RAA 2015 Vol. **15** No. **3**, 305–312 doi: 10.1088/1674–4527/15/3/001
http://www.raa-journal.org *http://www.iop.org/journals/raa*

Research in Astronomy and Astrophysics

Estimation of transition redshift based on Reinsch splines *

Guo-Dong Lü[1], De-Zi Liu[2], Shuo Yuan[2] and Tong-Jie Zhang[2]

[1] College of Information System and Management, National University of Defense Technology, Changsha 410073, China
[2] Department of Astronomy, Beijing Normal University, Beijing 100875, China; *adzliu@gmail.com*

Received 2013 November 4; accepted 2014 June 11

Abstract Many schemes have been proposed to define a model-independent constraint on cosmological dynamics, such as the nonparametric dark energy equation of state $\omega(z)$ or the deceleration parameter $q(z)$. These methods usually contain derivatives with respect to observational data with noise. However, there can be large uncertainties when one estimates values with numerical differentiation, especially when noise is significant. We introduce a global numerical differentiation method, first formulated by Reinsch, which is smoothed by cubic spline functions, and apply it to the estimation of the transition redshift $z_{\rm t}$ with a simulated expansion rate $E(z)$ based on observational Hubble parameter data. We also discuss some deficiencies and limitations of this method.

Key words: cosmological parameters — methods: data analysis — numerical

1 INTRODUCTION

The central task of modern cosmology is to uncover the dynamic evolution and geometric structure of the Universe. According to many cosmological observations, such as distant type Ia supernovae (SN Ia), cosmic microwave background (CMB) and so forth, the recent Universe, dominated by so-called dark energy, is undergoing accelerated expansion. The Lambda CDM model (ΛCDM) based on general relativity can provide an explanation that is consistent with observational data. However, the nature of dark energy still remains mysterious. In order to study the characteristics of the accelerating Universe, many different methods have been proposed to overcome the limitations of concrete cosmological models, e.g. the direct reconstruction of the equation of state of dark energy $\omega(z)$ (Clarkson & Zunckel 2010; Holsclaw et al. 2010) or the expansion rate $E(z)$ and deceleration parameter $q(z)$ (Daly & Djorgovski 2003).

Observations show that the Universe has undergone a "dynamic phase transition" from decelerating to accelerating expansion, which leads to a change in the sign of the deceleration parameter $q(z)$. The transition redshift $z_{\rm t}$, defined as $q(z_{\rm t}) \equiv 0$, has a profound impact on the evolution of the Universe. It is subject to different kinds of components and cosmological models (Lima et al. 2012). For example, in the ΛCDM model, the transition redshift can be derived by $z_{\rm t} = (2\Omega_\Lambda/\Omega_M)^{1/3} - 1$. Therefore, an accurate constraint on $z_{\rm t}$ will provide insights into how the Universe evolves. Different

∗ Supported by the National Natural Science Foundation of China.

groups have measured z_t with observations of SN Ia, baryon acoustic oscillations and CMB (Riess et al. 2007; Komatsu et al. 2011). In order to constrain z_t model-independently, various parametric expressions of $q(z)$ have been presented. Such parametric expressions, e.g. $q(z)$ and $\omega(z)$, are extremely convenient and effective for cosmological research.

Recently, observational Hubble parameter data (OHD) $H(z)$ have attracted much attention on the constraints related to cosmological parameters (Yi & Zhang 2007; Chen & Ratra 2011). Ma & Zhang (2011) and Zhang et al. (2010) have summarized the power and potential of OHD from a statistical point of view. At present, there are a total of 30 independent measurements of $H(z)$ (Simon et al. 2005; Gaztañaga et al. 2009; Stern et al. 2010; Moresco et al. 2012; Blake et al. 2012; Zhang et al. 2012; Chuang & Wang 2013; Busca et al. 2013). The relation between deceleration parameter $q(z)$ and $H(z)$ is $q(z) = \mathrm{d}H^{-1}(z)/\mathrm{d}t - 1$, so we can derive the transition redshift z_t as

$$z_\mathrm{t} = \left[\frac{\mathrm{d}\ln H(z)}{\mathrm{d}z}\right]^{-1}_{z=z_\mathrm{t}} - 1 = \left[\frac{\mathrm{d}\ln E(z)}{\mathrm{d}z}\right]^{-1}_{z=z_\mathrm{t}} - 1\,, \tag{1}$$

where $E(z) = H(z)/H_0$.

