Cosmological tests using the angular size of galaxy clusters

Jun-Jie Wei,1,2★ Xue-Feng Wu1,3,4★ and Fulvio Melia1,5★†

1Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China
2University of Chinese Academy of Sciences, Beijing 100049, China
3Chinese Center for Antarctic Astronomy, Nanjing 210008, China
4Joint Center for Particle, Nuclear Physics and Cosmology, Nanjing University-Purple Mountain Observatory, Nanjing 210008, China
5Department of Physics, The Applied Math Program, and Department of Astronomy, The University of Arizona, Tucson, AZ 85721, USA

Accepted 2014 November 21. Received 2014 November 20; in original form 2014 September 1

ABSTRACT
We use measurements of the galaxy-cluster angular size versus redshift to test and compare the standard model (ΛCDM) and the $R_h = ct$ Universe. We show that the latter fits the data with a reduced $\chi^2_{dof} = 0.786$ for a Hubble constant $H_0 = 72.6^{+13.8}_{-3.4}$ km s$^{-1}$ Mpc$^{-1}$, and $H_0$ is the sole parameter in this model. By comparison, the optimal flat $\Lambda$ cold dark matter (ΛCDM) model, with two free parameters (including $\Omega_m = 0.50$ and $H_0 = 73.9^{+10.5}_{-9.5}$ km s$^{-1}$ Mpc$^{-1}$), fits the angular-size data with a reduced $\chi^2_{dof} = 0.806$. On the basis of their $\chi^2_{dof}$ values alone, both models appear to account for the data very well in spite of the fact that the $R_h = ct$ Universe expands at a constant rate, while ΛCDM does not. However, because of the different number of free parameters in these models, selection tools, such as the Bayes Information Criterion, favour $R_h = ct$ over ΛCDM with a likelihood of $\sim 86$ per cent versus $\sim 14$ per cent. These results impact the question of galaxy growth at large redshifts. Previous work suggested an inconsistency with the underlying cosmological model unless elliptical and disc galaxies grew in size by a surprisingly large factor $\sim 6$ from $z \sim 3$ to 0. The fact that both ΛCDM and $R_h = ct$ fit the cluster-size measurements quite well casts some doubt on the suggestion that the unexpected result with individual galaxies may be due to the use of an incorrect expansion scenario, rather than astrophysical causes, such as mergers and/or selection effects.

Key words: galaxies: clusters: general – galaxies: evolution – cosmological parameters – cosmology: theory – distance scale – large-scale structure of Universe.

1 INTRODUCTION
The use of standard rods – distance scales with no evolution in linear size over the lifetime of the Universe – to carry out geometric tests of cosmological models was first proposed by Hoyle (1959), but the application of this idea to real data took a long time, specifically because of the difficulty in finding suitable objects or structures for this purpose. The earliest tests of cosmological models using the observed dependence of the angular size of galaxies or kpc-scale radio sources could not easily define a true metric rod and were subject to unknown evolutionary effects (e.g. Sandage 1988). There was uncertainty about whether the observed size–redshift relation in radio galaxies was an indication of an actual evolution in size (Kapahi 1987; Barthel & Miley 1988; Neeser et al. 1995), or whether it was due to selection effects (Nilsson et al. 1993; Singal 1993). This was partially addressed in a subsequent study using an enlarged sample of double-lobed quasars at $z > 0.3$ (Buchalter et al. 1998), which showed no change in apparent angular size within the range $1.0 < z < 2.7$, consistent with standard cosmology without significant evolution.

Separate investigations specifically to study the cosmic deceleration (or acceleration) were carried out by Gurvits (1993, 1994, using VLBI visibility data obtained at 13 cm by Preston et al. 1985). This analysis also provided estimates of the dependence of the apparent angular size of compact sources on their luminosity and rest-frame frequency. Similar studies were also carried out by Kellermann (1993) and Wilkinson et al. (1998). Within a few years, a much enlarged sample of 330 compact radio sources distributed over a broad range of redshifts $0.011 < z < 4.72$ started to demonstrate that the angular size–redshift relation for compact radio sources is consistent with the predictions of standard Friedmann–Robertson–Walker (FRW) models without the need to consider evolutionary or selection effects (Gurvits, Kellermann & Frey 1999).

Other groups carried out their own cosmological tests using powerful radio-lobed radio galaxies, presumed to be reasonable standard yardsticks to determine global cosmological parameters (Daly 1994, 1995). The method was applied and discussed by Guerra & Daly (1996, 1998), Guerra (1997), and Daly, Guerra & Wan (1998, 1999),...
who reported that the data at that time strongly favoured a low (matter) density Universe. Vishwakarma (2001) and Lima & Alcaniz (2002) used the Gurvits et al. (1999) compact radio source angular size versus redshift data to set constraints on cosmological parameters. This analysis was extended by Chen & Ratra (2003), who used these data to place constraints on cosmological model parameters for a variety of cosmological constant scenarios. And a more focused study using Fanaroff–Riley type IIb radio galaxy redshift–angular size data to derive constraints on the parameters of a spatially flat cosmology with a dark-energy scalar field was carried out by Podariu et al. (2003).

