The isotropic blackbody CMB as evidence for a homogeneous universe

Clifton, T; Clarkson, C; Bull, P

The isotropic blackbody CMB as evidence for a homogeneous universe

Timothy Clifton,[1] Chris Clarkson,[2] and Philip Bull[3]

[1]*School of Physics and Astonomy, Queen Mary University of London, UK.*
[2]*Department of Mathematics and Applied Mathematics, University of Cape Town, South Africa.*
[3]*Department of Astrophysics, University of Oxford, UK.*

The question of whether the Universe is spatially homogeneous and isotropic on the largest scales is of fundamental importance to cosmology, but has not yet been answered decisively. Surprisingly, neither an isotropic primary CMB nor combined observations of luminosity distances and galaxy number counts are sufficient to establish such a result. The inclusion of the Sunyaev-Zel'dovich effect in CMB observations, however, dramatically improves this situation. We show that even a solitary observer who sees an isotropic blackbody CMB can conclude that the universe is homogeneous and isotropic in their causal past when the Sunyaev-Zel'dovich effect is present. Critically, however, the CMB must either be viewed for an extended period of time, or CMB photons that have scattered more than once must be detected. This result provides a theoretical underpinning for testing the Cosmological Principle with observations of the CMB alone.

The current concordance model of cosmology is based on the homogeneous and isotropic Friedmann-Lemaître-Robertson-Walker (FLRW) solutions of Einstein's equations. The high degree of symmetry assumed in these solutions makes them sufficient to explain the near perfect isotropy of the Cosmic Microwave Background (CMB) and other astrophysical observables, but it remains to be demonstrated whether or not they are the *only* spacetime geometries that are compatible with the data. This question is particularly pertinent due to the apparent necessity that more than 95% of the matter content of the Universe must be in the form of dark energy and dark matter in order for the concordance model to be made compatible with observations. The inferred existence of these substances holds such profound consequences for our understanding of basic physics that establishing the validity of the assumed FLRW geometry is now an imperative. So, what observables are required in order to prove the universe is FLRW on large scales?

An important step toward answering this question was provided by Ehlers, Geren and Sachs (EGS) [1], and later fleshed out by others [2, 3] (see [4] for a review). These authors used the Copernican Principle, that we are typical observers, to show that isotropy of the CMB about every point in a region of spacetime is only possible if the geometry of spacetime in that region is spatially homogeneous and isotropic. This result is perturbatively stable in the sense that *near* isotropy of the CMB implies *near* homogeneity of spacetime, although this requires extra assumptions about unobservable quantities [5–8]. An alternative proof of spatial homogeneity using luminosity distances, that also relies on the Copernican Principle, was found by Hasse and Perlick [9]. While compelling these theorems all require observations to be made at *all* points in a region of spacetime to make definite conclusions. Isotropy of the CMB on our own sky is not even sufficient to determine that our local region of space is isotropic around us [10].

Alternatively, the authors of [11] have shown that in order to determine whether the Universe is isotropic around us it is necessary and sufficient to have isotropic observations of luminosity distances, number counts, lensing, and angular peculiar velocities at every redshift, and in every direction. To then determine spatial homogeneity requires an extra independent observable beyond these four, unless one is prepared to specify the value of Λ *a priori* (assuming dark energy is due to the cosmological constant [33]) [12]. While this prescription for determining spatial homogeneity and isotropy has the important quality of relying solely on directly observable quantities, rather than the Copernican Principle, it also requires large amounts of information from a number of different observables.

Here we show that inclusion of the the Sunyaev-Zel'dovich (SZ) effect when considering CMB observations allows one to retain the minimal observational requirements of EGS, while removing the assumption of the Copernican Principle. The SZ effect is due to the scattering of CMB photons by charged matter, and has already been shown to be a powerful tool for constraining radial inhomogeneity within the class of cosmological models constructed from the Lemaître-Tolman-Bondi solutions [13–22]. We extend these previous studies to consider the potential of the SZ effect to act as proof of FLRW geometry, rather than simply as a tool for constraining particular deviations away from it. This results in a stronger statement than that of EGS, as it requires observations made by only a single observer, rather than from all observers in a region of spacetime. It is a much less demanding statement than the result of [11], as it requires observations of the CMB only (although the CMB must be viewed for an extended period of time, or photons that have scattered more than once must be detectable).

