COSMOLOGY

Tensions between the early and late Universe

A Kavli Institute for Theoretical Physics workshop in July 2019 directed attention to the Hubble constant discrepancy. New results showed that it does not appear to depend on the use of any one method, team or source. Proposed solutions focused on the pre-recombination era.

Nearly a century of cosmological research has led us to a standard model of cosmology, the Λ cold dark matter model, or ΛCDM, with its six free parameters and several ansatzes. This model is dominated by dark components (energy and matter) with still uncertain physics. With some simplifying assumptions about the uncertain bits, the version of the model calibrated on physics prior to recombination (63% dark matter, 15% photons, 10% neutrinos and 12% atoms) is used to predict the physical size of density fluctuations in the plasma — that is, how far sound, or any perturbation in the photon–baryon fluid, could have travelled from the beginning of the Universe to recombination — the sound horizon and its overtones, as well as the primordial baryon density. By comparing the fluctuation spectrum predicted by the model to the angular spectrum observed in the cosmic microwave background (CMB), the six free parameters are set and the ansatzes are tested. An alternative to the use of the CMB for setting the sound horizon may be derived by relating measurements of the primordial deuterium abundance to the predicted baryon density. The evolving form of the model (68% dark energy, 27% dark matter and 5% atoms) is then used to predict the expansion history of the Universe from redshift $z = 1{,}000$ to $z = 0$. Uncalibrated high-redshift type-Ia supernovae (SNe) and baryon acoustic oscillations (BAOs) provide 'guard rails' between $z \approx 2$ and 0; they do not tell us if we are on the 'right road' but they make sure we do not miss the curves in the model's road (for instance, the cosmic acceleration must be consistent with $w = -1$, where the equation of state parameter w is model-dependent and $w = -1$ corresponds to a cosmological constant) along the way. The model calibrated on early-Universe observations predicts the present-day value of several cosmological parameters, some of which can be empirically measured locally (for $z < 1$) with little or no model dependence. In particular, the model calibrated with data from the Planck mission predicts the Hubble constant, today's expansion rate, to a remarkable 1% precision, 67.4 ± 0.5 km s^{-1} Mpc^{-1}. Is this whole story right?

The simplest test of this paradigm, from end to end, is to compare the absolute scale provided through the application of early-Universe physics (for example, the physical size of the sound horizon used to interpret the CMB and BAOs) to the absolute scale measured by the Hubble constant in the local, late-time Universe. Because the Universe has only one true scale, and in light of the uncertain physics of the dark sector, comparing the two calibrated at opposite ends of the Universe's history is natural and potentially insightful. (To determine whether a measurement is truly derived from the 'early' or 'late' Universe it is necessary to trace back its chain of calibration — a useful check is to determine whether or not it depends on, for example, the number of neutrinos assumed in the standard model.) During 15–17 July 2019, 108 attendees of the workshop 'Tensions between the Early and the Late Universe' gathered at the Kavli Institute for Theoretical Physics (KITP) to consider growing tensions between the early-Universe predictions and the late-Universe measurements and how they might be explained. More details about the workshop, including online presentations, are available here: https://www.kitp.ucsb.edu/activities/enervac-c19.

The early Universe

The early Universe probes were discussed at length. The two key questions were: (1) What kind of cross-checks can be used to identify unknown systematic errors that may affect the predictions for H_0? (2) Is there any hint of tension in early-Universe data that may perhaps reveal systematic errors or shortcomings of the standard six-parameter model?

Several talks addressed the first question. In addition to the well-known small difference between the inference of H_0 from low- and high-angular-resolution Planck and Wilkinson Microwave Anisotropy Probe data, all of the early-Universe data seem to be consistently predicting a low value of H_0. The Atacama Cosmology Telescope (ACT) and the South Pole Telescope (SPT) are in agreement with Planck, and any CMB data used to calibrate the sound horizon and subsequently the BAOs leads to a low H_0 of ~67–68.5 km s^{-1} Mpc^{-1}, even without Planck. A completely independent and statistically consistent value of H_0 can be obtained by using light-element abundances to calibrate the sound horizon, BAOs and other lower-redshift probes.

