WHAT ARE THE $\Omega m h^2 (z_1, z_2)$ AND Om $(z_1, z_2)$ DIAGNOSTICS TELLING US IN LIGHT OF $H(z)$ DATA?

XIAOGANG ZHENG$^{1,2}$, XUHENG DING$^{1,3}$, MAREK BIESIADA$^{1,2}$, SHUO CAO$^1$, AND ZONG-HONG ZHU$^1$

$^1$ Department of Astronomy, Beijing Normal University, Beijing 100875, China
$^2$ Department of Astrophysics and Cosmology, Institute of Physics, University of Silesia, Uniwersytecka 4, 40-007, Katowice, Poland
$^3$ Department of Physics and Astronomy, University of California, Los Angeles, CA 90095-1547, USA

Received 2016 February 17; revised 2016 April 5; accepted 2016 April 21; published 2016 June 27

ABSTRACT

The two-point diagnostics $Om(z_1, z_2)$ and $Om h^2(z_1, z_2)$ have been introduced as an interesting tool for testing the validity of the $\Lambda$ cold dark matter ($\Lambda$CDM) model. Recently, Sahni et al. combined two independent measurements of $H(z)$ from baryon acoustic oscillation (BAO) data with the value of the Hubble constant $H_0$, and used the second of these diagnostics to test the $\Lambda$CDM (a constant equation-of-state parameter for dark energy) model. Their result indicated a considerable tension between observations and predictions of the $\Lambda$CDM model. Since reliable data concerning the expansion rates of the universe at different redshifts $H(z)$ are crucial for the successful application of this method, we investigate both two-point diagnostics on the most comprehensive set of $N = 36$ measurements of $H(z)$ from BAOs and the differential ages (DAs) of passively evolving galaxies. We discuss the uncertainties of the two-point diagnostics and find that they are strongly non-Gaussian and follow the patterns deeply rooted in their very construction. Therefore we propose that non-parametric median statistics is the most appropriate way of treating this problem. Our results support the claim that $\Lambda$CDM is in tension with $H(z)$ data according to the two-point diagnostics developed by Shafieloo, Sahni, and Starobinsky. However, other alternatives to the $\Lambda$CDM model, such as the wCDM or Chevalier–Polarski–Linder models, perform even worse. We also note that there are serious systematic differences between the BAO and DA methods that ought to be better understood before $H(z)$ measurements can compete with other probes methods.

Key words: cosmology: observations – dark energy – methods: statistical

1. INTRODUCTION

Soon after the discovery of the accelerating expansion of the universe (Riess et al. 1998; Perlmutter et al. 1999), the $\Lambda$ cold dark matter ($\Lambda$CDM) model was proposed as the simplest explanation for this phenomenon. Since then it has survived increasingly stringent tests, not only related to the late accelerating phase of expansion but also as a framework in which the precise cosmic microwave background (CMB) data acquired up to the present time could be best understood. However, many researchers have raised serious concerns with the claim that the $\Lambda$CDM model was an ultimate solution. Initially because of conceptual problems like fine tuning, but also because of some discrepancies like small-scale anomalies and the recently reported tension between Planck and Canada–France–Hawaii–Telescope Lensing Survey measurements—see e.g., Macaulay et al. (2013), Ade et al. (2014), Raveri (2016). This has motivated many researchers to take on the challenge of testing the very foundations of the $\Lambda$CDM. For example, Zunckel & Clarkson (2008) formulated a ”litmus test” for the $\Lambda$CDM model. Others challenged even more fundamental aspects like the Copernican principle (Uzan et al. 2008; Valkenburg et al. 2014).

However, the most popular probe used to test the $\Lambda$CDM model and to seek evidence of an evolving cosmic equation of state is the one initiated by Sahni et al. (2008) after they introduced the one point $Om(z)$ diagnostic and generalized it to the two-point case $Om(z_1, z_2) = Om(z_1) - Om(z_2)$. Later, they developed this further, introducing in Shafieloo et al. (2012) the improved two-point diagnostic $Om h^2(z_1, z_2)$, which they subsequently used in Sahni et al. (2014) to perform this test on three accurately measured values of $H(z)$ from baryon acoustic oscillations (BAOs). These were: the $H(z = 0)$ measurement by Riess et al. (2011) and Ade et al. (2014), the $H(z = 0.57)$ measurement from Sloan Digital Sky Survey Data Release 9 (SDSS DR9) (Samushia et al. 2013), and the most recent $H(z = 2.34)$ measurement from the Ly$\alpha$ forest in SDSS DR11 (Delubac et al. 2015). They found that all three values of the two-point diagnostics $Om h^2(z_1, z_2)$ were in strong tension with the $\Omega_{m0}h^2$ reported from Planck (Ade et al. 2014). It has also been noted (Sahni et al. 2014; Delubac et al. 2015) that the Ly$\alpha$ forest measurement at $z = 2.34$ could be in tension not only with the $\Lambda$CDM model but also with other dark energy models based on general relativity. Because such a conclusion could be of paramount importance for dark energy studies, in our recent paper (Ding et al. 2015), we performed this test with a larger sample of $H(z)$ comprising six BAO measurements and 23 data points from cosmic chronometers (the differential ages (DAs) of passively evolving galaxies). Essentially, our conclusion was that the tension between the $H(z)$ data and the $\Lambda$CDM model exists. In this paper we study the performance of the $Om h^2(z_1, z_2)$ and $Om(z_1, z_2)$ two-point diagnostics in more detail. In Section 2 we briefly review the concepts of $Om h^2(z_1, z_2)$ and $Om(z_1, z_2)$. Section 3 reviews the $H(z)$ data. A detailed analysis of the statistical properties of both two-point diagnostics and the results obtained with them on $H(z)$ data is the subject of Section 4. Finally we conclude in Section 5.