The SZ effect is often divided into two different contributions; the thermal SZ effect (tSZ) [23] and the kinematic SZ effect (kSZ) [24]. The former of these describes

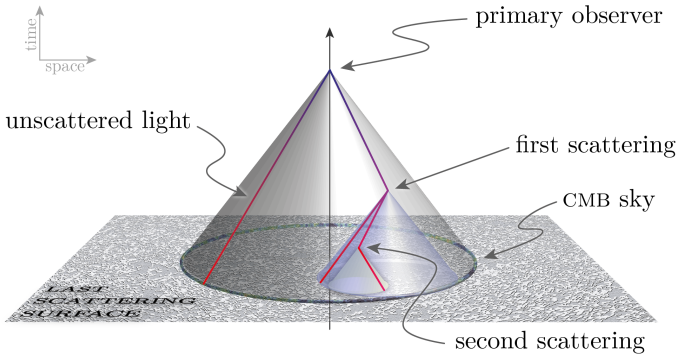

Figure 1: The kSZ effect provides information about the CMB sky at other points on our past lightcone.

the transfer of energy from the hot electrons in the intracluster medium to the cooler photons of the CMB. This is the easiest of the two effects to detect observationally, but is the least important for our current considerations. We will assume here that the tSZ effect is well understood, and can be removed from the CMB signal along with other unwanted foreground sources (a process that will certainly be complicated by relativistic corrections).

The kSZ effect also alters the spectrum of the scattered light, but this is due to the anisotropy seen in the CMB sky of the scatterer rather than any transfer of energy from baryons to photons in the baryon rest frame. In an FLRW spacetime, any such anisotropy is due to the peculiar motion of the scatterer. In the rest frame of the scatterer, the re-scattered light then maintains the same distribution function it had before the scattering event (all other changes being encapsulated in the tSZ), so that an observer in the rest frame of the CMB must see radiation that undergoes a Lorentz boost after scattering. For the case of blackbody radiation this corresponds solely to a change in temperature of the scattered radiation. This mechanism therefore provides, in effect, a set of mirrors that allows us to view the CMB from different locations. We shall therefore refer to the light scattered into our line of sight as being *reflected* by the scatterer (which we will refer to as a cluster, for simplicity).

The picture described above is valid in an FLRW universe with an isotropic radiation field, but in the present study this is exactly the thing we want to prove the existence of. It is therefore necessary to generalise the existing concept of the kSZ effect.

The picture we have for this generalised scenario is illustrated in Fig. 1. As we look back along our past null cone we will see the reflecting clusters, whose own past null cones coincide with ours in one direction, but otherwise crosses the last scattering surface within our causal past. If there exist sufficiently many clusters, we will receive photons from every part of the last scattering surface that we are causally connected to, rather than from just the single sphere that we observe directly. We assume that the formation of the last scattering surface proceeds in thermal equilibrium, so that the emitted radiation is blackbody, and that the Universe is optically thin at all times and everywhere after last scattering. Scattering off the clusters can then result in a possible temperature change, but the spectrum must remain a blackbody as Lorentz transformations at the point of reflection, and cosmological evolution in an arbitrary spacetime, both preserve the form of a blackbody spectrum. Also illustrated in Fig. 1 is the possibility of photons being scattered off two clusters before they reach us, which we will return to later.

