

DRAFT VERSION AUGUST 3, 2017  
Preprint typeset using L<sup>A</sup>T<sub>E</sub>X style emulatej v. 12/16/11

## THE SLOAN DIGITAL SKY SURVEY REVERBERATION MAPPING PROJECT: COMPOSITE LAGS AT $z \lesssim 1$

JENNIFER I-HSIU LI<sup>1</sup>, YUE SHEN<sup>1,2,\*</sup>, KEITH HORNE<sup>3</sup>, W. N. BRANDT<sup>4,5,6</sup>, JENNY E. GREENE<sup>7</sup>, C. J. GRIER<sup>4,5</sup>, LUIS C. HO<sup>8,9</sup>, CHRIS KOCHANEK<sup>10</sup>, DONALD P. SCHNEIDER<sup>4,5</sup>, JONATHAN R. TRUMP<sup>11</sup>, KYLE S. DAWSON<sup>12</sup>, KAIKE PAN<sup>13</sup>, DMITRY BIZYAEV<sup>13,14</sup>, DANIEL ORAVETZ<sup>13</sup>, AUDREY SIMMONS<sup>13</sup>, ELENA MALANUSHENKO<sup>13</sup>

*Draft version August 3, 2017*

### ABSTRACT

We present composite broad-line region (BLR) reverberation-mapping lag measurements for H $\alpha$ , H $\beta$ , He II  $\lambda$ 4686 and Mg II for a sample of 144,  $z \lesssim 1$  quasars from the Sloan Digital Sky Survey Reverberation Mapping (SDSS-RM) project. Using only the 32-epoch spectroscopic light curves in the first 6-month season of SDSS-RM observations, we compile correlation-function measurements for individual objects and then coadd them to allow the measurement of the average lags for our sample at mean redshifts of 0.4 (for H $\alpha$ ) and  $\sim 0.65$  (for the other lines). At similar quasar luminosities and redshifts, the sample-averaged lag decreases in the order of Mg II, H $\alpha$ , H $\beta$  and He II. This decrease in lags is accompanied by an increase in the mean line width of the four lines, and is roughly consistent with the virialized motion for BLR gas in photoionization equilibrium. These are among the first RM measurements of stratified BLR structure at  $z > 0.3$ . Dividing our sample by luminosity, H $\alpha$  shows clear evidence of increasing lags with luminosity, consistent with the expectation from the measured BLR size-luminosity relation based on H $\beta$ . The other three lines do not show a clear luminosity trend in their average lags due to the limited dynamic range of luminosity probed and the poor average correlation signals in the divided samples, a situation that will be improved with the incorporation of additional photometric and spectroscopic data from SDSS-RM. We discuss the utility and caveats of composite-lag measurements for large statistical quasar samples with reverberation-mapping data.

*Keywords:* black hole physics – galaxies: active – line: profiles – quasars: general – surveys

### 1. INTRODUCTION

Reverberation mapping (RM) is a technique used to infer the size of the broad-line region (BLR) in active galactic nuclei (AGN) and quasars by measuring the time delay between the continuum and broad-line flux variations (e.g., Blandford & McKee 1982; Peterson 1993). Combining the BLR size with the virial velocity inferred from the width of the broad lines, one can derive a virial estimate of the mass of the black hole. This is the primary technique used to measure BH masses in AGN and quasars, and anchors secondary

methods of active BH mass estimation based on single-epoch spectroscopy (for a recent review, see, Shen 2013).

Over the past two decades, RM measurements have been performed for dozens of low-redshift ( $z < 0.3$ ) AGN and quasars (e.g., Peterson et al. 1998a; Kaspi et al. 2000, 2005; Peterson et al. 2002; Peterson et al. 2004; Bentz et al. 2009, 2010a, 2013; Denney et al. 2009, 2010; Rafter et al. 2011, 2013; Barth et al. 2011a,b, 2013, 2015; Grier et al. 2012a; Du et al. 2014, 2015, 2016; Hu et al. 2015), and the feasibility and potential of this technique for measuring BH masses and understanding the inner structure of AGN and quasars has been well demonstrated. In recent years, RM has been attempted at higher redshifts ( $z > 0.3$ ) and for high-luminosity quasars as well (e.g., Kaspi et al. 2007; Trevese et al. 2014; Jiang et al. 2016), but successful measurements are still rare in these regimes. An interesting new approach to RM is to use multi-object spectrographs to study hundreds of quasars simultaneously (MOS-RM, e.g., Shen et al. 2015a; King et al. 2015). In addition to the much improved observational efficiency, MOS-RM programs aim to detect time lags for uniformly selected samples of AGN and quasars at substantially higher redshifts and luminosities with multi-year time baselines to sample the slow variability patterns of these high-redshift objects.

We are conducting one of the first MOS-RM programs using the SDSS Baryon Oscillation Spectroscopic Survey (BOSS) spectrograph (Dawson et al. 2013; Smee et al. 2013) on the 2.5-m SDSS telescope (Gunn et al. 2006), accompanied by dedicated photometric monitoring using a number of ground-based wide-field imagers. Details of the SDSS-RM project are presented in Shen et al. (2015a). Analyses of the initial season (2014) of spectroscopic data have led to the first robust detections of BLR lags at  $z > 0.3$ , and a subset of 15 individual detections were reported by Shen et al. (2016).

<sup>1</sup> Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

<sup>2</sup> National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

\* Alfred P. Sloan Research Fellow

<sup>3</sup> SUPA Physics/Astronomy, Univ. of St. Andrews, St. Andrews KY16 9SS, Scotland, UK

<sup>4</sup> Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA, 16802, USA

<sup>5</sup> Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

<sup>6</sup> Department of Physics, 104 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA

<sup>7</sup> Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

<sup>8</sup> Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

<sup>9</sup> Department of Astronomy, School of Physics, Peking University, Beijing 100871, China

<sup>10</sup> Department of Astronomy, The Ohio State University, 140 West 18th Avenue, Columbus, OH 43210, USA

<sup>11</sup> Department of Physics, University of Connecticut, 2152 Hillside Rd Unit 3046, Storrs, CT 06269, USA

<sup>12</sup> Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA

<sup>13</sup> Apache Point Observatory and New Mexico State University, P.O. Box 59, Sunspot, NM, 88349-0059, USA

<sup>14</sup> Sternberg Astronomical Institute, Moscow State University, Moscow

The large sample of objects with RM data from the SDSS-RM project also enables an investigation of composite lag detections, where one may boost the signal-to-noise ratio (SNR) of the lag detection (inferred from the cross-correlation between continuum and line light curves) by stacking the results from individual objects (e.g., Fine et al. 2012, 2013). As illustrated by these authors, this technique is useful for measuring the *average* lag for a sample of quasars, even if the individual light curves are of insufficient quality to measure a reliable lag or if the correlated variability is buried in the random, uncorrelated intrinsic variability of each light curve. Composite lag measurements provide a complementary approach to individual lag measurements to perform RM studies for high- $z$  samples.

In this work we test the feasibility of composite lag detections using a subset of SDSS-RM data from our first-season observations conducted in 2014. We use the well-calibrated first-season (6-month) spectroscopic light curves *alone* to create the coadded correlation function. This slightly differs from Fine et al. (2013) which combined high-cadence continuum light curves with a few spectroscopic epochs, but the spirit is the same in both studies. This approach is carried out in parallel to our ongoing effort on individual lag detections, and it is particularly useful for lag detections with weak broad lines (such as He II). The motivations for performing such an exercise with the SDSS-RM data are the following: (1) with composite lags we will attempt to measure average lags for different line species, in particular the weak broad line He II  $\lambda 4686$ , in the same sample of quasars, allowing us to explore the stratified structure of quasar BLRs; (2) ongoing and future MOS-RM programs will produce light-curve data for large statistical quasar samples. Composite lag measurements then offer a promising way to boost the signal-to-noise ratio, and to measure the average lags for quasars binned by different physical properties. This work serves as a demonstration of concept for this approach using the first-season SDSS-RM spectroscopic-only data. Nevertheless, the lag measurements we present here are at mean redshifts  $> 0.3$ , where RM measurements are rare, and include the first measurement of (composite) He II lags at such high redshifts.

This paper is organized as follows. In §2 we describe the sample and light-curve data used, and in §3 we describe the technical details of stacked lag detection, with several tests to demonstrate its robustness in §3.2. We present our results in §4, and summarize our findings in §5, with an outlook for future improvement. Throughout the paper we adopt a flat  $\Lambda$ CDM cosmology with  $\Omega_{\Lambda} = 0.7$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  in calculating luminosities. Unless stated otherwise, all stacked lag measurements are performed in the observed frame for the reasons discussed in §3.

## 2. DATA

SDSS-RM simultaneously monitors 849 broad-line quasars at  $0.1 < z < 4.5$  with a flux limit of  $i_{\text{psf}} = 21.7$  (Shen et al. 2015a). The spectroscopic data used in this work are from the 32 epochs taken in 2014 (from January to July) as part of the SDSS-RM project within the SDSS-III (Eisenstein et al. 2011) Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al. 2013). The wavelength coverage of BOSS spectra is  $\sim 3650 - 10,400 \text{ \AA}$ , with a spectral resolution of  $R \sim 2000$ . Each epoch had a typical exposure time of 2 hrs, resulting a typical SNR of  $\sim 4.5$  per  $69 \text{ km s}^{-1}$  pixel at  $g_{\text{psf}} = 21.2$  averaged over the  $g$  band. The epoch-by-epoch spectra were

pipeline-processed as part of the SDSS-III Data Release 12 (Alam et al. 2015), followed by a custom flux calibration scheme and improved sky subtraction as described by Shen et al. (2015a). The improved spectrophotometry has a nominal absolute accuracy of  $\sim 5\%$ .