2. THE $Om(z)$ METHODOLOGY IN BRIEF

The so-called $Om(z)$ diagnostic has been introduced as an alternative to the common approach of testing models of accelerated expansion of the universe using the phenomenological assumption of a perfect fluid with an equation of state $p = w \rho$ filling the universe (in addition to pressureless matter and now dynamically negligible radiation). Cosmological constant $\Lambda$ corresponds formally to $w = -1$. The model
The expansion rate is sometimes denoted as $E(z)$, but a clever observation that the Friedmann equation in this model: $H(z)^2 = H_0^2 [\Omega_{m,0}(1+z)^3 + 1 - \Omega_{m,0}]$ can be rearranged as

$$\Omega_m(z) \equiv \frac{\tilde{H}^2(z) - 1}{(1+z)^3 - 1} = \Omega_{m,0},$$

where $\tilde{H}(z) \equiv H(z)/H_0$. In the literature this dimensionless expansion rate is sometimes denoted as $E(z)$. We retain the notation reminiscent of the Hubble function $H(z)$ and use a tilde when it is normalized by the Hubble constant $H_0$ (present expansion rate). We will also use a similar quantity $h(z) \equiv H(z)/100$ km s$^{-1}$ Mpc$^{-1}$. To finish our remarks on the nomenclature conventions, let us recall that for historical reasons it is commonly accepted to use the notation $h \equiv H_0/100$ km s$^{-1}$ Mpc$^{-1}$ for the dimensionless Hubble constant. What is remarkable about Equation (1) is the fact that the left-hand side is a function of redshift and the right-hand side is a number, so the falsifying power of Equation (1) is strong. If we knew, from observations, the expansion rates at different redshifts we would be able to differentiate between $\Lambda$CDM and other dark energy models (including evolving dark energy). Although very attractive from the theoretical point of view, this test was not easy to perform because there were no accurate direct measurements of $H(z)$ at the time of its formulation, so the researchers willing to use it were forced to reconstruct $H(z)$ from distance measurements of Type I supernovae (SNe Ia) and this resulted in an increased uncertainty. Currently we are in much better position, having at our disposal a considerable amount of $H(z)$ measurements obtained from BAO and DA techniques, as will be discussed later. Another issue was that the $\Omega_m(z)$ diagnostic in the $\Lambda$CDM model should not only be constant but exactly equal to the present matter density parameter $\Omega_{m,0}$ which is not easy to measure directly, and its indirectly inferred value from CMB or SN Ia data was also a subject of debate.

Therefore Shafieloo et al. (2012) developed this method further by noting that the two-point diagnostics:

$$\Omega_m(z_i, z_j) \equiv \Omega_m(z_i) - \Omega_m(z_j)$$

$$= \frac{\tilde{H}^2(z_i) - 1}{(1+z_i)^3 - 1} - \frac{\tilde{H}^2(z_j) - 1}{(1+z_j)^3 - 1}$$

should always vanish in the $\Lambda$CDM model:

$$\Omega_m(z_i, z_j)_{\Lambda CDM} = 0$$

for all $i, j$. If we just knew the expansion rates at different redshifts, we would be able to tell whether these data are consistent with the $\Lambda$CDM or not without any need to know the matter density parameter. Compared to the original $\Omega_m(z)$ diagnostic, this two-point diagnostic has another advantage: a sample of $n$ measurements offers us $\frac{n(n-1)}{2}$ different values of two-point diagnostics. As we will see later, this happens at the cost of creating complex statistical properties of two-point diagnostics. Moreover, vanishing $\Omega_m(z_i, z_j)_{\Lambda CDM}$ is again just the litmus test. If we want to distinguish between different dark energy models, we need to write down the corresponding theoretical expression expected for the right-hand side. For the simplest phenomenology of dark energy with a constant equation of state parameter $w = \text{const.}$, the theoretical expression for Equation (2) should be

$$\Omega_m(z_i, z_j)_{\text{wCDM}} = \left(1 - \Omega_{m,0}\right) \left[\frac{(1+z_i)^3(1+w) - 1}{(1+z_j)^3 - 1} - \frac{(1+z_j)^3(1+w) - 1}{(1+z_i)^3 - 1}\right].$$

Therefore, assuming the redshift ordering $z_j > z_i$, inequality $\Omega_m(z_i, z_j) > 0$ implies quintessence ($w > -1$) while $\Omega_m(z_i, z_j) < 0$ implies the phantom scenario ($w < -1$). Similarly, for the evolving equation of state (Chevalier & Polarski 2001; Linder 2003) modeled by the Chevalier–Polarski–Linder (CPL) parameterization, the expression should be

$$\Omega_m(z_i, z_j)_{\text{CPL}} = \left(1 - \Omega_{m,0}\right) \left[\frac{(1+z_i)^3(1+w_0 + w\eta) \exp(-3w_0 z_i/(1+z_i)) - 1}{(1+z_j)^3 - 1} - \frac{(1+z_j)^3(1+w_0 + w\eta) \exp(-3w_0 z_j/(1+z_j)) - 1}{(1+z_i)^3 - 1}\right].$$

In their recent paper, Sahni et al. (2014) used a slightly different version of a two-point diagnostic

$$\Omega m h^2(z_i, z_j)_{\text{wCDM}} = \frac{\tilde{H}^2(z_i) - \tilde{H}^2(z_j)}{(1+z_j)^3 - (1+z_j)^3},$$

which again should be equal to $\Omega_{m,0} h^2$ in the framework of the $\Lambda$CDM model. For dark energy with a constant equation of state $w = \text{const.}$, the theoretical expression of Equation (5) should be

$$\Omega m h^2(z_i, z_j)_{\text{wCDM}} = \Omega_{m,0} h^2 + \left(1 - \Omega_{m,0}\right) h^2 \left[\frac{(1+z_i)^3(1+w) - 1}{(1+z_j)^3 - 1} - \frac{(1+z_j)^3(1+w) - 1}{(1+z_i)^3 - 1}\right].$$

and for the CPL parameterization, one can expect that

$$\Omega m h^2(z_i, z_j)_{\text{CPL}} = \Omega_{m,0} h^2 + \left(1 - \Omega_{m,0}\right) h^2 \left[\frac{((1+z_i)^3(1+w_0 + w\eta) e^{-3w_0 z_i}/(1+z_i))}{((1+z_j)^3(1+w_0 + w\eta) e^{-3w_0 z_j}/(1+z_j))}/(1+z_j)^3 - (1+z_j)^3\right].$$