Let us denote the incident temperature in each direction on the reflecting cluster's sky as $T_i = T_i(\theta, \phi, z)$, where θ and ϕ are spherical polar coordinates on their sky, chosen such that $\theta = \pi$ is the direction of the eventual observer (us), and z is the redshift of the cluster on the eventual observer's sky. The occupation number of photons received from a particular direction (θ, ϕ) on the cluster's sky can be written as $N_i = \mathcal{B}(\nu, T_i)$, where $\mathcal{B}(\nu, T) = (e^{\nu/T} - 1)^{-1}$, are the occupation numbers of a blackbody spectrum with frequency ν, and where we have set $k_B = h = 1$. The fraction of light that is reflected towards the observer from every direction on the cluster's sky is given by the Thomson cross-section, which, after the effects of the tSZ effect have been removed, gives the occupation number of the reflected light in the rest frame of a particular cluster as $N_r(\nu, T_i, z) = \frac{3}{16\pi} \int \tau(1+\cos^2\theta)\mathcal{B}(\nu, T_i) \sin\theta d\theta d\phi$, where $\tau \ll 1$ is the electron-scattering optical depth of the cluster, which is assumed to fill the telescope beam. We now want to know the conditions on incident radiation on the cluster, $T_i = T_i(\theta, \phi, z)$, for the sum of the reflected light and the unscattered light to have a blackbody spectrum when it is observed at $z = 0$. Recall that blackbody spectra are unchanged after propagating though spacetimes with arbitrary curvature, up to a change in temperature by one factor of redshift [25], so we can write the observed temperature of any blackbody distribution as $\bar{T} = T/(1 + z)$ (where z is the redshift at which it had temperature T). For a continuous distribution of matter we can also write the reflected radiation in some interval Δz along one of our own past-directed null geodesics as $N_r(\nu, T_i, z)\Delta z$. The distribution function of photons that make it to us is then given by $N_{\rm tot} = \mathcal{B}(\nu, \bar{T}_c) + \int N_r(\nu, \bar{T}_i, z)dz - \int N_r(\nu, \bar{T}_c, z)dz$, where $T_c(z) = T_i(0, 0, z)$ is the temperature of the unscattered light at redshift z [34]. The first term on the RHS of this equation is the contribution from the unscattered CMB, the second term is CMB radiation that is scattered towards us (that we would otherwise not be able to observe), and the third term is the CMB radiation that is scattered away from us (that would otherwise reach us in the absence of any scattering). For $N_{\rm tot}$ to be a blackbody with some temperature T_0 we then require $N_{\rm tot} = \mathcal{B}(\nu, T_0)$. If we change variables to

$x = -\cos\theta - \frac{1}{3}\cos^3\theta$, and expand each term in a power series around its temperature, we get

$$\sum_{n,k} c_n \binom{n}{k} (-1)^k \nu^{n-k} \left[\left(\bar{T}_c^{k-n} - T_0^{k-n}\right) + \int \frac{3\tau}{16\pi} \left(\bar{T}_i^{k-n} - \bar{T}_c^{k-n}\right) dx d\phi dz \right] = 0, \quad (1)$$

where $c_n = (-1)^n e A_n(e)/n!(e-1)^{n+1}$, and $A_n(x)$ are the Euler polynomials. It can be shown that $A_n(e)$ are positive definite, so that the c_n have sign $(-1)^n$. Now, Eq. (1) must be true for each value of $n-k$, as nothing here is a function of ν except ν^{n-k} itself. For $n-k=j$ we therefore have

$$\sum_{n=0}^{\infty} c_n \binom{n}{j} (-1)^{(n-j)} \left[\left(\bar{T}_c^{-j} - T_0^{-j}\right) + \int \frac{3\tau}{16\pi} \left(\bar{T}_i^{-j} - \bar{T}_c^{-j}\right) dx d\phi dz \right] = 0, \quad (2)$$

for all $j \neq 0$. In the $j=1$ and $j=2$ cases, we also have $\sum_{n=0}^{\infty} n c_n (-1)^{n-1} \neq 0$ and $\sum_{n=0}^{\infty} n(n-1) c_n (-1)^{n-2} \neq 0$, as $c_n(-1)^{n-1} < 0$ and $c_n(-1)^{n-2} > 0$ for all n. It must then be the case that

$$\left[1 - \frac{3}{16\pi}\int \tau dx d\phi dz\right] \left(\frac{1}{\bar{T}_c} - \frac{1}{T_0}\right)^2 + \int \frac{3\tau}{16\pi} \left(\frac{1}{\bar{T}_i} - \frac{1}{T_0}\right)^2 dx d\phi dz = 0. \quad (3)$$