As far as the second question is concerned, some curiosities among high-redshift probes at the level of ~2σ were identified. The most compelling ones appear to be the departure from unity of the nuisance parameter A_{lens}, which is used to match CMB anisotropies (temperature fluctuations around $z \approx 1{,}000$) and CMB-lensing data (from the deflection of CMB photons by gravitational masses, such as clumps of dark matter). If confirmed, this departure from unity represents evidence that something is not well understood in the relationship between CMB anisotropies and the growth of structure, and thus could perhaps hint at new physics. The other 2σ curiosities that were discussed were: (1) the tension between the two-dimensional constraints in the S_8–Ω_m plane inferred from the CMB and those inferred by cosmic shear data, where $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$, with σ the present-day linear theory root-mean-square amplitude of the matter fluctuation spectrum averaged in spheres of radius 8 h^{-1} Mpc, and Ω_m is the present-day matter density in units of the critical density; (2) the tension between the cosmological parameters inferred from the BAO signal in galaxies at $z < 1$ and those of the Ly-α line of hydrogen at higher redshifts; and (3) drifts of the model parameters with the CMB fluctuation scale used to determine the model. The statistical errors of these methods are expected to shrink in the next few years, and will reveal whether the tension is a statistical fluke of the kind that one may expect when considering of order dozens of true and nuisance parameters, or whether it is indicative of some yet-to-be discovered systematic or new physics. Nevertheless, many wondered if a solution to the late- versus early-Universe discrepancy may be more