More recently, in one of the better known studies using this method, based on the average linear size of galaxies with the same luminosity (see e.g. Barden et al. 2005; McIntosh et al. 2005; Trujillo et al. 2006; López-Corredoira 2010), the measurements do not appear to be consistent with an expanding cosmology, unless galaxies have grown in size by a surprisingly large factor 6 from redshift $z = 3.2$ to 0.

Perhaps the cosmology itself is wrong or, more simply, there is still some ambiguity concerning the use of galactic size as a standard rod. Indeed, if another more reliable scale could be found, and shown to be consistent with, say, the standard model (ΛCDM), the contrast between this result, and the disparity emerging through the use of galaxies, could be useful in affirming the need for stronger evolution in galactic growth than is predicted by current theory.

Our focus in this paper will be galaxy clusters, which can also be used as standard rulers under appropriate conditions. These are the largest gravitationally collapsed structures in the Universe, with a hot diffuse plasma ($T_e \sim 10^7$–$10^8$ K) filling the intergalactic medium. Their angular size versus redshift can be measured using a combination of the Sunyaev–Zel’dovich effect (SZE) and X-ray surface brightness observations. The SZE is the result of high-energy electrons distorting the cosmic microwave background radiation (CMB) through inverse Compton scattering, during which low-energy CMB photons (on average) receive an energy boost from the high-energy electrons in the cluster (Sunyaev & Zel’dovich 1970, 1972). The same hot gas emits X-rays primarily via thermal bremsstrahlung. While the SZE is a function of the electron number density $n_e$ and temperature $T_e$ along the line of sight, the X-ray emission scales as $S_X \propto \int n_e^2 T_e \, dl$ (in terms of the electron cooling function $N_{e}$). This dependence on density, along with a suitable model for the cluster gas, enables a direct distance determination to the galaxy cluster. This method is independent of the extragalactic distance ladder and provides distances to high-redshift galaxy clusters.

SZE/X-ray distances have been previously used to constrain some cosmological parameters and to test the distance duality (DD) relation per se but, rather, to address two related issues. First, we will use the newer and larger sample of galaxy-cluster angular size versus redshift measurements from Bonamente et al. (2006) to constrain cosmological models. In particular, we wish to see if the aforementioned tension between galaxy growth and the conventional expansion scenario is supported by the cluster data, or whether the latter confirm the basic theoretical predictions, thus reinforcing the need for a stronger galactic evolution at high redshifts. Secondly, we wish to use this relatively new probe of the Universe’s expansion to directly test the $R_0 = ct$ Universe (Melia & Shevchuk 2012) against the data and to see how its predictions compare with those of a cold dark matter (ΛCDM).

The outline of this paper is as follows. In Section 2, we will briefly summarize the galaxy-cluster angular-size sample at our disposal. We will present theoretical fits to the data in Section 3, and constrain the cosmological parameters – both in the context of ΛCDM and the $R_0 = ct$ Universe – in Section 4. We will end with a discussion and conclusion in Section 5.

2 THE CLUSTER ANGULAR-SIZE SAMPLE

In addition to the luminosity distance, $D_L$, which is necessary for measurements involving standard candles such as Type Ia supernovae (Type Ia SNe), the angular-diameter distance (ADD), $D_A$, is also used in astronomy for objects whose diameter (i.e. the standard ruler) is known. The luminosity distance and ADD can be measured independently using different celestial objects, but they are related via Etherington’s reciprocity relation:

$$\frac{D_L}{D_A} (1 + z)^{-2} = 1. \quad (1)$$

This relation, sometimes referred as the DD relation, is completely general and is valid for all cosmological models based on Riemannian geometry. That is, its validity is independent of Einstein’s field equations for gravity and the nature of the matter energy content of the universe. It requires only that the source and observer be connected via null geodesics in a Riemannian space–time and that the number of photons be conserved.

X-ray observations of the intracluster medium, combined with radio observations of the galaxy cluster’s SZE, allow an estimate of the ADD to be made. Recently, Bonamente et al. (2006) determined the distance to 38 clusters of galaxies in the redshift range $0.14 \leq z \leq 0.89$ using X-ray data from Chandra and SZE data from the Owens Valley Radio Observatory and the Berkeley-Illinois-Maryland Association interferometric arrays. The data shown in Table 1 are reproduced from the compilation of Bonamente et al. (2006).

