Cosmology at a crossroads

by David Wiltshire

David Wiltshire is a professor of theoretical physics at the University of Canterbury. His research interests broadly cover general relativity, cosmology and quantum gravity. He obtained his PhD in the 1980s in the group of Stephen Hawking at the University of Cambridge, UK, and after a variety of research and teaching appointments in Italy, UK and Australia returned home to Christchurch in 2001

Since the mid-2000s his particular interest has been the challenge to theoretical cosmology posed by the apparently accelerated expansion of the Universe, or so-called "dark energy". While revisiting old assumptions about cosmological observations and the way the mass distribution in the Universe is averaged, he posits that "dark energy" may in fact be a misidentification of "quasi-local gravitational energy", an aspect of Einstein's theory that we have yet to fully understand. His approach is called "timescape cosmology" and may lead to a reinterpretation of "dark energy" and also possibly of "dark matter" as a modified geometrical theory of gravity on the largest scales in the Universe.

What is the Universe made of?

Cosmology today stands at a unique point in its history. Astronomical observations of ever increasing variety and detail have revealed a vast cosmic web of large complex structures. Yet the theoretical model by which we extract the Universe's expansion history from such data is increasingly challenged as observational precision improves. Estimates of the present expansion rate of the Universe, the Hubble constant, H_0, are now discrepant by up to 5 standard deviations (“5 sigma”) – the gold standard for a discovery in particle physics. Along with other tensions in measures of the expansion history, there are anomalies in the spectrum of tiny ripples in the Cosmic Microwave Background (CMB) – the relic radiation of the Big Bang. Furthermore, images from the James Webb Space Telescope now reveal a Universe with an abundance of galaxies that formed much earlier than expected.

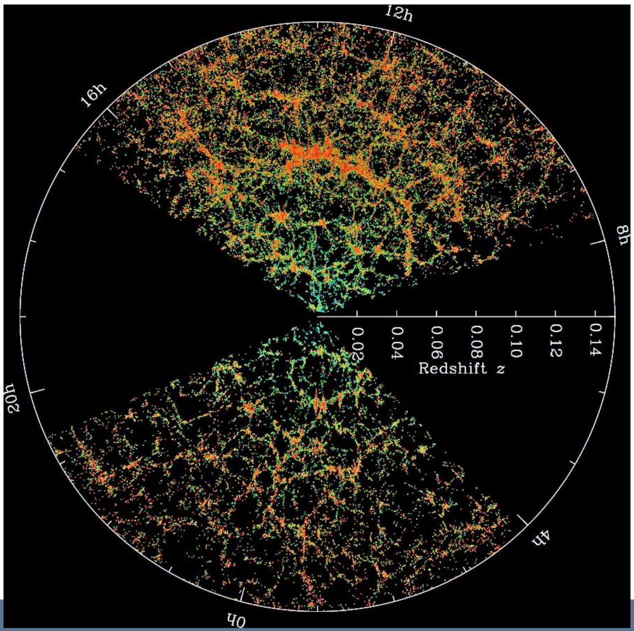

Fig 1. Cosmic web showing the late epoch Universe: voids are surrounded by sheets and threaded by thin filaments of galaxy clusters. Each point represents a galaxy. [Image credit: Sloan Digital Sky Survey]

Cosmology at a crossroads (cont'd)

These expectations are all built on the standard ΛCDM cosmology which – despite the cracks mentioned above – has passed numerous independent observational tests. The standard cosmology does, however, contain two fundamental mysteries at its core:

I. 25% Cold Dark Matter (CDM) which cannot be ordinary `baryonic' matter of which atoms are made, which does not interact electromagnetically, has not been directly detected and is only inferred via indirect gravitational effects as missing mass; and

II. 70% dark energy – a mysterious pressure in the vacuum of space, as exemplified by a cosmological constant, Λ. This is needed to counter the attractive force of gravity, thereby allowing cosmic expansion to accelerate at late epochs, explaining supernova observations.