We then performed a spectral-refinement procedure, designated “PrepSpec”, on the custom flux-calibrated multi-epoch spectra, as described in detail in Shen et al. (2016). PrepSpec further improves the relative flux measurements by adjusting the flux levels of the individual epochs using the fluxes of narrow emission lines, assumed to remain constant over the monitoring period. The PrepSpec procedure derived continuum and broad-line light curves for each object in our sample, as well as measurements of the emission line widths from both the mean and the root-mean-squared (rms) spectra. These *spectroscopic-only* light curves form the basis of our composite-lag analysis.

Occasionally a particular epoch will appear as an outlier in the light curve. These events are usually caused by the poor SNR of that particular epoch (e.g., Epoch 3 and Epoch 7; see Shen et al. 2016), but occasionally may be due to unknown systematics in the PrepSpec procedure. To build a uniform set of light-curve data, which is crucial when carrying on to further correlation function and coadding calculations, we identify outlier epochs with differences larger than 3 standard deviations from the linear interpolations of the original light curves and replace their flux measurements with linear interpolations from the rest of the light curve. Similarly, we interpolate the light curve at rare epochs with missing data (mostly due to bad spectrograph fibers) and assign the flux error to be the average of the remaining data points. The fraction of these “fixed” light curve data points is 9.7% of total light curve points. Figure 1 presents an example of our light curve adjustment. With this adjustment, all objects have exactly the same cadence in their light curves, allowing a uniform binning of the correlation functions.

As discussed in Shen et al. (2016), one concern of our analysis is that PrepSpec may underestimate the light-curve uncertainties, and an empirical upper limit on the flux uncertainties is 3%. We have tested our calculations with the original PrepSpec errors and the inflated light curve errors, and found consistent results. To be conservative with the light-curve errors, all analyses presented in this paper will adopt the 3% inflated errors.

## 3. COMPOSITE LAG MEASUREMENTS

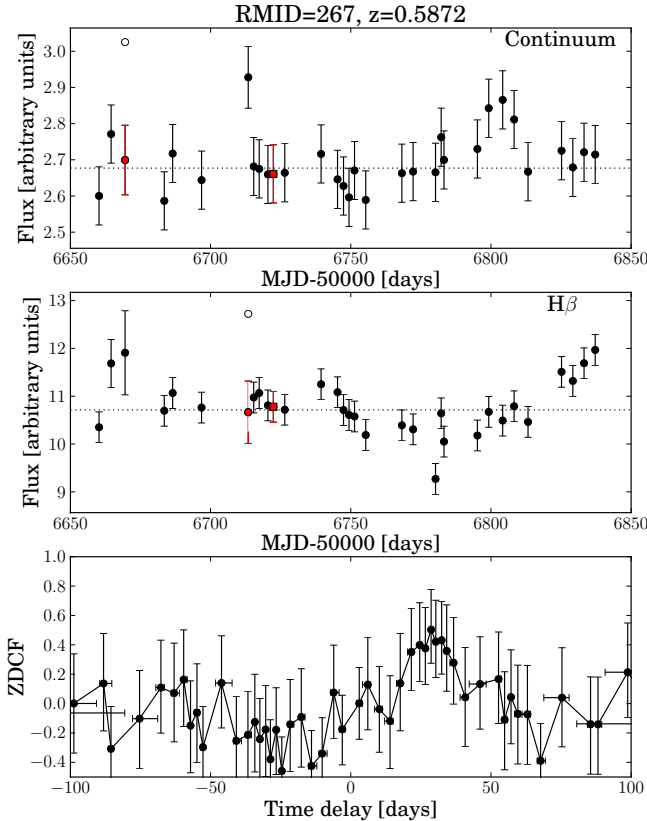
In this work we focus on the low- $z$  subset of our sample for which a lag should be detectable given the 6-month observed period of our first-season monitoring. Specifically we require that a continuum light curve at rest-frame  $5100 \text{ \AA}$  is available, which is the commonly adopted reference continuum for low- $z$  RM work. This requirement limits our sample to  $z \lesssim 1$  and 190 objects. We consider four broad lines that are of primary interest in this redshift regime and for which we have available light curves from PrepSpec:  $H\alpha$ ,  $H\beta$ , He II  $\lambda 4686$ , and Mg II.

Furthermore, we remove objects whose expected observed-frame  $H\beta$  lags (following the  $R-L$  relation in Bentz et al. 2013) exceed the maximum lag range searched for correlation ( $\sim 120$  days), given the total baseline of  $\sim 180$  days of our light curves. These objects will not contribute correlation signals to the coadded correlation function, but will add noise instead. In addition, these objects are typically at higher luminosity and redshift, and thus will bias the sample-averaged luminosity upwards. Indeed, we have tested including these

**Table 1**  
Composite Lags

Line	$N_{\text{obj}}$	$\langle z \rangle$	$\langle \log L_{5100} \rangle$ [ $\text{ergs}^{-1}$ ]	$\langle \tau \rangle$ (Simple Median)									$\langle \tau \rangle$ (Weighted Mean)									(line width)				$\langle \tau_{\text{rest}} \rangle$ (Rest Frame)		
				polynomial fit			direct centroid			polynomial fit			direct centroid			$\sigma_{\text{mean}}$	$\sigma_{\text{rms}}$	FWHM <sub>mean</sub>	FWHM <sub>rms</sub>	polynomial fit								
				median	16%	84%	median	16%	84%	median	16%	84%	median	16%	84%					median	16%	84%						
H $\alpha$ all	56	0.39	43.66	26.4	17.7	35.0	27.9	19.1	35.1	23.7	17.8	29.5	23.3	17.7	29.0	2251	2266	2668	4281	19.7	15.3	24.6						
other	43	0.39	43.64	22.9	13.8	41.0	24.3	14.4	38.9	21.7	13.2	34.4	21.8	13.5	31.4	2473	2479	2681	4662	19.2	11.9	28.2						
first lags	13	0.39	43.74	29.0	17.4	37.0	30.6	20.5	38.9	26.5	21.2	31.4	27.4	22.1	31.3	1464	1511	2621	2932	19.5	16.0	22.7						
low- $L$	28	0.32	43.29	17.3	6.5	27.1	17.0	8.5	25.4	16.1	7.1	23.3	15.4	6.5	23.3	2306	2284	2842	4424	14.9	8.6	20.0						
high- $L$	28	0.45	44.02	52.1	33.0	63.7	47.6	34.8	57.6	41.6	31.9	58.8	38.2	30.1	48.5	2196	2247	2493	4138	27.3	19.1	36.0						
H $\beta$ all	144	0.62	43.90	22.7	18.4	26.2	24.2	18.7	26.9	19.7	16.9	22.1	19.9	16.4	22.7	3915	3481	4471	7216	12.7	10.6	15.2						
other	129	0.64	43.92	18.4	11.3	24.5	19.1	12.4	25.5	15.9	11.7	19.5	15.0	11.0	19.0	4153	3696	4553	7628	9.3	6.5	11.9						
first lags	15	0.43	43.81	26.4	23.8	29.8	27.8	25.5	31.3	27.4	24.1	30.3	27.7	24.9	30.2	2107	1842	3839	4076	20.0	18.2	22.1						
low- $L$	72	0.51	43.59	21.5	17.1	24.7	21.7	16.0	25.1	19.0	16.6	21.6	18.3	15.1	21.8	3680	3363	4707	6983	12.4	9.9	14.1						
high- $L$	72	0.73	44.21	24.9	11.7	44.4	28.0	17.5	41.4	24.6	12.1	52.9	26.4	15.5	48.3	4148	3597	4237	7448	14.5	9.1	36.0						
He II all	144	0.62	43.90	12.7	6.4	18.3	12.4	5.3	18.7	8.2	2.0	15.6	9.8	3.2	16.3	6064	5163	5728	11753	3.8	1.3	7.2						
other	129	0.65	43.91	12.7	3.2	22.0	12.7	2.5	22.7	4.5	-2.2	16.0	6.1	-0.8	18.0	6411	5432	6006	12313	3.4	-0.2	8.0						
first lags	15	0.43	43.80	13.8	8.2	18.3	14.2	10.0	18.9	13.9	9.5	16.7	13.6	10.2	16.7	3374	3067	3570	7402	6.5	3.5	8.9						
Mg II all	127	0.69	43.99	38.2	24.6	56.8	39.0	28.3	49.1	45.9	35.4	64.0	38.4	30.4	46.3	3166	2969	3145	5985	35.9	20.9	42.9						
other	117	0.71	43.99	34.2	21.8	52.7	36.9	25.4	49.6	30.3	16.5	49.2	30.9	19.8	43.7	3294	3086	3198	6209	28.9	16.3	41.7						
first lags	10	0.53	43.93	52.7	40.8	58.8	43.2	30.4	52.8	51.9	44.4	59.4	44.7	39.3	52.1	1695	1621	2544	3421	33.6	23.9	38.0						

NOTE. — All lag measurements are in units of days in the observed frame except for the last three columns. We take these results based on the weighted mean and the polynomial fit as our fiducial results. For each line, the “other” sample excludes the 15 objects reported in Shen et al. (2016), and the “high- $L$ ” and “low- $L$ ” samples are the full sample divided at the median luminosity (for H $\alpha$  and H $\beta$  only). The sample-averaged redshift, (host-corrected) quasar continuum luminosity and line widths (denoted by “ $\langle \rangle$ ”) are the weighted means calculated using the same weights as for the weighted mean lags. The measurement uncertainties of the average line widths are negligible and not reported here (see text). The last three columns show the results for the stacks performed in the rest frame of individual quasars.