3. DATA

Our data comprise 36 measurements of $H(z)$ acquired by means of two different techniques. The first part of the data comes from cosmic chronometers (Jimenez & Loeb 2002), i.e., massive, early-type galaxies evolving passively on a timescale longer than their age difference. Certain features of their spectra, such as the D4000 break at 4000 Å indicative of the evolution of their stellar populations, enable us to measure the age difference of such galaxies. Hence, we use the abbreviation DA to denote the cosmic chronometer technique. The most recent results obtained with this technique on the very rich data from the Baryon Oscillation Spectroscopic Survey DR9 have
been published by Moresco et al. (2016). Therefore we use 30 measurements of \(H(z)\) via the DA technique: 23 of these are the same as we used in Ding et al. (2015), supplemented with two high redshift DA data points from Moresco (2015) and five more \(H(z)\) data points from Moresco et al. (2016). The second part of our data comes from the analysis of BAOs. The BAO data comprise six measurements. Table 1 summarizes our data and also provides references for the original sources.

Previous papers by Sahni et al. (2014) and Ding et al. (2015) suggested a tension between \(\Omega_m h^2\) calculated from \(H(z)\) data and \(\Omega_m h^2 = 0.1426 \pm 0.0025\) from the Planck satellite (Ade et al. 2014). In the first step of the current study we will readdress this issue using the larger data set of Table 1. We will go a step further, considering also the \(\Omega_m(z_r, z_i)\) diagnostic, and for this purpose we need to assume a specific value of the Hubble constant \(H_0\). We take the value \(H_0(\text{Planck}) = 67.4 \pm 1.4\) suggested by (Ade et al. 2014). Moreover, we will also consider two more parameterizations for the dark energy, other than the \(\Lambda\)CDM model, namely wCDM and CPL. Therefore, in order to calculate the theoretically expected values of the \(\Omega_m(\Omega_m, z)\) and \(\Omega_m(z_r, z_i)\) two-point diagnostics, we will use \(\Omega_m(\Omega_m, H_0)\), and equation of state parameters in the wCDM and CPL models as reported by Betoule et al. (2014) (their Tables 14 and 15). These parameters have been constrained by a combination of the Planck and WMAP satellite measurements of the CMB temperature fluctuations used jointly with the characteristic scale of the BAO and the SN Ia Joint Light Analysis (JLA) compilation. The parameters are summarized in Table 2.

Because the \(H(z)\) data set we used is inhomogeneous we performed our analysis of the two-point diagnostics not only on a full sample of \(N = 36\) combined DA+BAO measurements, but also on the DAs (\(N = 30\)) and BAOs (\(N = 6\)) separately. Moreover, since the \(z = 2.34\) measurement (Delubac et al. 2015) turns out to have a large effect on BAO results, we have also considered an \(N = 35\) sub-sample by excluding this measurement from the full DA+BAO sample. The above mentioned effect can be seen in Figure 1 where we have used different samples of \(H(z)\) to constrain the \((\Omega_m, H_0)\) parameters in the spatially flat \(\Lambda\)CDM model where \(H(z) = H_0\sqrt{\Omega_m(1 + z)^3 + 1 - \Omega_m}\). One can see that inclusion of the \(z = 2.34\) data point dramatically improves the BAO fit but there is still a mismatch between the BAO and DA 68% confidence regions.

### Table 1

Data of the Hubble Parameter \(H(z)\) at Different Redshifts \(z\); \(H(z)\) and \(\sigma_H\) Are in Units of \((\text{km s}^{-1} \text{Mpc}^{-1})\)

| \(z\) | \(H(z)\) | \(\sigma_H\) | Method | Reference |
|------|----------|-------------|--------|-----------|
| 0.07 | 69       | 19.6        | DA     | Moresco et al. (2014) |
| 0.09 | 69       | 12          | DA     | Jimenez et al. (2003) |
| 0.12 | 68.6     | 26.2        | DA     | Moresco et al. (2014) |
| 0.17 | 83       | 8           | DA     | Simon et al. (2005)   |
| 0.1791 | 75 | 4           | DA     | Moresco et al. (2012) |
| 0.1993 | 75 | 5           | DA     | Moresco et al. (2012) |
| 0.2   | 72.9     | 29.6        | DA     | Blake et al. (2005)   |
| 0.27  | 77       | 14          | DA     | Moresco et al. (2005) |
| 0.28  | 88.8     | 36.6        | DA     | Moresco et al. (2014) |
| 0.35  | 82.7     | 8.4         | BAO    | Chuang & Wang (2013)  |
| 0.3519 | 83 | 14          | DA     | Moresco et al. (2012) |
| 0.3802 | 83 | 13.5        | DA     | Moresco et al. (2016) |
| 0.4   | 95       | 17          | DA     | Simon et al. (2005)   |
| 0.4004 | 77 | 10.2        | DA     | Moresco et al. (2016) |
| 0.4247 | 87.1| 11.2        | DA     | Moresco et al. (2016) |
| 0.44  | 82.6     | 7.8         | BAO    | Blake et al. (2012)   |
| 0.4497 | 92.8| 12.9        | DA     | Moresco et al. (2016) |
| 0.4783 | 80.9| 9           | DA     | Moresco et al. (2016) |
| 0.48  | 97       | 62          | DA     | Stern et al. (2010)   |
| 0.57  | 92.9     | 7.8         | BAO    | Anderson et al. (2014) |
| 0.5929 | 104| 13          | DA     | Moresco et al. (2012) |
| 0.6   | 87.9     | 6.1         | BAO    | Blake et al. (2012)   |
| 0.6797 | 92 | 8           | DA     | Moresco et al. (2012) |
| 0.73  | 97.3     | 7           | BAO    | Blake et al. (2012)   |
| 0.7812 | 105| 12          | DA     | Moresco et al. (2012) |
| 0.8754 | 125| 17          | DA     | Moresco et al. (2012) |
| 0.88  | 90       | 40          | DA     | Stern et al. (2010)   |
| 0.9   | 117      | 23          | DA     | Simon et al. (2005)   |
| 1.037 | 154      | 20          | DA     | Moresco et al. (2012) |
| 1.3   | 168      | 17          | DA     | Simon et al. (2005)   |
| 1.363 | 160      | 33.6        | DA     | Moresco et al. (2015) |
| 1.43  | 177      | 18          | DA     | Simon et al. (2005)   |
| 1.53  | 140      | 14          | DA     | Simon et al. (2005)   |
| 1.75  | 202      | 40          | DA     | Simon et al. (2005)   |
| 1.965 | 186.5    | 50.4        | DA     | Moresco et al. (2015) |
| 2.34  | 222      | 7           | BAO    | Delubac et al. (2015) |