What 95% of the present epoch Universe is made of still eludes physical explanation.

There is a bigger elephant in the room, however: a 100-year old simplifying assumption put into Einstein's equations – that the Universe expands on average exactly as if all the cosmic structures of Fig. 1 are smoothed into a featureless fluid. At any instant of time the fluid is assumed to be identical everywhere in space – *homogeneous* – and in all directions – *isotropic*. The CMB reveals that the Universe was indeed very smooth when it was 380,000 years old. However, at the present epoch evidence for an average isotropic expansion law is only found on scales larger than 450 million light years, three times the diameter of the most typical voids.

Inhomogeneous cosmology, backreaction and the Timescape

General Relativity (GR) is directly tested in *few-body systems* from the Solar System to supermassive black holes. GR is based on Einstein's tensor field equations in which the curved spacetime geometry is proportional to the energy-momentum generated by matter fields. As John Wheeler said: *Matter tells space how to curve, space tell matter how to move*. In few-body systems the energy-momentum is obtained by an average or coarse-graining of non-gravitational forces only, by well understood procedures: we do not need to worry about equations describing the composition of the Earth or Sun to treat one as a point particle moving in the gravitational field of the other. However, moving from a few bodies to an average geometry for the whole Universe involves fitting one geometry inside another in a complex hierarchy of structures: from stars to galaxies, to galaxy clusters, to filaments and voids, to the Universe.

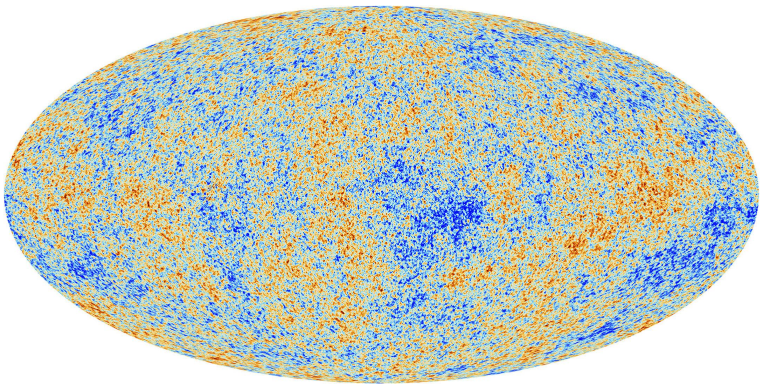

Fig 2*. Tiny ripples in the cosmic microwave background of a few parts in 100,000 of the average temperature. These remain after subtracting the 100 times larger dipole anisotropy (Fig 5), and galactic and point source foregrounds. The temperature fluctuations arise from density fluctuations from which all structures (stars, galaxies, clusters, voids…) grew via gravitational instability.*
[Image credit: ESA, Planck collaboration]

Cosmology at a crossroads (cont'd)

Since Einstein's equations are nonlinear we cannot separate the coarse-graining of geometry from the coarse-graining of matter. The *fitting problem* for general relativistic cosmology is hard and unsolved. However, the basic principles of GR must be part of the solution. Since Einstein's equations directly relate matter and curvature on small scales, limited by the finite speeds of propagation of light and sound, there is no reason to expect the Universe to evolve on average exactly as if curvature is the same everywhere in space. Yet that is exactly the simplifying assumption built into the 100-year old spatially homogeneous and isotropic Friedmann-Lemaître-Robertson-Walker (FLRW) geometries of the standard cosmology.

In the early 2000s Thomas Buchert introduced a pioneering formalism for the cosmological fitting problem in GR. He showed that small scale inhomogeneities may grow to significantly affect average cosmic expansion, giving differences from FLRW evolution – called backreaction. In this setting, cosmic acceleration may actually be a misinterpretation of observations, explaining a coincidence that acceleration appears to begin only at epochs when vast structures come to dominate the cosmic web. In other words, dark energy may be an illusion. Unsurprisingly, backreaction has been much debated. The debate is muddied by the fact that some theorists consider alternative averaging schemes to Buchert's which automatically guarantee backreaction to be insignificant.

