




# Application of a Space-based Optical Interferometer Toward Measuring Cosmological Distances of Quasars

Ying-Ke Huang<sup>1</sup>, Yue-Dong Fang<sup>1</sup>, Kai-Xing Lu<sup>2</sup> , Zhi-Xiang Zhang<sup>3</sup>, Ji-Lin Liu<sup>1</sup>, Sha-sha Li<sup>2</sup>, Bao-Rui Luo<sup>1</sup>, Qin Lin<sup>1</sup>, and Zhuo-Xi Huo<sup>1</sup>

<sup>1</sup> Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology, Beijing 100081, China; [huangyingke@qxslab.cn](mailto:huangyingke@qxslab.cn), [huozhuoxi@qxslab.cn](mailto:huozhuoxi@qxslab.cn)

<sup>2</sup> Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China

<sup>3</sup> Department of Astronomy, Xiamen University, Xiamen, Fujian 361005, China

Received 2021 September 7; revised 2021 December 26; accepted 2021 December 28; published 2022 February 18

## Abstract

Measuring quasar distance through joint analysis of spectroastrometry and reverberation mapping observations is a new method for driving the development of cosmology. In this paper, we carry out detailed simulation and analysis to study the effect of four basic observational parameters (baseline length, exposure time, equivalent diameter and spectral resolution) on the data quality of differential phase curves (DPCs), and furthermore on the accuracy of distance measurement. In our simulation, we adopt an axisymmetrical disk model of a broad line region (BLR) to generate differential phase signals. We find that the differential phases and their Poisson errors could be amplified by extending the baseline, while the influence of optical path difference errors can be reduced during fitting the BLR model. Longer exposure time or larger equivalent diameter helps reduce the absolute Poisson error. Therefore, the relative error of DPCs could be reduced by increasing any of the above three parameters, then the accuracy of distance measurement could be improved. In contrast, the uncertainty of absolute angular distances ( $D_A$ ) could be improved with higher spectral resolution, although the relative error of DPCs would be amplified. We show how the uncertainty of distance measurement varies with the relative error of DPCs. For our specific set of model parameters, without considering more complicated structures and kinematics of BLRs in our simulation, it is found that the relative error of DPCs  $< 20\%$  is a limit for accurate distance measurement. The relative error of DPCs has a lower limit (roughly 5%) and the uncertainty in distance measurement can be better than 2%.

**Key words:** (galaxies:) quasars: emission lines – techniques: interferometric – (cosmology:) distance scale

## 1. Introduction

One of the basic open questions of modern astrophysics is to accurately measure the cosmological distances of extragalactic objects to understand the increasing  $H_0$  tension (Peacock 1999; Freedman & Madore 2010; Weinberg et al. 2013; Freedman 2017; Riess et al. 2019). Active galactic nuclei (AGNs), which are known as the brightest objects in the universe, have been utilized to measure cosmological distances since they were discovered (Sandage 1965; Hoyle 1966; Longair & Scheuer 1967; Baldwin 1977; Elvis & Karovska 2002; Quercellini et al. 2009; Wang et al. 2013; Marziani & Sulentic 2014). Nevertheless, due to lack of understanding of AGN physics, many attempts have been made but none of them were proven to be bias free. Recently, a joint analysis (Wang et al. 2020) of spectroastrometry (SA: Petrov 1989; Bailey 1998; Gravity Collaboration et al. 2017) and reverberation mapping (RM: Blandford & McKee 1982; Peterson 1993; Kaspi et al. 2000; Bentz et al. 2013; Du et al. 2018) observations provided a direct method to measure absolute angular distances ( $D_A$ ) of AGNs. This method does not need calibration using cosmic ladders so it has the potential to

provide absolute distance measurements for quasars in the high- $z$  universe.

The high-energy output of an AGN is argued to be the result of accreting matter by a supermassive black hole at the center of a galaxy (Lynden-Bell & Rees 1971). A hallmark of AGN spectra is the existence of broad emission lines (Osterbrock & Mathews 1986; Osterbrock 1989; Ho 2008). These broad emission lines are thought to originate from photoionization of the broad line region (BLR) clouds by high-energy continuum photons from the central accretion disk. The revolution of these clouds is dominated by Kepler motion around the central black hole due to gravity, which has been confirmed by multiple campaigns of several AGNs (Peterson et al. 2004). The emission lines are broadened up to several thousand  $\text{km s}^{-1}$  (Antonucci 1993; Urry & Padovani 1995). However, it is impractical to spatially resolve the BLR by direct photometry, because the angular size of the BLR is smaller than the existing spatial resolution limit ( $< 0.1$  mas; Blandford & McKee 1982; Elvis & Karovska 2002).

In the past few decades, a promising approach to directly constrain BLR geometry is the widely used RM technique. The

underlying principle of the RM technique is photoionization theory (Wandel et al. 1999; Kaspi et al. 2000; Bentz et al. 2006). By analyzing the variation of continuum and emission lines after measuring the time delay between the variable continuum and the responsive broad emission line, it is possible to constrain the geometry and kinematics of the BLR and further estimate the characteristic size of the BLR as well as the black hole mass. In 2018, Gravity Collaboration et al. (2018) proposed another method. By measuring the photocenters of emission lines in different wavelength channels, the SA method (Beckers 1982; Petrov 1989; Petrov et al. 1992; Gnerucci et al. 2010) proved to have great advantage in providing information on the spatial structure at scale much smaller than the spatial resolution of an interferometer. An intuitive picture of how the SA method can spatially resolve the bulk motion of the BLR clouds is illustrated as follows: there are two clouds located at a distance significantly smaller than the spatial resolution, so they are seen as a non-resolved source with a global angular size, but one cloud is moving toward the observers and the other moving away from observers. As the line-of-sight (LOS) velocities of the two clouds are different, the photons from the same emission line but from different clouds are shifted to different wavelengths. By measuring the interferometric phases (providing position information on-sky) over wavelengths, we can obtain the photocenter position of clouds with different velocities and further separate the motion of BLR clouds. The SA method demonstrates the capability of overcoming the spatial resolution limit, thus being able to constrain the angular scale, geometric structure and dynamics of the BLR.

In fact, RM observations provide the linear size of BLR and are more sensitive to the direction along the LOS. SA measures the angular size of BLR and is more sensitive to a direction perpendicular to the LOS (Wang et al. 2020). Relying on the geometric and dynamical BLR model, we can extract information from RM and SA data to measure the cosmological distance. However, up to now, RM campaigns measure the region of optical emission line (mainly using  $H\beta$ ) while the SA campaigns measure the near-infrared (NIR) emission lines (using  $Pa\alpha$  in Gravity Collaboration et al. 2018 and Wang et al. 2020). Since different emission lines come from different areas of the BLR (Clavel et al. 1991; Dietrich & Kollatschny 1995; Kollatschny 2003), if RM and SA rely on the same emission line, the systematic error (mentioned in Wang et al. 2020) of distance measurement in joint analysis could be significantly reduced. It is necessary to carry out SA observations in optical bands or RM observations in infrared bands. Considering the extinction and jitter of the Earth's atmosphere, the interference process in space shows advantages in long exposures. A space-based optical interferometer will be the future direction.

As a necessary step, we aim at studying the uncertainty of distance measurement caused by basic observational parameters of a space-based optical interferometer in this paper. The structure of this paper is as follows. In Section 2, we

describe the geometric and dynamical model of the BLR adopted. In Section 3 we simulate the expected spectro-astrometric signals and describe the analysis method in Section 4. Results are provided in Section 5, and we show our summary and discussions in the last section.

## 2. Parameterized BLR Model

In the optical band, an AGN spectrum consists of continuum emission and line emission. The continuum emission is considered to originate chiefly from the accretion disk. For nearby galaxies, continuum emission from the host galaxy could also not been ignored. For simplicity, we assume the region that produced continuum emission has an axisymmetrical structure and the photocenter of continuum emission is located at the black hole. The broad emission lines are considered be generated from photoionization of BLR clouds by the ultraviolet (UV) continuum photons from the accretion disk. The emission line used in our simulation is  $H\beta$ , which is a strong optical emission line and widely used for RM observation to estimate the linear size of the BLR. The continuum underneath emission lines is often assumed to have a linear form. This hypothesis is widely utilized in RM to calculate the integral flux of  $H\beta$ , where a continuous spectrum is required to be subtracted off (Du et al. 2014; Huang et al. 2019). In order to estimate the photocenter position of a broad  $H\beta$  emission line and simulate the SA observations, we rely on a simple BLR model that is characterized by a flat disk with circular Kepler rotation.

In recent years, thanks to many studies combining RM observation and the parameterized BLR model, we have gained a better understanding of BLR physics (Li et al. 2013, 2018; Pancoast et al. 2014a, 2014b; Grier et al. 2017; Williams et al. 2018). In this section we give an overview of the parameterized model we used for BLR. It consists of many isolated BLR clouds, which are modeled as a large number of non-interacting point particles. These particles reprocess the continuum photons originating from the accretion disk into the emission line photons instantaneously. The wavelength of the emission line photons is determined by the particles' velocity and the time lag for the response is determined by their position. For simplicity, the accretion disk is regarded as a point-like geometry so that the UV ionizing continuum is isotropic and the flux of continuum falls off with the square of the distance.

