Dark energy, a cosmological constant, and type Ia supernovae

Lawrence M Krauss¹, Katherine Jones-Smith¹
and Dragan Huterer²

¹ CERCA, Case Western Reserve University, Cleveland, OH 44106-7079, USA
² Kavli Institute for Cosmological Physics and Department of Astronomy and
Astrophysics, University of Chicago, Chicago IL 60637, USA
E-mail: krauss@cwru.edu

Received 29 January 2007
Published 21 May 2007
Online at http://www.njp.org/
doi:10.1088/1367-2630/9/5/141

Abstract. We focus on uncertainties in supernova measurements, in particular
of individual magnitudes and redshifts, to review to what extent supernovae
measurements of the expansion history of the universe are likely to allow us
to constrain a possibly redshift-dependent equation of state of dark energy, \( w(z) \).
We focus in particular on the central question of how well one might rule out
the possibility of a cosmological constant \( w = -1 \). We argue that it is unlikely
that we will be able to significantly reduce the uncertainty in the determination of
\( w \) beyond its present bounds, without significant improvements in our ability
to measure the cosmic distance scale as a function of redshift. Thus, unless
the dark energy significantly deviates from \( w(z) = -1 \) at some redshift, very
stringent control of the statistical and systematic errors will be necessary to
have a realistic hope of empirically distinguishing exotic dark energy from a
cosmological constant.
Contents

1. Introduction 2
2. Mapping the expansion history 3
   2.1. Key questions and parametrizations 3
   2.2. Direct reconstruction 4
   2.3. Principal components (PCs) 5
   2.4. Uncorrelated estimates of DE evolution 7
3. Lambda or not? 8
4. Conclusions 11
References 12

1. Introduction

Eight years ago two teams observing distant type Ia supernovae (SNe Ia) announced evidence that the expansion of the universe is speeding up [1, 2]. The distant supernovae appear dimmer than they would be in a matter-only universe. If this is a true distance effect, it implies that about 70% of the energy density of the universe resides in a smooth component with negative pressure, leaving only about 25% in dark matter and 5% in baryonic matter. While compelling evidence for precisely this combination was pointed out several years earlier [3], based on measurements of the clustering of galaxies, age of the universe, measurements of the baryon and dark matter densities, and the Hubble constant, SNe Ia were the ‘shot heard around the world’ as they provided explicit evidence for the largest contributor to the energy density of the universe. This constituent became known as dark energy.

Since that time the constraints have improved significantly: a combination of cosmic microwave background experiments [4], large-scale structure surveys [5, 6] and, particularly, new SNe Ia observations [7]–[11], now constrain the dark energy equation of state, $w \equiv p/\rho$ to be within about $-1 \pm 0.1$.

Unfortunately, our understanding of dark energy is as murky as ever. The simplest model for dark energy is provided by the cosmological constant term in Einstein’s equations, and is still an excellent fit to observations as it predicts $w = 1$ identically and at all times. However, even if observations were to pin down the value of $w$ to be $-1$, this gives us essentially no insight into the possible source of dark energy. While a cosmological constant may be the best bet, it is simple to imagine other sources of dark energy, including the energy density associated with a false vacuum metastable scalar field, that would produce a similar value. The only way to get any new theoretical handle on dark energy is to be able to unambiguously determine a deviation from $-1$, if such a deviation indeed exists.

Various proposals for dynamical dark energy have been put forth which might produce such deviations. Most notably, a scalar field rolling down its effective potential can provide the necessary energy density and acceleration of the universe [12]–[17]. Generically many of these possibilities are already ruled out by the data. Among those that are not, none is particularly compelling as they also typically do not naturally address the problem why the observed dark energy density ($\sim 10^{-120} M_{\text{Pl}}^4$) is so small, nor why dark energy starts to dominate the expansion of the universe only at recent times—i.e. a redshift $z \lesssim 1$. Nevertheless, before we can address
such puzzles we need to know empirically if the dark energy is measurably distinguishable from a cosmological constant.

Here we will be concerned with two specific observational factors, as well as one overriding theoretical constraint. Observationally we need to determine both the magnitude and redshift of individual objects in order to map the universe’s expansion history. Theoretically we have to account for the fact that dark energy has, a priori, no predetermined time dependence, and thus our analyses must allow for arbitrary time variations.

