Department of Physics and Astronomy
University of Canterbury, Private Bag 4800
Christchurch, New Zealand

PHYS 690 MSc Thesis

submitted in partial fulfilment of the requirements for

The Degree of Master of Science in Physics

The cosmological rest frame

---

by

James H. McKay

Supervisor: Prof. David L. Wiltshire

May 4, 2015

Contents

Acknowledgements

I would like to thank Prof. David Wiltshire for his supervision of this project and helpful suggestions. I also wish to thank Rick Watkins and Ahsan Nazer for discussions on the technical aspects of this project.

Abstract

The analysis of the uniformity of a spherically averaged Hubble expansion in the Local Group frame of reference by Wiltshire, Smale, Mattsson and Watkins (2013) is extended. We carry out an investigation to constrain the frame of reference from which the spherically averaged Hubble expansion is the most uniform by applying arbitrary Lorentz boosts to the data. The proposition of a systematic boost offset between the Hubble expansion in the Local Group and CMB reference frames is verified within statistical uncertainties. This evidence further supports the claim that the Local Group is closer to the frame of reference in which Hubble expansion should be considered. We subsequently carry out a statistical analysis in search of a frame of minimum expansion variation and find consistent results with the systematic boost offset analysis. However, there is a considerable degeneracy to perform boosts in the plane of the galaxy, which may be a consequence of a lack of constraints from the Zone of Avoidance where data is absent. The COMPOSITE sample of 4,534 galaxies is used primarily, with the key results repeated with the recently released *Cosmicflows*-2 sample of 8,162 galaxies.

The treatment of Malmquist distance bias is investigated in the context of the *Cosmicflows*-2 and COMPOSITE samples. We find systematic differences in the inclusion of the large SFI++ subsample into these catalogues. These differences are explored and the origin of Malmquist distance bias reviewed. We find the *Cosmicflows*-2 data produces results which naïvely suggest more variation of cosmic expansion than would be expected in any cosmological model when the methods of Wiltshire *et al.* are applied. We trace this discrepancy to the fact that the distribution Malmquist biases have not been corrected for in the *Cosmicflows*-2 survey.

CHAPTER 1

Introduction

Before the 16th century the notion that our civilization was not fixed at centre of the universe was preposterous. With the progress of astronomy the more familiar concept of the Earth *moving* in space soon overtook this geocentric view. With the question resolved in our own solar system, we now face similar questions on a cosmological scale. The notion of cosmological rest and the consequent assumptions involved have implications throughout observational cosmology. Is any observer truly *at rest* in our vast and dynamic universe, and if so, who? In this thesis we will attempt to address the problem of rest on a cosmological scale.

The perspective we view the universe from is crucial in modern cosmology. If one was to measure the speed of vehicles on a motorway while travelling oneself, one's results would certainly be different to those of a stationary observer on the road side. Likewise our telescopes see the universe differently to the view we would have if our planet, and even our galaxy, were thrust on a wildly different course relative to nearby matter. Fortunately, if the travelling observer on the motorway knew his own velocity relative to the stationary observer it would be possible to work out the true speed of the passing vehicles relative to the road. When one records cosmological observations from a telescope a similar process is followed until the ambiguity of "what is stationary in the universe" gives us infinitely many possible outcomes.

The *cosmic standard of rest* lacks a unique definition from the fundamental principles of physics. The discovery of the Cosmic Microwave Background (CMB) radiation gave cosmologists the most widely accepted definition to date. Early observations measured

a CMB that was isotropic yet as measurement precision improved angular variations were detected. The most significant of these is a temperature dipole. If we assume that there is a frame of reference in which the temperature of the CMB appears perfectly isotropic and homogeneous, then we should be able to infer the velocity of our motion with respect to this frame from the departures from isotropy. The observed dipole then arises due to an effective "scooping up" of radiation in the direction of our movement exactly as one expects from using special relativity alone. By calculating the velocity required to create the same anisotropy using special relativity, a velocity was obtained for the motion of the Earth relative to the CMB radiation [1–3].

The frame of reference in which the CMB radiation has no dipole anisotropy is defined as the CMB frame. In the standard cosmology this is believed to be the absolute rest frame for a Friedmann-Lemaître-Robertson-Walker (FLRW) expanding universe, and thus the frame in which cosmological observations are generally considered with respect to. If the CMB is both cosmological and isotropic, Stewart and Sciama claimed in 1967 that this will only be so in the rest frame of distant matter which last scattered the radiation [4]. This implies that any frame of reference with a velocity relative to the surface of average homogeneity must have a kinematic dipole anisotropy.

The discovery of a background radiation agreed well with the phenomenology of Einstein's general relativity. Not only was it crucial for establishing the hot big bang theory as the basis for standard cosmology, but it also gave a means of defining a comoving observer. The standard cosmology makes the assumption of homogeneity and isotropy of space at all times, thus allowing cosmic expansion to be characterised by a single scale factor, $a(t)$, in the FLRW geometry

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a(t)^2\left(\frac{\mathrm{d}r^2}{1-kr^2} + \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right) \tag{1.1}$$

where $k = -1, 0, 1$ determines the spatial curvature. With a dust source this solution describes either an expanding or contracting universe. An observer at fixed spatial coordinates in this geometry is then defined as comoving. In the standard cosmology, a comoving observer is expected to see the CMB as isotropic and homogeneous. This has since provided a reference with which cosmologists can model the universe's largest structures and their evolution. Thus, if anything is to be the currently defined *cosmic standard of rest* it is the CMB frame.

As measurements of the background radiation and the structure in the universe become more precise inconsistencies begin to arise. If one assumes that the CMB dipole

anisotropy has a kinematic origin anomalies are found, as seen in the PLANCK analysis [5]. The details of these inconsistencies are discussed in §3.4. In addition these simplified models of the universe rely on the assumption that space is on average homogeneous and isotropic. While this assumption holds when we average over huge scales, on smaller scales the complicated nature of spacetime can bring into question the simplicity of a standard of rest which applies everywhere and always. We will discuss these problems in the following sections.

1.1 The expansion of the universe

The realisation that we live in an expanding universe has been fundamental to both observational and theoretical cosmology. The initially crude observations of receding galaxies have advanced to unveil the greatest mysteries of modern science. From Hubble's law to dark energy, the study of expansion has been a principle ingredient for defining new problems in cosmology.

The earliest conjectures regarding the nature of our universe were highly speculative and lacked reliable observational evidence. William Herschel was one of the first astronomers to study the distribution of matter in the universe in detail. In 1785 Herschel [6] produced a disc-shaped model galaxy based on star counts. However, he eventually conceded to his fault in assuming that all stars have the same luminosity [7]. As Herschel discovered increasingly many nebulae, away from the bounds of his initial disc-shaped model, he was unable to determine an edge for the galaxy and lost faith in his model. It was not established that these "nebulae" were in fact galaxies at vast distances, and thus it continued to be assumed that the universe was infinite in space and had the density of our galaxy throughout.

The lack of a definitive description of the matter distribution within our universe enabled a diverse range of theoretical cosmologies to be explored. Einstein's general relativity had been highly successful in explaining gravitational interactions in the solar system, such as the gravitational bending of light and the precession of Mercury. The next application was to gravitational interactions on the scale of the known universe.

In his first attempts at a cosmological model for the universe Einstein was driven mainly by philosophical principles. He demanded that there be no beginning in time and that the universe did not change on average. This in addition to the commonly accepted view of the matter distribution constrained Einstein to a non-evolving universe that had

the density of our galaxy everywhere. However, the original equations based purely on gravitational interactions would not admit such a solution. If dense enough, gravitationally attracting matter would eventually cause the universe to contract. Likewise if the density of matter was low enough, it would expand forever. Thus in order to achieve the desired model Einstein added an extra term to the field equations to effectively oppose gravity. The resultant "Einstein static universe" [8] is in fact inherently unstable; given a small change in the model parameters results in either infinite expansion or contraction towards a "big crunch". Despite the lack of success of this model, the additional cosmological constant term remained in future applications of the general relativistic field equations.

Not long after Einstein's static universe was postulated Friedmann [9,10] and Lemaître [11] independently discovered the general $k = \pm 1$ solutions described by (1.1). These solutions model a universe of expanding or contracting gravitational matter. These dynamical solutions in which the universe possibly had a beginning or an end challenged the commonly accepted notion that the universe was static. This negated Einstein's original purpose in introducing a cosmological constant, although both authors allowed for it as a possible non-zero parameter. The observational evidence in support of an expanding universe became widely established shortly after this discovery.

Up until the late 1920s the work of theoretical cosmologists in applying general relativity to the universe attracted little attention in the wider scientific community. This changed with the discovery of the distance-redshift relation by Hubble[1]. Hubble observed Cepheid variable stars in nebulae at distances far beyond the bounds of the known galaxy. From these observations Hubble discovered an, albeit rough, linear relationship between the distance to the object and the amount that the radiation it emits had been redshifted due the recessional motion [12]. This also led to the understanding that these "nebulae" are in fact other galaxies. Figure 1.1 gives an example of the linear relationship Hubble discovered and a typical amount of scatter from this linear trend which will becomes a focal point of our investigation.

The redshift observed from distant celestial objects was readily converted into a corresponding "recessional velocity"[2] by observationalists using the simple radial redshift

[1] In fact, Lemaître was first in using Hubble's data to publish an article accounting "for the radial velocity of extra-galactic nebulae" [11]. However, his work was overlooked at first.

[2] This choice of wording must be treated with care as the effect is not purely due to a recessional velocity but rather the expansion of space between the observer and the object. For example, on large scales the "recessional velocity" can become in excess of the speed of light, which is simply due to the

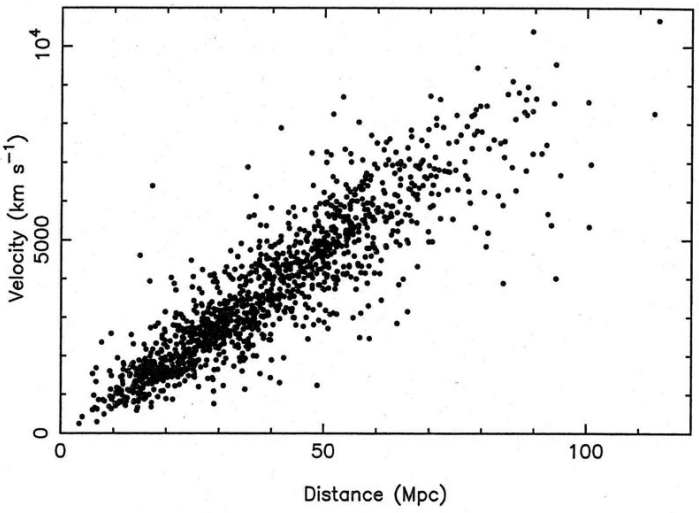


Figure 1.1: The linear Hubble relationship is visible in this plot of 1355 galaxies' apparent recessional velocity and distance. The considerable scatter is assumed to be due to statistical uncertainty and the random motion of galaxies.

law from special relativity. For an object receding in flat space we have the relation

$$1 + z = \sqrt{\frac{1+\beta}{1-\beta}} \tag{1.2}$$

where $\beta = v/c$ and v is the recessional velocity. Although the use of (1.2) for cosmological applications was a rather naïve assumption, as we will come to understand, it served as strong evidence in support of expanding universe models. However, the mechanism behind this apparent expansion raised still further questions. The Milne universe was one alternative to the metric expansion of space as proposed by Friedmann and Lemaître. In the Milne universe the geometry was not time dependent and the receding galaxies were the result of an explosion of matter which did not affect the geometry [13]. On the other hand, in response to Hubble's discovery, Einstein and de-Sitter advocated for a spatially flat Friedman universe ($k = 0$) with the cosmological constant set to zero [14]. For the next seven decades the need for the cosmological constant was questioned while the evidence of an expanding universe continued to mount. Eventually as this evidence become increasingly more accurate and focused to larger and larger redshifts a new discovery was made that has seen the cosmological constant make a return.

expansion of the intervening space rather than a violation of the laws of physics.

The standard cosmology that we have today has been shaped by a range of independent observations over the past two decades. Galaxy surveys in the early 1990s indicated a matter density of only 15-30% [15] which implied that the universe is either spatially open or that a large amount of the density in the universe was not accounted for. The observations of Type Ia supernovae became the tipping point for the proposal of a dark energy dominated universe. These stellar explosions are visible from several thousand megaparsecs away and are believed to be an incredibly accurate distance indicator [16]. The study of the redshifts of these very distant objects is consistent with a universe that was expanding at a *slower* rate in the past. This observation, contrary to what one would expect in a universe of gravitationally attracting matter, has provided evidence for the existence of a cosmological constant [16] or "dark energy"[3] component to make up the total density of the universe. The existence of a cosmological constant is also consistent with recent studies of the CMB power spectrum [17], the Baryon Acoustic Oscillations [18–20] and large scale structure [21–23].

The observations we have briefly discussed, along with the work of numerous other scientists, have culminated in the ΛCDM standard cosmology. This model consists of a cosmological constant (Λ), associated with dark energy, and cold dark matter, along with ordinary baryonic matter and radiation. While consistent with observations to date, the standard cosmology is still relatively young and many challenges still exist for observational cosmology. One such issue is the inhomogeneity of space, which we will explore in the following section.

1.2 The inhomogeneous universe

The *Copernican principle* states that we should not occupy a privileged spatial position in the universe. The consequence of this in cosmological models is that many researchers use this as a justification for the *Cosmological principle* that the geometry of the universe must be spatially homogeneous and isotropic. However, in reality we have only a *statistical* Copernican principle due to the large amount of structure visible in our local universe. In such a way, we occupy a planet in a typical galaxy in a typical cluster in a typical filament. Thus in some statistical sense the universe is indeed homogeneous

[3]Technically, if dark energy is assumed to be a perfect fluid with equation of state $P = w\rho c^2$, then $w < -1/3$ is the defining requirement that makes cosmic acceleration possible. A cosmological constant has $w = -1$, and such a value is favoured observationally.

and isotropic when averaged over large enough scales. This coarse graining over suitably large scales enables one to model the universe with a single scale factor as in (1.1). Thus the statistical Copernican principle is fundamental to the ΛCDM cosmology but does not lead uniquely to it, as there is no priori reason why the geometry that describes statistical averages must be an exact solution of Einstein's equations.

In the standard model any variations from the homogeneous background are modeled using perturbation theory or N-body Newtonian simulations. However, perturbation theory only holds if the perturbations are not too large (in the realm of linear theory) and Newtonian simulations are only valid for non-relativistic speeds and weak gravitational fields assuming a given FLRW background. We will now discuss the relevance of inhomogeneity in the study of cosmology and further highlight the inadequacies of the standard model to address this problem.

The scale that we must average over to claim homogeneity is a matter of some debate [24–27]. Indeed, although the universe started out very smooth we now observe a high level of structure even beyond the scale of galaxy clusters. The 7th release of the Sloan Digital Sky Survey [30] and subsequent analysis [31] found that 62% of the volume studied was occupied by voids, typically of diameter $30\,h^{-1}$Mpc [28, 29], surrounded and threaded by filaments of galaxies. Here h is related to the Hubble constant, H_0, by $H_0 = 100h$ km s^{-1} Mpc. Similar studies such as the 2dF [32] and 6dF [33] galaxy redshift survey have provided further insight into the large scale structure in the universe (see Figure 1.2). From this we can see why statistical homogeneity is only achieved at scales of order $100\,h^{-1}$Mpc. This is sufficiently larger than the largest typical structures, which is consistent with the most conservative estimates of $70\,h^{-1}$Mpc [34]. Although larger structures do exist, such as the $320\,h^{-1}$Mpc long Sloan Great Wall [35] and the $350\,h^{-1}$Mpc long Large Quasar Group [27], Scimgeour *et al.* have argued that these fluctuations are still compatible with a homogeneity scale below $100\,h^{-1}$Mpc, especially since structures such as this tend to be filamentary, which will be averaged over in volume statistics [34].

The study of inhomogeneous cosmology has produced many alternatives to homogeneous models. These alternatives include exact inhomogeneous solutions to Einstein's field equations. For example, the Lemaître-Tolman model [36, 37] describes a universe with spherically symmetric but radial inhomogeneity. This model has been successful in explaining the accelerated expansion seen in the supernovae data without the need for

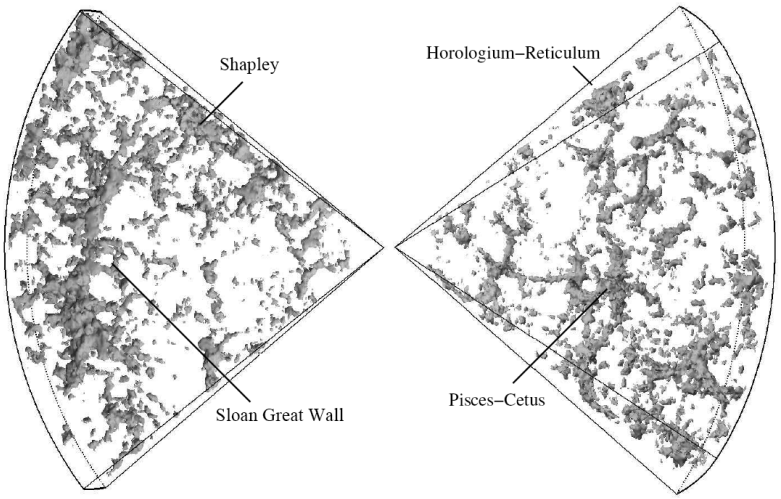


Figure 1.2: Three dimensional reconstruction of the inner parts of the 2dF Galaxy Redshift Survey. Figure reconstructed from Colless *et al.* [32] by the Delaunay Tessellation Field Estimator [40].

dark energy [38]. However, by demanding spherical symmetry one must abandon the Copernican principle in order to define some origin for the universe. The Szekeres [39] model is a generalisation of the Lemaître-Tolman model in which spherical symmetry of the metric is no longer required. However, the fact that these models strongly violate the Copernican principle means that they are unlikely to be physical candidates for a model of the actual universe. Other models comprise of exact solutions like the Lemaître-Tolman model stitched together on a FLRW background. These Swiss-Cheese models satisfy the Copernican principle, but still demand that the global average geometry is exactly a FLRW geometry, which the first principles of general relativity do not require.

Another alternative to the standard homogeneous models is to challenge the assumption of average FLRW evolution. The scalar averaging approach of Buchert is one such proposition [41, 42]. By averaging scalars, such as those describing the density, expansion and shear of spacetime, Buchert avoided the mathematically complicated problem of averaging tensors[4], such as the spacetime metric. Such averaging techniques also highlight the problem that the process of averaging and time evolution are non-commutative.

[4]There is no mathematically well-defined way to average tensors, as any approach breaks general covariance in some fashion. Zalaletdinov has proposed a scheme for averaging tensors [45–47]. However, this requires the introduction of additional mathematical structures and physical assumptions which may not be relevant for the specific case of cosmological averages [48].

The early universe was almost homogeneous and isotropic, but once deviations from homogeneity become nonlinear with respect to perturbation theory about a homogeneous universe, they can influence the average global expansion. This effect was first called *back reactions of the inhomogeneities* and was found to act as an additional energy density [43].

The use of these averaged inhomogeneous cosmological models has become an alternative to the standard model containing dark energy. Along with the possibility that backreaction could account for some of the cosmic acceleration (see [44] for a complete discussion on dark energy and the averaging problem) inhomogeneity can also affect the path travelled by light rays and the calibration of clocks and rods (that is, how we compare time and length scales between different observers in the universe). Because of this, inhomogeneity can skew measurements of redshifts and distances [50] which may be misconceived as cosmic acceleration.

The timescape cosmology [50–52] is one such inhomogeneous model based on a particular physical interpretation of the Buchert averaging formalism [41, 42]. As a viable alternative to the standard model the timescape cosmology does not require dark energy in order to be consistent with observations. By approaching the problems of gravitational energy and average cosmic evolution [49] the timescape model has been developed into an alternative to the standard homogeneous cosmology and has remained viable in tests to date [50, 58–61]. The phenomenology of the timescape model motivates the investigation in this thesis, and is consistent with our results. However, the analysis performed is model independent and thus remains free of theoretical bias.

1.3 Peculiar velocities and differential expansion

The exact scale of statistical homogeneity may still be a matter of debate, yet the existence of inhomogeneity below this scale is not in question. Nevertheless the assumption of a homogeneous universe is prevalent in analysis well within this range. In the literature it is generally assumed that any deviations from a linear Hubble law are due to gravitational clustering of matter on a FLRW background. These random peculiar velocities are incorporated into Hubble's law, $cz = H_0 r$, as

$$v_{\rm pec} = cz - H_0 r \tag{1.3}$$

where cz is the redshift times the speed of light and r is the luminosity distance to the object of interest. For large enough redshifts the peculiar velocity is significantly smaller than the magnitude of the motion due to the expansion of space, and in such a case the original Hubble law is sufficient.

When considering the peculiar velocity of each galaxy in a particular survey a "field" of velocities can be constructed. The study of the peculiar velocity field is a large part of observational cosmology. Finding patterns in this field of velocities, and tendencies towards a particular direction have become known as bulk flow studies. In conventional understanding, all peculiar velocities are driven by the peculiar gravitational potential, the infall of objects towards higher density regions. This is simply the result of gravitational clustering. Kashlinsky, Atrio-Barandela, Kocevski and Ebeling report to have found a "strong and coherent" bulk flow – the so-called "dark flow" – of order 600-1000 km s^{-1} on scales of at least $300\,h^{-1}$Mpc, which is at the limit of their data [62, 63] by using the kinematic Sunyaev-Zel'dovich effect. This result has led to claims of a tilted universe, in which inhomogeneities exist beyond our horizon that are so large they affect motions over massive scales. Such a finding is in disagreement with the standard model which requires homogeneity on scales far smaller than this. These claims have meet with criticism, including the PLANCK 2013 results which report to show no statistically significant dark flow [64]. The reason for the difference in these claims in particular may be the result of the statistical techniques used, with Keisler [65] finding that Kashlinsky *et al.* did not properly account for primary anisotropies in the CMB, which play an important part in their study. Atrio-Barabdela has produced a "minority report" [66], in which he claims the dark flow is present in the PLANCK results, with the apparent discrepancy with Kashlinsky *et al.* [62, 63] not being an issue about a signal, but rather its statistical significance.

Due to the variety of data sets and methodologies used to find patterns in the peculiar velocity field there are a range of contradictions even amongst claims of bulk flows. Using the large COMPOSITE data set Watkins, Feldman and Hudson find a bulk flow of 407 ± 81 km s^{-1} towards $(l, b) = (287°, 8°) \pm (9°, 6°)$ in galactic coordinates [67], with 90% of this sample within $107\,h^{-1}$Mpc. However with the *Cosmicflows*-2 [68] data set Watkins and Feldman using the same method found a bulk flow of 262 ± 60 km s^{-1} on the $100\,h^{-1}$Mpc scale, citing treatment of bias in the distances as the explanation for this difference [69]. Different methodologies also create varied results. For example, Davis

and Nusser [70] find a bulk flow of 333 ± 38 km s^{-1} towards $(l, b) = (279°, 10°) \pm (3°, 3°)$ within a $100\,h^{-1}$Mpc sphere using large subsamples of the COMPOSITE catalogue, while on the other hand the work of Turnbull *et al.* [71] finds a bulk flow of 249 ± 76 km s^{-1} in the direction $(l, b) = (319°, 7°) \pm (18°, 14°)$. Although this is consistent with the larger bulk flow of Watkins, Feldman and Hudson, it is inconsistent with that found by Kashlinsky *et al.* [62, 63].

The inconsistencies that seem to plague the peculiar velocity framework bring us to question the assumptions at its basis. According to general relativity the notion of a boost only makes sense at a point. The concept of objects thousands of parsecs away having a "velocity" relative to us is conceptually flawed[5]. While gravitational clustering is an inevitable cause of scatter in the data, and indeed a result of inhomogeneity, more care is required when the universe is also expanding. The peculiar velocity framework makes a strong geometrical assumption over and above what is demanded by general relativity. In particular, the quantity $v_{\rm pec}$ defined by (1.3) 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 FLRW model with a single cosmic scale factor, $a(t)$, whose derivative defines a single global Hubble constant, $H_0 = \dot{a}/a|_{t_0}$. 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. When space is inhomogeneous, regions of different density will expand at different rates. The expansion of regions that have a high density of gravitationally attracting matter will be retarded by gravitational attraction, whilst the expansion of the far less dense void regions will take longer to slow down. On scales below that of statistical homogeneity we expect the motions of celestial objects to be more complex than those in a homogeneous universe, even if we could ignore the random local velocities due to gravitational clustering. For example, if a void lies between us and an object of interest, that object will appear to be receding faster

[5]One can meaningfully ascribe a local velocity to any galaxy in terms of its local Lorentz boost with respect to a canonical local observer, and then parallely propagate this quantity along null geodesics to any other observer if the exact geometry of the universe is known. However, velocity is an intrinsically local notion, whereas its common use by astronomers is global, implicitly assuming that a global set of spatial axes exists – something general relativity does not require.

than if a dense wall of galaxies was in its place.