How to treat the numerical differentiation correctly is one common problem appearing in these works. The accuracy of numerical differentiation applied to noisy observational data is difficult to control. It encompasses many subtleties and pitfalls that may cause a large error in the actual computation. From a mathematical point of view (Ahnert & Abel 2007), the derivative of a given function is obtained by infinitesimal calculus; it is, however, impractical for real experimental data due to its property of being discrete. Additionally, for most cases, we must take the measured error or noise into consideration. An effective estimation of the derivative for a noisy sample has to tackle these two limitations.

To compute the numerical differentiation of a given sample without knowing the underlying function, typically one needs to obtain the approximation with some basic functions. Then one can hope that the derivative of the approximation is good enough to represent the actual situation. Many approaches have been developed, such as Principal Component Analysis (PCA) (Shapiro & Turner 2006; Benitez-Herrera et al. 2013; Nesseris & García-Bellido 2013), Gaussian Processes (Seikel et al. 2012a) and so forth. Reinsch (1967) developed another optimal algorithm to perform numerical differentiation with spline functions, hereafter referred to as Reinsch splines. The method can smooth the data globally by cubic spline functions. Therefore, the final differentiation can be derived analytically. In this work, we apply Reinsch splines to estimate z_t based on the current OHD, of which 28 measurements are assembled by Farooq & Ratra (2013). The other two are 79.69±3.32 and 86.45±3.27 at z = 0.24 and 0.43, respectively (Gaztañaga et al. 2009).

This paper is organized as follows: the fundamental algorithm of Reinsch splines and the corresponding errors are furnished in Section 2. Then we utilize this algorithm to estimate the transition redshift z_t using a simulated cosmological expansion rate $E(z)$ in Section 3. Finally, a brief summary and discussion will be given in Section 4.

2 METHOD OF REINSCH SPLINES

Reinsch splines were first proposed by Reinsch (1967) to replace strict interpolation by some kind of smoothing. The appropriate trial functions to estimate the experimental data are spline functions.

Given a sample (x_i, y_i) $(i = 1, ..., n)$ which satisfies

$$x_1 < x_2 < ... < x_n\,, \tag{2}$$

the problem can then be stated as a special instance of the *Tikhonov regularization* method, which looks for the minimum in the functional

$$\Phi(f) = \alpha\left\{\sum_{i=1}^{n}\left(\frac{y_i - f(x_i)}{\sigma_i}\right)^2 - S\right\} + ||f''(x)||^2\,, \tag{3}$$

where σ_i is the noise and $f(x)$ is the optimal square integrable function over the domain. $||f''(x)||$ denotes the L^2-norm

$$||f''(x)|| = \left(\int_{x_1}^{x_n} f''(x)^2 dx \right)^{1/2} . \tag{4}$$

S is a given positive constant, allowing for an implicit rescaling of the quantities σ_i, which controls the extent of smoothing. If σ_i is the estimate of the standard deviation of y_i, the value of S will lie within

$$n - \sqrt{2n} \le S \le n + \sqrt{2n} . \tag{5}$$

α is the Lagrangian parameter satisfying $\mathrm{d}\Phi(f)/\mathrm{d}\alpha = 0$ and $\alpha \neq 0$.