Essentially all of the consequences of dark energy follow from its effect on the expansion rate:

$$H^2 = \frac{8\pi G}{3} (\rho_M + \rho_{DE}),$$

(1)

$$H^2(z)/H_0^2 = \Omega_M (1 + z)^3 + \Omega_{DE} \exp \left[ 3 \int_0^z [1 + w(z')]d\ln(1 + z') \right],$$

where $\Omega_M$ and $\Omega_{DE}$ are the dark matter and dark energy density relative to critical, respectively, and we have ignored the relativistic components. SNe Ia effectively measure the luminosity distance

$$d_L(z) = (1 + z)r(z) = \int_0^z \frac{dz'}{H(z')},$$

(2)

where $r(z)$ is the comoving distance and we have assumed a flat universe. Since observations of supernovae allow us, at least in principle, to map the expansion history of the universe, one can use this data to constrain the nature of dark energy.

This paper is organized as follows. In section 2, we review a variety of parametrizations of the expansion history of the universe, and therefore ways to measure the properties of dark energy. In section 3, we study the extent to which an increased statistical error in supernova distances affects determination of the equation of state of dark energy and tests of its consistency with the vacuum energy value of $-1$. We conclude in section 4.

2. Mapping the expansion history

2.1. Key questions and parametrizations

At the present time, it has become clear that there are two major goals that upcoming dark energy probes should address.

1. Is dark energy consistent with the vacuum energy scenario—that is, is $w(z) = -1$?
2. Is the equation of state $w$ constant in time (or redshift)?

Violation of either of these two hypotheses would be a truly momentous discovery: the former would rule out a pure cosmological constant, while the latter would provide further understanding of the nature of dark energy via its dynamics.
With this in mind, the most obvious approach is to parametrize the equation of state of dark energy as a constant piece plus a redshift-varying one [18]–[20]

\[ w(z) = w_0 + w'z, \]

\[ w(z) = w_0 + \frac{w_\alpha z}{1 + z}, \]

\[ w(z) = w_0 + \frac{w_f - w_0}{1 + \exp[(z - z_i)/\Delta]}, \]

where the first equation assumes linear evolution with redshift, the second is linear with scale factor, and the third allows for the transition between two asymptotic constant values of the equation of state, with the transition occurring at redshift \( z_i \) having width in redshift \( \Delta \).

Equation (4), in particular, has become commonly used to plan probes of dark energy as it retains the minimally required two parameters and does not diverge at high \( z \), while still allowing for the low-redshift dynamics (e.g. variation in the value of \( w(z) \)).

Many other simple parametrizations of the equation of state have been suggested (e.g. [21, 22]); similar proposals have been extended to the Hubble parameter (e.g. [23]). Finally, specific combinations of the expansion history parameters, such as the ‘statefinder’ [24], have been proposed as good discriminators between phenomenological dark energy descriptions. It is well worth emphasizing that all of the above parametrizations are ad hoc, and may lead to biases as the true DE model may not follow the form imposed by these functions. Their advantage, however, is in simplicity and the fact that two or three additional parameters in the dark energy sector will be, in the near future, measured to a good accuracy, at least when the data from the various cosmological probes is combined.

### 2.2. Direct reconstruction

Going in the opposite direction, the most general way to probe the background evolution of dark energy has been proposed in [25]–[27]: the equation of the distance versus redshift can be inverted to obtain \( w(z) \) as a function of the first and second derivatives of the (SNe Ia-inferred, for example) comoving distance

\[ 1 + w(z) = \frac{1 + z}{3} \frac{3H_0^2\Omega_M(1 + z)^2 + 2(d^2r/dz^2)/(dr/dz)^3}{H_0^2\Omega_M(1 + z)^3 - (dr/dz)^2}. \]

Similarly, assuming that dark energy is due to a single rolling scalar field, one can reconstruct the potential of the field exactly

\[ V[\phi(z)] = \frac{1}{8\pi G} \left[ \frac{3}{(dr/dz)^2} + (1 + z) \frac{d^2r/dz^2}{(dr/dz)^3} \right] - \frac{3\Omega_M H_0^2(1 + z)^3}{16\pi G}, \]

where the upper (lower) sign applies if \( \phi > 0 \) (< 0). The sign is arbitrary, as it can be changed by the field redefinition, \( \phi \leftrightarrow -\phi \).