Differential expansion is a feature of general exact solutions to the cosmological Einstein equations, such as the Lemaître–Tolman–Bondi (LTB) [36,37,72] and Szekeres [39] 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, Smale, Mattsson and Watkins [73] examined the variation of the Hubble expansion from a fresh perspective, by generalising the earlier approaches of [101] and [103]. In particular, given that there is a notion of statistical homogeneity on large ($\gg 100\,h^{-1}$Mpc) scales, then an average expansion law characterised by a single asymptotic Hubble constant, $\bar{H}_0$, is applicable on such scales. However, from the first principles of general relativity one should make no geometrical assumptions about cosmic expansion below the statistical homogeneity scale. One can nonetheless perform radial and angular averages of the distance versus redshift of a large sample of galaxies in spherical shells, and compare the results with the asymptotic Hubble constant in order to quantify the variation of the Hubble expansion.

Wiltshire *et al.* [73] conducted such an analysis on the COMPOSITE sample of 4,534 cluster, group and galaxy distances [74,75], with the following results:

- A linear Hubble law with a spherically averaged Hubble constant, H_s, which is statistically indistinguishable from the asymptotic Hubble constant, $\bar{H}_0$, is found to emerge in independent radial shells with mean distances in the range $\bar{r}_s > 70\,h^{-1}$Mpc.

- On scales $r \lesssim 65\,h^{-1}$Mpc the spherically averaged value, H_s, in independent shells is greater than the asymptotic value, $\bar{H}_0$. However, the difference is significantly larger in the standard rest frame of the cosmic microwave background (CMB) radiation when compared to either the rest frame of the Local Group (LG) of galaxies or the Local Sheet (LS). In other words, the spherically averaged Hubble expansion is more uniform in the LG rest frame than in the CMB rest frame, with very strong Bayesian evidence $\ln B \gg 5$. The uniformity of expansion in the LG and LS frames is statistically indistinguishable.

- By a variety of angular tests, the residual variation of the spherical (monopole) Hubble expansion is found to be correlated with structures in the range $32\,h^{-1}$ – $62\,h^{-1}$Mpc, which give a Hubble expansion dipole with a markedly different character in the CMB and LG frames.
- A skymap of angular variation of the Hubble expansion in the LG frame (constructed by Gaussian window averaging [103]) has a very strong dipole. The angular expansion skymap has a correlation coefficient of -0.92 with the residual CMB temperature dipole in the LG frame.

The first of the results above is consistent with other observations which find that a notion of statistical homogeneity emerges at scales of order $70\,h^{-1}$ – $100\,h^{-1}$Mpc [24, 34]. Furthermore, the fact that $H_s > \bar{H}_0$ on the $\lesssim 65\,h^{-1}$Mpc scales that the Hubble expansion is nonlinear agrees well with the observation that the largest typical structures in the late epoch Universe are voids of diameter $30\,h^{-1}$Mpc [28, 29]. Pan *et al.* [31] found that voids occupy 62% of the volume studied in the 7th release of the Sloan Digital Sky Survey [30]. If one constructs averages in spherical shells (see the conceptual schematic in Figure 1.3) then once the shells are a few times larger than the diameter of the largest typical nonlinear structures, a well defined average, $\bar{H}_0$, 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 $\bar{H}_0$.

Wiltshire approaches the problem of differential expansion in the timescape model [50–52]. By making the appropriate choice of rest frame Wiltshire postulates that we can reduce the scatter in the Hubble expansion on scales below that of statistical homogeneity. In the standard cosmology, since comoving observer are assumed to be those at rest on the surfaces of average homogeneity, it follows that the cosmic rest frame is the one in which the Hubble expansion is most uniform with the smallest statistical fluctuations. Given the evidence of a strong CMB dipole, which is of order 10^2 larger than higher order multipoles, ever since the CMB dipole was first detected [53] it has been assumed that the CMB frame should also be the frame of minimum Hubble expansion variation.

Given the potential importance of such a result, it is important to try to characterise the frame of minimum Hubble expansion variation in purely observational terms. Wiltshire *et al.* [73] compared the LG and LS frames with that of the CMB, motivated by the

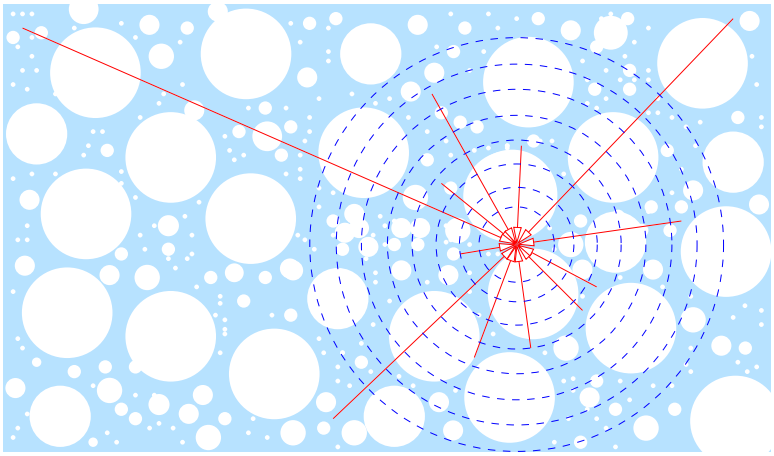


Figure 1.3: From Wiltshire *et al.* [73]. Schematic diagram of spherical averaging. The universe is described as ensemble of filaments, walls and voids: expanding regions of different density which have decelerated by different amounts and therefore experience different local expansion rates at the present epoch. If one averages cz/r in spherical shells (dotted lines) about a point then once the shells are a few times larger than the typical nonlinear structures, an average Hubble law with small statistical scatter is obtained,whereas there are considerable deviations for shells on scales comparable to the typical nonlinear structures.

fact a frame close to the LG frame would be the natural standard of rest according to the "Cosmological Equivalence Principle" [49] which underlies Wiltshire's approach [50–52] to the averaging problem in inhomogeneous cosmology [54–57].

It is the aim of this thesis to determine in a model–independent fashion whether a rest frame of minimum Hubble expansion variation can be found among all the frames boosted by arbitrary amounts with respect to the LG and CMB rest frames, and what the frame is. It will also be necessary to consider the different ways to characterise such a minimum variation frame. While our principlal results will follow Wiltshire *et al.* [73] by analysing the COMPOSITE sample [74, 75], we have also considered the recently published *Cosmicflows*-2 sample of Tully *et al.* [68]. We find that the 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 discussion highlights how the implicit assumption of a FLRW expansion law below the scale of statistical homogeneity via (1.3) subtly influences the manner in which such biases are treated in practice.

1.4 Summary of research

The original work in this thesis is presented in chapters 4 and 5, and §2.2. Chapter 4 focuses on the search for a frame of minimum Hubble expansion variation, as a standard of rest. We consider a range of methods to constrain the parameters for a such a frame of reference. The hypothesis that structure in the range of $40h^{-1} - 60\,h^{-1}$Mpc [73] has important implications on our cosmological observations is further supported with our findings. In chapter 5 we consider Hubble expansion variation in the recently released *Cosmicflows-2* [68] catalogue of distances and repeat the key results from the previous chapter. Systematic differences in the treatment of distance biases in this data set and the COMPOSITE sample cause fundamental problems in our analysis. A detailed review of these biases is given in chapter 2, along with a brief independent investigation into the various differences between the distance catalogues in §2.2. In chapter 3 we introduce the relevant theory and review the recent work of Wiltshire *et al.* [73] on variation in the Hubble expansion, thus providing the necessary framework for our extension of this work.

CHAPTER 2

Bias in distance determination

The assignment of a distance to astrophysical objects has been one of the greatest challenges in astronomy. While redshift can be determined directly only the smallest astronomical distances can be measured using the direct parallax method, with all others requiring indirect measurements. Even so, using a variety of indirect distance indicators large catalogues of galaxies have been compiled with known distances and redshifts. Yet once a "raw" distance is obtained further corrections are often required. This is due to biases which arise in the techniques used to assign distances. Of particular interest in this investigation is Malmquist bias.

The study of large scale bulk flows has motivated the compilation of dense, far reaching surveys. We will focus on two catalogues in particular, the COMPOSITE sample compiled by Watkins, Feldman and Hudson [74] and updated by Watkins and Hudson [75] and the *Cosmicflows*-2 (CF2) sample compiled by Tully *et al.* [68]. These catalogues both make use of a range of distance indicators across multiple independent surveys and provide redshift, galactic latitude and longitude, distance and distance uncertainty. Because of the dependence of our work on accurate distance determination we investigate the systematic differences that arise in the construction of these catalogues.

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[1]. The COMPOSITE sample consists of distances for

[1] *Cosmicflows*-2 distances retrieved 14/10/14 from http://edd.ifa.hawaii.edu/.

4,534 galaxies, galaxy groups and clusters. These compilations have objects in common, including the large SFI++ sample of Springob et al. [76] which we will focus on in our investigation.

2.1 Malmquist bias

The term Malmquist bias originated in 1922 in reference to a particular type of statistical selection bias [77]. This refers to a bias in the luminosity of objects in magnitude limited samples. In its modern use Malmquist bias also refers to biases that occur as a geometric consequence of performing observations in three dimensional space and the inhomogeneous spatial distribution of objects. Correcting for these biases is an important part of any astronomical survey. Unfortunately the process is not trivial and there is no standardized treatment. We will explore the modern definition of Malmquist bias and one method of correction with particular attention to the differences between the COMPOSITE and *Cosmicflows*-2 distance catalogues. For a complete review of Malmquist bias the reader is referred to Teerikorpi [78]. Malmquist bias can be separated into 3 main types.

To begin with we must define what we mean by *bias*. If $\hat{x}$ is an estimator of some quantity, x_0 is the true value, and $E(\hat{x}|x_0)$ is the expected value then the bias is defined as

$$B(x) = E(\hat{x}|x_0) - x_0. \tag{2.1}$$

Thus only if the expected value is equal to the true value is an estimator *unbiased*.

2.1.1 Selection bias

The first type of Malmquist bias is a systematic error in the average derived distance for a group of objects such as a galaxy cluster. This is due to an upper magnitude cut off at large distances since distant objects appear fainter. If in a cluster of interest only the brightest galaxies are given attention we will have a bias towards these more luminous objects. Consequentely if the galaxies in a cluster have an average flux $\langle F\rangle_0$, since we can only observe the brightest of these, we measure a biased average flux $\langle F\rangle_b > \langle F\rangle_0$. 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 (and thus is unchanged). Unless an observer is able to detect even the faintest galaxies, this bias is always going to be present.

Selection bias becomes an issue when using the Tully-Fisher relation (TFR). We will discuss the selection bias in the TFR in more technical detail in §2.1.4, given its importance in the SFI++ and CF2 distance calculations.

2.1.2 Homogeneous distribution bias

The second type of bias is categorized into two subordinate types, the homogeneous and inhomogeneous distribution Malmquist biases. These arise due to the distribution of objects in three dimensional space. The homogeneous Malmquist bias is a systematic average error for objects around the same derived distance.

The homogeneous Malmquist bias can be understood in terms of statistical *scatter*. If we assume a standard Gaussian scatter, σ, 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. So at a given derived distance distance, more galaxies will have been scattered by the errors down from larger distances than up from smaller ones [79, 80]. In other words, a galaxy is more likely to be scattered by error towards the observer, than scattered away. This effect is equivalent to giving more 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.

The adjustment required for homogeneous Malmquist bias can be derived following the method of Butkevich, Berdyugin and Teerikorpi [82] . We begin with the *fundamental equation of stellar statistics*

$$a(m) = \omega \int_0^\infty \phi(M)\rho r^2 \mathrm{d}r \tag{2.2}$$

which is the distribution of stars of apparent magnitude m within a solid angle ω, where $\phi(M)$ is the luminosity function and ρ is the number density of objects at luminosity distance r.

Now we compute the *the mean absolute magnitude, M, for stars of a given apparent*

magnitude m to be

$$E(M|M_0) = \frac{\int_0^\infty M\phi(M)\rho r^2 dr}{\int_0^\infty \phi(M)\rho r^2 dr} \tag{2.3}$$

which can be expressed in terms of (2.2) as

$$E(M|M_0)a(m) = \omega \int_0^\infty M\phi(M)\rho r^2 \mathrm{d}r. \tag{2.4}$$

If we assume a Gaussian luminosity function

$$\phi(M) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(M-M_0)^2}{2\sigma^2}\right] \tag{2.5}$$

then it follows that $M\phi(M) = M_0\phi(M) - \sigma^2\phi'(M)$. Substitution into (2.4) gives

$$E(M|M_0)a(m) = \omega \int_0^\infty \left(M_0\phi(M) - \sigma^2\phi'(M)\right)\rho r^2 \mathrm{d}r \tag{2.6}$$

$$= -\sigma^2\omega \int_0^\infty \phi'(M)\rho r^2 \mathrm{d}r + M_0 a(m) \tag{2.7}$$

and thus

$$\left(E(M|M_0) - M_0\right) a(m) = -\sigma^2\omega \int_0^\infty \phi'(M)\rho r^2 \mathrm{d}r. \tag{2.8}$$

If one now notes that the derivative of $a(m)$ is

$$\frac{\mathrm{d}a(m)}{\mathrm{d}m} = \omega \int_0^\infty \phi'(M)\rho r^2 \mathrm{d}r \tag{2.9}$$

then we may rewrite (2.8) as

$$B(M) = (E(M|M_0) - M_0) = -\sigma^2 \frac{\mathrm{d}\ln(a(m))}{\mathrm{d}m} \tag{2.10}$$

In the special case that the spatial distribution of objects is homogeneous the right hand side of (2.10) is simplified. In particular, if ρ is a constant one obtains the Eddington-Malmquist formula [81], $E(M|M_0) = M_0 - 1.382\sigma^2$, or expressed in terms of distance modulus

$$E(\mu_{\mathrm{true}}|\mu_{\mathrm{der}}) = \mu_{\mathrm{derived}} + 1.382\sigma^2 \tag{2.11}$$

where $\mu \equiv 5\log r + 25$ and r is the corresponding luminosity distance in Mpc. Thus in the most simple case the homogeneous distribution bias results in an *underestimate* of the true distance. However, as we will discuss in §2.1.4 when the distribution of objects is more complicated this is not always the case.

2.1.3 Inhomogeneous distribution bias

The inhomogeneous bias is analogous to the homogeneous one in that errors scatter measurements out of higher into lower density regions [83]. Thus the inhomogeneous effect arises from the 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.

It is apparent that Malmquist bias in all its forms can pose significant issues for distance determination. Depending on the techniques used, treating this bias is essential in obtaining the true distance to other galaxies. However, no standardized treatment is available that will simultaneously account for all three types of Malmquist bias, nor is there a universal understanding of when all three treatments are required. In the following subsection we review one technique for treatment of selection bias in the TFR, as this is of relevance to the catalogues we use. In the subsequent section we investigate how a particular subsample of the *Cosmicflows-2* catalogue was incorporated with respect to the treatment of Malmquist bias.

2.1.4 Malmquist bias in the Tully-Fisher relation

We will demonstrate the problem of Malmquist bias in the context of the TFR for distance determination. In order to calculate the distance to a galaxy one must know its absolute luminosity. This is simplified if a standard candle exists. Standard candles are a class of objects which have known absolute luminosity due to a characteristic quality that may be determined independently of observed brightness. In 1977 Tully and Fisher [84] devised a means of estimating the absolute luminosities of certain types of spiral galaxies. This was achieved by measurement of the maximum speed of rotation of the galaxy, which is correlated with the mass and in turn the absolute luminosity [85]. The TFR still requires calibration from primary standard candles and is thus classed as a secondary standard candle.

The measurement of the maximum speed of rotation was through the observation of the absorption line due to hydrogen atom energy level transitions. The transition of interest corresponds to the 21cm absorption line, and the width of this line is widened by the Doppler effect, thus giving an indication of the galaxies' maximum speed of rotation. This principle can be applied to other absorption lines, such as a radio frequency

transition in the carbon monoxide molecule [86], so is not restricted to the 21cm line, although this is the most commonly used.

By using a sample of galaxies with known distances it is possible to form a linear relationship between an estimator for the absolute magnitude, $\hat{M}$, and the 21cm line width, P,

$$\hat{M} = \alpha P + \beta \tag{2.12}$$

where α and β are constants to be determined by a linear regression on our calibrating sample. We use the convention that an estimator is denoted with a caret, a true value is denoted with a subscript "0" and a random variable without bold face. The way this linear regression is approached is crucial, and demonstrates the effect selection bias can have in the application of the TFR.

There are two options for performing the linear regression required to determine α and β in (2.12). The first is referred to as the *direct* TFR, which involves linear regression of M on P, which gives us a relation for the expected value of M at a given P, which is typically how one makes use of the TFR. On the other hand, one may perform linear regression of P on M, which is known as the *inverse* TFR. The resulting equations of the best fit direct and inverse lines are [87]

$$E(M|P) = M_0 + \frac{\rho\sigma_M}{\sigma_P}(P - P_0) \tag{2.13}$$

$$E(P|M) = P_0 + \frac{\rho\sigma_P}{\sigma_M}(M - M_0) \tag{2.14}$$

where M_0 and P_0 are the mean values, σ_P and σ_M are the standard deviations of the absolute magnitude and line width respectively and ρ is the correlation coefficient for the bivariate distribution. In Figure 2.1 we plot the best fit for the direct (2.13) (blue line) and inverse (2.14) (red line) TFR. Both of these can rearranged to the form of (2.12) as required. The definition of $\hat{M}$ in each case is slightly different; for the direct case $\hat{M}$ is the mean absolute magnitude at the observed line width value. In the inverse case, the value of $\hat{M}$ is such that the observed line width is equal to its expected value, $P = P_0$, when $M = \hat{M}$. Although these two relations will yield different coefficients in our TFR they are both unbiased estimators when there are no selection effects present [87].

For the purpose of an example we will use a bivariate normal distribution to produce a random galaxy sample of absolute magnitudes and 21cm line widths. In Figure 2.1 we plot this random sample, which simulates the scatter expected if the cluster does or does

not suffer from an upper magnitude cutoff. The issue of selection bias in the TFR and possible corrections for it become apparent when we apply an upper magnitude cutoff to our example. This type of bias was realized by Teerikorpi in 1987 and is known as cluster incompleteness bias [88]. Figure 2.1 shows the resultant direct and inverse TFR linear regression on the same sample with and without an upper bound applied to the absolute magnitudes. Since this is an identical sample the distribution of lines widths is not affected. We can immediately see the change in the slope, which will result in wildly different distance estimates.

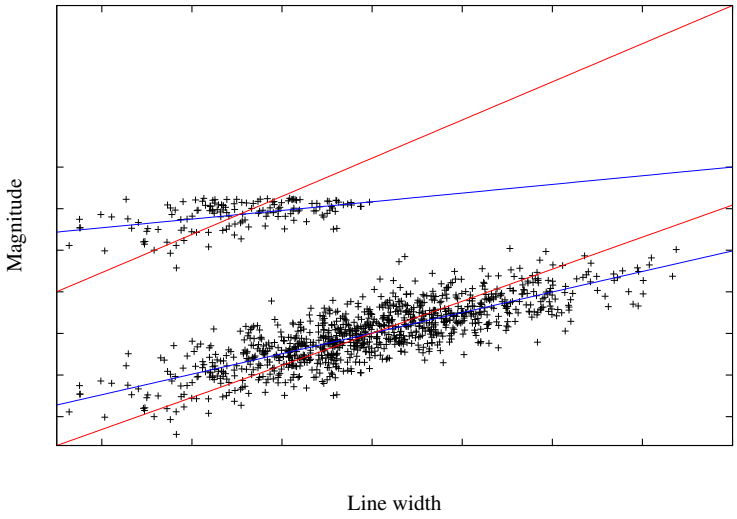


Figure 2.1: The application of the Tully Fisher relation to a cluster of galaxies with (upper sample on plot) and without (lower sample on plot) an upper magnitude cutoff. The blue and the red lines display the result of a linear regression of the form given in (2.12) for the direct and inverse cases respectively. The sample in this example was produced at random and left unchanged in one case and shifted in magnitude and had an upper limit applied in the other.

As we can see an upper magnitude cutoff produces several effects for the TFR.

(1) The derived slope is changed.

(2) The apparent zero-point of the relation in the cluster is too bright.

(3) The scatter becomes apparently low.

Therefore, if we used a nearby cluster to calibrate the TFR, and then apply the relation to a more distant cluster, we would systematically *underestimate* the distance. This is

because, although we would still observe the same distribution of P values with respect to magnitude, as is seen by the increased slope of the direct TFR these are on average at brighter magnitudes. So the expected value of M at a given P is going to be systematically brighter than it would be in the nearby cluster, due to the more distant galaxies fading out.

The inverse relation is commonly used over the TFR. This is because in an ideal situation the distribution of line widths is not restricted by observational selection. As we can see in Figure 2.1 the red line corresponding to the linear regression of the inverse TFR, is unchanged by our magnitude limited sample. This observation was first made by Schechter in 1980 [89]. Willick [90] showed that there does still exist a small bias, yet that this was still a factor of 6 smaller than that from the direct relation in using a concrete example [91].

There are of course limitations to the inverse relation that hamper its practical use as a distance indicator. For example, the line width is often restricted by some upper and lower cut-offs. The distribution of the line width can also differ between the calibrator cluster and that of distant clusters, with a small change in the dispersion causing problems [92]. These concerns aside, this is the primary way the TFR is used, and is the method adopted in the compilation of the sample considered in the following section.

2.2 *Cosmicflows*-2 Malmquist treatment

The *Cosmicflows*-2 (CF2) and COMPOSITE catalogues deal with Malmquist distance bias in different ways. This is particularly evident in the way that Tully *et al.* [68] have incorporated the SFI++ sample. SFI++ consists of TFR derived distances for 4,861 field and cluster galaxies [76]. As the SFI++ data is presented with and without a global correction for all forms of Malmquist bias in the distances, it is an ideal candidate with which to study the systematic differences between the CF2 and COMPOSITE catalogues.

In publishing corrected distances Springob *et al.* [76] provide a word of caution. While the inhomogeneous and homogeneous distribution Malmquist bias are "straightforward" to correct (given access to a reconstruction of the local density field), Springob *et al.* [76] are limited in their treatment of selection bias. This is because the selection criteria used in the survey are very inhomogeneous. Thus an ad hoc selection criteria is constructed

that is designed to mimic the observational properties of the catalogue. Springob *et al.* state that their method, explained in detail in Springob *et al.* 2007 [76], is adopted out of necessity. The authors go on to advise that if possible one should adopt a different Malmquist bias-correction if a subsample of SFI++ is used for which a homogeneous selection function can be applied.

The COMPOSITE sample incorporates the corrected SFI++ distances whereas CF2 does not. However, in doing so Tully *et al.* [68] "make no adjustments for the distribution Malmquist effects" in their reported CF2 distances, arguing to account for the selection bias only through their implementation of the TFR. In particular, they use an inverse TFR procedure to reduce the selection bias only, stating that only a small subsequent correction for residual bias is required [68]. The calibration carried out for this relation follows the procedures of Tully and Pierce [93], Courtois and Tully [94] and Source *et al.* [95]. Thus, although apparently removing speculation about the treatment of selection bias by Springob *et al.* in SFI++, they have chosen to disregard the distribution Malmquist bias entirely. We are interested here in comparisons between subsets of the SFI++ and CF2 catalogues in order to understand the implications of the way this data is incorporated.

Tully *et al.* [68] 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

$$\Delta\mu_1 = 0.492(\pm 0.011) + 0.000031(\pm 0.000002)cz_{\rm LS} \tag{2.15}$$

where $\Delta\mu_1 \equiv \mu_{cf2} - \mu_{sfi}$, μ_{cf2} is the CF2 distance modulus with the zero point established by Courtois and Tully [94], μ_{sfi} is the Springob *et al.* [76] unadjusted modulus with a nominal zero point consistent with $H_0 = 100h$ km s^{-1}Mpc^{-1} and where $z_{\rm LS}$ is the raw[2] redshift in the rest frame of the Local Sheet, which is close to the Local Group frame [96]. We independently confirm the slope in (2.15) using the appropriate zero point[3], and plot this in Figure 2.2. Note that this comparison is 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

[2]The correction of Eqn. (13) in Tully *et al.* not applied [68].

[3]The intercept in (2.15) is determined by a scaling of the data so we are not interested in confirming this for our investigation.

