Mon. Not. R. Astron. Soc. **380,** 865–876 (2007) doi:10.1111/j.1365-2966.2007.12000.x

Reconstructing the history of dark energy using maximum entropy

Caroline Zunckel★ and Roberto Trotta★
Oxford University, Astrophysics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH

Accepted 2007 May 18. Received 2007 May 8; in original form 2007 February 27

ABSTRACT

We present a Bayesian technique based on a maximum-entropy method to reconstruct the dark energy equation of state (EOS) $w(z)$ in a non-parametric way. This Maximum Entropy (MaxEnt) technique allows to incorporate relevant prior information while adjusting the degree of smoothing of the reconstruction in response to the structure present in the data.

After demonstrating the method on synthetic data, we apply it to current cosmological data, separately analysing Type Ia supernova measurement from the *HST*/GOODS programme and the first-year Supernovae Legacy Survey (SNLS), complemented by cosmic microwave background and baryonic acoustic oscillation data. We find that the SNLS data are compatible with $w(z) = -1$ at all redshifts $0 \leqslant z \lesssim 1100$, with error bars of the order of 20 per cent for the most-constraining choice of priors. The *HST*/GOODS data exhibit a slight (about 1σ significance) preference for $w > -1$ at $z \sim 0.5$ and a drift towards $w > -1$ at larger redshifts which, however, is not robust with respect to changes in our prior specifications. We employ both a constant EOS prior model and a slowly varying $w(z)$ and find that our conclusions are only mildly dependent on this choice at high redshifts.

Our method highlights the danger of employing parametric fits for the unknown EOS, that can potentially miss or underestimate real structure in the data.

Key words: methods: data analysis – dark matter.

1 INTRODUCTION

With the confirmation of the accelerated expansion of the Universe (Riess et al. 1998; Perlmutter et al. 1999; Lange et al. 2001; Hoekstra, Yee & Gladders 2002; Riess et al. 2004; Cole et al. 2005; Astier et al. 2006; Riess et al. 2007; Spergel et al. 2007) comes the inference that the cosmic dynamics are today dominated by a component that competes *against* gravitational collapse of matter and thus must have negative pressure (Frieman et al. 1995; Peebles & Ratra 2003): this has been dubbed dark energy. All observations are presently compatible with dark energy being in the form of Einstein's cosmological constant Λ, a new form of matter energy with equation of state (EOS) $w = \rho/p = -1$. However, it has been shown that an EOS which changes with redshift, $w(z)$, can mimic a cosmological constant and fit the current data if the parametrization of w is assumed to be a constant (Linder 2004; Simpson & Bridle 2006). At the same time, an explanation based on the cosmological constant still suffers from the so-called 'coincidence' and 'fine tuning' problems, and it remains unclear whether selection effects of the kind embodied by anthropic arguments can offer a solution (Starkman & Trotta 2006).

Determining whether dark energy is constant in time or has dynamical properties is one of the most-pressing outstanding questions in cosmology, as witnessed by the multiplication of observational proposals trying to elucidate the question (see e.g. Trotta & Bower 2006, for an overview). We are thus led to question the simple solution of a constant $w(z)$, especially in view of the fact that several recent works have highlighted the dangers of fitting current data by assuming a specific parametrization for the dark energy EOS (Bassett, Corasaniti & Kunz 2004; Linder 2004).

The purpose of this paper is to explore the use of Bayesian statistical techniques based on a maximum-entropy method to investigate the time dependence of dark energy by imposing minimal assumptions on the functional form of $w(z)$. Whenever external (prior) information is used, its impact is clearly expressed by our formalism, making the reconstruction totally transparent. To do so we require information about the expansion of the universe. The quality and quantity of observational data of cosmological relevance is rapidly increasing: Type 1a supernova (SNIa) (see e.g. Riess et al. 2004; Astier et al. 2006) can be calibrated to serve as standard candles (see Nugent et al. 2006, for a recent proposal of using SNeII-P instead). SNIa observations can thus measure the luminosity distance as a function of redshift, $D_{\rm L}(z)$,

$$D_{\rm L}(z) = \frac{c}{100\,h}\frac{1}{(1+z)}\int_0^z \frac{1}{H(x)}\,{\rm d}x \quad ({\rm Mpc}), \tag{1}$$

where the present-day Hubble constant is $H_0 = 100\,h\,{\rm km\,s^{-1}\,Mpc^{-1}}$, c is the speed of light in km s^{-1} and the redshift-dependent Hubble function $H(z)$ can be expressed in terms of the present-day matter energy content of the Universe as (here and in the following

★E-mail: clz@astro.ox.ac.uk (CZ); rxt@astro.ox.ac.uk (RT)

we assume a spatially flat Universe):

$$H^2(z) = (1 - \Omega_{\rm m} - \Omega_{\rm DE})(1+z)^2 + \Omega_{\rm m}(1+z)^3 + \Omega_{\rm DE} \exp\left[3\int_0^z \frac{1+w(z)}{1+z}\,{\rm d}z\right]. \quad (2)$$

Here, $\Omega_{\rm m}$ is the density of matter in units of the critical density today and $\Omega_{\rm DE}$ is the (present-day) dark energy density in units of the critical density. The dark energy density time-evolution is determined by its EOS $w(z)$,

$$\rho_{\rm DE}(z) = \rho_{\rm DE}(0) \exp\left[3\int_0^z \frac{1+w(z)}{1+z}\,{\rm d}z\right]. \quad (3)$$

Observations of the luminosity distance find a powerful geometric complement in the use of 'standard rulers' such as the position of the acoustic peaks in the cosmic microwave background (CMB) power spectrum (Bennett et al. 2003; Spergel et al. 2007) and the (transversal) baryon acoustic oscillation (BAO) signature recently measured in the galaxy matter power spectrum (Cole et al. 2005; Eisenstein et al. 2005). Such data can be used to constrain the angular diameter distance $D_{\rm A}(z)$

$$D_{\rm A}(z) = (1+z)^2 D_{\rm L}(z). \quad (4)$$

Apart from its geometrical impact on the angular diameter and luminosity distance relations, the properties of dark energy also influence the growth of structures and can therefore be constrained through weak lensing (see e.g. Hu 2002; Hoekstra et al. 2006; Jarvis et al. 2006) and cluster counts data. In order to constrain $w(z)$ from measurements of either $D_{\rm L}(z)$ or $D_{\rm A}(z)$, we need to perform two derivatives, as it is evident from equations (2) and (3). This is problematic if we consider the increase in the noise that accompanies each derivative. In addition to information loss through this indirect determination of $w(z)$, the current data tend to be sparse with a large sample variance.

For these reasons, it seems timely and relevant to shift attention to establishing more powerful statistical methods to extract in the most-efficient and faithful way the information on $w(z)$ contained in present and upcoming large data sets. The development of a technique that can cope with the patchy distribution in redshift space while making minimal assumptions on the time properties of dark energy is the logical next step towards improving our understanding of dark energy. In this paper, we apply a modified Maximum Entropy (MaxEnt) technique that has been diversely used to successfully reconstruct images and spectra under unfavourable conditions (for applications to astrophysical problems, see e.g. Bridle et al. 1998; Marshall et al. 2002; Maisinger, Hobson & Lasenby 2004). With a firm basis in probability theory, the method can be tailored to the needs of dark energy reconstruction from present data. Our application of MaxEnt aims at reconstructing the dark energy EOS while minimizing our assumptions regarding the form of $w(z)$.

This paper is organized as follows. We firstly outline the statistical framework of our technique in Section 2. We then proceed to test our reconstruction method on synthetic data in Section 3 and then apply it to present cosmological data in Section 4. We offer our conclusions in Section 5.

2 MAXIMUM-ENTROPY RECONSTRUCTION TECHNIQUE

2.1 Motivation

When attempting to constrain the nature of dark energy, a procedure common in the literature is to Taylor expand the quantity $\rho_{\rm DE}(z)$ or $w(z)$ around $z = 0$, and then constraining the expansion coefficients through the data (Efstathiou 1999; Huterer & Turner 2001; Weller & Albrecht 2002). An example of such a parametrization that is commonly employed is $w(z) = w_0 + (1-a)\,w_1$ where $a = \frac{1}{(1+z)}$ (Chevallier & Polarski 2001; Linder 2003). Alternatively, one might prefer to parametrize the time dependence of the EOS using some smooth function (such as the ones used in Dick, Knox & Chu 2006), that is hoped will encapsulate the essential features of the dynamics one wishes to constrain. Both procedures are not free from the risk of giving misleading results, since they impose artificial assumptions on the form of the EOS, which often have no basis in any physical mechanism. Bassett et al. (2004) highlight the dangers of implementing such parametrizations. Moreover, this will only be sensitive to departures from a constant density within a restricted set of models (Dick et al. 2006).

Huterer & Starkman (2003) introduced a Principal Component Analysis (PCA) of the function $w(z)$, with the aim of adopting a parametrization appropriate to the data sets used (see also Dick et al. 2006; Simpson & Bridle 2006 for related issues). These PCA modes are argued to form a natural basis in which to characterize dark energy evolution and by using only the first few well-determined eigenvectors in the reconstruction one tries to exclude noisy modes and thereby gain accuracy in the reconstruction. However, the method has the disadvantage of introducing an ill-controlled bias at high redshifts, that is, the removal of strongly oscillating (and noisy) modes may mislead one to the conclusion that the EOS reverts to the fiducial model at large redshift with artificially small error bars. While we recognize the merits of the PCA method, we wish to improve on it in this last respect by making the assumptions that will control the behaviour of the reconstructed $w(z)$ at large redshifts explicitly clear.

