Nonparametric Reconstruction of the Dark Energy Equation of State

Tracy Holsclaw,[1] Ujjaini Alam,[2] Bruno Sansó,[1] Herbie Lee,[1] Katrin Heitmann,[2] Salman Habib,[3] and David Higdon[4]

[1]*Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064*
[2]*ISR-1, MS D466, Los Alamos National Laboratory, Los Alamos, NM 87545*
[3]*T-2, MS B285, Los Alamos National Laboratory, Los Alamos, NM 87545*
[4]*CCS-6, MS F600, Los Alamos National Laboratory, Los Alamos, NM 87545*

(Dated: September 17, 2009)

The major aim of ongoing and upcoming cosmological surveys is to unravel the nature of dark energy. In the absence of a compelling theory to test, a natural approach is to first attempt to characterize the nature of dark energy in detail, the hope being that this will lead to clues about the underlying fundamental theory. A major target in this characterization is the determination of the dynamical properties of the dark energy equation of state w. The discovery of a time variation in $w(z)$ could then lead to insights about the dynamical origin of dark energy. This approach requires a robust and bias-free method for reconstructing $w(z)$ from data, which does not rely on restrictive expansion schemes or assumed functional forms for $w(z)$. We present a new nonparametric reconstruction method for the dark energy equation of state based on Gaussian Process models. This method reliably captures nontrivial behavior of $w(z)$ and provides controlled error bounds. We demonstrate the power of the method on different sets of simulated supernova data. The GP model approach is very easily extended to include diverse cosmological probes.

PACS numbers: 98.80.-k, 02.50.-r

I. INTRODUCTION

The discovery of the accelerated expansion of the Universe [1, 2] poses perhaps the greatest puzzle in fundamental physics today. A solution of this problem will profoundly impact cosmology and could also provide key insights in reconciling gravity with quantum theory. Driven by these motivations, the fundamental aim of ground and space based missions such as the the Sloan Digital Sky Survey III, the Dark Energy Survey, the Joint Dark Energy Mission (JDEM), the Large Synoptic Survey Telescope – to name just a few – is to unravel the secret of cosmic acceleration. In search of the underlying explanation, theoretical approaches fall into two main categories: (i) dark energy – invoking a new substance, the simplest being a cosmological constant, and (ii) modified gravity – invoking new dynamics of space-time.

A fundamental difficulty in dark energy investigations is the absence of a single *compelling* theory to test against observations. Data analysis efforts therefore focus on characterizing w and its time-dependence. A key objective of upcoming surveys is to determine the evolution of the dark energy equation of state $w = -p/\rho$ (p=pressure, ρ=density). Observations are consistent with a cosmological constant, Λ, ($w = -1$), at the 10% level, the time-variation being unconstrained (for recent constraints on w, see e.g. Ref. [3]). The implied value of Λ is in utter disagreement with estimates of the vacuum energy, being too small by a factor $> 10^{60}$. It is therefore an adhoc addition with no hint of a possible origin, hence the focus on dynamical explanations, e.g., field theory models or modified gravity. The dynamical imprints on observations must necessarily be subtle, otherwise they would have been discovered already.

It is essential to constrain the behavior of w in a nonparametric way and avoid biasing of results due to specific assumptions regarding its functional form. It was first pointed out in Ref. [4] that a reconstruction program for dark energy directly from observational data is indeed possible. This work was followed by a large number of papers suggesting many different ways of reconstructing diverse properties of dark energy. For a recent review on dark energy reconstruction methods, see, e.g., Ref. [5].

The common method currently used to constrain the evolution of w is to employ simple parametrizations, e.g., $w = w_0 + w_1 z$ [6] or $w = w_0 - w_1 z/(z+1)$ [7, 8]. These have obvious shortcomings due to lack of generality and error control. Recently, principal component analysis (PCA) has become popular (see, e.g., Refs. [9, 10]). In this approach, $w(z)$ is written in terms of a compact set of (uncorrelated) principal components, the number of the components depending on the data quality (better data usually implies more components).

In the current paper, we propose a new, nonparametric reconstruction approach for $w(z)$ based on Gaussian Process (GP) models. GP modeling is a nonparametric regression approach particularly well suited for interpolation of smooth functions. The GP is simply a generalization of the Gaussian probability distribution, extending the notion of a Gaussian distribution over scalar or vector random variables to function spaces. While a Gaussian distribution is specified by a scalar mean μ or a mean vector and a covariance matrix, the GP is specified by a mean function and a covariance function. GPs have been successfully applied in astrophysics and cosmology to construct prediction schemes for the dark matter power spectrum and the cosmic microwave background (CMB) temperature angular power spectrum [11–

13], to model asteroseismic data [14], and to derive photometric redshift predictions [15]. Here we will use the GP modeling approach to directly model $w(z)$ from supernova data. Supernova data hold by far the most information about possible time dependence of $w(z)$, though baryon acoustic oscillation and CMB measurements hold complementary information (see, e.g., Ref. [16] for a recent combined reconstruction analysis). Our approach can be very easily extended to accommodate more than one observational probe. For clarity we will restrict ourselves in this paper to supernova measurements only, a more inclusive approach will be developed in future work.

Since current data quality does not allow to place strong constraints on a possible redshift dependence of $w(z)$, we create a set of simulated data of JDEM quality to demonstrate our new method. We consider a constant equation of state and two models with varying $w(z)$. Our approach performs extremely well in capturing nontrivial deviations from a constant equation of state.

The paper is organized as follows. In Section II we provide a brief overview on how supernova data are used to constraint the equation of state of dark energy. We describe the simulated data sets and their error properties in Section III. In Section IV we introduce different reconstruction methods and describe in detail our new approach. We present our results in the same section, contrasting our method with the most commonly used parametric approach by Chevallier & Polarski and Linder [7, 8]. We conclude in Section V.

II. MEASURING THE EXPANSION HISTORY OF THE UNIVERSE WITH SUPERNOVAE

Currently, supernova measurements are the best source of information about possible deviation of $w(z)$ from a constant. Supernovae allow us to measure the luminosity distance D_L which is directly connected to the expansion history of the Universe described by the Hubble parameter $H(z)$. For a spatially flat Universe, the relation is given by:

$$H(z) = \frac{\dot{a}}{a} = \frac{1}{c}\left[\frac{d}{dz}\left(\frac{D_L(z)}{1+z}\right)\right]^{-1}, \qquad (1)$$

where $a(z)$ is the scale factor and c is the speed of light. The assumption of spatial flatness is well justified from CMB observations. It is convenient to scale out H_0 ($h(z) = H(z)/H_0$) leading to:

$$h(z) = \frac{H_0}{c}\left[\frac{d}{dz}\left(\frac{D_L(z)}{1+z}\right)\right]^{-1}, \qquad (2)$$

or equivalently:

$$D_L(z) = \frac{c(1+z)}{H_0}\int_0^z \frac{ds}{h(s)}. \qquad (3)$$

Instead of $D_L(z)$, the data is usually given in terms of the distance modulus μ as function of redshift. The relation between μ and the luminosity distance is given by

$$\mu(z) = m - M = 5\log_{10} D_L(z) + 25. \qquad (4)$$

Writing out the expression for the Hubble parameter $h(z)$ in Eqn. (3) explicitly in terms of a general equation of state leads to

$$\mu(w(z), z) = 25 + 5\log_{10}\left\{\frac{c(1+z)}{H_0}\int_0^z ds(1+s)^{-\frac{3}{2}} \times \left[\Omega_m + (1-\Omega_m)e^{3\int_0^s \frac{w(u)}{1+u}du}\right]^{-\frac{1}{2}}\right\}. \qquad (5)$$

Currently, the quality of supernova data is not good enough to reconstruct the equation of state beyond a cosmological constant. The error bars have to be improved and the number of supernovae especially at high redshifts z needs to be larger to get firm constraints on a possible variation in w (for a recent discussion, see, e.g. [17]). Nevertheless, future supernova surveys, especially space based, hold the promise to change this. In the following we explore a new method on constraining possible redshift dependence in $w(z)$ using simulated data. These data mimic the expected quality of future space based observations. Following Ref. [17], we will use Eqn. (5) as foundation for our analysis.

