




Constraining the Spatial Curvature of the Local Universe with Deep Learning

Liang Liu^{1,2}, Li-Juan Hu¹, Li Tang^{1,2} , and Ying Wu^{1,2}

¹Department of Math and Physics, Mianyang Teachers' College, Mianyang 621000, China; tang@mtc.edu.cn

²Research Center of Computational Physics, Mianyang Teachers' College, Mianyang 621000, China

Received 2023 July 18; revised 2023 August 21; accepted 2023 August 27; published 2023 October 31

Abstract

We use the distance sum rule method to constrain the spatial curvature of the Universe with a large sample of 161 strong gravitational lensing systems, whose distances are calibrated from the Pantheon compilation of type Ia supernovae using deep learning. To investigate the possible influence of mass model of the lens galaxy on constraining the curvature parameter Ω_k , we consider three different lens models. Results show that a flat Universe is supported in the singular isothermal sphere (SIS) model with the parameter $\Omega_k = 0.049_{-0.125}^{+0.147}$. While in the power-law (PL) model, a closed Universe is preferred at the $\sim 3\sigma$ confidence level, with the parameter $\Omega_k = -0.245_{-0.071}^{+0.075}$. In the extended PL model, the 95% confidence level upper limit of Ω_k is < 0.011 . As for the parameters of the lens models, constraints on the three models indicate that the mass profile of the lens galaxy could not be simply described by the standard SIS model.

Key words: (cosmology:) cosmological parameters – (cosmology:) distance scale – (stars:) supernovae: general

1. Introduction

The standard cosmological model (Λ CDM model), regarded as a cornerstone solution derived from the homogeneous and isotropic Friedmann–Robertson–Walker (FRW) metric, presents a comprehensive framework in cosmology. This model postulates the existence of radiation, ordinary baryonic matter, non-luminous dark matter, and enigmatic dark energy as constituents of the Universe. Its validity and credibility find strong support from a plethora of cosmological observations (Ade et al. 2016; Aghanim et al. 2020). Especially, the most recent findings derived from the conclusive full-mission analysis of the cosmic microwave background (CMB) anisotropies by the Planck mission exhibit remarkable agreement with the prevailing spatially flat 6-parameter Λ CDM cosmological model. These results not only validate the standard framework but also provide stringent constraints on the cosmological parameters with an exceptional level of precision (Aghanim et al. 2020). However, some challenges come following the success of the Λ CDM model. Recently, the “ H_0 tension problem,” i.e., the measured value of Hubble constant H_0 from the local type Ia supernovae (SNe Ia) observation is inconsistent with the result from the Planck observation of CMB, has attracted great attention (Freedman 2017; Riess et al. 2016, 2019; Di Valentino et al. 2021). This discrepancy is possibly caused by either unknown systematic uncertainties or new physics beyond the standard Λ CDM cosmology.

In the context of the FRW metric, the spatial curvature parameter plays a pivotal role in elucidating the geometric nature of the Universe. By incorporating measurements from

the CMB and baryon acoustic oscillation (BAO), it has been established that the Universe can be reasonably modeled as spatially flat. This conclusion is supported by the constraint on the curvature parameter, $\Omega_k = 0.001 \pm 0.002$ (Aghanim et al. 2020). Considering the intricate degeneracy between the curvature parameter and the equation of state of dark energy, the assumption of a flat Universe is commonly adopted during the analysis of dark energy properties. Small deviation of spatial curvature from zero would generate enormous effects on the reconstruction of dark energy and on the evolution of the Universe (Ichikawa & Takahashi 2006; Clarkson et al. 2007; Gong & Wang 2007; Virey et al. 2008). Although the Planck CMB data constrain the spatial curvature at a very high precision, the predicted evolution not only depends on a certain cosmological model (the Λ CDM model), but also on the evolution of the early Universe. Recently, the reanalysis of Planck data showed that a closed Universe is favored against a flat Universe (Di Valentino et al. 2019, 2021). The presence of the so-called “ H_0 tension problem” suggests the possibility of deviations between the actual state of the Universe and the predictions of the standard Λ CDM model. Specifically, it implies that the cosmological parameters derived from CMB measurements may differ from those obtained through local data. Consequently, it becomes crucial to ascertain the spatial curvature of the local Universe in a manner that is independent of specific theoretical models.

The measurement of spatial curvature is generally the by-product of the validity test of the FRW metric. A model-independent approach was introduced by Clarkson et al. (2007, 2008) to scrutinize the validity of the FRW metric.

This method involves a comparative analysis of the cosmic expansion rate and cosmological distance, and has since been widely employed to examine the FRW metric and impose constraints on the spatial curvature (Mortsell & Jonsson 2011; Sapone et al. 2014; Cai et al. 2016). Bernstein (2006) put forth an alternative model-independent geometric approach to constrain spatial curvature. This methodology revolves around the fundamental sum rule of distances along null geodesics within the FRW metric framework. Räsänen et al. (2015) employed the distance sum rule (DSR) to evaluate the accuracy of the FRW metric. By combining data from SNe Ia and strong gravitational lensing (SGL), they examined the validity of the FRW metric. Their analysis confirmed the overall validity of the FRW metric, although the obtained constraint on the spatial curvature parameter was relatively weak or loosely constrained. Considering the interdependence between the curvature parameter and the parameters of the lensing model, Xia et al. (2017) adopted more intricate lensing models in their analysis and attained constraints on the spatial curvature by leveraging a substantial data set comprising 118 SGL systems (Cao et al. 2015, 2016). Following this line, there were a series of works devoted to constraining the spatial curvature with updated observational data (Li et al. 2018; Qi et al. 2019; Liu et al. 2020; Cao et al. 2022). It should be noted that the constraints on the spatial curvature derived from the aforementioned studies suffer from limitations arising from the relatively small size of the available SGL sample, as well as uncertainties stemming from unknown systematic effects. Moreover, the methods to calibrate the distances of lenses and sources within SGL systems rely on a polynomial approximation that is assumed to fit the SNe Ia sample. Further research and advancements in data acquisition and analysis techniques are necessary to address these limitations and improve the precision of spatial curvature measurements in cosmology.

To alleviate the above shortcomings, Wang et al. (2020) made significant advancements in constraining the spatial curvature. They employed the DSR method and combined data from the Pantheon SN Ia compilation with a data set comprising 161 SGL systems. Notably, they avoided assumptions regarding the parametric form of the distance-redshift relation of SNe Ia. Instead, they employed a Gaussian Process (GP) method to reconstruct the dimensionless comoving distance based on the Pantheon compilation. Without the prior of H_0 , the constraints on spatial curvature are $\Omega_k = 0.57^{+0.20}_{-0.28}$ in the singular isothermal sphere (SIS) model, $\Omega_k = -0.246^{+0.078}_{-0.100}$ in the power-law (PL) model, and $\Omega_k = 0.25^{+0.16}_{-0.23}$ in the extended power-law (EPL) model. Previous studies have indicated that a larger data set is beneficial to achieve a tighter constraint on Ω_k (Xia et al. 2017; Li et al. 2018; Qi et al. 2019). However, the GP method is unable to extrapolate the curve beyond the available data region, and its accuracy diminishes significantly in regions where data points are sparse.

Consequently, in their analysis, only the SGL systems with redshifts lower than the maximum redshift of the SNe Ia data could be utilized. This constraint resulted in a reduction in the number of available SGL systems from the initial 161–135. Therefore, although the constraint of Wang et al. (2020) is tighter than previous works, the method to reconstruct the distance-redshift relation can be further improved so that all SGL systems can be used to constrain the spatial curvature.