The MaxEnt technique we employ has parallels with the well-known maximum-likelihood (ML) approach, but introduces new features ensuring that in the case where insufficient information is available the most likely distribution is the most uniform, that is, the one with *maximum entropy* (or minimum information content). For an overview of the connection between entropy and information content, see, for example, Trotta (2005) and Kunz et al. (2006). In other words, where ML merely maximizes the likelihood, often unnecessarily overfitting the noise in the data, MaxEnt seeks the optimum trade-off between a smooth, maximally entropic distribution and the rough distribution mapped out by the data. The most-important characteristic is that the MaxEnt method is auto-regulating, that is, the amount of smoothness (or raggedness) in the reconstruction is consistently determined through the data themselves (see Section 2.2.4 below). In our Bayesian perspective, extra information coming from prior beliefs or theoretical prejudice can be naturally incorporated in the reconstruction via the MaxEnt prior. As we show below, this gives MaxEnt the power to cope with situations where the dimensionality of the parameter space potentially exceeds the number of data points, a difficult reconstruction problem that is ill defined under ML techniques. This feature clearly makes MaxEnt highly applicable to the case of dark energy reconstruction.

2.2 The MaxEnt formalism

The task at hand is to determine the EOS of dark energy from sparse data on $D_{\rm L}$, $D_{\rm A}$ and, to a limited extent, on $H(z)$ (as encapsulated by today's detection of BAO). We consider a piecewise constant $w(z)$ in N bins out to a maximum redshift $z_{\rm max}$. Let w_j be the value of the EOS in the jth bin, $1 \leqslant j \leqslant N$. In analogy with the treatment given in Skilling (1989) for the case of image reconstruction, we

gather all the EOS bins values in an 'image vector' $\boldsymbol{w}$. We seek to determine the posterior distribution of $\boldsymbol{w}$, given the observed data D, $\Pr(\boldsymbol{w} \mid \boldsymbol{D})$. This is obtained through Bayes' theorem as

$$\Pr(\boldsymbol{w} \mid \boldsymbol{D}) = \frac{\Pr(\boldsymbol{D} \mid \boldsymbol{w})\Pr(\boldsymbol{w})}{\Pr(\boldsymbol{D})}. \tag{5}$$

The quantity $\Pr(\boldsymbol{w})$ is the prior probability representing all the information about the distribution $\boldsymbol{w}$ before the data $\boldsymbol{D}$ have been collected; $\Pr(\boldsymbol{D} \mid \boldsymbol{w})$ is the likelihood and describes the underlying statistical process and $\Pr(\boldsymbol{D})$ is the model likelihood ('evidence'), which is relevant for model selection questions but that is unimportant in this case. We will therefore neglect this proportionality constant in the following.

2.2.1 *The MaxEnt prior*

The principle of MaxEnt is employed to determine a prior $\Pr(\boldsymbol{w})$ that encapsulates all the external information about $w(z)$ we wish to specify in the absence of the data. Following Skilling (1989), we adopt the principle that the least-biased model that encodes any given prior information is the one which maximizes the entropy of the distribution while remaining consistent with the information. This prior is appealing for its characteristic of maximizing the uncertainty (entropy) of the distribution thus making minimal assumptions. The MaxEnt prior (Skilling 1989) takes the form

$$\Pr(\boldsymbol{w} \mid \alpha, \boldsymbol{m}) = \frac{\exp(\alpha S(\boldsymbol{w}, \boldsymbol{m}))}{Z_S} \tag{6}$$

where $S(\boldsymbol{w}, \boldsymbol{m})$ is the entropy of $\boldsymbol{w}$ relative to the model $\boldsymbol{m}$ and α is a regularizing constant. The model $\boldsymbol{m}$ defines the image vector to which $\boldsymbol{w}$ reverts in the absence of any data, and as such it defines a measure on the DE parameter space. The information entropy is analogous to the thermodynamic entropy in statistical mechanics, which is given by the logarithm of the number of states by which one can arrive at a given macroscopic constraint. In the same way, the information entropy can be described as the logarithm of the number of ways in which one can arrive at a particular $\boldsymbol{w}$ in a Poisson process when starting with the model $\boldsymbol{m}$ (Silver, Sivia & Gubernatis 1990). The entropy S for an N-dimensional discrete parameter space is (Skilling 1989):

$$S(\boldsymbol{w}, \boldsymbol{m}) = \sum_{j=1}^{N} \left[w_j - m_j - w_j \log\left(\frac{w_j}{m_j}\right) \right]. \tag{7}$$

The log term is reminiscent of the Kullback–Leibler divergence between $\boldsymbol{w}$ and $\boldsymbol{m}$, encoding the amount of information present in $\boldsymbol{w}$ with respect to the model $\boldsymbol{m}$. In our case, we do not apply the entropy to the values of $\boldsymbol{w}$ directly, but rather to the space of coefficients of an expansion of $\boldsymbol{w}$ in a series of basis functions (that we choose to be top-hat functions in redshift space, see Section 2.3 for details). We can think of the coefficients of the expansion as a series of weights that encode how much each basis function contributes to the total $w(z)$. We can then apply the MaxEnt prior on the space of these weights, by thinking of them as expressing the relative contribution of each basis function – in other words, in a phenomenological way we take the weights to represent relative probabilities for the presence of each basis function in the final $w(z)$. Below, we will use the notation $\boldsymbol{w}$ as a shortcut to indicate the vector of weights of the dark energy expansion. The same applies to the model $\boldsymbol{m}$, that in the entropy term is represented by its expansion coefficients in the chosen basis functions.

Evidently, $S(\boldsymbol{w})$ (for a fixed $\boldsymbol{m}$) is a convex function which reaches a maximum for $\boldsymbol{w} = \boldsymbol{m}$ with a value $S = 0$. Thus in the absence of any information from the data, the entropy term reverts to the model. The normalizing partition function for the entropy is given by

$$Z_S = \int \exp \alpha S(\boldsymbol{w}) \det[g]^{1/2} \, \mathrm{d}^N \boldsymbol{w}. \tag{8}$$

The measure is defined as the invariant volume $\det g^{1/2}$ of the metric g defined on the space where $g_{ii} = 1/w_i$ and $g_{ij} = 0$ for $i \neq j$ (also known as the Fisher information matrix). By expanding to second order around the model $\boldsymbol{w} = \boldsymbol{m}$ (at the maximum $S = 0$), we obtain the partition function in the Laplace approximation:

$$Z_S = \left(\frac{\alpha}{2\pi}\right)^{N/2}, \tag{9}$$

where N is the number of parameters, in our case the number of expansion coefficients for $w(z)$.

2.2.2 *The likelihood*

The likelihood is defined as the probability of the *data*, *given the parameters*,

$$\Pr(\boldsymbol{D} \mid \boldsymbol{w}) = \frac{\exp(-\mathcal{L}(\boldsymbol{w}))}{Z_{\mathrm{L}}}, \tag{10}$$

and is the probability that the observed data $\boldsymbol{D}$ could have been generated from a given $\boldsymbol{w}$. For data $\boldsymbol{D}$ subject to Gaussian noise the likelihood function is

$$\mathcal{L}(\boldsymbol{w}) = \frac{1}{2}(\boldsymbol{D} - f(\boldsymbol{w}))^{\mathrm{T}}[C^{-1}](\boldsymbol{D} - f(\boldsymbol{w})), \tag{11}$$

where C is the data covariance matrix and $f(\boldsymbol{w})$ denotes the functional dependence of the observable on the DE parameters in our case, $f = H$ or $f = D_{\mathrm{A}}$ (D_{A} and D_{L} being simply related through the redshift, see 4). In the case of independent data points with uncorrelated noise, the covariance matrix is diagonal with the non-zero elements being the variances of each measurement, denoted by σ_i^2, $i = 1, \ldots, N_D$. The normalizing partition function for $\mathcal{L}$ is

$$Z_{\mathrm{L}} = \int \exp(\mathcal{L}(\boldsymbol{w})) \, \mathrm{d}^{N_D} D. \tag{12}$$

Using the identity for the normalized probability distribution, we obtain

$$Z_{\mathrm{L}} = \frac{(2\pi)^{N_D/2}}{\sqrt{\det[C^{-1}]}}. \tag{13}$$

2.2.3 *The posterior probability*

From the likelihood in equation (10) and the prior in equation (6), we obtain from Bayes' theorem the posterior probability for the DE parameters $\boldsymbol{w}$:

$$\Pr(\boldsymbol{w} \mid \boldsymbol{D}, \alpha, \boldsymbol{m}) \propto \exp(\alpha S(\boldsymbol{w}) - \mathcal{L}(\boldsymbol{w})). \tag{14}$$

Given that $\mathcal{L}(\boldsymbol{w})$ is quadratic in $\boldsymbol{w}$ and $S(\boldsymbol{w})$ is a convex function, the above is well constrained with the peak of the posterior for $\boldsymbol{w}$ being determined by a competition between S and $\mathcal{L}$, mediated by the value of α. We thus see that the MaxEnt prior will be useful in the case where the parameter space dimensionality exceeds the size of the data set in that the entropy is incorporated as a regularization to avoid overfitting, while retaining maximum flexibility in the underlying parametrization of $w(z)$. S penalizes the excess 'structure' in the data, with the regularizing parameter α dictating the degree of this smoothing. The choice of α is thus very important: a small value

of α produces little smoothing and the likelihood will dominate, resulting in a reconstructed distribution where the noise might be mistaken as real structure. Alternatively, too large a value for α leads to information loss, with the entropic prior overriding the information coming from the data.