III. DESCRIPTION OF THE SYNTHETIC DATA SETS

In this section we introduce three synthetic data sets which we will use to gauge how the GP approach compares to more conventional methods to estimate $w(z)$. Synthetic data sets have three major advantages: (i) We know the truth and therefore can get a quantitative measure on how well each method performs. (ii) We can control the data quality. The errors for current supernova data are too large to allow us to distinguish models of dark energy other than a cosmological constant. We can mimic the data quality which will be available from future space-based supernova surveys. (iii) We can investigate dark energy models with very different equations of state $w(z)$.

All of the data sets have space mission quality, we assume the measurement of $n \simeq 2000$ supernovae, distributed over a redshift range of $z = 0 - 1.7$ with larger concentration of supernovae in the mid-range redshift bins ($z = 0.4 - 1.1$). For each supernova, we provide a measurement for the distance modulus μ_i and we assume a statistical error of $\tau_i = 0.06$, as expected from future surveys such as JDEM [18]. We represent the measured points in the following form:

$$\mu_i = \alpha(z_i) + \epsilon_i. \qquad (6)$$

In this notation, ϵ_i encapsulates the distribution of the error terms which is in our case a normal distribution

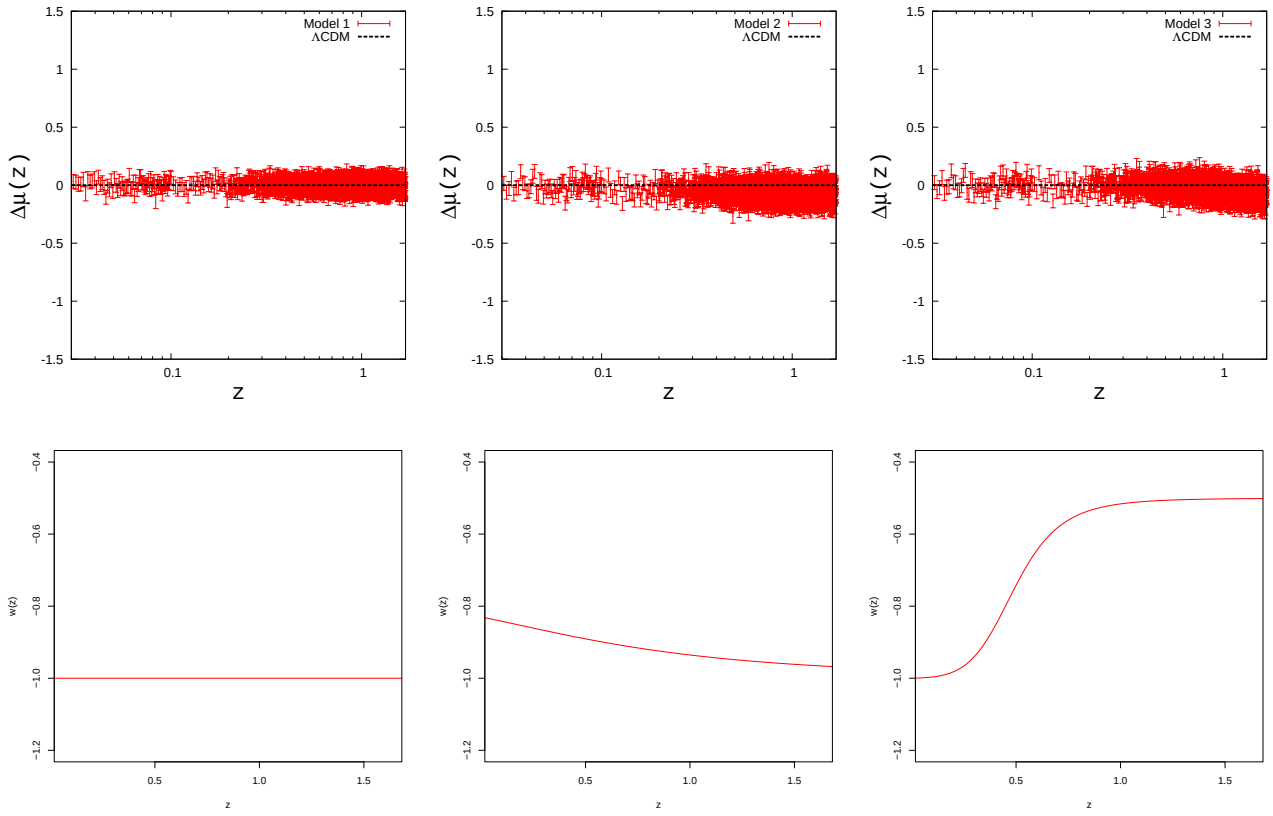

FIG. 1: Three simulated data sets. The upper row shows $\Delta\mu$ (the data itself minus the corresponding value for a ΛCDM model) as a function of redshift z, the lower panels show the corresponding behavior of the equation of state $w(z)$ as a function of redshift. The first model we consider is a cosmological constant, the second model is based on a quintessence model. The third data set, also based on a quintessence model, has been chosen to test our method on a non-trivial equation of state.

with mean zero and a standard deviation $\tau_i\sigma$. τ_i is the observed error and σ accounts for a possible rescaling of the error. The observations μ_i follow also a normal distribution with mean $\alpha(z_i)$ and standard deviation $\tau_i\sigma$. In addition, we assume that the errors are independent. For each of the data sets we fix $\Omega_0 = 0.27$ and $H_0 = 72.0$ (km/s)/Mpc. The three simulated data sets with errors and the corresponding equation of state are shown in Figure 1.

Data Set 1: The first data set is simply a constant equation of state, $w = -1$.

Data Set 2: The second data set is based on a quintessence model with a minimally coupled scalar field with the equation of motion $\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0$ and the potential $V(\phi) = V_0\phi^{-2}$ [19]. The equation of state is given by

$$w = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}. \tag{7}$$

This model leads to a slight variation in the equation of state as a function of z as can be seen in the middle panel in the lower row in Figure 1.

Data Set 3: The third model is a variable dark energy model with the following equation of state [20]:

$$\begin{aligned} w(z) &= w_0 + (w_m - w_0)\frac{1 + e^{\frac{1}{\Delta_t(1+z_t)}}}{1 - e^{\frac{1}{\Delta_t}}} \\ &\times \left[1 + \frac{e^{\frac{1}{\Delta_t}} - e^{\frac{1}{\Delta_t(1+z_t)}}}{e^{\frac{1}{\Delta_t(1+z)}} + e^{\frac{1}{\Delta_t(1+z_t)}}}\right], \end{aligned} \tag{8}$$

with the values $w_0 = -1.0$, $w_m = -0.5$, $z_t = 0.5$, $\Delta_t = 0.1$. Model 3 has $w > -1$ everywhere, so it can be realized by a quintessence field. We choose this third data set because it cannot easily be fit by any of the currently used parametric reconstruction methods. The equation of state has an S-shaped form shown in the right lower panel in Figure 1.

The upper panels in Figure 1 demonstrate impressively the difficult task ahead: shown are the differences $\Delta\mu$ for each data set with respect to a ΛCDM model with $w = -1$. The deviations from the straight line are marginal. We will demonstrate in the following that the GP model is an excellent approach which enables us to pick out these marginal differences and reconstruct the dark energy equation of state reliably.

IV. RECONSTRUCTION OF THE DARK ENERGY EQUATION OF STATE

Measuring the equation of state of dark energy seems to be currently the most promising venue to gain at least some understanding of the nature of dark energy. Almost any dynamical origin of dark energy, such as a quintessence field, would lead to a time variation in $w(z)$. Unfortunately, we cannot measure the equation of state directly. Instead, we measure the luminosity distance redshift relation given in Eqn. (5) from supernovae which contains information about $w(z)$. In order to extract the temporal behavior of the equation of state, one approach would be to fit the measurements of μ as a function of z and extract $w(z)$ by taking two derivatives. Due to the noise in the data, this approach is inapplicable and even if the data is smoothed before fitting them, results are usually unsatisfactory.

