Fables of reconstruction: controlling bias in the dark energy equation of state

Robert G. Crittenden,[1] Gong-Bo Zhao,[1] Levon Pogosian,[2] Lado Samushia,[1,3] and Xinmin Zhang[4,5]

[1] *Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*
[2] *Department of Physics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada*
[3] *Abastumani Astrophysical Observatory, Ilia State University, 2A kazbegi Ave, Tbilisi, GE-0160, Georgia*
[4] *Theoretical Physics Division, Institute of High Energy Physics,*
Chinese Academy of Science, P.O.Box 918-4, Beijing 100049, P.R.China
[5] *Theoretical Physics Center for Science Facilities (TPCSF),*
Chinese Academy of Science, Beijing 100049, P.R.China

We develop an efficient, non-parametric Bayesian method for reconstructing the time evolution of the dark energy equation of state $w(z)$ from observational data. Of particular importance is the choice of prior, which must be chosen carefully to minimise variance and bias in the reconstruction. Using a principal component analysis, we show how a correlated prior can be used to create a smooth reconstruction and also avoid bias in the mean behaviour of $w(z)$. We test our method using Wiener reconstructions based on Fisher matrix projections, and also against more realistic MCMC analyses of simulated data sets for Planck and a future space-based dark energy mission. While the accuracy of our reconstruction depends on the smoothness of the assumed $w(z)$, the relative error for typical dark energy models is $\lesssim 10\%$ out to redshift $z = 1.5$.

I. INTRODUCTION

The nature of dark energy (DE), the source driving the acceleration of the universe in the framework of general relativity, has remained a mystery since the acceleration was first discovered [1]. With the accumulating observational data, including supernovae (SN), cosmic microwave background radiation (CMB) and large scale structure (LSS) data, we hope to understand whether DE is a cosmological constant or whether it is dynamical. Dynamical dark energy is characterised by having an equation-of-state $w \neq -1$, and which in general could be a function of redshift $w(z)$ or, equivalently, the scale factor $w(a)$. One of the key goals in dark energy observations is to constrain the equation of state and its evolution.

The time evolution of w can, in principle, be reconstructed from data using either parametric or non-parametric methods (e.g. [2]). Non-parametric methods have the advantage that they do not assume an *ad hoc* functional form of $w(z)$, which could lead one to miss evidence for a different kind of evolution. Unfortunately, non-parametric methods generically contain many degrees of freedom which can lead to parameter degeneracies, making them difficult to explore.

Many efforts have been made in the literature to develop non-parametric methods, which can vary greatly depending on the nature of the data and the choices of the functions to be reconstructed. One particular focus has been using SN and other standard candle data to reconstruct various ways of paramterizing the background cosmology; the techniques either fit directly to SN magnitudes or their luminosity distances $D_L(z)$, or to more indirect quantities such as dark energy density, $\rho_{DE}(z)$, the expansion history, $H(z)$, and the equation of state $w(z)$, or much more indirect quantities such as the potential of the DE scalar field(s) [3–13].

Unfortunately, many of these techniques are specific to luminosity distance measurements, or more generally to measures of the background expansion rate, making it difficult to include other types of data. The background measurements can be limited by our knowledge of other parameters, in particular the value of the dark energy density today, Ω_{DE} [5, 8]. Thus, it is useful to include measurements which could be sensitive to the growth of structure; such measurements, including weak lensing, galaxy clustering and the integrated Sachs-Wolfe effect (ISW), can help to break the degeneracies and can potentially be the basis for consistency tests of the DE paradigm [14, 15].

When comparing background and structure evolution data, it is useful to focus on more indirect parameterisations like the equation of state, $w(z)$. Non-parametric approaches typically would expand $w(z)$ in terms of bins or basis functions. A powerful tool is the principal component approach, which uses forecasts of the future experiments in the form of Fisher matrix projections to find linear combinations of these basis functions which are independent and well determined [13, 16–22].

One reconstruction approach is to make a truncation of the best constrained principal components, but the number of modes kept is an open question. This often is decided without reference to the data, potentially missing evidence because it is not expected. This simple truncation can lead to significant biases and unrealistically small errors in the reconstruction where the data are poor or are absent entirely [16]. It is clear that some bias in reconstructions is inevitable, particularly when one attempts to reconstruct $w(z)$ in regimes where the data are poor or are absent entirely. However, it is worth investigating whether a better reconstruction method can be found.

In a fully Bayesian approach, one can explicitly specify a prior on the behaviour of $w(z)$, which then determines which modes are kept in the reconstruction, based on where the evidence from the data outweighs the prior. But the crucial question is then to understand the best

way to specify priors in order to minimise the bias, while at the same time reducing the impact of data noise in the reconstruction. Here we extend previous work, where we assumed that $w(z)$ could be treated as a Gaussian random field with a proposed correlation function [17]; this correlation function effectively enforces a smoothness on the reconstruction. Our approach is close in spirit to recently proposed Gaussian Process (GP) method [23, 24], which applied a similar prior to reconstructions of $w(z)$ using SN data, though there are important differences as described below.

In this paper we discuss how to choose a prior and implement it in a computationally efficient method to reconstruct $w(z)$ non-parametrically from the observed data. As we will show, it is very accurate (with relative error $\lesssim 10\%$ for a range of dark energy models), computationally cheap and fast, straightforward to implement, and can be used to analyse any kind of cosmological data. After discussing general issues regarding the choice of prior correlation functions, we test our reconstruction accuracy using the simulated mock data for Planck [27] and a future space-based mission. We then conclude with a discussion of how to extend this method to other parameterization and how to calculate more realistic priors.

II. RECONSTRUCTION METHODS

Reconstruction methods in general have been well explored; here we focus on a Bayesian method, so to make our assumptions as explicit as possible. Whatever the method, one needs some metric to evaluate the quality of the reconstructions. One natural choice of metric, or risk function, is the mean squared error (MSE) of the binned equation of state, between the true model and the reconstructed one:

$$\text{MSE} \equiv \sum_i (w_i^{\text{true}} - w_i^{\text{recon}})^2. \qquad (1)$$

The MSE is not the only possible choice, but it is common and we adopt it here.

Our Bayesian method must assume an *a priori* probability distribution in the space of models. In the absence of data, the reconstruction will return *some* fiducial model, $\mathbf{w}^{\text{fid}}$, which is the peak of the prior distribution; if the fiducial model differs from the true one, the mean reconstructed model will be biased, where $\mathbf{w}^{\text{mean}} \equiv \langle \mathbf{w}^{\text{recon}} \rangle$ is averaged over the ensemble of possible data consistent with the true model. The ensemble average of the MSE can be shown to be the sum of two terms: the variance with which the mean model will be reconstructed and the bias between the true model and the mean one:

$$\langle \text{MSE} \rangle = \sum_i \langle (w_i^{\text{mean}} - w_i^{\text{recon}})^2 \rangle + (w_i^{\text{true}} - w_i^{\text{mean}})^2. \quad (2)$$

The challenge of reconstruction is to keep both types of errors small; a stronger prior will reduce the variance of the reconstruction, but will increase the bias if the fiducial and true models do not match.

Here we focus on a fully Bayesian reconstruction method. In it, we simply assume a prior probability distribution on the function we wish to reconstruct, and our reconstructed function is the maximum of the posterior probability distribution, which is the product of the data likelihood and the prior. Thus, the reconstruction method is entirely specified by the definition of the prior and this makes our reconstruction assumptions fully explicit.

We try to choose the prior to minimise the bias; however, this is intrinsically subjective because it depends on the expected 'true' model. In principle, this should be determined by theoretical considerations and/or previous observations. Given that we are making projections based on the most optimal future data, previous observations are not expected to have significant impact, making theoretical considerations key. Theoretical priors can be made by considering ensembles of possible models, either using a combination of parametric models for the equation of state, or by using physical models; e.g., by putting a prior on the space of possible quintessence potentials. However, it is clear that this choice could vary wildly between theorists.

One of the problems of using a non-parametric models is the sensitivity of answers to the binning scheme assumed. If few bins are assumed, then the answers are well determined, but the resulting reconstruction has unphysical discrete structures resulting from the binning. But if many bins are assumed, some degrees of freedom (d.o.f.) will not be constrained by the data, resulting in large variance and slow convergence of Monte Carlo Markov chain (MCMC) methods. An advantage of the Bayesian approach is that we can use many bins, but then constrain the residual degrees of freedom by imposing a prior on the space of functions. This prior tends to make the functions smoother, reflecting our preconceptions of how the equation of state should evolve.

Another common approach to reconstructions is to simply smooth the data, e.g. by using some implementation of a low-pass filter, on the assumption that the models will not have high frequency variations. This equates to an infinitely strong prior, which does not allow high frequency modes in the reconstruction, no matter how strong the evidence for them may be. While this is possible to implement with the methods we describe, it seems more reasonable to set some strong but finite prior based on theoretical considerations; that way, high frequency modes could, in principle, enter the reconstruction if the evidence for them becomes sufficiently strong (where the definition of sufficient is based purely on theoretical grounds.)