3 THEORETICAL FITS

The theoretical ADD $D_A$ is a function of the cluster’s redshift $z$, and is different for different cosmological models. Both ΛCDM and $R_0 = ct$ are FRW cosmologies, but the latter includes the additional constraint $p = -\rho/3$ on the overall equation of state. The densities are often written in terms of today’s critical density, $\rho_c \equiv 3c^2 H_0^2/8\pi G$, represented as $\Omega_m \equiv \rho_m/\rho_c$, $\Omega_\Lambda \equiv \rho_\Lambda/\rho_c$, and $\Omega_{de} \equiv \rho_{de}/\rho_c$. $H_0$ is the Hubble constant, and $\Omega_\Lambda$ is simply $\Omega_\Lambda$ when dark energy is assumed to be a cosmological constant.

3.1 ΛCDM

In a flat ΛCDM Universe with zero spatial curvature, the total scaled energy density is $\Omega \equiv \Omega_m + \Omega_\Lambda + \Omega_{de} = 1$. When dark energy is included with an unknown equation of state, $\rho_{de} = \omega_{de\rho_{de}}$, the most general form of the ADD is given by the expression

$$D^\Lambda_{\text{CDM}}(z) = \frac{1}{H_0} \frac{c}{|\Omega_\Lambda|^{1/2}} \sin n \left\{ \frac{\Omega_\Lambda}{|\Omega_\Lambda|^{1/2}} \right\} \times \int_0^z \frac{dz}{\sqrt{\Omega_m(1 + z)^3 + \Omega_\Lambda(1 + z)^2 + \Omega_{de}(1 + z)^{3(1 + \omega_{de})}}} . \quad (2)$$
Table 1. ADD of galaxy clusters.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>$z$</th>
<th>$D_A$(Mpc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abell 1413</td>
<td>0.142</td>
<td>780$^{+130}_{-130}$</td>
</tr>
<tr>
<td>Abell 2204</td>
<td>0.152</td>
<td>610$^{+50}_{-50}$</td>
</tr>
<tr>
<td>Abell 2259</td>
<td>0.164</td>
<td>580$^{+250}_{-250}$</td>
</tr>
<tr>
<td>Abell 586</td>
<td>0.171</td>
<td>520$^{+150}_{-120}$</td>
</tr>
<tr>
<td>Abell 1914</td>
<td>0.171</td>
<td>440$^{+40}_{-40}$</td>
</tr>
<tr>
<td>Abell 2218</td>
<td>0.176</td>
<td>660$^{+140}_{-110}$</td>
</tr>
<tr>
<td>Abell 665</td>
<td>0.182</td>
<td>660$^{+100}_{-100}$</td>
</tr>
<tr>
<td>Abell 1689</td>
<td>0.183</td>
<td>650$^{+90}_{-90}$</td>
</tr>
<tr>
<td>Abell 2163</td>
<td>0.202</td>
<td>520$^{+40}_{-40}$</td>
</tr>
<tr>
<td>Abell 773</td>
<td>0.217</td>
<td>980$^{+170}_{-140}$</td>
</tr>
<tr>
<td>Abell 2261</td>
<td>0.224</td>
<td>730$^{+130}_{-130}$</td>
</tr>
<tr>
<td>Abell 2111</td>
<td>0.229</td>
<td>640$^{+200}_{-200}$</td>
</tr>
<tr>
<td>Abell 267</td>
<td>0.23</td>
<td>600$^{+110}_{-90}$</td>
</tr>
<tr>
<td>RX J2129.7+10005</td>
<td>0.235</td>
<td>460$^{+110}_{-80}$</td>
</tr>
<tr>
<td>Abell 1835</td>
<td>0.252</td>
<td>1070$^{+20}_{-20}$</td>
</tr>
<tr>
<td>Abell 68</td>
<td>0.255</td>
<td>630$^{+100}_{-100}$</td>
</tr>
<tr>
<td>Abell 697</td>
<td>0.282</td>
<td>880$^{+230}_{-230}$</td>
</tr>
<tr>
<td>Abell 611</td>
<td>0.288</td>
<td>780$^{+180}_{-180}$</td>
</tr>
<tr>
<td>ZW 3146</td>
<td>0.291</td>
<td>830$^{+20}_{-20}$</td>
</tr>
<tr>
<td>Abell 1995</td>
<td>0.322</td>
<td>1190$^{+140}_{-140}$</td>
</tr>
<tr>
<td>MS 1358.4+6245</td>
<td>0.327</td>
<td>1130$^{+90}_{-90}$</td>
</tr>
<tr>
<td>Abell 370</td>
<td>0.375</td>
<td>1080$^{+100}_{-200}$</td>
</tr>
<tr>
<td>MACS J2228.5+2036</td>
<td>0.412</td>
<td>1220$^{+240}_{-240}$</td>
</tr>
<tr>
<td>RX J1347.5−1145</td>
<td>0.451</td>
<td>960$^{+60}_{-60}$</td>
</tr>
<tr>
<td>MACS J2214.9−1359</td>
<td>0.483</td>
<td>1440$^{+270}_{-270}$</td>
</tr>
<tr>
<td>MACS J1311.0−0310</td>
<td>0.49</td>
<td>1380$^{+370}_{-370}$</td>
</tr>
<tr>
<td>CL 0016+1609</td>
<td>0.541</td>
<td>1380$^{+220}_{-220}$</td>
</tr>
<tr>
<td>MACS J1149.5+2223</td>
<td>0.544</td>
<td>1380$^{+160}_{-160}$</td>
</tr>
<tr>
<td>MACS J1423.8+2404</td>
<td>0.545</td>
<td>1490$^{+60}_{-60}$</td>
</tr>
<tr>
<td>MS 0451.6−0305</td>
<td>0.55</td>
<td>1420$^{+260}_{-230}$</td>
</tr>
<tr>
<td>MACS J2129.4−0741</td>
<td>0.57</td>
<td>1330$^{+370}_{-370}$</td>
</tr>
<tr>
<td>MS 2053.7−0449</td>
<td>0.583</td>
<td>2480$^{+410}_{-410}$</td>
</tr>
<tr>
<td>MACS J0647.7+7015</td>
<td>0.584</td>
<td>770$^{+210}_{-180}$</td>
</tr>
<tr>
<td>MACS J0744.8+3927</td>
<td>0.686</td>
<td>1680$^{+380}_{-380}$</td>
</tr>
<tr>
<td>MS 1137.5+6625</td>
<td>0.784</td>
<td>2850$^{+520}_{-520}$</td>
</tr>
<tr>
<td>RX J1716.4+6708</td>
<td>0.813</td>
<td>1040$^{+510}_{-430}$</td>
</tr>
<tr>
<td>MS 1054.5−0321</td>
<td>0.826</td>
<td>1330$^{+260}_{-260}$</td>
</tr>
<tr>
<td>CL J1226.9+3332</td>
<td>0.89</td>
<td>1080$^{+280}_{-280}$</td>
</tr>
</tbody>
</table>