Buchert's mathematical formalism can be interpreted in many different ways, and in itself it offers no explanation as to why the Universe should have a close to average isotropic expansion despite observed inhomogeneities.

This fact, which the standard cosmology assumes rather than explains, demands an explanation. I offered one in 2007 by revisiting the first principles of General Relativity in extending Einstein's classic 1907 thought experiments to the case of expanding space. The result is an extension of Einstein's Strong Equivalence Principle to cosmological averages which I call the Cosmological Equivalence Principle, and an observationally viable cosmology without dark energy, the Timescape.

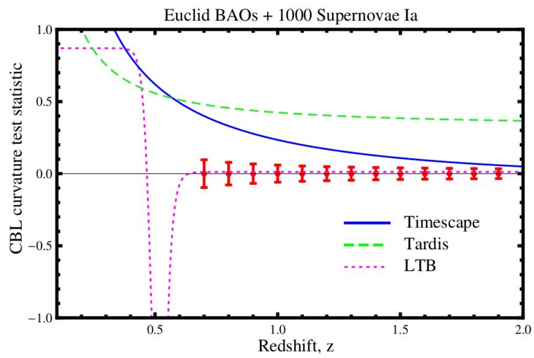

***Fig. 3**. FLRW curvature null test. Simulated ΛCDM data (red points) for the Euclid satellite compared to 2 falsifiable backreaction model predictions – Timescape, Tardis – and an unfalsifiable inhomogeneous Lemaître-Tolman-Bondi (LTB) model. [Image credit: D Sapone et al, 2014]*

In the presence of spatially varying curvature generated by growing density gradients at late epochs, I claim there are choices of regional rulers and clocks that nonetheless maintain an average isotropic expansion. The concepts of elapsed time, and age of the Universe, vary between observers in bound structures and the statistical volume average. This is the notion of a Timescape. The Timescape cosmology has passed many of the same key tests that the standard ΛCDM cosmology does. Since ΛCDM is empirically a good fit in many tests, any successful model will inevitably have differences that are small; for Timescape the differences in the average expansion compared to ΛCDM models are 1-3% at any given distance.

Cosmology at a crossroads (cont'd)

ESA's Euclid satellite will have the precision to test these differences over the next 6 years. It will determine the angular size of a cosmic standard ruler, the *Baryon Acoustic Oscillation* (BAO) scale: the echo of the peak density of sound waves in the primordial plasma, as reflected in numbers of galaxies that grew from that initial excess density. Looking back in time over many epochs of cosmic history, Euclid will determine how this angular scale evolved as structures grew. It will have the precision to directly test whether average curvature is spatially constant as in the FLRW models, or not as in backreaction scenarios including the Timescape and Tardis models (Fig. 3).

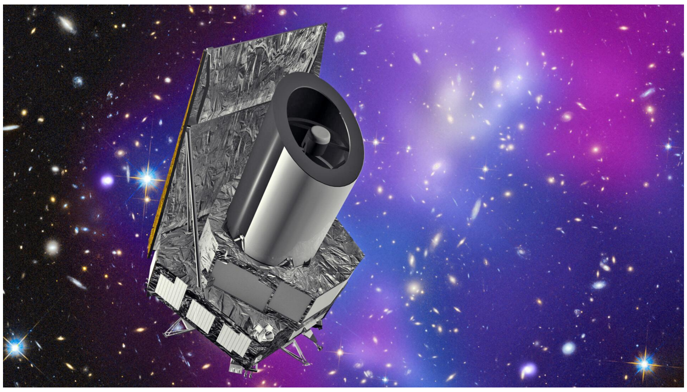

Fig. 4. *ESA's Euclid satellite was originally set to launch in 2022 on a Soyuz rocket, but due to world events will now be launched on a Falcon 9 rocket by SpaceX in 2023. [Image credit: ESA, Euclid Consortium]*

