Physics of the Dark Universe 32 (2021) 100824

Contents lists available at ScienceDirect

Physics of the Dark Universe

journal homepage: www.elsevier.com/locate/dark

Strongly lensed supernovae as a self-sufficient probe of the distance duality relation

Fabrizio Renzi [a,*], Natalie B. Hogg [b], Matteo Martinelli [c], Savvas Nesseris [c]

[a] Institute Lorentz, Leiden University, PO Box 9506, Leiden 2300 RA, The Netherlands
[b] Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Road, Portsmouth, PO1 3FX, United Kingdom
[c] Instituto de Física Téorica UAM-CSIC, Campus de Cantoblanco, E-28049 Madrid, Spain

ARTICLE INFO

Article history:
Received 20 October 2020
Received in revised form 19 April 2021
Accepted 20 April 2021

Keywords:
Cosmology
Strong lensing
Forecasts

ABSTRACT

The observation of strongly lensed Type Ia supernovae enables both the luminosity and angular diameter distance to a source to be measured simultaneously using a single observation. This feature can be used to measure the distance duality parameter $\eta(z)$ without relying on multiple datasets and cosmological assumptions to reconstruct the relation between angular and luminosity distances. In this paper, we show how this can be achieved by future observations of strongly lensed Type Ia systems. Using simulated datasets, we reconstruct the function $\eta(z)$ using both parametric and non-parametric approaches, focusing on Genetic Algorithms and Gaussian processes for the latter. In the parametric approach, we find that in the realistic scenario of $N_{\rm lens} = 20$ observed systems, the parameter ϵ_0 used to describe the trend of $\eta(z)$ can be constrained with the precision achieved by current SNIa and BAO surveys, while in the futuristic case ($N_{\rm lens} = 1000$) these observations could be competitive with the forecast precision of upcoming LSS and SN surveys. Using the machine learning approaches of Genetic Algorithms and Gaussian processes, we find that both reconstruction methods are generally well able to correctly recover the underlying fiducial model in the mock data, even in the realistic case of $N_{\rm lens} = 20$. Both approaches learn effectively from the features of the mock data points, yielding 1σ constraints that are in excellent agreement with the parameterised results.

1. Introduction

The HOLISMOKES project recently demonstrated that the exciting possibility of using strongly lensed Type Ia supernovae (SNIa) as a precision probe in cosmology could soon become a reality [1]. Strong gravitational lensing occurs when a massive object lies along the line of sight between a luminous source and an observer. The gravitational field of the lens distorts the spacetime along the line of sight, bending the light path of photons coming from the source which results in a remapping of the source light into multiple images [2,3].

Due to the different light paths taken by photons coming from the source, these images arrive at the observer at different times and are therefore delayed with respect to one another. The time delay between images, which can be measured up to an arbitrary length of time [4,5], is a typical lensing observable which is only sensitive to the mass profile of the lens and to a combination of the source and lens angular diameter distances, the so-called time delay distance [2,3]. Provided that one can properly reconstruct the lens mass profile, the strong lensing time delay can then be used as a tracer of the distance–redshift relation, and to infer constraints on cosmological parameters [6–9].

While lensing can happen at all scales (*i.e.* the lens can be as small as a star or as big as a galaxy cluster), for cosmological inference one typically relies on galaxy–galaxy lensing events [6]. This is because galaxies are believed to have simple mass profiles that can be effectively parameterised as a power law, and a larger lensing probability, making them more abundant in the sky, although it has been shown that uncertainties in the mass profiles play a significant role in constraining cosmological parameters [10]. Furthermore, by combining measurements of the velocity dispersion of the stars orbiting the lensing galaxy with the strong lensing time delay, it is possible to obtain a measurement of the angular diameter distance to the lens, which breaks the degeneracy between different lens mass profiles [11–14].

However, strong lensing observables are significantly affected by the specific alignment between the lens and the source, making only a fraction of the observed lensing events suitable for the extraction of cosmological information [15,16]. To achieve percentage accuracy on time delay measurements, the image separation is required to be $> 1''$, the magnitude of the faintest image $m_i < 21$ in the i-band, and the lensing galaxy magnitude $m_i < 22$ [17]. Therefore, it is clear that, along with

* Corresponding author.
E-mail address: renzi@lorentz.leidenuniv.nl (F. Renzi).

https://doi.org/10.1016/j.dark.2021.100824
2212-6864/

good source–lens alignment, one needs sources with a typical brightness comparable to a galaxy to accurately distinguish the lens galaxy from the lensed images. This has led to the use of lensed quasars as the major cosmological probe in the context of lensing, an approach which has been proven by the H0LiCOW collaboration to be highly successful in deriving cosmological constraints [18–24].

There exists another family of astrophysical objects that have luminosities comparable to that of a galaxy: supernova explosions. The concept of using strong lensing of SNIa as a cosmological probe was pioneered in 1964 by Refsdal [25], who showed that the strong lensing time delays can be used to directly measure the Hubble parameter, $H(z)$. However, since lensed supernovae are thought to be far rarer than lensed quasars, the idea of using them for cosmology has long been considered a fruitless endeavour. This changed with the recent observations of two lensed supernova events (the core collapse supernova "Refsdal" [26] in 2014 and the Type Ia supernova iPTF16geu [27] in 2016), which reinvigorated the field [28]. As highlighted by HOLISMOKES [1], cosmology with strongly lensed SNIa will soon be possible with surveys like LSST, which is expected to measure around a thousand such events [29–32].

As previously mentioned, gravitational lensing remaps the source light from the source plane to the lens plane. While the source surface brightness is conserved in the process, the area on the lens plane in which source photons are remapped is not conserved. In other words the flux of the lensed images is different from the source flux, their ratio defining the magnification factor. From lensing observations, one typically measures the ratio of magnification between the images by comparing their measured fluxes, but the total magnification is not directly measurable because the unlensed source brightness (*i.e.* the unlensed source flux) is unknown. So, despite their relative rarity in comparison to lensed quasars, lensed SNIa have one compelling advantage: they allow the source brightness to be measured independently from lensing observations [33].

By assuming that SNIa are standardisable candles, the brightness (and brightness decay after the explosion) can be inferred from the light curves of the lensed events, which are well known from unlensed supernovae observations. The total magnification can then be tightly constrained, reducing the uncertainties in the lens mass profile and improving the possible cosmological constraints [33]. Since this enables us to measure the luminosity distance to these events, they can be used to test more fundamental aspects of the standard cosmological model.

We note that microlensing and other lensing effects related to substructures (such as dust clouds and subhalos) in the deflector galaxy can significantly affect the standardisable nature of SNIa, leading to large uncertainties in the inferred unlensed flux [31–35]. However, it is expected that a significant fraction of lensed SNIa will be standardisable: around 20% from an LSST-like survey [32,35,36]. In the following, we assume the effect of microlensing and other effects related to substructures in the lensing galaxy to be negligible.

The distance duality relation (DDR), which relates luminosity distances to angular diameter distances, is one example of a fundamental component of cosmology which is accessible with strongly lensed SNIa. Combining information from the velocity dispersion of stars in the lensing galaxy with lensing observations and supernova light curves, lensed SNIa can provide both measurements of angular diameter and luminosity distance, making these events particularly well-suited to probing the DDR and investigating any possible deviations from it, which could indicate the presence of new physics.

In this paper, we aim to reconstruct a function related to the DDR using mock datasets of strongly lensed SNIa. We create the mock datasets for an LSST-like survey, testing three cases: realistic (20 useful lensed SNIa as expected by LSST after 10 years of observations [1]), optimistic (100 lenses corresponding to the total number of spatially-resolved lensed SNIa by LSST [29]) and futuristic (1000 lenses representing the number of events we expect to observe in the next few decades). Using both parametric and non-parametric approaches for our reconstructions, we investigate whether violations of the distance duality relation could be detected with datasets of this size, finding that the realistic LSST-like survey would be competitive with other more traditional probes of the DDR such as the combination of SNIa and BAO observations.

We note that a similar analysis, involving strong lensing in the context of constraining the DDR, was performed in [37–39]. However, our approach in this paper differs significantly to those previous works. The main difference is that in those works it was shown that it is possible to obtain angular diameter distance measurements from strong lensing events in place of other observations able to provide this quantity (such as BAO), but additional distance luminosity measurements were still needed to constrain the DDR. Instead, we show that both the luminosity and angular diameter distances can be measured from strongly lensed SNIa, exploiting the standardisable nature of supernovae explosions along with the "standard ruler" nature of strong lensing events. This makes strongly lensed SNIa a self-sufficient probe of the DDR.

The structure of our paper is as follows: in Section 2 we present some theoretical aspects of the distance duality relation, in Section 3 we discuss the physics of the strongly lensed supernovae and the details of the mock data, while in Section 4 we present our methodology, with the parameterised and non-parametric approaches, and our results. Finally, in Section 5 we summarise our conclusions.

2. The distance duality relation

The distance duality relation is given by [40]

$$d_L(z) = (1+z)^2 d_A(z), \tag{1}$$

where $d_L(z)$ is the luminosity distance and $d_A(z)$ is the angular diameter distance. It holds under the conditions that photons travel along null geodesics in an pseudo-Riemannian spacetime, and that the number of photons is conserved [41].

The first condition is a fundamental statement about the geometry of spacetime and the photon mass and is only violated in theories of gravity with a non-Riemannian geometry, or in theories where photons do not propagate on null geodesics due to coupling with other fields (see e.g. [42–46]). It is easier to imagine deviations from DDR occurring due to non-conservation of the photon number, for example by absorption or scattering by dust as they propagate to the observer, or via more exotic mechanisms such as the conversion of photons to axions as they interact with cosmic magnetic fields [47].

In order to investigate these possible deviations from DDR, a function $\eta(z)$ can be defined from Eq. (1) as

$$\eta(z) = \frac{d_L(z)}{(1+z)^2 d_A(z)}, \tag{2}$$

which is equal to unity if the DDR is not altered. DDR violation mechanisms are integrated effects, where photons interact with intervening components along the line of sight. Thus, one can expect that for a photon at redshift zero, such an effect does not have time to take place and no violation of the relation is present, meaning that $\eta(z=0)=1$. This is also clear from Eq. (2), whose limit for $z=0$ is $\lim_{z\to 0}\eta(z)=1$. For this reason, we impose that

$\eta(z)$ is equal to 1 at vanishing redshifts, for both our parametric and non-parametric reconstructions.