Here we focus on phenomenological choices for the prior, and make recommendations based on simple considerations, but allow that this will be up to individual choice.

III. THE CORRELATED PRIOR

A. The correlation function

The prior distribution could be an arbitrarily complex multi-variate probability distribution for the non-parametric amplitudes. We assume this distribution is Gaussian, meaning it is specified by a covariance matrix describing fluctuations around a mean or fiducial model, $\mathbf{w}_{\mathrm{fid}}$. Here for simplicity, we further assume that the covariance matrix is a function only of the scale factor and in particular that it is translation independent, depending only on the difference in scale factor $|a - a'|$.

These assumptions allow us to specify the prior using a simple one-dimensional function, defined in [17] as

$$\xi_w(|a - a'|) \equiv \left\langle [w(a) - w^{\mathrm{fid}}(a)][w(a') - w^{\mathrm{fid}}(a')] \right\rangle, \quad (3)$$

which can be used to generate the likelihood for any $w(a)$. This has the effect of reducing the degrees of freedom of the function, effectively binding together neighboring bins below some specified correlation length a_c. This is precisely the kind of prior assumed in the Gaussian process approach [23, 24], though the shape and parametrization of the correlation function can differ considerably. In Ref. [17], we applied this smoothness prior to the Fisher forecasts, and found it a natural way to quantify how much we could learn from future data.

Here we choose the scale factor, a, as our independent variable, which is somewhat arbitrary and in our earlier work [17] instead used redshift. A function which is translation invariant in one variable will not be precisely translation invariant in another, meaning the priors cannot be equivalent. However, if our answers strongly depend on such differences, then likely the priors have been made too strong.

B. Implementation of the correlation prior

We start by discretising $w(a)$ at $a \geq a_{\min}$ using bins uniform in scale factor a, and use a wide bin for $a < a_{\min}$. We have checked that using uniform bins in $\ln(a)$ or z returns a consistent result when the number of bins N is large enough, namely, $N \geq 20$ with $a_{\min} = 0.4$, which we adopt in this work. Note that we use `tanh` bins rather than tophat bins so that the time derivative of w is stable at any a, which is an important issue for calculating the dark energy perturbations when using CMB and LSS data.

Given some correlation shape and choice of fiducial model, it is straight forward to discretize them given our choice of binning. Let us assume the ith bin is from a_i to $a_i + \Delta$, and for simplicity we will assume that all bins have the same width $\Delta = a_{i+1} - a_i$. The equation of state averaged over each bin is given by

$$w_i = \frac{1}{\Delta} \int_{a_i}^{a_i+\Delta} da \, w(a). \quad (4)$$

We can write the variation from the true model averaged over the bin as, $\delta w_i = w_i^{\mathrm{true}} - w_i^{\mathrm{fid}}$. Calculating the covariance matrix of the binned equation of state is then straightforward:

$$C_{ij} \equiv \langle \delta w_i \delta w_j \rangle = \frac{1}{\Delta^2} \int_{a_i}^{a_i+\Delta} da \int_{a_j}^{a_j+\Delta} da' \xi_w(|a - a'|). \quad (5)$$

Given this, the prior for the model can be written as

$$\mathcal{P}_{\mathrm{prior}}(\mathbf{w}) \propto e^{-(\mathbf{w} - \mathbf{w}^{\mathrm{fid}})^T \mathbf{C}^{-1} (\mathbf{w} - \mathbf{w}^{\mathrm{fid}})/2}. \quad (6)$$

Subsequently, in the MCMC, we minimize the total posterior χ^2 defined as

$$\chi^2 = \chi^2_{\mathrm{data}} + \chi^2_{\mathrm{prior}}, \quad (7)$$

where

$$\chi^2_{\mathrm{prior}} = -2 \ln \mathcal{P}_{\mathrm{prior}} = (\mathbf{w} - \mathbf{w}^{\mathrm{fid}})^T \mathbf{C}^{-1} (\mathbf{w} - \mathbf{w}^{\mathrm{fid}}). \quad (8)$$

In Sec. IV we show that one can marginalize over possible choices of $\mathbf{w}^{\mathrm{fid}}$, or define a procedure by which $\mathbf{w}^{\mathrm{fid}}$ is defined using an averaging of $\mathbf{w}$. In such a case, the prior probability can be written as

$$\mathcal{P}(\mathbf{w}) \propto e^{-\mathbf{w}^T \tilde{\mathbf{C}}^{-1} \mathbf{w}/2}, \quad (9)$$

where $\tilde{\mathbf{C}}$ is a modified version of the correlation matrix, meaning that

$$\chi^2_{\mathrm{prior}} = \mathbf{w}^T \tilde{\mathbf{C}}^{-1} \mathbf{w} . \quad (10)$$

The correlation function typically provides a frequency dependent prior: high frequency oscillations are suppressed by the prior, while the low frequency modes are largely unaffected and are so dependent on the data. Providing a prior stabilizes the high frequency variances and allows us to focus on the more interesting low frequency modes. It also significantly improves the MCMC convergence, as the flat directions are now curtailed by the prior. Also, as long as there are sufficient bins compared to the correlation length, the prior largely wipes out dependence on the precise choice of binning.

C. The Wiener filter

To reconstruct $w(a)$ from real data, we simply fold the theoretical prior into our MCMC search using Eq. (7) and search for the optimal model. One can gain analytical insight into reconstructions using the correlated prior by exploiting similar arguments used for *Wiener filtering*, a well understood methodology for reconstructing a signal in the presence of noise. The Wiener approach can be applied in this context once we recognise that the Fisher matrix (or its inverse) describes the expected noise covariance in the $w(a)$ reconstruction, while our theoretical prior describes our expected signal covariance.

We assume that we have some noisy data vector which is the best fit to the observed data, and we wish to estimate the best actual $w(a)$ given the data has noise described by the Fisher matrix, F_{ij}. The expected likelihood is given by

$$\mathcal{P}(\mathbf{w}^{\text{obs}}|\mathbf{w}) \propto e^{-(\mathbf{w}^{\text{obs}}-\mathbf{w})^T \mathbf{F}(\mathbf{w}^{\text{obs}}-\mathbf{w})/2}. \qquad (11)$$

Combining this with the theoretical prior defined above, we find the posterior distribution $\mathcal{P}(\mathbf{w}|\mathbf{w}^{\text{obs}})$. We define the reconstructed equation of state as that which maximizes the posterior; we find

$$\mathbf{F}(\mathbf{w}^{\text{recon}} - \mathbf{w}^{\text{obs}}) + \mathbf{C}^{-1}(\mathbf{w}^{\text{recon}} - \mathbf{w}^{\text{fid}}) = 0, \qquad (12)$$

which has the solution,

$$\mathbf{w}^{\text{recon}} = \mathbf{F}^{-1}(\mathbf{C} + \mathbf{F}^{-1})^{-1}\mathbf{w}^{\text{fid}} + \mathbf{C}(\mathbf{C} + \mathbf{F}^{-1})^{-1}\mathbf{w}^{\text{obs}}. \qquad (13)$$

Assuming that the noise dominates in the higher frequencies, this is effectively a low-pass filtering of the data combined with a high-pass filtering of the fiducial model. Given most choices of fiducial models will be smooth, the latter should have minimal effect.

As mentioned above and detailed in Sec. IV, it is possible to eliminate the explicit dependence on $\mathbf{w}^{\text{fid}}$ by introducing a modified correlation matrix $\tilde{\mathbf{C}}$ with the prior probability given by Eq. (9). Then the maximum posterior solution becomes simply

$$\mathbf{w}^{\text{recon}} = \tilde{\mathbf{C}}(\tilde{\mathbf{C}} + \mathbf{F}^{-1})^{-1}\mathbf{w}^{\text{obs}}. \qquad (14)$$

This analytic solution is useful for projections, which we present in Sec. V B 2. When working with real data, it is very difficult to estimate $\mathbf{w}^{\text{obs}}$, which optimises the observations alone, because flat directions in parameter space make the MCMC convergence impossible when the number of bins is large. But when searching for $\mathbf{w}^{\text{recon}}$, such flat directions are curtailed by the prior.