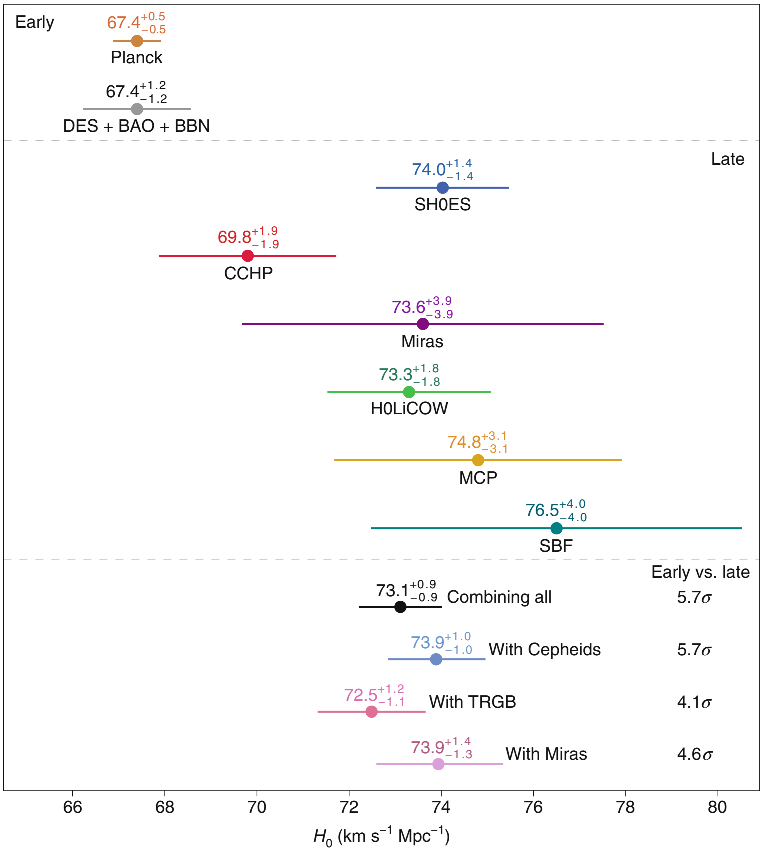

Fig. 1 | Compilation of Hubble constant predictions and measurements taken from the recent literature and presented or discussed at the meeting. Two independent predictions based on early-Universe data[28,29] are shown in the top left (additional results, utilizing other CMB experiments, have been presented with similar findings). DES, Dark Energy Survey; BBN, Big Bang nucleosynthesis. The middle section shows late-Universe measurements and the bottom section shows combinations of the late-Universe measurements and lists the tension with the early-Universe predictions, using Gaussian approximations to the posterior distribution functions of each method. We stress that the three variants of the local distance ladder method — Cepheids (SH0ES), TRGB (CCHP) and Miras — share some SN Ia calibrators and cannot be considered statistically independent. Likewise, the SBF method is calibrated based on Cepheids or TRGB and so it cannot be considered fully independent of the local distance ladder method. Thus, the 'combining all' value should be considered for illustration only, since its derivation neglects covariance between the data. The three combinations of Cepheids, TRGB and Miras are based on statistically independent datasets and therefore the significance of their discrepancy with the early-Universe prediction is accurate — even though, of course, separating the probes results in a loss of precision. In summary, the difference is more than 4σ, less than 6σ and is robust to the exclusion of any one method, team or source. The software that was used to make the figure is publicly available here: https://github.com/vbonvin/. Credit: courtesy of Vivien Bonvin and Martin Millon.

credible if it also ameliorated one or more of these lesser tensions.

The late Universe

A summary talk from the SH0ES (Supernovae H_0 for the Equation of State) team presented the status of a fifteen-year effort to build a consistently measured distance ladder using geometric distances to calibrate Cepheids, followed by 19 hosts to both type-Ia supernovae (SNe Ia) and Cepheids, followed by hundreds of SNe Ia in the Hubble flow. Highlights included the use of near-infrared Hubble Space Telescope (HST) photometry for all Cepheid data and five independent sources of geometric distances including three types of Milky Way parallaxes, Large Magellanic Cloud (LMC) detached eclipsing binaries and the masers in NGC 4258. The result was $H_0 = 74.0 \pm 1.4$ km s^{-1} Mpc^{-1} (ref. [1]).

The H0LiCOW (H_0 Lenses in COSMOGRAIL's Wellspring) team described new results from strong lensing time delays between multiple images of background quasars. Six such systems have now been measured leading to $H_0 = 73.3^{+1.7}_{-1.8}$ km s^{-1} Mpc^{-1} (ref. [2]). Robustness was demonstrated by the use of double- and quadruple-lensed quasars, long and short time delays, imaging from HST or Keck adaptive optics[3] and different mass-modelling software packages[4]. Because each lens model must be constructed individually for each system, the lensing analysis was carried out blindly to avoid experimenter bias in the model construction and is completely independent of the local distance ladder method. The SH0ES and H0LiCOW results were known only shortly before the meeting and together provide a 5.3σ difference from the early-Universe value.

New results from the Megamaser Cosmology Project (MCP)[5] were also presented, which uses very long baseline interferometry (VLBI) observations of water masers in circumnuclear orbits around supermassive black holes to measure geometric distances. A much-improved measurement of the distance to the nearby NGC 4258 (similar distance but full error reduced from 2.6% to 1.5%) was presented. In addition, a longer timespan of VLBI measurements and improved analysis of the distances to four other masers in the Hubble flow — UGC 3789, CGCG 074-064, NGC 5765b and NGC 6264 — were presented and together yielded $H_0 = 74.8 \pm 3.1$ km s^{-1} Mpc^{-1}. This determination does not require a distance ladder.

The Carnegie–Chicago Hubble Program (CCHP) collaboration, which used tip of the red giant branch (TRGB) measurements in the LMC to calibrate 18 SNe Ia (across 14 hosts) in lieu of Cepheids to connect the distance ladder[6], had new results as well. The pros and cons of TRGB and Cepheids were extensively discussed across three talks. TRGB was recognized as a valuable independent tool, well-understood from first principles and observable on simple backgrounds. The CCHP result of $H_0 = 69.8 \pm 1.9$ km s^{-1} Mps^{-1}, based on a new calibration of the TRGB I-band luminosity of $M_I = -4.05$ in the LMC, was presented[6]. The source of the reduction in this value compared to the prior TRGB result of $M_I = -3.97$ from ref. [7] was extensively discussed and identified as traceable to a 0.08 mag increase in the estimate of the LMC TRGB I-band luminosity, of which 0.06 ± 0.02 mag was attributed to a different method for estimating TRGB extinction in the LMC, which is 3σ greater than the values given by the Optical Gravitational Lensing Experiment (OGLE) reddening maps (the TRGB colour method yields $A_I = 0.16 \pm 0.02$, whereas the reddening