The function $\eta(z)$ is also commonly parameterised in the literature (e.g. [48,49]) as

$$\eta(z) = (1+z)^{\epsilon(z)}, \tag{3}$$

where $\epsilon(z) \neq 0$ is equivalent to $\eta(z) \neq 1$, thus indicating a deviation from the standard DDR. To probe this relation and search for violations of DDR, objects for which both a luminosity distance and angular diameter distance are available are needed. This motivates the use of strongly lensed SNIa, which amply fulfil these criteria.

3. Strongly lensed supernovae

A survey of strongly lensed SNIa will observe the distance modulus of the supernovae, *i.e.* the difference between its apparent and absolute magnitude, which is given by

$$\mu(z_s) = 5\log_{10}\left(\frac{d_L(z_s)}{\text{Mpc}}\right) + 25, \tag{4}$$

and the time delay distance (see e.g. [8]),

$$d_{\Delta t}(z_l) = (1+z_l)(1+z_s)d_A(z_l)d_A(z_s) \times \left[(1+z_s)d_A(z_s) - (1+z_l)d_A(z_l)\right]^{-1}, \tag{5}$$

where z_s is the redshift of the source and z_l the redshift of the lens. Notice that Eq. (5) only holds under the assumption of flat space, i.e. $\Omega_k = 0$, in the context of a flat Friedmann–Lemaître–Robertson–Walker metric. In curved space, the second term on the right hand side would become $d_A(z_s, z_l)$. In this paper we want to obtain measurements of $d_A(z_s)$ and therefore the assumption of a flat Universe allows us to isolate this term in the time delay distance expression. We leave the investigation of more general cases for future work. Under this assumption we can invert Eq. (5) and obtain $d_A(z_s)$, and we can write our parameterisation of the distance duality relation in terms of the distance modulus, the angular diameter distance at the lens and the time delay distance as

$$\eta(z_s) = \frac{10^{-5+\mu(z_s)/5}}{(1+z_l)(1+z_s)}\left[\frac{1}{d_A(z_l)} - \frac{1+z_l}{d_{\Delta t}}\right](\text{Mpc}). \tag{6}$$

The number of currently detected lensed SNIa is insufficient for any precise cosmological application, so we turn to mock datasets to forecast our future ability to probe the distance duality relation with these events.

3.1. Mock dataset

To generate our mock datasets, we focus on lensed SNIa for which measurements of the kinematics of the lens galaxy are available, along with time delay observations. In this scenario, strong lensing will provide two independent distance measures at the same time [11,13,17]: $d_{\Delta t}(z_l)$ and $d_A(z_l)$. The measurements of the time delay distance of a lens are obtained by combining the observation of time delays between the light curves of multiple images, a lens mass model for the lensing galaxy and a reconstruction of the mass environment along the line of sight [18–24]. We therefore consider only these contributions to the uncertainties of $d_{\Delta t}$.

As in [1], to estimate the precision on $d_{\Delta t}$ we conservatively adopt a 5% uncertainty for the time delay and a 3% uncertainty for both the mass profile and the lens environment. Summing these in quadrature we obtain a cumulative uncertainty on $d_{\Delta t}$ of 6.6%, in agreement with current constraints from lensed quasars[1] [24]. For the angular diameter distance to the lens, $d_A(z_l)$, we assume a scenario where spatially-resolved observations of the kinematics of the lens galaxy are available, so that the uncertainties of d_A are essentially dominated by the time delay uncertainties. These measurements are expected to be obtained easily after all the SNIa images have faded. We therefore adopt a 5% precision for d_A.

The missing ingredient of our mock dataset is now the distance modulus $\mu(z_s)$ of the lensed SNIa. This quantity must be reconstructed starting from the lensed distance modulus of four lensed images. For standardisable candles this implies fitting the lensed light curves, with exactly the same procedure used for unlensed SNIa, to provide an estimate of the lensed distance modulus $\hat{\mu}$ without any cosmological assumption or knowledge of the lens model. The unlensed distance modulus is then related to the lensed one by the following relation:

$$\mu = \hat{\mu} + \frac{5}{2}\log_{10} A \tag{7}$$

where A is the magnification factor of the lensed event, defined as the ratio of the lensed to unlensed flux, *i.e.*

$$2.5\log_{10} A = 2.5\log_{10}\left(\frac{f_{\text{lensed}}}{f_{\text{unlensed}}}\right) \tag{8}$$

This delensing procedure to infer the unlensed distance modulus can be summarised in two simple steps:

1. Estimate the lensed magnitude, $\hat{\mu}$, from the observed light curves of the lensed SNIa.
2. Assume a mass profile to estimate the lensing magnification,[2] delens the SNIa and obtain the unlensed modulus distance, μ.

Assuming this approach to be feasible for all the systems in our catalogues to infer the unlensed $\mu(z_s)$, we model its error budget due to the SNIa brightness uncertainties following [50] and to this we add in quadrature the magnification uncertainty:

$$\sigma[\mu(z_s)]^2 = \delta\mu(z_s)^2 + \sigma^2_{\text{flux}} + \sigma^2_{\text{scat}} + \sigma^2_{\text{intr}} + \frac{25\sigma^2_{\log A}}{4} \tag{9}$$

where the systematic uncertainties due to flux calibration are given by $\sigma_{\text{flux}} = 0.01$, the intrinsic scatter of SNe at fixed colour, also known as colour smearing, is given by $\sigma_{\text{scat}} = 0.025$, the intrinsic distance scatter is $\sigma_{\text{intr}} = 0.12$ and finally, we also include an irreducible distance modulus error, which we assume affects all events coherently and varies linearly with redshift in the form $\delta\mu(z_s) = e_M z_s$ with e_M drawn from a normal distribution $\mathcal{N}(0, 0.01)$ [50]. For the error on the lensing magnification we assume a $\sim 20\%$ fractional uncertainties, *i.e.* $\sigma_{\log A} = 0.09$ [33].

To generate the mock, we assume the lens distribution to be uniform in the range $0.1 \leq z \leq 0.9$ and the source redshift to be twice the lens redshift *i.e.* $z_s = 2z_l$ for simplicity. Even though there will be a distribution for the redshifts of the sources this has a small impact on cosmological inference [51,52].

[1] The assumed uncertainties correspond to having a perfect knowledge of the lens mass profile and its environment. As detailed in [10], a hierarchical analysis of the lensing observables may lead to higher uncertainties in the time delay distance.

[2] As the unlensed flux is not measured in lensing observations, lensing magnification has to be determined from the lens mass profile. However, the same mass profile is needed to infer the angular and time delay distances and can be found by studying the lens galaxy and its environment [14,18,20,22,24]. Another possibility is to get the unlensed magnitude from an external catalogue of unlensed SNIa and estimate the magnification from Eq. (8) [27]. In this case one can still estimate the distance modulus but it would be the same as the one being inferred from the actual SNIa catalogue, spoiling the information of the lensed event except for the redshift z_s

Assuming a ΛCDM fiducial cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_m = 0.3$ (with $\Omega_k = 0$), we calculate the angular diameter distance $d_A(z)$ at the given z_l and z_s. From this we can obtain $d_{\Delta t}(z)$ using Eq. (5), while to compute the fiducial distance modulus $\mu(z)$ we use Eq. (4), obtaining the luminosity distance from $d_A(z)$ through Eq. (2), which implies choosing a fiducial $\eta(z)$. We rely on the parameterised expression of $\eta(z)$ of Eq. (3), and we choose for our fiducial a constant $\epsilon(z) = \epsilon_0$. We focus on three different choices for this parameter, in order to be able to test the precision of future observations in different scenarios. We choose the standard DDR value $\epsilon_0 = 0$, and two fiducials with different degrees of departure from DDR, with $\epsilon_0 = 0.01,\ 0.05$.

Once the fiducial trends for our observables are computed, we obtain the mock datasets by drawing a random Gaussian shift around the fiducial, using the estimated 1σ uncertainties for $d_A(z_l)$, $d_{\Delta t}(z_l)$ and $\mu(z_s)$:

$$D_{i,\mathrm{mock}} = D_{\mathrm{mock}}(z_i) = D_{\mathrm{true}}(z_i) + \delta D(z_i), \tag{10}$$

with $i = 1 \ldots N_{\mathrm{lens}}$, D_{true} representing the fiducial value of either d_A, $d_{\Delta t}$ and μ, and δD being the corresponding Gaussian deviate. From this we get our mock distances as $D_{i,\mathrm{mock}} \pm \sigma_{D(z_i)}$ where $\sigma_{D(z_i)}$ are the 1σ uncertainties of the distance considered. Finally we use Eq. (6) to obtain a mock catalogue for $\eta(z_i)$ from the mock datasets of $d_A(z_l)$, $d_{\Delta t}(z_l)$ and $\mu(z_s)$. To obtain the error on each of the data points of the mock of $\eta(z_i)$, we employ an MCMC-like approach, detailed as follows:

1. We construct the distribution of each of the $D_{i,\mathrm{mock}}$ distances at each redshift z_i of the catalogue, drawing 10,000 random samples from the assumed distribution for $D_{i,\mathrm{mock}}$.
2. We combine each of the 10,000 random samples using Eq. (6) to obtain 10,000 realisations of the distribution of $\eta(z_i)$ at each redshift z_i.
3. We calculate the mean and standard deviation of $\log_{10}\eta(z_i)$ from the $\eta(z_i)$ distributions at each redshift to construct our final mock datasets.

A more detailed explanation of the procedure followed to construct the mock datasets can be found in Appendix.