For $\tau \neq 0$ and $\int \tau dx d\phi dz < 16\pi/3$ we therefore have that $\bar{T}_i^{-1} = \bar{T}_c^{-1} = T_0^{-1}$. The first of these conditions is that there should exist scatterers everywhere, and the second is that the amount of reflected radiation must be less than the amount of the incident radiation (from the Beer-Lambert law). We therefore have that $T_i(\theta,\phi,z) = T_c$ for every θ and ϕ, at every z where $\tau \neq 0$. We also have that $T_c = (1+z)T_0$, so that the temperature of the observed CMB must be the same as the emitted CMB, up to a factor of redshift. This result shows that the CMB must be isotropic about every reflecting cluster, and is essentially due to the fact that blackbody spectra of different temperatures cannot be summed to give another blackbody spectrum [26].

This result tells us that if the CMB was emitted from a thermal process as a blackbody, and is observed as a blackbody today, then the CMB sky at every point on our past lightcone must be isotropic. Any anisotropies at any point on our past light cone would cause distortions in the spectrum of radiation we observe. Surprisingly, this is not yet restrictive enough to require FLRW geometry.

Up to this point, we have only considered the CMB sky of observers at other points on our past null cone. This is not sufficient to establish either homogeneity or isotropy of space around any point, however, as we also require information about derivatives of geometric quantities and the matter content of the Universe in order to propagate information off our past null cone.

The starting point for this is the Boltzmann equation for photons, which in general involves a collisional term for the Thomson scattering. This term is proportional to changes in the distribution function [27], and as we have shown that a vanishing kSZ effect implies isotropy of the CMB about every cluster, this means that the collision term must vanish at every cluster where the kSZ effect vanishes. Hence, it is sufficient for us to consider the collisionless Liouville equation. This tells us that if every time-like observer following a congruence u^a sees an isotropic radiation field then u^a must be (parallel to) a conformal Killing vector, and the spacetime must be conformally stationary [1]. The anisotropic pressure evolution equation that is derived from the quadrupole of the Liouville equation then tells us that u^a must be shear-free, and that the acceleration $A_a = u^b\nabla_b u_a$ and expansion rate $H = \frac{1}{3}\nabla_a u^a$ must satisfy $\nabla_{[a}\left(A_{b]} - H u_{b]}\right) = 0$. One now needs to make assumptions about the matter content in order to make further statements. For an irrotational, geodesic perfect fluid, it follows that the spacetime must be FLRW [28]. The radiation-only case is the original EGS result [1]. In the case of a mixture of dust, radiation and dark energy in the form of a scalar field (for which Λ is a special case) a little more work is required [3] because one cannot assume *ab initio* that the gradient of the scalar field is aligned with the dust observers, but isotropy implies it must be and so the result still holds. The EGS theorem holds in general scalar-tensor theories of gravity [29].

The final stage of our result therefore requires us to show that the CMB must be isotropic inside our past lightcone, as well as on it. We can see two possible ways of doing this:

(1). If we observe the CMB for a finite interval of time [30]. This would allow us to receive information about the CMB sky of all observers in the 4-dimensional region of spacetime swept out by our past null cone over this interval. If no kSZ effect is measured at any time, then one can infer that the entire region is filled with clusters that see isotropic CMB radiation. The region must therefore have FLRW geometry, and taking any surface within it as an initial Cauchy surface, we can establish that our entire causal past must also be FLRW.

