Geometric Distances of Quasars Measured by Spectroastrometry and Reverberation Mapping: Monte Carlo Simulations

Yu-Yang Songsheng1,2, Yan-Rong Li1, Pu Du1, and Jian-Min Wang1,2,3
1 Key Laboratory for Particle Astrophysics, Institute of High Energy Physics, Chinese Academy of Sciences, 19B Yuquan Road, Beijing 100049, People’s Republic of China; wangjm@ihep.ac.cn
2 University of Chinese Academy of Sciences, 19A Yuquan Road, Beijing 100049, People’s Republic of China
3 National Astronomical Observatories of China, Chinese Academy of Sciences, 20A Datun Road, Beijing 100020, People’s Republic of China

Received 2020 November 25; revised 2021 January 28; accepted 2021 February 9; published 2021 April 13

Abstract

Recently, GRAVITY on board the Very Large Telescope Interferometer (VLTI) first spatially resolved the structure of the quasar 3C 273 with an unprecedented resolution of ~10 μas. A new method of measuring parallax distances has been successfully applied to the quasar through joint analysis of spectroastrometry (SA) and reverberation mapping (RM) observation of its broad-line region (BLR). The uncertainty of this SA and RM (SARM) measurement is about 16% from real data, showing its great potential as a powerful tool for precision cosmology. In this paper, we carry out detailed analyses of mock data to study impacts of data qualities of SA observations on distance measurements and establish a quantitative relationship between statistical uncertainties of distances and relative errors of differential phases. We employ a circular disk model of the BLR for the SARM analysis. We show that SARM analyses of observations generally generate reliable quasar distances, even for relatively poor SA measurements with error bars of 40% at peaks of phases. Inclinations and opening angles of BLRs are the major parameters governing distance uncertainties. It is found that BLRs with inclinations ≥10° and opening angles ≤40° are the most reliable regimes from SARM analysis for distance measurements. Through analysis of a mock sample of AGNs generated by quasar luminosity functions, we find that if the GRAVITY/GRAVITY+ can achieve a phase error of 0°1 per baseline for targets with magnitudes K ≤ 11.5, the SARM campaign can constrain H₀ to an uncertainty of 2% by observing 60 targets.

Unified Astronomy Thesaurus concepts: Distance measure (395); Reverberation mapping (2019); Optical interferometry (1168); Hubble constant (758)

1. Introduction

The distance of a celestial object can be in principle measured through the geometric relation \( D_\lambda = \Delta R / \Delta \theta \), where \( \Delta R \) and \( \Delta \theta \) are its linear and angular size, respectively. However, it is extremely hard to measure both \( \Delta R \) and \( \Delta \theta \) for the same object, in particular those at cosmological distances. Either \( \Delta R \) is too large, or \( \Delta \theta \) is too small. At cosmological distances, only active galactic nuclei (AGNs) and quasars can be feasibly measured for both \( \Delta R \) and \( \Delta \theta \) of their broad-line regions (BLRs) owing to the breakthrough progress of high spatial resolution and reverberation mapping (RM) of AGNs and quasars nowadays, respectively. Thanks are given to GRAVITY, an interferometric instrument operating in the K band at the Very Large Telescope Interferometer (VLTI), for its unprecedented high spatial resolution through spectroastrometry (SA; Eisenhauer et al. 2008; Gravity Collaboration et al. 2017), making it feasible to measure angular sizes of BLRs of AGNs. Recently, a direct measurement of angular diameter of the BLR of quasar 3C 273 through SA of VLTI (Gravity Collaboration et al. 2018) has reached a spatial resolution of ~10 μas, successfully revealing a flattened and Keplerian rotating disk-like structure of the BLR. Along with the measurement of its linear size by a long-term RM campaign of 3C 273 (Zhang et al. 2019), Wang et al. (2020) made a joint analysis of SA and RM (SARM) data and obtained the first parallax distance for the quasar with 16% precision. This is a compelling effort for quasar distances and shines a light on a new way for cosmology of the Hubble constant, which gives rise to an intensive debate currently known as H₀ tension (Riess et al. 2019).

Broad emission lines with FWHMs ranging from \( 10^3-10^4 \) km s⁻¹ are the prominent features of type I AGN and quasar spectra. They are from fast-moving clouds in BLRs photoionized by ionizing radiation from accretion disks around central supermassive black holes (SMBHs; Lynden-Bell 1969; Rees 1984). Variation in the strength and profile of the broad line will follow the ionizing continuum, but with a delay because of different paths of the broad line and ionizing photons. The delay is approximately equal to the light-travel time from the central source to the BLR. This is known as the RM of the BLR in AGNs (Blandford & McKee 1982). Physical sizes of BLRs can be simply estimated by cross-correlation functions (CCFs) between light curves of continuum and broad emission lines (Peterson 1993). RM campaigns spectroscopically monitoring AGNs can further probe geometries and dynamics of BLRs, as well as SMBH masses, by more advanced techniques such as velocity-resolved CCF (Bentz et al. 2010), transfer function recovery (Horne 1994), and dynamical modeling (Pancoast et al. 2011). Over the past few decades, about 120–150 AGNs with high-quality data have been measured for their sizes and black hole mass through AGN Watch (Peterson et al. 1998; Bentz et al. 2013), Bok2.4 (Kaspi et al. 2000), SEAMBHs (Super-Eddington Accreting Massive Black Holes; Du & Wang 2019), MAHA (Monitoring AGNs with Hβ Asymmetry; Du et al. 2018; Brotherton et al. 2020), and SDSS-RM projects (Shen et al. 2016; Grier et al. 2017; Fonseca Alvarez et al. 2020). More targets are being monitored and expected to be reported in the next few years.

For most AGNs, the sizes of their BLRs range from a few light-days to a few hundred light-days, and their angular diameters
hardly exceed $\sim 100 \mu\text{as}$, well below the imaging resolution of currently available facilities. Fortunately, due to the bulk motion (e.g., rotation, inflow or outflow) of the BLR gas, the gas moving toward observers and the gas moving away from observers are distributed in different spatial positions, making it possible to apply the super-resolution capability of SA to BLRs (Petrov et al. 2001; Marconi et al. 2003). For a source whose global angular size is smaller than the interferometer resolution, the interferometric phase is proportional to its angular displacements along the projected baseline from the direction where the optical path difference to the pair of telescopes vanishes (Petrov 1989). SA measures the interferometric phase as a function of wavelength near the broad emission line, thus obtaining angular displacements of clouds with different line-of-sight (LOS) velocities. The angular size, geometry, and dynamics of BLRs can be probed in this way.

