Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **524,** 1885–1892 (2023) https://doi.org/10.1093/mnras/stad1984
Advance Access publication 2023 July 3

Analysing the large-scale bulk flow using cosmicflows4: increasing tension with the standard cosmological model

Richard Watkins,[1]★ Trey Allen,[1] Collin James Bradford,[1] Albert Ramon, Jr,[1] Alexandra Walker,[1] Hume A. Feldman ,[2]★ Rachel Cionitti,[2] Yara Al-Shorman,[2] Ehsan Kourkchi[3] and R. Brent Tully[3]★

[1]*Department of Physics, Willamette University, Salem, OR 97301, USA*
[2]*Department of Physics & Astronomy, University of Kansas, Lawrence, KS 66045, USA*
[3]*Institute for Astronomy, University of Hawaii, Honolulu, HI 96822, USA*

Accepted 2023 June 27. Received 2023 June 27; in original form 2023 February 7

ABSTRACT

We present an estimate of the bulk flow in a volume of radii 150−200 h^{-1} Mpc using the minimum variance method with data from the *CosmicFlows-4* (CF4) catalogue. The addition of new data in the CF4 has resulted in an increase in the estimate of the bulk flow in a sphere of radius 150 h^{-1} Mpc relative to the *CosmicFlows-3* (CF3). This bulk flow has an ~ 0.015 per cent chance of occurring in the standard cosmological model with cosmic microwave background derived parameters. Given that the CF4 is deeper than the CF3, we were able to use the CF4 to accurately estimate the bulk flow on scales of 200 h^{-1} Mpc (equivalent to 266 Mpc for Hubble constant $H_0 = 75$ km s^{-1} Mpc^{-1}) for the first time. This bulk flow is in even greater tension with the standard model, having $\sim 1.5 \times 10^{-4}$ % probability of occurring. To estimate the bulk flow accurately, we introduce a novel method to calculate distances and velocities from distance moduli that is unbiased and accurate at all distances. Our results are completely independent of the value of H_0.

Key words: galaxies: kinematics and dynamics – galaxies: distances and redshifts – galaxies: statistics – (cosmology:) cosmological parameters – (cosmology:) large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

In an expanding universe, the observed redshift of an object at cosmological distances arises from two separate and independent effects. The first is due to the expansion of the universe and is proportional to the distance to the object (Hubble 1929). The other is due to the local (peculiar) velocity that is determined only by the mass distribution around the object. Peculiar velocities (Rubin & Ford 1970; Zeldovich & Sunyaev 1980; Andernach & Zwicky 2017) have been used as a probe of large-scale structure, as they provide information about density perturbations, and hence the mass distribution, on scales of and larger than the effective depths of surveys. To study the mass distribution, one must take into account redshift distortions (Hamilton 1998) that arise from peculiar velocities and thus bias results from redshift surveys. Furthermore, peculiar velocity studies provide a mechanism to identify the possible sources of gravitation attraction in large volumes (e.g. Jacoby et al. 1992; Willick 1994; Strauss & Willick 1995). Many groups have surveyed (e.g. Rubin & Ford 1970; Rubin et al. 1976; Dressler et al. 1987; Lauer & Postman 1994; Riess, Press & Kirshner 1995; Zaroubi et al. 2001; Kashlinsky et al. 2008; Tully et al. 2013; Springob et al. 2014; Tully, Courtois & Sorce 2016) and analysed (e.g. Feldman & Watkins 1994, 1998, 2008; Nusser & Davis 1995, 2011; Watkins & Feldman 1995, 2007, 2015a, b; Feldman et al. 2003; Hui & Greene 2006; Sarkar, Feldman & Watkins 2007; Tully et al. 2008; Abate & Erdoğdu 2009; Watkins, Feldman & Hudson 2009; Feldman, Watkins & Hudson 2010; Davis et al. 2011; Nusser, Branchini & Davis 2011; Agarwal, Feldman & Watkins 2012; Macaulay et al. 2012; Turnbull et al. 2012; Rathaus, Kovetz & Itzhaki 2013; Nusser 2014; Carrick et al. 2015; Nusser 2016; Hellwing et al. 2017; Peery, Watkins & Feldman 2018; Wang et al. 2018) peculiar velocity catalogues in the last half a century.

To construct a peculiar velocity catalogue, one must estimate both the redshift of an object and its distance, usually in the form of the distance modulus. The redshift to galaxies is a reasonably straightforward and accurate measurement, usually using spectroscopic data (e.g. Geller & Huchra 1989). However, estimating the distance modulus to faraway galaxies is difficult, expensive, and requires a great deal of telescope time. The most common distance indicators, the Tully–Fisher (TF, Tully & Fisher 1977) and the Fundamental Plane (FP, Dressler 1987) relations, have large uncertainties, around 0.4 in the distance modulus for individual galaxies. Other, more accurate, distance indicators exist, such as Type Ia supernova (SNIa, Phillips 1993), Cepheids (Leavitt & Pickering 1912), tip of the red giant branch (Lee, Freedman & Madore 1993), and surface brightness fluctuations (Tonry & Schneider 1988) among others. However, objects with distances measured with these more accurate methods typically make up only a small fraction of most peculiar velocity catalogues.

★ E-mail: rwatkins@willamette.edu (RW); feldman@ku.edu (HAF); rtully@hawaii.edu (RBT)

Peculiar velocities of individual objects in a catalogue can be combined to find the average velocity (bulk flow) of large volumes that contain some or all of the catalogue objects. The value of the bulk flow of a volume of some radius estimated from a catalogue can then be compared to that predicted by the standard model of cosmology (ΛCDM, Planck Collaboration 2016). It is important to remember that theoretical models predict only the variance of the bulk flow components, since in a homogeneous and isotropic universe the average bulk flow should vanish. This means that adding new distances in order to measure the bulk flow on a given scale more accurately does not usually improve the constraint that the bulk flow can put on models unless the bulk flow is either increased or already larger than expected. However, acquiring more data in order to make an accurate estimation of the bulk flow in a larger volume has the potential to strengthen constraints, since the ΛCDM model suggests that the variance for the bulk flow of a volume decreases with radius. At a large enough radius, even a modest bulk flow can be in tension with the expected scale of the bulk flow.

In Section 2, we discuss our unbiased distance and velocity estimators; in Section 3, we present our methodology for estimating the bulk flow in a spherical volume using radial peculiar velocities; in Section 4, we describe the *CosmicFlows-4 (CF4)* catalogue that we are using in our analysis; in Section 5, we discuss how we locate catalogue objects in space; in Section 6, we show how we estimate the bulk flow from the CF4 using the minimum variance (MV) method; in Section 7, we present the results; we discuss our results in Section 8.