### Table 2

The Best-fitted Values of Parameters for Three Dark Energy Models Obtained from Joint Analysis of Planck+WP+BAO+JLA Data (Betoule et al. 2014)

| Model | \(\Omega_m(0)\)    | \(H_0\)   | \(w\) | \(w_0\) | \(w_a\) |
|-------|-------------------|-----------|-------|--------|--------|
| \(\Lambda\)CDM | 0.305 ± 0.010 | 68.34 ± 1.03 | ... | ... | ... |
| wCDM  | 0.303 ± 0.012 | 68.50 ± 1.27 | −1.027 ± 0.055 | ... | ... |
| CPL   | 0.304 ± 0.012 | 68.59 ± 1.27 | ... | −0.957 ± 0.124 | −0.336 ± 0.552 |
4. RESULTS

In order to gain insight concerning the $\Omega m_h^2(z_i, z_j)$ and $\Omega m(z_i, z_j)$ two-point diagnostics calculated for every combination of pairs taken from the 36 $H(z)$ data points, i.e., 630 pairs of $(z_i, z_j)$ in total, Figure 2 displays these diagnostics together with their uncertainties as a function of redshift difference $\Delta z = |z_i - z_j|$. There are some interesting features regarding the uncertainties of the two-point diagnostics. One can see that they are apparently non-Gaussian and the two-point diagnostics—in particular $\Omega m_h^2(z_i, z_j)$—are heteroscedastic. The reasons for this can be understood by looking at the formulae for the corresponding uncertainties. Namely, by applying the error propagation formula to the definitions of $\Omega m_h^2(z_i, z_j)$, i.e., Equation (5), and $\Omega m(z_i, z_j)$, i.e., Equation (2), one obtains, respectively,

$$\sigma_{\Omega m_h^2, ij}^2 = \frac{4(h^2(z_i)\sigma_{h,z_i}^2 + h^2(z_j)\sigma_{h,z_j}^2)}{(1 + z_i)^3 - (1 + z_j)^3} \left(\frac{1}{\Omega m(z_i, z_j)}\right)^2$$  \hspace{1cm} (8)

where $\sigma_{h,zi}$ denotes the uncertainty of the $i$th Hubble parameter measurement in units of (100 km s$^{-1}$ Mpc$^{-1}$), i.e.,

$$\sigma_{h,z_i} = 0.01 \sigma_{H(z_i)},$$

and

$$\sigma_{\Omega m, ij}^2 = \frac{4h^2(z_i)\sigma_{H(z_i)}^2 + 4h^2(z_j)\sigma_{H(z_j)}^2}{((1 + z_i)^3 - 1)^2 + (1 + z_j)^3 - 1)^2}$$  \hspace{1cm} (9)

where in this case, because of normalizing to the actual Hubble constant $H_0$ one has:

$$\sigma_{H(z_i)}^2 = \left(\frac{\sigma_{h(z_i)}}{H_0}\right)^2 + \left(\frac{H(z_i)\sigma_{H_0}}{H_0^2}\right)^2$$  \hspace{1cm} (10)

and $\sigma_{H(z_i)}$ denotes the uncertainty of the $i$th Hubble parameter in units of (km s$^{-1}$ Mpc$^{-1}$). Now one can see from Equation (8) that the uncertainty of $\Omega m_h^2(z_i, z_j)$ is large whenever the redshifts $z_i$ and $z_j$ are close to each other, whereas the uncertainty of $\Omega m(z_i, z_j)$ is large whenever one of the redshifts in the pair is close to zero.

Two-point diagnostics used as tests of the $\Lambda$CDM model are supposed to provide just a constant numerical value for this model, therefore one should first produce a summary statistics of their values calculated on the data sets. Because of the statistical properties discussed above, we used two approaches.
Table 3

Results of $\Omega m(z_i, z_j)$ and $\Omega mh^2(z_i, z_j)$ Two-point Diagnostics Calculated on Different Sub-samples Using the Weighted Mean and the Median Statistics

|                  | $|N_i|<1$ | $|N_i|<1$ |
|------------------|---------|---------|
|                  | $\Omega m(z_i, z_j)_{w.m.}$ | $\Omega m(z_i, z_j)_{m.s.}$ | $\Omega mh^2(z_i, z_j)_{w.m.}$ | $\Omega mh^2(z_i, z_j)_{m.s.}$ |
| Full sample ($n=36$) | $-0.0061 \pm 0.0111$ | $91.90\%$ | $-0.0199 \pm 0.0077$ | $92.22\%$ |
| $z = 2.34$ excluded ($n=35$) | $-0.0137 \pm 0.0123$ | $92.61\%$ | $-0.0257 \pm 0.0046$ | $92.61\%$ |
| DA only ($n=30$) | $-0.0019 \pm 0.0165$ | $92.87\%$ | $-0.0305 \pm 0.0077$ | $93.10\%$ |
| BAO only ($n=6$) | $0.0058 \pm 0.0351$ | $100\%$ | $0.0326 \pm 0.0063$ | $100\%$ |
| $\Omega m^2(z_i, z_j)_{w.m.}$ | $0.1259 \pm 0.0019$ | $83.49\%$ | $0.1501 \pm 0.0049$ | $79.37\%$ |
| $\Omega m^2(z_i, z_j)_{m.s.}$ | $0.1404 \pm 0.0040$ | $82.02\%$ | $0.1586 \pm 0.0029$ | $85.04\%$ |
| DA only ($n=30$) | $0.1437 \pm 0.0046$ | $81.61\%$ | $0.1729 \pm 0.0072$ | $87.82\%$ |
| BAO only ($n=6$) | $0.1231 \pm 0.0045$ | $100\%$ | $0.1218 \pm 0.0002$ | $100\%$ |

Note. For the $\Omega m(z_i, z_j)$ diagnostic the Hubble constant value of $H_0 = 67.4 \pm 1.4$ km s$^{-1}$ Mpc$^{-1}$ was assumed. The results of the $\Omega mh^2(z_i, z_j)$ diagnostic should be compared to the Planck result $\Omega_m h^2_{\text{Planck}} = 0.1426 \pm 0.0025$. The percentage of residual distribution falling within $|N_i|<1$ for the main sample and different sub-samples is shown as an indicator of non-Gaussianity.