In this paper, we will maintain the advantages of Wang et al. (2020), i.e., the large data set and model-independence, and employ a deep learning method to reconstruct the distance-redshift relation based on the Pantheon data set, extending it up to the maximum redshift of the available SGL systems. Deep learning is a realm dedicated to the research of various Artificial Neural Networks (ANNs), which are composed of layers of neurons modeled after the biological neurons in a human brain. Hence, deep learning is fantastic to deal with large and highly complex tasks, such as classification, clustering, generation and so on. Deep learning has emerged as a powerful tool in various cosmological research areas, demonstrating its effectiveness in tasks such as predicting galaxy morphology (Dieleman et al. 2015), constraining dark energy (Escamilla-Rivera et al. 2019), and calibrating gamma-ray bursts (GRBs) (Luongo & Muccino 2021; Tang et al. 2021b). In our recent work (Tang et al. 2021a), we applied deep learning techniques to reconstruct the distance-redshift relation of SNe Ia without making any assumptions about the cosmological model or the parametric form of the relation. Furthermore, we utilized this reconstructed relation to investigate potential redshift dependencies in the luminosity corrections of GRBs. Unlike the GP method, which is constrained to reconstruct the curve within the data region, deep learning has the capacity to extend the reconstruction far beyond the available data region. Thus all of the SGL systems can be used and the constraint on the spatial curvature would be tighter.

The structure of the remaining sections of this paper is as follows: Section 2 provides an overview of the DSR method and the lens mass models utilized in constraining the spatial curvature. Section 3 outlines the observational data sets employed in the analysis and details of the procedure for reconstructing the distance-redshift relation using deep learning techniques. The obtained results are presented in Section 4. Lastly, Section 5 contains the discussion and summary.

2. Methodology

In the context of a homogeneous and isotropic Universe, the spacetime can be described by the FRW metric, given by

$$ds^2 = -c^2 dt^2 + \frac{a(t)^2}{1 - Kr^2} dr^2 + a(t)^2 r^2 d\Omega^2, \quad (1)$$

where c represents the speed of light, and K is a constant that denotes the spatial curvature of the Universe. Specifically,

when $K < 0$, $K = 0$, and $K > 0$, it corresponds to an open, flat, and closed Universe, respectively. The scale factor $a(t)$ represents the expansion of the Universe with respect to cosmic time, and its derivative $\dot{a} \equiv \frac{da}{dt}$ defines the Hubble parameter $H \equiv \frac{\dot{a}}{a}$. To quantify the spatial separation between a source at redshift z_s observed from redshift z_l , the dimensionless comoving distance is expressed as

$$d(z_l, z_s) = \frac{1}{\sqrt{|\Omega_k|}} S_k \left(\sqrt{|\Omega_k|} \int^{z_s} z_l \frac{dz'}{E(z')} \right), \quad (2)$$

where $\Omega_k \equiv -\frac{Kc^2}{H_0^2 a_0^2}$ represents the normalized curvature parameter. The reduced Hubble parameter is denoted as $E(z) \equiv \frac{H(z)}{H_0}$, where H_0 represents the present-day value of the Hubble parameter. The function S_k is defined as follows

$$S_k(x) = \begin{cases} \sinh(x), & (\Omega_k > 0), \\ x, & (\Omega_k = 0), \\ \sin(x), & (\Omega_k < 0). \end{cases} \quad (3)$$

For simplicity, we introduce the notation $d(z) \equiv d(0, z)$, $d_l \equiv d(0, z_l)$, $d_s \equiv d(0, z_s)$, and $d_{ls} \equiv d(z_l, z_s)$. Under the assumption that cosmic time t and redshift z have a one-to-one correspondence, and with the condition that the derivative of $d(z)$ with respect to z satisfies $d'(z) > 0$, the three-dimensional distances (d_l , d_s , and d_{ls}) are connected through the DSR relation (Räsänen et al. 2015)

$$\frac{d_{ls}}{d_s} = \sqrt{1 + \Omega_k d_l^2} - \frac{d_l}{d_s} \sqrt{1 + \Omega_k d_s^2}. \quad (4)$$

If the Universe is accurately described by the FRW metric, the curvature parameter Ω_k should be a constant. Therefore, if the validity of the FRW metric is confirmed, the DSR relation provides a means to constrain the value of Ω_k . By analyzing the relation between the dimensionless distances, we can obtain valuable insights into the spatial curvature of the Universe.

The dimensionless comoving distances d_l and d_s can be obtained through the analysis of SN Ia data. On the other hand, the distance ratio d_{ls}/d_s is determined using the data from SGL observations. In terms of the Einstein radius and the velocity dispersion associated with the lens mass profile, the expression for the distance ratio can be formulated. For certain gravitational lens systems, the mass distribution of the lens has been observed to closely approximate an isothermal profile (Cohn et al. 2001; Munoz et al. 2001; Rusin et al. 2002; Treu & Koopmans 2002; Rusin & Kochanek 2005). Consequently, the SIS model has emerged as a prevalent and straightforward choice for describing the lens mass profile. This model effectively emulates the flat rotation curves characteristic of galaxies, featuring a density inversely proportional to the square of the galaxy's radius. Additionally, the structure of galaxies has been extensively explored through N -body simulations (Navarro et al. 1996; Moore et al. 1998). Navarro et al. (Navarro et al. 1996)

discovered that the halo profile of a galaxy exhibits an approximate isothermal behavior across a wide range of radii. However, it deviates from the r^{-2} PL near the central region, transitioning to a steeper profile than r^{-2} as the distance approaches the virial radius. To adequately address the specific features of the halo profile, a density function incorporating a PL dependence on radius has been introduced to describe the lens profile. This density function resembles the generalized Navarro–Frenk–White (NFW) profile (Navarro et al. 1996). Among these descriptions, the PL model stands out, characterized by a variable PL index denoted as γ . Remarkably, when γ is set to 2, the profile aligns with a configuration akin to an SIS. It is noteworthy that these profile descriptions do not inherently distinguish between luminosity density and total mass density. When accounting for the presence of dark matter, the luminosity density may diverge from the overall galaxy profile. This prompts the introduction of models like the EPL model, which accommodates the complexities stemming from both luminosity density and dark matter distribution within the lens mass. Hence, we consider three distinct lens models: the SIS model, the PL model, and the EPL model, to comprehensively investigate the influence of various lens galaxy mass profiles on the constraints imposed on the spatial curvature.

Within the SIS model, the mass density distribution of the lens galaxy follows a scaling relation of $\rho \propto r^{-2}$. This leads to an expression for the distance ratio as follows (Mollerach & Roulet 2002)

$$\frac{d_{ls}}{d_s} = \frac{c^2 \theta_E}{4\pi \sigma_{\text{SIS}}^2}, \quad (5)$$

where θ_E represents the Einstein radius, and σ_{SIS} is the velocity dispersion associated with the lens mass profile. It is worth noting that the equivalence between the observed stellar velocity dispersion σ_0 and σ_{SIS} within the context of the SIS model is not an absolute requirement (Khedekar & Chakraborti 2011). There may be potential deviations between the observed stellar velocity dispersion and the characteristic velocity dispersion associated with the SIS model. Consequently, to account for such deviation, a phenomenological parameter f is introduced, yielding $\sigma_{\text{SIS}} = f\sigma_0$ (Kochanek 1992; Ofek et al. 2003; Cao et al. 2012; Li et al. 2019). Notably, the free parameter f is anticipated to fall within the range of $0.8 < f^2 < 1.2$ (Ofek et al. 2003). In practice, the velocity dispersion is typically measured within the aperture radius θ_{ap} in actual SGL data. To convert the measurement to σ_0 , an aperture correction formula (Jorgensen et al. 1995) can be employed, given by the equation

$$\sigma_0 = \sigma_{\text{ap}} \left(\frac{\theta_{\text{eff}}}{2\theta_{\text{ap}}} \right)^\eta, \quad (6)$$

where σ_{ap} represents the luminosity weighted average of the line-of-sight velocity dispersion within the aperture radius, θ_{eff}

corresponds to the effective angular radius, and η is the correction factor fixed to -0.066 (Cappellari et al. 2006; Chen et al. 2019). It is important to consider that the uncertainty associated with σ_{ap} propagates to σ_0 , subsequently impacting σ_{SIS} . Additionally, the uncertainty in the distance ratio d_{ls}/d_s is derived from the uncertainties in θ_E and σ_{SIS} . In this work, we adopt a fractional uncertainty of 5% for θ_E (Liao et al. 2016).

