PROBING THE EXPANSION HISTORY OF THE UNIVERSE BY MODEL-INDEPENDENT RECONSTRUCTION FROM SUPERNOVAE AND GAMMA-RAY BURST MEASUREMENTS

CHAO-JUN FENG AND XIN-ZHOU LI
Shanghai United Center for Astrophysics (SUCA), Shanghai Normal University, 100 Guilin Road, Shanghai 200234, P.R. China; fengcj@shnu.edu.cn, kyzh@shnu.edu.cn
Received 2015 September 17; accepted 2016 February 29; published 2016 April 5

ABSTRACT

To probe the late evolution history of the universe, we adopt two kinds of optimal basis systems. One of them is constructed by performing the principle component analysis, and the other is built by taking the multidimensional scaling approach. Cosmological observables such as the luminosity distance can be decomposed into these basis systems. These basis systems are optimized for different kinds of cosmological models that are based on different physical assumptions, even for a mixture model of them. Therefore, the so-called feature space that is projected from the basis systems is cosmological model independent, and it provides a parameterization for studying and reconstructing the Hubble expansion rate from the supernova luminosity distance and even gamma-ray burst (GRB) data with self-calibration. The circular problem when using GRBs as cosmological candles is naturally eliminated in this procedure. By using the Levenberg–Marquardt technique and the Markov Chain Monte Carlo method, we perform an observational constraint on this kind of parameterization. The data we used include the “joint light-curve analysis” data set that consists of 740 Type Ia supernovae and 109 long GRBs with the well-known Amati relation.

Key words: cosmological parameters – methods: data analysis

1. INTRODUCTION

One of the major targets for present observations is to learn the evolution history of the universe through the cosmic expansion rate. Observations of Type Ia supernovae (SNe Ia) have indicated that the universe is currently accelerating (Riess et al. 1998; Perlmutter et al. 1999). For a lack of deeper understanding, the cause of this acceleration is usually explained by introducing an exotic energy component called dark energy. There are many dark energy models that are based on different physical origins; see Li et al. (2011) for a recent review. A specific dark energy model is usually characterized by a small set of parameters. One can constrain these parameters by observational data to obtain the expansion rate of the universe. Although this approach is reasonable, the result is often dependent on which model one used. Thus, an interesting question is how to probe the cosmic evolution history from observations without any reference to a specific dark energy model.

All such researches are often called the cosmological-model-independent reconstruction of the cosmic expansion rate from observations, and it has been largely discussed in the literature. Most of them are based on a smoothing procedure in redshift bins (Huterer & Starkman 2003). Crittenden et al. (2009) and Simpson & Bridle (2006) have also performed principle component analysis (PCA) to reconstruct the dark energy equation of state. Mignone & Bartelmann (2008) have expanded the luminosity distance into a series of orthonormal functions as a basis to reconstruct the cosmic expansion rate. Maturi & Mignone (2009) have optimized this basis system to be capable to describe cosmologies independently of their background physics. The quality of the estimation of the luminosity distance is also improved. Li et al. (2014) have applied this method to determine the curvature parameter.

Gamma-ray bursts (GRBs) are the most intense explosions in the universe, and they can potentially be another standard candle living in the high redshifts. There are many GRBs observed at $0.1 < z < 8.1$, whereas the maximum redshift could be 10 or even larger in future observations. Thus, GRBs are a complementary probe to SNe Ia; see Schaefer (2007) for a review on the so-called GRB cosmology. However, there is a circularity problem when using GRBs as cosmological candles, because low-redshift GRBs at $z < 0.1$ are too few to calibrate the correction relation in a model-independent way. Hence, an input cosmology is needed to obtain the relation, but it leads to the circular problem when constraining cosmological parameters. To alleviate the circularity problem, some statistical methods have been proposed in Ghirlanda et al. (2004), such as the scatter method, the luminosity distance method, and the Bayesian method in Firmani et al. (2005). Liang et al. (2008) and Kodama et al. (2008) have suggested calibrating GRBs by using the SN Ia data; see also Wei (2010) for a relevant work. Another interesting approach was proposed by Li et al. (2008), in which they have treated the parameters involved in GRBs as free parameters and determined them simultaneously with other cosmological parameters by global fitting.

In this paper, we adopt two kinds of optimal basis system to probe the evolution history of the universe. One of them is constructed by performing the PCA following the method of Mignone & Bartelmann (2008) and Maturi & Mignone (2009). But there are some differences, which will be discussed in the next section. The other kind of optimal basis is built by taking the multidimensional scaling (MDS) approach (Borg et al. 2013, chap. 5), which is another powerful method to reconstruct the cosmic expansion rate. These basis systems have been optimized for different kinds of cosmological models that are based on different physical assumptions, even for a mixture model of them. Therefore, the so-called feature space that is projected from the basis systems is cosmological model independent, and it provides a parameterization for studying and reconstructing the Hubble expansion rate from the SN luminosity distance and even GRB data with self-calibration. By using the Levenberg–Marquardt (LM) technique and
the Markov Chain Monte Carlo (MCMC) method, we perform an observational constraint on this kind of parameterization. The data we used include the “joint light-curve analysis” (JLA) data set, which consists of 740 SNe Ia and 109 long GRBs with the well-known Amati relation. The circular problem when using GRBs as cosmological candles is naturally eliminated in this procedure. This may look like the global fitting method proposed by Li et al. (2008), but here we do not assume any cosmological models in advance.