**Figure 1.** An example of the light curve adjustment process described in §2. The top two panels show the continuum (estimated at rest-frame 5100 Å) and broad-line light curves. The open points are the epochs we identified as outliers, and the red points are the “fixed” light curve data (circles are for outliers and squares are for missing data). The bottom panel shows the ZDCF. This object is one of the 15 first-lag objects reported in Shen et al. (2016) where a lag detection is possible with the spectro-only LCs.

objects in the coadded correlation functions, and found the resulting average lags are almost identical, but the sample-averaged luminosity is higher compared to the results excluding them. This additional requirement reduces the total num-

bers of objects used in this study to 144 (roughly  $\sim 25\%$  of sources are removed).

Table 1 lists the composite-lag measurements for these four lines and the corresponding number of objects in each coadd.

### 3.1. Methodology of composite lags

To measure the cross-correlation signal in individual objects, we use the  $z$ -transformed discrete correlation function (ZDCF, Alexander 2013). ZDCF is a more robust method to measure time lags for sparse and unevenly-sampled light curves than the traditional discrete correlation function (DCF, Edelson & Krolik 1988) and the interpolated cross-correlation function (ICCF, Gaskell & Peterson 1987). By using equal population binning and Fisher’s  $z$ -transform, ZDCF provides accurate estimates for light curves with as few as  $\sim 15$  epochs. Compared with the ICCF and DCF, ZDCF provides more reliable error estimation (whereas ICCF does not provide errors on the correlation), which is useful in weighting the data points in the coadded correlation; more importantly, unlike the ICCF, the ZDCF does not suffer from correlated errors in the continuum and broad-line light curves as measured from the same spectrum. Finally, given the uniform sampling of light curves used here, the automatic time-lag binning of ZDCF yields the same binning for all objects, allowing a straightforward coaddition of points in each time-lag bin. These properties of the ZDCF make it the ideal tool to measure the average time lag for a given sample in our database. The bottom panel of Figure 1 shows the ZDCF for an example object in our sample.

One drawback of using the ZDCFs in the stacking is that the individual cross correlations are stacked in the observed frame (unless some sort of interpolation and rebinning of the individual ZDCFs is used). Therefore there is an additional broadening of the stacked ZDCF due to different time-dilation factors in different objects. Fortunately, as further discussed in §4.1, when scaled by the sample-averaged  $(1+z)$  factor, observed-frame coadded lags (without rebinning) are consistent with the rebinned rest-frame coadded lags, without the complications of data interpolation. Given the many advantages of ZDCFs discussed above, we will use ZDCFs and the observed-frame stacks to demonstrate our methodology be-

low.

Once we have measured the individual ZDCF $_{i,j}$  for object  $i$  and time-lag bin  $j$ , and the associated errors  $\sigma_{i,j}$ , we create a composite ZDCF in two different ways. The first approach is a simple median coadd, where we take the median of the ZDCF distribution in each time-lag bin. In the second approach, we assign the individual ZDCF points in each time-lag bin an object-based weight

$$W_i = \text{median} \left[ \frac{1}{\sigma_{i,*}^2} \right], \quad (1)$$

and calculate the weighted mean and uncertainty in each time-lag bin. In this definition, the weight is on an object-by-object basis and is identical across all time-lag bins for the same object.<sup>16</sup> The weighted mean provides a better measurement of the coadded ZDCF by up-weighting signals from high-quality light curves, and we take these results based on the weighted mean as our fiducial results. Conversely, both the simple-median approach and the weighted-mean approach account for contributions from all objects and hence the resulting coadded correlation is not dominated by a few objects.

To quantify the inferred *average* lags, we use a second-order polynomial to iteratively fit the coadded ZDCF points within  $\pm 50$  days around the model peak until the peak position converges to within 10% of itself. We also measure a centroid by directly using the coadded ZDCF data points within  $\pm 50$  days of the best-fit polynomial peak and found consistent results. Bootstrap resampling of the objects contributing to the coadded ZDFs is adopted to estimate the values and uncertainties in the measured average lags. We adopt the median of the lag distribution from the bootstrap samples as the measured lag to provide more robust estimates than using the peak/centroid measured from the original coadded ZDCF.

To assess visually the statistical significance of the coadded lag detection, we generate mock light curve pairs by shuffling the real light curve epochs for each object and measure the coadded ZDCF from the individual ZDFs derived from the mock data. This Monte Carlo procedure was repeated for 100 realizations, and the median and the 16th and 84th percentiles of the distribution of the coadded ZDCF are recorded as the expected  $1\sigma$  range of coadded signals expected from uncorrelated light curves. By randomly shuffling the real light curve epochs we destroy any intrinsic correlation.<sup>17</sup> For well detected composite lags, the coadded ZDCF should lie above the expected uncertainties from uncorrelated light curves.

As a consistency check, we coadded the ICCFs of individual objects and found consistent signals. However, there is often extra signal near zero lag due to correlated errors between the continuum and broad-line light curves (see discussions in Shen et al. 2016). The calculation of the ZDCF avoids such complications, and provides a cleaner coadded correlation.

A coadded correlation function is a diluted and broadened version of the individual correlation functions because different objects have different intrinsic lags and redshift dilation, and because low-variability light curves contribute both signal

<sup>16</sup> We adopt this weighting scheme instead of using different weights  $1/\sigma_{i,j}^2$  for the same object across time-lag bins to avoid “glitches” in the  $1/\sigma_{i,j}^2$  distribution due to unknown systematics in our light curves. The individual-epoch weighting scheme can lead to large fluctuations in the coadded ZDCF across time-lag bins.

<sup>17</sup> This shuffled-epoch test (e.g., Shen et al. 2016) also removes the characteristic red noise of quasar variability, but such variability will not introduce correlated continuum and line variability on the timescales of interest here.

and noise. If appropriately weighted, this averaging process should reduce the noise from individual correlation measurements, and boost the SNR of the average lag enough to make a detection possible. We intentionally use all available objects so that the average lag represents the entire quasar sample, even though the inclusion of the low-variability subset of the sample may lead to noisier measurements.

In addition to coadding all objects, we also coadded the ZDCF for all sources excluding the 15 first-lag sources reported in Shen et al. (2016) (the “other” sample). For H $\alpha$  and H $\beta$ , we are able to further divide the sources into low and high luminosity bins since the coadded ZDCF peaks are more significant for these two lines. We calculated 16 sets of coadded ZDCF in total, and the results are presented in Table 1 and discussed in § 4.

Although our default results are based on the coadded ZDCF in the observed frame, we also compute average lags from the coadded ZDCF in the rest-frame of the quasars with rebinning and interpolation of the light curves. The results based on the rest-frame coadds are listed in the last two columns of Table 1.

### 3.2. The robustness of composite lags

Before we present the composite lag results for our SDSS-RM sample, we perform a series of additional tests to demonstrate the feasibility and robustness of this coadding technique.

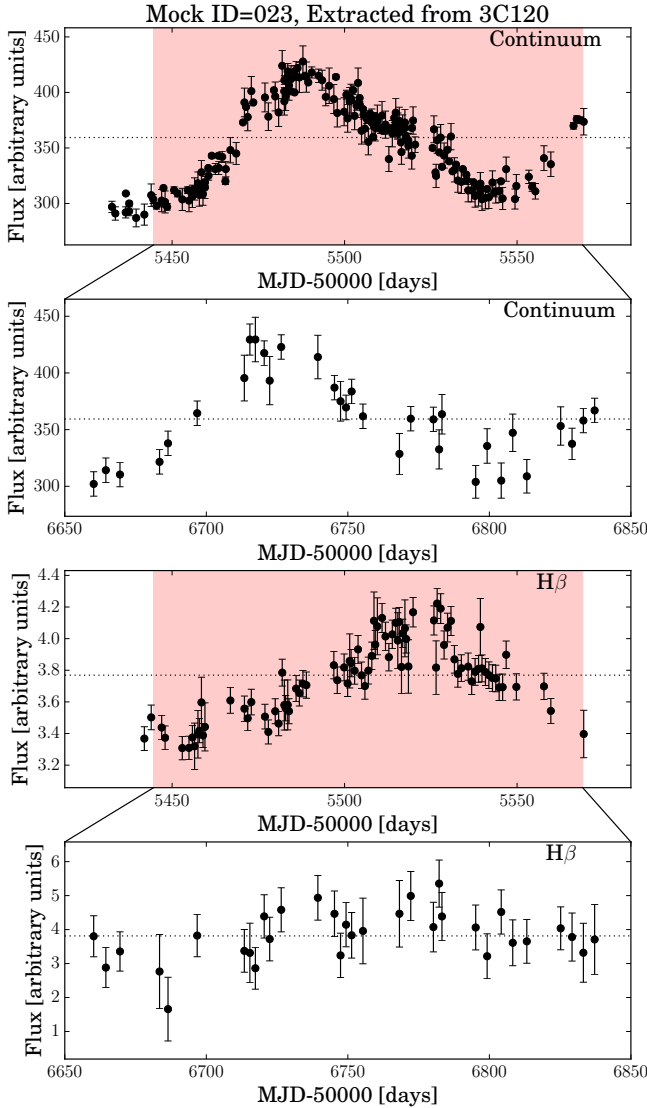
#### 3.2.1. Test on the local RM sample

The first test is to perform exactly the same coadding technique as described in §3.1 on the Bentz et al. (2013) sample, which includes 71 pairs of LCs from 41 local sources. This is a crucial test to evaluate the reliability of the coadding technique with a sample of known lags.