(2) wCDM – a flat universe with a constant dark-energy equation of state, but with a $w_{de}$ that is not necessarily equal to $-1$. Here $\Omega_m = 0$, so the free parameters may be chosen from the following: $H_0$, $\Omega_{\Lambda}$, and $w_{de}$.

(3) For comparison, we also consider a model with $\Omega_m = 1.0$, i.e. the Einstein-de Sitter (E-deS) cosmology.

### 3.2 The $R_0 = ct$ Universe

In the $R_0 = ct$ Universe, the ADD is given by the much simpler expression

$$D_A^{R_0=ct}(z) = \frac{c}{H_0} \ln(1+z).$$

The factor $c/H_0$ is in fact the gravitational horizon $R_0(t_0)$ at the present time, so we may also write the ADD as

$$D_A^{R_0=ct}(z) = \frac{R_0(t_0)}{H_0} \frac{\ln(1+z)}{1+z}.$$  

A detailed account of the differences between $\Lambda$CDM and $R_0 = ct$ is provided in Melia & Shevchuk (2012), Melia & Maier (2013), and Wei, Wu & Melia (2013). A more pedagogical description may also be found in Melia (2012). An important distinction between these two models is that whereas the $R_0 = ct$ Universe expands at a constant rate, $\Lambda$CDM predicts an early phase of deceleration, followed by a current acceleration. Therefore, an examination of the cluster angular-size data, spanning the redshift range $(-0.9 < z < 0.9)$ within which the transition from deceleration to acceleration is thought to have occurred, could in principle help to distinguish between these two cosmologies.

Briefly, the $R_0 = ct$ Universe is a FRW cosmology that has much in common with $\Lambda$CDM, but includes an additional ingredient motivated by several theoretical and observational arguments (Melia 2007; Melia & Abdelqader 2009; Melia & Shevchuk 2012; Melia 2013). Like $\Lambda$CDM, it adopts an equation of state $p = w\rho$, with $p = \rho_m + \rho_i + \rho_{de}$ and $\rho = \rho_m + \rho_i + \rho_{de}$, but goes one step further by specifying that $w = (\rho_i/3 + w_{de}\rho_{de})/\rho = -1/3$ at all times. Here, $\rho$ is the pressure and $\rho$ is the energy density, and subscripts m, i, and de refer to radiation, matter, and dark energy, respectively.

One might come away with the impression that this equation of state cannot be consistent with that (i.e. $w = (\rho_i/3 - \rho_{de})/\rho$) in the standard model. But in fact nature is telling us that if we ignore the constraint $w = -1/3$ and instead proceed to optimize the parameters in $\Lambda$CDM by fitting data, the resultant value of $w$ averaged over a Hubble time is actually $\approx -1/3$ within the measurement errors (Melia 2007; Melia & Shevchuk 2012). In other words, though $w = (\rho_i/3 - \rho_{de})/\rho$ in $\Lambda$CDM cannot be equal to $-1/3$ from one moment to the next, its value averaged over the age of the Universe is equal to what it would have been in $R_0 = ct$.