The structure of this paper is as follows. In Section 2, we present the essential parts of the model-independent method and show how efficient it is when these methods are applied to optimize the basis for different kinds of cosmological models. The description of data and application of the method to reconstruct the evolution of the universe are given in Section 3. The discussions and conclusions are presented in Section 4.

2. MODEL-DEPENDENT METHOD

In this section, we will present applications of PCA and MDS to the JLA and GRB data to probe the evolution of the universe. At first, we will give the basic formulæ and expand cosmological observables into a finite sum of functions as bases. Then, the bases are optimized by using the PCA and MDS methods, respectively. In fact, the basic ideas of PCA and MDS are very similar, and one can finally obtain the most important components that could be used to describe the observables. It should be noted that we will directly focus on the cosmological observables such as distances instead of the physical quantities within a specific cosmological model, such as the equation of state of dark energy.

2.1. Basic Formulae for the Cosmic Expansion

In the Friedmann–Robertson–Walker metric, the luminosity distance is given by

\[ D_L(z) = \frac{c}{H_0} \frac{1 + z}{\Omega_k} \sin\left(\sqrt{\Omega_k} \int_0^z dz' \frac{E(z')}{E(z)}\right). \]  

(1)

with \( E(z) = H(z)/H_0 \), and \( \sin(x) = \sin(x), \sin(x) \) for \( k = 1, 0, -1 \), respectively. Here \( c \) is the speed of light, and \( \Omega_k \equiv -k/(a_0 H_0)^2 \) denotes the density of the spatial curvature at present. By taking the derivative of Equation (1) with respect to the redshift \( z \), one can obtain

\[ E(z)^{-1} = \frac{D'(z)}{\sqrt{1 + \Omega_k D^2(z)}}, \]  

(2)

where \( D(z) \) is the \( H_0 \)-independent comoving angular diameter distance that relates to the luminosity distance as

\[ D(z) = \frac{H_0}{c} \frac{D_L(z)}{1 + z}, \]  

(3)

and where the prime denotes the derivative with respect to the redshift \( z \). For a flat universe, Equation (2) could be written as

\[ E(z)^{-1} = \frac{H_0}{c} \left[ \frac{D_L(z)}{1 + z} - \frac{D_L(z)}{(1 + z)^2} \right]. \]  

(4)

Obviously, if the behaviors of both \( D(z) \) (or \( D_L(z) \)) and its derivative could be dug up from some observational data, one can obtain the evolution history of the universe from Equations (2) or (4). Although the data from observations of SNe Ia provide measurements of the distance modulus and redshifts, it is not a convenient way to take the derivative of the luminosity distance directly from the data, because the result would be extremely noisy and unreliable. Therefore, we need to first properly smooth the data by fitting an adequate function \( D(z) \) to the measurements in a model-independent way. The derivative can then be approximated by the derivative of \( D(z) \). This can be achieved through an expansion of \( D(z) \) into a finite sum of suitable functions \( p_i(z) \) such as

\[ D(z) = \sum_{i=1}^{M} c_i p_i(z). \]  

(5)

The \( M \) coefficients \( c_i \) can be determined by fitting the data, namely, \( c_i \) are those that minimize the \( \chi^2 \) statistic function. The number of terms to be included in the expansion depends on the choice of the orthonormal basis and the quality of the data. The basis \( \{p_i(z)\} \) could be arbitrary with ideal data, but it will not be in practice. Benitez-Herrera et al. (2012) have used the Gram–Schmidt orthonormalization to decompose the luminosity distance, and they found a systematic trend on the slope of the reconstructed cosmic expansion rate. It indicated that a randomly chosen system of orthonormal basis functions may not be well adapted to the behavior of the measured data. Maturi & Mignone (2009) have suggested optimizing the basis system by using PCA to reduce the number of coefficients \( M \) in Equation (5), and the possible bias introduced by the choice of the basis is also removed (see Benitez-Herrera et al. 2013). In this paper, we will make use of two optimal basis systems: one derived from the PCA and the other from the MDS approach. The number of coefficients required is minimized by either the PCA or MDS method. Besides, it also removes any bias introduced by the choice of the basis.

2.2. The Optimal Basis

2.2.1. The Training Set and Its Generator

To obtain the optimal basis, we start by writing the \( H_0 \)-independent comoving angular diameter distance in a column vector \( D = (D(z_1), D(z_2), \ldots, D(z_n))^T \in \mathbb{R}^n \), which can be regarded as a single point in an \( n \)-dimensional space. In the literature, \( n \) is often taken to be the number of data points from observations, but we will expand the variables at redshifts in a certain range with a small interval, say, 0.1. Then, we apply the spline interpolation method to calculate the distance to data points. Also, the range of redshifts is enlarged to cover that of the GRBs.

