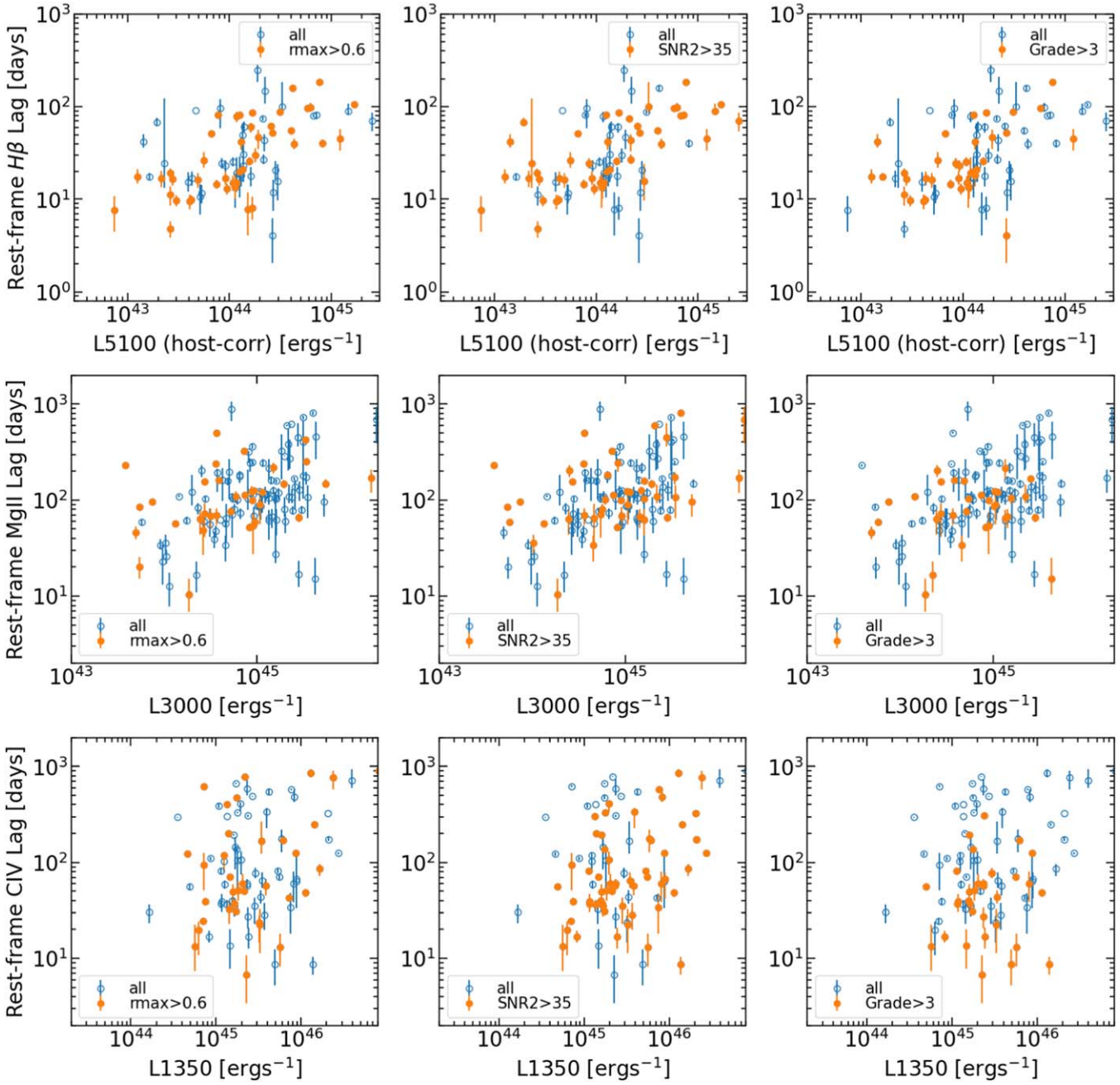

Figure 17. Tests of lag detection with more stringent quality cuts. In each panel, the open symbols represent the fiducial lag sample, and the filled symbols are for the subset. The left column shows results with a $r_{\rm max} > 0.6$ cut; the middle column shows results with an SNR2 > 35 cut; the right column shows the results with the manual inspection cut. Details regarding these quality cuts and their impact on the measured *R–L* relations are presented in Appendix C.

