Defining the frame of minimum Hubble expansion variance

James H. McKay* & David L. Wiltshire†
Department of Physics & Astronomy, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

23 March 2015

ABSTRACT
We characterize a cosmic rest frame in which the variation of the spherically averaged Hubble expansion is most uniform, under local Lorentz boosts of the central observer. Using the COMPOSITE sample of 4534 galaxies, we identify a degenerate set of candidate minimum variance frames, which includes the rest frame of the Local Group (LG) of galaxies, but excludes the standard Cosmic Microwave Background (CMB) frame. Candidate rest frames defined by a boost from the LG frame close to the plane of the galaxy have a statistical likelihood similar to the LG frame. This may result from a lack of constraining data in the Zone of Avoidance. We extend our analysis to the Cosmicflows-2 (CF2) sample of 8,162 galaxies. While the signature of a systematic boost offset between the CMB and LG frames averages is still detected, the spherically averaged expansion variance in all rest frames is significantly larger in the CF2 sample than would be reasonably expected. We trace this to the CF2 distances being reported without a correction for inhomogeneous distribution Malmquist bias. Systematic differences in the inclusion of the large SFI++ subsample into the COMPOSITE and CF2 catalogues are analysed. Our results highlight the importance of a careful treatment of Malmquist biases for future peculiar velocities studies, including tests of the hypothesis of Wiltshire et al. (2013) that a significant fraction of the CMB temperature dipole may be nonkinematic in origin.

Key words: cosmology: observations — cosmology: theory — distance scale

1 INTRODUCTION
Although the Universe is spatially homogeneous in some statistical sense, at the present epoch it exhibits a complex hierarchical structure, with galaxy clusters forming knots, filaments and sheets that thread and surround voids, in a complex cosmic web (Forero–Romero et al. 2009; Bilicki et al. 2014; Einasto 2014). Deviations from homogeneity are conventionally treated in the framework of peculiar velocities, by which the mean redshift, z, and luminosity distance, r, of a galaxy cluster are converted to a peculiar velocity according to

\[ v_{\text{pec}} = cz - H_0 r \]  

where \( c \) is the speed of light and \( H_0 = 100 h \text{ km sec}^{-1} \text{ Mpc}^{-1} \) the Hubble constant.

The peculiar velocity framework makes a strong geometrical assumption over and above what is demanded by general relativity. In particular, the quantity \( v_{\text{pec}} \) defined by (1) only has the physical characteristics of a velocity if one implicitly assumes the spatial geometry on all scales larger than those of bound systems is exactly described by a homogeneous isotropic Friedmann-Lemaître-Robertson-Walker (FLRW) model with a single cosmic scale factor, \( a(t) \), whose derivative defines a single global Hubble constant, \( H_0 = \dot{a}/a \bigg|_{t_0} \). Deviations from the uniform expansion are then ascribed to local Lorentz boosts of each galaxy cluster with respect to the spatial hypersurfaces of average homogeneity.

It is a consequence of general relativity, however, that inhomogeneous matter distributions generally give rise to a differential expansion of space that cannot be reduced to a single uniform expansion plus local boosts. This is a feature of general exact solutions to the cosmological Einstein equations, such as the Lemaître–Tolman–Bondi (LTB) (Lemaître 1933; Tolman 1934; Bondi 1947) and Szekeres (1975) models. Any definition of the expansion rate in such models depends on the spatial scale relative to that of the inhomogeneities. Although one can define scale dependent Hubble parameters for specific exact solutions – for example, given the spherical symmetry of the LTB model – the actual cosmic web is sufficiently complex that in reality one must deal with spatial or null cone averages in general relativity.

In recent work Wiltshire et al. (2013) examined the variation of the Hubble expansion from a fresh perspec-
A linear Hubble law with a spherically averaged Hubble expansion applies on scales larger than 100 h⁻¹Mpc, then once the shells are a few times larger than the diameter of the largest typical nonlinear structures, a well-defined asymptotic average, $H_{n}$, is obtained which does not change when shells are further enlarged. When shells are $1 - 2$ times the diameter of the typical nonlinear voids, however, a variation in expansion rate is seen and since the faster expanding voids dominate by volume then the average, $H_{s}$, is increased relative to $H_{p}$.

While the results just discussed may not be surprising to those familiar with the statistics of the cosmic web, the remaining results of Wiltshire et al. (2013) listed above are not at all expected in the conventional approach to peculiar velocities in observational cosmology. In particular, the cosmic rest frame should be the one in which the variation of cosmic expansion is a minimum in some sense. The results of Wiltshire et al. (2013) show that such a frame of minimum Hubble expansion variance is not the conventional CMB frame. In particular, a substantial fraction of the CMB dipole usually attributed to a local boost of the Local Group at 635 ± 38 km sec⁻¹ (Tully et al. 2008) in the direction $(\ell, b) = (276.4, 29.3)$ would be due to a nonkinematic differential expansion of space. Wiltshire et al. (2013) estimate that a 0.5% anisotropy in the distance-redshift relation on $\lesssim 65 h^{-1}$Mpc scales would be required to achieve the observed properties of the CMB dipole. Furthermore, using ray-tracing simulations on exact solutions of Einstein’s equations for nonlinear inhomogeneities on such scales it is found that it is possible to produce realistic values of higher order CMB multipoles while accounting for the residual dipole in the LG frame (Bodeiko, Nazer & Wiltshire 2015). This may potentially provide an explanation for the angular scale dependence observed in attempts to measure the boost to the CMB frame using the effects of special relativistic aberration and modulation on the CMB anisotropy spectrum (Aghanim et al. 2014). Since subtraction of a kinematic dipole is a key step in the map making process, this may also have a direct impact on various large angle anomalies in the CMB anisotropy spectrum (Ade et al. 2014). Furthermore, given that a transformation to the CMB frame is a step that is taken in many observational procedures without question, the possibility that a substantial fraction of the CMB dipole may be nonkinematic would have a large impact in other areas of observational cosmology.

Given the potential importance of such a result, it is important to try to characterize the frame of minimum Hubble expansion variation in purely observational terms. Wiltshire et al. (2013) compared the LG and LS frames with that of the CMB, motivated by the starting point that a frame close to the LG frame would be the natural standard of rest according to the “Cosmological Equivalence Principle” (Wiltshire 2008) which underlies the approach of Wiltshire (2007, 2008) to the averaging problem in inhomogeneous cosmology (Ellis 1984, Buchert 2000, 2008, Wiltshire 2013, Nazer & Wiltshire 2015).

It is the aim of this paper to determine in a model-independent fashion whether a rest frame of minimum Hubble expansion variance can be found among all the frames boosted by arbitrary amounts with respect to the LG and LS frames.

### References

1. Larger structures such as the $320 h^{-1}$Mpc long Sloan Great Wall (Gott et al. 2005) and the $350 h^{-1}$Mpc long Large Quasar Group (Clowes et al. 2013) are known, but it is arguable that these are not typical.

2. See Wiltshire et al. (2013) Fig. 1 for a useful conceptual diagram.
CMB rest frames, and what the frame is. It will also be necessary to consider the different ways to characterize such a minimum variation frame. While our principal results are determined from an analysis of the COMPOSITE sample (Watkins et al. 2009, Feldman et al. 2010), we have also considered the recently published Cosmicflows-2 (CF2) sample of Tully et al. (2013). We find that issue of the different treatments of Malmquist bias in the two datasets at present prevents as detailed analysis as we perform for the COMPOSITE sample. However, our investigation highlights how the implicit assumption of a FLRW expansion law below the scale of statistical homogeneity via (1) subtly influences the manner in which such biases are treated in practice, and raises concerns about remaining distribution biases in the CF2 distances.

2 METHODOLOGY

In this paper we will use the methodology of Wiltshire et al. (2013), who considered both radial and angular averages of the Hubble expansion in the COMPOSITE sample, adapting techniques that Li & Schwarz (2008) and McClure & Dyer (2007) had previously applied to radial and angular averages respectively in the much smaller Hubble Telescope Key Project dataset (Freedman et al. 2001). We are primarily interested in the radial spherical averages, as it was by this method that Wiltshire et al. (2013) found the most decisive evidence that the Hubble expansion is significantly more uniform in the LG or LS frames, as compared to the standard CMB rest frame. Our first aim is to repeat this analysis for rest frames which are boosted arbitrarily with respect to a given frame, and to search for a frame in which the variation of the Hubble expansion is a minimum in some sense, using the COMPOSITE sample.

Following Wiltshire et al. (2013), we determine the best fit linear Hubble law by standard linear regression in independent radial shells, minimizing the quantity

\[
\chi^2_s = \sum_{i=1}^{N_s} \left[ \sigma_i^{-1} (r_i - cz_i/H) \right]^2 ,
\]

with respect to \(H\), where \(z_i, r_i\) and \(\sigma_i\) are respectively the redshift, distance and distance uncertainty of each object, and \(N_s\) is the number of data points in a given shell, \(s\). This method, with distance as function of redshift, is chosen because all uncertainties in the COMPOSITE sample have been included as distance uncertainties. Any corrections that would be required due to noise arising from peculiar motion within galaxy clusters have been accounted for by assigning a distance and uncertainty to the cluster itself rather than individual galaxies (Watkins et al. 2009). The values of \(z_i\) are taken to be exact, whereas the distance uncertainties are large, approximately 15% for most COMPOSITE sample objects. Fortunately, the large number of data points within each shell nonetheless lead to statistically significant results.

The value of the Hubble constant in the \(s\)th shell is then determined as

\[
H_s = \left( \sum_{i=1}^{N_s} \frac{c(z_i)^2}{\sigma_i^2} \right)^{-1} \left( \sum_{i=1}^{N_s} \frac{cz_i r_i}{\sigma_i^2} \right)^{-1} ,
\]

while its weighted average luminosity distance is

\[
\bar{r}_s = \left( \sum_{i=1}^{N_s} \frac{r_i}{\sigma_i^2} \right) \left( \sum_{i=1}^{N_s} \frac{1}{\sigma_i^2} \right)^{-1} .
\]

The radial averages are computed in two different shell configurations of 11 shells, for luminosity distances with \(r_s < r \leq r_{s+1}\). Both configurations use shells of width 12.5 \(h^{-1}\)Mpc in most cases, but start from a different innermost shell cutoff of either 2 \(h^{-1}\)Mpc or 6.25 \(h^{-1}\)Mpc. The two shell configurations are labeled using unprimed and primed integers respectively, as given in Table 1 of Wiltshire et al. (2013). Given less data at large distances, the shells 10 and 10′ are taken to be wider than the rest so as to include approximately the same number of data points as the inner shells. They both have an outer cutoff at 156.25 \(h^{-1}\)Mpc. The outermost shell 11 has an outer bound given by the largest distance in the sample, and is the same in both configurations, to provide an anchor for the asymptotic value of the Hubble constant, \(\bar{H}_0\).