### 2.1. Geometry

In our simulation, the radial distribution of the point particles is described as a shifted  $\Gamma$ -distribution. The distance  $r$  of a point particle from the black hole is given by

$$r = \Gamma_0(1 - F)\beta^2 R_{\text{BLR}} + R_s + F R_{\text{BLR}}. \quad (1)$$

Here  $\Gamma_0 = p(x|1/\beta^2, 1)$  is drawn randomly from a Gamma distribution

$$p(x|\alpha, \theta) = \frac{x^{\alpha-1} \exp(-x/\theta)}{\theta^\alpha \Gamma(\alpha)}, \quad (2)$$

and the point particle is then shifted radially by a Schwarzschild radius  $R_s = 2GM_{\text{BH}}/c^2$ , plus a minimum radius  $R_{\text{min}}$  of BLR. Here  $R_{\text{BLR}}$  is the mean radius and  $F = R_{\text{min}}/R_{\text{BLR}}$  is a ratio of the minimum radius to the mean radius. Within  $R_{\text{min}}$ , the line-emitting particles are not allowed to exist.  $\beta$  is the shape parameter of the radial distribution: small values of  $\beta$  mean narrow normal distributions while large values mean exponential distributions. We then define a half-opening angle  $\theta_o$  for the overall geometry and the spatial distribution of point particles. Here  $\theta_o = 0^\circ$  defines a thin disk (ring) while  $\theta_o = 90^\circ$  represents a spherical distribution. An observer views the BLR from an inclination angle  $\theta_i$ , where  $\theta_i = 0^\circ$  corresponds to a face-on orientation.

For each point particle, the emission is weighted by a parameter  $\omega(\phi)$ ,

$$\omega(\phi) = \frac{1}{2} + k \cos(\phi), \quad (3)$$

where  $k$  is a free parameter between  $-0.5$  and  $0.5$ . When  $k=0.5$ , the inner side of BLR is contributing more line emission. Here  $\phi$  is the angle between the observer's LOS to the black hole and the point particle's line to the black hole. The particles could be clustered and  $\theta$  is the angle of a point particle from the disk

$$\theta = \arccos(\cos \theta_o + (1 - \cos \theta_o)U^\gamma), \quad (4)$$

where  $U$  is a random number drawn uniformly between 0 and 1, and  $\gamma$  is a free parameter drawn uniformly between 1 and 5. When  $\gamma=1$ , the point particles are clustered toward the disk.

We also adopt a usual assumption that BLR clouds have the same size and no shadowing among them (Li et al. 2013). We use the traditional linear response of the emission lines to the continuum. For a BLR that has a given geometry and cloud distribution, the emission-line flux  $f_i$  at time  $t$  is estimated by summing over the emissions from all the clouds

$$f_i = A \sum_i \omega_i \frac{f_c}{r_i^2}. \quad (5)$$

Here  $\omega_i$  is the weight of the  $i$ th cloud calculated by Equation (3),  $A$  is the response coefficient and  $f_c/r_i^2$  describes the ionizing flux received by the  $i$ th cloud at time  $t$ .

## 2.2. Dynamics

We assume a flattened disk-like BLR with Keplerian rotation. Due to the gravity of the central black hole, the point particles rotate in a circular orbit. The Keplerian velocities  $V_K$  of the point particles are drawn from a distribution that depends

**Table 1**  
Parameters Used in the BLR Model

Parameter	Meaning	Value	Prior Range
$F$	fractional inner radius of BLR	0.49	[0, 1]
$\beta$	radial distribution of BLR clouds	1.09	[0, 4]
$\theta_o(^{\circ})$	half opening angle of the BLR	39.96	[0, 90]
$\theta_i(^{\circ})$	inclination angle of the BLR	8.41	[0, 90]
PA(^{\circ})	position angle	210.99	[0, 520]
$R_{\text{BLR}}(\text{lt-day})$	averaged linear sizes	184.17	[1, 10 <sup>3</sup> ]
$M_{\text{BH}}(10^8 M_{\text{Sun}})$	supermassive black hole mass	5.78	[10 <sup>-2</sup> , 10]
$D_A(\text{Mpc})$	absolute angular distance	551.50	[10, 10 <sup>4</sup> ]

**Note.** The parameters were used in the BLR model and the values obtained from Wang et al. (2020).

upon the black hole mass  $M_{\text{BH}}$  and the distance  $r$ , which could be estimated as

$$V_K = \left( \frac{GM_{\text{BH}}}{r} \right)^{1/2}. \quad (6)$$

For a thick Keplerian disk with a half opening angle  $\theta_o$ , there will be an angle  $\theta$  between the orbital plane and the equatorial plane. Thus, the LOS velocities  $v_{\text{line}}$  are related to the Keplerian velocities, half-opening angle and the inclination angle. For cloud particles with an emitted wavelength  $\lambda_{\text{emit}}$ , considering the relativistic effect, Doppler broadening and gravitational redshift which will affect the emission line profile, the observed wavelength  $\lambda_{\text{obs}}$  can be written as

$$\lambda_{\text{obs}} = \left( 1 + \frac{v_{\text{line}}}{c} \right) \left( 1 - \frac{V_K^2}{c^2} \right)^{-1/2} \left( 1 - \frac{R_S^2}{r} \right)^{-1/2} \lambda_{\text{emit}}. \quad (7)$$

In order to explore the influence of basic observational parameters on the uncertainty of distance measurement, we adopt the same set of BLR model parameters as shown in Wang et al. (2020), which has been summarized in Table 1.

## 3. The Spectroastrometric Signal of the BLR

The SA method (see Bailey 1998; Rakshit et al. 2015) shows advantages in providing information on the spatial structure of the object, especially when the global angular size  $\Lambda$  of a non-resolved source is smaller than the interferometer limit  $\lambda/B$ , here  $\lambda$  is the observed wavelength and  $B$  is the length of the interferometer baseline. By measuring the photocenters of different wavelength channels, this method has been applied to observe AGNs (Petrov et al. 2001; Marconi et al. 2003; Gnerucci et al. 2010, 2011; Rakshit et al. 2015) in NIR and successfully constrained the size of the BLR for 3C 273 (Gravity Collaboration et al. 2018; Wang et al. 2020). In this section we give a description of how we simulate the SA signal.

For a given object, the observed interferometric phase  $\Phi$  at wavelength  $\lambda$  can be written as

$$\Phi(\lambda) = \Phi_*(\lambda) + \Phi_t(\lambda), \quad (8)$$

where  $\Phi_*(\lambda)$  is the phase contributed from the real astrophysical signal and  $\Phi_t(\lambda)$  is the phase contributed from the atmosphere, telescopes, interferometric delay lines and the various instrumental parts down to the detector (see Vannier et al. 2006). Petrov (1989) shows that, for an interferometer with a baseline  $\mathbf{B}$ , the interferometric phase for the object is

$$\Phi_*(\lambda) = -2\pi \frac{\mathbf{B}}{\lambda} \cdot \boldsymbol{\epsilon}(\lambda). \quad (9)$$

Here  $\boldsymbol{\epsilon}(\lambda)$  is the photocenter of the source at wavelength  $\lambda$  and could be written as

$$\boldsymbol{\epsilon}(\lambda) = \frac{\int \int \mathbf{r} O(\mathbf{r}, \lambda) d^2 \mathbf{r}}{\int \int O(\mathbf{r}, \lambda) d^2 \mathbf{r}}. \quad (10)$$

$O(\mathbf{r}, \lambda)$  is the brightness distribution of the source. If defining the fraction of the flux of the emission line to total as  $f_l(\lambda) = F_l(\lambda)/F_{\text{tot}}(\lambda)$ , where  $F_l(\lambda)$  is the flux of the emission line at wavelength  $\lambda$  and  $F_{\text{tot}}(\lambda)$  is total flux at wavelength  $\lambda$ , we have  $\boldsymbol{\epsilon}(\lambda) = f_l(\lambda) \boldsymbol{\epsilon}_l(\lambda)$ ,  $\mathbf{r}$  is the vector to the central black hole and

$$\boldsymbol{\epsilon}_l(\lambda) = \frac{\int \int \mathbf{r} O_l d^2 \mathbf{r}}{\int \int O_l d^2 \mathbf{r}}. \quad (11)$$

We use the bold letters to signify the vectors  $\mathbf{B}$ ,  $\boldsymbol{\epsilon}(\lambda)$  and  $\mathbf{r}$ .