Our choice to construct the catalogue using $\log_{10}\eta(z_i)$ is motivated by the fact that the distribution of $\eta(z_i)$ are almost log-normal and therefore $\log_{10}\eta(z_i)$ is almost Gaussian distributed around zero *i.e.* $\log_{10}\eta(z_i) \approx \mathcal{N}(0, \sigma_{\log_{10}\eta(z_i)})$. This allows us to derive constraints from our mock catalogues by employing an MCMC approach with a Gaussian likelihood of the form:

$$-2\ln\mathcal{L} = \sum_{i=1}^{N_{\mathrm{lens}}} \frac{\left[\log_{10}\eta(z_i) - \log_{10}\eta^{\mathrm{th}}(z_i)\right]^2}{\sigma^2_{\log_{10}\eta(z_i)}} \tag{11}$$

where $\log_{10}\eta^{\mathrm{th}}(z_i)$ is the theoretical value of $\log_{10}\eta(z_i)$.

Furthermore, the choice of constructing the catalogue for $\log_{10}\eta(z)$ is also useful for the application of Gaussian processes that we describe in Section 4.3 below; this approach requires the choice of a mean prior for the reconstructed function, which is usually assumed to be zero in standard applications. The choice of reconstructing $\log_{10}\eta(z)$ allows us to keep this assumption without significantly biasing the results.

4. Methodology and results

In this section we describe the methodology we use in our analysis and our corresponding results. We first use a simple parameterisation of the DDR violation function $\eta(z)$, forecasting the constraints that can be achieved with realistic ($N_{\mathrm{lens}} = 20$), optimistic ($N_{\mathrm{lens}} = 100$) and futuristic ($N_{\mathrm{lens}} = 1000$) mock datasets. We then focus only on the realistic and optimistic datasets and we apply machine learning approaches, namely Genetic Algorithms (GA) and Gaussian processes (GP), to reconstruct $\eta(z)$.

Table 1
Mean values and 68% confidence level intervals for the ϵ_0 parameter, using mock data with different number of lenses and fiducial values for ϵ_0.

$N_{\mathrm{lens}} = 20$	$N_{\mathrm{lens}} = 100$	$N_{\mathrm{lens}} = 1000$
Fiducial $\epsilon_0 = 0.0$		
0.0098 ± 0.057	$0.015^{+0.019}_{-0.023}$	0.0038 ± 0.0065
Fiducial $\epsilon_0 = 0.01$		
0.022 ± 0.056	0.025 ± 0.021	0.0127 ± 0.0064
Fiducial $\epsilon_0 = 0.05$		
0.056 ± 0.057	0.066 ± 0.022	0.0534 ± 0.0065

4.1. Parameterised approach

We first adopt a simple parameterised approach to forecast the constraints achievable on DDR violation with future strongly lensed SNIa observations. We use the parameterisation of Eq. (3), and we assume the function $\epsilon(z)$ to be constant, with its value ϵ_0 the free parameter that we want to constrain with our mock dataset.

We build a likelihood module interfaced with the publicly available MCMC sampler `Cobaya` [53] which compares the prediction for

$$\log_{10}\eta^{\mathrm{th}}(z) = \epsilon_0 \log_{10}(1+z), \tag{12}$$

with the mock dataset we described in Section 3.1.

The improvement brought by strongly lensed SNIa observations to this analysis is evident. In most previous constraints of DDR violations, predictions of both $d_L(z)$ and $d_A(z)$, which enter in the definition of $\eta(z)$ in Eq. (2), were required, as the two observables are compared independently with data (see e.g. [54–56]). Such an approach is intrinsically dependent on the assumptions made about the expansion history of the Universe, and in particular on the assumed dark energy model driving the late time accelerated expansion. Here, such an assumption is not necessary, as the distances entering Eq. (6) are obtained at each redshift from a single observation, and therefore there is no need to assume a cosmological model to reconstruct the luminosity and angular distances.

However, it is important to note that we assume that $\eta(z)$ as defined in Eq. (2) is a valid description of DDR violation, which implies that the Universe is to first approximation homogeneous and isotropic. Finally, for Eq. (6) to hold, we further assume that the contributions to the total energy density by curvature are negligible ($\Omega_k = 0$).

For these reasons, the only free parameter in this analysis is ϵ_0, for which we use a flat prior. The constraints we obtain on this are shown in Table 1 and the posterior distributions in Fig. 1. We find that the realistic case ($N_{\mathrm{lens}} = 20$) would achieve the same constraining power of current constraints obtained through the combination of SNIa and BAO observations [55], while the futuristic case ($N_{\mathrm{lens}} = 1000$) reaches a sensitivity similar to the one that can be achieved by the combination of the *Euclid* BAO survey with the full LSST SNIa survey [55].

The optimistic case ($N_{\mathrm{lens}} = 100$) sits somewhere in the middle, but given the reduced number of assumptions made on the cosmological model in the analysis of strongly lensed SNIa, using this approach could allow DDR violation to be disentangled from other cosmological mechanisms [56].

4.2. Genetic algorithms

Here we describe a non-parametric reconstruction of the duality parameter $\eta(z)$, which is based on a machine learning approach called the Genetic Algorithms (GA) and is complementary

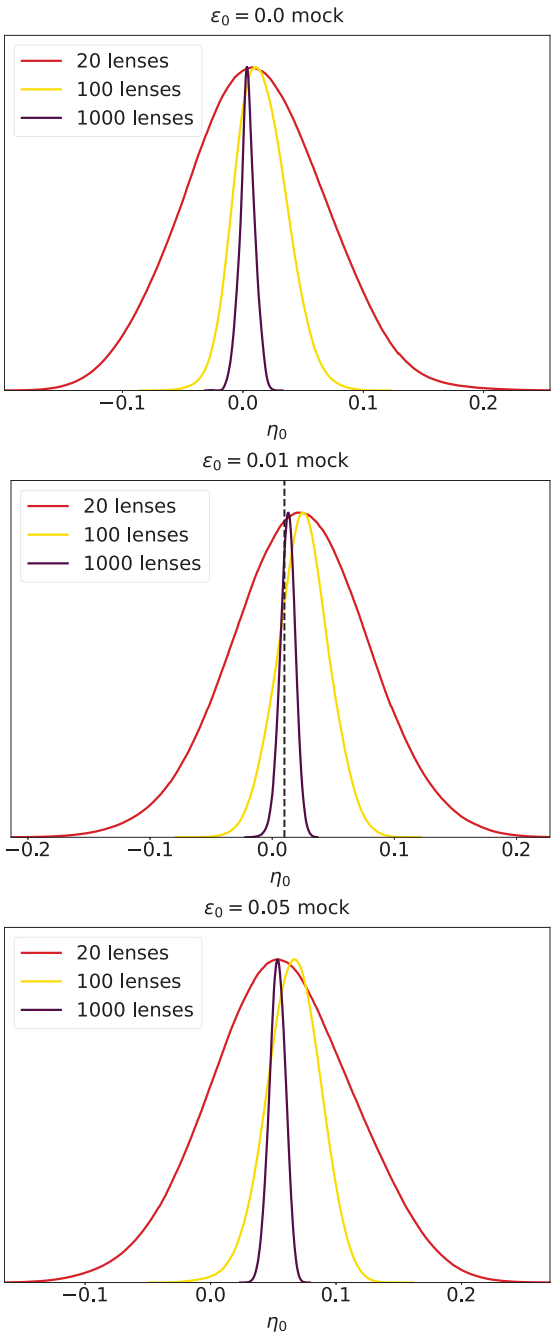

Fig. 1. Posterior distributions for the DDR violation parameter ϵ_0. The three panels refer to the different fiducial values considered to build the mock data: $\epsilon_0 = 0$ (top), $\epsilon_0 = 0.01$ (centre) and $\epsilon_0 = 0.05$ (bottom). In all panels the different lines show the posterior distribution for the realistic (red), optimistic (yellow) and futuristic (purple) cases. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to the parameterised analysis of the previous section. The GA are a particular stochastic optimisation approach, loosely inspired from the theory of evolution and mimicking the stochastic operations of mutation, *i.e.* the merging of different individuals to form descendants, and crossover, a random change in the chromosomes of an individual. This is achieved by emulating natural selection, *i.e.* in a given environment, a population (in our case a set of test functions) will evolve and adapt under the pressure of the operators of mutation and crossover.

In general, the reproductive success of every member of the population is assumed to be proportional to their fitness, which is a measure of how well they fit the data in question. Here we implement a standard χ^2 statistic as described in the previous sections. For more details on the GA and their applications to cosmology see Refs. [57–64].

A quick overview of the fitting process is as follows. During the initialisation of the code a set of test functions is formed using a group of orthogonal polynomials, called the grammar. This is a crucial step as it has been shown that the choice of the grammar may significantly affect the convergence rate of the GA code [57]. Using then this initial population, we encode the duality parameter $\eta(z)$ in every member of the population and we also require that $\eta(z)$ satisfies a set of physical priors and initial conditions. In our analysis we remain completely agnostic regarding the DDR deviation mechanism, so we only assume that the duality parameter satisfies $\eta(z = 0) = 1$, but we make no assumption of a dark energy model.

After preparing the initial population, we then estimate the fitness of every member using the χ^2 and then we apply the stochastic operators of crossover and mutation to a subset of the best-fitting functions chosen via tournament selection [57]. We then repeat this process thousands of times, so as to make certain the GA code has converged, and we also use several different random seeds, in order to avoid biasing the run due to a specific random seed.

The errors in the reconstruction are calculated using the path integral approach of Refs. [58,60]. In this approach the error regions are estimated by integrating the likelihood over all functions of the functional space scanned by the GA. This method has been validated by comparing its error estimates against bootstrap Monte Carlo and Fisher matrix errors [58]. Finally, here we use the publicly available code `Genetic Algorithms`.[3]

The results of the GA reconstruction can be seen in Fig. 2. In the left column we show the reconstructions for 20 lenses, while in the right column we show the case for 100 lenses. The mocks in the top row were made with $\epsilon = 0$, the ones in the middle row with $\epsilon = 0.01$, while the ones in the bottom row with $\epsilon = 0.05$. As can be seen, in both cases of the 20 and 100 lenses, the GA is able to correctly recover within the errors the underlying fiducial model $\eta_{\rm fid}(z) = \log_{10}(1+z)^{\epsilon_0}$, shown with a dashed line in each of the panels.

Specifically, we find that in the case of the 20 lenses the GA is able to predict the fiducial model very well across all redshifts, albeit with a small tension at high redshifts ($z \gtrsim 1.4$) due to the lack of points. On the other hand, in the case of the 100 lenses the GA reconstruction remains very close to the fiducial model at all redshifts.