The uncertainty in \(\bar{H}_0\) is taken as

\[
\sigma_s = \sqrt{\sigma_{0s}^2 + \sigma_{1s}^2} ,
\]

where

\[
\sigma_{0s} = H_s \sigma_0 / \bar{r}_s
\]

and

\[
\sigma_{1s} = \left( \sum_{i=1}^{N_s} \frac{(cz_i)^2}{\sigma_i^2} \right)^{3/2} \left( \sum_{i=1}^{N_s} \frac{cz_i r_i}{\sigma_i^2} \right)^{-1} ,
\]

A zero point uncertainty \(\sigma_0\) is added in quadrature to the standard uncertainty \(\sigma_s\) determined through error propagation in (3), to yield the total uncertainty \(\sigma_s\) for the Hubble constant in each shell. The zero point uncertainty arises from the fact that a linear Hubble law must necessarily pass through the origin. However, there is an uncertainty in determining the origin, due to a 20 km sec\(^{-1}\) uncertainty in the heliocentric peculiar velocity of the LG (and LS) frames (Tully et al. 2008) and a 0.4% uncertainty in the magnitude of the CMB dipole (Fixsen et al. 1996), which combine to give \(\sigma_0 = 0.201 h^{-1}\)Mpc (Wiltshire et al. 2013). For each shell, the weighted zero point uncertainty \(\sigma_{0s}\) is the uncertainty in the linear Hubble law for a shell with mean distance, \(\bar{r}_s\), due to the uncertainty in the origin alone. This uncertainty is more significant for shells with a small mean distance compared to those at a larger distance, due to the shorter lever arm of the linear Hubble law for shells with smaller mean distances.

As a measure of the difference in the local Hubble expansion from its asymptotic value, we use the quantity

\[
\delta H_s = \left( H_s - H_0 \right) / \bar{H}_0
\]

where \(\bar{H}_0\) is the mean asymptotic value of the Hubble constant. In our case, \(\bar{H}_0\), and its uncertainty are calculated from the data points at distances \(r > 156.25 h^{-1}\)Mpc (shell 11 in both unprimed and primed configurations). Wiltshire et al. (2013) choose the distance scale for the outermost shell as one which is statistically reliable for the COMPOSITE data, while including only data at distances

© 2015 RAS, MNRAS.
larger than the baryon acoustic oscillation scale, the largest scale on which one might expect to see the effects of inhomogeneity on the local Hubble expansion. They verify that a linear Hubble law with a very high goodness of fit is found for \( r > 156.25 \, h^{-1} \) Mpc, with a Hubble constant \( \dot{H}_0 = (100.1 \pm 1.7) \) km sec\(^{-1}\) Mpc\(^{-1}\) in the CMB frame or \( \dot{H}_0 = (101.0 \pm 1.7) \) km sec\(^{-1}\) Mpc\(^{-1}\) in the LG/LS frames, consistent with the normalization \( H_0 = 100 \) km sec\(^{-1}\) Mpc\(^{-1}\) used in the COMPOSITE sample.

### 3 MONOPOLE EXPANSION VARIATION DUE TO SYSTEMATIC BOOST OFFSETS

Wiltshire et al. (2013) propose that the much larger monopole variation of the linear Hubble law observed in the CMB frame, as compared to the LG/LS frames, may have a systematic origin. This arises from the nonlinear dependence of \( H_s \) in (9) on the individual \( c_z \) values, which change when performing a local Lorentz boost to a different frame. By calculating the result of an arbitrary boost on the individual \( H_s \) values Wiltshire et al. (2013) obtain an explicit form for this systematic variation.

Consider redshifts observed in a frame of reference in which the variation of the spherically averaged Hubble expansion is minimized. Under an arbitrary local boost the observed redshifts transform as

\[
\Delta c_z \rightarrow \Delta c'_z = c_z + v \cos \phi_i
\]

where \( v \) is the boost magnitude and \( \phi_i \) is the angle between each data point and the boost direction. In (9) this results in the changes \((c_z)^2 \rightarrow (c'_z)^2 = (c_z)^2 + 2c_zv \cos \phi_i + v^2 \cos^2 \phi_i\) in the numerator and \( c_zr_i \rightarrow c_zr_i + r_i v \cos \phi_i \) in the denominator.

The linear contributions to the transformed quantities in the denominator and numerator of (9) should be approximately self-cancelling for data which is uniformly distributed on the celestial sphere in each shell. When taking a spherically symmetric average on such data then on average each positive contribution from the term \( v \cos \phi_i \) will cancel with a negative contribution from a data point on the opposite side of the sky. The lack of data in the Zone of Avoidance does not pose a problem as this absence of data is symmetrical on opposite sides of the sky. A rough self-cancellation of the linear contributions would only be invalid when one side of the sky has a significant lack of data as compared to the opposite side of the sky. The COMPOSITE sample does indeed have sufficient sky coverage to satisfy this requirement, with the exception of the first unprimed shell 1, with \( 2 < r \leq 12.5 \, h^{-1} \) Mpc (Wiltshire et al. 2013), which is excluded in the data analysis.

With such a cancellation assumed we are left with the difference

\[
\Delta H_s = H'_s - H_s \sim \left( \sum_{i=1}^{N_s} \frac{(v \cos \phi_i)^2}{\sigma_i^2} \right)^{-1} \left( \sum_{i=1}^{N_s} \frac{c_zr_i}{\sigma_i^2} \right)^{-1}
\]

\[
\approx \frac{v^2}{2H_0 \langle r^2 \rangle_s}
\]

(10)

where

\[
\langle r^2 \rangle_s \equiv \frac{1}{\sum_{i=1}^{N_s} \sigma_i^2} \left( \sum_{i=1}^{N_s} \frac{1}{\sigma_i^2} \right)^{-1}
\]

(11)

is a weighted average of the squared luminosity distance in each shell. The second line of (10) follows by assuming that:

(i) \( \langle v \cos \phi_i \rangle^2 = v^2 \langle \phi_i^2 \rangle \sim v^2 \) is roughly constant from shell to shell; (ii) the leading order linear Hubble approximation \( c_z \sim \dot{H}_0 r_i \) is made in the denominator. The uncertainty in (10) is given by

\[
\delta \langle r^2 \rangle_s = 2 \left( \sum_{i=1}^{N_s} \frac{r_i^2}{\sigma_i^2} \right)^{1/2} \left( \sum_{i=1}^{N_s} \frac{1}{\sigma_i^2} \right)^{-1}.
\]

(12)

Our first goal is to verify that the difference in the spherically averaged Hubble expansion between the LG and CMB frames of reference is statistically consistent with a systematic variation of the form (10). To achieve this we fit a power law to the observed data. As there is a correlated uncertainty in both the independent and dependent variables, \( \langle r^2 \rangle_s \) and \( \delta H_s \), respectively, a standard least squares method is not appropriate. Instead we use a total least squares fit, or “error in variables” method with a model of the form

\[
\delta H = A \langle r^2 \rangle_s^p.
\]

(13)

In comparison with (10), we expect \( p \approx -1 \) and \( A \approx v^2/(2\dot{H}_0) \). The details of this method are presented in the Appendix.

Carrying out this analysis, we do indeed find a difference consistent with a systematic boost offset between the LG and CMB frames of reference. Systematic uncertainties arise in the choice of shell boundaries. Considering only primed shells gives a value of \( p = -1.01 \pm 0.27 \). If we take the unprimed shells then we obtain \( p = -0.79 \pm 0.16 \) if shell 1, with \( 2 < r \leq 12.5 \, h^{-1} \) Mpc, is included and \( p = -0.61 \pm 0.31 \) if this first shell – which may have insufficient sky coverage – is excluded. The data in the range \( 6.25 < r \leq 12.5 \, h^{-1} \) Mpc common to both the first primed and unprimed shells is important in establishing the boost offset which is more pronounced at small \( r \). To account for systematic uncertainties, we have therefore applied a continuous variation of the first shell boundary in the range \( 2 \rightarrow 3.65 \, h^{-1} \) Mpc, while keeping the widths of the shells fixed. This leads to a value of \( p = -0.88 \pm (0.25)_{\text{stat}} \pm (0.13)_{\text{sys}} \), where the first uncertainty is the statistical and the second systematic. For the case of the primed shells we also note the corresponding velocity calculated from the best fit value of \( a \) is \( v = 646 \pm 545 \) km sec\(^{-1}\), which is indeed close to the actual boost magnitude of \( 635 \pm 38 \) km sec\(^{-1}\), albeit with a very large uncertainty.

We repeated the analysis using 8 shells rather than 11 to smooth out variations that could interfere with the systematic boost offset. The second configuration uses shells of width \( 18.75 \, h^{-1} \) Mpc, starting from an inner cutoff of \( 2 \, h^{-1} \) Mpc and \( 9.375 \, h^{-1} \) Mpc for unprimed and primed shells respectively. We find \( p = -0.89 \pm 0.34 \) for the primed shells and \( p = -0.96 \pm 0.26 \) for the unprimed shells. With a continuous variation of the inner shell boundary from \( 2 \rightarrow 9.375 \, h^{-1} \) Mpc, we arrive at a value of \( p = -0.87 \pm (0.33)_{\text{stat}} \pm (0.09)_{\text{sys}} \). Fig. (1b) shows the resultant best fit curves.

In Fig. (1b) we note a discrepancy between the best fit power law and the negative values on shells in the range of \( 40 \rightarrow 60 \, h^{-1} \) Mpc (or \( \langle r^2 \rangle_s = 1600\rightarrow 3300 \, (h^{-1} \text{Mpc})^2 \)). This is understood to be the result of structures in this particular range, which give rise to both a residual monopole and dipole variation of the Hubble expansion in the LG frame.
Wiltshire et al. (2013) show, there is evidence that the boost to the CMB frame somewhat compensates for structures in this range. One finds that in the range $40 < r < 60 \text{ h}^{-1}\text{Mpc}$ (and only in this range) the monopole variation is less in the CMB frame, $(\delta H_s)_{\text{CMB}} < (\delta H_s)_{\text{LG}}$, while the dipole magnitude is also less in the CMB frame, becoming consistent with zero in the middle of the range. If the boost to the CMB frame exactly compensated for structures in the range $40 < r < 60 \text{ h}^{-1}\text{Mpc}$ then the dipole magnitude should remain close to zero in shells at larger distances. However, the magnitude of the CMB frame dipole increases to a residual offset at large distances, while the LG frame dipole is consistent with zero in most outer shells.

Thus it appears that the dipole almost – but not entirely – has the character of a Lorentz boost dipole. Given that there are structures that give rise to a residual nonlinear Hubble expansion in the range $40 < r < 60 \text{ h}^{-1}\text{Mpc}$ we cannot expect a perfect power law fit. However, the deviation from a power law is consistent with the observation that $(\langle H_s \rangle)_{\text{CMB}} < (\langle H_s \rangle)_{\text{LG}}$ in the range over which the boost almost compensates for nonlinear structures.

Now that we have verified the power law nature of the difference $H_{\text{CMB}} - H_{\text{LG}}$ we must check whether this result is unique for the boost to the LG frame. To investigate this we determine the Hubble constant in radial shells for frames boosted arbitrarily with respect to the CMB, denoted by frame “X”, and then fit to the resulting $\Delta H = H_{\text{CMB}} - H_{X}$ curve. We vary the direction of the boost to frame X while holding the magnitude constant, thus producing a sky map. We first choose a magnitude of 635 km sec$^{-1}$ corresponding to the boost from the CMB to LG frame of reference.