In terms of the expansion dynamics, $\Lambda$CDM must guess the constituents of the Universe and their individual equations of state, and then predict the expansion rate as a function of time. In contrast, $R_0 = ct$ acknowledges the fact that no matter what these constituents are, the total energy density in the Universe gives rise to a gravitational horizon coincident with the better known Hubble radius. But because this radius is therefore a proper distance, the application of Weyl's postulate forces it to always equal $ct$. Thus, on every time slice, the energy density $\rho$ must partition itself among its various constituents ($\rho_m$, $\rho_i$, and $\rho_{de}$) in such a way as to always adhere to this constraint, which also guarantees that the expansion rate be constant in time.
3.3 Optimization of the parameters

For each model, the best fit is obtained by minimizing the function

$$
\chi^2_{\text{ADD}} = \sum_{i=1}^{38} \frac{(D^i_{\text{obs}}(z_i) - D^i_{\text{th}}(z_i))^2}{\sigma_{\text{tot}}^2(z_i)},
$$

where $\xi$ stands for all the cosmological parameters that define the fitted model, $z_i$ is the redshift of the observed galaxy cluster, $D^i_{\text{obs}}$ is the predicted value of the ADD in the cosmological model under consideration, and $D^i_{\text{th}}$ is the measured value. There are three sources of uncertainty in the measurement of $D_A$: the cluster modelling error $\sigma_{\text{mod}}$, the statistical error $\sigma_{\text{stat}}$, and the systematic error $\sigma_{\text{sys}}$. The modelling errors are shown in Table 1 and the statistical and systematic errors are presented in Table 3 of Bonamente et al. (2006). In our analysis, we combine these errors in quadrature. Thus, the total uncertainty $\sigma_{\text{tot}}$ is given by the expression

$$
\sigma_{\text{tot}} = \sigma_{\text{mod}}^2 + \sigma_{\text{stat}}^2 + \sigma_{\text{sys}}^2.
$$

4 RESULTS

Though the number of free parameters characterizing ΛCDM can be as large as six or seven, depending on the application, here we take the minimalist approach and use the four most essential ones: the Hubble constant $H_0$, the matter energy density $\Omega_m$, the dark-energy density $\Omega_{de}$, and the dark-energy equation of state parameter $w_{de}$. By comparison, the $R_0 = ct$ Universe has only one free parameter—the Hubble constant $H_0$.

In this section, we optimize the fit for several dark-energy models (which we call ΛCDM and $w$CDM), one without dark energy (the E-deS Universe), and for the $R_0 = ct$ Universe. The outcome for each model is described and discussed in subsequent subsections.

4.1 ΛCDM

In ΛCDM, the dark-energy equation of state parameter, $w_{de} = w$, is exactly −1. Assuming a flat universe, $\Omega_{de} = 1 - \Omega_m$, there are only two free parameters: $\Omega_m$ and $H_0$. Type Ia SN measurements (see e.g. Garnavich et al. 1998; Perlmutter et al. 1998, 1999; Riess et al. 1998; Schmidt et al. 1998), CMB anisotropy data (see e.g. Ratra et al. 1999; Podariu et al. 2001; Spergel et al. 2003; Komatsu et al. 2009, 2011; Hinshaw et al. 2013), and baryon acoustic oscillation peak length-scale estimates (see e.g. Percival et al. 2007; Gaztañaga, Cabrè & Hui 2009; Samushia & Ratra 2009), strongly suggest that we live in a spatially flat, dark-energy-dominated universe with concordance parameter values $\Omega_m \approx 0.3$ and $\Omega_{de} \approx 0.7$. We will first adopt this concordance model, though to improve the fit we keep $H_0$ as a free parameter. The 38 cluster distances are fitted with this theoretical $D_A(z)$ function and the constraints on $H_0$ are shown in Fig. 1 (solid curve). For this fit, we obtain $H_0 = 77.5^{+3.0}_{-3.1} \text{ km s}^{-1} \text{ Mpc}^{-1}$ (68% per cent confidence interval). The $\chi^2$ per degree of freedom for the concordance model with an optimized Hubble constant is $\chi^2_{\text{tot}} = 29.21/37 = 0.789$, remembering that all of its parameters, save for $H_0$, are assumed to have prior values.

If we relax the priors, and allow both $\Omega_m$ and $H_0$ to be free parameters, we obtain best-fitting values ($\Omega_m, H_0$) = (0.50, 73.93 $\pm 0.08 \text{ km s}^{-1} \text{ Mpc}^{-1}$). Fig. 2 shows the $1\sigma$–$3\sigma$ constraint contours of the probability in the ($\Omega_m, H_0$) plane. These contours show that at the $1\sigma$ level, $64.4 < H_0 < 84.5 \text{ km s}^{-1} \text{ Mpc}^{-1}$, but that $\Omega_m$ is poorly constrained; only a lower limit of $\sim 0.08$ can be set at this confidence level. The cross indicates the best-fitting pair ($\Omega_m, H_0$) = (0.50, 73.9 $\text{ km s}^{-1} \text{ Mpc}^{-1}$). We find that the $\chi^2$ per degree of freedom is $\chi^2_{\text{tot}} = 29.01/36 = 0.806$.

