REVIEW ARTICLE

A buyer's guide to the Hubble constant

Paul Shah[1] · Pablo Lemos[2] · Ofer Lahav[1]

Abstract

Since the expansion of the universe was first established by Edwin Hubble and Georges Lemaître about a century ago, the Hubble constant H_0 which measures its rate has been of great interest to astronomers. Besides being interesting in its own right, few properties of the universe can be deduced without it. In the last decade, a significant gap has emerged between different methods of measuring it, some anchored in the nearby universe, others at cosmological distances. The SH0ES team has found $H_0 = 73.2 \pm 1.3$ kms^{-1} Mpc^{-1} locally, whereas the value found for the early universe by the Planck Collaboration is $H_0 = 67.4 \pm 0.5$ kms^{-1} Mpc^{-1} from measurements of the cosmic microwave background. Is this gap a sign that the well-established ΛCDM cosmological model is somehow incomplete? Or are there unknown systematics? And more practically, how should humble astronomers pick between competing claims if they need to assume a value for a certain purpose? In this article, we review results and what changes to the cosmological model could be needed to accommodate them all. For astronomers in a hurry, we provide a buyer's guide to the results, and make recommendations.

Keywords Cosmology · Distance scale · Hubble constant · Cepheids · Supernovae · Cosmic background radiation

✉ Ofer Lahav
o.lahav@ucl.ac.uk

[1] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

[2] Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

Contents

1 Introduction

In 1917, Einstein was the first to combine the assumptions of homogeneity and isotropy with his new theory of general relativity, and produce a solution for the universe as a whole (Einstein 1917). Einstein imposed his belief in a static universe, and famously introduced the cosmological constant Λ, to make his equations compatible with this assumption. Friedmann (1922) derived a solution for an expanding (or contracting) universe, but his work remained largely unknown until

after his death. Establishing expansion as an observational fact was very challenging with the technology of the day. George Lemaître published the first estimate of the expansion rate in Lemaître (1927). Two years later, Edwin Hubble[1] combined his observations of stellar magnitudes using the Mount Wilson telescope with Shapley's, Humanson's and Slipher's redshifts z to a similar result (Hubble 1929) (Fig. 1). Hubble's constant, as it later became known, is then the constant of proportionality between recession speed v and distance d:

$$v = H_0 \, d \, . \tag{1}$$

Surprisingly perhaps, it was not until 1958 that the first recognisably modern value $H_0 \simeq 75 \;\; \mathrm{kms}^{-1} \, \mathrm{Mpc}^{-1}$ was published (Sandage 1958). Sandage made several corrections to Hubble's earlier results. First, he noted the population of Cepheid variable stars was not as homogeneous as first thought. This added both scatter and bias to distance estimates, compounded by the low numbers of Cepheids observed. Second, and more seriously, Hubble had mistaken (far brighter) HII regions as bright stars, and, therefore, his estimate of the distances to galaxies was too low.

Fast forward to today, and these historical developments seem to echo some present day debates. H_0 is measured in a number of ways, which produce inconsistent values. In particular, in 2018, the Planck Collaboration used the Cosmic Microwave Background (CMB) temperature and polarization anisotropies and the ΛCDM cosmological model, to find $H_0 = 67.4 \pm 0.5 \;\; \mathrm{kms}^{-1} \mathrm{Mpc}^{-1}$ (Planck Collaboration 2020) whereas in 2021 the SH0ES Collaboration (Riess et al. 2021) used Cepheids and supernovae to find $H_0 = 73.2 \pm 1.3 \;\; \mathrm{kms}^{-1} \mathrm{Mpc}^{-1}$ (here, and for the rest of the review we quote 68% confidence limits). We show the main modern results in Fig. 2.

Why should this disagreement matter? First, it may be a sign that the standard cosmological model is incomplete and new physics is required. All H_0 results place some reliance on a background cosmology (for example to obtain peculiar velocity adjustments from a model), but the sensitivity is large when comparing results projected over large distances. Second, other cosmological parameters such as matter densities, curvature and large scale structure are often degenerate with H_0 in observational data; knowing H_0 more accurately helps resolve their values too. Third, knowing H_0 accurately improves astrophysics, as distances in the universe are $\propto H_0^{-1}$. Its dimensionless cousin $h = H_0/100 \;\; \mathrm{kms}^{-1} \, \mathrm{Mpc}^{-1}$ is ubiquitous in formulae.

We organise our review as follows. Section 1 defines H_0 and distances; it may be skipped by a reader familiar with cosmological models. In Sect. 2, we show how H_0 is calculated from observational data, and the type of problems that might generally arise. We also briefly discuss the use of Bayesian methods as a tool to discriminate between competing observations and models. In Sect. 3, we discuss ways in which H_0 has been recently estimated. There is a rich literature on the subject, and it is difficult to cover all papers. Our approach is to cite for each topic a seminal paper,

---

[1] The historical timeline and attribution of credit for the discovery of the expansion of the universe is examined in Way (2013).

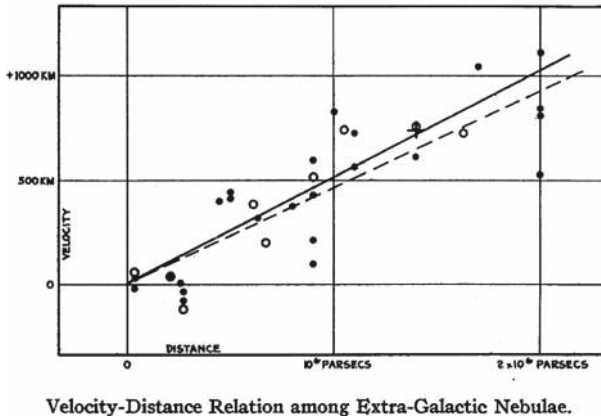Velocity-Distance Relation among Extra-Galactic Nebulae.

Fig. 1 Hubble's original diagram from Hubble (1929). Despite the typo on the labelling of the y-axis, which should read kms^{-1} and not km, it is still easy to read off $H_0 \simeq 500$ kms^{-1} Mpc^{-1}. Hubble and Humanson were using the largest telescope in the world at the time. M31 is recognisable as the lowest black dot in the bottom left; Humanson had determined it is velocity as 220 kms^{-1} towards the Milky Way (the modern value is 110 kms^{-1})

and the recent most significant developments. In Sect. 4, we discuss the possibility that measurements are correct, and it is our understanding of cosmology that is wrong. In Sect. 5 we conclude, and in the spirit of our guide for consumers, we provide our buyer's advice and recommendations. We aim to be impartial, and the views expressed here are solely our own. Busy readers could review Sects. 2, 4 and 5, and dip into Sect. 3 if more detail is needed.

For the remainder of this review, we adopt $c = 1$ throughout the text.

1.1 Linking H_0 to $H(z)$

Hubble's "constant" (Eq. 1) is not fixed when we observe beyond our local cosmological neighbourhood; which is to say it is not fixed in time. Therefore, we write $H_0 \equiv H(t_0)$, where t_0 is the present day and in general $H = H(t)$. The expansion has slowed down in the past, but the universe is now accelerating (since $z \sim 0.6$) and has been dark energy dominated since $z \sim 0.3$. The change in $H(t)$ may be written using the phenomenological *deceleration parameter* $q(t)$ as

$$\frac{dH}{dt} = -H^2(t)(1 + q(t)). \tag{2}$$

In our local universe $q(t)$ is approximately constant, and hence some authors adopt $q_0 \equiv q(t_0)$ as a constant (see for example Riess et al. 2016 or Freedman et al. 2019). As redshift is a monotonically decreasing function of time, we can write $H = H(z)$ and approximate

$$H(z) \simeq H_0[1 + (1 + q_0)z]. \tag{3}$$

As $q_0 \simeq -0.55$, light travelling to us from more than 100 Mpc away will have been

Fig. 2 Summary of recent H_0 measurements. We have intentionally selected a limited number of results to show those that are as independent from each other as possible, in the sense that they use different photometric data, distance calibrators and so on. More comprehensive versions of this plot can be found for example in Di Valentino et al. (2021). Comments: (1) CCHP and Yuan et al. share a common distance to LMC as a calibrator. (2) TDCOSMO is a re-analysis of almost the same data as H0LiCOW, but with changes to the galactic potentials. (3) BOSS and DES share a prior constraint on baryon densities from BBN. (4) The results of Blakeslee et al. use new SBF observations, whereas Khetan et al. use archival SBF distances to calibrate SN Ia. The code use to generate this figure is publicly available at https://github.com/Pablo-Lemos/plot1d
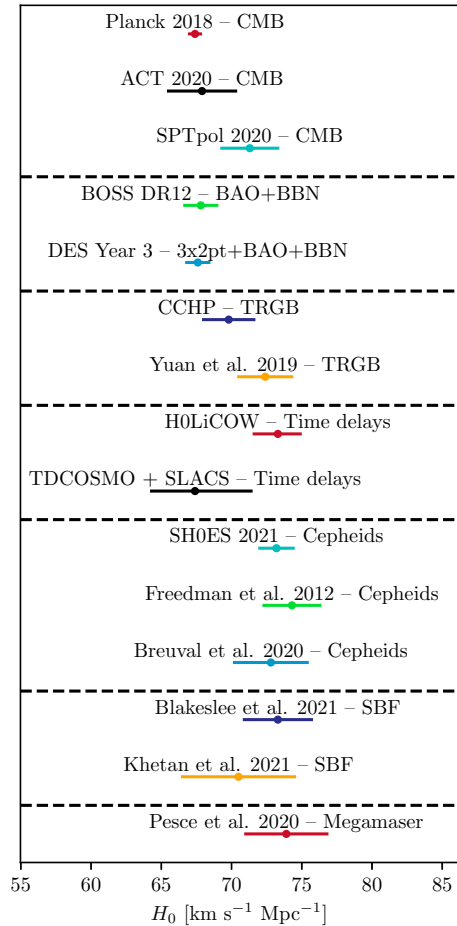
emitted when the Hubble constant was more than 1% different to its current value. The purpose of phrasing cosmography in this way is to avoid explicit assumptions on the matter or energy content of the universe, and it is of course possible to use more general parametrisations to expand $H(z)$. While this approximation is reasonable for $z \ll 0.1$, if we wish to go further, or link phenomenological parameters to physical quantities, we need a cosmological model.

Cosmological models usually assume the universe is homogenous and isotropic on large scales, and has a space-time metric. Under those assumptions, the Friedmann–Robertson–Walker (FRW) metric is

$$ds^2 = dt^2 - a^2(t)\left(\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin\theta^2 d\phi^2)\right), \qquad (4)$$

where $a(t)$ is a scale factor defining how physical distances evolve with cosmological time t, and (r, θ, ϕ) are comoving polar coordinates centered on ourselves. k is a curvature parameter, which here has units of inverse area as we wish to set $a(t_0) = 1$

(an alternative convention is to set $k = 0, \pm1$ but it is not in general possible to do both). Results from Planck (Planck Collaboration 2020) for the CMB in isolation show a preference for mildly closed universe where $k > 0$, and allowing k to vary from zero lowers the CMB derived H_0 to $63.6^{+2.1}_{-2.3}$ kms^{-1} Mpc^{-1}. However, most other observational evidence points to a flat universe (for example galaxy survey data or gravitational lensing of the CMB—see for example Efstathiou and Gratton 2020). Our discussion would not be materially affected by including spatial curvature, and for this review we will assume a flat universe where $k = 0$. We return to the point in Sect. 4.

The Hubble parameter is then defined as

$$H(t) = \frac{\dot{a}}{a}, \tag{5}$$

where $\dot{a} \equiv da/dt$. The scale factor and redshift are related by

$$a(t) = \frac{1}{1 + z_{\text{cos}}(t)}. \tag{6}$$

By z_{cos}, we mean the redshift that would be seen if both the observer and emitter were stationary in comoving coordinates, which we take to be the frame in which the CMB has no dipole. Peculiar velocity is then the velocity with respect to this frame. We know the solar peculiar velocity relative to the CMB from our observed dipole, but to estimate z_{cos} from the observed redshift z_{obs} we also need the peculiar velocity of the emitter. For example, a 10% error in a peculiar velocity of 300 kms^{-1} of a galaxy lying at 50 Mpc away would result in a 1% error in H_0 were we to calculate it solely from that galaxy. For that reason, astronomers seek large numbers of objects distributed across the sky, deep into the "Hubble flow", meaning their peculiar velocities are small compared to Hubble expansion and are assumed to average out. From here, we write $z = z_{\text{cos}}$ unless indicated otherwise.

The ΛCDM cosmological model is defined as

$$\left(\frac{H}{H_0}\right)^2 = \Omega_{k,0}(1+z)^2 + \Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{\Lambda,0}. \tag{7}$$

This can be generalised with an equation of state parameter $p = w\rho$ for dark energy where p is pressure, and the above corresponds to the cosmological constant $w = -1$. The fictitious curvature density is $\Omega_k = 1 - \Omega_m - \Omega_r - \Omega_\Lambda$, and we have assumed spatial flatness $\Omega_k = 0$.

The present day density fractions for matter, radiation and dark energy $\Omega_{i,0} \equiv \Omega_i(z = 0)$ are defined in terms of the physical densities ρ_i as

$$\Omega_i(z) = \frac{\rho_i(z)}{\rho_{\text{crit}}(z)},$$
$$\rho_{\text{crit}}(z) = \frac{3H^2(z)}{8\pi G}. \tag{8}$$

It is straightforward to expand $H(a(z))$ as a Taylor series in z and obtain Eq. (3) to first order where

$$q_0 = -\left.\frac{a\ddot{a}}{\dot{a}^2}\right|_{t=t_0}. \tag{9}$$

So, in a narrow sense, a cosmological model is a function to derive $H(z)$ from H_0 and Ω_i or vice versa (provided nothing converts energy from one type to another). This is how H_0 is calculated from $H(z)$ at $z \sim 1100$ for the CMB. Hence, one way to reconcile the Hubble tension is to change the function $H(z, \Omega_i)$ by adding new energy components or novel interactions, and we return to this later.

1.2 Distances from angles and fluxes

Luminosity distance is defined to recover the standard inverse square law ratio between the bolometric luminosity L and flux F that would hold in a flat, static universe:

$$d_L^2 \equiv \frac{L}{4\pi F}. \tag{10}$$

In a homogeneous and isotropic universe we find

$$d_L = \frac{(1 + z_{\mathrm{obs}})}{H_0} \int_0^z \frac{\mathrm{d}z'}{E(z')}, \tag{11}$$

where it is conventional to make the dependence on H_0 explicit by setting $H(z) \equiv H_0 E(z)$. Substituting in Eq. (7), for flat ΛCDM we have

$$E(z) = \sqrt{\Omega_{\Lambda,0} + \Omega_{m,0}(1 + z)^3 + \Omega_{r,0}(1 + z)^4}. \tag{12}$$

Angular diameter distance is the ratio between the physical size l of a distant object, and the small angle $\delta\theta$ it subtends on the sky:

$$d_A = \frac{l}{\delta\theta}. \tag{13}$$

The Etherington relation (Etherington 1933)

$$d_L = (1 + z_{\mathrm{obs}})^2 d_A \tag{14}$$

is a useful way to convert between the two.[2]

To link ΛCDM to our local universe, we expand Eq. (7) to second order in z:

---

[2] The relation uses z_{obs} rather than z_{cos} as it is due to the total time dilation and redshift between the source and observer. See Davis et al. (2019) for an informative review of redshifts in cosmology. We also recommend Hogg (1999) for a thorough discussion of distance definitions.

$$H(z) = H_0[1 + (1 + q_0)z + (j_0 - q_0^2)\frac{z^2}{2}],$$

$$d_L = \frac{z}{H_0}[1 + \frac{1}{2}(1 - q_0)z - \frac{1}{6}(1 - q_0 - 3q_0^2 + j_0)z^2], \qquad (15)$$

where the *jerk parameter* j_0 is defined as

$$j_0 = -\frac{a\,\ddot{a}}{\dot{a}^3}\Big|_{t=t_0}. \qquad (16)$$

Eq. 15 is now a reasonable approximation to Eqs. (7) and (11) out to $z \sim 0.6$.

Setting $\Omega_{r,0} \simeq 0$, we then obtain $q_0 = \frac{1}{2}(\Omega_{m,0} - 2\Omega_{\Lambda,0})$ and $j_0 = \Omega_{m,0} + \Omega_{\Lambda,0} \simeq 1$.

Hubble's law $v = H\,d$ is implicit in Eq. (11). Expanding the integral as a Taylor series in z, we see that $v = cz$ as used by Hubble and Lemaître (and re-introducing c here for clarity) is only valid as a low-z approximation.

1.3 The age of the universe

The age of the universe t_0 is inversely proportional to H_0, with dependence on other cosmological parameters. From the definition of $H(t) = \dot{a}/a$ we can write

$$t_0 = \int_0^1 \frac{da}{aH(a)}. \qquad (17)$$

In the special case of a flat, radiation-free universe, where we set $\Omega_{m,0} + \Omega_{\Lambda,0} = 1$, it can be written analytically as

$$t_0 = \frac{2}{3}\frac{1}{H_0}\frac{1}{\sqrt{1 - \Omega_{m,0}}} \ln\left[\frac{1 + \sqrt{1 - \Omega_{m,0}}}{\sqrt{\Omega_{m,0}}}\right]. \qquad (18)$$

For a flat universe with $\Omega_{m,0} = 0.3$ this gives $H_0 t_0 \approx 0.96$. Historically (in the 1980s), it was believed that the universe was Einstein–de Sitter ($\Omega_{m,0} = 1.0$ and $\Omega_{\Lambda,0} = 0.0$), which yields the product $H_0 t_0 = \frac{2}{3}$. Then, a low value of $H_0 \approx 50$ kms^{-1} Mpc^{-1} was required to ensure that the universe is older than the oldest globular clusters. The existence of $\Lambda > 0$ makes the universe naturally older. The Planck estimate is $t_0 = 13.797 \pm 0.023$ Gyr within ΛCDM (Table 2 for Planck alone, 68% CL; Planck Collaboration 2020). This value for the age of the universe is comfortably larger than the age of any known astronomical object.

2 Inference of H_0

2.1 Standard candles and the nearby distance ladder

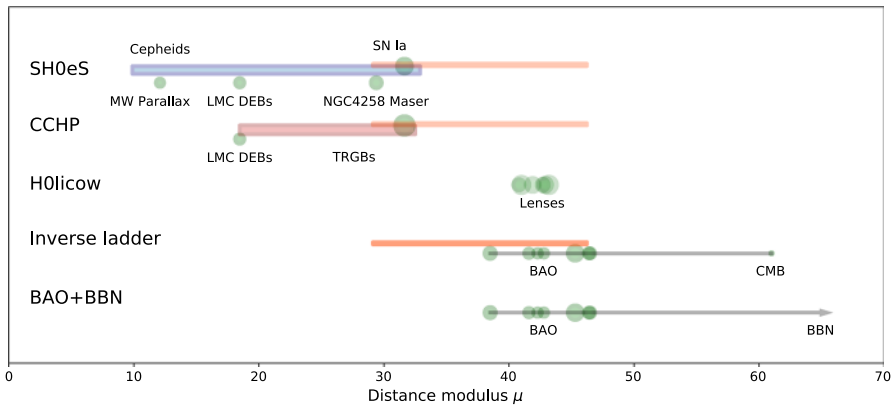A *standard candle* is any population of stars or events which

Fig. 3 Our schematic illustration of the construction of distance ladders. Green circles represent the calibration of the distance ladder (either the base or overlap of each rung), and the bars are the rungs of the distance ladder. The size of dots or thickness of bars are in proportion to their contribution to the error budget of H_0; in the case of the CMB and BBN we have opted for a thin line to show the dependence on ΛCDM

– can be reliably identified
– have the same characteristics wherever they are seen
– have an established luminosity law specifying the absolute magnitude in terms of observable quantities.

Although the luminosity law is determined empirically by calibration (except in the case of gravitational waves—see Sect. 2.3), there is an advantage if there is also a solid understanding of the underlying physics of the standard candle as in that case the calibration can be cross-checked against a theoretically-derived one.

The nearby distance ladder starts with a choice of standard candle, and a calibration of the absolute magnitude M, for example using parallax distances and apparent magnitudes m. m is defined by

$$m_X = -2.5 \log_{10}\left(\frac{F_X}{F_{X,0}}\right), \tag{19}$$

where F_X is the energy flux per unit area per second across the wavelength range of the band X, and $F_{X,0}$ is the fixed reference flux for the magnitude system being used (for example the Vega system). M is the apparent magnitude the object would have if it were at a distance of 10 parsecs. Distance is conveniently quoted as the *distance modulus*

$$\mu = m - M, \tag{20}$$

and then the luminosity distance (10) becomes

$$d_L = 10^{0.2(\mu+5)} \text{ parsecs }, \tag{21}$$

which can then be substituted into Eqs. (11), (15) or similar relations to obtain H_0.

A given standard candle seen over a range of distances is termed a "rung", and in turn used as calibrators for the next rung. For example, the SH0ES team (Riess et al. 2019) calibrate Classical Cepheids (CC) using parallaxes, a maser distance to NGC4258, and detached eclipsing binaries (DEBs) in the LMC as their first rung. Their next rung is Type Ia supernovae (SN Ia), calibrated using the 19 galaxies in which both Cepheids and SN Ia have been observed. We illustrate their construction in Fig. 3.

The quality of the standard candle depends on a number of considerations. Are there enough with good distances to accurately calibrate the absolute magnitude? Can we clearly identify them at large distances? Can they be observed out to a sufficient distance to reach the next rung? Are the objects observed at large distances of the same type as local ones used as calibrators? How to correct for extinction, reddening, metallicity effects and crowded starfields? Which band has the most reliable magnitudes? If data have been combined from different telescopes, have the right adjustments been made to convert photometry? How is magnitude to be defined for variable stars? Each rung depends on the previous one, and errors will propagate up the ladder.

A more subtle issue is that the conversion of observational data to H_0 is non-linear. As the expectation $\mathbb{E}[d^n] \neq \mathbb{E}[d]^n$ in general, scatter in observational data will introduce systemic bias. Bias can also be introduced by sample selection unless care is taken as $\mathbb{E}[d_1 < d < d_2] \neq \mathbb{E}[d]$: we average our selected sample but wish to know the expectation of the unselected one. Selection may be overt (for example by cutting outliers) or due to our telescope seeing only up to $m < m_0$ (known as Malmquist bias; Malmquist 1922). Both of these can be (and usually are) corrected for, but require some assumptions and a careful analysis of the data and reduction pipeline.

2.2 Standard rulers and the inverse distance ladder

A *standard ruler* is a feature on the sky of a known physical size l, which enables us to calculate the angular diameter distance d_A defined in Eq. (13) from their angular size. Parallax is an obvious example, and also the size of orbits of masers and detached eclipsing binaries can be determined from their positions, light curves and spectroscopy. In the early universe, acoustic pressure waves in the primordial charged particle and radiation plasma set a physical size called the *sound horizon* r_s, which is how far they travel from the initial seeds[3] that generate them. The universe was not that dense at that time, so Thomson scattering by charged particles was necessary to propagate the waves, and hence they are frozen-in by recombination. The sound horizon is then imprinted on the CMB as peaks in the power spectrum of temperature fluctuations, and in the later spatial distribution of galaxies (known as *baryon acoustic oscillations*, or BAO for short).[4] The sound horizon may be

---

[3] These are thought to be density perturbations originating from quantum fluctuations in an inflationary era.

[4] The physical size of the sound horizons defined by the CMB and BAO are different by $\sim 4\%$. Loosely speaking, the CMB sound horizon is defined by the redshift $z_\star$ where the "photons stop caring about the

calculated in a cosmological model, and depends both on the expansion rate $H(z)$ (the waves are carried along by expanding spacetime) and the matter-energy content of the early universe (determining the sound speed).

The inverse distance ladder, as its name suggests, works in the opposite direction to the nearby distance ladder but on the same principle. It can use the sound horizon for a starting d_A, a background cosmology, and the Etherington relation (14) to calibrate the luminosity distances d_L of standard candles. For example, Lemos et al. (2019) calibrate BAO at $z \simeq 1$ and SN Ia at $z < 1$ using the CMB sound horizon. They replace the standard ΛCDM formula for $H(z)$ with a parametric form. $H(z)$ is extrapolated to today to find $H_0 = 68.42 \pm 0.88$ kms^{-1} Mpc^{-1}. Thus, it is shown that discrepancy between the CMB value of Planck Collaboration (2020) and Riess et al. (2019) is not caused by assuming the late-universe is ΛCDM.

We can express the difference between early and late universe H_0 in terms of ruler size or luminosity differences. Planck implies $r_s = 147.27 \pm 0.31$ Mpc in ΛCDM; it would need to be $\simeq 10$ Mpc lower (Knox and Millea 2020) to bring consistency with Riess et al. (2019). Alternatively, Eqs. (21) and (11) show that Cepheids or SN Ia would need to be $\simeq 0.2$ mag brighter than thought to bring consistency with Planck.

In summary, we see the two ways of constructing ladders, propagated toward each other by the "guard rails" of SN Ia, do not meet! Hence, the H_0 tension is sometimes characterised as "early" versus "late", from which follows the question "Is ΛCDM right?". This may be premature: in fact, few late-universe results in isolation are fully inconsistent with early universe ones, as we discuss later.

2.3 No-ladder H_0

We have seen distance ladders require calibration, whether they are nearby or inverse. However, there are some self-calibrating observations from which H_0 may be calculated directly.

One example is the CMB. The detailed shape of the power spectrum determines other parameters in ΛCDM, and the value of H_0 derived from it is best understood as just one part of the simultaneous inference of *all* cosmological parameters. A further example is maser emission systems, which occur in the nucleus of certain galaxies and are bright enough to be seen at cosmological distances. The emission spots appear to follow Keplerian orbits, and so with some disk modelling, the size of the orbit and hence the angular diameter distance can be deduced directly.

Gravitational waves assume general relativity. The masses of merging compact objects and luminosity can be obtained from the shape of the waveform. That is to say, there is no need for an empirical calibration of their intrinsic luminosities, and instead the observational challenge is to determine the redshift of the source. Often referred to as *standard sirens* rather than standard candles, a gravitational wave event whose source galaxy has been identified (by locating the optical counterpart)

Footnote 4 continued

baryons" (otherwise known as the surface of last scattering), and for BAO by the redshift z_{drag} where "the baryons stop caring about the photons". The difference between the two is due to their number densities.

is referred to as a "bright siren", otherwise it is called a "dark siren". Most gravitational wave events are dark, but progress can be made statistically with them given sufficient numbers.