To display these sky maps in a meaningful fashion we cannot simply plot the value of $p$. Suppose that the CMB is boosted from a frame which has $\delta H_s = A (r) p$ with $p = -1$ and $A > 0$, representing the best fit boost offset. If one now boosts in the opposite direction by 635 km sec$^{-1}$ then one finds a best fit power law with $\delta H_s = A (r) p$ with $p \approx -1$ but $A < 0$ since the CMB frame necessarily has the smaller value of $H_s$ on average. In each case we must first of all determine whether $f_p$ gives a better overall fit with $A > 0$ or $A < 0$ – given that some data points will always be opposite to the overall trend. In Fig. we plot

$$f_p = \begin{cases} p + 1, & A > 0 \\ 2 - |p + 1|, & A < 0 \end{cases}$$

which takes the value $f_p = 0$ at the best fit with $A > 0$ and $f_p = 2$ at the best fit with $A < 0$. The latter point turns out to be in the opposite direction, but not exactly opposite the best fit direction, reflecting the uncertainties in the method.

Fig. shows that for a boost magnitude of 635 km sec$^{-1}$ the Local Group is indeed contained in a set of frames that display strong evidence of being a minimum Hubble expansion variance frame as $p \approx -1$. More precisely, the monopole Hubble expansion in the CMB frame compared to frames boosted at 635 km sec$^{-1}$ relative to the CMB has the mathematical character of a systematic boost offset for directions close to that of the LG. In Fig. we can see a distinctive difference between the directions for which the boosted frame has the lesser variation ($A > 0$, values plotted closest to 0), and the directions for which the CMB frame has the lesser variation ($A < 0$, values plotted closest to 2).

We have verified that the additional monopole variation of the Hubble expansion seen in the CMB frame does have the character of what is expected by a boost from the LG frame, if the LG frame is close to the frame in which the monopole variation is minimized. However, as yet we have not varied the magnitude of the boost. Given the uncertainties we have already noted, it is clear that any systematic boost offset will become very hard to confirm statistically if the boost magnitude is small, since the other uncertainties will then become dominant. The systematic boost offset method can only give a rough indication of the boost direction for the large boosts rather than defining a precise

\footnote{Both primed and unprimed shells are used (in the 11 shell case), to produce a smoothed sky map without determining systematic uncertainties.}
“minimum Hubble expansion variance frame”. There are in fact many degenerate frames.

4 MINIMIZING THE AVERAGE SPHERICAL HUBBLE EXPANSION VARIANCE

We will now investigate arbitrary variations of the boost magnitude and direction to determine to what extent we can define a frame of reference in which the monopole variation in the Hubble expansion is minimized.

4.1 Variations in the “nonlinear regime”

Initially we will consider monopole variations with respect to a uniform $\delta H = 0$ expectation below the scale of statistical homogeneity ($\lesssim 100 \, h^{-1} \text{Mpc}$). This is quantified by summing the mean square differences of (5), to give a statistic

$$
\chi^2(n_f, n_i) = \sum_{i=n_i}^{n_f} \frac{\bar{H}_0^2 \delta H^2}{\bar{H}_0^2 \sigma^2_h + H_0^2 \sigma^2_{h0}},
$$

(15)

where $n_f$ and $n_i$ define the upper and lower shells included in the range of the calculation respectively. We will take the primed shells with, $n_i = i$' and $n_f = 8'$, covering the range $6.25 < r < 106.25 \, h^{-1} \text{Mpc}$. This includes all data potentially in the regime of nonlinear Hubble expansion, while excluding the innermost unprimed shell which may have incomplete sky coverage.

The frame of reference with the minimum monopole variation is found using a downhill optimization with (15). This reveals a global minimum for a boost in the direction $(\ell, b) = (59.3^\circ, 16.6^\circ)$ with magnitude $740.6^{+522}_{-728.8} \, \text{km sec}^{-1}$ with respect to the LG frame. Even though the boost is large, the uncertainties are also large so that the boost is only just over $1\sigma$ from zero. However, the corresponding boost from the CMB reference frame to the global minimum frame has a magnitude $1203^{+522}_{-625} \, \text{km sec}^{-1}$, which is even larger and now significantly nonzero. To make sense of these results we will consider in more detail the confidence intervals associated with this minimum.

To visualize the distribution of $\chi^2_{\ell}$ in the 3-dimensional parameter space $\{v, \ell, b\}$ we show two angular slices at fixed values of $v$ in Fig. 3 and a slice along the locus of $(\ell, b)$ values for which $\chi^2_{\ell}$ is minimized for fixed $v$ in Fig. 4. The angular distribution in Fig. 3 remains similar for all nonzero velocities. However, the confidence intervals grow large as $v$ is decreased, eventually taking up the whole sky for very small velocities. Only for large boosts is there a well-defined boost direction that reduces monopole variation.

In Fig. 4 for each boost magnitude we locate the direction of minimum $\chi^2_{\ell}$ and plot the corresponding value. The angular coordinates $(\ell, b)$ of the minimum are different in each case. The distribution of $\chi^2_{\ell}$ relative to the 68.3% confidence interval (dashed horizontal line on Fig. 4a) reveals the main problem in constraining a “minimum variation frame” by this technique. The near flat distribution of $\chi^2_{\ell}$ values within $1\sigma$ of the global minimum are found to lie on a locus of boost directions $(\ell, b)$ which all lie close to the galactic plane. This is of course the Zone of Avoidance region, where the COMPOSITE sample lacks data, due to the Milky Way obscuring distant galaxies. Evidently, we are free to perform large boosts in the plane of the galaxy, as the data is not constrained there. This hypothesis could be checked by simulating data with the same characteristics as the COMPOSITE sample, using exact solutions of Einstein’s equations (Bolejko et al. 2015). Such an investigation is beyond the scope of this paper. For now, with the available data we can only conclude that the LG frame is not ruled out as the standard of rest by this criterion, being only just over one standard deviation from the minimum.

We have found a set of degenerate frames of reference which might be taken as the minimum average monopole

Figure 2. The best fit parameters to a systematic boost offset (14) for frames boosted from the CMB frame at 635 km sec$^{-1}$. The black cross denotes the boost to the LG frame and the white cross denotes the boost to the frame with minimum variation in the spherically averaged Hubble law for a boost of this magnitude, which will be discussed in (11). In all figures, the galactic longitudes $\ell = 0^\circ, 180^\circ, 360^\circ$ are on the right edge, centre and left edge respectively.
Defining the frame of minimum Hubble expansion variance

3. Variations in the “linear regime”

The goodness of fit, $Q_s$, of a linear Hubble law in the innermost shells of Table 1 is poor, as we would expect since these shells are in the nonlinear regime. Beyond approximately $75 \, h^{-1}\text{Mpc}$ we expect to pass into the linear regime (Scrimgeour et al. 2012), and this is seen with the decreasing values of $\chi^2_s/\nu$, where $\chi^2_s$ is given by (2) for each shell.

However, contrary to expectation, the MV frame deter-

4.2 Variations in the “linear regime”

The goodness of fit, $Q_s$, of a linear Hubble law in the innermost shells of Table 1 is poor, as we would expect since these shells are in the nonlinear regime. Beyond approximately $75 \, h^{-1}\text{Mpc}$ we expect to pass into the linear regime (Scrimgeour et al. 2012), and this is seen with the decreasing values of $\chi^2_s/\nu$, where $\chi^2_s$ is given by (2) for each shell.

However, contrary to expectation, the MV frame deter-

Figure 3. Contour maps of angular variation of $\chi^2_a$ for two choices of boost magnitude with respect to the Local Group: (a) 740 km sec$^{-1}$, (b) 630 km sec$^{-1}$. The solid contours show the 68.3% and 90% confidence intervals for the direction of the boost to the minimum $\chi^2_a$ frame of reference in each case. The white cross on (b) shows the direction of the boost to the CMB (also of magnitude $\approx$ 630 km sec$^{-1}$).

In all figures, the galactic longitudes $\ell = 0^\circ$, $180^\circ$, $360^\circ$ are on the right edge, centre and left edge respectively.

Figure 4. Variation of the minimum $\chi^2_a$ with fixed boost velocity, $v$. The locus of $\ell$ and $b$ values for which this minimum is found lies within $\approx 20^\circ$ of the galactic plane for cases in which $\chi^2_a$ is within 1 $\sigma$ of the global minimum, as indicated by the dashed line.

variance frame. Although large uncertainties still exist, we are able to see that the boost to the CMB frame, shown by a white cross in Fig. 3(a), is far from our degenerate set of possible boosts to the minimum variance frame. Table 1 gives a comparison between the LG frame and the absolute minimum $\chi^2_a$ frame, denoted the minimum variance (MV) frame, similar to Table 1 of Wiltshire et al. (2013). In particular, using the complementary incomplete gamma function for the $\chi^2_a$ distribution we directly calculate the probabilities $P_{\text{LG}}$, $P_{\text{MV}}$ of the spherically averaged Hubble law giving a value that coincides with $\bar{H}_0$, when all shells larger than a given shell are included. If we exclude the innermost unprimed shell 1, then the Bayesian evidence for the MV frame having a more uniform average Hubble expansion than the LG frame, lies at most in the range $1 < \ln B < 3$, which represents positive but not strong Bayesian evidence (Kass & Raftery 1995). By contrast, the Bayesian evidence that the LG frame has a more uniform spherically average Hubble expansion than the CMB frame is very strong with $\ln B > 5$ (Wiltshire et al. 2013).
Table 1. Hubble expansion variation in radial shells in minimum Hubble expansion variance (MV) and LG frames. Spherical averages are computed for two different choices of shells, $r_s < r \leq r_{s+1}$, the second choice being labeled by primes. In each case we tabulate the inner shell radius, $r_s$; the weighted mean distance, $\bar{r}_s$; the shell Hubble constants, $(H_s)_{LG}$ and $(H_s)_MV$, in the LG and MV frames, and their uncertainties determined by linear regression within each shell, together with its “goodness of fit” probability $Q_s$, and reduced $\chi^2$ (for $v = N_s - 1$); $\ln B$ where $B$ is the Bayes factor for the relative probability that the MV frame has more uniform $\delta H_s = 0$ than the LG frame when $\chi^2$ is summed in all shells with $r > r_s$. $H_s$ and $\sigma_s$ are given in units $h$ km sec$^{-1}$ Mpc$^{-1}$.