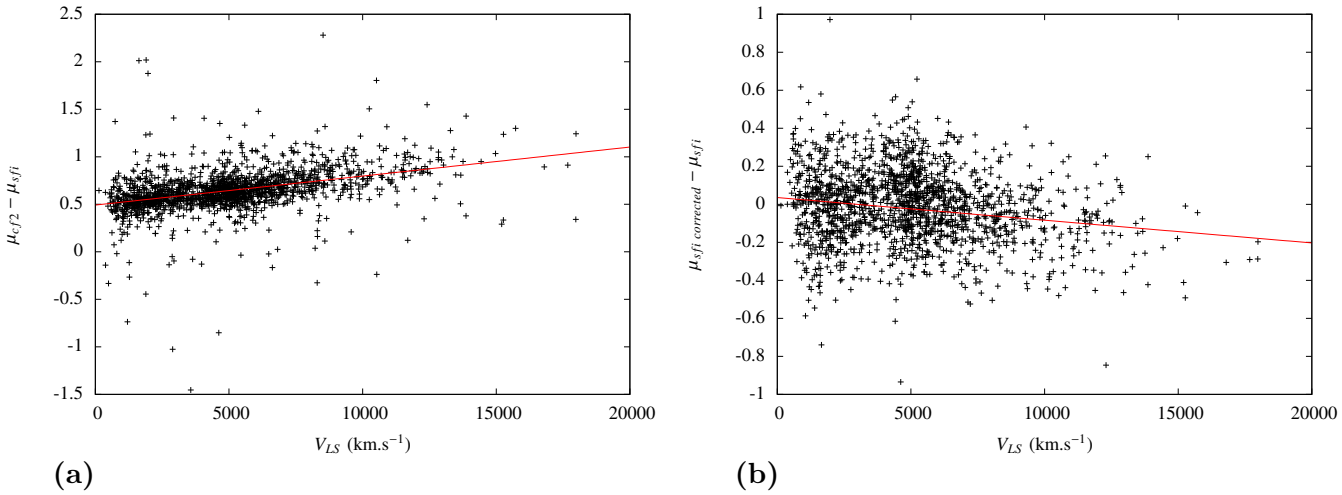


Figure 2.2: The difference in distance modulus between: (a) the CF2 and biased raw SFI++ distances for the points common to the two surveys, as presented in [68]. (b) the SFI++ biased and unbiased (corrected) catalogues for the same points. Where $V_{LS} = cz_{LS}$.

final analysis Tully *et al.* [68] use averages of the CF2 and SFI++ distances with double weight given to the CF2 distances.

The intercept in Figure 2.2 (a) simply reflects the different normalizations of H_0 in the two samples. The positive slope of the linear relation in Figure 2.2 (a) is consistent with the CF2 inverse TFR distance having a correction that accounts for the raw SFI++ distances being increasingly underestimated due to the Malmquist selection bias.

We are now interested in comparing (2.15) to the correction used by Springob *et al.* [76] in the SFI++ sample, and subsequently adopted in the COMPOSITE sample. Thus we repeat this type of analysis on the SFI++A subsample but now comparing the cases with and without all three Malmquist bias distance corrections, finding a linear relationship of the form

$$\Delta\mu_2 = 0.0356(\pm 0.0063) - 0.000012(\pm 0.000001)cz_{\rm LS} \tag{2.16}$$

where $\Delta\mu_2 \equiv \mu_{sfi,corrected} - \mu_{sfi}$, μ_{sfi} are the raw distance moduli and $\mu_{sfi,corrected}$ have been corrected for Malmquist bias by Springob *et al.* [76]. The data and best fit line are displayed in Figure 2.2. We can immediately see from (2.15) and (2.16) that there is a difference between the corrections. For small redshifts the correction is positive indicating raw distances are underestimated, while for large redshifts the

correction is negative indicating raw distances are overestimated. Thus adjusting the intercept of Figure 2.2 (a) to zero we would have in the limit of large redshift a hierarchy $\mu_{sfi,corrected} < \mu_{sfi} < \mu_{cf2}$. Thus this subset is consistent with the observation of Watkins and Feldman that "the distances are systematically larger in the *Cosmicflows*-2 catalogue [than in the COMPOSITE catalogue] due to a different approach to bias correction" [69].

The fact that 2.16 (b) has a negative slope is difficult to reconcile with the conventional understanding of Malmquist homogeneous and selection bias. As discussed in §2.1 we expect such biased distances to be underestimates. However, Feast [97] argue that the use of the TFR can result in *overestimates* for distances in particular cases. Feast shows that when the spatial density of objects at a given 21cm line width is constant, the required Malmquist correction is the classical one in (2.11). Yet he goes on to argue that when the this is not the case it is possible to obtain overestimated distances. Springob *et al.* [76] explain that due to "very inhomogeneous selection criteria" they followed an ad hoc construction to deal with the Malmquist bias in their sample. Thus the slope of the correction in Figure 2.16 (b) is certainly not inconsistent with that which includes the inhomogeneous distribution Malmquist bias.

The 1970 points in the SFI++ sample that are not also contained in the original CF2 survey, henceforth SFI++B, are not incorporated into CF2 using the correction (2.15). 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.* state that these points, if corrected using (2.15) cause a "highly significant decrease in the Hubble parameter with increasing velocity". We independently verified this result[4]. Thus Tully *et al.* 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 [98, 99]. 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

[4]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.* [73]. 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 Figures 2.3 and 2.4, as opposed to binning in distance, as will be carried out in §3.3.1 (Figure 3.1), so careful analysis is required to make sense of the different approaches. The direct calculation of the Hubble constant in each bin is also different to that described in §3.3.1, as it is not clear which method Tully *et al.* use we apply a simple weighted average of cz_i/r_i values and obtain consistent results (although as no Table of values is provided by Tully *et al.* we can only verify by inspection).

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 if they do show different characteristics. Considering $\Delta\mu_2$ for SFI++B we find a correction

$$\Delta\mu_2 = 0.0417(\pm 0.0061) - 0.000012(\pm 0.000001)cz_{\rm LS} \tag{2.17}$$

which has an intercept consistent within uncertainties and identical slope to (2.16). Since there is no apparent difference using this test we repeat a similar analysis to that performed by Tully *et al.* to test for bias (Figure 10 of [68]). This analysis is based on the fact that selection bias is manifest by an increase in the Hubble parameter with redshift [100], for data binned by redshift. In Figures 2.3 and 2.4 we repeat the same type of calculation of the Hubble constant in redshift bins performed by Tully *et al.* [68] for the subsets of interest to us and compare the results.

In Figure 2.3 we produce plots equivalent to Figure 10 in Tully *et al.* [68] 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 0.03σ to 1.8σ in individual bins. The weighted mean of these differences is 0.84σ, and thus we do not see a significant difference between SFI++A and SFI++B.

Unfortunately, this means we must question conclusions drawn using the the full CF2 catalogue, as we note that two mutually consistent halves of the SFI++ sample have been incorporated into CF2 in different ways.

On the other hand the COMPOSITE sample [75] uses a much larger subset of SFI++ distances corrected for Malmquist biases (after rejection of outliers). While it appears there are 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.* [76] included both corrected and raw distances, to allow others to take on the challenging task of Malmquist bias corrections. However, in doing so one must also account for the distribution homogeneous and inhomogeneous Malmquist bias, which was not done in the CF2 catalogue.

As another test of differences between the SFI++A and SFI++B subsamples we have also repeated the analysis of Fig. 2.3, but now to compare the raw and corrected distances within each subsample. Fig. 2.4 shows the comparison for the SFI++B subsample. In this case the difference in the Hubble parameters in each bin vary from a minimum

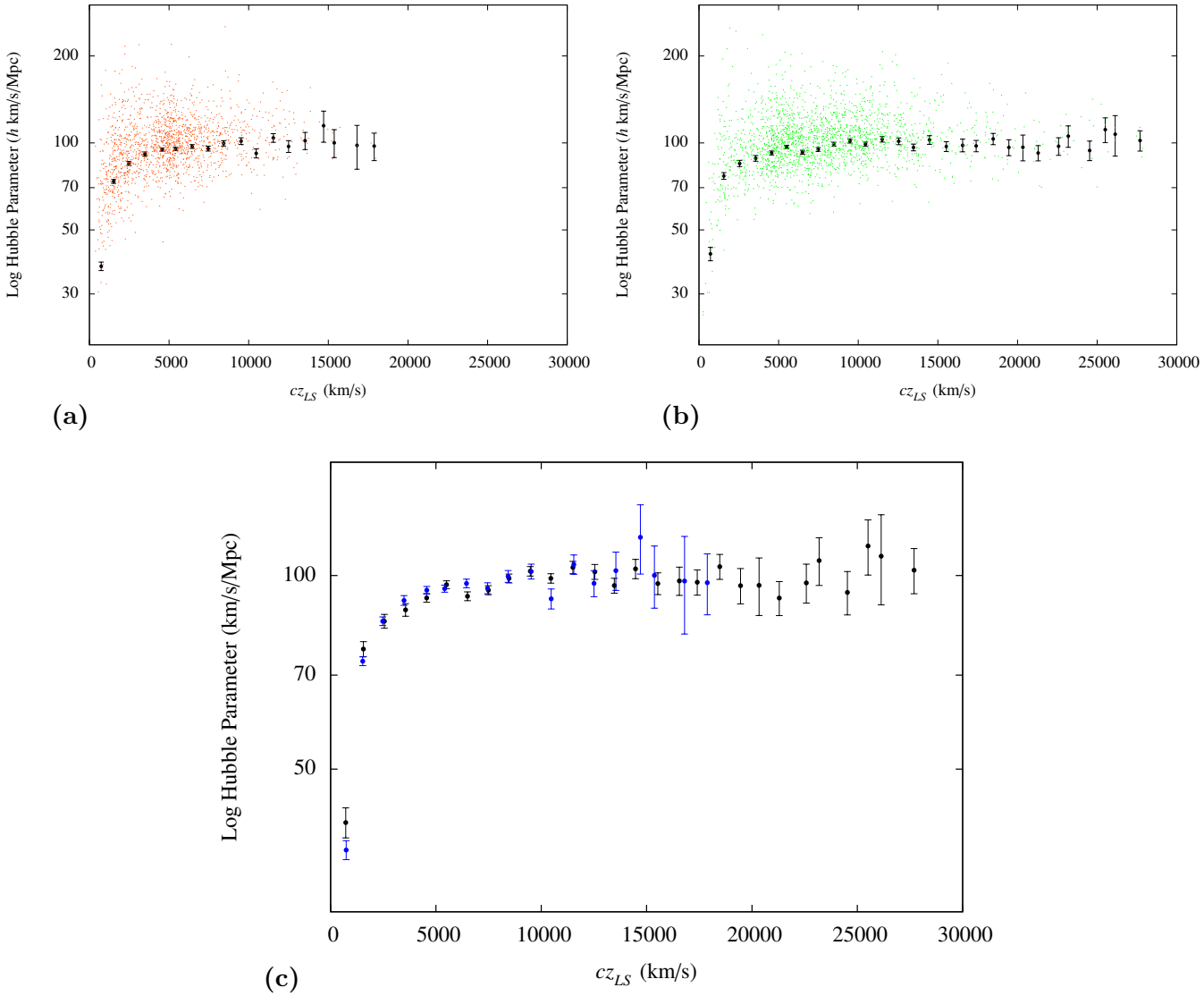


Figure 2.3: The Hubble parameter, $H_i = cz_i/r_i$, computed for each individual data point (coloured points) and from averaging in 1000 km s^{-1} bins (black points and error bars) using the (a): SFI++A subsample. (b): SFI++B subsample. (c) Comparison of the averaged points in (a) and (b) with blue points being from SFI++A and black points from SFI++B.

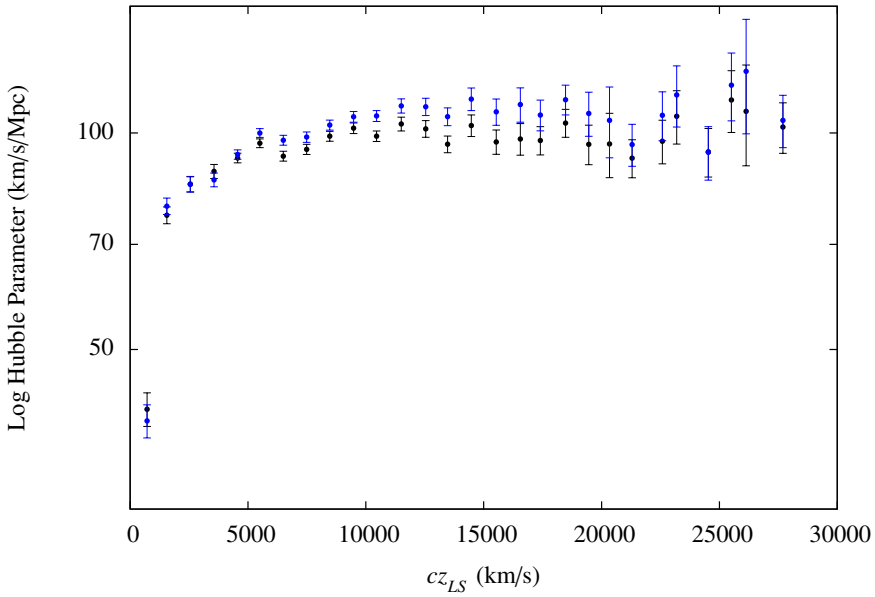


Figure 2.4: The Hubble parameter, $H_i = cz_i/r_i$, computed from averaging in 1000 km s^{-1} bins using the SFI++B subsample with corrections for Malmquist bias (blue crosses) and without corrections (black filled circles).

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, which reinforces our concerns about the way Tully *et al.* [68] have used this data.

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.* [68] 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 §4.2. Rather we follow Wiltshire *et al.* [73] 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.

CHAPTER 3

Hubble expansion variation

In a recent publication the assumptions of the standard cosmological rest frame have been challenged [73]. Wiltshire, Smale, Mattsson and Watkins (henceforth WSMW13) [73] show that without referring to peculiar velocities we can reduce the variation observed in the Hubble expansion on scales below the scale of statistical homogeneity (SSH). This is achieved by transforming the redshift data so as to analyse it from the perspective of an observer at rest with respect to the Local Group (LG) of galaxies, a rest frame boosted at 635±38 km s^{-1} with respect to the CMB rest frame [96]. In this chapter we review this work along with the earlier investigation by Li and Schwarz [101] (henceforth LS08).

3.1 Spherical averages

LS08 use a subset of 54 distances from the Hubble Telescope Key Project data [102] to study the effect of cosmological back reaction as a function of averaging scale[1]. The data analysed was restricted to the range $31.3 - 185.7$ Mpc ($22.5h^{-1} - 130\,h^{-1}$Mpc with $h = 0.72$). In order to study the effect averaging on different scales has, LS08 average in spherical shells consisting of data in a range $r_{\min} < r < r_{\max}$ where $r_{\min}$ remains fixed and $r_{\max}$ is increased for each subsequent shell. This technique is useful when data is limited, yet it does result in shells being correlated with all smaller shells contained within. An alternative to this method was used by WSMW13. In this case the

[1]Backreaction refers to the non-commutativity of taking averages and dynamical evolution, and the implications this has for the average expansion history of the universe.

much larger COMPOSITE sample was split into completely independent shells. That is, $r_{\min} = r_s$ and $r_{\max} = r_{s+1}$ are both varied from an inner cutoff of $r_{\min} = 2\,h^{-1}$Mpc up to $156.25\,h^{-1}$Mpc, in steps of $12.5\,h^{-1}$Mpc.

Both LS08 and WSMW13 calculate the mean distance and Hubble constant for each shell. LS08 only considered the CMB rest frame, whereas WSMW13 also considered the LG and Local Sheet (LS) frames. We will outline the details of the analysis by WSMW13 as this is the style we will follow in the next chapter.

Radial averages are computed in two different shell configurations. Both configurations use shells of radii $12.5\,h^{-1}$Mpc but start from a different inner shell cutoff of either $2\,h^{-1}$Mpc or $6.25\,h^{-1}$Mpc, where these shells are labeled using s and s' respectively. The 10 and $10'$ shells are larger than the rest so as to include approximately the same number of data points as the inner shells, and thus both have an outer cutoff at $156.25\,h^{-1}$Mpc. The outermost shell has no outer bound and is the same in both configurations.

The COMPOSITE sample distance uncertainties are large, approximately 15% for most sample objects. Fortunately, the large number of data points within each shell provide statistically meaningful results when averaged appropriately. We denote the individual distance uncertainties by σ_i, the distance to each object by r_i and the redshift by z_i. The weighted average distance for the s^{th} spherical shell is computed as

$$\bar{r}_s = \left(\sum_{i=1}^{N_s} \frac{r_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{N_s} \frac{1}{\sigma_i^2}\right)^{-1} \tag{3.1}$$

where N_s is the number of data points in the s^{th} shell. The Hubble constant in each shell is determined by standard linear regression, treating r and a function of z the quantity

$$\chi_s^2 = \sum_i \left[\sigma_i^{-1}(r_i - cz_i/H)\right]^2 \tag{3.2}$$

is minimized to give

$$H_s = \left(\sum_{i=1}^{N_s} \frac{(cz_i)^2}{\sigma_i^2}\right)\left(\sum_{i=1}^{N_s} \frac{cz_i r_i}{\sigma_i^2}\right)^{-1} \tag{3.3}$$

where the values of z_i are taken to be exact. This method is chosen because all uncertainties in the COMPOSITE sample have been included as distance uncertainties. Any corrections that would be required because of noise arising from peculiar motion within galaxy clusters have been accounted for in this sample by assigning a distance and uncertainty to the cluster itself rather than individual galaxies [74].

The uncertainty in the Hubble constant for each shell is determined by the individual distance uncertainties added in quadrature to a zero point uncertainty. This zero point uncertainty arises from the fact that a linear Hubble law must pass through the origin, yet the position of such an origin is not itself exact. This arises from a 20 km s^{-1} uncertainty in the heliocentric peculiar velocity of the LG (and LS) [96] and a 0.4% uncertainty in the magnitude of the CMB dipole [3]. This zero point uncertainty is given by WSMW13 as $\sigma_0 = 0.201h^{-1}$ Mpc. For each shell we compute the weighted zero point uncertainty as

$$\bar{\sigma}_{0s} = H_s \frac{\sigma_0}{\bar{r}_s} \tag{3.4}$$

which can be understood as an uncertainty in the mean distance for each shell. This uncertainty is significant for shells with a small mean distance compared to those at a larger distance. This is effectively due to the longer lever arm of the Hubble law for shells with large mean distances. Next we add (3.4) in quadrature to the uncertainty from the individual galaxy distances,

$$\bar{\sigma}_{1s} = \left(\sum_{i=1}^{N_s} \frac{(cz_i)^2}{\sigma_i^2} \right)^{3/2} \left(\sum_{i=1}^{N_s} \frac{cz_i r_i}{\sigma_i^2} \right)^{-1}, \tag{3.5}$$

to give the total uncertainty $\bar{\sigma}_s$ of H_s. We will follow the notation of WSMW13 and adopt the convention that all quantities that are averaged (angularly or radially) will be donated with an over-bar to distinguish these from values of individual data points.

In addition to the expressions in WSMW13 it will be useful in our investigation to note the uncertainty in the weighted average of the squared distance of each data point, which we calculate to be

$$\bar{\sigma}_{\langle r_i^2 \rangle_s} = 2 \left(\sum_{i=1}^{N_s} \frac{r_i^2}{\sigma_i^2} \right)^{1/2} \left(\sum_{i=1}^{N_s} \frac{1}{\sigma_i^2} \right)^{-1}, \tag{3.6}$$

through standard error propagation.

Finally we must define a measure of variation in the local Hubble expansion from the asymptotic value. LS08 and WSMW13 use the quantity

$$\delta H_s = \frac{\left(H_s - \bar{H}_0\right)}{\bar{H}_0} \tag{3.7}$$

where $\bar{H}_0$ is the mean asymptotic value of the Hubble constant. This asymptotic value and its uncertainty are calculated from the data points beyond $r = 156.25\,h^{-1}$Mpc

(shell 11 for both configurations). As this distance is larger than the BAO scale, which is the largest scale on which we should expect to see the effects of inhomogeneity on the local Hubble expansion [73], it is appropriate to calculate H_0 from these data points. The result, $\bar{H}_0 = (100.1 \pm 1.7)h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the CMB frame and $\bar{H}_0 = (101.0 \pm 1.7)h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the LG/LS frames, is consistent with the choice of $H_0 = 100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ which is used in the COMPOSITE sample to convert velocity uncertainties (as is standard in the peculiar velocity framework) into the distance uncertainties used here.

3.2 Angular averages

As the distribution of foreground structure is certainly not isotropic we expect the effect on the local Hubble expansion will not be either. In order to study this effect we must consider angular variation through appropriate averaging techniques. McClure and Dyer [103] (henceforth MD07) analyzed the HST Key data [102] in the CMB reference frame and found a 13% angular variation in the Hubble constant. WSMW13 use the following techniques, similar to those of MD07, to produce angular averages of the Hubble expansion.

3.2.1 Gaussian window averages

The COMPOSITE sample can only provide information on the Hubble expansion along specific lines of sight. In order to quantify angular Hubble expansion variation we assign a weighted Hubble constant at every point on the sky, on a spherical grid up to the desired resolution. The technique used by MD07, known as Gaussian window averaging, is used to compute a mean H_α by weighting each value of cz_i/r_i by its angular separation from the point of interest, where α represents the angular coordinates of this point.

MD07 use a Gaussian weighting to smear the data. The weights are given by

$$W_{i\alpha} = \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp\left(\frac{-\theta_i^2}{2\sigma_\theta^2}\right) \tag{3.8}$$

where σ_θ determines the smoothing scale and θ_i is the angular separation from the point of interest on the grid to the relevant data point, such that $\cos\theta_i = \vec{r}_{\mathrm{grid}} \cdot \vec{r}_i$.

The lower bound for the smoothing scale is determined by the angular width of the Zone of Avoidance (ZoA), the region of the sky obstructed by our own galaxy.

The COMPOSITE sample lacks data across approximately 30° centred on the plane of the galaxy due to this ZoA. Thus the smoothing scale must be sufficiently large that grid points within this zone still provide reliable results. Therefore the diameter of the Gaussian window function, $2\sigma_\theta$, must not be less than the width of the ZoA. Therefore we use a smoothing scale no less than 15°.

On the other hand if the smoothing scale is too large angular resolution will be lost. For this reason both MD07 and WSMW13 use a value of $\sigma_\theta = 25°$. This smoothing angle subtends an area of 0.59 steradians which corresponds to 4.8% of the entire sky. WSMW13 have also explored the effect of varying the smoothing scale between 15° and 40° finding no significant changes to the results [73]. We will therefore use the same smoothing scale in our investigation.

MD07 and WSMW13 approach the calculation of H_α in two different ways. For MD07 the individual distance uncertainties in the HST Key data set are relatively small compared to those in the COMPOSITE sample. Therefore the results of MD07 are not grossly dependent on how the uncertainties are treated. In contrast WSMW13 show that the COMPOSITE sample is far more sensitive to the treatment of uncertainties.

Since the uncertainties in the COMPOSITE sample are predominately in the distance measurements it is preferable to calculate the inverse Hubble constant at each spherical polar grid point. This average is given by

$$H_\alpha^{-1} = \frac{\sum_{i=1}^{N} W_{i\alpha} r_i (cz_i)^{-1}}{\sum_{j=1}^{N} W_{j\alpha}} \tag{3.9}$$

with a variance

$$\bar{\sigma}^2_{H_\alpha^{-1}} = \frac{\sum_{i=1}^{N} W_{i\alpha}^2 \sigma^2_{H_i^{-1}}}{\left(\sum_{j=1}^{N} W_{j\alpha}\right)^2} \tag{3.10}$$

where

$$\sigma_{H_i^{-1}} = \frac{\sigma_i}{cz_i} \tag{3.11}$$

is the uncertainty in the inverse Hubble constant at each data point. From standard error propagation we find that

$$\bar{\sigma}_\alpha = \bar{\sigma}_{H_\alpha^{-1}} H_\alpha^2 \tag{3.12}$$

is the uncertainty in the weighted Hubble constant at α.

The final technical detail is in the choice of the weights. The weights given by (3.8) are non-linear, and are used in what would otherwise be a linear regression. For this reason WSMW13 introduce an inverse variance (IV) weighting. These weights are given by

$$W_{i\alpha} = \frac{1}{\sigma^2_{H_i^{-1}}\sqrt{2\pi}\sigma_\theta} \exp\left(\frac{-\theta_i^2}{2\sigma_\theta^2}\right) \tag{3.13}$$

which depend on a variance which is linear in the measurement uncertainties. WSMW13 show that using this weighting when calculating H_α by its inverse gives results close to that given with the original weighting. However, they observe different results if IV weightings are used for the alternative and less suitable method of linear regression to give H_α directly. We will make use of (3.13) when calculating angular variation statistics in the following chapter.