Table 6
Regression Results for Higher-quality Lags

$Y = a + bX +$ Scatter							
Cut	$Y\|X$	X	X Range	a	b	$\sigma_{\rm int}$	$N_{\rm fit}$
$r_{\rm max} > 0.6$	$\log \tau_{\rm rest,H\beta}$	$\log L_{5100,{\rm host-corr}}$	[42.87, 45.22]	$1.432^{+0.044}_{-0.047}$	$0.54^{+0.08}_{-0.09}$	$0.28^{+0.04}_{-0.03}$	44
SNR2 > 35	$\log \tau_{\rm rest,H\beta}$	$\log L_{5100,{\rm host-corr}}$	[42.87, 45.40]	$1.450^{+0.041}_{-0.041}$	$0.43^{+0.07}_{-0.07}$	$0.27^{+0.03}_{-0.03}$	48
Grade $= 4/5$	$\log \tau_{\rm rest,H\beta}$	$\log L_{5100,{\rm host-corr}}$	[43.10, 45.08]	$1.418^{+0.047}_{-0.048}$	$0.42^{+0.10}_{-0.11}$	$0.28^{+0.04}_{-0.03}$	37
$r_{\rm max} > 0.6$	$\log \tau_{\rm rest,MgII}$	$\log L_{3000}$	[43.58, 46.23]	$2.079^{+0.060}_{-0.061}$	$0.21^{+0.10}_{-0.09}$	$0.32^{+0.05}_{-0.04}$	35
SNR2 > 35	$\log \tau_{\rm rest,MgII}$	$\log L_{3000}$	[43.58, 46.28]	$2.111^{+0.053}_{-0.053}$	$0.22^{+0.09}_{-0.09}$	$0.33^{+0.04}_{-0.04}$	44
Grade $= 4/5$	$\log \tau_{\rm rest,MgII}$	$\log L_{3000}$	[43.69, 45.63]	$1.935^{+0.056}_{-0.058}$	$0.09^{+0.11}_{-0.11}$	$0.27^{+0.05}_{-0.04}$	32
$r_{\rm max} > 0.6$	$\log \tau_{\rm rest,CIV}$	$\log L_{1350}$	[44.66, 46.95]	$1.752^{+0.121}_{-0.119}$	$0.47^{+0.16}_{-0.16}$	$0.54^{+0.08}_{-0.06}$	36
SNR2 > 35	$\log \tau_{\rm rest,CIV}$	$\log L_{1350}$	[44.69, 46.95]	$1.665^{+0.085}_{-0.086}$	$0.47^{+0.12}_{-0.12}$	$0.43^{+0.05}_{-0.04}$	57
Grade $= 4/5$	$\log \tau_{\rm rest,CIV}$	$\log L_{1350}$	[44.69, 46.14]	$1.634^{+0.113}_{-0.112}$	$0.005^{+0.219}_{-0.219}$	$0.41^{+0.07}_{-0.05}$	32

Note. Same as Table 4 but for subsets of higher-quality lags as described in Appendix C. Because these more stringent cuts impose additional selection functions on the resulting lag sample, we caution that these best-fit *R–L* relations are most likely biased, and therefore are not for practical uses.

Appendix D
Lag Examples

We provide several additional examples in Figures 18–20 for Hβ, Mg II, and C IV lags in our fiducial lag sample that represent the range of detection qualities based on our visual inspection. We caution that our visual inspection is subjective on individual cases, but on average, higher-grade lags are more robustly detected. Lower-grade lags do not necessarily mean the lag is spurious, and may reflect the complexity of the BLR and the limitations of measuring an average lag.