Through a joint analysis of SARM data, the linear and angular size of the BLR can be determined simultaneously, providing a parallax measurement of an AGN’s distance. Such a measurement does not need calibration of any cosmic distance ladder or correction of extinction and reddening, providing a geometric way to measure the Hubble constant. The first SARM analysis of a single object, 3C 273, obtains $H_0 = 71.5_{-1.5}^{+1.9} \text{km s}^{-1} \text{Mpc}^{-1}$ with a precision of 16% (Wang et al. 2020). Moreover, based on the current capabilities of GRAVITY, about 50 AGNs are expected to be observed through SA, bringing the uncertainty of the current capabilities of GRAVITY, about 50 AGNs are expected to be observed through SA, bringing the uncertainty of $H_0$ below 2.5% (Wang et al. 2020). Future SA observations of fainter AGNs across a wider redshift range through powerful facilities (GRAVITY+/VLTI, Extremely Large Telescope) will be feasible. Not only can the Hubble constant be determined more precisely, but also a wider distance-redshift relation can be established to look back at the expansion history of our universe. Cosmology will be greatly advanced by this effort.

As a natural step to approach quasar distances through SARM observations, it is necessary to simulate mock data using the Monte Carlo method and quantify uncertainties of distance measurements under varied data qualities and object properties. It will provide guidance for arranging the observation campaign and selecting appropriate targets to minimize uncertainties. This paper is structured as follows. Section 2 presents the framework used to generate mock data and estimate the input model parameters. Section 3 analyzes and summarizes simulation results. Discussions are provided in Section 4, and conclusions are given in Section 5.

2. Methodology

In order to study impacts of various factors on the constraint ability of the SARM campaign on cosmology, we will use a parameterized BLR model to generate mock data of SARM observations. Appendix A presents details of mock data generations. Reconstruction of the BLR model can be done through mock data fitting via the diffusive nested sampling (DNest) algorithm\(^4\) (Brewer & Foreman-Mackey 2018), obtaining the probability distribution of model parameters, including the angular distance of the AGN. It is our goal to analyze how uncertainties and degeneracies of reconstructed parameters change with relative errors of data and values of input parameters. The result can provide guidance for the object selection and observation strategy in future SARM campaigns to meet the needs of cosmology. The method to generate and fit SARM data is implemented in BRAINS.\(^3\)

2.1. Parameterized BLR Model

Over the past few decades, a sample of more than 100 AGNs with RM observations have advanced our understanding of BLRs a lot (Peterson et al. 1998; Kaspi et al. 2000; Bentz et al. 2013; Du et al. 2018). Particularly, geometries and dynamics of BLRs can be well studied through velocity-resolved delay maps obtained by RM (Grier et al. 2013). A disk-like BLR is common in many broad-line Seyfert 1 galaxies (Bentz et al. 2013; Grier et al. 2013; Du et al. 2016; Lu et al. 2016; Xiao et al. 2018), and even in some narrow-line Seyfert 1 galaxies (Du et al. 2016). Multiple RM observations of several objects, such as NGC 5548, 3C 390.3, and NGC 7469, show that the FWHM of H$\beta$ line and its lag $\tau_{H\beta}$ follows $\tau_{H\beta} \propto \text{FWHM}^{-1/2}$, indicating Keplerian rotation of the BLR gas (Peterson et al. 2004). A flattened disk-like BLR in 3C 273 has also been detected by GRAVITY. A disk-like BLR with Keplerian rotation could be common.

A parameterized BLR model has been applied widely in RM data fitting in order to obtain physical quantities, such as BLR size and black hole mass, in a self-consistent way (Pancoast et al. 2014b; Li et al. 2018; Williams et al. 2018). A comprehensive model with $\sim 30$ parameters has been developed to include complex geometries and dynamics for perfect fitting of the observation data (Pancoast et al. 2014a). But as a preliminary guidance for SARM, we reasonably assume the Keplerian rotating disk model for the BLR and include only the fundamental parameters for concise estimations of parameter impacts. This model also works quite well in Gravity Collaboration et al. (2018) and Wang et al. (2020) when fitting SA and SARM data of 3C 273, respectively.

The BLR is composed of a large amount of line-emitting clouds on circular orbits around the central black hole with Keplerian velocities, as shown in Figure 1. The radial distribution of clouds is described by a shifted $\Gamma$-distribution. The distance of BLR clouds to the SMBH is generated by

$$r = R_S + \mathcal{F} R_{BLR} + \Gamma_0 \beta^2 (1 - \mathcal{F}) R_{BLR},$$  \hspace{1cm} (1)

where $R_S$ is the Schwarzschild radius, $R_{BLR}$ is the mean radius, $\mathcal{F} = R_{in}/R_{BLR}$ is the fraction of the inner to the mean radius, $\beta$ is the shape parameter, and $\Gamma_0 = p(x|\beta^{-2}, 1)$ is a random number drawn from a $\Gamma$-distribution

$$p(x|\alpha, \chi_0) = \frac{x^{\alpha-1} \exp(-x/\chi_0)}{\chi_0^{\alpha} \Gamma(\alpha)},$$  \hspace{1cm} (2)

where $x_0$ is a scale factor ($x_0 = 1$ here), $\alpha = \beta^{-2}$, and $\Gamma(\alpha)$ is the $\Gamma$-function. Symbols and meanings of all free parameters used in the BLR model are summarized in Table 1.

Our model in this work differs a little from that in Gravity Collaboration et al. (2018) and Wang et al. (2020) in the orientation distribution of orbital planes. We assume that the cosine of the angle between the direction of orbital angular momentum and $Z$-axis is uniformly distributed over $[\cos \theta_{\text{tan}}, 1]$ to reach a uniform distribution of clouds. In previous work, the angle between the direction of angular momentum and $Z$-axis is uniformly distributed over $[0, \theta_{\text{tan}}]$, and so clouds will

---

\(^3\) The DNest algorithm was proposed by Brewer et al. (2011), and an implementation package developed by the authors is available at https://github.com/eggplantbren/DNest4. In this work, we use our own DNest package CDNest (Li 2020), which is written in C language and enables the standardized parallel message passing interface, which is available at https://github.com/LiyrAstroph/CDNest.

\(^4\) https://github.com/LiyrAstroph/BRAINS
be accumulated near the equatorial plane. We make the change for a more well-defined definition of opening angles. Clouds are randomly distributed along a given orbit by assigning their orbital phases uniformly over \([0, 2\pi]\).