Figure 3. The $\Omega m(z_i, z_j)$ two-point diagnostic displayed as the weighted mean (left panels) and as the median value (right panels) indicated by dashed lines surrounded by color bands denoting 68% confidence regions. The long solid line shows the $\Omega m(z_i, z_j) = 0$ level expected for the $\Lambda$CDM. The four panels correspond to the four respective sub-samples: $N = 6$ BAO data, $N = 30$ DA data, the $N = 35$ combined BAO+DA sample with the $H(z = 2.34)$ data point excluded, and the full $N = 36$ combined BAO+DA data.

Figure 4. Histograms of the $\Omega m(z_i, z_j)$ two-point diagnostic calculated with different samples: $N = 6$ BAO data, $N = 30$ DA data, the $N = 35$ combined BAO+DA sample with the $H(z = 2.34)$ data point excluded, and the full $N = 36$ combined BAO+DA data.

The first was to calculate the weighted mean, since this is the most popular way of summarizing measurements encountered in the literature, unfortunately sometimes without checking the validity of such an approach. The weighted mean formula for the $\Omega m(z_i, z_j)$ diagnostic reads:

$$\Omega m_{w.m.} = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} \Omega m(z_i, z_j) / \sigma_{\Omega m,ij}^2}{\sum_{i=1}^{n} \sum_{j=i+1}^{n} 1 / \sigma_{\Omega m,ij}^2}$$ (11)

and its variance is:

$$\sigma_{\Omega m, w.m.}^2 = \left( \sum_{i=1}^{n} \sum_{j=i+1}^{n} 1 / \sigma_{\Omega m,ij}^2 \right)^{-1}$$ (12)

with $\sigma_{\Omega m,ij}^2$ given by Equation (9). Similarly, the weighted mean formula for the $\Omega mh^2(z_i, z_j)$ diagnostic is:

$$\Omega mh^2_{w.m.} = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} \Omega mh^2(z_i, z_j) / \sigma_{\Omega mh^2,ij}^2}{\sum_{i=1}^{n} \sum_{j=i+1}^{n} 1 / \sigma_{\Omega mh^2,ij}^2}$$ (13)

and its variance is:

$$\sigma_{\Omega mh^2, w.m.}^2 = \left( \sum_{i=1}^{n} \sum_{j=i+1}^{n} 1 / \sigma_{\Omega mh^2,ij}^2 \right)^{-1}$$ (14)

with $\sigma_{\Omega mh^2,ij}^2$ given by Equation (8).

The second approach is the “median statistics” method, which was pioneered by Gott et al. (2001). It is based on a very well known property of the median which, being a non-
parametric measure, is robust and can be used without any prior assumption about the underlying distribution, in particular without assuming its Gaussianity. From the definition of the median, the probability that any particular measurement, one of the $N$ independent measurements, is higher than the true median is 50%. Consequently, the probability that the $n$th observation out of the total number $N$ is higher than the median follows the binomial distribution: $P = \frac{2^{-N}}{n!(N-n)!}$. This property allows us to calculate the 68% confidence intervals of the median.

The results for the $\Omega m(z_i, z_j)$ and $\Omega m^{h2}(z_i, z_j)$ diagnostics obtained from the full sample and its different sub-samples are listed in Table 3 and shown in Figures 3 and 5. The weighted mean approach is meaningful only under the assumption of statistical independence of the data, a lack of systematic effects, and a Gaussian distribution of errors. Hence, in order to test the Gaussianity of error distributions, we follow the approach of Chen et al. (2003), Crandall & Ratra (2014), and Crandall et al. (2015). Their idea was to construct an error distribution, a histogram of measurements as a function of $N_m$, the number of standard deviations that a measurement deviates from a central estimate. For example, $N_m$ for the $\Omega m(z_i, z_j)$ observable with respect to its weighted mean value would be: $N_m,k = (\Omega m(z_i, z_j) - \Omega m_{w(m.)})/\sigma_{\Omega m,i,j}$ where the $k$-index identifies the pair $(i,j)$. In a similar manner we calculate $N_m$ with respect to the median value $N_m,k = (\Omega m(z_i, z_j) - \Omega m_{w(m.)})/\sigma_{\Omega m,i,j}$. The percentage of measurements having $|N_m| < 1$ is a convenient measure of deviation from the Gaussian distribution, for which it should be equal to 68.3%. Therefore, in Table 3 (and also later in Tables 4 and 5) we report the corresponding percentage of the distribution falling within $\pm 1\sigma$, i.e., $|N_m| < 1$. One clearly sees that they strongly deviate from the Gaussian expectation. One can also see this non-Gaussian intuitively in Figures 4 and 6 where the histograms of the calculated $\Omega m(z_i, z_j)$ and $\Omega m^{h2}(z_i, z_j)$ are shown. We also performed the Kolmogorov–Smirnov test which strongly rejected the hypothesis of Gaussianity in each sub-sample (with $p$-values ranging from $10^{-4}$ to $10^{-7}$). Therefore we can conclude that the weighted average scheme is not appropriate here and the median statistics method is more reliable.