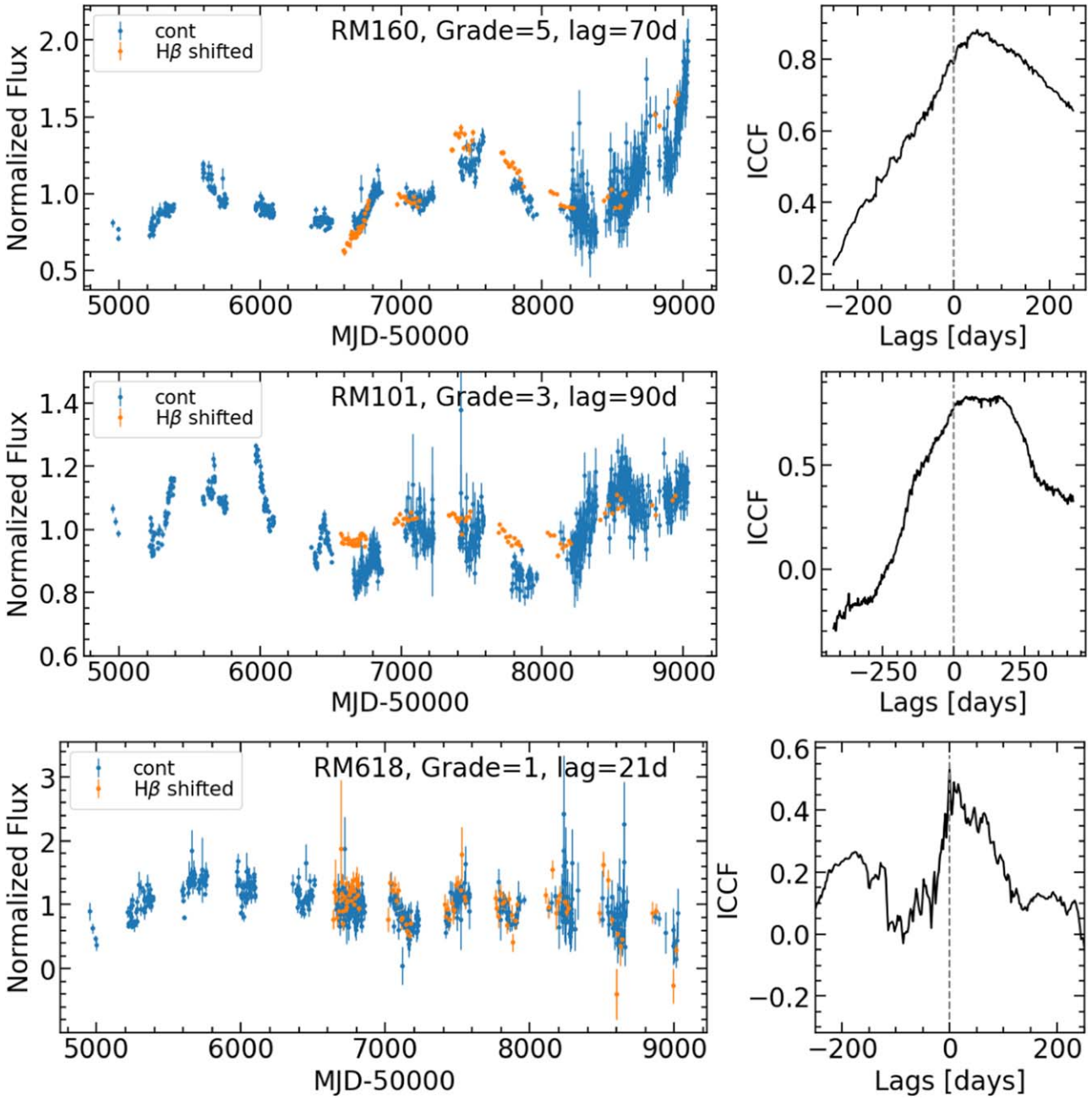

Figure 18. Additional examples of Hβ lags (light curves in the left panel, and ICCF in the right panel). Higher (visual inspection) grades correspond to better measured lags. The continuum and broad-line light curves are normalized to have unity mean within their respective baseline, and therefore, they do not necessarily line up after shifting the broad-line light curve by the observed-frame `PyROA` lag.