For gravitational lenses, general relativity links the time delay caused by the lensing of the background source (a combination of both the longer path and time dilation) and the mass distribution of the lensing galaxies. The absolute time delay is not known, but if a rapidly varying source like a quasar can be seen in multiple images, the relative time delay between images allows the angular diameter distance of the lens to be calculated. In this case, the challenge is to obtain enough constraints on the mass distribution of both the lensing galaxies and the general concentration of matter along the line of sight, using for example the velocity dispersions and surface brightness of the lensing galaxies, and imaging data.

2.4 What could cause the tension?

Using the word "tension", cosmologists mean the discrepancy in measurements of H_0 is at a level which is large compared to the reported errors. This means that if the values and errors are correct, this is very unlikely to be the result of chance.

Measuring the Hubble constant has always pushed the limits of the available telescopes of the day, and as a consequence observers have to be very careful to avoid bias in their derivations of H_0, and have accounted for all errors. Hubble believed he had one population of Cepheids, whereas we know today he had two, and had also confused nebulae with bright stars. Alternatively, many researchers interpret the tension in the spirit of the precession of the perihelion of Mercury: something is wrong with the (Newtonian) model and a new one is needed (general relativity). We can categorise explanations as follows:

– **Observational bias**. An observational bias is an error in mean photometry that would be expected to increase with magnitude or distance. To give some examples, consider first crowding. For distant stars, resolving them from their neighbours becomes harder, therefore, their photometry will be progressively more blended with other stars the more distant they are. Blending increases the apparent magnitude, and changes the colour (see for example the discussion in Sect. 4.2 in Javanmardi et al. 2021). For very faint stars, the response of the detector may be non-linear (Riess et al. 2019), and needs to be corrected. Another issue is combining observations between ground and space telescopes, as in general fainter stars will be observed from space, but nearby ones more cheaply from the ground. Aside from atmospheric extinction, each instrument has different passbands, detector response and resolution, meaning the magnitudes of the same star observed in each telescope will be different. Photometry must be transformed to a common system (see for example Eqs. 10–12 in Riess et al. 2016), and if not done (or done incorrectly) some bias will likely have been introduced. Any parameters derived from photometry—such as photometric redshifts—would inherit the same propensity to bias.
– **Astrophysical bias.** An astrophysical bias occurs when the properties of the object being studied are not fully resolved, and those properties differ with

distance. For example, consider Cepheids in the LMC and SN Ia hosts. The LMC is close so Cepheids with a full range of periods can be seen, whereas for distant galaxies only the brighter Cepheids with longer periods are seen. Additionally, the LMC is metal-poor compared to a typical spiral galaxy, the Cepheids there may be expected to be relatively metal-poor compared to those in SN Ia hosts. Hence, any curvature in the Leavitt law, or mis-calibrated metallicity dependency could bias distances (see for example Freedman and Madore 2010). A second example is the step-like link between SN Ia magnitudes and properties of the host galaxy (Smith et al. 2020 and references therein). This could indicate there are two distinct populations of SN Ia. If then the SN Ia in the 19 galaxies where both Cepheids and SN Ia were of mostly one type, whereas the rest a blend of both types, H_0 would be biased by the calibration of SN Ia.

– **Statistical bias.** The main causes of statistical bias will be selection effects and scatter as we discussed in the introduction. Statistical biases can be corrected by running random simulated observations through the same selection and analysis pipeline, but the simulations will themselves need some physical parameter choices, perhaps determined from previous surveys, or fixed in advance to "reasonable" levels. In Bayesian analysis, residual dependence on the choice of prior is a feature of sparse observations. Statistical bias correction is subtle and difficult, as we see later in the sections on parallax and SN Ia.

– **Physics of ΛCDM.** Before invoking new physics, could the explanation be found within ΛCDM? Cosmological formulae such as Eq. (7) are derived from a homogeneous, isotropic universe. Could corrections allowing for inhomogeneities be large enough to explain the tension? A specific example is the "Hubble Bubble" or *local void* proposal (for a recent example, see Shanks et al. 2019), in which we are by chance located in an under-dense region, and our local H_0 is different from the "universal" one. Additionally, inhomogeneities mean we must correct redshifts for peculiar velocities, and further the propagation of light through overdense or underdense regions might bias our analysis (Kaiser and Peacock 2016).

– **New Physics.** If the expansion history of the universe were different to ΛCDM, H_0 inferred from the CMB or BAO might be brought into alignment with the local value. Alternatively, performing our analysis of a non-ΛCDM universe using ΛCDM formulae may have confused us. For example, an extra particle species would increase the pre-recombination expansion speed, so reduce the size of the sound horizon: sound waves have less time to propagate before they are frozen in. To keep the same observed angular size of the CMB temperature fluctuations, the value of H_0 calculated from the CMB will increase (see Eqs. 11, 13, 14). However, as we see later ΛCDM makes many other successful predictions, such as the CMB spectrum itself or primordial element abundances, and is not lightly tampered with.
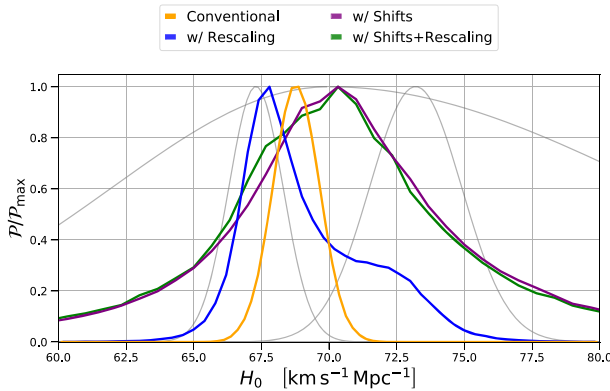
Fig. 4 An illustration of BACCUS applied to Planck, H0LiCOW and SH0ES data, whose posteriors are shown as thin grey lines. Conventional (orange) is the standard Bayesian combination assuming equal weighting, w/Rescaling (blue) is equivalent to hyper-parameters, w/Shifts (brown) adds an unknown systemic error offset to each dataset, and w/Shifts + Rescaling (green) combines hyper-parameters and systemic offsets. Image reproduced with permission from Bernal and Peacock (2018), copyright by IOP/SISSA

2.5 Is the tension significant?

The tension is often quoted as the number of standard deviations "$m\sigma$". In particle physics, the meaning is clear: there are millions of collisions and probability is frequentist. There is no need to work in a Bayesian framework, with some prior assumption of a parameter to update with new data. The law of large numbers drives distributions to Gaussian normal shape, and we can translate $m\sigma$ to a probability of occurrence by chance. At the 5σ level, it is unanimously agreed new physics has been detected. None of the above applies in cosmology!

Nevertheless, that the tension is significant should not be in doubt: see Fig. 2. But we are looking for more from our data analysis, and there are three ways in which Bayesian statistics are helpful, which we now briefly survey.

Bayes' theorem states that the posterior probability distribution is

$$P(\theta \mid D, \mathcal{M}) = \frac{P(D \mid \theta, \mathcal{M})P(\theta \mid \mathcal{M})}{P(D \mid \mathcal{M})}, \tag{22}$$

where D is the data and θ are the parameters of the model $\mathcal{M}$ of interest. For example $\mathcal{M}$ might represent ΛCDM with its associated parameters including H_0 and $\mathcal{M}_1$ some extension of it with additional parameters. $P(\theta \mid \mathcal{M})$ is any prior belief in the parameters of the model and $P(D \mid \theta, \mathcal{M})$ is called the *likelihood*, typically specified by the team analysing the data. The denominator

$$P(D \mid \mathcal{M}) = \int_{\Theta} P(D \mid \theta, \mathcal{M})P(\theta \mid \mathcal{M})\mathrm{d}\theta, \tag{23}$$

is called the *evidence*. An extended cosmological model will have a smaller evidence if there exist large values of the parameter space with low likelihood, even if

it agrees better with the data. Bayesian evidence then naturally embodies Ockham's razor: a simpler model will have a larger evidence, unless the extended model has a significantly better fit to the data.

Second, Bayesian statistics can also help in re-analysis, in the hope that the data itself may reveal issues. A Bayesian hierarchical analysis was used by Feeney et al. (2018) to test relaxing distributional and outlier assumptions embedded in the χ^2 fit used in SH0ES analyses. Cardona et al. (2017) have re-analysed SH0ES data using "hyperparameters", which are weightings of datasets proposed as a measure of credibility by Lahav et al. (2000). An agnostic prior for the weights is set, and the hyperparameters are marginalised over. Both results are consistent with SH0ES. Bernal and Peacock (2018) extend the hyperparameter method by adding a free parameter shift in the mean of each dataset to account for unknown systematics, which they dub "BACCUS". Using this to combine Planck, SH0ES and other datasets, produces a compromise. As shown in Fig. 4, the posterior middles the two with much larger error bars, which perhaps is unsurprising given the agnosticism of the method. Such types of analysis need to be taken with a grain of salt: data that is in tension should not be combined, and BACCUS is not a substitute for a critical analysis of why the tension has happened. However, if one *demands* a method to merge disparate results in a Bayesian framework, BACCUS is a way to achieve that.

Third, we may want to know how valid a combination of data sets is. If two posteriors for the same parameter barely overlap, a Bayesian analysis will seriously mislead with error bars that are too small, as shown by the yellow line in Fig. 4. We seek a statistic that is symmetric, (reasonably) independent of prior assumptions and models, and straightforward to calculate and interpret. The R statistic

$$R = \frac{P(D_A, D_B)}{P(D_A)P(D_B)} = \frac{P(D_A \mid D_B)}{P(D_A)} = \frac{P(D_B \mid D_A)}{P(D_B)} \tag{24}$$

compares the evidence of dataset D_A in light of knowing D_B to that of D_A alone, but is dependent on the prior and so is not usually comparable between different papers. Handley and Lemos (2019) define a new statistic called "suspiciousness" as

$$\log S = \log R - \log I, \tag{25}$$

where I is the information ratio $\log I = \mathcal{D}_A + \mathcal{D}_B - \mathcal{D}_{AB}$ which quantifies the information gain between prior and posterior. $\mathcal{D}_A$ is defined as

$$\mathcal{D}_A = \int_\Theta P(\theta \mid D_A) \log \frac{P(\theta \mid D_A)}{P(\theta)} d\theta, \tag{26}$$

and $\mathcal{D}_B$ and $\mathcal{D}_{AB}$ are defined similarly by replacing $A \to B, AB$, respectively. This is independent of the prior and (being an integral) the choice of parameters. We can interpret $\log S \ll 0$ as the two data sets being in tension: loosely speaking, the evidence of combining them does not exceed the information of considering them separately. Therefore, suspiciousness fits the criteria of simplicity and interpretation we outlined above.

Finally, even in light of the tension, we cannot dodge the question posed in our introduction: "All this debate is interesting, but which value for H_0 should I use, and

is it valid to use that in ΛCDM?". We suspend judgement until after we have surveyed the data and potential new models.

3 Measuring H_0

3.1 Parallax

Parallax is both the oldest astrometric technique, and the easiest to understand. Hold out your thumb at arm's length, relative to some fixed point on the wall, and alternately close one eye and then the other. Relative to the fixed wall, the apparent position of your thumb will change, and this is how our depth perception works: the smaller the change in position, the longer your arm must be. The same principle works with stellar distance, where now our "binoculars" correspond to the Earth's position on opposite sides of its orbit. The change in a fixed star's position $2\varpi = \theta_1 - \theta_2$ arcseconds, due to the change in the Earth's position by 2 a.u. over 6 months leads to the distance $d = 1/\varpi$ parsecs.[5] Our measurement of position must be very precise. Although the nearest star, Proxima Centauri, has a parallex of approximately 0.77 arcseconds, modern measurements target a remarkable 10 μas, which is the size of a thumbnail on the Moon as seen from Earth.

Modern parallax measurements began with the satellite Hipparcos (the name alludes to the ancient Greek astronomer Hipparchus, who measured the distance to the Moon). Launched by the European Space Agency in 1989, it measured the parallaxes of 100,000 stars at an accuracy of up 0.5 milliarcsecond (mas), the fixed background frame now being extragalactic sources such as quasars. Although undoubtedly impressive, at 1000 light years an error of 0.5 mas would still be a distance error of 15%. A further drawback is Cepheids, an important part of the distance ladder we discuss shortly, are relatively rare stars and only a handful are located in our neighbourhood of the Milky Way.

Moving on to the present day, our two best current sources of parallax are Gaia and the Fine Guidance Sensor/Wide Field Cameras (FGS/WFC3) aboard the Hubble Space Telescope (HST). Gaia was launched in 2013 and the mission goal is to measure over a billion stars (including 9,000 Cepheids and half a million reference quasars), both in our galaxy and satellites like the LMC. The mission-expectation precision is 7 μas at $m = 12$, rising to 26 μas at $m = 20$ (Gilmore 2018). Gaia does this by slowly scanning the sky with two telescopes set at relative angles of 106.5°, to make a one-dimensional measurement of the time and position of each star that slowly drifts across the CCD. Up to 70 measurements will be made for each star, which allows the additional calculation of proper motions, and even small changes in position caused by the gravitational tug of planets orbiting the star. The HST operates on similar principles in "spatial scanning" mode; although it cannot survey like Gaia, when focused on nearby Cepheids its errors appear competitive (Riess et al. 2018a).

---

[5] Many authors use π to denote parallax. We aim to avoid any potential confusion using ϖ.
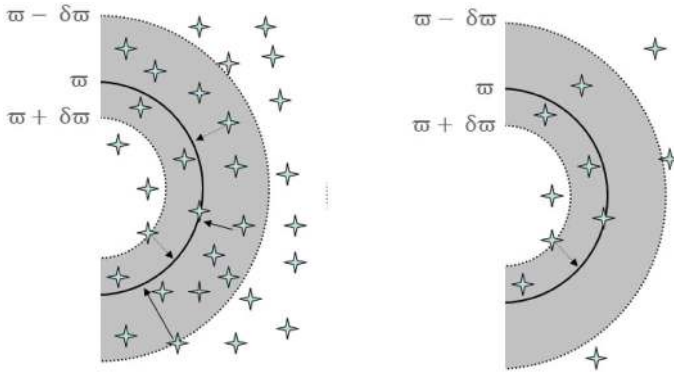
Fig. 5 An illustration of Lutz–Kelker–Hanson population bias. Assume each parallax measurement can be in error by up to $\pm\delta\varpi$. In the figure on the left, there is a constant spatial density of stars. Then, the region $(\varpi, \varpi - \delta\varpi)$ has a greater number of stars that can scatter to the observed parallax ϖ than the region $(\varpi, \varpi + \delta\varpi)$ and parallaxes are biased too low. Conversely in the figure on the right, the stellar density drops sharply beyond ϖ due to either the edge of the population or magnitude limitations. More stars are available to scatter out than in, and parallaxes are biased too high

Gaia's high precision is dependent on a very stable mechanical structure of the spacecraft. A variation in the angle between the two fields (which might be caused by thermal expansion) could cause spatial variations in apparent parallax, or if synchronous with the scan period even a fixed systematic offset. Indeed, such a variation has been inferred from the interim Data Release 2 (DR2) (Gaia Collaboration 2018). The average of the quasar parallaxes in it is $-29\,\mu$ as (negative parallaxes can happen when position measurement error is larger than the parallax), and there were indications this "zero point" may vary with stellar colour, luminosity and position on the sky (Arenou et al. 2018). Riess et al. (2018b) compared HST Cepheid parallaxes to Gaia, simultaneously solving for the Gaia zero point and Cepheid calibration. They found a difference between them $46 \pm 15\,\mu$ as, with Gaia parallaxes again appearing too low. As the typical parallax of a Milky Way Cepheid is $400\,\mu$ as, this is very material to distance estimates. Breuval et al. (2020) creatively replaced DR2 Cepheid parallaxes with those of resolved bound binary companions (where available), which being dimmer are closer to the ideal magnitude range for Gaia.

During the preparation of this review, Gaia Early Data Release 3 was made available, and has already by used by Riess et al. (2021) to revisit Gaia parallaxes. Gaia EDR3 is not intended to be the final word, but indeed Cepheid zero points seem now to be reduced below $10\,\mu$ as. A calibration of Cepheids using 75 Gaia parallaxes only gives $H_0 = 73.0 \pm 1.4$ kms^{-1} Mpc^{-1}, which is slightly lower, but consistent with their previous result based on HST parallaxes.

Before moving on to discuss alternative calibrations, we will first digress into a discussion of parallax bias. The potential for bias occurs in any astrophysical observation, and is often the subject of lengthy analysis in H_0 papers. As it is most easily understood in the context of parallax, it is helpful to discuss it here.

Parallax bias is usually referred to as Lutz–Kelker–Hanson (LKH) bias (Lutz and Kelker [1973](#); Hanson [1979](#)). This is a summation of three quite different effects: non-linearity of the desired variable (distance) with respect to the observed variable (parallax), population bias (have we observed the object we intended to, or did we confuse it with something else?), and selection bias (our surveys are normally magnitude limited, so we will only "select" objects for which $m < m_0$).

To explain non-linear bias, imagine we have a *symmetric* error in our parallax measurement, such as might be caused by an instrumental point spread function. To be concrete, suppose the likelihood of measuring $150\,\mu$ as is the same as measuring $50\,\mu$ as when $100\,\mu$ as is the true value. If we average the distance, we will obtain $\bar{d} = 6666.67 + 20000 = 13,333$ which is biased with respect to the true distance of 10,000 parsecs. In mathematical terms, because $d = 1/\varpi$, then $\mathbb{E}[d] \neq 1/\mathbb{E}[\varpi]$.

Population bias arises in parallax when we consider a broad survey of stars at a given distance r from our position. Assuming a roughly constant spatial density of similar stars, there are more stars in the shell $(r, r + \Delta r)$ than there are in the shell $(r - \Delta r, r)$ for some finite Δr. Hence, there are more (further) stars whose parallaxes may be over-estimated to place them at r than (closer) stars whose parallaxes may be over-estimated. Taken to extremes, there are huge numbers of stars with effective parallaxes of zero, waiting for their small but finite chance to "crowd in" to a given measured parallax. This will bias observed parallaxes too low. Note that if we were certain of our identification of the star (as we would be for a Milky Way Cepheid close to us), we need not consider population bias: it would stand out from the crowd (Fig. [5](#)).

Conversely, suppose we were observing close to our magnitude limitations. Now the opposite bias would occur: we cannot see the further stars, so they cannot crowd in. But the same number of closer stars are available to crowd out, so our observed parallaxes will now be biased too high. This is the well-known Malmquist bias (Malmquist [1922](#)), and is a major issue for surveys as naturally we will try to see as far as we can!

To deal with bias then involves *modelling* of the instrumental error, the selection function, the population scatter and so forth. In modern surveys, this is normally done by constructing simulated catalogs with known physical parameters and some assumptions, and putting those catalogs through the same analysis pipeline as the real data to see what biases emerge. For example, Riess et al. ([2018a](#)) compute distance modulus biases of between 0.03 and 0.12 mag for MW Cepheids using a model for galactic stellar distributions. If working in a Bayesian framework, a posterior distribution for the distance may be derived using the method of Schönrich et al. ([2019](#)). An alternative is to work directly with the parallaxes, instead converting Cepheid magnitudes to predicted parallaxes as done by Riess et al. ([2018b](#)). As the magnitudes are measured considerably more accurately than the parallaxes, bias corrections to the magnitude to parallax conversion are not necessary. To check the predictions of LKH bias, Oudmaijer et al. ([1998](#)) compared ground-based to Hipparcos parallaxes, finding a bias towards brighter magnitudes up to relative error $\approx 30\%$, and dimmer magnitudes for larger error when Malmquist bias is dominant, as one would expect from the discussion above.

We end our digression on biases here and move on to discuss other geometrical distances.

3.2 Detached eclipsing binaries

Imagine we had a star for which we knew the surface radiant flux density J, and the physical size R. The stellar luminosity would straightforwardly follow as $L = 4\pi R^2 J$, and we would have a standard candle.

Cool, stable, helium-burning giants (that is, red clump) are sufficiently bright to be seen outside the Milky Way. For these "late-type" stars, just such an empirical relationship can be established for the surface brightness S_V defined as

$$S_V \equiv V + 5\log(\phi) \quad \text{mag} , \tag{27}$$

where ϕ is the stellar angular diameter, and V the visual band magnitude. This relationship has been calibrated by angular diameters obtained from optical interferometry of nearby stars, and is given by

$$S_V = (1.330 \pm 0.017) \times ((V - K) - 2.405) + (5.869 \pm 0.003) \quad \text{mag} , \tag{28}$$

where $V - K$ is the colour difference between magnitudes in the V and 2.2 μm near-infrared K band (Pietrzyński et al. 2019). The scatter is just 0.018 mag. Rearranging the definition of surface brightness and with $\phi = 2R/d$ it then follows that

$$d = 9.2984 \times \left(\frac{R}{R_\odot}\right)(10^{0.2(V-S_V)} \quad \text{mas}) \quad \text{parsecs} , \tag{29}$$

where R and ϕ have been converted to solar radii and milli–arcseconds, respectively. The pre-factor is purely geometric.

But how can we know the radius of distant stars? Eclipsing binaries allow just that. If the stars are well separated enough to spectroscopically resolve each one,[6] but close enough to eclipse each other, we can obtain their individual surface brightnesses, colours, radial velocities, eclipse depths and shapes, and the orbital period. With this data, the radius (and other physical parameters such as mass, eccentricity and inclination of the orbital plane) of each star can be solved for. Such an alignment of the stars is of course rare, but sufficient numbers do exist! By painstakingly observing 20 systems in the LMC over more than 20 years (covering many eclipses) Pietrzyński et al. (2019) determine the stellar radii to 0.8% accuracy. Crowding can be easily spotted in the light curve and removed. They derive $\mu_{\text{LMCCentre}} = 18.477 \pm 0.004\,(\text{stat}) \pm 0.026\,(\text{sys})$, where the main contribution to the systemic error budget is the S_V relation above.

This is the most accurate measurement of the distance to the LMC to date. As we shall see shortly when we discuss standard candles, this result has become key to

---

[6] We should correctly refer to these systems as "double-lined", as "detached" means physically separate rather than resolvable. However, the literature on H_0 almost always uses "detached" so we stick with it here.

many recent H_0 results: every standard candle can be found in the LMC, and the low error budget allows for a very accurate calibration.

Looking forwards, although DEBs have been found in M31 they are "early-type" stars with hot atmospheres, for which a reliable surface brightness to colour relation has not been established. The hope is future 30 m-class telescopes will have sufficient spectroscopic resolution to extend this to late-types in M31 and other local group galaxies (Beaton et al. 2019).

3.3 Masers

Maser emission occurs when thermal collisions in warm gas in an accretion disk around the central black hole of a galaxy drive a population inversion of molecular energy levels. Such systems are rare: the disk must be "just right", not too hot, and not too cold, and have suitable local molecular abundance. The Type 2 Seyfert galaxy NGC 4258 at a distance of 7.5 Mpc is just such a system. Isolated bright spots of 22.235 GHz maser emission (from a hyperfine transition of H_2O) can be
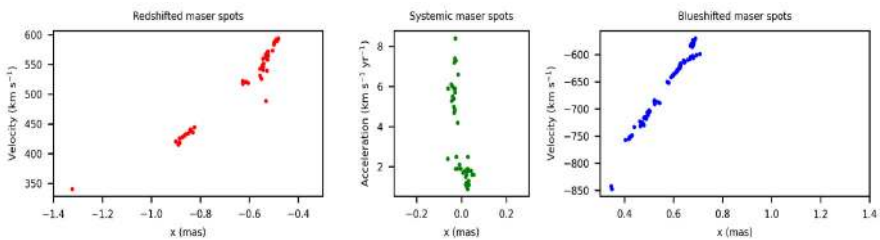
(a) Maser positions, velocities and acceleration
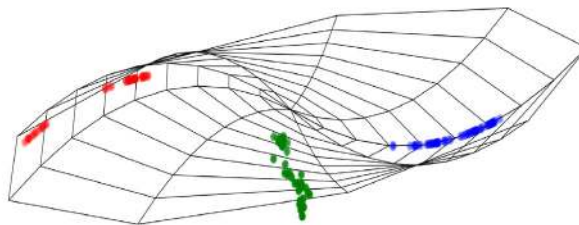
(b) The accretion disk of UGC 3789 (maser spot positions are schematic, not actual).

Fig. 6 The UGC 3789 maser system, showing the warped disk, velocities and acceleration of maser spots. The top left and top right figures show the LOS velocities in kms^{-1} relative to the central blackhole, offset by position in milliarcseconds; the *GM/r* curve shape is apparent. The top centre panel shows the LOS acceleration in kms^{-1} per year, calculated from the velocity drifts over 3 years of observations. In the bottom panel, we schematically show the warped disk and projected spots. UGC 3789 is at a distance of 46 Mpc so 1 mas $\simeq$ 0.24 pc. The disk and maser spot data is available in Reid et al. (2013)

seen in three regions on the sky stretching 20 mas (~ 0.1 ly) long, with the overall shape of a warped line. After subtracting the overall system redshift, the spots on one side are blue-shifted by $\sim 1000 \, \mathrm{kms^{-1}}$, the other side is red-shifted by the same amount, and the spots in the middle are low velocity (Argon et al. 2008). If the spots are observed for long enough, their accelerations and proper motion can be obtained from the steady drift of their Doppler velocities and positions. The central spots show the lowest l.o.s. velocity and highest acceleration, and the outer spots having the highest velocity and lowest l.o.s. acceleration.