| Shell $s$ | 1′ | 2′ | 3′ | 4′ | 5′ | 6′ | 7′ | 8′ | 9′ | 10′ | 11′ |
|-----------|----|----|----|----|----|----|----|----|----|-----|-----|
| $N_s$     | 321| 513 |553 |893 |681 |485 |343 |273 |164 |206 | 91 |
| $r_s$ ($h^{-1}$Mpc) | 6.25| 18.75 |31.25 |43.75 |56.25 |68.75 |81.25 |93.75 |106.25 |118.75 |156.25 |
| $\bar{r}_s$ ($h^{-1}$Mpc) | 12.26 | 23.46 |37.61 |49.11 |61.74 |73.92 |87.15 |99.12 |111.95 |131.49 |182.59 |
| $(H_s)_{LG}$ | 103.5 | 103.5 |103.9 |106.6 |103.9 |102.0 |103.2 |103.6 |101.6 |102.7 |101.0 |
| $(\sigma_s)_{LG}$ | 1.8 | 1.1 | 0.9 | 0.7 | 0.8 | 0.8 | 0.9 | 0.9 | 1.0 | 0.9 | 1.7 |
| $(Q_s)_{LG}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.998 | 0.940 | 1.000 | 1.000 | 1.000 | 0.993 | 0.999 |
| $(\chi^2/\nu)_{LG}$ | 13.56 | 7.767 |2.185 |1.419 |0.864 |0.909 |0.594 |0.542 |0.622 |0.803 |0.590 |
| $(H_s)_MV$ | 181.5 | 102.9 |106.7 |104.5 |104.8 |102.9 |102.6 |103.9 |104.9 |102.7 |102.0 |
| $(\sigma_s)_MV$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.330 | 0.887 | 1.000 | 1.000 | 1.000 | 0.964 | 0.999 |
| $(Q_s)_MV$ | 29.130 | 13.230 |3.037 |2.005 |1.021 |0.928 |0.682 |0.600 |0.667 |0.854 |0.603 |
| $(\chi^2/\nu)_MV$ | 3.53 | 2.85 |2.79 |1.99 |0.85 |0.33 |0.36 |0.14 |0.07 |0.45 | |
| $\ln B$ ($r \geq r_s$) | 1.85 | 1.15 | 0.920 | 0.682 | 0.600 | 0.667 | 0.854 | 0.603 | 1.00 | 0.999 |

Any true candidate for a minimum Hubble expansion variance frame should also clearly demonstrate the emergence of a linear Hubble law consistent with the existence of a statistical homogeneity scale. The $\chi^2_s$ statistic involves minimizing the variation $\delta H_s = (H_s - \bar{H}_0)/\bar{H}_0$ relative to the asymptotic Hubble constant. But a boost can also alter $\bar{H}_0$ in a way which makes for a worse goodness of fit to a linear Hubble law. Therefore it is not a completely suitable candidate statistic.

In order to quantify the emergence of a linear Hubble law we will therefore alternatively minimize the quantity

$$\chi^2_h = \frac{\sum_{s=7}^{11} \nu_s (\bar{H}_s - H_0)^2}{\sum_{s=7}^{11} \nu_s}$$

(16)

where $\chi^2_s$ is given by (2) in the $s$th shell and $\nu_s$ denotes the degrees of freedom in each shell. This sum is performed over the unprimed configuration of shells as shell 7 has an inner cutoff near the boundary of the nonlinear and linear regimes. Thus (16) gives a measure of the goodness of fit to a linear Hubble law averaged over the outer 5 shells, without normalizing the asymptotic Hubble constant. We find $\chi^2_h = 0.631$ for the LG frame, $\chi^2_h = 0.692$ for the MV (minimum $\chi^2_h$) frame, and $\chi^2_h = 0.653$ for the CMB frame. Thus even the CMB frame show a clearer emerging linear Hubble law than the minimum $\chi^2_h$ frame.

To determine whether there is any frame with a more
Defining the frame of minimum Hubble expansion variance

In Fig. 5 we locate the direction of minimum $\chi^2_b$ at each boost magnitude and plot the corresponding value, analogously to Fig. 4. We find a best fit boost of $222.3 \text{ km sec}^{-1}$ in the direction $(\ell, b) = (241.84^\circ, 70.53^\circ)$ with respect to the LG frame, with a value of $\chi^2_b = 0.621$. This is $45^\circ$ from the direction of the residual CMB temperature dipole in the LG frame, and so does not appear related.

In Fig. 6 for each boost magnitude we calculate the value of $\chi^2_a$ in an $(\ell, b)$ direction determined by minimizing with respect to $\chi^2_b$, and vice versa. Thus, we compute the locus of $(\ell, b)$ values in the $\{v, \ell, b\}$ parameter space that minimize $\chi^2_b$ for each fixed $v$, and then compute $\chi^2_a$ at these parameter values, and vice versa. It is apparent that making boosts of the order of $100 \text{ km sec}^{-1}$ along the locus of $(\ell, b)$ values which minimize $\chi^2_b$ results in an increase in $\chi^2_a$ beyond its 68.3% confidence interval (horizontal dashed line). This is because making improvements in $\chi^2_b$ requires boosts in $(\ell, b)$ directions away from our minimum $\chi^2_a$.

However, we note that all values of $\chi^2_b$ shown in Fig. 5 and 6 are consistent with very probable fits, given the low values of $\chi^2_b$. Furthermore, local boosts have a relatively small effect on the values of $cz_i$ for shells in the linear regime. These facts mean we should also exercise caution about drawing strong conclusion solely from the minimization of $\chi^2_b$. While the statistic is not potentially biased by the outermost shell as the $\chi^2_a$ statistic is, there are less data points per se in those outer shells used in (16), as compared to the inner shells which are most important in the $\chi^2_a$ statistic. Much more data is needed to use the $\chi^2_b$ statistic in a reliable way.

Finding a joint minimum for $\chi^2_a$ and $\chi^2_b$ would be feasible if the $(\ell, b)$ values of each global minimum were close on the sky. However, this is not the case – the global minimum for $\chi^2_a$ is at an angle of $93^\circ$ from the global minimum of $\chi^2_b$, i.e., they are roughly orthogonal. As there is no unbiased way to weight these two statistics we cannot set out to determine a weighted minimum as any result would be highly sensitive to our choice of weighting.
(a) \( f_p = \lfloor p + 1 \rfloor \) when \( A \geq 0 \) and \( 2 - |p + 1| \) when \( A < 0 \)

(b) \( f_p = \lfloor p + 1 \rfloor \) when \( A \geq 0 \) and \( 2 - |p + 1| \) when \( A < 0 \)

Figure 7. The best fit parameters to a systematic boost offset for boosts from the Local Group of magnitude: (a) \( 740 \) km sec\(^{-1} \); (b) \( 200 \) km sec\(^{-1} \). The thick blue contours denote the corresponding \( \chi^2 \) distribution. In (a) both the 68.3\% and 90\% confidence intervals are displayed, being \( \chi^2_1 = 2.26 \) and \( \chi^2_2 = 3.97 \) respectively, while in (b) only the 68.3\% confidence interval is visible (\( \chi^2_2 = 2.90 \)). In all figures, the galactic longitudes \( \ell = 0^\circ, 180^\circ, 360^\circ \) are on the right edge, centre and left edge respectively.

4.3 Systematic boost offsets from the Local Group

Neither statistic \( \chi^2_a \) or \( \chi^2_b \) appears entirely satisfactory for establishing a global minimum expansion variance frame. The \( \chi^2_a \) statistic is the better measure of Hubble expansion variation in the nonlinear regime but is also affected by potential bias in the anchoring of \( H_i \). The most we can say is that there is a freedom to perform large boosts in the plane of the galaxy, given the lack of data in the Zone of Avoidance.

If \( \chi^2_a \) is taken as the better statistic, then a criterion for breaking the boost degeneracy may be possible by returning to the systematic boost offset analysis of McKay et al. Any true best fit frame should show a clear signal of a boost offset \( \ell, b \) with respect to the Local Group. The “best” boost offset can be characterized in 3 ways, each with its own challenges.

(1) Determine the boost for which \( p = -1 \). This is hindered by the fact that there are many boosts that satisfy this criterion, at almost every magnitude from the LG.

(2) From the value of \( A \) in McKay et al. determine a derived boost velocity, \( v_{\text{der}} \). Any boost offset should have \( v_{\text{der}} \) consistent with the true boost magnitude \( v_{\text{true}} \) within uncertainties. However, this is difficult due to the large uncertainties associated with the value of \( A \).

(3) Determine a measure of variance in the fit of the boost offset, given by (AS) in the Appendix. This is also problematic since all fits are extremely good due to the large uncertainties in the \( H_i \).

We will therefore use method (1) to determine the direction of the boost on the sky, and then consider (2) and (3) to constrain the magnitude\(^5\).

First we check for a systematic boost offset for the global \( \chi^2 \) minimum frame determined in McKay et al. A sky map of \( f_p \) values as given by McKay et al. for boosts of magnitude \( 740 \) km sec\(^{-1} \) is given in Fig. 7(a), with the 1\( \sigma \) and 2\( \sigma \) confidence intervals for \( \chi^2_a \) displayed. We note that there are in fact boosts with values of \( p \approx -1 \) consistent with the \( \chi^2_a \) minimum. However, these directions are far more constrained and do not align with the exact minimum. In addition, the \( \ell, b \) direction at this magnitude with the best fit to McKay et al. has an inconsistent value of the derived velocity. Thus we do not see a clear systematic boost offset between the Local Group and the frame corresponding to the global minimum \( \chi^2_a \), further ruling this out as a potential candidate for the standard of rest we are looking for.

The next step is a global search for the best systematic boost offset from the LG. In order to further understand the angular distribution of \( f_p \) values for boosts from the LG frame, we arbitrarily choose a boost magnitude of \( 200 \) km sec\(^{-1} \) and plot \( f_p \) with respect to \( \ell, b \) in Fig. 7(b). We have found that for all interpolating velocities between the \( 200 \) km sec\(^{-1} \) and \( 740 \) km sec\(^{-1} \) cases displayed in Figure 7 there is a region of \( \ell, b \) values for which \( f_p \approx 0 \). Thus in order to use this method to find a realistic systematic boost offset we must use an additional criterion.

In Fig. 8 we plot the values of \( v_{\text{der}} / v_{\text{true}} \) and \( S/(n - 2) \) (for \( n \) data points) from (AS) with respect to the boost magnitude, where for each magnitude the \( \ell, b \) direction is that for which \( p \) is closest to \( -1 \) and \( A > 0 \) (i.e. \( f_p \approx 0 \)). Thus, we can use these additional quantities to constrain a systematic boost offset along the locus of \( \ell, b \) directions in the 3-dimensional \( \{ v, \ell, b \} \) parameter space. The expected value of \( S \) has a \( \chi^2 \) distribution for \( n - 2 \) degrees of freedom, and thus \( S/(n - 2) \) has an expectation value of unity (York et al. 2015).
Clearly, the values of $S/(n - 2)$ in Fig. 8(b) are consistent with a very good fit to (13) for all boosts. Our inability to tightly constrain the boost magnitude is no doubt due to the lack of data in the Zone of Avoidance and large uncertainties in the values of $H_{LG} - H_X$. Although (13) it is not useful for constraining the boost magnitude, we nonetheless see that the ratio of derived and true velocities in Fig. 8 does show a meaningful difference on this interval.

Using both primed and unprimed shells in our calculation we find that the $v_{\text{der}}/v_{\text{true}} = 1$ at $v_{\text{true}} = 122.5 \text{ km sec}^{-1}$ in a direction $(\ell, b) = (60^\circ, -4^\circ)$. For this boost, we find the result $p = -1.13 \pm (1.16)_{\text{stat}} \pm (0.29)_{\text{sys}}$, where the systematic uncertainty is determined as in (13) consistent with $p = -1$. The difference between the spherically averaged Hubble expansion in this frame, which we denote by $X$ and the LG is shown in Fig. 9. One can see that a systematic boost offset is apparent here. However, due to the large uncertainties that arise when taking differences of the $H_s$ the result is also statistically consistent with zero. (Thus the question of whether the first unprimed shell 1 should be included in the analysis, due to its incomplete sky cover, is immaterial.) This frame $X$ also lies within $1\sigma$ of the global minima of $\chi^2_\text{a}$ and $\chi^2_\text{b}$. Our choice of frame $X$ above is based on taking $v_{\text{der}}/v_{\text{true}} = 1$, a condition which may only be approximately matched in reality, given our huge uncertainties in the coefficient $A$. We again have a degeneracy in the choice of minimum Hubble expansion variance frame that satisfies the two conditions $p = -1$ and $v_{\text{der}} = v_{\text{true}}$.