Now, we select a group of models that are believed to space the set of variable cosmologies and calculate \( D \) for each model to generate a set of vectors \( D_i \) with \( i = 1, 2, \ldots, M \), where \( M \) is the number of models. The ensemble of models \( T = (D_1, D_2, \ldots, D_M) \in \mathbb{R}^{n \times M} \) is called the training set, introduced by Maturi & Mignone (2009). In principle, the training set could be constructed from any model with arbitrary functions, but it is convenient to consider the models to be at least weakly resembling the data set (Maturi & Mignone 2009). In other words, one can choose any models, as long as the data set is tightly enclosed in the distribution of the \( M \)-point cluster in the \( n \)-dimensional space (Maturi & Mignone 2009; Benitez-Herrera et al. 2013; Li et al. 2014). To avoid confusion with a
specific cosmological model that determines the evolution of the universe, we would like to call these models the training set generators (TSGs), which means that they are only responsible for building the training set.

In the literature, the ΛCDM model with parameters uniformly sampled in the parameter space is often considered as a TSG to build the training set, but of course other kinds of cosmological models can be used as well, such as the dynamical dark energy models, modified gravity models, or even a mixture of them. However, the resulting optimal basis system is independent of any TSGs. To see this, we will take the nonflat ΛCDM, the wCDM, the Chevallier–Polarski–Linder (CPL) parametrization model (Chevallier & Polarski 2001; Linder 2003), the FSLL parametrization model without divergence (Feng et al. 2012), the holographic dark energy (HDE) model (Li 2004), the Dvali-Gabadadze-Porrati (DGP) model (Dvali et al. 2000; Deffayet 2001; Deffayet et al. 2002), the new agegraphic dark energy (NADE) model (Wei & Cai 2008), the Ricci dark energy (RDE) model (Feng & Li 2009; Gao et al. 2009), and their mixture as the TSGs. This also shows that no matter which kind of TSG we used, the dimensionality of the training set could be reduced efficiently by using either the PCA or the MDS analysis.

2.2.2. Building the Optimal Basis with PCA

PCA is a very useful statistical tool to reduce the dimensionality of an initially large training set space. Taking the mean of the training set, we obtain a reference model \( D_{\text{ref}} \) that defines the origin of the n-dimensional space:

\[
D_{\text{ref}} = \langle D_i \rangle = \frac{1}{M} \sum_{i=1}^{M} D_i \in \mathbb{R}^{n \times 1}.
\]  

Then, one can define the so-called covariance matrix by

\[
S = \frac{1}{M} \Delta \Delta^T, \quad \text{with} \quad \Delta = T - D_{\text{ref}} A \in \mathbb{R}^{n \times M},
\]

where \( A = (1, 1, 1, \ldots, 1) \in \mathbb{R}^{1 \times M} \). Therefore, the principle components (PCs) are the eigenvectors of the \( S \) matrix, which can be obtained by solving the eigenvalue problem \( S w_i = \lambda_i w_i \). In the following, the eigenvalues \( \lambda_i (i = 1, 2, \ldots, n) \) are sorted in a descendent sequence \( \lambda_1 > \lambda_2 > \ldots \), and the corresponding eigenvectors \( w_i \) are called the first PC \((w_1)\), the second PC \((w_2)\), and so on. This gives us the components in order of significance, and we can decide to ignore the components of lesser significance. For instance, if we choose only the first \( p \) eigenvectors, then the information content of the training set can be optimized via a linear transformation \( W: \mathbb{R}^n \rightarrow \mathbb{R}^p \) mapping the training set vectors into a so-called feature space:

\[
t_i = W^T D_i \in \mathbb{R}^p, \quad (i = 1, 2, \ldots, M).
\]

Here \( t_i \) are called the feature vectors, while the linear transformation is given by \( W = (w_1, w_2, \ldots, w_p) \). We do lose some information by ignoring \((w_{p+1}, w_{p+2}, \ldots, w_p)\), but if their eigenvalues are small enough, we do not lose much. Then, one could expand \( D(z) \) into the optimal basis

\[
D(z) = D_{\text{ref}} + \sum_{i=1}^{p} c_i w_i,
\]

with some coefficients \( c_i \) that will be determined by fitting data through \( \chi^2 \) minimization. The \( D'(z) \) is derived by taking the derivative with respect to the redshift on both sides of Equation (8). Since the eigenvalue \( \lambda_i \) is just the variance of \( \Delta \) along the vector \( w_i \), the percentage of variance we are willing to consider will then determine the number of PCs to be included in the reconstruction matrix \( W \), i.e., the value of \( p \). For example, we define the cumulative percentage of total variation (Jolliffe 2002; Benitez-Herrera et al. 2013 Section 6.1.1) as

\[
r_p = \frac{\sum_{i=1}^{p} \lambda_i}{\sum_{i=1}^{n} \lambda_i},
\]

and after setting a threshold, e.g., \( r_p > 99\% \), it will return the value of \( p \).