For space-based interferometers, the phase contributed from the atmosphere could be ignored but the phase contributed from the instrument could not. In principle, if we add a spatial filter before the beam recombination, the optical effects left after this spatial filtering only include differences in intensity between beams and an optical path difference (OPD). The previous one can be calibrated by photometric channels, so only the phase shift caused by the OPD is taken into consideration in our simulation. The OPD has two origins: one originates exclusively from piston and chromatic dispersion along the path before the filtering, while another originates from wave front corrugations induced by imperfect adaptive optics (Tubbs 2005; Vannier et al. 2006). For a given spectral channel, the phase shift caused by OPD is

$$\Phi_{\text{OPD}}(\lambda) = (2\pi/\lambda) \text{OPD}. \quad (12)$$

Hence, the differential phase between a channel  $\lambda_i$  and the reference channel  $\lambda_r$  will be

$$\begin{aligned} \Delta\Phi(\lambda_i, \lambda_r) &= [\Phi_*(\lambda_i) - \Phi_*(\lambda_r)] \\ &+ [\Phi_{\text{OPD}}(\lambda_i) - \Phi_{\text{OPD}}(\lambda_r)]. \end{aligned} \quad (13)$$

By inserting Equations (9), (10) and (12) into Equation (13), we have

$$\begin{aligned} \Delta\Phi(\lambda_i, \lambda_r) &= -2\pi \mathbf{B} \cdot \left[ f_l(\lambda_i) \frac{\boldsymbol{\epsilon}_i(\lambda_i)}{\lambda_i} \right. \\ &\quad \left. - f_l(\lambda_r) \frac{\boldsymbol{\epsilon}_r(\lambda_r)}{\lambda_r} \right] + 2\pi \cdot \text{OPD} \cdot \left( \frac{1}{\lambda_i} - \frac{1}{\lambda_r} \right). \end{aligned} \quad (14)$$

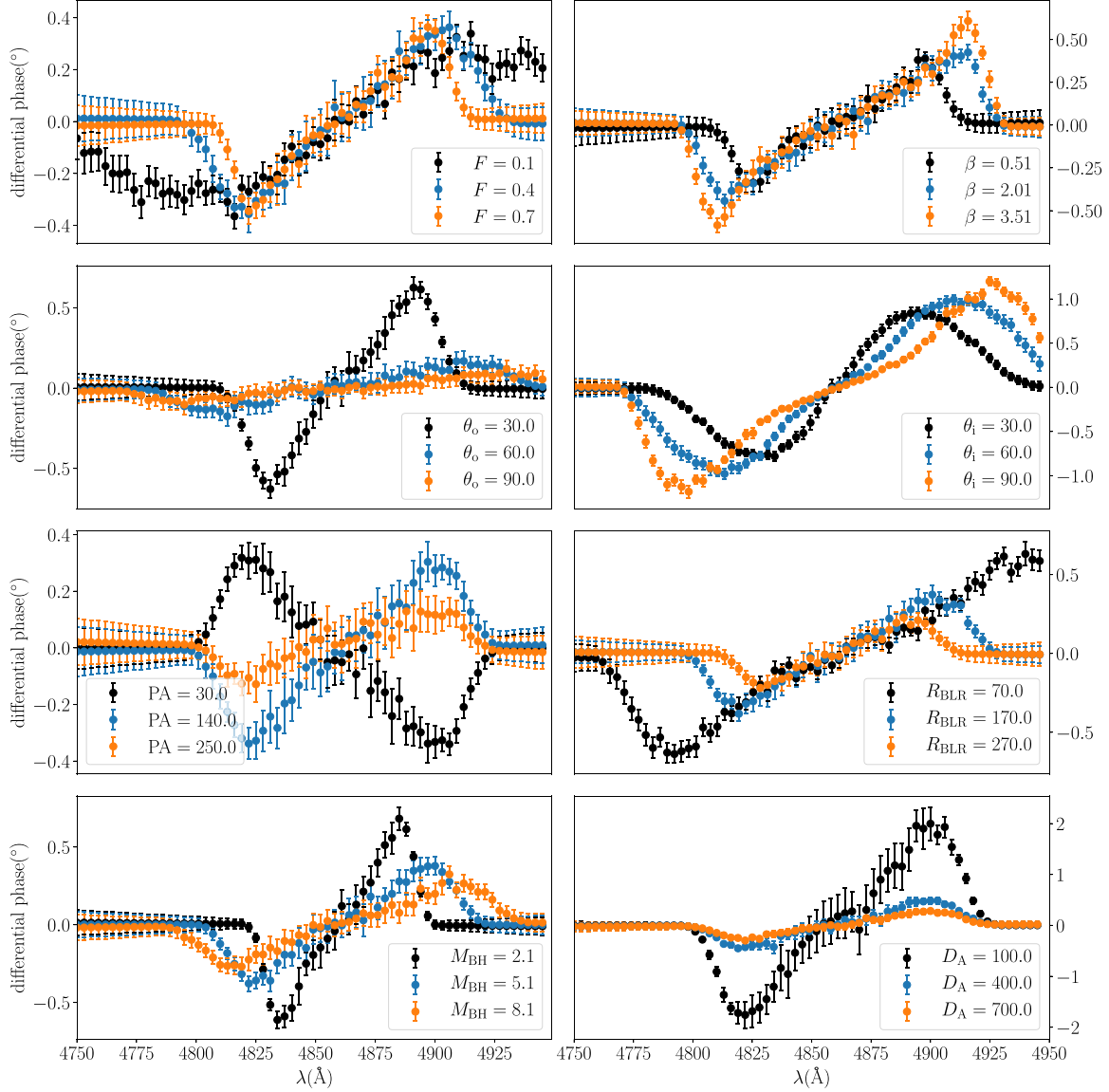
If the BLR model is specified, the surface brightness distribution of the regions is mainly related to the total number of collected photons, which is linearly dependent on the number of telescopes ( $n_T$ ), the overall quantum efficiency ( $\eta$ ), the collecting area of telescopes  $S$  and the exposure time  $t$ . For each wavelength channel, the observed photocenter is further determined by the choice of spectral resolution  $R$ . If we also know the configuration of the projected baseline  $\mathbf{B}$ , we could predict the actual phase signal by Equation (9). By giving the limit OPD (rms) control errors during an observation, using Equation (14), we could obtain wavelength-dependent differential phase after DPCs. For simplicity, the equivalent diameter ( $D$ ) is used to estimate the number of photons per unit time in our simulation,  $S = \pi(D/2)^2$ .

In our simulation, uncertainty of the signal consists of two parts, one part comes from Poisson noise and the other is caused by the uncertainty in OPD control. Each set of scientific exposures is assumed to be the average over  $n_{\text{obs}}$  (the number of observations) frames of  $t$ -s integration (NDIT =  $n_{\text{obs}}$  and DIT =  $t$  s). We use the average value of  $n_{\text{obs}}$  observations at wavelength  $\lambda$  as the differential phase  $\Delta\Phi(\lambda)$  and the dispersion of  $n_{\text{obs}}$  observations as the error. We show the dependence of mock DPCs on eight parameters used in the BLR model in Figure 1.

#### 4. Analysis

In this section, we describe the analytical method of obtaining the BLR model parameters from mock data. The joint analysis proposed by Wang et al. (2020) uses data from three parts: light curves of optical continuum and emission lines, DPCs and the profiles of the emission lines. The light curves and emission profiles depend on optical spectral observation, which has been a very mature technology, so we assume that the quality of the light curves and the profiles of the emission lines are good enough. Thus, the uncertainty in the joint analysis depends mainly on the data quality of DPCs.

As mentioned above, we generate the mock DPCs from Equation (14). We use the DPCs to obtain the posterior distribution of the BLR model parameters. Suppose that probability distributions for the measurement errors of the DPCs are Gaussian and uncorrelated, then the likelihood



**Figure 1.** Dependence of DPCs on eight parameters. The parameters for the interferometer are fixed at the values listed in Table 2. When varying one parameter, other parameters are fixed at the values listed in Table 1. The parameters and their values are shown in the legend. The units of  $\theta_o$ ,  $\theta_i$  and PA are degrees. The units for  $R_{\text{BLR}}$ ,  $M_{\text{BH}}$  and  $D_A$  are lt-day,  $10^8 M_{\text{Sun}}$  and Mpc, respectively.

function can be written as

$$P(\mathcal{D}|\Theta) = \prod_{i=1}^{N_{\text{obs}}} \prod_{j=1}^{N_{\lambda}} \frac{1}{\sqrt{2\pi\sigma_{\phi_{ij}}^2}} \exp \left\{ -\frac{[\Phi_{\text{obs}} - \Phi_{\text{mod}}(\Theta)]^2}{2\sigma_{\phi_{ij}}^2} \right\}. \quad (15)$$

Here  $\mathcal{D}$  represents the measured data,  $\Theta$  signifies the BLR model parameters,  $\Phi_{\text{obs}}$  is the interferometric phase of the emission line with the uncertainties  $\sigma_{\phi_{ij}}$ ,  $\Phi_{\text{mod}}(\Theta)$  is the corresponding predicted values from the BLR model,  $N_{\text{obs}}$  is the number of observations and  $N_{\lambda}$  is the corresponding

number of wavelength bins. According to Bayes' theorem, if we know the prior distribution of the model parameter, the posterior probability distribution for  $\Theta$  should be given by

$$P(\Theta|\mathcal{D}) = \frac{P(\Theta)P(\mathcal{D}|\Theta)}{P(\mathcal{D})}, \quad (16)$$

where  $P(\mathcal{D})$  is a normalization factor. We list the parameters of the BLR model and the prior ranges in Table 1.

The EMCEE package is used to sample the parameters efficiently (Foreman-Mackey et al. 2013), which is an MIT licensed pure-Python implementation of Goodman & Weare's



Affine Invariant Markov chain Monte Carlo (MCMC) Ensemble sampler. The sampling is run with 1000 “walkers” independently and converges within about 2000 trials.

## 5. Results

Utilizing the BLR model described in Section 2, we generate mock observations according to a set of varying basic observational parameters ( $B$ ,  $D$ ,  $t$ ,  $R$ ) described in Section 3. Combined with the method mentioned in Section 4, we first study the effect of inclination angle and the projected angle on the photocenter shift, then we study the influence of the basic observational parameters on the data quality of DPCs and the accuracy of distance measurement.