The key assumption these observations support is that all the spots form an orbital system with shared parameters, as it is anticipated that disk viscosity will have reduced the orbits to close to circular. The disk shape is then modelled (including such parameters as inclination, warp, residual eccentricity and so forth), fitted and the orbital parameters of the spots derived. The outer dot velocities v show a Keplerian behaviour with radius: $v^2 \propto GM/r$, and as the acceleration is then $\dot{v} \propto GM/r^2$, the physical radius of the system can be determined. The angular diameter distance is then

$$d_A = \frac{v^2}{\dot{v} \times \theta}, \qquad (30)$$

where θ is the angular impact parameter. Humphreys et al. (2013) have observed NGC 4258 for 10 years at 6 monthly intervals using VLBI interferometry, which resolves the relative position of the maser spots to $<3 \, \mu$as accuracy (even the tectonic drift of the ground radio telescopes must be corrected for). The Doppler shifts are measured to within an accuracy of $1 \, \mathrm{kms^{-1}}$, and the central spots accelerate by $\sim 10 \, \mathrm{kms^{-1}yr^{-1}}$. The proper motions $\dot{\theta} \, \mu$ as $\mathrm{yr^{-1}}$ provide additional information as $d_A = v/\dot{\theta}$ Mpc . This pattern fits an accretion disk orbiting a $10^7 \, M_\odot$ black hole.

Reid et al. (2019) derive the distance as $d = 7.576 \pm 0.082(\mathrm{stat}) \pm 0.076(\mathrm{sys})$ Mpc, an accuracy of 1.4%, competitive with the LMC distance error. The main statistical contribution to the error budget is the positional error of the ~ 300 spots.

Only 8 megamaser galaxies have been detected, with the furthest being NGC 6264 at 141 Mpc, and it is unlikely more will now be found at usable distances. We illustrate the data for UGC 3789 from Reid et al. (2013) in Fig. 6a, b. The Megamaser Cosmology project (Pesce et al. 2020) have used six of these galaxies to find $H_0 = 73.9 \pm 3.0 \, \mathrm{kms^{-1} \, Mpc^{-1}}$, independently of distance ladders.

Although it is too close to determine H_0 directly (as its peculiar motion could be a large fraction of its redshift), NGC 4258 has become a key calibrator of the distance ladder, owing to its low error budget. Its particular usefulness is that, unlike the LMC or SMC, it is a fairly typical barred spiral galaxy, similar in morphology and environmental conditions (metallicity, star-formation rate and so forth) to the ones in which Cepheids and Type Ia supernovae are seen at greater distances. Using Cepheid and SN Ia data from Riess et al. (2016), Reid et al. (2019) find $H_0 = 72.0 \pm 1.9 \, \mathrm{kms^{-1} \, Mpc^{-1}}$ using solely NGC 4258 as the geometric calibrator.

3.4 Cepheids

Having discussed calibrators, we can now talk about the "engine room" of distance ladders. Cepheids form two classes, but it is the younger, population I, classical Cepheids which are of interest. These are yellow bright giants and supergiants with masses 4–20 $M_\odot$ and their brightness cycles over a regular period, between days and months, by around 1 magnitude. They are bright, up to 100,000 $L_\odot$, and they can be seen out to 30 Mpc with the HST. Although Milky Way Cepheids had been observed and catalogued since the 18th century, it was first discovered by Henrietta Swan Levitt in 1908 that there was linear relation between the logarithm of their oscillation period and absolute magnitudes, the *Leavitt period-luminosity law* (we use the term *P-L law* for Miras). She had been observing Cepheids in the SMC and LMC, using the Harvard College Observatory telescope, and decided to rank them in order of magnitude. As stars in the LMC will have roughly the same distance from the Earth, the Leavitt law was immediately clear from their apparent magnitudes. In fact, one could say modern extragalactic astrometry began with her groundbreaking discovery.

Fig. 7 Illustration the Leavitt law in four galaxies used for the local distance ladder. The LMC and N4258 are the two main calibrators of Cepheid distances. For the two example SN Ia hosts N4536 and N1015, it is harder to observe the fainter, shorter-period Cepheids. The slope is fixed at −3.26, corresponding to the best estimate global fit in Riess et al. (2016), and magnitudes are m_H^W. We have inverted the normal decreasing magnitude axis used in the literature for presentation purposes
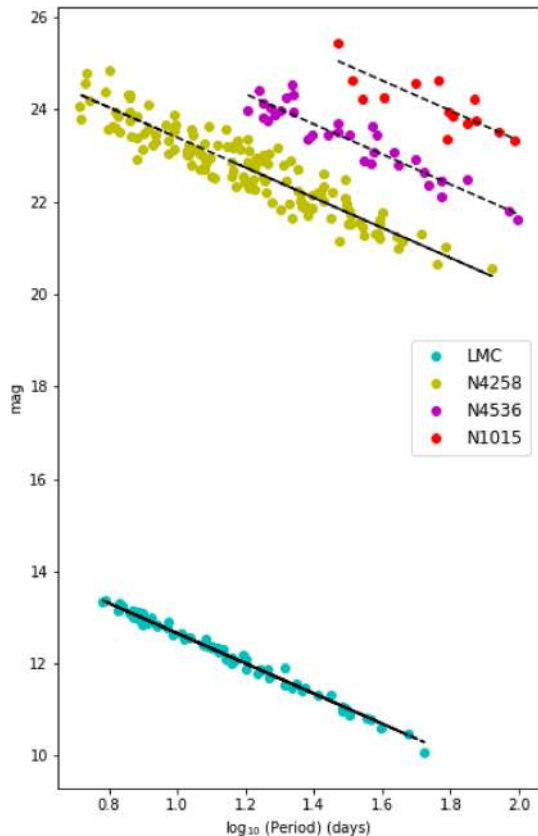
That stars can pulsate is not so surprising; after all, a star is in local equilibrium, and would be expected to oscillate about the equilibrium given any perturbation. However, something must *drive* the oscillation otherwise it would dissipate. For Cepheids, the driver is heat-trapping by an opaque layer of doubly-ionised He surrounding a He-burning core. The trapped heat increases pressure, which expands the star, cooling the ionised He to the point where it can recombine and, therefore, becomes transparent. As the radiation escapes, the core cools and re-contracts, which in turn releases gravitational energy into the He layer. The He heats, re-ionises and the cycle repeats, with period proportional to the energy released. The Cepheid population lies in an instability strip in the horizontal branch of the Hertzsprung–Russell diagram; the cool (red) edge of the population is thought to be due to the onset of convection in the He layer, and the hot (blue) edge by the He ionisation layer being too far into the atmosphere for pulsations to occur. They, therefore, form a well-defined population.

A straightforward understanding of the origin of the Leavitt law can be found in thermodynamic and dynamic arguments. The luminosity of a Cepheid will depend on its surface area via the Stefan–Boltzmann law $L = 4\pi R^2 \sigma T^4$, which expressed in bolometric magnitudes is

$$M = -5 \log_{10} R - 10 \log_{10} T + \text{ const. } . \tag{31}$$

The stellar radius can be mapped to the period by writing an equation of motion for the He layer as

$$m\ddot{r} = 4\pi r^2 p - \frac{GMm}{r^2}, \tag{32}$$

where r denotes the radial position of the layer, p is the pressure and m the mass of the layer. For adiabatic expansion, it is then straightforward to solve for the period P and we find $P\sqrt{\bar{\rho}} = \text{ const. }$ where $\bar{\rho}$ is the mean density (for further details see Cox 1960). With $\bar{\rho} \propto R^{-3}$ and temperature mapped to colour $B - V$, we obtain the Leavitt law as

$$M = \alpha \log_{10} P + \beta(B - V) + \gamma, \tag{33}$$

where P is the period in days, α, β are empirically calibrated from the Cepheid data, and for the zero-point γ we need a distance measurement.

It is preferable to use so-called Wesenheit magnitudes, which are constructed to be reddening-free, and have a reduced colour dependence (intrinsically redder Cepheids are fainter). Let's see how this works in practice. Riess et al. (2019) define an NIR Wesenheit magnitude $m_H^W \equiv m_{\text{F160W}} - 0.386(m_{\text{F555W}} - m_{\text{F814W}})$ using HST filter magnitudes, where the numerical constant is derived from a reddening law. They observed 70 LMC Cepheids with long periods, and after setting $\beta = 0$, they derive $\alpha = -3.26$ with intrinsic scatter 0.075 mag. By comparison, the scatter using solely optical F555W magnitudes is 0.312 mag. After subtracting the DEB distance modulus (Pietrzyński et al. 2019), we obtain $\gamma = -2.579$. The formal error in the Cepheid sample mean is 0.0092 mag, equivalent to 0.42% in distance (Fig. 7).
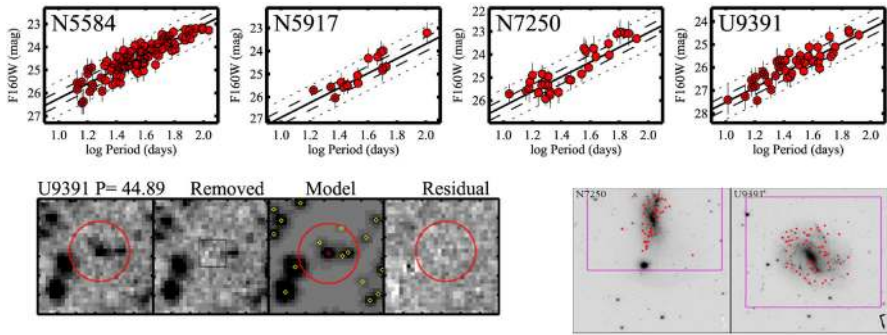
Fig. 8 Illustration of Cepheid photometry from SN Ia host galaxies from Riess et al (Riess et al. 2016). U9391 is the most distant with a distance modulus of 32.92 ± 0.06. The association of Cepheids with spiral arms is clearly visible, and in the bottom left is an example crowding correction for point sources and background flux. The scatter around the Leavitt law is ~ 0.7 mag, thought to be due to residual crowding effects

We would now like to feel that we can deduce the distance to any galaxy we can find Cepheids in, by applying this law to convert periods to absolute magnitudes, and comparing to the apparent magnitudes we observe. As is usual, though, things are not that simple! There are three principle objections:

– **Are Cepheids in the LMC the same as the ones in distant galaxies?** The LMC has a lower metallicity than a typical spiral galaxy, so can we expect the Cepheids found there to have the same brightness? Unfortunately, it is hard to determine the metallicity of a Cepheid from its colour alone, and studies on the effects of metallicity are variable (Ripepi et al. 2020). One might try extending the Leavitt law to add a metallicity term κ [Fe/H]. Riess et al. (2019) have estimated Cepheid metallicity in the LMC based on optical spectra of nearby HII regions, and find $\kappa = -0.17 \pm 0.06$, which is consistent with an earlier estimate from Freedman and Madore (2010) that LMC Cepheids are 0.05 mag dimmer than galactic Cepheids, and later results (Gieren et al. 2018; Breuval et al. 2021).

– **Is the Leavitt law linear?** This matters for H_0, because the Cepheids seen in distant galaxies have longer periods than those used to calibrate the Leavitt law: they are brighter and more easily observed at distance. So, $\log P \sim 1.5$ for Cepheids in SN Ia host galaxies, whereas the average for the LMC is $\log P \sim 0.8$. If the Leavitt law is curved, a bias would be introduced. This can be dealt with by either introducing a separate calibration for short and long period Cepheids, or selecting only long period nearby Cepheids for calibration.

– **Can we obtain clean astrometry of very distant Cepheids?** We want to observe Cepheids as far away as we can, to maximise the overlap with the SN Ia observations. But pushing the limits of resolution brings the risk of crowding, whereby the Cepheid photometry is blended with nearby dimmer, redder and cooler stars. Indeed, this seems to be the main cause of the increased scatter in residuals for distant galaxies. There are some ways to deal with crowding: Riess et al. (2016) (hereafter R16) add random Cepheids to images of the same galaxy

and put them through the same analysis pipeline, to see how well input parameters are recovered. Figure 8 shows an example of this. Cepheids which are outliers in colour, indicative of high blending, may be discarded (removing the colour cut lowers H_0 by 1.1 kms^{-1} Mpc^{-1}). Another way to test for a crowding bias is to look for compression of the relative flux difference from peak to trough (a more blended Cepheid will be more compressed) as is done in Riess et al. (2020). Using an optical Wesenheit magnitude, in which stars may be less crowded than the NIR one, reduces H_0 by 1.7 kms^{-1} Mpc^{-1} (R16), although the argument can be made this is due to higher metallicity effects in the optical (Riess et al. 2019). As with metallicity, the matter of crowding continues to debated.

Systematic error analysis is provided in Sect. 4 and Table 8 of R16, where the effect of some different analytical choices such as breaks in the Leavitt law, different assumed values for the deceleration parameter q_0, and methods of outlier rejection are shown to be ±0.7%. This may not cover all potential systematics. In a recent talk, Efstathiou (2020) re-examined the Leavitt laws of galaxies presented in R16. Calibrating each galaxy individually, the slopes of SN Ia host galaxies are generally shallower than M31 or the LMC, which should not be the case if Cepheids are a single population. A change in the slope will alter the zero-point, as would a change in the distance calibration. It is then noted that forcing the slope of the Leavitt law to -3.3 (the M31 value), in combination with using only NGC4258 as the anchor, lowers the R16 H_0 value to 70.3 ± 1.8 kms^{-1} Mpc^{-1} (Equation 4.2b). Additionally, there appears to be tension between the relative magnitudes of Cepheids in the LMC and NGC4258, and the distance modulus implied by Masers and DEBs; calibrating with solely the LMC gives a higher H_0 value by 4.4 kms^{-1} Mpc^{-1}. It is a feature of χ^2 fits (as used in R16) that the joint solution will be drawn to the data with the lowest error, which in this case is the LMC value. But if two subsets of the data are in tension, it is uncertain that the one with the lowest χ^2 would have the least systematics.

Such analyses do not show one value is preferable over another, nor can they show where any discrepancy may lie—it might be metallicity effects, crowded Cepheid photometry, the NGC4258 distance, or some other systematic. Re-analyses of the results of R16 by various authors (Feeney et al. 2018; Cardona et al. 2017; Zhang et al. 2017; Follin and Knox 2018; Dhawan et al. 2018) use the same Cepheid photometric reduction data so are not independent as such.

Research has accelerated to close down these potential issues. We have already mentioned replacing LMC and NGC4258 Cepheids with Milky Way Cepheids in the section on parallax. In a recent paper, Riess et al. (2020) show crowding effects can be detected as a light curve amplitude compression, and that their correction method has been robust. Finally, Javanmardi et al. (2021) fully re-derive the Cepheid periods and luminosities from the original HST imaging for NGC 5584 (which is face-on to the line of sight), intentionally using different analytical choices, finding no systematic difference in the light curve parameters.

All this said, it would be clearly preferable if we had some other candles to check against Cepheids. We now turn to two possibilities, Miras and the Tip of the Red Giant Branch.

3.5 Miras

Miras are variable stars that have reached the tip of the Asymptotic Giant Branch (AGB), comprising an inert C-O core, and a He-burning shell inside a H-burning shell. Their mass is 0.8–8 $M_\odot$, although they are typically at the lower end of this range. Their large size of ~ 1 a.u. means they are actually brighter by 2–3 mag than Cepheids in the NIR. That they follow a *P-L* law was first established in 1928 (Gerasimovic 1928), but being tricky to categorise and observe, were not extensively studied until the demand for cross-checks on Cepheids has brought renewed interest in them.

Miras form two distinct populations with O-rich and C-rich spectra, with the O-rich ones exhibiting somewhat less scatter than the C-rich (Feast 2004), but more than Cepheids. Miras have very long periods $90 < P < 3000$ days and their light curves have many peaks of variable amplitude superimposed (see for example Fig. 3 in Yuan et al. 2017), so investment in observation time is needed to reliably determine P. The *P-L* law curves upwards at around $P \sim 400$ days, where it seems likely the extra luminosity has been due to episodes of Hot-Bottom-Burning, so an extra quadratic term is required. Miras have high mass loss rates, and the surrounding dust cloud means Wesenheit magnitudes will be less reliable at subtracting reddening compared to Cepheids (because a standard reddening law is assumed in constructing them). Lastly, a size of 1 a.u. would mean their angular diameter is comparable to their parallax, and in addition the photocentre moves around the star, making parallax measurements challenging.

So why bother with them? Their advantages versus Cepheids is that they are (a) more numerous, and, therefore, easier to find in halos where there is less crowding (b) older, so can be found in all types of galaxies including SN Ia hosts with no Cepheids (c) $2 - 3$ magnitudes brighter than Cepheids in the near infrared. Given imaginative observation strategies and careful population analysis, the issues above can be addressed, and modern studies now exist for Miras in the LMC (Yuan et al. 2017), NGC 4258 (Huang et al. 2018), and the SN Ia host NGC 1559 (Huang et al. 2020).

In the first of the above, Yuan et al. (2017) establish a *P-L* curve using 600 Miras in the LMC. They have sparse *JHK*-band observations from the LMC Near Infrared Synoptic Survey (LMCNISS, Macri et al. 2015), which on their own would not be enough to establish the period, but using much denser Optical Gravitational Lens Experiment (OGLE) I-band observations (Szymański et al. 2011), they are able to establish a regression rule for the relationship between passbands, and use OGLE to "fill in" the light curves. The classification of Miras into O-rich or C-rich is also obtained from OGLE. The reference magnitude is defined as the median of maxima
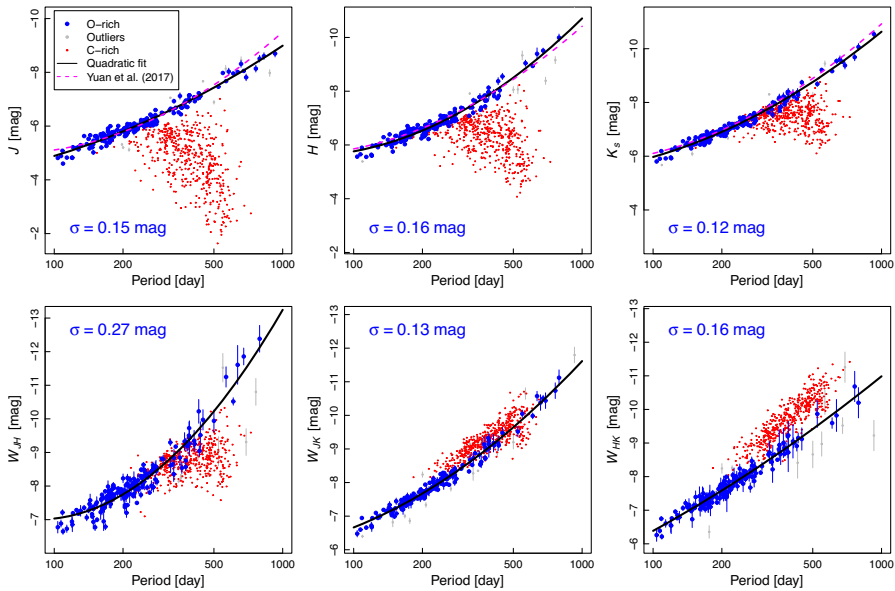
Fig. 9 A PL calibration for Miras in the LMC. While non-linearity can straightforwardly be fitted, contamination must be kept under control. As seen in the bottom three panels, Wesenheit magnitudes have more scatter compared to standard magnitudes (the opposite of applying them to Cepheids), and they also hinder separating O-rich from C-rich stars. Image reproduced with permission from Yuan et al. (2017), copyright by AAS

and minima of the fitted light curves. The period is obtained by fitting to a sum of sine functions, progressively adding harmonics if supported by a Bayesian Information Criterion.[7] They fit a law[8]

$$M = a_0 + a_1(\log_{10} P - 2.3) + a_2(\log_{10} P - 2.3)^2. \qquad (34)$$

The K-band for O-rich Miras shows the lowest scatter of 0.12 mag, with $a_0 = -6.90 \pm 0.01$, $a_1 = -3.77 \pm 0.08$, $a_2 = -2.23 \pm 0.2$. The zero-point is obtained using the DEB distance to the LMC given by Pietrzyński et al. (2013). By comparison, the m_H^W scatter for LMC Cepheids obtained by SH0ES was 0.075 mag (Riess et al. 2019) (Fig. 9).

Huang et al. (2018) observed NGC4258 with the HST WFC in 12 epochs over 10 months. Mira candidates were identified by their V and I amplitude variation, and it was possible to use LMC data to show that using this method contamination between O-rich and C-rich could be kept to a minimum. Their "Gold" sample size was 161 Miras, and fitting Eq. (34) for apparent magnitudes gave $a_0 = 23.24 \pm 0.01$ for F160W with scatter 0.12. Adjusting the ground-based photometry of LMCNISS to F160W equivalent, the authors calculate the relative distance modulus

---

[7] Whilst there are well known issues with the Bayesian Information Criteria, mainly an unnecessary penalty to fits with a large a number of data points, Yuan et al. (2017) uses this statistic, and it is beyond the scope of this review to repeat their analysis using a different statistic.

[8] It is good practice to center around the mean P to reduce correlation between coefficients.
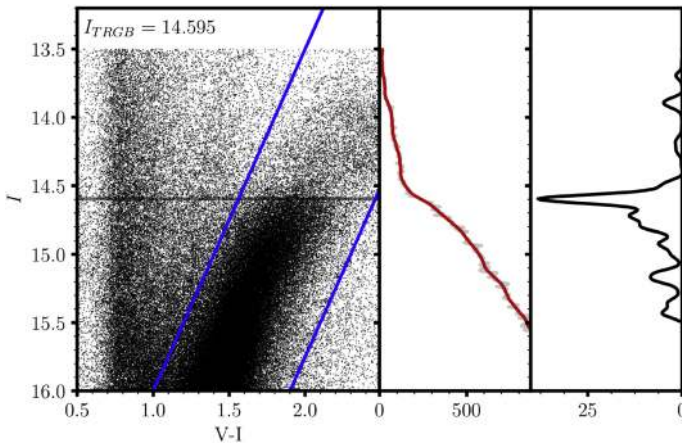
Fig. 10 Fitting the TRGB edge from the colour-magnitude diagram of Freedman et al. (2019). The middle panel shows the horizontally-summed (between the blue lines) and smoothed number count, from which the edge is determined by the maximum (vertical) gradient, which is the peak in the right-hand panel. The background AGB population is visible in the stars above the tip

between the LMC and NGC4258 of $\Delta\mu = 10.95 \pm 0.01\text{(stat)} \pm 0.06\text{(sys)}$. This is consistent with the Cepheid value of $\Delta\mu = 10.92 \pm 0.02$ (Riess et al. 2016) and the Maser-DEB value of $\Delta\mu = 10.92 \pm 0.041$ (Reid et al. 2019; Pietrzyński et al. 2019). Similar methods are used by Huang et al. (2020) for observations of NGC 1559, with the addition of a low period cut to deal with potential incompleteness bias of fainter Miras. Using the LMC DEB and NGC 4258 distances as calibrators, they obtain $H_0 = 73.3 \pm 4.0$ kms^{-1} Mpc^{-1}. So Miras are consistent with Cepheids, but due to their larger error budgets, they are also consistent with Planck results.

3.6 Tip of the Red Giant branch

Red Giant branch stars are stars of mass $\sim 0.5-2.0 M_\odot$ that are at a late stage in their evolution, having moved off the main sequence towards lower temperatures and higher luminosities. The core of the star is degenerate He, and as He ash rains down from the H-burning shell surrounding it, it contracts and heats up. Since the temperature range for He fusion ignition is rather narrow, and the degenerate matter means there is a strong link between core mass and temperature, the core in effect forms a standard candle inside the star, just prior to ignition. As soon as the helium flash ignites, the star moves over the course of 1 My or so (almost instantaneously in stellar evolution terms) to the red clump at a higher temperature and somewhat lower luminosity. Therefore, there is a well-defined edge in a colour-magnitude diagram that can be used to identify the tip of the red giant population just before the
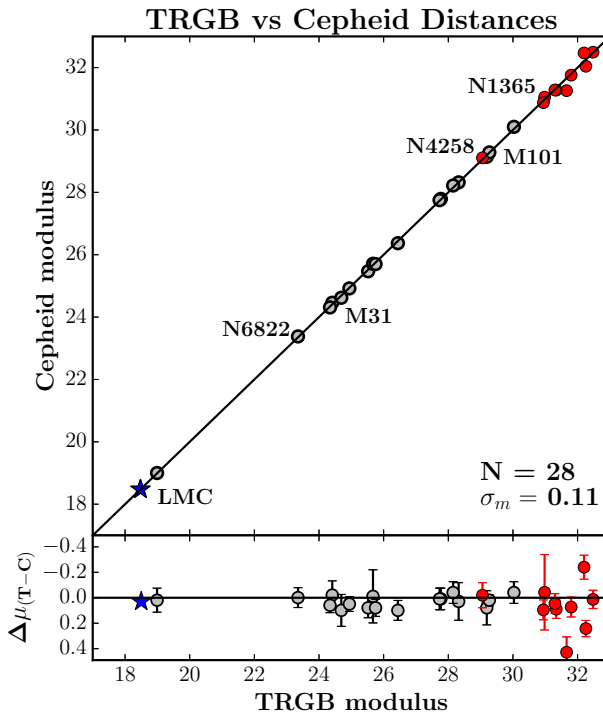
Fig. 11 Comparison of TRGB and Cepheid distance moduli, from which it can be seen TRGB distances are consistently larger across a range of galaxies. The red dots are SN Ia host galaxies, the grey dots other galaxies, and the blue star is the LMC TRGB calibrator. Image reproduced with permission from Freedman et al. (2019), copyright by AAS

flash (TRGB). The TRGB is then not one single star, but a statistical average for a suitably large population.

The TRGB offers many advantages compared to other standard candles. Although red giants are fainter than Cepheids in the optical, in the K-band they are ~ 1.6 magnitudes brighter. Stars at or near the TRGB are relatively common, and can be observed readily in uncrowded and dust-free galactic halos. They are also abundant in the solar neighbourhood, meaning great numbers are available for calibration by parallax (by contrast, Cepheids with good parallaxes in Gaia DR3 will number in the hundreds). Sufficient numbers to resolve the tip can be seen in globular clusters such as ω Centuari, where well-formed colour-magnitude diagrams can resolve metallicity and extinction effects, and an absolute magnitude calibration can be made either from Gaia DR3 parallaxes (Soltis et al. 2021) or DEB distances (Cerny et al. 2021).

Like Miras, they are found in all galactic types. Observation is efficient: one does not need to revisit fields to determine periods. The James Webb Space Telescope is capable of observing red giants in the near IR at distances of ~ 50 Mpc, which is comparable to Cepheid distances obtainable from the HST.