2.2.3. Building the Optimal Basis with MDS

MDS is another useful statistical tool to reduce the dimensionality of the training set. There are many types of MDS (Borg et al. 2013, chap. 5), which can be classified according to whether the similarities in data are qualitative (called nonmetric MDS) or quantitative (called metric MDS). In this paper, we will take the algorithms of the so-called classical MDS (CMDS), a special kind of metric MDS. In CMDS, a single Euclidean distance matrix is often used. From the training set \( T \) built before, one can easily construct a square-distance matrix \( Q \) (Borg et al. 2013, chap. 5), whose components are given by

\[
Q_{ij} = \sum_{k=1}^{n} (D_{ik} - D_{jk})^2 \in \mathbb{R}^{M \times M},
\]

with \( i = 1, 2, \ldots, M \). The matrix \( Q \) describes the dissimilarity of a pair of Ds. Centering the matrix \( Q \), we obtain the Gram matrix of \( Q \):

\[
G = -\frac{1}{2} Q Q^T,
\]

where \( Z = I_M - M^{-1} 1 1^T \), with \( I_M \) the identity matrix of order \( M \), and \( 1 \) a vector with a 1 in each of its entries. Then, we compute the eigenvalues and eigenvectors of the matrix \( G = \lambda_i \gamma_i \). And as before, \( \lambda_i \) are sorted in a descending sequence \( \lambda_1 > \lambda_2 > \lambda_3 > \ldots \). Therefore, by taking the first \( p \) positive eigenvalues and the corresponding first \( p \) eigenvectors, we get the MDS configuration with low dimension \( p < M \) as \( X = Y_i \Lambda_i^{1/2} \in \mathbb{R}^{M \times p} \), where \( Y_i = (y_{i1}, y_{i2}, \ldots, y_{ip}) \) and \( \Lambda_i = \text{diag} (\lambda_1, \lambda_2, \ldots, \lambda_p) \). Here \( X' \in \mathbb{R}^{p \times M} \) plays the same role as the feature space in the PCA mapping, i.e.,

\[
t = W^T T \in \mathbb{R}^{p \times M}.
\]

Finally, one could expand \( D(z) \) into

\[
D(z) = \sum_{i=1}^{p} c_i \tilde{w}_i,
\]

where the optimal bases are given by

\[
W = (\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_p) = T T^T Y_i \Lambda_i^{1/2} \Sigma \in \mathbb{R}^{n \times p}.
\]

Here \( \Sigma \) is a diagonal matrix to rescale the basis, such that the maximum absolute value of each \( \tilde{w}_i \) is equal to 1. We shall find that this model is also well consistent with observations.
Defining the following two cumulative quantities,

\[ r_p^{(1)} = \frac{\sum_{i=1}^{p} \lambda_i}{\sum_{i=1}^{n} \lambda_i}, \quad r_p^{(2)} = \frac{\sum_{i=1}^{p} \lambda_i^2}{\sum_{i=1}^{n} \lambda_i^2}, \] (14)

we can determine the value of \( p \) if either of the thresholds is satisfied, e.g., \( r_p^{(1)} > 99\% \), or \( r_p^{(2)} > 99\% \).

### 2.3. Efficiency of Different TSGs

For each TSG mentioned before, we construct the training set 20 times and average the values of \( r_1 \) and \( r_2 \) for the PCA method and the values of \( r_1^{(1,2)} \) and \( r_1^{(2,2)} \) for the MDS method. In each training set, there are 20 models besides the mixture one. In each model, the range of redshifts is \( z \in [0, 10] \) with an interval of 0.1, and the parameters of the TSGs are uniformly sampled, with boundaries listed in the last column of Table 1 to calculate the distance \( D_L \). \((i = 1, 2, \ldots, 20)\), so that the training set \( T \in \mathbb{R}^{100 \times 20} \). Results are summarized in Table 1. It should be noted that the training set from the mixture TSG contains 100 models that are randomly chosen from the other TSGs in Table 1 with the same parameter ranges. From Table 1, one can see that the value of \( p \) satisfying Equation (9) or Equation (14) could be very small, say, \( p = 1, 2 \), depending on which TSG is used.

We have chosen the ΛCDM model as the TSG to build the training set 10,000 times, and we have plotted the distribution of the values of \( r_1 \) in the top left panel of Figure 1 for the PCA analysis, and for each time the number of models is uniformly sampled, with boundaries listed in the last column of Table 1 to calculate the distance \( D_L \). \((i = 1, 2, \ldots, 20)\), so that the training set \( T \in \mathbb{R}^{100 \times 20} \). Results are summarized in Table 1. It should be noted that the training set from the mixture TSG contains 100 models that are randomly chosen from the other TSGs in Table 1 with the same parameter ranges. From Table 1, one can see that the value of \( p \) satisfying Equation (9) or Equation (14) could be very small, say, \( p = 1, 2 \), depending on which TSG is used.