2.2.4 *The regularization parameter α*

In order to select the correct value of α, we add it to the hypothesis space as an additional parameter and let the data select the optimum value. Using once more Bayes' theorem we have

$$\Pr(\alpha \mid \boldsymbol{D}, \boldsymbol{m}) \propto \Pr(\alpha) \Pr(\boldsymbol{D} \mid \alpha, \boldsymbol{m}), \tag{15}$$

where $\Pr(\alpha)$ is the prior on α. The joint posterior probability is then (up to irrelevant constants):

$$\begin{aligned}\Pr(\boldsymbol{w}, \alpha, \boldsymbol{D}|\boldsymbol{m}) &\propto \Pr(\alpha)\Pr(\boldsymbol{w} \mid \alpha)\Pr(\boldsymbol{D} \mid \boldsymbol{w}) \\ &= \Pr(\alpha)\frac{\exp(\alpha S(\boldsymbol{w}) - \mathcal{L}(\boldsymbol{w}))}{Z_S(\alpha) Z_{\rm L}}.\end{aligned} \tag{16}$$

We adopt a Jeffreys' prior on α, which is flat in $\gamma = \log\alpha$, reflecting ignorance of the scale of the variable, within the range $-10 \leqslant \gamma \leqslant 10$. This corresponds to choosing $\Pr(\alpha) \propto 1/\alpha$. In the final inference on $\boldsymbol{w}$ we marginalize over the nuisance parameter α, even though the distribution of α is usually fairly strongly peaked and thus marginalization is almost equivalent to maximization (i.e. simply fixing α to the value of the peak of the posterior).

2.2.5 *Model specification*

The joint posterior in equation (16) is conditional on the specific choice of model $\boldsymbol{m}$, to which the reconstructed $\boldsymbol{w}$ defaults in the absence of data. The entropic prior distribution is introduced to penalize the posterior for unwarranted complexity. Given that the model is the vector to which $\boldsymbol{w}$ should revert in the absence of data, it must represent maximal smoothness. We need to establish what distribution $\boldsymbol{m}$ will encompass this in the context of the EOS of dark energy. In image reconstruction the default model is usually taken to be a flat surface equal to the mean of the data. When the data are then included via the likelihood, variation about this mean is introduced. In our case, this means choosing a constant model, $\boldsymbol{m}$ = constant that is *uniform* in redshift space. There is, however, no obvious choice for the magnitude of this constant. There are various possible choice of $\boldsymbol{m}$, reflecting different prior beliefs about the dark energy EOS. One can thus usefully think of $\boldsymbol{m}$ as encapsulating a fiducial, reference model we want to test the observations against. One possibility is to set $\boldsymbol{m} = -1$ at all redshifts, thus representing a cosmological constant. This is recommended if we are testing for deviations from $w(z) = -1$: if significant deviations from the model are found in the reconstructed EOS, then this is an indication that the data are informative enough to override the entropic pull towards the model, and thus that such deviations are likely to reflect real structure in the data. A more skeptical attitude towards dark energy might be encapsulated by choosing a constant model $\boldsymbol{m} = 0$, which corresponds to a pressureless, dust-like fluid. In this case, if the reconstructed $w(z)$ assumes values below 0, this can be interpreted as a strong indication for the presence of a fluid with negative pressure, with data being strong enough to dominate the entropic tendency for a pure matter Universe. In principle, a theoretical prejudice in the form of a particular redshift dependence of $\boldsymbol{m}$ could also be implemented easily in the same fashion.

Finally, one can also employ Bayes' theorem to take $\boldsymbol{m}$ into the joint posterior, by writing

$$\Pr(\boldsymbol{w}, \boldsymbol{D}, \alpha, \boldsymbol{m}) \propto \Pr(\boldsymbol{m})\Pr(\boldsymbol{w}, \boldsymbol{D}, \alpha \mid \boldsymbol{m}) \tag{17}$$

and marginalizing over $\boldsymbol{m}$ in the left-hand-side, after specifying a prior over the model space, $\Pr(\boldsymbol{m})$. In this work we take all the model vectors to be constant over the whole redshift range; thus, the specification of the model amounts to the choice of the value of the constant. We restrict our considerations to the range $-1 \leqslant \boldsymbol{m} \leqslant 0$, and when performing a marginalization over the model we will take a flat prior in this range, that is, $\Pr(\boldsymbol{m}) =$ constant.

2.3 Dark energy parametrization and reconstruction

As motivated above, we decompose $w(z)$ into a weighted sum of orthogonal functions in redshift space, with the parameters being given by the weights encoding the amount that each function contributes to the overall $w(z)$. There are, of course, several different meaningful expansion functions such (e.g. principal components, Chebychev functions, etc.) but we make use of the simplest option, decomposing $w(z)$ is into a series of N step-functions $\Phi_i(z)$:

$$w(z) = -2 + \sum_{i=1}^{N} C_i \Phi_i(z), \tag{18}$$

where $\Phi_i(z) = 1$ for $z_i - \Delta z/2 < z < z_i + \Delta z/2$ and $\Phi_i(z) = 0$ everywhere else. Since the least-stringent limits we will impose on $w(z)$ are $-2 \leqslant w_i \leqslant 0$, the above ensures that the expansion coefficients $C_i (i = 1, \ldots, N)$ are positive numbers, a necessary requirement for our entropic prior. The parameter space $\boldsymbol{w}$ is thus constructed from the coefficients C_i themselves, which are allowed to vary within the range $0 \leqslant C_i \leqslant 2$. An advantage of this piecewise parametrization of $w(z)$ is that it will be possible to capture a sharp change in the EOS, provided the binning is sufficient. In order to capture different features of the time-evolution of the EOS, other expansion functions may be more appropriate. We experimented with Chebychev functions and found that their oscillatory behaviour was not helpful in reconstructing sharp changes in the EOS. Such smooth functions might be more useful if one wants to test quintessence models exhibiting a gentle evolution of $w(z)$.

In the bulk of recent analyses, the limit $-1 \leqslant \boldsymbol{w} \leqslant 0$ is imposed; the lower limit stemming from the null energy condition which must be satisfied for dark energy to be stable (Alcaniz 2004). Although models of dark energy that allow $w < -1$ violate the weak energy condition in the context of general relativity, these 'phantom' components have been studied by many authors (Caldwell 2002). There have been claims that such phantom behaviour is unstable when regarded as a quantum field theory (Carroll, Hoffman & Trodden 2003). From a phenomenological perspective, it makes sense both to restrict the range of our reconstruction to lie within $-1 \leqslant \boldsymbol{w} \leqslant 0$, and to extend the parameter space to values below $\boldsymbol{w} = -1$ to check the stability of the reconstruction. We will thus present results also for the case where the EOS can attain values as low as $\boldsymbol{w} = -2$.

In this work we assume flat spatial sections, and thus $\Omega_{\rm m} = 1 - \Omega_{\rm DE}$. The parameters included in the hypothesis space are summarized in Table 1. These are $\Omega_{\rm DE}$, the Hubble parameter today, H_0 in km s^{-1} Mpc^{-1}, and the coefficients of the DE decomposition, C_i, as described by equation (18). This generic characterization requires the number of expansion functions (which can be effectively characterized as top-hat bins) to be sufficiently large for this to be a suitable description of $w(z)$. As described above, we also include

Table 1. Priors on cosmological and nuisance parameters used in the analysis. We employ a Jeffreys' prior on α, that is, we take the prior to be flat in $\gamma = \log\alpha$ to reflect ignorance on the scale of the regularizing parameter α.

Parameter	Prior	
$\Omega_{\rm DE}$	0.0...1.0	Top hat
Ω_κ	0	Flatness imposed
C_i	0...2	Top hat
$\gamma = \log(\alpha)$	$-10 \ldots 10$	Top hat
$\boldsymbol{m}$	$-1 \ldots 0$	Top hat

the nuisance parameters $\gamma = \log\alpha$ and the value of the model EOS, $\boldsymbol{m}$, whenever this is marginalized over.