To match our sample size and observing period, we first generate 144 pairs of mock LCs by randomly choosing LC segments of individual objects from the Bentz et al. (2013) sample, which are assigned redshifts drawn from the redshift distribution of SDSS-RM quasars. The LC segments are required to span 180 days (in the observed-frame with the assigned redshift) with both continuum and line observations. We then degrade the mock LCs by increasing the measurement errors such that the RMS variability of the LCs (normalized by measurement errors) matches that of the SDSS-RM sample on average. This is because the local RM AGN typically have larger variability amplitudes than SDSS-RM quasars at higher luminosities. Finally, we match the exact cadence of SDSS-RM on the mock LCs with linear interpolation. See Fig. 2 for an example of the original and degraded LCs.

For sources with short observing durations and high cadences (i.e., corresponding to short lags), we deliberately assign high redshifts ( $z > 1$ ) to ensure they can be observed with our observing cadence. For the few sources with long lags or sparse LCs, we do not include them in our test as the lags cannot be detected within the period of 180 days. 51 out of the 71 LC pairs are used for generating the mock LCs and the coadding. The resulting sample of mock LCs has a similar size, cadence and variability/noise ratio to our SDSS-RM sample. The mock LCs are of too poor quality to formally detect a lag in individual objects.

We then perform the coadding technique as described previously in §3.1 on all mock LCs and two luminosity-divided



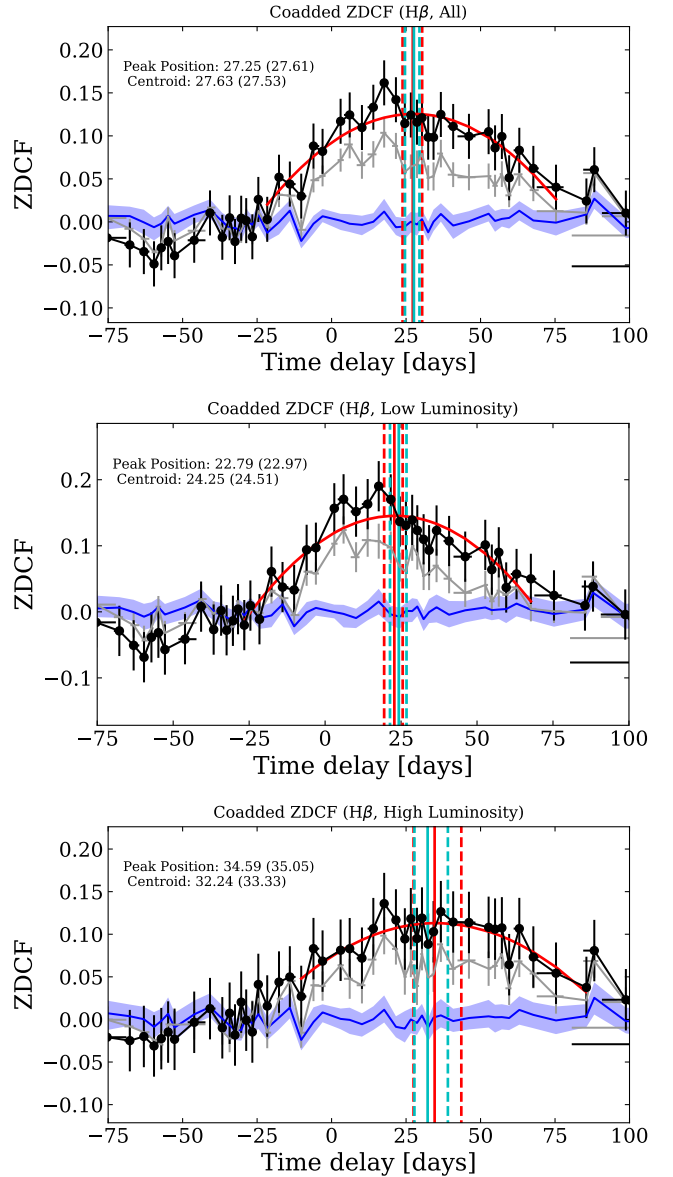
**Figure 2.** An example of a mock LC generated as described in §3.2.1. From the top to bottom panels are the original continuum LC, mock continuum LC, original line LC and mock line LC. The red shaded areas show the segments of the original LCs that are used to generate the mock LCs.

subsets. The results are shown in Fig. 3. Objects in the Bentz et al. (2013) sample have larger variability than SDSS-RM objects, and the coadded ZDCF yields a well-defined peak. We show the coadded results (after de-redshifting to the rest-frame) on the original  $R-L$  plot from Bentz et al. (2013) in Fig. 4. The results agree with the Bentz et al. (2013)  $R-L$  relation nicely for the full sample and the luminosity-divided samples. It is worth noting that the coadded lags are measured from the coadded correlation function, rather than from the simple average of individual lags, which are usually difficult to measure in low-quality LCs.

This test demonstrates that the coadding technique can yield consistent, sample-averaged lags.

### 3.2.2. SNR Dependence

We further test the robustness of the coadding technique at different SNR of the light curves. Again, we start from the mock sample generated from the Bentz et al. (2013) sample as described in §3.2.1. Next, we further degrade the sample SNR by applying a constant inflating factor to the flux errors

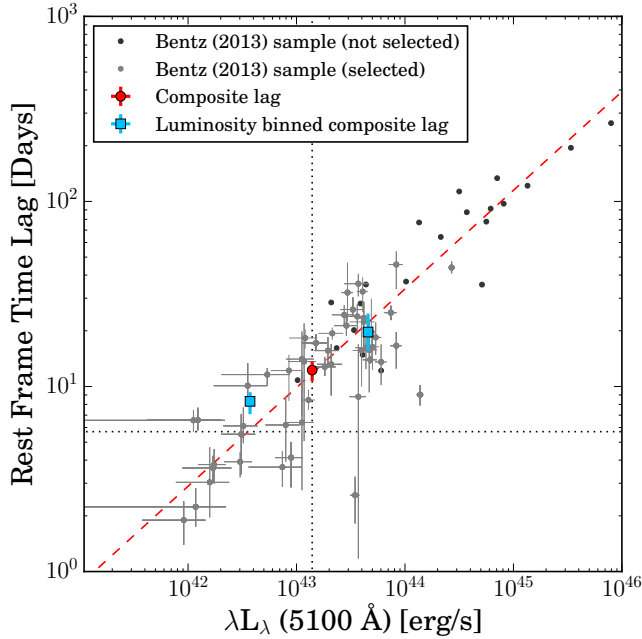


**Figure 3.** Coadded ZDCF using the LCs from the Bentz et al. (2013) sample, following the methodology detailed in §3.1. From top to bottom, the panels show the full sample, and the low and high luminosity subsets. In each panel, the blue line and the shaded band are the expected signal and uncertainty (16th and 84th percentiles of the distribution) generated from mock light curves with no intrinsic lags. The black data points show the coadded ZDCF with the weighted mean. Red curves are the polynomial fits to the coadded ZDCF within  $\pm 50$  days around the peak. Bootstrap resampling is used to estimate the uncertainty in the peak measurements. The estimated peak of the correlation (median of the bootstrap distribution) and its  $1\sigma$  uncertainties are indicated by the red (polynomial fit) and cyan (direct centroid) vertical lines. The values measured from the median of the bootstrap distribution are marked in the upper-left corner, with the values measured from the original ZDCF indicated in the parentheses. For comparison, the grey line shows the coadded ZDCF with the simple median approach

of all the LCs in the sample (but keeping the original LC flux unchanged). These inflated LC errors are propagated to the calculation of the individual ZDCF. This way, the individual ZDCF will have lower significance and larger uncertainties. Finally, we measure the coadded ZDCF and lags on the degraded samples following the same methodology as before.

Fig. 5 shows the coadded ZDCFs and measured average lags as a function of SNR (i.e., the error inflating factor). As





**Figure 4.**  $H\beta$  BLR radius and luminosity scatter plot from Bentz et al. (2013). The grey dots shows the data points used for generating mock LCs in our test and the black dots were excluded due to insufficient coverage or lags exceeding our baseline. The red dashed line is the  $R-L$  relation derived by Bentz et al. (2013). The red circle and the cyan squares are the average lags of the whole and luminosity-divided samples calculated by the coadding technique, with negligible error bars. The average cadence of the SDSS-RM observation and the median luminosity of the mock sample are indicated by the horizontal and vertical dotted lines, respectively.