Within the framework of the PL spherical model, the mass density distribution of the lensing galaxy is characterized by a spherically symmetric PL behavior, expressed as $\rho \propto r^{-\gamma}$, where γ represents the PL index. The distance ratio in the PL model can be described as (Koopmans et al. 2006)

$$\frac{d_{\text{ls}}}{d_s} = \frac{c^2 \theta_E}{4\pi \sigma_{\text{ap}}^2} \left(\frac{\theta_{\text{ap}}}{\theta_E} \right)^{2-\gamma} f^{-1}(\gamma), \quad (7)$$

where

$$f(\gamma) = -\frac{1}{\sqrt{\pi}} \frac{(5-2\gamma)(1-\gamma)}{3-\gamma} \frac{\Gamma(\gamma-1)}{\Gamma(\gamma-3/2)} \left[\frac{\Gamma(\gamma/2-1/2)}{\Gamma(\gamma/2)} \right]^2. \quad (8)$$

It is worth noting that when γ takes the value of 2, the PL model reduces to the standard SIS model. To account for the potential redshift evolution of the mass density profile, we introduce a parameterization for γ , expressed as $\gamma(z) = \gamma_0 + \gamma_1 z$, where γ_0 and γ_1 represent two independent free parameters.

Within the EPL model, the luminosity density profile $\nu(r)$ can differ from the total mass density profile $\rho(r)$, accounting for the presence of a dark matter halo. We adopt the following functional forms for the PL mass density profile and the luminosity density of stars,

$$\rho(r) = \rho_0 \left(\frac{r}{r_0} \right)^{-\alpha}, \quad \nu(r) = \nu_0 \left(\frac{r}{r_0} \right)^{-\delta}, \quad (9)$$

where α and δ correspond to the PL index parameters, r_0 represents the characteristic length scale, and ρ_0 and ν_0 are normalization constants. The distance ratio in the EPL model is expressed as (Birrer et al. 2019; Lee 2021)

$$\frac{d_{\text{ls}}}{d_s} = \frac{c^2 \theta_E}{2\sigma_0^2 \sqrt{\pi}} \frac{3-\delta}{(\xi-2\beta)(3-\xi)} \left(\frac{\theta_{\text{eff}}}{\theta_E} \right)^{2-\alpha} \times \left[\frac{\lambda(\xi) - \beta\lambda(\xi+2)}{\lambda(\alpha)\lambda(\delta)} \right], \quad (10)$$

where $\xi = \alpha + \delta - 2$, $\lambda(x) = \Gamma\left(\frac{x-1}{2}\right)/\Gamma\left(\frac{x}{2}\right)$, and β represents an anisotropy parameter that characterizes the anisotropic distribution of the three-dimensional velocity dispersion. In accordance with Wang et al. (2020), we consider β as a nuisance parameter and marginalize over it with a Gaussian prior of $\beta = 0.18 \pm 0.13$. Simultaneously, we treat α and δ as

free parameters. It is worth noting that when $\alpha = \delta = 2$ and $\beta = 0$, the EPL model reduces to the standard SIS model.

3. Observational Data and Deep Learning

The distance ratios d_{ls}/d_s are obtained from the observations of SGL systems. In a recent study, Chen et al. (2019) compiled a new SGL sample by combining data from various galaxy surveys, including the Lenses Structure and Dynamics (LSD) survey (Treu & Koopmans 2004), the Sloan Lens ACS (SLACS) survey (Bolton et al. 2006), the CFHT Strong Lensing Legacy Survey (Cabanac et al. 2008), and the BOSS Emission-Line Lens Survey (Brownstein et al. 2012). This compiled sample consists of 161 galaxy-scale SGL systems, covering a redshift range of $z_l \in [0.0624, 1.004]$ for the lens galaxies and $z_s \in [0.197, 3.595]$ for the source galaxies. In Figure 1, we illustrate the distribution of the SGL sample, derived from diverse survey sources, as depicted in the z_l - z_s plane. Additionally, we provide the redshift distribution of the lens objects, with a predominant concentration of lenses residing at an approximate redshift of $z_l \sim 0.2$.

The dimensionless comoving distances d_l and d_s are derived from the luminosity distance D_L of SNe Ia using the relation

$$d(z) = \frac{H_0 D_L(z)}{c(1+z)}. \quad (11)$$

The luminosity distance D_L can be obtained from the light curve of SNe Ia. Considering a specific redshift z , the distance modulus of SNe Ia can be expressed as

$$\mu = 5 \log_{10} \frac{D_L(z)}{\text{Mpc}} + 25 = m_{B,\text{corr}} - M_B, \quad (12)$$

where M_B represents the absolute magnitude and $m_{B,\text{corr}}$ denotes the corrected apparent magnitude observed in the B -band, reported in the largest and most recent Pantheon data set (Scolnic et al. 2018). The redshift range of the SN Ia sample used in our work, i.e., the Pantheon data set, is $z \in [0.01, 2.30]$.

To obtain the comoving distances d at the redshifts of the lens and source for all the SGL systems, it is necessary to reconstruct a continuous curve of the distance-redshift relation $d(z)$ based on the Pantheon sample. Previous work by Wang et al. (2020) employed the GP method to reconstruct a smooth curve of $d(z)$ from SN Ia data. However, the reconstructed uncertainty of the GP method tends to be large in regions where the data points are sparse, and it becomes even more challenging to estimate distances beyond the observed redshift range. Consequently, SGL systems with source redshifts larger than 2.3 could not be utilized in their analysis.

In this paper, we adopt a deep learning method to reconstruct the distance-redshift curve without any specific assumption about its parametric form. This approach allows us to reconstruct the distance-redshift relation using a wide range of redshifts, covering the entire redshift range of the SGL

sample. Specifically, we can extend the reconstruction up to a redshift of $z=4$, thus ensuring that we encompass the full redshift range of the SGL systems under consideration. This utilization of deep learning enables us to overcome the limitations associated with the sparse data points and extrapolate the distance-redshift relation to regions beyond the direct observational range.

Deep learning has emerged as a powerful methodology for analyzing complex and intricate data sets. One common approach involves the utilization of ANNs as underlying models, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Bayesian Neural Networks (BNNs), among others. These neural networks typically consist of multiple layers of interconnected processing units, where each layer receives information from the previous layer, transforms it and then propagates it to the subsequent layer. Through training, these networks aim to learn and represent the underlying patterns and structures within the data. In the context of our research, we employ RNNs as a key component of our deep learning approach. RNNs are well-suited for handling sequential data and making predictions based on learned data representations. By feeding the Pantheon data set into the RNN, we can effectively capture the relationship between the distance modulus μ and the redshift z . This enables us to predict distances at arbitrary redshifts, even beyond the range covered by the observational data. However, RNNs alone are insufficient for providing uncertainty estimates for these predictions. To address this limitation, we incorporate BNNs into our network architecture. BNNs serve as a complementary component to the RNNs and allow us to calculate the uncertainty associated with the distance predictions. Our previous work (Tang et al. 2021a) had incorporated both RNNs and BNNs to model the distance

modulus-redshift relationship based on the Pantheon data set, while this current research emphasizes the reconstruction of the distance curve $d(z)$ using the deep learning approach.