3.3 The Local Group as the frame of minimum Hubble expansion variation

3.3.1 Monopole variation

Through the spherical averaging techniques outlined in §3.1 WSMW13 compared monopole Hubble expansion variation in the CMB, Local Group and Local Sheet frames. The variation from a linear Hubble law can be measured by considering the mean square differences from a uniform expectation ($\delta H = 0$). Therefore we are interested in the statistic

$$\chi^2(r_s) = \sum_{i=s}^{12} \frac{(\delta H_j - 0)^2}{\sigma^2_{\delta H_j}} \tag{3.14}$$

where $\sigma_{\delta H_j}$ is the variance in δH_j which we calculate as

$$\sigma^2_{\delta H_j} = \left(\frac{\partial}{\partial \bar{H}_0}\delta H_j\right)^2 \bar{\sigma}^2_{\bar{H}_0} + \left(\frac{\partial}{\partial H_j}\delta H_j\right)^2 \bar{\sigma}^2_{H_j} \tag{3.15}$$

$$= \left(-\frac{H_j}{\bar{H}_0^2}\right)^2 \bar{\sigma}^2_{\bar{H}_0} + \frac{1}{\bar{H}_0^2}\bar{\sigma}_{H_j^2} \tag{3.16}$$

which gives the explicit form

$$\chi^2(r_s) = \sum_{i=s}^{12} \frac{\bar{H}_0^4 \delta H_i^2}{\bar{H}_0^2\sigma^2_{H_i} + H_i^2\sigma^2_{H_0}}. \tag{3.17}$$

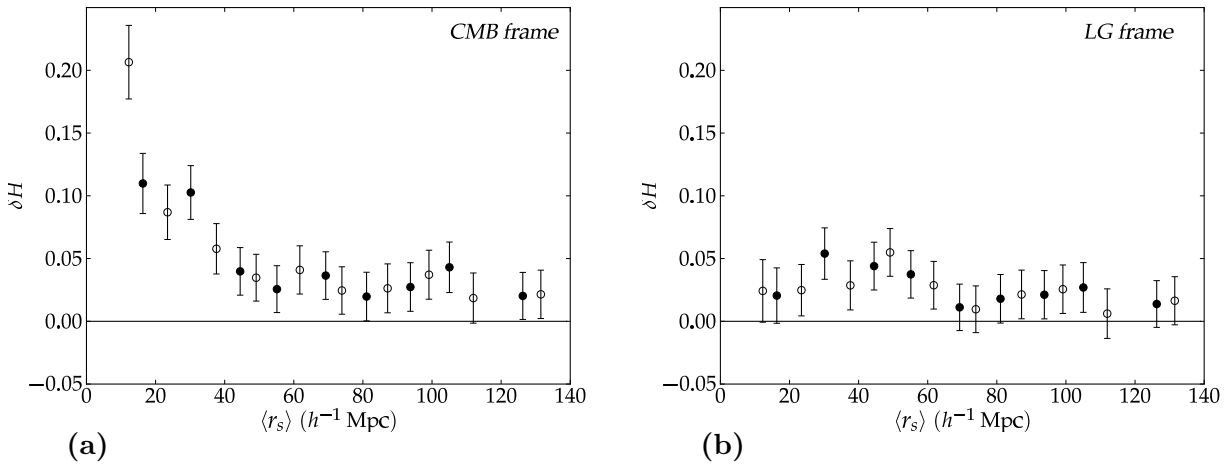


Figure 3.1: Variation in the Hubble expansion $\delta H_s = (H_s - \bar{H}_0)/\bar{H}_0$ in spherical shells as a function of weighted mean shell distance: (a) CMB frame, (b) LG frame. In each case the filled data points represent the first choice of shells, and the unfilled circles the alternative choice of shells. The data point for the first shell has been omitted from the plots since δH is so large in the CMB frame that it is off-scale. For this omitted shell, with a mean weighted distance of $\langle r_s \rangle = 5.43\, h^{-1}$Mpc, we have $\delta H_{\rm CMB} = 0.737 \pm 0.029$ and $\delta H_{\rm LG} = 0.168 \pm 0.007$. Figure obtained from WSMW13 [73].

where r_s is the inner cutoff distance. This χ^2 is calculated using the averaged Hubble constants in all shells with $r > r_s$. Using this statistic WSMW13 calculate the probability of a uniform Hubble expansion for each rest frame and choice of r_s directly from the χ^2 probability distribution

$$P(\chi^2, \nu) = \frac{\Gamma(\nu, \chi^2/2)}{\Gamma(\nu)} = \int_{\chi^2/2}^{\infty} t^{\nu-1} e^{-t}\, \mathrm{d}t \Big/ \int_0^{\infty} t^{\nu-1} e^{-t} \mathrm{d}t \tag{3.18}$$

which is computed numerically. The values of $P_{\rm LG}(r_s)$, $P_{\rm CMB}(r_s)$ and the Bayes factor $B(r_s) = P_{\rm LG}(r_s)/P_{\rm CMB}(r_s)$ are computed in order to compare the relative uniformity of the Hubble expansion.

WSMW13 consider the Hubble expansion in the Local Group, Local Sheet and CMB frames. Figure 3.1 from WSMW13 [73] shows (3.7) in each shell for the CMB and Local Group frames. It is immediately clear that the Hubble expansion is far more uniform in the LG frame. Using (3.17) as described above WSMW13 [73] show with decisive Baysian evidence that the Hubble expansion is more uniform when analysed in the LG frame compared to the CMB.

3.3.2 Angular variation

It is found that when considered from the Local Group (and Local Sheet) the variations in the Hubble expansion at distances greater than $15\,h^{-1}$Mpc exhibit a dipole nature. In fact, there is a strong correlation between the fraction of the Hubble expansion variation dipole and the CMB dipole anisotropy which is usually considered a result of the motion of the LG.

To compare the CMB temperature and Hubble expansion variation dipoles an artificial residual CMB temperature map is produced in each frame of reference. This is achieved by subtracting a Lorentz boosted CMB sky with temperature

$$T' = \frac{T_0}{\gamma\left(1 - (v/c)\cos(\theta)\right)} \tag{3.19}$$

from the corresponding observed pure temperature monopole plus dipole maps using the values of Fixsen *et al.* [104].

After digitizing the Hubble expansion variation sky map and the CMB residual temperature map we can use the Pearson correlation coefficient to give a measure of the correlation between these maps. This correlation coefficient is given by

$$\rho_{\nu T} = \frac{\sqrt{N_p}\sum_\alpha \bar{\sigma}_\alpha^{-2}\left(H_\alpha - \bar{H}\right)\left(T_\alpha - \bar{T}\right)}{\sqrt{\left[\sum_\alpha \bar{\sigma}_\alpha^{-2}\right]\left[\sum_\alpha \bar{\sigma}_\alpha^{-2}\left(H_\alpha - \bar{H}\right)^2\right]\left[\sum_\alpha \left(T_\alpha - \bar{T}\right)^2\right]}} \tag{3.20}$$

where α denotes the angular coordinates of each pixel in the digitized sky maps, T_α is the temperature at α and $\bar{T}$ is the mean temperature. The Hubble parameter is calculated using a weighted average at each point on the sky map as given in §3.2.1.

WSMW13 calculate the Pearson correlation coefficient, (3.20), in the LG frame to be $\rho_{\nu T} = -0.9240$ for a $r > 15\,h^{-1}$Mpc sphere, with IV weighting and a smoothing scale of $\sigma_\theta = 25^\circ$ [73] (see Figure 3.2). Similar results are obtained for the LS and different choices of the smoothing scale, but not in the CMB frame. A further analysis was undertaken by WSMW13 to compare the strength of the Hubble expansion variation dipole in the CMB and LG frames, using a least squares fit of a dipole law in the same radial shells used to investigate the spherical (monopole) variations. These results led WSMW13 to the conclusion that "the boost from the LG to the CMB frame is compensating for structures in the range $30h^{-1} \lesssim r \lesssim 62\,h^{-1}$Mpc" [73].

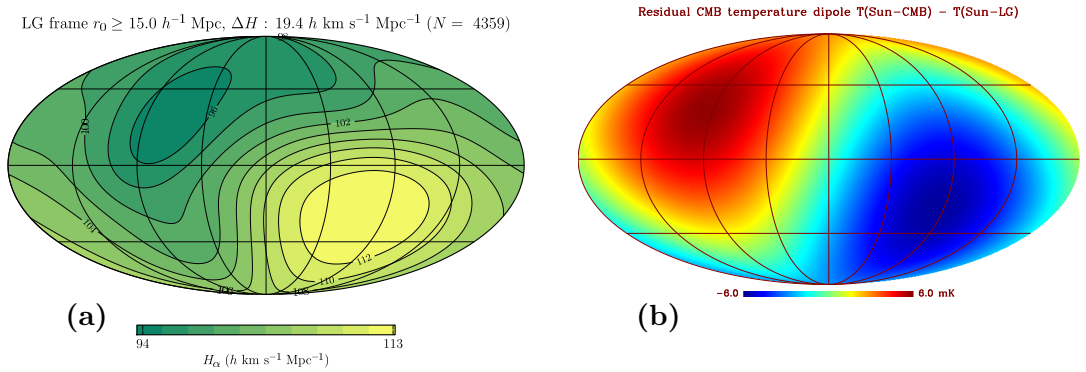


Figure 3.2: LG frame Hubble expansion variation map for $r > 15\,h^{-1}$Mpc with IV weightings (panel(a)) compared to residual CMB temperature dipole in the LG rest frame (panel (b)). In all figures, the galactic longitudes $\ell = 0°, 180°, 360°$ are on the right edge, centre and left edge respectively.

3.4 The cosmic microwave background radiation

In challenging the standard of cosmic rest one must address the origin of the CMB dipole anisotropy. As discussed in the introduction, the dipole anisotropy in the CMB has become the de facto reference for defining cosmic rest. It has been claimed for over 40 years that this dipole is the result of Lorentz boosting of the CMB due to our motion. If we are to argue for a standard of rest in which a dipole anisotropy exists we must simultaneously seek a non-kinematic origin of this dipole.

The CMB dipole as a result of special relativity is straightforward to derive. Consider an observer at rest in an isotropic background of radiation and an observer boosted at a velocity v. The stationary observer sees photons of momentum $|\boldsymbol{p}|$ and by a standard Lorentz boost the other observer sees photons with momentum

$$|\boldsymbol{p}| = \gamma\,(1 + \beta\cos\theta)\,|\boldsymbol{p'}| \tag{3.21}$$

where $\beta = v/c$ and θ depends on the angle between the photon momenta $\boldsymbol{p}$ and the velocity of the Earth, and $\gamma = \left(1-\beta^2\right)^{-1/2}$. Next consider the photon density $N_\gamma(\boldsymbol{p})$ for the stationary observer and $N'_\gamma(\boldsymbol{p'})$ for the boosted observer. This photon density is a Lorentz invariant scalar given by

$$N_\gamma(\boldsymbol{p}) = \frac{1}{h^3}\frac{1}{\exp(|\boldsymbol{p}|c/kT_0) - 1} \tag{3.22}$$

and thus Lorentz invariance of the number of photons in a given momentum space volume demands

$$\frac{1}{\exp(|\boldsymbol{p}|c/kT_0)-1} = \frac{1}{\exp(\gamma\,(1+\beta\cos\theta)\,|\boldsymbol{p'}|c/kT')-1} \tag{3.23}$$

and thus (3.19) is obtained, where $T_0 = 2.7255 \pm 0.0006$ K is the CMB mean temperature [105]. Eqn. (3.19) can be expanded in powers of β to give

$$\Delta T = T' - T = T\left[-\frac{\beta^2}{6} - \beta P_1(\cos\theta) + \frac{2\beta^2}{3}P_2(\cos\theta) + \mathcal{O}(\beta^3)\right] \tag{3.24}$$

where $P_1(x) = x$ and $P_2(x) = \frac{1}{2}\left(3x^2-1\right)$ are the 1st and 2nd order Legendre polynomials which represent a dipole and quadrapole respectively.

The frame in which no dipole of the form in (3.19) is observed is defined as the CMB frame. Based on this analysis the LG of galaxies is claimed to be moving at a velocity of 635 ± 38 km s^{-1} relative to the CMB frame [96]. Studies of the CMB anisotropies commonly assume that the dipole is due to precisely the mechanism derived here, and subsequently a kinematic dipole is subtracted from the data. However, as can be seen in (3.24) there are subtle effects associated with a kinematic dipole which are not guaranteed to match the physical reality. Thus in assuming a kinematic dipole problems can arise in the resultant temperature map when adjusted to the CMB frame.

The subtle inconsistencies in the CMB temperature map, arising after the subtraction of a kinematic dipole and the foreground effects of our own galaxy, are known as anomalies. In the PLANCK 2013 results [5] [106] anomalies that were previously found in the WMAP data [107] are confirmed once again at levels of significance of about 3σ. These anomalies are associated with large angle multipoles in the CMB temperature anisotropy spectrum. Some of these anomalies are

(1) The power asymmetry between northern and southern hemispheres [108–111],

(2) The low quadrupole power [108, 112],

(3) The alignment of the quadrupole and octupole [112–115] and

(4) The parity asymmetry [116].

3.4.1 CMB dipole as the result of foreground structure

A non-kinematic origin for the CMB dipole is based on consideration of local structure and differential expansion. This differential expansion can cause differences in the distance to the surface of last scattering, and thus the temperature of the observed photons that have travelled for almost the age of the universe to reach us.

Wiltshire *et al.* [73] describe how inhomogeneous foreground structure can produce a dipole like anisotropy in the CMB. On scales below $100\,h^{-1}$Mpc the universe is sufficiently inhomogeneous that differential expansion can be expected to give differences in the distance to the surface of last scattering. Although photons will see similar inhomogeneity throughout their entire journey to us, the effect on large scales is averaged out. On scales below that of statistical homogeneity this influence on photon paths will not be averaged out. The magnitude of the required difference in comoving distance to the surface of last scattering across the sky is calculated by Wiltshire *et al.* to be $\delta D = \mp(0.33 \pm 0.02)\,h^{-1}$Mpc for a temperature dipole of $\delta T/T \approx 10^{-3}$ in a spatially flat ΛCDM model with $\Omega_{M0} = 0.25$, $\Omega_{R0} = 4.15h^{-2} \times 10^{-5}$ and $h = 0.72$. The ΛCDM model is used phenomenologically for distance estimates; a similar result is obtained using the timescape model [50].

If the structure responsible for the dipole anisotropy was within $65\,h^{-1}$Mpc then the required difference in distance between opposite points on the sky would be less than 0.5%, which is entirely plausible given the amount of angular Hubble expansion variation seen by Wiltshire *et al.* [73]. In fact, there is evidence that foreground structure within $65\,h^{-1}$Mpc is a viable explanation for the observed CMB dipole in the Local Group [73]. This is further extended by our investigation in chapter 4.

The fact that anomalies arise *after* the subtraction of a kinematic dipole demonstrate that our proposition for a non-kinematic dipole is certainly not in contradiction with current understanding. The PLANCK collaboration have recently claimed to observe Doppler boosting of the CMB [5]. By measuring aberration and modulation in the CMB temperature fluctuations PLANCK find a multipole direction consistent with the CMB dipole for small angular scales. However, on large angular scales the direction of the dipole shifts towards the modulation dipole anomaly direction found for WMAP [107] and confirmed by PLANK [106]. These larger angular scales are comparable with the projected angular size of inhomogeneities on scales $< 65\,h^{-1}$Mpc. Thus this does *not* confirm that the observed dipole is due Doppler boosting, in fact this scale dependence of aberration and modulation many be a signature of a non-kinematic dipole.

CHAPTER 4

In search of a cosmological rest frame

4.1 Introduction

In the previous chapter we explored the evidence that the spherically averaged Hubble expansion is more uniform when considered in the Local Group frame compared to the CMB frame. The evidence is certainly very decisive and points to the Local Group as a likely candidate for the standard of rest. However, such a result is not complete unless we can show that this is the only desirable frame of reference for uniformity of Hubble expansion. Before any definition of standard of rest is made we must systematically rule out all other possibilities.

While minimizing the variation in the Hubble expansion is the primary aim in defining a cosmological rest frame it is not the only factor considered. Depending on the choice of analysis we use and due to our finite sample of data defining this frame will certainly be subject to uncertainties and possible ambiguities.

In our investigation we focus predominately on the systematic boost offset observed between the spherically averaged Hubble law in the CMB and Local Group frames and how this property can be used to constrain the standard of rest. We will also attempt to compute a minimum Hubble expansion variation frame for radial averages, and carry out an investigation into angular variation. The spherical averaging techniques are identical to that of WSMW13, as outlined in the previous chapter, unless otherwise stated. In

chapter 5 we will repeat key parts of this analysis with the *Cosmicflows*-2 sample of Tully *et al.* [68].

4.2 Hubble expansion variation as a systematic boost offset

WSMW13 [73] propose that monopole variation from a linear Hubble law may have a systematic origin. This arises through the non-linear dependence of H_s in (3.3) on the individual cz_i values. By calculating the result of an arbitrary boost on the individual H_s values WSMW13 obtain an explicit form for this systematic variation.

Consider redshifts taken with respect to a frame of reference in which the spherically averaged Hubble expansion variation is minimized. Then make an arbitrary boost such that redshifts transform as

$$cz_i \rightarrow cz_i' = cz_i + v\cos\phi_i \tag{4.1}$$

where v is the boost magnitude and ϕ_i is the angle between the data point and the boost direction. In (3.3) this results in the changes $(cz_i)^2 \rightarrow (cz_i')^2 = (cz_i)^2 + 2cz_i v\cos\phi_i + v^2\cos^2\phi_i$ in the numerator and $cz_i r_i \rightarrow cz_i r_i + r_i v\cos\phi_i$ in the denominator.

The linear contributions to the transformed quantities in the denominator and numerator of (3.3) should be approximately self cancelling. When taking a spherically symmetric average on data which is distributed uniformly over the celestial sphere then on average each positive contribution from the term $v\cos\phi$ 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 in data is symmetrical on opposite sides of the sky. This assumption would only be invalid when one side of the sky has a significant lack of data, which is not seen on the opposite side. The COMPOSITE sample does indeed have sufficient sky coverage to satisfy this requirement [73]. With such a cancellation assumed we are left with the difference

$$\Delta H_s = H_s' - H_s \approx \frac{v^2}{2\bar{H}_0 < r_i^2 >_s} \tag{4.2}$$

where $< r_i^2 > \equiv \left(\sum_{i=1}^{N_s} \frac{r_i^2}{\sigma_i^2}\right)\left(\sum_{i=1}^{N_s} \frac{1}{\sigma_i^2}\right)^{-1}$ is a weighted average in each shell and $\bar{H}_0$ is the asymptotic value of the Hubble constant.

We present a verification 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 (4.2). To achieve this we must fit a power law to the observed data. As there is a correlated uncertainty in both the independent and dependent variables, $\langle r_i^2 \rangle_s$ and ΔH_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

$$\Delta H = A \langle r_i^2 \rangle^p . \tag{4.3}$$

In comparison with (4.2), we expect $p \approx -1$ and $A \approx v^2/(2\bar{H}_0)$. 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 [73] – 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 – $6.35\,h^{-1}$Mpc, while keeping the widths of the shells fixed. This leads to a value of $p = -0.88 \pm (0.25)_{\rm stat} \pm (0.13)_{\rm sys}$ 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 $v = 646 \pm 545\,\mathrm{km\,s^{-1}}$, which is indeed close to the actual boost magnitude of $635 \pm 38\,\mathrm{km\,s^{-1}}$, 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 – $9.375\,h^{-1}$Mpc, we arrive at a value of $p = -0.87 \pm (0.33)_{\rm stat} \pm (0.09)_{\rm sys}$. Fig. 4.1(b) shows the resultant best fit curves.

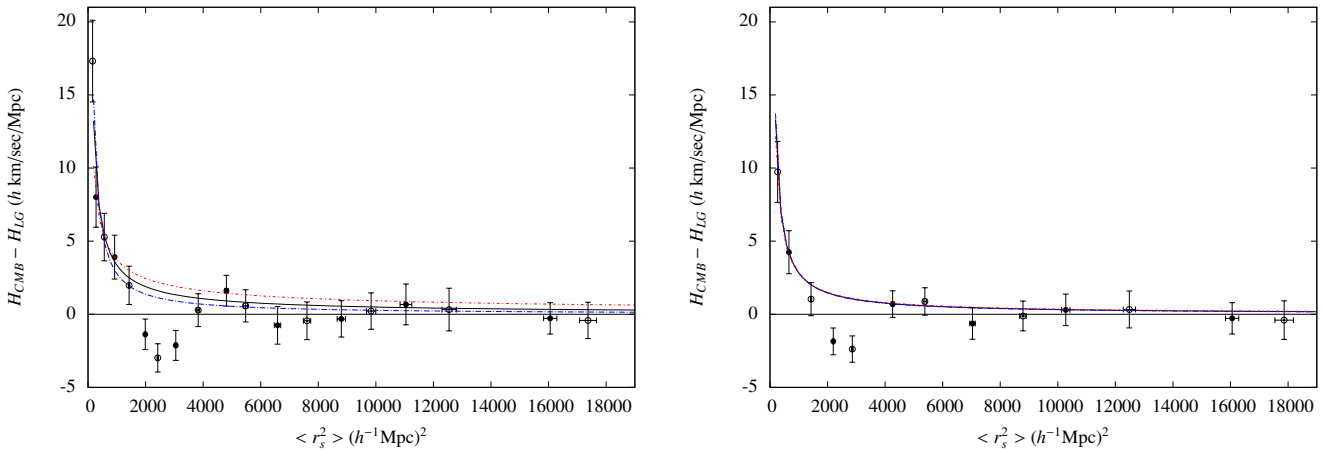


Figure 4.1: Best fit power law to the radial variation in the spherically averaged Hubble law in 2 configurations of: **(a)** 11 shells; **(b)** 8 shells. The dashed blue (lower) curve indicates the best fit to primed shells only (empty circles), the dotted red (upper) curve indicates the best fit to unprimed shells only (filled circles). The solid black curve is the combined best fit using both primed and unprimed shells. The first data point – corresponding to unprimed shell 1 – is omitted in each plot, as it is off the scale.

In Fig. 4.1 we note a discrepancy between the best fit power law and the negative values on shells in the range of 40 – 60 h^{-1}Mpc (or $\langle r_s^2 \rangle = 1600$–$3600(\,h^{-1}\text{Mpc})^2$). 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. As WSMW13 show there is evidence for structure in this range being responsible for a dipole seen when taking angular averages of the Hubble expansion, with this dipole being more pronounced in the Local Group frame.

The boost to the CMB frame almost compensates for the dipole variation. One finds that in the range $40h^{-1} \lesssim r \lesssim 60\,h^{-1}$Mpc (and only in this range) for monopole variation $(\delta H_s)_{\text{CMB}} < (\delta H_S)_{\text{LG}}$, while the Hubble variation dipole in the CMB frame becomes consistent with zero. If the boost to the CMB frame exactly compensated for structures in the range $40h^{-1} - 60\,h^{-1}$Mpc then this should remain true at larger distances. However the magnitude of the CMB frame dipole increases to a residual offset showing that the dipole almost – but not entirely – has the character of a Lorentz boost dipole. Given that there are nonlinear structures we cannot obtain a perfect power law fit in Figure 4.1. However, the deviation from a power law is consistent with the observation

that $(H_s)_{\rm CMB} < (H_s)_{\rm LG}$ in the range over which the boost almost compensates for the nonlinear structures.

Now that we have verified the power law nature of the difference $H_{CMB} - H_{LG}$ we must check whether this result is unique for the boost to the Local Group. To investigate this we determine the Hubble constant in radial shells for many other frames boosted with respect to the CMB, denoted by frame "x", and then fit (4.3) to the resulting $\Delta H = H_{CMB} - H_x$ curve. We vary the direction of the boost to frame x while holding the magnitude constant, thus producing a sky map. We choose a magnitude of 635 km s^{-1} corresponding to the boost to the Local Group 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 $\Delta H_s = A\left\langle r_i^2\right\rangle^p$ with $p = -1$ and $A > 0$, representing the best fit boost offset. If one now boosts in the *opposite* direction by $635\,\mathrm{km\,s^{-1}}$ then one finds a best fit power law with $\Delta H_s = A\left\langle r_i^2\right\rangle^p$ 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 (4.3) 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. 4.2 we plot[1]

$$f_p = \begin{cases} |p+1|, & A \geq 0 \\ 2 - |p+1|, & A < 0 \end{cases} \tag{4.4}$$

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. 4.2 shows that for a boost magnitude of $635\,\mathrm{km\,s^{-1}}$ 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 $635\,\mathrm{km\,s^{-1}}$ relative to the CMB has the mathematical character of a systematic boost offset for directions close to that of the LG. In Fig. 4.2 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).

[1]Both primed and unprimed shells are used (in the 11 shell case), to produce a smoothed sky map without determining systematic uncertainties.