Another approach is to assume a certain parametric form for $w(z)$. For example, if we assume w=const., the integral over $w(z)$ in Eqn. (5) can be solved analytically and the best-fit value for w can then be determined from measurements of μ via e.g. a χ^2-minimization. Current data is in good agreement with a constant w at the 10% level (for the most recent analysis see Ref. [3] and references therein for earlier results). The next step is to assume a weak redshift dependence of $w(z)$. One way to realize this is an expansion of $w(z)$ in its redshift evolution of the form $w = w_0 + w_1 z$, which was suggested in, e.g., Ref. [6]. It was pointed out in Ref. [8] that this parametrization is not well suited for $z > 1$ which is the regime that holds the most promise to distinguish different models of dark energy. Ref. [8] therefore proposed a different parametrization for w of the form $w = w_0 - w_1 z/(z+1)$ which had been already suggested two years earlier in Ref. [7]. This parametrization has several nice features: it is well behaved beyond $z = 1$, it still has only two parameters and therefore is relatively easy to constrain, and it captures in general the behavior of different classes of dynamical dark energy models. The major disadvantage is that the parametrization will allow only to reconstruct a monotonic behavior of $w(z)$. More involved parametrizations have been suggested to address this problem. For an overview we refer the reader to the reviews [5, 21] and references therein.

Nonparametric reconstruction methods have been studied less so far, in part because the current data quality does not require more sophisticated methods for reconstruction. Nevertheless, with future data quality in mind, nonparametric method will be much more powerful to extract information about $w(z)$. Nonparametric models will be able to capture more complex behavior of $w(z)$ and should prevent bias due to a restricted parametrization. Currently, the most popular method is to use a principal component analysis. This method has been used recently by the JDEM Figure of Merit Science Working Group [22] to access the performance of JDEM with respect to constraining the dark energy equation of state and for analyzing recent supernova, large scale structure, and CMB measurements (see Ref. [23] for the latest analysis of currently available data and references therein).

In this paper we will study the ansatz w=const. and the parametrization suggested in Refs. [7, 8] to compare them with the GP model approach. We will show results for the reconstructed equation of state as a function of redshift. We first assume that we know the values for Ω_0 and H_0 exactly. While this would be of course not true for real data, this is mainly to simplify the analysis. In the next step, we drop this assumption and include these parameters in our estimations.

A. Parametric Reconstruction

For our parametric reconstruction study we use a Bayesian analysis approach following Ref. [24]. We focus the analysis on two of the most popular models for w: $w = \text{const.} = a$ and $w(z) = a - b\frac{z}{z+1}$. We use Markov Chain Monte Carlo (MCMC) algorithms to perform the analysis [25]. This results in posterior estimates and probability intervals for Ω_m, H_0, and any variables that are needed in the parametric form being assumed for $w(z)$. We use consistent priors in all of our models (including the GP model described in the next section) so the results are readily comparable:

$$\pi(a) \sim U(-25, 1), \tag{9}$$
$$\pi(b) \sim U(-25, 25), \tag{10}$$
$$\pi(H_0) \sim N(72, 1^2), \tag{11}$$
$$\pi(\Omega_m) \sim N(0.27, 0.03^2), \tag{12}$$
$$\pi(\sigma^2) \sim IG(10, 9), \tag{13}$$

and the likelihood

$$L(\sigma, \theta) \propto \left(\frac{1}{\tau_i \sigma}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\mu_i - \mu(z_i,\theta)}{\tau_i \sigma}\right)^2}, \tag{14}$$

where θ encapsulates the cosmological parameters to be constraint, i.e. a subset or all of $\{a, b, H_0, \Omega_m\}$. Some comments about the notation: the "$\sim$" simply means "distributed according to". U is a uniform prior, N is a Gaussian (or normal distributed) prior. The square at the second parameter in $N(a, b^2)$ is used to indicate that b is the standard deviation. (There is no standard convention about the second parameter after the comma, it could be also the variance. The square makes it very clear that we mean the standard deviation.) IG is an inverse Gamma distribution prior. (The mean of a Gamma distribution $\Gamma(a, b)$ is given by a/b and the standard deviation by $\sqrt{a/b^2}$.)

1. A Constant Equation of State

The simplest assumption we can make for $w(z)$ is that it is redshift independent. In this case, Eqn. (5) simplifies

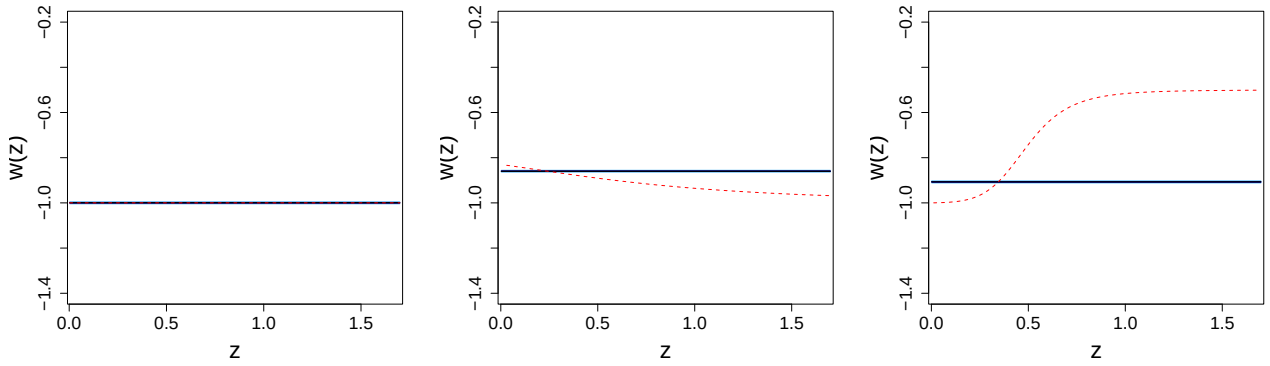

FIG. 2: Reconstruction results for w for data sets 1-3 (left to right) assuming w =const. The red dashed line shows the truth, the black line the reconstruction results, the dark blue shaded region indicates the 68% convidence level, while the light blue shaded region extends to 95%. The assumption w =const. makes it obviously impossible to capture any time dependence in w in data sets 2 and 3. It is interesting to note that the best fit model is highly influenced by the value of w at lower redshifts.

