Local Gravity versus Local Velocity: Solutions for β and nonlinear bias

Marc Davis[1]*, Adi Nusser[2], Karen L. Masters[4], Christopher Springob[5]
John P. Huchra[6], Gerard Lemson[7]

[1] *Departments of Astronomy & Physics, University of California, Berkeley, CA. 94720*
[2] *Physics Department and the Asher Space Science Institute-Technion, Haifa 32000, Israel*
[3] *Smithsonian Astrophysical Observatory, Cambridge, MA 02138*
[4] *Institute for Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Burnaby Road, Portsmouth PO1 3FX, UK*
[5] *Anglo-Australian Observatory, P.O. Box 296, Epping, NSW 1710, Australia*
[6] *Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[7] *Max-Planck Institute of Astrophysics, Garching DE*

17 January 2011

ABSTRACT

We perform a reconstruction of the cosmological large scale flows in the nearby Universe using two complementary observational sets. The first, the SFI++ sample of Tully-Fisher (TF) measurements of galaxies, provides a direct probe of the flows. The second, the whole sky distribution of galaxies in the 2MASS redshift survey (2MRS), yields a prediction of the flows given the cosmological density parameter, Ω, and a biasing relation between mass and galaxies. We aim at an unbiased comparison between the peculiar velocity fields extracted from the two data sets and its implication on the cosmological parameters and the biasing relation. We expand the fields in a set of orthonormal basis functions, each representing a plausible realization of a cosmological velocity field smoothed in such a way as to give a nearly constant error on the derived SFI++ velocities. The statistical analysis is done on the coefficients of the modal expansion of the fields by means of the basis functions. Our analysis completely avoids the strong error covariance in the smoothed TF velocities by the use of orthonormal basis functions and employs elaborate mock data sets to extensively calibrate the errors in 2MRS predicted velocities. We relate the 2MRS galaxy distribution to the mass density field by a linear bias factor, b, and include a luminosity dependent, $\propto L^{\alpha}$, galaxy weighting. We assess the agreement between the fields as a function of α and $\beta = f(\Omega)/b$, where f is the growth factor of linear perturbations. The agreement is excellent with a reasonable χ^2 per degree of freedom. For $\alpha = 0$, we derive $0.28 < \beta < 0.37$ and $0.24 < \beta < 0.43$, respectively, at the 68.3% and 95.4% confidence levels (CLs). For $\beta = 0.33$, we get $\alpha < 0.25$ and $\alpha < 0.5$, respectively, at the 68.3% and 95.4% CLs. We set a constraint on the fluctuation normalization, finding $\sigma_8 = 0.66 \pm 0.10$, which is only 1.5σ deviant from WMAP results. It is remarkable that σ_8 determined from this local cosmological test is close to the value derived from the CMB, an indication of the precision of the standard model.

Key words: Cosmology: large-scale structure of the Universe, dark matter, cosmological parameters

* E-mail: mdavis@berkeley.edu

1   INTRODUCTION

For 15 years, the problem of large-scale flows of galaxies has seen little attention relative to other probes of the large scale structure in the Universe. The data on peculiar velocities has been difficult to obtain, and the results had contradictory conclusions (Strauss & Willick 1995; Zaroubi 2002). They are limited to small redshifts ($\sim 100 h^{-1}$Mpc) at which distance indicators can reliably be used. These earlier forays into the subject led to disagreements that few people wanted to sift through. But in the interval, the data has improved dramatically, thus stirring recent activity in the subject.

Peculiar velocities are unique in that they provide explicit information on the three dimensional mass distribution, and measure mass on scales of $20 - 50h^{-1}$ Mpc, a scale untouched by alternative methods. Local peculiar velocity data is, in principle, affluent in cosmological information. Power spectra and correlation functions could be derived from the data by direct calculation or by maximum likelihood techniques (e.g. Gorski et al. 1989; Jaffe & Kaiser 1995; Freudling et al. 1999; Juszkiewicz et al. 2000; Bridle et al. 2001; Abate & Erdoğdu 2009). Direct calculation of low order moments of the flow, such as the bulk motion and the shear could also analyzed within the framework of cosmological models (e.g. Feldman et al. 2010). Some authors claim the bulk flow of mixed catalogs of galaxies argues there are problems with ΛCDM (Watkins et al. 2009), but recent results (Nusser & Davis 2011) show that the SFI++ catalog by itself has a large scale bulk flow that is consistent with ΛCDM, and this analysis has smaller error bars. Other, perhaps more ambitious, applications could involve an assessment of the statistical nature of the initial cosmological large scale fluctuations, i.e. whether gaussian or otherwise (Nusser & Dekel 1993; Bernardeau et al. 1995). All these analyses could be performed with peculiar velocity measurements alone.