(2). If we can observe CMB radiation that has been scattered more than once [31, 32], as was suggested in the original paper by Sunyaev and Zel'dovich [24]. This situation is illustrated by the existence of the 'second scatterer' in Figure 1. If such scattering is observable the CMB sky of the second scatterers must also be isotropic if we are still to observe the re-scattered CMB photons as being a blackbody (the proof of this can be found in the Appendix). Note that only two scatterings are

required, as this is sufficient to show that the CMB must be isotropic around every point in our causal past.

The situation we have considered here is of course highly idealized: The CMB is neither exactly isotropic nor blackbody, and our treatment of the scattering events themselves is also idealized. In reality, we see only a limited number of scatterings that must necessarily only happen at relatively late times (when structure has formed). Also, the removal of the tSZ effect will undoubtedly always be imperfect, as will the subtraction of other foreground sources. Furthermore, we have been somewhat optimistic in considering that it may be possible to observe the CMB for an extended period of time, or that double scattering events can be detected. Nevertheless, we have demonstrated what is required to show that the Universe is FLRW using the CMB alone, without assuming anything about the symmetries of spacetime on the largest scales. Our result holds provided dark energy can be described as a scalar field, and holds for general scalar-tensor theories of gravity too.

To make these ideas more realistic they need to be shown to be perturbatively stable, which is non-trivial [5–7]. An application to the real universe will also require careful consideration of the consequences of imperfect observations and noncontinuous scatterers. We leave this for future work.

Appendix: Multiple scatterings.

Let us denote quantities evaluated at the primary observer (us) with a subscript 0, and those evaluated at the first and second scatterings with A and B, respectively. The redshift of a first scatterer, as measured by the primary observer, is then z_A, and the redshift of a second scatterer, as measured by a first scatterer, is z_B. Angular coordinates on the sky of the first and second scatterers will be written as (θ_A, ϕ_A) and (θ_B, ϕ_B). The temperature of CMB radiation measured on these scatterers' skies are then $T_A = T_A(\theta_A, \phi_A, z_A)$ and $T_B = T_B(\theta_A, \phi_A, z_A, \theta_B, \phi_B, z_B)$. Using this notation, the occupation number of CMB photons at the primary observer is

$$N_{\rm tot} = \mathcal{B}(\nu, \bar{T}_c) + \frac{3}{16\pi}\int \tau(Y_A)\left[\mathcal{B}(\nu,\bar{T}_A) - \mathcal{B}(\nu,\bar{T}_c)\right] dY_A + \frac{9}{(16\pi)^2}\int \tau(Y_A)\tau(Y_B)\left[\mathcal{B}(\nu,\bar{T}_B) - \mathcal{B}(\nu,\bar{T}_A)\right] dY_A dY_B,$$

where we have written $Y_A = \{x_A, \phi_A, z_A\}$, $dY_A = dx_A d\phi_A dz_A$, etc. The first term on the RHS of this equation is from the unscattered CMB, and the second term is from the CMB light scattered toward and away from us by photons that are scattered once. The third term on the RHS is new, and corresponds to photons that are scattered towards and away from us by double scatterings. The calculation now proceeds as in the single scattering case, and results in

$$\begin{aligned} 0 &= \left[1 - \frac{3}{16\pi}\int \tau(Y_A) dY_A\right]\left(\frac{1}{\bar{T}_c} - \frac{1}{T_0}\right)^2 \\ &+ \frac{9}{(16\pi)^2}\int \tau(Y_A)\tau(Y_B)\left(\frac{1}{\bar{T}_B} - \frac{1}{T_0}\right)^2 dY_A dY_B \\ &+ \frac{3}{16\pi}\int \tau(Y_A)\left[1 - \frac{3}{16\pi}\int \tau(Y_B) dY_B\right]\left(\frac{1}{\bar{T}_A} - \frac{1}{T_0}\right)^2 dY_A. \end{aligned}$$

For $\tau \neq 0$ and $\int \tau dY < 16\pi/3$ we therefore have that $T_A = T_B = T_c = (1+z)T_0$. The inclusion of third and higher order scatterings will proceed in an analogous way.