### 5.1. The Effect of Inclination Angle on The Photocenter Shift

The SA method can spatially resolve the velocity gradient perpendicular to the LOS direction, the size of which varies with the inclination angle. Therefore, we first study how inclination angle affects the distribution of photocenters. We used the hypothesized BLR model mentioned above. The parameters of the model are fixed and listed in Table 1, except for the inclination angle. We vary the inclination angle ( $\theta_i$ ) from  $0^\circ$  to  $180^\circ$ . Then we estimate the corresponding distribution of photocenters which are shown in the left panel of Figure 2. By changing the inclination angle, we find that for a BLR with Keplerian motion, the more “edge-on” the LOS is, the larger the spatial distribution size of wavelength-dependent photocenters. When viewed in a direction perpendicular to the disk plane (face-on:  $\theta_i = 0^\circ$ ), the photocenters of different wavelengths could not be distinguished (as shown in the lower right panel of Figure 2). This conclusion can be understood as when viewed from angle  $\theta_i$ , the velocity components along the LOS direction are proportional to  $\sin \theta_i$ . Therefore, when viewed from the edge-on direction ( $\theta_i = 90^\circ$ ), the velocity along the LOS direction achieved a maximum such that the observed wavelength gradient is the largest (depicted in the upper right panel and lower right panel of Figure 2). We must emphasize that shielding of the dust torus is not taken into account in our simulation.

### 5.2. The Influence of Projected Angle on the Photocenter Shift

In fact, the angle between the baseline and the plane of BLR will also affect the observed distribution of photocenters. This angle is defined as the projected angle (displayed in Figure 3(b) as  $\theta_B$ ) in our simulation. Colored points represent the photocenters of different wavelengths, and the black solid line signifies the baseline direction. Each baseline can only measure the distribution of photocenters projected to the baseline direction, which could be indicated from Equation (9). To study the impact of the angle  $\theta_B$ , we use the same BLR model as mentioned above and

fix  $\theta_i = 90^\circ$ , then vary the angle  $\theta_B$  to measure the displacement of photocenters at different wavelengths. Results are displayed in Figure 3(a) and the relation between the maximum offset of photocenters and  $\theta_B$  is plotted in Figure 3(c). We see that the maximum offset of photocenters is proportional to the absolute value of  $\cos \theta_B$ . When the baseline is perpendicular to the plane, the photocenters of different wavelengths along the baseline direction are indistinguishable.

In actual observations, the direction of the BLR plane (so the projected angle  $\theta_B$ ) is unknown to the observer. To ensure that the mock data contain information about spatial position, we adopted an observation strategy of rotating the baseline. In our simulation, we rotate the baseline regularly with a cadence of  $10^\circ$ . Thus we get 19 mock data for each set of observational parameters.

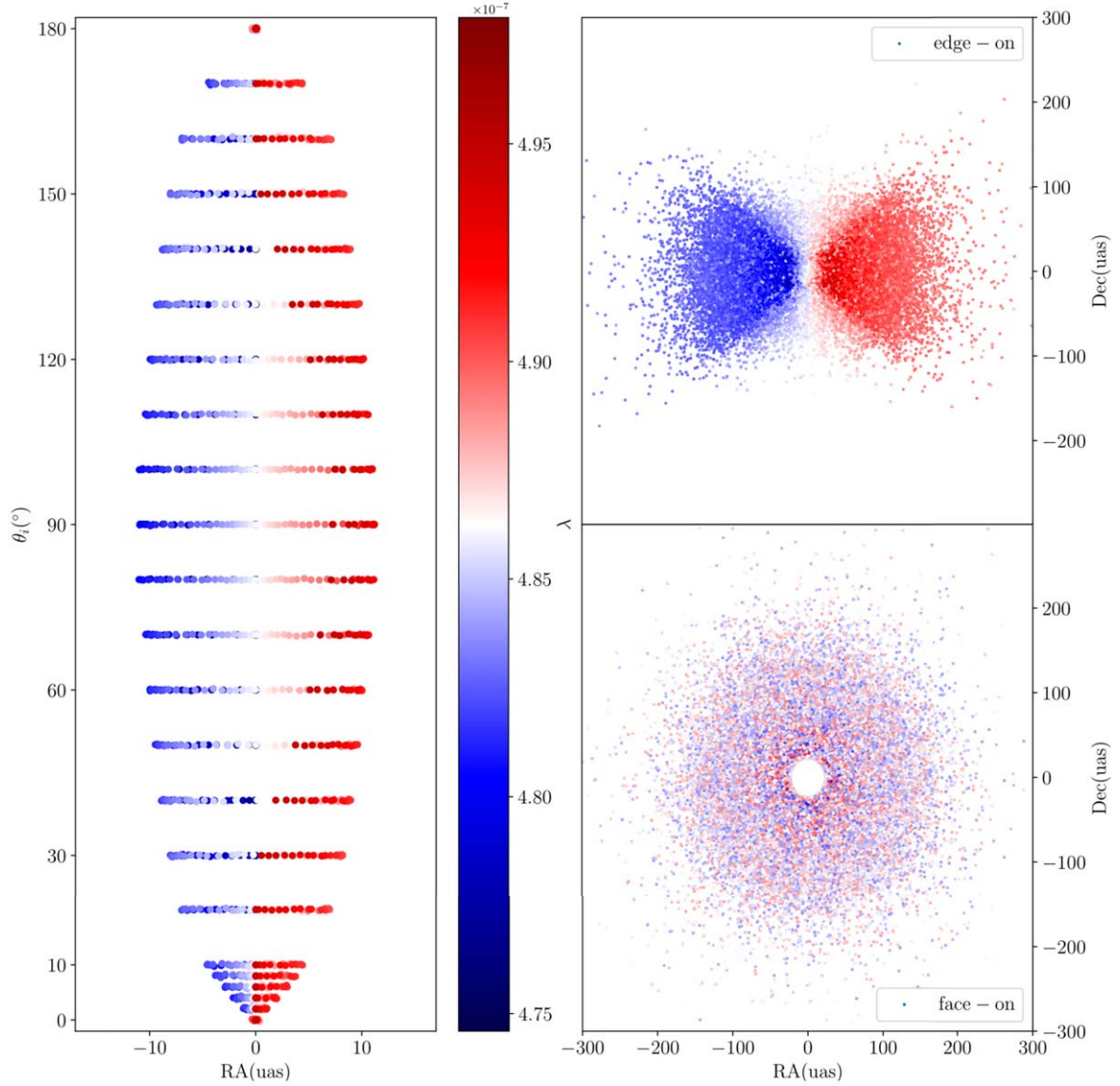
### 5.3. The Influence of Basic Observational Parameters on Data Quality

Furthermore, we study the effects of the basic observational parameters by varying one parameter at a time (the ranges of variations are summarized in Table 2). In addition, the projected angle  $\theta_B$  was fixed at  $\theta_B = 0^\circ$ .

In order to quantify the impact of basic observational parameters mentioned above, we define two metrics to describe the data quality: the largest amplitude and the relative error of DPCs. The largest amplitude of DPCs is defined as the maximum absolute amplitude of a set of DPCs. The relative error of DPCs is quantified as the ratio of phase error to phase value where phase is the largest amplitude of DPCs. We use the relative Poisson error of DPCs, which is defined as the ratio of the Poisson error to the largest amplitude of DPCs, to quantify the impact of Poisson noise. The results are depicted in Figure 4.

The impact of baseline length on data quality is shown in Figure 4(a). We feature three DPCs generated under three different baseline lengths in the left panel. These DPCs display an obvious S-shape. The value of the differential phase increases as the baseline gets longer, as does the absolute error. We find that the largest amplitude of DPCs increases linearly with the baseline length (shown in the upper right panel), which has been implied from Equation (9). As the baseline gets longer, the relative error tends to be smaller (shown in the bottom right panel). We also find that the Poisson error becomes dominating while the baseline lengthens.

The impact of the exposure time is shown in Figure 4(b). We also display three DPCs generated under three different exposure times in the left panel. It is easy to see that as the exposure time increases, the absolute error of DPCs decreases significantly, but the value of the differential phase remains essentially the same. As featured in the upper right panel, the largest amplitude of DPCs does not change with exposure time. However, relative error and relative Poisson error of the DPCs decrease as the exposure time increases.



**Figure 2.** The effect of inclination angle on the photocenter shift. Left: The calculated distribution of photocenters that varies with the inclination angle. Upper right: The position distribution of photocenters at different wavelengths as seen by an observer from the edge-on direction ( $\theta_i = 90^\circ$ ). Lower right: The position distribution of photocenters at different wavelengths as seen by an observer from the face-on direction ( $\theta_i = 0^\circ$ ).