IV. EXPLORING SPECIFIC CORRELATED PRIORS

A. The choice of correlation function

The choice of amplitude and shape of the prior can critically effect the reconstructed $w(a)$ and other conclusions, such as the number of dark energy parameters which can be constrained. The prior should incorporate both our theoretical prejudices and other earlier cosmological data which have not been explicitly included in the present analysis. Theoretically, one might examine the equation of state dynamics of the quintessence field, beginning with some measure on the space of quintessence potentials (e.g. [28], also see [29].) Earlier data can also be used to decide on the prior. However, it is not clear how useful this is, as the earlier data will naturally be weaker than the data being considered, so it would be surprising

if it resulted in stronger constraints on any degree of freedom. If the experiment provides independent constraints, it should be incorporated explicitly rather than using a heuristic prior model which can misrepresent the distribution of the information. Here we attempt a more pragmatic approach, by using simulated data and attempting to reconstruct a number of 'typical' dark energy models.

One choice is to assume that the prior is diagonal in the binning, effectively assuming the correlation is a delta function [20]. If all bins have the same prior, the correlation matrix is proportional to the identity matrix, which has the advantage that the principal components of data remain unchanged when the prior is included. However, such a prior lacks the bin independence and smoothing properties of a more realistic prior choice.

In previous work [17], we worked with redshift z as the variable, and modeled the correlation function using a form $\xi_w(\delta z) = \xi_w(0)/[1 + (\delta z/z_c)^2]$. Here, we adopt the same form expressed in terms of the scale factor:

$$\xi_w(\delta a) = \frac{\xi_w(0)}{1 + (\delta a/a_c)^2} , \qquad (15)$$

where a_c describes the typical smoothing distance, and $\xi_w(0)$ which is a normalising factor that relates to the amplitude of the expected variance of $w(a)$. In what follows, we will refer to this model as CPZ. The normalisation can also be chosen by specifying the allowed variance of the average equation of state,

$$\sigma_{\bar{w}}^2 \equiv \int_{a_{\min}}^1 \int_{a_{\min}}^1 \frac{da \, da' \, \xi_w(a - a')}{(1 - a_{\min})^2} \simeq \frac{\pi \xi(0) a_c}{1 - a_{\min}}. \qquad (16)$$

The precise shape assumed was somewhat arbitrary, chosen for simplicity and ease of use; while the correlation is expected to decrease, the decline could be steeper a steeper power, or exponential rather than a power law. The Gaussian process [23, 24] work instead typically assumes an exponential fall off, with 'hyper-parameters' determining the shape; some distribution is assumed for these hyper-parameters and then they are marginalised over.

B. Comparing different correlation shapes

We compare a few different shapes for the correlation functions, to see the impact of their qualitative features. In addition to the form discussed above, $\xi_{\text{CPZ}}(\delta a) = \xi_w(0)/[1 + (\delta a/a_c)^2]$, we also consider an exponential fall off, $\xi_{\exp}(\delta a) = \xi_w(0)e^{-\delta a/a_c}$, and a general power law form, $\xi_{\text{pow}}(\delta a) = (\delta a/a_c)^{-n}$. We normalise each to the same mean variance, $\sigma_{\bar{w}}^2$ and set $a_c = 0.06$, apart from the power law form where specifying the variance fixes the value of a_c. We then compare the discretised correlation matrices over the range $a = [0.4, 1.]$ The different shapes are compared in Figure 1.

We first consider the behaviour of the functions at small separations, where it can either approach a constant or diverge for the power law models. If the slope
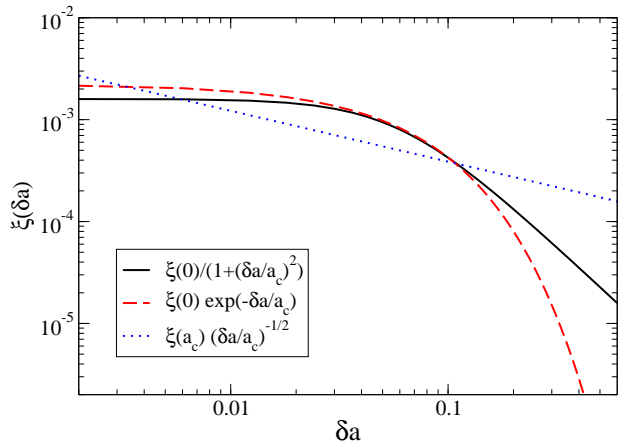
FIG. 1. Different prior shapes normalised to the same mean variance and $a_c = 0.06$: CPZ (black, solid), exponential (red, dashed), power law (blue, dotted).
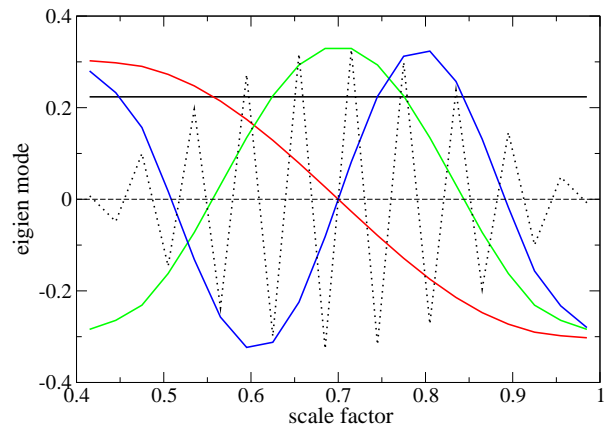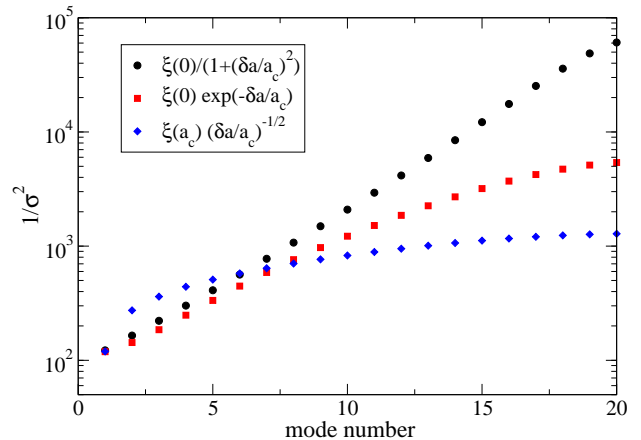

FIG. 3. Different prior eigenvalues normalised to the same mean variance: CPZ (black, circles), exponential (red, squares), power law (blue, diamonds). The normalisation results in the same lowest eigenvalue, and modes increase in frequency.

FIG. 2. Some of the normalised eigenvalues of the prior. Changes to the prior description have little effect on the modes themselves. The modes are largely ordered in frequency (black, red, green, blue), with the lowest frequencies the least constrained by the prior. The highest frequency mode (dashed) is the most constrained. The deviation from the simple Fourier transform is due to the boundaries on w.

is too steep, the discretised diagonal correlation also diverges unless a small scale cut-off is imposed. In the limit that this cut-off becomes small, the correlation matrix effectively becomes diagonal which we consider as a separate case. Below, we focus on the power law case only for $n = 1/2$, a shape shallow enough to avoid the divergence.

The impact of these different ways of defining priors is perhaps most clearly seen in the eigenvalues of the associated prior matrices. To illustrate this, we choose a binning (20 bins between $a = 0.4$ and $a = 1.$) and calculate the corresponding correlation matrices, and find the eigenmodes and their eigenvalues in this binning.

Perhaps surprisingly, virtually all the ways of defining the prior result in similar prior eigenmodes, and in identical ordering for the eigenvalues of these modes. A few of the modes are shown in Fig. 2, where it is clear that the modes correspond closely with the Fourier basis, apart from minor differences which arise from the scale factor boundaries. Because correlations fall off with distance, the most constrained modes are those with the highest frequencies. Slowly varying modes are least constrained by the prior. (In the limit of a diagonal correlation matrix, the eigenvalues are degenerate and the mode definitions are arbitrary.)

However, the different correlation shapes significantly change the associated eigenvalues, as can be seen in Fig. 3. Because the models are normalised to the same mean variance, the lowest eigenvalues are the same. However, the spacing of the higher values is significantly different. For the CPZ model, the eigenvalues are nearly equally spaced in log, and are steepest of these three models for a given a_c, leading to the most constrained high frequency modes. This spacing directly depends on a_c: the longer the correlation distance, the steeper the spectrum. In the limit of small a_c, the matrix becomes diagonal. Because of its relatively simple behaviour and transparent dependence on its parameters, we adopt the CPZ form as our default model below.

C. Changing the fiducial model

The prior distribution function, as defined in (6), depends explicitly on the choice of the fiducial equation of state $w^{\text{fid}}(a)$. For simplicity, we could assume the fiducial model to be $w = -1$, but arguably this is *ad hoc*. Formally the priors should be chosen from fundamental physics; a Gaussian prior centred at $w = -1$ is unphysical in many simple models that do not allow the so called 'phantom' [30] region, $w < -1$, which violates the null energy condition. Dynamical models like quintessence may asymptote to cosmological constant behaviour at early or late times (thawing or freezing models), but are always constrained to have $w \geq -1$. While phantom models are possible, models which cross the 'phantom divide' (dubbed 'quintom' models [31]) can also be a struggle to build in a natural way. In our present phenomenological approach, it makes more sense to choose a prior which does not favour a particular value. Since we wish to test ΛCDM, we prefer to use a prior which does not favor $w = -1$ over another model, e.g., $w = -0.9$.