It is also beneficial that the modelling of late-stage stellar evolution is quite well understood. Empirical calibrations of absolute magnitude and residual dependence on metallicity, mass and age can be checked against the results of stellar codes. For example, the dependency on metallicity is somewhat complex: the presence of metals dims optical passbands by absorption in stellar atmospheres, and brightens the NIR. Serenelli et al. (2017) express the TRGB as a curve with linear and quadratic terms in colour. The authors find an I-band median $M_I^{\text{TRGB}} = -4.07$ with a variation with colour of ± 0.08, and a downward slope in the colour-magnitude diagram as metallicity increases. This median value is consistent with empirical calibrations.

Fitting the TRGB is equivalent to finding the mode of the gradient of the (noisy) stellar distribution in the colour-magnitude diagram, and techniques are borrowed from image processing to do this. The tip has a background of AGB stars, and the edge is sensitive to the distribution of them close to it. If the contamination is large or variable, there is a risk the mode can shift. It is, therefore, important to have a dense population of stars, and enough fields to test the robustness of the fitting process. For example, Hoyt et al. (2021) bin data by separation from the galactic centre and fit the tip separately for each bin.

The Carnegie-Chicago Hubble program (CCHP) have determined TRGB distances for 10 SN Ia host galaxies that also have Cepheid distances, using HST I-band imaging (Freedman et al. 2019). They calibrate the zero point from the TRGB of the LMC and SMC using the DEB distance, finding $M_I = -4.049 \pm 0.022$ (stat) ± 0.039 (sys) , which is consistent with the theoretical value above (Fig. 10).

We show in Fig. 11 their comparison of TRGB and Cepheid distances, covering a range of distances from 7 Mpc to almost 20 Mpc. By expanding their sample to non-SN Ia host galaxies, the authors show the TRGB has a lower scatter versus their Hubble diagram than do Cepheids, by a factor of 1.4. Calibrating the Pantheon SN Ia sample on TRGB distances alone gives $H_0 = 70.4 \pm 1.4 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$, with the difference to SH0ES results due to the TRGB distances to the SN Ia host galaxies being further than Cepheid distances.

Currently, this result remains the subject of debate. Yuan et al. (2019) find $M_I = -3.99$, which gives $H_0 = 72.4 \pm 2.0 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$. They attribute their result to a different methodology to determine the LMC extinction. Freedman et al. 2019 determine the extinction applicable to the TRGB by comparison of LMC and SMC photometry in $VIJHK$ bands, and using the lower SMC extinction as a reference point. Yuan et al. 2019 adopt the standard reddening law of Fitzpatrick 1999). They also revise the blending corrections of the older, lower resolution photometry of the SMC. Conversely, a calibration of the TRGB in the outskirts of NGC 4258 to the maser distance by Jang et al. (2021) gives a value fully consistent with Freedman et al. (2019). In a followup paper, Madore and Freedman (2020) present an upgraded methodology in which stellar photometry is itself used to separate metallicity and extinction effects (by exploiting their different sensitivities in J, H and K-band magnitudes), confirming their earlier results. Other results (Soltis et al. 2021; Reid et al. 2019; Capozzi and Raffelt 2020; Cerny et al. 2021—see for

example Table 3 in Blakeslee et al. 2021) cluster evenly around these two values. We also note the difference between SN Ia zero points (Fig. 6 of Freedman et al. 2019) of the TRGB vs Cepheids appears to increase with distance, and is particularly large for one galaxy, N4308.

More recently, an updated value of $H_0 = 69.8 \pm 0.6$ (stat) $\pm$ 1.6 (sys) kms^{-1} Mpc^{-1} was published by Freedman (2021). New and updated values for M_I based on the anchors of the LMC, SMC, NGC 4258 and galactic globular clusters are derived. Additionally, potential zero-point systematic problems with the direct application of Gaia EDR3 data to globular clusters (as done by Soltis et al. 2021) are discussed. Each anchor is consistent with each other (although the closeness of the agreement suggests the errors may be over-estimated), and the paper brings the TRGB method on par with Cepheids in terms of number of anchors used.

The TRGB continues to attract attention because the CCHP result is very interesting: it is a late universe distance ladder that is in reduced 2σ tension with the Planck value of $H_0 = 67.4 \pm 0.5$ kms^{-1} Mpc^{-1}. While it cannot by itself fully resolve the Hubble tension, in combination with some other systematic—perhaps in the calibration of SN Ia (see next section)—it may offer a resolution of it that does not involve new physics.

What is also clear is that quality of a distance ladder is less a matter of *quantity* of objects, but rather the *accuracy* of calibration and *control* of systematics. The TRGB seems very promising, however. Once the calibration is agreed among the community, they offer the tantalising prospect of either confirming the H_0 tension through two independent data sets, or reducing it to a lower statistical significance.

3.7 Type Ia supernovae

It is hard to overstate the impact of Type Ia supernovae in cosmology. At $M \sim -19$, they are both bright enough to be seen well into the Hubble flow (the furthest to date is SN Wilson at $z = 1.914$), and have a standardisable luminosity. The mainstream view is that they originate from the accretion of material from a binary companion onto a carbon-oxygen white dwarf. When the white dwarf reaches the maximum mass that can be sustained by degeneracy pressure, the Chandrasekhar mass $M_{ch} = 1.4M_\odot$, a runaway fusion detonation occurs destroying the white dwarf. Approximately $\sim 0.6M_\odot$ fuses into heavier elements during the first few seconds of the explosion, and the observed light curve, which peaks at ~ 20 days and lasts for ~ 60 days is powered by radioactive decay of ^{56}Ni to ^{56}Co and then to ^{56}Fe. Hence, a SN Ia is a "standard bomb", primed to explode when it reaches critical mass.

However, there is surprisingly little observational evidence to confirm the accretion theory. Both the lack of H lines and computer modelling suggest the donor star should survive the explosion, but a search of the site of the nearby SN2011fe both pre- and post- explosion have revealed no evidence of a companion. This may support an alternative "double degenerate" theory that some, or all, SN Ia originate from the merger of two white dwarves (for a review, see Maoz et al. 2014). But if

Fig. 12 Pantheon SN Ia distance moduli and residuals versus a best fit ΛCDM cosmology. Image reproduced with permission from Scolnic et al. (2018), copyright by AAS
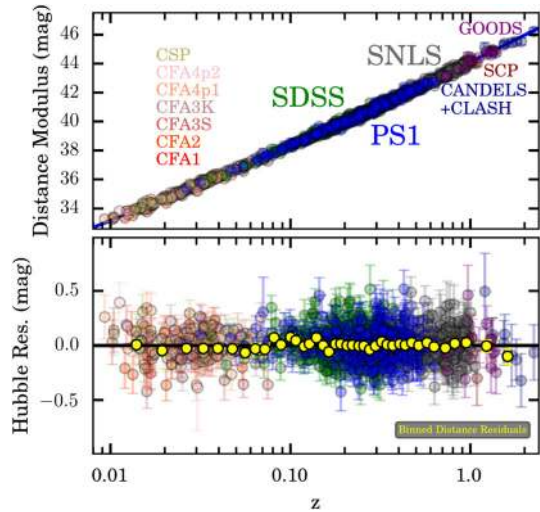
that is the case, why are their luminosities so uniform? Whether SN Ia have one or two types of progenitors has important implications for the Hubble constant, which we return to shortly.

Individual SN Ia luminosities can vary by up to a factor of 2, but they are empirically standardisable by the Tripp estimator (Tripp and Branch 1999):

$$\mu = m_b - M_{\text{fid}} + \alpha x - \beta c + \Delta_M + \Delta_B, \tag{35}$$

where μ is the distance modulus, m_b ia the apparent AB magnitude, and M_{fid} is the absolute magnitude of a typical SN Ia. The parameter x is the "stretch" of the light curve, which is a dimensionless measure of how long the bright peak of the light curve lasts: longer duration SN Ia are brighter. c is a measure of colour: redder SN Ia are dimmer. Δ_M is a correction for the environment of the SN Ia, and Δ_B a statistical bias correction for selection effects, analogous to the LKH bias of parallax we described earlier. M_{fid} can be calibrated by finding SN Ia in galaxies for which there are Cepheid or TRGB luminosity distances, or in the inverse distance ladder approach by combination with angular diameter distances from baryon acoustic oscillations, or the CMB.

SN Ia are rare events: approximately 1 per galaxy per century, but by scanning a reasonably sized patch of sky, supernova surveys can detect hundreds per year.[9] Hence, supernovae catalogs are not uniformly distributed on the sky. However, SN Ia close enough to calibrate their magnitudes are few: there are just 19 SN Ia in galaxies with Cepheid distances from HST observations, although more are expected soon from new cycles (see for example HST proposal 16,198).

To maximise statistics and sky coverage, it is common to use SN Ia from multiple surveys in cosmological analyses. Scolnic et al. (2018) have compiled the Pantheon

---

[9] Supernovae that are photometric candidates to be SN Ia outnumber those that are confirmed by a factor of 10; the limitation is the availability of spectroscopy.

sample of 1,048 SN Ia from CfA, CSP, PS1, SDSS, SNLS and HST surveys in a uniform light curve calibration, representing almost 40 years of observational data. The sample divides into 180 mostly older Low-z ($z < 0.1$) SN Ia which are uniformly distributed on the sky, and 868 recent High-z SN Ia concentrated in the survey fields, notably the thin "Stripe 82" along the celestial equator. The sets have a small overlap at $z \sim 0.1$, and cross-checks on their spectral and other parameters show they are likely to be from the same underlying population (that is, High-z SN Ia are not intrinsically different to Low-z). However, the mean values of light curve parameters like stretch and colour do drift with z, possibly due to selection effects or changes in the host galaxy properties at higher redshift. Large numbers of Type Ia supernova have also been found, or are in the process of being surveyed, by the Dark Energy Survey (Abbott et al. 2019), All-Sky Automated Survey for Supernovae (Holoien et al. 2018), Foundation Survey (Jones et al. 2019), and Zwicky Transient Facility (Yao et al. 2019).

The process of fitting a SN Ia lightcurve is straightforward. *ugriz* photometry is corrected for the background sky obtained from imaging the site after the nova has faded (a bonus of SN Ia being transient), and the light curve is fitted to a template, outputting the parameters x, c and m_b. However, the bias correction Δ_B is challenging and subtle: it depends on the astrophysical modelling of sources of scatter and contamination by mis-classified supernovae, and the selection function of survey. For example, more distant SN Ia may be preferentially targeted for spectroscopic confirmation when their host galaxy is faint, so if their luminosities do depend on the galactic environment, this could bias the High-z population relative the to Low-z. The selection function of early Low-z surveys are not easy to model, and their biases can be up to 0.06 magnitudes. This could account for a somewhat elevated scatter of Pantheon residuals around $z \sim 0.1$ in the Hubble diagram where the two sets join (Fig. 12). After fitting α and β, the residual intrinsic scatter of the SN Ia population is $\sigma_{int} \sim 0.1$ magnitudes, which compares favourably with other standard candles.

We return now to the question of whether SN Ia have additional environmental correlations that affect their luminosity (that is, what the nature of the Δ_M term in Tripp estimator is), and why this may matter for H_0. It is generally accepted that SN Ia in galaxies with stellar mass $M_* > 10^{10} M_\odot$ are intrinsically brighter by $\Delta_{M,*} \sim 0.04 - 0.06$ than those in lower mass galaxies, and that the transition is sharp, rather than a gradual evolution (see Smith et al. 2020 and references therein). It has, therefore, become common to use a step-function for $\Delta_{M,*}$ in fits for H_0. Riess et al. (2016) calculate the effect is to lower H_0 by 0.7%, as the mass of the calibrator galaxies is slightly lower than the mass of the Hubble flow set. The key point is then this: given that Cepheids are young, bright stars formed in active galaxies, is $\Delta_{M,*}$ sufficiently discriminating to ensure the 19 supernovae in galaxies with Cepheid distances are representative of the thousands in the Hubble flow?

Host galaxy mass itself cannot matter to an individual SN Ia, so it must be a proxy for some other environmental condition. Metallicity does vary with galactic mass, but the step-like feature suggests a connection with star formation: galaxies with mass below the threshold tend to be active, whereas those above are mostly

passive. One hypothesis is that there is a prompt/bright SN Ia type that is continuously renewed by star formation, and an delayed/dimmer type originating from stars formed when the galaxy was young (Rigault et al. 2015, 2013; Maoz et al. 2014). But any link between this and SN Ia formed by accretion onto or mergers of white dwarfs is speculative. Also, with estimates of "global" properties being effectively light-weighted, they are biased to galactic centres rather than the outskirts where most supernovae are seen.

If some SN Ia are associated with sites of star-formation, we can expect a better proxy to be local specific star formation rates (that is, star formation normalised by local stellar mass or LsSFR for short), on the grounds that the younger SN Ia population will not have had enough time to disperse from their birth region. Because the pixel resolution of typical SN Ia surveys can resolve a 2 kpc aperture only out to $z \sim 0.1$, studies concentrate on the Low-z sample. Star formation can be estimated from photometry, but is best tracked by H_α lines from ionised gas or UV imaging, when available.

Rigault et al. (2020) have examined 141 SN Ia from the Nearby Supernova Factory sample (which has high resolution spectroscopy) and find a significant correlation between LsSFR and α, the stretch slope, and no correlation for β. This suggests there is indeed an intrinsic difference between SN Ia originating in active areas versus passive ones. The size of the LsSFR step is $\Delta_{M,LsSFR} = 0.163 \pm 0.029$ mag, and including it in the Tripp luminosity estimator eliminates the need for a further global host mass step. The size of the step is interesting as it is significantly larger than $\Delta_{M,*}$, and because most SN Ia in the Cepheid calibrators are in regions of high LsSFR, whereas in the rest of the Low-z sample the fraction is about a half. Hence, a potential bias to H_0 could be as much as 3%. In contrary results, Jones et al. (2018) find a local stellar mass step the same size as $\Delta_{M,*}$ and little effect on H_0, and Riess et al. (2016) found just $0.1 \, \mathrm{km s^{-1} \, Mpc^{-1}}$ difference in their H_0 calculation after restricting to SN Ia found in spiral galaxies. In a recent paper, Brout and Scolnic (2021) argue the true factor is host galaxy extinction (correlated to star formation and hence mass), and present evidence of a distribution separated by colour into blue "clean" SN Ia with lower scatter than their red "dusty" counterparts. The host mass step is absent in the blue sample, and they suggest to use solely these in cosmological analyses to avoid bias (another possibility is to use near-infrared photometry; Dhawan et al. 2018).

In summary, while there is strong evidence SN Ia depend on local specific star formation rates, there is not yet a consensus on the cause, whether they constitute one or two populations, and the level of bias (if any) to H_0. They are numerous and, therefore, high *relative* precision beyond $z > 0.1$, and have been dubbed the "guard rails" of the Hubble diagram. However, their *absolute* accuracy in distance ladders is limited by the small numbers available to calibrate their luminosities, and some uncertainty about the underlying astrophysics. Work is underway to address both of these issues: for example, the Foundation and Zwicky Transient Facility surveys (Jones et al. 2019; Yao et al. 2019) are targeting a new high-quality Low-z set with local spectroscopy to test environmental effects or extensions to the Tripp estimator. New HST observation cycles aim to extend the number of Cepheid calibrators up

from 19, and the JWST will access a greater volume, perhaps increasing the calibration set beyond a hundred for both TRGBs and Cepheids. This would make a 1% calibration from at least two independent sources achievable.

3.8 Time delay of lensed quasars

Quasars are bright and can be variable on short timescales. Some are seen to be gravitationally lensed by a foreground galaxy, and each image will have a time delay caused by the extra path length and the gravitational time dilation of general relativity:

$$\Delta t_i = D_{\Delta t}\left[\frac{1}{2}(\theta_{\mathbf{i}} - \beta)^2 - \psi(\theta_{\mathbf{i}})\right], \tag{36}$$

where $\theta_{\mathbf{i}}$ is the sky position of the image, β the (unknown) source position, and $\psi(\theta_{\mathbf{i}})$ the potential of the lensing galaxy integrated along the line of sight. The relevant quantity for H_0 in the above equation is

$$D_{\Delta t} \equiv (1 + z_l)\frac{D_l D_s}{D_{ls}} \propto H_0^{-1}, \tag{37}$$

where D_l, D_s, D_{ls} is the angular diameter distance to the lens, source, and between the lens and the source, respectively, as specified by Equations (11, 14)—in the case of D_{ls} the integration would be done between the redshifts of the lens and source. Then, if we have a system with multiple images, the relative time delay $\Delta t_{ij} = \Delta t_i - \Delta t_j$ of the quasar variability between the images scales with H_0^{-1}. The proportionality follows from Eqs. (11, 14) with a weak dependence on other cosmological parameters. Acquiring the time delays Δt_i requires extensive observations: relative delays range from weeks to a few months. Errors are typically $\sim 6\%$, although these are reducing over time. A typical lensing system may have $z_s \sim 1.5$ and $z_l \sim 0.5$, so we are probing intermediate cosmological distances.

Each single lens system may provide an independent estimate of H_0. But to do so, we need an idea of the lensing potential $\psi(\theta_i)$ above. There is also a tricky degeneracy to deal with, known as the *mass sheet transformation*. For our purposes, it is sufficient to understand that the same observed time delays may be produced by simultaneously scaling the line-of-sight surface density $\Sigma(\theta)$ of the lensing system by a constant factor $\Sigma(\theta) \to \epsilon\Sigma(\theta)$ (the "mass-sheet") and $H_0 \to \epsilon H_0$. For a clear account of the details, see Falco et al. (1985).

The ingredients for modelling are then this: (a) a range of plausible mass profiles for the lensing galaxy with associated nuisance parameters to be marginalised over (b) a way to deal with the mass sheet degeneracy, through independent constraints on the mass distribution of the lensing galaxy and along the line-of-sight.

This is a challenging, time-consuming and model dependent process for each system! Some other choices must be made, such as which of the galaxies along the line of sight will be modelled and which will be absorbed into the line of sight mass average (Chen et al. 2019). Galactic stellar velocity dispersions provide information on the lensing potential, but as a function of which galactic mass profile is assumed,

and with an anistropy nuisance parameter. Also, the accuracy of velocity dispersions at such distances are $\sim 10\%$. If there are visible background galaxies, they may also be lensed and the shape distortions provide extra information on the mass profile. The line of sight mass may be estimated from the density of foreground galaxies, by searching in computer simulations of large scale structure for similar lines of sight, and calculating the line of sight density from them (lenses are associated with over-densities and, therefore, focusing is more likely). But a major advantage is that this modelling can be done blind to final value of H_0, eliminating the risk of confirmation bias on the part of the experimenter.

The H0LiCOW team have analysed six such systems so far (Wong et al. 2020), and we show their results below. The main contributions to the errors are uncertainty in the time delay, and velocity dispersion of the lens. H0LiCOW use flat priors of $\Omega_M \in (0.05, 0.5)$ and $H_0 \in (0, 150)$ kms^{-1} Mpc^{-1}, to avoid any depen-dence on other cosmological datasets (although they do assume a flat universe). The results are:

Lens name	$D_{\Delta t}$ (Mpc)	D_l (Mpc)	H_0 (kms^{-1} Mpc^{-1})
PG1115+080	1470^{+137}_{-127}	697^{+186}_{-144}	$81.1^{+8.0}_{-7.1}$
RXJ1131-1231	2096^{+98}_{-83}	804^{+141}_{-112}	$78.2^{+3.4}_{-3.4}$
HE0435-1223	2707^{+183}_{-168}	-	$71.7^{+4.8}_{-4.5}$
WFI2033-4723	4784^{+399}_{-248}	-	$71.6^{+3.8}_{-4.9}$
B1608+656	5156^{+296}_{-236}	1228^{+177}_{-151}	$71.0^{+2.9}_{-3.3}$
SDSS1206+4332	5769^{+589}_{-471}	1805^{+555}_{-398}	$68.9^{+5.4}_{-5.1}$

Their combined result is $H_0 = 73.3 \pm 1.8$ kms^{-1} Mpc^{-1}, a final error budget of 2.4%. However, a key assumption in the combination is independence, and there are many modelling features in common between the lenses, in particular the specific mass profiles used. There is also an apparent drift in H_0 values with lensing distance, although the statistical significance is not strong enough to say this is not just random chance.

A re-analysis has been undertaken by the TD Cosmo team (Birrer et al. 2020). By treating the mass sheet degeneracy as a fully unknown parameter at the population level, their analysis reduces reliance on specific mass profiles. For same set of systems, they find $H_0 = 74.5^{+5.6}_{-6.1}$ kms^{-1} Mpc^{-1}. The central value is consistent with H0LiCOW, but the wider confidence intervals reflect the greater freedom allowed on lens mass distribution, combined with the small number of time-delay lenses. To re-narrow the uncertainty range, the paper investigates what effect introducing a set of non-time delay lensing systems from the SDSS survey may have. The purpose of this larger set is to use their imaging to constrain the mass-profile of lensings galaxy, on the assumption they are drawn from the same parent population as the H0LiCOW sample. TD Cosmo then find $H_0 = 67.4^{+4.1}_{-3.2}$ kms^{-1} Mpc^{-1}. However,

the H0LiCOW sample is selected as those systems capable of producing a clear time-delay signal, which may not have the same characteristics as those selected as having clean shear images. So the assumption they share the same mass profile as each other may not be valid.

It is, therefore, premature to claim lensed quasars are consistent with either SH0ES or Planck. More data is needed: upcoming surveys by the Vera Rubin Observatory, Euclid and Nancy Grace Roman telescope will result in several hundred new time delay systems being discovered (so improvements in analysis efficiency will certainly be needed!). Adaptive optics can provide spatially resolved velocity dispersions to improve the mass models. TD Cosmo estimate 200 extra systems will be needed to constrain H_0 within 1.2%, which may be within reach in the next five years.

3.9 Gravitational waves

The gravitational wave signal emitted by the merger of two compact objects can be used as a self-calibrating standard candle. There are now operational detectors at LIGO Hanford and LIGO Livingston in the USA, Virgo in Italy, and KAGRA in Japan, with a further planned LIGO India (Abbott et al. 2020). The detectors measure the strain amplitude of a gravitational wave using laser interferometry to detect the miniscule changes in the length of perpendicular beams as a wave passes by. The purpose of the two sites in the USA is to filter out local seismic vibrations. The wave amplitude is related to the *chirp mass* $\mathcal{M}$ which is in turn derivable from the waveform calculated for a merger. A simplified form of the relevant equations are

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \tag{38}$$

$$= \frac{1}{G} \left[\frac{5}{96} \pi^{-8/3} f^{-11/3} \dot{f} \right]^{3/5} \tag{39}$$

$$\mathcal{M}_z = (1 + z_{\text{obs}}) \mathcal{M} \tag{40}$$

$$h(t) = \frac{\mathcal{M}_z^{5/3} f(t)^{2/3}}{d_L} F(\theta, i) \cos \Phi(t) , \tag{41}$$

where f is the frequency, m_1 and m_2 the merging masses, $\Phi(t)$ the phase, and $h(t)$ the measured dimensionless strain of the strongest harmonic (Abbott et al. 2016). The rest-frame chirp mass is redshifted by z_{obs}, and F is a function of the angle between the sky position of the source and detector arms, and the inclination i between the binary orbital plane and line of sight (Arun et al. 2009). Note that as we are measuring amplitude rather than energy flux, $h(t) \propto d_L^{-1}$.

In a loose sense, every wave cycle is a measurement of the chirp mass $\mathcal{M}$ as it sets $f, \dot{f}$, although in practice the full waveform is fitted. The relative precision is
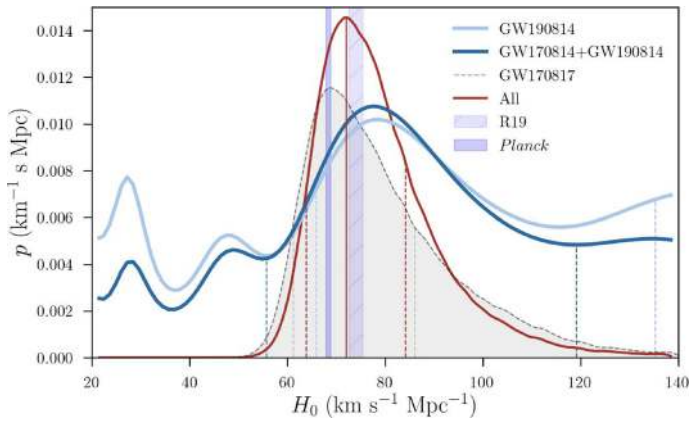
Fig. 13 H_0 posterior distribution for the events GW190814 and GW170814, which was localised to a region in the DES survey footprint containing ~ 1800 possible host galaxies. The light and dark blue lines represent the dark approach for each siren, which can be compared to the gray dashed line for GW170814 which was localised to one galaxy by its electromagnetic counterpart. The estimate for this bright standard siren can be improved by the addition of the dark sirens, as shown in dark red line. The posteriors of Planck (Planck Collaboration 2020) and SH0ES (Riess et al. 2019) are shown in purple boxes as a guide. Image reproduced with permission from Palmese et al. (2020), copyright by AAS

floored by sensitivity of the strain measurement, currently around 5%. If the redshift and angles were known, d_L and hence H_0 would be determined to the same precision.

When a binary neutron star (BNS) system merges, there is an accompanying burst of light from matter outside the combined event horizon. For this reason, it is known as a "bright siren". If the flash can be observed, the host galaxy is identified and one can use its redshift in Eq. (41). The event GW170817 was just such a BNS merger (Abbott et al. 2017b). The gravitational wave was measured in Hanford and Livingston, which was enough to locate the sky position to 28 $\deg^2$ (see next paragraph for how). Given the search region, an optical counterpart was found in NGC 4993 at a distance of ~ 40 Mpc.[10] Around 3000 cycles of the wave resolved the chirp mass in the detector frame as $\mathcal{M} = 1.197 M_\odot$ to accuracy of 1 part in 10^3, consistent with a BNS merger. The main remaining uncertainty is then the inclination angle i. Using a flat prior for $\cos i$, Abbott et al. (2017a) obtain $H_0 = 70.0^{+12}_{-8}$ kms^{-1} Mpc^{-1}. Hotokezaka et al. (2019) improved this to $H_0 = 70.3^{+5.3}_{-5.0}$ kms^{-1} Mpc^{-1} by constraining i with observations of the interaction of the ejecta with the ISM.