Direct reconstruction is the most general method for inferring the dark energy history, and it is the only approach that is truly model-independent (despite some claims in the literature to the contrary). However, direct reconstruction comes at a steep price—it calls for taking the
Figure 1. Example of the direct reconstruction, simulated for future SNe Ia data and assuming the true equation of state is \( w(z) = -1 \). The left, middle and right panels show results when the luminosity distance data are fit with a third-, fourth- and fifth-order polynomial in redshift, respectively. Note that, depending on which order polynomial is used to fit the data, significant bias, or statistical error, or both are introduced. Adopted from [28].

second derivative of noisy data. In order to take the second derivative, one essentially must fit the luminosity distance (or SNe Ia apparent magnitude) data with a smooth function—a polynomial, Padé approximant, spline with tension etc. Unfortunately, the parametric nature of the fitting process introduces systematic biases. After valiant attempts to do this using a variety of methods for smoothing or fitting the data (e.g. [23], [28]–[30]), various authors found that direct reconstruction is simply too challenging and not robust even with SNe Ia data of excellent quality. Figure 1 shows an example of the direct reconstruction, simulated for future SNe Ia data and assuming the true equation of state is \( w(z) = -1 \) (adopted from [28]). Note that, depending on which order polynomial is used to fit the data, significant bias, or statistical error, or both are introduced. For an excellent review of the dark energy reconstruction and related issues, see [31].

2.3. Principal components (PCs)

The next most general method, introduced in [32] (see also [33, 34]), is to compute the PCs of the quantity that we are measuring—the equation of state or energy density of dark energy. PCs are the redshift weights (or window functions) of the function in question, and are uncorrelated by construction. In this scheme, one simply lets the data decide which weights of the function are measured best, and which ones are measured most poorly. The PCs form a natural basis that parametrizes the measurements of any particular survey.

To compute the PCs, let us parametrize \( w(z) \) (same arguments follow for \( \rho(z) \) or \( H(z) \)) in terms of piecewise constant values \( w_i \) (\( i = 1, \ldots, N \)), each defined in the redshift range \( z_i = [(i-1)\Delta z, i\Delta z] \) where \( \Delta z = z_{\text{max}}/N \). In the limit of large \( N \) this recovers the shape of an arbitrary dark energy history (in practice, \( N \gtrsim 20 \) is sufficient). We then proceed to compute the covariance matrix for the parameters \( w_i \), plus any other cosmological parameters such as \( \Omega_M \), then marginalize over the latter. We then have the covariance matrix, \( C \), for the \( w_i \).
It is then a simple matter to find a basis in which the parameters are uncorrelated; this is achieved by simply diagonalizing the inverse covariance matrix (which is in practice computed directly and here approximated with $F$). Therefore

$$F = C^{-1} = W^T \Lambda W,$$

(8)

where the matrix $\Lambda$ is diagonal and the rows of the decorrelation matrix $W$ are the eigenvectors $e_i(z)$, which define a basis in which our parameters are uncorrelated [35]. The original function can be expressed as

$$w(z) = \sum_{i=1}^{N} \alpha_i e_i(z),$$

(9)

where $e_i$ are the ‘PCs’ (see figure 2). Using the orthonormality condition, the coefficients $\alpha_i$ can be computed as

$$\alpha_i = \sum_{a=1}^{N} w(z_a) e_i(z_a).$$

(10)

Diagonal elements of the matrix $\Lambda$, $\lambda_i$, are the eigenvalues which determine how well the parameters (in the new basis) can be measured; $\sigma(\alpha_i) = \lambda_i^{-1/2}$. We have ordered the $\alpha$s so that $\sigma(\alpha_1) \leq \sigma(\alpha_2) \leq \cdots \leq \sigma(\alpha_N)$.

PCs have one distinct advantage over fixed parametrizations. They allow the data to determine which weight of the cosmological function $w(z)$ (or $\rho_{DE}(z)$, or $H(z)$) is best determined. In fact, the PCs depend on the cosmological probe, on the specifications of the probe (redshift and sky coverage etc), and (more weakly) on the true cosmological model. These dependencies are features and not bugs, and they make the PCs a useful tool in survey design. For example, one can design a survey that is most sensitive to the dark energy equation of state at some specific redshift, or study how many independent parameters are measured by any given combination of cosmological probes (e.g. [22]).