The results of $\Omega m(z_i, z_j)$ shown in Table 3 and in Figure 3 suggest that the weighted mean of this diagnostic is compatible with the $\Lambda$CDM irrespective of the sample used. However, as we argued, the weighted mean is not an appropriate measure in light of the non-Gaussian error distribution. On the other hand, at the level of the median statistics, the results are incompatible with the $\Lambda$CDM. This conclusion seems much more justified than the previous one drawn from the weighted mean. However, one can also see that the median of $\Omega m(z_i, z_j)$ from BAOs is positive and the median from DAs is negative. In other words, the BAO median statistics of the $\Omega m(z_i, z_j)$ two-point diagnostic suggest quintessence ($w > -1$) while the DA median suggests phantom behavior ($w < -1$). Of course, the combined data inherit the DA behavior because the median is

### Table 4

Results of $\Omega m(z_i, z_j)$ Two-point Diagnostics Residuals Calculated for the Three Cosmological Models, $\Lambda$CDM, wCDM, and CPL, on Different Sub-samples Using the Weighted Mean and the Median Statistics

| Sample/R$_{f(m.i)}$ | $R_{(m,i)}$ (ACDM) | $|N_m| < 1$ | $R_{(m,i)}$ (wCDM) | $|N_m| < 1$ | $R_{(m,i)}$ (CPL) | $|N_m| < 1$ |
|---------------------|-------------------|------------|-------------------|------------|-------------------|------------|
| Full sample (n = 36) | -0.0150 ± 0.0107 | 92.06%     | -0.2354 ± 0.0108 | 68.10%     | -0.1773 ± 0.0122 | 75.24%     |
| $z = 2.34$ excluded (n = 35) | -0.0226 ± 0.0118 | 92.27%     | -0.2550 ± 0.0119 | 69.41%     | -0.1956 ± 0.0133 | 76.30%     |
| DA only (n = 30) | -0.0124 ± 0.0159 | 93.10%     | -0.2803 ± 0.0160 | 68.74%     | -0.1994 ± 0.0179 | 74.94%     |
| BAO only (n = 6) | -0.0013 ± 0.0335 | 100%       | -0.1657 ± 0.0339 | 66.67%     | -0.1358 ± 0.0426 | 93.33%     |

### Table 5

Results of $\Omega m^{h2}(z_i, z_j)$ Two-point Diagnostics Residuals Calculated for the Three Cosmological Models, $\Lambda$CDM, wCDM, and CPL, On Different Sub-samples Using the Weighted Mean and the Median Statistics

| Sample/R$_{f(m.i)}$ | $R_{(m,i)}$ (ACDM) | $|N_m| < 1$ | $R_{(m,i)}$ (wCDM) | $|N_m| < 1$ | $R_{(m,i)}$ (CPL) | $|N_m| < 1$ |
|---------------------|-------------------|------------|-------------------|------------|-------------------|------------|
| Full sample (n = 36) | -0.0157 ± 0.0021 | 83.65%     | -0.0140 ± 0.0022 | 83.81%     | 0.1063 ± 0.0041 | 67.83%     |
| $z = 2.34$ excluded (n = 35) | -0.0016 ± 0.0040 | 82.52%     | 0.0006 ± 0.0040 | 82.52%     | 0.1268 ± 0.0047 | 76.64%     |
| DA only (n = 30) | 0.0018 ± 0.0046 | 81.84%     | 0.0039 ± 0.0047 | 82.30%     | 0.1270 ± 0.0055 | 75.40%     |
| BAO only (n = 6) | -0.0194 ± 0.0053 | 100%       | -0.0186 ± 0.0057 | 100%       | 0.0597 ± 0.0197 | 80%        |
| Full sample (n = 36) | 0.0076 ± 0.0049 | 79.37%     | 0.0099 ± 0.0056 | 80%        | 0.1654 ± 0.0045 | 70.79%     |
| $z = 2.34$ excluded (n = 35) | 0.0162 ± 0.0044 | 85.04%     | 0.0189 ± 0.0031 | 85.55%     | 0.1733 ± 0.0043 | 75.29%     |
| DA only (n = 30) | 0.0304 ± 0.0076 | 87.82%     | 0.0335 ± 0.0070 | 88.51%     | 0.1733 ± 0.0045 | 74.48%     |
| BAO only (n = 6) | -0.0207 ± 0.0011 | 100%       | -0.0196 ± 0.0003 | 100%       | 0.1256 ± 0.0017 | 60%        |

Note. The percentage of residual distribution falling within $|N_m| < 1$ for the main sample and different sub-samples is shown as an indicator of non-Gaussianity.
Figure 5. The $Omh^2(z, \theta)$ two-point diagnostic displayed as the weighted mean (left panels) and as the median value (right panels) indicated by dashed lines surrounded by color bands denoting 68% confidence regions. The middle panels show the Planck result of $\Omega_m h^2_{\text{Planck}} = 0.1426 \pm 0.0025$ — this is the value expected for $Omh^2(z, \theta)$ within $\Lambda$CDM model. The four panels correspond to the four respective sub-samples: $N = 6$ BAO data, $N = 30$ DA data, the $N = 35$ combined BAO+DA sample with the $H(z = 2.34)$ data point excluded, and the full $N = 36$ combined BAO+DA data.

Robust against “outliers” (here, the less numerous BAO sample). One should treat these diverging conclusions as an indication of a systematic difference between BAO and DA data concerning $H(z)$ measurements.

Table 3 and Figure 5 also show the results of the $Omh^2(z, \theta)$ diagnostics. Here one can clearly see incompatibility with the $\Lambda$CDM when the $\Omega_m h^2$ value suggested by Planck is taken as a reference. The DA and BAO+DA combined data with the $H(z = 2.34)$ data point excluded are compatible with the $\Lambda$CDM for the weighted mean, but our previous comments raising doubts about the appropriateness of this approach are valid here as well. One can also see the difference between BAO and DA: the $Omh^2(z, \theta)$ inferred from BAOs is lower and the one inferred from DAs is higher than the reference value. So we can conclude that even though there are systematic differences between the BAO and DA values, both these data sets of $H(z)$ measurements are not consistent with the $\Lambda$CDM.