Hanke & Scherzer (2001) provided a rigorous proof that the minimizer of Equation (3) is a natural cubic spline. Following the notation of Reinsch (1967), we express $f(x)$ as

$$f(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, \tag{6}$$

where

$$x_i \le x \le x_{i+1}, \qquad i = 1, ..., n - 1 . \tag{7}$$

The spline functions satisfy specified smoothing and boundary conditions, but there is no need to exactly agree with the experimental data y_i on the node x_i. Once we obtain the coefficients a_i, b_i, c_i and d_i, the spline functions will be uniquely determined. A constructive algorithm for calculating the splines has been given by Reinsch (1967). In the remaining part of this section, we will focus on error analysis related to Reinsch splines.

Three sources of errors will impact the final evaluation of the numerical differentiation: experimental errors resulting from individual data points, truncation errors between the optimal $f'(x)$ and the true derivatives, and rounding errors due to the narrow length of a subinterval. The truncation errors greatly depend on the choice of the algorithm and the fitting accuracy. Hanke & Scherzer (2001) provided a formula for the truncation errors, and a rigorous proof can also be found. As for the experimental errors, the propagation of them is dealt with using the conventional expression

$$\sigma_m^2 = \sum_{i=1}^{N} \sigma_i^2 \left(\frac{\partial m}{\partial y_i} \right)^2 , \tag{8}$$

where m denotes the coefficients of the spline functions, and N is the number of the points in a subinterval.

3 ESTIMATION OF THE TRANSITION REDSHIFT

From Equation (1), the technique of numerical differentiation needs to be used to determine $z_{\rm t}$ if an $H(z)$ sample is available. Unfortunately, unlike the SN Ia dataset, the recent OHD are too penurious to provide a significant estimation of the transition redshift using Reinsch splines. There are some other reasons that make it difficult. First, the apparent uncertainty in the existing $H(z)$ will greatly increase the possibility of failure. Moreover, the sparse sample will lead optimal functions to seriously deviate from actuality, as well as the corresponding differentiation. Lima et al. (2012) suggest three different means to obtain $H(z)$, so it is expected that such shortcomings could be overcome by ongoing and future observations that enlarge the volume of the sample.

To extend the applications of OHD further and test the Reinsch splines, a simulated sample of $H(z)$ will be helpful. Furthermore, from Equation (1), the dimensionless expansion rate $E(z)$ will be a better choice for the present analysis. Therefore, we will concentrate on the technique of generating a sample of $E(z)$ in this section.

3.1 Simulated $E(z)$

Our simulation is based on the spatially flat ΛCDM model with $\Omega_m = 0.28$ and $\Omega_\Lambda = 0.72$. A Gaussian prior of $H_0 = 74.2 \pm 3.6$ km s^{-1} Mpc^{-1} suggested by Riess et al. (2009) is adopted. The expansion rate $E(z)$ in the fiducial model can be written as

$$E_{\rm fid}(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}\,. \tag{9}$$

Next we describe the simulation procedure in detail. Some similar but slightly different techniques have also been proposed by Ma & Zhang (2011), Wang & Zhang (2012) and Seikel et al. (2012b) to simulate the Hubble parameter.

Generally, the simulated $E(z)$ sample should take the fiducial model value $E_{\rm fid}$ as its expectation and follow the same systematic information and characteristic of the 'observed' $E(z)$ ($E_{\rm o}$) derived directly from OHD. However, this goal is difficult to reach, because we have quite a few real data points and very little knowledge about $E_{\rm o}$, e.g. the distribution of data points along the redshift axis. Therefore, our $E(z)$ simulation includes the following two procedures.

3.1.1 The offset estimation: $\varepsilon(z)$

To generate the $E(z)$ sample at any given redshift value, we introduce a variable $\varepsilon(z)$ satisfying

$$E_{\rm sim}(z) = E_{\rm fid}(z) + \varepsilon(z)\,, \tag{10}$$

which represents the offset between the fiducial value and simulation at a given z. In order to make $\langle E_{\rm sim}\rangle = E_{\rm fid}$, $\varepsilon(z)$ should be a random variable with respect to z.