Figure 2. The four best-determined and two worst-determined PCs of $w(z)$ for a future SNe Ia survey such as SNAP [36]. Adopted from [32].
2.4. Uncorrelated estimates of DE evolution

A useful extension of the PC formalism is to compute the band powers in redshift of the dark energy function, \( w(z) \) (or \( \rho_{\text{DE}}(z) \) or \( H(z) \)). These band powers can be made 100\% uncorrelated by construction, and the price to pay is a small leakage in the sensitivity of each band power outside of its redshift range. For details on how this is implemented, see [37].

The left panel of figure 3 shows the constraints on the four band powers of \( w(z) \), adopted from [37], assuming the Riess04 SNe Ia data [7] and a prior on \( \Omega_M \). The right panel of the figure shows the same method to constrain three band powers, using the newest SNe Ia data from the Hubble Space Telescope, adopted from [10], in combination with the baryon acoustic oscillation measurements [6]. The slight preference for \( w(z) \) increasing with redshift is seen in both panels; however, significant systematics still affect the data, in particular because of the heterogeneity of the SNe Ia datasets used. Therefore, it is too early to claim any evidence for the departures from \( \Lambda \)CDM (the standard flat cosmology in which the cold dark matter model is supplemented by dark energy), and it remains to be seen how the results change once we have a systematically more homogeneous and statistically more powerful data set (for the requirements on the SNe Ia systematics, see [38]).

It interesting to see in figure 3 how good the constraints are, given that this is current data and that interesting constraints are obtained on 3–4 band powers. A superior data set with an excellent control of the systematics, such as that expected from a dedicated space telescope such as SuperNova/Acceleration Probe (SNAP; [36]) will significantly improve the constraints, and also allow a finer resolution (i.e. more band power parameters) in redshift.

Nevertheless, it is clear from the existing set that either significant redshift evolution in \( w \) or significant improvement in the data will be required before either redshift evolution or deviation from \(-1\) could be unambiguously inferred using this technique. We now turn to considering this question in more detail.
3. Lambda or not?

As stressed earlier, one of the most important outstanding questions in cosmology is whether dark energy is consistent with a cosmological constant. This hypothesis can be tested in a variety of ways. The simplest approach is to compute a simple likelihood comparison between the data and the vacuum energy model \( w(z) = -1 \). A much more sophisticated (but admittedly less robust) approach would be to perform some type of reconstruction of the energy density \( \rho_{\text{DE}}(z) \) and check whether or not it is consistent with a constant value (or similarly, to reconstruct the equation of state \( w(z) \) and check whether it is consistent with \(-1\)).

Here we consider the \( \Lambda \) hypothesis test in terms of PCs. We consider models in the \( w_0-w_a \) plane (see equation (4)) and as we outline below, we use the best measured PCs to determine if we can statistically distinguish the model from \( \Lambda \text{CDM} \). In this section, we assume future SNe Ia data with 3000 SNe distributed uniformly in \( 0.1 < z < 1.7 \); this corresponds roughly (but not exactly) to what is expected from the SNAP space telescope [36].

For a fixed dark energy model described by some values of the cosmological parameters \( w_0 \) and \( w_a \), let the PCs take values \( \alpha_i \), with associated errors \( \sigma(\alpha_i) \) (note, we marginalize the results over the values of \( \Omega_M \), thus enlarging the errors in the PCs; in this way we can talk about models in the \( w_0-w_a \) plane without further recourse to \( \Omega_M \)). Further, from equation (10) it follows that the PC coefficients for the \( w(z) = -1 \) model are

\[
\bar{\alpha}_i = (-1) \sum_{a=1}^{N} e_i(z_a),
\]

where \( e_i(z) \) is the shape of the \( i \)th PC in redshift. We can now perform a simple \( \chi^2 \) test to determine whether \( w \) is constant:

\[
\chi^2 = \sum_{i=1}^{M} \frac{(\alpha_i - \bar{\alpha}_i)^2}{\sigma^2(\alpha_i)},
\]

where we have chosen to keep only the first \( M \) PCs, since the best-determined modes will contribute the most to the sum (note that the test is valid regardless of the value of \( M \); in practice, our \( M \) is typically \( 3-5 \)). Given that we have \( M \) degrees of freedom, each model \( (w_0, w_a) \) will lead to \( \chi^2 \) which may or may not be inconsistent with the \( \Lambda \text{CDM} \) model at some fixed confidence level.