One way to eliminate the dependence on $w^{\text{fid}}(a)$ is to marginalise over it. For a constant w^{fid} this can be done analytically; we can write the fiducial model as $\mathbf{w}^{\text{fid}} = w^{\text{fid}}\mathbf{u}$, where $\mathbf{u} = (1, 1, 1, ...)$ and we assume a weak Gaussian prior on w^{fid} with variance given by σ_{fid}^2. Marginalising over w^{fid} has the effect of changing the effective correlation matrix:

$$\mathcal{P}_{\text{prior}} \propto \int dw^{\text{fid}} e^{-(\mathbf{w}-w^{\text{fid}}\mathbf{u})\mathbf{C}^{-1}(\mathbf{w}-w^{\text{fid}}\mathbf{u})/2-(w^{\text{fid}})^2/2\sigma_{\text{fid}}^2}$$
$$\propto e^{-\mathbf{w}^T \tilde{\mathbf{C}}^{-1}\mathbf{w}/2} , \tag{17}$$

where

$$\tilde{C}_{ij}^{-1} = C_{ij}^{-1} - \frac{C_{ik}^{-1}u_k u_l C_{lj}^{-1}}{\sigma_{\text{fid}}^{-2} + u_l C_{lk}^{-1}u_k} . \tag{18}$$

Such expressions are a common result when marginalising over nuisance parameters (e.g. [32]) and can be simply inverted using the Sherman-Morrison formula [33], a special case of the Woodbury formula [34], to find:

$$\tilde{C}_{ij} = C_{ij} + \sigma_{\text{fid}}^2 u_i u_j. \tag{19}$$

Here we see the utility of introducing a prior on the fiducial value, however weak – a flat prior corresponds to an infinite variance, which makes $\tilde{C}$ undefined. Most importantly, marginalising assigns large noise to the flat mode in the prior and ensures that for reconstructions this mode is determined by the data alone.

An alternative approach is to define the fiducial model to be a function of the assumed equation of state ($\mathbf{w}^{\text{fid}}(\mathbf{w})$) rather than being fixed; for example, it could float to a constant value given by the average of the equation of state,

$$w^{\text{fid}} = \bar{w} \equiv \frac{1}{a_f - a_i}\int_{a_i}^{a_f} da\, w(a) = \frac{1}{N}\Sigma_i w_i. \tag{20}$$
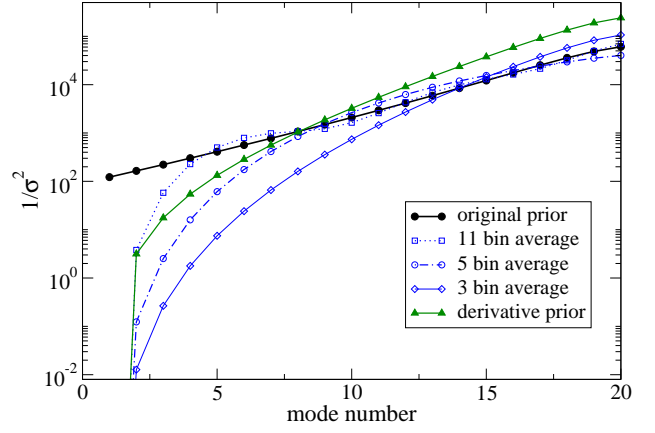

FIG. 4. Impact on the eigenspectrum of various transforms considered. The original CPZ spectrum is in black (solid circles). Marginalising over the mean only zeroes out the first eigenvalue. Using a running average tends to reduce the lowest frequency modes while leaving the higher frequency constraints unchanged. (11 bin average - open circles, 5 bin average - open squares, 3 bin average - open diamonds.) Taking the CPZ prior form for the derivative results in the green triangles; the spectrum steepens and the first mode is zeroed.

Such a change, where the fiducial model is a linear transformation of the underlying model, $\mathbf{w}^{\text{fid}} = S\mathbf{w}$ is also equivalent to simply changing the effective correlation matrix. The explicit dependence of the prior on the fiducial model is absorbed into a modified correlation matrix,

$$\tilde{C}_{ij}^{-1} = (\delta_{ik} - S_{ik}^T)C_{kl}^{-1}(\delta_{lj} - S_{lj}). \tag{21}$$

In the case above where the fiducial model is the average over all bins, we have $S_{ij} = u_i u_j/N$, implying

$$\tilde{C}_{ij}^{-1} = C_{ij}^{-1} - \frac{1}{N}(u_i u_k C_{kj}^{-1} + C_{ik}^{-1}u_k u_j) + \frac{1}{N^2}u_i u_j u_k C_{kl}^{-1}u_l. \tag{22}$$

This is similar to the result for marginalisation and, like that case, can result in a degenerate direction unless an additional prior is assumed; without this, the matrix cannot be inverted.

The implementation of a 'floating' fiducial model can be based on other smoothing schemes. We might instead define the fiducial model as a more local average of the true model, which reduces the prior penalty for the longer wavelength modes, making them more responsive to the data. For example, we could have

$$w_i^{\text{fid}} = \sum_{|a_j - a_i|\leq a_c} w_j^{\text{true}}/N_j \tag{23}$$

where N_j denotes the number of the neighboring w bins around the ith bin w_i within the 'correlation' radius a_c. We tried other smoothing methods, such as a Gaussian kernel scheme, and found very similar results.

These modifications to the fiducial model significantly affect the eigenvalues, particularly the one for the constant w mode, which is effectively reset to zero (or the

prior that it has been explicitly given.) Both approaches, marginalising over the mean value or redefining the fiducial model to the global average, do this while leaving the other eigenvalues unchanged, just as expected. Using a more local definition of the fiducial model also reduces the constraints on other long wavelength modes while leaving the shorter ones mostly unchanged, as can be seen in Fig. 4. We adopt this local averaging method of Eq. (23) when performing the reconstructions in Sec. V

D. Correlations of $w(a)$ derivatives

Another way to ensure that the average of the fiducial model does not impact the reconstruction is to apply a prior instead to the derivative of the $w(a)$ function. Essentially, we specify a correlation matrix for the difference of the equation of state between bins,

$$\langle (w_i - w_{i+1})(w_j - w_{j+1}) \rangle = D_{ij}. \qquad (24)$$

This has one less dimension than the correlation of the w bins itself. Again, we can define a Gaussian prior based on this correlation matrix.

$$\mathcal{P}(\mathbf{w}) \propto e^{-(w_i - w_{i+1})^T D_{ij}^{-1}(w_j - w_{j+1})/2}. \qquad (25)$$

This can be thought of as a prior on the bin values themselves with an alternative inverse correlation defined by,

$$\tilde{C}_{ij}^{-1} = D_{ij}^{-1} - D_{i-1,j}^{-1} - D_{i,j-1}^{-1} + D_{i-1,j-1}^{-1} \qquad (26)$$

The indices of $\tilde{C}$ span $[1, N]$, while those for D span $[1, N-1]$; where an index for D exceeds its range (i.e. $i = 0, N$), the inverse matrix is taken to be zero.

While $\tilde{C}_{ij}^{-1}$ is well defined in this way, it is not invertible because the constant w mode is not constrained in any way. If required, this can be combined with a weak prior on the constant w mode to yield a correlation matrix which is invertible.

Assuming the fiducial CPZ correlation for the derivative of w (rather than for w itself) results in a new spectrum of eigenvalues. As expected, the constraint on the mean mode is reset to zero. However, it also results in a steeper spectrum than the original one, as can be seen in Fig. 4. This is roughly equivalent to using a smaller correlation length in the original prescription, and so would not qualitative change the analysis below.

V. RECONSTRUCTING THE EQUATION OF STATE

Next we combine the priors we have defined with the data to see how effective they will be in reconstructing $w(a)$. We do this in two ways: first, using calculations of the Fisher matrices for a set of observations to test Wiener reconstructions of some sample $w(a)$ functions,
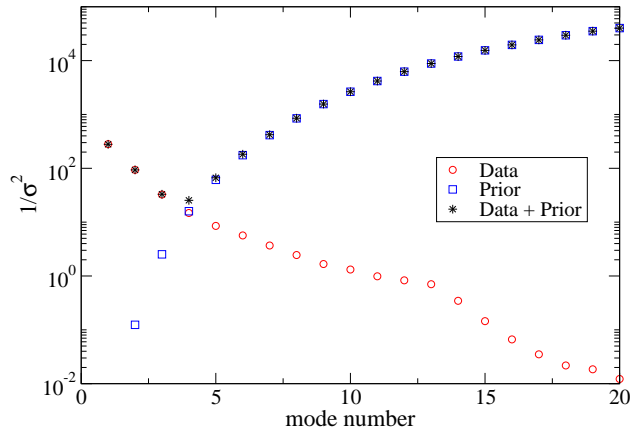

FIG. 5. Eigenvalues from the data (red circles) and prior independently (blue squares) and combined (black stars). The modes are arranged roughly from low frequency to high frequency, though this correspondence is not exact when the data are included. The low frequency modes are constrained by the data, while the high frequency modes are constrained by the prior.

effectively assuming a simple Gaussian approximation to the likelihood. Second, we create more realistic simulations of the expected CMB, SN and $H(z)$ data and reconstruct $w(a)$ by finding the maximum posterior model using MCMC techniques.