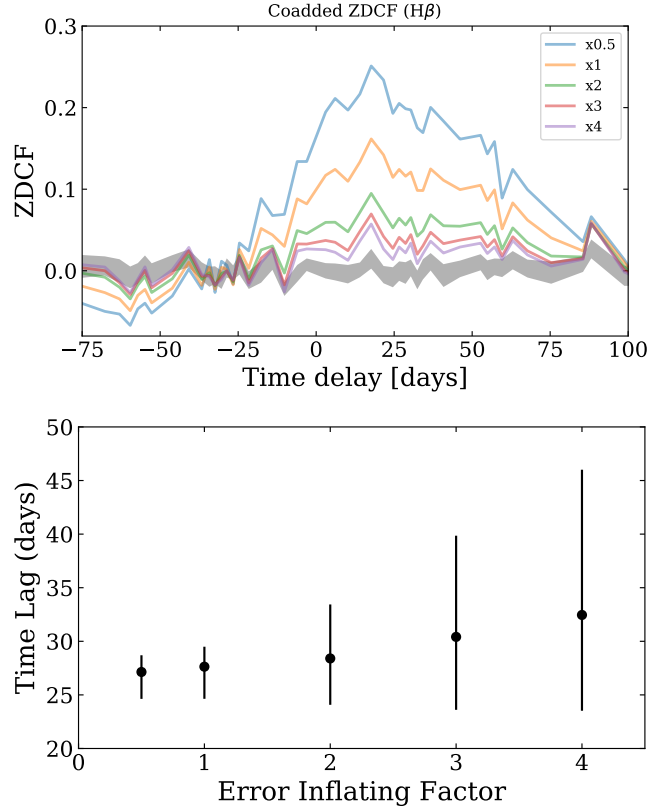
SNR decreases, the significance of the coadded ZDCF drops and the uncertainty<sup>18</sup> in the measured coadded lag increases, as expected. However, there is no significant bias in the inferred average lags when the SNR decreases, indicating that the coadding technique is stable against SNR. In addition to lowering the SNR of the LCs, we also tested a case where we reduce the LC errors by 50% and show the results in Fig. 5. As expected, the resulting coadded ZDCF has a stronger signal and the composite lag has a smaller uncertainty than before. This demonstrates the general utility of the coadding technique to strengthen a detection.

We note that lowering the SNR of the light curves is equivalent to decreasing the intrinsic variability of the light curves while keeping the SNR fixed. Lower-variability light curves require better SNR to be able to detect the correlated signals. On the other hand, the composite method cannot boost the signal-to-noise of the coadded lags indefinitely. If the individual light curves are of too poor quality, then even the composite method will not be able to measure a meaningful average lag of the sample.

### 3.2.3. Composite lags based on alternative flux measurements

Measuring fluxes from a quasar spectra can be a difficult task, which is particularly true for weak broad lines (such as  $He\ II$ ) and for cases where the decomposition of the continuum, broad- and narrow-line emission is ambiguous. As described in detail in Shen et al. (2016), PrepSpec models each individual spectrum with multiple components to account for

<sup>18</sup> The uncertainty in the measured average lag includes both the measurement uncertainty from the noise in the LCs and the systematic uncertainty from the sample variance and the coadding technique.

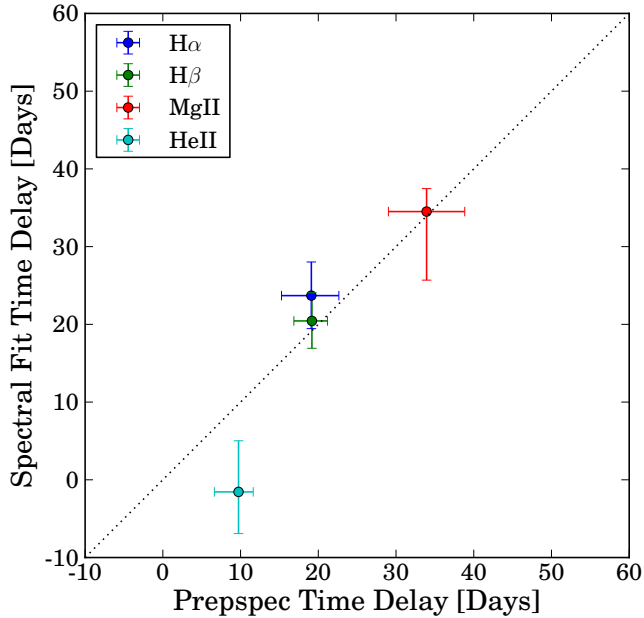


**Figure 5.** Effects of the SNR on estimates of the composite lag. The top panel shows the coadded ZDFs from the mock sample generated from the Bentz et al. (2013) local RM sample (see §3.2.1) as a function of the constant error inflating factor applied to the sample. The gray shaded band shows the expected signal from random, uncorrelated light curves for the original SNR. The bottom panel shows the average lag measured from the coadded ZDCF as a function of the error inflating factor. As the SNR decreases, the significance of the coadded ZDCF decreases and the uncertainty in the measured average lag increases, as expected. Nevertheless, there is no significant bias in the measured average lag as SNR decreases, suggesting that the coadding technique is robust even for low SNR data.

the continuum, the broad lines and the narrow lines. It is possible that fluxes measured with alternative approaches differ from the PrepSpec fluxes in a systematic way. Fortunately, the spirit of the composite lag technique is to average out these potential systematic uncertainties in the flux measurements, and provide unbiased average results. Therefore we expect our results are insensitive to systematic uncertainties in the light curve flux measurements in individual objects.

To test the above statement, we use continuum and line light curves measured with the independent spectral fitting approach of Shen et al. (2011). A full description of the spectral fits to SDSS-RM quasars will be presented elsewhere (Shen et al., in preparation). In short, we fit the continuum and  $Fe\ II$  emission underneath the broad lines, as well as the adjacent narrow lines, and extract the broad-line flux from functional fits. The main difference between the alternative spectral fits and PrepSpec is the explicit inclusion of the  $Fe\ II$  emission in the former approach, but these two approaches also differ in many technical details in how the continuum and lines are modeled. In general, the alternative spectral fits have larger measurement uncertainties in the fluxes due to additional model components (e.g.,  $Fe\ II$  emission).

We then apply the same coadding method on the set of light curves based on the alternative flux measurements and



**Figure 6.** Comparison of the observed composite lags based on light curve fluxes measured with a different spectral fitting approach with the fiducial results based on PreSpec. Due to the more complicated decomposition in the alternative spectral fits, the flux measurements have larger uncertainties, which led to the non-detection of the He II lag.

measure the average lags. The results for the 15 first-lag objects reported in Shen et al. (2016) are compared with the fiducial results based on PreSpec fluxes, as shown in Figure 6. Coadding all objects yields similar results. We did not detect the He II lag with the alternative flux measurements (i.e., the lag is consistent with zero) due to larger measurement uncertainties in the light curve fluxes. Nevertheless, this test demonstrates that the average lags measured from the coadding technique are robust against the details of the flux measurements, as expected from the nature of averaging.

#### 4. RESULTS

With the extensive tests in §3.2, we have demonstrated that the composite lag technique is meaningful, and robust against details in the individual correlation function measurements. We now present the composite lag measurements for our SDSS-RM sample.

##### 4.1. Coadded lags and errors

Figure 7 presents the coadded ZDCFs for the four lines using all objects with available PreSpec light curves. In all cases, a peak with the line lagging the continuum is present in the ZDCF. However, since the SDSS-RM quasars generally have lower variability amplitudes than those of the local RM AGN sample, the uncertainties in the average lags are substantially larger than in our tests using the local RM sample.

To demonstrate that the coadded ZDCF is not dominated by a few objects with well-detected individual lags, we remove the 15 objects reported in Shen et al. (2016) and calculate the coadded ZDCF for the remaining objects. As these 15 objects were the ones with individual lag detections, they have the ZDCFs with the highest SNR. The results are listed in Table 1 as the “other” samples<sup>19</sup>. We still detect statistically signifi-

<sup>19</sup> We note that some of the objects in the “other” sample can have individually detectable lags if the light curves are of better quality. For example,

cant lags that are consistent with the full sample, demonstrating the effectiveness of the coadding technique.

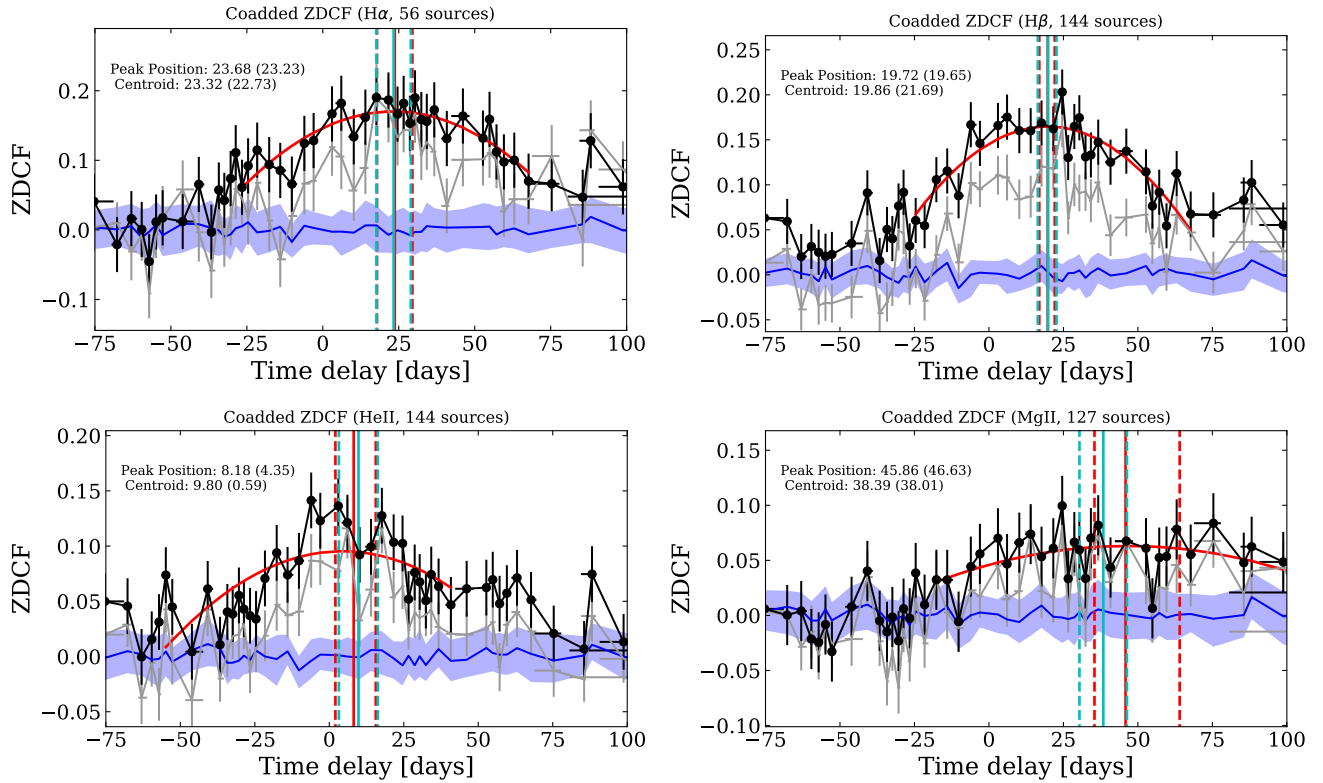
As a sanity check, coadding the 15 objects reported in Shen et al. (2016) produces much more significant signals (higher ZDCF peak amplitudes and narrower peaks) given the better quality of individual ZDCFs for these objects than for the rest of our sample (see Figure 8). The composite H $\beta$  and Mg II lags are also consistent with the median of the individual lags reported in Shen et al. (2016). The relative durations of the composite lags for different lines are consistent with results based on the full sample, as discussed below.