While Euclid will enable precision measurements of the expansion history that definitively test the FLRW assumption, if the standard FLRW model is incorrect then there should be other definitive signatures of inhomogeneous cosmology. Indeed, there is such an anomalous signature, recognised as a 2 to 3 sigma tension for over a decade, which new observations have pushed over the crucial 5 sigma threshold last year.

Motion versus expansion: the anomalous radio galaxy and quasar dipole

The standard cosmology accounts for inhomogeneities by small perturbations of the average FLRW solution, and their nonlinear evolution using large N-body computer simulations, almost exclusively with only Newtonian gravity. It is assumed that all spatial variations of cosmic expansion can be reduced to uniform FLRW expansion in a *cosmic rest frame*, plus relative local peculiar motions, so-called *boosts*, of all galaxies. Local boosts are calculated purely in special relativity, and do not require full GR. Einstein's Strong Equivalence Principle guarantees that in GR we can always perform arbitrary boosts at a source and observer. What is not required by GR, however, is that after subtracting the boosts the average propagation of light from source to observer should follow paths predicted by a FLRW solution. That is the 100-year old ad hoc assumption that backreaction models challenge.

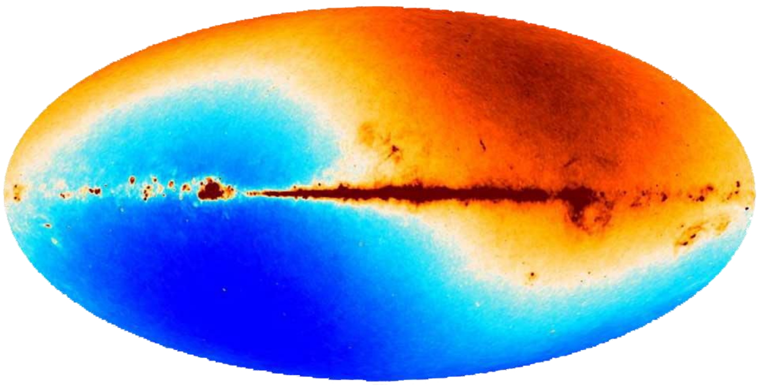

Fig. 5. *Dipole anisotropy of the cosmic microwave background, conventionally interpreted as purely due to our motion towards the hotspot. The horizontal dark stripe is emission from our Galaxy. [Image credit: ESA, Planck collaboration]*

Cosmology at a crossroads (cont'd)

The remarkable isotropy of the CMB is conventionally taken as evidence for a cosmic rest frame. The CMB radiation has an ideal black body spectrum peaked at microwave frequencies with a temperature of 2.725 Kelvin, with the same value anywhere on our sky up to parts of one in a thousand. Furthermore, the milli-Kelvin fluctuations have the characteristic form of a dipole, just as the special relativistic Doppler effect would predict if our motion with respect to the cosmic rest frame is 371 km/s in the direction of the constellation Leo. This must include components of motion within the largest structure to which we are gravitationally bound: the Local Group of galaxies. Space does not expand within the Local Group. By vector addition, once we have subtracted the motion of the Sun within the Milky Way, and of the Milky Way within the Local Group, we find that the Local Group of galaxies must be moving at 635 km/s in the direction of the constellation Hydra. Such putative Local Group motion is on a scale on which the Universe is expanding.

A catch in GR is that differential cosmic expansion on small cosmological scales (less than 400 million light years) combined with our known motion within the Local Group, can also produce a leading CMB dipole anisotropy very much like the one observed. For realistic models differences between a non-kinematic anisotropy and a purely kinematic anisotropy only show up at the level of a few percent of the dipole, on large angular scales. But this is precisely the amplitude of the primordial fluctuations seen in Fig. 2. Furthermore, the CMB quadrupole and other large angle multipoles have puzzling anomalies whose statistical significance increased with increasing observational precision, culminating in the results of the Planck collaboration.