A. Data assumptions

We test our reconstruction algorithms by combining the CMB distance prior [35] from Planck [27] with supernovae (SN) and galaxy-galaxy correlation measurements potentially possible for a future space-based dark energy mission. Note that our assumptions are close to those that could be achieved from a next generation survey such as Euclid, but do not exactly match those now expected for the default Euclid mission [25, 26]. For the CMB, we use a Fisher matrix forecast based on the Planck data. For the SN data, we assume 4000 SN distributed in 14 redshift bins from $z = 0.15$ to $z = 1.55$ for a deep SN survey [36]. We also include 300 low-z SN from the Nearby Supernova Factory [37] to improve the constraint and assume a Gaussian noise with a variance $\sigma = 0.13$ for all the data points.

The measurement of the expansion rate at different redshifts $H(z)$ can tighten constraints on $w(a)$, since the DE energy density is directly related to the Hubble rate at a given redshift (e.g., [38]). One of the most promising ways of estimating $H(z)$ from future spectroscopic surveys is to observe the Alcock-Paczynski effect in the power spectrum or correlation function of galax-

ies [39], exploiting features such as the baryon acoustic oscillations (BAO) and the power spectrum peak. The BAO feature in the correlation function is at a fixed comoving scale which is determined by the physics at recombination; measuring the redshift difference of this feature in the radial correlation function provides a direct local measurement of $H(z)$. We divide the galaxies $(0.35 < z < 2.5)$ into bins of width $\Delta z = 0.1$ and predict the measurement errors $\sigma_{H(z)}$ in each redshift bin, using the Fisher matrix technique of [40]. We assume that the spaced-based survey would cover 20,000 deg^2 surveying area and will detect emission line galaxies with the redshift precision of $\sigma_z = 0.001(1 + z)$. For our projections, we adopt the values derived in [41] for the number density of emission line galaxies and the bias model of [42].

Given these data models, we consider the following set of cosmological parameters,

$$\mathbf{P} \equiv (\omega_b, \omega_c, \Theta_s, \tau, n_s, A_s, w_1, ..., w_{20}, \mathcal{N}), \qquad (27)$$

where ω_b and ω_c are the physical baryon and cold dark matter densities relative to the critical density respectively, Θ_s is the ratio of the sound horizon to the angular diameter distance at decoupling, τ denotes the optical depth to reionization, n_s and A_s are the primordial power spectral index and amplitude respectively and $w_1, ..., w_{20}$ denote the 20 w bins uniform in a. We also include and marginalize over one nuisance parameter $\mathcal{N}$ for supernovae, which accounts for the calibration uncertainty in measuring the supernova intrinsic luminosity. The details of the Fisher forecast and the related survey parameters can be found in [43] and [44]. After marginalising over the cosmological and nuisance parameters, the projected constraint on a constant equation of state model is $\sigma_{\bar{w}} = 0.015$.

B. Reconstruction Forecasts

A Fisher matrix analysis provides a means of projecting the constraints possible for the experiments, under the optimistic assumption that the resulting likelihood is Gaussian (Eq. 11.) Combined with our assumed Gaussian prior, we can analytically forecast the expected reconstructions when assuming a particular underlying $w(a)$ using the Wiener filter approach described in Section III C. This approach also allows us to quantify the expected bias and variance for a given input model.

1. Combined eigenmodes

When combining data with the prior, the effective resulting Fisher matrix changes from $\mathbf{F} \to \mathbf{F} + \tilde{\mathbf{C}}^{-1}$. Many of the general aspects of reconstructions can be understood by considering the eigenvalues of this combined matrix. Roughly speaking, the data constrain the low frequency modes the best, while the prior constrains the
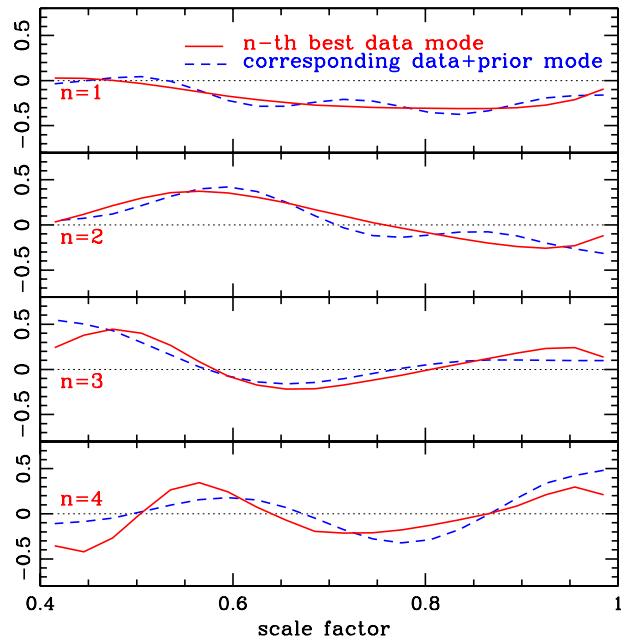

FIG. 6. The first four best constrained eigenmodes of the data covariance matrix (solid lines), along with the corresponding modes of the data+prior covariance (dashed lines). As seen in Fig. 5, the data eigenvalues become comparable to the prior eigenvalues at the $n = 4$ mode, and the $n > 4$ modes are completely determined by the prior.

high frequencies, and the most constrained modes of each survive in the combined matrix. However, this is only approximate since the data and prior have different eigenmodes; in particular, the data modes are more weighted to the lower redshifts.

Fig. 5 shows the impact of combining the data and prior on the eigenvalues. The cross over point is crucial, as it determines the reconstruction bias, while the slope of the prior eigenvalues will determine the variance of the reconstructions. With the data described above and our fiducial prior choice, only 4 data modes survive to play a role in the reconstructions. The surviving modes are shown in Fig. 6. One can see that, while the shapes are distorted by the prior, their basic features remain in tact.

More modes would survive if the data were better, or if the prior were weaker (e.g. reducing the prior normalisation $\sigma_{\bar{w}}$ or the slope a_c.) The benefit of a weaker prior is that more modes would be used to reconstruct the data, giving smaller potential bias. The cost would be significantly more variance in the reconstructions. The key issue is whether the true underlying $w(a)$ will require more modes to achieve a small bias.
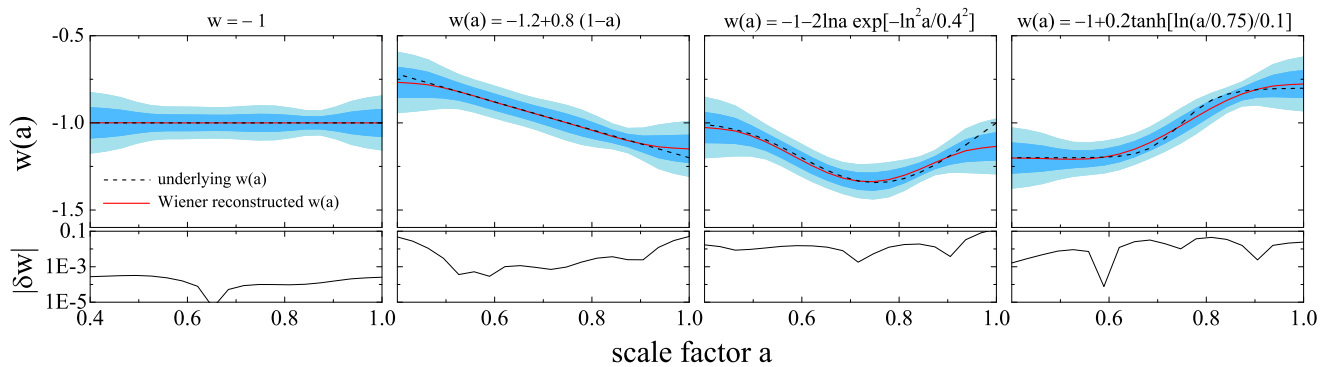
FIG. 7. Wiener reconstructions of the various test functions, based on the Fisher projections for a joint data set of SN + $H(z)$ + CMB distance prior (Planck), and using our fiducial prior assumptions.