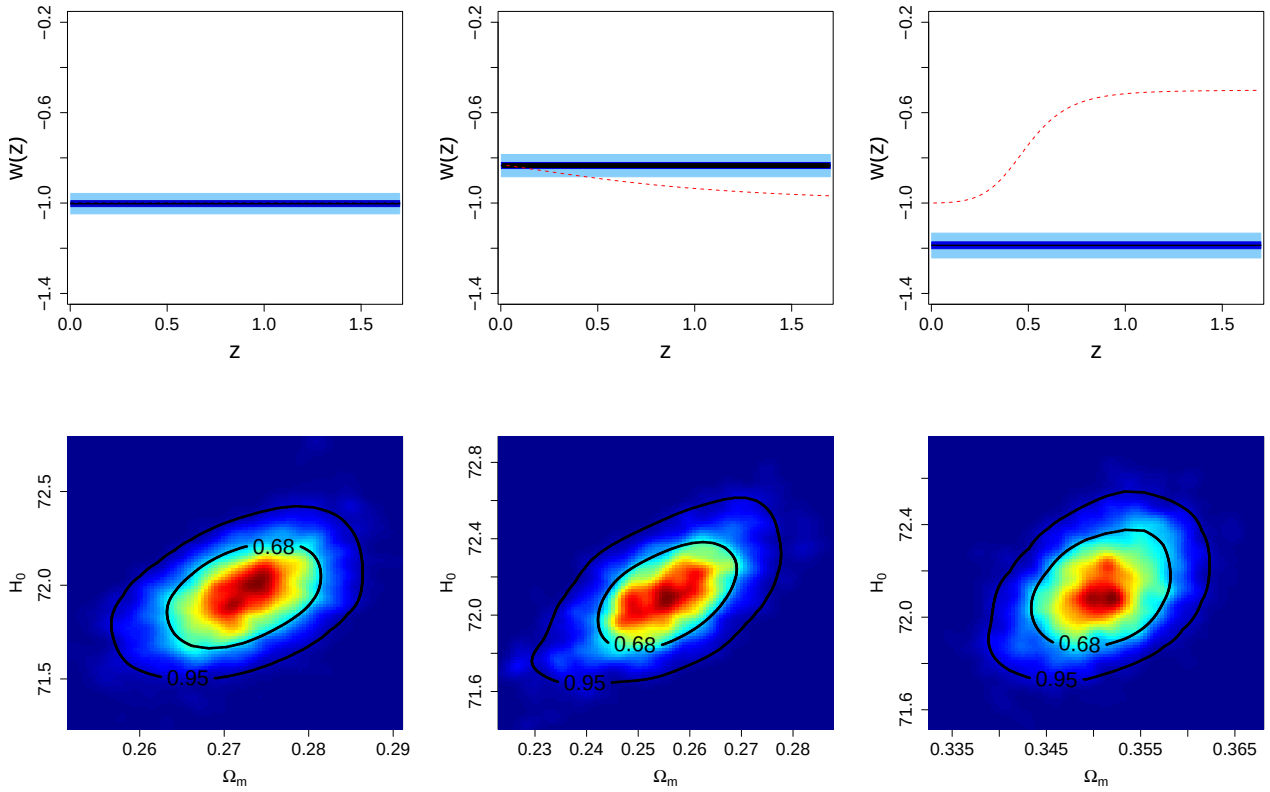

FIG. 3: Upper row: same as in Figure 2, but this time we include uncertainties in the knowledge of Ω_m and H_0. The lower row shows the 68% and 95% confidence levels for the two cosmological parameters Ω_m and H_0 for the three data sets. The result for data set 1 is very accurate – the predictions for w, Ω_m, and H_0 are close to the truth. As before, the predictions for data sets 2 and 3 are not very good for w and Ω_m is biased towards an incorrect value too.

to

$$\mu(w,z) = 25 + 5\log_{10}\left\{\frac{c(1+z)}{H_0}\int_0^z ds\left[\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^3(1+s)^{3w}\right]^{-\frac{1}{2}}\right\}. \quad (15)$$

Current data are in very good agreement with this assumptions. We will use the ansatz for w=const.=a as a first test and attempt to reconstruct all three data sets. As discussed earlier we use an MCMC algorithm for our analysis. We run the chain about 10,000 times reaching

TABLE I: w=const - 95% PIs

Set	a	Ω_m	H_0	σ^2
1	$-1.000^{+0.006}_{-0.005}$	0.27	72	$1.00^{+0.05}_{-0.05}$
2	$-0.860^{+0.005}_{-0.005}$	0.27	72	$1.02^{+0.05}_{-0.05}$
3	$-0.907^{+0.005}_{-0.005}$	0.27	72	$1.13^{+0.05}_{-0.06}$
1	$-1.003^{+0.031}_{-0.033}$	$0.272^{+0.010}_{-0.010}$	$71.95^{+0.32}_{-0.32}$	$1.00^{+0.05}_{-0.05}$
2	$-0.834^{+0.049}_{-0.050}$	$0.255^{+0.016}_{-0.016}$	$72.10^{+0.31}_{-0.33}$	$1.02^{+0.06}_{-0.05}$
3	$-1.187^{+0.054}_{-0.056}$	$0.351^{+0.008}_{-0.008}$	$72.12^{+0.30}_{-0.28}$	$1.02^{+0.05}_{-0.05}$

TABLE II: $w = a - bz/(1+z)$ - 95% PIs

Set	a	b	Ω_m	H_0	σ^2
1	$-1.009^{+0.026}_{-0.027}$	$-0.056^{+0.150}_{-0.149}$	0.27	72	$1.00^{+0.05}_{-0.05}$
2	$-0.830^{+0.022}_{-0.022}$	$0.171^{+0.128}_{-0.124}$	0.27	72	$1.02^{+0.05}_{-0.05}$
3	$-1.099^{+0.022}_{-0.023}$	$-1.047^{+0.118}_{-0.117}$	0.27	72	$1.02^{+0.06}_{-0.05}$
1	$-1.001^{+0.050}_{-0.047}$	$-0.061^{+0.571}_{-0.484}$	$0.266^{+0.031}_{-0.037}$	$71.97^{+0.37}_{-0.36}$	$1.00^{+0.05}_{-0.05}$
2	$-0.832^{+0.048}_{-0.046}$	$0.044^{+0.429}_{-0.366}$	$0.257^{+0.034}_{-0.042}$	$72.09^{+0.36}_{-0.37}$	$1.02^{+0.06}_{-0.05}$
3	$-1.165^{+0.069}_{-0.072}$	$-1.055^{+0.442}_{-0.313}$	$0.285^{+0.039}_{-0.042}$	$72.35^{+0.30}_{-0.30}$	$1.02^{+0.06}_{-0.05}$

convergence very quickly within the first one hundred iterations.

Figure 2 shows the results for the case where we fix Ω_m and H_0 to their known values. As to be expected, the reconstruction works extremely well for the model where w=const. (left panel). The best fit value for a is given in Table I and is very close to the truth with small error bars. Also not surprising, the results for the models which have a time varying w are rather inaccurate. The best fit value for a in both cases seems to be more guided by the value of $w(z)$ at low redshift, while one might have expected it to settle more on the average value for $w(z)$. On the other hand, there are many more data points at low z which most likely led to this result.

In the next step, we do not fix Ω_m and H_0 at their known values but include them in the analysis. The assumed priors are given in Eqs. (12) and (11). The results for w (including the truth) and the confidence levels for Ω_m and H_0 are shown in Figure 3. The best fit values including error bars are given in Table I. Since Ω_m and a are highly correlated they have to be sampled jointly with a covariance structure obtained after running the process for some time. As in the case of Ω_m and H_0 fixed, the analysis works very well for the case of w =const. Not surprising, the error bands are increased but the best fit value is well predicted for all parameters. In the two cases of varying w the strong degeneracy between w and Ω_m becomes very apparent, see the middle and right panel in Figure 3. For the second data set, the estimate for w is higher over the whole z range than the truth, which in turn lowers the prediction for Ω_m considerably. For the third case, the situation is reversed, the prediction for w is much too low leading to an overestimation of Ω_m. In both cases, the prediction for H_0, which is mainly anchored by the amplitude of the measurements for μ is close to the truth. We note that the "truth" for Ω_m and H_0 is of course not exact since we are working with one realization for each data set. For example, for the first data set, the best fit value for H_0 is slightly lower in this realization than the input value.

2. *Chevallier-Polarski-Linder Parametrization*

Next, we investigate a commonly used parametrization of the dark energy equation of state which is given by

$$w(z) = a - b\frac{z}{1+z}. \tag{16}$$

The parametrization was introduce independently by Refs. [7] and [8].

As for the case w=const., one integral in Eqn. (5) can be solved analytically and the expression simplifies to:

$$\mu(a,b,z) = 25 + 5\log_{10}\left\{\frac{c(1+z)}{H_0}\int_0^z ds\left[\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^3(1+s)^{3(a-b+1)}e^{\frac{3bs}{1+s}}\right]^{-\frac{1}{2}}\right\}. \tag{17}$$

This parametrization allows for a weak monotonic time dependence in w and should be able to capture the behavior of our second model reasonably well. Of course with the introduction of a new parameter compared to the first parametrization, error bars will increase. As for the previous case, we first fix Ω_m and H_0 to their known values. Again, 10,000 simulation runs lead to an acceptance for $w(z)$ within 10-40%. The results are summarized in Figure 4 and Table II. For the w=const. data set the parametrization picks up a very small variation in w but the prediction $w = -1$ is well within error bars. The mild variation with z in the second data set is captured rather well with this parametrization. For the third data set the parametrization is not quite flexible enough. While the overall behavior (the rise at high redshift) is captured, the S-shape of the underlying equation of state cannot be extracted from the data. The parametrization is reliably finding a time dependence in w in this case but not the specific form of w which would be important for distinguishing different models of dark energy.