maps used in ref. [7] gave $A_I = 0.10$).
Yuan et al. presented a poster showing that blending in ground-based observations of TRGB in the Magellanic Clouds, particularly for the Magellanic Cloud Photometric Survey (MCPS), can bias the calibration of TRGB. The corresponding paper[8], which was subsequently submitted to *The Astrophysical Journal*, showed that the use of the MCPS data impacted the measurement of LMC extinction and its fix would change H_0 as determined from TRGB to 72.4 ± 1.9 km s^{-1} Mpc^{-1}.

New infrared surface brightness fluctuation (SBF) measurement results on two new sets of HST imaging data of early-type hosts from the massive sample and a sample of SNe Ia hosts were presented, raising the total sample size from $N = 15$ to $N = 54$ galaxies out to 100 Mpc (ref. [9]). The result was $H_0 = 76.5 \pm 4.0$ km s^{-1} Mps^{-1}. Important features of this measurement are that it is fully independent of the use of SNe Ia (which other distance ladders use) and was shown to vary within the error by altering the source of the calibration of the SBF luminosity from Cepheids, TRGB and stellar population models.

New results were presented using oxygen-rich variable stars at the tip of asymptotic giant branch stars (Miras) observed in the near-infrared in lieu of TRGB or Cepheids to connect the distance ladder. The pros of using Miras are that they are brighter than TRGB and offer an older population than Cepheids that are present in elliptical galaxies and haloes of spiral galaxies. The cons include the potential confusion between C-rich Miras and those undergoing hot bottom burning, but this issue can be mitigated by using the period range $P < 400$ d. Results from the SH0ES Team[10] were presented using Mira measurements with HST filter F160W in NGC 4258 and in the halo of the farthest Mira host and first SN Ia host to date, NGC 1559, at a distance near 20 Mpc. A distance ladder from Miras to connect the geometric distances to the LMC and NGC 4258 and then to calibrate SNe Ia yielded $H_0 = 73.6 \pm 3.9$ km s^{-1} Mps^{-1} with the error dominated by the 5% uncertainty in the calibrated luminosity of a single SN Ia, but observations for three other SNe Ia were reported to be in progress.

These two recent results and four new results are shown in a summary plot (Fig. 1) with approximate data combinations yielding a 4σ to 6σ discrepancy with the early-Universe result. The use of all data yielded a 6σ difference and a value of $H_0 = 73.3 \pm 0.8$ km s^{-1} Mps^{-1}. However, there is some overlap in data between the three ladders (Cepheids, TRGB and Miras) that connects geometrical distances to SNe Ia, and therefore they cannot be simply averaged without accounting for their covariance. To give the reader a set of truly independent datasets and a feel for the impact of removing some experiments, combinations using only one of these at a time (with the other two eliminated) are also shown. The combination including Cepheids yields $H_0 = 73.9 \pm 1.0$ km s^{-1} Mps^{-1} and a 5.8σ difference. Miras yield a similar combined result but with lower (4.4σ) significance due to the small number of Mira/SN hosts. TRGB (without Cepheids or Miras) combined with the other measurements gives $H_0 = 72.5 \pm 1.2$ km s^{-1} Mps^{-1} and a 4σ difference. A more careful combination of all of the data accounting for all covariance would thus be expected to give a result between 4σ and 6σ. The use of all methods that do not use SNe Ia (masers, SBF and strong lensing) also yield more than a 4σ discrepancy with respect to the early-Universe value.

Extensive use was made of the recent 1% measurement of the distance to the LMC from 20 detached eclipsing binaries[11] and there was some discussion on how to extend the method to other hosts.

Three talks discussed the present status of Gaia data release 2 parallaxes and approaches to calibrate their zero-point uncertainty value, which was seen to increase with brightness and redness (away from the faint, blue quasar sample used to provide the initial value). There was much optimism that this issue would be closer to being settled by the time of Gaia data release 3.

Many talks and ensuing Q&A sessions discussed future prospects of each method and the potential to reach sub-per-cent precision on H_0 from each individual method.