2. Wiener filter projections

If the Gaussian approximation to the likelihoods are reasonable, then the Fisher matrix should allow a good approximation to the expected errors on the measured w_i bins. We can make realisations of the observed binned w_i data by assuming some true underlying model and adding to it random noise derived using the principal components of the Fisher matrix. Each principal component is assigned a random Gaussian amplitude with the appropriate variance, and these are summed to get the total noise in the binned data. These realisations are then Wiener smoothed using Eq. (14) to get the reconstructed $w(a)$. (Note that when $\tilde{\mathbf{C}}$ makes negligible constraint on the average, then the term $\mathbf{F}^{-1}(\tilde{\mathbf{C}} + \mathbf{F}^{-1})^{-1}\mathbf{w}^{\mathrm{fid}}$ can be neglected if the fiducial model itself is constant, making Eqns. (13) and (14) equivalent.)

With these assumptions, we can derive analytically the mean reconstructed models, as well as the variance and bias expected. For a given $\mathbf{w}^{\mathrm{true}}$, the average reconstructed model is

$$\mathbf{w}^{\mathrm{mean}} = \tilde{\mathbf{C}}(\tilde{\mathbf{C}} + \mathbf{F}^{-1})^{-1}\mathbf{w}^{\mathrm{true}}. \quad (28)$$

This implies that that bias contribution to the MSE is

$$\sum_i (w_i^{\mathrm{true}} - w_i^{\mathrm{mean}})^2 = ||\mathbf{F}^{-1}(\tilde{\mathbf{C}} + \mathbf{F}^{-1})^{-1}\mathbf{w}^{\mathrm{true}}||^2. \quad (29)$$

We can similarly predict the variance contribution under these assumptions to find,

$$\sum_i \langle (w_i^{\mathrm{mean}} - w_i^{\mathrm{recon}})^2 \rangle = \mathrm{Tr}[(\mathbf{F} + \tilde{\mathbf{C}}^{-1})^{-2}\mathbf{F}]. \quad (30)$$

Note that in this approximation, the variance is independent of the assumed model. We have confirmed these analytic results using 20,000 realisations for each input model.

We consider four basic underlying phenomenological models ($\mathbf{w}^{\mathrm{true}}$) to test the ability of our algorithm to capture different kinds of evolving features in $w(a)$:

$$w_{\mathrm{const}} = -1.0$$
$$w_{\mathrm{lin}} = -1.2 + 0.8(1-a)$$
$$w_{\mathrm{feat}} = -1 + 2\ln a \cdot \exp(-\ln^2 a/0.4^2)$$
$$w_{\mathrm{trans}} = -1 + 0.2\tanh(\ln(a/a_t)/\Delta) . \quad (31)$$

We refer to these as the constant w model, the linear model [45], the feature model (see discussions in [46]) and the transition model (e.g. [47]), respectively. For this section, we also consider a fifth 'thawing' model based on the parametrization of [29] for a thawing quintessence which smoothly goes from $w = -1$ at high redshifts to $w = -0.8$ today. The transition model changes from $w = -1.2$ at high redshifts to $w = -0.8$ at low redshifts, and we chose the transition to occur at $a_t = 0.75$ with width of $\Delta = 0.1$.

These functions and their reconstructions are shown in Fig. 7. For these reconstructions we use our fiducial prior, which we define as the CPZ prior ($\sigma_{\bar{w}} = 0.02, a_c = 0.06$) with a fiducial model as the average of bins within $\Delta a = 0.06$ given by Eq. (23) (effectively the five bin average shown in Fig. 4.) Note that $\sigma_{\bar{w}}$ primarily is used to normalise the high frequency priors; in fact, the local averaging means that effectively no prior is placed on $\bar{w}$.

One can see that on average the reconstructions are fairly unbiased, with practically no bias in the case of a constant w. This is expected, as the smoother functions are well represented by the modes which survive the prior. Models with stronger evolution, such as the feature model and the fast transition model, have the most bias. However, even for these, the average MSE is dominated by the variance for this choice of prior.

The variance can be decreased by strengthening the prior, but this effectively allows fewer data modes to survive and so increases the potential bias of the reconstructions. Thus the MSE cannot be dramatically reduced except in those cases where the bias is intrinsically small. For example, the constant w modes are unbiased in this prescription, and so the variance can be made very small.

The MSE, and its contributions from bias and variance for each input model are shown in Table I. One can
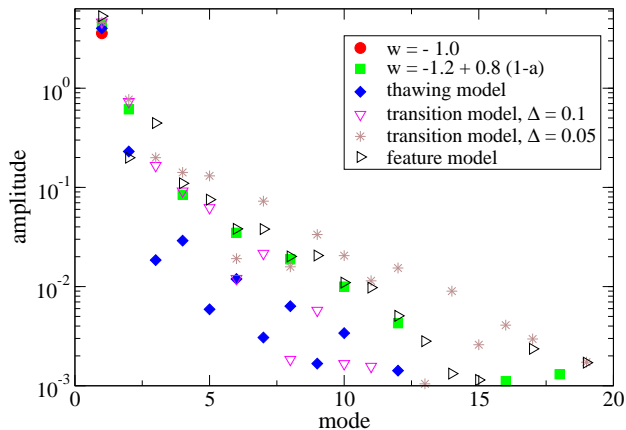
FIG. 8. Amplitudes of the various test functions expanded in the prior eigenmodes. Those functions which are best reconstructed are those dominated by the low mode numbers. Those with higher frequency features are reconstructed with more bias.

Model	MSE	Bias	Variance
$w = -1$	0.064	-	0.064
Linear	0.068	0.004	0.064
Feature	0.081	0.017	0.064
Transition ($\Delta = 0.1$)	0.071	0.007	0.064
Transition ($\Delta = 0.05$)	0.095	0.031	0.064
Thawing	0.065	0.001	0.064

TABLE I. A comparison of MSE results for different assumed models, broken down into bias and variance contributions.

see that the MSE figures are dominated by the variance, with the bias term being comparable in the transition model when the change is very quick. This is intentional; a stronger prior does not significantly reduce the MSE for most of the models we assume (the exception being the constant model), and given the choice it seems better to have a slightly noisier reconstruction than one which could be significantly biased.

The relative biases can be easily understood by looking at the input models expanded in the eigenmodes of the prior, as shown in Fig. 8. We see that the models which are reconstructed with least bias (constant, thawing model, slow transition model) have little support beyond the fifth prior mode, while the others have some higher frequency components which are wiped out by the prior.

C.   Reconstructions from simulated data

We next explore how our reconstruction method works for more realistic data, because Fisher forecasts tend to be optimistic and ignore potential non-Gaussianities in the likelihood. Here we look at single realisations of the potential data sets based on the models in Eq. (31). For

these we then perform a blind MCMC reconstruction, adding to the data χ^2_{data} a similar term for the prior χ^2_{prior} as described in Sec III B.

To generate the simulated data, we set the fiducial values of other cosmological parameters to values favoured by the WMAP7 observations [35], i. e. $\Omega_m = 0.263$ and $H_0 = 71$ km/s/Mpc. We then simulate the SN luminosity data by randomly generating the SN redshifts and errors, while respecting their expected redshift density. The simulated SN distance modulus μ of four phenomenological dark energy models (offset by that of the ΛCDM model) are shown in Fig. 10. We similarly randomly generate $H(z)$ measurements in the redshift bins described above consistent with their expected errors.

For the CMB data, we generate realisations for the three CMB distance parameters [35] consistent with their expected covariances expected from Planck. Note that for the real data analysis, one should use the full CMB spectrum data, which is also sensitive to dark energy through the integrated Sachs-Wolfe effect. For such calculations, one must include dark energy perturbations, otherwise the derived constraints on cosmological parameters will be biased [48–50]. Furthermore, one should treat dark energy perturbations self-consistently when w crosses $w = -1$ [51, 52]. One must be careful even with the CMB distance prior, which can be biased if the model is significantly different from the ΛCDM for which it has been derived [53].

We simultaneously fit the 20 bins with Ω_m and H_0 to our combined mock dataset using a modified version of CAMB/CosmoMC [54], with the smoothness prior implemented. In Fig. 11 we show the reconstructed $w(a)$ from a joint mock dataset of SN + $H(z)$ +CMB distance prior (Planck). As shown, our algorithm can successfully capture all the DE models we hide in the data, and the reconstruction is very accurate – the absolute value of the relative difference $\lesssim 10\%$ in all cases. The calculation of the prior χ^2_{prior} adds virtually no overhead to the likelihood calculations, making the method quite efficient.