### 2.2. Spectroastrometry

A detailed mathematical formulation of SA of BLRs can be found in Rakshit et al. (2015) and Songsheng et al. (2019a). We summarize it here for the reader’s convenience. For an interferometer with a baseline \(B\), the differential interferometric phase for a nonresolved source is

\[
\phi_{\text{d}}(\lambda) = -2\pi u \cdot [\epsilon(\lambda) - \epsilon(\lambda_0)],
\]

where \(u = B / \lambda\) is the spatial frequency, \(\epsilon\) is the photocenter of the source at wavelength \(\lambda\), and \(\lambda_0\) is the wavelength of a reference channel. Here the bold letters represent vectors. Given the surface brightness distribution \(O(\alpha, \lambda)\) of the source, we have

\[
\epsilon(\lambda) = \int \alpha O(\alpha, \lambda) d^2\alpha / \int O(\alpha, \lambda) d^2\alpha,
\]

where \(\alpha\) is the angular displacement on the celestial sphere. For AGNs, \(O = O_\ell + O_c\), where \(O_\ell\) and \(O_c\) are the surface brightness distributions contributed by the BLR and continuum region, respectively. Once the BLR model is set up, \(O_\ell\) can be calculated through

\[
O(\alpha, \lambda) = \int \Xi_x F(\mathbf{r}, \mathbf{V}) \delta(\alpha - \alpha')\delta(\lambda - \lambda') d^3r d^3V,
\]

where \(\lambda = \lambda_{\text{cen}} (1 + V \cdot \mathbf{n}_{\text{obs}} / c) (1 - R_s / r)^{-1/2}\) is the shifted wavelength of the photon from the broad emission line centered at \(\lambda_{\text{cen}}\) by Doppler effect and gravitation; \(\gamma_0 = (1 - V^2 / c^2)^{-1/2}\) is the Lorentz factor; \(\alpha' = [r - r \cdot \mathbf{n}_{\text{obs}}] \mathbf{n}_{\text{obs}} / D_A\); \(r\) is the displacement to the central black hole; \(\Xi_x\) and \(F(\mathbf{r}, \mathbf{V})\) are the reprocessing coefficient and velocity distribution of the clouds at position \(r\), respectively; \(F_c\) are ionizing fluxes received by an observer; \(\mathbf{n}_{\text{obs}}\) is the unit vector pointing from the observer to the source; and \(D_A\) is the angular size distance of the AGN. Introducing the fraction of the emission line to total (\(\ell_\lambda\)), we have

\[
\epsilon(\lambda) = \ell_\lambda \epsilon_{\ell}(\lambda),
\]

where

\[
\epsilon_{\ell}(\lambda) = \int r O_\ell d^2\alpha / \int O_\ell d^2\alpha = \ell_\lambda F_\ell(\lambda) / F_\text{tot}(\lambda),
\]

\[
= \int O_c d^2\alpha, F_\text{tot}(\lambda) = F_\ell(\lambda) + F_c(\lambda).
\]
The differential phase curve (DPC) can be obtained by inserting Equations (6), (5), and (4) into Equation (3). If the DPC whose amplitudes are a few degrees can be measured with SA, the effective spatial resolution will be 100 times the resolution limit \( \lambda / B \).

2.3. Reverberation Mapping

A detailed mathematical formulation of one-dimensional RM can be found in Li et al. (2013, 2018). We also summarize it here for the reader’s convenience. A damped random walk (DRW) model is used to describe continuum variations in order to interpolate and extrapolate light curves of the continuum (Kelly et al. 2009; Zu et al. 2013). We first express the continuum light curve \( y \) by \( y = s + n + E_q \), where \( s \) denotes the underlying signal of the variation, \( n \) represents the measurement noise, \( q \) is the mean value of the light curve, and \( E \) is a vector whose elements are all 1. In the DRW model, the covariance function between \( s \) and \( q \) is given by

\[
S(t_i, t_j) = \sigma_d^2 \exp \left( -\frac{|t_i - t_j|}{\tau_d} \right),
\]

where \( t_i \) and \( t_j \) are times for signal \( s_i \) and \( s_j \), respectively; \( \sigma_d \) is the long-term standard deviation of the variation; and \( \tau_d \) is the typical timescale.

We further assume that both \( s \) and \( n \) are Gaussian and uncorrelated. Given \( y \), the posterior distribution of \( \sigma_d, \tau_d \), and \( q \) can be obtained by Bayesian analysis with likelihood function

\[
P(y|\sigma_d, \tau_d, q) = \frac{1}{\sqrt{(2\pi)^m|C|}} \exp \left[ -\frac{(y - E_q)^T C^{-1} (y - E_q)}{2} \right],
\]

where superscript “\( T \)” denotes transposition; \( C = S + N \); \( S \) and \( N \) are the covariance matrix of \( s \) and \( n \), respectively; and \( m \) is the number of data points.

Given parameters \( (\sigma_d, \tau_d, q) \), a typical realization for the observed continuum light curve will be

\[
f_c = (u_s + \hat{s}) + E(u_q - \hat{q}),
\]

where \( \hat{q} = E^T C^{-1} y / E^T C^{-1} E \), \( \hat{s} = S^T C^{-1} (y - E_q) \) and \( u_s \) and \( u_q \) are Gaussian processes with zero mean and covariance matrices \( Q = (S^T + N^{-1})^{-1} \) and \( C_q = (E^T C^{-1} E)^{-1} \), respectively. Next, \( u_s \) and \( u_q \) are treated as free parameters and further constrained by light curves of the emission line.

Given the BLR model and the realization of continuum light curves, the variation of the emission line is calculated by

\[
f(t) = \int_{-\infty}^{+\infty} d\tau d\nu \frac{\Xi_e f_c(t')}{4\pi \tau^2} n(r) \delta(t' - t + \tau),
\]

where \( \tau = (r - r \cdot n_{\text{obs}})/c \), \( \Xi_e \) is the reprocessing coefficient, and \( n(r) \) is the number density of the clouds.

2.4. Joint Analysis

The goal of a fully joint analysis is to obtain the posterior probability distribution of the model parameters consistently using the combined data from SARM observations. SARM observations are conducted independently; thus, we assume that the probability distributions for the measurement errors of light curves, profiles, and DPCs are uncorrelated. The joint likelihood function can be written as

\[
P(\mathcal{D}|\Theta) = \prod_{i=1}^{N_{\mathcal{D}}} \frac{1}{\sqrt{2\pi\sigma}\varepsilon} \exp \left\{ -\frac{[f_{\text{obs}} - f_{\text{mod}}(\Theta)]^2}{2\sigma^2} \right\}
\]

\[
\times \prod_{j=1}^{N_{\mathcal{D}}} \prod_{k=1}^{N_{\mathcal{D}}} \frac{1}{\sqrt{2\pi\sigma_{\phi}\varepsilon}} \exp \left\{ -\frac{[\phi_{\text{obs}} - \phi_{\text{mod}}(\Theta)]^2}{2\sigma_{\phi}^2} \right\}
\]

\[
\times \prod_{j=1}^{N_{\mathcal{D}}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{[f_{\text{obs}} - f_{\text{mod}}(\Theta)]^2}{2\sigma_{\phi}^2} \right\},
\]

(11)

where \( \mathcal{D} \) represents the data set; \( \Theta \) represents all the model parameters; \( f_{\text{obs}}, f_{\text{mod}} \) are the continuum data; \( f_{\text{obs}}, f_{\text{mod}} \) and \( \phi_{\text{obs}}, \phi_{\text{mod}} \) are the linear flux, line profile, and the DPC of the emission line with measurement uncertainties \( \sigma_{\phi}, \sigma_{\phi} \), respectively; and \( f_{\text{mod}}, f_{\text{mod}}, \phi_{\text{mod}} \) are the corresponding predicted values from the BLR model.