The results of this section confirm the finding of (13) that our determination of a suitable cosmic rest frame is limited in the COMPOSITE sample by a degeneracy under boosts close to the plane of the galaxy. The consistency between the methods of this section and (13) may be less significant, as they are not completely independent. In the LG frame the primary source of the monopole variation is the increased value of $H_s$ in the innermost shells, while the more distant shells closer to the linear regime show closer to asymptotic values. Thus boosting to a frame with a reduced $H_s$ in the innermost shells will give the most significant improvement to $\chi^2_a$, relative to which small changes in the more distant shells are negligible. This is precisely the type of difference we model with a power law of the form (13) with $p \approx -1$. Consequently, if our hypothesis concerning (13) is correct then is not surprising that we see the consistency in the angular directions that minimize $\chi^2_a$ on one hand, and which give values of $p \approx -1$ with $A > 0$ on the other.

### 4.4 Angular Hubble expansion variation

Wiltshire et al. (2013) also explore the extent to which angular averages of the Hubble expansion offer an independent characterization of a minimum Hubble variation frame of reference. When one takes angular Gaussian window averages of the Hubble expansion a dipole becomes apparent. Wiltshire et al. (2013) show that this dipole is strongly correlated with the residual CMB temperature dipole when both are referred to the LG (or LS) rest frame. If we are to define a new cosmic standard of rest, within which we still observe a residual CMB temperature dipole, then Wiltshire et al. (2013) argue that such a dipole must have a nonkinematic origin. The correlation of the residual CMB temperature dipole and Hubble expansion dipole supports the proposal that structures in the nonlinear regime of expansion are simultaneously responsible for these effects. Thus we are interested in finding a frame of reference in which this correlation is maximized.

We have investigated this question, and find that correlation of the residual CMB temperature and Hubble variation dipoles under arbitrary boosts does not offer a viable characterization of the minimum variation frame. Kay (2013). By making boosts in the appropriate direction from the LG we are able to artificially increase both the magnitude of the Hubble expansion dipole and the CMB residual temperature dipole simultaneously, at the expense of also increasing the monopole expansion variance. By calculating the Hubble expansion dipole and higher multipole coeffi-

© 2015 RAS, MNRAS
Figure 9. The variation in the spherically averaged Hubble law: (a) $\delta H_s$ in the LG frame; (b) $\delta H_s$ in frame $X$; (c) the systematic boost offset between the LG and frame $X$. (Note: the first unprimed shell 1 is shown.)

Figure 9.

5 HUBBLE EXPANSION VARIATION IN THE COSMICFLOWS-2 CATALOGUE

Thus far our investigation has been based entirely on the COMPOSITE catalogue of galaxy data. In this section we aim to repeat the monopole Hubble expansion variation analysis on the recently released Cosmicflows-2 (CF2) catalogue. Systematic differences become apparent in this analysis which we will investigate in §6.

5.1 Spherically averaged Hubble expansion and treatment of biases

We repeat our earlier analysis to calculate $\delta H_s$ in the same two configurations of 11 spherical shells. The CF2 data is presented with a modified “recession velocity” ($cz$) and

CF2 is a compilation of distances and redshifts from both new and existing sources of observational data. The entire catalogue consists of over 8000 galaxies both locally and extending beyond the scale of statistical homogeneity. This compilation has a large subset of galaxies and galaxy clusters in common with the COMPOSITE sample, including the large SFI++ sample of Springob et al. (2007). In total, the distances are determined by six different methods and compiled together by Tully et al. (2013). The CF2 data is presented in two sets, one with all individual galaxies included, and one condensed into galaxy groups, including groups consisting of one galaxy. We will use the entire data set of 8162 galaxy redshifts and distances, freely available from the extragalactic distance database.

Note: the first unprimed shell 1 is shown.

C⃝ 2015 RAS, MNRAS,
a raw observed redshift. Given the prevalence of the use of such modifications, particularly in bulk flow studies, it is worthwhile to briefly investigate the effect such modifications can have on monopole Hubble expansion variation. Tully et al. (2013) define an adjustment assuming a FLRW cosmology with $\Omega_M = 0.73$ and $\Omega_\Lambda = 0.27$. This adjustment is given by a Taylor expansion to $O(z^3)$ of a homogeneous isotropic expansion law,

$$cz_{\text{mod}} = cz [1 + \frac{1}{6}(1 - q_0)z - \frac{1}{6}(1 - q_0 - 3h_0^2 + 1)z^2]$$ \(17\)

where $q_0 = 0.5(\Omega_M - 2\Omega_\Lambda)$ and $z$ is the redshift in the CMB frame.

Since we wish to deal with cosmological model–independent quantities, this is not the type of adjustment that should be made. In particular, a homogeneous isotropic expansion law cannot be assumed below the scale of statistical homogeneity if the conclusions of Wiltshire et al. (2013) are correct. Nor should such an expansion law be assumed in the CMB rest frame. However, for completeness we consider the adjustment \(17\) in order to rule it out as the cause of much larger systematic differences we will discuss shortly.

Fig. 10 shows $\delta H$ using the CF2 sample. It is immediately evident that these plots are very different to those of Wiltshire et al. (2013) for the COMPOSITE sample. To study the nature of this systematic difference we make use of the SFI++ sample, which is a subset of both the COMPOSITE and CF2 samples. SFI++ (Springob et al. 2007) consists of Tully-Fisher (1977) Relation (TFR) derived distances for 4,861 field and cluster galaxies. Since Springob et al. (2007) supply the SFI++ sample with and without corrections for Malmquist biases, it makes it an ideal candidate with which to study the effects that the treatment of such biases has on the spherically averaged Hubble expansion.

To understand the effect of the Malmquist bias corrections applied by Springob et al. (2007), we calculate the monopole variation of the Hubble expansion for the SFI++ sample with and without corrections. Figure 11 shows the significant difference in $\delta H_s$ between these two treatments. Since uncorrected SFI++ data points are included in the CF2 catalogue, we can determine whether there is any systematic difference between this subsample and the remainder of the CF2 catalogue. If we take the CF2 catalogue and remove the 3625 points in common with our SFI++ sample, we arrive at Fig. 12. We find that $\delta H_s$ does not change to any statistically significant extent, by removal of this potentially biased data. This is an indication that the systematic bias present in the SFI++ raw distances – uncorrected for Malmquist bias – is likely to also be present in the rest of the CF2 data. Thus we are confident that the discrepancy seen between the monopole variation of Fig. 11 for CF2 and Fig. 3 of Wiltshire et al. (2013) for the COMPOSITE sample is a systematic issue, arising from the the treatment of Malmquist bias in the CF2 catalogue, as we discuss further in \[13\].

5.2 The systematic boost offset revisited

One may ask whether, despite the obvious problems with a systematic bias in the CF2 data, any of the analyses applied to the COMPOSITE can nonetheless yield meaningful results.

\[\text{Fig. 10. The monopole Hubble expansion variation for the CF2 sample without the FLRW “correction” \[17\] (black filled circles) and with the FLRW “correction” \[17\] (blue crosses) in the: (a) Local Group frame of reference; (b) CMB frame of reference.}\]
Table 2. Hubble expansion variation in radial shells in CMB and LG frames for the CF2 data. Spherical averages (3) are computed for two different choices of shells, \( r_s < r \leq r_{s+1} \), the second choice being labeled by primes. In each case we tabulate the number of data points per shell, the weighted mean distance, \( \bar{r}_s \); the shell Hubble constants, \( (H_s)_{\text{LG}} \) and \( (H_s)_{\text{CMB}} \) and their associated uncertainties in the LG and CMB frames for both the raw redshifts and those adjusted with (17).

| Shell s  | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| \( N_s \) | 579 | 946 | 834 | 936 | 959 | 794 | 739 | 670 | 497 | 825 | 333 |
| \( r_s (h^{-1}\text{Mpc}) \) | 2.00 | 12.50 | 25.00 | 37.50 | 50.00 | 62.50 | 75.00 | 87.50 | 100.00 | 112.50 | 156.25 |
| \( \langle r_s \rangle (h^{-1}\text{Mpc}) \) | 3.41 | 16.67 | 30.07 | 43.49 | 55.59 | 67.99 | 80.40 | 93.57 | 105.34 | 128.00 | 186.90 |
| \( (H_s)_{\text{CMB}} \) | 177.3 | 110.6 | 110.8 | 102.4 | 102.3 | 100.9 | 99.4 | 96.9 | 94.5 | 90.5 |
| \( (\sigma_s)_{\text{CMB}} \) | 10.5 | 1.5 | 1.0 | 0.8 | 0.8 | 0.8 | 0.8 | 0.9 | 0.7 | 0.9 |
| \( (H_s)_{\text{CMB, adjusted}} \) | 177.7 | 111.3 | 111.9 | 104.1 | 103.2 | 102.0 | 99.6 | 97.7 | 94.9 | 94.9 |
| \( (\sigma_s)_{\text{CMB, adjusted}} \) | 10.5 | 1.5 | 1.1 | 0.8 | 0.8 | 0.8 | 0.8 | 0.9 | 0.7 | 1.0 |
| \( (H_s)_{\text{LG}} \) | 112.2 | 103.6 | 110.0 | 108.4 | 103.7 | 101.8 | 100.9 | 99.5 | 96.5 | 94.9 | 90.4 |
| \( (\sigma_s)_{\text{LG}} \) | 6.6 | 1.4 | 1.0 | 0.8 | 0.8 | 0.8 | 0.8 | 0.9 | 0.7 | 0.9 |
| \( (H_s)_{\text{LG, adjusted}} \) | 112.6 | 104.2 | 111.1 | 109.8 | 105.3 | 103.8 | 103.2 | 102.1 | 99.2 | 98.0 | 94.8 |
| \( (\sigma_s)_{\text{LG, adjusted}} \) | 6.7 | 1.4 | 1.1 | 0.8 | 0.8 | 0.8 | 0.8 | 0.9 | 0.7 | 1.0 |