Using the ZDCF and coadding in the observed frame avoids interpolation and rebinning of the individual correlation functions, which will lead to correlated errors and complicate the interpretation. However, a caveat of this approach is that objects at different redshifts will have different time dilations of their intrinsic lags. This will result in an additional broadening of the coadded correlation function compared to stacking in the rest-frame of these quasars. We have tested coadding the individual ZDCFs in the rest frame of the quasars by shifting and interpolating the original ZDCFs onto a common rest-frame time-lag grid. We found consistent results (to within  $\sim 1\sigma$ ) and uncertainties compared to the stacks performed in the observed frame after scaling the latter by a factor of  $1 + \langle z \rangle$ . The results for the rest-frame stacks are provided in the last three columns in Table 1 for completeness. Stacking in the rest frame does not produce significantly smaller error bars for the average lags, which may be due to the possibility that objects with different luminosities (which presumably have different lags) already significantly broaden the stacked ZDCF, and the additional broadening due to redshift is subdominant. If we assume the  $R \propto L^{0.5}$  relation from Bentz et al. (2013), the dispersion in the observed-frame lags due to luminosity is a factor of  $\sim 2-3$  larger than that due to the  $(1+z)$  factor in the four line samples.

##### 4.2. Lags for different line species

Figure 9 shows the average lags as a function of sample-averaged continuum luminosity for the four lines, where the lags have been shifted to the rest frame using the sample-averaged redshift. To remove host-starlight contamination from the rest-frame 5100 Å continuum-luminosity measurements, we adopted a spectral-decomposition technique detailed in Shen et al. (2015b) to derive the quasar-only luminosity. We use the same weights as for the coadded ZDCFs to compute a weighted-mean luminosity for the sample. The H $\alpha$  sample has a relatively lower average redshift than other three line samples, and thus has the lowest average luminosity. For the other three lines, their samples have similar continuum luminosities and redshifts, allowing a fair comparison of their average lags. If we extrapolate the average H $\alpha$  lag to higher continuum luminosity assuming  $\tau \propto L^{0.5}$  (i.e., consistent with the measured  $R-L$  relation for H $\beta$ , Bentz et al. 2013), we derive the open data point in Figure 9 at the sample averaged luminosity of the H $\beta$  sample. Although the error bars overlap, Figure 9 reveals some evidence that He II has, on average, shorter lags than H $\beta$ , and H $\beta$  has shorter average lags than H $\alpha$ . While having large error bars, the composite Mg II lag appears to be slightly longer than that for H $\alpha$ . These results

our recent SDSS-RM work incorporating additional photometric light curves reported additional H $\alpha$  and H $\beta$  lags (Grier et al. 2017). However, these lags are non-detections with the spectroscopic-only light curves used here, and the composite method is the only way to recover an average lag for these objects.



**Figure 7.** Coadded ZDCFs (points) for the four broad emission lines considered in this work. In each panel, the blue line and the shaded band are the expected signal and uncertainty (16th and 84th percentiles of the distribution) generated from mock light curves with no intrinsic lags. The black data points show the coadded ZDCF with the weighted mean, which we take as the fiducial coadded results. Red curves demonstrate polynomial fits to the coadded ZDCF  $\pm 50$  days around the peak. Bootstrap resampling is used to estimate the uncertainty in the peak measurements. The estimated peak of the correlation (median of the bootstrap distribution) and its  $1\sigma$  uncertainties are indicated by the red (polynomial fit) and cyan (direct centroid) vertical lines. The values measured from the median of the bootstrap distribution are marked in the upper-left corner, with the values measured from the original ZDCF indicated in the parentheses. For comparison, the grey line shows the coadded ZDCF with the simple median approach, which has a lower but still statistically significant amplitude for the lag detection.

are qualitatively consistent with earlier individual RM measurements of different line species in the same objects (e.g., Peterson & Wandel 1999; Peterson et al. 2004; Bentz et al. 2010a; Grier et al. 2012b, 2013). In addition, the tentative evidence that Mg II has on average longer lags than  $H\beta$  is also consistent with the finding that Mg II varies less than  $H\beta$  on the same timescales (e.g., Sun et al. 2015), suggesting that the Mg II gas may be located slightly further out than the  $H\beta$  gas.

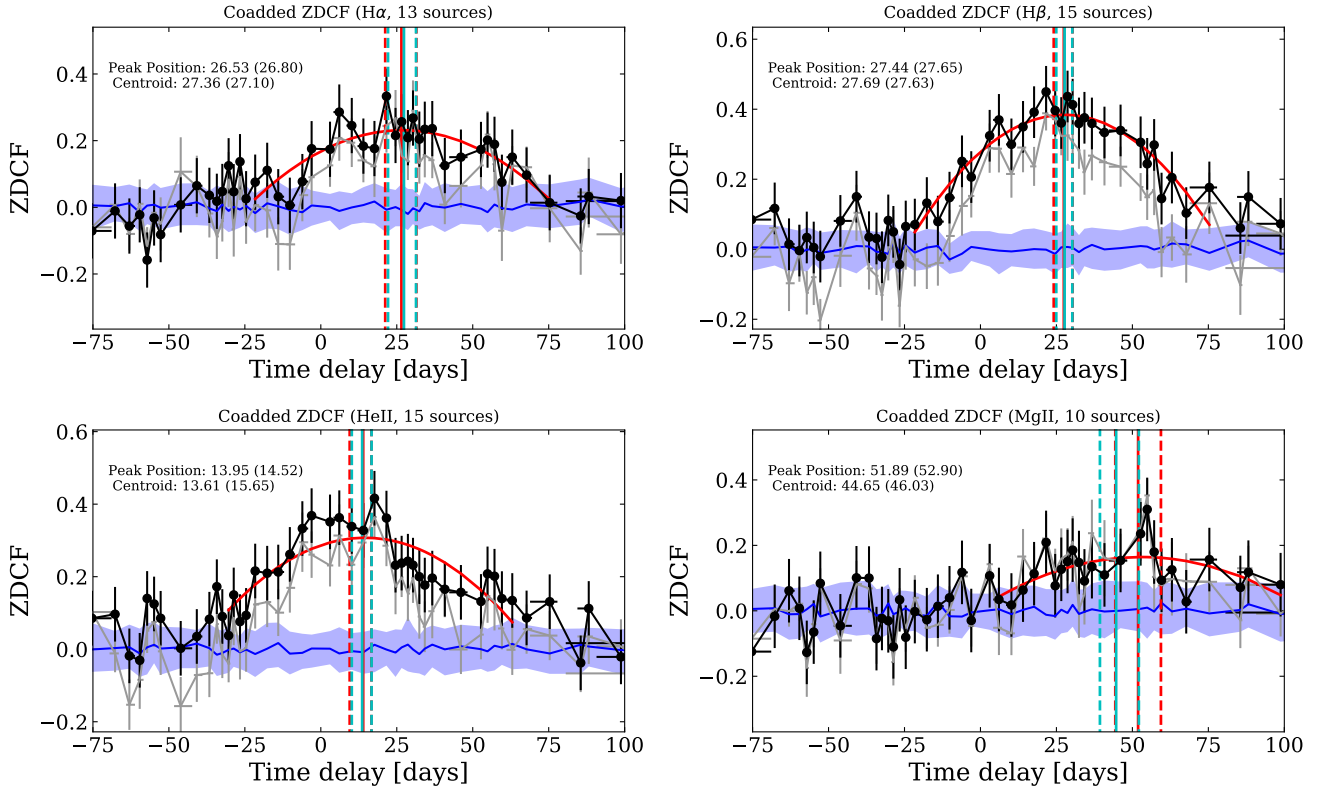
However, Figure 9 suggests that the average  $H\beta$  lag is significantly shorter than the expected lag from the measured  $R-L$  relation based on individual RM measurements in the low- $z$  sample (e.g., Bentz et al. 2013). We speculate that the main reason for this discrepancy is due to imperfectly weighted mean redshifts and luminosities for the sample. When averaging the ZDCF for a sample covering a wide redshift and luminosity range, objects that have more obvious (higher amplitude) correlations contribute more power to the peak location in the final coadded ZDCF in a non-trivial way. Lower- $z$  and lower-luminosity objects are likely to have a stronger influence on the determination of the composite lag than higher- $z$  and higher-luminosity objects, given the higher S/N in the light curves and stronger intrinsic variability. As a result, the sample-averaged redshift and luminosity using the weights based on measurement errors in the individual ZDCFs (instead of the amplitude of the ZDCFs) may overestimate the true sample-averaged redshift and luminosity. Unfortunately, without a robust weighting scheme to account for this complication, we can only qualitatively explain this dis-

crepancy, but cannot correct the bias quantitatively. The solution would be to coadd objects roughly at the same redshift and luminosity, which should become possible when we have better individual ZDCF measurements (see §5) or with future, larger MOS-RM samples. Nevertheless, this caveat should affect all lines for the same set of objects, and hence the relative lags of different lines should be robust.