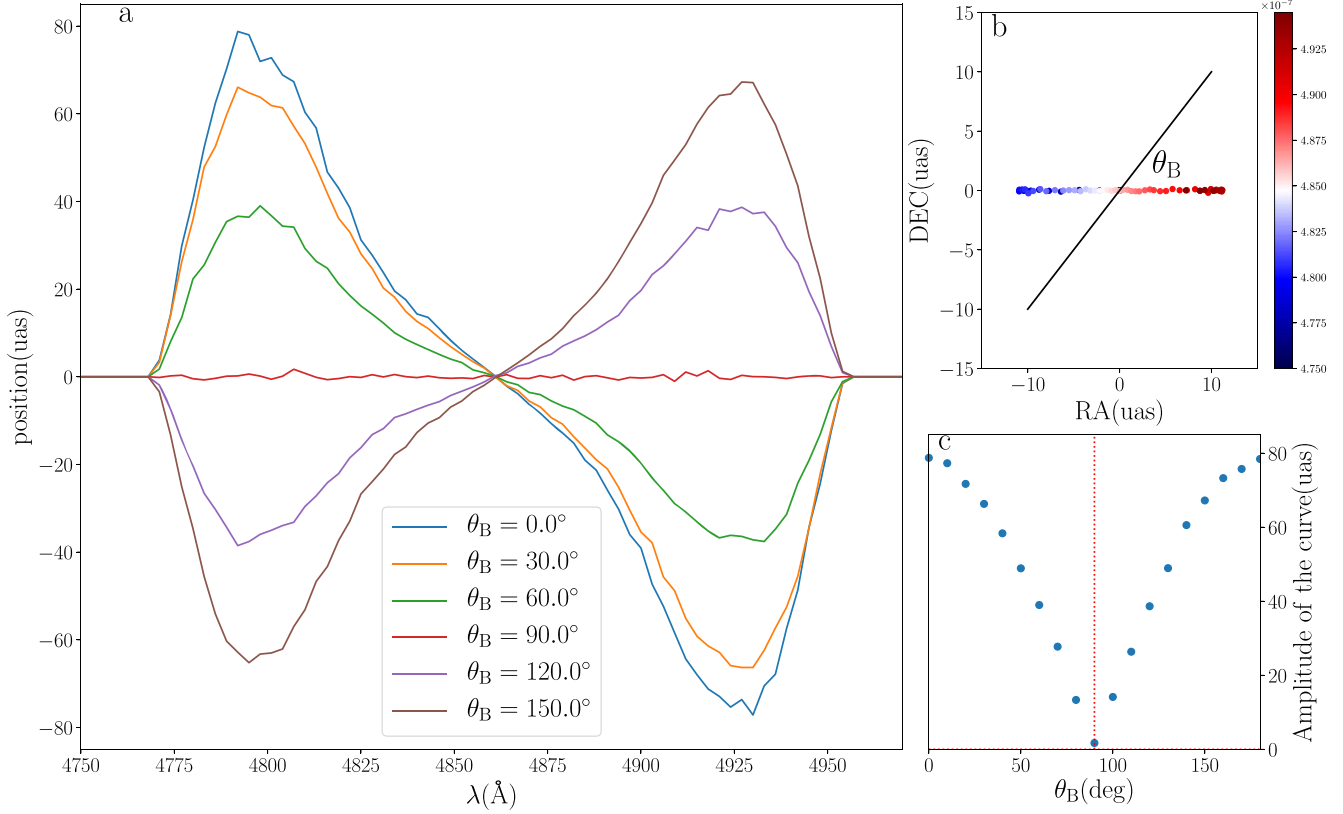
The influence of the equivalent diameter is basically the same as that of exposure time, and the results are depicted in Figure 4(c). As the equivalent diameter becomes larger, absolute error, relative error and relative Poisson error of the DPCs are significantly reduced, but the largest amplitude of DPCs remains the same. For our simulation, we find that by changing the equivalent diameter, relative error of the DPCs can be reduced to less than 5%, which is more effective than increasing the exposure time to roughly 10%.

The results of how the spectral resolution impacts the DPCs are shown in Figure 4(d). As the resolution increases, the number of data points on the DPC increases significantly (shown in the left

panel), while the largest amplitude of DPCs increases slightly (as shown in the upper right panel) and the relative error of DPCs becomes larger (as displayed in the bottom right panel). This conclusion can be understood as the larger the spectral resolution, the more wavelength channels there are, but the less number of photons in each wavelength channel.

#### 5.4. The Effect of Basic Observational Parameters on the Accuracy of Distance Measurement

In order to quantify the accuracy of distance measurement, we define the median value of the posterior distribution as the best value for distance inferred from MCMC, with the 16% and



**Figure 3.** The influence of projected angle on the photocenter shift. Left: Simulated position of photocenters at different wavelengths under different projected angle  $\theta_B$ . Upper right: The black solid line shows the baseline direction, colored points represent the photocenters of different wavelengths and  $\theta_B$  is the projected angle. Lower right: The maximum offset of photocenters varies with  $\theta_B$ .

**Table 2**  
Parameters Used in the Interferometry Model

Parameter	Meaning	Value	Range
$B(m)$	the interferometer baseline	100	[20, 1000]
$D(cm)$	the diameter of the telescope	100	[20, 1000]
$t(s)$	the time of exposure	2000	[200, 20000]
$R$	the resolution of the spectrum	1000	[250, 10000]
$n_{obs}$	the number of observations	30	...
OPD(nm)	the accuracy of OPD control	5	...

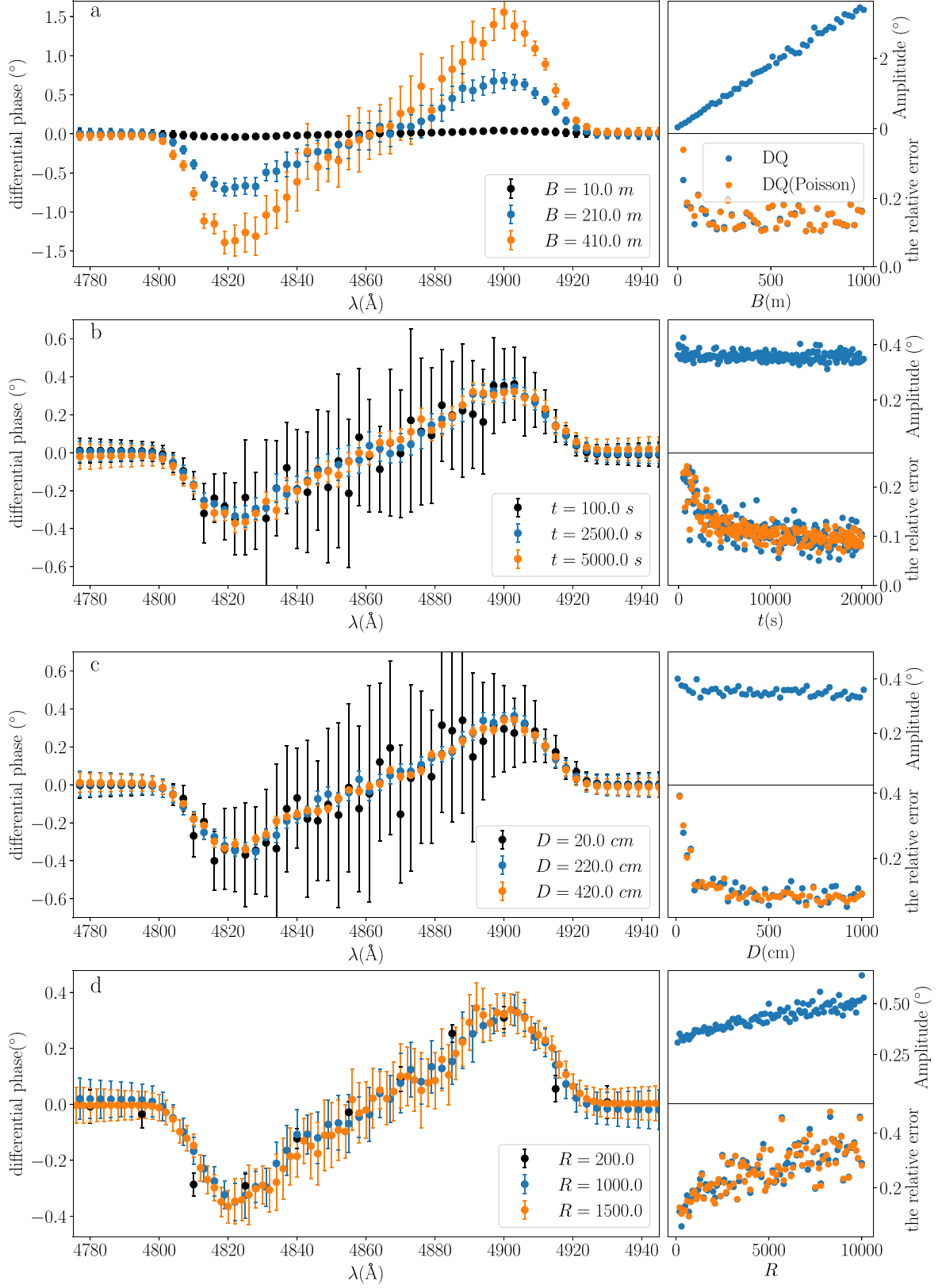
**Note.** The parameters were used in the interferometer model and the range of the change. In our simulation, the OPD is a random value chosen from a Gaussian distribution with dispersion 5 nm (rms), which is equal to an absolute phase shift of  $\sim 0.06$  rad ( $3.7^\circ$ ) at 4861 Å.