Ten years ago we found in an analysis of 4534 galaxies that average "local" cosmic expansion is actually significantly more uniform in the rest frame of the Local Group rather than in the putative rest frame of the CMB.

The kinematic nature of the CMB dipole, or indeed of any isotropic background of very distant sources, can be directly tested via the predictions of special relativistic aberration and modulation. For the CMB the direction of the dipole determined by the Planck team in 2013 using this method was found to be consistent with the observed dipole when the analysis was confined to angles less than 1 degree across. However, the aberration dipole direction moves across the sky to point along an axis associated with the known anomalies when only large angles are considered.

An equivalent test on the distribution of distant radio galaxies and quasars has now reached a very high sensitivity with an analysis of 1.36 million quasars and 0.5 million radio galaxies, in a study led by Nathan Secrest of the US Naval Observatory. These reveal dipoles of amplitudes 2 and 3 times larger than the kinematic expectations, pointing 26° and 45° away from the CMB dipole direction. The two results combined give a 5.1 sigma disagreement from the expectation of the standard cosmology. Published last October in *Astrophysical Journal Letters* under the title *"A Challenge to the Standard Cosmological Model"* this is perhaps the strongest evidence yet of the need for a paradigm shift.

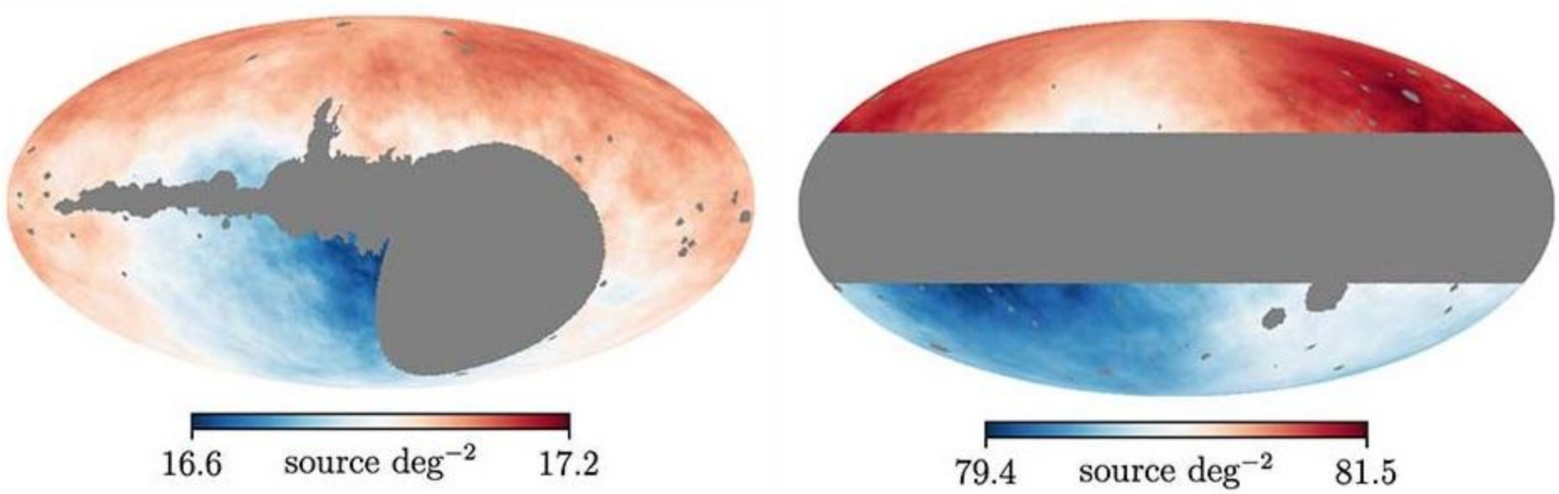

Fig. 6. The smoothed sky map of NVSS radio sources (left) and WISE quasars (right) exhibiting the characteristic dipole anisotropy due to aberration which is expected due to our local motion. The amplitude and direction relative to the CMB dipole are anomalous, however. Some parts of the sky have been masked to ensure uniformity of the source counts and to block out foregrounds. [Image credit: N Secrest et al, 2022]