In light of Bayes’s theorem, the posterior probability distribution for \( \Theta \) is given by

\[
P(\Theta|\mathcal{D}) = \frac{P(\Theta|\mathcal{D}) P(\mathcal{D}|\Theta)}{P(\mathcal{D})},
\]

(12)

where \( P(\Theta) \) is the prior distribution of the model parameter and \( P(\mathcal{D}) \) is a normalization factor. The DNest algorithm can be applied to sample the distribution Equation (12). A brief introduction to the DNest algorithm is given in Appendix C.

3. Results

3.1. Impacts of Data Qualities

SARM data consist of light curves of optical continuum and emission lines, profiles of near-IR (NIR) emission lines, and their DPC. Observation of light curves and line profiles are relatively mature, and data quality for most SARM targets can be quiet good. On the contrary, techniques for obtaining DPCs are still in development, and data quality largely depends on seeing and target brightness. In this subsection, we mainly study the impact of DPC data quality on the measurement of distance, black hole mass, and BLR geometry.

To generate typical mock data, we take the fiducial values in Table 1 for the BLR model. The light curves of continuum and emission lines last for 200 days with a 1-day cadence. The continuum light curves are generated by the DRW model with \( \tau_d = 60 \) days and \( \sigma_d = 0.25 \). The relative errors of continuum and line variations are 0.5% and 1%, respectively. We generate mock RM data with higher qualities than those in typical RM campaigns in order to focus on impacts of data qualities of SA observations on distance measurements. The redshift of the target is assumed to be 0.01, and the NIR emission lines for SA are Brackett \( \gamma \) (Br\( \gamma \)) centered at 2.166 \( \mu \)m in rest frame. The equivalent width of Br\( \gamma \) is 40 \( \kappa \). We follow the spectral resolution of GRAVITY (\( \lambda / \Delta \lambda / 500 \)) and broaden the profile by a Gaussian with FWHM = 4 nm (\( \sim 600 \) km s\(^{-1}\)). The relative error of the profile is 0.5%. The configuration of the
baseline we used for generating the DPC is the same as that of Extended Data Figure 1 (b) in Gravity Collaboration et al. (2018) (which means that the target is observed four times in total). We also assume that the absolute error of phase is the same in all wavelength channels and baselines for simplification.

To quantify data qualities of DPCs, we define the relative error as the ratio of phase error to the peak of DPCs with largest amplitudes. We vary the relative error of the DPC to be 5%, 10%, 20%, and 40%. For each data set, we obtain the posterior probability distribution of model parameters through Bayesian analysis introduced in Section 2.4. An example of generation and fitting of mock data is illustrated in Appendix A.

We define the inferred value of the model parameters as the median of the posterior distribution, and the lower/upper bound as 16% and 84% quantile. Uncertainties of the parameters are defined as half of the difference between the upper and lower bounds, and bias as the difference between the inferred and input value. Finally, for dimensional quantities, we divide the uncertainty by inferred value and bias by input value to get the relative uncertainty and bias; for angles, we divide them by 1 rad instead. We emphasize here that bias is not systematic uncertainty. It is the deviation of inferred values of model parameters from input ones caused by the finite width of the posterior probability distribution (uncertainty). The absolute value of bias is comparable to uncertainty usually. A much larger bias than uncertainty indicates the failure of Bayesian analysis, mostly caused by degeneracy between model parameters.

Relative uncertainties and bias of part of the model parameters under different SA data qualities are shown in Figure 2. As we can see in the top panel of Figure 2(a), relative uncertainties of $R_{\text{BLR}}$ are 2.58% and independent of errors of DPCs, since it is only constrained by light curves of continuum and emission lines. Uncertainties of PA are proportional to errors of DPCs, and the proportional coefficient is approximately 0.14. Since $D_A = R_{\text{BLR}} / \xi_{\text{BLR}}$, where $\xi_{\text{BLR}}$ is the mean angular size, we have

$$\delta D_A = \sqrt{\delta R_{\text{BLR}}^2 + \delta \xi_{\text{BLR}}^2},$$

where $\delta_x$ means the relative uncertainty of $x$. If we assume $\delta_{\text{BLR}} = k \delta_{\text{DPC}}$ and let $\delta_{R_{\text{BLR}}} = 2.58\%$, we can get through fitting

$$\delta D_A = \sqrt{(2.58\%)^2 + (0.23\delta_{\text{DPC}})^2}.$$  

We must emphasize that this relation is valid only for this set of parameters, since relative uncertainties of $D_A$ also depend on
inclinations and opening angles, as shown in the next subsection.

There are no systematic uncertainties for mock data analysis, but they should be considered in real observations. The bottom panel of Figure 2(a) presents the relation between bias and data error. Median values of posterior distributions are obtained with greater randomness than uncertainties, leading to a more irregular pattern. The bias of $R_{\text{BLR}}$ changes little with data error, while the bias of $D_A$ increases with data error, in agreement with results in the top panel. The bias of PA fluctuates when $\delta_{\text{DPC}}$ increases to just below 20%, but increases a lot when $\delta_{\text{DPC}} = 40%$.

As shown in Figure 2(b), for $M_*$, $i$, and $\theta_{\text{opn}}$, there is no systematic change of uncertainty or bias when the error of DPC data varies, because these parameters are mainly constrained by profiles of emission lines. But we note that there is a dramatic increase of bias for $M_*$ when $\delta_{\text{DPC}} = 20%$. It may be caused by strong correlation between $M_*$, $i$, and $\theta_{\text{opn}}$, as shown in Figure 3. The correlation comes from the fact that the FWHM of the emission line can remain unchanged if we increase black hole mass and decrease opening or inclination angle simultaneously. In the case of strong correlation, input values have exceeded the $1\sigma$ range in one-dimensional distributions, though they are still on the boundary of the $1\sigma$ range in two-dimensional distributions. Furthermore, the sampling algorithm we used also performs worse when correlations between parameters get stronger. Thus, strong correlation between parameters can lead to a large bias in fitting.

### 3.2. Amplitudes of DPCs

In order to estimate the precision of the distance measurement for a sample of targets, we need to predict amplitudes of DPCs for given magnitudes, redshifts, and profiles of emission lines. From Equation (3), the differential phase depends on the photocenter $\ell(\lambda)$ of the target, as well as its projection to the baseline $B$ of the interferometer.