84% quantiles as the lower and upper bounds respectively. The relative uncertainty of the distance measurement is defined as the ratio of the error (mean of the lower and upper bounds) to the best value mentioned above. The relative bias is defined as the fraction of the bias (difference between the best value and the input value) to the input value. We simulate the actual DPCs using the BLR model mentioned in Section 2 and the

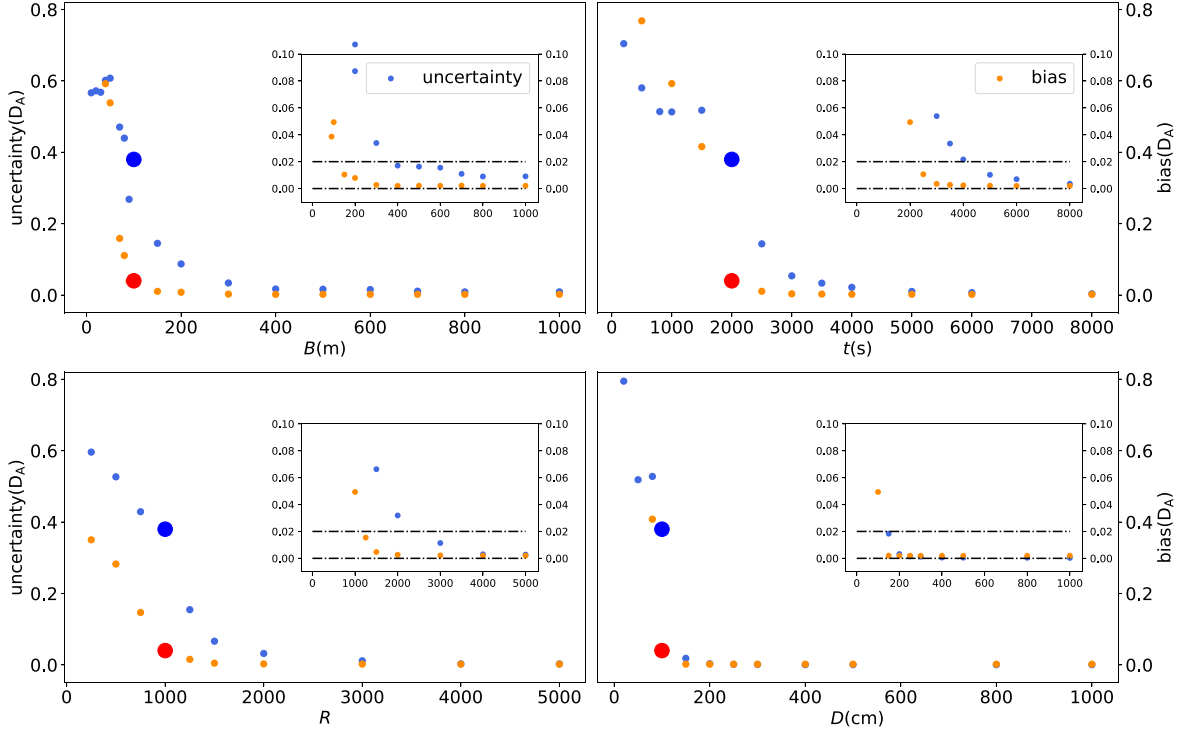
basic observational parameters described in Section 3. Then we apply the MCMC method described in Section 4 to obtain probability density distributions of BLR model parameters. In the Appendix, we list the simulated DPCs (Figure A1) and the mock H $\beta$  line profile (Figure A2) generated by using the instrument parameters listed in Table 2, as well as the probability density distributions (Figure A3) of model parameters. We obtain the relative uncertainty (blue points) and bias (orange points) of distance measurements ( $D_A$ ) at different baseline lengths, exposure times, equivalent diameters and spectral resolutions. The results are depicted in Figure 5.

We see that as the parameter values increase, the relative uncertainty of distance measurement ( $D_A$ ) can be reduced to 2%. There is an optimum value for each parameter. Above this optimum value, the accuracy of  $D_A$  will improve slowly with the parameter value. Extending the baseline can proportionally amplify the difference phases and their Poisson errors, which means that the influence of OPD errors can be reduced during the fitting process. Therefore, extending the baseline can improve the accuracy of distance measurement. Either increasing the exposure time or the equivalent diameter results in





**Figure 4.** Comparison of DPCs and data quality under different parameters. When varying one parameter, other parameters are fixed at the values listed in Table 2. (a) baseline length  $B$ ; (b) total integration time  $t$ ; (c) equivalent diameter  $D$ ; (d) spectral resolution  $R$ . Left column: Comparison of simulated DPCs with error under different values of the parameter. Right: data quality (DQ) for DPCs differing only by the parameter (upper: amplitudes of the curves; lower: the relative error of DPCs, blue for total error and orange for Poisson error).



**Figure 5.** The relative uncertainty (blue points) and bias (orange points) of distance measurement change with each parameter. Upper left panel is for baseline length. Upper right panel is for exposure time. Lower left panel is for spectral resolution and lower right panel is for equivalent diameter. The large point is the result corresponding to the typical value listed in Table 2.

improving the accuracy of distance measurement, since the more photons that are collected, the smaller the absolute Poisson error. Interestingly, the larger the resolution is, the larger the relative error of DPCs, but the smaller the uncertainty of  $D_A$  will be. This is because as the resolution increases, the number of photons in each wavelength interval decreases and thus the Poisson error of the difference phase increases, meanwhile the amount of data on the DPC increases. As expected, the relative bias of  $D_A$  becomes smaller as the parameter becomes larger, and is even more sensitive than the relative error.

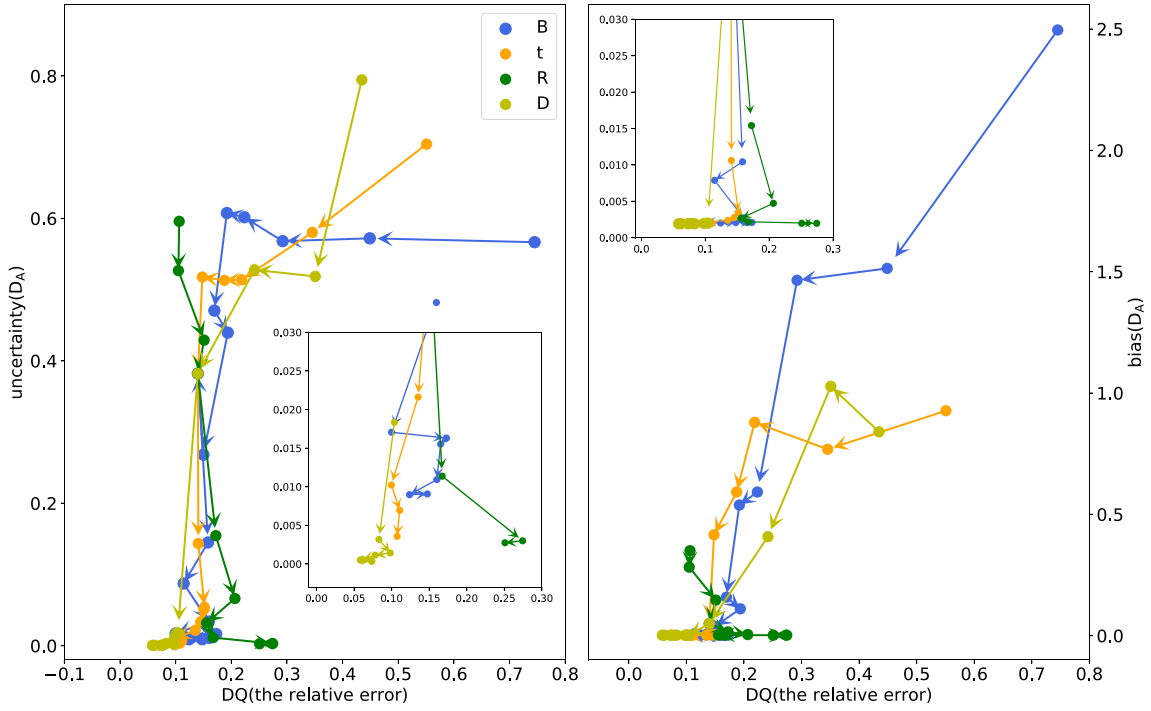
### 5.5. The Effect of Data Quality on the Accuracy of Distance Measurement

We also study how the uncertainty of distance measurement depends on the data quality of DPCs. Figure 6 visualizes the relationship resulting from our simulation. We see an obvious turnaround near the data quality around 20%. Only when the data quality is better than this turnaround point is the accuracy of  $D_A$  highly sensitive to the data quality. Different colors in Figure 6 represent different observational parameters. The direction of the arrow between data points represents parameters becoming larger and larger.

As the baseline increases (blue points), data quality gets better first, but there is no effect on the accuracy of distance measurement. When the relative error of DPC reaches better than 20%, extending the baseline can significantly improve the accuracy of  $D_A$ , although the data quality remains basically the same. Similar results are found by increasing  $t$  and  $D$ . However, increasing the spectral resolution (green points) gives more complex results. As the spectral resolution increases, data quality slowly deteriorates, but the accuracy of  $D_A$  is increasing. This may be due to the fact that the larger the spectral resolution, the more data points are on the DPCs, but the fewer photons there are in each wavelength interval. In actual observations, when DPCs with higher spectral resolution but lower data quality are obtained, it is customary to rebin the DPCs to reduce the resolution but increase the data quality. However, using DPCs that have not been rebinned is a better option for improving the accuracy of  $D_A$ .

## 6. Summary and Discussion

In this paper, we have studied the application of the SA method using a space-based optical interferometer to measure cosmological distance of quasars, and we have discussed how the basic observational parameters affect the accuracy of distance measurements. Our main results are as follows:



**Figure 6.** The relationship between the relative error of DPCs and the accuracy of distance measurement. Left: The relative uncertainty of distance measurement and the relative error of DPCs under different parameters. Right: The relative bias of distance measurement and the relative error of DPCs under different parameters. The arrow direction is the direction in which the parameters increase.