2 UNBIASED DISTANCE AND VELOCITY ESTIMATORS

Distances to galaxies are measured via distance moduli, μ, with Gaussian distributed measurement errors σ_μ. [Extensive discussions of errors in the most common distance indicators in the CF4 can be found in (Kourkchi et al. 2020a, b, 2022) and (Howlett et al. 2022) for the Tully–Fisher and Fundamental Plane estimators, respectively.] The relationship between measured distance modulus and actual distance is

$$\mu = 5\log_{10}(d) + 25 + \delta, \tag{1}$$

where d is the distance in Mpc and δ is a measurement error drawn from a Gaussian distribution of width σ_μ. We can use equation (1) to express an estimated distance $d_{\rm est}$, which includes the effect of measurement noise, in terms of the true distance d and the measurement noise δ,

$$\begin{aligned} d_{\rm est} &= 10^{\frac{\mu}{5}-5} = 10^{5\log_{10}(d)/5}10^{\delta/5} \\ &= de^{\delta\ln(10)/5} = de^{\delta\kappa}, \end{aligned} \tag{2}$$

where $\kappa \equiv \ln(10)/5$. Averaging over measurement errors δ we find the well-known result that $d_{\rm est}$ is biased, in that

$$\langle d_{\rm est}\rangle = d\langle e^{\delta\kappa}\rangle \neq d. \tag{3}$$

Using biased distance estimates results in similarly biased peculiar velocities. Several methods have been proposed to deal with this bias (e.g. Watkins & Feldman 2015b). Here, we take the novel approach of calculating the bias exactly so that it can be corrected for. Correcting for this bias is important in that it could manifest as a spurious large-scale flow that could affect our bulk flow estimates.

Assuming that δ is drawn from a Gaussian distribution of width σ_μ, the probability density $P(\delta)$ of the measurement error δ is given by

$$P(\delta) = \frac{1}{\sqrt{2\pi}\sigma_\mu}e^{-\delta^2/2\sigma_\mu^2}. \tag{4}$$

We can thus write the average in equation (3) as

$$\begin{aligned} \langle d_{\rm est}\rangle &= d\frac{1}{\sqrt{2\pi}\sigma_\mu}\int_{-\infty}^{\infty} e^{\delta\kappa}e^{-\delta^2/2\sigma_\mu^2}\,d\delta \\ &= d\frac{1}{\sqrt{2\pi}\sigma_\mu}\int_{-\infty}^{\infty} \exp\left(\delta\kappa - \delta^2/2\sigma_\mu^2\right)\,d\delta. \end{aligned} \tag{5}$$

The integral can easily be evaluated by completing the square, giving

$$\langle d_{\rm est}\rangle = d\ \exp\left((\kappa\sigma_\mu)^2/2\right). \tag{6}$$

Thus, we can correct for the bias by multiplying our distance estimates calculated from the distance modulus by a factor, so that the corrected distance estimate $d_{\rm c}$ given by

$$d_{\rm c} = 10^{\frac{\mu}{5}-5}\ \exp\left(-(\kappa\sigma_\mu)^2/2\right), \tag{7}$$

is unbiased, in that $\langle d_{\rm c}\rangle = d$.

To find the uncertainty in the distance estimate we need to find $\langle d_{\rm c}^2\rangle$. This is given by

$$\begin{aligned} \langle d_{\rm c}^2\rangle &= d^2\frac{f^2}{\sqrt{2\pi}\sigma_\mu}\int_{-\infty}^{\infty} e^{2\delta\kappa}e^{-\delta^2/2\sigma_\mu^2}\,d\delta \\ &= d^2\frac{f^2}{\sqrt{2\pi}\sigma_\mu}\int_{-\infty}^{\infty} \exp\left(2\delta\kappa - \delta^2/2\sigma_\mu^2\right)\,d\delta, \end{aligned} \tag{8}$$

where we have defined the factor f to be

$$f \equiv \exp\left(-(\kappa\sigma_\mu)^2/2\right). \tag{9}$$

Completing the square as we did before gives

$$\langle d_{\rm c}^2\rangle = d^2f^2f^{-4} = d^2f^{-2}, \tag{10}$$

thus, the uncertainty in the distance estimate $d_{\rm c}$ is given by

$$\sigma_{\rm c}^2 = \langle d_{\rm c}^2\rangle - \langle d_{\rm c}\rangle^2 = d^2(f^{-2} - 1). \tag{11}$$

The argument of the exponential in the factor f is much less than 1 for typical uncertainties σ_μ, so we can use the Taylor series expansion $e^x \approx 1 + x$, giving

$$f^{-2} = \exp\left(2(\kappa\sigma_\mu)^2/2\right) \approx 1 + (\kappa\sigma_\mu)^2, \tag{12}$$

so that

$$\sigma_{\rm c} \approx \kappa\sigma_\mu d. \tag{13}$$

Thus, distance uncertainties grow approximately linearly with distance and can be expressed as a percentage of distance.

The unbiased distance estimate $d_{\rm c}$ can be used to calculate peculiar velocities using

$$v_{\rm c} = cz_{mod} - H_0d_{\rm c}; \tag{14}$$

where H_0 is the Hubble constant, c is the speed of light, and $z_{\rm mod}$ is the redshift modified to account for the deviation from a linear Hubble's Law

$$z_{\rm mod} = z\left(1 + \frac{1}{2}(1-q_0)z - \frac{1}{6}(1 - q_0 - 3q_0^2 + j_0)z^2\right), \tag{15}$$

where z is the measured redshift and q_0 and j_0 are the deceleration and jerk parameters. The use of modified redshift is important only for the more distant objects in the CF4 catalogue. Peculiar velocities calculated in this way are also unbiased, in that

$$\begin{aligned} \langle v_{\rm c}\rangle &= \langle cz_{\rm mod} - H_0d_{\rm c}\rangle = cz_{\rm mod} - H_0\langle d_{\rm c}\rangle \\ &= cz_{\rm mod} - H_0d = v. \end{aligned} \tag{16}$$

However, the errors in peculiar velocities estimated in this way are not Gaussian distributed.

An alternative approach to calculating unbiased peculiar velocities was given in Watkins & Feldman (2015b). In terms of the distance modulus μ, that estimator can be written as

$$v_{\rm f} = cz_{\rm mod} \ln(10) \left(\log(cz_{\rm mod}/H_0) - \frac{\mu - 25}{5} \right). \quad (17)$$

This formula gives peculiar velocities with Gaussian distributed errors; however, it has the disadvantage of only being unbiased for galaxies where the true velocity v is much less than the redshift cz. The uncertainty in this estimator is the same as that given in equation (13).