It is important to emphasise that Figs. 7 and 11, while apparently very similar, are actually showing very different things. Fig. 7 shows the average reconstructed $w(a)$ over the ensemble of data realisations consistent with the experiments, and the errors show the distribution of the reconstructed $w(a)$ for these different data realisations. Fig. 11 instead shows the reconstructed $w(a)$ for a single simulated data set (similar to what we expect to get from real observations) and its error bars represent the uncertainty in the determination of $w(a)$ for this realisation. This best fit differs from the input model both because of the average bias seen in Fig. 7 and because of the randomness of the realised data. However, the error regions are similar because they are related to very similar $\Delta\chi^2$ functions.
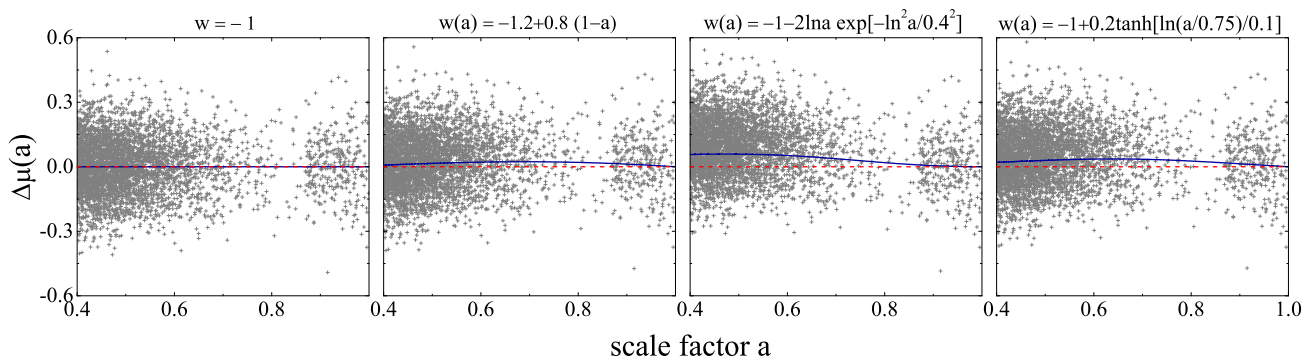
FIG. 9. The simulated SN magnitudes for a future deep survey as described in [36] for four phenomenological dark energy models listed in Eq (31). The blue solid line shows the relative SN magnitude with respect to that of the ΛCDM model. The red dashed line shows $\Delta\mu = 0$ to guide eyes.
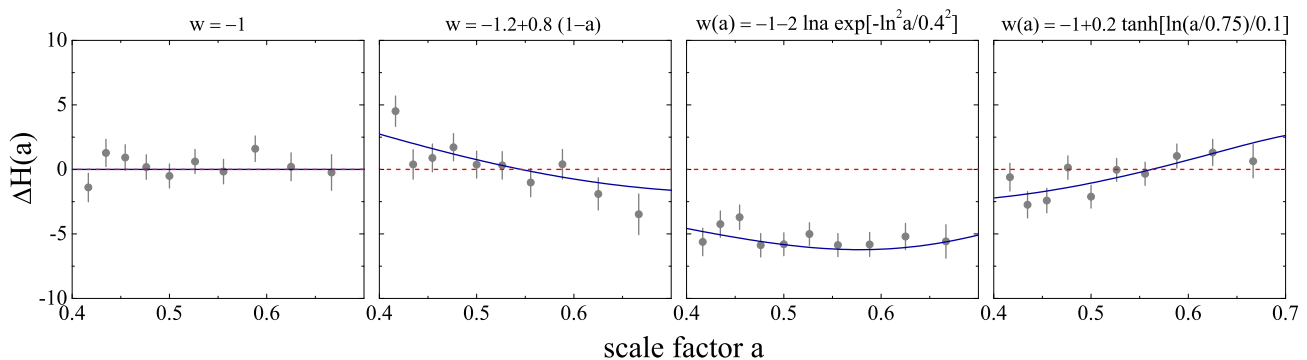

FIG. 10. The simulated $H(a)$ measurements for a future half-sky survey as described in [40] for four phenomenological dark energy models listed in Eq (31). Again, the blue solid line shows the difference in $H(a)$ with respect to that of the ΛCDM model. The red dashed line shows $\Delta H(a) = 0$ to guide eyes.

VI. DISCUSSION

Our reconstructions are very good, but some bias remains; this primarily takes the form of smoothing out the more quickly varying $w(a)$ behaviours. While this can be made smaller by weakening the prior, the cost would be to increase the variance in the reconstruction. The bias simply reflects our assumption that smoother $w(a)$ models are more likely.

This is not an intrinsic limitation of the method, but reflects our choice to assume that $w(a)$ is smooth. While we do not have to make this assumption, we do require some theoretical basis where smoothness (or a measure) can be assumed, such as the DE scalar field potential $V(\phi)$. However, this is essentially the requirement that the given model is well enough posed to be falsifiable. Any model description should allow for the calculation of the prior in some effective basis, such as bins of $w(a)$.

We could either perform the reconstruction directly using the parameters where the behaviour is expected to be smooth, or we could reconstruct in a more phenomenological basis, like $w(a)$, but use a prior in this basis derived assuming smooth behaviour in another basis. Direct re-

construction of something like the scalar field potential is better in the sense that one keeps in contact with the theoretical context in which the prior makes sense; however, it makes it harder to compare the conclusions arising from different assumptions. Alternatively, we can derive a prior on $w(a)$, either arising from more fundamental physics or purely phenomenological models by calculating the resulting w correlation matrices C by marginalising over whatever parameters appear in the description (e.g. [28]).

Inevitably, the simplifying assumptions we have made above for the form of these correlations will not hold when realistic models are considered. In particular, we assumed that the prior is Gaussian, and that the correlations are a function of δa and are translation independent. The latter is certainly not expected, for example 'freezing' models of quintessence will have larger variance at high redshifts, while 'thawing' models will have larger variance at low redshifts. Even translating the simple linear parameterisation, $w(a) = w_0 + w_a(1-a)$ with Gaussian priors on w_0 and w_a, into a correlation matrix on $w(a)$ turns out not to be translation invariant; the resulting variance is smallest at $a = 1$, and increases with
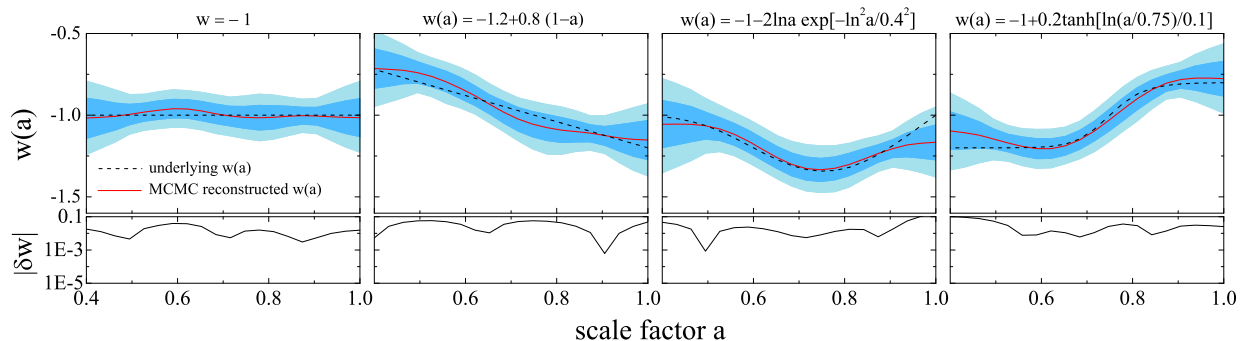
FIG. 11. The reconstructed $w(a)$ from a joint mock datasets of SN + $H(z)$ + CMB distance prior (Planck). We show the best fit model (red solid), the true model we put in (black dashed) with 68 and 95% CL. error (dark and light shaded bands) in the upper panels, and the absolute value of the relative difference in the lower panels. Unlike the Wiener results, these are based on a single realisation of the data.

$1 - a$.

It is possible that including many different classes of models together, such as thawing and freezing models, might homogenise the correlations to some extent. Alternatively, one could hope that such deviations of the covariance from our simplifying assumptions will not be significant enough to greatly impact the reconstructions. However, this requires further exploration and we will investigate this in future work.

VII.   CONCLUSIONS

Here we have outlined a Bayesian method for reconstructing $w(a)$, which essentially involves multiplying the likelihood of a non-parametric description of $w(a)$ by a suitably defined prior. This prior is defined in such a way as to make the reconstructions smooth and independent of the binning choice, eliminating flat directions in parameter space which are a generic problem for non-parametric approaches.

With the right choice of prior, we can also eliminate the dependence on the fiducial model; reconstructions where the data are poor or non-existent are instead determined by the assumed smoothness of the function, tending towards the mean of the measured $w(a)$. However, application of a prior always implies some bias; in this case, the bias is largest for models which are quickly changing in redshift or scale factor.