Assuming there is no bias in the sign of the offset $\hat{\varepsilon}(z)$ = $E_{\rm o}(z) - E_{\rm fid}(z)$, we can make the offset satisfy $|\hat{\varepsilon}(z)| \leq E(z)\eta$ for most $E_{\rm o}(z)$ values, where η is a constant. The best estimate of η is 0.1320. As a result, we can denote

$$\varepsilon_{\pm}(z) = \pm E_{\rm fid}(z)\eta \tag{11}$$

as the boundaries of the offset values of a simulated expansion rate. As a random variable, $\varepsilon(z)$ follows the Gaussian distribution $N(0, \eta E_{\rm fid}(z)/2)$ so that the probability of $\varepsilon(z)$ falling within the offset domain is 95.4%.

3.1.2 The uncertainty estimation: $\sigma(z)$

There is an apparent trend in the errors of $E_{\rm o}(z)$: The uncertainty $\sigma_{E(z)}$ becomes larger as the redshift z increases except for six outliers of which four are from Zhang et al. (2012) and two are from Stern et al. (2010). In order to trace the global change in the errors, a conservative way is to ignore these points. Finally, we find the rest are right within the region between the lines

$$\sigma_+ = 0.2324z + 0.1365\,, \qquad \sigma_- = 0.1091z + 0.0393\,. \tag{12}$$

The midline between the two boundaries is

$$\sigma_m = \frac{1}{2}(\sigma_+ + \sigma_-)\,, \tag{13}$$

which represents the expected value of uncertainty $\sigma_{\rm sim}$ in the simulated expansion rate. Meanwhile, we make the $\sigma_{\rm sim}$ follow a Gaussian distribution $N(\sigma_m(z), \rho(z))$ at any given redshift, where

$$\rho(z) = \frac{\sigma_+ - \sigma_-}{4} \tag{14}$$

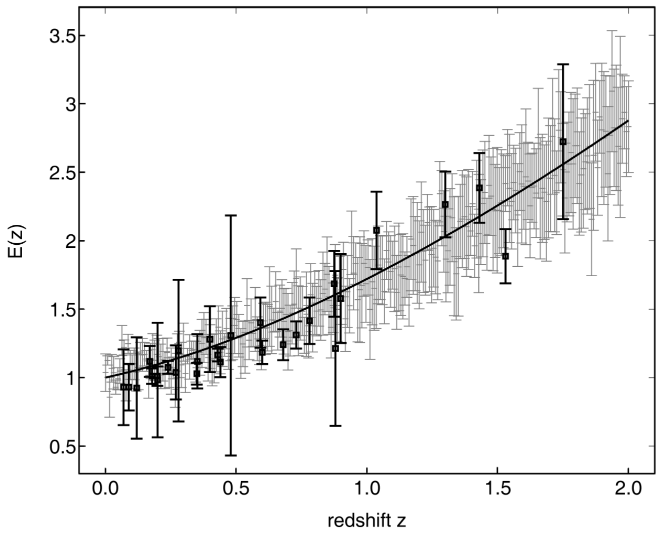

Fig. 1 Simulated $E(z)$ sample based on the fiducial ΛCDM model. The 'observed' $E_o(z)$ points are also displayed as black squares with error bars (E_o at $z = 2.3$ is not shown). The solid line represents the $E(z)$ in ΛCDM with $\Omega_m = 0.28$ and $\Omega_\Lambda = 0.72$.

ensuring that the simulated uncertainty $\sigma_{\rm sim}$ falls within the region between lines $\sigma_+(z)$ and $\sigma_-(z)$ with a probability of 95.4%.

Following the above two procedures, the method of generating a simulated $E(z)$ sample can be completed. First of all, we can calculate the fiducial values $E_{\rm fid}(z)$ from Equation (9). A random variable $\varepsilon(z)$ can be drawn from a Gaussian distribution, so $E_{\rm sim}(z)$ is generated via Equation (10) at a given z. Finally, the corresponding uncertainty $\sigma_{\rm sim}(z)$ is also estimated through another Gaussian distribution.