Given the redshift of an AGN, its angular and luminosity distance can be estimated easily using current values of cosmological parameters. The luminosity of the AGN can then be obtained from its magnitude. The widely used $R-L$ relation (Bentz et al. 2013; but see Du & Wang 2019 for its revised version) now comes in and predicts the average size of the BLR. However, a small inclination angle or large opening angle would make the system more symmetric, leading to a photocenter displacement much smaller than the average angular size of the BLR. Meanwhile, when the emission line is weak compared to the continuum, the peak of $\epsilon(\lambda)$ will decrease significantly from Equation (6). In order to quantify all these effects, we choose $i$, $\theta_{\text{opn}}$, $F$, and $\beta$ randomly from the ranges shown in Table 1 to generate a large sample of $\epsilon(\lambda)$ and corresponding line profiles. Dimensional parameters $D_A$, $R_{\text{BLR}}$, and $M_*$ will be fixed since they can only change the overall amplitude and width rather than the shape of the DPC and profile. We divide the sample
into two categories according to the number of peaks in their line profiles. For each category, we calculate the median value and 1σ limit of the ratio $\xi_{\text{BLR}}$ under different $f_{\text{max}}$, where $\xi_{\text{BLR}}$ is the average angular size of the BLR, and $f_{\text{max}}$ and $f_{\text{max}}$ are the maximum values of $|\epsilon(\lambda)|$ and $f_\ell(\lambda)$, respectively. The result is shown in Figure 4(a). Clearly, the displacement of the photocenter is roughly proportional to the ratio of line flux to total flux at line center, and photocenters of targets with multiple peaks in their profiles are systematically larger than those with single peaks.

In order to predict the peak amplitude of the actual phase signal, we must know the projected baseline configuration and position angle of the target. For simplicity, we assume that targets are at the zenith when observed and that the configuration of baselines is the same as that of VLTI. Position angles of targets are uniformly distributed between $0^\circ$ and $360^\circ$. The result is shown in Figure 4(b). Similarly, phase amplitudes of targets with multiple peaks in their profiles are larger than those with a single peak.

3.3. Impacts of Inclination and Opening Angles

Even if the data qualities are the same, uncertainties and accuracies of distance measurements can vary with shapes of DPCs and profiles. Dimensional parameters $D_A$, $R_{\text{BLR}}$, and $M$, can only change the overall angular width and width rather than the shape of the DPC and profile. Their impacts on the distance measurement can be converted to impacts of relative error of the data. Among remaining parameters, $i$ and $\theta_{\text{opn}}$ can change the shape of the DPC and profile drastically (Rakshit et al. 2015; Songsheng et al. 2019b), thus altering degeneracies among all parameters, consequently further changing the uncertainty of parameter inference. In order to study their impacts, we vary $i$ and $\theta_{\text{opn}}$ from $5^\circ$ and $45^\circ$, respectively, and keep other parameters the same as those in Section 3.1 and the relative error of the DPC at 20%. For each data set, we perform the Bayesian analysis and calculate relative uncertainty and bias for each parameter. The results are shown in Figure 5.

Generally, uncertainties of $D_A$ become large when $i$ decreases, as illustrated in Figure 5(a). For low inclinations (such as $i \lesssim 10^\circ$), the profile of the broad emission line usually possesses a single peak (slightly depends on opening angles), as shown in Figure 4(c). In this case, the inclination is hard to determine accurately owing to degeneracies with other parameters. Furthermore, the LOS approaching the symmetry axis will significantly increase the symmetry of the system, thus decreasing the amplitude of the DPC drastically. Thus, the correlation between $i$ and $D_A$ becomes much stronger, as shown in Figure 6, leading to larger uncertainties of $D_A$.

For moderate inclinations ($i \sim 15^\circ$–$25^\circ$), uncertainties of $D_A$ also increase with opening angles of BLRs owing to increasing symmetries. There are obvious leaps of correlations between $D_A$ when $\theta_{\text{opn}} = 25^\circ$ ($i = 15^\circ$) and $\theta_{\text{opn}} = 35^\circ$ ($i = 25^\circ$), since profiles become single peaked when $\theta_{\text{opn}} > i$. If inclinations become larger ($i \sim 35^\circ$–$45^\circ$), systems will keep asymmetry and degeneracies between $D_A$ and $i$ are weak, contributing to much smaller uncertainties of $D_A$.

The variation of the $D_A$-bias with inclinations and opening angles is roughly similar. We emphasize that the Bayesian inference becomes very unstable for a system with low inclinations and large opening angles because the system will be highly symmetric in the view of the observer. In such a case, extremely large biases ($\gt 100\%$) appear, even much larger than uncertainties, and the inferred value of distance cannot be trusted.

In Figure 5(b), the evolution of the uncertainty of $i$ with opening angles is similar to that of $D_A$. There is no systematic increase of $\delta_i$ as $i$ decreases since our definition of $\delta_i$ is $\Delta_i/1 \text{ rad}$ rather than $\Delta_i/i$. In Figure 5(c), uncertainties of PA measurements increase slightly when inclinations decrease, but they can be determined much more accurately compared to
other parameters. In Figure 5(d), uncertainties of measurements of $M_*$ also tend to be large for face-on targets, since a slight change of $i$ can cause a significant variation of width of the DPC or profile when inclination is small. In Figure 5(e), inference of $R_{\text{BLR}}$ is less affected by the inclination and opening angle, as one-dimensional RM is insensitive to them. In Figure 5(f), uncertainties of $\theta_{\text{opn}}$ vary with inclinations and opening angles similarly to those of $i$ owing to correlations between $i$ and $\theta_{\text{opn}}$. Generally, values of $\Delta i$ and $\Delta \theta_{\text{opn}}$ are relatively close.
4. Discussion

4.1. Error Budget

The statistical relative uncertainties of distance measurements can be written as

$$\delta_{D_x} = \sqrt{\delta_{R_{BLK}}^2 + \alpha^2 \delta_{DPC}^2},$$

(15)

where the coefficient $\alpha$ depends on inclination and opening angles of BLRs. Multiple observations ($N$) of a target can significantly reduce the statistical uncertainty. Considering these influences, we have

$$\delta_{D_x} = \sqrt{\delta_{R_{BLK}}^2 + \alpha^2 \delta_{DPC}^2}/N.$$  

(16)

From simulations shown in Figure 5, we can estimate the value of $\alpha$ for different inclination and opening angles. The result is shown in Figure 7. When the inclination is very small ($i \sim 5^\circ$), the $\alpha$ is around 2, and the bias of inference is too large to obtain a credible measurement. For moderate inclinations ($i \sim 15^\circ$–$25^\circ$), we have $\alpha \sim 0.5$ when $i \geq \theta_{opn}$ and $\alpha \sim 1.5$ when $i \leq \theta_{opn}$. For large inclinations ($i \sim 35^\circ$–$45^\circ$), $\alpha$ is about 0.4.