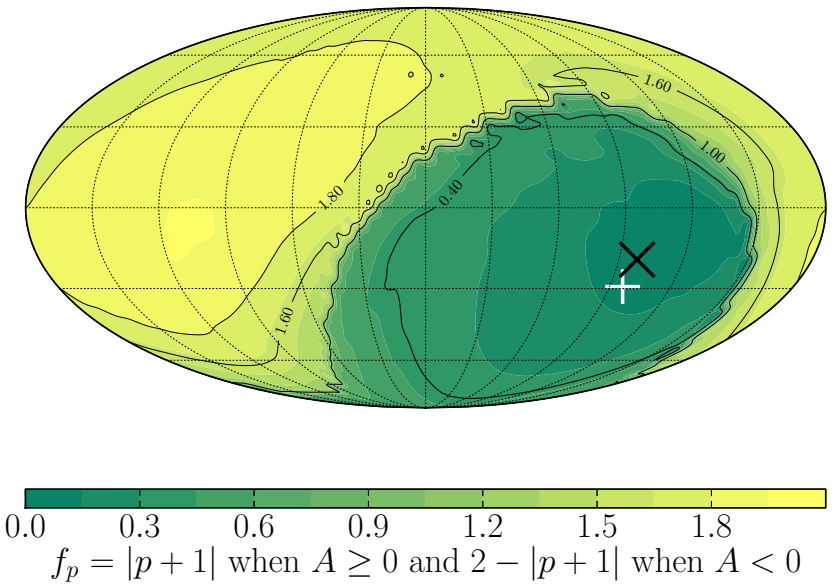


Figure 4.2: he best fit parameters to a systematic boost offset (4.3) for frames boosted from the CMB frame at $635\,\mathrm{km\,s^{-1}}$. The black cross denotes the boost to the LG frame and the white cross denotes the boost to the frame with minimum variation in the spherically averaged Hubble law for a boost of this magnitude, which will be discussed in §4.3.1. In all figures, the galactic longitudes $\ell = 0°, 180°, 360°$ are on the right edge, centre and left edge respectively.

We have verified that there is indeed a systematic boost offset in the spherically averaged Hubble expansion between the CMB and LG frames. We have also checked that this is not the result of random effects by studying the same property for all possible boosts of the same magnitude across the sky. We see the existence of the correct power law in the ΔH data is consistent with a boost in the direction of the Local Group. Therefore the existence of a systematic boost offset has given an independent verification in the search for a frame of minimum Hubble expansion variation.

However, we have as yet not varied the amplitude of the boost. Given the uncertainties we have already noted, it is clear that the systematic boost offset will become very hard to confirm statistically if the boost amplitude 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 "minimum Hubble expansion variation frame". There are in fact many degenerate frames. We will investigate varying the boost amplitude in the next section.

4.3 Minimizing the average spherical Hubble expansion

4.3.1 Variations in the "nonlinear regime"

Defining the frame of reference which gives the minimum monopole variation in the Hubble expansion is the first consideration in a search for a cosmological standard of rest. Furthermore any final definition of a standard of rest should be consistent within one standard deviation of this minimum.

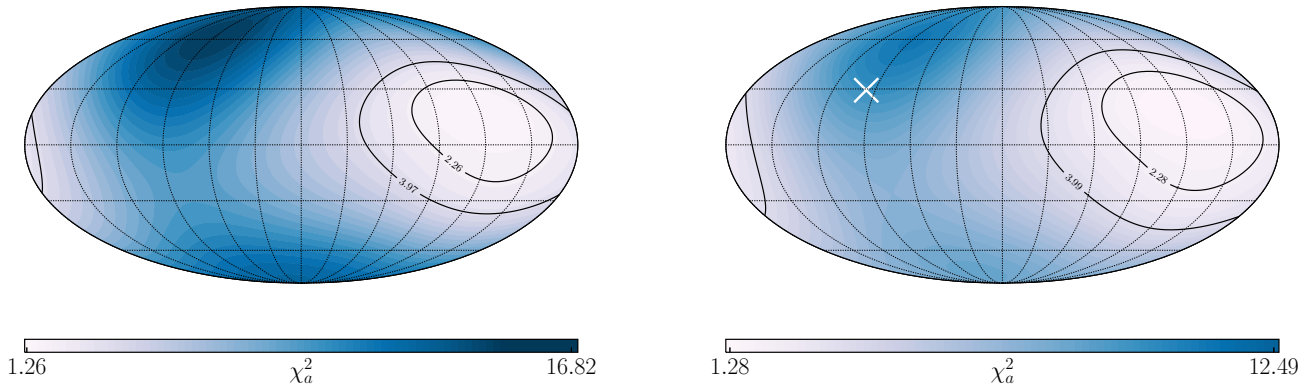
Initially we treat monopole variation with respect to a $\delta H = 0$ expectation within the scale of statistical homogeneity ($\approx 100h^{-1}$Mpc). This is quantified with the statistic

$$\chi^2_a(n_f, n_i) = \sum_{i=n_i}^{n_f} \frac{\bar{H}_0^4 \delta H_i^2}{\bar{H}_0^2 \sigma^2_{H_i} + H_i^2 \sigma^2_{H_0}} \tag{4.5}$$

where n_f and n_i define the upper and lower shells included in the range of the calculation respectively. We will minimize on shells $1'$ through $8'$ covering a range of $6.25\,h^{-1}$Mpc to $106.25\,h^{-1}$Mpc. This covers the range of interest and avoids the data in the inner most shell which has incomplete sky coverage and the outermost shells where Hubble expansion is in the linear regime.

The frame of reference with the minimum monopole variation is found using a downhill optimization with (4.5). This reveals a global minimum at a boost in the direction $(l, b) = (59.3°, 16.6°)$ with a magnitude of $740.6^{+515.4}_{-728.8}$ km s^{-1} with respect to the Local Group frame of reference. This corresponds to a very large boost of 1203 km s^{-1} from the CMB reference frame. To make sense of this large boost velocity we consider in more detail the confidence intervals associated with this minimum.

To give an idea about the distribution of χ^2_a in the 3-dimensional parameter space $\{v, l, b\}$ we show two angular slices at fixed values of v in Figure 4.3, and a slice along the locus of (l, b) values for which χ^2_a is minimized for fixed v in Figure 4.4. The angular distribution in Figure 4.3 remains similar for all non-zero velocities. However, the confidence intervals become increasingly large as v is decreased, eventually taking up the whole sky for very small velocities. Only for large boosts do we have a well-defined



(a) **(b)**

Figure 4.3: Contour maps of angular variation of χ^2_a for two choices of boost amplitude with respect to the Local Group. (a) 740 km s^{-1} (b) 630 km s^{-1}. The inner and outer contours show the 68.3% and 90% confidence intervals respectively for the direction of the boost to the minimum χ^2_a frame of reference for each magnitude. The white cross on (b) shows the direction of the boost to the CMB (also of magnitude $\approx$ 630 km s^{-1}).

direction in which to boost to reduce monopole variation. The significance of this reduction in χ^2_a becomes apparent when we consider the distribution with respect to the magnitude.

In Figure 4.4 for each boost magnitude we locate the direction of minimum χ^2_a and plot the corresponding value. The distribution of χ^2_a relative to the 68.3% confidence interval (dashed horizontal line on Figure 4.4 (a)) reveals the primary issue in constraining a "minimum variation frame" through this technique. The near flat distribution of χ^2_a values within 1σ of the global minimum are found for a locus (l, b) of boost directions which all lie close to the galactic plane, the same region of the sky in which the COMPOSITE sample has incomplete sky coverage. The Zone of Avoidance, where the Milky Way obstructs the study of more distant objects, may be the cause of this degeneracy. We are free to perform large boosts in the plane of the galaxy as the data is not contained there. This hypothesis could be checked by simulating data with the same characteristics as the COMPOSITE sample, using exact solutions of Einstein's equations. Such an investigation is beyond the scope of this thesis. For now, with the available data we can only conclude that the Local Group is not ruled out as the standard

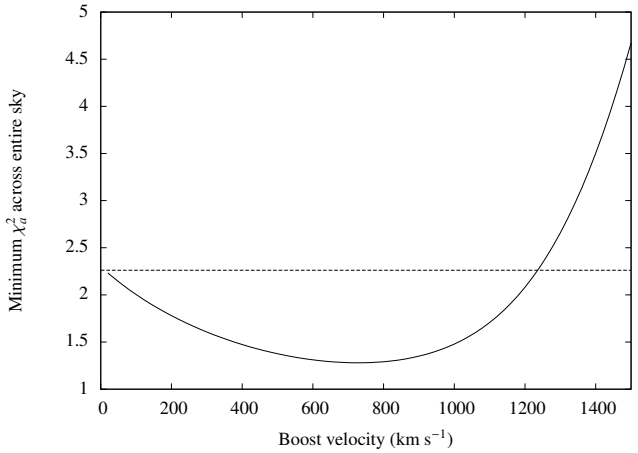


Figure 4.4: Variation of the minimum χ_a^2 with fixed boost velocity, v. The locus of (l, b) values for which this minimum is found lies within $\approx 20^\circ$ degrees of the galactic plane for cases in which χ_a^2 is within 1σ of the global minimum, as indicated by the dashed line.

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 consistent with the minimum average monopole variation frame. Although large uncertainties still exist, we are able to see that the boost to the CMB frame is far from our degenerate set of possible boosts to a definition of the standard of rest. See the location of the white cross in Figure 4.3. Table 4.1 gives a comparison between the Local Group and the absolute minimum χ_a^2 frame (labeled the Minimum Variation (MV) frame), where we have calculated the probabilities of a uniform Hubble expansion for each frame directly from the complementary incomplete gamma function for the χ_a^2 distribution and the Bayes factor is calculated as $B(r_s) = P_{\mathrm{MV}}/P_{\mathrm{LG}}$.

4.3.2 Variation in the "linear regime"

As seen from the value of χ^2 from (3.2) in Table 4.1 the fit of a linear Hubble law in the inner most shells is poor, as we would expect since these shells are in the *non-linear* regime. Beyond approximately $75\,h^{-1}$Mpc we expect to pass into the linear regime [34] and this is seen with the decreasing values of χ^2/ν (where χ^2 with no subscript corresponds to (3.2)). Now, the frame identified to be the minimum with (4.5) shows an increase in the asymptotic Hubble constant in the outermost shell of almost 1%, and

Shell s	1	2	3	4	5	6	7	8	9	10	11
N_s	92	505	514	731	819	562	414	304	222	280	91
r_s (h^{-1}Mpc)	2.00	12.50	25.00	37.50	50.00	62.50	75.00	87.50	100.00	112.50	156.25
$\bar{r}_s$ (h^{-1}Mpc)	5.43	16.33	30.18	44.48	55.12	69.24	81.06	93.75	105.04	126.27	182.59
$(H_s)_{LG}$	117.9	103.1	106.5	105.5	104.8	102.1	102.8	103.2	103.7	102.4	101.0
$(\bar{\sigma}_s)_{LG}$	4.6	1.4	1.0	0.7	0.7	0.7	0.9	0.9	1.0	0.8	1.7
$(Q_s)_{LG}$	0.000	0.000	0.000	0.000	0.998	0.940	1.000	1.000	1.000	0.993	0.999
$(\chi^2/\nu)_{LG}$	23.656	7.767	2.185	1.419	0.864	0.909	0.594	0.542	0.622	0.803	0.590
$(H_s)_{MV}$	118.5	102.9	106.7	104.5	104.8	102.9	102.6	103.9	104.9	102.7	102.0
$(\bar{\sigma}_s)_{MV}$	4.6	1.4	1.1	0.7	0.7	0.7	0.9	0.9	1.0	0.8	1.7
$(Q_s)_{MV}$	0.000	0.000	0.000	0.000	0.330	0.887	1.000	1.000	1.000	0.964	0.999
$(\chi^2/\nu)_{MV}$	29.130	12.320	3.037	2.005	1.021	0.928	0.682	0.600	0.667	0.854	0.603
$\ln B$ ($r \geq r_s$)	3.53	2.85	2.79	1.99	0.85	0.33	0.36	0.14	0.07	0.45	

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
$(\bar{\sigma}_s)_{LG}$	1.8	1.1	0.9	0.7	0.8	0.8	0.9	0.9	1.0	0.9	1.7
$(Q_s)_{LG}$	0.000	0.000	0.000	0.031	0.960	1.000	1.000	1.000	0.996	0.999	0.999
$(\chi^2/\nu)_{LG}$	11.427	3.246	1.792	1.090	0.907	0.701	0.592	0.608	0.728	0.711	0.590
$(H_s)_{MV}$	102.7	104.3	103.3	106.1	104.2	102.9	102.9	104.7	102.7	102.9	102.0
$(\bar{\sigma}_s)_{MV}$	1.8	1.1	0.9	0.7	0.8	0.8	0.9	0.9	1.0	0.9	1.7
$(Q_s)_{MV}$	0.000	0.000	0.000	0.000	0.481	1.000	1.000	1.000	0.967	0.997	0.999
$(\chi^2/\nu)_{MV}$	18.547	4.940	2.429	1.428	1.002	0.734	0.704	0.613	0.807	0.752	0.603
$\ln B$ ($r \geq r_s$)	2.38	2.32	2.34	1.93	0.55	0.33	0.41	0.14	0.24	0.50	

Table 4.1: Hubble expansion variation in radial shells in minimum Hubble expansion variation (MV) and LG frames. Spherical averages (3.3) 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 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 χ^2 (for $\nu = N_s - 1$); $\ln B$ 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 χ^2 is summed in all shells with $r > r_s$. H_s and $\bar{\sigma}_s$ are given in units $h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

a worse fit to a linear Hubble law in all shells beyond this non-linear regime. However, any true candidate for a minimum Hubble expansion variation frame should also clearly demonstrate the emergence of a linear Hubble law consistent with the existence of a statistical homogeneity scale. The χ^2_a statistic (4.5) 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 conduct a similar search as conducted for χ^2_a above. We will minimize the quantity

$$\chi^2_b = \frac{\sum_{s=7}^{11} \nu_s \left(\frac{\chi^2_s}{\nu_s}\right)}{\sum_{s=7}^{11} \nu_s} \tag{4.6}$$

where χ^2_s is given by (3.2) in the s^{th} shell and ν_s is 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 this "linear scale". Thus (4.6) gives a measure of the goodness of fit to a linear Hubble law across the outer 5 shells. This choice is made without fitting a single Hubble law across this whole range to allow for the Hubble constant to continue to approach its asymptotic value, as can be seen in Table 4.1. We find $\chi^2_b = 0.631$ for the LG frame, $\chi^2_b = 0.692$ for the minimum χ^2_a frame, and $\chi^2_b = 0.653$ in the CMB frame. So even the CMB frame has a more suitable emerging Hubble law than the minimum χ^2_a frame. To understand if there is any improvement on the LG we consider the distribution of (4.6) with respect to boosts from the LG.

In Figure 4.5 we locate the direction of minimum χ^2_b at each boost magnitude and plot the corresponding value, analogously to Figure 4.4. We find a best fit boost of 222.3 km s^{-1} in the direction $(l, b) = (241.84°, 70.53°)$ with respect to the LG frame, with a value of $\chi^2_b = 0.621$. This is 45° from the direction of the residual CMB temperature dipole in the LG frame and so does not appear related.

In Figure 4.5 for each boost magnitude we calculate the value of χ^2_a in an (l, b) direction determined by minimizing with respect to χ^2_b, and vice versa. Thus, we compute the locus of (l, b) values in the $\{v, l, b\}$ parameter space that minimize χ^2_b for each fixed v, and then compute χ^2_a at these parameter values, and vice versa. It is apparent that making boosts of the order of 100 km s^{-1} along the locus of (l, b) values which minimize χ^2_b results in an increase in χ^2_a beyond its 68.3% confidence interval (horizontal dashed

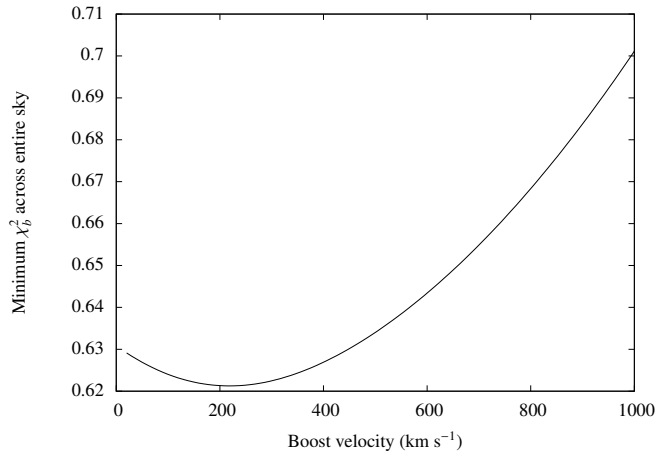
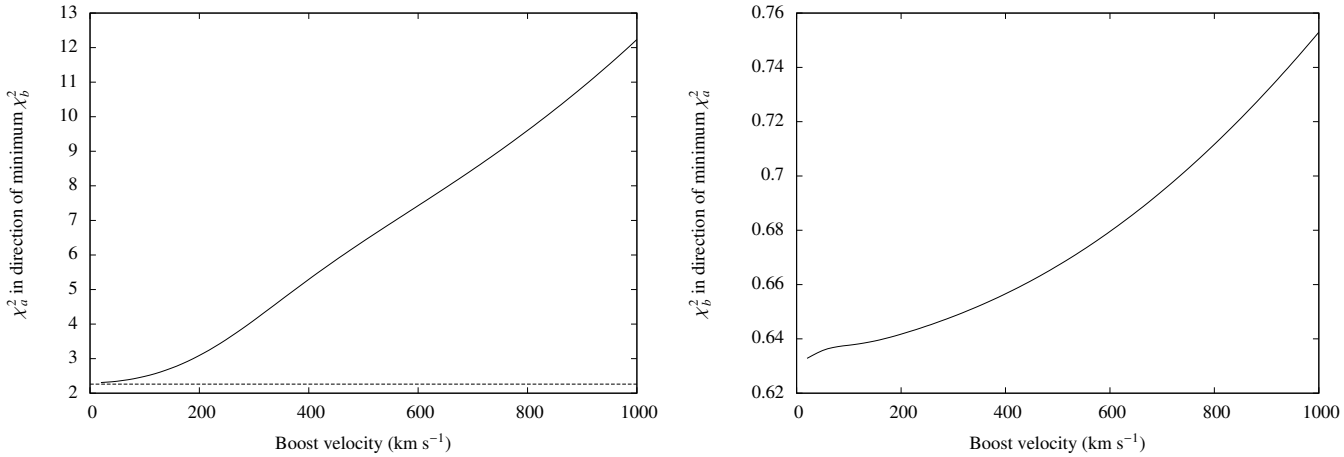


Figure 4.5: Variation of the minimum χ_b^2 with fixed boost velocity, v.



(a) **(b)**

Figure 4.6: (a): Variation of χ_a^2 along a locus of (l, b) values for which χ_b^2 is minimized at each fixed boost velocity v. (a) Variation of χ_b^2 along a locus of (l, b) values for which χ_a^2 is minimized at each fixed boost velocity v.

line). This is because making improvements on χ_b^2 requires boosts in (l, b) directions away from our minimum χ_a^2.

The Hubble expansion is indeed very close to linear in the outer shells, as indicated by the extremely low values of χ_b^2. Additionally local boosts have a relatively small effect on the values of cz_i in the scale of linearity. These facts mean we should also exercise caution about drawing conclusion solely from the minimization of χ_b^2. While the statistic

is not potentially biased by the outermost shell as the χ^2_a statistic is, there are less data points per se in those outer shells used in (4.6), as compared to the inner shells which are important in the χ^2_a statistic. All values of χ^2_b shown in Figure 4.5 and 4.6 are consistent with extremely probable fits. Much more data is needed to use the χ^2_b in a reliable way.

Finding a joint minimum for χ^2_a and χ^2_b would be feasible if the (l, b) values of each global minima were close on the sky. However, this is not the case – the global minimum for χ^2_a is at an angle of 93° from the global minimum of χ^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.

4.3.3 Systematic boost offsets from the Local Group

Neither statistic χ^2_a or χ^2_b appears entirely satisfactory for establishing a global minimum expansion variation frame. The χ^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 $\bar{H}_0$. The most that 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 χ^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 §4.2. Any true best fit frame should show a clear signal of a boost offset (4.2) with respect to the Local Group. The "best" boost offset can be characterised 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 (4.3) determine a derived boost velocity, $v_{\rm der}$. Any boost offset should have $v_{\rm der}$ consistent with the true boost magnitude v_{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 (7.8) in the Appendix. This is also problematic since all fits are extremely good due to the large uncertainties in the H_s.

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[2].

First we check for a systematic boost offset for the global $\chi^2_{\rm a}$ minimum frame determined in §4.3.1. A sky map of f_p values as given by (4.4) for boosts of magnitude $740\,{\rm km\,s^{-1}}$ is given in Fig. 4.8(a), with the $1\,\sigma$ and $2\,\sigma$ confidence intervals for $\chi^2_{\rm a}$ displayed. We note that there are in fact boosts with values of $p \approx -1$ consistent with the $\chi^2_{\rm a}$ minimum. However, these directions are far more constrained and do not align with the exact minimum. In addition, the (ℓ, b) direction at this magnitude with the best fit to (4.3) 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 $\chi^2_{\rm a}$, 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 (4.4) for boosts from the LG frame, we arbitrarily choose a boost magnitude of $200\,{\rm km\,s^{-1}}$ and plot f_p with respect to (ℓ, b) in Fig. 4.8(b). We have found that for all interpolating velocities between the $200\,{\rm km\,s^{-1}}$ and $740\,{\rm km\,s^{-1}}$ cases displayed in Figure 4.8 there is a region of (ℓ, 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. 4.9 we plot the values of $v_{\rm der}/v_{\rm true}$ and $S/(n-2)$ (for n data points) from (7.8) with respect to the boost magnitude, where for each magnitude the (ℓ, 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 (ℓ, b) directions in the 3-dimensional $\{v, \ell, b\}$ parameter space. The expected value of S has a χ^2 distribution for $(n-2)$ degrees of freedom, and thus $S/(n-2)$ has an expectation value of unity [117]. Clearly, the values of $S/(n-2)$ in Fig. 4.9(b) are consistent with a very good fit to (4.3) 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 $H_{\rm LG} - H_{\rm X}$. Although (7.8) it is not useful for constraining the boost magnitude, we nonetheless see that the ratio of derived and true velocities in Fig. 4.9 does show a meaningful difference on this interval.

[2]The choice of shell boundaries introduces systematic uncertainties in this analysis. In the production of the sky maps in Fig. 4.8 we calculate both primed and unprimed shells and fit the power law (4.3) to all points. For Fig. 4.9 we consider fits to primed and unprimed only, and to both.

Using both primed and unprimed shells in our calculation we find that the $v_{\rm der}/v_{\rm true} = 1$ at $v_{\rm true} = 122.5\,{\rm km\,s^{-1}}$ in a direction $(\ell, b) = (60°, -4°)$. For this boost, we find the result $p = -1.13 \pm (1.16)_{\rm stat} \pm (0.29)_{\rm sys}$, where the systematic uncertainty is determined as in §4.2, 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. 4.7. 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 $\chi^2_{\rm a}$ and $\chi^2_{\rm b}$.

Our choice of frame X above is based on taking $v_{\rm der}/v_{\rm true} = 1$, 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 $v_{\rm der} = v_{\rm true}$.

4.3.4 The degenerate boost direction

The methods of §4.3.1 and §4.3.3 have determined the existence of a degenerate boost direction in the plane of the galaxy. The consistency between these methods may be less significant as they are not completely independent. In the Local Group frame of reference the primary source of the Hubble 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 χ^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 (4.3) with $p \approx -1$. Thus it is not surprising that we see the consistency in the direction of minimum χ^2_a and the values of $p \approx -1$ and $A > 0$.