| Shell s' | 1'  | 2'  | 3'  | 4'  | 5'  | 6'  | 7'  | 8'  | 9'  | 10' | 11' |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| \( N_s \) | 869 | 867 | 846 | 989 | 889 | 777 | 643 | 648 | 412 | 625 | 333 |
| \( r_s (h^{-1}\text{Mpc}) \) | 6.25 | 18.75 | 31.25 | 43.75 | 56.25 | 68.75 | 81.25 | 93.75 | 106.25 | 118.75 | 156.25 |
| \( \langle r_s \rangle (h^{-1}\text{Mpc}) \) | 10.76 | 23.54 | 36.85 | 49.29 | 61.86 | 74.59 | 87.01 | 99.37 | 111.95 | 133.10 | 186.90 |
| \( (H_s)_{\text{CMB}} \) | 126.1 | 109.2 | 109.6 | 103.6 | 101.8 | 102.2 | 99.2 | 99.4 | 95.6 | 94.0 | 90.5 |
| \( (\sigma_s)_{\text{CMB}} \) | 2.5 | 1.2 | 0.9 | 0.8 | 0.7 | 0.8 | 0.8 | 0.8 | 0.9 | 0.8 | 0.9 |
| \( (H_s)_{\text{CMB, adjusted}} \) | 126.7 | 110.0 | 110.9 | 105.1 | 103.6 | 104.4 | 101.6 | 102.2 | 98.4 | 97.3 | 94.9 |
| \( (\sigma_s)_{\text{CMB, adjusted}} \) | 2.5 | 1.2 | 0.9 | 0.8 | 0.8 | 0.8 | 0.8 | 1.0 | 0.8 | 1.0 |
| \( (H_s)_{\text{LG}} \) | 109.0 | 103.8 | 111.6 | 105.6 | 102.5 | 101.6 | 99.5 | 99.0 | 95.6 | 94.5 | 90.4 |
| \( (\sigma_s)_{\text{LG}} \) | 2.1 | 1.1 | 1.0 | 0.8 | 0.7 | 0.8 | 0.8 | 0.9 | 0.9 | 0.9 |
| \( (H_s)_{\text{LG, adjusted}} \) | 109.5 | 104.6 | 112.9 | 107.1 | 104.3 | 103.8 | 101.8 | 101.7 | 98.4 | 97.8 | 94.8 |
| \( (\sigma_s)_{\text{LG, adjusted}} \) | 2.1 | 1.1 | 1.0 | 0.8 | 0.8 | 0.8 | 0.8 | 1.0 | 0.8 | 1.0 |

Figure 11. The monopole Hubble expansion variation for the SFI++ sample without corrections for Malmquist bias (black filled circles) and with corrections (blue crosses) in: (a) the Local Group frame of reference; (b) the CMB frame of reference.
We applied the analysis of §4.1 in the CF2 catalogue, but found no reference frame in which $\chi^2$ approaches unity, or even within the same order of magnitude. While some decrease in the variation of Hubble expansion was found for boosts in the galactic plane, the uncertainties were too large to give any statistically significant results. Such investigations must be abandoned until the bias problems in the CF2 catalogue are dealt with.

By contrast, we found that in spite of the bias problem, the signature of a systematic boost offset studied in §3 is nonetheless evident in CF2, as this involves the difference of the $H_0$ values in the LG and CMB frames from Table 2 as plotted in panels (a) and (b) of Fig. 10.

Fig. 12 shows the results of repeating the analysis used in §3 using the unadjusted CF2 distances. The best fit value for $p$ in §3 is found to be $p = -0.83 \pm 0.17$ for unprimed and $p = -0.86 \pm 0.26$ for primed shells. Varying the shell boundaries as in §3 we find a value of $p = -0.84 \pm (0.21)_{\text{stat}} \pm (0.06)_{\text{sys}}$. However, if we compare Fig. 13 with Fig. 10 we see that there are more data points with $(H_0)_{\text{CMB}} < (H_0)_{\text{LG}}$, which do not conform to the power law §3. However, the range of distances of the shells for which this is true coincides in Fig. 10 and Fig. 13 being $40 h^{-1} \lesssim r \lesssim 60 h^{-1}\text{Mpc}$ in the COMPOSITE sample and $30 h^{-1} \lesssim r \lesssim 67 h^{-1}\text{Mpc}$. This is consistent with the hypothesis that aside from the systematic boost offset, there are structures responsible for nonlinear deviations in the monopole Hubble expansion in the range identified in the COMPOSITE sample, but untreated biases in the CF2 catalogue have broadened the range of distances attributed to the same structures.

The systematic boost offset is still evident in the innermost shells of the CF2 sample. However, the fit is somewhat worse than in the COMPOSITE sample due to more data points lying in the increased range which deviates from §3.

Nonetheless, we can still check if the boost offset signature is unique to the angular direction of residual CMB dipole in the LG frame. Fig. 14 shows the value of $f_\ell$ from §4 which represent the best fit parameters for a systematic boost offset for boosts of 635 km sec$^{-1}$ across the entire sky. The results are consistent with those found for the COMPOSITE sample, providing further evidence that this is indeed not a random statistical outcome but is consistent with out hypothesis.

We have also repeated the analysis of §4.5 for the best fit to a systematic boost offset from the LG frame, as shown in Figure 15. The angular directions found are consistent with our results for the COMPOSITE sample. For example, on the 200 km sec$^{-1}$ sky map the best fit value is in the direction $(\ell, b) = (55^\circ, -5^\circ)$, close to that found earlier. However, the value of $p = -0.92 \pm 0.75$ has a far greater statistical uncertainty, which may well be due to the untreated biases.

In conclusion, the greater number of data points in the CF2 catalogue may potentially yield statistically more accurate results than the COMPOSITE sample. However, at present this is prevented on account of untreated biases, to which we now turn.

6 COSMICFLOWS-2 MALMQUIST TREATMENT

The Cosmicflows-2 (CF2) and COMPOSITE catalogues deal with Malmquist distance biases in different ways. The results of the last section indicate that the treatment of such biases is crucial in establishing the actual nature of the variation of cosmic expansion below the statistical homogeneity scale. Therefore we will perform further analyses to better understand these differences.

Since the treatment of Malmquist bias is complex, we will first briefly remind the reader that in its current usage this term refers to at least three distinct biases that affect the average derived distance of a galaxy cluster:

(i) Selection bias This is the simple homogeneous systematic
error that results from using objects in a magnitude limited sample (Malmquist 1922). If the galaxies in a cluster have a true average flux $\langle F \rangle_0$, since we can only observe the brightest of these, we measure a biased average flux $\langle F \rangle_b > \langle F \rangle_0$, with the effect of bias increasing with distance. Thus the luminosity distance $d_L^2 = \langle L \rangle_0 / (4\pi \langle F \rangle)$ will be less when using a biased average flux, where $\langle L \rangle_0$ is the average luminosity as determined from standard candles calibrated using nearby objects.

(ii) **Homogeneous distribution bias** This is a systematic average error for objects around the same derived distance which can be understood in terms of statistical scatter. If we assume a standard Gaussian scatter, $\sigma$, in the derived distances about an estimated mean, then – since the radial number density grows as $N(r) \propto r^3$ – there are more objects with true distances larger than the estimated distance, than smaller. Thus at a given derived distance distance, more galaxies will have been scattered by the errors down from larger true distances than up from smaller ones.

Lynden-Bell et al. (1988), Hanski (1999). This is equivalent to giving more statistical weight to more distant values and thus the probability distribution for the true distance is no longer Gaussian along the line of sight, centered on the measured distance, instead being skewed towards greater distances. For the special case of a constant density distribution, for example, one arrives at Eddington’s formula (Eddington 1914)

$$E(\mu_{\text{true}} | \mu_{\text{der}}) = \mu_{\text{derived}} + 1.382\sigma^2,$$

where $\mu \equiv 5 \log r + 25$ is the standard distance modulus when $r$ is given in Mpc.

(iii) **Inhomogeneous distribution bias** The inhomogeneous bias is analogous to the homogeneous one in that involves a systematic error in the statistical scatter of objects due to their distribution in 3-dimensional space. However, in this case it arises as number counts are higher in regions of greater density, resulting in a systematic scatter of measurements out of higher into lower density regions (Strauss & Willick 1995).
Thus the inhomogeneous effect is very sensitive to large variations in large-scale structure along the line of sight. Failure to account for this type of bias can give spurious infall signatures onto high density regions. This bias is of course far more difficult to account for, requiring accurate density fields for structure along the line of sight for each observation.

The selection and homogeneous distribution biases typically lead to underestimates of distances that increase as the distance grows, so that a plot of $\mu_{\text{true}} - \mu_{\text{h}}$ versus redshift has a positive slope. However, the inhomogeneous distribution bias can lead to the opposite effect. For example, Feast (1987) shows that in applying the TFR method when the spatial density of objects at a given 21cm line width is constant, then the required Malmquist correction is the classical one given by Eddington (1914). However, when this is not the case it is possible to obtain overestimated distances (Feast 1987).

Moreover, regardless of the details of any inhomogeneous matter distribution, one just needs the spherically averaged $N(r)$ to decrease sufficiently quickly for the direction of scatter in the standard homogeneous distribution bias to be reversed, giving overestimated distances.

Some biases can be dealt with in the data reduction. In particular, applying the inverse TFR method rather than the direct method one can effectively eliminate the selection bias (Schechter 1980), leaving only a considerably smaller bias (Willick 1994, 1995).

In the CF2 catalogue (Tully et al. 2013) use an inverse TFR procedure to reduce the selection bias only, stating that only a small subsequent correction for residual bias is required. In particular, they “make no adjustments for the distribution Malmquist effects” in their reported CF2 distances (Tully et al. 2013). Their calibration carried out for this relation follows the procedures of Tully & Pierce (2000), Courtois & Tully (2012) and Sorce et al. (2013).

On the other hand, the SFI++ catalogue (Springob et al. 2007) which forms the major part of the COMPOSITE sample includes corrections to account for all homogeneous and inhomogeneous biases in their data. In their view their own treatment of the homogeneous and inhomogeneous distribution biases was “straightforward”, given access to a reconstruction of the local density field (Erdoğdu et al. 2006). However, they stated that their treatment of the selection bias was ad hoc because the selection criteria used are designed to mimic the observational properties of the survey as closely as possible, and so are very inhomogeneous. They therefore provided both the raw and corrected distances, should other researchers adopt alternative methods for dealing with selection bias.

The COMPOSITE sample incorporates the SFI++ distances with the Malmquist bias corrections of Springob et al. (2007), whereas CF2 uses a subset of uncorrected SFI++ distances. While we do not have the data to independently repeat any of the bias corrections, we are able to test the consistency of assumptions made by Tully et al. (2013).

6.1 The SFI+++ subsamples of CF2

Tully et al. (2013) find that for 2071 common points between their own survey and the SFI+++ survey (excluding 5 points judged to be “bad”) there was a “correction” of the form

$$\Delta \mu_1 = 0.492(\pm 0.011) + 0.000031(\pm 0.000002)cz_{\text{LS}}$$

where $\Delta \mu_1$ is the CF2 distance modulus with the zero point established by Courtois & Tully (2012), and $cz_{\text{LS}}$ is the unadjusted modulus with a nominal zero point consistent with $H_0 = 100\frac{\text{km}}{\text{sec}}\frac{\text{Mpc}}{\text{Mpc}}$, and $z_{\text{LS}}$ is the raw redshift in the

\[ \text{FLRW adjustment} \] is not applied.

---

Figure 15. Best fit parameters to (13) using unadjusted CF2 data for boosts from the LG of: (a) 450 km sec$^{-1}$; (b) 200 km sec$^{-1}$. In all figures, the galactic longitudes $\ell = 0^\circ, 180^\circ, 360^\circ$ are on the right edge, centre and left edge respectively.
rest frame of the Local Sheet, which is close to the Local Group frame (Tully et al. 2008). We independently confirm the slope in (19) using the appropriate zero point \( z = 0.06 \) and plot this in Figure 10. Note that this comparison is made for a subset of the SFI++ sample, henceforth SFI++A, consisting only of objects that are common between the SFI++ survey and the independently obtained CF2 distances. In their final analysis Tully et al. (2013) use averages of the CF2 and SFI++ distances with double weight given to the CF2 distances.