Another possible explanation of the discrepancy is that most of the objects included in the coadded ZDCF have different accretion rates and hence different spectral energy distributions than those for the local sample used to measure the  $R-L$  relation for  $H\beta$ . Recently, Du et al. (2015, 2016) suggested that objects with higher accretion rates have significantly shorter BLR lags compared to lower accretion-rate objects. Quasars with higher Eddington ratios also vary less than those with lower Eddington ratios (e.g., Ai et al. 2013), leading to a potential selection bias in the sample with robust lag detections or target samples for RM campaigns. The local sample used to measure the  $R-L$  relation is dominated by relatively low accretion-rate objects (e.g., Figure 1 of Shen et al. 2015a) compared to the general population of SDSS-RM quasars. Although it is unlikely that this is the complete explanation, we will investigate this possibility further in future work to determine if it can account for at least some of the discrepancy seen here.

It is worth noting that our recent work on individual lag measurements based on both spectroscopic and photometric data (Grier et al. 2017) measured  $\sim 40 H\beta$  lags and 18  $H\alpha$  in





**Figure 8.** Same format as Figure 7, but for the 15 first-lag objects reported in Shen et al. (2016).

the SDSS-RM sample. Complementary to the composite approach, these individual  $H\beta$  lags are also significantly shorter than the expectation from the local  $R-L$  relation on average. We have tested by coadding the objects with detections in Grier et al. (2017) using spectro-only LCs. We got rest-frame composite lags of  $16^{+5.3}_{-4.6}$  days for  $H\alpha$  and  $14^{+1.8}_{-1.8}$  days for  $H\beta$ , fully consistent with the average lags from Grier et al. (2017). This again demonstrates that our composite approach is robust in recovering a sample-averaged lag.

#### 4.3. Luminosity and line-width dependence of time lags

We further divide each sample at the median luminosity and investigate the luminosity dependence of the average lags. For  $H\alpha$ , a clear luminosity dependence of the average lag is seen (time lags are shorter for the lower luminosity subsample) after the  $(1+z)$  time dilation is included (see Table 1), and is roughly consistent with the measured  $R-L$  relation for  $H\beta$  (e.g., Bentz et al. 2013). However, the luminosity trend is poorly constrained for  $H\beta$ , He II and Mg II, both due to the weak coadded ZDCF signals in the divided samples, and to the narrow dynamic range in luminosity covered by the full samples. We further tested dividing the sample according to the expected lags in the observed frame using the  $R-L$  relation for  $H\beta$  in the local sample (Bentz et al. 2013), but the resulting composite lags are again too noisy to reveal any conclusive trends for  $H\beta$ , He II and Mg II. We plan to investigate the luminosity dependence further with the inclusion of photometric data in our analysis in the future.

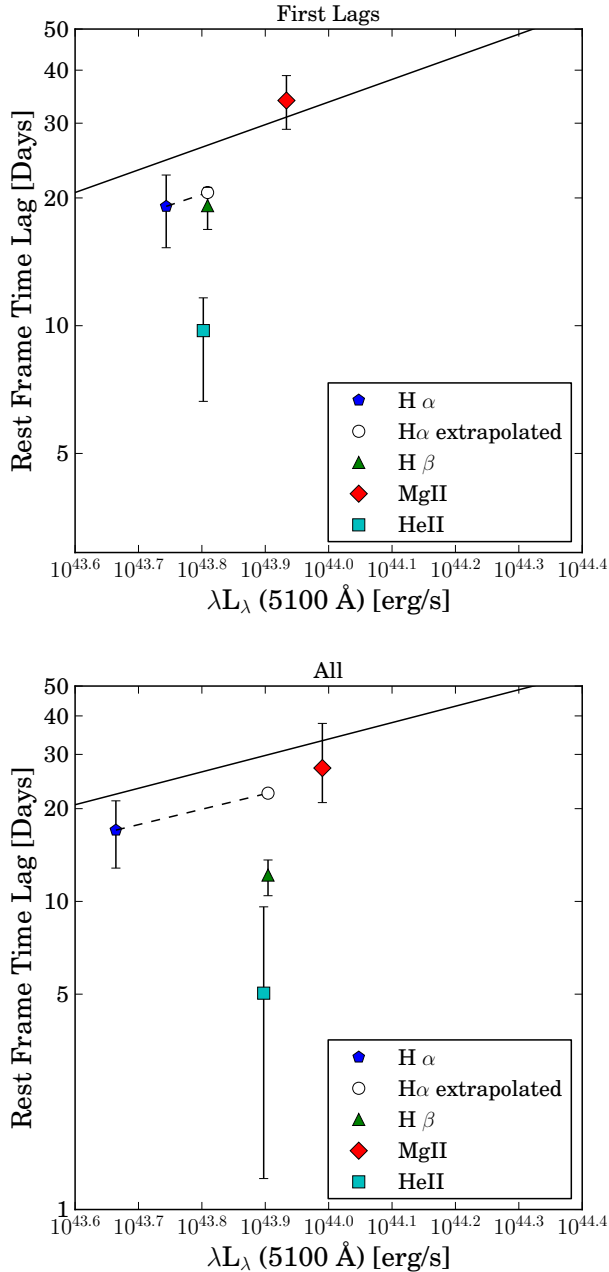
Figure 10 shows the relationship between the average lag and the sample-averaged (weighted-mean) line width for the four lines. We use four definitions of line width: the line dispersion (square root of the second moment of the line)  $\sigma_{\text{line}}$  measured from the mean and rms spectra, and the full-width-

at-half-maximum (FWHM) measured from the mean and rms spectra. The rms spectra are produced by PrepSpec after decomposing the line emission and the continuum emission. All four line-width definitions are consistent with the virial prediction, and the luminosity-extrapolated  $H\alpha$  point matches as well. Lines with shorter lags in general have larger velocity widths (e.g., Peterson & Wandel 1999; Grier et al. 2013), consistent with the scenario that there is a stratification of the BLR for different lines with different ionization potentials such that high-ionization lines (e.g., He II) are on average closer to the BH and hence have larger line widths.

## 5. CONCLUSIONS

Using the large sample of quasars with RM data from the SDSS-RM project, we have tested the feasibility of measuring composite lags for statistical samples. We applied the ZDCF method on the  $z \lesssim 1$  subset of the SDSS-RM quasar sample, and measured composite lags for  $H\alpha$ ,  $H\beta$ , He II and Mg II at average redshifts  $> 0.3$ . Compared to the earlier work on composite RM (Fine et al. 2013), our work focused on a different regime of redshift and luminosity with the unique SDSS-RM sample, and provided composite-lag measurements for broad lines other than Mg II and C IV. The findings from this work are the following:

1. When luminosity and redshift are matched, the sample-averaged lag decreases in the order of Mg II,  $H\alpha$ ,  $H\beta$ , and He II, suggesting that there is a stratification of the BLR, with high-ionization lines closer to the ionizing continuum than low-ionization lines. In particular, Mg II may have slightly longer lags on average than  $H\beta$ , suggesting that the Mg II gas may be located at a larger distance from the black hole than the  $H\beta$  gas.



**Figure 9.** The measured average lags of the four broad emission lines from the polynomial fit to the coadded ZDCF calculated using the weighted mean method, as a function of the weighted mean luminosity of the objects in the coadd (top panel: 15 first-lag objects reported in Shen et al. (2016); bottom panel: all sources in the sample). For  $H\alpha$ , we also show the expected lag at the same average redshift and luminosity of the  $H\beta$  sample as the open symbol (without error bars). The solid line is the measured  $R-L$  relation for  $H\beta$  in the local RM sample (Bentz et al. 2013).

2. Lines with shorter average lags have larger average line widths. The relation between average line width and lag is roughly consistent with the virial relation.

These results are in qualitative agreement with earlier RM studies based on individual objects and at lower redshifts (e.g., Peterson & Wandel 1999; Peterson et al. 2004; Bentz et al. 2010a; Grier et al. 2012b, 2013), and are among the first results on stratified BLR structure at  $z > 0.3$  (e.g., Fine et al. 2013), particularly for He II lags.