Assuming uncorrelated Gaussian noise, the log-likelihood of a point in parameter space is given by

$$-2\log\mathcal{L}(\boldsymbol{w}) = \sum_{i=1}^{N_D}\left[\frac{D_i - f(z_i,\boldsymbol{w})}{\sigma_i}\right]^2, \tag{19}$$

where for each datum i at redshift z_i we have $f(z_i,\boldsymbol{w}) \equiv H(z_i)$ for future radial baryonic oscillation measurements, $f(z_i,\boldsymbol{w}) \equiv D_{\rm A}(z_i)$ for present and future transversal baryonic oscillation data and CMB data and $f(z_i,\boldsymbol{w}) \equiv D_{\rm L}(z_i)$ for SNIa data. Furthermore, σ_i^2 is the measurement variance. The Hubble parameter as a function of redshift is obtained via equation (2), where the energy density is calculated using, for z_a falling within the ith bin,

$$\frac{\rho_{\rm DE}(z_a)}{\rho_{\rm DE}(0)} = \left(\frac{1+z_a}{1+z_i-\Delta z_i/2}\right)^{3(1+w_i)} \times \Pi_{j=1}^{i-1}\left(\frac{1+z_j+\Delta z_j/2}{1+z_j-\Delta z_i/2}\right)^{3(w_j+1)}. \tag{20}$$

From the above, the angular diameter distance $D_{\rm A}(z_a)$ can then be computed using equations (1) and (2). For piecewise constant $w(z)$, we employ the trapezoid rule to approximate the integral, obtaining

$$D_{\rm A}(z_a) = \frac{1}{1+z_a}\left\{\frac{\delta z}{4}\frac{1}{H_0} + \frac{1}{H(z_a)} + \sum_{j=2}^{a-1}\frac{\delta z}{2}\left[\frac{1}{H(z_{j-1})} + \frac{1}{H(z_j)}\right]\right\}. \tag{21}$$

The binning of the integral, defined by δz, is determined based on a fixed level of fractional accuracy for the integration, that we set to 10^{-11} as determined by the extrapolation error estimate. Finally, the entropy of a vector $\boldsymbol{w}$ with respect to a model $\boldsymbol{m}$ is given by

$$S(\boldsymbol{w}) = \sum_{i=1}^{N}\left[C_i - M_i - C_i\log\left(\frac{C_i}{M_i}\right)\right], \tag{22}$$

where the M_i, $i = 1, \ldots, N$ are the coefficients of the model $\boldsymbol{m}$ in expansion (18).

To sample the posterior probability distribution efficiently, we use a Monte Carlo Markov Chain (MCMC) which employs a Metropolis algorithm. For more details about MCMC, see, for example, Neal (1992) and Lewis & Bridle (2002). Since the MaxEnt method is designed to achieve the optimal reconstruction independently of the number of degrees of freedom in the parametrization of $w(z)$, we expect that the number of basis functions N will not affect greatly the reconstruction, as long as N is chosen large enough to capture the possible structure in the data. In the following we choose $N = 10$ but we have checked that the results do not vary much if one uses $N = 5$ or 20 instead.

When using actual data, we divide the redshift range spanned by either the Supernovae Legacy Survey (SNLS) or the *HST*/GOODS SN measurements into $N = 10$ equally spaced bins, corresponding to the N basis functions for $w(z)$. We then extend the last bin to cover all of the redshift range to last scattering when computing the angular diameter distance to the CMB. In other words, we take $w(z)$ to be constant (but not fixed to -1) between the redshift of the highest SN in the samples and $z = 1089$. This extrapolation is weaker than the 'strong' prior used in the analysis of Riess et al. (2007), which assumed that $w = -1$ at $z > 1.8$.

3 DEMONSTRATION OF THE MAXENT METHOD

We now proceed to test our MaxEnt reconstruction method with synthetic data. Our benchmark data set consists of $N_D = 10$ measurements of $H(z)$ and $N_D = 10$ of $D_{\rm A}(z)$ [or equivalently, $D_{\rm L}(z)$] distributed uniformly in the redshift range $0 \leqslant z \leqslant 1$. Although the actual measurements will in reality be less homogeneous, this does not represent a problem for our reconstruction algorithm, as we show below. Existing measurements of $D_{\rm L}(z)$ out to $z \sim 2$ will be vastly improved when future surveys such as DES or LSST will be able to observe thousands of SNeIa per year (Abbott et al. 2005; Tyson 2006) and space-based projects such as *SNAP* (Aldering et al. 2004), *ADEPT* or *DUNE* will provide observations at redshifts beyond $z \sim 0.8$. Future spectrographic surveys such as the Wide-Field Multi-Object Spectrograph (WFMOS) or HETDEX ought to deliver constraints on $D_{\rm A}(z)$ of 1 per cent at $z \sim 1$ and 1.5 per cent at $z \sim 3$ (1σ) and $H(z \sim 1)$ to 1.2 per cent (Glazebrook et al. 2005; Kelz et al. 2006), with better performance still to be expected when the Square Kilometer Array will come online (Blake et al. 2004).

We add Gaussian noise to our synthetic data as a fixed percentage of the true value of the measured quantity. We use an optimistic noise level of 1 per cent. Although it is beyond the scope of this paper to make quantitative predictions about the performance of future surveys in reconstructing the EOS, our benchmark data set roughly reflects the potentiality of future observations. We also test the performance of the method when the signal-to-noise ratio level is degraded by a factor of 10, in order to check for bias in the reconstruction due to our entropic prior when the quality of the data are poor. In this case, we use a noise level of 10 per cent in the luminosity distance and Hubble function measurements.

We show in Fig. 1 the reconstructed EOS for our benchmark scenario with high-quality observations ($\sigma = 1$ per cent, $N_D = 20$ observations). In all three panels, we have marginalized over the prior model $\boldsymbol{m}$. We note that the reconstruction is satisfactory in all three cases. We have checked that the $\boldsymbol{w}$ extracted when marginalizing over the prior model $\boldsymbol{m}$ has comparable accuracy to the case of a fixed model $\boldsymbol{m}$. The bottom panel shows how the method deals with gaps in the redshift range of the observations: the smoothing effect of the entropic term enlarges the errors in the region where no data are available, while the reconstructed EOS tracks the true value at small and large redshifts, where the data are clustered. Here we have employed a 10-dimensional $\boldsymbol{w}$, but we have checked that increasing the number of elements to $N = 20$ does not lead to any significant change in the reconstruction. As expected, the error bars in the regions where $w(z)$ is well constrained by observations are considerably smaller.

We now turn to the case where the data are noisy, and hence we expect our entropic prior to play a more important role in the

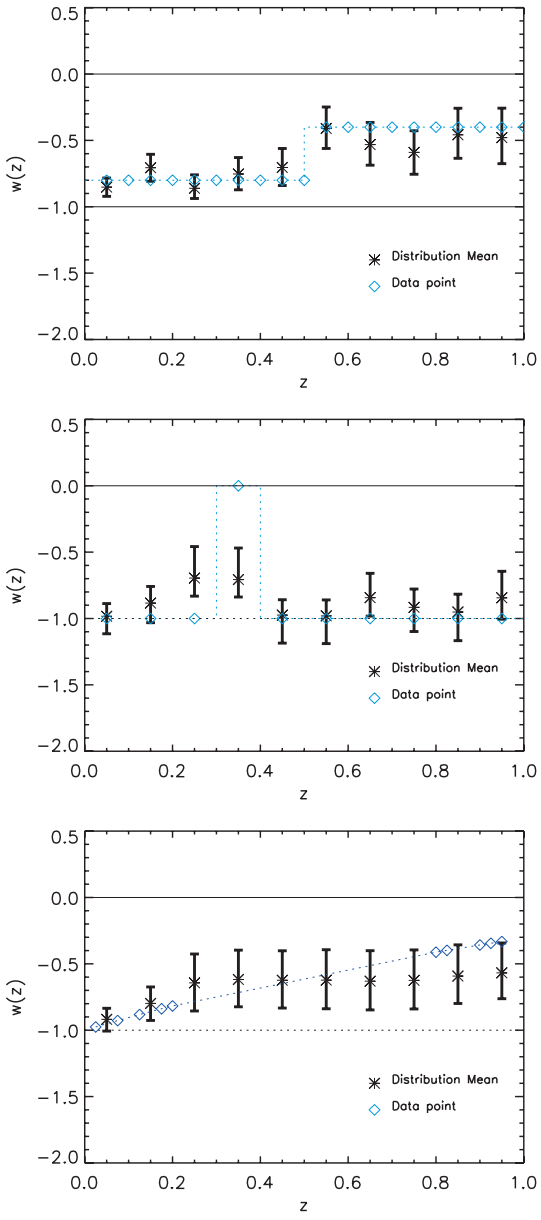

Figure 1. Reconstructed EOS w (black error bars showing 1σ posterior constraints) using our MaxEnt method for high-quality synthetic data ($N_D = 20, \sigma = 1$ per cent, location shown by the blue diamonds) in the redshift range $0 \leqslant z \leqslant 1$. Top panel: the true EOS (blue, dotted line) is taken to be a step function. Middle panel: the true EOS shows a sharp peak at $z \sim 0.4$. Lower panel: the true EOS is slowly evolving with redshift, and the synthetic data are now clustered at low and high redshift. Despite the absence of data points at intermediate redshifts, the high-z reconstruction tracks the true function with reasonable accuracy, while the intermediate-redshift errors increase correspondingly. In all three cases, the value of the prior model $\boldsymbol{m}$ for $w(z)$ has been marginalized over and the MaxEnt reconstruction satisfactorily recovers the true EOS. We have plotted horizontal lines at $w = 0$ and -1 to guide the eye.