For black hole mergers, no optical counterpart is generated, and these are called "dark sirens". However, it is still possible to constrain H_0 using them if a probable redshift can be estimated. To do this, the relative amplitudes and time delay between detectors located around the globe are used to approximately determine the sky

---

[10] As the host galaxy was relatively close, the peculiar velocity adjustment was significant at $\sim 10\%$ of the Hubble flow velocity and the adjustment needs some care (Mukherjee et al. 2021; Howlett and Davis 2020; Nicolaou et al. 2020).

position (Soares-Santos et al. 2019). For example, if 3 detectors observe the wave with perfect accuracy, the two independent time delays and three measured amplitudes will in turn determine the two sky position angles, two polarisation amplitudes, and one phase lag between polarisations. Hence, simultaneous detections of an event are essential in narrowing the size of the region on the sky where the source is located. A reasonable prior for H_0 will constrain the redshift range, and hence determine a localisation volume. Given a suitably complete galaxy catalog with sufficient sky coverage (the best available currently being SDSS and DES), galaxies within this volume can be averaged over with a suitable weighting to determine a value for H_0 (Schutz 1986). Although each individual event is not very accurate, there are many more black hole mergers than neutron star mergers, so the errors are competitive in aggregate. It is also straightforward to mix dark and bright sirens to produce a combined result (Palmese et al. 2020) (Fig. 13).

As each event's inclination angle is independent, averaging over N events will improve the errors relative to a single event by a factor of $N^{-1/2}$. The main limiting factor is then detector calibration and sensitivity, which determines the event rates. Work is underway at the LIGO sites to improve this, using quantum engineering techniques such as squeezed light.[11] Chen et al. (2018) estimate a 5yr observing run by the upgraded LIGO, Virgo, KAGRA and LIGO India detectors will be enough to measure H_0 to 1% precision by 2030.

In summary, the appeal of gravitation waves is that if one believes General Relativity correct, they are self-calibrating systems independent of the distance ladder and with minimal astrophysics input. Event rates and improvements in detector calibration are sufficient to converge to a $\sim 2\%$ precision within the next decade.

3.10 Baryon acoustic oscillations

Baryon acoustic oscillations (BAO) are the relic of density fluctuations propagating in the pre-recombination universe. The baryons in an initial localised seed of overdensity (such as would originate from inflation) are subject to radiation pressure and expand outwards at the speed of sound c_s, leaving the dark matter behind. At recombination, temperatures have dropped enough so that protons and electrons can combine to neutral hydrogen and the Thomson scattering rate quickly drops below the Hubble expansion rate. At the drag redshift z_{drag} shortly after, baryons are released from radiation pressure and the baryon overdensity shell is "frozen in" at a characteristic distance relative to the central dark matter overdensity. This distance

$$r_d = \int_{z_{\text{drag}}}^{\infty} \frac{c_s(z)}{H(z)} \, \mathrm{d}z, \tag{42}$$

in comoving coordinates is what is referred to as the sound horizon. As the universe

---

[11] Squeezed light is a state of light where the distribution of the Heisenberg uncertainty between observables is controlled for a specific application. In this case, the aim is to reduce photon shot noise at the expected frequency of the gravitational waves, and it is a nice crossover from techniques developed for quantum engineering to fundamental physics.
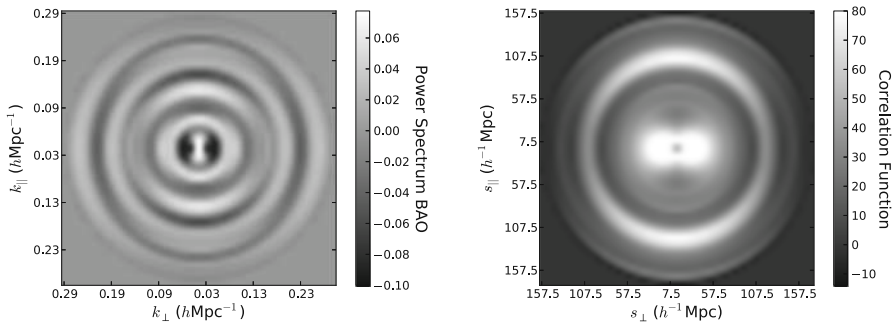
Fig. 14 A beautiful illustration of the BAO peak in the redshift bin $0.4 < z < 0.6$ of BOSS DR12. The left-hand panel shows the power spectrum along and transverse to the line of sight, and the right-hand panel is the two-point correlation function. The anisotropy visible in the graphs breaks the $d_A^2 H^{-1}$ degeneracy. Image reproduced with permission from Alam et al. (2017), copyright by the authors

continues to evolve, both the baryon shell and the central dark matter overdensity attract further gravitational infall of matter, forming galaxies. Thus, the sound horizon becomes visible as a characteristic (statistical) physical separation between galaxies. The sound speed $c_s = 1/\sqrt{3(1 + R)}$ depends on the baryon-to-photon ratio with $R = 3\rho_b/4\rho_\gamma$. z_{drag} has weak cosmological dependence, and Planck data gives $z_{drag} = 1059 \pm 0.3$. The key point is that given a cosmological model, r_d is calculable and its scale is large enough that the evolution of structure from z_{drag} to $z \sim 1$ is nearly linear. Thus we can compare observations of galaxy clustering in the late universe to r_d. Armed with data for r_d we can use Eq. (42) above to solve for H_0 in our model.

The theory of measuring BAO was largely worked out by the mid 1990s (Feldman et al. 1994). One assumes galaxies are a Poisson sample of the relative matter density field $1 + \delta(\mathbf{r})$ and so

$$P(\text{ Volume element } \delta V \text{ contains a galaxy}) = \delta V \bar{n}(1 + \delta(\mathbf{r})), \qquad (43)$$

where $\bar{n}$ is the expected mean space density of galaxies. It is possible to work either with the two-point correlation function $\xi(\mathbf{r}) \equiv \langle \delta(\mathbf{r}')\delta(\mathbf{r}' + \mathbf{r}) \rangle$ or equivalently the power spectrum

$$P(\mathbf{k}) \equiv \int d^3 r \xi(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}). \qquad (44)$$

It might seem preferable to use the power spectrum, as we are looking for a characteristic wavelength, and the modes of the power spectrum are independent if the density field is Gaussian. Also, the galaxy survey footprint in real space becomes an easily understood point spread function in Fourier space. In practice though, because of issues with binning such as cut-offs in Fourier space, most research papers calculate both the correlation function and power spectrum. Adjusting $\bar{n}$ above can correct for bias introduced by edge effects of angular and redshift cuts of the survey.

Expanding the power spectrum in spherical harmonics, the monopole determines $d_A^2 H^{-1}(z)$, and the quadrupole determines $d_A H(z)$ (Padmanabhan and White [2008]) in a given redshift bin centred at z. With $\Delta\theta$ and Δz, the density peak separation transverse to and along the line of sight, we can solve for d_A and H. We start with a mock galaxy catalog constructed in a fiducial cosmology seeded with density perturbations. This is matched to our real data with stretches $\alpha_\perp, \alpha_\parallel$:

$$\alpha_\perp = \frac{d_M(z) r_{d,\mathrm{fid}}}{d_{M,\mathrm{fid}(z) r_d}} \quad \alpha_\parallel = \frac{H_{\mathrm{fid}}(z) r_{d,\mathrm{fid}}}{H(z) r_d} \tag{45}$$

$$\Delta\theta = r_d / d_M(z) \tag{46}$$

$$\Delta z = H(z) r_d, \tag{47}$$

where $d_M = (1+z) d_A$ is the comoving angular distance. Here $\alpha = \alpha_\perp^{2/3} \alpha_\parallel^{1/3}$ is the monopole term and $\epsilon = \alpha_\perp / \alpha_\parallel$ is the quadrupole.

There is the suggestion of circularity in this: a random catalog constructed in ΛCDM (with small scale power suppressed) is then "tweaked" to compare to the real sky to justify ΛCDM. However, this method is essentially perturbative. If the real cosmology is far from the fiducial one, a bias might be introduced, but if it is close it should work well.

A technique called *reconstruction* is used to sharpen the BAO signal (Eisenstein et al. [2007]; Anderson et al. [2012]), and we describe it briefly here as we will mention it in Sect. [4], when we discuss modified gravity. The observed positions of galaxies will be somewhat moved from their original positions "on" the sound horizon, due to their subsequent infall towards overdensities. This has the effect of blurring the peak of the correlation function, and reduces precision. Reconstruction reverses the infall in the following way: (a) take a volume and smooth out small scale structure < 20 Mpc (b) embed it into a larger scale random structure to remove edge effects (c) estimate the displacement $\mathbf{q}$ of each galaxy by the continuity equation $\nabla \cdot \mathbf{q} = -\delta_{\mathrm{gal}} / b_{\mathrm{gal}}$ where δ_{gal} is the fractional galaxy overdensity and b_{gal} the galaxy bias (d) shift the galaxies by $-\mathbf{q}$ plus an additional shift for redshift space distortions (e) do the same to the mock catalog.

Historically, the first galaxy surveys to report a BAO detection were the Two Degree Field survey (Cole et al. [2005]), and the Sloan Digital Sky Survey (Eisenstein et al. [2005]), followed by WiggleZ (Blake et al. [2011], [2012]) and others. We discuss here results from Baryon Oscillation Spectroscopic Survey (BOSS) (Anderson et al. [2012]). BOSS has spectroscopically surveyed 1.5 million luminous galaxies over 10,000 deg^2 between $0.2 < z < 0.7$, with angular separations greater than $62''$ (which is 0.5 Mpc at $z \sim 0.7$). Alam et al. ([2017]) analyse the BAO signal in Data Release 12. The galaxies are divided into three bins each of redshift and $\cos\theta$. From a fiducial cosmology with $h = 0.676$, they find $\alpha = 1.042 \pm 0.016$ after reconstruction. Their result is summarised as a set of three values for $d_m(z_i)$ and $H(z_i) \times r_d / r_{d,\mathrm{fid}}$ for $z_i = 0.38, 0.51, 0.61$, with precision $\sim 2.5\%$ (Fig. [14]).

The construction of mock galaxy catalogs is a key part of this analysis. An important parameter is b_{gal} which is the (assumed linear) ratio of the galaxy

overdensity to the matter overdensity. Constructing a mock galaxy catalog by brute force of N-body simulations is prohibitively expensive. Instead, construction starts with a fast, approximate gravity solver for the matter density field seeded with initial fluctuations, and resolution down to halo size. Galaxies are inserted into the matter halos using the bias ratio, while preserving the two- and three- point correlation functions. BOSS assumed b_{gal} is constant, but more sophisticated models exist. For example, bias tends to be higher for red galaxies (see for example Zehavi et al. 2011). Attention also needs to be paid to what types of galaxies are being counted: massive luminous galaxies of the type measured by BOSS reside in high density nodes of the cosmic web, whereas emission line galaxies are strung out along the filaments. For further details, see Kitaura et al. (2016). BOSS DR12 uses two independent mock catalogs to mitigate against biases introduced by simulation effects.

BAO in isolation do not constrain H_0, but rather the combination of $H(z)$ and the sound horizon r_d described above. To obtain H_0, we must either fit a cosmological model with constraints on its density parameters supplied by other datasets (for example SN Ia, CMB, or even the full shape of the matter-power spectrum), or we must supply a prior for r_d (which is often obtained from the CMB, but see also below). We describe three independent results below.

Alam et al. (2017) used ΛCDM calibrated to a combination of the Planck power spectrum and the JLA sample of SN Ia. They find $H_0 = 67.6 \pm 0.5$ kms^{-1} Mpc^{-1}. This value is consistent with the Planck result (Planck Collaboration 2020), but is not independent of it.

Lemos et al. (2019) use a parametric formula for $H(z)$, in conjuction with a WMAP-derived prior for r_d and the Pantheon SN Ia sample, to derive $H_0 = 67.9 \pm 1.0$ kms^{-1} Mpc^{-1}. This result is now independent of Planck, and their parametric $H(z)$ allows for deviation from ΛCDM by separating early and late universe expansion histories.

Addison et al. (2018) use data (almost) independent of CMB in ΛCDM. They determine $\Omega_m = 0.292 \pm 0.02$ by combining galaxy and Lyα BAO.[12] Ω_r is derived from the CMB temperature as measured by COBE/FIRAS and numerous balloon experiments (Fixsen 2009). Setting $\Omega_k = 0$ determines $\Omega_\Lambda = 1 - \Omega_m - \Omega_r$. To get the sound horizon, we now just need the baryon density Ω_b (see Eq. (42)). Without using the CMB power spectrum, the best available way to get it is from primordial deuterium abundances. In Big Bang nucleosynthesis (BBN), deuterium is burned to create ^{4}He, with some uncertainty due to the reaction rate of $d(p,\gamma)^3$He. The reaction rate increases with physical baryon density so [D/H] decreases with $\Omega_b h^2$. The authors use the primordial [D/H] abundance measured from metal-poor damped Ly-α systems, which is $2.547 \pm 0.033 \times 10^{-5}$. Putting this all together, they derive $H_0 = 66.98 \pm 1.18$ kms^{-1} Mpc^{-1}. The only CMB data that has been used is the

---

[12] In the analysis of Addison et al. (2018) the two data sets are in mild tension, however, recent data have a greater level of compatibility—see for example Fig. 5 of Alam et al. (2021).

temperature, and the result corroborates the Planck value.[13] Repeating the analysis with galaxy data from the Dark Energy Survey (DES) Year 1 produces a consistent result (Abbott et al. 2018).

These results are consistent with more recent ones. DES Y3 survey data has been used to derive a set of three correlation measures (referred to as "3x2pt"). These measure the angular correlation of galaxy positions with each other, with the tangential shear of their shapes caused by weak lensing, and the cross-correlation in the two components of their ellipticities also due to weak lensing. Combining this with Planck, SN Ia data, BAO and redshift space distortions, the DES Collaboration (Abbott et al. 2021) find $H_0 = 68.0^{+0.4}_{-0.3}$ kms^{-1} Mpc^{-1} when fitting to the ΛCDM model. The eBOSS survey, the culmination of more than 20 years of survey work at the Apache Point Observatory, find $H_0 = 68.19 \pm 0.37$ kms^{-1} Mpc^{-1} when fitting to a ΛCDM model using their data plus the above additional probes to also fit to ΛCDM (Alam et al. 2021). These seem to be the tightest constraints on H_0 yet published.

We also note that the logic explained above to derive H_0 can be reversed: BAO can be combined with a H_0 prior from nearby universe results, instead treating the sound horizon r_d as a free parameter (Bernal et al. 2016). The r_d values so inferred are in tension with the CMB, and we return to this point in Sect. 4.2.

In summary, BAO serve as a standard ruler in the late universe which are almost independent of astrophysical assumptions. They may be used in extended models beyond ΛCDM, albeit with the caveat they have been derived perturbatively against a ΛCDM background simulation. The two main data sets of BOSS and DES are soon to be joined by others. The Dark Energy Spectroscopic Instrument is operational, the ESA survey satellite Euclid is scheduled for launch in 2022 which is also when science operations on the Vera Rubin Observatory LSST in Chile commence. All of these surveys are complementary and will increase the precision and range of BAO measurements. Combined with deuterium abundances, BAO are an important corroboration of the Planck and WMAP results for H_0.

3.11 Cosmic microwave background

The CMB power spectrum carries the imprint of the same acoustic oscillations we described for BAO, sampled on the surface of last photon scattering. The Planck 2018 value of $H_0 = 67.4 \pm 0.5$ kms^{-1} Mpc^{-1} (Planck Collaboration 2020) is notable for its precision, and is consistent with previous results such as Boomerang 64 ± 10 kms^{-1} Mpc^{-1} (Percival et al. 2002) and WMAP 70.0 ± 2.2 kms^{-1} Mpc^{-1} (Hinshaw et al. 2013). More recently, the Atacama Cosmology Telescope (ACT) finds 67.9 ± 1.5 kms^{-1} Mpc^{-1} (Aiola et al. 2020). There is, however, a moderate tension between Planck and the South Pole Telescope (SPT) result of

---

[13] Standard ΛCDM is assumed, and in particular that the expansion rate around the time of BBN as constrained by He abundances (parametrised by the effective number of relativistic species in the early universe N_{eff}) is consistent with the CMB.

71.6 ± 2.0 kms^{-1} Mpc^{-1}, and the origin of the difference is not yet clear (Henning et al. 2018; Handley and Lemos 2021).

Two questions are immediately raised: how is H_0 derived from a power spectrum sampled on the two-dimensional sky of the early universe, and why is the Planck value so precise?

To answer the former, we need to explain how the CMB power spectrum determines cosmological parameters. The spectrum is created by primordial fluctuations that evolve as a function of their scale and the sound speed, until recombination. Whilst this radiation is highly isotropic, we can obtain cosmological results from measurements of small anisotropies in both the temperature and the polarization of CMB photons. The former of these provides the largest amount of cosmological information, and can in fact measure the Hubble constant at high enough accuracy to be in tension with direct measurements without relying on polarization. These temperature anisotropies ($\Delta T \equiv T(\mathbf{n}) - T_0$) can be expanded in a basis of spherical harmonics:

$$\Delta T(\mathbf{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_\ell^m(\mathbf{n}). \tag{48}$$

The power spectrum of the coefficients is commonly expressed as $\mathcal{D}_\ell = \ell(\ell+1)C_\ell/2\pi$, where

$$C_\ell \equiv \frac{1}{2\ell+1} \sum_m a_{\ell m} a_{\ell\,-m} \tag{49}$$

is the estimator of the power spectrum (we have only one CMB, so averaging over m is a way to estimate the true "universe average" we want). The spectrum is shown in Fig. 16 and contains all the statistical information of the temperature anisotropies, as long as these are Gaussian (no evidence supporting non-Gaussianity in the CMB has been detected at the time of writing). This power spectrum has three main features:

– **Large-scale plateau** at scales ($\ell < 100$) that are not affected by post-recombination physics, and, therefore, reflect the primordial power spectrum. As noted above, we have just one "sampling" of the random seeds, so the amount of information we can extract from low multipoles is limited by cosmic variance $\propto (2\ell+1)^{-1/2}$.
– **Acoustic oscillations** which depend on the sound horizon size

$$r_s(z_\star) = \frac{1}{H_0 \Omega_m^{1/2}} \int_0^{a_\star} \frac{c_s}{\sqrt{a + a_{\mathrm{eq}}}} \mathrm{d}a \tag{50}$$

in flat ΛCDM where $a_{\mathrm{eq}} = \Omega_r/\Omega_m$, $\Omega_b h^2$ determines c_s and both determine $a_\star = a(z_\star)$, the scale factor at the surface of last scattering. These create the peaks in the spectrum.
– **Silk damping** which is photon diffusion from hotter regions to colder ones, suppressing small scale power (Silk 1968). The mean free path of a photon is
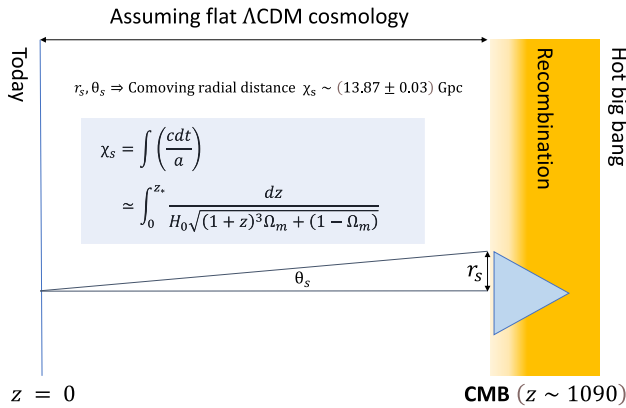
Fig. 15 A schematic of how the sound horizon determines H_0. We observe the angular peak spacing θ_s, and calculate r_s from the pre-combination universe given a model for the sound speed and expansion history there. The comoving distance χ_s is then used to solve for H_0 given how χ_s evolves in the post-recombination universe. The box gives the flat ΛCDM formula for this. This figure is our version of Antony Lewis' original at https://cosmologist.info

$\lambda = 1/(\sigma_T n_e)$ where n_e is the number density of charged particles and σ_T the Thomson cross-section, so photons can diffuse a length $r_{\text{Silk}} = \sqrt{\lambda/H}$. As the sound horizon $r_s \propto 1/H$, then $r_{\text{Silk}}/r_s \propto \sqrt{H/n_e}$. Silk damping creates the downward slope in the spectrum, so this is sensitive to $n_e \propto \Omega_b$.

With all of this, how does the CMB constrain the Hubble parameter? First, we can measure the acoustic scale length r_s, which serves as a standard ruler. As described in Eq. 50, this can be calculated as a function of Ω_r, $\Omega_m h^2$ and $\Omega_b h^2$: The radiation density Ω_r is determined by the temperature of the CMB. $\Omega_m h^2$ can be estimated from the effect of the integrated Sachs–Wolfe effect (Sachs and Wolfe 1967)—the redshift of CMB photons through wells or hills of gravitational potential during eras of non-matter domination—in the low multipoles of the power spectrum. Finally, the relative amplitudes of the acoustic peaks in the CMB serve to measure the ratio Ω_b/Ω_m. Once we have computed the acoustic scale, we need a measurement of its comoving angular diameter distance θ_s, which is given by the distance between the acoustic peaks.[14] In a flat ΛCDM model, this provides a measurement of the matter density Ω_m, which combined with the measurement of $\Omega_m h^2$ gives an estimate on the Hubble parameter. This is illustrated graphically in Fig. 15.

Planck measured full-sky temperature T and polarisation E fluctuations in 9 frequencies, with angular resolution up to 5 arcmins ($\ell \simeq 2000$) and sensitivity $2 \times 10^{-6}\,\mu K$. Before calculating the TT, TE and EE power spectra, foreground effects must be calculated and removed, and much of the work done by the Planck

---

[14] This is to speak loosely—if we just knew the location of the peaks to 1% accuracy then the accuracy of H_0 would be only $\sim 7\%$. Using the entire spectrum is the driver for the precision. For a very accessible account of the CMB power spectrum with semi-analytic equations we recommend Mukhanov (2004).
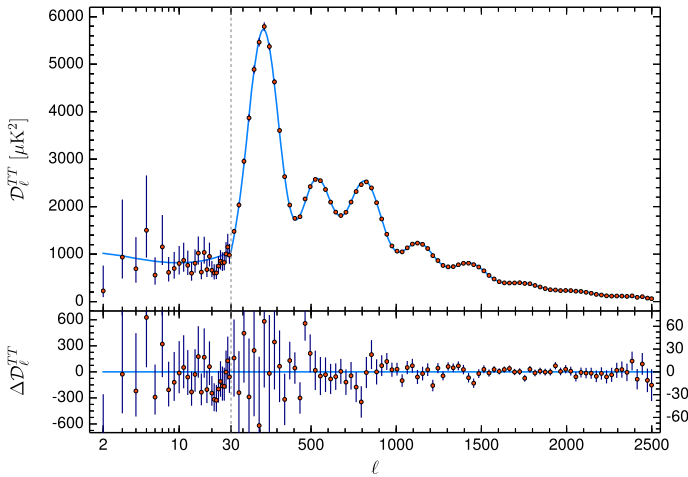
Fig. 16 The binned Planck power spectrum for TT and residuals to the best fit ΛCDM model. The vertical line delineates the different methodology used to resolve the power spectrum at $l < 30$. Note that the figure shows the power spectrum of the temperature T, not of the anisotropies Θ. Image reproduced with permission from Planck Collaboration (2020), copyright by ESO

collaboration between successive data releases has been to develop the foreground model, and reduce the uncertainties it may introduce. Foreground effects include:

– **Peculiar velocity of the Earth-Sun system** induces a dipole which is easily subtracted. The annual motion of the Earth is actually helpful to calibrate the detector response.
– **Galactic synchrotron and free-free emission** caused by cosmic rays interacting with the magnetic field and ISM of the MW. This has a non-thermal frequency spectrum (in effect its colour is different to the CMB).
– **Galactic dust emission** whose contribution to the spectrum is also subtracted by estimating its mapped "colour" difference to the CMB.
– **Point sources** of microwave emission are masked. Sky masking introduces some correlation between multipoles.
– **Sunyaev–Zel'dovich (SZ) effect** where CMB photons interact with hot intracluster gas and are Doppler-shifted (the kinetic SZ effect) by bulk motion or up-scattered (the thermal SZ effect) (Sunyaev and Zel'dovich 1972). The former effect does not alter the black-body spectrum, but the latter does. The kinetic SZ effect is sensitive to epoch of re-ionisation at $z \sim 8$ (Planck Collaboration 2014b). The thermal SZ effect may be exploited to produce a cluster map from which information on inhomogeneity can be extracted (Planck Collaboration 2014a). Modelling of intracluster gas predicts a spectral template (which peaks around $\ell \simeq 2000$) to subtract.
– **Weak lensing** of the CMB by matter. Weak lensing peaks between $1 < z < 2$, and may be used as a probe of the distribution of foreground matter. The effect increases as we move to smaller scales, particularly $\ell > 2000$ (Zaldarriaga and Seljak 1998). There are two signals of lensing: (a) its non-gaussianity, which can

be measured by a four-point correlation function (b) a smoothing effect on the peaks and troughs of the spectrum, both of which can be calculated as a function of Ω_M. However, there is more smoothing apparent in the spectrum than is predicted by the correlation function so a phenomenological parameter $\Omega_{m,\text{spectrum}=A_{\text{lens}}\Omega_{m,\text{correlation}}}$ has been introduced to capture their relative amplitude. $A_{\text{lens}} \simeq 1.2 \pm 0.1$ is seen in both Planck and South Polar Telescope (SPT) data (Bianchini et al. 2020), but appears to be absent in the Atacama Cosmology Telescope (ACT) data (Aiola et al. 2020).

The latter two are perhaps better characterised as secondary anisotropies, as they have been used to derive cosmological information in their own right. The secondary anisotropies then provide a consistency check of certain parameters derived from the primary anisotropies, if the modelling is correct. We examine A_{lens} in greater depth below.

Cross-checks on the effectiveness of foreground modelling are done by switching frequency bands, using alternative astrophysical models, and numerical simulations of parameter recovery from random CMB backgrounds (Planck Collaboration et al. 2016, 2020). The spectrum is then compared to one computed in a standard code such as CMBFAST (Seljak and Zaldarriaga 1996), CAMB (Lewis et al. 2000) or CLASS (Blas et al. 2011) and a posterior for each cosmological parameter is computed. In each check, no serious discrepancy with the main results was found. Although each C_ℓ individually has $\simeq 7\%$ noise, there are 2000 of them and the error is reduced further by the lensing signal. The binned spectrum and residuals is shown in Fig. 16. We now consider two questions regarding the data.