To use our method, we also need to assign a peculiar velocity to each group. For nearby objects, the estimate $v_{\rm c}$, (equation 14), is preferred over $v_{\rm f}$ (equation 17), since the formula used to calculate $v_{\rm f}$ is not accurate at small distances. However, at large distances the estimate $v_{\rm f}$ might be preferred since it has Gaussian distributed errors and thus is a better match to the assumptions behind our method. However, we experimented with using $v_{\rm c}$ at small distances and then transitioning to $v_{\rm f}$ at larger distances, and it made virtually no difference to our results. It appears that, at least for the current work, that as long as the velocities are unbiased, velocity errors average out effectively regardless of their distribution. Thus in this work, we use equation (14) exclusively to estimate peculiar velocities.

3 ESTIMATING THE BULK FLOW FOR A SPHERICAL VOLUME USING RADIAL PECULIAR VELOCITIES

The bulk flow is one of the most basic ways to characterize the peculiar velocity field in the local Universe. The bulk flow is defined as the average velocity in a region, usually taken to be a spherical volume V surrounding the Milky Way galaxy with radius R:

$$U_i = \frac{1}{V} \int_V v_i \, d^3x, \quad (18)$$

where v_i for $i = 1, 2, 3$ are the cartesian components of the full three-dimensional velocity field and $V = \frac{4}{3}\pi R^3$ is the volume of the sphere. If we imagine the peculiar velocity field as being the sum of waves of various wavelengths λ, then the bulk flow averages out waves with $\lambda \lesssim R$ and thus mostly reflect the amplitudes of waves with $\lambda > R$. Since homogeneity and isotropy requires the power spectrum to go to zero as λ goes to infinity, the bulk flow should decrease with increasing radius R. Thus, the bulk flow is a probe of the power spectrum on scales that are difficult to probe using redshift surveys, and the measurement of the bulk flow provides an important test of the standard cosmological model.

One difficulty in measuring the bulk flow as defined in equation (18) is that we can only measure the radial component of the velocity field. However, if we make the assumption that the velocity field is curl-free ($\vec{\nabla} \times \vec{\mathbf{v}} = 0$), as velocities generated by gravity must be, then the radial velocity carries the same information as the full three-dimensional field, and the bulk flow can be written as a weighted average of the *radial peculiar velocity*. The following derivation is taken from Nusser (2014, 2016). We begin by writing the velocity field in terms of a scalar potential ϕ,

$$\vec{\mathbf{v}} = -\vec{\nabla}\phi. \quad (19)$$

In terms of ϕ the bulk flow becomes

$$U_i = -\frac{1}{V} \int_V \nabla_i \phi \, d^3x = -\frac{1}{V} \int_V \vec{\nabla} \cdot (\hat{\mathbf{x}}_i \phi) \, d^3x, \quad (20)$$

where $\hat{\mathbf{x}}_i$ are the cartesian unit vectors. The divergence theorem can then be used to convert the volume integral to an integral over the surface S of the region,

$$U_i = -\frac{1}{V} \int_S \hat{\mathbf{r}} \cdot \hat{\mathbf{x}}_i \phi \, d\Omega = -\frac{1}{V} \int_S \hat{\mathbf{r}}_i \phi \; R^2 \, d\Omega, \quad (21)$$

where $\hat{\mathbf{r}}$ is the radial unit vector with $\hat{r}_i$ being its ith component. This integral picks out the dipole contribution to ϕ. We can make this explicit by expanding ϕ's angular dependence in terms of real-valued spherical harmonics

$$\phi(\mathbf{r}) = \phi_o(r) + \sum_i \phi_i(r) \hat{\mathbf{r}}_i + \sum_{l>1,m} \phi_{l,m}(r) Y_{l,m}(\theta, \phi), \quad (22)$$

where the second term is a sum over the $l = 1$ real valued spherical harmonics. Plugging this into equation (21) and using the orthogonality of the spherical harmonics we obtain a simple expression for the bulk flow in terms of the value of ϕ_i on the surface of the region,

$$U_i = -\phi_i(R)/R. \quad (23)$$

Nusser (2014) showed that the bulk flow can also be written as an integral of the radial peculiar velocity, s, times $\hat{r}_i$ and a radial weight function $w(r)$,

$$\frac{1}{V} \int_V w(r) s \, \hat{r}_i \, d^3x = -\frac{1}{V} \int_V w(r) \frac{\partial}{\partial r} \phi \, \hat{r}_i \, r^2 dr d\Omega \quad (24)$$

$$= -\frac{1}{R^3} \int_0^R w(r) \frac{\partial}{\partial r} \phi \; r^2 dr, \quad (25)$$

where we have plugged in equation (22) and used the orthogonality of the spherical harmonics to do the integral over angles (see also Peery et al. 2018). We see that if $w = R^2/r^2$, then we can use the fundamental theorem of calculus to show that

$$\frac{1}{V} \int_V w(r) s \, \hat{r}_i \, d^3x = -\phi(R)/R = U_i, \quad (26)$$

thus showing that the bulk flow can be expressed as a weighted average of the *radial* velocity with weights proportional to $\hat{r}_i/r^2$.

While the interpretation that the $\hat{r}_i/r^2$ weighted averages of the radial velocities are equivalent to the average of the full three-dimensional velocity is interesting, the validity of our analysis does not depend on the assumption of a curl-free velocity field. Even if we do not make this assumption, our bulk flow component estimates are still orthogonal linear combinations of radial peculiar velocities whose expectations in the standard cosmological model can be calculated using linear theory. Beyond their interpretation, these moments have been shown to probe the power spectrum in a way that cleanly separates large-scale and small-scale motions (Peery et al. 2018).