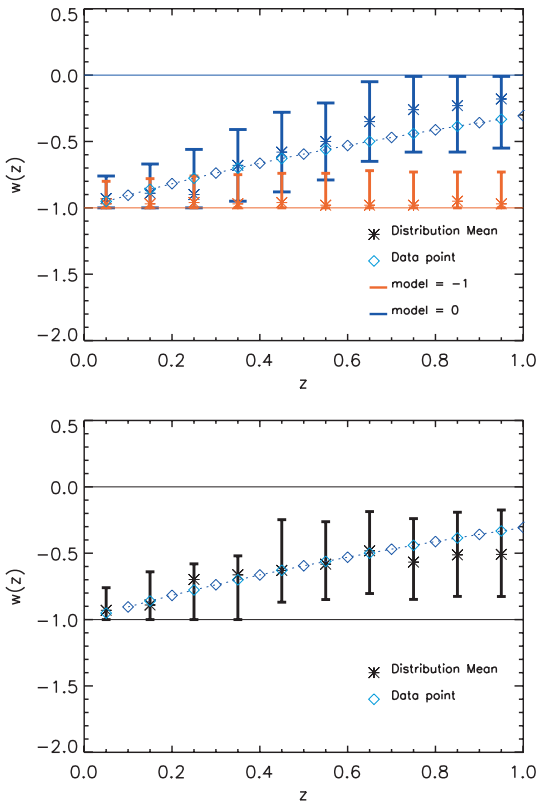

Figure 2. Reconstructed EOS w (error bars showing 1σ posterior constraints) for noisy synthetic data ($N_D = 20, \sigma = 10$ per cent, location shown by the blue diamonds). Top panel: the reconstructed EOS for a model $\boldsymbol{m} = -1$ (red error bars) has collapsed towards the model due to the entropic prior for redshifts $z \gtrsim 0.3$, while the reconstruction with $\boldsymbol{m} = 0$ (blue error bars) tracks better the true EOS (blue, dotted line). Bottom panel: after marginalization over the model value, the bias in the reconstruction has disappeared, but the error bars have become suitably larger.

reconstruction. The top panel of Fig. 2 shows the reconstructed EOS for a noisy data set of $N_D = 20$ measurements with noise $\sigma = 10$ per cent and a slowly evolving true EOS, $\boldsymbol{w}\,(z) = 1 - \ln(1 + z)$. In the top panel, we show the result when employing as prior model values the two liming cases of w with strong theoretical prejudice: $\boldsymbol{m} = -1$ (cosmological constant, red error bars) and $\boldsymbol{m} = 0$ (Einstein–de Sitter universe, blue error bars). In the high-z bins the reconstruction becomes increasingly mismatched with the underlying true EOS. Because the dependence of the data on the EOS requires integrating the EOS over redshift, any error in the reconstruction at low redshift is accumulated as z increases. As a result the entropy tends to dominate over the likelihood and the mean parameter values collapse towards the model at higher redshift, especially for the case where $\boldsymbol{m} = -1$. Even though the reconstruction has appeared to degrade for the $\boldsymbol{m} = 0$ case, it is encouraging that the mean values of the parameters in the lower redshift bins ($z \lesssim 0.3$) are reasonably close to the true values.

Evidently the choice of prior model $\boldsymbol{m}$ does have some bearing on the reconstructed value of the parameters at high redshift (given that the entropy dominates the posterior for poorly informative data) and this must be kept in mind when interpreting the results when dealing with noisy data. This example highlights the problem of distinguishing a genuine affinity for a certain function that happens to closely resemble the model from a strong default towards the model on account of noisy observations, that is, how will we interpret a result very close to $w(z) = -1$? In this case, there are two options: the observational evidence is noisy and uninformative, leading to entropy domination, or the data are good and favour an actual value close to -1. In the latter case, attempting a reconstruction with $\boldsymbol{m} = 0$ will allow to test the strength of the data in pulling $w(z)$ towards the cosmological constant value. An alternative means of

recovering $w(z)$ in a truly model-independent way is to include the elements of the model vector in the hypothesis space, as discussed in Section 2.2.5. Marginalizing over $\boldsymbol{m}$ then amounts to selecting the optimum distribution at each sampled point. The model still represents the most entropic distribution given that it is uniform in z-space. The result of this procedure is shown in the bottom panel of Fig. 2, where marginalization over the model has cured the skew in the reconstruction observed above, albeit at the price of delivering larger error bars.

4 APPLICATION TO PRESENT-DAY DATA

We now apply our reconstruction procedure to actual data, encompassing CMB observations, baryonic acoustic oscillation measurements, determination of the present value of the Hubble parameter and two different SNIa samples.

4.1 Data sets

4.1.1 Baryon acoustic oscillations

The comoving sizes along the line of sight, $r_{||}$ (in redshift space), and in the traverse directions, $r_{\perp}$, of a feature sitting at a redshift z are related to the redshift range Δz covered and the angle subtended $\Delta\Theta$, respectively, by

$$r_{||} = \frac{c\Delta z}{H(z)} \quad \text{and} \quad r_{\perp} = (1+z)D_{\rm A}(z)\Delta\Theta. \tag{23}$$

If the absolute values of $r_{\perp}$ and $r_{||}$ are known, they become standard rulers giving us a handle on $H(z)$ and $D_{\rm A}(z)$. If only the relative sizes are known, then the standard rulers are expressed in units of H_0. If only the ratio $r_{||}/r_{\perp}$ is known, this becomes the Alcock–Paczynski test. The BAO phenomenon can be used as such a standard ruler.

After recombination, when the Universe becomes neutral and photons free stream from the cosmic plasma, the driving force of the harmonic oscillation is removed and the sound speed of the now-neutral medium essentially falls off to zero, ending wave propagation. The spatial distribution of the baryons at this stage will then reflect the characteristic scale of the acoustic waves. Seeing as the perturbations in the baryon and cold dark matter distributions seed the formation of large-scale structure, we expect to see acoustic peaks in the late-time matter power spectrum (Eisenstein et al. 2005). This becomes a standard ruler because the scale of these acoustic oscillations is self-calibrated under standard recombination (Hu 2005). It depends solely on the photon–baryon ratio and radiation–matter ratio at recombination which are determined with excellent precision in the CMB power spectrum from the CMB peak morphology (Eisenstein & White 2004). The change in the apparent size of this scale from recombination to the present will depend on the expansion history of the universe through projection effects. These acoustic features appear as rings in angular and redshift space (Hu & Haiman 2003). The actual measurement from the SDSS LRG sample is of the dilation factor, defined as

$$D_{\rm V}(z) = \left[D_{\rm A}^2(z)\frac{cz}{H(z)}\right]^{1/3}, \tag{24}$$

where the comoving angular diameter distance $D_{\rm A}$ is taken as the transverse dilation and the line-of-sight measurement of this scale $\frac{cz}{H(z)}$ is taken to be the radial dilation. The observed correlation scale constrains a function of the dilation factor, and a single data point is reported in Eisenstein et al. (2005):

$$A \equiv D_{\rm V}(0.35)\frac{\sqrt{\Omega_{\rm m}H_0^2}}{0.35c} = 0.469 \pm 0.017. \tag{25}$$

(see also Cole et al. 2005, for a similar detection of the acoustic feature in the 2dF catalogue). The above assumes $\Omega_{\rm b}h^2 = 0.027$. The log-likelihood for the BAO data is given by

$$\chi^2_{\rm BAO} = \frac{(A-0.469)^2}{0.017^2}. \tag{26}$$

4.1.2 The Type Ia supernova data

SNeIa are good candidates for standard candles and are useful in determining distances on extragalactic scales. Due to the complexity of the physics involved, the SNeIa are not perfect standard candles, having a dispersion of 0.3–0.5 mag in their peak magnitudes (Straumann 2006). However, the peak brightnesses appear to be tightly correlated to the time-scale of their brightening and fading and one can extract an empirical relation between absolute peak luminosity and the morphology of the light curves to constrain the absolute brightnesses, and thus obtain measurements of $D_{\rm L}(z)$.

The first group of SNe, termed the 'gold' set, are from the *HST*/GOODS programme (Riess et al. 2004), complemented by the recently discovered higher redshift SNe, reported in Riess et al. (2007), while the second sample is taken from the SNLS (Astier et al. 2006). As discussed in, for example, Wang & Mukherjee (2006), it appears that there are systematic differences between these two data sets that arise from differences in the data processing. It is therefore necessary to consider the two data sets separately, and compare the results as a consistency check.

(i) *The 'gold' sample.* The distance modulus μ is defined as the difference between the apparent magnitude m and the absolute magnitude M:

$$\mu = m - M = 5\log_{10}\left[\frac{D_{\rm L}(z)}{10\ {\rm pc}}\right]. \tag{27}$$

Given that the absolute magnitude M is a unknown, we consider the following distance modulus $\mu_i(\delta M)$:

$$\mu_i(\delta M) = \mu_i^{\rm d} - \delta M. \tag{28}$$

Here δM is the difference between the mean of the true absolute magnitudes and the estimated absolute magnitude, while $\mu_i^{\rm d}$ are the observed magnitudes after dust corrections and recalibration through the shape of the luminosity evolution function. The quantity δM is the difference between the true mean absolute magnitude and the estimated absolute magnitude of the SNe and is marginalized over.

(ii) *The SNLS sample.* The SNIa data from the SNLS are reduced in a different manner in that the light curves provide constraints on various parameters which are then used to calculate the effective apparent magnitude. For a description of the calculation of μ_i, see Dick et al. (2006). From the observed μ_i with variances $\sigma^2_{\mu,i}$ for each set of SNeIa, we perform an analytical marginalization over the absolute magnitude M. Defining the quantities

$$f_0 = \sum_{i=1}^{N}\frac{5\log_{10}D_{\rm L} - \mu_i}{\sigma^2_{\mu,i}}, \tag{29}$$

$$c = \sum_{i=1}^{N}\frac{1}{\sigma^2_{\mu,i}} \tag{30}$$

and

$$f_1 = \sum_{i=1}^{N} \frac{(5\log_{10} D_L(z_i) - \mu_i)^2}{\sigma^2_{\mu,i}}, \quad (31)$$

the M-independent log-likelihood for the SNeIa is calculated as

$$\chi^2_{SN} = f_1 - \frac{f_0^2}{c}. \quad (32)$$

4.1.3 *CMB and HST data*

The *WMAP* three-year measurement of the CMB shift parameter describing the location of peaks in the CMB power spectrum serves to constrain the angular diameter distance to last scattering (Spergel et al. 2007). This is independent of most assumptions of the form of dark energy, and is given by Wang & Mukherjee (2006):

$$R = \Omega_m^{1/2} H_0 \int_0^{z_{CMB}} \frac{dz}{H(z)} = 1.70 \pm 0.03, \quad (33)$$

where z_{CMB} is the redshift to last scattering, taken in our case to be 1089.