The intercept in (19) is determined by a scaling of the data due to the Malmquist selection bias. Let us now compare (16) to the Malmquist correction used by Springob et al. (2007) in the SFI++ sample, and subsequently adopted in the COMPOSITE sample. We repeat the analysis of Fig. 10 on the SFI++A subsample but now using the distances as corrected by Springob et al. (2007), which include both selection and distribution bias corrections. We find a linear relationship of the form

\[
\Delta \mu_2 = 0.0356(\pm 0.0063) - 0.000012(\pm 0.000001)cz_{\text{LS}}
\]

(20)

where \( \Delta \mu_2 \equiv \mu_{\text{SFI++ corrected}} - \mu_{\text{SFI++}} \) is the difference in distance moduli between the corrected and raw distances. The data and best fit line are displayed in Figure 15.

We can immediately see from (19) and (20) that there is a significant difference between the corrections. For small redshifts the Springob et al. (2007) correction is positive indicating raw distances are underestimated, while for large redshifts the correction is negative indicating raw distances are overestimated. Adjusting the intercept of Fig. 15(a) to zero we then find a hierarchy \( \mu_{\text{SFI++ corrected}} < \mu_{\text{SFI++}} < \mu_{\text{CF2}} \) in the limit of large redshifts. This is consistent with the observation of Watkins & Feldman (2013) that: “the distances are systematically larger in the Cosmicflows-2 catalogue [than in the COMPOSITE catalogue] due to a different approach to bias correction”.

The fact that Fig. 20(b) has a negative slope means that the dominant correction cannot arise solely from selection and homogeneous distribution biases, since as noted above both of these effects underestimate true distances. The difference therefore must be due to the treatment of the inhomogeneous distribution bias, which Springob et al. (2007) have included but Tully et al. (2013) have not.

The 1970 points in the SFI++ sample that are not also contained in the original CF2 survey, henceforth SFI++B, have been incorporated into CF2 without using their correction (19). This data covers a larger range, up to redshifts of almost \( z = 0.1 \), whereas SFI++A only covers up to \( z = 0.06 \). Tully et al. (2013) state that these points, if corrected using (19) cause a “highly significant decrease in the Hubble parameter with increasing velocity”. We independently verified this result. Thus Tully et al. (2013) do not adjust these distances, instead claiming that they are of a different nature altogether, the main difference being that these consist of cluster samples from a different survey (Dale et al. 1999a). In these samples rotation information for the galaxies was obtained from optical spectroscopy rather than the standard 21cm Hydrogen line widths. However, Tully et al. state that “it is not clear to us why this component of SFI++ does not manifest the selection Malmquist bias”.

Since it appears two halves of the SFI++ have been incorporated into CF2 in different ways we determine whether they do actually show different characteristics. Considering \( \Delta \mu_2 \) for SFI++B we find a correction

\[
\Delta \mu_2 = 0.0417(\pm 0.0061) - 0.000012(\pm 0.000001)cz_{\text{LS}}
\]

(21)

which has an intercept consistent within uncertainties and identical slope to (20). Since there is no apparent difference using this test we repeat a similar analysis to that performed by Tully et al. (2013, Fig. 10) to test for bias. This analysis is based on the fact that selection bias is manifest by an increase in the Hubble parameter with redshift (Teerikorp 1995), for data binned by redshift. In Figs. 17 and 18 we repeat the analysis of the Hubble constant in redshift bins performed by Tully et al. (2013) for the subsets of interest, and compare the results.

In Fig. 17 we produce plots equivalent to Tully et al. (2013, Fig. 10) for the SFI++A and SFI++B subsamples, both using raw distances. We subsequently find that the difference in the Hubble constant in individual redshift bins for the SFI++A and SFI++B ranges from 0.03\( \sigma \) to 1.8\( \sigma \) in individual bins. The weighted mean of these differences is 0.84\( \sigma \), and thus we do not see a significant difference between SFI++A and SFI++B.

We note that two mutually consistent halves of the SFI++ sample have been incorporated into CF2 in different ways, but are uncertain as to what impact this may have on the full CF2 catalogue.

On the other hand the COMPOSITE sample (Feldman et al. 2010) uses a much larger subset of SFI++ distances corrected for Malmquist biases (after rejection of outliers). While it appears that there may be inconsistencies in the inclusion of the raw SFI++ distances into the CF2 catalogue, it is possible that the SFI++ corrected distances are subject to systematic error also. It is for this reason that Springob et al. (2007) included both corrected and raw distances, to allow others to take on the challenging task of Malmquist bias corrections. However, the CF2 catalogue is only corrected for selection bias, leaving the correction of the distribution homogeneous and inhomogeneous Malmquist biases as a task for the user.

The intercept in (19) is determined by a scaling of the data so we are not interested in independently confirming this for our investigation.

A decreasing Hubble constant below the scale of statistical homogeneity is, to a limited extent, what is expected from the analysis of Wiltshire et al. (2013). Thus trends which appear anomalous as compared to a standard FLRW expectation should not automatically be regarded as a signal of unaccounted observational bias. However, there are also systematic differences that occur when binning in redshift, as in Figs. 17 and 18, as opposed to binning in distance with \( \delta \), so careful analysis is required to make sense of the different approaches. The direct calculation of the Hubble parameters in each bin is also different to that described in \( \delta \), as it is not clear which method Tully et al. (2013) use. We apply a simple weighted average of \( cz_i/r_i \) values and obtain consistent results (although as no values are tabulated by Tully et al. 2013) we can only verify by inspection.
Defining the frame of minimum Hubble expansion variance

Figure 16. The difference in distance modulus between: (a) the CF2 and raw SFI++ distances for the points common to the two surveys, as presented in Tully et al. (2013). (b) the SFI++ raw and corrected catalogues for the same points.

Figure 17. The Hubble parameter, $H_i = c z_i / r_i$, computed for each individual data point (coloured points) and from averaging in 1000 km sec$^{-1}$ bins (black points) using the: (a) SFI++A subsample; (b) SFI++B subsample. (c) Comparison of the averaged points in (a) and (b) with blue crosses being from SFI++A and black filled circles from SFI++B.
As another test of differences between the SFI++A and SFI++B subsamples we have also repeated the analysis of Fig. 13 but now to compare the raw and corrected distances within each subsample. Fig. 18 shows the comparison for the SFI4++B subsample. In this case the difference in the Hubble parameters in each bin vary from a minimum of 0.01σ to a maximum 2.2σ in individual bins, with a weighted mean difference of 1σ. For SFI++A the difference in the Hubble parameters vary from 0.05σ to 1.9σ in individual bins, with a weighted mean difference of 1.0σ. Thus again we do not see a significant difference between the subsamples.

It may appear surprising that the raw and uncorrected data only differ by 1σ on average when binned by redshift. However, once inhomogeneous Malmquist bias is accounted for the sign of the correction is different at large redshifts as compared to low redshifts, meaning that for an intermediate range the correction is small. The approach by Tully et al. (2013) of binning in redshift is not an appropriate one to use when performing a parameter minimization that involves boosts to rest frames in which the redshift is changed, as in §3. Rather we followed Wiltshire et al. (2013) in binning by distance. This led to differences between the raw and corrected data sets which are statistically much more marked than is evident if one bins by redshift.

7 DISCUSSION

We have investigated the extent to which it is possible to define a standard of cosmological rest based on a frame in which variance of the spherically averaged (monopole) Hubble expansion is a minimum. Such averages do not make any assumptions about the geometry of space below the scale of statistical homogeneity (≤100h⁻¹Mpc) of the sort implicit in the standard peculiar velocities framework, which assumes a Euclidean spatial geometry.

We studied the systematic variation that arises when an arbitrary boost is made from a frame of reference in which the spherically averaged Hubble expansion is most uniform. We found that such a systematic variation is indeed detected to a statistically significant extent between the CMB and LG frames of reference, using the COMPOSITE sample (Watkins et al. 2009; Feldman et al. 2010).

This strengthens the proposition made by Wiltshire et al. (2013) that the Local Group may be a more suitable cosmic rest frame than the standard CMB frame, and that consequently a significant fraction of the CMB dipole may be nonkinematic. Given that the CMB frame is still the defacto choice for the cosmic rest frame, this conclusion would have a far reaching impact for many areas of cosmology (Wiltshire et al. 2013), including the question of large angle CMB anomalies (Ade et al. 2014).

We extended the analysis to search for an improvement on the Local Group as the standard of rest, using a variety of tests. Very large boosts from the LG frame can be excluded if we simultaneously demand that while the residual variance of the Hubble expansion should be small in the regime of nonlinear expansion at small distances, a clear signature of an emerging linear Hubble law should also be found at larger distances. However, we found that in applying all possible tests there is still freedom to perform quite large boosts close to the plane of the galaxy, presumably because the lack of data in the Zone of Avoidance leads to a lack of suitable constraints there. This hypothesis could potentially be tested on simulated data using exact solutions of Einstein’s equations (Boleiko et al. 2015).

Since our conclusions depend on the COMPOSITE sample, we repeated our analysis insofar as it was possible for the recently released Cosmicflows-2 sample (Tully et al. 2013). This catalogue of 8,162 galaxy, groups and cluster distances is considerably larger than the the COMPOSITE sample, and potentially could provide more accurate results although it is also of course limited in the Zone of Avoidance.

We have found very significant differences in the results for the CF2 and COMPOSITE samples, as can be seen by comparing Fig. 3 of Wiltshire et al. (2013) and our Fig. 10. These result from differences in the treatment of the Malmquist bias between the SFI++ catalogue and CF2 catalogue, as previously noted by Watkins & Feldman (2013). We also found apparent inconsistencies in the manner of inclusion of subsamples of the SFI++ catalogue into the CF2 catalogue, with respect to the treatment of Malmquist bias. More significantly, since the reported CF2 distances do not include corrections for the inhomogeneous distribution Malmquist bias they may be of limited use until such corrections are applied.

The conclusions of Wiltshire et al. (2013) are dependent on the treatment of the Malmquist bias in the SFI++ catalogue being accurate, as this constitutes the largest part of the COMPOSITE sample. Naturally, one might question whether any systematic procedure of Springob et al. (2007) could somehow spuriously lead to an unusually uniform Hubble expansion in the LG frame through an error in the Malmquist bias procedure.

We find no grounds for this. In particular, our analysis shows that the difference between the CMB and LG frames has the distinctive signature of a systematic boost offset noted by Wiltshire et al. (2013). Nothing in the Malmquist bias correction procedure of Springob et al. (2007) could obviously introduce this signature through a systematic error. Their analysis does not single out the LG, or LS, frame in any way; indeed all their redshifts are referred to the CMB.
frame. Furthermore, although the remaining bias means that the CF2 sample is currently unusable from the point of view of determining a frame of minimum spherically averaged Hubble expansion variance in the nonlinear regime, we observed in §5.2 that the difference of the CMB and LG frame spherical averages nonetheless still shows the signature of the systematic boost offset in the CF2 data.