4 DATA

In this paper we analyse the group version of the *CosmicFlows-4* catalogue (Tully et al. 2023, hereafter CF4). The catalogue gives modified redshift $cz_{\rm mod}$, distance modulus μ, uncertainty in distance modulus σ_μ, Galactic longitude l, and Galactic latitude b, for over 38 000 groups and individual galaxies. The majority of distance moduli in the CF4 are estimated using the TF (Tully & Fisher 1977) and the FP (Dressler 1987) relations, with typical uncertainties of around 0.4 for individual galaxies. The remaining distance moduli in CF4 are estimated using more accurate methods such as SNIa (Phillips 1993), Cepheids (Leavitt & Pickering 1912), or tip of the red giant branch (Lee et al. 1993). When there are multiple measured

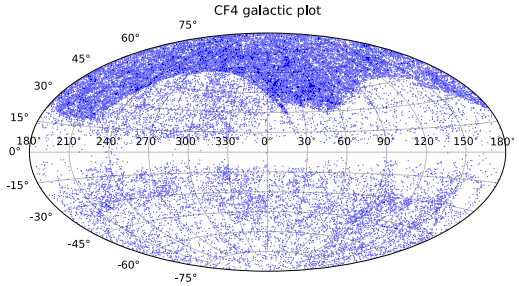

Figure 1. Angular distribution of CF4 galaxies in Aitoff–Hammer projection Galactic coordinates. Note the large number of galaxies in the Northern Galactic hemisphere from the SDSS galaxy sample.

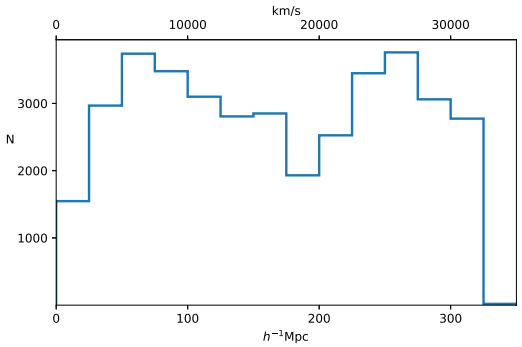

Figure 2. The radial distribution of CF4 galaxies and groups.

distance moduli for galaxies in a group, the group distance modulus is a weighted average and can have much smaller uncertainties. The zero-point calibration of the CF4 catalogue gives distances consistent with a value of $H_0 = 74.6$ km s^{-1} Mpc^{-1} (Tully et al. 2023); however, it is important to note that our results are independent of the choice of zero-point of the catalogue.

The CF4 adds two major new data sets to the *CosmicFlows-3* (CF3, Tully et al. 2016): (a) distance moduli of $\simeq$ 34 000 galaxies from the Sloan Digital Sky Survey (SDSS; York et al. 2000; Howlett et al. 2022) using the Fundamental Plane method and (b) $\simeq$ 10 000 distance moduli of spiral galaxies obtained using the Baryonic Tully Fisher Relation (Kourkchi et al. 2022). The addition of the SDSS galaxies greatly expands the depth of the CF4 relative to the CF3; however, the SDSS galaxies are all in the Northern Galactic hemisphere, so that the increase in depth is highly anisotropic. In Fig. 1, we show the angular distribution of the CF4 groups and galaxies in Galactic coordinates. In Fig. 2, we show the radial distribution of the CF4 objects. In Fig. 3 we show the radial distribution, but this time subdivided into objects in the positive and negative sides of each Galactic cartesian coordinate; this plot shows that while in the x and y directions groups and galaxies are fairly evenly distributed, in the z direction the distribution of objects on the Northern Galactic hemisphere is quite different than that in the Southern hemisphere. In particular, there are many more objects in the north and they are generally much deeper. The anisotropic distribution of the CF4 galaxies makes it particularly important to use a formalism such as the MV method that allows us to estimate a physically relevant bulk flow; otherwise the bulk flow would be difficult to interpret and not comparable to other results.

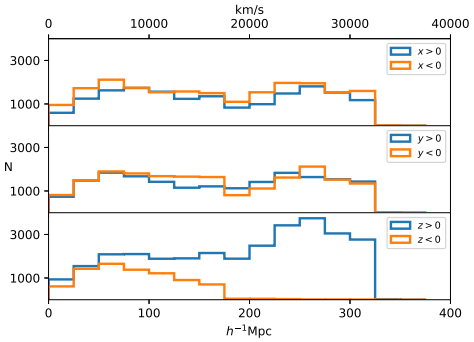

Figure 3. The redshift distribution of CF4 galaxies in Galactic coordinates, dividing galaxies into those with positive and negative values of each cartesian coordinate. We see that while galaxies are distributed fairly evenly in x and y, in the z direction the distribution is markedly different in the positive and negative directions. This is mostly due to the SDSS galaxy sample.

5 LOCATING GALAXIES WITHIN THE VOLUME

In order to theoretically model the data in the CF4 we need to first locate the groups in space. While both galaxies and groups have accurate angular positions, distance estimates to galaxies are less accurate and require more thought. For nearby objects, equation (7) provides an unbiased, accurate distance estimate. However, as mentioned above, distance uncertainties grow linearly with distance, so that distances become increasingly uncertain as we consider more distant groups. An alternative distance estimate d_z is obtained from redshift information,

$$d_z = cz_{\rm mod}/H_0. \tag{27}$$

Redshifts have small measurement uncertainties; the main source of error in d_z is due to peculiar velocities, which contribute to the redshift through the Doppler effect. For distant galaxies, the error in redshift distance due to peculiar velocities is smaller than the error in the distance obtained from the distance modulus.

In previous work we have used d_z to locate galaxies in space in order to avoid the large errors in object positions that would result from using distance estimates for distant objects; this does not break the self-consistency of our measurements of peculiar velocity, since only measured distances are used in those calculations. However, the use of d_z has the disadvantage of being less accurate for nearby objects and not making sense for objects with with negative redshifts. Ideally, we would like to use distance estimates at nearby distances and redshift distance for farther away objects.

To address this challenge, we use a novel way of determining the distance to galaxies to locate their positions in space. We use a distance estimate that is a weighted average of the unbiased distance d_c calculated from the distance modulus, given in equation (7), and the redshift distance d_z given in equation (27), using weights that minimize the uncertainty in the average. Our distance estimates are thus given by

$$d_{\rm est} = \frac{d_c\sigma_c^{-2} + d_z\sigma_z^{-2}}{\sigma_c^{-2} + \sigma_z^{-2}}, \tag{28}$$

where the distance uncertainty σ_c is given in equation (13) and $\sigma_z = \sigma_v/H_0$, where σ_v is the velocity dispersion of galaxies, which is about 300 km s^{-1}. Practically speaking, $d_{\rm est}$ corresponds to d_c at small distances and to d_z at large distances, making a smooth

transition between the two for intermediate galaxies. The exact value chosen for σ_v (and thus σ_z) does not have a significant effect on our results; indeed, since we are looking at large scale flows, this treatment in general does not have a significant effect on our results beyond allowing us to include a small number of objects with negative redshifts.