1. For the same target, the observed spatial distribution of the photocenters at different wavelengths will be affected by inclination angle  $\theta_i$  and the projected angle  $\theta_B$ .
2. Four basic observational parameters will affect the data quality of DPCs. As the baseline gets longer, the value of differential phase increases, as does the absolute error. When the exposure time or equivalent diameter increases, the absolute error, the relative error and the relative Poisson error decrease significantly. However, the larger the spectral resolution is, the larger the amplitude and relative error of DPCs are.
3. Four basic observational parameters will affect the accuracy of distance measurement. Extending the baseline will amplify the difference phases and the Poisson errors but will reduce the impact of OPD errors during the fitting process, so the uncertainty and bias of distance measurement both decrease. Increasing the exposure time or effective aperture can reduce the Poisson error due to the increase in the number of photons, thus improving the accuracy of distance measurement. With the increase of spectral resolution, the Poisson error increases due to the decrease of photon number in each wavelength channel, but the number of data on DPC increases, so although the relative error of DPC becomes larger, the uncertainty of distance measurement becomes smaller.

By assuming a parameterized BLR model and specifying the value of parameters, we obtained the surface brightness distribution of the BLR and continuum regions, then calculated the photocenters at different wavelengths. Combined with the basic observational parameters, we conduct an extensive set of simulations to get simulated DPCs, and then we utilized the MCMC method to obtain the posterior probability distributions of BLR model parameters from the mock data. In our simulation, we can eventually limit the relative uncertainty of data to 10%–20% by extending the baseline. The exposure time eventually limits the relative uncertainty to 5%–15% while the equivalent diameter can be limited to 5%–10%. Moreover, the results show that the SA method makes sense for distance measurement only if the relative error of DPCs is less than 20%. In this case, further increases in baseline length, exposure time and equivalent diameter will not greatly reduce the relative error of DPCs, but the uncertainty of distance measurement can be sharply reduced. However, it should be emphasized that these thresholds of the relative error are only valid for our simulation, since they depend on the specific set of BLR model and simulated observations. The simplest model for BLR is used in our simulation. Systematic errors, which are caused by complicated structures and kinematics of the BLR, such as non-disk structure, non-Keplerian kinematics and the disordered motion of these clouds, are not taken into account. So, the

relative error of DPCs would be underestimated during simulation then the accuracy of model parameters would be overestimated. On the other hand, we did not consider the uncertainty of RM data in the fitting process, which also resulted in overestimation of the accuracy of model parameters.

It also needs to be emphasized that both RM and SA are required to measure  $D_A$  in reality. The reason that we can rely on SA to determine  $R_{\text{BLR}}$  (without using the light curves from RM) is because a specific BLR model is used in our simulation to simulate DPCs, and then the same BLR model is used to perform the DPC fitting to retrieve all the model parameters. This is also why the degeneracy among the eight parameters is so weak. However, in reality the geometric distribution, dynamics and reprocessing coefficient of BLR clouds are much more complex, so it is not possible to have a perfect BLR model for DPC fitting, hence it still requires RM to measure  $R_{\text{BLR}}$ .

### Acknowledgments

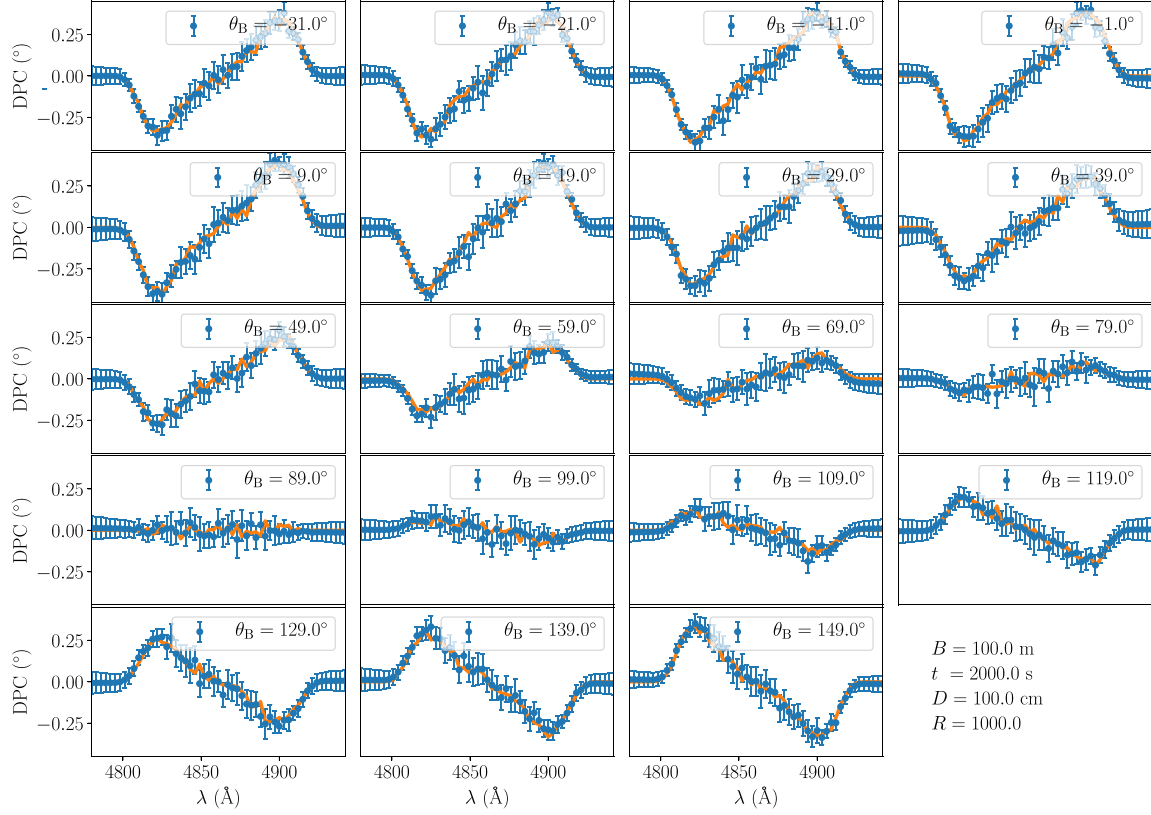
We are grateful to the referee for constructive suggestions that improved the manuscript. We acknowledge the financial support of the National Natural Science Foundation of China (Grants Nos. 12003077, 11703077, 12073068) and the Yunnan Province Foundation (202001AT070069).

### Appendix Mock Data

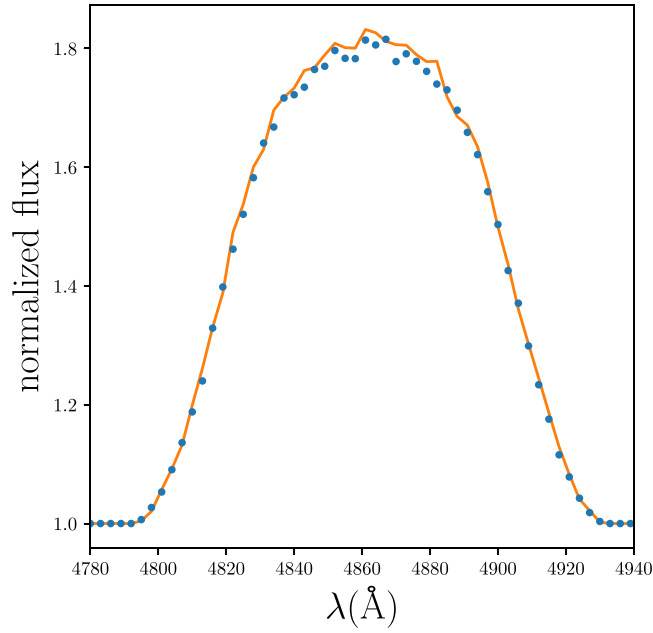
In this part, we use an example to illustrate the process of generating simulated data and the process of fitting. By adopting the parameterized BLR model described in Section 2 and the values of parameters listed in Table 1, we calculate the surface brightness distribution of BLR. We used  $2 \times 10^5$  (same as Gravity Collaboration et al. 2018) clouds in the model. Combined with the basic observational parameters and the values listed in Table 2, we simulate the  $\text{H}\beta$  emission line and DPCs. The spectral resolution used is 1000, so there are 67 wavelength bins between 4750 and 4950 Å. For each wavelength bin, we obtain the number of photons for each cloud that belongs to this bin. Combined with the position of the cloud, we calculate the total flux and the photocenter of the bin, and then calculate the total flux and the photocenter of this bin. The mock  $\text{H}\beta$  line profile (blue points) is shown in Figure A2. We generate 19 mock DPCs by changing projected angle from  $0^\circ$  to  $180^\circ$ . Blue points with errorbars in Figure A1 are the mock data.