In order to determine the ability of future SNe surveys to make this distinction, we focus here on the sensitivity of the results to the measured magnitude uncertainty. We focus on this factor for two reasons. Firstly, we believe it will be the single most important determinant of the ability of future surveys to possibly rule out a cosmological constant, and secondly, because our examination of past surveys suggests that one should consider the possibility of redshift–dependent magnitude measurement errors. This latter possibility is not unexpected. It is systematically harder to determine the luminosity of ever fainter and more distant objects.

Indeed, it is this possibility that originally motivated [39] the current study. Measuring supernovae at even higher redshift provides a useful lever-arm to distinguish between different equations of state, unless the magnitude uncertainty increases more quickly with redshift than the redshift–distance relation diverges for differing values of the dark energy equation of state. For purposes of this analysis, we used 217 SNe Ia, from the initial large SNe surveys [1, 2],
supplemented by several more recent measurements, fitting the quoted magnitude uncertainty as a function of redshift to a linear relation. Although the oft-claimed quoted magnitude uncertainty per low-redshift supernova is 0.15 we found that a free fit to the data tended to prefer a slightly larger value. Motivated by these fits and for purposes of comparison, we considered both redshift independent uncertainties of 0.15 or 0.21 magnitudes, as well as redshift dependent uncertainties 0.15 + 0.079z or 0.21 + 0.079z mag.

Figure 4 shows the regions in the $w_0-w_a$ that are ruled out at 68% CL (region outside of the innermost closed curve) 95% CL (region outside of the middle curve) and 99.7% CL (region outside of the outermost curve, coinciding with the blue shaded region). In other words, the blue region corresponds to dark energy scenarios where the $\Lambda$CDM model can be ruled out at $>3\sigma$ confidence. This test is performed by ‘measuring’ the PCs from each $(w_0, w_a)$ model and comparing them to those from the $\Lambda$CDM model. We show cases when the error per SNe is 0.15 or 0.21 magnitudes (top and bottom left panels), and when it is increasing with redshift as 0.15 + 0.079z or 0.21 + 0.079z mag (top and bottom right panels).

region corresponds to dark energy scenarios where the $\Lambda$CDM model can be ruled out at $> 3\sigma$ confidence.

In the four panels of figure 4, we show cases when the error per SNe is 0.15 or 0.21 magnitudes, and when it is increasing with redshift as $0.15 + 0.079z$ or $0.21 + 0.079z$ mag. Clearly, increasing magnitude uncertainties with redshift can significantly affect the size of the indistinguishable parts of model space. Note too that the shape of the region that cannot be distinguished from $\Lambda$CDM is of characteristic shape, is elongated roughly along the $7(w_0 + 1) + w_a \approx 0$ direction—this is a direct consequence of the fact that the physically relevant quantity is $w(z) = w_0 + w_a z/(1 + z)$ whose average value is constrained to be around $-1$ and is best determined at redshift $z \approx 0.2$, see figure 2.

One other way to get a handle on whether $w \neq -1$ is to explore the most precise constraint on $w$ that might be obtainable at any single redshift. For, if $w \neq -1$ at any redshift, this will establish unambiguously that the dark energy is not a cosmological constant.

To explore the sensitivity of SNe for this purpose we parametrize the dark energy equation of state as in equation (4), and study the accuracy in the equation of state $w(z)$ at the best determined, or pivot, point. The pivot value is given by [29, 40]

$$w_{\text{pivot}} = w_0 - \frac{\text{Cov}(w_0, w_a)}{\text{Cov}(w_a, w_a)} w_a,$$

where $\text{Cov}(x, y)$ stands for elements of the covariance matrix element, while the pivot redshift at which $w(z)$ is best determined is given by

$$z_{\text{pivot}} = -\frac{\text{Cov}(w_0, w_a)}{\text{Cov}(w_0, w_a) + \text{Cov}(w_a, w_a)}.$$

Other parameters we marginalize over are the matter density relative to critical, $\Omega_M$, and the offset in the SNe Ia Hubble diagram, $M$. Throughout we assume a flat universe.