6 ESTIMATING THE BULK FLOW USING THE MV METHOD

Estimating the bulk flow is complicated by the fact that galaxies are both unevenly distributed in the volume and have varying uncertainties that generally increase with distance. Simply calculating the weighted average of the radial velocities of galaxies within a given volume using weights proportional to $\hat{r}/r^2$ will not in general result in a good approximation to the integral over the velocity field given in equation (26). Instead, we imagine a more general weighted average

$$u_i = w_{i,n} s_n, \tag{29}$$

where repeated indices are summed over, the s_n are radial peculiar velocities of groups or individual galaxies in a survey, and $w_{i,n}$ are weights designed so that u_i gives the best possible estimate of the bulk flow U_i in a volume of radius R. Here the weights $w_{i,n}$ should account for both the distribution of galaxies and the uncertainties of their peculiar velocity measurements. To determine the optimal weights we will use the MV method, developed in Watkins et al. (2009); Feldman et al. (2010); Peery et al. (2018). The idea of the MV method is to calculate the weights that minimize the theoretical average square difference $\langle (u_i - U_i)^2 \rangle$ between the estimated bulk flow components in equation (29), u_i, and the bulk flow components calculated for a theoretical ideal survey, U_i, consisting of exact peculiar velocity measurements measured at uniformly distributed points weighted by $1/r^2$. In practice, this means generating a random set of N_p points selected so that there are equal number of points per radial shell (see Peery et al. 2018, for details).

The method allows for constraints to be placed on the weights so that estimated bulk flow components are independent of the value of the Hubble constant. This is particularly important given that we are in an era where the results of different methods of measurement of the Hubble constant are in tension with each other (see e.g. Freedman 2017; Dainotti et al. 2021; Valentino et al. 2021; Dainotti et al. 2022; Riess et al. 2022).

Once the weights for the bulk flow component estimates are determined, we can calculate the tensor angle-averaged window function $W^2_{ij}(k)$ for the components, given by

$$W^2_{ij}(k) = w_{i,n} w_{j,m} f_{nm}(k), \tag{30}$$

where repeated indices are summed over, and

$$f_{nm}(k) = \int \frac{d^2\hat{k}}{4\pi} (\hat{\mathbf{r}}_n \cdot \hat{\mathbf{k}})(\hat{\mathbf{r}}_m \cdot \hat{\mathbf{k}}) e^{ik\,\hat{\mathbf{k}}\cdot(\mathbf{r}_n - \mathbf{r}_m)} \tag{31}$$

(Watkins et al. 2009). The diagonal elements of the tensor window function $W^2_{ii}(k)$ tell us which scales contribute to the bulk flow components. In addition, a comparison of the window function for the bulk flow estimate u_i with the window function for the ideal bulk flow U_i can indicate how well our data can estimate the bulk flow of an ideal survey. We discuss this comparison in more detail when we present our results in Section 7.

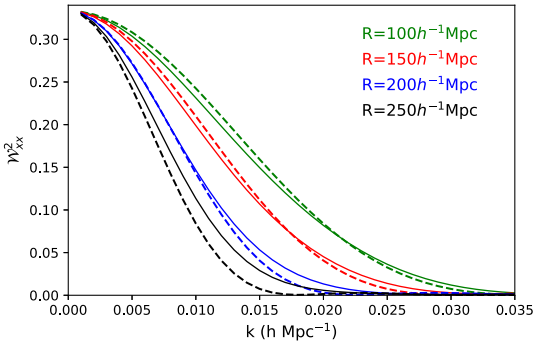

Figure 4. Window functions for estimates of the x component of the bulk flow for various radii R calculated for the CF4 catalogue. The dashed lines show the window functions for an ideal survey of the same radius.

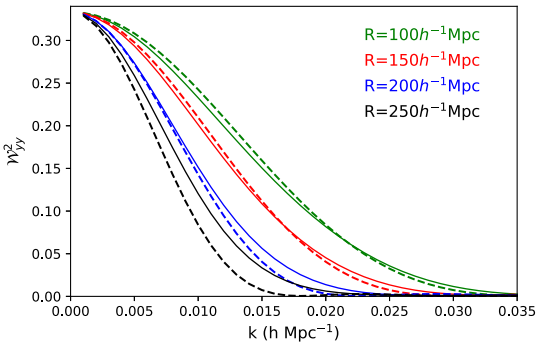

Figure 5. Same as Fig. 4 for the y component.

Given the window functions in equation (30) we can calculate the theoretical covariance matrix for the components of the bulk flow

$$R_{ij} = \langle u_i u_j \rangle = \frac{H_0^2 \Omega_m^{1.1}}{2\pi^2} \int dk\; P(k) W^2_{ij}(k) + \sum_n w_{i,n} w_{j,n} \left(\sigma_n^2 + \sigma_*^2\right), \tag{32}$$

where the window function $W^2_{ij}(k)$ is given in equation (30) (Kaiser 1988). Note that if distances are expressed in units of h^{-1} Mpc, the covariance matrix is independent of the value of H_0.

The covariance matrix can be used together with the bulk flow component estimates u_i to calculate a χ^2 for the three component degrees of freedom,

$$\chi^2 = u_i R^{-1}_{ij} u_j, \tag{33}$$

where repeated indices are summed over (Kaiser 1988). From this χ^2 distribution for three degrees of freedom we can find the probability of finding a value of χ^2 that is as large or larger than our calculated value.

7 RESULTS

As mentioned above, the CF4 is deeper than the CF3, allowing us to measure the bulk flow at greater depth, but also more anisotropic. In order to determine the maximum depth to which we can measure the bulk flow accurately, we need to examine the window functions of our bulk flow estimates. In Figs 4, 5, and 6 we show the window functions for the Galactic cartesian coordinates x, y, and z bulk flow estimates for various radii R. In the figures we see that the bulk flow estimate window functions match well with the ideal window

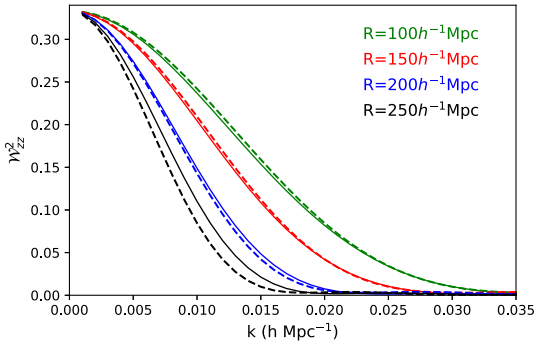

Figure 6. Same as Fig. 4 for the z component.