4.3. Gaussian processes

The classic definition of a Gaussian process (GP) is "a collection of random variables, any finite number of which have a joint Gaussian distribution" [65]. A GP can be thought of as a generalisation of a Gaussian probability distribution, but whereas a probability distribution describes finite-dimensional random variables, a stochastic process governs the properties of functions. In our case, this function that we use a GP to reconstruct is $\log_{10}\eta(z)$, with the redshifts being the input fed to the GP. In general, the GP is completely specified by its mean and covariance functions, though the mean function is usually taken to be zero for the sake of simplicity and a baseline value of zero is hard-coded into many of the popular GP regression packages.

[3] https://github.com/snesseris/Genetic-Algorithms.

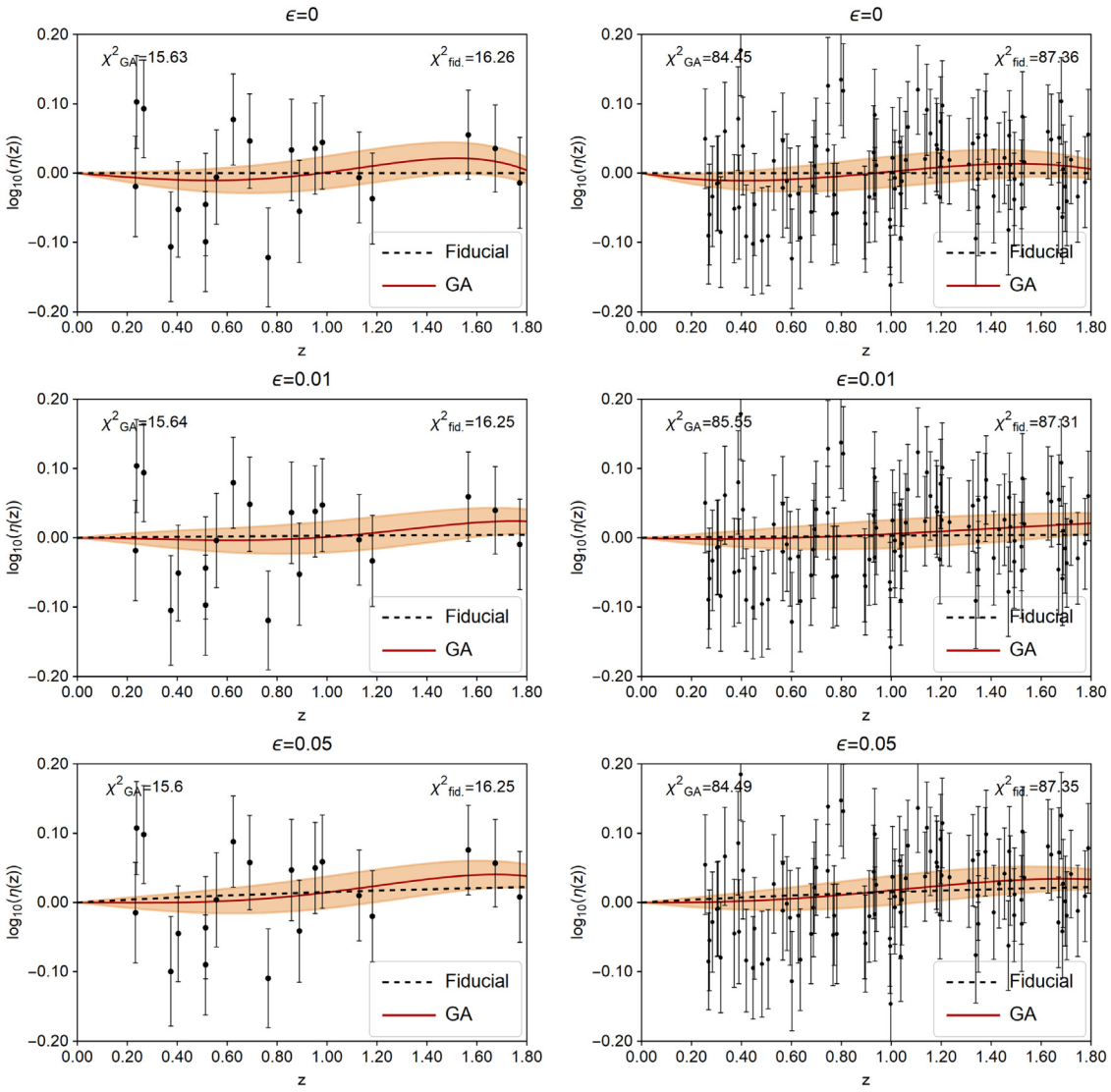

Fig. 2. The GA reconstructions for the 20 lenses (left column) and for 100 lenses (right column). The mocks in the top row were created with $\epsilon = 0$, the ones in the middle row with $\epsilon = 0.01$, while the ones in the bottom row with $\epsilon = 0.05$. The orange shaded regions show the 1σ error for the GA, while the dashed black lines show the fiducial model, $\log_{10}(1+z)^{\epsilon_0}$, in each case.

There are many options for the covariance function, or kernel, $k(z, \tilde{z})$. GPs have been applied to reconstruct a wide variety of functions in cosmology (see e.g. [66–74]) and there is still some debate over the best choice of kernel, as the choice can strongly influence the resulting GP reconstruction. In this work, we choose to proceed by tailoring the kernel to one supporting a reconstruction that finds an increasing trend in redshift, as this is what we expect the fiducial models to produce.

It was found in [75] that the Matérn class of kernels performed best when reconstructing the equation of state of dark energy, $w(z)$, using SNIa data. This class of kernels take the following form [65]:

$$k(z,\tilde{z}) = \sigma_M^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} d(z,\tilde{z})}{\ell} \right)^{\nu} \times K_\nu \left(\frac{\sqrt{2\nu} d(z,\tilde{z})}{\ell} \right), \tag{13}$$

where $d(z, \tilde{z})$ represents the Euclidean distance between the inputs z and $\tilde{z}$, $\Gamma(\nu)$ is the gamma function, K_ν is a modified Bessel function and ν controls the shape of the covariance function, tending to the Gaussian limit as $\nu \to \infty$. The hyperparameters ℓ and σ_M correspond to the approximate length scale over which the function varies and the magnitude of those variations respectively. The choice of a half-integer value for ν is made in order to remove the dependence on the Bessel function [75]. The larger the value of ν, the smoother the resulting GP, although for $\nu \geq 7/2$, the results become hard to distinguish from one another [65]. Overall, this makes $\nu = 5/2$ a good choice.

In the course of our analysis, we found that when a Matérn kernel is used alone, the GP struggles to follow the trend in redshift introduced by the fiducial models of $\epsilon = 0.01$ and $\epsilon = 0.05$. We therefore create a custom kernel that better suits our problem, by adding a dot product kernel to a Matérn ($\nu = 5/2$) kernel. The dot product kernel takes the general form

$$k(z,\tilde{z}) = \sigma_d + z \cdot \tilde{z}, \tag{14}$$

where the hyperparameter σ_d acts on the dot product kernel in a similar way to how σ_M acts on the Matérn kernel. For the Matérn class of kernels, σ_M acts to rescale the GP covariance, whereas for the dot product kernel, σ_d acts as a constant offset of the covariance of the GP. We note that the dot product kernel is non-stationary, meaning that the resulting GP depends not only on the

relative positions of the points, but on their absolute positions. A translation in the input space (i.e. shifting the mock data points in redshift) will therefore result in a different GP prediction from the dot product kernel even if the kernel hyperparameter is kept fixed [76].

We use the Gaussian process regressor provided by the Python package `scikit-learn` [77] to perform our reconstruction of $\log_{10}\eta(z)$ with the custom kernel described above. The package also allows for optimisation of the value of any hyperparameters in the kernel by maximising the log-likelihood of the GP output. We list the optimised values of ℓ, σ_M and σ_d in Table 2 to give an idea of the general behaviour of our custom kernel.

Note that we do not fix these values by hand in the kernel. The only information we give to the kernel is the upper and lower bound that the optimiser explores between for the value of the length scale ℓ. This choice of bound can have an effect on the resulting reconstruction, as there may be multiple values of the hyperparameters that maximise the log-likelihood. However, the optimisation routine will only be able to find one of the maximal values each time the procedure is run. The bounds can therefore be manually shrunk to eliminate all but one of the maximal values of each of the hyperparameters, forcing the GP to use that particular combination.

The value of the hyperparameter ℓ corresponds to the average variation in the z-direction of the data, and is expected to be of order of the average distance between each mock data point. Therefore,to select the upper and lower bounds for the length scale in the Matérn kernel, we considered the approximate average distance between each mock data point in the catalogue, roughly 0.08 in terms of the redshift in the case of 20 lenses. Since it is squared, we then expect the learned length scale to be of the order 10^{-3}. In the case of 100 lenses, the mock data points are spaced closer together, leading us to expect a learned length scale on the order of 10^{-4}. We therefore set the bounds of the Matérn kernel as 10^{-5} and 10^{-1} to safely incorporate these expected values.

The value of σ_M instead corresponds to the typical variation in amplitude of the function, which is expected to be of the order of the average error of the data points *i.e.* ~ 0.05. Finally, the dot product kernel is equivalent to a linear regression in which σ_d is the intercept of the fit. From Eq. (12) it is straightforward to see that $\sigma_d \approx \epsilon_0^2 = O(10^{-4})$. We therefore see that the expected values for σ_d and σ_M fall well within the imposed bounds for the GP hyperparameters. While at first glance this "recipe" used to build the kernel appears somewhat naïve, its validity is confirmed by the optimised hyperparameter values reported in Table 2.

The results of the GP reconstruction using the custom kernel are shown in Fig. 3. The left column shows the reconstructions of $\log_{10}\eta(z)$ for the realistic case of 20 strongly lensed SNIa, and the right column shows the optimistic case of 100 lenses. The mock data in the top row was created with no deviation from ΛCDM or the standard DDR, i.e. $\epsilon = 0.0$, while the middle row shows the mock data for which $\epsilon = 0.01$ and the bottom row $\epsilon = 0.05$.