For the traditional Cepheid-based local distance ladder, future precision was expected to reach 1.3% with caution expressed that a notional goal of 1.0% would be hard to reach. Improvements in the TRGB- and Miras-based distance ladders are also expected with the launch of the James Webb Space Telescope and the upcoming Gaia data releases. The precision of time-delay cosmography is currently limited by sample size. With the recent explosion of discoveries of quadruply imaged quasars in wide-field imaging surveys[12] and the discovery of the first lensed SNe[13] it is clear that sample size is not a limitation anymore. The current limitation is follow-up and the scientist-time required for high-precision lens models. There is no known source of systematic error that would prevent reaching sub-per-cent precision with this method, even though of course this has to be demonstrated in practice. The precision of the MCP is also limited by sample size and although the number of masers at the correct distance is finite, there appears to be room for further improvements in precision. Likewise, there is room for further improvement in the precision of the SBF method, even though it is unclear whether either method can ultimately reach sub-per-cent precision.

A (still blind) time delay for the multiply imaged SN Refsdal was presented en route to a determination of H_0 to an anticipated statistical precision of ~7% (ref. [14]). There was also much excitement about the prospects of gravitational waves and standard sirens to contribute to the conversation. Even though the current sample and corresponding precision is not competitive with other methods[15], based on the forecasts shown, the method will soon be another powerful and independent tool at our disposal.

Completely independent of the distance determination methods is the cosmic chronometers method discussed at this meeting. In contrast to everything that we have summarized so far, the method has the advantage of measuring age (as opposed to distance) and thus it is perhaps the only one that can directly answer the existential question "how old is the Universe?" with minimal cosmological assumptions. Ref. [16] reports an age of the Universe $t_U = 13.2 \pm 0.44$ Gyr from 22 globular clusters[17]. In a ΛCDM model, the ages of these objects implies $H_0 = 71.0 \pm 2.8$ km s^{-1} Mps^{-1}. On the other hand, relative ages of suitably selected old passively evolving elliptical galaxies (see, for example, refs. [18,19] and references therein) yield an estimate of $H(z)$ with the most statistical power at $z = 0.43$, $H(z = 0.43) = 91.8 \pm 5.3$ km s^{-1} Mps^{-1}. The next step in these measurements should involve a better characterization of model uncertainties, perhaps through a survey of the model space — such efforts are underway. As the statistical precision of the method progresses, comparison with other probes and internal consistency checks will give the systematic noise floor related to our ability to determine the ages of stellar populations.

The initially reported tension between the uncalibrated BAO distances at $z_{\rm eff} = 2.3$ via the Ly-α forest and the one measured at $z = 0.75$ has decreased to below 2σ with the latest data[20,21], and is therefore not significant enough to be considered a true tension.

The tension in the σ_8 parameter (or equivalently S_8) as inferred by the Planck mission and measured by weak gravitational lensing surveys, is below the 3σ level, although the exact significance depends on the lensing dataset chosen and the assumptions made in the analysis. While investigation is important, this 'tension' is not as dramatic as that in the H_0 parameter. Two considerations are called for, however. (1) This may be related to the Planck internal consistency test offered by the parameter A_{lens}. This parametrizes the gravitational lensing amplitude in CMB data and of course depends very closely on the amplitude of perturbations (that is, σ_8). When inferred from the smoothing of the high-ℓ angular temperature power spectrum peaks, its value is $\gtrsim 2\sigma$ away from that inferred from the CMB lensing signal. (2) Any new physics introduced to explain the H_0 discrepancy should not make the σ_8 tension significantly worse.

Ideas to reconcile the two

This leaves us with the question: how can the H_0 discrepancy be solved?