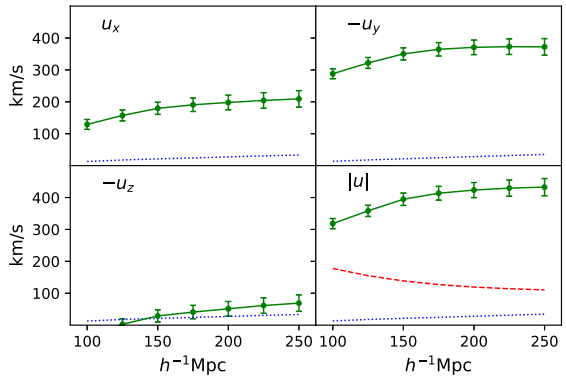

Figure 7. The green points with error bars show the bulk flow components and magnitude estimated from the CF4 catalogue as a function of radius R. The error bars indicate the uncertainty in the estimates due to measurement noise. The dotted blue lines show the theoretical standard deviation of the expected differences between the bulk flow estimates and the bulk flow from an ideal survey calculated using the cosmological standard model (not including measurement noise). The red dashed line indicates the theoretical expectation for the magnitude of the bulk flow calculated using the cosmological standard model.

functions out to a radius of about $R = 200\,h^{-1}$ Mpc, indicating that the estimates are probing the power spectrum in the same way. Beyond this radius, we start to see a deviation (albeit small) of the estimated window function from the ideal window function, indicating that the catalogue has insufficient information at the outer edge of the volume to accurately estimate the bulk flow on this scale. In this work we focus on estimating the bulk flow for $R = 150\,h^{-1}$ Mpc and $R = 200\,h^{-1}$ Mpc.

In Fig. 7 we show the bulk flow components and magnitude (with measurement noise error bars) calculated using the CF4 data as a function of the radius R. Included in the plots (dotted blue lines) is the standard deviation of the expected difference between the bulk flow estimates and the bulk flow from an ideal survey calculated using the cosmological standard model with the *Planck* central parameters (Planck Collaboration 2016); the total deviation from the ideal bulk flow is given by a quadrature sum of this value and the measurement noise. The dashed red line on the magnitude plot indicates the expectation for the bulk flow magnitude calculated using the cosmological standard model.

In Fig. 8 we show the χ^2 for the three bulk flow component degrees of freedom as a function of R calculated using the theoretical covariance matrix (equation 32). We also show the probability of finding a χ^2 value that is as large or larger. While we have focused on $R = 150\,h^{-1}$ Mpc and $R = 200\,h^{-1}$ Mpc in this paper, we can

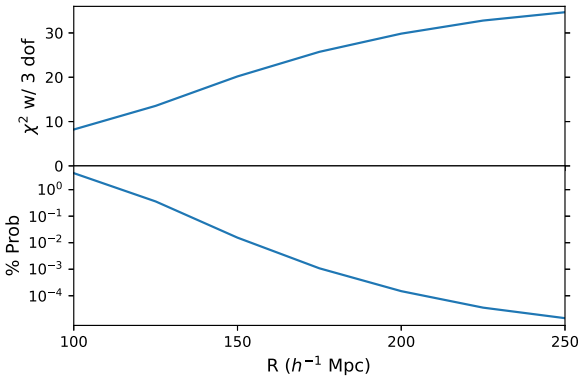

Figure 8. The χ^2 for the three bulk flow degrees of freedom as a function of radius R. Also shown in the corresponding probability to find a χ^2 value that is as large or larger.

Table 1. Summary of bulk flows for $R = 150\,h^{-1}$ Mpc and $R = 200\,h^{-1}$ Mpc. The uncertainties include both the theoretical difference between the estimate and the bulk flow from an ideal survey and the measurement noise.

	$R = 150\,h^{-1}$ Mpc	$R = 200\,h^{-1}$ Mpc
Expectation (km s^{-1})	139	120
Bulk flow (km s^{-1})	395 ± 29	427 ± 37
Direction	$l = 297^\circ \pm 4^\circ$	$l = 298^\circ \pm 5^\circ$
	$b = -4^\circ \pm 3^\circ$	$b = -7^\circ \pm 4^\circ$
χ^2 with 3 degrees of freedom	20.19	29.84
Probability	1.54×10^{-4}	1.49×10^{-6}

see from the figure that the probability of obtaining a bulk flow as large or larger continues to decrease as R increases beyond $200\,h^{-1}$ Mpc. However, as discussed above, one must take our bulk flow results beyond $200\,h^{-1}$ Mpc with a grain of salt, since at these radii the window functions for the bulk flow estimate are beginning to deviate from the ideal window functions. This is mostly due to the lack of data at large distances; without much new information, bulk flow estimates on these scales will seem to change more slowly with radius than the actual bulk flow.

8 DISCUSSION

As shown in Table 1, our bulk flow estimates in spheres of radii $R = 150\,h^{-1}$ Mpc and $R = 200\,h^{-1}$ Mpc have magnitudes in excess of 380 km s^{-1} and directions that are $\sim 20^\circ - 30^\circ$ away from the Shapley Concentration. We see that the probability of observing a bulk flow as large or larger for $R = 150\,h^{-1}$ Mpc is small, only about 0.015 per cent. This is significantly smaller than the probability of the CF3 bulk flow at this radius of 2.2 per cent (Peery et al. 2018); the addition of new data has not only decreased the uncertainty but also increased the estimate of the bulk flow, resulting in a tension with the standard cosmological model at this radius that is significantly stronger. Additionally, the fact that the CF4 is deeper than the CF3 allows us to accurately measure the bulk flow at larger radii. As shown in Table 1, the probability of obtaining a bulk flow as large or larger for $R = 200\,h^{-1}$ Mpc is even smaller, 0.00015 per cent. While this percentage does not quite reach 5σ significance, it does present a significant challenge to the standard model in addition to the tension seen in the value of the Hubble constant (see e.g. Riess et al. 2022).

Most of the CF4 distances are obtained using either the FP or the TF distance indicators. In order to test whether these galaxies whose distances have been measured by different distance indicators

give significantly different results, we have analysed two smaller data sets: (1) where all groups and individual galaxies that contain only FP galaxies (26 800 groups and individuals out of 38 058) are removed; and (2) where groups and individual galaxies that contain only TF galaxies (9 060 groups and individuals out of 38 058) are removed. Note that there are a few thousand groups containing galaxies with distance measurements from a mix of distance indicators that are common to both sub-catalogues. The bulk flows estimated from both of these sub-catalogues do not differ significantly from that of the entire catalogue (though the errors increase), giving us confidence that it is not a problem with one or the other of these main distance indicators that is causing the larger than expected bulk flow that we observe.