The architecture of our network is illustrated in Figure 2. The central component is the RNN, which consists of three layers: an input layer that receives the redshift z as the feature, a hidden layer that processes information from the previous layer and passes it to the next layer, and an output layer that generates the target output, which in this case is the comoving distance d . The RNN is designed to capture the temporal dependencies and patterns in the input data. To overcome the challenges associated with training RNNs on long sequential data and to address the issue of information retention over long periods, we employ Long Short-Term Memory (LSTM) cells as the basic units of our network. LSTM cells enhance RNNs by incorporating explicit memory mechanisms, allowing the network to selectively store, discard, and retrieve information. The input and hidden layers of our network consist of 100 LSTM cells each.

In the training process, the RNN is fed with the Pantheon data to learn and represent the relationship between the comoving distance d and the redshift z . This is achieved by minimizing a loss function that quantifies the discrepancy between the network's predictions and the observed distances. In this work, we utilize the mean-squared-error (MSE) function as the loss function, and we employ the Adam optimizer to find the minimum of this function. To enhance the network's performance, we introduce a non-linear activation function denoted as A_f . In our previous research on reconstructing the distance modulus $\mu(z)$, we found that the hyperbolic tangent (tanh) function outperformed other activation functions such as ReLU, ELU, and SELU. However, since we are now reconstructing the comoving distance $d(z)$ instead of the

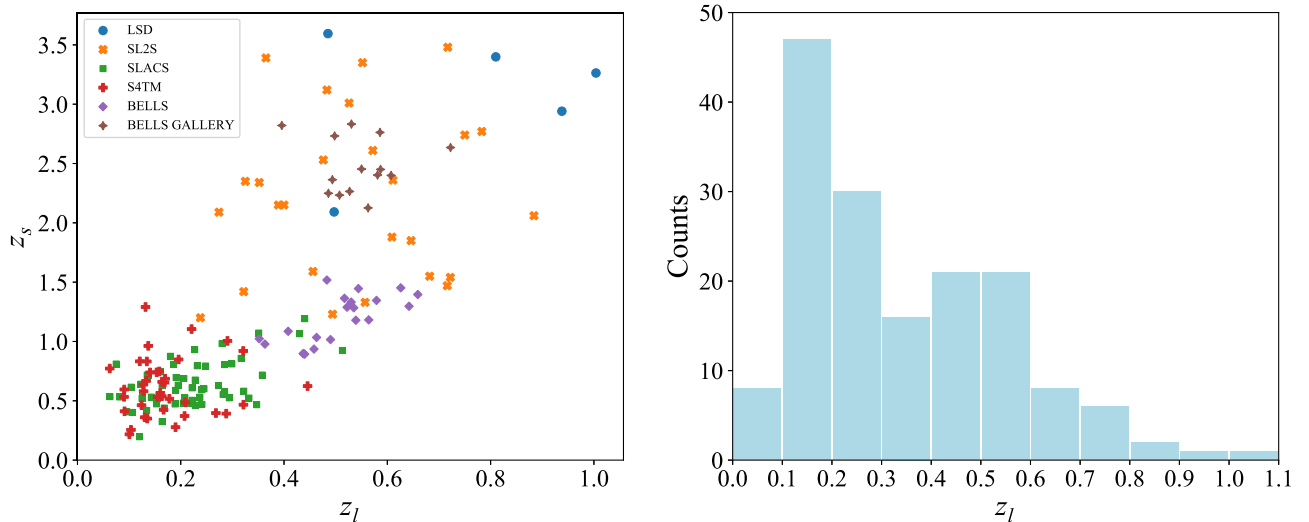


Figure 1. Left: The distribution of the SGL sample obtained from various surveys in the z_l - z_s plane. Right: the redshift distribution of lenses.

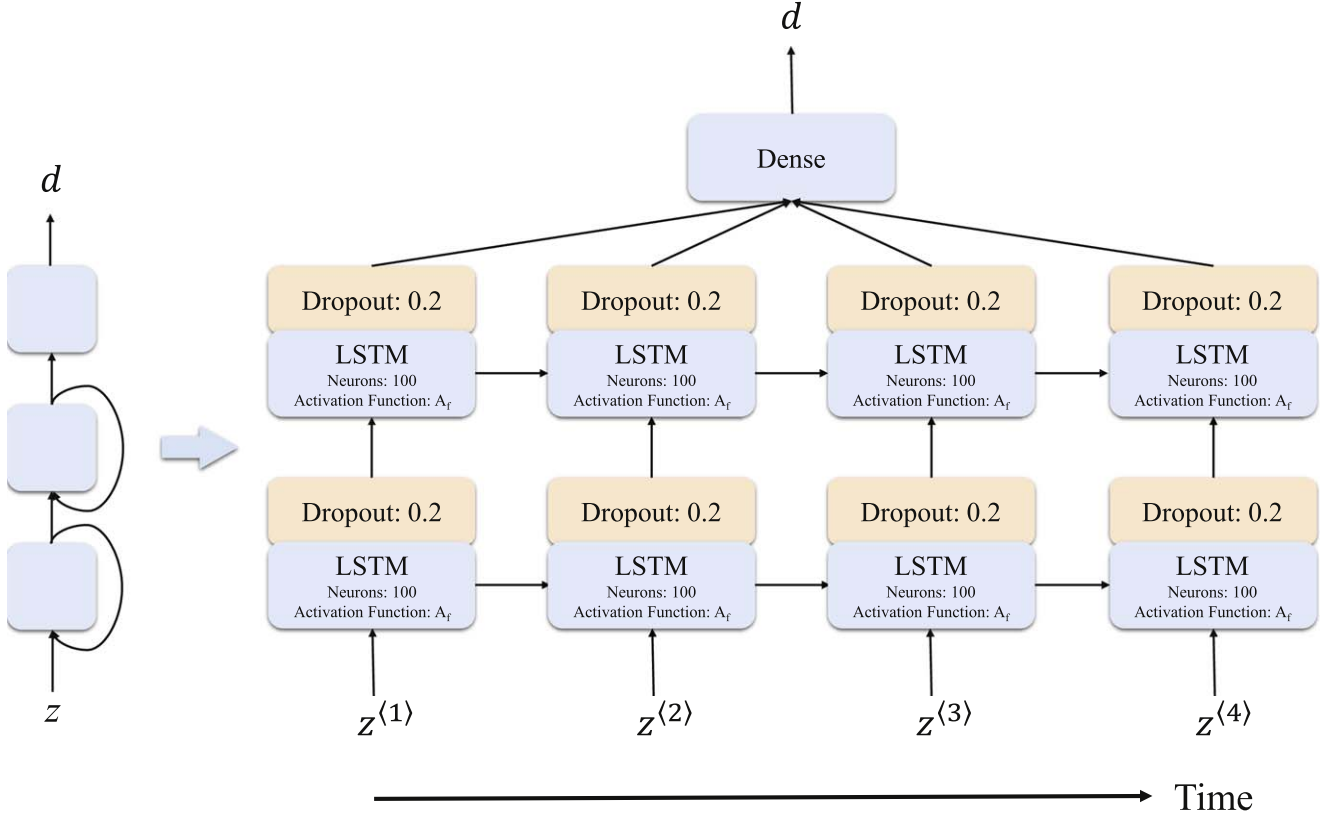


Figure 2. Left: The network architecture comprising a single hidden layer is illustrated. Right: The network unfolded up to time step $t=4$, denoted as $Z^{(i)}$ representing the i th time step. In our network, both the input layer and hidden layer are composed of LSTM cells, housing 100 neurons each. The output layer is a fully connected (dense) layer. To mitigate overfitting, the dropout technique is implemented between each LSTM cell and its subsequent layer.

distance modulus, we compare the performance of all four activation functions (tanh, ReLU, ELU, and SELU) to determine the most suitable choice. By setting the time step to $t=4$ and utilizing the LSTM-based RNN architecture with appropriate activation functions, our network aims to learn the complex relationship between the redshift z and the comoving distance d . Through training and optimization, we obtain a model that can predict distances at arbitrary redshifts, including those beyond the range covered by the Pantheon data set.