Inclinations and opening angles can be in principle inferred through the number of peaks in line profiles. As shown in Figure 4(c), if inclinations are large and opening angles are small ($i > \theta_{opn}$), simulated profiles usually possess multiple peaks, and vice versa. So selecting objects with multiple peaks in their profiles can reduce uncertainties in distance measurements notably. In practice, due to the blend of narrow-line emission and instrument broadening effect, the proportion of objects with multiple peaks in profiles in real samples is much smaller than that in our simulated samples (see, e.g., Liu et al. 2019, Table 2). A quick way to estimate inclinations and opening angles of a large sample of objects is to compare their profiles with template profiles simulated by BLR models. Objects with $i < 10^\circ$ or $\theta_{opn} > 40^\circ$ can be excluded at the first step. Before making observation plans, we can fit profiles of specific objects to the BLR model to obtain a reliable estimation for inclination and opening angles (e.g., Raimundo et al. 2019, 2020). Then, we can predict the expected DPC for each object and choose appropriate ones for the SARM campaign. Other ways to determine inclinations of AGNs include narrow-line region kinematics (Fischer et al. 2013) and parsec-scale radio jets (e.g., Kun et al. 2015).

Systematic errors of distance measurement through SARM have been discussed thoroughly in Wang et al. (2020). First, the RM campaign measures the region of optical emission lines (usually H$\beta$ for objects with low redshifts), while SA conducted by GRAVITY can only measure the region of NIR emission lines (usually Br$\gamma$ or Pa$\alpha$). Their size may be slightly different owing to the different optical depths. But this can be reduced by comparing widths of different emission lines in joint analysis of velocity-resolved SARM observations. For 3C 273, the difference between the size of the H$\beta$- and Pa$\alpha$-emitting region is estimated to be 13%, which is slightly smaller than the statistical error. An RM campaign using the same emission line as that in SA observation is highly needed to quantify this effect. Second, the RM campaign measures the variable part of a BLR, while SA observes the entire region, resulting in systematic errors in the SARM analysis. By comparing the shape and width of the mean with rms spectra across the whole RM campaign, the error can also be assessed. For 3C 273, the variable part of its BLR shows little difference with the entire BLR. Finally, signatures of inflow or outflow have been found recently by analyzing high-quality RM data...
The Astrophysical Journal Supplement Series, 253:57 (15pp), 2021 April

Songsheng et al.

The general shape of the differential phase curve for an inflow BLR is similar to that of a Keplerian BLR, except that the displacement of the photocenter would be parallel to the axis rather than perpendicular to it (Rakshit et al. 2015). The angular size of a BLR can still be well constrained by SA observation, and so the uncertainty of distance through SARM analysis will not change a lot. However, the relation between clouds’ velocities and black hole mass will be much more ambiguous, increasing the uncertainty of black hole mass measurement. Based on this limited information and the fact that the angular distance of 3C 273 measured by SARM is consistent with that from the current cosmological model, we draw a conclusion that systematic errors are at most comparable to statistical errors.

4.2. Model Tests

In our mock data analysis, the correctness of the BLR model is always guaranteed, and all the fittings can reach the $\chi^2 \approx 1$ level. However, BLRs are diverse individually. Our simplest model works quite well for 3C 273 since its emission-line profiles are symmetric. When selecting targets, objects with simple and regular line profiles should be prioritized. The response of emission line to continuum also needs to be clear to avoid complications by long-term trending. BLR “holiday,” multiple lags, or fast breathing. Before Bayesian analysis, model-independent methods, such as reconstruction of a velocity-delay map through the maximum entropy method or calculation of the centroid position of the photocenter in velocity-delay map through the maximum entropy method or model-independent methods, such as reconstruction of a multiple lags, or fast breathing. Before Bayesian analysis, model-independent methods, such as reconstruction of a velocity-delay map through the maximum entropy method or calculation of the centroid position of the photocenter in several wavelength channels, should be applied to obtain the basic geometry and dynamics of the BLR. The appropriate model with a minimum number of parameters will be tested first. Subsequently, more can be added until a good fitting is reached.

Actually, the radial distribution of BLR clouds described by Equations (1) and (2) can be alternatively assumed by other forms, such as Gaussian distributions (Pancoast et al. 2011). This leads to different results of distances from the fittings, which could be regarded as one of systematical errors. This can be addressed by evaluating the Bayesian evidence of the model, which is a prime result of the nested sampling method (Skilling 2006; Shaw et al. 2007). Moreover, SMBH mass can be simultaneously obtained by the SARM analysis. We will discuss these issues in a separate paper.

4.3. H$_0$ Tension

Since the beginning of the previous century, advances in the distance measurement of extragalactic objects have been driving the development of cosmology. In particular, the cosmic distance ladder using Cepheid and Type Ia supernovae as standard candles has made great achievements, including the discovery of the expansion of the universe (Hubble 1929), as well as its acceleration (Riess et al. 1998; Perlmutter et al. 1999). Currently, the Equation of State of Dark energy (SH0ES) program achieves a measurement of the Hubble constant ($H_0$) with 1.91% precision based on the distance ladder, providing a value of $H_0 = 74.03 \pm 1.42$ km s$^{-1}$ Mpc$^{-1}$ (Riess et al. 2019). However, observations of the cosmic microwave background (CMB) radiation from the early universe enable an indirect inference of $H_0$ under the assumption of the flat $\Lambda$CDM cosmological model, giving $H_0 = 67.4 \pm 0.5$ km s$^{-1}$ Mpc$^{-1}$ (Planck Collaboration et al. 2020). Tensions between the early and late universe measurements are at the level of $\sim 4.4\sigma$ (Verde et al. 2019), indicating that either unaccounted-for systematic errors might bias at least one of these measurements (see discussion of this possibility in Davis et al. 2019; Rameez & Sarkar 2019; Rose et al. 2020), or the flat $\Lambda$CDM cosmological model needs to be extended to include new physics, such as dynamical/interacting dark energy (e.g., Di Valentino et al. 2016, 2017; Huang & Wang 2016; Zhao et al. 2017), dark radiation (e.g., Buen-Abad et al. 2015; Ko & Tang 2016), neutrino self-interaction (e.g., Blinov et al. 2019; Kreisch et al. 2020), and so on.