Given the current tension in the value of the Hubble constant (e.g. Valentino et al. 2021), it is important to consider how a different value of H_0 might effect our results. First, changing the value of H_0 introduces a spurious radial inflow or outflow into our data. While a radial flow does not contribute to the bulk flow estimate in a completely isotropic survey, one might be concerned that in a survey such as the CF4 with different radial distributions in different directions, a radial flow may bleed into a bulk flow estimate. However, as discussed in Section 6, our method contains an explicit constraint making our bulk flow estimate independent of the value of H_0. The second place that the value of H_0 might effect our results is by changing the scale of redshift distances d_z used to locate groups in space, thus changing how the bulk flow probes the power spectrum. Here, as is also discussed in section 6, the power spectrum scales with H_0 in just such a way that if distances are measured in units of h^{-1} Mpc, theoretical bulk flow estimates are independent of the Hubble constant. Taken together, these two considerations allow us to make bulk flow estimations, and more importantly, estimates of the probability of our bulk flows in the standard model of cosmology, that are completely independent of the Hubble constant.

Our results have clarified features of the bulk flow dependence on radius that were hinted at in previous studies (e.g. Peery et al. 2018). In particular, we see from Fig. 7 that, contrary to expectations, all three bulk flow components increase in magnitude as the radius of the volume increases beyond 100 h^{-1} Mpc. This behaviour is difficult to explain in the standard cosmological model, particularly since higher order moments of the velocity field do not appear to be larger than expected (Feldman et al. 2010). While this behaviour is difficult to explain in the standard model, it could be an indication that the particle rest frame of the Universe is not the same as that inferred from the dipole in the cosmic microwave background (CMB) radiation (see e.g. Kashlinsky et al. 2008; Ma, Gordon & Feldman 2011; Migkas et al. 2021; Kashlinsky & Atrio-Barandela 2022). Indeed, significant evidence that the CMB frame is not the rest frame of the Universe can also be seen in the analyses of the distribution of Quasars (Secrest et al. 2021; Dam, Lewis & Brewer 2022; Secrest et al. 2022) and voids (Haslbauer, Banik & Kroupa 2020), for an interesting review see (Kumar Aluri et al. 2023). We should note that the flattening out of the curves in Fig. 7 is not necessarily due to a convergence to a rest frame; there is little information in the CF4 beyond 200 h^{-1} Mpc that could change the values of the bulk flow components.

One might be concerned that our large bulk flow result is somehow arising from the anisotropic distribution of objects in the CF4 catalogue. There are two reasons to think that this is not the case. First, the MV analysis method weights objects in such a way as to compensate for their distribution; regions with more objects are down-weighted and regions with fewer objects are up-weighted by the method. The aim of the MV method is to generate weights to make the resulting bulk flow as close as possible to what would be obtained if there was a uniform distribution of objects in the volume. We have checked that our method has not resulted in a small subset of objects having undue influence on our bulk flow result. Secondly, Fig. 3 shows that the anisotropy of the CF4 catalogue is mainly in the z direction; the sample is much more balanced in the x and y directions. Since we only have radial peculiar velocities, objects along a particular direction make the largest contribution to the bulk flow in that direction; this suggests that if anisotropy were effecting our result, it would have the greatest effect on the z-component of the bulk flow. However, we see from Fig. 7 that the z-component of the bulk flow is quite consistent with expectations; it is the y-component that is unexpectedly large in magnitude. Altogether it is difficult to see how an anisotropy that is primarily in the z direction could cause an unusually large bulk flow in the y direction.

Given the greater context of our result being one of several that is contributing to an increasing tension with the standard model of cosmology, it is important to continue to improve our data set and our analysis methods. Unlike in the case of the Hubble constant, where resolving questions around the zero point of the distance ladder is essential, our result is relatively insensitive to the precise value of the Hubble constant and hence also the zero point. Thus the accuracy of our result is likely to be improved only through the collection of more distance measurements. Given that velocity errors increase with distance, finding distance estimators that give smaller percentage errors would greatly increase the value of new measurements.

The χ^2 analysis we have used has the advantage of being simple and extendable; very few steps separate the model, in particular the power spectrum, from the results. Other power spectra could be substituted and tested with very little effort. An alternative approach would be to create mock CF4 catalogues drawn from simulations using standard model parameters. These mock catalogues could be run through the same analysis that we used for the real data, and the statistics of the resulting bulk flows could be studied. While the minimum variance method has already been tested on earlier peculiar velocity data sets (Agarwal et al. 2012), the analysis of mock catalogues could firmly eliminate the possibility that the appearance of a large bulk flow is the consequence of some peculiarity of the CF4 or the particular analysis used here. However, this approach is challenging owing to the complex geometry of the CF4 and its mixture of groups and individual galaxies. In addition, our results suggest catalogues with bulk flows as large as the actual CF4 will be extremely rare, necessitating the creation of a large number of mock catalogues in order to assess the significance of our result. Further evaluations with mock catalogues require separate detailed studies and will likely be the basis of future papers.

While this manuscript was under review we became aware of Whitford, Howlett & Davis (2023), which carries out a similar analysis to that presented here except with the addition of mock catalogues. Their results are consistent with ours, although the level of tension they find with the standard model is smaller due to their larger uncertainty estimates.

ACKNOWLEDGEMENTS

RW, TA, CJB, AR, and AW were partially supported by the US National Science Foundation grant AST-1907365. HAF, RC, and YA-S were partially supported by US National Science Foundation grant AST-1907404. Funding for the *Cosmicflows* project was provided by the US National Science Foundation grant AST09-08846, the National Aeronautics and Space Administration grant

NNX12AE70G, and multiple awards to support observations with *HST* through the Space Telescope Science Institute.