In the context of BNN, it is worth noting that designing a traditional BNN is a challenging task due to its inherent complexity. Fortunately, Gal & Ghahramani (2016a, 2016b, 2016c) had demonstrated that dropout, commonly used in deep neural networks as a regularization technique to address the issue of overfitting, can be viewed as an approximation to Bayesian inference in deep GPs. This means that a network incorporating dropout can be considered mathematically equivalent to a Bayesian model. In this study, we incorporate the dropout technique within the RNN to emulate the characteristics of a BNN. By executing the trained network multiple times, we can generate multiple predictions for the

comoving distance at different redshifts. This process allows us to obtain a range of possible predictions and, consequently, estimate the confidence region associated with these predictions. This approach effectively mimics the behavior of a BNN, where the network models the posterior distribution over the parameters. In our research, we employ a dropout rate of 0.2 between the LSTM layer and its subsequent layer.

To begin the reconstruction of the comoving distance $d(z)$, we first normalize the comoving distance data obtained from the Pantheon compilation according to Equations (11) and (12) with the chosen parameters $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $M_B = -19.36$ (Scolnic et al. 2014). Next, we sort the normalized data points (z_i, d_i) in ascending order of redshift z_i and reorganize them into four sequences. In each sequence, the redshifts and the corresponding normalized distances are used as input and output vectors, respectively, for training the network. Subsequently, we train the network constructed as described above using TensorFlow³ for a total of 1000 iterations. The well-trained network is saved for later use. In the final step, we execute the trained network 1000 times to

³ <https://www.tensorflow.org>

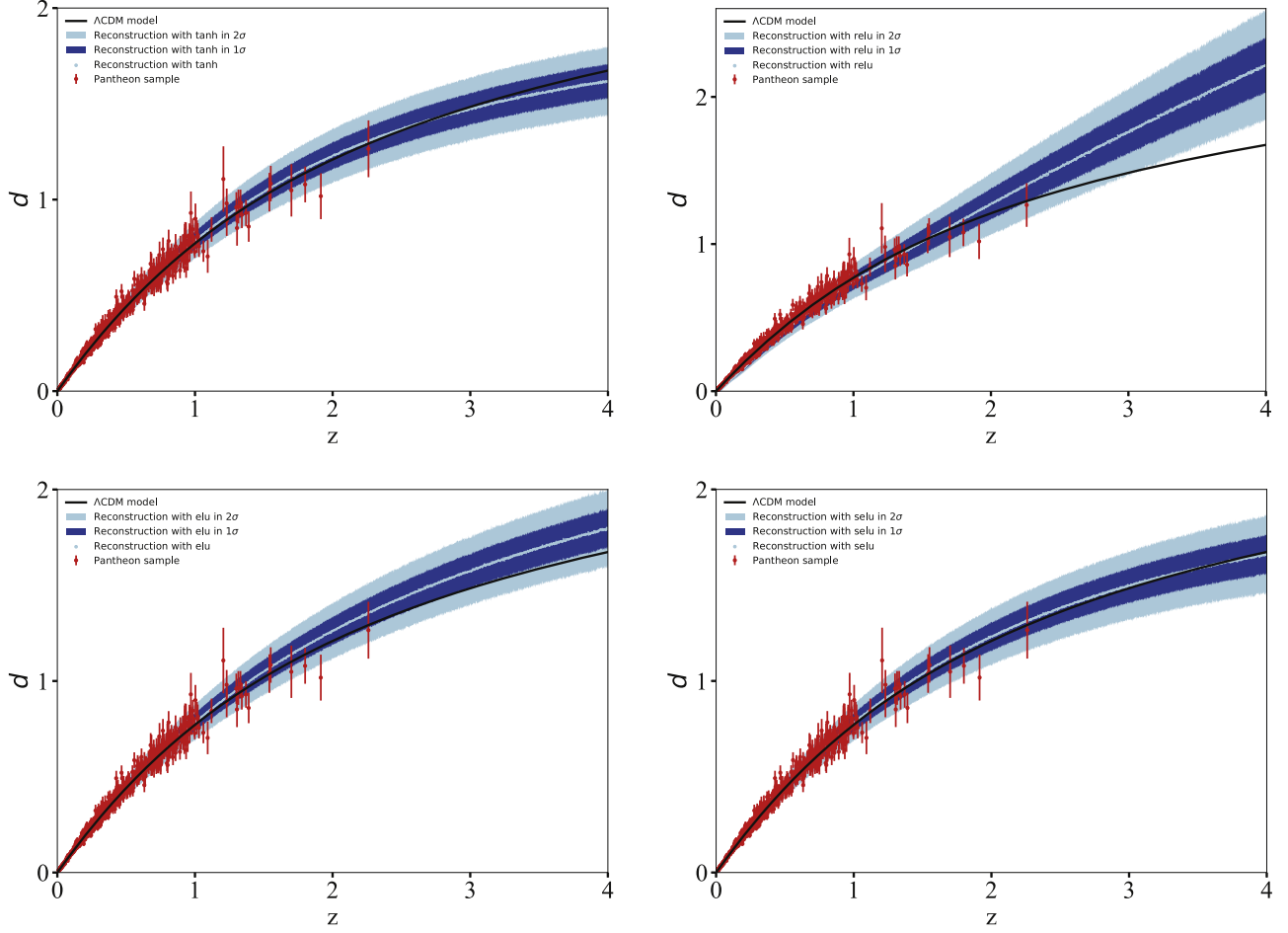


Figure 3. The reconstructions of $d(z)$ from the Pantheon data are presented, employing four distinct activation functions. Top-left: tanh; top-right: ReLU; bottom-left: ELU; bottom-right: SELU.

predict the distance d over the redshift range $z \in [0, 4]$. The distribution of the predicted distances is obtained as a Gaussian distribution.

The results of the distance reconstruction using the four activation functions (tanh, ReLU, ELU, and SELU) are plotted in Figure 3. For comparison, we also include the best-fitting curve of the Λ CDM model (represented by the black line). It is worth noting that while the uncertainty in the reconstructed curve using deep learning may be slightly larger than that obtained using the GP method within the data region, the advantage of deep learning lies in its ability to reconstruct the curve beyond the data region. This enables us to leverage the full sample of SGL systems.

As depicted in the results, it is observed that only the reconstructions using the tanh and SELU activation functions are consistent with the flat Λ CDM model within the 1σ confidence level. Considering that most of the current cosmological probes favor the Λ CDM model, and the reconstructed curves using ReLU and ELU functions deviate

from the Λ CDM model too much at high redshift, these two activation functions are excluded in the following calculation. Therefore, we will derive the dimensionless comoving distances of the SGL systems from the reconstruction with the tanh and SELU functions. We emphasize that the reconstructed curves using deep learning are independent of cosmological model. The Λ CDM curves plotted in Figure 3 are just for comparison.