We also include the constraint on the present value of the Hubble constant obtained by the *HST* Key Project (Freedman et al. 2001), by using a Gaussian likelihood with mean and standard deviations given by $H_0 = 72 \pm 8$ km Mpc s^{-1}.

4.2 Results

We plot in Fig. 3 the results of our reconstruction from the SNIa data from the SNLS and from the *HST*/GOODS programmes. In both cases we have added the CMB, *HST* and BAO measurements, and we have marginalized over the model, in order to be as conservative as possible. Furthermore, the maximum range for the EOS has been taken to be $-2 \leqslant \boldsymbol{w} \leqslant 0$. We plot regions encompassing 68 per cent of posterior probability for each $\boldsymbol{w}$ bin – note that since these are marginalized values, their magnitude is independent of the correlations between reconstructed points (the issue of correlations is addressed in detail below, see Fig. 4).

We do not find any significant deviation from a cosmological constant behaviour from the SNLS data (see top panel of Fig. 3) at all redshifts. Our posterior constraints in this case are rather conservative, as a consequence of the assumptions made in the reconstruction (i.e. large $\boldsymbol{w}$ range and marginalization over the model). When using the *HST*/GOODS data, the recovered $\boldsymbol{w}$ in the first bin agrees with that found for the SNLS data. Given that a number of the SNIa in this bin are common to both surveys, this provides a consistency check. The reconstructed EOS from the *HST*/GOODS data are, however, found to prefer slightly higher values in the third bin ($z \simeq 0.5$), excluding the cosmological constant value to a little bit more than 1σ significance. The significance of this rise, however, has to be assessed with care, especially if we recall that the constraining power of the SNIa data degrades in this region (Simpson & Bridle 2006). At redshifts above $z \sim 0.5$, the mean of the reconstruction settles around $w \sim -0.7$, although we note that the best-fitting points remain very close to $w = -1$ (red triangles in Fig. 3). The implied early-time behaviour of dark energy is consistent with the result of $w = -0.8^{+0.6}_{-1.0}$ found in Riess et al. (2007) for $z > 1$, using what they call their 'strong' prior. If instead the EOS is assumed to be constant over the entire redshift space (i.e. if we reduce the number of w components to $N = 1$), then we obtain from the *HST*/GOODS data $w = -0.89 \pm 0.07$, in agreement with usual results (see e.g. Riess et al. 2007). This clearly demonstrates the danger of assuming $w(z)$ to be time-independent, as one would miss in this way possible features in the data.

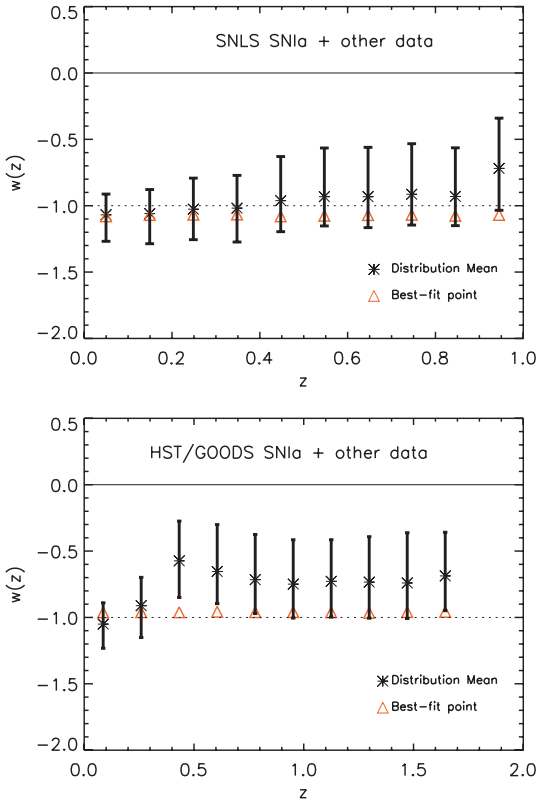

Figure 3. Reconstructed EOS (marginalized error bars encompassing 68 per cent of posterior probability) from SNLS data (top panel) and *HST*/GOODS data (bottom panel), including CMB, *HST* and BAO measurements, as well (note that the redshift range is different for the two panels). The prior model $\boldsymbol{m}$ is constant in redshift and has been marginalized and the assumed range of the EOS is $-2 \leqslant \boldsymbol{w} \leqslant 0$ (both are conservative choices). The horizontal lines indicate the upper bounds of the allowed $\boldsymbol{w}$ range (solid) and position of $w = -1$ (dotted) in order to guide the eye. The SNLS data do not show significant deviations from $w = -1$ over the whole redshift range. The reconstruction from the *HST*/GOODS data is also consistent with a cosmological constant, although it appears to slightly prefer a higher value at redshift $z \sim 0.5$. The best-fitting points are all very close to $w = -1$.

The use of the entropic prior introduces correlations among the reconstructed points (see Huterer & Cooray 2005, for a technique to extract uncorrelated band power estimates of the EOS). The correlation coefficients from the posterior distribution over the $\boldsymbol{w}$ parameters are shown for both data sets in Fig. 4. We note that correlations are in general relatively mild, flattening around the level of ∼20 per cent for correlations with bins at larger redshifts, where the entropic prior becomes more important. The strongest correlations (at the level of ∼50 per cent) are observed among parameters in the second and third redshift bins, where the BAO measurement strongly constrains the EOS and due to the fact that the observables are integrated over redshift, we expect a negative correlation among the well-constrained value at the position of the BAO measurement and the $\boldsymbol{w}$ values at lower redshift.

We now investigate the case where we impose that $w \geqslant -1$ on our parameter space. The results are shown in Fig. 5, where the reconstructed EOS using the SNLS tracks the cosmological constant value at all redshifts, with 1-tail 1σ errors of the order of 0.2 at all z values. Because of the reduced freedom in $\boldsymbol{w}$, the reconstruction collapses

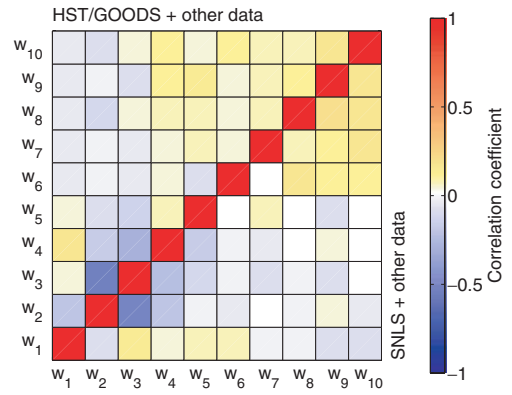

Figure 4. Correlation matrix for the reconstructed EOS parameters from *HST*/GOODS (upper left-hand panel) and SNLS (lower right-hand panel) data sets, including CMB, *HST* and BAO measurements. The strongest (anti)correlations are between values in bins 2 and 3 which roughly coincide with the redshift position of the BAO measurement.

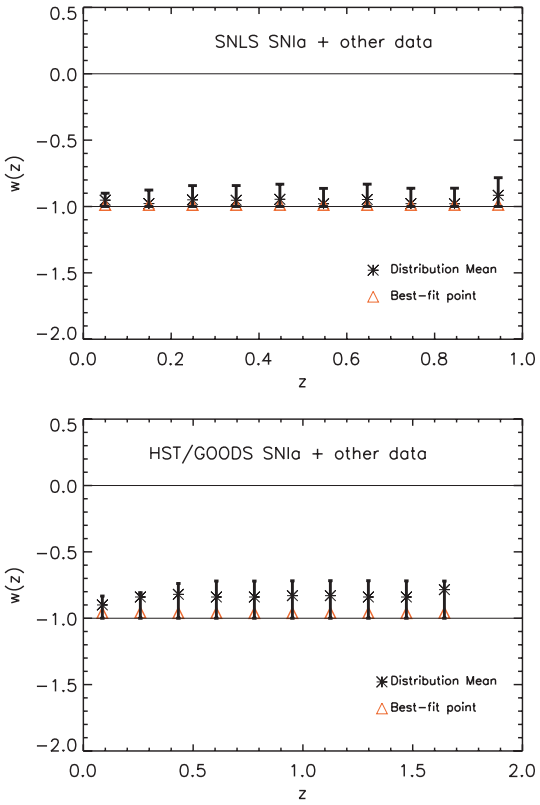

Figure 5. Reconstructed EOS (marginalized error bars encompassing 68 per cent of posterior probability) from SNLS data (top panel) and *HST*/GOODS data (bottom panel), marginalizing over a constant prior model but restricting the w range to $-1 \leqslant \boldsymbol{w} \leqslant 0$. The horizontal lines indicate the upper and lower bounds of the allowed $\boldsymbol{w}$ range (solid) in order to guide the eye. We find no significant deviation from $w = -1$ for the SNLS data set. For the *HST*/GOODS sample the reconstruction tends to drift to larger values at higher redshift.