[1] Ehlers, J., Geren, P., & Sachs, R. K., *J. Math. Phys.* **9** 1344 (1968).
[2] Ferrando, J. J., Morales, J. A., & Portilla, M., *Phys. Rev. D* **46** 578 (1992).
[3] Clarkson, C. & Barrett, R., *Class. Quant. Grav.* **16** 3781 (1999).
[4] Clarkson, C. & Maartens, R., *Class. Quant. Grav.* **27** 124008 (2010).
[5] Maartens, R., Ellis, G. F. R., & Stoeger, W. R., *Phys. Rev. D* **51** 5942 (1995).
[6] Stoeger, W. R., Maartens, R., & Ellis, G. F. R., *Astrophys. J.* **443** 1 (1995).
[7] Maartens, R., Ellis, G. F. R., & Stoeger, W. R., *Phys. Rev. D* **51** 1525 (1995).
[8] Räsänen, S., *Phys. Rev. D* **79** 123522 (2009).
[9] Hasse, W. & Perlick, V., *Class. Quantum Grav.* **16** 2559 (1999).
[10] Maartens, R., *Phil. Trans. R. Soc. A* **369**, 5115 (2011).
[11] Maartens, R. & Matravers, D. R., *Class. Quant. Grav.* **11** 2693 (1994).
[12] Mustapha, N., Hellaby, C., & Ellis, G. F. R., *Mon. Not. Roy. Astron. Soc.* **292** 817 (1997).
[13] García-Bellido, J. & Haugbølle, T., *JCAP* **9** 16 (2008).
[14] Garfinkle, D., *Class. Quant. Grav.* **27** 065002 (2010).
[15] Zhang, P. & Stebbins, A., *Phys. Rev. Lett.* **107** 041301 (2011).
[16] Zibin, J. P. & Moss, A., *Class. Quant. Grav.* **28** 164005 (2011).
[17] Bull, P., Clifton, T., & Ferreira, P. G., *Phys. Rev. D* **85**, 024002 (2012).
[18] Yoo, C. M., Nakao, K-I., & Sasaki, M., *JCAP* **10** 11 (2010).
[19] Clarkson, C. & Regis, M., *JCAP* **2** 13 (2011).
[20] Goodman, J., *Phys. Rev. D* **52** 1821 (1995).
[21] Caldwell, R. R. & Stebbins, A., *Phys. Rev. Lett.* **100** 191302 (2008).
[22] Marra, V. & Notari, A., *Class. Quant. Grav.* **28** 164004 (2011).
[23] Sunyaev, R. A. & Zel'dovich, Ya. B., *Astrophys. Sp. Sci.* **7** 3 (1970).
[24] Sunyaev, R. A. & Zel'dovich, Ya. B., *Mon. Not. Roy. Astron. Soc.* **190** 413 (1980).
[25] Ellis, G. F. R., in *General Relativity and Cosmology*, Proc. Int. School of Physics "Enrico Fermi" (Varenna), p. 104 (R. K. Sachs, Ed.), Academic Press, San Diego (1971).

[26] Chluba, J. & Sunyaev, R. A., *A. & A.* **424**, 389 (2004).
[27] Maartens, R., Gebbie, T., & Ellis, G. F. R., *Phys. Rev. D* **59**, 083506 (1999).
[28] Krasiński, A., *Inhomogeneous Cosmological Models*, p. 333, Cambridge University Press, Cambridge (1997).
[29] Clarkson, C. A., Coley, A. A., & O'Neill, E. S. D., *Phys. Rev. D* **64**, 063510 (2001).
[30] Quercellini, C. *et al.*, arXiv: 1011.2646.
[31] Itoh, N. *et al.*, *Mon. Not. Roy. Astron. Soc.* **327**, 567 (2001).
[32] Dolgov, A. D. *et al.*, *Astrophys. J.* **554**, 74 (2001).
[33] If dark energy is a dynamical field then we actually have an extra functional degree of freedom to determine observationally, complicating the situation even more.
[34] The integrals in this expression are understood to be along the past nullcone, which may not be monotonic in z.