### 3. THE COSMIC EXPANSION

#### 3.1. Data Descriptions

#### 3.1.1. JLA Supernova Data

The latest large SNe Ia data set is the JLA sample, which contains 740 spectroscopically confirmed SNe Ia covering the redshift range \( 0.01 < z < 1.3 \) with high-quality light curves. The distance estimator in this analysis assumes that SNe with identical color, shape, and galactic environment have on average the same intrinsic luminosity for all redshifts. This hypothesis is quantified by a linear model, yielding a standardized distance modulus (Betoule et al. 2014; Shafer 2015)

\[
\mu_{\text{obs}} = m_B - (M_B - A \cdot s + B \cdot C + P \cdot \Delta_M),
\] (15)

where \( m_B \) is the observed peak magnitude in rest-frame B band, \( M_B, s, C \) are the absolute magnitude, stretch, and color measures, respectively, which are specific to the light-curve fitter employed, and \( P(M_\text{iso} > 10^{10} M_\odot) \) is the probability that the SN occurred in a high-stellar-mass host galaxy. The stretch, color, and host-mass coefficients \( (A, B, \Delta_M, \text{respectively}) \) are nuisance parameters that should be constrained along with other cosmological parameters. On the other hand, the distance modulus predicted from a cosmological model for an SN at redshift \( z \) is given by

\[
\mu_{\text{model}}(z, \theta) = 5 \log_{10} \left[ \frac{D_L(z)}{10 \text{ pc}} \right],
\] (16)

where \( \theta \) are the cosmological parameters in the model and \( D_L(z) \) is the luminosity distance. For a given pair of the heliocentric-frame and the CMB-frame redshifts \( (z_{\text{hel}}, z_{\text{cmb}}) \) from the JLA data,

\[
D_L(z = z_{\text{cmb}}) = \frac{c}{H_0} \frac{1 + z_{\text{hel}}}{\sqrt{1 + z_{\text{hel}}}} \sinh \left( \frac{\sqrt{1 + z_{\text{hel}}}}{H_0} \int_0^{z_{\text{cmb}}} \frac{dz'}{E(z')} \right)
\]

\[
= (1 + z_{\text{hel}})r_A(z_{\text{cmb}}),
\] (17)

where \( r_A(z) \) is the comoving angular diameter distance. The \( \chi^2 \) statistic is then calculated in the usual way:

\[
\chi^2_{\text{SN}} = (\mu_{\text{obs}} - \mu_{\text{model}})^T C_{\text{SN}}^{-1} (\mu_{\text{obs}} - \mu_{\text{model}}),
\] (18)

with \( C_{\text{SN}} \) the covariance matrix of \( \mu_{\text{obs}} \).

#### 3.1.2. GRB Data

The GRB data we will use is compiled by Amati et al. (2008, 2009) and Amati (2010), in which there are 109 long GRBs with measured redshift \((0.1 < z < 8.1)\) and spectral peak energy. There are 50 GRBs at \( z < 1.4 \) and 59 GRBs at \( z > 1.4 \) in this data set (see Wei 2010, Tables 1 and 2). The well-known Amati correlation (Amati et al. 2002) in GRBs is given by

\[
\log_{10} E_{\text{iso}} = \lambda + b \log_{10} \left( \frac{E_{\gamma}}{300 \text{ KeV}} \right)
\] (19)
where $E_{\text{iso}}$ is the isotropic-equivalent radiated energy, while $E_{\text{p,i}}$ is the cosmological rest-frame spectral peak energy. Here $\lambda$ and $b$ are constants to be determined by observations; see Wei & Cai (2008). The isotropic-equivalent radiated energy $E_{\text{iso}}$ is related to the bolometric fluence $S_{\text{bolo}}$ of gamma rays in the GRB at redshift $z$:

$$E_{\text{iso}} = 4\pi D_L^2 S_{\text{bolo}} (1 + z)^{-1}.$$  \hspace{1cm} (20)

Then, from the GRB data one can obtain the distance modulus as

$$\mu_g = \frac{5}{2} \log_10 \left[ (1 + z) \left( \frac{E_{\text{p,i}}}{300 \text{ KeV}} \right)^b \frac{S_{\text{bolo}}^{-1}}{100 \text{ pc}^2} \right] + \frac{5\lambda}{2}, \hspace{1cm} (21)$$

with uncertainties

$$\sigma_{\mu_g}^2 = \left( \frac{5}{2 \ln 10} \right)^2 \left[ b^2 \left( \frac{\sigma_{E_{\text{p,i}}}}{E_{\text{p,i}}} \right)^2 + \left( \frac{\sigma_{S_{\text{bolo}}}}{S_{\text{bolo}}} \right)^2 + \sigma_{\text{sys}}^2 \right] \hspace{1cm} (22)$$

The $\chi^2$ statistic is then calculated by

$$\chi^2_g = \sum_{i=1}^{N} \frac{(\mu_g - \mu_{\text{model}})^2}{\sigma_{\mu_g}^2}, \hspace{1cm} (23)$$

with $N$ data points. Here $\sigma_{\text{sys}}$ in Equation (22) denotes the systematic error, which accounts for the extra scatter of the luminosity relation.