4.2 $w$CDM and E-deS

For the $w$CDM model, $w_{de}$ is constant but possibly different from −1. For a flat universe ($\Omega_m = 0$), there are therefore possibly three free parameters: $\Omega_m$, $w_{de}$, and $H_0$, though to keep the discussion as simple as possible, we will here treat $\Omega_m$ as fixed at the value 0.3, and allow $H_0$ and $w_{de}$ to vary.

Fig. 3 shows the constraints using the Bonamente et al. (2006) ADD data with the dark-energy model. These contours show that at the $1\sigma$ level, $62.3 < H_0 < 97.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$, but that $w_{de}$ is poorly constrained; only an upper limit of $\sim 0.47$ can be set at this confidence level. The cross indicates the best-fitting pair ($w_{de}, H_0$) = (−0.55, 73.1 $\text{ km s}^{-1} \text{ Mpc}^{-1}$) with $\chi^2_{\text{tot}} = 29.04/36 = 0.807$. With the current level of precision, we see that the cluster ADD data can be fitted very well with this variation of ΛCDM, but the optimized parameter $w_{de}$ is only marginally consistent with that of the concordance model. We emphasize however that the scatter seen in the best-fitting diagram (Fig. 4) is still significant. Future
measurements of the ADD may provide much tighter constraints on the cosmological parameters.

For the E-deS model with \( \Omega_m = 1 \), the only free parameter is \( H_0 \). As shown in Fig. 1 (dotted curve), the best fit corresponds to \( H_0 = 67.5^{+3.5}_{-3.2} \text{ km s}^{-1} \text{ Mpc}^{-1} \) (68 per cent confidence interval). With 38 – 1 = 37 degrees of freedom, the reduced \( \chi^2 \) for the E-deS model with an optimized Hubble constant is \( \chi^2_{\text{ dof}} = 29.49/37 = 0.797 \).

4.3 The \( R_h = ct \) Universe

The \( R_h = ct \) Universe also has only one free parameter, \( H_0 \). The results for the \( R_h = ct \) Universe are shown in Fig. 1 (dashed line). We see here that the best fit corresponds to \( H_0 = 72.6^{+3.8}_{-3.4} \text{ km s}^{-1} \text{ Mpc}^{-1} \) (68 per cent confidence interval). With 38 – 1 = 37 degrees of freedom, the reduced \( \chi^2 \) in \( R_h = ct \) is \( \chi^2_{\text{ dof}} = 29.07/37 = 0.786 \).

To facilitate a direct comparison between \( \Lambda \text{CDM} \) and \( R_h = ct \), we show in Fig. 4 the 38 Chandra/SZE cluster distance measurements, together with the best-fitting theoretical curve in the \( R_h = ct \) Universe (corresponding to \( H_0 = 72.6^{+3.8}_{-3.4} \text{ km s}^{-1} \text{ Mpc}^{-1} \)), and similarly for \( \Lambda \text{CDM} \) (with \( H_0 = 73.9^{+10.6}_{-3.4} \text{ km s}^{-1} \text{ Mpc}^{-1} \) and \( \Omega_m = 0.50 \)). In this figure, we also show the ADDs of nearby clusters from Mason et al. (2001). These are not included in the fits, but demonstrate that the best-fitting curves are in agreement with low-redshift X-ray/SZE measurements. Strictly based on their \( \chi^2_{\text{ dof}} \) values, \( \Lambda \text{CDM} \) and \( R_h = ct \) appear to fit the data comparably well. However, because these models formulate their observables (such as the ADD in equations 2 and 3) differently, and because they do not have the same number of free parameters, a comparison of the likelihoods for either being closer to the ‘true’ model must be based on model selection tools. As we shall see, the results of our analysis in this paper tend to favour \( R_h = ct \) over \( \Lambda \text{CDM} \), which we now demonstrate quantitatively.

4.4 Model selection tools

The use of model selection tools in one-on-one cosmological model comparisons has already been discussed extensively in the literature (see e.g. Liddle, Mukherjee & Parkinson 2006; Liddle 2007). More recently, the successful application of model selection tools, such as the Akaike Information Criterion (AIC), the Kullback Information Criterion (KIC), the Bayes Information Criterion (BIC), has been reported by Shi, Huang & Lu (2012) and Melia & Maier (2013).