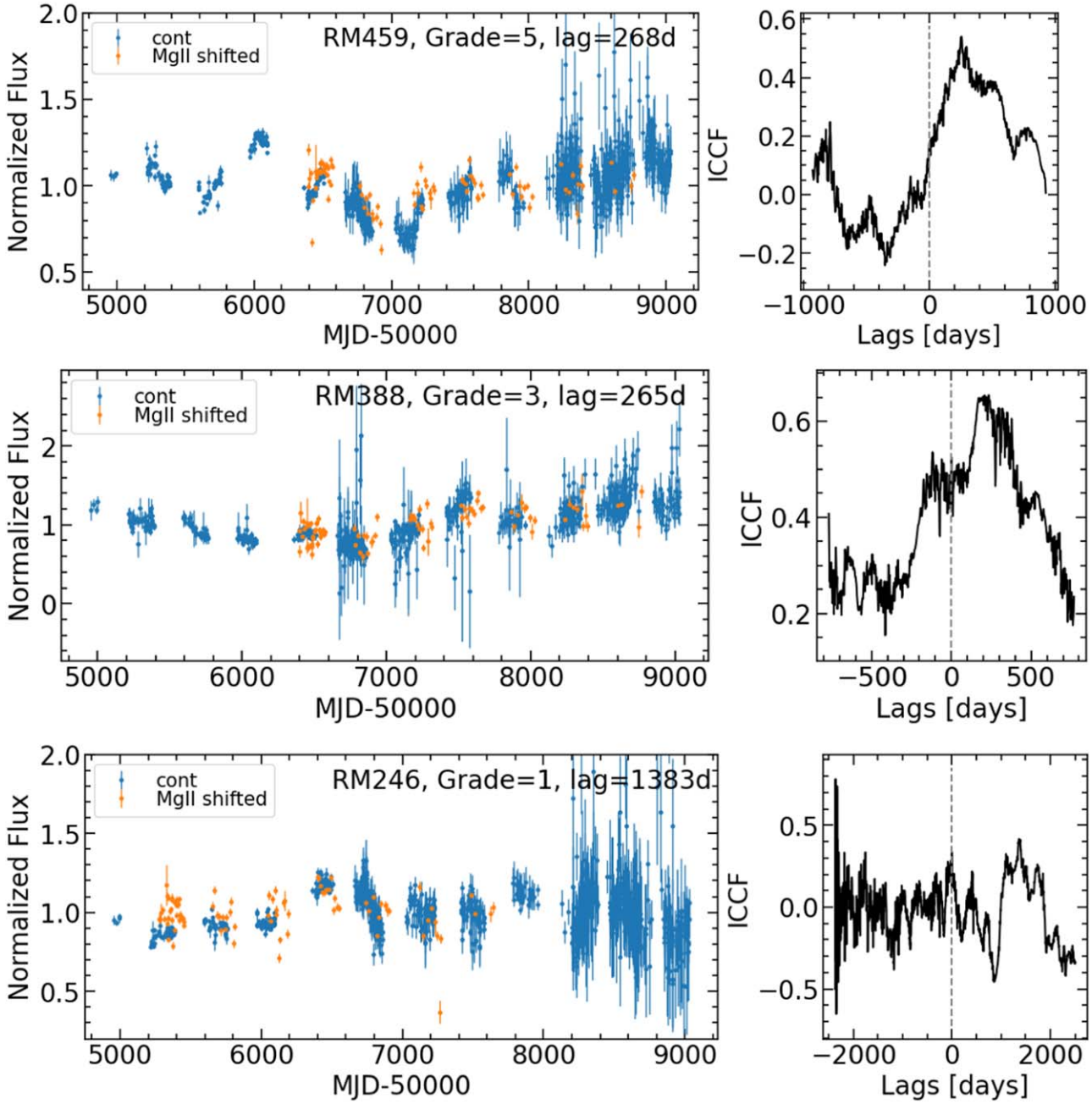

Figure 19. Additional examples of Mg II lags. Higher (visual inspection) grades correspond to better measured lags.