The advantages of this approach are that it makes one's

priors explicit and it is easy to implement; once one decides on the prior parameters, e.g. the correlation length and the strength of the prior, it is straightforward and fast to evaluate the prior likelihood for whatever choice of binning is used. This can be combined with the data in an MCMC analysis, searching for the highest posterior solution for the w bin amplitudes; indeed, by breaking the parameter degeneracies, the prior will speed up MCMC analyses.

With the accumulating high-quality data of SN, CMB and LSS, our method is an ideal tool to help reveal the nature of dark energy in the near future. Here we have focused on the methodology only, but in future work we will apply this prior to present data.

[1] A. Riess *et al.*, Astron. J. **116**, 1009 (1998); S. Perlmutter *et al*, Ap. J. **517**, 565 (1999).

[2] V. Sahni and A. Starobinsky, Int. J. Mod. Phys. D **15** (2006) 2105 [arXiv:astro-ph/0610026].

[3] D. Huterer, M. S. Turner, Phys. Rev. **D60** (1999) 081301. [astro-ph/9808133].

[4] T. D. Saini, S. Raychaudhury, V. Sahni, A. A. Starobinsky, Phys. Rev. Lett. **85** (2000) 1162-1165. [astro-ph/9910231].

[5] T. Chiba, T. Nakamura, Phys. Rev. **D62** (2000) 121301. [astro-ph/0008175].

[6] D. Huterer, M. S. Turner, Phys. Rev. **D64** (2001) 123527. [arXiv:astro-ph/0012510 [astro-ph]].

[7] J. Weller, A. Albrecht, Phys. Rev. **D65** (2002) 103512. [astro-ph/0106079].

[8] U. Alam, V. Sahni, T. D. Saini, A. A. Starobinsky, Mon. Not. Roy. Astron. Soc. **344** (2003) 1057. [astro-ph/0303009].

[9] R. A. Daly, S. G. Djorgovski, Astrophys. J. **597** (2003) 9-20. [astro-ph/0305197].

[10] A. Shafieloo, U. Alam, V. Sahni, A. A. Starobinsky, Mon. Not. Roy. Astron. Soc. **366** (2006) 1081-1095. [astro-ph/0505329].

[11] J. Dick, L. Knox, M. Chu, JCAP **0607** (2006) 001. [astro-ph/0603247].

[12] A. Shafieloo, Mon. Not. Roy. Astron. Soc. **380** (2007) 1573-1580. [astro-ph/0703034 [ASTRO-PH]].

[13] C. Clarkson, C. Zunckel, Phys. Rev. Lett. **104**, 211301 (2010). [arXiv:1002.5004 [astro-ph.CO]].

[14] M. J. Mortonson, W. Hu, D. Huterer, Phys. Rev. **D79** (2009) 023004. [arXiv:0810.1744 [astro-ph]].

[15] M. J. Mortonson, [arXiv:1005.3716 [astro-ph.CO]].

[16] D. Huterer, G. Starkman, Phys. Rev. Lett. **90** (2003) 031301. [astro-ph/0207517].

[17] R. G. Crittenden, L. Pogosian and G. B. Zhao, JCAP **0912** (2009) 025 [arXiv:astro-ph/0510293].

[18] A. Albrecht, G. Bernstein, Phys. Rev. **D75** (2007) 103003. [astro-ph/0608269].

[19] J. Tang, F. B. Abdalla, J. Weller, Mon. Not. Roy. Astron. Soc. **416** (2011) 2212. [arXiv:0807.3140 [astro-ph]].

[20] A. Albrecht, L. Amendola, G. Bernstein, D. Clowe, D. Eisenstein, L. Guzzo, C. Hirata, D. Huterer *et al.*, [arXiv:0901.0721 [astro-ph.IM]].

[21] T. D. Kitching, A. Amara, [arXiv:0905.3383 [astro-ph.CO]].

[22] M. J. Mortonson, D. Huterer, W. Hu, Phys. Rev. **D82** (2010) 063004. [arXiv:1004.0236 [astro-ph.CO]].

[23] T. Holsclaw, U. Alam, B. Sanso, H. Lee, K. Heitmann, S. Habib and D. Higdon, Phys. Rev. D **82** (2010) 103502 [arXiv:1009.5443 [astro-ph.CO]].

[24] T. Holsclaw, U. Alam, B. Sanso, H. Lee, K. Heitmann, S. Habib and D. Higdon, Phys. Rev. Lett. **105** (2010) 241302 [arXiv:1011.3079 [astro-ph.CO]].

[25] http://sci.esa.int/euclid

[26] R. Laurejis *et al.*, "Euclid Definition Study Report," arXiv:1110.3193 [astro-ph.CO].

[27] http://www.rssd.esa.int/index.php?project=Planck

[28] D. Huterer, H. V. Peiris, Phys. Rev. **D75** (2007) 083503. [astro-ph/0610427].

[29] R. Crittenden, E. Majerotto and F. Piazza, Phys. Rev. Lett. **98** (2007) 251301 [arXiv:astro-ph/0702003].

[30] R. R. Caldwell, Phys. Lett. B **545**, 23 (2002).

[31] B. Feng, X. L. Wang and X. M. Zhang, Phys. Lett. B **607**, 35 (2005).

[32] J. R. Bond and R. G. Crittenden, NATO Adv. Study Inst. Ser. C. Math. Phys. Sci. **565** (2001) 241 [arXiv:astro-ph/0108204].

[33] J. Sherman, W. J. Morrison, Ann. Math. Stat. 20: 621 (1949); Ann. Math. Stat. 21 (1): 124127 (1950).

[34] M. A. Woodbury, Memorandum Rept. 42, Statistical Research Group, Princeton University (1950).

[35] E. Komatsu *et al.* [WMAP Collaboration], arXiv:1001.4538 [astro-ph.CO].

[36] P. Astier, J. Guy, R. Pain and C. Balland, arXiv:1010.0509 [astro-ph.CO].

[37] W. M. Wood-Vasey *et al.*, New Astron. Rev. **48** (2004) 637 [arXiv:astro-ph/0401513].

[38] D. Stern, R. Jimenez, L. Verde, M. Kamionkowski and S. A. Stanford, JCAP **1002** (2010) 008 [arXiv:0907.3149 [astro-ph.CO]].

[39] N. Kaiser, Mon. Not. Roy. Astron. Soc. **227** (1987) 1

[40] L. Samushia, et al., Mon. Not. Roy. Astron. Soc. (2010) [arXiv:1006.0609 [astro-ph.CO]]

[41] J. E. Geach, et al., Mon. Not. Roy. Astron. Soc. **402** (2010) 1330

[42] A. Orsi, et al., Mon. Not. Roy. Astron. Soc. **405** (2010) 1006

[43] L. Pogosian, P. S. Corasaniti, C. Stephan-Otto, R. Crittenden, R. Nichol, Phys. Rev. **D72** (2005) 103519. [astro-ph/0506396].

[44] G. -B. Zhao, L. Pogosian, A. Silvestri, J. Zylberberg, Phys. Rev. **D79** (2009) 083513. [arXiv:0809.3791 [astro-ph]].

[45] M. Chevallier, D. Polarski, Int. J. Mod. Phys. **D10** (2001) 213-224. [gr-qc/0009008]; E.V. Linder, Phys. Rev. Lett. **90**, 091301 (2003) [arXiv:astro-ph/0208512].

[46] J. -Q. Xia, G. -B. Zhao, H. Li, B. Feng, X. Zhang, Phys. Rev. **D74** (2006) 083521. [astro-ph/0605366]; B. Feng, M. Li, Y. -S. Piao, X. Zhang, Phys. Lett. **B634** (2006) 101-105. [astro-ph/0407432].

[47] P. S. Corasaniti, E. J. Copeland, Phys. Rev. **D67** (2003) 063521. [astro-ph/0205544].

[48] J. Weller and A. M. Lewis, Mon. Not. Roy. Astron. Soc. **346** (2003) 987 [arXiv:astro-ph/0307104].

[49] R. Bean and O. Dore, Phys. Rev. D **69** (2004) 083503 [arXiv:astro-ph/0307100].

[50] D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **170** (2007) 377 [arXiv:astro-ph/0603449].

[51] G. B. Zhao, J. Q. Xia, M. Li, B. Feng and X. Zhang, Phys. Rev. D **72** (2005) 123515 [arXiv:astro-ph/0507482].

[52] W. Fang, W. Hu and A. Lewis, Phys. Rev. D **78** (2008) 087303 [arXiv:0808.3125 [astro-ph]].

[53] H. Li, J. Q. Xia, G. B. Zhao, Z. H. Fan and X. Zhang, Astrophys. J. **683** (2008) L1 [arXiv:0805.1118 [astro-ph]].

[54] A. Lewis and S. Bridle, Phys. Rev. D **66** (2002) 103511 [arXiv:astro-ph/0205436].