In the realistic case of 20 lenses, we see that the relatively small number of points does not prevent the GP from correctly recovering the fiducial model (dashed line in all three panels of Fig. 3) to within 1σ for all the fiducial cases.

In the optimistic case of 100 lenses, the error of the GP at high redshift is decreased with respect to the 20 lens case, due to the increased information given to the GP by the additional mock data points. However, for this particular mock dataset realisation, the reconstruction does not recover the fiducial model as well as the 20 lens case, with a slight overestimation of the $\log_{10}\eta(z)$ function at higher redshifts for all three values of ϵ_0. However, even with this overestimation, the reconstruction is again never more than 1σ away from the true fiducial model.

In all cases we report the χ^2 statistic for the fiducial model and the GP reconstruction in the legend of the plots.

Table 2
Values of the kernel hyperparameters after optimisation.

ϵ_0	ℓ	σ_M	σ_d
$N_{\rm lens} = 20$			
0.0	1.00×10^{-3}	3.16×10^{-3}	1.47×10^{-6}
0.01	1.00×10^{-3}	3.16×10^{-3}	1.12×10^{-6}
0.05	1.00×10^{-3}	3.16×10^{-3}	1.04×10^{-6}
$N_{\rm lens} = 100$			
0.0	1.00×10^{-3}	3.16×10^{-3}	1.25×10^{-2}
0.01	1.00×10^{-3}	3.16×10^{-3}	1.13×10^{-2}
0.05	1.00×10^{-3}	3.16×10^{-3}	4.56×10^{-3}

5. Conclusions

In this paper we investigated the possibility of using future observations of strongly lensed Type Ia supernovae to constrain deviations from the standard distance duality relation. A departure from the DDR could be a significant smoking gun for deviations from the standard cosmological model, as it would signal that fundamental assumptions are violated, which we discussed in Section 2.

Such violations are usually investigated in the literature by combining different observations together; this allows the luminosity and angular distances to be reconstructed separately and the function $\eta(z)$, equal to unity in the standard model, to be constrained. In Section 3 we discussed how the observation of strongly lensed SNIa can instead directly provide the two distances at the redshift of the source, and can therefore be used to obtain measurements of $\eta(z)$, avoiding the need to reconstruct the two distances. Notice however that such a measurement is possible only under certain assumptions; one needs to be able to obtain the luminosity distance of the lensed supernovae and remove any possible magnification due to the lens, while the measurement of the angular distance at the source redshift can be obtained from the time delay distance only through the assumption of a flat Universe and if kinematic measurements of the lens galaxy are available.

Other than these assumptions, the use of such observations allows us to obtain our results without any further dependence on the cosmological model, even in the parametric approach that we discuss in Section 4. For this case we find that, as expected, the results strongly depend on the number of systems that will be observed by future surveys; for a realistic number of strongly lensed SNIa ($N_{\rm lens} = 20$) the constraints we obtain on ϵ_0 are of the order of those obtained through the combination of currently available SNIa and BAO surveys, while in our most futuristic case ($N_{\rm lens} = 1000$) bounds on DDR violation obtained through strong lensing are expected to be competitive with those forecast for upcoming LSS surveys.

The results of the Genetic Algorithm reconstruction for both cases of 20 and 100 lenses for $\epsilon_0 = (0.0, 0.01, 0.05)$ were shown in Section 4.2 and in Fig. 2. In all cases the GA was able to correctly recover the underlying fiducial model within the errors.

In Section 4.3, we presented the results of our Gaussian process reconstruction. We reconstructed $\log_{10}\eta(z)$ for the fiducial models of $\epsilon_0 = 0.0$, $\epsilon_0 = 0.01$ and $\epsilon_0 = 0.05$ using both 20 lenses and 100 lenses, finding that the GP was well able to correctly recover the underlying fiducial in the mock data.

In summary, we have shown how strongly lensed SNIa will be a powerful probe of distance measures in cosmology in the upcoming LSST era. We have discussed how these systems are uniquely able to provide measurements of both luminosity and angular diameter distances, allowing excellent constraints to be placed on the distance duality relation. If any deviations from this relation were to be detected it would be an exciting hint at possible new physics easily accessible to other next-generation surveys.

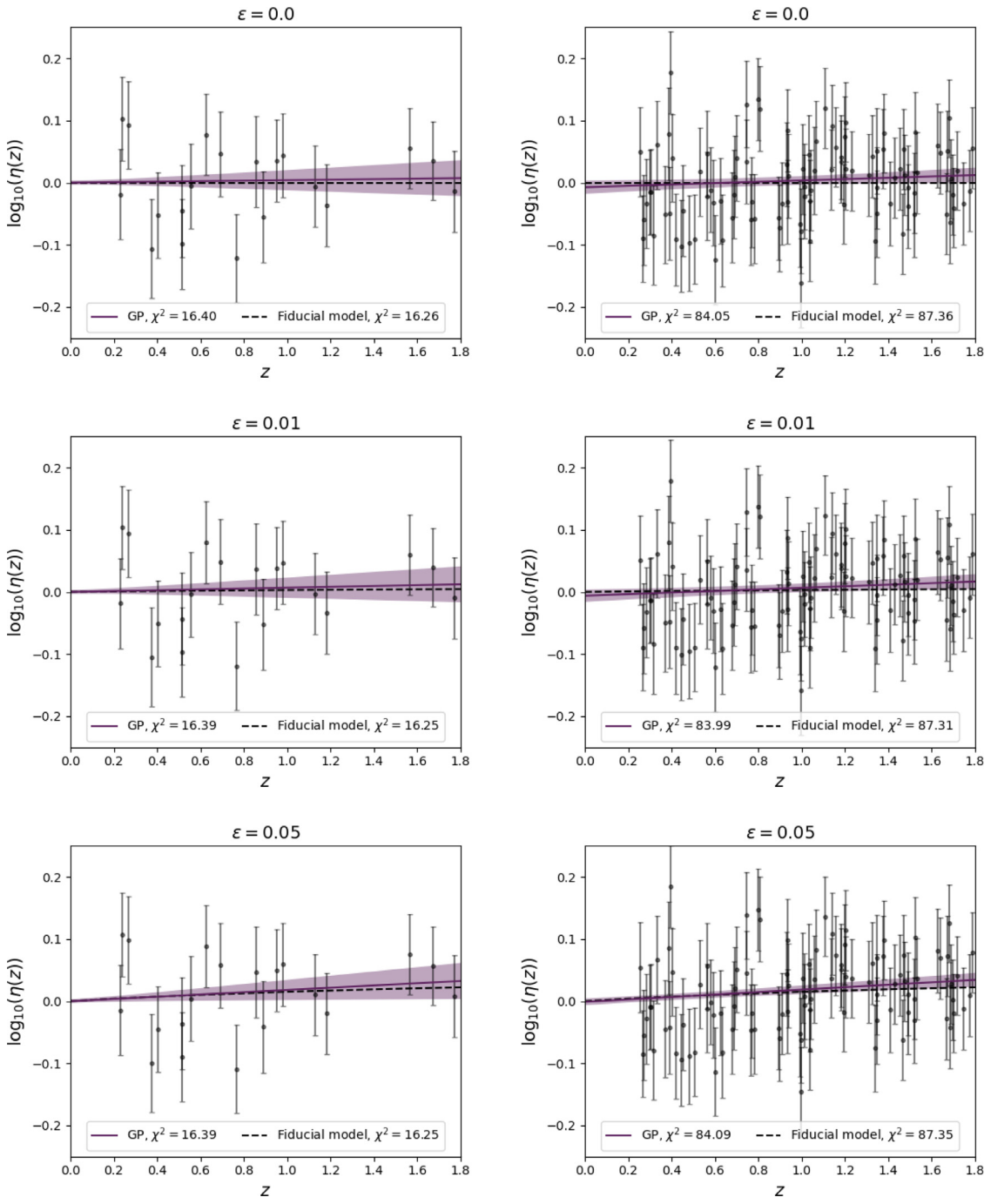

Fig. 3. The GP reconstructions for the 20 lenses (left column) and for 100 lenses (right column). The mocks in the top row were created with $\epsilon_0 = 0$, the ones in the middle row with $\epsilon_0 = 0.01$, while the ones in the bottom row with $\epsilon_0 = 0.05$. The shaded regions show the 1σ error for the GP, while the dashed black lines show the fiducial model, $\log_{10}(1+z)^{\epsilon_0}$, in each case.

CRediT authorship contribution statement

Fabrizio Renzi: Methodology, Software, Formal analysis, Validation, Writing - original draft. **Natalie B. Hogg:** Software, Formal analysis, Writing - original draft, Writing - review & editing. **Matteo Martinelli:** Conceptualization, Software, Formal analysis, Validation, Writing - original draft, Supervision. **Savvas Nesseris:** Conceptualization, Software, Formal analysis, Writing - original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Mike Shengbo Wang for a useful discussion regarding the statistical validity of the mock data. Numerical computations were done on the Hydra HPC Cluster of the Instituto de Física Teórica UAM/CSIC. FR acknowledges support from

the NWO, The Netherlands and the Dutch Ministry of Education, Culture and Science (OCW), The Netherlands, and from the D-ITP consortium, a program of the NWO that is funded by the OCW, The Netherlands. NBH is supported by UK STFC studentship ST/N504245/1. MM has received the support of a fellowship from "la Caixa" Foundation (ID 100010434), with fellowship code LCF/BQ/PI19/11690015, and the support of the Spanish Agencia Estatal de Investigacion through the grant "IFT Centro de Excelencia Severo Ochoa SEV-2016-0597". MM also wants to thank the Big Star Bar for providing a work space and an internet connection during this period of remote work. SN acknowledges support from the research projects PGC2018-094773-B-C32, the Centro de Excelencia Severo Ochoa Program, Spain SEV-2016-059 and the Ramón y Cajal program, Spain through Grant No. RYC-2014-15843.

In this work we made use of the following Python packages that are not mentioned in the text: GetDist [78], a tool for the analysis of MCMC samples, Matplotlib [79], for the realisation of the plots in the paper, NumPy [80], for numerical linear algebra, and SciPy [81], for numerical sampling of the statistical distributions involved in our data analysis.