We fit the BLR model to these mock DPCs. Bayesian statistics are used to measure confidence intervals of the model parameters. The priors are listed in Table 1. Markov Chain Monte Carlo code EMCEE is used to sample the posterior and obtain the posterior probability distribution of BLR model

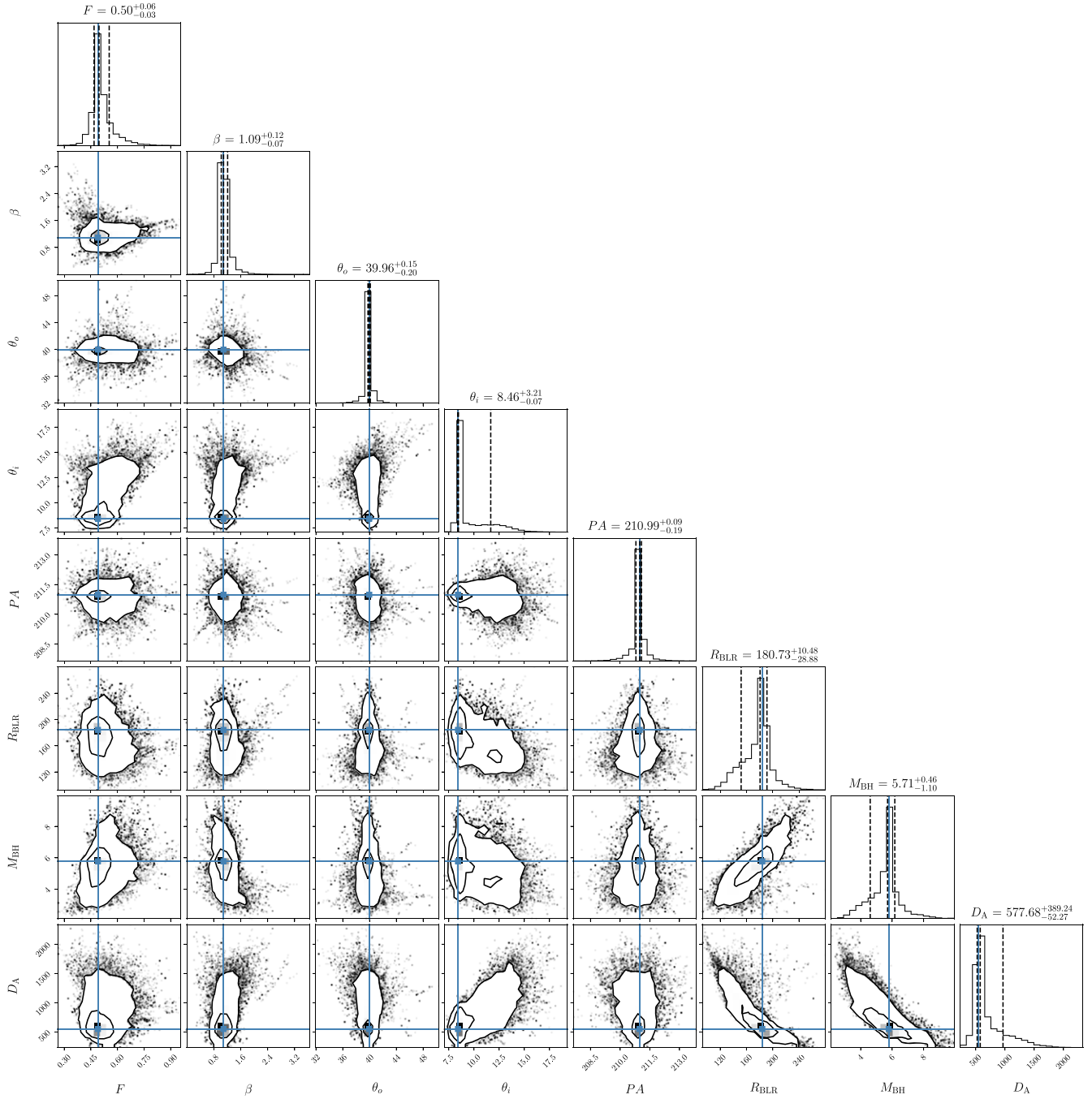




**Figure A1.** An example of DPCs we generated differing only in baseline projected angle  $\theta_B$ . The values of necessary interferometer parameters used are listed in the lower right part. Blue points with error bars are the differential phase. The thick solid lines are the best fitting using model parameters drawn from the probability distribution.



**Figure A2.** An example of  $H\beta$  profile we generated (blue points). The thick solid line is the best fitting using model parameters drawn from the probability distribution.



**Figure A3.** Probability density distributions of BLR parameters. The median values and error bars ( $1\sigma$  level) of the parameters are given on the tops of panels. The blue lines represent input values. The contours are at  $1\sigma$  and  $2\sigma$ . The dashed lines in the one-dimensional distributions are the 16%, 50% and 84% quantiles.

parameters. The results are shown in Figure A3. The median values of parameters are given on the tops of panels.

### ORCID iDs

Kai-Xing Lu  <https://orcid.org/0000-0002-2310-0982>

### References

- Antonucci, R. 1993, *ARA&A*, **31**, 473
- Bailey, J. A. 1998, *Proc. SPIE*, **3355**, 932
- Baldwin, J. A. 1977, *LicOB*, **748**, 1
- Beckers, J. M. 1982, *AcOpt*, **29**, 361
- Bentz, M. C., Denney, K. D., Grier, C. J., et al. 2013, *ApJ*, **767**, 149

- Bentz, M. C., Peterson, B. M., Pogge, R. W., et al. 2006, *ApJ*, **644**, 133
- Blandford, R. D., & McKee, C. F. 1982, *ApJ*, **255**, 419
- Clavel, J., Reichert, G. A., Alloin, D., et al. 1991, *ApJ*, **366**, 64
- Dietrich, M., & Kollatschny, W. 1995, *A&A*, **303**, 405
- Du, P., Hu, C., Lu, K.-X., et al. 2014, *ApJ*, **782**, 45
- Du, P., Zhang, Z.-X., Wang, K., et al. 2018, *ApJ*, **856**, 6
- Elvis, M., & Karovska, M. 2002, *ApJL*, **581**, L67
- Foreman-Mackey, D., Hogg, D. W., Lang, D., et al. 2013, *PASP*, **125**, 306
- Freedman, W. L. 2017, *NatAs*, **1**, 0121
- Freedman, W. L., & Madore, B. F. 2010, *ARA&A*, **48**, 673
- Gnerucci, A., Marconi, A., Capetti, A., et al. 2010, *A&A*, **511**, A19
- Gnerucci, A., Marconi, A., Capetti, A., et al. 2011, *A&A*, **536**, A86
- Gravity Collaboration, Abuter, R., Accardo, M., et al. 2017, *A&A*, **602**, A94
- Gravity Collaboration, Sturm, E., Dexter, J., et al. 2018, *Natur*, **563**, 657
- Grier, C. J., Pancoast, A., Barth, A. J., et al. 2017, *ApJ*, **849**, 146
- Ho, L. C. 2008, *ARA&A*, **46**, 475
- Hoyle, F. 1966, *Natur*, **210**, 1346
- Huang, Y.-K., Hu, C., Zhao, Y.-L., et al. 2019, *ApJ*, **876**, 102
- Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, *ApJ*, **533**, 631
- Kollatschny, W. 2003, *A&A*, **407**, 461
- Li, Y.-R., Songsheng, Y.-Y., Qiu, J., et al. 2018, *ApJ*, **869**, 137
- Li, Y.-R., Wang, J.-M., Ho, L. C., et al. 2013, *ApJ*, **779**, 110
- Longair, M. S., & Scheuer, P. A. G. 1967, *Natur*, **215**, 919
- Lynden-Bell, D., & Rees, M. J. 1971, *MNRAS*, **152**, 461
- Marconi, A., Maiolino, R., & Petrov, R. G. 2003, *Ap&SS*, **286**, 245
- Marziani, P., & Sulentic, J. W. 2014, *AdSpR*, **54**, 1331
- Osterbrock, D. E. 1989, *S&T*, **78**, 491
- Osterbrock, D. E., & Mathews, W. G. 1986, *ARA&A*, **24**, 171
- Pancoast, A., Brewer, B. J., & Treu, T. 2014a, *MNRAS*, **445**, 3055
- Pancoast, A., Brewer, B. J., Treu, T., et al. 2014b, *MNRAS*, **445**, 3073
- Peacock, J. A. (ed.) 1999, *Cosmological Physics* (Cambridge: Cambridge Univ. Press), 704
- Peterson, B. M. 1993, *PASP*, **105**, 247
- Peterson, B. M., Ferrarese, L., Gilbert, K. M., et al. 2004, *ApJ*, **613**, 682
- Petrov, R. G. 1989 in *Proc. NATO Advanced Study Institute Vol. 274* (Dordrecht: Kluwer), 249
- Petrov, R. G., Balega, Y. Y., Blazit, A., et al. 1992, in *European Conf. Proc. High-Resolution Imaging by Interferometry II*, 435
- Petrov, R. G., Malbet, F., Richichi, A., et al. 2001, *CRPhy*, **2**, 67
- Quercellini, C., Cabella, P., Amendola, L., et al. 2009, *PhRvD*, **80**, 063527
- Rakshit, S., Petrov, R. G., Meiland, A., et al. 2015, *MNRAS*, **447**, 2420
- Riess, A. G., Casertano, S., Yuan, W., et al. 2019, *ApJ*, **876**, 85
- Sandage, A. 1965, *ApJ*, **141**, 1560
- Tubbs, R. 2005, *ApOpt*, **44**, 6253
- Urry, C. M., & Padovani, P. 1995, *PASP*, **107**, 803
- Vannier, M., Petrov, R. G., Lopez, B., et al. 2006, *MNRAS*, **367**, 825
- Wandel, A., Peterson, B. M., & Malkan, M. A. 1999, *ApJ*, **526**, 579
- Wang, J.-M., Du, P., Valls-Gabaud, D., et al. 2013, *PhRvL*, **110**, 081301
- Wang, J.-M., Songsheng, Y.-Y., Li, Y.-R., et al. 2020, *NatAs*, **4**, 517
- Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, *PhR*, **530**, 87
- Williams, P. R., Pancoast, A., Treu, T., et al. 2018, *ApJ*, **866**, 75