DATA AVAILABILITY

Data available on request. The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abate A., Erdoğdu P., 2009, MNRAS, 400, 1541
Agarwal S., Feldman H. A., Watkins R., 2012, MNRAS, 424, 2667
Andernach H., Zwicky F., 2017, preprint (arXiv:1711.01693)
Carrick J., Turnbull S. J., Lavaux G., Hudson M. J., 2015, MNRAS, 450, 317
Dainotti M. G., De Simone B., Schiavone T., Montani G., Rinaldi E., Lambiase G., 2021, ApJ, 912, 150
Dainotti M. G., De Simone B. D., Schiavone T., Montani G., Rinaldi E., Lambiase G., Bogdan M., Ugale S., 2022, Galaxies, 10, 24
Dam L., Lewis G. F., Brewer B. J., 2022, preprint (arXiv:2212.07733)
Davis M., Nusser A., Masters K. L., Springob C., Huchra J. P., Lemson G., 2011, MNRAS, 413, 2906
Dressler A., 1987, ApJ, 317, 1
Dressler A., Faber S. M., Burstein D., Davies R. L., Lynden-Bell D., Terlevich R. J., Wegner G., 1987, ApJ, 313, L37
Feldman H. A., Watkins R., 1994, ApJ, 430, L17
Feldman H. A., Watkins R., 1998, ApJ, 494, L129
Feldman H. A., Watkins R., 2008, MNRAS, 387, 825
Feldman H. A., Watkins R., Melott A. L., Chambers S. W., 2003, ApJ, 599, 820
Feldman H. A., Watkins R., Hudson M. J., 2010, MNRAS, 407, 2328
Freedman W. L., 2017, Nat. Astron., 1, 0121
Geller M. J., Huchra J. P., 1989, Science, 246, 897
Hamilton A. J. S., 1998, in Hamilton D.ed., Astrophysics and Space Science Library, Vol. 231, The Evolving Universe. Kluwer, Dordrecht, p. 185
Haslbauer M., Banik I., Kroupa P., 2020, MNRAS, 499, 2845
Hellwing W. A., Nusser A., Feix M., Bilicki M., 2017, MNRAS, 467, 2787
Howlett C., Said K., Lucey J. R., Colless M., Qin F., Lai Y., Tully R. B., Davis T. M., 2022, MNRAS, 515, 953
Hubble E., 1929, Proc. Natl. Acad. Sci. USA, 15, 168
Hui L., Greene P. B., 2006, Phys. Rev. D, 73, 123526
Jacoby G. H. et al., 1992, PASP, 104, 599
Kaiser N., 1988, MNRAS, 231, 149
Kashlinsky A., Atrio-Barandela F., 2022, MNRAS, 515, L11
Kashlinsky A., Atrio-Barandela F., Kocevski D., Ebeling H., 2008, ApJ, 686, L49
Kourkchi E., Tully R. B., Anand G. S., Courtois H. M., Dupuy A., Neill J. D., Rizzi L., Seibert M., 2020a, ApJ, 896, 3
Kourkchi E. et al., 2020b, ApJ, 902, 145
Kourkchi E., Tully R. B., Courtois H. M., Dupuy A., Guinet D., 2022, MNRAS, 511, 6160
Kumar Aluri P. et al., 2023, Class. Quantum Gravity, 40, 094001
Lauer T. R., Postman M., 1994, ApJ, 425, 418
Leavitt H. S., Pickering E. C., 1912, Harv. Coll. Obs. Circ., 173, 1
Lee M. G., Freedman W. L., Madore B. F., 1993, ApJ, 417, 553
Ma Y.-Z., Gordon C., Feldman H. A., 2011, Phys. Rev. D, 83, 103002
Macaulay E., Feldman H. A., Ferreira P. G., Jaffe A. H., Agarwal S., Hudson M. J., Watkins R., 2012, MNRAS, 425, 1709
Migkas K., Pacaud F., Schellenberger G., Erler J., Nguyen-Dang N. T., Reiprich T. H., Ramos-Ceja M. E., Lovisari L., 2021, A&A, 649, A151
Nusser A., 2014, ApJ, 795, 3
Nusser A., 2016, MNRAS, 455, 178
Nusser A., Davis M., 1995, MNRAS, 276, 1391
Nusser A., Davis M., 2011, ApJ, 736, 93
Nusser A., Branchini E., Davis M., 2011, ApJ, 735, 77
Peery S., Watkins R., Feldman H. A., 2018, MNRAS, 481, 1368
Phillips M. M., 1993, ApJ, 413, L105
Planck Collaboration, 2016, A&A, 594, A13
Rathaus B., Kovetz E. D., Itzhaki N., 2013, MNRAS, 431, 3678
Riess A. G., Press W. H., Kirshner R. P., 1995, ApJ, 445, L91
Riess A. G. et al., 2022, ApJ, 934, L7
Rubin V. C., Ford W. K., Jr, 1970, ApJ, 159, 379
Rubin V. C., Thonnard N., Ford W. K., Jr, Roberts M. S., 1976, AJ, 81, 719
Sarkar D., Feldman H. A., Watkins R., 2007, MNRAS, 375, 691
Secrest N. J., von Hausegger S., Rameez M., Mohayaee R., Sarkar S., Colin J., 2021, ApJ, 908, L51
Secrest N. J., von Hausegger S., Rameez M., Mohayaee R., Sarkar S., 2022, ApJ, 937, L31
Springob C. M. et al., 2014, MNRAS, 445, 2677
Strauss M. A., Willick J. A., 1995, Phys. Rep., 261, 271
Tonry J., Schneider D. P., 1988, AJ, 96, 807
Tully R. B., Fisher J. R., 1977, A&A, 54, 661
Tully R. B., Shaya E. J., Karachentsev I. D., Courtois H. M., Kocevski D. D., Rizzi L., Peel A., 2008, ApJ, 676, 184
Tully R. B. et al., 2013, AJ, 146, 86
Tully R. B., Courtois H. M., Sorce J. G., 2016, AJ, 152, 50
Tully R. B. et al., 2023, ApJ, 944, 94
Turnbull S. J., Hudson M. J., Feldman H. A., Hicken M., Kirshner R. P., Watkins R., 2012, MNRAS, 420, 447
Valentino E. D. et al., 2021, Class. Quantum Gravity, 38, 153001
Wang Y., Rooney C., Feldman H. A., Watkins R., 2018, MNRAS, 480, 5332
Watkins R., Feldman H. A., 1995, ApJ, 453, L73
Watkins R., Feldman H. A., 2007, MNRAS, 379, 343
Watkins R., Feldman H. A., 2015a, MNRAS, 447, 132
Watkins R., Feldman H. A., 2015b, MNRAS, 450, 1868
Watkins R., Feldman H. A., Hudson M. J., 2009, MNRAS, 392, 743
Whitford A. M., Howlett C., Davis T. M., 2023, preprint (arXiv:2306.11269)
Willick J. A., 1994, ApJS, 92, 1
York D. G. et al., 2000, AJ, 120, 1579
Zaroubi S., Bernardi M., da Costa L. N., Hoffman Y., Alonso M. V., Wegner G., Willmer C. N. A., Pellegrini P. S., 2001, MNRAS, 326, 375
Zeldovich Y. B., Sunyaev S. R. A., 1980, Pis'ma Astron. Zh., 6, 737

This paper has been typeset from a TEX/LATEX file prepared by the author.