Since the boost offset is detectable in the independently reduced CF2 data, it cannot be an artefact of the Malmquist bias treatment of Springob et al. (2007). Furthermore, the departure of the nonlinear expansion from the boost power law (13) that is seen when comparing Figs. 1 and 13 is precisely what is to be expected if there are additional unaccounted-for uncertainties in individual CF2 distances as compared to the SFI++ ones; the distance range of structures counted uncertainties in individual CF2 distances as compared to the SFI++ ones: the distance range of structures.

If Fig. 10 was based on accurate distances, it would imply that the Hubble expansion in all frames of reference is far less uniform than might reasonably be expected in any viable cosmological model; in particular, there is a monopole or “Hubble bubble” variation of order 15–20% in the range $20 < r < 60$ h$^{-1}$Mpc in the CF2 sample, as compared to 4–5% in the COMPOSITE sample. The largest “Hubble bubble” variation that has ever been claimed on such scales using more accurate Type Ia supernovae distances in the CMB frame is $6.5 \pm 2.2\%$ (Zehavi et al. 1998).

Tully et al. (2013) chose not to correct for the distribution biases, as they wished to separate “the issues of distance measurements and velocity field inferences”. Indeed, in the peculiar velocities approach the distribution bias may be much less significant. In new work Hoffman, Courtois & Tully (2015) use the CF2 catalogue to reconstruct large scale structure by means of the Wiener filter and constrained realizations of Gaussian fields assuming a WMAP constrained ΛCDM model as a Bayesian prior. They observe that “the Malmquist bias introduces a spurious strong monopole term into the reconstructed velocity field but is expected to hardly affect the bulk velocity which is associated with the dipole of the velocity field” Hoffman et al. (2013). This would appear to be the counterpart of the large monopole we observe in Fig. 10 in our analysis, Hoffman et al. (2013) corrected for the bias but noted that the bulk velocity analysis is “virtually unaffected by the Malmquist bias”.

In our case, bulk flows on scales $\gtrsim 100$ h$^{-1}$Mpc may be an artefact of using the CMB rest frame as the standard when it does not coincide with the frame of minimum Hubble expansion variance [Wiltshire et al. 2013]. Thus large scale bulk flows are not our primary interest. Rather, we are interested in detecting the systematic monopole variance [10]. In distinction to the peculiar velocity approach our method by necessity is sensitive to a monopole bias. In fact, our method of binning in radial shells by distance with an anchoring to $H_0$, is particularly sensitive to any distribution bias which follows from a number density, $N(r)$, with strong gradients. The bias effect in Fig. 10 can be largely reproduced by applying a uniform Hubble law to the CF2 redshifts, adding Gaussian scatter to create a mock distance catalogue, and then applying our binning strategy (R. Watkins, private communication).

It may be possible to construct the 635 km sec$^{-1}$ velocity attributed to the LG within the ΛCDM model, as recently done by Hess & Kitaura (2014) who used constrained N-body simulations and nonlinear phase space reconstructions to arrive at a value $v_{LG} = 685 \pm 137$ km sec$^{-1}$. However, this itself does not constitute a proof of the standard kinematic interpretation, but rather a verification within the 20% uncertainty of a computer simulation. The ΛCDM model is certainly phenomenologically very successful, and any competing model can only be viable insofar as many of its predictions are close to the standard model, as is the case, for example, in the timescape cosmology [Wiltshire 2004, Nazer & Wiltshire 2013]. What is important in testing the standard model is to seek observations which are not expected in its framework. Although the signature of the systematic monopole boost offset [10] between the CMB and LG frames should be checked in ΛCDM simulations, it is not an observation that should obviously arise if we have purely a FLRW geometry with local Lorentz boosts. Properly characterizing and determining a frame of minimum Hubble expansion variance is therefore a fundamental question open to more precise observational tests in future.

In conclusion, we have defined a degenerate set of frames that are candidates for a cosmic standard of rest based on minimum spherical averaged Hubble expansion variance. The set of frames includes the LG frame but excludes the CMB frame. The degeneracy in the definition of the rest frame is associated with a freedom to boost in the plane of the galaxy, probably due to a lack of constraining data in the Zone of Avoidance.

The larger CF2 sample may potentially tighten the constraints on the definition of the minimum Hubble variance frame. However, it is first necessary to reduce the data in the manner of Springob et al. (2007) to remove the inhomogeneous distribution bias which appears to be the source of the large discrepancies found in [5]. The Universe is very inhomogeneous below the scale of statistical homogeneity, and ironically it is only once the biases associated with such inhomogeneities are removed that a picture of a remarkably uniform average Hubble expansion actually emerges.

A careful treatment of inhomogeneous Malmquist bias is therefore key to for the future progress of our understanding of the nature of cosmic expansion as the surveys grow ever larger.

ACKNOWLEDGMENTS

We thank Ahsan Nazer for discussions, and Rick Watkins, Brent Tully and Hélène Courtois for helpful correspondence.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543.
Ade P.A.R., et al., 2014, A&A, 571, A23.
Aghanim N. et al., 2014, A&A, 571, A27.
Bilicki M., Peacock J.A., Jarrett T.H., Cluver M.E., Steward L., 2014, arXiv:1408.0799.
Bolejko K., Nazer M.A., Wiltshire D.L., 2015, in preparation.
Bondi H., 1947, MNRAS, 107, 410
Buchert T., 2000, Gen. Relativ. Grav., 32, 105.
Buchert T., 2008, Gen. Relativ. Grav., 40, 467.
APPENDIX A: LINEAR REGRESSION VIA TOTAL LEAST SQUARES

Consider a general linear model with errors in both the dependent and independent variables. We can express such a model as

$$y_t = y_0 + \beta x_t$$

where $(Y_t, X_t) = (y_t, x_t) + (e_t, u_t)$

The measurement errors to be normally distributed with a covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{ee} & \sigma_{eu} \\ \sigma_{ue} & \sigma_{uu} \end{pmatrix}.$$  (A2)

We can extend this analysis to allow for different errors at each point, which we will refer to as weights given by $\omega(X_t) = 1/\sigma_{uu,t}$, $\omega(Y_t) = 1/\sigma_{ee,t}$, and the correlation coefficient between the errors given by $\gamma_t = \sigma_{eu,t}\sqrt{\omega(X_t)\omega(Y_t)}$.

To carry out a least squares minimization we must calculate the statistical distance from an observation to the true value. In standard least squares this is the vertical distance from the data point to the model, since the values of the independent variable are assumed to be exact. Now, in the most simple case we would have the squared Euclidean distance from the observed data points to true value in the model as

$$[Y_t - (\beta_0 + \beta_1 x_t)]^2 + (X_t - x_t)^2 = \epsilon_t^2 + u_t^2$$  (A3)
but if the variances of $e_i$ and $u_i$ are different from unity
this statistical distance becomes $\sigma_{ee}^2 e_i^2 + \sigma_{uu}^2 u_i^2$, and if these
variances are correlated we must use the covariance matrix
to give the “statistical” distance

$$ (Y_i - \beta_0 - \beta_1 x_i, X_i - x_i) \Sigma^{-1} (Y_i - \beta_0 - \beta_1 x_i, X_i - x_i)^T. \quad (A4) $$

In our case we must determine the values of $\beta_0$ and $\beta_1$ that
minimize $(A4)$. That is, we must find the values $(\hat{\beta}_0, \hat{\beta}_1)$ that minimize this sum for the given observations.
First we fix the $x_i$ values by treating them as unknown con-
tstants in a standard linear regression of the form

$$ \begin{bmatrix} Y_i - \beta_0 \\ X_i \end{bmatrix} = \begin{bmatrix} \beta_1 \\ 1 \end{bmatrix} x_i + \begin{bmatrix} e_t \\ u_t \end{bmatrix} \quad (A5) $$

for which the generalized least squares estimator gives

$$ \hat{x}_t = [\beta_1, 1] \Sigma^{-1} (\beta_1, 1)^T \quad (A6) $$

Substitution of $\hat{x}_t$ into $(A4)$ gives

$$ \frac{(Y_i - \beta_0 - \beta_1 X_i)^2}{(\sigma_{ee} - \beta_1 \sigma_{eu} + \beta_1^2 \sigma_{uu})} \quad (A7) $$

so that after summing over all $N$ points we obtain

$$ S = \sum_{i=1}^N \frac{(Y_i - \beta_0 - \beta_1 X_i)^2}{(\sigma_{ee} - \beta_1 \sigma_{eu} + \beta_1^2 \sigma_{uu})}, \quad (A8) $$

which is the expression to be minimized. In (York 1969) the
linear equation that minimizes $(A8)$ is given by

$$ \beta_1 = \frac{\sum_{i=1}^N Z_i^2 V_i \left[ \frac{U_i}{\omega(Y_i)} + \frac{\beta_1 V_i}{\omega(X_i)} - \frac{\alpha_i V_i}{\alpha_i} \right]}{\sum_{i=1}^N W_i^2 U_i \left[ \frac{U_i}{\omega(Y_i)} + \frac{\beta_1 V_i}{\omega(X_i)} - \frac{\alpha_i V_i}{\alpha_i} \right]} \quad (A9) $$

where

$$ \alpha_i^2 = \omega(X_i) \omega(Y_i), \quad U_i = X_i - \bar{X}, \quad V_i = Y_i - \bar{Y}, \quad \bar{X} = \sum_{i=1}^N Z_i X_i / \sum_{i=1}^N X_i \quad \text{and} \quad \bar{Y} = \sum_{i=1}^N Z_i Y_i / \sum_{i=1}^N Y_i, \quad Z_i = \frac{\omega(X_i) \omega(Y_i)}{\omega(Y_i) + b^2 \omega(Y_i) - 2 b \gamma_i \alpha_i}. $$

Clearly $(A9)$ requires an iterative process to find $\beta_1$ which
begins with an initial guess which may be found from per-
forming a standard linear regression assuming the $X_i$ to be
exact. After $\beta_1$ is obtained the value of $\beta_0$ is found from the
fact that the mean must be on the best fit line and thus

$$ \beta_0 = \bar{Y} - \beta_1 \bar{X}. $$

The uncertainties in the parameter values, $\sigma_{\beta_0}$ and $\sigma_{\beta_1}$, are
(Titterington & Halliday 1979)

$$ \sigma_{\beta_0}^2 = \frac{\sum Z_i x_i^2}{(\sum Z_i^2)(\sum x_i^2) - (\sum Z_i x_i)^2}, \quad (A10) $$

$$ \sigma_{\beta_1}^2 = \frac{\sum Z_i}{(\sum Z_i^2)(\sum x_i^2) - (\sum Z_i x_i)^2}. \quad (A11) $$

We now return to the transformed model from $(A9)$ which takes on the form

$$ \log(\delta H_i) = p \log(\langle r_i^2 \rangle) + \log \left( \frac{v_i^2}{2H_0} \right) \quad (A12) $$

such that we may identify $y_i = \log(\delta H_i)$, $x_i = \langle r_i^2 \rangle$, and

$$ (\beta_0, \beta_1) = \left( p, \frac{v_i^2}{2H_0} \right). $$

© 2015 RAS, MNRAS