At such a crossroads, new approaches to determining the Hubble constant without calibration of the local distance ladders or assumption of the flat $\Lambda$CDM cosmological model are particularly important as arbitrators. Gravitational waves emitted by binary neutron star merger can be used as a “standard siren” to determine the luminosity distance of the binary (Schutz 1986). The recent observation of the merger signal GW180817 (Abbott et al. 2017c), along with its optical counterpart (Abbott et al. 2017a; Soares-Santos et al. 2017), provided the first standard siren measurement of $H_0$ with precision 14% (Abbott et al. 2017b). Its constraint ability of cosmological parameters, including the Hubble constant, has been extensively studied (Cai & Yang 2017; Chen et al. 2018). Another promising approach relies on strong gravitational lenses with measured time delays between the multiple images (Refsdal 1964). The program $H_0$ Lenses in COSMOGRAIL’s Wellspring (H0LiCOW) (Suyu et al. 2017) obtained a 2.4% measurement of $H_0$ from a joint analysis of six gravitationally lensed quasars, in agreement with measurements of $H_0$ by the local distance ladder (Wong et al. 2020).

Systematic errors caused by BLR size discrepancy of different emission lines are not included in Figure 8, making it an optimistic estimation of $H_0$ constraints by the SARM project. However, if we use the velocity-resolved RM data by including profiles of optical emission lines, BLR sizes of optical and NIR lines can be different in the model. The systematic errors can be alleviated.

In order to study the constraint ability on $H_0$ by joint analysis of SARM data, we generate a mock sample of type I AGNs for GRAVITY/GRAVITY+. Details on the generation and properties of the sample are presented in Appendix B. For each object in the sample, the profile and DPC are simulated using BLR models. Once we obtain the maximum amplitudes of phase curves $\phi_{\text{max}}$, uncertainties of distance measurements for each object can be estimated by Equation (16). If other cosmological parameters are fixed except $H_0$, the uncertainty of the $H_0$ measurement using the whole sample can be obtained.

Since 200-day light curves with 1-day cadence and no gaps can hardly be achieved in typical RM campaigns, we would assume that the relative uncertainty of the BLR size measurement through the RM campaign is 10% rather than 2% shown in Figure 2(a). (For example, a recent RM campaign conducted by Bentz et al. 2021 for NGC 3783, which is also a suitable target for GRAVITY, obtained a time lag measurement with 10% uncertainty.) We also assume that the uncertainty of the phase measurement for each baseline is 0.1°. Note that highly face-on objects ($i \lesssim 10°$) are discarded owing to their large biases in distance measurements. Objects with maximum phases less than 0°1 are also excluded since they help little to

---

\textsuperscript{7} If orientations of BLRs are randomly distributed and the largest inclination of a type I AGN is 45°, the fraction of objects with $i \lesssim 10°$ is less than 5%.
constrain the Hubble constant. The result is shown in Figure 8. As we can see, by observing 60 targets with $K$-band magnitudes better than 11.5 and maximum phase larger than $0^\circ.1$ through GRAVITY, the uncertainty of $H_0$ will be less than 2%. Currently, GRAVITY can achieve a phase error of $0^\circ.1$ per baseline in observations of bright ($K \sim 10–11$) AGNs (Gravity Collaboration et al. 2018; Dexter et al. 2020; Gravity Collaboration et al. 2020). But the integration time for each object may reach 10 hr owing to the limit of fringe tracking. Fortunately, GRAVITY+, the upgraded version of GRAVITY, aims to achieve on-axis fringe tracking for AGNs as faint as $K \sim 14–15$ in the near future (Dexter et al. 2020). If we assume that GRAVITY+ requires 1–2 hr of integration time to reach a phase error of $0^\circ.1$ per baseline for these bright targets, 60–120 hr of observation time can measure the Hubble constant with better than 2% accuracy.

Finally, we would like to point out that a sample composed of 53 AGNs for future SARM campaigns has been selected by Wang et al. (2020) from the current catalogs (mainly from Veron-Cetty and Verson’s catalog, and 2dF and 6dF etc.). AGNs surveyed by 4MOST in the future will be greatly increased in the southern hemisphere; moreover, powerful GRAVITY+ onboard VLTI will conveniently observe fainter AGNs and quasars at cosmic noon between $z = 2$ and 3, offering opportunities for measuring cosmic expansion history from the local universe to deeper regions.

5. Conclusion

In this paper, we conduct a mock data analysis for cosmology through the SARM campaign. We have tested how the relative uncertainties of distance measurements depend on errors of DPCs, as well as the roles of inclinations and opening angles of BLRs. For BLRs with inclinations $\geq 10^\circ$ and opening angles $\leq 40^\circ$, analyses of SARM data can generate reliable quasar distances even for relatively poor SA measurements with relative error of 40% for GRAVITY-like facilities. If the limiting magnitude of the GRAVITY reaches 11.5 in $K$ band and errors of phase measurements are as low as $0^\circ.1$, the SARM campaign can constrain $H_0$ to an uncertainty of 2% by observing 60 targets.

We thank Jinyi Shangguan for useful comments on the manuscript. We are also grateful to the members of the IHEP AGN group for enlightening discussions. J.-M.W. thanks the support by National Key R&D Program of China through grant 2016YFA0400701; by NSFC through grants NSFC-11991050, NSFC-11991054, NSFC-11833008, and NSFC-11690024; and by grant Nos. QYZDJ-SSW-SLH007 and XDB23010400.

Appendix A
Mock Data

Here we present an example of mock data generation and its fitting for illustration. We use the parameterized BLR model described in Section 2.1, and values of its parameters are taken to be the fiducial values listed in Table 1. The continuum variation is generated by the DRW model with timescale $\tau_d = 60$ days and amplitude $\sigma_d = 0.25$. Then, it is convolved with the velocity-resolved transfer function determined by the BLR model to obtain the light curve of the emission line. Both light curves are sampled with 1-day cadence and last for 200 days. The relative uncertainties of each data point are 0.5% for continuum and 1% for emission line. Blue points with error bars in Figure 9 are mock light curves we generated.

---

For detailed information see [https://www.mpe.mpg.de/ir/gravityplus](https://www.mpe.mpg.de/ir/gravityplus).

[https://www.4most.eu/cms/](https://www.4most.eu/cms/)
The profile and DPCs are simulated using the same BLR model. The emission line used for SA is Brγ. To normalize the profile, we assume that the equivalent width of the line is 40 Å in rest frame. The redshift of the object is 0.01, shifting the central wavelength of the Brγ to 2.18766 μm. There are 40 wavelength bins between 2.14 and 2.24 μm, and the profile and DPC are broadened by a Gaussian with FWHM = 4 nm. The relative uncertainties of the profile and DPC are 0.5% and 20%, respectively. The object is observed for four times with six baselines, and the configuration of baselines is the same as that of Extended Data Figure 1(b) in Gravity Collaboration et al. (2018). Points with error bars in Figure 10 are mock profile and differential phase curves we generated.

Bayesian analysis is applied to fit the combined mock data jointly to the BLR model and obtain the posterior probability distribution of model parameters, as described in Section 2.4. The DNest algorithm is used to sample the posterior distribution (12) (Brewer & Foreman-Mackey 2018; Li 2020), generating a posterior sample of model parameters and associated light curves, profiles, and DPCs. The reconstructed curves with smallest χ² (the best fitting) are thick solid lines shown in Figures 9 and 10, while those randomly drawn from the posterior sample are thin gray lines.