The statistical significance of this result is naturally being debated. In astrophysics there can always be unaccounted for systematic biases. All we actually observe are wavelengths and intensities of radiation, their time series and angles on the sky. There are always statistical selection biases arising from the limitations of our telescopes. Furthermore, the light was produced by processes that are often impossible to recreate in any terrestrial laboratory. It then travels to us across the vastest distances in the Universe through a heap of mess in between. When presented with a puzzle, a conservative observer will naturally seek an explanation in an unaccounted observational bias, as that is so often the cause.

As a conservative theorist, however, I am deeply aware that Einstein did not provide final foundational answers in applying his theory to cosmology. The FLRW assumption is an ad hoc one theoretically, and a non-kinematic dipole is the natural result of any cosmology that breaks this assumption.

Computational cosmology with numerical General Relativity

While backreaction models can make predictions about average expansion (Fig 3) and while a non-kinematic anisotropy is expected in such models, calculations of the precise amplitude and direction of a feature such as Fig 6 require numerical simulations using full GR.

Large numerical simulations have been undertaken in the standard cosmology since the 1990s. While sophisticated in treating matter these simulations still use Newton's gravity theory, rather than GR, with cosmic expansion scaled by the FLRW solutions. Although Einstein's field equations are 107 years old, it took 90 years to fully implement them numerically. Computational challenges due to the complex nonlinearities of the field equations are compounded by intrinsic physical ambiguities about how we split space and time. The 2-body problem was only solved in the mid-2000s. Cosmology in GR was computationally too hard until very recently, even with the largest supercomputers.

Cosmology at a crossroads (cont'd)

Cosmology poses new challenges for numerical relativity over and above those of few-body systems. It is only in the last 7 years that new techniques implementing the full Einstein equations in cosmological simulations have been pioneered by a handful of researchers worldwide. One of these pioneers, Hayley Macpherson, was awarded the Charlene Heisler Prize by the Astronomical Society of Australia for her PhD thesis from Monash University in 2020. Now a NASA Einstein Fellow at the University of Chicago, Macpherson is a co-supervisor (by distance) of my PhD student Michael Williams at the University of Canterbury. He is currently investigating the self-consistency of the standard ΛCDM cosmology within GR. There are still many open questions as to whether structure grows in the same way in the standard cosmology when we go beyond Newtonian gravity.

Such questions need to be understood before we can begin to tackle the larger challenges posed by a complete paradigm shift.

Simulations begin with an initial spectrum of density fluctuations consistent with the CMB anisotropies (Fig 2) which are then evolved forward in time by Einstein's equations (Fig 7). Once we have resolved a slew of technical questions concerning simulations with standard model initial conditions, our ultimate aim is to change the initial conditions. Removing dark energy while including an initial very small amount backreaction at the CMB epoch, will allow us to validate or refute the Timescape scenario and to directly tackle questions such as those posed by the anomalous dipole in the radio galaxies and quasars.