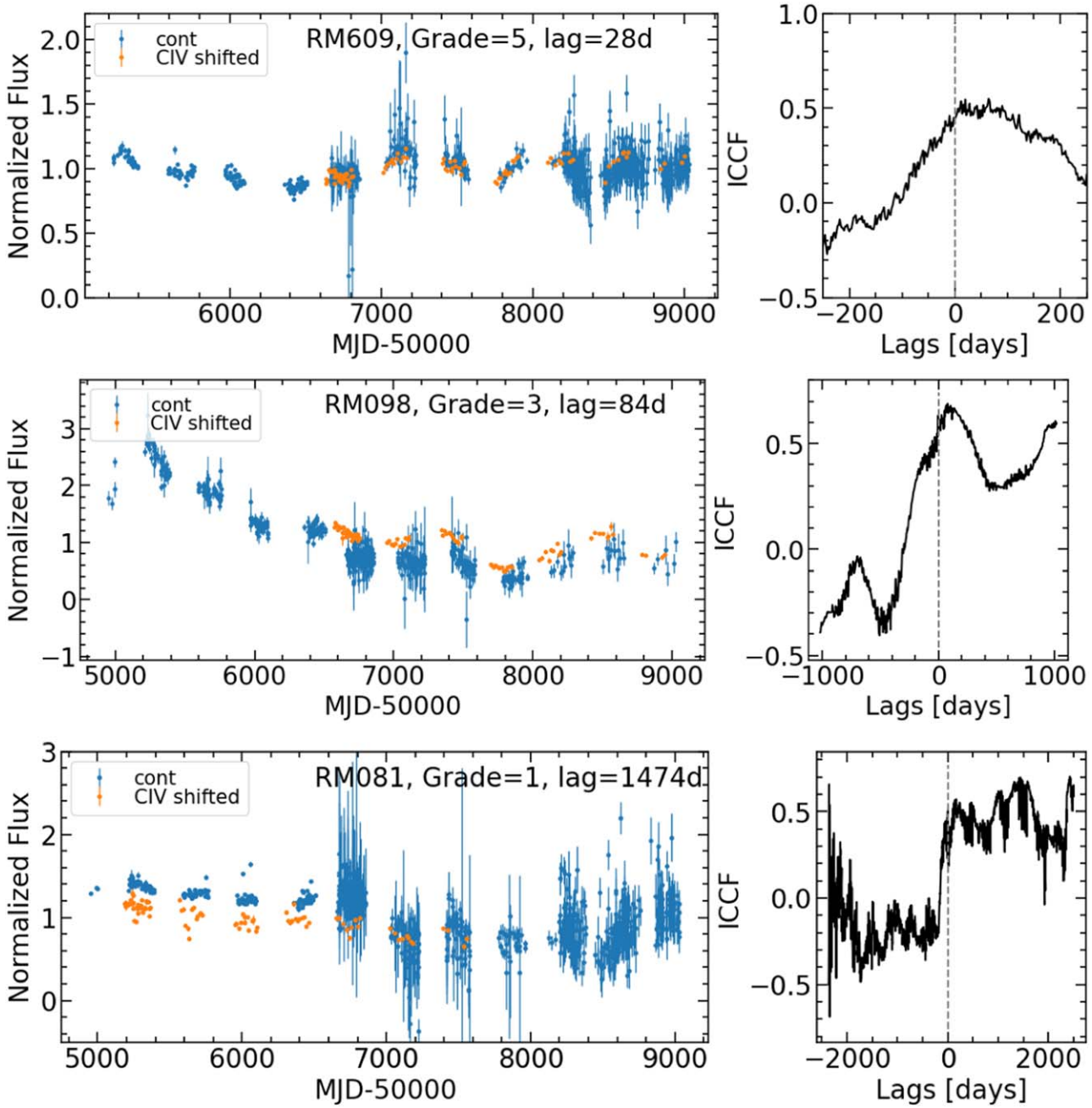

Figure 20. Additional examples of C IV lags. Higher (visual inspection) grades correspond to better measured lags.

ORCID iDs

Yue Shen https://orcid.org/0000-0003-1659-7035
Catherine J. Grier https://orcid.org/0000-0001-9920-6057
Keith Horne https://orcid.org/0000-0003-1728-0304
Zachary Stone https://orcid.org/0000-0002-8501-3518
Jennifer I. Li https://orcid.org/0000-0002-0311-2812
Qian Yang https://orcid.org/0000-0002-6893-3742
Yasaman Homayouni https://orcid.org/0000-0002-0957-7151
Jonathan R. Trump https://orcid.org/0000-0002-1410-0470
Scott F. Anderson https://orcid.org/0000-0002-6404-9562
W. N. Brandt https://orcid.org/0000-0002-0167-2453
Patrick B. Hall https://orcid.org/0000-0002-1763-5825
Luis C. Ho https://orcid.org/0000-0001-6947-5846
Linhua Jiang https://orcid.org/0000-0003-4176-6486
Donald P. Schneider https://orcid.org/0000-0001-7240-7449
Charling Tao https://orcid.org/0000-0001-7961-8177
Yusra AlSayyad https://orcid.org/0009-0008-9216-7516
Matthew A. Bershady https://orcid.org/0000-0002-3131-4374
Michael R. Blanton https://orcid.org/0000-0003-1641-6222
Dmitry Bizyaev https://orcid.org/0000-0002-3601-133X
Kevin Bundy https://orcid.org/0000-0001-9742-3138
Yuguang Chen https://orcid.org/0000-0003-4520-5395
Megan C. Davis https://orcid.org/0000-0001-9776-9227
Kyle Dawson https://orcid.org/0000-0002-0553-3805
Xiaohui Fan https://orcid.org/0000-0003-3310-0131
Jenny E. Greene https://orcid.org/0000-0002-5612-3427
Hannes Gröller https://orcid.org/0000-0001-6327-2757
Yucheng Guo https://orcid.org/0000-0002-6137-0422
Héctor Ibarra-Medel https://orcid.org/0000-0002-9790-6313

Ryan P. Keenan https://orcid.org/0000-0003-1859-9640
Cassandra Lejoly https://orcid.org/0000-0003-0165-7701
Zefeng Li https://orcid.org/0000-0001-7373-3115
Jundan Nie https://orcid.org/0000-0001-6590-8122
Paul S. Smith https://orcid.org/0000-0002-5083-3663
Wei Leong Tee https://orcid.org/0000-0003-0747-1780
Anne-Marie Weijmans https://orcid.org/0000-0002-5908-6852
Jiachuan Xu https://orcid.org/0000-0003-0871-8941
Minghao Yue https://orcid.org/0000-0002-5367-8021