The results including estimations for Ω_m and H_0 are similar. As for the w=const. parametrization, the parameters are all sampled jointly because of their strong correlations. The strong correlations between the parameters degrade the accuracy of the prediction for w somewhat. For the first data set, the prediction for Ω_m is slightly low which in turn amplifies a time dependence in the best fit w which is does not exist in the data set. Again, the error bars are large and clearly $w = -1$ is well

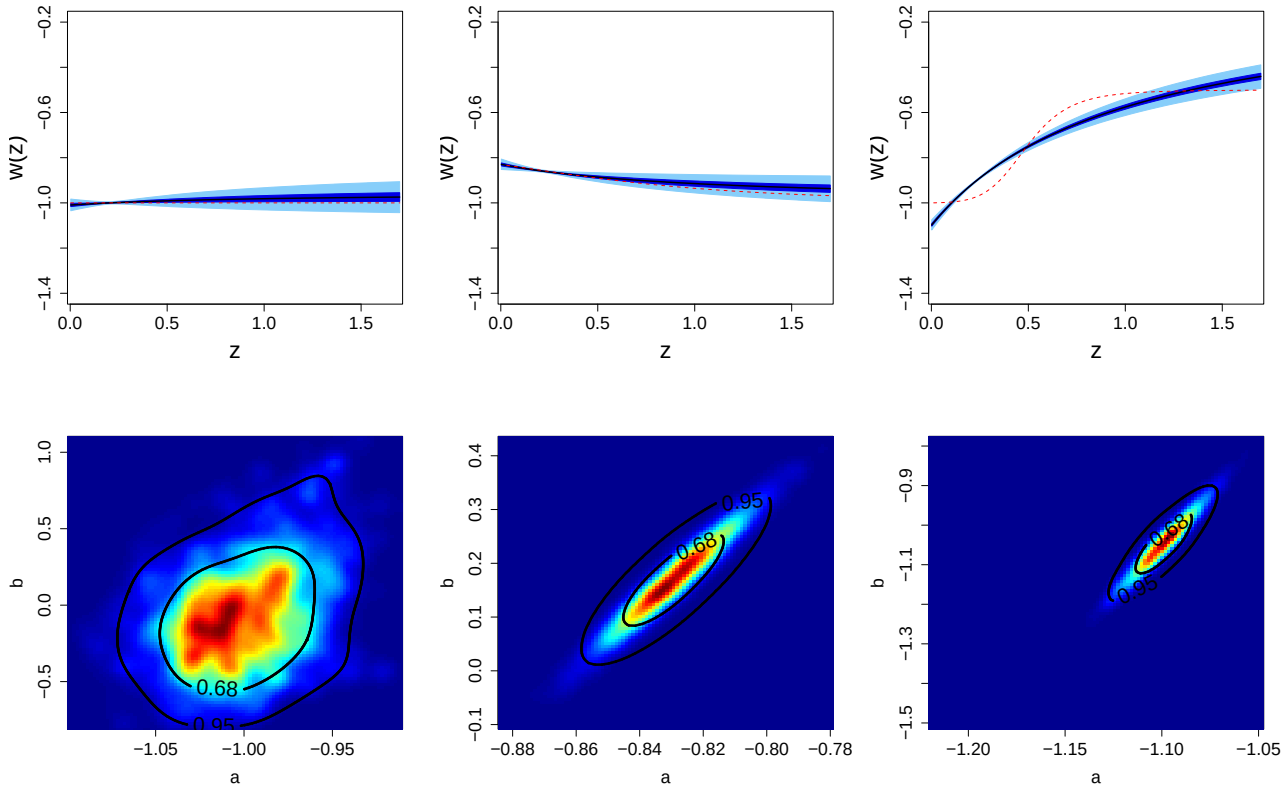

FIG. 4: Upper row: same as in Figure 2, but this time the reconstruction is based on Eqn. (16). The parametrization captures the variation in data set 2 reasonably well, but is still not flexible enough to reconstruct an equation of state with less smooth changes as in data set 3. The lower row shows the 68% and 95% confidence levels for the fitting parameters w_0 and w_1 in Eqn. (16) for the three data sets.

within them. For the second data set, the prediction for Ω_m is even lower. The prediction for b which captures the time dependence of w is too low and the overall time variation of $w(z)$ is underpredicted. For data set 3, Ω_m is overpredicted which leads to a slight degradation in the prediction for w itself.

Overall, the parametrization works rather well, especially for moderately varying w, as to be expected. The draw backs are obvious: sudden changes in w cannot be captured and the data quality has to be very good in order to get reasonable constraints.

TABLE III: GP model - 95% PIs

Set	Ω_m	H_0	σ^2
1	0.27	72	$1.00^{+0.05}_{-0.05}$
2	0.27	72	$1.02^{+0.06}_{-0.05}$
3	0.27	72	$1.13^{+0.06}_{-0.05}$
1	$0.272^{+0.018}_{-0.022}$	$71.94^{+0.40}_{-0.41}$	$1.00^{+0.05}_{-0.05}$
2	$0.260^{+0.020}_{-0.017}$	$72.18^{+0.41}_{-0.44}$	$1.02^{+0.06}_{-0.05}$
3	$0.257^{+0.022}_{-0.020}$	$72.23^{+0.45}_{-0.42}$	$1.02^{+0.06}_{-0.05}$

B. Nonparametric Reconstruction: Gaussian Process Model

After having explored the standard parametric reconstruction methods, we now turn to our new, nonparametric method based on GP modeling [26]. A Gaussian process is a stochastic process such that when sampled at any finite collection of points, the values jointly follow a multivariate normal (MVN) distribution. Thus the process can be defined by its mean and correlation functions. This model is advantageous in that it allows for a flexible fit to a function, $w(z)$, based on probability theory rather than assuming a parametric form like the previous models. We still assume that the errors of the data follow a Gaussian distribution as in the previous models and we use the same likelihoods. We are using Bayesian techniques including the MCMC algorithm which allows us to estimate the parameters of the correlation function along with the other parameters in the model at the same time [25].

For the GP, we assume that $w(z_1), ..., w(z_n)$ for any collection of $z_1, ..., z_n$ follow multivariate Gaussian distributions with mean negative one and exponential co-