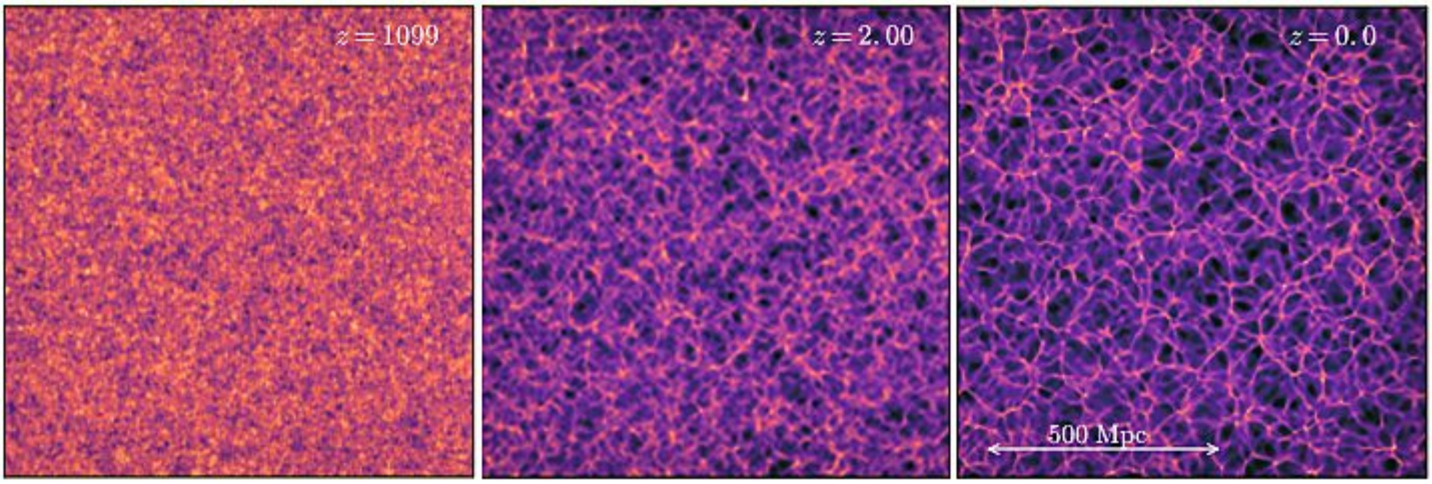

Fig. 7. Emergence of a cosmic web in a cosmological simulation using General Relativity. Panels show a 2--dimensional slice of the simulated evolving density distribution: from left, 300,000 years after the Big Bang to right, a universe similar to ours today. Dark regions are void of matter, and lighter purple regions are more dense. [Image credit: H J Macpherson et al, 2019]

Cosmology at a crossroads (cont'd)

The outlook

Changing the foundations of a cosmological model on which many decades of research and thousands of careers are built is not easy when it comes to securing funding, even with a proposal which seeks a deeper understanding of unsolved open questions in our best theory of gravity, General Relativity. To actually shift the paradigm requires convincing many more theorists to take up these difficult questions. Unfortunately, many theorists find it easier to either invent new matter fields that have never been observed, or to even modify Einstein's gravity in ad hoc ways, while keeping the FLRW assumption because it is simple. If observational tensions and anomalies continue to grow, there is likely to be a tipping point. If my assessment is correct, that tipping point will come within the next decade.

What would a new paradigm look like? As with any foundational revolution in science, much of that is beyond our present imagination. One thing is clear, however. A non-kinematic differential expansion affecting the CMB dipole (Fig. 5) at the level of 1% would require us to redraw the CMB sky map (Fig. 2) since the primordial fluctuations are of the same order. Differences would only occur on the very largest angles subtended on our sky by nearby voids, sheets and filaments in our cosmic "back yard" (less than 400 million light years away). CMB temperature differences on angles less than 1 degree apart would still be the same statistically, with little effect on many parameters that are important for modelling the formation of the first stars and galaxies. However, to the eye the biggest blobs in Fig 2 could well be in different places! This sky map is sometimes referred to as the "baby photo of the Universe". Consequently, we can reasonably expect that a new paradigm would change the face of the Universe in quite a literal sense.

Fo further information see:
http://www2.phys.canterbury.ac.nz/~dlw24/