Figure 1 shows the final simulation in which we truncate the redshift to 2.0, which is enough for an estimate of the transition redshift. The readers are referred to Ma & Zhang (2011) for more detailed discussions about properties of the reconstruction.

3.2 Numerical Results

In this section, the method of Reinsch splines is applied to estimate the transition redshift using the simulated expansion rate $E(z)$.

In order to obtain the optimal function $f(x)$ defined in Equation (6), we need to minimize the functional $\Phi(f)$ (Eq. (3)). For such noisy data, constant S plays an important role. A proper choice of the constant will substantially boost the fitting accuracy of numerical differentiation. If the underlying function to be fitted is known, the classical least-squares fitting can be used. However, how to determine it without any model information, as done for the model-independent tests in cosmology, and how to judge whether the chosen S can approach the real situation remain to be solved. The general practice is to set $S \simeq n$ which satisfies Equation (5). It works fairly well when we test the method on different analytic functions with small errors. As an example, we test the method on an oscillating function $f(x) = 1 - 1.02\exp(-0.2x)\sin(0.98x + 4.9) + O(x)$, where $O(x)$ indicates a small Gaussian fluctuation. The result of the numerical differentiation is portrayed in Figure 2. We divide the differentiation into 14 equal bins and calculate the mean of each bin as the final result. Such reduction can mitigate the impact of singularities that arise from the initialization

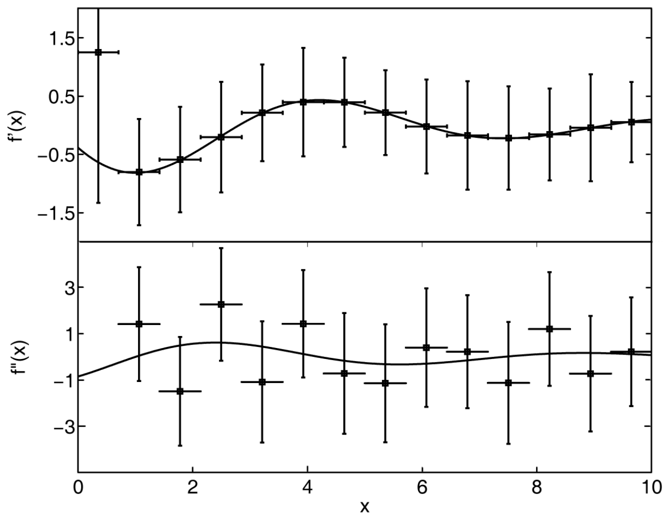

Fig. 2 *Upper panel*: numerical differentiation of function $f(x) = 1 - 1.02\exp(-0.2x)\sin(0.98x + 4.9) + O(x)$. The result is divided into 14 equal bins. The squares with error bars are the mean of each bin. The solid line is the theoretical derivative of $f(x)$. *Bottom panel*: second derivative of the same function. For clarity, the errors shown in the figure are ten times smaller than the actual errors.

and smoothing procedures. This is particularly important for our method, so the same consideration is also adopted for the following analysis of transition redshift.

However, if we increase the Gaussian fluctuation $O(x)$ to the case in which the relative error reaches ∼10%, we find that a simple assumption of $S \simeq n$ cannot reproduce the expected results, but makes the fitting fluctuant. It partly stems from the degeneracy between S and σ in Equation (3). As a result, constant S is sensitive to the relative error estimate. In our study, the relative errors of $E(z)$, i.e. $\sigma_{E(z)}/E(z)$, are up to 10% for most measurements. In order to constrain S quantitatively, here we introduce one possible estimation to minimize

$$\kappa = \sum_{i=1}^{n-1} ||\mathbf{v_{i+1}} - \mathbf{v_i}|| \,, \tag{15}$$

where $\mathbf{v_i} = (x_i, f'(x_i))$ and $||\cdot||$ represents the distance between two adjacent points. With this condition, we can suppress significant fluctuations arising from large relative errors and make the numerical differentiation as smooth as possible. Though such a minimization may not be optimal, it can weaken the contributions from large errors. We iteratively calculate S based on the above equation. For our sample, the best value is $S = 101.40$ and the corresponding $\kappa = 5.98$.