4.4 Angular Hubble expansion variation

WSMW13 postulate that consideration of angular Hubble expansion averages could offer an independent test of the minimum variation frame of reference. When one takes angular averages of the Hubble expansion a dipole becomes apparent. WSMW13 show that this dipole is strongly correlated with CMB residual temperature dipole when both

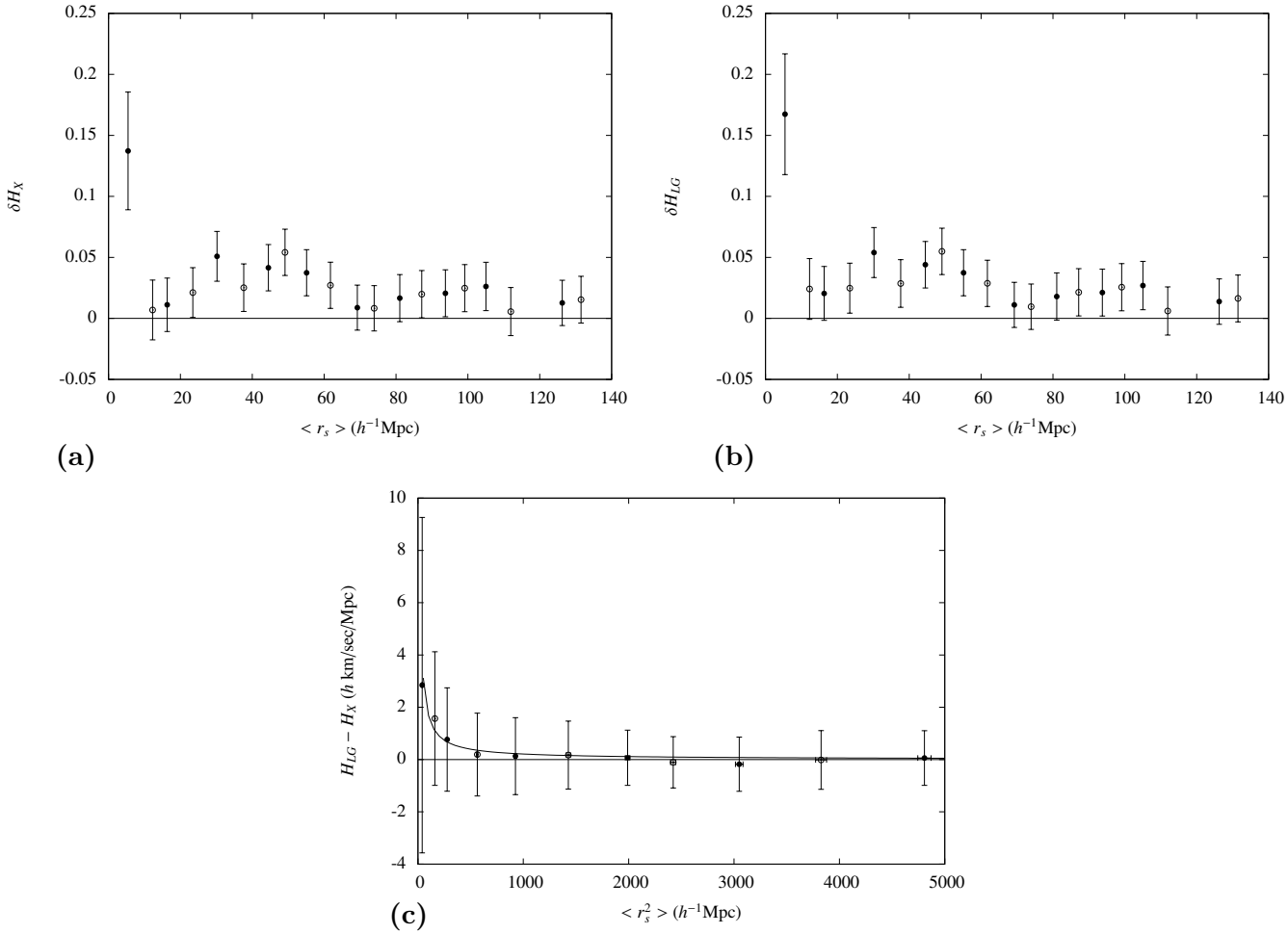


Figure 4.7: The variation (3.7) in the spherically averaged Hubble law in (a): the Local Group frame (b): frame X. (c): The systematic boost offset between frame X and the LG.

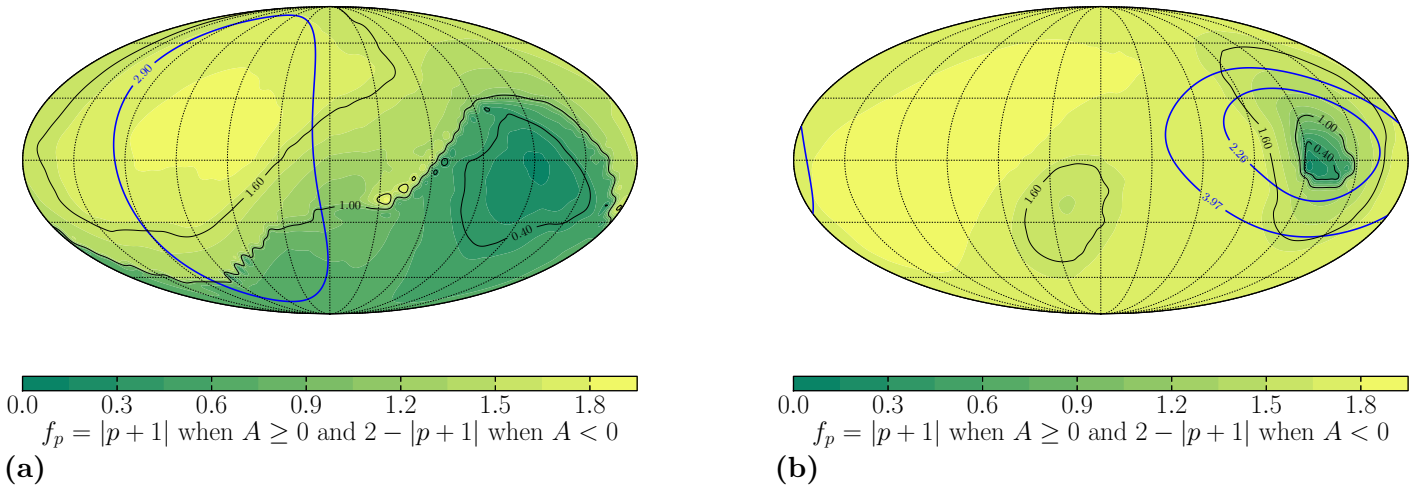
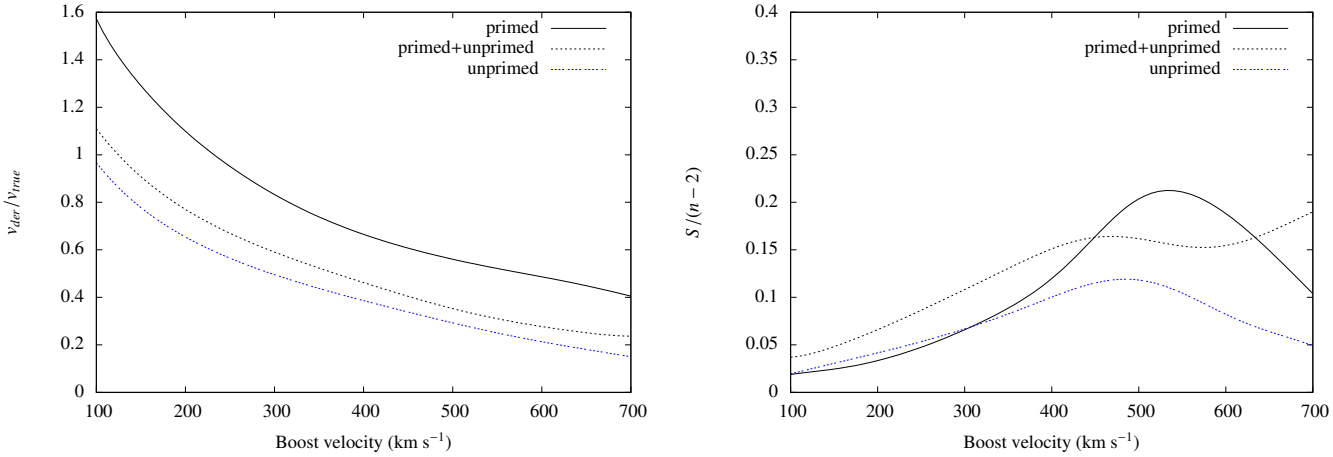


Figure 4.8: The best fit parameters to a systematic boost offset for boosts from the Local Group of magnitude (a): 740 km s^{-1} (b): 200 km s^{-1}. The thick blue contours denote the corresponding χ^2_a distribution. In (a) only the 68.3% confidence interval is visible ($\chi^2_a = 2.90$) and on the right both the 68.3% and 90% confidence intervals are displayed, being $\chi^2_a = 2.26$ and $\chi^2_a = 3.97$ respectively. In all figures, the galactic longitudes $\ell = 0°, 180°, 360°$ are on the right edge, centre and left edge respectively.

are considered with respect to the Local Group. If we are to define a new standard of rest, within which we still observe a residual CMB temperature dipole, then this dipole must have a non-kinematic origin. The correlation of the CMB residual temperature dipole and Hubble expansion dipole is a necessary condition for the foreground structure to be simultaneously responsible for these effects[3]. Thus we are interested in finding a frame of reference in which this correlation is maximized.

To compare the CMB temperature and Hubble expansion variation dipoles an artificial residual CMB temperature map is produced in each frame of reference. This is achieved by subtracting a boosted CMB sky with temperature (3.19) from the corresponding observed pure temperature monopole plus dipole maps using the values of

[3] A correlation alone is not a sufficient condition for the dipole to be non-kinematic. However, a boost to the CMB frame would eliminate a kinematic dipole in the angular variation map once one averaged on scales larger than the structures responsible for the dipole. WSMW13 show on the contrary that for large spherical shells the fit of a dipole Hubble law gives a dipole consistent with zero in the LG frame but not in the CMB frame.



(a) **(b)**

Figure 4.9: **(a)** The ratio of the derived velocity from the best fit power law and the true boost velocity, at each v the direction is determined by the best fit for which p is closest to -1. **(b)** The statistical variation in the same direction as determined by $S/(n-2)$ from (7.8), where n is the number of data points being fit.

Fixsen *et al.* [104]. The value of v is that of the boost we are interested in, and $\gamma \approx 1$ since we are not dealing with relativistic velocities. After digitizing the Hubble expansion variation sky map and the CMB residual temperature map we can use the Pearson correlation coefficient to give a measure of the correlation between these maps using (3.20).

4.4.1 A possible constraint on the minimum Hubble expansion variation frame

The correlation between the Hubble expansion variation dipole and residual temperature dipole found in the Local Group frame by WSMW13 is not perfect. This is certainly to be expected, since the angular Hubble variation on any scale contains higher order multipoles and is not a pure dipole. However, we will investigate if a frame of reference does exist in which the the value of $\rho_{\nu T}$ is closer -1.

First we calculate the $\rho_{\nu T}$ for each boost magnitude from the LG along the locus of (l, b) directions corresponding to the minimum χ^2_a at each fixed magnitude. More precisely, for each boost magnitude we take the direction with the lowest χ^2_a as calculated on shells $1' - 8'$. By this approach we are placing greater value on minimizing monopole

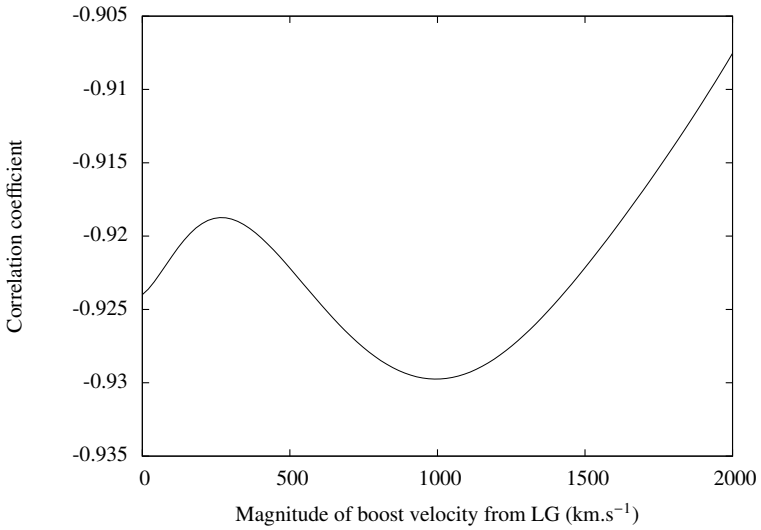


Figure 4.10: Pearson correlation coefficient for the correlation of the residual CMB temperature and Hubble expansion variation dipoles for $r \geq 15h\,h^{-1}$Mpc in the direction determined by minimum χ^2_a with respect to the boost magnitude.

variation with respect to direction. This is reasonable considering the nature of the well defined χ^2_a distribution and associated confidence intervals, as opposed to the near flat distribution with respect to magnitude. Figure 4.10 shows the value of the correlation coefficient across a range of boost velocities using this method.

It is evident that there is a large window of boost velocities from zero to over 1000 km s^{-1} that show no significant improvement on the correlation in the LG frame (see the vertical scale on Figure 4.10). Therefore we conclude that we cannot use this angular analysis to remove the degeneracy as we see no notable decrease in $\rho_{\nu T}$ compared to the Local Group along the direction of minimum χ^2_a.

When we consider the $\rho_{\nu T}$ in a range of different *directions* it is apparent that boosts do exist from the Local Group that give significant improvements in the correlation.

By making boosts in the appropriate direction from the Local Group we are able to artificially increase the strength of the Hubble expansion dipole *and* CMB residual temperature dipole simultaneously at the expense of also increasing the monopole variation. By calculating the Hubble expansion dipole and quadrupole coefficients using HEALPIX[4] in the *frame of maximum correlation* for a given boost magnitude from the

[4]http://healpix.jpl.nasa.gov/ [118]

Local Group we can observe this artificial increase in the strength of the dipole relative to other variations. Boosting in directions which make the dipoles more pronounced will naturally increase the $\rho_{\nu T}$ in (3.20), but this is irrelevant if the monopole variation is also increased. Therefore as this artificially induced increase in the correlation cannot be distinguished from a physically meaningful increase this method does not offer a viable criterion for the minimum variation frame and is thus abandoned as a line of investigation.

4.4.2 Structures in the range $40h^{-1} - 60\,h^{-1}$Mpc

The method of §4.4.1 does not provide a useful means of further constraining a global standard of rest. However, if we restrict the range of shells used in the χ^2_a statistic we can further test the improvement of the CMB frame over the LG frame in the range $40h^{-1} - 60\,h^{-1}$Mpc noted in WSMW13.

We begin with a naïve investigation through visual inspection of sky maps. We calculate $\rho_{\nu T}$ for 164 points on a grid across the entire sky at three different boost velocities, using (3.20) in a $r > 15\,h^{-1}$Mpc sphere. These boosts are of magnitude 200 km s^{-1}, 740 km s^{-1} (corresponding to the global minimum χ^2_a found in section 4.3.1) and an intermediate value of 450 km s^{-1}. Since the value of $\rho_{\nu T}$ will continue to achieve values ≈ -1 for larger and larger boosts, there is no unambiguous way to choose a specific magnitude to consider. Thus we arbitrarily choose three different magnitudes to plot in the first instance.

The resulting maps, produced using an interpolation function, are presented in Figures 4.11 and 4.12. In these plots we overlay contours for the distribution of $\chi^2_a = \chi^2_a(n_i, n_f)$ calculated on different ranges of shells (different values of n_i and n_f in (4.5)). For the 200 km s^{-1} case we do not plot this distribution as well, allowing one to clearly see the nature of $\rho_{\nu T}$ with respect to direction.

The distribution of (3.20) across the sky does show an interesting property when considered with χ^2_a computed on particular ranges of shells. It is clear that the distribution of $\chi^2_a(1', 8')$, across shells $1' - 8'$ ($6.25h^{-1} - 106.25\,h^{-1}$Mpc), does not align with the most correlated Pearson correlation coefficient values. In contrast, the distribution of $\chi^2_a(4', 6')$, on shells $4' - 6'$ ($43.75h^{-1} - 81.25\,h^{-1}$Mpc), and $\chi^2_a(4', 5')$, shells $4' - 5'$ ($43.75h^{-1} - 68.75\,h^{-1}$Mpc) show increased values towards the direction of $\rho_{\nu T} \approx -1$. Therefore, we expect to see an increased monopole variation of the shells in this range,

at the same time as an improvement in the alignment of the CMB residual temperature and angular Hubble expansion variation dipoles.

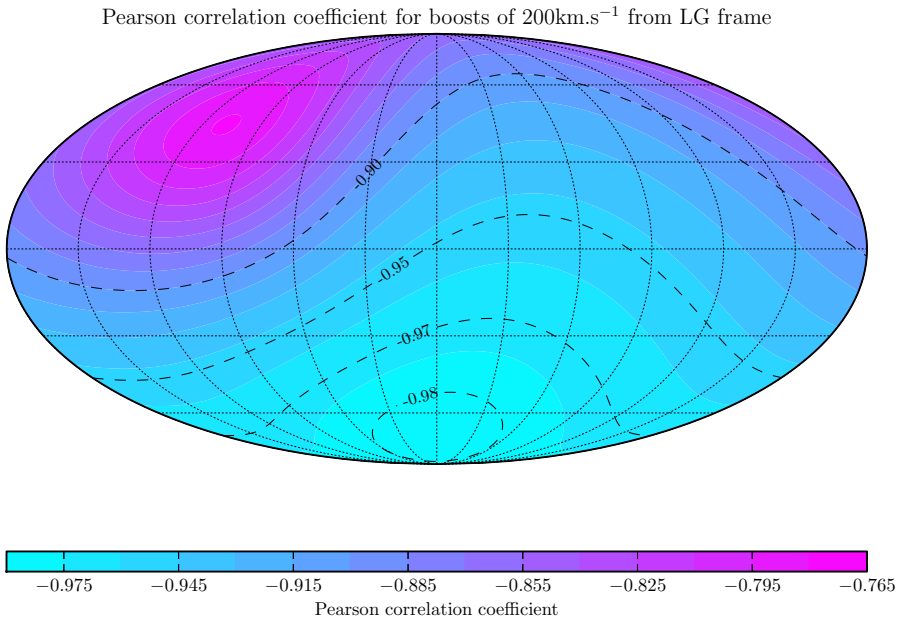
This expectation is observed when we consider δH in a frame boosted at 450 km s^{-1} with a value of $\rho_{\nu T} = -1.00$. In Figure 4.13 this frame clearly shows an increased monopole variation over the $40h^{-1} - 60\,h^{-1}$Mpc range compared to the same plot in the LG or CMB frames. We note that a similar effect is evident when comparing the CMB and Local Group frames, where although the LG has an overall decrease in monopole variation, the LG does display an increased variation over this same range (see Figure 3.1). Therefore we make the preliminary observation that boosting to a frame with improved correlation increases the variation over the $40h^{-1} - 60\,h^{-1}$Mpc range.

We are interested in determining a more precise range over which increases in monopole variation are correlated with increases in $\rho_{\nu T}$. Thus instead of dealing with fixed spherical shell boundaries, we allow the boundaries to vary. We will retain the shell width of $12.5\,h^{-1}$Mpc, and primarily use two consecutive shells to calculate $\chi^2_a(1,2)$ while the inner cutoff of shell 1 is varied from $0\,h^{-1}$Mpc to $100\,h^{-1}$Mpc. The alignment of the resulting χ^2_a distribution and $\rho_{\nu T}$ distribution is measured using the appropriate Pearson correlation coefficient

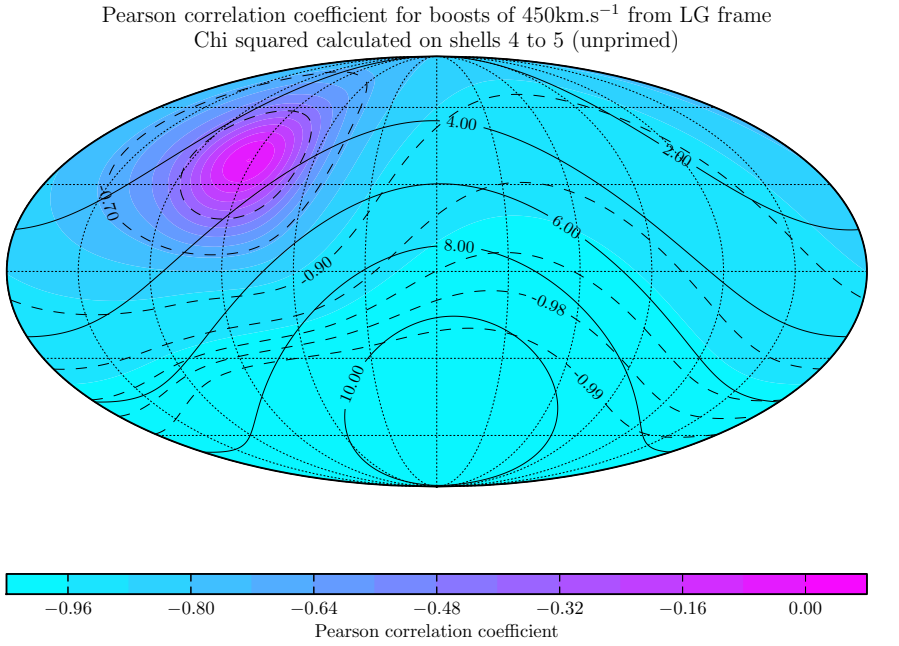
$$\rho_{\chi^2} = \frac{\sqrt{N_p}\sum_\alpha \left(\chi^2_a(\alpha) - \overline{\chi^2_a}\right)\left(\rho_{\nu T}(\alpha) - \overline{\rho_{\nu T}}\right)}{\sqrt{\left[\sum_\alpha \left(\chi^2_a(\alpha) - \overline{\chi^2_a}\right)^2\right]\left[\sum_\alpha \left(\rho_{\nu T}(\alpha) - \overline{\rho_{\nu T}}\right)^2\right]}} \tag{4.7}$$

where $\overline{\chi^2_a}$ and $\overline{\rho_{\nu T}}$ are the mean values and $\chi^2_a(\alpha)$ and $\rho_{\nu T}(\alpha)$ denote the value of χ^2_a and $\rho_{\nu T}$ at each individual point respectively. In this case we use a HEALPIX equal spacing grid of 3072 points.

We are specifically interested in determining the range over which an increase in monopole variation correlates with a decrease in $\rho_{\nu T}$ and thus the case where ρ_{χ^2} is closet to -1. Figure 4.13 gives the value of (4.7) with respect to the inner shell cutoff as this is varied. We see that increased variation in the 2 shells (covering a range of $25\,h^{-1}$Mpc) is most correlated with a decrease in $\rho_{\nu T}$ when these shells are on the range of approximately $41h^{-1} - 66\,h^{-1}$Mpc. We repeat this analysis on 3 consecutive shells and find consistent results yet the range of best anti-correlation becomes less well defined since the distribution is affected by monopole variations that do not show the same characteristics. This is what one would expect if the structure associated with this effect is also confined to an interval $\lesssim 25\,h^{-1}$Mpc in radial width.

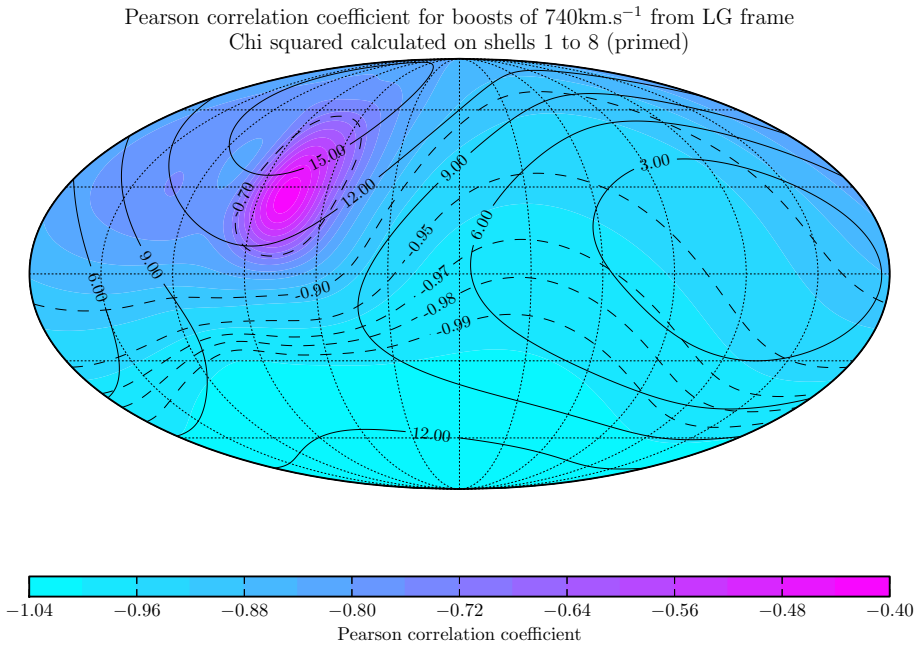


(a)

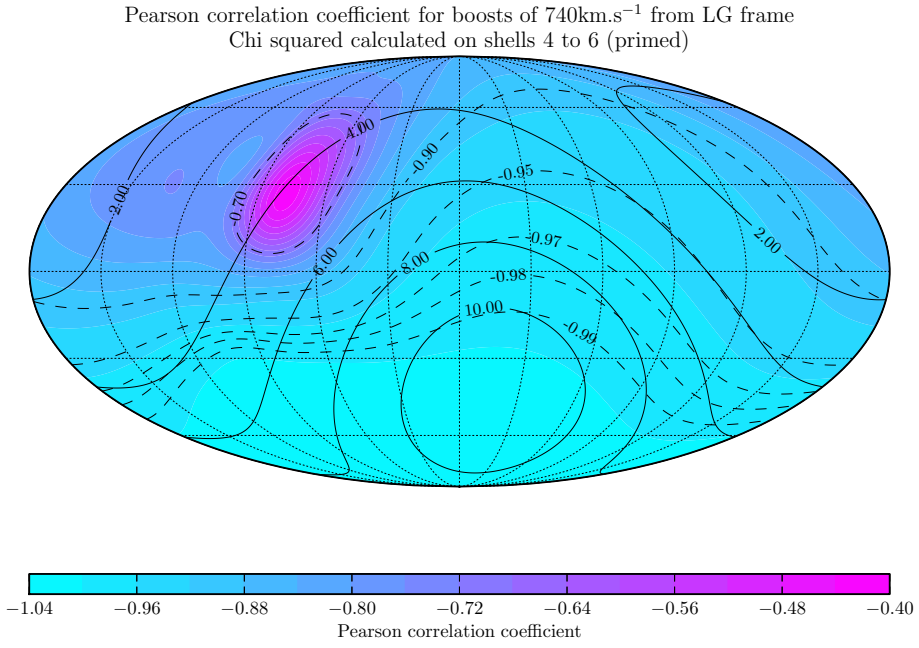


(a)

Figure 4.11: (a): The distribution of $\rho_{\nu T}$ with respect to boost direction for boosts of magnitude 200 km s^{-1}. (b): The distribution of $\rho_{\nu T}$ with respect to boost direction for boosts of magnitude 450 km s^{-1} given by the dashed contours and the background colour. The distribution of χ^2_a is given by the solid contours, where χ^2_a is computed over shells 4' to 5' ($43.75h^{-1} - 68.75\,h^{-1}$Mpc).