In the literature, the value of $\sigma_{\text{sys}}$ is often estimated by finding the value such that a $\chi^2_g$ fit to the luminosity calibration curve...
produces a reduced $c^2_g$ of unity (see Schaefer 2007). In fact, the systematic error should not depend on the number of data points $N$. Based on this assumption, we randomly choose a subset of the whole data set of 109 GRBs, i.e., $N = 20, 30, \ldots, 100$. Then, we find the value of $\sigma_{sys}$ such that the reduced $\chi^2$ is unity. We have performed this procedure 100 times and averaged the value of $\sigma_{sys}$, presented in Table 2. Also, the standard deviations of $\sigma_{sys}$ are given in the same table. Finally, we obtained averaging systematic error (weighted by the standard deviations) as

$$\sigma_{sys} = 0.7571,$$

which will be used in the next fitting procedures. Besides, from Table 2, it is clear that $\sigma_{sys}$ depends on the model through the $\chi^2_g$.

### 3.1.3. Fitting Results

During the fitting procedure, we have set the threshold in Equation (9) to be $r_p > 99.99\%$ for the PCA-model. For the MDS-model, we require either $r_p^{(1)} > 99.99\%$ or $r_p^{(2)} > 99.99\%$ to be satisfied; see Equation (14). Then, we get two parameters, $c_1$ and $c_2$, for the PCA-model and one parameter, $c_1$, for the MDS-model. For comparison, these two models are fitted to observations by using both the LM technique and the MCMC method. The current value of the Hubble parameter is fixed to be $H_0 = 70.0 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

At first, only JLA data are used to fit models. After marginalizing the nuisance parameters of JLA, we obtain $c_1 = 10.37 \pm 1.35$, $c_2 = 0.3617 \pm 0.3157$ with $\chi^2_{min}/\text{dof} = 683.001/738$ (LM) and $c_1 = 10.40^{+1.39}_{-1.36}$, $c_2 = 0.3540^{+0.3213}_{-0.3254}$.

Figure 2. MDS method. Top: histograms of the first eigenvalue (left) and second eigenvalue (right) in the percentage of total absolute eigenvalues from the $\Lambda$CDM TSG. Bottom left: the first four bases for the comoving angular diameter distances. To get a better visualization to see how they differ from each other, we plot their difference with their average and then divided by their number; $(\bar{w}_i - \langle \bar{w} \rangle)/i$ for $i = 1, 2, 3, 4$. Bottom right: the scree plot. All the values of parameters are uniformly sampled with boundaries $0.1 < \Omega_m < 0.9, 0.1 < \Omega_{\Lambda} < 0.9$ and $-0.1 < \Omega_k < 0.1$ as listed in Table 1. The range of redshifts is $z \in [0, 10]$ with an interval of 0.1.
Table 1
The Efficiency of PCA and MDS for Different TSGs

| TSGs  | PCA  | MDS  | Parameters                          |
|-------|------|------|--------------------------------------|
|       | $(r_1)$ | $(r_2)$ | $(r_1^{(1)})$ | $(r_1^{(2)})$ | $(r_2^{(1)})$ | $(r_2^{(2)})$ | $(r_1^{(2)})$ |
| ACDM  | 99.834% | 99.996% | 99.692% | 99.998% | 99.998% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, -0.1 < $\Omega_K < 0.1$ |
| wCDM  | 99.859% | 99.995% | 99.413% | 99.992% | 99.999% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, -0.1 < $\Omega_K < 0.1$, -1.5 < $w$ < -0.5 |
| CPL   | 99.906% | 99.998% | 99.701% | 99.997% | 99.998% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, -1.5 < $w_0$ < -0.5, -0.5 < $w_1$ < 0.5 |
| FSLL-Y | 99.895% | 99.997% | 99.767% | 99.997% | 99.999% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, -1.5 < $w_0$ < -0.5, -0.5 < $w_1$ < 0.5 |
| FSLL-Y | 99.905% | 99.998% | 99.795% | 99.998% | 99.998% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, -1.5 < $w_0$ < -0.5, -0.5 < $w_1$ < 0.5 |
| HDE   | 99.922% | 99.999% | 99.460% | 99.996% | 99.997% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, 0.1 < $C$ < 1.5 |
| DGP   | 99.804% | 99.999% | 99.260% | 99.998% | 99.999% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$ |
| NADE  | 99.897% | 99.999% | 99.706% | 99.998% | 99.998% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, 1.5 < $n$ < 3.5 |
| RDE   | 99.947% | 99.998% | 99.848% | 99.955% | 99.998% | 99.999% | 0.1 < $\Omega_M < 0.9$, 0.1 < $\Omega_V < 0.9$, 0.1 < $\alpha$ < 1.0 |
| Mixture | 99.780% | 99.995% | 99.250% | 99.998% | 99.995% | 99.999% | Take the same ranges as above. |

Notes.

a Chevallier & Polarski (2001); Linder (2003).
b Feng et al. (2012).
c Li (2004).
d Dvali et al. (2000); Deffayet (2001); Deffayet et al. (2002).
e Wei & Cui (2008).
f Gao et al. (2009); Feng & Li (2009).