Appendix. Details of the mock catalogue creation

In this Appendix we describe in more detail the MCMC-like approach used to construct our mock catalogues of $\eta(z_i)$ with $i = 1 \dots N_{\rm lens}$. As discussed in the main text, the methodology followed to generate our mock catalogues has three distinct steps. We start by constructing the probability distribution function (PDF) of the distances involved in the DDR. For a given redshift z_i we start drawing random Gaussian deviates, $\delta D(z_i)$, from a Gaussian distribution of the form:

$$\mathcal{N}(0, \sigma_{D_i} D_i^{\rm true}) \tag{A.1}$$

with $D_i^{\rm true}$ being the true value of the distance for an assumed cosmological model (for this work, $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_m = 0.3$ along with the three chosen values of ϵ_0) at z_i and σ_{D_i}, the observational error on this distance. We then construct the PDF of $D_{i,\rm mock}$ by extracting 10,000 samples from a Gaussian distribution with mean $\bar{D}_i = D_i^{\rm true} + \delta D_i$ and standard deviation σ_{D_i} *i.e.*

$$D_{i,\rm mock} = \mathcal{N}(\bar{D}_i, \sigma_{D_i} \bar{D}_i)\,. \tag{A.2}$$

A comparison of the true and mock PDFs is plotted in Fig. A.4 for the angular diameter distance.

With the PDFs of d_A, d_Δ and μ in hand, we proceed in an MCMC-like fashion. We assume the PDFs of d_A, $d_{\Delta t}$ and μ to be the posteriors of a hypothetical MCMC run with the three distances as independent parameters, so that at each redshift z_i, each triplet $\{d_{A,n}, d_{\Delta t,n}, \mu_n \mid n = 1 \dots 10^4\}$ constitutes a sample of an MCMC chain. Therefore at each n we combine the triplet values, using Eq. (6) to obtain a sample of the posterior of $(\log_{10}\eta(z_i))_n$, *i.e.* we treat $\log_{10}\eta(z_i)$ as a derived parameter of the MCMC. We apply this procedure to all 10,000 samples to construct the distribution of $\log_{10}\eta(z_i)$.

A comparison of the true and mock PDFs of $\log_{10}\eta(z_i)$ is plotted in Fig. A.5 while in Fig. A.6 we show a sample mock for $N_{\rm lenses} = 20$. As we can see from Fig. A.5, the assumption $\log_{10}\eta(z_i) \approx \mathcal{N}(0, \sigma_{\log_{10}\eta(z_i)})$ is very much in agreement with the numerical distributions of $\log_{10}\eta(z_i)$ constructed with our methodology.

From the PDFs of $\log_{10}\eta(z)$, we can also perform some sanity checks. First of all, assuming that $\log_{10}\eta(z) = const$, we can multiply the PDFs of all the $\log_{10}\eta(z_i)$ to obtain a combined posterior and therefore the mock best fit for $\log_{10}\eta(z)$. We show the combined PDFs of $\log_{10}\eta$ for two mocks of $N_{\rm lens} = 20$, 100 plotted against the combined true PDFs of $\log_{10}\eta$ for $N_{\rm lens} = 20$ in Fig. A.7. While this best-fit value will not be as accurate as the one obtained from a full MCMC sampling, it can signal inconsistency in the mock dataset without the need for a complex analysis. Furthermore, we can construct the χ^2 distribution, testing 10,000 realisations of a mock against the hypothesis $\log_{10}\eta(z_i) \approx \mathcal{N}(0, \sigma_{\log_{10}\eta(z_i)})$ as an additional sanity check. In Fig. A.8 we show the comparison between the distribution of χ^2 values for the 20 lens mock dataset and the theoretical χ^2 distribution for 20 degrees of freedom. We can see that the mock distribution follows the theoretical one extremely well.

So far, we found that our mocks are generally within the 1σ bounds of the true combined PDF, even though a significant deviation from the fiducial might happen in correspondence with the higher/lower tail of the χ^2 distribution for the mocks. In summary, this procedure has two main advantages: (1) it exposes the PDFs of the data points of the mocks, allowing them to be used for sanity checks and eventually for a full MCMC sampling similar to what has been done for the analysis of the H0LiCOW lenses (see e.g. [23]) and (2) it allows us to reconstruct the errors of the data points directly from their posteriors, removing any assumptions coming from the standard error propagation formula.

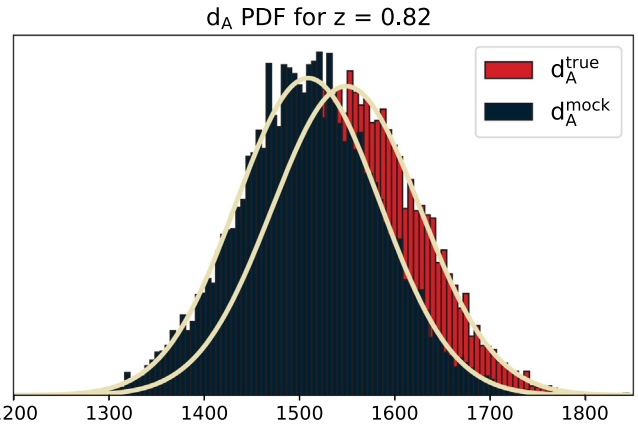

Fig. A.4. Comparison of the PDFs obtained from the truth value of the angular diameter distance $d_A^{\rm true}$ and the corresponding distribution for the mock value $d_A^{\rm mock}$ at fixed redshift. The solid lines show the corresponding theoretical Gaussian PDFs with $\sigma_{D_i} = 0.05$.

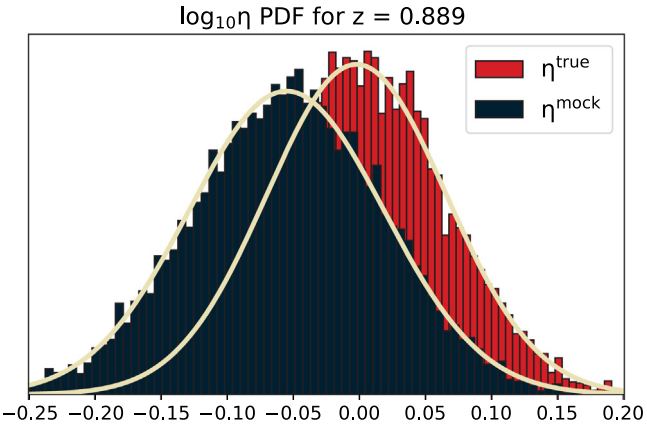

Fig. A.5. Comparison of the PDFs obtained from the truth value of the DDR function $\log_{10}\eta^{\rm true}$ and the corresponding distribution for the mock value $\log_{10}\eta^{\rm mock}$ at fixed redshift. The solid lines show the corresponding theoretical Gaussian PDFs with $(\sigma_{\log_{10}\eta})_{\rm mock} = 0.068$ and $(\sigma_{\log_{10}\eta})_{\rm true} = 0.074$.

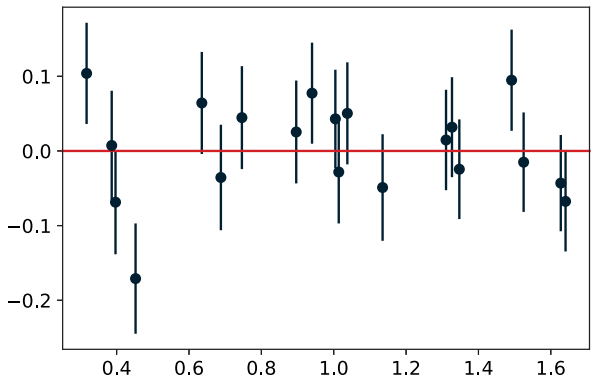

Fig. A.6. A sample mock for the 20 lenses catalogue constructed with the methodology described in Appendix.

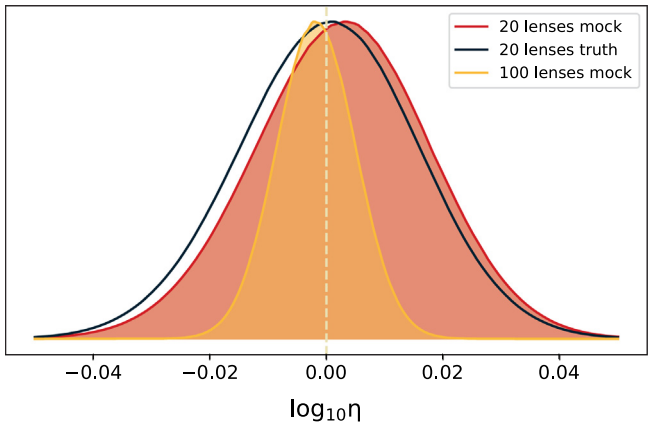

Fig. A.7. The combined PDFs of $\log_{10}\eta$ for mocks of $N_{\rm lens} = 20$, 100 plotted against the combined true PDFs of $\log_{10}\eta$ for $N_{\rm lens} = 20$.

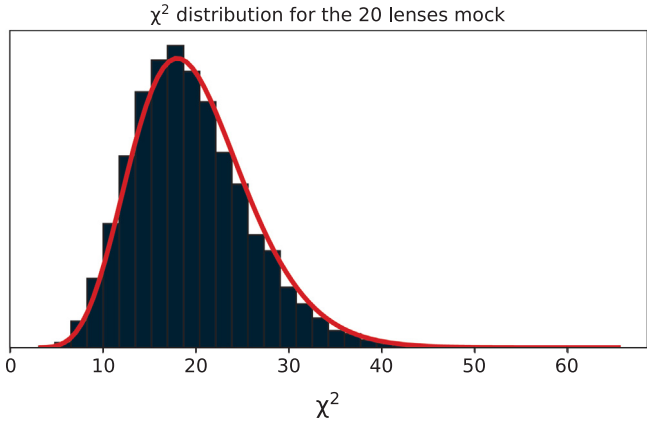

Fig. A.8. Comparison between the distribution of χ^2 values for the 20 lens mock dataset (dark histogram) with the theoretical χ^2 distribution for 20 degrees of freedom (red solid line).

References

[1] S. Suyu, et al., HOLISMOKES – I. Highly optimised lensing investigations of supernovae, microlensing objects, and kinematics of ellipticals and spirals, 2020, arXiv:2002.08378.