The most sceptical approach is to invoke systematic errors in the data. However, given the size of the discrepancy and the number of independent routes finding it, a single systematic error cannot be the explanation. It should also be said that to follow this approach too strongly is to lose the ability to make fresh discoveries. A more formal way to invoking multiple, unknown systematic errors follows the BACCUS (BAyesian Conservative Constraints and Unknown Systematics) approach[22]. This was not presented at the meeting, but since then one of us (L.V.) has experimented with this approach using only shifts to combine the late-Universe H_0 determinations, marginalizing over (unknown) possible systematic shifts for each measurement. The shifts are assumed to be drawn from the same prior distribution: a Gaussian with width σ_a, which gets marginalized with a uniform hyperprior $-10 < \sigma_a < 10$ km s^{-1} Mpc^{-1}. As expected, the BACCUS combination widens the tails of the resulting posterior distribution compared to the conventional (Gaussian) combination, but does not single out any local-Universe measure as atypically shifted compared to the others. The combination of all independent measurements (SH0ES, H0LiCOW, SBF, MCP, TRGB and globular cluster ages) yields a grand average of $H_0 = 72.7 \pm 1.2(2.9)$ km s^{-1} Mps^{-1} at a confidence level (CL) of 68(95)%, nominally still 4σ from the early-Universe value, but the widening of the posterior tails implies that this is 2.6(2.8)σ or a CL of 99.1(99.5)% away from a H_0 value of 68(67.4) km s^{-1} Mpc^{-1}. However, it should be noted that the conventional high bar of ~5σ (one in a million experiments) as a discovery threshold already invokes the pessimistic approach of expecting the presence of unknown systematic errors by requiring margin. The BACCUS approach provides an alternative method to invoke this pessimist's prior so a >99% CL after employing it is significant. A sceptic could require a high discovery threshold or use the BACCUS approach but presumably not both without a strong prior against the possibility of new physics.

After a thorough re-analysis and cross checks of multiple CMB observations (based on Planck, SPT, ACT observations and so on), it is clear that systematic errors in CMB data alone cannot explain the tension.

Moreover, a suite (SH0ES, Miras, H0LiCOW, MCP and SBF) of low-redshift, different, truly independent H_0 measurements, affected by completely different possible systematics, agree with each other; it seems improbable that completely independent systematic errors affect all of these measurements by shifting them all by about the same amount and in the same direction.

An obvious but important caveat is that, if this tension is an indication of new physics beyond ΛCDM, the new model should not do worse than standard ΛCDM in describing all other cosmological observations.

For example, there is not much freedom to change the expansion history from that of a standard ΛCDM model below $z \approx 2$: the guardrails offered by SNe and BAO do not allow this. Moreover, model changes away from ΛCDM are tightly constrained by CMB data[23,24].

The early-Universe H_0 determination relies on angular scales such as the sound horizon (at radiation drag) and matter–radiation equality. These angular scales are extremely well determined by CMB data, but they depend on a ratio of two qualitatively different quantities: the physical scale (which depends on early-time physics and background parameters, such as the physical densities of matter, baryons and so on) and the angular distance to the CMB (which depends on H_0 as well as other background parameters). To keep the angular scales fixed while increasing H_0, both the physical scales and the distance must decrease. To reconcile the H_0 values, the CMB-inferred sound horizon at the epoch of radiation drag should be lowered by ~7%, but any new physics should only affect the decade of expansion before recombination; changes from ΛCDM in other windows would worsen the fit to existing data. In particular, any change in background parameters (physical densities) should be mostly via H_0 and not via the density parameters themselves. Few examples to achieve this were presented. One possibility is a scalar field acting as an early dark energy component[25]. The dynamics of the field are constructed so that the energy density of this early dark energy component is relevant only over a narrow epoch in the expansion history of the Universe: after matter–radiation equality but before recombination. Such a model yields a higher value for the CMB-inferred H_0, greatly alleviating the tension and, notably, preserving a good fit to all relevant observations (CMB, BAO, SNe and so on). The epoch immediately preceding recombination is favoured because it is the time when the bulk of the sound horizon accrues. However, more data, especially CMB polarization measurements or very low multipole moment ℓ measurements at greater precision, are needed to test other predictions of the model and to determine whether an additional early component (and the extra parameters that this model introduces) is actually favoured over a ΛCDM model.