References

Alexander, T. 2013, arXiv:1302.1508
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bailey, J. 1998, MNRAS, 301, 161
Barth, A. J., Bennert, V. N., Canalizo, G., et al. 2015, ApJS, 217, 26
Barth, A. J., Nguyen, M. L., Malkan, M. A., et al. 2011, ApJ, 732, 121
Barth, A. J., Pancoast, A., Bennert, V. N., et al. 2013, ApJ, 769, 128
Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002
Bentz, M. C., Denney, K. D., Grier, C. J., et al. 2013, ApJ, 767, 149
Bentz, M. C., Walsh, J. L., Barth, A. J., et al. 2009, ApJ, 705, 199
Bentz, M. C., Williams, P. R., Street, R., et al. 2021, ApJ, 920, 112
Bentz, M. C., Williams, P. R., & Treu, T. 2022, ApJ, 934, 168
Blandford, R. D., & McKee, C. F. 1982, ApJ, 255, 419
Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
Bosco, F., Hennawi, J. F., Stern, J., & Pott, J.-U. 2021, ApJ, 919, 31
Brotherton, M. S., Du, P., Xiao, M., et al. 2020, ApJ, 905, 77
Busch, G., Zuther, J., Valencia, S. M., et al. 2014, A&A, 561, A140
Cackett, E. M., Bentz, M. C., & Kara, E. 2021, iSci, 24, 102557
Coatman, L., Hewett, P. C., Banerji, M., et al. 2017, MNRAS, 465, 2120
Collin, S., Kawaguchi, T., Peterson, B. M., & Vestergaard, M. 2006, A&A, 456, 75
Czerny, B., Olejak, A., Rałowski, M., et al. 2019a, ApJ, 880, 46
Czerny, B., Wang, J.-M., Du, P., et al. 2019b, ApJ, 870, 84
Dalla Bontà, E., Peterson, B. M., Bentz, M. C., et al. 2020, ApJ, 903, 112
De Rosa, G., Fausnaugh, M. M., Grier, C. J., et al. 2018, ApJ, 866, 133
De Rosa, G., Peterson, B. M., Ely, J., et al. 2015, ApJ, 806, 128
Denney, K. D. 2012, ApJ, 759, 44
Dexter, J., Xin, S., Shen, Y., et al. 2019, ApJ, 885, 44
Dix, C., Matthews, B., Shemmer, O., et al. 2023, ApJ, 950, 96
Donnan, F. R., Horne, K., & Hernández Santisteban, J. V. 2021, MNRAS, 508, 5449
Du, P., Lu, K.-X., Zhang, Z.-X., et al. 2016, ApJ, 825, 126
Du, P., & Wang, J.-M. 2019, ApJ, 886, 42
Du, P., Zhang, Z.-X., Wang, K., et al. 2018, ApJ, 856, 6
Dyer, J. C., Dawson, K. S., du Mas des Bourboux, H., et al. 2019, ApJ, 880, 78
Edelson, R. A., & Krolik, J. H. 1988, ApJ, 333, 646
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
Elvis, M., & Karovska, M. 2002, ApJL, 581, L67
Fausnaugh, M. M., Denney, K. D., Barth, A. J., et al. 2016, ApJ, 821, 56
Fonseca Alvarez, G., Trump, J. R., Homayouni, Y., et al. 2020, ApJ, 899, 73
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Fries, L. B., Trump, J. R., Davis, M. C., et al. 2023, ApJ, 948, 5
Gaskell, C. M., & Peterson, B. M. 1987, ApJS, 65, 1
Goad, M. R., Korista, K. T., & Ruff, A. J. 2012, MNRAS, 426, 3086
Graham, A. W., Onken, C. A., Athanassoula, E., & Combes, F. 2011, MNRAS, 412, 2211
GRAVITY+ Collaboration, Abuter, R., Allouche, F., et al. 2022, A&A, 665, A75
GRAVITY Collaboration, Amorim, A., Bauböck, M., et al. 2020, A&A, 643, A154
GRAVITY Collaboration, Amorim, A., Bauböck, M., et al. 2021, A&A, 654, A85
Gravity Collaboration, Sturm, E., Dexter, J., et al. 2018, Natur, 563, 657
Greene, J. E., & Ho, L. C. 2005, ApJ, 630, 122
Grier, C. J., Hall, P. B., Brandt, W. N., et al. 2015, ApJ, 806, 111
Grier, C. J., Pancoast, A., Barth, A. J., et al. 2017a, ApJ, 849, 146
Grier, C. J., Peterson, B. M., Horne, K., et al. 2013, ApJ, 764, 47
Grier, C. J., Peterson, B. M., Pogge, R. W., et al. 2012, ApJ, 755, 60
Grier, C. J., Shen, Y., Horne, K., et al. 2019, ApJ, 887, 38
Grier, C. J., Trump, J. R., Shen, Y., et al. 2017b, ApJ, 851, 21
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
Guo, H., Barth, A. J., & Wang, S. 2022, ApJ, 940, 20
Guo, H., Shen, Y., He, Z., et al. 2020, ApJ, 888, 58
Hemler, Z. S., Grier, C. J., Brandt, W. N., et al. 2019, ApJ, 872, 21
Ho, L. C., & Kim, M. 2014, ApJ, 789, 17
Homayouni, Y., Trump, J. R., Grier, C. J., et al. 2019, ApJ, 880, 126
Homayouni, Y., Trump, J. R., Grier, C. J., et al. 2020, ApJ, 901, 55
Hönig, S. F., Watson, D., Kishimoto, M., et al. 2014, Natur, 515, 528
Hoormann, J. K., Martini, P., Davis, T. M., et al. 2019, MNRAS, 487, 3650