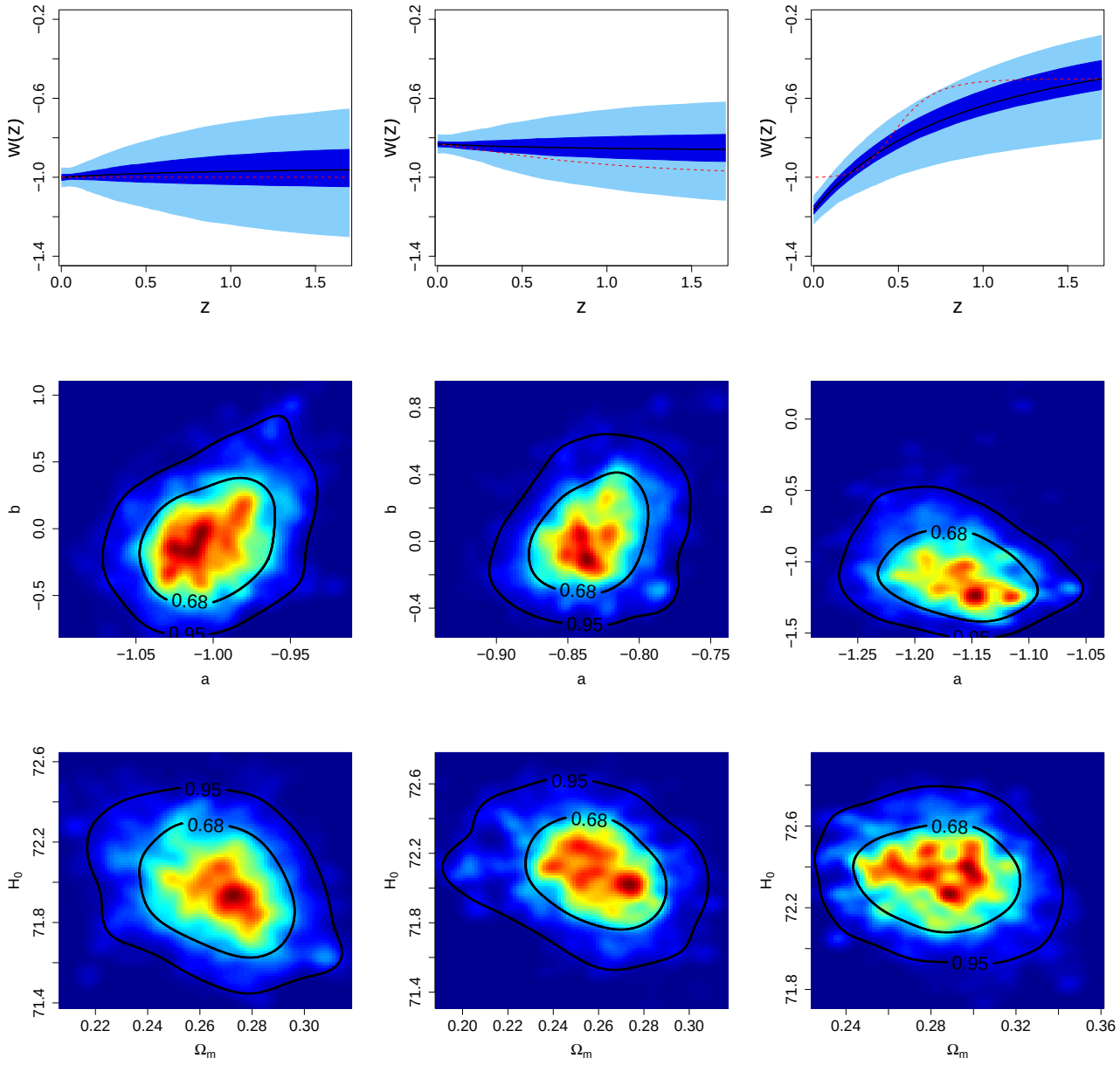

FIG. 5: Upper row: same as in Figure 4 but this time we include uncertainties in Ω_m and H_0. The results are very similar to Figure 4 – though the error bands are as to be expected larger.

variance function written as

$$K(z, z') = \kappa^2 \rho^{|z-z'|^\alpha}. \tag{18}$$

The value for α influences the smoothness of the prediction: for $\alpha = 2$ we will obtain very smooth realizations with infinitely many derivatives, $\alpha = 1$ leads to rougher realizations which are good for modeling continuous non-differentiable functions. In the current paper we will use $\alpha = 1$ to allow for maximum flexibility in reconstructing w. The mean of the GP is fixed in our results but we explored other means and found very similar results; the final choice is made to improve the stability of the MCMC. ρ has a prior of $Beta(6, 1)$ and κ^2 has a vague prior $IG(25, 9)$. As in the other models, Ω_m and H_0 are given priors based on currently available estimates.

Following the notation of Eqn. (5) we set up the following GP for w:

$$w(u) \sim \mathrm{GP}(-1, K(u, u')). \tag{19}$$

Recall that we have to integrate over $w(u)$ (Eqn. 5):

$$y(s) = \int_0^s \frac{w(u)}{1+u} du. \tag{20}$$

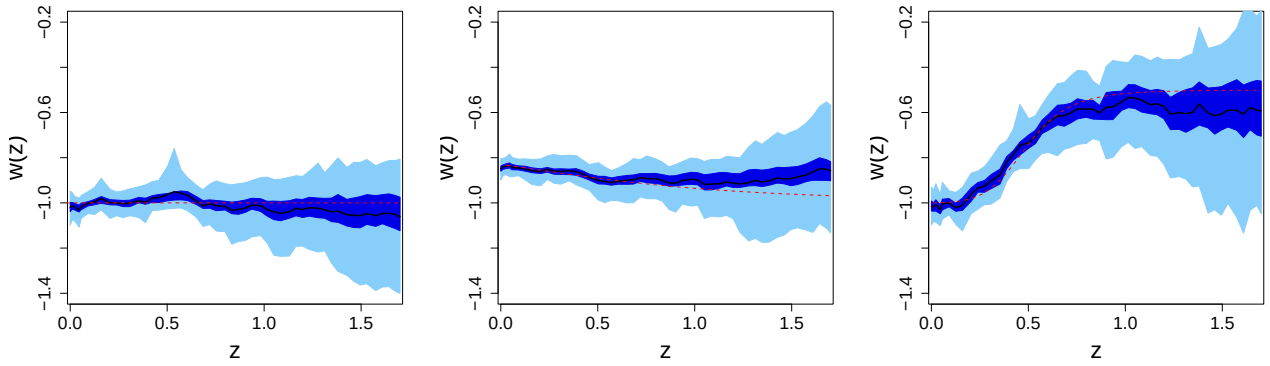

FIG. 6: Same as in Figure 2, but this time the reconstruction is based on the GP model approach. For all three data sets the GP model captures the true behavior of w very well. The results at higher z are slightly worse due to the sparser supernova sampling beyond $z = 1.1$.

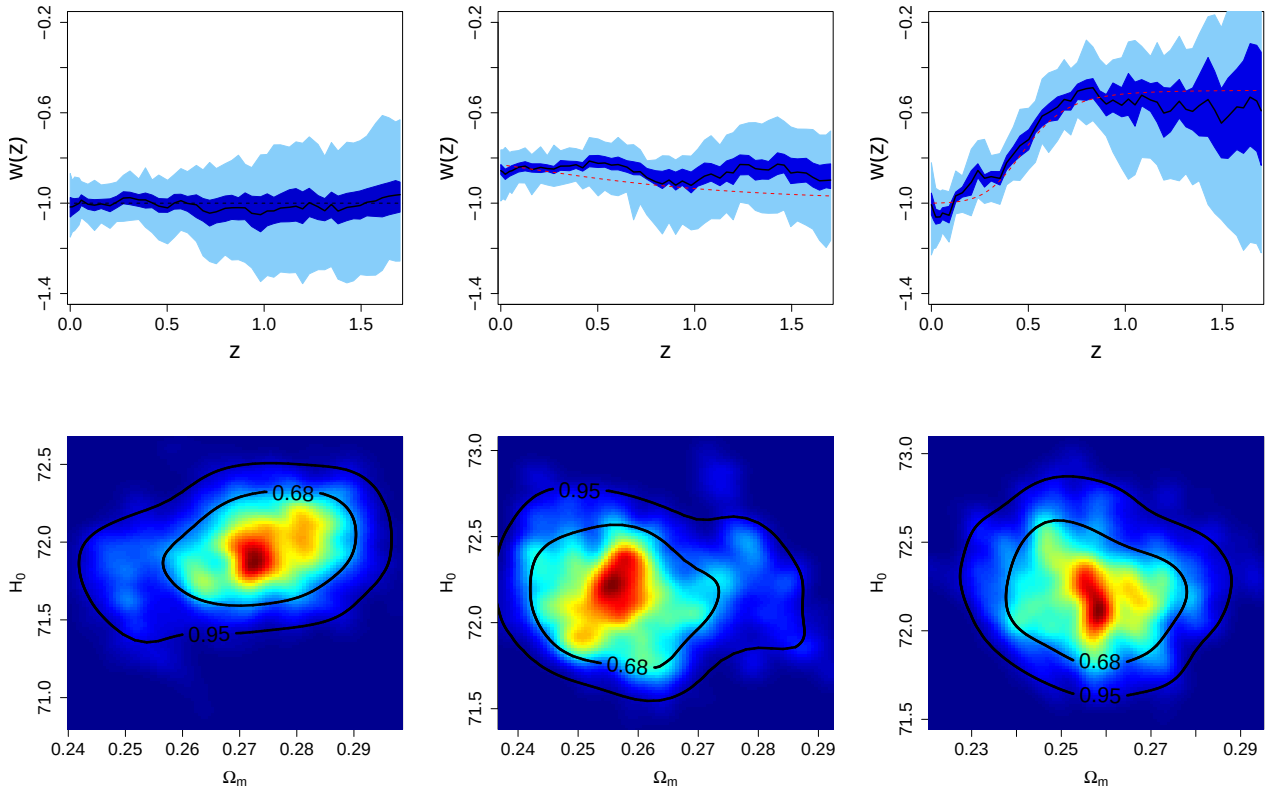

FIG. 7: Upper row: same as in Figure 6 including uncertainties in Ω_m and H_0. The reconstruction again works very well for all three cases.