Therefore, we can ask if some other parameterization of dark energy can perform better. In particular, we consider the simplest extensions of the $\Lambda$CDM, i.e., the wCDM and CPL parameterizations. In these models the expected values of the two-point diagnostics are no longer constant, but rather the functions of redshifts given by Equations (3), (4), (6), (7), hence we have evaluated the theoretically expected values $Om(z, \theta)_{\text{wCDM}}$ and $Om(z, \theta)_{\text{CPL}}$ (i.e., the right-hand sides of the respective equations) assuming the cosmological parameters reported in Table 2, and then we calculated the residuals $R_{\text{w.m.}}(z, \theta) = Om(z, \theta) - Om(z, \theta)_{\text{obs}}$ (similarly $R_{\text{omh^2}}(z, \theta)$ for the second two-point diagnostic). In principle the residuals should be zero (or rather compatible with zero in a statistical sense). If for a given model they deviate from zero more than for the $\Lambda$CDM model, it means that this model is less supported by $H(z)$ data in terms of two-point diagnostics. We have summarized the residuals as the weighted mean:

$$R_{\text{w.m.}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} R(z_i, \theta_j) / \sigma_{R_{ij}}^2}{\sum_{i=1}^{n} \sum_{j=1}^{n} 1 / \sigma_{R_{ij}}^2}$$

with the variance

$$\sigma_{R_{\text{w.m.}}}^2 = \left( \sum_{i=1}^{n} \sum_{j=1}^{n} 1 / \sigma_{R_{ij}}^2 \right)^{-1}$$

and the median. The results are listed in Tables 4 and 5 and are shown in Figures 7 and 9. Figures 8 and 10 display the histograms of the residuals. Let us recall that the cosmological model parameters used for calculating the theoretically expected diagnostics were taken after the JLA study (Betoule et al. 2014) as indicated in Table 2. Therefore, here the expected value of $Omh^2(z, \theta)$ in the $\Lambda$CDM model was not $\Omega_m h^2$ after Planck (Ade et al. 2014), but the respective value...
The two-point diagnostics $O_m(z_i, z_j)$ and $Omh^2(z_i, z_j)$ have been introduced as an interesting tool for testing the validity of the $\Lambda$CDM model. Reliable data concerning expansion rates of the universe at different redshifts $H(z)$ are crucial for their successful application. Currently we are at a moment in time when fairly reliable data of this kind are being obtained from the DA and BAO techniques. Therefore, in this paper we examined both diagnostics on the comprehensive set comprising data compiled in Ding et al. (2015) supplemented by the most recent DA measurements by Moresco (2015) and Moresco et al. (2016). An important motivation for this study was the paper by Sahni et al. (2014) where, based on three $H(z)$ measurements from BAOs (including the $z = 2.34$ measurement by Delubac et al. 2015) they claimed that recent precise

suggested by the Table 2. Similarly, $O_m(z_i, z_j)$, which is expected to vanish in the $\Lambda$CDM model, was calculated with the $H_0$ suggested by the JLA study, not by Planck.

One can see from Figure 7 that the residuals $R_{O_m}(z_i, z_j)$ for the $\Lambda$CDM model are closer to zero than for the wCDM or CPL models, irrespective of the sample. From Table 4 one can see that the residuals of the wCDM or CPL models summarized in the weighted mean scheme are more than 15σ away from the expected value of zero (for the full sample). This deviation in terms of median statistics is even bigger. One can also see this clearly in Figure 8 where the histograms of $R_{O_m}(z_i, z_j)$ are shown. In the case of the $Omh^2(z_i, z_j)$ diagnostics the performance of the $\Lambda$CDM and wCDM models is similar: using the full sample, the weighted mean of $R_{Omh^2}(z_i, z_j)$ residuals is at 7σ away from zero. As shown in Figure 10, the bulk of the $R_{Omh^2}(z_i, z_j)$ distributions for the $\Lambda$CDM or wCDM models contain zero in their tails, while the distribution for the CPL model is considerably away from zero (it corresponds to 26σ for the weighted mean). It seems that despite the presence of some tension between the $\Lambda$CDM model and the two-point diagnostics evaluated on the most recent $H(z)$ data, as noted, e.g., in Sahni et al. (2014) and Ding et al. (2015), this model still performs better than its immediate extensions, wCDM or CPL, in particular the latter. It should be stressed that the above mentioned performance of different models refers only to the two-point diagnostics considered. Therefore, this cannot be treated as a decisive ranking of competing models. A major obstacle for using two-point diagnostics for cosmological models other than $\Lambda$CDM is that in such cases it ceases to be such a strong “screening test”, because its expected value is no longer a number, but a function of redshift involving cosmological model parameters, which should somehow be assessed prior to using this test.

5. CONCLUSIONS

The two-point diagnostics $O_m(z_i, z_j)$ and $Omh^2(z_i, z_j)$ have been introduced as an interesting tool for testing the validity of the $\Lambda$CDM model. Reliable data concerning expansion rates of the universe at different redshifts $H(z)$ are crucial for their successful application. Currently we are at a moment in time when fairly reliable data of this kind are being obtained from the DA and BAO techniques. Therefore, in this paper we examined both diagnostics on the comprehensive set comprising data compiled in Ding et al. (2015) supplemented by the most recent DA measurements by Moresco (2015) and Moresco et al. (2016). An important motivation for this study was the paper by Sahni et al. (2014) where, based on three $H(z)$ measurements from BAOs (including the $z = 2.34$ measurement by Delubac et al. 2015) they claimed that recent precise
measurements of expansion rates at different redshifts suggest a severe tension with the $\Lambda$CDM model. Our study (Ding et al. 2015) confirmed this claim, however this was based only on one particular two-point diagnostic $Omh^2(z_i, z_j)$ which is expected to be equal to $\Omega_m h^2$ in the $\Lambda$CDM model. In this paper we not only used a larger data set—enriched by the most recent DA data—but we also considered the $Om(z_i, z_j)$ two-point diagnostic which is expected to be zero in the $\Lambda$CDM. Therefore this diagnostic does not depend on our knowledge of the matter density parameter and the uncertainty about its value does not propagate into the inference. Being aware that the BAO and DA techniques are prone to different systematic uncertainties, and because of the large effect of the $z = 2.34$ data point, we have analyzed not only full combined sample of $N=36$ BAO+DA data, but also different sub-samples. It turned out that both two-point diagnostics have non-Gaussian distributions and therefore the median statistic is a more appropriate way to describe them than the weighted mean scheme. The median statistic results support the claim that the $H(z)$ data seem to be in conflict with the $\Lambda$CDM model. However, the two-point diagnostics evaluated on the BAO and DA data deviate in different directions from the expectations for the $\Lambda$CDM model. This indicates that there are serious systematic effects in these two approaches. The DA method is very simple and transparent in its design. The major source of systematics is the adopted population synthesis model which quantifies the relation between $D4000$ spectral break, metallic- city, star formation history, and the age of the galaxy (Moresco et al. 2016). In contrast, in spite of its huge statistical power, the BAO technique is much more complex. In order to derive $H(z)$ from the large-scale clustering patterns of galaxies, one has to not only determine the baryon acoustic peak in the angle-averaged clustering pattern, but also measure the Alcock–Paczynski effect from the two-point statistics of galaxy clustering. This requires a good understanding of redshift–space distortions and sophisticated statistical methods. This suggests that the BAO data on $H(z)$ should be treated with caution when used for constraining cosmological models, much more than in the case of using the more directly observable “dilation scale” distance $D_i(z_i)$.