Table 2
System Error Evaluations with Their Standard Deviations

| # of Data | PCA | MDS |
|-----------|-----|-----|
|           | $(\sigma_{\mu_i})$ | Std. | $(\sigma_{\mu_i})$ | Std. |
| 20        | 0.7910 | 0.1361 | 0.7637 | 0.1321 |
| 30        | 0.7445 | 0.0828 | 0.7492 | 0.0849 |
| 40        | 0.7588 | 0.0690 | 0.7794 | 0.0721 |
| 50        | 0.7568 | 0.0495 | 0.7561 | 0.0602 |
| 60        | 0.7627 | 0.0417 | 0.7546 | 0.0445 |
| 70        | 0.7558 | 0.0351 | 0.7608 | 0.0360 |
| 80        | 0.7509 | 0.0322 | 0.7581 | 0.0303 |
| 90        | 0.7498 | 0.0251 | 0.7595 | 0.0242 |
| 100       | 0.7531 | 0.0161 | 0.7586 | 0.0167 |

with $\chi^2_{\text{min}}/\text{dof} = 683.001/738$ (MCMC) for the PCA-model. We obtain $c_1 = 2.342 \pm 0.0163$ with $\chi^2_{\text{min}}/\text{dof} = 683.942/739$ (LM) and $c_2 = 2.342 \pm 0.0167$ with $\chi^2_{\text{min}}/\text{dof} = 683.942/739$ (MCMC) for the MDS-model. The contours for parameters $c_1$, $c_2$ of the PCA-model and their 1D histograms are plotted in Figure 3.

Next, both JLA and GRB data are used. The nuisance parameters of JLA are also marginalized since we do not have interest in them. However, the parameters $\lambda$ and $b$ in the Amati correlation (19) are kept free to see how good the calibration is. We obtain $c_1 = 11.52 \pm 0.85$, $c_2 = 0.0434 \pm 0.1546$, $\lambda = 52.850 \pm 0.041$, $b = 1.600 \pm 0.071$ with $\chi^2_{\text{min}}/\text{dof} = 787.592/845$ (LM) and $c_1 = 11.58 \pm 0.83$, $c_2 = 0.0329 \pm 0.1509$, $\lambda = 52.852 \pm 0.0399$, $b = 1.606 \pm 0.070$ with $\chi^2_{\text{min}}/\text{dof} = 787.601/845$ (MCMC) for the PCA-model. We obtain $c_1 = 2.231 \pm 0.016$, $\lambda = 52.841 \pm 0.037$, $b = 1.593 \pm 0.070$ with $\chi^2_{\text{min}}/\text{dof} = 787.767/846$ (LM) and $c_1 = 2.231 \pm 0.016$, $\lambda = 52.843 \pm 0.036$, $b = 1.590 \pm 0.074$ with $\chi^2_{\text{min}}/\text{dof} = 787.769/846$ (MCMC) for the MDS-model. The contours for parameters $c_1$, $c_2$, $\lambda$, $b$ and their 1D histograms are plotted in Figure 4 for the PCA-model. The contours for parameters $c_1$, $c_2$, $\lambda$, $b$, and their 1D histograms are plotted in Figure 5 for the MDS-model.

The calibration of data for the 109 GRBs is also shown in Figure 6, in which the propagated uncertainties of $\log_{10}E_{\text{iso}}$ and $\log_{10}E_{\text{p,i}}$ are estimated by

$$
\sigma_{\log_{10}E_{\text{iso}}} = \sqrt{\sigma_b^2 \left[ \log_{10} \left( \frac{E_{\text{p,i}}}{300 \text{ KeV}} \right) \right]^2 + b^2 \sigma_{\log_{10}E_{\text{p,i}}}^2},
$$

$$
\sigma_{\log_{10}E_{\text{p,i}}} = \frac{1}{\ln 10} \sigma_{E_{\text{p,i}}},
$$

It is clear that the calibration in this work is well consistent with data.
3.2. Reconstruction of History

Now, we are ready to reconstruct the history of the universe. The cosmic expansion rates \( E(z) = H(z)/H_0 \) with different spatial curvatures are plotted in Figures 7 and 8. The relative errors of \( E(z) \) are estimated by

\[
\frac{\sigma_E}{E} = \sqrt{\left( \frac{\Omega_k D^2}{1 + \Omega_k D^2} \right)^2 \left( \frac{\sigma_{\Omega_k}^2}{4 \Omega_k^2} + \frac{\sigma_D^2}{D^2} \right) + \frac{\sigma_{D'}^2}{D'^2}},
\]

where \( \sigma_D = \sqrt{\mathcal{W}^2 \sigma_{\hat{c}_i}^2 + \mathcal{W}^2 \sigma_{\hat{c}_j}^2} \), \( \sigma_{D'} = \sqrt{\mathcal{W}^2 \sigma_{\hat{c}_i}^2} \) for the PCA-model and \( \sigma_D = \sqrt{\mathcal{W}^2 \sigma_{\hat{c}_i}^2} \), \( \sigma_{D'} = \sqrt{\mathcal{W}^2 \sigma_{\hat{c}_i}^2} \) for the MDS-model.

From Figures 7 and 8, one can see that the relative error of \( E(z) \) in the MDS-model is about 10 times less than that in the PCA-model. And that is as it should be, because there is one parameter \( \hat{c}_i \) in the MDS-model, while there are two parameters \( c_1 \) and \( c_2 \) in the PCA-model. In both models, the relative error of \( E(z) \) is small at low redshifts, say, \( 0.5 < z < 1.0 \), since most of the data points belong to this range of redshifts.