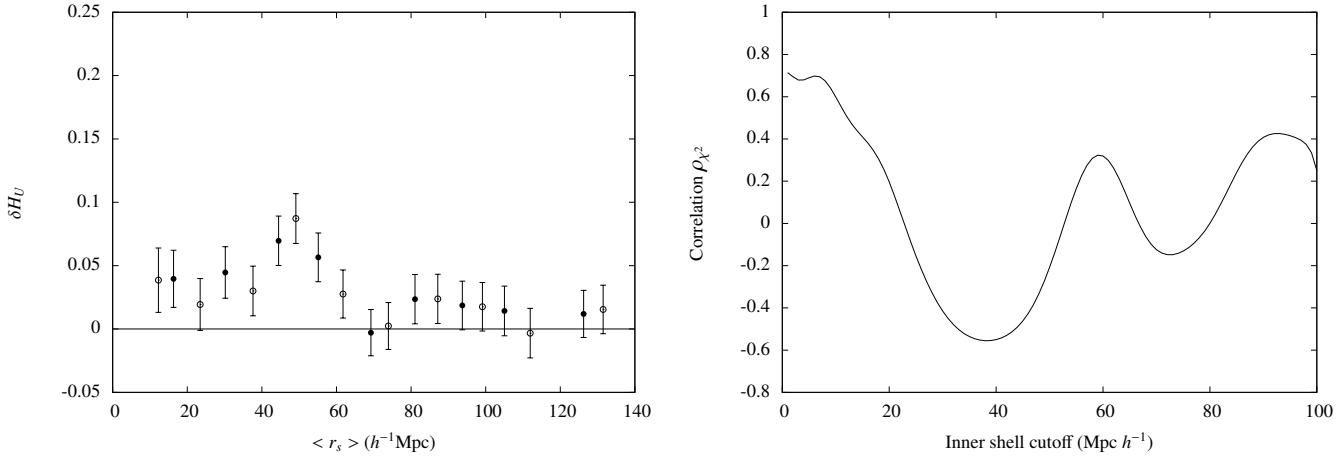


(a)



(a)

Figure 4.12: The distribution of $\rho_{\nu T}$ with respect to boost direction for boosts of magnitude 740 km s^{-1} given by the dashed contours and the background colour. The distribution of χ^2_a is given by the solid contours, where χ^2_a is computed over shells (a): 1' to 8' ($6.25h^{-1} - 106.25\,h^{-1}$Mpc). (b): 4' to 6' ($43.75h^{-1} - 81.25\,h^{-1}$Mpc).



(a) **(b)**

Figure 4.13: (a): The monopole Hubble variation in a frame of reference boosted at 450 km s^{-1} (chosen arbitrarily) with respect to the LG in a direction $(l, b) = (140°, -40°)$ for which $\rho_{\nu T} = -1.00$. (b): The correlation (4.7) of the χ^2_a distribution for two 12.5 h^{-1}Mpc shells and the $\rho_{\nu T}$ distributions as a function of the inner shell cutoff, for boosts of 450 km s^{-1}.

Therefore boosts that increase the power of the Hubble variation dipole simultaneously increase the monopole variation over the $41h^{-1} - 66\,h^{-1}$Mpc range[5]. This further supports the notion that structure in this particular range is associated with the dipole as originally postulated by WSMW13 [73].

[5]We also note a strong positive correlation for shells with very low inner cutoff radii, which on further investigation is the result of an increase in χ^2_a *and* $\rho_{\nu T}$ (towards values close to 0), both in the direction of the boost to the CMB frame. Thus it is not relevant to this discussion.

CHAPTER 5

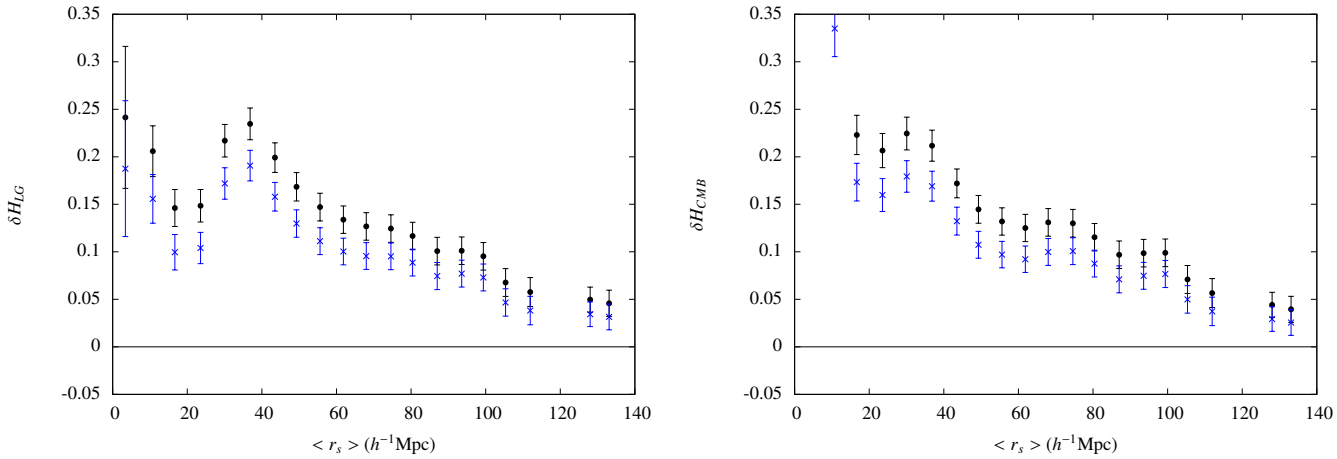
Hubble expansion variation in the *Cosmicflows*-2 catalogue

The previous chapter was 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 introduced in chapter 2. We aim to verify our previous results by unravelling the systematic differences in these data sets that become evident through our analysis.

The same analysis is repeated to calculate δH from (3.7) in the same two configurations of 11 spherical shells used in chapter 4. The CF2 data is presented with a modified "recession velocity" (cz) and a raw observed redshift. Given the prevalence of the use of these type of modifications, particularly in bulk flow studies, it is worthwhile to briefly investigate the effect these can have on monopole Hubble expansion variation. Tully *et al.* [68] define an adjustment assuming a FLRW cosmology with $\Omega_m = 0.27$ and $\Omega_\lambda = 0.73$. This adjustment is given by a Taylor expansion to $\mathcal{O}(z^3)$ of a homogeneous isotropic expansion law,

$$cz_{\text{mod}} = cz\left[1 + 0.5(1 - q_0)z - (1/6)(1 - q_0 - 3q_0^2 + 1)z^2\right] \tag{5.1}$$

where z is the redshift in the CMB frame and $q_0 = 0.5(\Omega_m - 2\Omega_\lambda)$. As we wish to deal with model independent quantities this is not the type of adjustment we want to make. In particular, a homogeneous isotropic expansion law should not be assumed below the scale of statistical homogeneity if the conclusions of WSMW13 are correct. However,



(a) **(b)**

Figure 5.1: The monopole Hubble expansion variation for the CF2 sample without the FLRW "correction" (5.1) (*black* filled circles) and with the FLRW "correction" (5.1) (*blue* crosses) in the: **(a)** Local Group frame of reference; **(b)** CMB frame of reference.

for completeness we consider this in order to rule it out as the cause of much larger systematic differences we will discuss shortly.

Figure 5.1 shows δH using the CF2 sample. It is immediately evident these plots are very different to that found with the COMPOSITE sample, as given in Figure 3.1. In fact the uniformity of Hubble expansion in spherical shells is considerably worse than in the COMPOSITE sample, in both rest frames. While there is a small shift introduced with the redshift modification, it is certainly not type of systematic we expect to explain this marked difference.

Note that the systematic discrepancy seen in Figure 5.1 can be misleading. Recall that due to the form of δH, the vertical shift seen in these values is primarily due to a change in the asymptotic value, $\bar{H}_0$, since data at smaller distances are barely affected by the correction in (5.1). Table 5.1 gives the numerical values of the Hubble constant with adjusted and raw redshifts, along with the number of objects in each shell and the mean shell radii.

5.1 CF2 and Malamquist bias in SFI++

The issue of Malmquist bias, explored in chapter 2, is the most likely cause of the systematic difference we see in Figure 5.1 compared with the same plot using the COMPOSITE

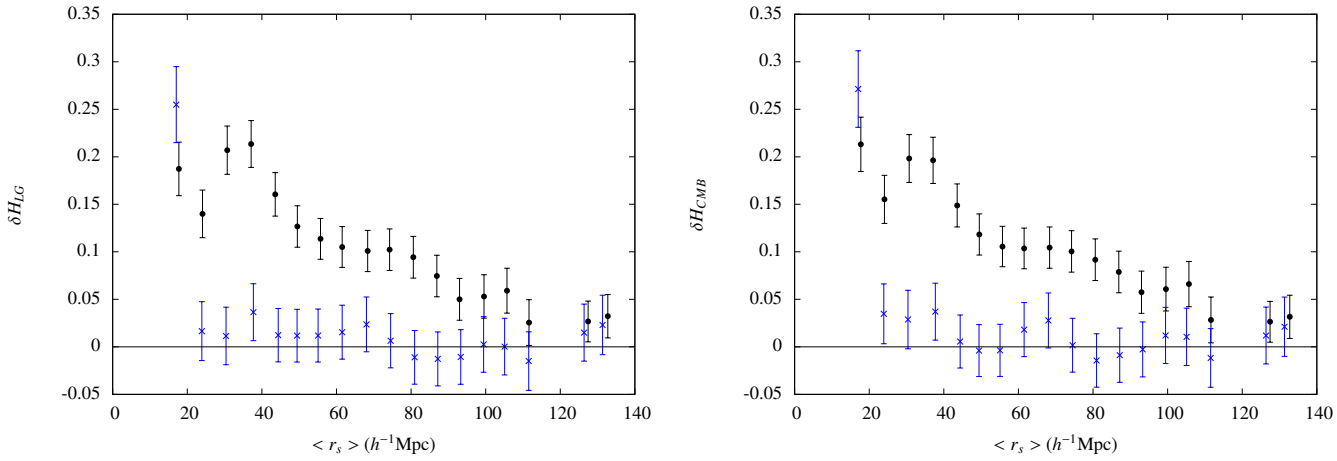
Shell s	1	2	3	4	5	6	7	8	9	10	11
N_s	579	946	834	936	959	794	739	670	497	825	333
r_s (h^{-1}Mpc)	2.00	12.50	25.00	37.50	50.00	62.50	75.00	87.50	100.00	112.50	156.25
$\langle r \rangle_s$ (h^{-1}Mpc)	3.41	16.67	30.07	43.49	55.59	67.99	80.40	93.57	105.34	128.00	186.90
$(H_s)_{\rm CMB}$	177.3	110.6	110.8	106.0	102.4	102.3	100.9	99.4	96.9	94.5	90.5
$(\sigma_s)_{\rm CMB}$	10.5	1.5	1.0	0.8	0.7	0.8	0.8	0.8	0.9	0.7	0.9
$(H_s)_{\rm CMB,adjusted}$	177.7	111.3	111.9	107.4	104.1	104.3	103.2	102.0	99.6	97.7	94.9
$(\bar{\sigma}_s)_{\rm CMB,adjusted}$	10.5	1.5	1.1	0.8	0.8	0.8	0.8	0.8	0.9	0.7	1.0
$(H_s)_{\rm LG}$	112.2	103.6	110.0	108.4	103.7	101.8	100.9	99.5	96.5	94.9	90.4
$(\sigma_s)_{\rm LG}$	6.6	1.4	1.0	0.8	0.8	0.8	0.8	0.8	0.9	0.7	0.9
$(H_s)_{\rm LG,adjusted}$	112.6	104.2	111.1	109.8	105.3	103.8	103.2	102.1	99.2	98.0	94.8
$(\sigma_s)_{\rm LG,adjusted}$	6.7	1.4	1.1	0.8	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}Mpc)	6.25	18.75	31.25	43.75	56.25	68.75	81.25	93.75	106.25	118.75	156.25
$\langle r \rangle_s$ (h^{-1}Mpc)	10.76	23.54	36.85	49.29	61.86	74.59	87.01	99.37	111.95	133.10	186.90
$(H_s)_{\rm CMB}$	126.1	109.2	109.6	103.6	101.8	102.2	99.2	99.4	95.6	94.0	90.5
$(\sigma_s)_{\rm CMB}$	2.5	1.2	0.9	0.8	0.7	0.8	0.8	0.8	0.9	0.8	0.9
$(H_s)_{\rm CMB,adjusted}$	126.7	110.0	110.9	105.1	103.6	104.4	101.6	102.2	98.4	97.3	94.9
$(\sigma_s)_{\rm CMB,adjusted}$	2.5	1.2	0.9	0.8	0.8	0.8	0.8	0.8	1.0	0.8	1.0
$(H_s)_{\rm LG}$	109.0	103.8	111.6	105.6	102.5	101.6	99.5	99.0	95.6	94.5	90.4
$(\sigma_s)_{\rm LG}$	2.1	1.1	1.0	0.8	0.7	0.8	0.8	0.8	0.9	0.8	0.9
$(H_s)_{\rm LG,adjusted}$	109.5	104.6	112.9	107.1	104.3	103.8	101.8	101.7	98.4	97.8	94.8
$(\sigma_s)_{\rm LG,adjusted}$	2.1	1.1	1.0	0.8	0.8	0.8	0.8	0.8	1.0	0.8	1.0

Table 5.1: Hubble expansion variation in radial shells in CMB and LG frames for the CF2 data. Spherical averages (3.3) 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, $(H_s)_{LG}$ and $(H_s)_{CMB}$ and there associated uncertainties in the LG and CMB frames for both the raw redshifts and those adjusted with (5.1).

sample in Figure 3.1. Since Springob *et al.* [76] supply a data set with and without corrections for Malmquist bias, it makes it an ideal set of data to consider the effects that this bias has on a spherically averaged Hubble expansion.

To understand the effect of Malmquist bias corrections, as applied by Springob *et al.* [76], we calculate the monopole variation of the Hubble expansion for the SFI++ sample with and without corrections. Figure 5.2 shows the significant difference in δH between these two treatments of distance. Thus we are interested in the effect that the uncorrected SFI++ data points have on the entire CF2 catalogue. We find that there are 3625 points in common with our version of the SFI++ sample and CF2, and repeat the analysis with these data points removed. We can see that removal of this potentially



(a) **(b)**

Figure 5.2: 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.

biased data does not affect the nature of δH to a statistically significant extent (Figure 5.3). This is an indication that the systematic bias present in the SFI++ raw distances (not corrected for Malmquist bias) is likely to also be present in the rest of the CF2 data. Thus we are confident that the discrepancy seen in these monopole variations is a systematic one, arising from the the treatment of Malmquist bias in the CF2 catalogue discussed in chapter 2.

5.2 The systematic boost offset revisited

If the CF2 data did not show evidence of a systematic bias in the distance then we should expect to see a systematic boost offset in the CF2 data between the spherically averaged Hubble expansion in the CMB and LG frames. Although we are aware that there is an unresolved systematic issue, we nonetheless explore whether such an effect is still observable to any extent. We found that in spite of the bias problem, the signature of a systematic boost offset studied in §4.2 is nonetheless evident in CF2, as this involves the *difference* of the H_s values in the LG and CMB frames from Table 5.1, as plotted in panels (a) and (b) of Fig. 5.1.

Fig. 5.4 shows the results of repeating the analysis used in §4.2, using the unadjusted CF2 distances. The best fit value for p in (4.3) is found to be $p = -0.83 \pm 0.17$ for

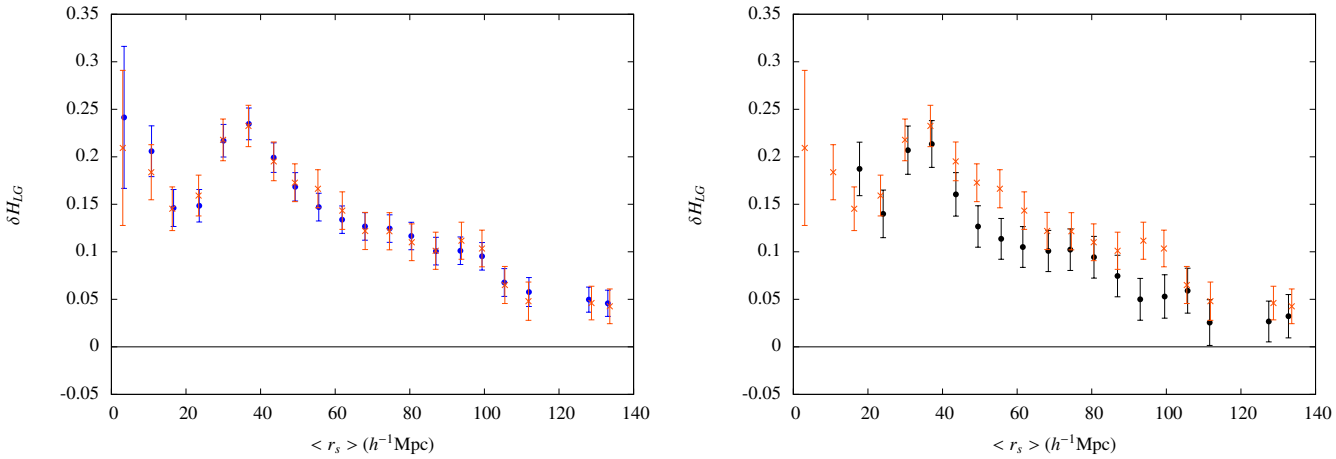


Figure 5.3: The monopole Hubble expansion variation in the LG frame for: **(a)** all CF2 data (*blue* filled circles) and CF2 data without SFI++ data (*red* crosses); **(b)** CF2 data without SFI++ data (*red* crosses) and SFI++ only without Malmquist corrections (*black* filled circles). For the uncorrected SFI++ data the first unprimed and primed values are not shown, these are 2.7 and 0.65 respectively.

unprimed and $p = -0.86 \pm 0.26$ for primed shells. Varying the shell boundaries as in §4.2 we find a value of $p = -0.84 \pm (0.21)_{\mathrm{stat}} \pm (0.06)_{\mathrm{sys}}$. However, if we compare Fig. 5.4 with Fig. 4.1, we see that there are more data points with $(H_s)_{\mathrm{CMB}} < (H_s)_{\mathrm{LG}}$, which do not conform to the power law (4.3). However, the range of distances of the shells for which this is true coincides in Fig. 4.1 and Figure 5.4, being $40\,h^{-1} \lesssim r \lesssim 60\,h^{-1}$Mpc in the COMPOSITE sample and $30\,h^{-1} \lesssim r \lesssim 67\,h^{-1}$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[1].

The systematic boost offset is still visible in the innermost shells for the CF2 sample, yet the fit becomes worse than with the COMPOSITE sample due to a increased number of outliers. However, we will still check if the boost offset signature is unique to this direction. Figure 5.5 shows the value of f_p from (4.4) representing the best fit parameters for a systematic boost offset for boosts across the entire sky, with magnitude

[1]These data points are necessarily disregarded when we perform the logarithmic transformation to fit the power law (4.3), and so do not contribute to the stated uncertainties.

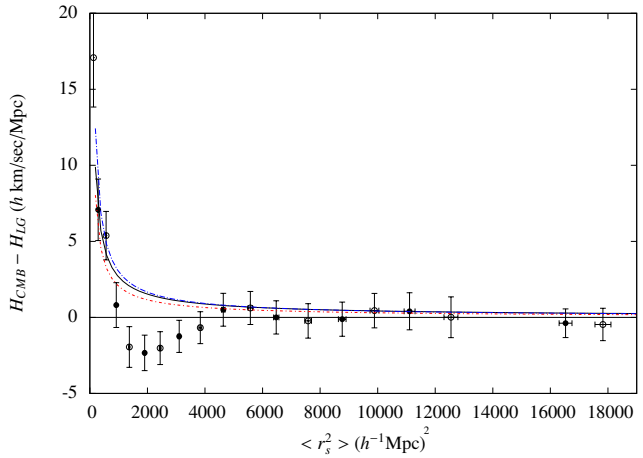


Figure 5.4: Best fit power law to the radial variation in the spherically averaged Hubble law in 11 shells for the CF2 galaxies data. The dashed blue curve is the best fit to primed shells only (empty circles), the dotted red curve is the best fit to unprimed shells only (filled circles) and the solid black curve is the best fit to all data points. The first data point – corresponding to unprimed shell 1 – is omitted, as it is off the scale.

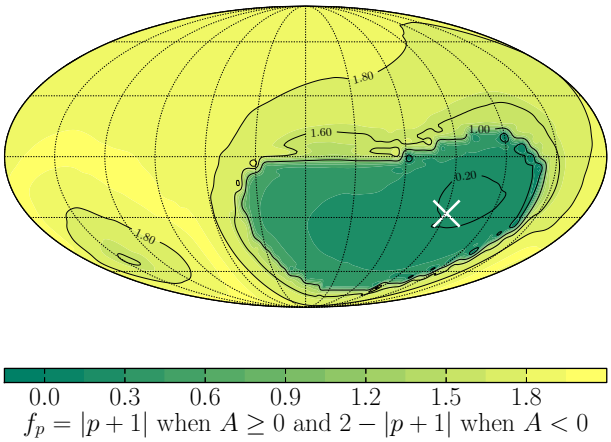


Figure 5.5: Best fit power law parameters to (4.3) across entire sky for boosts of $635\,\mathrm{km\,s^{-1}}$ from the CMB frame. The cross indicates the direction of the boost to the LG, which is also of magnitude $635\,\mathrm{km\,s^{-1}}$. In all figures, the galactic longitudes $\ell = 0°, 180°, 360°$ are on the right edge, centre and left edge respectively.

equal to that required for the boost from the CMB to the LG. We see results consistent with that found with the COMPOSITE sample, giving evidence that this is indeed not a random statistical outcome but a more physically meaningful observation.

5.3 Constraining a minimum Hubble expansion variation frame with CF2

We attempt to repeat a χ^2_a analysis on the CF2 catalogue. However, due to the huge variation evident in Figure 5.1 this is not possible. In fact, we find no frame in which χ^2_a approaches unity, or even within the same order of magnitude. We do observe a decrease in variation for boosts in the galactic plane. However, having begun such an investigation we found huge uncertainties such that little was to be gained. Thus without proper consideration of bias in the distances, we are not able to use the CF2 catalogue to constrain a frame of minimum Hubble expansion variation using the χ^2_a analysis of §4.3.1.

As seen in §5.2 we still see a signature of a systematic boost offset between the spherically averaged Hubble expansion in the LG and CMB frames of reference. This direction is similar to that found with the COMPOSITE sample. Calculating the best fit to the difference in Hubble expansion for frames boosted from the LG gives the sky maps in Figure 5.6. We see the direction which shows a best fit closest to the form of (4.2) is more well defined with this larger data set, and still aligns with that found using the COMPOSITE sample.