Horne, K., Peterson, B. M., Collier, S. J., & Netzer, H. 2004, PASP, 116, 465
Kaiser, N., Burgett, W., Chambers, K., et al. 2010, Proc. SPIE, 7733, 77330E
Kara, E., Mehdipour, M., Kriss, G. A., et al. 2021, ApJ, 922, 151
Kaspi, S., Brandt, W. N., Maoz, D., et al. 2007, ApJ, 659, 997
Kaspi, S., Brandt, W. N., Maoz, D., et al. 2021, ApJ, 915, 129
Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, ApJ, 533, 631
Kelly, B. C. 2007, ApJ, 665, 1489
Kelly, B. C., & Shen, Y. 2013, ApJ, 764, 45
Kinemuchi, K., Hall, P. B., McGreer, I., et al. 2020, ApJS, 250, 10
King, A. L., Martini, P., Davis, T. M., et al. 2015, MNRAS, 453, 1701
Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv:1711.03234
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Li, I.-H. J., Shen, Y., Brandt, W. N., et al. 2019, ApJ, 884, 119
Li, J. I. H., Shen, Y., Ho, L. C., et al. 2021, ApJ, 906, 103
Li, J. I.-H., Shen, Y., Ho, L. C., et al. 2023, ApJ, 954, 173
Li, Y.-R., Songsheng, Y.-Y., Qiu, J., et al. 2018, ApJ, 869, 137
Li, Y.-R., & Wang, J.-M. 2023, ApJ, 943, 36
Li, Y.-R., Wang, J.-M., Hu, C., Du, P., & Bai, J.-M. 2014, ApJL, 786, L6
Li, Y.-R., Wang, J.-M., Songsheng, Y.-Y., et al. 2022, ApJ, 927, 58
Linzer, N. B., Goulding, A. D., Greene, J. E., et al. 2022, ApJ, 937, 65
Lira, P., Kaspi, S., Netzer, H., et al. 2018, ApJ, 865, 56
Liu, T., Merloni, A., Simm, T., et al. 2020, ApJS, 250, 32
MacLeod, C. L., Ivezić, Ž., Sesar, B., et al. 2012, ApJ, 753, 106
Malik, U., Sharp, R., Penton, A., et al. 2023, MNRAS, 520, 2009
Martínez-Aldama, M. L., Czerny, B., Kawka, D., et al. 2019, ApJ, 883, 170
Matsuoka, Y., Strauss, M. A., Shen, Y., et al. 2015, ApJ, 811, 91
McConnachie, A., Babusiaux, C., Balogh, M., et al. 2016, arXiv:1606.00043
Mejía-Restrepo, J. E., Lira, P., Netzer, H., Trakhtenbrot, B., & Capellupo, D. M. 2018, NatAs, 2, 63
Onken, C. A., Ferrarese, L., Merritt, D., et al. 2004, ApJ, 615, 645
Pancoast, A., Brewer, B. J., Treu, T., et al. 2014, MNRAS, 445, 3073
Park, D., Barth, A. J., Woo, J.-H., et al. 2017, ApJ, 839, 93
Park, D., Kelly, B. C., Woo, J.-H., & Treu, T. 2012, ApJS, 203, 6
Pei, L., Fausnaugh, M. M., Barth, A. J., et al. 2017, ApJ, 837, 131
Peterson, B. M. 1993, PASP, 105, 247
Peterson, B. M., Ferrarese, L., Gilbert, K. M., et al. 2004, ApJ, 613, 682
Proga, D., Stone, J. M., & Kallman, T. R. 2000, ApJ, 543, 686
Runnoe, J. C., Brotherton, M. S., Shang, Z., Wills, B. J., & DiPompeo, M. A. 2013, MNRAS, 429, 135
Shen, Y. 2012, ApJ, 757, 152
Shen, Y. 2013, BASI, 41, 61
Shen, Y., Brandt, W. N., Dawson, K. S., et al. 2015a, ApJS, 216, 4
Shen, Y., Greene, J. E., Ho, L. C., et al. 2015b, ApJ, 805, 96
Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008, ApJ, 680, 169
Shen, Y., Grier, C. J., Horne, K., et al. 2019a, ApJL, 883, L14
Shen, Y., Hall, P. B., Horne, K., et al. 2019b, ApJS, 241, 34
Shen, Y., & Ho, L. C. 2014, Natur, 513, 210
Shen, Y., Horne, K., Grier, C. J., et al. 2016, ApJ, 818, 30
Shen, Y., & Kelly, B. C. 2012, ApJ, 746, 169
Shen, Y., & Liu, X. 2012, ApJ, 753, 125
Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45
Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
Songsheng, Y.-Y., Li, Y.-R., Du, P., & Wang, J.-M. 2021, ApJS, 253, 57
Starkey, D. A., Horne, K., & Villforth, C. 2016, MNRAS, 456, 1960
Stone, Z., & Shen, Y. 2023, MNRAS, 524, 4521
Stone, Z., Shen, Y., Burke, C. J., et al. 2022, MNRAS, 514, 164
Suberlak, K. L., Ivezić, Ž., & MacLeod, C. 2021, ApJ, 907, 96
Sun, M., Grier, C. J., & Peterson, B. M. 2018, PyCCF: Python Cross Correlation Function for reverberation mapping studies, Astrophysics Source Code Library, ascl:1805.032
Sun, M., Trump, J. R., Shen, Y., et al. 2015, ApJ, 811, 42
Swann, E., Sullivan, M., Carrick, J., et al. 2019, Msngr, 175, 58
U, V., Barth, A. J., & Vogler, H. A. 2022, ApJ, 925, 52
van Groningen, E., & Wanders, I. 1992, PASP, 104, 700
Vanden Berk, D. E., Shen, J., Yip, C.-W., et al. 2006, AJ, 131, 84