We use the fact that the integral of a GP is also a GP with mean and correlation dependent on the original GP [26]. The integral of a GP can be found by integrating the correlation function. We therefore set up a second GP for $y(s)$:

$$y(s) \sim \mathrm{GP}\left(-\ln(1+s), \kappa^2 \int_0^s \int_0^{s'} \frac{\rho^{u-u'} du du'}{(1+u)(1+u')}\right). \tag{21}$$

The mean value for this GP is simply obtained by solving the integral in Eqn. (20) for the mean value of the for $w(u)$, negative one. We can now construct a joint GP for

$y(s)$ and $w(u)$:

$$\begin{bmatrix} y(s) \\ w(u) \end{bmatrix} \sim \mathrm{MVN}\left[\begin{bmatrix} -\ln(1+s) \\ -1 \end{bmatrix}\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right], \quad (22)$$

with

$$\Sigma_{11} = \kappa^2 \int_0^s \int_0^{s'} \frac{\rho^{u-u'} du du'}{(1+u)(1+u')}, \quad (23)$$

$$\Sigma_{22} = \kappa^2 \rho^{|z-z'|^\alpha}, \quad (24)$$

$$\Sigma_{12} = \kappa^2 \int_0^s \frac{\rho^{u-u'} du}{(1+u)}. \quad (25)$$

The mean for $y(s)$ given $w(u)$ can be found through the following relation:

$$y(s)|w(u) = -\ln(1+s) + \Sigma_{12}\Sigma_{22}^{-1}\left(w(u) - (-1)\right). \quad (26)$$

Now only the outer integral is left to be solved in Eqn. (5) which can be computed by standard numerical methods. Note that we have never to calculate the double integral in Σ_{11} which would be numerically costly. In addition, the method does not require the inversion of one large covariance method and is therefore rather efficient. More details about each step in the GP model algorithm are given in Appendix A.

As for the parametrized reconstruction methods, we first apply our new method assuming we know the exact values for Ω_m and H_0. The results are shown in Figure 6. The predictions from the GP model for $w(z)$ are remarkably accurate for all three data sets. The slight noise in the predictions is due to the choice of the functional form of the covariance function. In particular, the prediction for the third data set, which was not captured very well by the other two methods, is very good. The GP model approach is able to capture the true behavior of w very well.

Last, we study the results from the GP model including uncertainties in Ω_m and H_0. As for the two parametrizations, degeneracies degrade the results slightly. For the case of w=-1 the prediction for Ω_m is slightly high, leading to a value smaller than $w = -1$. For the second model, we find the opposite: the best fit value for Ω_m is slightly low and the prediction for w above the truth. While for the third data set the best fit value for Ω_m is also on the low side, the GP model approach captures the overall behavior of the true $w(z)$ well. All these results will certainly improve if we include different data sets, e.g., CMB or baryon acoustic oscillation measurements to break the degeneracies. The main point here is to show that the GP model is able to capture non-trivial behavior in $w(z)$ extremely well.

V. CONCLUSIONS

The nature and origin of dark energy is currently a complete mystery. With the lack of a compelling theory to test against, the major aim is to first characterize dark energy w by measuring its equation of state. Field theoretical models of dark energy predict a slight time variation in w and if future surveys could capture such a time dependence we hope to be able to understand or at least learn something about the nature of dark energy. Supernova measurements are a very promising probe for w and future surveys like JDEM promise to measure $w(z)$ with high accuracy.

In order fully exploit such measurements, we need a reliable and robust reconstruction method. In this paper we have introduced a new reconstruction approach based on GP modeling. This approach is a nonparametric method and the modeling parameters are constraint directly from the data. We have demonstrated that we can extract non-trivial behavior of w as a function of redshift with data of JDEM quality. We have also contrasted our new method versus two standard parametrizations. While the z-dependent parametrization leads to good results for capturing small variations in $w(z)$ it cannot reconstruct complex forms of $w(z)$.

The GP model approach outlined in this paper for the analysis of supernova measurements can very easily be extended to include different cosmological probes. This analysis will be carried out in future work. A paper on applying our new method on currently available data is in preparation.

APPENDIX A: GP MODEL ALGORITHM

In this appendix we give a detailed description of the implementation of the algorithm of the GP to extract $w(z)$.

1. Initialize all variables: $\theta = \theta_1$, $\rho = \rho_1$, $\kappa^2 = \kappa_1^2$, and $w^o(u) = w^o_{m,1}(u)$. $w(u)$ will be a vector with m points in our GP and $y(s)$ will have $m \cdot h$ points. We will run this algorithm $q = 1, ..., Q$ times. Set all tuning parameters, $\delta_{1,2,3,4}$, which will need to be tuned until good mixing occurs. Also, all proposals used are symmetric and will not need a jumping function in α_{MH}.
2. Propose $\rho^* = Unif(\rho_1 - \delta_1, \rho_1 + \delta_1)$
   (a) Compute the covariance matrix $K_{22\rho^*} = \rho^{*|u_j - u_i|^\alpha}$

(b) Compute the Cholesky decomposition for $K_{22\rho^*} = U'_{\rho^*}U_{\rho^*}$

(c) Compute the special $K_{12\rho^*} = \int_0^{s'} \frac{\rho^{*|u-s|^\alpha}}{1+u}du$ with Chebyshev-Gauss quadrature.

(d) We want $y_{\rho^*}(s) = \theta_{q-1}\ln(1+s) + [\kappa^2_{q-1}K_{12*}][\kappa^2_{q-1}K^{-1}_{22*}](w_{\rho^*}(u) - \theta_{q-1})$
where: $w_{\rho^*}(u) = [\kappa_{q-1}U'_{\rho^*}]w^o_{m,q-1} + \theta_{q-1}$
$y_{\rho^*}(s) = \theta_{q-1}\ln(1+s) + [\kappa^2_{q-1}K_{12*}][\kappa^2_{q-1}K_{22*}]^{-1}(\left(\kappa_{q-1}U'_{\rho^*}w^o_{m,q-1} + \theta_{q-1}\right) - \theta_{q-1})$
$y_{\rho^*}(s) = \theta_{q-1}\ln(1+s) + \kappa_{q-1}K_{12*}[(U'_{\rho^*}U_{\rho^*})^{-1}U'_{\rho^*}]w^o_{m,q-1}$
$y_{\rho^*}(s) = \theta_{q-1}\ln(1+s) + \kappa_{q-1}K_{12*}[U^{-1}_{\rho^*}]w^o_{m,q-1}$

(e) $L(z_i,\mu_i,\tau_i|w_{\rho^*},\sigma^2_{q-1}) = e^{-\frac{1}{2}\sum\left(\frac{\mu_i - T(z_i,w_{\rho^*}(u))}{\tau_i\sigma_i}\right)^2}$ where the definite integrations in $T(z_i, w_{\rho^*}(u))$ are done numerically through summations of the trapezoid algorithm.