Figure 9. Example of mock light curves we generated and its fitting to the BLR model. The first row is continuum variation, and the second is emission-line variation. Error bars for data points reflect 1σ uncertainties. The thick red lines are the best fitting, while the thin gray lines are fittings using model parameters drawn from the probability distribution of model parameters.
In order to study the constraint ability on $H_0$ by joint analysis of SARM data from a large sample of type I AGNs, we need to generate a mock sample using realistic luminosity functions of quasars. We first divide the redshift (up to $z = 1$) and bolometric luminosity (from $10^{41.6}$ to $10^{51.5}$ erg s$^{-1}$) into short bins. Then, we estimate the number of quasars in each bin using the redshift-dependent bolometric quasar luminosity function$^{10}$ in Hopkins et al. (2007). Assuming that the inclinations of AGNs are isotropic and the broad emission line can be observed when $i < 45^\circ$, only 29.3% of them are type I. The distribution of AGNs on the celestial sphere is also isotropic. When the difference between the decl. of the target and latitude of VLTI is smaller than 45$^\circ$, the object can be observed by GRAVITY. Thus, only 64.3% of the remaining type I AGNs are selected. For each object in our mock list, we calculate the $L_{5100}$ luminosity and $K$-band magnitude using the bolometric-luminosity-dependent spectral energy distribution provided by Hopkins et al. (2007). Objects whose redshifts are smaller than 0.01 are discarded since their peculiar motions may be comparable to the Hubble flow. We also remove those with bolometric luminosities larger than $10^{48}$ erg s$^{-1}$ because their variations are too slow for a typical RM campaign and those with $K$-band magnitude larger than 13 because they are too faint. The distribution of $K$-band magnitudes of our mock sample is shown in Figure 11(a).

Before simulating line profiles and DPCs for objects in our mock sample, we need to know values of model parameters listed in Table 1 of each BLR. The standard $R - L$ relation in

---

$^{10}$ A quasar luminosity function calculator script is available at http://www.tapir.caltech.edu/~phopkins/Site/qlf.html.
Benz et al. (2013) is applied to estimate the size of the BLR, while luminosity-dependent Eddington ratios of type I AGNs measured by Lusso et al. (2012) are used to infer the black hole mass. Note that dispersions in those relations are also included. Other model parameters are assigned randomly according to the last column of Table 1.

The spectral coverage of GRAVITY is 1.98–2.40 μm. In hydrogen emission lines, only Brγ (2.166 μm), Paα (1.875 μm), and Paβ (1.282 μm) are able to be observed by GRAVITY for objects with z < 1. Landt et al. (2008) studied NIR broad emission line properties of 23 well-known type I AGNs. We assume that distributions of equivalent widths of Brγ, Paα, and Paβ in our mock samples are Gaussian with means and standard deviations determined by samples in Landt et al. (2008). For each object we calculate line profiles in the observer’s frame of those three emission lines. We adopt the emission line with largest equivalent width within the spectral coverage of GRAVITY as the target line. Objects without appropriate emission lines are discarded. The distribution of redshifts and K-band magnitude of objects with emission lines within the spectral coverage is shown in Figure 11(b). When z ≤ 0.1, Brγ can be observed by GRAVITY, and K-band magnitudes can be as bright as 9. When 0.1 < z < 0.3, Paα is chosen, and the K-band magnitude of the brightest one is about 10.5. When 0.6 < z < 0.8, Paβ is chosen, and K-band magnitudes are around 14. Note that objects are removed if FWHMs of their profiles are less than 1500 km s⁻¹ owing to the limited spectral resolution of GRAVITY (~500 when observing AGNs) or if inclinations are less than 10° because of high uncertainties in distance measurements of face-on objects.

Finally, to determine the projected lengths of baselines and calculate the DPC for each object, we assume that all of them are observed when reaching their zeniths for simplicity. The distribution of line intensities ℓmax and phase signal amplitudes ϕmax is shown in Figure 11(c). For Brγ lines, ℓmax is mostly less than 0.1 and ϕmax is less than 0°2, which is compatible with the latest observation of IRAS 09149–6206 by GRAVITY (Gravity Collaboration et al. 2020). For Paα lines, ℓmax is about 0.2–0.4 and ϕmax is about 0°2–0°6, which is compatible with the recent observation of 3C 273 by GRAVITY (Gravity Collaboration et al. 2018).

**Appendix C**

**Diffusive Nested Sampling**

Nested sampling was proposed by Skilling (2004) to evaluate the evidence Z of a model M:

\[ Z = P(\mathcal{D}|M) = \int P(\mathcal{D}|\Theta, M) P(\Theta|M) d\Theta, \]

where \( \mathcal{D} \) and \( \Theta \) represent the data set and model parameters, respectively, \( P(\Theta|M) \) is the prior probability distribution of parameters \( \Theta \) in model \( M \), and \( P(\mathcal{D}|\Theta, M) \) is the likelihood function.

Nested sampling starts with \( n \) points \( \Theta_i \) sampled from prior \( P(\Theta|M) \). The likelihood of each point \( L(\Theta_i) \equiv P(\mathcal{D}|\Theta_i, M) \) is evaluated. The minimum of likelihoods \( L_1 \) and the corresponding particle is saved. Then, this particle will be replaced by a new one drawn from the prior probability distribution but under a constraint \( L(\Theta_i) > L_1 \) via the Markov Chain Monte Carlo method. Again, the minimum of likelihoods is saved and iteration continues. As nested likelihood levels \( L_1 < L_2 < \cdots \) are created, particles move progressively toward higher likelihoods. The posterior distribution of \( \Theta \) can be obtained as a by-product by recording positions and likelihoods of particles in the process.

There are several variants of nested sampling, such as CosmoNest (Parkinson et al. 2011) and MultiNest (Feroz et al. 2009). Diffusive nested sampling (Brewer et al. 2011) is the one that makes improvements to the original method to overcome its drawbacks when sampling multimodal or highly correlated distributions. In the process of creating levels, particles at high levels can diffuse to lower levels. If a distribution has isolated islands with high likelihoods, particles in classic nested sampling may be stuck in one island and fail to explore other islands. However, particles in diffusive nested sampling can diffuse to lower levels where the distribution is sufficiently broad without isolated islands so that particles can easily move over the whole parameter space.

Diffusive nested sampling performs well to overcome high dimensions, multimodal distributions, and phase changes. It has been successfully applied to fit the BLR model to SARM data to measure the BLR size, black hole mass, and distance to the quasar (e.g., Pancoast et al. 2014b; Li et al. 2018; Wang et al. 2020).