to the lower limit of the allowed w range, even after marginalization over the model values. In the case of the *HST*/GOODS data, a gentle rise of $w(z)$ away from -1 is again observed. The larger error bars suggest that the entropy becomes important and that the

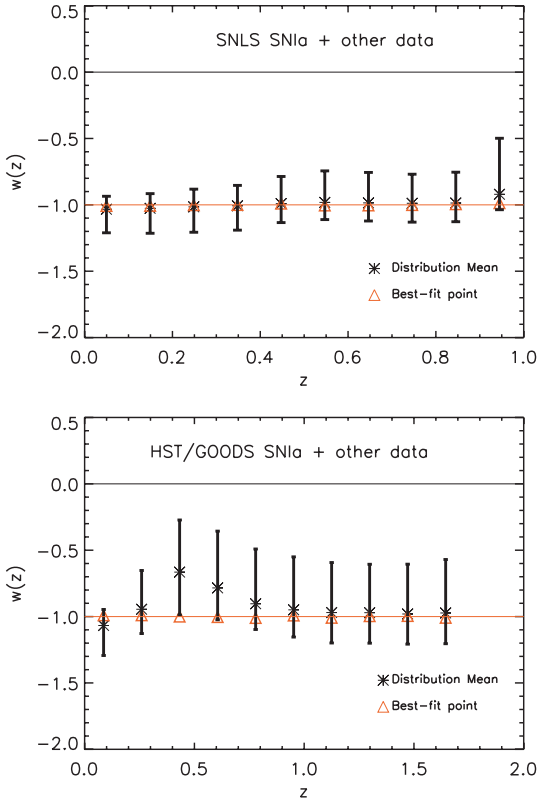

Figure 6. Reconstructed EOS from SNLS data (top panel) and *HST*/GOODS data (bottom panel), assuming an entropic prior model $\boldsymbol{m} = -1$ and a $\boldsymbol{w}$ range $-2 \leqslant \boldsymbol{w} \leqslant 0$. The horizontal lines indicate the upper bound of the allowed $\boldsymbol{w}$ range (black lines) and the model $\boldsymbol{m}$ (red lines). The SNLS data are compatible with the model and show error bars of the order of 20 per cent at all redshifts. The slight bump at $z \sim 0.5$ for the *HST*/GOODS data survives the entropic prior.

value $\boldsymbol{w} \sim -0.8$ to which the reconstruction tends at redshifts $z \gtrsim 0.5$ is mediated by the mean value of the prior model $\boldsymbol{m}$. As before, the best-fitting points remain very close to $w = -1$ at all redshifts.

We can increase the amount of prior information by considering the case where a constant prior model value $\boldsymbol{m} = -1$ is used, see Fig. 6. This is helpful in assessing whether the structure observed in the *HST*/GOODS sample is strong enough to override our entropic prior. The reconstruction from the SNLS data remains close to $\boldsymbol{w} = -1$ with error bars of the order of 20 per cent at all redshifts. One has, however, to keep in mind that the tightness of the errors is partially helped by the supplementary information provided by the entropic prior. This demonstrates how the use of cosmological constant as the model can be problematic as one cannot say conclusively whether this indicates that the data are very strong or alternatively overridden by the entropy if no significant deviations from the model are observed. The result from the *HST*/GOODS data shows again a high value of $w(z)$ being favoured in the third bin and subsequent collapse towards to the model $\boldsymbol{m}$ at higher redshifts. The persistence of a deviation towards $\boldsymbol{w} > -1$ at redshift $z \sim 0.5$ in the presence of the strong prior favouring the cosmological constant suggests it is a real feature of the data.

Finally, we also investigated the stability of the reconstruction against a change in the number of reconstructed components. Since the MaxEnt technique is designed to automatically deal with the structure in the data by adjusting the degree of smoothness of the

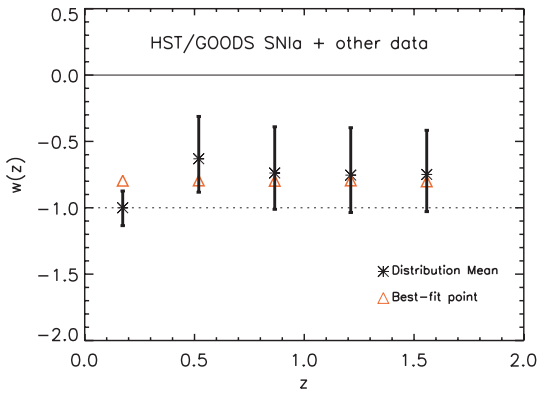

Figure 7. Reconstructed EOS from the *HST*/GOODS data with five components in the reconstructed $\boldsymbol{w}$. Compare with the bottom panel of Fig. 3.

reconstruction, we do not expect that a change in the number of bins would make a large difference in the reconstructed EOS. This is demonstrated in Fig. 7, where the reconstruction analogous to the bottom panel of Fig. 3 has been performed by halving the number of bins to $N = 5$, without appreciable differences in the end result.

Lastly, we investigate the dependency of our results on the model one chooses for the entropic prior. Up until this point we have assumed the prior model $\boldsymbol{m}$ to be constant with redshift such that it introduces a suitable degree of smoothing of time-dependent noisy features in the data. This choice reflects a specific belief in the true form of the EOS and given the large range of dark energy models on the market, it is important to assess the impact of a different prior model $\boldsymbol{m}$. Another popular class of models are given by an EOS that is a smooth varying function of redshift, such as

$$m(z) = w_0 + w_1 \frac{z}{(1+z)}. \tag{34}$$

Here the assumption is that the true EOS is a function of time and evolves sufficiently slowly such that it may be effectively characterized in a phenomenological way by the two parameters w_0 and w_1 (Sahni & Starobinsky 2006). This particular function is a good approximation to many dark energy models but clearly it is limited to how well it can cope with a rapidly evolving EOS (Liddle et al. 2006). Following, for example, Ichikawa & Takahashi (2006), we impose the further constraint that the early Universe is matter-dominated, that is, we impose the condition $w_0 + w_1 < 0$ on the prior model. In order to be as conservative as possible, we again marginalize over both prior model parameters, w_0 and w_1, as follows. The expression for the entropy is again given by equation (22), but the model coefficients are now given by

$$C_i = \left(C_0^{\rm m} - 2\right) + \left(C_1^{\rm m} - 2\right) \frac{z_i}{(1+z_i)} + 2. \tag{35}$$

Here $C_0^{\rm m}$ and $C_1^{\rm m}$ are the coefficient representing the prior model parameters. They are included in the hypothesis space, and then marginalized over.

We display our results using this slowly-evolving prior model in Fig. 8 (with the model parameters marginalized over). This is to be contrasted with the analogous case of Fig. 3, where the entropic model $\boldsymbol{m}$ is constant in redshift space. At low redshifts the reconstructions for different prior models do not differ appreciably, with the most-notable difference being a shift towards slightly higher values of w for the case of the SNLS data set. For the GOODS/*HST* programme data set, the peak in the third bin is recovered but is

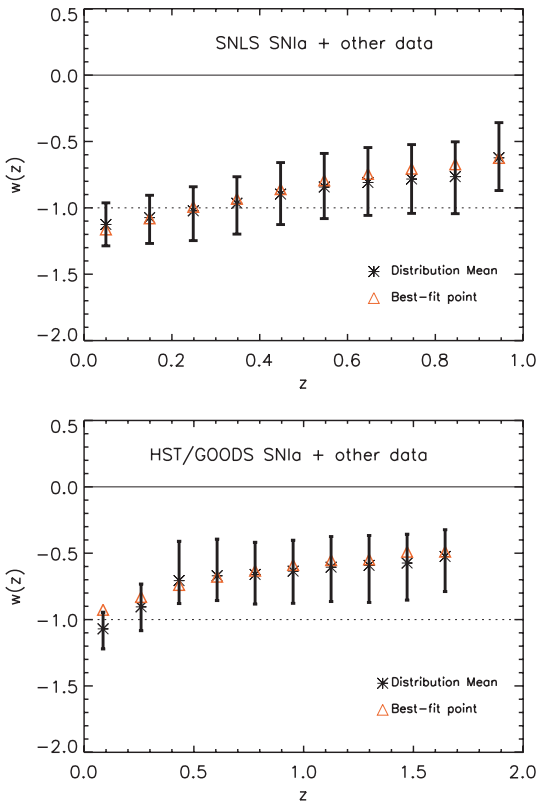

Figure 8. Reconstructed EOS (marginalized error bars encompassing 68 per cent of posterior probability) from the SNLS data (top panel) and the *HST*/GOODS data (bottom panel) marginalizing over a prior model of the type (w_0, w_1). The horizontal lines indicate the upper bound of the allowed $\boldsymbol{w}$ range (solid lines) and position of $w = -1$ (dotted lines) in order to guide the eye. A comparison with Fig. 3, where a constant prior model was used, indicates that the results from both data sets are robust against the choice of the prior model while exhibiting a mild dependence on this choice at higher redshift, where the constraining power of the data degrades and entropic domination sets in.

lower than in the previous constant $\boldsymbol{m}$ case. Again, the reconstructed EOS prefers higher values in the last few bins. At high redshifts, due to the accumulation of error, we expect more entropic domination and hence a stronger weight of the prior model choice. However, in this region dark energy becomes progressively less important and the ability of data to constrain the time-depedence of the EOS thus degrades considerably. It is encouraging, however, that the low redshift behaviour of the reconstructed EOS agrees well for both marginalized prior models. This further strengthens our conclusions regarding the power and robustness of our technique.