For each fitted model, the AIC is given by

\[
\text{AIC} = \chi^2 + 2\eta, \tag{6}
\]

where \( \eta \) is the number of free parameters. If there are two models for the data, \( M_1 \) and \( M_2 \), and they have been separately fitted, the one with the least resulting AIC is assessed as the one more likely to be ‘true’. If \( AIC_0 \) is associated with model \( M_0 \), the unsymmetrized confidence that \( M_a \) is true is the ‘Akaike weight’ \( \exp(-\text{AIC}_a/2) \), with likelihood

\[
\mathcal{L}(M_a) = \frac{\exp(-\text{AIC}_a/2)}{\exp(-\text{AIC}_0/2) + \exp(-\text{AIC}_2/2)} \tag{7}
\]

of being closer to the correct model. Thus, the difference \( \text{AIC}_2 - \text{AIC}_1 \) determines the extent to which \( M_1 \) is favoured over \( M_2 \). KIC takes into account the fact that the probability density functions of the various competing models may not be symmetric. The unbiased estimator for the symmetrized version (Cavanaugh 1999, 2004) is given by

\[
\text{KIC} = \chi^2 + 3\eta. \tag{8}
\]

It is very similar to the AIC, but strengthens the dependence on the number of free parameters (from 2\( \eta \) to 3\( \eta \)). The strength of the evidence in KIC favouring one model over another is similar to that for AIC; the likelihood is calculated using the same equation (7), though with \( AIC_0 \) replaced with \( KIC_0 \). The BIC is an asymptotic \( (N \to \infty) \) approximation to the outcome of a conventional Bayesian
inference procedure for deciding between models (Schwarz 1978), defined by
\[ \text{BIC} = \chi^2 + (\ln N)g, \]
where \( N \) is the number of data points. It suppresses overfitting very strongly if \( N \) is large, and has now been used to compare several popular models against \( \Lambda \text{CDM} \) (see e.g. Shi et al. 2012).

With the optimized fits we have reported in this paper, our analysis of the cluster ADDs shows that the AIC does not favour either \( R_0 = ct \) or the concordance model when we assume prior values for all of its parameters (except for the Hubble constant). The calculated AIC likelihoods in this case are \( \approx 51.7 \) per cent for the former, versus \( \approx 48.3 \) per cent for the latter. However, if we relax some of the priors, and allow both \( \Omega_m \) and \( h_0 \) to be optimized in \( \Lambda \text{CDM} \), then \( R_0 = ct \) is favoured over the standard model with a likelihood of \( \approx 72.5 \) per cent versus \( 27.5 \) per cent using AIC, \( \approx 81.3 \) per cent versus \( \approx 18.7 \) per cent using KIC, and \( \approx 85.7 \) per cent versus \( \approx 14.3 \) per cent using BIC. The ratios would be much greater for variants of \( \Lambda \text{CDM} \) that contain more free parameters than the basic \( \Lambda \text{CDM} \) model. Since we fixed \( \Omega_m \) in the \( w \text{CDM} \) model, there are two free parameters when we are fitting the ADD data in \( w \text{CDM} \). We find that \( R_0 = ct \) is favoured over \( w \text{CDM} \) by a likelihood of \( \approx 72.8\text{--}85.9 \) per cent versus \( 27.2\text{--}14.1 \) per cent using these three model selection criteria.

5 DISCUSSION AND CONCLUSIONS

In this paper, we have used the cluster angular-diameter size versus redshift data to compare the predictions of several cosmological models. We have individually optimized the parameters in each case by minimizing the \( \chi^2 \) statistic. We have found that the current cluster ADD constraints are not very restrictive for \( w \text{CDM} \), though its optimized parameter values appear to be consistent at the 1\( \sigma \) level with those of the concordance model.

A comparison of the \( \chi^2_{\text{od}} \) for the \( R_0 = ct \) Universe and \( \Lambda \text{CDM} \) shows that the cluster ADD data favour the former over the latter though, on the basis of these \( \chi^2_{\text{od}} \) values, one would conclude that both provide good fits to the observations. The \( R_0 = ct \) Universe fits the data with \( \chi^2_{\text{od}} = 0.786 \) for a Hubble constant \( h_0 = 72.6_{-0.8}^{+0.6} \text{ km s}^{-1} \text{ Mpc}^{-1} \), and \( h_0 \) is the sole parameter in this model. By comparison, the optimal flat \( \Lambda \text{CDM} \) model, which has two free parameters (including \( \Omega_m = 0.50 \) and \( h_0 = 73.9_{-0.6}^{+0.5} \text{ km s}^{-1} \text{ Mpc}^{-1} \)), fits the angular-size data with a reduced \( \chi^2_{\text{od}} = 0.806 \). However, statistical tools, such as the AIC, KIC, and BIC, tend to favour the \( R_0 = ct \) Universe. Since \( \Lambda \text{CDM} \) (with assumed prior values for \( k \) and \( w \) ) has one more free parameter than \( R_0 = ct \), the latter is preferred over \( \Lambda \text{CDM} \) with a likelihood of \( \approx 72.5 \) per cent versus \( \approx 27.5 \) per cent using AIC, \( \approx 81.3 \) per cent versus \( \approx 18.7 \) per cent using KIC, and \( \approx 85.7 \) per cent versus \( \approx 14.3 \) per cent using BIC. The ratios would be greater for the other variants of \( \Lambda \text{CDM} \) because they each have more free parameters than the basic \( \Lambda \text{CDM} \) model. Since we fixed \( \Omega_m \) in the \( w \text{CDM} \) model, there are two free parameters when we are fitting the ADD data in \( w \text{CDM} \). We find that \( R_0 = ct \) is favoured over \( w \text{CDM} \) by a likelihood of \( \approx 72.8\text{--}85.9 \) per cent versus \( 27.2\text{--}14.1 \) per cent using these three model selection criteria.