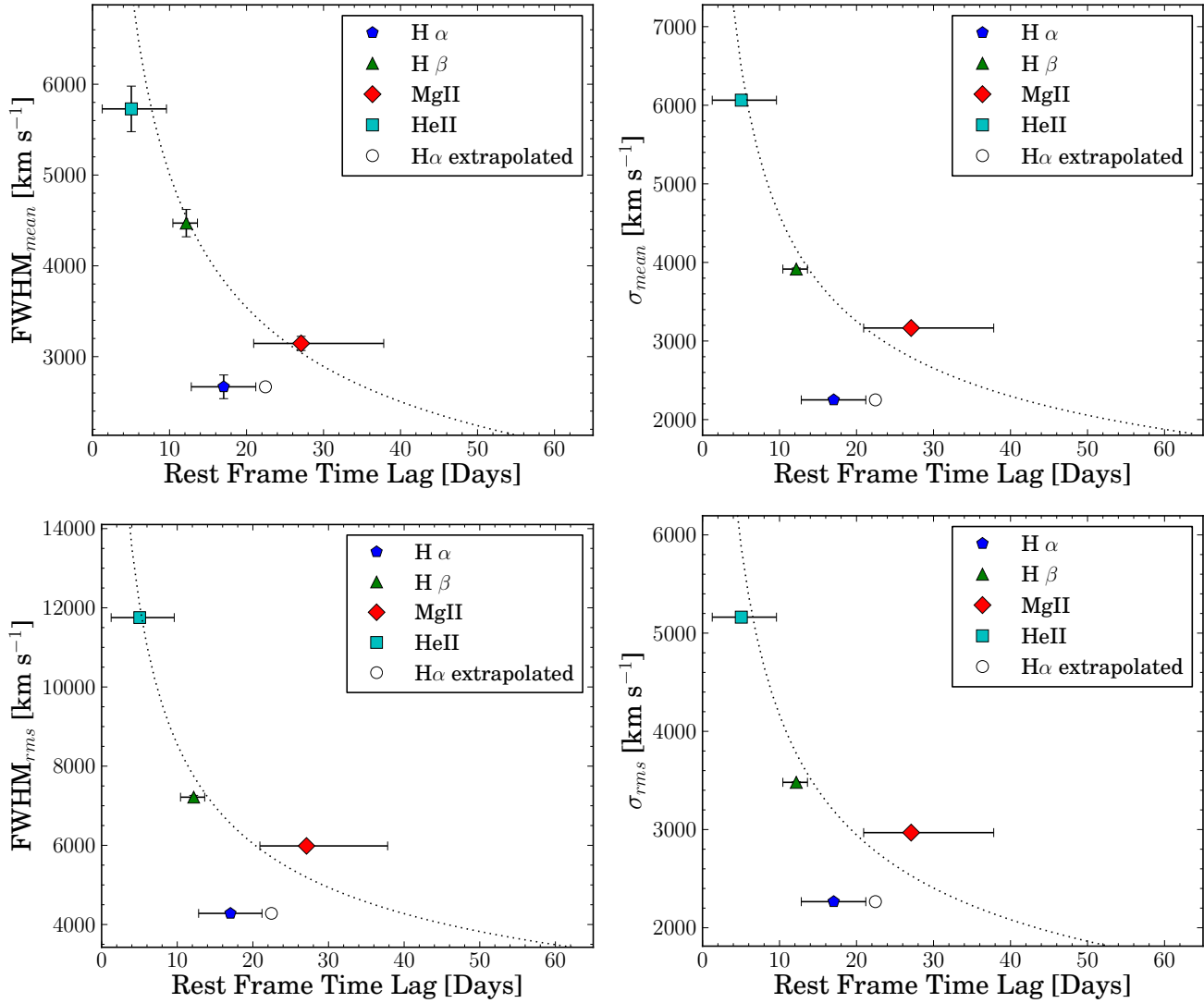
However, as discussed in §4, it is challenging to assign an average redshift and luminosity to the sample, due to the non-trivial contributions from individual objects to the coadded ZDCF. For flux-limited samples that cover broad ranges of redshift and luminosity, the simple median or weighted mean redshift and luminosity may be overestimated. This caveat can be avoided by focusing on samples within a narrow redshift-luminosity range, at the cost of significantly degrading the sample statistics. Finally, as pointed out in earlier work (Fine et al. 2013; Brewer & Elliott 2014), it is somewhat ambiguous to interpret the stacked correlation function for the underlying sample other than the indication of a typical “average” lag from the peak of the stacked correlation. It is possible to deploy more sophisticated statistical inferences to extract information about the intrinsic-lag distribution of the underlying sample from the stacked analysis (e.g., Brewer & Elliott 2014).

SDSS-RM continues to monitor the same quasar sample to extend the time baseline for the detection of long lags at higher redshifts. In addition, we are incorporating the more densely sampled photometric light curves into our time-series analyses. These photometric light curves will improve the spectrophotometry of our spectroscopic epochs with overlapping photometric epochs. With the addition of photometric light curves to enhance the correlation signals in individual objects and the extended time baseline, we plan to expand the redshift coverage to  $z > 1$ , to include additional broad lines (such as  $C\text{III}]$  and  $C\text{IV}$ ), as well as to improve greatly the quality of composite lag measurements in subsamples divided by luminosity and other quasar parameters.

We thank Brad Peterson for useful comments on the draft, and the anonymous referee for suggestions that significantly improved this work. JIL and YS acknowledge support from an Alfred P. Sloan Research Fellowship and NSF grant AST-1715579. WNB, CJG, DPS and JRT acknowledges support from NSF grant AST-1517113. KH acknowledges support from STFC grant ST/M001296/1. LCH was supported by the National Key Program for Science and Technology Research and Development (2016YFA0400702) and the National Science Foundation of China (11303008, 11473002).

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

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 Astrofísica 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



**Figure 10.** Relations between the average lag (from the polynomial fit) and line width for the four lines. The average lags have been shifted to rest frame using the sample-averaged redshifts in Table 1. The adopted line widths are placed at the weighted mean line width for each sample, the uncertainty of which is estimated using propagated individual measurement uncertainties, and is almost negligible. We use four definitions of the line width. In each panel, the dotted line indicates the virial relation  $V \propto \tau^{-0.5}$ , fit to all data points except for that of  $H\alpha$ , for which the sample average luminosity is quite different from those of the other three lines. The open symbol in each panel is the expected lag (without error bars) for  $H\alpha$  at the same average redshift and luminosity of the  $H\beta$  sample.

University.

## REFERENCES

- Ai, Y. L., Yuan, W., Zhou, H., et al. 2013, *AJ*, 145, 90  
 Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, 219, 12  
 Alexander, T. 2013, arXiv:1302.1508  
 Barth, A. J., et al. 2011a, *ApJ*, 743, L4  
 Barth, A. J., et al. 2011b, *ApJ*, 732, 121  
 Barth, A. J., et al. 2013, *ApJ*, 769, 128  
 Barth, A. J., Bennert, V. N., Canalizo, G., et al. 2015, *ApJS*, 217, 26  
 Bentz, M. C., et al. 2009, *ApJ*, 705, 199  
 Bentz, M. C., et al. 2010a, *ApJ*, 716, 993  
 Bentz, M. C., et al. 2013, *ApJ*, 767, 149  
 Blandford, R. D., & McKee, C. F. 1982, *ApJ*, 255, 419  
 Brewer, B. J., & Elliott, T. M. 2014, *MNRAS*, 439, L31  
 Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10  
 Denney, K. D., et al. 2009, *ApJ*, 704, L80  
 Denney, K. D., et al. 2010, *ApJ*, 721, 715  
 Du, P., Hu, C., Lu, K.-X., et al. 2014, *ApJ*, 782, 45  
 Du, P., Hu, C., Lu, K.-X., et al. 2015, *ApJ*, 806, 22  
 Du, P., Lu, K.-X., Zhang, Z.-X., et al. 2016, *ApJ*, in press; arXiv:1604.06218  
 Edelson, R. A., & Krolik, J. H. 1988, *ApJ*, 333, 646  
 Eisenstein, D. J., et al. 2011, *AJ*, 142, 72  
 Fine, S., et al. 2012, *MNRAS*, 427, 2701  
 Fine, S., et al. 2013, *MNRAS*, 434, L16  
 Gaskell, C. M., & Peterson, B. M. 1987, *ApJS*, 65, 1  
 Grier, C. J., et al. 2012a, *ApJ*, 755, 60  
 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. 2012b, *ApJ*, 744, L4  
 Grier, C. J., Trump, J. R., Shen, Y., et al. 2017, *ApJ*, submitted  
 Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, *AJ*, 131, 2332  
 Hu, C., Du, P., Lu, K.-X., et al. 2015, *ApJ*, 804, 138  
 Jiang, L., et al. 2016, *ApJ*, 818, 137  
 Kaspi, S., et al. 2005, *ApJ*, 629, 61  
 Kaspi, S., et al. 2007, *ApJ*, 659, 997  
 Kaspi, S., Smith, P. S., Netzer, H., Maoz, D., Jannuzi, B. T., & Giveon, U. 2000, *ApJ*, 533, 631  
 King, A. L., et al. 2015, *MNRAS*, 453, 1701  
 Peterson, B. M. 1993, *PASP*, 105, 247  
 Peterson, B. M., & Wandel, A. 1999, *ApJ*, 521, L95  
 Peterson, B. M., et al. 2002, *ApJ*, 581, 197  
 Peterson, B. M., et al., 1998a, *ApJ*, 501, 82  
 Peterson, B. M., et al. 2004, *ApJ*, 613, 682  
 Rafter, S. E., Kaspi, S., Behar, E., Kollatschny, W., & Zetzl, M. 2011, *ApJ*, 741, 66

- Rafter, S. E., Kaspi, S., Chelouche, D., Sabach, E., Karl, D., & Behar, E. 2013, *ApJ*, 773, 24
- 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., Horne, K., Grier, C. J., et al. 2016, *ApJ*, 818, 30
- 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
- Sun, M., Trump, J. R., Shen, Y., et al. 2015, *ApJ*, 811, 42
- Trevese, D., Perna, M., Vagnetti, F., Saturni, F. G., & Dadina, M. 2014, *ApJ*, 795, 164