Is Planck self-consistent? A feature which is (barely) visible by eye in Fig. 16 is an oscillating residual pattern lining up to the spectral peaks, indicating the Planck spectrum is slightly tilted versus theory (see Fig. 24 of Planck Collaboration (2020) for an enlarged view). As each multipole is nearly independent, we can split the spectrum and run the analysis on each half. Addison et al. (2016) find that A_{lens} is equivalent to a drift of cosmological parameters with different scales in the CMB. They fix $A_{\text{lens}} = 1$ and compare H_0 for $\ell < 1000$ and $\ell > 1000$ finding $H_0 = 69.7 \pm 1.7$ kms^{-1} Mpc^{-1} and $H_0 = 64.1 \pm 1.7$ kms^{-1} Mpc^{-1}, respectively. Allowing $A_{\text{lens}} = 1.3$, or a tilt parameter restores concordance.

We caution that A_{lens} is not a physical parameter connected to lensing; it is a "fudge factor" that resolves internal inconsistency in CMB data. It does not resolve the H_0 tension, but does resolve matter power amplitude σ_8 tension with galaxy survey data such as the Dark Energy Survey and Kilo-Degree Survey. Could it be a foreground effect? Its value is relatively stable for various band channels and sky masks, which argues that it is not. The Planck Collaboration followed up on this curiosity (Planck Collaboration 2020). They note a dip in the power spectrum for $l < 30$ pulls low-ℓ H_0 higher; this may account for the lower-resolution WMAP measurement being somewhat higher than Planck. The dip between $1420 \leq \ell \leq 1480$ mimics lensing but may be an unaccounted foreground effect. Potential explanations such as negative curvature density $\Omega_k < 0$ and modified gravity are investigated but no convincing evidence is found in favour of these

models when BAO data are added to the fit. A re-analysis of the Planck data using a different foreground subtraction methodology (Efstathiou and Gratton 2019) found that A_{lens} decreased from 1.26 to 1.06 as the data were broadened from the temperature-only power spectrum to include polarisation and lensing data. Further, Planck data passes other consistency tests with the baseline ΛCDM model where A-type parameters may be introduced, such as the expected magnitude of the lensing correlation function and the Sachs–Wolfe effect. Hence, the preferred explanation of the Planck Collaboration is that A_{lens} is a statistical fluke. Nevertheless, while A_{lens} does seem to vary over the sky, it is perhaps of concern that it seems to be larger for full-sky Planck data is and is a roughly 2σ effect.

Is Planck consistent with other current CMB measurements? The SPT covers a 6% patch of the southern sky at a resolution $6\times$ that of Planck, finding $H_0 = 71.3 \pm 2.1$ kms^{-1} Mpc^{-1} (Henning et al. 2018). SPT is limited to $\ell > 650$ due its low sky coverage and atmospheric effects and has greater noise, but when the results are compared on a like-for-like basis ("Planck-in-patch" with SPT for $650 < \ell < 1800$), they are consistent with Planck. The analyses are independent and so argue against systematic errors (in this patch and multipole range) in either experiment. However, including higher SPT multipoles in the 150 GHz spectrum causes H_0 to drift higher: $H_0 \simeq 74 \pm 3$ kms^{-1} Mpc^{-1} for $650 < \ell < 3000$ (Aylor et al. 2017). While this is the opposite direction to Planck data, this may be due to the $l < 30$ effect in Planck noted above. SPT confirm the Planck result of a greater lensing effect than predicted, and also note a spectral tilt versus the best-fit ΛCDM. However, the trend is not apparent in 143 GHz data, so it is not at all clear if this is a physical effect, rather than merely chance or systematics. The ACT Data Release 4 derives cosmological parameters from a $6,000 \text{deg}^2$ patch of the southern sky up to $\ell \sim 4000$, finding $H_0 = 67.9 \pm 1.5$ kms^{-1} Mpc^{-1} (Aiola et al. 2020), however, ACT appears to be in 2.6σ tension with Planck Handley and Lemos (2021); this appears to be caused by differing physical baryon densities.

In summary, Planck 2018 results for H_0 are consistent with previous and current independent CMB data. The high precision for H_0 is a consequence of Planck's high resolution, lack of atmospheric interference, and full sky coverage. Checks have been performed on the codes, the methods used to remove foregrounds, and beam effects. Internal inconsistencies hint at new physics or foreground effects at small angular scales, manifesting either as a tilt or a smoothing of the spectrum, and non-physical parameters like A_{lens} have been introduced to capture them. However, after their introduction the results for H_0 are not materially affected, and remain in tension with late universe results. Nevertheless, the persistence of these effects across differing sky patches, multipole ranges and experiments make them less likely to be a statistical fluke. It is hoped next generation CMB experiments with resolution up to $\ell = 5000$ will shed further light on this.

3.12 Other methods

Promising methods that we did not have space to discuss are Surface Brightness Fluctuations (Tonry and Schneider 1988; Blakeslee et al. 2021; Khetan et al. 2021),

Cosmic Chronometers (Jimenez and Loeb 2002; Moresco et al. 2018), the Tully–Fisher relation (Tully and Fisher 1977; Kourkchi et al. 2020; Schombert et al. 2020), Type II supernovae (de Jaeger et al. 2020), HII galaxies (Terlevich and Davies 1981; Arenas et al. 2018), and Galaxy Parallax (Croft 2021). Quasars (Risaliti and Lusso 2019) and GRBs (Schaefer 2007) offer the prospect of extending the Hubble diagram up to $z \sim 5$, further testing ΛCDM. However, despite improvements in characterising their intrinsic luminosity from observables (for example X-ray and UV flux in the case of quasars), it remains challenging at present to regard them as standard candles. We refer the reader to the above for the latest work in these fields.

4 Could new physics explain the tension?

We have now reviewed the main recent results on H_0, and while in the spirit of scientific scepticism we have identified potential sources of systematic error, we hope it is also clear how much effort has been made to root out potential biases. Therefore, a possibility that should be taken seriously is that the tension is real, and a sign of new physics.

Some authors have interpreted the disagreement between CMB results and Cepheids as "early versus late" tension. In this view, ΛCDM with its fluids of radiation, standard model neutrinos, cold dark matter, baryons and dark energy is not the right model to derive $H(z = 0)$ from the apparent $H(z \sim 1100)$. If late-time physics were changed, H_0 derived from the early universe might be reconciled to the local one. Alternatively, a change to pre-recombination physics would alter our calculation of the sound horizon r_s, and hence we would need to change distances to retrieve the same angular sizes of the CMB temperature fluctuations. More radically, some authors have argued our local H_0 is different from an average over randomly placed observers.

The essential problem for model builders is this: in other respects, ΛCDM fits the data very well across a huge span of the history of the universe. The CMB is not merely a "snapshot" of the universe, but carries the imprint of several epochs. The positions and heights of the peaks are sensitive to the content[15] between $z \sim 10^5$–10^3, the high-ℓ slope depends on the baryon-to-photon ratio close to recombination at $z \sim 1100$, and CMB lensing by matter peaks between $z \sim 1$–2. At earlier times, ΛCDM makes an accurate prediction of primordial element abundances, especially deuterium. At late times, galaxy surveys are consistent with the evolution of perturbations in a ΛCDM background,[16] and place tight limits on deviation from spatial flatness in the mid to late universe. The shape of $H(z)$ can be read off SN Ia luminosities between $0.01 < z < 1.4$, and is also consistent with ΛCDM.

What we want to have is a model capable of modifying H_0 by the right amount, without disrupting ΛCDM's other successful predictions. The model must be

---

[15] In ΛCDM, this is parametrised by r_{eq}, the size of modes crossing the horizon at matter/radiation equality $z \sim 3300$.

[16] With the caveat of a moderate tension between the CMB and BAO values of the amplitude of fluctuations in the matter-power spectrum, parametrized by σ_8.

testable, compatible with particle physics results, and ideally not fine-tuned. Two serious problems must be avoided: by dint of extra parameters, most models increase the range of allowed H_0 values. When a prior of H_0 from the local universe is convolved with the expanded likelihood, the posterior will—by construction—overlap with the prior. So to claim a resolution of the Hubble tension in this way is to use the same data twice: once to construct the posterior, and once to compare it. This can be avoided if the extra parameters have some preferred values—for example, a range that is predicted by microphysics (Vagnozzi 2020). A second problem with the use of H_0 priors has been discussed by Efstathiou (2021), who points out the difference between it and the actual data analysis performed by the SH0ES team, which is a calibration of the SN Ia fiducial absolute magnitude M_B, followed by a conversion from M_B to H_0. To use an H_0 prior is then to transform and compress the data away from its true source; it would instead be better to use M_B directly and priors for this are given in Camarena and Marra (2021). Another prevalent issue is the selective use of datasets—in particular, BAO provide strong constraints on the late universe as we shall see shortly, and to omit them is again to risk a misleading analysis.

A large number of creative proposals have been put forward, so much so that unfortunately we do not have space to introduce them all. Our preference here is to review selected models in broad classes that we hope will be informative, and readers are referred to extensive reviews by Mörtsell and Dhawan (2018), Knox and Millea (2020), Di Valentino et al. (2021), and Vagnozzi (2020) if they would like more detail.

4.1 Late-universe physics

In some respects, it is easier to propose changes to the late universe: one does not have to "negotiate" with the CMB power spectrum! It is straightforward to generalise the expansion history into any number of "cosmology-independent" forms, that do not rely on any specific form of matter-energy, or general relativity: all that is retained are the Copernican principles of homogeneity and isotropy, and the existence of a space-time metric (for example, see Eq. (15)). The latter is important, as it implies the Etherington relation: $d_L = (1 + z)^2 d_A$, which locks angular diameter distances against luminosity distances.

4.1.1 Modified gravity

Traditional models of large-scale modified gravity (normally referred to as MOND) would seem to have a low chance of solving the Hubble tension; the success of using reconstruction to sharpen the BAO peak argues that standard gravity works as we expect on these scales. Desmond et al. (2019) have proposed that small-scale modified gravity could distort the calibration of Cepheids in large galaxies like the Milky Way and NGC4258, however, this appears to be disfavoured by the fact that using the LMC as a sole calibrator does not materially change H_0.

4.1.2 Inhomogeneities and redshifts

Clearly our universe is not completely uniform, so could we by chance live in an underdense region, measuring a local H_0 above the universal one?

Shanks et al. (2019) have investigated the "local hole" idea, initially noted in galaxy survey data by Keenan et al. (2013). They cite peculiar velocity outflows from $0 < z < 0.1$ in the 6dFGS galaxy survey as evidence for a local underdensity, finding the universal (i.e. CMB) H_0 to be lower than our local one by 1.8%. However, we would expect to see the effect of a line of sight exiting our underdense region as a kink in the residuals of SN Ia luminosities fitted to ΛCDM, or alternatively as an anistropy in H_0 depending on the footprint of the SN Ia survey used. While the sky distribution of low-z SN Ia versus Hubble flow SN Ia are very different, and there is an uptick in magnitude residuals on the boundary between the two (see for example Fig. 11 of Scolnic et al. 2018), the lack of evidence for such a kink places limits on any local under-density (Kenworthy et al. 2019). Also, large-scale structure simulations find the probability for us to be in an underdensity of such magnitude is less than 1% (Odderskov et al. 2017; Macpherson et al. 2019) (although a different likelihood of a void may be obtained in a revision of ΛCDM). That said, the observational tension between claims of local under-density based on galaxy surveys, and the lack evidence for it in SN Ia data, remains unexplained.

Density fluctuations also necessitate the adjustment of redshifts for peculiar velocities. Although heliocentric redshifts can be measured up to an accuracy of 10^{-7}, care is needed in the conversion to cosmological ones. Davis et al. (2019) have analysed the effect of redshift biases on H_0. Changing z by $+0.001$ for a set of standard candles between $0.01 < z < 0.15$ causes a bias in H_0 of $\sim 3\%$. Systematics might arise in peculiar velocity estimates or corrections for residual stellar motion (for example, if observing on one side of a spiral galaxy). Rameez (2019) has pointed out differences between the redshifts for SN Ia in common between the JLA and Pantheon catalogs; there are 58 with differences > 0.0025, concentrated in an arc opposed to the CMB dipole. Nevertheless, this does not appear to be biasing any H_0 estimates.

4.1.3 Modifying late-time ΛCDM

In ΛCDM, dark energy is a property of the vacuum, with equation of state $\rho = wP$ where $w = -1$. This distinguishes it from standard inflation models, where $-1 < w < 0$ is a time-varying function of the potential and kinetic energy of a scalar field, but makes its microphysical origin rather obscure. This has caused some theorists to speculate with alternative models.

In principle then, modifying dark energy in the late universe may be considered as a solution. The CMB has little to say about late-time dark energy: at early times, its physical density is much smaller than matter and radiation. The late time Integrated Sachs–Wolfe effect does imprint on the spectrum, but only as a large scale secondary anisotropy once the universe has entered the dark-energy dominated phase at around $z \sim 0.3$.
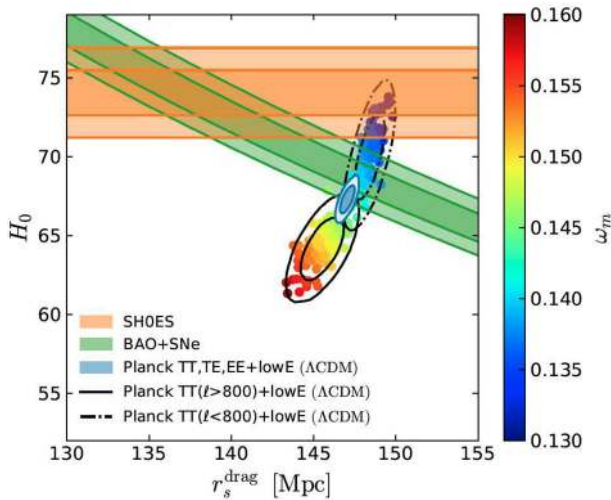
Fig. 17 ΛCDM tensions in the $r_d - H_0$ plane. ω_m is the physical matter density $\Omega_m h^2$ and $r_s^{drag} \equiv r_d$ is the sound horizon from the end of the baryon drag epoch. There is no combination of values in standard, flat ΛCDM that can explain observations. Image reproduced with permission from Knox and Millea (2020), copyright by APS

A simple modification of ΛCDM would be to allow w to take on an arbitrary (constant) value. However, as what is required is a boost to late-time expansion to match the local universe H_0, we would need $w < -1$ (known as phantom dark energy) to solve the tension. Phantom dark energy does not occur in standard single scalar field models, but is possible in models with more complex field configurations. If sustained, phantom dark energy leads to a "Big Rip" in a finite time, where all matter is pulled apart by accelerated expansion of the universe.

Alternatively, dark energy may be considered a late-universe emergent phenomenon by allowing its equation of state to vary with time. This may be done writing $w(a) = w_0 + (1 - a)w_a$, or by replacing $\Omega_{\Lambda,0}$ in Eq. (12) with a phenomenological $\Omega(z)$. This could be motivated by the action of scalar fields, dark matter decaying to dark energy, or some other novel microscopic theory of dark energy (see for example Di Valentino et al. 2021). However, it can be shown that such models have a generic problem: the sound horizon. The sound horizon r_d is normally introduced via its definition from early universe physics (Eq. 42). But it is also imprinted on the late-time universe, as the peak in the angular correlation function of galaxy number densities, which constrains the product $r_d H_0$. All that is needed is a late-time distance measurement to convert the angular size to a physical size at a given redshift, SN Ia to constrain the late-time expansion history, and r_d may be calibrated *independently* of the CMB.

Arendse et al. (2019) have used the H0LiCOW lens distances, BAO and Pantheon supernovae to obtain just such a late-time r_d. They show that two modified late-time dark energy models applied to Planck data can resolve its H_0 tension with the late-universe, but not its r_d tension. Knox and Millea (2020) investigated this in ΛCDM with the same result.

Hence, if all data including BAO are considered—as they should be—it appears that late-time modifications to ΛCDM will not be satisfactory as a solution to the Hubble tension.

4.2 Early-universe proposals

We have mentioned how the CMB is a set of rulers whose scale is fixed at different epochs. To recap, the observed angular scale of the sound horizon at a given redshift z is

$$\theta(z) = \frac{r_s}{d_A} = \frac{\int_{z_D}^{\infty} \frac{c_s(z')dz'}{H(z')}}{\frac{1}{1+z}\int_0^z \frac{dz'}{H(z')}} \ , \tag{51}$$

where $c_s(\omega_b, z)$ is the sound speed which depends on the physical density of baryons, given by $\omega_b = \Omega_b h^2$. Decoupling happens when the mean free path of photons approaches the Hubble radius $c/H(z)$, hence z_D is also a function of all of the densities $z_d \equiv z_d(\omega_b, \omega_m, \omega_\gamma)$. Loosely, the goal is to reduce the sound horizon from (42) by $\simeq 7\%$, as shown in Fig. 17.
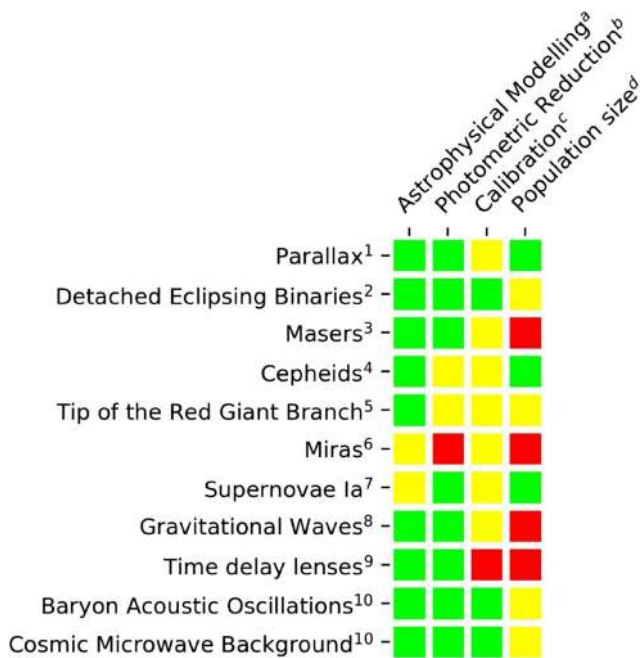
We may isolate the free parameters of this formula, from those that are fixed by data as follows. As discussed, we may omit dark energy ω_Λ in the numerator. In the late universe the radiation density is not significant, and using the spatial flatness implied by BAO, we may write $\Omega_\Lambda = 1 - \Omega_m$.

Ω_m is determined by the heights of peaks in the CMB spectrum. Successive peaks represent modes that have crossed the horizon successively earlier in the universe when the ratio of matter to radiation was lower, and so will have had their growth suppressed by radiation pressure to a greater degree. Ω_m determined in this way from the CMB is consistent at a variety of redshifts with values from the late universe: Lyman$-\alpha$ absorption lines of quasars, BAO, galaxy lensing and the SN Ia Hubble diagram (see Fig. 3).

Further, Ω_b is determined by the Silk damping of high-ℓ modes, and also by the relative heights of odd and even peaks: these are, respectively, compression and rarefaction modes, sensitive to the pressure to density ratio. Ω_b from the CMB is consistent with the deuterium fraction predicted by BBN. Then, substituting for $H(z)$ using (7), Eq. (51) becomes

$$\theta(z) = \frac{\int_{z_D(\omega_b, \Omega_m h^2)}^{\infty} \frac{c_s(\omega_b, z')dz'}{[(1+z')^3 + \frac{\omega_r}{\Omega_m h^2}(1+z')^4]^{1/2}}}{\frac{1}{1+z}\int_0^z \frac{dz'}{[\frac{1-\Omega_m}{\Omega_m} + (1+z')^3]^{1/2}}} \ . \tag{52}$$

Hence we see that for fixed ω_b, ω_r and Ω_M—all parameters whose CMB values are corroborated by non-CMB datasets—h only appears in the numerator. So to change H_0 we must either accept a departure from the Friedmann equations, change ω_r, allow conversion between energy types, or add a new non-baryonic, non-dark matter component to the mix of the early universe.

	Astrophysical Modelling [a]	Photometric Reduction [b]	Calibration [c]	Population size [d]
Parallax [1]	green	green	yellow	green
Detached Eclipsing Binaries [2]	green	green	green	yellow
Masers [3]	green	green	yellow	red
Cepheids [4]	green	yellow	yellow	green
Tip of the Red Giant Branch [5]	green	yellow	yellow	green
Miras [6]	yellow	red	yellow	red
Supernovae Ia [7]	green	green	yellow	green
Gravitational Waves [8]	green	green	yellow	red
Time delay lenses [9]	green	green	red	red
Baryon Acoustic Oscillations [10]	green	green	green	yellow
Cosmic Microwave Background [10]	green	green	green	yellow

Analysis Inputs

a. **Astrophysical Modelling** is whether the object being observed has a solid theoretical understanding.
b. **Photometric Reduction** are uncertainties that may be introduced converting imaging into parameters.
c. **Calibration** is uncertainty about the absolute magnitudes or physical size of the object.
d. **Population Size** characterises the quantity of independent data sets, objects or events that are available.

Probes of H_0

1. **Parallaxes** in Gaia DR3 have considerably improved the zero point, but not yet fully to science mission goals.
2. **Detached eclipsing binaries** provide a solid distance to the LMC and it is hoped in future to M31.
3. **Masers** are rare astrophysical systems and it is unlikely many more useful ones will be found, and reliance must be placed on the modelling of the accretion disk.
4. **Cepheid** crowding, blending and the calibrated slope of the Leavitt law has been the subject of recent work.
5. **TRGBs** are well understood stars which can be observed in uncrowded fields, but the calibration is debated. The magnitude is a statistical fit to a stellar population, rather than individual stars in the case of Cepheids.
6. **Miras** relatively new to distance ladders so not many galaxies have Mira distances. Their uncertainties may reduce quickly in future work.
7. **SN Ia** progenitor theory of an accreting white dwarf lacks observational evidence, hence many models of environmental or progenitor effects have been proposed. It is not clear if the host mass step used in SH0eS analyses is the best one.
8. **Gravitational waves** have ongoing improvements with detector calibration and sensitivity at LIGO and it is anticipated that more binary neutron star and binary black hole mergers will be detected.
9. **Time delay lenses** have mass modelling uncertainties, and a larger number of systems will need to be analysed to provide reliable convergence.
10. **BAO** and **CMB** future observations will extend the resolution of current data and help discriminate between new cosmological model proposals.

4.2.1 Extra relativistic species

Extra relativistic species, such as additional neutrinos beyond the Standard Model flavours, are well-motivated in extensions to the standard model in particle physics. There is a well-known constraint of 3 on the number of light neutral particles from the decay width of the Z-boson, but this bound may be avoided if the new particles do not couple to the Z (for example, if they are sterile neutrinos which only interact gravitationally). The number of species are parametrised by $N_{\rm eff}$ which is defined by

$$\omega_r = \left\{1 + N_{\rm eff}\frac{7}{8}(\frac{4}{11})^{4/3})\right\}\omega_\gamma \, , \tag{53}$$

where the fractions are due to the fermionic nature of neutrinos, and in the standard model $N_{\rm eff} = 3.045$.[17] Adding extra species would mean H_0 in Eq. (52), would need to be increased to keep θ unchanged (Bernal et al. 2016).

But if we increase ω_r, we have altered the redshift of matter-radiation equality $z_{\rm eq}$ which influences the relative heights of the peaks. If we restore the heights by also increasing Ω_m to keep $z_{\rm eq}$ unchanged, we conflict with Ω_m determined from SN Ia, BAO and lensing unless we accept dark matter decay between the CMB and then. Additionally, $N_{\rm eff}$ is constrained by He abundance to be <4 at the 3σ level during BBN (Aver et al. 2015) (see also Fig. 39 in Planck Collaboration 2020).

One way in which these constraints can be avoided is via self-interacting neutrinos. The new species couples to itself, but not the standard model neutrinos (although it mixes with them). Free streaming neutrinos have a small but measurable impact on the spectrum as they travel faster than the sound speed c_s and drag some of photon-baryon fluid with them, resulting in a sound horizon that would be larger than a neutrinoless universe. Self-interactions slow the speed of the neutrinos, and by tuning the strength of the interaction it is possible to arrange for the CMB spectral peak heights to be unchanged (with some small changes to other parameters such as the primordial power spectrum tilt n_s and Ω_A). By adding an additional coupling to a new scalar particle, their mixing may be suppressed at BBN, in which case they evade the He abundance constraints. Kreisch et al. (2020) have analysed such a model and found $H_0 = 72.3 \pm 1.4$ kms^{-1} Mpc^{-1} for $N_{\rm eff} = 4.02 \pm 0.29$.

The drawback is the self-interaction must be extremely strong. That is, the effective coupling $G_{\rm eff}$ required between neutrinos is $\sim 10^{10} G_{\rm Fermi}$ where the Fermi constant of weak interactions in the Standard Model is $G_{\rm Fermi} = 1.17 \times 10^{-5}$ GeV^{-2}. Such strength contradicts calculations of the decay time of muons and tau leptons, unless the self-interactions are restricted to the tau neutrino, or weakened to

---

[17] The slight deviation from 3 is due to non-thermal spectral distortions caused by electron-positron annihilation on the neutrino energy spectrum.

a level where they do not fully resolve the tension (Das and Ghosh 2021). So this proposal does seem highly fine-tuned. However, it is accessible to testing by CMB S4 experiments, as it predicts a stronger decrease of the damping tail for modes $\ell > 2000$.

4.2.2 Early dark energy

Early dark energy (EDE) refers to a boost to the expansion of the pre-recombination universe caused by a scalar field potential $V(\phi) \sim \phi^n$ in a manner similar to (but milder than) inflation. This dark energy must be present at around $z \sim 10^4$, as that is the time from which most of the growth of the sound horizon occurs. BBN constraints mean it must be absent earlier. The dark energy must also decay by the time of recombination, not to disrupt the damping tail of the CMB spectrum. EDE then provides a short-term boost to the expansion speed, which decreases the sound horizon by modifying the numerator of Eq. (52).