4. Results

With the reconstructed $d(z)$ curve, we can obtain the dimensionless comoving distance and the corresponding uncertainty at z_l and z_s , then calculate the distance ratio $R_{\text{SNe}} \equiv d_l/d_s$ according to Equation (4), and the uncertainty $\sigma_{R_{\text{SNe}}}$ propagates from the uncertainties of d_l and d_s . In addition, the distance ratios $R_{\text{SGL}} \equiv d_{ls}/d_s$ can also be obtained from SGL systems using Equations (5), (7) or (10), according to different mass models of a lens galaxy. The corresponding

Table 1

The Best-fitting Parameters in the Framework of SIS model Using the Distance Reconstructed with Tanh and SELU Functions

	Ω_k	f
tanh	$0.049^{+0.147}_{-0.125}$	$1.038^{+0.009}_{-0.008}$
SELU	$0.082^{+0.125}_{-0.130}$	$1.039^{+0.009}_{-0.009}$

uncertainty $\sigma_{R_{SGL}}$ propagates from the uncertainties of the observations of SGL. To compare the distance ratio obtained from SNe Ia and SGL systems, we determine the best-fitting parameters by maximizing the likelihood function, which is proportional to the exponential of the negative chi-square statistic, i.e., $\mathcal{L} \propto \exp(-\chi^2/2)$, where

$$\chi^2(\mathbf{p}, \Omega_k) = \sum_{i=1}^{161} \frac{(R_{SNe} - R_{SGL})^2}{\sigma_{total}^2}. \quad (13)$$

Here, \mathbf{p} represents the set of parameters for the lens mass profile, where $\mathbf{p} = f$ for the SIS model, $\mathbf{p} = (\gamma_0, \gamma_1)$ for the PL model, and $\mathbf{p} = (\alpha, \delta)$ for the EPL model. The term σ_{total} represents the total uncertainty, which includes contributions from the uncertainty in the reconstruction and the uncertainty propagated from the SGL observations

$$\sigma_{total}^2 = \sigma_{R_{SNe}}^2 + \sigma_{R_{SGL}}^2. \quad (14)$$

Assuming a flat prior on all free parameters, we calculate the posterior Probability Density Function (PDF) of the parameter space using the Python package `emcee` (Foreman-Mackey et al. 2013). It is worth noting that the prior on the spatial curvature parameter Ω_k is set to $\Omega_k \geq -0.39$ to ensure that $1 + \Omega_k d_l^2 \geq 0$ and $1 + \Omega_k d_s^2 \geq 0$ are within the redshift range $z \leq 4$.

In the context of the SIS lens model, we present the best-fitting parameters obtained using the reconstructions with the tanh and SELU activation functions in Table 1. Additionally, we provide the 1σ and 2σ confidence contours as well as the marginalized PDFs for the parameter space in Figure 4. For the spatial curvature, it is constrained to be $\Omega_k = 0.049^{+0.147}_{-0.125}$ with tanh function and $\Omega_k = 0.082^{+0.125}_{-0.130}$ with SELU function. The constraints on Ω_k in both of the two functions support a flat Universe within the 1σ confidence level, consistent with the Planck results (Aghanim et al. 2020). The constraint on the parameter f is rather tight, $1.038^{+0.009}_{-0.008}$ with the tanh function and $1.039^{+0.009}_{-0.009}$ with the SELU function. Both of them exclude the standard SIS model ($f=1$) at more than the 4σ confidence level. This indicates that the lens mass profile slightly but with strong evidence deviates from the standard SIS model.

In the context of the PL lens model, the parameters are presented in Table 2. Additionally, the contours and PDFs for the parameter space are plotted in Figure 5. Similar to the SIS model, the constraints with tanh and SELU activation functions are consistent with each other at the 1σ confidence level.

SIS model

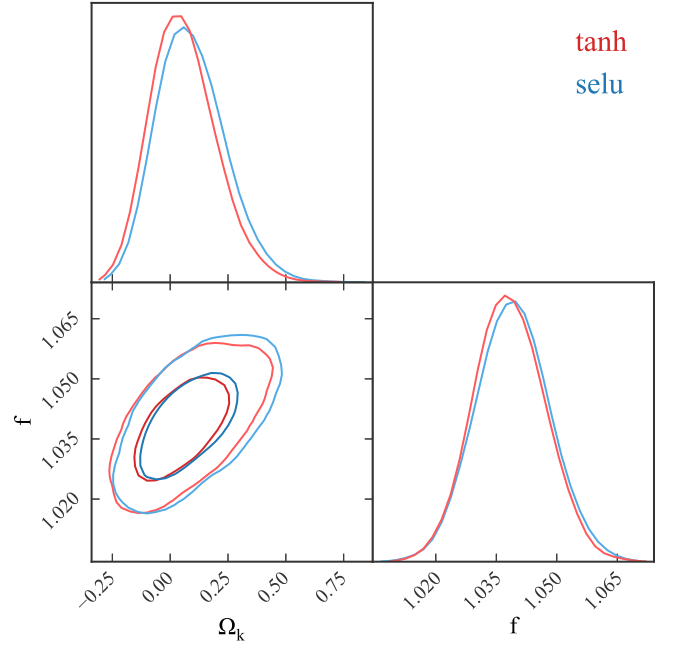


Figure 4. The two-dimensional confidence contours and one-dimensional PDFs for the parameters within the SIS model framework are depicted. The results obtained using the distance reconstructed with the tanh and SELU are represented by the red and blue lines, respectively.

Table 2

The Best-fitting Parameters in the Framework of PL Model Using the Distance Reconstructed with Tanh and SELU Functions

	Ω_k	γ_0	γ_1
tanh	$-0.245^{+0.075}_{-0.071}$	$2.076^{+0.028}_{-0.030}$	$-0.309^{+0.114}_{-0.091}$
SELU	$-0.232^{+0.076}_{-0.068}$	$2.074^{+0.028}_{-0.029}$	$-0.307^{+0.123}_{-0.093}$

However, the constraint on curvature parameter in the PL model is totally different from that in the SIS model. The spatial curvature is constrained to be $\Omega_k = -0.245^{+0.075}_{-0.071}$ with tanh function and $\Omega_k = -0.232^{+0.076}_{-0.068}$ with SELU function. The constraints on Ω_k in the PL model prefer a closed Universe at the $\sim 3\sigma$ confidence level. For the lens parameters, they are constrained to be $(\gamma_0, \gamma_1) = (2.076^{+0.028}_{-0.030}, -0.309^{+0.114}_{-0.091})$ with the tanh function, and $(\gamma_0, \gamma_1) = (2.074^{+0.028}_{-0.029}, -0.307^{+0.123}_{-0.093})$ with the SELU function. The results deviate from the standard SIS model ($\gamma_0 = 2, \gamma_1 = 0$) at more than the 2σ confidence level, demonstrating that the total mass density profile of the lens galaxy possibly evolves with cosmic time.

In the context of the EPL lens model, the results obtained with two activation functions are shown in Table 3 and Figure 6. The constraint in the EPL model is looser than that in

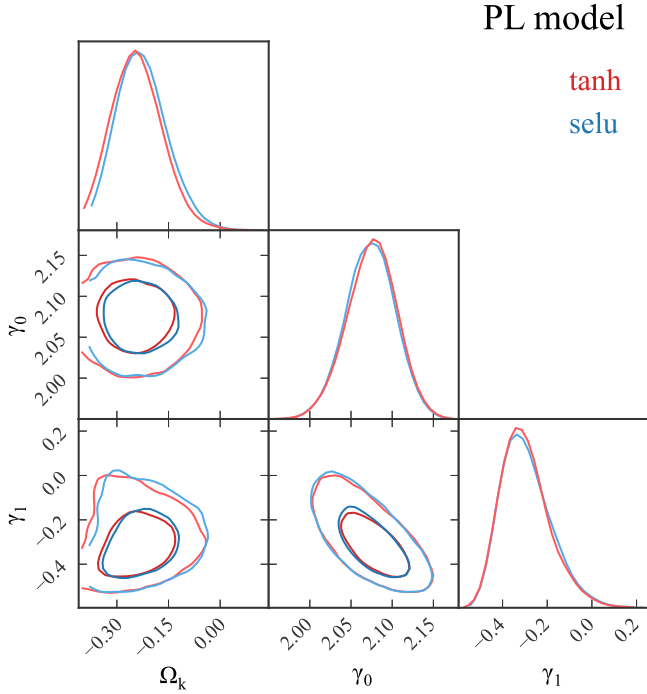


Figure 5. The two-dimensional confidence contours and one-dimensional PDFs for the parameters within the PL model framework are depicted. The results obtained using the distance reconstructed with the tanh and SELU are represented by the red and blue lines, respectively.