[2] P. Schneider, J. Ehlers, E. Falco, Gravitational Lenses, Springer, 1992.

[3] P. Schneider, C.S. Kochanek, J. Wambsganss, Gravitational Lensing: Strong, Weak and Micro, Springer, 2006.

[4] E.F. Borra, Detection of gravitational lenses and measurement of time delays from classical electromagnetic radiation fluctuations, Mon. Not. R. Astron. Soc. 289 (1997) 660–664, URL: http://arxiv.org/abs/astro-ph/9704074. arXiv:Arxiv:astro-ph/9704074v1.

[5] E.F. Borra, Observations of time delays in gravitational lenses from intensity fluctuations: The coherence function, Mon. Not. R. Astron. Soc. 389 (2008) 364–370, URL: http://arxiv.org/abs/0806.2252. arXiv:Arxiv:0806.2252v1.

[6] T. Treu, Strong lensing by galaxies, Ann. Rev. Astron. Astrophys. 48 (2010) 87–125, http://dx.doi.org/10.1146/annurev-astro-081309-130924, arXiv:1003.5567.

[7] T. Treu, P.J. Marshall, Time delay cosmography, Astron. Astrophys. Rev. 24 (1) (2016) 11, http://dx.doi.org/10.1007/s00159-016-0096-8, arXiv:1605.05333.

[8] S.H. Suyu, T.-C. Chang, F. Courbin, T. Okumura, Cosmological distance indicators, Space Sci. Rev. 214 (5) (2018) 91, http://dx.doi.org/10.1007/s11214-018-0524-3, arXiv:1801.07262.

[9] B. Shiralilou, M. Martinelli, G. Papadomanolakis, S. Peirone, F. Renzi, A. Silvestri, Strong lensing time delay constraints on dark energy: a forecast, J. Cosmol. Astropart. Phys. 04 (2020) 057, http://dx.doi.org/10.1088/1475-7516/2020/04/057, arXiv:1910.03566.

[10] S. Birrer, et al., TDCOSMO IV: Hierarchical time-delay cosmography – joint inference of the hubble constant and galaxy density profiles, 2020, arXiv:2007.02941.

[11] S. Suyu, P. Marshall, M. Auger, S. Hilbert, R. Blandford, L. Koopmans, C. Fassnacht, T. Treu, Dissecting the gravitational lens B1608+656. II. Precision measurements of the hubble constant, spatial curvature, and the dark energy equation of state, Astrophys. J. 711 (2010) 201–221, http://dx.doi.org/10.1088/0004-637X/711/1/201, arXiv:0910.2773.

[12] D. Paraficz, J. Hjorth, Gravitational lenses as cosmic rulers: density of dark matter and dark energy from time delays and velocity dispersions, Astron. Astrophys. 507 (2009) L49, http://dx.doi.org/10.1051/0004-6361/200913307, arXiv:0910.5823.

[13] I. Jee, E. Komatsu, S.H. Suyu, Measuring angular diameter distances of strong gravitational lenses, J. Cosmol. Astropart. Phys. 11 (2015) 033, http://dx.doi.org/10.1088/1475-7516/2015/11/033, arXiv:1410.7770.

[14] S. Suyu, et al., H0LiCOW – I. H0 Lenses in COSMOGRAIL's Wellspring: program overview, Mon. Not. R. Astron. Soc. 468 (3) (2017) 2590–2604, http://dx.doi.org/10.1093/mnras/stx483, arXiv:1607.00017.

[15] T. Treu, et al., Dark energy with gravitational lens time delays, in: Community Summer Study 2013: Snowmass on the Mississippi, 2013, arXiv:1306.1272.

[16] K. Liao, et al., Strong lens time delay challenge: II. Results of TDC1, Astrophys. J. 800 (1) (2015) 11, http://dx.doi.org/10.1088/0004-637X/800/1/11, arXiv:1409.1254.

[17] I. Jee, E. Komatsu, S.H. Suyu, D. Huterer, Time-delay cosmography: Increased leverage with angular diameter distances, J. Cosmol. Astropart. Phys. 04 (2016) 031, http://dx.doi.org/10.1088/1475-7516/2016/04/031, arXiv:1509.03310.

[18] K.C. Wong, et al., H0LiCOW – IV. Lens mass model of HE 0435−1223 and blind measurement of its time-delay distance for cosmology, Mon. Not. R. Astron. Soc. 465 (4) (2017) 4895–4913, http://dx.doi.org/10.1093/mnras/stw3077, arXiv:1607.01403.

[19] V. Bonvin, et al., H0LiCOW – V. New COSMOGRAIL time delays of HE 0435−1223: H_0 to 3.8 per cent precision from strong lensing in a flat ΛCDM model, Mon. Not. R. Astron. Soc. 465 (4) (2017) 4914–4930, http://dx.doi.org/10.1093/mnras/stw3006, arXiv:1607.01790.

[20] O. Tihhonova, et al., H0LiCOW VIII. A weak-lensing measurement of the external convergence in the field of the lensed quasar HE 0435−1223, Mon. Not. R. Astron. Soc. 477 (4) (2018) 5657–5669, http://dx.doi.org/10.1093/mnras/sty1040, arXiv:1711.08804.

[21] S. Birrer, et al., H0LiCOW - IX. Cosmographic analysis of the doubly imaged quasar SDSS 1206+4332 and a new measurement of the Hubble constant, Mon. Not. R. Astron. Soc. 484 (2019) 4726, http://dx.doi.org/10.1093/mnras/stz200, arXiv:1809.01274.

[22] C.E. Rusu, et al., H0LiCOW XII. Lens mass model of WFI2033-4723 and blind measurement of its time-delay distance and H_0, 2019, http://dx.doi.org/10.1093/mnras/stz3451, arXiv:1905.09338.

[23] K.C. Wong, et al., H0LiCOW XIII. A 2.4% measurement of H_0 from lensed quasars: 5.3σ tension between early and late-universe probes, 2019, http://dx.doi.org/10.1093/mnras/stz3094, arXiv:1907.04869.

[24] G.C.-F. Chen, et al., A SHARP view of H0LiCOW: H_0 from three time-delay gravitational lens systems with adaptive optics imaging, Mon. Not. R. Astron. Soc. 490 (2) (2019) 1743–1773, http://dx.doi.org/10.1093/mnras/stz2547, arXiv:1907.02533.

[25] S. Refsdal, On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect, Mon. Not. R. Astron. Soc. 128 (4) (1964) 307–310, http://dx.doi.org/10.1093/mnras/128.4.307.

[26] P.L. Kelly, et al., Multiple images of a highly magnified supernova formed by an early-type cluster galaxy lens, Science 347 (2015) 1123, http://dx.doi.org/10.1126/science.aaa3350, arXiv:1411.6009.

[27] A. Goobar, et al., IPTF16geu: A multiply imaged, gravitationally lensed type Ia supernova, Science 356 (2017) 291–295, http://dx.doi.org/10.1126/science.aal2729, arXiv:1611.00014.

[28] J.R. Pierel, S.A. Rodney, Turning gravitationally lensed supernovae into cosmological probes, Astrophys. J. 876 (2) (2019) 107, http://dx.doi.org/10.3847/1538-4357/ab164a, arXiv:1902.01260.

[29] LSST Science Collaboration, P.A. Abell, J. Allison, S.F. Anderson, J.R. Andrew, et al., LSST science book, version 2.0, 2009, arXiv e-prints. arXiv:0912.0201.

[30] P. Marshall, et al., (LSST), Science-driven optimization of the LSST observing strategy, 2017, http://dx.doi.org/10.5281/zenodo.842713, arXiv:1708.04058.

[31] D.A. Goldstein, P.E. Nugent, A. Goobar, Rates and properties of supernovae strongly gravitationally lensed by elliptical galaxies in time-domain imaging surveys, Astrophys. J. Suppl. 243 (1) (2019) 6, http://dx.doi.org/10.3847/1538-4365/ab1fe0, arXiv:1809.10147.

[32] S. Huber, et al., (LSST Dark Energy Science), Strongly lensed SNe Ia in the era of LSST: observing cadence for lens discoveries and time-delay measurements, Astron. Astrophys. 631 (2019) A161, http://dx.doi.org/10.1051/0004-6361/201935370, arXiv:1903.00510.

[33] M. Oguri, Y. Kawano, Gravitational lens time delays for distant supernovae: break the degeneracy between radial mass profiles and the hubble constant, Mon. Not. R. Astron. Soc. 338 (2003) L25–L29, http://dx.doi.org/10.1046/j.1365-8711.2003.06290.x, arXiv:astro-ph/0211499.

[34] D.A. Yahalomi, P.L. Schechter, J. Wambsganss, A quadruply lensed SN Ia: Gaining a time-delay...losing a standard candle, 2017, arXiv:1711.07919.

[35] M. Foxley-Marrable, T.E. Collett, G. Vernardos, D.A. Goldstein, D. Bacon, The impact of microlensing on the standardization of strongly lensed type ia supernovae, Mon. Not. R. Astron. Soc. 478 (4) (2018) 5081–5090, http://dx.doi.org/10.1093/mnras/sty1346, arXiv:1802.07738.

[36] V. Bonvin, O. Tihhonova, M. Millon, J. Chan, E. Savary, S. Huber, F. Courbin, Impact of the 3D source geometry on time-delay measurements of lensed type-Ia Supernovae, Astron. Astrophys. 621 (2019) A55, http://dx.doi.org/10.1051/0004-6361/201833405, arXiv:1805.04525.

[37] R. Holanda, V. Busti, J. Alcaniz, Probing the cosmic distance duality with strong gravitational lensing and supernovae Ia data, J. Cosmol. Astropart. Phys. 02 (2016) 054, http://dx.doi.org/10.1088/1475-7516/2016/02/054, arXiv:1512.02486.

[38] R. Holanda, V. Busti, F. Lima, J. Alcaniz, Probing the distance-duality relation with high-z data, J. Cosmol. Astropart. Phys. 09 (2017) 039, http://dx.doi.org/10.1088/1475-7516/2017/09/039, arXiv:1611.09426.