Once S is fixed, the optimal $f(x)$ can also be obtained with the algorithm presented in Section 2. Figure 3 shows the final numerical results. The redshift bins are [0.00, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00]. As a whole, we can fit the derivative fairly well, especially when $z > 0.5$. However, at low redshift ($z \leq 0.5$), there exists a slight deviation with respect to the expectation because of the default initialization of Reinsch splines and the significant error in our simulated sample. The best estimate of the transition redshift is $z_{\rm t} = 0.69^{+0.06}_{-0.14}$ which is approximately consistent with observational and theoretical results. Note that we just take the statistical errors and truncation errors into consideration due to the fact that the rounding errors are negligible in our sample.

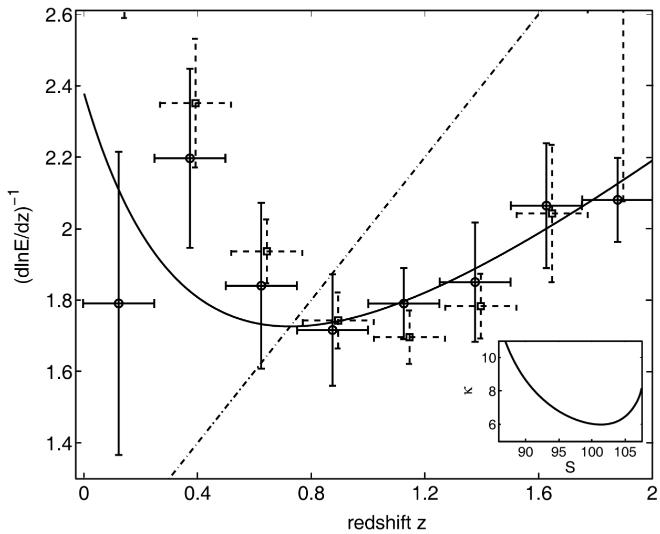

Fig. 3 The determination of the transition redshift using simulated $E(z)$. The redshift is divided into eight bins to estimate the average differentiation. Open circles indicate the result by our method, while open squares (slightly shifted) are derived by GaPP for comparison. The solid line is the derivative of $E(z)$ in the ΛCDM model (see Eq. (1)), while the dash-dotted line is $1 + z$. The intersection of the lines represents the transition redshift in the ΛCDM model. The inset at the lower right is an illustration of the variation of κ with respect to S. The transition redshift $z_{\rm t} \sim 0.69$ is estimated by using Reinsch splines.

For comparison, we also apply GaPP[1] (Seikel et al. 2012a), which is based on Gaussian processes, to our sample to estimate the numerical differentiation. Open squares in Figure 3 show the result in the same redshift bins without any prior information. Clearly, the method of Reinsch splines performs better than GaPP, especially at low redshift, $z < 0.8$, where the GaPP method significantly deviates from the expected result. Meanwhile, we conclude that both methods fail to fit the lower redshift range ($z < 0.5$) due to the impact of large relative errors. This was also discussed by Seikel et al. (2012a), who argued that large measured errors tend to reduce the fitting by the Gaussian process. However, by constraining the behavior of S, we can partly handle the impact of large errors. In addition, because our purpose in this work is to develop a completely model-independent method without any additional or prior assumptions, it should still be noted that for a fair comparison we do not add a prior in GaPP, though such a treatment will not cause GaPP to perform optimally.

4 DISCUSSION AND CONCLUSIONS

In the present work, we introduce a general method of numerical differentiation referred to as Reinsch splines, and preliminarily apply it to the analysis of the transition redshift $z_{\rm t}$ with a simulated expansion rate $E(z)$. The determination of constant S is essential to approach the optimal $f(x)$, and we provide an effective recipe to evaluate it when the experimental error σ is significant. This demonstrates that this method can estimate $z_{\rm t}$ using observational Hubble parameter data under current observation accuracy if we have enough measurements. Compared with other parametric or statistical methods (such as the Bayesian estimate), one merit of Reinsch splines is that we can avoid the potential impact of model assumptions.