But as we have found in other one-on-one comparisons, fits to the data using \( \Lambda \text{CDM} \) often come very close to those of \( R_0 = ct \), which lends some support to our inference that the standard model functions as an empirical approximation to the latter. Though it lacks the essential ingredient in \( R_0 = ct \) — the equation of state \( p = -\rho/3 \) — it none the less has enough free parameters one can optimize to produce a comparable fit to the observations. There is no better example of this than the comparison of curves in Fig. 4. The optimized \( \Lambda \text{CDM} \) prediction tracks that of the \( R_0 = ct \) Universe almost identically. For this particular data set, the difference in outcome using model selection tools is therefore almost entirely due to the fact that \( R_0 = ct \) has fewer parameters than the standard model.

With this result, we can now address the second goal of our paper – to provide some evidence in support of, or against, the strong evolution in galaxy size implied by earlier studies assuming \( \Lambda \text{CDM} \). The analysis of the average linear size of galaxies with the same luminosity, though over a range of redshifts (see e.g. Barden et al. 2005; McIntosh et al. 2005; Trujillo et al. 2006), had indicated that the standard model is consistent with these data only if these galaxies were \( \approx 6 \) times smaller at \( z = 3.2 \) than at \( z = 0 \). López-Corredoira (2010) reconsidered this question, pointing out that current ideas invoked to explain this effect probably cannot adequately account for all of the deficit of large objects at high redshifts. For example, the main argument in favour of the evolution in size for a fixed luminosity is that younger galaxies tend to be brighter, and we should expect to see younger galaxies at high redshift. Therefore, for a given luminosity corresponding to some radius at present, galaxies in the past could have produced the same luminosity with a smaller radius. López-Corredoira estimates that this effect can contribute a factor \( \approx 2\text{--}3 \) difference in size, depending on whether one is looking at elliptical or disc galaxies. Other factors contributing to the evolution in size include mergers, and observational selection effects, such as extinction.

It appears that none of these explanations, on their own, can account for the required overall growth. However, our use of galaxy clusters in this paper has shown that the cluster ADD data are quite consistent with both \( \Lambda \text{CDM} \) and \( R_0 = ct \). Thus, the fact that both of these models provide an excellent explanation for the ADD measurements using clusters casts some doubt on the suggestion that the inferred galaxy-size evolution is due to the adoption of an incorrect cosmology, rather than effects such as those discussed above. If it turns out that the combination of growth factors together cannot account for the factor 6 difference in size between \( z = 3.2 \) and 0, we would conclude that some other astrophysical reasons are responsible, rather than the adoption of an incorrect expansion scenario.

In conclusion, even though the current galaxy-cluster angular size versus redshift data are still somewhat limited by relatively large uncertainties, we have demonstrated that they can none the less already be used to carry out meaningful one-on-one comparisons between competing cosmologies. The ADD measurements for these structures appear to be consistent with both \( \Lambda \text{CDM} \) and the \( R_0 = ct \) Universe, suggesting that the inconsistent results obtained with similar work using individual galaxy sizes are probably not an indication of gross deficiencies with the assumed cosmological model.

ACKNOWLEDGEMENTS

We acknowledge Yun Chen for his kind assistance. This work is partially supported by the National Basic Research Program (‘973’ Program) of China (Grants 2014CB845800 and 2013CB834900), the National Natural Science Foundation of China (grant nos 113222328 and 11373068), the One-Hundred-Talents Program, the Youth Innovation Promotion Association, and the Strategic Priority Research Program ‘The Emergence of Cosmological Structures’ (Grant No. XDB09000000) of the Chinese Academy of Sciences, and the Natural Science Foundation of Jiangsu Province. FM is
grateful to Amherst College for its support through a John Woodruff Simpson Lectureship, and to Purple Mountain Observatory in Nanjing, China, for its hospitality while this work was being carried out. This work was partially supported by grant 2012T110011 from the Chinese Academy of Sciences Visiting Professorships for Senior International Scientists, and grant GDJ20120491013 from the Chinese State Administration of Foreign Experts Affairs.

REFERENCES

Barthel P. D., Miley G. K., 1988, Nature, 333, 319
Komatsu E. et al., 2011, ApJS, 192, 18
Vishwakarma R. G., 2001, Class. Quantum Gravity, 18, 1159

This paper has been typeset from a TeX file prepared by the author.