Agrawal et al. (2019) have shown $H_0 \sim 72$ kms^{-1} Mpc^{-1} for a fraction $\Omega_{\mathrm{EDE}} \sim 5\%$ of the energy content and the onset of its decay at redshift $z_c \sim 5000$, for a range of potentials. It may be said that this feels like fine-tuning again (though it should be said that the same point may be made about late dark energy—the "why now" problem)! The model shows a small residual oscillatory signal in the CMB spectrum versus a standard ΛCDM fit, so if H_0 is calculated in ΛCDM in an EDE universe, it is progressively biased lower for higher angular resolution (Knox and Millea 2020). Although this trend is already seen in WMAP and Planck, it is opposite to the trend in SPT.

EDE can then be tested by higher resolution in CMB S4 experiments. It may also be testable before then: it is necessarily present around the time of matter-radiation equality, and, therefore, will leave an imprint in the present-day matter power spectrum $P(k)$. The spectrum captures the strength of matter density fluctuations of wavenumber k and decreases for $k > k_{\mathrm{eq}}$, the wavelength corresponding to matter-radiation equality, with a gradient weakly proportional to k_{eq}. This spectrum is measured in galaxy surveys, and the greater precision of the Vera Rubin Observatory LSST and Euclid survey will constrain the presence of any additional components of the universe at z_{eq} when EDE needs to be close to its maximum effect.

4.2.3 T_0 tension

In Eq. (52), ω_r is determined by the temperature of today's CMB, measured by the COBE instrument FIRAS in 1996 to be $T_0 = 2.726 \pm 0.0013 K$ (Fixsen 2009) and the equation of state for radiation energy $T(t) = T_0(1 + z(t))^4$. Changes to ω_r in (52) can be reabsorbed into changes to h, so there is some degeneracy between the temperature of universe and H_0. A thought experiment can then be done: what if T_0 was allowed to vary outside FIRAS bounds, keeping other quantities fixed? Ivanov et al. (2020) have considered this, opting to keep fixed ω_i/T_0^3, which corresponds to

the absolute energy scale of components. On combining Planck data with SH0ES, they find the H_0 tension is resolved within in ΛCDM, but with $T_0 = 2.582 \pm 0.033$.

While this is not new physics unless one proposes a new equation of state for radiation, it is a helpful recasting of the tension. First, no balloon or other measurement of T_0 has produced such a low T_0 value (Fixsen 2009). Second, the temperature of the universe may be estimated at earlier epochs via the Sunyaev–Zel'dovich effect, providing an independent probe of $T(z)$. Luzzi et al. (2015) use the Planck SZ cluster catalog to find that deviations from $T(z) = T_0(1 + z)$ are limited to $\sim 3\%$ at 2σ confidence out to $z = 0.94$.

5 Conclusions and buyer's advice

Much effort has been expended by all the teams involved to reduce photometric biases, environmental effects, calibration error, lens mass modelling biases, CMB foreground effects and so on. Nevertheless, there remain some areas where a degree of healthy scientific scepticism might be focused, or improvements might be forthcoming. In the spirit of our Buyer's Guide, we offer our view on these areas in Fig. 18.

The last 20 years have seen the error bars on H_0 shrink from 8% to 1%–2% for the highest precision results.[18] For the late universe, the SH0ES team has calibrated the Cepheid luminosity zero point to 1.0% precision, and the values are consistent whether parallaxes, the DEB distance to the LMC, or the maser distance to NGC4258 are used. The calibration of SN Ia luminosity adds another 1.3% uncertainty, resulting in H_0 to 1.8% accuracy (Riess et al. 2021). For the early universe, the high resolution and sky coverage of Planck results in 0.75% accuracy (Planck Collaboration 2020). This has been characterised as "early versus late" tension, but this would be to ignore the results from the CCHP calibration of TRGBs which are a late universe result intermediate between Planck and SH0ES (Freedman et al. 2019). Additionally, the size of the BAO sound horizon in the late universe when calibrated with early universe BBN constraints is consistent with the CMB value (Addison et al. 2018).

5.1 Systematics or beyond ΛCDM?

Not too surprisingly there is at present no clear answer yet to the question "What is the true value of H_0?". We do, however, think there is a compelling cosmological model, and that is ΛCDM. It has been the best model around for the last 30 years, and no new model has yet presented a convincing case to replace it. Having said that, it is still disturbing that the two main ingredients in ΛCDM, dark matter and dark energy, are not understood. The H_0 derived from the CMB is, therefore, part of a "package deal" which involves other cosmological parameters within the ΛCDM paradigm. Fortunately, the underlying physics of the CMB is well defined—the CMB fluctuation spectrum is a solution of Boltzmann's equation. In contrast, the

---

[18] For a nice illustration of the evolution of Cepheid, TRGB and CMB distances, see Fig. 17 of Freedman et al. (2019).

cosmic ladder measurements of H_0 probe directly this parameter and no others, and the astrophysics of the standard candles used is not fully understood.

Although many plausible extensions of ΛCDM can resolve the tension, none enjoy majority support. They either do not resolve other tensions related to H_0 such as the sound horizon, or seem somewhat contrived. Thus, at best they can be seen as effective theories of some unknown, and possibly more natural, microphysics. Given the lack of a compelling explanation from theory, and a greater understanding of how hard it is to "break" ΛCDM, opinion in the community seems to be shifting. At a conference called "Beyond ΛCDM" in Oslo in 2015, a poll suggested 69% of participants believed new physics the most likely explanation (although we suspect some bias in the voter views given the title of the conference!). However, at a conference in 2021 entitled "Hubble Tension Headache" held virtually by Southampton University, UK, over 50% of participants favoured the explanation that there were still systematics in the data.

5.2 Open issues

Here we comment on some frequently asked questions:

– **Is SH0ES H_0 supported?** Results from Miras, Megamasers and HST Key Project have central values that are consistent with SH0ES, but due to their larger error bars are only in tension with Planck at the $1.5 \sim 3.0\sigma$ level. Results from lensed quasars, TRGBs and surface brightness fluctuations vary between papers, and cannot be regarded as convincingly supportive of SH0ES.

– **Is Planck H_0 supported?** The Planck value for H_0 is corroborated by the Atacama Cosmology Telescope to within 1%, and is also consistent with the earlier WMAP satellite. However, it is in 2.1σ tension with the South Pole Telescope. Planck's H_0 is also supported by BAO from the BOSS and DES galaxy surveys combined with an early universe prior for the baryon density.

– **Are systematics still present?** Regarding Cepheids, considerable recent progress has been made in demonstrating crowding is under control and photometric reduction is free from bias. Nevertheless, potential weaknesses remain in the calibration and environmental corrections of SN Ia, and the calibration of TRGBs has been disputed. Re-analyses of lensed quasars show systematics are not yet under control. The nature of CMB analyses do not lend themselves easily to part-by-part decomposition or independent review, but a full re-analysis of Planck by Efstathiou and Gratton (2019) did not alter cosmological parameters.

– **Is characterisation of "early versus late" helpful?** The strongest evidence in favour of this are the Cepheid results and Planck, as while other results are broadly consistent with this view they are individually more uncertain. However, if the calibration of TRGB were resolved in favour of the CCHP value, it would weaken the case for this view. It is, however, helpful to consider new cosmological models in terms of those modifying the post- or pre-recombination universe, as at present those modifying the pre-recombination universe appear more viable.

- **Is there a convincing theoretical solution?** Simply put, ΛCDM is very hard to break. Proposals need to clear three main hurdles: fitting a wide range of cosmological data, consistency with other predictions of ΛCDM, and providing a convincing Bayesian argument that the model is preferable to ΛCDM. None yet do so. Nevertheless, there are strong hints what a solution might look like, and it is fair to point out that the innovation of Λ in ΛCDM itself solved a previous tension between a low matter density and flat universe.

- **Are TRGBs and gravitational waves offering a way forwards?** It is likely that a consensus on TRGB calibration will be reached in the near-future, and TRGBs promise cleaner fields and better theoretical modelling than Cepheids. They are also ideal for JWST observations, which has four times the J band resolution of the HST/WFC3 H band. Miras are interesting, but are more expensive to observe and harder to calibrate. Gravitational waves offer an inherently transparent and accurate way to measure H_0, once enough bright and dark siren data have been collected by the middle of this decade. If these two were to corroborate SH0ES, it would be strong evidence in favour of the breakdown of ΛCDM. Alternatively, if they agree with Planck, one would have to conclude SH0ES is the outlier.

5.3 Buyer's advice

We would suggest the following in response to the question of "Which H_0 should I use?", depending on the desired application.

- **For cosmological inference**: As the tension in the Hubble parameter is between local direct estimates, and the best-fitted ΛCDM model, the only two possible explanations are unknown systematic effects in the local direct estimate, or that our universe is described by a model different from ΛCDM. If it turns out that there are systematics in the local estimate, it is clear that we should use Planck's value $H_0 = 67.4$ $\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ (Planck Collaboration 2020). If the universe is not ΛCDM, it remains the case that ΛCDM is a great fit to cosmological data with all other parameters consistent with their CMB-derived values. We, therefore, recommend the full package of Planck or Planck-like cosmological parameters should be used, so one can test the the ΛCDM model self consistently. A specific example is in choosing parameters for N-body simulations. It makes sense to select the Planck set of parameters (H_0, Ω_m and σ_8, etc.) so one can test the growth of structure with cosmic time given ΛCDM fit to Planck. In contrast, the ΛCDM model with $H_0 = 73.2$ $\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ is a poor fit to most existing cosmological observations (CMB and galaxy surveys). Another application of using H_0 and other parameters from Planck is when using them as priors for analyses of new data sets, in a Bayesian framework. Here the errors bars of parameters or the full Planck posterior of the probability distribution are important, as they propagate through the parameter marginalisation process. We also maintain that it is not correct to inflate H_0 confidence intervals to "hedge bets", unless one does that in a non-ΛCDM model and fully rederives the posteriors for all other parameters in a Bayesian fashion.

- **For the local universe:**. The SH0ES value of $H_0 = 73.2 \pm 1.3$ kms^{-1} Mpc^{-1} (Riess et al. 2021) using Gaia parallaxes and Cepheids is a good choice, as it is consistent with independent measurements such as Megamasers and Miras and re-analyses of the data (albeit generally starting with the photometric parameters, and not the original imaging) have produced mostly consistent results. Therefore, the SH0ES value should be used for calculations that require an expansion rate or distances alone, but are local enough to have minimal cosmological model dependence. Having said that, the CCHP TRGB result of $H_0 = 69.8 \pm 1.9$ kms^{-1} Mpc^{-1} (Freedman et al. 2019) is noteworthy for its consistency with Planck and the advantages TRGB observations may offer relative to Cepheids. It is very much a case of "watch this space" for developments.
- **For pedagogical purposes:** For the purpose of teaching cosmology, popular talks, or back-of-the-envelope calculations it would make sense to use the intermediate round value of $H_0 = 70$ kms^{-1} Mpc^{-1}. Wherever possible we recommend indicating the dependence of your key results on H_0 either in formulae, or as a method to adjust the key results. Anticipating a future resolution of the tension, this will be of great assistance to future researchers.

5.4 Future prospects

Looking to the future, the Zwicky Transient Facility and Foundation surveys of nearby SN Ia will be very helpful in reducing potential calibration issues. They will do this by resolving the underlying population characteristics, having cleaner selection functions, and providing more galaxies in which to calibrate the distant Hubble flow sample. The early signs of Gaia Early Data Release 3 is that it provides a considerable reduction, but not elimination, of the bias apparent in Data Release 2. But it is not the last word, and the next release is scheduled for 2022. The key point of Gaia is that it addresses lingering concerns about how the low metallicity environment of the LMC, or the maser disk modelling and crowded fields of NGC4258 may affect the calibration of standard candles. In particular, it will be useful to have accurate Gaia parallaxes to Milky Way globular clusters such as ωCen to provide additional calibrators of the TRGB. The JWST will greatly expand the range of TRGB observation, and provide continuity in the case of any further degradation of the ageing HST. The Extremely Large Telescope in Chile, scheduled for first light in 2025, can extend the range of DEB distances to M31 allowing it to contribute its Cepheid and TRGB populations to calibrations.

In the field of gravitational waves, it is perhaps disappointing that there have been no more observations of "bright sirens" like the neutron star merger GW170817 so far. But the commencement of operations at the VIRGO detector in Italy, and recently the KAGRA detector in Japan offer hope that the more frequent event detections and better sky localization, combined with an improved instrumental calibration, will provide a 2% measurement of H_0 within this decade. Many more time-delay lensing systems will be seen in future galaxy surveys, so there is a strong

incentive to resolve remaining systematics in the modelling and speed up the analysis pipeline.

The Simons Telescope in Chile aims to map the polarisation spectrum of the CMB to an order of magnitude higher than Planck, and will start taking data in the next two to three years. Close to 2030, two new CMB Stage 4 telescopes will be operational in Chile and the South Pole, which will further extend the spectral resolution. The depth of these surveys will be able to support or rule out many pre-combination modifications of ΛCDM. Surveys conducted by the Dark Energy Spectroscopic Instrument (DESI), Vera C. Rubin Observatory (previously named LSST), Euclid satellite and Nancy Grace Roman Space Telescope (previously named WFIRST), combined with theoretical progress in the quasi-linear regime of structure formation, will enable the matter-power spectrum to be compared with ΛCDM predictions with much greater depth and resolution.

References

Abbott BP et al (2016) Observation of gravitational waves from a binary black hole merger. Phys Rev Lett 116(6):1–16. https://doi.org/10.1103/PhysRevLett.116.061102. arXiv:1602.03837

Abbott BP et al (2017) A gravitational-wave standard siren measurement of the Hubble constant. Nature 551(7678):85–98. https://doi.org/10.1038/nature24471. arXiv:1710.05835

Abbott BP et al (2017) GW170817: observation of gravitational waves from a binary neutron star inspiral. Phys Rev Lett 119(16):30–33. https://doi.org/10.1103/PhysRevLett.119.161101

Abbott BP et al (2020) Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA. Living Rev Relat 23:3. https://doi.org/10.1007/s41114-020-00026-9. arXiv:1304.0670

Abbott TMC et al (2021) Dark Energy Survey year 3 results: Cosmological constraints from galaxy clustering and weak lensing. https://doi.org/10.1103/PhysRevD.98.043526. arXiv:2105.13549

Abbott TMC, Allam S, Andersen P et al (2019) First cosmology results using type Ia supernovae from the Dark Energy Survey: constraints on cosmological parameters. ApJ 872(2):L30. https://doi.org/10.3847/2041-8213/ab04fa. arXiv:1811.02374

Abbott TMC et al (2018) Dark Energy Survey year 1 results: a precise H_0 estimate from DES Y1, BAO, and D/H data. MNRAS 480(3):3879–3888. https://doi.org/10.1093/MNRAS/STY1939. arXiv:1711.00403

Addison GE, Huang Y, Watts DJ et al (2016) Quantifying discordance in the 2015 Planck CMB spectrum. ApJ 818(2):132. https://doi.org/10.3847/0004-637x/818/2/132. arXiv:1511.00055

Addison GE, Watts DJ, Bennett CL et al (2018) Elucidating ΛCDM: impact of Baryon acoustic oscillation measurements on the Hubble constant discrepancy. ApJ 853(2):119. https://doi.org/10.3847/1538-4357/aaa1ed. arXiv:1707.06547

Agrawal P, Cyr-Racine FY, Pinner D, Randall L (2019) Rock 'n' roll solutions to the Hubble tension. arXiv:1904.01016

Aiola S, Calabrese E, Maurin L et al (2020) The Atacama Cosmology Telescope: DR4 maps and cosmological parameters. J Cosmol Astropart Phys 12:047. https://doi.org/10.1088/1475-7516/2020/12/047. arXiv:2007.07288 [astro-ph.CO]

Alam S, Ata M, Bailey S et al (2017) The clustering of galaxies in the completed SDSS-III Baryon oscillation spectroscopic survey: cosmological analysis of the DR12 galaxy sample. MNRAS 470(3):2617–2652. https://doi.org/10.1093/mnras/stx721. arXiv:1607.03155

Alam S, Aubert M, Avila S et al (2021) Completed SDSS-IV extended Baryon oscillation spectroscopic survey: cosmological implications from two decades of spectroscopic surveys at the Apache point observatory. Phys Rev D 103(8):40. https://doi.org/10.1103/PhysRevD.103.083533. arXiv:2007.08991

Anderson L, Aubourg E, Bailey S et al (2012) The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample. MNRAS 427(4):3435–3467. https://doi.org/10.1111/j.1365-2966.2012.22066.x. arXiv:1203.6594

Arenas DF, Terlevich E, Terlevich R et al (2018) An independent determination of the local Hubble constant. MNRAS 474(1):1250–1276. https://doi.org/10.1093/mnras/stx2710. arXiv:1710.05951

Arendse N, Wojtak RJ, Agnello A et al (2019) Cosmic dissonance: new physics or systematics behind a short sound horizon? A&A 57(639):1

Arenou F, Luri X, Babusiaux C et al (2018) Gaia Data Release 2. A&A 616:A17. https://doi.org/10.1051/0004-6361/201833234

Argon AL, Greenhill LJ, Reid MJ, Moran JM, Humphreys EML (2008) Toward a new geometric distance to the active Galaxy NGC 4258. I. VLBI monitoring of water maser emission. ApJ 659(2):1040. https://doi.org/10.1086/524728

Arun KG, Buonanno A, Faye G, Ochsner E (2009) Higher-order spin effects in the amplitude and phase of gravitational waveforms emitted by inspiraling compact binaries: ready-to-use gravitational waveforms. Phys Rev D 79(10):104023. https://doi.org/10.1103/PhysRevD.79.104023. arXiv:0810.5336

Aver E, Olive KA, Skillman ED (2015) The effects of He i λ10830 on helium abundance determinations. J. Cosmology Astropart. Phys. 2015(7). https://doi.org/10.1088/1475-7516/2015/07/011. arXiv:1503.08146

Aylor K, Hou Z, Knox L et al (2017) A comparison of cosmological parameters determined from CMB temperature power spectra from the south pole telescope and the planck satellite. ApJ 850(1):101. https://doi.org/10.3847/1538-4357/aa947b. arXiv:1706.10286

Beaton RL, Birrer S, Dell'Antonio I, et al (2019) Measuring the Hubble constant near and far in the Era of ELT's. BAAS 51(3):456. arXiv:1903.05035

Bernal JL, Peacock JA (2018) Conservative cosmology: combining data with allowance for unknown systematics. J Cosmol Astropart Phys 07:002. https://doi.org/10.1088/1475-7516/2018/07/002. arXiv:1803.04470

Bernal JL, Verde L, Riess AG (2016) The trouble with H_0. J Cosmol Astropart Phys 10:019. https://doi.org/10.1088/1475-7516/2016/10/019. arXiv:1607.05617

Bianchini F, Wu WLK, Ade PAR et al (2020) Constraints on cosmological parameters from the 500 deg^2 SPTPOL lensing power spectrum. ApJ 888(2):119. https://doi.org/10.3847/1538-4357/ab6082. arXiv:1910.07157

Birrer S, Shajib AJ, Galan A et al (2020) TDCOSMO IV. Hierarchical time-delay cosmography - joint inference of the Hubble constant and galaxy density profile. A&A 165:A165. https://doi.org/10.1051/0004-6361/202038861

Blake C, Kazin EA, Beutler F et al (2011) The WiggleZ Dark Energy Survey: Mapping the distance-redshift relation with baryon acoustic oscillations. MNRAS 418(3):1707–1724. https://doi.org/10.1111/j.1365-2966.2011.19592.x. arXiv:1108.2635

Blake C, Brough S, Colless M et al (2012) The WiggleZ dark energy survey: joint measurements of the expansion and growth history at $z < 1$. MNRAS 425(1):405–414. https://doi.org/10.1111/j.1365-2966.2012.21473.x. arXiv:1204.3674

Blakeslee JP, Jensen JB, Ma CP, Milne PA, Greene JE (2021) The Hubble constant from infrared surface brightness fluctuation distances. ApJ 911(1):65. https://doi.org/10.3847/1538-4357/abe86a. arXiv:2101.02221 [astro-ph.CO]

Blas D, Lesgourgues J, Tram T (2011) The cosmic linear anisotropy solving system (CLASS). Part II: Approximation schemes. J Cosmol Astropart Phys 07:034. https://doi.org/10.1088/1475-7516/2011/07/034. arXiv:1104.2933

Breuval L, Kervella P, Anderson RI et al (2020) The Milky Way Cepheid Leavitt law based on Gaia DR2 parallaxes of companion stars and host open cluster populations. A&A 643:A115. https://doi.org/10.1051/0004-6361/202038633. arXiv:2006.08763 [astro-ph.SR]

Breuval L, Kervella P, Wielgórski P et al (2021) The influence of metallicity on the Leavitt Law from geometrical distances of Milky Way and Magellanic Clouds Cepheids. arXiv:2103.10894

Brout D, Scolnic D (2021) It's dust: solving the mysteries of the intrinsic scatter and host-galaxy dependence of standardized type Ia supernova brightnesses. ApJ 909(1):17. arXiv:2004.10206

Camarena D, Marra V (2021) On the use of the local prior on the absolute magnitude of Type Ia supernovae in cosmological inference. MNRAS 504(4):5164–5171. https://doi.org/10.1093/mnras/stab1200. arXiv:2101.08641

Capozzi F, Raffelt G (2020) Axion and neutrino bounds improved with new calibrations of the tip of the red-giant branch using geometric distance determinations. Phys Rev D 102(8):083007. https://doi.org/10.1103/PhysRevD.102.083007. arXiv:2007.03694

Cardona W, Kunz M, Pettorino V (2017) Determining H_0 with Bayesian hyper-parameters. J Cosmol Astropart Phys 3:056. https://doi.org/10.1088/1475-7516/2017/03/056. arXiv:1611.06088 [astro-ph.CO]

Cerny W, Freedman WL, Madore BF et al (2021) Multi-wavelength, optical (VI) and near-infrared (JHK) calibration of the tip of the red giant branch method based on Milky Way globular clusters. A&A arXiv:2012.09701

Chen GCF, Fassnacht CD, Suyu SH et al (2019) A SHARP view of H0LiCOW: H_0 from three time-delay gravitational lens systems with adaptive optics imaging. MNRAS 490(2):1743–1773. https://doi.org/10.1093/mnras/stz2547. arXiv:1907.02533

Chen HY, Fishbach M, Holz DE (2018) A two per cent Hubble constant measurement from standard sirens within five years. Nature 562(7728):545–547. https://doi.org/10.1038/s41586-018-0606-0. arXiv:1712.06531

Cole S, Percival WJ, Peacock JA et al (2005) The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. MNRAS 362(2):505–534. https://doi.org/10.1111/j.1365-2966.2005.09318.x. arXiv:0501174 [astro-ph]

Cox JP (1960) Theory of stellar pulsation, Princeton series in astrophysics, vol 31, Princeton legacy library edition edn. Princeton University Press

Croft RA (2021) Direct geometrical measurement of the Hubble constant from galaxy parallax: predictions for the Vera C. Rubin Observatory and Nancy Grace Roman Space Telescope. MNRAS 501(2):2688–2703. https://doi.org/10.1093/mnras/staa3769

Das A, Ghosh S (2021) Flavor-specific interaction favours strong neutrino self-coupling. J Cosmol Astropart Phys 7:36. https://doi.org/10.1088/1475-7516/2021/07/038. arXiv:2011.12315

Davis TM, Hinton SR, Howlett C, Calcino J (2019) Can redshift errors bias measurements of the Hubble constant? MNRAS 2957:2948–2957. https://doi.org/10.1093/mnras/stz2652. arXiv:1907.12639

de Jaeger T, Stahl BE, Zheng W et al (2020) A measurement of the Hubble constant from Type II supernovae. MNRAS 496(3):3402–3411. https://doi.org/10.1093/MNRAS/STAA1801. arXiv:2006.03412

Desmond H, Jain B, Sakstein J (2019) Local resolution of the Hubble tension: the impact of screened fifth forces on the cosmic distance ladder. Phys Rev D 100(4):043537. https://doi.org/10.1103/PhysRevD.100.043537. arXiv:1907.03778

Dhawan S, Jha SW, Leibundgut B (2018) Measuring the Hubble constant with Type Ia supernovae as near-infrared standard candles. A&A 609:A72. https://doi.org/10.1051/0004-6361/201731501. arXiv:1707.00715 [astro-ph.CO]

Di Valentino E, Mena O, Pan S et al. (2021) In the Realm of the Hubble tension—a review of solutions. arXiv:2103.01183

Efstathiou G (2020) A Lockdown Perspective on the Hubble Tension. arXiv:2007.10716

Efstathiou G (2021) To H_0 or not to H_0? MNRAS 505(3):3866–3872. https://doi.org/10.1093/mnras/stab1588. arXiv:2103.08723

Efstathiou G, Gratton S (2020) The evidence for a spatially flat Universe. MNRAS Lett 496(1):L91–L95. https://doi.org/10.1093/mnrasl/slaa093. arXiv:2002.06892

Efstathiou GP, Gratton S (2019) A detailed description of the CamSpec likelihood pipeline and a reanalysis of the Planck high frequency maps. arXiv:1910.00483

Einstein A (1917) Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. Sitzungsber Königl Preuss Akad Wissenschaft pp 142–152. https://doi.org/10.1007/BF01591613

Eisenstein DJ, Zehavi I, Hogg DW et al (2005) Detection of the Baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies. ApJ 633(2):560–574. https://doi.org/10.1086/466512. arXiv:0501171 [astro-ph]

Eisenstein DJ, Seo HJ, Sirko E, Spergel DN (2007) Improving cosmological distance measurements by reconstruction of the Baryon acoustic peak. ApJ 664(2):675–679. https://doi.org/10.1086/518712. arXiv:0604362 [astro-ph]

Etherington IMH (1933) Republication of: LX. On the definition of distance in general relativity. Gen Relat Gravit 39(7):1055–1067. https://doi.org/10.1007/s10714-007-0447-x

Falco EE, Gorenstein MV, Shapiro II (1985) On model-dependent bounds on H_0 from gravitational images: application to Q0957+561 A, B. ApJ 289:L1–L4. https://doi.org/10.1086/184422

Feast M (2004) AGB variables as distance indicators. IAU Colloquium 193:304–312. https://doi.org/10.1017/s0252921100010836. arXiv:0309369 [astro-ph]