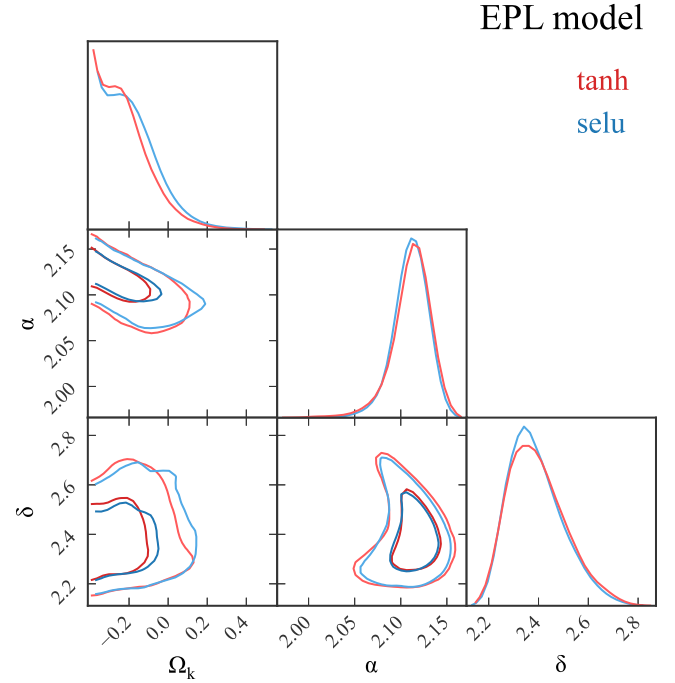


Figure 6. The two-dimensional confidence contours and one-dimensional PDFs for the parameters within the EPL model framework are depicted. The results obtained using the distance reconstructed with the tanh and SELU are represented by the red and blue lines, respectively.

the SIS and PL models. With two functions, the spatial curvature parameters are constrained to be $\Omega_k < 0.011$ in the tanh function and $\Omega_k < 0.051$ in the SELU function at the 95% confidence level. For the set of lens parameters, the results obtained with two activation functions are consistent with each other. We obtain $(\alpha, \delta) = (2.114^{+0.016}_{-0.019}, 2.383^{+0.128}_{-0.099})$ with the tanh function and $(\alpha, \delta) = (2.112^{+0.017}_{-0.018}, 2.375^{+0.125}_{-0.089})$ with the SELU function. Both results deviate from the SIS model ($\alpha = \delta = 2$) at more than the 1σ confidence level. Especially for α , it rules out $\alpha = 2$ at approximately the 3σ confidence level. These results affirm that the influence of dark matter in early-type galaxies should be considered and the total-mass profiles are not necessarily consistent with the luminosity profiles.

For enhanced clarity, Figure 7 showcases the optimal fitting outcomes concerning the curvature parameter Ω_k , accompanied by their corresponding 1σ uncertainties within the context of the SIS and PL models. Additionally, the upper and lower bounds of Ω_k within the EPL model are exhibited. Furthermore, to facilitate comprehensive comparison, we integrate the constraints on Ω_k originating from alternative cosmological methodologies, including outcomes from the Planck (Aghanim et al. 2020) and extended Baryon Oscillation Spectroscopic Survey (eBOSS) (Alam et al. 2021). Upon

Table 3

The Best-fitting Parameters in the Framework of EPL Model Using the Distance Reconstructed with Tanh and SELU Functions, the Constraints of Ω_k are Shown with the 95% Confidence Level Upper Limits

	Ω_k	α	δ
tanh	<0.011	$2.114^{+0.016}_{-0.019}$	$2.383^{+0.128}_{-0.099}$
SELU	<0.051	$2.112^{+0.017}_{-0.018}$	$2.375^{+0.125}_{-0.089}$

meticulous scrutiny, it becomes conspicuous that the constraints derived from the PL model exhibit noteworthy deviations from the outcomes of the Planck and eBOSS investigations. Meanwhile, the results stemming from the SIS and EPL models manifest congruence with the findings of the Planck and eBOSS initiatives. It is worth noting that the constraints associated with the EPL model, while consistent, display a marginally reduced stringency. Our findings underscore that, if the Universe is indeed flat, subtle deviations from the isothermal profile are discernible within the lens distribution. Moreover, it is imperative to duly consider variables such as the redshift evolution of lens profiles and the intricate interplay of dark matter in the broader landscape of cosmological research.

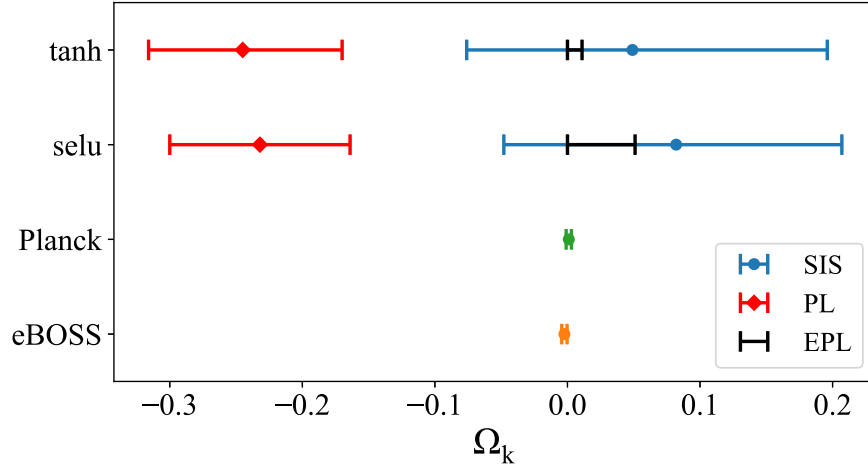


Figure 7. Constraint results of Ω_k in three different lens models using two activation functions in our work, compared with the constraints from other cosmological probes, Planck and eBOSS.

5. Discussion and Summary

Based on geometrical optics, the DSR offers a model-independent approach to testing the validity of the FRW metric in cosmology. The DSR has proven to be a valuable tool for constraining the spatial curvature of the Universe. Applying the DSR method, Wang et al. (2020) recently investigated the spatial curvature with the combination of an SGL sample and the latest Pantheon SNe Ia. Although the total number of SGL systems is 161, the available SGL systems in Wang et al. (2020) are just 135 due to the GP regression used to reconstruct the distance-redshift relation being unable to reconstruct the curve well beyond the data region. In this research, we use the same data samples but with a deep learning method to constrain the spatial curvature. In contrast to the GP method, deep learning exhibits enhanced capability in effectively reconstructing data beyond the observed range. Hence, we can make use of the full SGL systems and improve the precision of the constraints.