Another family of possibilities was presented that instead extends the radiation sector of the early Universe physics. A solution invoking extra free streaming neutrinos is penalized by a worse fit to high-ℓ CMB angular power spectra — where the specific gravitational coupling of free streaming neutrinos leaves its signature. However, this behaviour may be offset by allowing neutrino self-interactions so that neutrinos do not free-stream but rather behave like tightly coupled radiation[26]. A model that allows neutrino self-interactions and additional neutrino species, if compared to the standard dataset combination of CMB and BAO data, produces an allowed region in parameter space that is characterized by a high $H_0 \simeq 72$ km s^{-1} Mpc^{-1} and an extra effective neutrino species ($N_{eff} \approx 4$), but very strong coupling. Such a solution predicts specific signatures in the matter power spectrum that may be sought experimentally. However, it is difficult to achieve the strong interaction that this solution finds from a particle model-building perspective, while still evading other constraints on neutrino physics. Also, in this case future data may either find signatures of the other predictions of this model or rule it out.

Similarly, models with an extra scalar field that provides energy injection localized around matter–radiation equality[27] also improves the fit to late-time H_0 determinations (still being consistent with BAO data) at little or even no cost to the fit to CMB data. While none of these models may appear natural, similar or even greater tuning may already be required to explain the two other accepted episodes of dark energy, raising the question, is twice okay but three too many?

As precision increases, one may wonder if cracks may be appearing in the ΛCDM model. It may be possible to find models that are radically different from ΛCDM but that still provide good fits to the data — and fix the 'cracks'. These are likely to have their own specific signatures, be it in other cosmological observables or particle-physics experiments, which will be crucial to make further progress. To summarize, the final speaker (Francis-Yan Cyr-Racine) concluded: "We have yet to identify a complete solution that is palatable to both cosmologists and particle physicists, but have important clues about what a successful model would look like."

Conclusions

During the last talk, prior to hearing about possible theoretical solutions to the tension, a draft version of Fig. 1 was shown to the audience and the audience was asked to vote on their perception of the significance of the tension on the following scale: 2σ = curiosity; 3σ = tension; 4σ = discrepancy or problem; 5σ = crisis. This clearly tongue-in-cheek experiment was carried out to evaluate what kind of Bayesian prior the attendees applied to the evidence. Most attendees voted for a Hubble constant "problem" with tails in favour of both "tension" and "crisis" and with no support for something less. Therefore, it appears that the issue is serious, not only taking the uncertainties at face value, but also in the eyes of the community as represented by the KITP workshop participants.

It was also clear that a great deal of progress has been made recently using new methods to tackle previously difficult measurements. Little can be learned by simply invoking a chequered past without critical study of the developments in these new measurements.

Finally, going forward, the resolution to the 'problem' will likely require a coordinated effort from the sides of theory, interpretation, data analysis and observations. To streamline the interaction between these different communities and promote the transparent transfer of information, participants advocated adopting the following best practices:

- Model assumptions: one should always make clear where the cosmology dependence enters in a measurement or interpretation.
- Reproducibility: one should release data and non-trivial software publicly (for example, like the CLASS and CAMB codes for cosmology — whenever possible it would be useful to provide any new tools as public plug-ins for these codes). Requests for data from published results should be fulfilled promptly.
- Data transparency: one should release more low-level data products where the least (cosmological) assumptions have been made.
- Blinding: blind analysis should be done whenever possible but especially if analysis choices must be made. Alternatively, the impact of choices should be clearly presented in variants of the primary analysis.
- When combining CMB data and late-time data for models that address the Hubble problem, one should also present results from CMB data alone. In a successful model, the addition of the low-redshift data should not degrade the fit to the CMB data.
- Data challenges: whenever possible, organize mock data challenges designed to blindly test the accuracy and precision of the methods and hypotheses adopted by the community.

❐

Licia Verde[1]*, Tommaso Treu[2] and Adam G. Riess[3,4]

[1]ICREA and ICCUB, University of Barcelona, Barcelona, Spain. [2]University of California, Los Angeles, Los Angeles, CA, USA. [3]Johns Hopkins University, Baltimore, MD, USA. [4]Space Telescope Science Institute, Baltimore, MD, USA.