Vestergaard, M., & Osmer, P. S. 2009, ApJ, 699, 800
Vestergaard, M., & Peterson, B. M. 2006, ApJ, 641, 689
Villafaña, L., Williams, P. R., Treu, T., et al. 2022, ApJ, 930, 52
Villafaña, L., Williams, P. R., Treu, T., et al. 2023, arXiv:2304.06764
Wang, J.-G., Dong, X.-B., Wang, T.-G., et al. 2009, ApJ, 707, 1334
Wang, J.-M., Songsheng, Y.-Y., Li, Y.-R., Du, P., & Zhang, Z.-X. 2020a, NatAs, 4, 517
Wang, S., Shen, Y., Jiang, L., et al. 2019, ApJ, 882, 4
Wang, S., Shen, Y., Jiang, L., et al. 2020b, ApJ, 903, 51
Waters, T., Kashi, A., Proga, D., et al. 2016, ApJ, 827, 53
Williams, P. R., Pancoast, A., Treu, T., et al. 2018, ApJ, 866, 75
Williams, P. R., Pancoast, A., Treu, T., et al. 2020, ApJ, 902, 74
Williams, P. R., Treu, T., Dahle, H., et al. 2021, ApJ, 911, 64
Wills, B. J., & Browne, I. W. A. 1986, ApJ, 302, 56
Woo, J.-H., Schulze, A., Park, D., et al. 2013, ApJ, 772, 49
Wu, Q., & Shen, Y. 2022, ApJS, 263, 42
Yang, Q., Shen, Y., Chen, Y.-C., et al. 2020, MNRAS, 493, 5773
Yu, Z., Kochanek, C. S., Peterson, B. M., et al. 2020, MNRAS, 491, 6045
Yu, Z., Martini, P., Penton, A., et al. 2023, MNRAS, 522, 4132
Zajaček, M., Czerny, B., Martinez-Aldama, M. L., et al. 2020, ApJ, 896, 146
Zhang, Z.-X., Du, P., Smith, P. S., et al. 2019, ApJ, 876, 49
Zu, Y., Kochanek, C. S., & Peterson, B. M. 2011, ApJ, 735, 80