(f) If we accept $\alpha_{MH} = \frac{L_{\rho^*}\pi(\rho^*)}{L_{\rho_{q-1}}\pi(\rho_{q-1})}$ then we will let $\rho_q = \rho^*$

3. Draw $\kappa^{2*} = Unif(\kappa^2_{q-1} - \delta_2, \kappa^2_{q-1} + \delta_2)$

(a) Compute $y_{\kappa^{2*}}(s) = \theta_{q-1}\ln(1+s) + \kappa^* K_{12\rho_q}[U^{-1}_{\rho_{q-1}}]w^o_{m,q-1}$

(b) $L(z_i,\mu_i,\tau_i|w_{\kappa^{2*}},\sigma^2_{q-1}) = e^{-\frac{1}{2}\sum\left(\frac{\mu_i - T(z_i,w_{\kappa^2*}(u))}{\tau_i\sigma_i}\right)^2}$ where the definite integrations in $T(z_i, w_{\kappa^{2*}}(u))$ are done numerically through summations of the trapezoid algorithm.

(c) If we accept $\alpha_{MH} = \frac{L_{\kappa^{2*}}\pi(\kappa^{2*})}{L_{\kappa^2_{q-1}}\pi(\kappa^2_{q-1})}$ then we will let $\kappa^2_q = \kappa^{2*}$

4. We propose a non-standard w^*_m for the GP. We start by drawing a proposal for $w^{o*} \sim MVN(w^o_{q-1}, \delta_3 I_{mxm})$

(a) Compute $y^*(s) = \theta_{q-1}\ln(1+s) + \kappa_q K_{12q}[U^{-1}_q]w^{o*}_{m,}$

(b) $L_{z_i,\mu_i,\tau_i|w^*_{new}(u),\sigma^2_{q-1}} = e^{-\frac{1}{2}\sum\frac{\mu_i - T(z_i,w^*_{new}(u))}{\tau_i\sigma}^2}$

(c) If we accept $\alpha_{MH} = \frac{L_{w^*_{new}(u)}MVN(w^{o*}_m|0,I)}{L_{w_{q-1}}MVN(w_{m,q-1}|0,I)}$ then $w^o_{m,q}(u) = w^{o*}_m(u)$ and the GP realization is $w_{m,q}(u) = w^*_m(u)$

5. Draw $\theta^* = Unif(\theta_{q-1} - \delta_4, \theta_{q-1} + \delta_4)$

(a) Compute $y^*(s) = \theta^*\ln(1+s) + \kappa_q K_{12_q}[U^{-1}_q]w^o_{m,q}$

(b) $L(z_i,\mu_i,\tau_i|w_{\theta^*},\sigma^2_{q-1}) = e^{-\frac{1}{2}\sum\left(\frac{\mu_i - T(z_i,w_{\theta^*}(u))}{\tau_i\sigma_i}\right)^2}$ where the definite integrations in $T(z_i, w_{\theta^*}(u))$ are done numerically through summations of the trapezoid algorithm.

(c) If we accept $\alpha_{MH} = \frac{L_{\theta^*}\pi(\theta^*)}{L_{\theta_{q-1}}\pi(\theta_{q-1})}$ then we will let $\theta_q = \theta^*$

6. $\sigma^2_q|... \sim IG\left(\frac{n}{2} + 2.01, \frac{1}{2}\sum\left(\frac{\mu_i - T(z|...)}{\tau_i}\right)^2 + 1\right)$

7. Repeat steps 2-6, Q times and rerun the entire algorithm as needed after resetting the tuning parameters

Acknowledgments

We would like to thank the Institute for Scalable Scientific Data Management for supporting this work. Part of this research was supported by the DOE under contract W-7405-ENG-36. UA, SH, KH, and DH acknowledge support from the LDRD program at Los Alamos National Laboratory. KH was supported in part by NASA. SH and KH acknowledges the Aspen Center for Physics, where part of this work was carried out. We would like to thank Martin White and Michael Wood-Vasey for useful discussions.

[1] S. Perlmutter *et al.* [Supernova Cosmology Project Collaboration], Astrophys. J. **517**, 565 (1999) [arXiv:astro-ph/9812133].

[2] A. G. Riess *et al.* [Supernova Search Team Collaboration], Astron. J. **116**, 1009 (1998) [arXiv:astro-ph/9805201].
[3] M. Hicken *et al.*, Astrophys. J. **700**, 1097 (2009) [arXiv:0901.4804 [astro-ph.CO]].
[4] A. A. Starobinsky, JETP Lett. **68**, 757 (1998) [Pisma Zh. Eksp. Teor. Fiz. **68**, 721 (1998)] [arXiv:astro-ph/9810431].
[5] V. Sahni and A. Starobinsky, Int. J. Mod. Phys. D **15**, 2105 (2006) [arXiv:astro-ph/0610026].
[6] J. Weller and A. J. Albrecht, Phys. Rev. Lett. **86**, 1939 (2001) [arXiv:astro-ph/0008314].
[7] M. Chevallier and D. Polarski, Int. J. Mod. Phys. D **10**, 213 (2001) [arXiv:gr-qc/0009008].
[8] E. V. Linder, Phys. Rev. Lett. **90**, 091301 (2003) [arXiv:astro-ph/0208512].
[9] D. Huterer and G. Starkman, Phys. Rev. Lett. **90**, 031301 (2003) [arXiv:astro-ph/0207517].
[10] A. Albrecht et al. 2009, arXiv:0901.0721.
[11] K. Heitmann, D. Higdon, C. Nakhleh and S. Habib, Astrophys. J. **646**, L1 (2006) [arXiv:astro-ph/0606154].
[12] S. Habib, K. Heitmann, D. Higdon, C. Nakhleh and B. Williams, Phys. Rev. D **76**, 083503 (2007) [arXiv:astro-ph/0702348].
[13] K. Heitmann, D. Higdon, M. White, S. Habib, B. J. Williams and C. Wagner, arXiv:0902.0429 [astro-ph.CO].
[14] B.J. Brewer and D. Stello, Mon. Not. Roy. Astron. Soc. **395**, 226 (2009) [arXiv:0902.3907].
[15] M. J. Way, L. V. Foster, P. R. Gazis and A. N. Srivastava, arXiv:0905.4081 [astro-ph.IM].
[16] U. Alam, V. Sahni and A. A. Starobinsky, JCAP **0702**, 011 (2007) [arXiv:astro-ph/0612381].
[17] C. Genovese, P. Freeman, L. Wasserman, R. Nichol, and C. Miller Source: Ann. Appl. Stat. Volume 3, Number 1 (2009), 144-178.
[18] G. Aldering et al. [SNAP Collaboration], arXiv:astro-ph/0405232.
[19] B. Ratra and P. J. E. Peebles, Phys. Rev. D **37**, 3406 (1988).
[20] P. S. Corasaniti, B. A. Bassett, C. Ungarelli and E. J. Copeland, Phys. Rev. Lett. **90**, 091303 (2003) [arXiv:astro-ph/0210209].
[21] J. Frieman, M. Turner and D. Huterer, Ann. Rev. Astron. Astrophys. **46**, 385 (2008) [arXiv:0803.0982 [astro-ph]].
[22] A. J. Albrecht *et al.*, arXiv:0901.0721 [astro-ph.IM].
[23] P. Serra, A. Cooray, D. E. Holz, A. Melchiorri, S. Pandolfi and D. Sarkar, arXiv:0908.3186 [astro-ph.CO].
[24] A. Gelman, B. Carlin, H. Stern, and D.Rubin, *Bayesian Data Analysis*, New York: Chapman and Hall (2004).
[25] D. Gameran and H.F. Lopes, *Markov Chain Monte Carlo: Stochastic Simulation for Baysian Inference*, New York: Chapman and Hall (2006).
[26] S. Banerjee, *Hierarchical Modeling and Analysis for Spatial Data*, New York: Chapman and Hall (2004).