In this study, we developed a combined RNN and BNN architecture to accurately reconstruct the distance-redshift relation using the Pantheon sample. The RNN component of the network is specifically designed to predict the comoving distance at a given redshift, while the BNN component serves as a valuable complement, allowing for the calculation of uncertainties associated with these predictions. In the process of the distance reconstruction, we considered four activation functions and found that only the tanh and SELU functions can reproduce the Pantheon data well. Hence, we calibrated the distance of SGL systems with tanh and SELU functions. To investigate the possible influence of different lens models on constraining the spatial curvature, we considered three types of lens models, i.e., the SIS model, PL model and EPL model. In the SIS model, the spatial curvature is constrained to be

$\Omega_k = 0.049^{+0.147}_{-0.125}$ with the tanh function, and $\Omega_k = 0.082^{+0.125}_{-0.130}$ with the SELU function. Comparing with the result of Wang et al. (2020), $\Omega_k = 0.57^{+0.20}_{-0.28}$, which favors an open Universe at 2σ , our result favors a flat Universe with a higher accuracy due to the increase of available SGL data points. In the PL model, a closed Universe is favored, with the curvature parameter $\Omega_k = -0.245^{+0.075}_{-0.071}$ with the tanh function, and $\Omega_k = -0.232^{+0.076}_{-0.068}$ with the SELU function, which are consistent with $\Omega_k = -0.246^{+0.078}_{-0.100}$ obtained in Wang et al. (2020). In the EPL model, the spatial curvature is constrained to be $\Omega_k < 0.011$ with the tanh function, and $\Omega_k < 0.051$ with the SELU function. Comparing with the results $\Omega_k = 0.250^{+0.16}_{-0.23}$ in Wang et al. (2020), our constraint on the spatial curvature parameter is looser in the EPL model, but there is no strong evidence ruling out a flat Universe. On the other hand, for the set of parameters in three lens models, the results demonstrate that the lens galaxies cannot be simply described by the standard SIS model.

In summary, the lens mass models have a noticeable influence on the curvature parameter. In the SIS model, a spatially flat Universe is favored within 1σ uncertainty. In the PL model, a closed Universe is favored at the $\sim 3\sigma$ confidence level. In the EPL model, the constraint is relatively loose, but a flat Universe could not be excluded. More accurate modeling of the lens mass profile is necessary to further improve the constraint on the curvature parameter.

ORCID iDs

Li Tang  <https://orcid.org/0000-0002-0816-7740>

References

- Ade, P. A. R., Aghanim, N., Arnaud, M., et al. 2016, *A&A*, 594, A13
 Aghanim, N., Akrami, Y., Ashdown, M., et al. 2020, *A&A*, 641, A6

- Alam, S., Aubert, M., Avila, S., et al. 2021, *PhRvD*, **103**, 083533
- Bernstein, G. 2006, *ApJ*, **637**, 598
- Birrer, S., Treu, T., Rusu, C. E., et al. 2019, *MNRAS*, **484**, 4726
- Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., & Moustakas, L. A. 2006, *ApJ*, **638**, 703
- Brownstein, J. R., Bolton, A. S., Schlegel, D. J., et al. 2012, *ApJ*, **744**, 41
- Cabanac, R. A., Alard, C., Dantel-Fort, M., et al. 2008, *A&A*, **461**, 813
- Cai, R.-G., Guo, Z.-K., & Yang, T. 2016, *PhRvD*, **93**, 043517
- Cao, S., Biesiada, M., Gavazzi, R., Piórkowska, A., & Zhu, Z.-H. 2015, *ApJ*, **806**, 185
- Cao, S., Biesiada, M., Yao, M., & Zhu, Z.-H. 2016, *MNRAS*, **461**, 2192
- Cao, S., Liu, T., Biesiada, M., et al. 2022, *ApJ*, **926**, 214
- Cao, S., Pan, Y., Biesiada, M., Godlowski, W., & Zhu, Z.-H. 2012, *JCAP*, **03**, 016
- Cappellari, M., Bacon, R., Bureau, M., et al. 2006, *MNRAS*, **366**, 1126
- Chen, Y., Li, R., Shu, Y., & Cao, X. 2019, *MNRAS*, **488**, 3745
- Clarkson, C., Bassett, B., & Lu, T. H.-C. 2008, *PhRvL*, **101**, 011301
- Clarkson, C., Cortês, M., & Bassett, B. 2007, *JCAP*, **2007**, 011
- Cohn, J. D., Kochanek, C. S., McLeod, B. A., & Keeton, C. R. 2001, *ApJ*, **554**, 1216
- Di Valentino, E., Melchiorri, A., & Silk, J. 2019, *NatAs*, **4**, 196
- Di Valentino, E., Anchordoqui, L. A., Akarsu, Özgür, et al. 2021, *Aph*, **131**, 102605
- Dieleman, S., Willett, K. W., & Dambre, J. 2015, *MNRAS*, **450**, 1441
- Escamilla-Rivera, C., Quintero, M. A. C., & Capozziello, S. 2019, *JCAP*, **2020**, 008
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
- Freedman, W. L. 2017, *NatAs*, **1**, 0121
- Gal, Y., & Ghahramani, Z. 2016a, arXiv:1512.05287
- Gal, Y., & Ghahramani, Z. 2016b, arXiv:1506.02157
- Gal, Y., & Ghahramani, Z. 2016c, arXiv:1506.02142
- Gong, Y.-G., & Wang, A. 2007, *PhRvD*, **75**, 043520
- Ichikawa, K., & Takahashi, T. 2006, *PhRvD*, **73**, 083526
- Jorgensen, I., Franx, M., & Kjaergaard, P. 1995, *MNRAS*, **276**, 1341
- Khedekar, S., & Chakraborti, S. 2011, *PhRvL*, **106**, 221301
- Kochanek, C. S. 1992, *ApJ*, **384**, 1
- Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, *ApJ*, **649**, 599
- Lee, S. 2021, arXiv:2104.09690
- Li, X., Tang, L., & Lin, H.-N. 2019, *MNRAS*, **482**, 5678
- Li, Z., Ding, X., Wang, G.-J., Liao, K., & Zhu, Z.-H. 2018, *ApJ*, **854**, 146
- Liao, K., Li, Z., Cao, S., et al. 2016, *ApJ*, **822**, 74
- Liu, Y., Cao, S., Liu, T., et al. 2020, *ApJ*, **901**, 129
- Luongo, O., & Muccino, M. 2021, *MNRAS*, **503**, 4581
- Mollerach, S., & Roulet, E. 2002, *Gravitational Lensing and Microlensing* (Singapore: World Scientific)
- Moore, B., Governato, F., Quinn, T. R., Stadel, J., & Lake, G. 1998, *ApJL*, **499**, L5
- Mortsell, E., & Jonsson, J. 2011, arXiv:1102.4485
- Munoz, J. A., Kochanek, C. S., & Keeton, C. R. 2001, *ApJ*, **558**, 657
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, **462**, 563
- Ofek, E. O., Rix, H.-W., & Maoz, D. 2003, *MNRAS*, **343**, 639
- Qi, J.-Z., Cao, S., Zhang, S., et al. 2019, *MNRAS*, **483**, 1104
- Räsänen, S., Bolejko, K., & Finoguenov, A. 2015, *PhRvL*, **115**, 101301
- Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, *ApJ*, **876**, 85
- Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, *ApJ*, **826**, 56
- Rusin, D., Norbury, M., Biggs, A. D., et al. 2002, *MNRAS*, **330**, 205
- Rusin, D. J., & Kochanek, C. S. 2005, *ApJ*, **623**, 666
- Sapone, D., Majerotto, E., & Nesseris, S. 2014, *PhRvD*, **90**, 023012
- Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, *ApJ*, **859**, 101
- Scolnic, D., Rest, A., Riess, A., et al. 2014, *ApJ*, **795**, 45
- Tang, L., Li, X., Lin, H.-N., & Liu, L. 2021a, *ApJ*, **907**, 121
- Tang, L., Lin, H.-N., Li, X., & Liu, L. 2021b, *MNRAS*, **509**, 1194
- Treu, T., & Koopmans, L. 2002, *ApJ*, **575**, 87
- Treu, T., & Koopmans, L. V. E. 2004, *ApJ*, **611**, 739
- Virey, J. M., Talon-Esmieu, D., Ealet, A., Taxil, P., & Tilquin, A. 2008, *JCAP*, **12**, 008
- Wang, B., Qi, J.-Z., Zhang, J.-F., & Zhang, X. 2020, *ApJ*, **898**, 100
- Xia, J.-Q., Yu, H., Wang, G.-J., et al. 2017, *ApJ*, **834**, 75