Using a Fisher matrix formalism, we estimate errors for a future SNAP-type survey [36] with 2800 SNe distributed in redshift out to $z = 1.7$ as given by [36], and combined with 300 local supernovae uniformly distributed in the $z = 0.03$–0.08 range. We study how the errors of the parameter of most interest, $w_{\text{pivot}}$, change as the individual SNe errors vary. We again assume that the statistical error per SNe scales with redshift as

$$\sigma_m(z) = \sigma_m(0) + A z,$$

where $\sigma_m(0) = 0.10, 0.15$ or 0.20, and we now let $A$ vary in order to ascertain sensitivity to this parameter.

As can be seen in figure 5, the constraints on $w_{\text{pivot}}$ are dependent on both $\sigma_m(0)$ and $A$; we can fit the constraint via an approximate relation

$$\sigma(w_{\text{pivot}}) \approx 0.17 (2\sigma_m(0) + A).$$

This equation shows that the slope of the statistical error versus redshift relation, $A$, contributes to $\sigma(w_{\text{pivot}})$ one half as much as the intercept of the same relation, $\sigma_m(0)$. For example, a 20% increase in the SNe error between the redshift of 0 and 1 (so that $A = 0.2 \sigma_m(0)$) will lead to a 10% increase in error associated with the pivot value of the equation of state.

4. Conclusions

SNe Ia are currently the strongest cosmological probes of dark energy, and are likely to remain the most solid source of information in the near future. This implies that control of the statistical errors is crucial. In particular, if we are ever to address the central question of cosmology, ‘Is the dark energy due to a cosmological constant?’, we need to be able to unambiguously determine that $w \neq -1$ at some, or all redshifts. The extent to which SNe Ia measurements will allow such a determination has been our prime concern in this analysis.

Figures 4 and 5 represent our primary results in this regard. Figure 4 makes clear that the model-independent constraints on $w(z)$, for a two-parameter class of deviations around $w(z) = -1$ (the two parameters are $w_0$ and $w_a$; see equation (4)), are quite sensitive to the measurement uncertainty in supernova magnitudes. Moreover, both figures 4 and 5 demonstrate that it is particularly important to attempt to maintain control over measurement uncertainties for higher redshift supernovae. As figure 4 also demonstrates, allowing for possible variations in $w$ with redshift implies that reducing the 95% confidence limit uncertainty in $w(z)$ for a fit near $w_0 = -1$ and $w_a = 0$ to significantly below 10% in $w_0$ will be challenging.

This latter point is even more important when considering whether one might utilize the measured value of $w(z)$ at the pivot point to attempt to discern some time at which $w \neq -1$. Redshift-dependent uncertainties in supernova magnitudes can easily almost double the inferred uncertainty in $w$ at this point. Moreover, only if the planned large-scale SNe surveys can maintain a uniform magnitude uncertainty per supernova, $\sigma_m$, less than 0.10 can we hope to derive a 95% confidence limit uncertainty in $w_{\text{pivot}}$ of less than about 10% (see figure 5) which itself may not be sufficient to distinguish some non-standard dark energy models from a cosmological constant.

It is not clear whether resolving the nature of dark energy is a 10-year or a 100-year problem. The answer partly depends on how much information measurements of dark energy can provide. Here we have addressed the former problem by studying the requirements of magnitude errors in a SNe Ia survey. Our results may be viewed as discouraging, or they may be viewed as inspiring...
for those observers who enjoy a demanding challenge. It is already becoming clear that, in order to have hope of detecting a deviation from the cosmological constant scenario, we need nature to be kind by producing a non-negligible deviation in \(w(z)\) at some redshift, \(\text{and} \) we need to control SNe magnitude errors, and systematics in particular, to high precision. If either one of these requirements is not met, we may have to rely on theorists to understand the nature of dark energy, which is an option over which we may have much less control.

References

[27] Starobinsky A A. 1998 JETP Lett. 68 757
[34] Simpson F and Bridle S. 2006 Phys. Rev. D 73 083001
Tegmark M 1997 Phys. Rev. D 55 5895
[36] SNAP online at http://snap.lbl.gov