[1] *http://www.acgc.uct.ac.za/ seikel/GAPP/index.html*

However, such an analytical method still has some deficiencies. Firstly, it is difficult to accurately predict the second derivative, especially when the error is relatively large. The bottom panel of Figure 2 shows one example. As a result, we cannot apply it to SN Ia data to successfully predict a model-independent equation of state for dark energy due to the existence of a second derivative. Such a shortcoming is expected to be overcome by combining it with other statistical techniques. In addition, the estimate of initial values is very crude due to the assumption of boundary conditions. One feasible solution is to relax the boundary conditions based on physical facts or some prior knowledge when constructing the optimal spline functions $f(x)$.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 11173006), the National Basic Research Program of China (973 program, 2012CB821804) and the Fundamental Research Funds for the Central Universities.

References

Ahnert, K., & Abel, M. 2007, Computer Physics Communications, 177, 764
Benitez-Herrera, S., Ishida, E. E. O., Maturi, M., et al. 2013, MNRAS, 436, 854
Blake, C., Brough, S., Colless, M., et al. 2012, MNRAS, 425, 405
Busca, N. G., Delubac, T., Rich, J., et al. 2013, A&A, 552, A96
Chen, Y., & Ratra, B. 2011, Physics Letters B, 703, 406
Chuang, C.-H., & Wang, Y. 2013, MNRAS, 435, 255
Clarkson, C., & Zunckel, C. 2010, Physical Review Letters, 104, 211301
Daly, R. A., & Djorgovski, S. G. 2003, ApJ, 597, 9
Farooq, O., & Ratra, B. 2013, ApJ, 766, L7
Gaztañaga, E., Cabré, A., & Hui, L. 2009, MNRAS, 399, 1663
Hanke, M., & Scherzer, O. 2001, American Mathematical Monthly, 108, 512
Holsclaw, T., Alam, U., Sansó, B., et al. 2010, Physical Review Letters, 105, 241302
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Lima, J. A. S., Jesus, J. F., Santos, R. C., & Gill, M. S. S. 2012, arXiv:1205.4688
Ma, C., & Zhang, T.-J. 2011, ApJ, 730, 74
Moresco, M., Cimatti, A., Jimenez, R., et al. 2012, J. Cosmol. Astropart. Phys., 8, 006
Nesseris, S., & García-Bellido, J. 2013, Phys. Rev. D, 88, 063521
Reinsch, C. H. 1967, Numerische mathematik, 10, 177
Riess, A. G., Strolger, L.-G., Casertano, S., et al. 2007, ApJ, 659, 98
Riess, A. G., Macri, L., Casertano, S., et al. 2009, ApJ, 699, 539
Seikel, M., Clarkson, C., & Smith, M. 2012a, J. Cosmol. Astropart. Phys., 6, 036
Seikel, M., Yahya, S., Maartens, R., & Clarkson, C. 2012b, Phys. Rev. D, 86, 083001
Shapiro, C., & Turner, M. S. 2006, ApJ, 649, 563
Simon, J., Verde, L., & Jimenez, R. 2005, Phys. Rev. D, 71, 123001
Stern, D., Jimenez, R., Verde, L., Kamionkowski, M., & Stanford, S. A. 2010, J. Cosmol. Astropart. Phys., 2, 008
Wang, H., & Zhang, T.-J. 2012, ApJ, 748, 111
Yi, Z.-L., & Zhang, T.-J. 2007, Modern Physics Letters A, 22, 41
Zhang, C., Zhang, H., Yuan, S., et al. 2012, arXiv:1207.4541
Zhang, T.-J., Ma, C., & Lan, T. 2010, Advances in Astronomy, 2010, 184284