It is interesting to see that in the MDS-model, the relative error of \( E(z) \) is a constant for a spatial-flat universe (\( \Omega_k = 0 \)). This could be seen from Equation (27): \( \sigma_{E}/E = \sigma_{D}/D' = \sigma_{\hat{c}_i}/\hat{c}_i \), since there is only one parameter \( \hat{c}_i \) in the MDS-model. Taking the best-fitting value for \( \hat{c}_i \) and its uncertainty \( \sigma_{\hat{c}_i} \), we obtain \( \sigma_{E}/E \approx 0.7\% \).

In Figures 7 and 8, the spatial curvature \( \Omega_k \) is chosen to show the differences of the cosmic expansion rate under different...
space geometries. In fact, Li et al. (2014) have already taken a model-independent approach to determine the spatial curvature by using the recent baryon acoustic oscillation (BAO) measurements. According to their conclusions, the errors of \( \Omega_k \) decrease with increasing redshift, and the best constraint is \( \Omega_k = -0.05 \pm 0.06 \) (at \( z = 2.36 \)). However, the errors of curvature at low redshifts are nearly of order unity (see Li et al. 2014, Figure 2). Considering the future BAO measurements, an improvement of at least one order of magnitude of \( \Omega_k \) could be expected at both low and high redshifts (Li et al. 2014).

The ratios of the cosmic expansion rate that was predicted from the \( \Lambda \)CDM, the wCDM, and the CPL model with their best-fitting parameters in Benitez-Herrera et al. (2013) to that reconstructed from the PCA-model and the MDS-model, i.e., \( H(z)/H_{PCA}(z) \) and \( H(z)/H_{MDS}(z) \), are plotted in Figure 9 and Figure 10, respectively.

From Figure 9, one can see that in the \( \Lambda \)CDM model the expansion rate is always smaller than that in the PCA-model, while in the wCDM model \( H(z) \) is first smaller than \( H_{PCA}(z) \) at low redshifts, then becomes larger than \( H_{PCA}(z) \) at medium redshifts, and finally gets smaller than \( H_{PCA}(z) \) again at high redshifts. In the CPL model, the behavior of \( H(z) \) is almost like that in the wCDM model except that \( H(z) \) is first larger than \( H_{PCA}(z) \). From Figure 10, one can see that the behavior of \( H(z) \) in these three physical models is almost identical, except a small difference at very low redshifts, and they are larger than \( H_{MDS}(z) \) at a large range of the redshifts.

Owing to the precision limit, we cannot determine the differences discussed above from the present observations, since these differences are really quite small.

4. DISCUSSIONS AND CONCLUSIONS

Cosmological variables such as the luminosity distance can be decomposed into some suitable basis. In this paper, we have proposed two methods, PCA and MDS, to optimize this basis. The projected feature spaces that describe the luminosity distance could then retain most of the origin information in a low-dimensional space. We call them the PCA-model and the MDS-model, respectively. It should be noticed that the procedures used above do not depend on any specific cosmological models. After that, observational data, including the JLA data set that consists of 740 SNe Ia and 109 long GRBs with the well-known Amati relation, are used to constrain the parameters of these two models by using the LM technique and the MCMC method. Finally, we obtain the evolution history of the universe including both the cosmic expansion rate and its relative errors, and we also compare the results with those predicted from the \( \Lambda \)CDM, the wCDM, and the CPL model with their best-fitting parameters.

We notice that either the PCA-model or the MDS-model could be used to perform the calibration to GRB data without any prior assumptions of a specific cosmological model. We
also estimate the system errors of GRB data. We can say with confidence that the error bars will become smaller when more accurate GRB data are obtained in the future.

C.J.F. would like to thank Puxun Wu for helpful discussions. This work is supported by National Science Foundation of China grant Nos. 11105091 and 11047138, “Chen Guang” project supported by Shanghai Municipal Education Commission and Shanghai Education Development Foundation grant No. 12CG51, National Education Foundation of China grant No. 2009312711004, Shanghai Natural Science Foundation, China grant No. 10ZR1422000, Key Project of Chinese Ministry of Education grant No. 211059, and Shanghai Special Education Foundation, No. ssd10004, and the Program of Shanghai Normal University.

REFERENCES

Amati, L. 2010, arXiv:1002.2232
Amati, L., Frontera, F., & Guidorzi, C. 2009, A&A, 508, 173
Amati, L., Frontera, F., Tavani, M., et al. 2002, A&A, 390, 81
Amati, L., Guidorzi, C., Frontera, F., et al. 2008, MNRAS, 391, 577
Benitez-Herrera, S., Ishida, E. E. O., Maturi, M., et al. 2013, MNRAS, 436, 854
Benitez-Herrera, S., Röpke, F., Hillebrandt, W., et al. 2012, MNRAS, 419, 513
Betoule, M., SDSS Collaboration, et al. 2014, A&A, 568, A22
Borg, I., Groenen-Partick, J. F., & Mair, P. 2013, Applied Multidimensional Scaling (New York: Springer)