**e-mail: liciaverde@gmail.com*

Published online: 27 September 2019
https://doi.org/10.1038/s41550-019-0902-0

References

1. Riess, A. G., Casertano, S., Yuan, W., Macri, L. M. & Scolnic, D. *Astrophys. J.* **876**, 85 (2019).
2. Wong, K. C. et al. Preprint at https://arxiv.org/abs/1907.04869 (2019).
3. Chen, G. C.-F. et al. Preprint at https://arxiv.org/abs/1907.02533 (2019).
4. Birrer, S. et al. *Mon. Not. R. Astron. Soc.* **484**, 4726–4753 (2019).
5. Reid, M. J. et al. *Astrophys. J.* **695**, 287 (2009).
6. Freedman, W. L. et al. Preprint at https://arxiv.org/abs/1907.05922 (2019).
7. Jang, I. S. & Lee, M. G. *Astrophys. J.* **835**, 28 (2017).
8. Yuan, W., Riess, A., Macri, L., Casertano, S. & Scolnic, D. Preprint at https://arxiv.org/abs/1908.00993 (2019).
9. Potter, C. et al. In *Am. Astron. Soc. Meet. #232* 319.02 (2018).
10. Huang, C. D. et al. Preprint at https://arxiv.org/abs/1908.10883 (2019).
11. Pietrzyński, G. et al. *Nature* **567**, 200–203 (2019).
12. Shajib, A. J. et al. *Mon. Not. R. Astron. Soc.* **483**, 5649–5671 (2019).
13. Kelly, P. L. et al. *Science* **347**, 1123–1126 (2015).
14. Grillo, C. et al. *Astrophys. J.* **860**, 94 (2018).
15. Abbott, B. P. et al. *Nature* **551**, 85–88 (2017).
16. Jimenez, R., Cimatti, A., Verde, L., Moresco, M. & Wandelt, B. *J. Cosmol. Astropart. Phys.* **2019**, 043 (2019).
17. O'Malley, E. M., Gilligan, C. & Chaboyer, B. *Astrophys. J.* **838**, 162 (2017).
18. Moresco, M. et al. *Astrophys. J.* **868**, 84 (2018).
19. Moresco, M. et al. *J. Cosmol. Astropart. Phys.* **2016**, 014 (2016).
20. Addison, G. E. et al. *Astrophys. J.* **853**, 119 (2018).
21. Cuceu, A., Farr, J., Lemos, P. & Font-Ribera, A. Preprint at https://arxiv.org/abs/1906.11628 (2019).
22. Bernal, J. L. & Peacock, J. A. *J. Cosmol. Astropart. Phys.* **2018**, 002 (2018).
23. Knox, L. & Millea, M. Preprint at https://arxiv.org/abs/1908.03663 (2019).
24. Aylor, K. et al. *Astrophys. J.* **874**, 4 (2019).
25. Poulin, V., Smith, T. L., Karwal, T. & Kamionkowski, M. *Phys. Rev. Lett.* **122**, 221301 (2019).
26. Kreisch, C. D., Cyr-Racine, F.-Y. & Doré, O. Preprint at https://arxiv.org/abs/1902.00534 (2019).
27. Agrawal, P., Cyr-Racine, F.-Y., Pinner, D. & Randall, L. Preprint at https://arxiv.org/abs/1904.01016 (2019).
28. Abbott, T. M. C. et al. *Mon. Not. R. Astron. Soc.* **480**, 3879–3888 (2018).
29. Aghanim, N. et al. Preprint at https://arxiv.org/abs/1807.06209 (2018).

Acknowledgements

We thank KITP for hosting and supporting the workshop 'Tensions between the Early and the Late Universe'. We thank all of the participants for an extremely lively and productive workshop. We are also grateful to those who contributed to the scientific results presented at the meeting but could not attend in person. Finally, we are especially grateful to V. Bonvin and A. Shajib for producing and updating versions of Fig. 1 during the meeting, keeping up with the fast pace of the presentation of new results. L.V. acknowledges support from the European Union's Horizon 2020 research and innovation programme ERC (BePreSySe, grant agreement 725327). T.T. acknowledges support from the Packard Foundation through a Packard Research fellowship, from the National Science Foundation through grant AST-1714953, and from NASA through grants 10158 HST-GO-12889 and HST-14254.