We have also asked whether other cosmological models, alternatives to the $\Lambda$CDM, perform better. In particular we considered the $w$CDM and CPL models. However, the diagnostic test was not so simple: we had to compare $Omh(z_i, z_j)$ and $Om(z_i, z_j)$ diagnostics calculated from $H(z)$ data against theoretically expected (redshift dependent) counterparts. We performed this calculation in the “observed–expected” residuals. It turned out that despite the revealed mismatch between the data and the $\Lambda$CDM, this model is still in better agreement with the data than the wCDM or CPL models. There is one caveat in our approach, namely that in order to evaluate the theoretically expected counterparts of the two-point diagnostics, we have taken cosmological parameters best fitted by the joint JLA study (Betoule et al. 2014) as a reference point. It would be tempting and more consistent to perform the fit of cosmological parameters based on the two-point diagnostics. However, because of the error distribution revealed in Figure 2 it would not give results that are competitive with other techniques. On the other hand, since the systematics underlying this peculiar behavior of uncertainties has been partly recognized, it could be used to define and select suitable sub-samples better suited to cosmological inference.

This work was supported by the Ministry of Science and Technology National Basic Science Program (Project 973) under grant Nos. 2012CB821804 and 2014CB845806, the Strategic Priority Research Program “The Emergence of Cosmological Structure” of the Chinese Academy of Sciences (No. XDB09000000), the National Natural Science Foundation of China under grant Nos. 1153001, 11373014, and 11073005. S.C. also acknowledges support from the China Postdoctoral Science Foundation under grant No. 2015T80052. M.B. obtained the approval of the Foreign Talent Introducing Project in China and gained a special fund in support of the Foreign Knowledge Introducing Project. He also gratefully acknowledges the hospitality of Beijing Normal University.

REFERENCES

Ade, P. A. R., et al. [Planck Collaboration] 2014, A&A, 571, A16
Anderson, L., Aubourg, E., Bailey, S., et al. 2014, MNRAS, 439, 83
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Blake, C., Brough, S., Colless, M., et al. 2012, MNRAS, 425, 405
Chen, G., Gott, J. J., III, & Ratra, B. 2003, PASP, 115, 1269
Chevallier, M., & Polarski, D. 2001, JIMP, D, 10, 213
Chuang, C. H., & Wang, Y. 2013, MNRAS, 435, 255
Crandall, S., Houston, S., & Ratra, B. 2015, MPLA, 30, 1550123
Crandall, S., & Ratra, B. 2014, PSBL, 732, 320
Delubac, T., Bautista, J. E., Busca, N. G., et al. 2015, A&A, 574, A59
Ding, X., Biesiada, M., Cao, S., Li, Z. X., & Zhu, Z. H. 2015, ApJL, 803, L22
Farooq, O., & Ratra, B. 2013, ApJL, 766, L7
Gott, J. R., III, Vogley, M. S., Podariv, S., & Ratra, B. 2001, ApJ, 549, 1
Jimenez, R., & Loeb, A. 2002, ApJ, 573, 37
Jimenez, R., Verde, L., Treu, T., & Stern, D. 2003, ApJ, 593, 622
Linder, E. V. 2003, PhRvL, 90, 091301
Macaulay, E., Wehus, I. K., & Eriksen, H. K. 2013, PhRvL, 111, 161301
Moresco, M. 2015, MNRAS, 450, L16
Moresco, M., Cimatti, A., Jimenez, R., et al. 2012, JCAP, 08, 006
Moresco, M., Pozzetti, L., Cimatti, A., et al. 2016, arXiv:1601.01701
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Raveri, M. 2016, PhRvD, 93, 043522
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009
Riess, A. G., Macri, L., & Csetsato, S. 2011, ApJ, 730, 119
Sahni, V., Shaferleo, A., & Starobinsky, A. A. 2008, PhRvD, 78, 103502
Sahni, V., Shaferleo, A., & Starobinsky, A. A. 2014, ApJL, 793, L40
Samushia, L., Reid, B. A., White, M., et al. 2013, MNRAS, 429, 1514
Shaferleo, A., Sahni, V., & Starobinsky, A. 2012, PhRvD, 86, 103527
Simon, J. Verde, L., & Jimenez, R. 2005, PhRvD, 71, 123001
Stern, D., Jimenez, R., Verde, L., et al. 2010, JCAP, 02, 008
Uzan, J.-P., Clarkson, Ch., & Ellis, G. F. R. 2008, PhRvL, 100, 191303
Valkenberg, W., Marra, V., & Clarkson, C. 2014, MNRAS, 438, L6
Zhang, C., Zhang, H., Yuan, S., et al. 2014, RAA, 14, 1221
Zunckel, C., & Clarkson, C. 2008, PhRvL, 101, 181301