As discussed earlier, the systematic boost offset cannot be used to constrain a velocity due to the limited amount of spherically averaged values and large uncertainties in the COMPOSITE sample. We check that the direction of the boost with the best fit to a systematic boost offset is consistent with earlier findings for the COMPOSITE sample. For example, on the 200 km s^{-1} sky map the best fit value is in a direction $(l, b) = (55°, -5°)$, close to that found earlier, yet the value of $p = -0.92 \pm 0.75$ indicates that this result is subject to much greater statistical uncertainty. Thus, it appears the denser CF2 sample could potentially help to rule out frames that were thought possible minimum Hubble expansion variation frames with the COMPOSITE sample. Unfortunately, without a rigorous χ^2_a analysis to support this, and a resolution of the

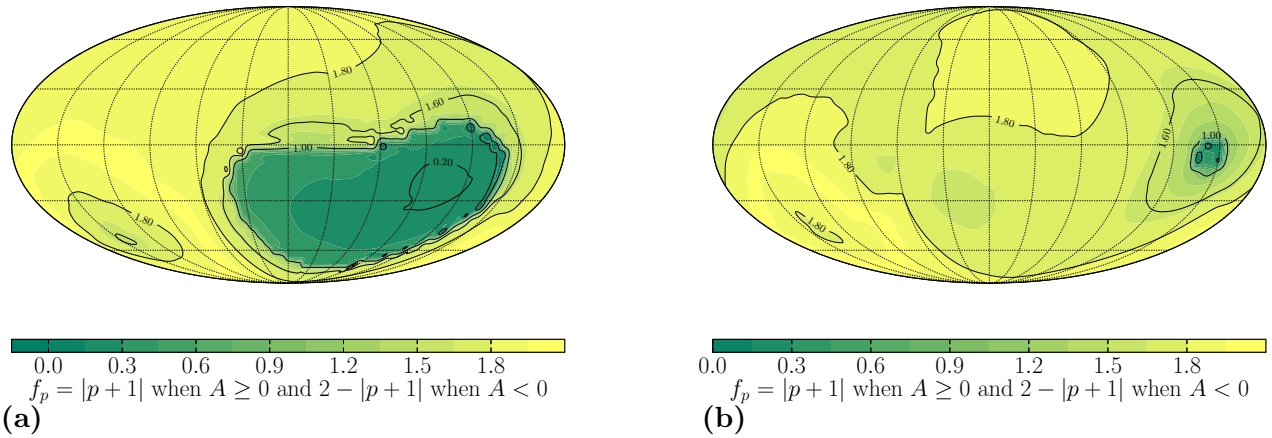


Figure 5.6: Best fit parameters to (4.3) for boosts from the LG of 200 km s^{-1} (a) and 450 km s^{-1} (b) with the CF2 data.

systematic problems in the distances, we cannot confirm that the LG is favoured as the standard of rest or any other frame using this raw data set. No doubt there is enough data present to make improvements over the COMPOSITE sample. However, all Malmquist biases first need to be treated in a manner similar to the treatment undertaken by Springob *et al.* [76].

CHAPTER 6

Conclusion

The aim of this investigation has been to determine the extent to which it is possible to define a standard of cosmological rest based on the variation of the Hubble expansion in a manner which does not make assumptions about the geometry of space below the scale of statistical homogeneity ($\lesssim 100\,h^{-1}$Mpc). We study the systematic variation that arises in a boost from a frame of reference in which the spherically averaged Hubble expansion would be most uniform. It is found that such a systematic variation exists between the spherically averaged Hubble expansion in the CMB and Local Group frames of reference. We extend this analysis to search for an improvement on the Local Group as the standard of rest. This is found to be consistent with a statistical analysis of the radial variations from a uniform expectation. The proposition made by Wiltshire *et al.* [73] that the Local Group is a more suitable rest frame than the CMB is consistent with our results. We have found that the Local Group is contained within a degenerate set of possible frames for the standard of rest. As the CMB frame is not contained within this degenerate set, we have been able to further rule it out as an appropriate standard of cosmic rest.

We also find the correlation of the CMB residual temperature dipole, conventionally associated with the motion of the Local Group, and the dipole observed in the angular Hubble expansion variation, is found to be optimal when monopole variation over the $40h^{-1}$Mpc to $60h^{-1}$Mpc range is greatest. This further supports the claim of Wiltshire *et al.* [73] that structure on this range is primarily responsible for the observed CMB temperature dipole anisotropy in a manner which cannot be reduced to a kinematic

origin.

The COMPOSITE [75] sample of 4,534 galaxies is used primarily, with the key tests repeated with the recently released *Cosmicflows*-2 sample of 8,162 galaxy, groups and cluster distances [68]. 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.* [73] and our Fig. 5.1. These result from differences in the treatment of the Malmquist bias between the SFI++ catalogue and CF2 catalogue, as previously noted by Watkins *et al.* [69]. 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.* [73] 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.* [76] 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 (4.2) noted by *et al.* [73]. Nothing in the Malmquist bias correction procedure of Springob *et al.* [76] 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.* [76]. Furthermore, the departure of the nonlinear expansion from the boost power law (4.3) that is seen when comparing Figures 4.1 and 5.4 is precisely what is to be expected if there are additional unaccounted uncertainties in individual CF2 distances as compared to the SFI++ ones:

the distance range of structures associated with the nonlinear expansion is broadened.

If Figures 5.1 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 $6.5 \pm 2.2\%$ [119].

Tully *et al.* [68] 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 [120] 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" [120]. This would appear to be the counterpart of the large monopole we observe in Figure 5.1 in our analysis. Hoffman *et al.* [120] 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 [73]. Thus large scale bulk flows are not our primary interest. Rather, we are interested in detecting the systematic monopole variance (4.2). 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 $\bar{H}_0$, is particularly sensitive to any distribution bias which follows from a number density, $N(r)$, with strong gradients. The bias effect in Figure 5.1 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 $635\,\mathrm{km\,s^{-1}}$ velocity attributed to the LG within the ΛCDM model, as has recently done by Hess & Kitaura [121] who used constrained

N-body simulations and nonlinear phase space reconstructions to arrive at a value $v_{\rm LG} = 685 \pm 137\,{\rm km\,s^{-1}}$. 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 [50, 122]. 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 monpole boost offset (4.2) 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. However, there is insufficient data in the COMPOSITE sample to break this degeneracy and thus either confirm or rule out the Local Group as this standard of rest. Since the direction of boosts between these frames is aligned with the plane of the galaxy, we suspect the degeneracy is due to a lack of data in the Zone of Avoidance. This hypothesis could potentially be tested on simulated data using exact solutions of Einstein's equations [123]. Although the larger *Cosmicflows*-2 sample will also suffer from a lack of data in the Zone of Avoidance, it is possible that the statistical constraints could be tightened given the increased size of the sample. However, it is first necessary to reduce the data in the manner of Springob *et al.* [76] to remove the inhomogeneous Malmquist bias which appears to be the source of the large discrepancies between *Cosmicflows*-2 and the COMPOSITE sample, which are evident in the SFI++ subsample as shown in Figure 5.2.

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.

At present we certainly cannot rule out the Local Group as a candidate for the frame of cosmic rest, but can offer a particular set of possible frames as alternatives. At present we have isolated a set of candidate reference frames for the frame of minimum Hubble expansion variation including the Local Group. The CMB rest frame is definitely ex-

cluded from the set of candidate rest frames. Our statistical analysis of the characteristic boost offset seen between the CMB and LG frames strengthens the claims of WSMW13. Given that the CMB frame is still the defacto choice for the cosmic rest frame, this conclusion would have a far reaching impact in many areas of cosmology.

Bibliography

[1] Kogut A, Lineweaver C, Smoot G F, *et al.* 1993 *Astrophys. J.* **419** 1

[2] Peebles P J E, D.T. Wilkinson D T 1968 *Phys. Rev.* **174** 2168

[3] Fixsen D J, Cheng E S, Gales J M, Mather J C, Shafer R A, Wright E L 1996 *Astrophys. J.* **473** 576

[4] Stewart J M, Sciama D W 1967 *Nature* **216**, 748

[5] Aghanim N *et al.*2014 *Astron. Astrophys.* **571** A27

[6] Herschel W 1785 *Phil. Transactions of the Royal Soc.* **75** 213

[7] Herschel W 1802 *Phil. Transactions of the Royal Soc.* **92** 477

[8] Einstein A 1917 *Sitz. Königlich Preuss. Akad. d. Wiss. Phys.-Math* **1** 142

[9] Friedmann A 1922 *Zeitschrift für Physik.* **10** 377

[10] Friedmann A 1924 *Zeitschrift für Physik.* **21** 326

[11] Lemaître, G 1927 *Ann. Soc. Sci. de Bruxelles* **A47** 29

[12] Hubble E 1929 *Proc. Natl. Acad. Sci.* **15** 168

[13] Milne E A 1935 *Nature* **130** 9

[14] Einstein A, de Sitter W 1932 *Proc. Nat. Acad. Sci. USA* **18** 213

[15] Estathiou G, Sutherland W J, Maddox S J 1990 *Nature* **348** 705

[16] Riess *et al.* 1998 *Astron. J.* **116** 1009

[17] Ade P A R *et al.* 2014 *Astron. Astrophys.* **571** A16

[18] Eisenstein D J, Zehavi D, Hogg D W, *et al.* 2005 *Astrophys. J.* **633** 560

[19] Cole S, Percival W J, Peacock J A, *et al.* 2005 *Mon. Not. R. Astr. Soc.* **362** 505

[20] Blake C. *et al.* 2011 *Mon. Not. R. Astr. Soc.* **415** 2876

[21] Reid B A, *et al.* 2010 *Mon. Not. R. Astr. Soc.* **404** 60

[22] Komatsu E, *et al.* 2009 *Astrophys. J. Suppl.* **180** 330

[23] Blake C. *et al.* 2011 *Mon. Not. R. Astr. Soc.* **415** 2892

[24] Hogg D W, Eisenstein D J, Blanton M R, Bahcall N A, Brinkmann J, Gunn J E, Schneider D P 2005 *Astrophys. J.* **624** 54

[25] Sylos Labini F, Vasilyev N L, Pietronero L, Baryshev Y V 2009 *Europhys. Lett.* **86** 49001

[26] Scrimgeour M, Davis T, Blake C, James J B, Poole G, Staveley-Smith L, Brough S, Colless M 2012 *Mon. Not. R. Astr. Soc.* **425** 116

[27] Clowes R G, Harris K S, Raghunathan S, Campusano L E, Soechting I K, Graham M J 2013 *Mon. Not. R. Astr. Soc.* **429** 2910

[28] Hoyle F, Vogeley M S 2002 *Astrophys. J.* **566** 641

[29] Hoyle F, Vogeley M S 2004 *Astrophys. J.* **607** 751

[30] Abazajian K N, Adelman-McCarthy J K, Agueros M A, Allam S S, Allende Prieto C, Anderson K S J, Anderson S F, Annis J, Bahcall N A *et al.* 2009 *Astrophys. J. Suppl.* **182** 543

[31] Pan D C, Vogeley M S, Hoyle F, Choi Y Y and Park C 2021 *Mon. Not. R. Astr. Soc.* **421** 926

[32] Colless M *et al.* 2001 *Mon. Not. R. Astr. Soc.* **328** 1039

[33] Campbell L A, Lucey J R, Colless M, *et al.* 2014 *Mon. Not. R. Astr. Soc.* **443** 1231

[34] Scrimgeour M, Davis T, Blake C *et al.* 2012 *Mon. Not. R. Astr. Soc.* **425** 116

[35] Gott J R, Juric M, Schlegel D, Hoyle F, Vogeley M, Tegmark M, Bahcall N, Brinkmann J 2005 *Astrophys. J.* **624** 463

[36] Lemaitre G 1933 *Ann. Soc. Sci. Bruxelles* **A53** 51

[37] Tolman R C, 1934 *Proc. Natl. Acad. Sci.* **20** 169

[38] Enqvist K, Mattsson T 2007 *J. Cosmol. Astropart. Phys.* **2** 019

[39] Szekeres P 1975 *Commun. Math. Phys.* **41** 55

[40] Schaap W 2007 PhD thesis (Rijksuniversiteit Groningen, The Netherlands)

[41] Buchert T 2000 *Gen. Relativ. Grav.* **32** 105

[42] Buchert T 2000 *Gen. Relativ. Grav.* **33** 1381

[43] Russ H and Soffel M H and Kasaiand M and Börner G 1997 *Phys. Rev.* **D 56** 2044

[44] Buchert T 2008 *Gen. Relativ. Grav.* **40** 467

[45] Zalaletdinov R M 1992 *Gen. Relativ. Grav.* **24** 1015

[46] Zalaletdinov R M 1993 *Gen. Relativ. Grav.* **25** 673

[47] Zalaletdinov R M 1997 *Bull. Astron. Soc. India* **25** 401

[48] Wiltshire D L, in S.E. Perez Bergliaffa and M. Novello (eds), *Proceedings of the XVth Brazilian School of Cosmology and Gravitation*, (Cambridge Scientific Publishers, 2015) in press; [arXiv:1311.3787]

[49] Wiltshire D L 2008 *Phys. Rev.* **D 78** 084032

[50] Wiltshire D L 2009 *Phys. Rev.* **D 80** 123512

[51] Wiltshire D L 2007 *New J. Phys.* **9** 377

[52] Wiltshire D L 2007 *Phys. Rev. Lett.* **99** 251101

[53] Partridge R B, Wilkinson D T 1967 *Phys. Rev. Lett.* **18** 557

[54] Ellis G F R in Bertotti B, de Felice F Pascolini A (eds) 1984 *General Relativity and Gravitation*, (Reidel, Dordrecht) pp. 215–288.

[55] Buchert T 2000 *Gen. Relativ. Grav.* **32** 105

[56] Buchert T 2008 *Gen. Relativ. Grav.* **40** 467

[57] Wiltshire D L 2011 *Class. Quantum Grav.* **28** 164006

[58] Leith B M, Ng S C C, Wiltshire D L 2008 *Astrophys. J.* **672** L91

[59] Smale P R, Wiltshire D L 2011 *Mon. Not. R. Astr. Soc.* **413** 367

[60] P R Smale 2011 *Mon. Not. R. Astr. Soc.* **418** 2779

[61] Duley J A G, Nazer M A, Wiltshire D L 2013 *Class. Quantum Grav.* **30** 175006

[62] Kashlinsky A, Atrio-Barandela F, Kocevski D and Ebeling H 2008 *Astrophys. J.* **686** L49

[63] Kashlinsky A, Atrio-Barandela F, Kocevski D and Ebeling H 2009 *Astrophys. J.* **691** 1479

[64] Ade P A R *et al.* 2014 *Astron. Astrophys.* **561** A97

[65] Keisler R 2009 *Astrophys. J.* **707** L42

[66] Atrio-Barandela F 2013 *Astron. Astrophys.* **557** A116

[67] Watkins R, Feldman H A, Hudson M J 2009 *Mon. Not. R. Astr. Soc.* **392** 743

[68] Tully R B, Courtois H M, Dophin A E, Fisher J R, Heraudeau P, Jacobs B A, Karachentsev I D, Makarov D, Makarova L, Mitronova S, Rizzi L, Shaya E J, Sorce J G, Wu P 2013 *Astron. J.* **146** 86

[69] Watkins R, Feldman H 2015 *Mon. Not. R. Astr. Soc.* **447** 132

[70] Nusser A, Davis M 2011 *Astrophys. J.* **736** 93

[71] Turnbull S J, Hudson M J, Feldman H A, Hicken M, Kirshner R P, Watkins R 2012 *Mon. Not. R. Astr. Soc.* **420** 447

[72] Bondi H 1947 *Mon. Not. R. Astr. Soc.* **107** 410

[73] Wiltshire D L, Smale P R, Mattsson T, Watkins R 2013 *Phys. Rev.* **88** 083529

[74] Watkins R, Feldman H A, Hudson M J 2009 *Mon. Not. R. Astr. Soc.* **392** 743

[75] Feldman H A, Watkins R, Hudson M J 2010 *Mon. Not. R. Astr. Soc.* **407** 2328

[76] Springob C M, Masters K L, Haynes M P, Giovanelli R, Marinoni C 2007 *Astrophys. J. Suppl.* **172** 599

[77] Malmquist K G 1922 *Medd. Lund Astron. Obs. Ser. II* **32** 64

[78] Teerikorpi P 1997 *Annual Rev. Aston. Astrophys.***35** 101

[79] Lynden-Bell D, Faber S M, Burstein D *et al.* 1988 *Astrophys. J.* **326** 19

[80] Hanski M 1999 *International Symposium on Astrophysics Research and Science Education* University of Notre Dame Press p.242

[81] Eddington A S 1914 *Stellar Movements and the Structure of the Universe* Macmillan, London

[82] Butkevich A G, Berdyugin A V and Teerikorpi P 2005 *Mon. Not. R. Astr. Soc.* **362** 321

[83] Strauss M A and Willick J A 1995 *Phys. Rep.* **261** 271

[84] Tully R B, Fisher J R 1977 *Astron. Astrophys.* **54** 661

[85] Aaronson M, Mould J R and Huchra J 1979 *Astrophys. J.* **229** 1

[86] Tutui Y *et al.* 2001 *Publ. Astron. Soc. Japan* **53** 701

[87] Hendry M A, Simmons J F L and Newsam A M 1993 *arXiv* 9310028v1

[88] Teerikorpi P 1987 *Astron. Astrophys.* **173** 39

[89] Schechter P 1980 *Astron. J.* **85** 801

[90] Willick J A 1994 *Astrophys. J. Suppl.* **92** 1

[91] Willick J A, Courteau S, Faber S M, Burstein D and Dekel A 1995 *Astrophys. J.* **446** 12

[92] Ekholm T, Teerikorpi P, Theureau G, Hanski M, Paturel G, Bottinelli L, Gouguenheim L 1999 *Astron. Astrophys.* **347** 99

[93] Tully R B and Pierce M J 2000 *Astrophys. J.* **533** 744

[94] Courtois H M and Tully R B 2012 *Astrophys. J.* **749** 174

[95] Sorce J G, Courtois H M, Tully R B, Seibert M, Scowcroft V, Freedman W L, Madore B F, Persson S E, Monson A, Rigby J 2013 *Astrophys. J.* **765** 94

[96] Tully R B, Shaya E J, Karachentsev I D, Courtois H M, Kocevski D D, Rizzi L, Peel A 2008 *Astrophys. J.* **676** 184

[97] Feast M W 1987 *The Observatory* **107** 185

[98] Dale D A, Giovanelli R, Haynes M P, Campusano L E, Hardy E 1999 *Astron. J.* **118** 1489

[99] Dale D A, Giovanelli R, Haynes M P, Hardy E, Campusano L E 1999 *Astron. J.* **118** 1468

[100] Teerikorpi P 1993 *Astron. Astrophys.* **280** 443

[101] Li N, Schwarz D J 2008 *Phys. Rev.* **D 78** 083531

[102] Freedman W L, Madore B F, Gibson B K, *et al.* 2001 *Astrophys. J.* **553** 47

[103] McClure M L, Dyer C C 2007 New Astron. **12** 533

[104] Fixsen D J, Cheng E S, Gales J M, Mather J C, Shafer R A, and Wright E L 1996 *Astrophys. J.* **473** 576

[105] Fixsen D J 2009 *Astrophys. J.* **707** 916

[106] Ade P A R 2014 *Astron. Astrophys.* **571** A23

[107] Hoftuft J *et al.* 2009 *Astrophys. J.* **699** 985

[108] Tegmark M, de Oliveira-Costa A, Hamilton A J S 2003 *Phys. Rev.* **D 68** 123523

[109] Eriksen H K, Banday A J, Gorski K M, Hansen F K, Lilje P B 2007 *Astrophys. J.* **660**, L81

[110] Hoftuft J, Eriksen H K, Banday A J, Gorski K M, Hansen F K, Lilje P B 2009 *Astrophys. J.* **699** 985

[111] Eriksen H K, Hansen F K, Banday A J, Gorski K M, Lilje P B 2004 *Astrophys. J.* **605** 14; (E) **609** 1198

[112] de Oliveira-Costa A, Tegmark M, Zaldarriaga M, Hamilton A 2004 *Phys. Rev.* **D 69** 063516

[113] Land K, Magueijo J 2005 *Phys. Rev. Lett.* **95** 071301

[114] Schwarz D J, Starkman G D, Huterer D, Copi C J 2004 *Phys. Rev. Lett.* **93** 221301

[115] Copi C J, Huterer D, Schwarz D J, Starkman G D 2006 *Mon. Not. R. Astr. Soc.* **367** 79

[116] Kim J, Naselsky P 2010 *Astrophys. J.* **714** L265

[117] York D, Evensen N M, Martnez M L, Delgado J D 2004 *Am. J. Phys.* **72** 367

[118] Gorski K M, Hivon E, Banday A J, Wandelt B D, Hansen F K, Reinecke M, Bartelman M 2005 *Astrophys. J.* **622** 759

[119] Zehavi I, Riess A G, Kirshner R P, Dekel A 1998 *Astrophys. J.* **503** 483

[120] Hoffman Y, Courtois H M, Tully R B 2015 arXiv:1503.05422

[121] Hess S, Kitaura F S 2014 arXiv:1412.7310.

[122] Nazer M A, Wiltshire D L 2015 *Phys. Rev.* **D**91 063519.

[123] Bolejko K, Nazer M A, Wiltshire D L In preparation

[124] York D 1969 *Earth and Planetary Science Letters* **5** 320

[125] Titterington D, Halliday A 1979 *Chem. Geol* **26** 183

CHAPTER 7

Appendix

7.1 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

$$\begin{aligned} y_t &= \beta_0 + \beta_1 x_t \\ (Y_t, X_t) &= (y_t, x_t) + (e_t, u_t) \end{aligned} \tag{7.1}$$

where (Y_t, X_t) are the observations, (y_t, x_t) are the true values and (e_t, u_t) are the measurement errors. We assume the measurement errors to be normally distributed with a covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{ee} & \sigma_{eu} \\ \sigma_{ue} & \sigma_{uu} \end{pmatrix}. \tag{7.2}$$

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

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 distance 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

$$[Y_t - (\beta_0 + \beta_1 x_t)]^2 + (X_t - x_t)^2 = e_t^2 + u_t^2 \tag{7.3}$$

but if the variances of e_t and u_t are different from unity this statistical distance becomes $\sigma_{ee}^{-1}e_t^2 + \sigma_{uu}^{-1}u_t^2$, and if these variances are correlated we must use the covariance matrix to give the "statistical" distance

$$(Y_t - \beta_0 - \beta_1 x_t, X_t - x_t)\,\Sigma^{-1}\,(Y_t - \beta_0 - \beta_1 x_t, X_t - x_t)^{\intercal}\,. \tag{7.4}$$

In our case we must determine the values of β_0 and β_1 that minimize (7.4). That is, we must find the values $(\hat{x}_t, \hat{y}_t)$ and $(\hat{\beta}_0, \hat{\beta}_1)$ that minimize this sum for the given observations. First we fix the x_t values by treating them as unknown constants in a standard linear regression of the form

$$\begin{bmatrix} Y_t - \beta_0 \\ X_t \end{bmatrix} = \begin{bmatrix} \beta_1 \\ 1 \end{bmatrix} x_t + \begin{bmatrix} e_t \\ u_t \end{bmatrix} \tag{7.5}$$

for which the generalized least squares estimator gives

$$\hat{x}_t = \left[(\beta_1, 1)\Sigma^{-1}(\beta_1, 1)^{\intercal}\right](\beta_1, 1)\Sigma^{-1}\,(Y_t - \beta_0, X_t)^{\intercal}\,. \tag{7.6}$$

Substitution of $\hat{x}_t$ into (7.4) gives

$$\frac{(Y_t - \beta_0 - \beta_1 X_t)^2}{\left(\sigma_{ee} - 2\beta_1\sigma_{eu} + \beta_1^2\sigma_{uu}\right)} \tag{7.7}$$

so that after summing over all N points we obtain

$$S = \frac{\sum_{i=1}^{N}(Y_t - \beta_0 - \beta_1 X_t)^2}{\left(\sigma_{ee} - 2\beta_1\sigma_{eu} + \beta_1^2\sigma_{uu}\right)}, \tag{7.8}$$

which is the expression to be minimized. York [124] gives the linear equation that minimizes (7.8) to be

$$\beta_1 = \frac{\sum_{i=1}^{N} Z_i^2 V_i \left[\frac{U_i}{\omega(Y_i)} + \frac{\beta_1 V_i}{\omega(X_i)} - \frac{\gamma_i V_i}{\alpha_i}\right]}{\sum_{i=1}^{N} W_i^2 U_i \left[\frac{U_i}{\omega(Y_i)} + \frac{\beta_1 V_i}{\omega(X_i)} - \frac{\beta_1\gamma_i U_i}{\alpha_i}\right]} \tag{7.9}$$

where

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

Clearly (7.9) requires an iterative process to find β_1 which begins with an initial guess which may be found from performing a standard linear regression assuming the X_i to be exact. After β_1 is obtained the value of β_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, σ_{β_0} and σ_{β_1}, are [125]

$$\sigma_{\beta_0}^2 = \frac{\sum Z_i x_i^2}{\left(\sum Z_i x_i^2\right)\left(\sum Z_i\right) - \left(\sum Z_i x_i\right)^2}, \tag{7.10}$$

$$\sigma_{\beta_1}^2 = \frac{\sum Z_i}{\left(\sum Z_i x_i^2\right)\left(\sum Z_i\right) - \left(\sum Z_i x_i\right)^2}. \tag{7.11}$$

We now return to the transformed model from (4.2) which takes on the form

$$\log(\delta H_i) = p \log(\langle r_i^2 \rangle) + \log\left(\frac{v^2}{2H_0}\right) \tag{7.12}$$

such that we may identify $y_i = \log(\delta H_i)$, $x_i = \langle r_i^2 \rangle$, and $(\beta_0, \beta_1) = \left(p, \frac{v^2}{2H_0}\right) = (p, A)$.