[39] A. Rana, D. Jain, S. Mahajan, A. Mukherjee, R. Holanda, Probing the cosmic distance duality relation using time delay lenses, J. Cosmol. Astropart. Phys. 07 (2017) 010, http://dx.doi.org/10.1088/1475-7516/2017/07/010, arXiv:1705.04549.

[40] I.M.H. Etherington, The definition of distance in general relativity, Phil. Mag. 15 (1933) 761–773, http://dx.doi.org/10.1080/14786443309462220.

[41] G.F.R. Ellis, On the definition of distance in general relativity: I. M. H. Etherington (Philosophical Magazine ser. 7, vol. 15, 761 (1933)), Gen. Relativity Gravitation 39 (7) (2007) 1047–1052, http://dx.doi.org/10.1007/s10714-006-0355-5.

[42] F. Hehl, P. Von Der Heyde, G. Kerlick, J. Nester, General relativity with spin and torsion: Foundations and prospects, Rev. Modern Phys. 48 (1976) 393–416, http://dx.doi.org/10.1103/RevModPhys.48.393.

[43] F.W. Hehl, J. McCrea, E.W. Mielke, Y. Ne'eman, Metric affine gauge theory of gravity: Field equations, Noether identities, world spinors, and breaking of dilation invariance, Phys. Rep. 258 (1995) 1–171, http://dx.doi.org/10.1016/0370-1573(94)00111-F, arXiv:gr-qc/9402012.

[44] R.T. Hammond, Torsion gravity, Rep. Progr. Phys. 65 (5) (2002) 599–649, http://dx.doi.org/10.1088/0034-4885/65/5/201.

[45] E. Gabrielli, K. Huitu, S. Roy, Photon propagation in magnetic and electric fields with scalar/pseudoscalar couplings: A new look, Phys. Rev. D 74 (2006) 073002, http://dx.doi.org/10.1103/PhysRevD.74.073002, arXiv:hep-ph/0604143.

[46] L.T. Santana, M.O. Calvão, R.R.R. Reis, B.B. Siffert, How does light move in a generic metric-affine background? Phys. Rev. D 95 (6) (2017) 061501, http://dx.doi.org/10.1103/PhysRevD.95.061501, arXiv:1703.10871.

[47] B.A. Bassett, M. Kunz, Cosmic distance-duality as a probe of exotic physics and acceleration, Phys. Rev. D 69 (2004) 101305, http://dx.doi.org/10.1103/PhysRevD.69.101305, arXiv:astro-ph/0312443.

[48] A. Avgoustidis, L. Verde, R. Jimenez, Consistency among distance measurements: transparency, BAO scale and accelerated expansion, J. Cosmol. Astropart. Phys. 2009 (06) (2009) http://dx.doi.org/10.1088/1475-7516/2009/06/012, arXiv:0902.2006.

[49] A. Avgoustidis, C. Burrage, J. Redondo, L. Verde, R. Jimenez, Constraints on cosmic opacity and beyond the standard model physics from cosmological distance measurements, J. Cosmol. Astropart. Phys. 2010 (10) (2010) http://dx.doi.org/10.1088/1475-7516/2010/10/024, arXiv:1004.2053.

[50] P. Astier, et al., Extending the supernova hubble diagram to z $\sim$ 1.5 with the euclid space mission, Astron. Astrophys. 572 (2014) A80, http://dx.doi.org/10.1051/0004-6361/201423551, arXiv:1409.8562.

[51] D. Coe, L.A. Moustakas, Cosmological constraints from gravitational lens time delays, Astrophys. J. 706 (1) (2009) 45–59, http://dx.doi.org/10.1088/0004-637x/706/1/45.

[52] E.V. Linder, Lensing time delays and cosmological complementarity, Phys. Rev. D 84 (2011) 123529, http://dx.doi.org/10.1103/PhysRevD.84.123529, arXiv:1109.2592.

[53] J. Torrado, A. Lewis, Cobaya: Code for Bayesian analysis of hierarchical physical models, 2020, arXiv e-prints. arXiv:2005.05290.

[54] A. Avgoustidis, C. Burrage, J. Redondo, L. Verde, R. Jimenez, Constraints on cosmic opacity and beyond the standard model physics from cosmological distance measurements, J. Cosmol. Astropart. Phys. 1010 (2010) 024, http://dx.doi.org/10.1088/1475-7516/2010/10/024, arXiv:1004.2053.

[55] M. Martinelli, et al., (EUCLID), Euclid: Forecast constraints on the cosmic distance duality relation with complementary external probes, 2020, arXiv:2007.16153.

[56] N.B. Hogg, M. Martinelli, S. Nesseris, Constraints on the distance duality relation with standard sirens, 2020, arXiv:2007.14335.

[57] C. Bogdanos, S. Nesseris, Genetic algorithms and supernovae type ia analysis, J. Cosmol. Astropart. Phys. 05 (2009) 006, http://dx.doi.org/10.1088/1475-7516/2009/05/006, arXiv:0903.2805.

[58] S. Nesseris, J. Garcia-Bellido, A new perspective on dark energy modeling via genetic algorithms, J. Cosmol. Astropart. Phys. 1211 (2012) 033, http://dx.doi.org/10.1088/1475-7516/2012/11/033, arXiv:1205.0364.

[59] S. Nesseris, A. Shafieloo, A model independent null test on the cosmological constant, Mon. Not. R. Astron. Soc. 408 (2010) 1879–1885, http://dx.doi.org/10.1111/j.1365-2966.2010.17254.x, arXiv:1004.0960.

[60] S. Nesseris, J. Garcia-Bellido, Comparative analysis of model-independent methods for exploring the nature of dark energy, Phys. Rev. D88 (6) (2013) 063521, http://dx.doi.org/10.1103/PhysRevD.88.063521, arXiv:1306.4885.

[61] D. Sapone, E. Majerotto, S. Nesseris, Curvature versus distances: Testing the FLRW cosmology, Phys. Rev. D 90 (2) (2014) 023012, http://dx.doi.org/10.1103/PhysRevD.90.023012, arXiv:1402.2236.

[62] R. Arjona, S. Nesseris, What can machine learning tell us about the background expansion of the universe?, 2019, arXiv e-prints. arXiv:1910.01529.

[63] R. Arjona, S. Nesseris, Hints of dark energy anisotropic stress using machine learning, 2020, arXiv e-prints. arXiv:2001.11420.

[64] R. Arjona, Machine learning meets the redshift evolution of the CMB temperature, 2020, arXiv:2002.12700.

[65] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.

[66] T. Holsclaw, U. Alam, B. Sansó, H. Lee, K. Heitmann, S. Habib, D. Higdon, Nonparametric reconstruction of the dark energy equation of state, Phys. Rev. D 82 (10) (2010) 103502, http://dx.doi.org/10.1103/PhysRevD.82.103502, arXiv:1009.5443.

[67] T. Holsclaw, U. Alam, B. Sansó, H. Lee, K. Heitmann, S. Habib, D. Higdon, Nonparametric dark energy reconstruction from supernova data, Phys. Rev. Lett. 105 (24) (2010) 241302, http://dx.doi.org/10.1103/PhysRevLett.105.241302, arXiv:1011.3079.

[68] T. Holsclaw, U. Alam, B. Sansó, H. Lee, K. Heitmann, S. Habib, D. Higdon, Nonparametric reconstruction of the dark energy equation of state from diverse data sets, Phys. Rev. Lett. 84 (8) (2011) 083501, http://dx.doi.org/10.1103/PhysRevD.84.083501, arXiv:1104.2041.

[69] A. Shafieloo, A.G. Kim, E.V. Linder, Gaussian process cosmography, Phys. Rev. D 85 (12) (2012) 123530, http://dx.doi.org/10.1103/PhysRevD.85.123530, arXiv:1204.2272.

[70] M. Seikel, C. Clarkson, M. Smith, Reconstruction of dark energy and expansion dynamics using Gaussian processes, J. Cosmol. Astropart. Phys. 6 (2012) 036, http://dx.doi.org/10.1088/1475-7516/2012/06/036, arXiv:1204.2832.

[71] M.-J. Zhang, H. Li, Gaussian processes reconstruction of dark energy from observational data, Eur. Phys. J. C 78 (2018) 460, http://dx.doi.org/10.1140/epjc/s10052-018-5953-3, arXiv:1806.02981.

[72] M. Martinelli, N.B. Hogg, S. Peirone, M. Bruni, D. Wands, Constraints on the interacting vacuum – geodesic CDM scenario, Mon. Not. R. Astron. Soc. 488 (3) (2019) 3423–3438, http://dx.doi.org/10.1093/mnras/stz1915, arXiv:1902.10694.

[73] F. Gerardi, M. Martinelli, A. Silvestri, Reconstruction of the dark energy equation of state from latest data: the impact of theoretical priors, J. Cosmol. Astropart. Phys. 1907 (2019) 042, http://dx.doi.org/10.1088/1475-7516/2019/07/042, arXiv:1902.09423.

[74] N.B. Hogg, M. Bruni, R. Crittenden, M. Martinelli, S. Peirone, Latest evidence for a late time vacuum – geodesic CDM interaction, Phys. Dark Univ. 29 (2020) 100583, http://dx.doi.org/10.1016/j.dark.2020.100583, arXiv:2002.10449.

[75] M. Seikel, C. Clarkson, Optimising Gaussian processes for reconstructing dark energy dynamics from supernovae, 2013, arXiv:1311.6678.

[76] D. Duvenaud, Automatic Model Construction with Gaussian Processes (Ph.D. thesis), University of Cambridge, 2014.

[77] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[78] A. Lewis, Getdist: a python package for analysing Monte Carlo samples, 2019, arXiv:1910.13970.

[79] J.D. Hunter, Matplotlib: A 2D graphics environment, Comput. Sci. Eng. 9 (3) (2007) 90–95, http://dx.doi.org/10.1109/MCSE.2007.55.

[80] C.R. Harris, K. Jarrod Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T.E. Oliphant, Array programming with NumPy, Nature 585 (2020) 357–362.

[81] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E.W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, S... Contributors, Scipy 1.0: Fundamental algorithms for scientific computing in python, Nat. Methods 17 (2020) 261–272, http://dx.doi.org/10.1038/s41592-019-0686-2.