Feeney SM, Mortlock DJ, Dalmasso N (2018) Clarifying the Hubble constant tension with a Bayesian hierarchical model of the local distance ladder. MNRAS 476(3):3861–3882. https://doi.org/10.1093/MNRAS/STY418. arXiv:1707.00007

Feldman HA, Kaiser N, Peacock JA (1994) Power spectrum analysis of three-dimensional redshift surveys. ApJ 426:23–37

Fitzpatrick E (1999) Correcting for the effects of interstellar extinction. PASP 111(755):63–75. https://doi.org/10.1086/316293. arXiv:9809387 [astro-ph]

Fixsen DJ (2009) The temperature of the cosmic microwave background. ApJ 707(2):916–920. https://doi.org/10.1088/0004-637X/707/2/916. arXiv:0911.1955

Follin B, Knox L (2018) Insensitivity of the distance ladder Hubble constant determination to Cepheid calibration modelling choices. MNRAS 477(4):4534–4542. https://doi.org/10.1093/mnras/sty720. arXiv:1707.01175

Freedman WL (2021) Measurements of the Hubble constant: tensions in perspective. ApJ 919(1):16. https://doi.org/10.3847/1538-4357/ac0e95. arXiv:2106.15656 [astro-ph.CO]

Freedman WL, Madore BF (2010) The Hubble constant. ARA&A 48:673–710. https://doi.org/10.1146/annurev-astro-082708-101829. arXiv:1004.1856

Freedman WL, Madore BF, Hatt D et al (2019) The Carnegie-Chicago Hubble Program. VIII. An independent determination of the Hubble constant based on the tip of the Red Giant Branch. ApJ 882(1):34. https://doi.org/10.3847/1538-4357/ab2f73. arXiv:1907.05922

Friedmann A (1922) Über die Krümmung des Raumes. Z Phys 10:377–386. Translation in Gen Relativ Gravit 31, 1991–2000 (1999). https://doi.org/10.1023/A:1026751225741

Gaia Collaboration, Brown AGA et al (2018) Gaia Data Release 2. Summary of the contents and survey properties. A&A 616. https://doi.org/10.1051/0004-6361/201833051. arXiv:1804.09365

Gerasimovic BP (1928) The absolute magnitudes of long period variable stars. PNAS 14(12):963–968. https://doi.org/10.1073/pnas.14.12.963

Gieren W, Storm J, Konorski P et al (2018) The effect of metallicity on Cepheid period-luminosity relations from a Baade-Wesselink analysis of Cepheids in the Milky Way and Magellanic Clouds. A&A 620. https://doi.org/10.1051/0004-6361/201833263

Gilmore G (2018) Gaia: 3-dimensional census of the Milky Way Galaxy. Contemp Phys 59(2):155–173. https://doi.org/10.1080/00107514.2018.1439700

Handley W, Lemos P (2019) Quantifying tensions in cosmological parameters: interpreting the DES evidence ratio. Phys Rev D 100(4):043504. https://doi.org/10.1103/physrevd.100.043504. arXiv:1902.04029

Handley W, Lemos P (2021) Quantifying the global parameter tensions between ACT, SPT, and Planck. Phys Rev D 103(6):063529. https://doi.org/10.1103/PhysRevD.103.063529. arXiv:2007.08496 [astro-ph.CO]

Hanson RB (1979) A practical method to improve luminosity calibrations from trigonometric parallaxes. MNRAS 186:875–896. https://doi.org/10.1093/mnras/186.4.875

Henning JW, Sayre JT, Reichardt CL et al (2018) Measurements of the temperature and e-mode polarization of the CMB from 500 square degrees of SPTpol data. ApJ 852(2):97. https://doi.org/10.3847/1538-4357/aa9ff4. arXiv:1707.09353

Hinshaw G, Larson D, Komatsu E, et al. (2013) Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Cosmological parameter results. ApJS 208(2). https://doi.org/10.1088/0067-0049/208/2/19. arXiv:1212.5226

Hogg DW (1999) Distance measures in cosmology. arXiv:astro-ph/9905116 [astro-ph]

Holoien TWS, Brown JS, Vallely PJ et al (2018) The ASAS-SN Bright Supernova Catalog—IV. MNRAS 484(2):1899–1911. https://doi.org/10.1093/mnras/stz073. arXiv:1811.08904

Hotokezaka K, Nakar E, Gottlieb O et al (2019) A Hubble constant measurement from superluminal motion of the jet in GW170817. Nat Astron 3(10):940–944. https://doi.org/10.1038/s41550-019-0820-1. arXiv:1806.10596

Howlett C, Davis TM (2020) Standard siren speeds: improving velocities in gravitational-wave measurements of H_0. MNRAS 3815(3):3803–3815. https://doi.org/10.1093/mnras/staa049. arXiv:1909.00587

Hoyt TJ, Beaton RL, Freedman WL et al (2021) The Carnegie Chicago Hubble Program X: tip of the Red Giant Branch Distances to NGC 5643 and NGC 1404. ApJ 915(1):34. https://doi.org/10.3847/1538-4357/abfe5a. arXiv:2101.12232 [astro-ph.GA]

Huang CD, Riess AG, Hoffmann SL et al (2018) A near-infrared period-luminosity relation for Miras in NGC 4258, an anchor for a new distance ladder. ApJ 857(1):67. https://doi.org/10.3847/1538-4357/aab6b3. arXiv:1801.02711

Huang CD, Riess AG, Yuan W et al (2020) Hubble Space Telescope observations of Mira variables in the SN Ia Host NGC 1559: an alternative candle to measure the Hubble constant. ApJ 889(1):5. https://doi.org/10.3847/1538-4357/ab5dbd. arXiv:1908.10883 [astro-ph.CO]

Hubble E (1929) A relation between distance and radial velocity amoung extra-galactic nebulae. PNAS 15(3):168–173. https://doi.org/10.1073/pnas.15.3.168

Humphreys EM, Reid MJ, Moran JM, Greenhill LJ, Argon AL (2013) Toward a new geometric distance to the active galaxy NGC 4258. III. Final results and the Hubble constant. ApJ 775(1). https://doi.org/10.1088/0004-637X/775/1/13

Ivanov MM, Ali-Haïmoud Y, Lesgourgues J (2020) H_0 tension or T_0 tension? Phys Rev D 102(6). https://doi.org/10.1103/PhysRevD.102.063515. arXiv:2005.10656 [astro-ph.CO]

Jang IS, Hoyt T, Beaton R, et al. (2021) The Carnegie-Chicago Hubble Program. IX. Calibration of the Tip of the Red Giant Branch Method in the Mega-Maser Host Galaxy, NGC4258 (M106). ApJ 906(2). arXiv:2008.04181

Javanmardi B, Mérand A, Kervella P et al (2021) Inspecting the Cepheid distance ladder: The Hubble Space Telescope distance to the SN Ia host galaxy NGC 5584. ApJ arXiv:2102.12489

Jimenez R, Loeb A (2002) Constraining cosmological parameters based on relative galaxy ages. ApJ 573(1):37–42. https://doi.org/10.1086/340549. arXiv:0106145 [astro-ph]

Jones DO, Riess AG, Scolnic DM et al (2018) Should Type Ia supernova distances be corrected for their local environments? ApJ 867(2):108. https://doi.org/10.3847/1538-4357/aae2b9. arXiv:1805.05911

Jones DO, Scolnic DM, Foley RJ et al (2019) The foundation supernova survey: measuring cosmological parameters with supernovae from a single telescope. ApJ 881:19. https://doi.org/10.3847/1538-4357/ab2bec. arXiv:1811.09286

Kaiser N, Peacock JA (2016) On the bias of the distance-redshift relation from gravitational lensing. MNRAS 455(4):4518–4547. https://doi.org/10.1093/mnras/stv2585. arXiv:1503.08506

Keenan RC, Barger AJ, Cowie LL (2013) Evidence for a ∼300 megaparsec scale under-density in the local galaxy distribution. ApJ 775(1). https://doi.org/10.1088/0004-637X/775/1/62. arXiv:1304.2884

Kenworthy WD, Scolnic D, Riess A (2019) The local perspective on the Hubble tension: local structure does not impact measurement of the Hubble constant. ApJ 875(2):145. https://doi.org/10.3847/1538-4357/ab0ebf. arXiv:1901.08681

Khetan N, Izzo L, Branchesi M et al (2021) A new measurement of the Hubble constant using Type Ia supernovae calibrated with surface brightness fluctuations. A&A 647:A72. https://doi.org/10.1051/0004-6361/202039196. arXiv:2008.07754 [astro-ph.CO]

Kitaura FS, Rodríguez-Torres S, Chuang CH et al (2016) The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Mock galaxy catalogues for the BOSS Final Data Release. MNRAS 456(4):4156–4173. https://doi.org/10.1093/mnras/stv2826. arXiv:1509.06400

Knox L, Millea M (2020) Hubble constant hunter's guide. Phys Rev D 101(4):043533. https://doi.org/10.1103/PhysRevD.101.043533. arXiv:1908.03663 [astro-ph.CO]

Kourkchi E, Tully RB, Eftekharzadeh S et al (2020) Cosmicflows-4: the Catalog of $\sim 10{,}000$ Tully-Fisher Distances. ApJ 902(2):145. https://doi.org/10.3847/1538-4357/abb66b. arXiv:2009.00733 [astro-ph.GA]

Kreisch CD, Cyr-Racine FY, Doré O (2020) Neutrino puzzle: anomalies, interactions, and cosmological tensions. Phys Rev D 101(12):123505. https://doi.org/10.1103/PhysRevD.101.123505. arXiv:1902.00534

Lahav O, Bridle SL, Hobson MP, Lasenby AN, Sodré L (2000) Bayesian 'hyper-parameters' approach to joint estimation: The Hubble constant from CMB measurements. MNRAS 315(4). https://doi.org/10.1046/j.1365-8711.2000.03633.x. arXiv:9912105 [astro-ph]

Lemaître G (1927) Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. Ann Soc Sci Bruxelles 47:49–59. Translation in Gen Relativ Gravit 45, 1635–1646 (2013) https://doi.org/10.1007/s10714-013-1548-3

Lemos P, Lee E, Efstathiou G, Gratton S (2019) Model independent $H(z)$ reconstruction using the cosmic inverse distance ladder. MNRAS 483(4):4803–4810. https://doi.org/10.1093/mnras/sty3082. arXiv:1806.06781

Lewis A, Challinor A, Lasenby A (2000) Efficient computation of cosmic microwave background anisotropies in closed Friedmann–Robertson–Walker models. ApJ 538(2):473–476. https://doi.org/10.1086/309179

Lutz TE, Kelker DH (1973) On the use of trigonometric parallaxes for the calibration of luminosity systems: theory. PASP 85(October):573

Luzzi G, Génova-Santos RT, Martins CJ, Petris MD, Lamagna L (2015) Constraining the evolution of the CMB temperature with SZ measurements from Planck data. J Cosmol Astropart Phys 2015(9). https://doi.org/10.1088/1475-7516/2015/09/011

Macpherson HJ, Price DJ, Lasky PD (2019) Einstein's Universe: Cosmological structure formation in numerical relativity. Phys Rev D 99(6):063522. https://doi.org/10.1103/PhysRevD.99.063522. arXiv:1807.01711

Macri LM, Ngeow CC, Kanbur SM, Mahzooni S, Smitka MT (2015) Large Magellanic Cloud near-infrared synoptic survey. I. Cepheid variables and the calibration of the Leavitt Law. AJ 149(4):117. https://doi.org/10.1088/0004-6256/149/4/117. arXiv:1412.1511

Madore BF, Freedman WL (2020) Mathematical underpinnings of the multiwavelength structure of the tip of the red giant branch. AJ 160(4):170. https://doi.org/10.3847/1538-3881/abab9a. arXiv:2008.00341 [astro-ph.GA]

Malmquist K (1922) On some relations in stellar statistics. Ark Mat Astron Fys 16(23)

Maoz D, Mannucci F, Nelemans G (2014) Observational clues to the progenitors of Type Ia Supernovae. ARA&A 52:107–170. https://doi.org/10.1146/annurev-astro-082812-141031. arXiv:1312.0628

Moresco M, Jimenez R, Verde L et al (2018) Setting the stage for cosmic chronometers. I. Assessing the impact of young stellar populations on Hubble parameter measurements. ApJ 868(2):84. https://doi.org/10.3847/1538-4357/aae829. arXiv:1804.05864

Mörtsell E, Dhawan S (2018) Does the Hubble constant tension call for new physics? J Cosmol Astropart Phys 2018(9). https://doi.org/10.1088/1475-7516/2018/09/025. arXiv:1801.07260

Mukhanov V (2004) "cMB-Slow" or how to determine cosmological parameters by hand? Int J Theor Phys 43(3):623–668. https://doi.org/10.1023/B:IJTP.0000048168.90282.db. arXiv:0303072 [astro-ph]

Mukherjee S, Lavaux G, Bouchet FR et al (2021) Velocity correction for Hubble constant measurements from standard sirens. A&A 646:1–11. https://doi.org/10.1051/0004-6361/201936724. arXiv:1909.08627

Nicolaou C, Lahav O, Lemos P, Hartley W, Braden J (2020) The impact of peculiar velocities on the estimation of the Hubble constant from gravitational wave standard sirens. MNRAS 495(1):90–97. https://doi.org/10.1093/mnras/staa1120. arXiv:1909.09609

Odderskov I, Hannestad S, Brandbyge J (2017) The variance of the locally measured Hubble parameter explained with different estimators. J Cosmol Astropart Phys 2017(3). https://doi.org/10.1088/1475-7516/2017/03/022. arXiv:1701.05391

Oudmaijer RD, Groenewegen MAT, Schrijver H (1998) The Lutz-Kelker bias in trigonometric parallaxes. MNRAS 294(3):L41–L46. https://doi.org/10.1111/j.1365-8711.1998.01409.x

Padmanabhan N, White M (2008) Constraining anisotropic baryon oscillations. Phys Rev D 77(12). https://doi.org/10.1103/PhysRevD.77.123540. arXiv:0804.0799

Palmese A, Devicente J, Pereira ME et al (2020) A statistical standard siren measurement of the Hubble constant from the LIGO/Virgo gravitational wave compact object merger GW190814 and Dark Energy Survey galaxies. ApJ 900(2):L33. https://doi.org/10.3847/2041-8213/abaeff. arXiv:2006.14961

Percival WJ, Sutherland W, Peacock JA et al (2002) Parameter constraints for flat cosmologies from cosmic microwave background and 2dFGRS power spectra. MNRAS 337(3):1068–1080. https://doi.org/10.1046/j.1365-8711.2002.06001.x

Pesce DW, Braatz JA, Reid MJ et al (2020) The Megamaser Cosmology Project. XIII. Combined Hubble constant constraints. ApJ 891(1):L1. https://doi.org/10.3847/2041-8213/ab75f0. arXiv:2001.09213

Pietrzyński G, Graczyk D, Gieren W et al (2013) An eclipsing-binary distance to the large magellanic cloud accurate to two per cent. Nature 495(7439):76–79. https://doi.org/10.1038/nature11878

Pietrzyński G, Graczyk D, Gallenne A et al (2019) A distance to the Large Magellanic Cloud that is precise to one per cent. Nature 567(7747):200–203. https://doi.org/10.1038/s41586-019-0999-4. arXiv:1903.08096

Planck Collaboration, Ade PAR, Aghanim N, et al. (2014a) Planck 2013 results. XX. Cosmology from Sunyaev–Zeldovich cluster counts. A&A 571:A20. https://doi.org/10.1051/0004-6361/201321521. arXiv:1303.5080 [astro-ph.CO]

Planck Collaboration, Ade PAR, Aghanim N et al (2014b) Planck 2013 results. XVI. Cosmological parameters. A&A 571:A16. https://doi.org/10.1051/0004-6361/201321591. arXiv:1303.5076 [astro-ph.CO]

Planck Collaboration, Ade PAR, Aghanim N et al (2016) Planck 2015 results. XV. Gravitational lensing. A&A 594:A15. https://doi.org/10.1051/0004-6361/201525941. arXiv:1502.01591 [astro-ph.CO]

Planck Collaboration, Aghanim N, Akrami Y et al (2020) Planck 2018 results. VI. Cosmological parameters. A&A 641:A6. https://doi.org/10.1051/0004-6361/201833910. arXiv:1807.06209 [astro-ph.CO]

Rameez M (2019) Concerns about the reliability of publicly available SNe Ia data. arXiv:1905.00221

Reid MJ, Braatz JA, Condon JJ, et al. (2013) The megamaser cosmology project. IV. A direct measurement of the Hubble constant from UGC 3789. ApJ 767(2). https://doi.org/10.1088/0004-637X/767/2/154

Reid MJ, Pesce DW, Riess AG (2019) An improved distance to NGC 4258 and its implications for the Hubble constant. ApJ 886(2):L27. https://doi.org/10.3847/2041-8213/ab552d. arXiv:1908.05625

Riess AG, Macri LM, Hoffmann SL, et al. (2016) A 2.4% determination of the local value of the Hubble constant. ApJ 826(1):56. https://doi.org/10.3847/0004-637x/826/1/56. arXiv:1604.01424

Riess AG, Casertano S, Yuan W et al (2018) New parallaxes of galactic Cepheids from spatially scanning the Hubble Space Telescope: implications for the Hubble constant. ApJ 855(2):136. https://doi.org/10.3847/1538-4357/aaadb7. arXiv:1801.01120

Riess AG, Casertano S, Yuan W et al (2018) Milky Way Cepheid standards for measuring cosmic distances and application to Gaia DR2: implications for the Hubble constant. ApJ 861(2):126. https://doi.org/10.3847/1538-4357/aac82e. arXiv:1804.10655

Riess AG, Casertano S, Yuan W, Macri LM, Scolnic D (2019) Large Magellanic Cloud cepheid standards provide a 1% foundation for the determination of the Hubble constant and stronger evidence for physics beyond ΛCDM. ApJ 876(1):85. https://doi.org/10.3847/1538-4357/ab1422. arXiv:1903.07603 [astro-ph.CO]

Riess AG, Yuan W, Casertano S, Macri LM, Scolnic D (2020) The accuracy of the Hubble constant measurement verified through cepheid amplitudes. ApJ 896(2):L43. https://doi.org/10.3847/2041-8213/ab9900. arXiv:2005.02445

Riess AG, Casertano S, Yuan W et al (2021) Cosmic distances calibrated to 1% precision with Gaia EDR3 parallaxes and Hubble Space Telescope photometry of 75 Milky Way Cepheids Confirm

Tension with ΛCDM. ApJ 908(1):L6. https://doi.org/10.3847/2041-8213/abdbaf. arXiv:2012.08534 [astro-ph.CO]

Rigault M, Copin Y, Aldering G et al (2013) Evidence of environmental dependencies of Type Ia supernovae from the Nearby Supernova Factory indicated by local Hα. A&A 560. https://doi.org/10.1051/0004-6361/201322104. arXiv:1309.1182

Rigault M, Aldering G, Kowalski M et al (2015) Confirmation of a star formation bias in type Ia supernova distances and its effect on the measurement of the Hubble constant. ApJ 802:20. https://doi.org/10.1088/0004-637X/802/1/20

Rigault M, Brinnel V, Aldering G et al (2020) Strong dependence of Type Ia supernova standardization on the local specific star formation rate. A&A 644:A176. https://doi.org/10.1051/0004-6361/201730404. arXiv:1806.03849 [astro-ph.CO]

Ripepi V, Catanzaro G, Molinaro R et al (2020) Period-luminosity-metallicity relation of classical Cepheids. A&A 642:1–7. https://doi.org/10.1051/0004-6361/202038714

Risaliti G, Lusso E (2019) Cosmological constraints from the Hubble diagram of quasars at high redshifts. Nat Astron 3(3):272–277. https://doi.org/10.1038/s41550-018-0657-z. arXiv:1811.02590

Sachs RK, Wolfe AM (1967) Perturbations of a cosmological model and angular variations of the microwave background. ApJ 147:73. https://doi.org/10.1086/148982

Sandage A (1958) Current problems in the extragalactic distance scale. ApJ 127(3):513–526. https://doi.org/10.1017/CBO9781107415324.004. arXiv:1011.1669v3

Schaefer BE (2007) The Hubble diagram to redshift > 6 from 69 gamma-ray bursts. ApJ 660(1):16–46. https://doi.org/10.1086/511742. arXiv:0612285 [astro-ph]

Schombert J, McGaugh S, Lelli F (2020) Using the baryonic Tully–Fisher relation to measure H_0. 60(2):71. https://doi.org/10.3847/1538-3881/ab9d88. arXiv:2006.08615

Schönrich R, McMillan P, Eyer L (2019) Distances and parallax bias in Gaia DR2. MNRAS 487(3):3568–3580. https://doi.org/10.1093/mnras/stz1451. arXiv:1902.02355

Schutz BF (1986) Determining the Hubble constant from gravitational wave observations. Nature 1:310–311. https://doi.org/10.1038/323310a0

Scolnic DM, Jones DO, Rest A et al (2018) The complete light-curve sample of spectroscopically confirmed SNe Ia from Pan-STARRS1 and cosmological constraints from the combined Pantheon sample. ApJ 859(2):101. https://doi.org/10.3847/1538-4357/aab9bb. arXiv:1710.00845

Seljak U, Zaldarriaga M (1996) A line-of-sight integration approach to cosmic microwave background analysis. ApJ 469:437–444. https://doi.org/10.1086/177793

Serenelli A, Weiss A, Cassisi S, Salaris M, Pietrinferni A (2017) The brightness of the red giant branch tip: theoretical framework, a set of reference models, and predicted observables. A&A 606:A33. https://doi.org/10.1051/0004-6361/201731004. arXiv:1706.09910

Shanks T, Hogarth LM, Metcalfe N (2019) Gaia Cepheid parallaxes and 'Local Hole' relieve H_0 tension. MNRAS Lett 484(1):L64–L68. https://doi.org/10.1093/mnrasl/sly239. arXiv:1810.02595

Silk J (1968) Cosmic Black-body radiation and galaxy formation. ApJ 151(February):459

Smith M et al [DES Collaboration] (2020) First cosmology results using type Ia supernovae from the Dark Energy Survey: the effect of host galaxy properties on supernova luminosity. MNRAS 494(3):4426–4447. https://doi.org/10.1093/mnras/staa946. arXiv:2001.11294

Soares-Santos M et al [DES Collaboration, LIGO Scientific Collaboration and Virgo Collaboration] (2019) First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo Binary-Black-hole Merger GW170814. ApJ 876(1):L7. https://doi.org/10.3847/2041-8213/ab14f1. arXiv:1901.01540 [astro-ph.CO]

Soltis J, Casertano S, Riess AG (2021) The parallax of ω Centauri measured from Gaia EDR3 and a direct, geometric calibration of the tip of the red giant branch and the Hubble Constant. ApJ 908:L5. https://doi.org/10.3847/2041-8213/abdbad. arXiv:2012.09196

Sunyaev RA, Zel'dovich YB (1972) The observation of relic radiation as a test of the nature of X-ray radiation from the clusters of galaxies. Comments Astrophys Space Phys 4:173

Szymański MK, Udalski A, Soszyński I, et al. (2011) The optical gravitational lensing experiment. OGLE-III Photometric Maps of the Galactic Bulge fields. Acta Astron 61(2):83–102

Terlevich R, Davies R (1981) The metallicities, velocity dispersions and true shapes of elliptical galaxies. MNRAS 196:381–395

Tonry J, Schneider D (1988) A new technique for measuring extragalactic distances. AJ 96(3):807–815. https://doi.org/10.1086/114847

Tripp R, Branch D (1999) Determination of the Hubble Constant using a two-parameter luminosity correction for Type Ia Supernovae. ApJ 525(1):209–214. https://doi.org/10.1086/307883

Tully RB, Fisher J (1977) A new method of determining distances to galaxies. A&A 54:661–673

Vagnozzi S (2020) New physics in light of the H_0 tension: an alternative view. Phys Rev D 102(2):023518. https://doi.org/10.1103/PhysRevD.102.023518. arXiv:1907.07569

Way MJ (2013) Dismantling Hubble's Legacy? In: Way MJ, Hunter D (eds) Origins of the Expanding Universe: 1912-1932. ASP Conference Series, vol 471. Astronomical Society of the Pacific, San Francisco, pp 97–134

Wong KC, Suyu SH, Chen GCF, et al. (2020) H0LiCOW – XIII. A 2.4 per cent measurement of H_0 from lensed quasars: 5.3 tension between early- and late-Universe probes. MNRAS 498(1):1420–1439. https://doi.org/10.1093/mnras/stz3094. arXiv:1907.04869 [astro-ph.CO]

Yao Y, Miller AA, Kulkarni SR et al (2019) ZTF early observations of Type Ia Supernovae. I. Properties of the 2018 sample. ApJ 886(2):152. https://doi.org/10.3847/1538-4357/ab4cf5. arXiv:1910.02967

Yuan W, Macri LM, He S et al (2017) Large Magellanic cloud near-infrared synoptic survey. V. Period-luminosity relations of Miras. AJ 154(4):149. https://doi.org/10.3847/1538-3881/aa86f1. arXiv:1708.04742

Yuan W, Riess AG, Macri LM, Casertano S, Scolnic DM (2019) Consistent Calibration of the Tip of the Red Giant Branch in the Large Magellanic cloud on the Hubble Space Telescope photometric system and a redetermination of the Hubble Constant. ApJ 886(1):61. https://doi.org/10.3847/1538-4357/ab4bc9. arXiv:1908.00993

Zaldarriaga M, Seljak U (1998) Gravitational lensing effect on cosmic microwave background polarization. Phys Rev D 58(2):6. https://doi.org/10.1103/PhysRevD.58.023003. arXiv:9803150 [astro-ph]

Zehavi I, Zheng Z, Weinberg DH, et al. (2011) Galaxy clustering in the completed SDSS redshift survey: the dependence on color and luminosity. ApJ 736(1). https://doi.org/10.1088/0004-637X/736/1/59. arXiv:1005.2413

Zhang BR, Childress MJ, Davis TM et al (2017) A blinded determination of H_0 from low-redshift Type Ia supernovae, calibrated by Cepheid variables. MNRAS 471(2):2254–2285. https://doi.org/10.1093/mnras/stx1600. arXiv:1706.07573