5 CONCLUSIONS

Given that the dark energy models on the market are predominantly phenomenological, a reconstruction technique that does not require specifying a parametric form for $w(z)$ would be an advantage. We presented such a method based on maximum entropy to reconstruct the EOS of dark energy within a Bayesian framework. The principle of maximum entropy is invoked when assigning the Bayesian prior. This means that in the absence of genuine signal, the model

$w(z)$ that is most smooth, or that maximizes the information entropy is favoured, with the extent of this bias being determined by a regularizing constant that is automatically adjusted to the data. In our analysis we decompose $w(z)$ into a sum of weighted orthogonal step-functions to facilitate the reconstruction of sharp features (provided binning is sufficient). Extensions of our analysis could easily include using alterative expansion functions to see whether other properties of the time-evolution of dark energy are detected or constrained.

We find that the reconstruction of a dynamical $w(z)$ using artificial data sets of $H(z)$ and $D_A(z)$ is very promising at low redshifts but suffers from a bias towards the chosen default model at higher z. To combat this effect, the prior model was incorporated into our hypothesis space allowing the reconstruction of a model-independent distribution of w, with a manageable loss of accuracy. Once the technique was established and demonstrated, it was applied to a combinations of the current cosmological data sets and two popular choices of prior models, namely a constant EOS and a mildly evolving $w(z)$.

Using a data set including the current *WMAP*3 measurement of the CMB shift parameter, the BAO measurement and the *HST* Key project measurement of the Hubble parameter in conjunction with the SNIa data from the SNLS project, we found that $w = -1$ in the redshift range $0 \leqslant z \leqslant 1100$, with error bars depending on the prior and model assumptions. In the most-optimistic case, where the data are supplemented by an entropic prior around $w = -1$, the error is of the order of 20 per cent at all redshifts. When the same data set was instead supplemented by the SN sample from the *HST*/GOODS program, the results agree at low ($z \lesssim 0.3$) redshift. We found, however, that the reconstruction tends to prefer a value $\boldsymbol{w} > -1$ around $z \sim 0.5$ with a significance between 1 and 2σ, depending on assumptions. This shows the dangers of fitting a parametric form of $w(z)$ to the data, in which case one is bound to miss possibly significant features in the measurements. The high-redshift behaviour of the EOS becomes increasingly dominated by the entropic prior and thus exhibits a mild dependence on the choice of prior model. We have investigated the correlation properties of our reconstruction, and identified a moderate anti-correlation among the first few redshift bins of our reconstructed points.

The MaxEnt technique presented here improves on other methods designed to minimize noise artefacts in that the amount of information taken from the data is not determined by the analyst but rather dictated by the data themselves. The presence of real structure rather than noise-induced complexity is indicated by the size of the error bars. The entropic prior adjusts the error bars when the information provided by the data is unreliable. Secondly, in the absence of real information the reconstruction tends towards our most-intuitive estimate of $w(z)$ with suitably large variance. In conclusion the merits of this technique are that it is self-regulating in the sense that it allows the data to determine the amount of structure that is included. More importantly, it does not require an inherent assumption of the functional form of the true EOS.

We hope that this technique will prove useful in deriving even stronger, model-independent constraints on the dark energy history from future, high-quality data.

ACKNOWLEDGMENTS

The authors wish to thank Bruce Bassett, Pedro Ferreira, Scott Dodelson, Josh Frieman, Dragan Huterer, Joe Silk and Glenn Starkman for stimulating discussions and useful suggestions. We are grateful to Sarah Bridle for suggesting to include the model in the parameter space and to an anonymous referee for many valuable comments. CZ is supported by a Domus A scholarship awarded by Merton College. RT is supported by the Royal Astronomical Society through the Sir Norman Lockyer Fellowship and by St Anne's College, Oxford. The use of the Glamdring cluster of the Oxford University is acknowledged.

REFERENCES

Abbott T. et al., 2005, preprint (astro-ph/0510346)
Alcaniz J. S., 2004, Phys. Rev. D, 69, 083521
Aldering G. et al., 2004, preprint (astro-ph/0405232)
Astier P. et al., 2006, A&A, 447, 31
Bassett B. A., Corasaniti P. S., Kunz M., 2004, ApJ, 617, L1
Bennett C. L. et al., 2003, ApJS, 148, 1
Blake C. A., Abdalla F. B., Bridle S. L., Rawlings S., 2004, New Astron. Rev., 48, 1063
Bridle S. L., Hobson M. P., Lasenby A. N., Saunders R., 1998, MNRAS, 299, 895
Caldwell R. R., 2002, Phys. Lett., B545, 23
Carroll S. M., Hoffman M., Trodden M., 2003, Phys. Rev. D, 68, 023509
Chevallier M., Polarski D., 2001, Int. J. Mod. Phys. D, 10, 213
Cole S. et al., 2005, MNRAS, 362, 505
Dick J., Knox L., Chu M., 2006, JCAP, 0607, 001
Efstathiou G., 1999, MNRAS, 310, 842
Eisenstein D. J. et al., 2005, ApJ, 633, 560
Eisenstein D. J., White M. J., 2004, Phys. Rev. D, 70, 103523
Freedman W. L. et al., 2001, ApJ, 553, 47
Frieman J. A., Hill C. T., Stebbins A., Waga I., 1995, Phys. Rev. Lett., 75, 2077
Glazebrook K., Eisenstein D., Dey A., Nichol B., 2005, preprint (astro-ph/0507457)
Hoekstra H., Yee H. K. C., Gladders M. D., 2002, ApJ, 577, 595
Hoekstra H. et al., 2006, ApJ, 647, 116
Hu W., 2002, Phys. Rev. D, 66, 083515
Hu W., 2005, in Wolff S. C., Lauer T. R., eds, ASP Conf. Ser. Vol. 339, Observing Dark Energy. Astron. Soc. Pac., San Francisco, p. 215
Hu W., Haiman Z., 2003, Phys. Rev. D, 68, 063004
Huterer D., Cooray A., 2005, Phys. Rev. D, 71, 023506
Huterer D., Starkman G., 2003, Phys. Rev. Lett., 90, 031301
Huterer D., Turner M. S., 2001, preprint (astro-ph/0103175)
Ichikawa K., Takahashi T., 2006, Phys. Rev. D, 73, 083526
Jarvis M., Jain B., Bernstein G., Dolney D., 2006, ApJ, 644, 71
Kelz A. et al., 2006, Proc. SPIE, 6273
Kunz M., Trotta R., Parkinson D., 2006, Phys. Rev. D, 74, 023503
Lange A. E. et al., 2001, Phys. Rev. D, 63, 042001
Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511
Liddle A. R., Mukherjee P., Parkinson D., Wang Y., 2006, Phys. Rev. D, 74, 123506
Linder E. V., 2003, Physical Rev. Lett., 90, 091301
Linder E. V., 2004, Phys. Rev. D, 70, 061302
Maisinger K., Hobson M. P., Lasenby A. N., 2004, MNRAS, 347, 339
Marshall P. J., Hobson M. P., Gull S. F., Bridle S. L., 2002, MNRAS, 335, 1037
Neal R. M., 1992, preprint (hep-lat/9208011)
Nugent P. et al., 2006, ApJ, 645, 841
Peebles P. J. E., Ratra B., 2003, Rev. Mod. Phys., 75, 559
Perlmutter S. et al., 1999, ApJ, 517, 565
Riess A. G. et al., 1998, ApJ, 116, 1009
Riess A. G. et al., 2004, ApJ, 607, 665
Riess A. G. et al., 2007, ApJ, 659, 89
Sahni V., Starobinsky A. A., 2006, Int. J. Mod. Phys. D, 15, 2105
Seo H.-J., Eisenstein D. J., 2003, ApJ, 598, 720
Shafieloo A., Alam U., Sahni V., Starobinsky A. A., 2006, MNRAS, 366, 1081

Silver R. N., Sivia D. S., Gubernatis J. E., 1990, Phys. Rev. B, 41, 2380
Simpson F., Bridle S., 2006, Phys. Rev. D, 73, 083001
Skilling J., 1989, Maximum Entropy and Bayesian Methods, 1st edn. Kluwer Academic Publishers, Cambridge
Spergel D. N. et al., 2007, ApJS, 170, 377
Starkman G. D., Trotta R., 2006, Phys. Rev. Lett., 97, 201301
Straumann N., 2006, Mod. Phys. Lett., A21, 1083
Trotta R., 2007, MNRAS, 378, 72
Trotta R., Bower R., 2006, Astron. Geophys., 47, 4
Tyson J. A., 2006, in Liss T. M., ed., AIP Conf. Proc. Vol. 870, Intersections of Particle and Nuclear Physics. Am. Inst. Phys., New York, p. 44
Wang Y., Mukherjee P., 2006, ApJ, 650, 1
Weller J., Albrecht A., 2002, Phys. Rev. D, 65, 103512

This paper has been typeset from a TEX/LATEX file prepared by the author.