Here we will be concerned with a comparison of the observed peculiar velocities on the one hand and the velocities derived from the fluctuations in the galaxy distribution on the other. The basic physical principle behind this comparison is simple. The large scale flows are almost certainly the result of the process of gravitational instability with overdense regions attracting material, and underdense regions repelling material. Initial conditions in the early universe might have been somewhat chaotic, so that the original peculiar velocity field (i.e. deviations from Hubble flow) was uncorrelated with the mass distribution, or even contained vorticity. But those components of the velocity field which are not coherent with the density fluctuations will adiabatically decay as the Universe expands, and so at late times one expects the velocity field to be aligned with the gravity field, at least in the limit of small amplitude fluctuations (Peebles

1980; Nusser et al. 1991). In the linear regime, this relation implies a simple proportionality between the gravity field $\mathbf{g}$ and the velocity field $\mathbf{v}_g$, namely $\mathbf{v}_g \propto \mathbf{g}\,t$ where the only possible time t is the Hubble time. The exact expression depends on the mean cosmological density parameter Ω and is given by (Peebles 1980),

$$\mathbf{v}_g(r) = \frac{2 f(\Omega)}{3 H_0 \Omega} \mathbf{g}(r) \ . \tag{1}$$

Given complete knowledge of the mass fluctuation field $\delta_\rho(\mathbf{r})$ over all space, the gravity field $\mathbf{g}(\mathbf{r})$ is

$$\mathbf{g}(\mathbf{r}) = G\bar{\rho} \int d^3 \mathbf{r}' \delta_\rho(\mathbf{r}') \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} \ , \tag{2}$$

where $\bar{\rho}$ is the mean mass density of the Universe. If the galaxy distribution at least approximately traces the mass on large scale, with linear bias b between the galaxy fluctuations δ_G and the mass fluctuations (i.e. $\delta_g = b\delta_\rho$), then from (1) and (2) we have

$$\mathbf{v}_g(r) = \frac{H_0 \beta}{4\pi\bar{n}} \sum_i \frac{1}{\phi(r_i)} \frac{\mathbf{r}_i - \mathbf{r}}{|\mathbf{r}_i - \mathbf{r}|^3} + \frac{H_0 \beta}{3} \mathbf{r} \ , \tag{3}$$

where $\bar{n}$ is the true mean galaxy density in the sample, $\beta \equiv f(\Omega)/b$ with $f \approx \Omega^{0.55}$ the linear growth factor (Linder 2005), and where we have replaced the integral over space with a sum over the galaxies in a catalog, with radial selection function $\phi(r)$[1]. The second term is for the uniform component of the galaxy distribution and would exactly cancel the first term in the absence of clustering within the survey volume. Note that the result is insensitive to the value of H_0, as the right hand side has units of velocity. We shall henceforth quote all distances in units of $\mathrm{km\,s^{-1}}$. The sum in equation (3) is to be computed in real space, whereas the galaxy catalog exists in redshift space. As we shall see in §3.1, the modified equation, which includes redshift distortions, maintains a dependence on Ω and b through the parameter β. Therefore, a comparison of the measured velocities of galaxies to the predicted velocities, $\mathbf{v}_g(r)$, gives us measure of β. Further, a detailed comparison of the flow patterns addresses fundamental questions regarding the way galaxies trace mass on large scales and the validity of gravitational instability theory.

In this paper we shall make this comparison using the best presently available data for both the velocity and gravity fields. The direct comparison of the peculiar velocities is fraught with difficulty. Distances to individual galaxies are typically uncertain at the 20% level and are furthermore subject to considerable Malmquist bias. We shall elaborate

---

[1] $\phi(r)$ is defined as the fraction of the luminosity distribution function observable at distance r for a given flux limit; see (e.g. Yahil et al. 1991).

galaxies were observed as part of the 6dFGS (Jones et al. 2005, 2009), which used the 6dF multi-fiber spectrograph on the 1.2m UK Schmidt in Siding Spring, Australia. Their ultimate product was a map of 110,256 2MASS galaxies in the Southern sky to a magnitude limit of K=12.75 mag and to within 10 deg of the Galactic plane. This survey is far deeper than the stated goal of 2MRS, but also has a higher Galactic latitude limit. In the Northern hemisphere the 2MRS builds on a strong tradition of redshift surveys at the CfA: the CfA redshift survey and ZCAT (Davis et al. 1982; de Lapparent et al. 1986). In the absence of a northern hemisphere equivalent to the 6dF, new redshift observations are done galaxy by galaxy using the 1.2m telescope at the Fred Lawrence Whipple Observatory on Mt. Hopkins, AZ. The average density of galaxies at the magnitudes 2MRS is observing is about 1 per degree, so without a wide-field multi-object spectrograph in the Northern hemisphere this remains the most efficient way to get new redshifts. Lower Galactic latitude galaxies in the Southern hemisphere ($|b| > 5$) were added to 2MRS from observations at CTIO.

The version of the 2MRS which is complete to K=11.25 (consisting of 23,200 galaxies; Huchra et al. 2005, Westover 2007) has been used in to calculate the acceleration on the Local Group by (Erdoğdu et al. 006a). The dipole estimate seems to converge to the CMB result within $60h^{-1}$ Mpc, suggesting that the bulk of the motion of the Local Group comes from structures within that distance. They also have done a dipole analysis, weighting the sample by its luminosity, rather than the counts, and find relatively minor changes. Density and velocity fields have been calculated by (Erdoğdu et al. 2006) for the K=11.25 sample. All major local superclusters and voids are successfully identified, and backside infall onto the "Great Attractor" region (at $50h^{-1}$ Mpc) is detected.

The 2MRS catalog appears to be a fair tracer of the underlying mass distribution. The real-space correlation lengths, r_0 is best fit by a regression $r_0 = (7.5 \pm 0.5) - (3.0 \pm 0.6)\log_{10}n$, where n is the cumulative number density in $10^{-3}\ h^3$ Mpc (Westover 2007). In contrast, (Zehavi et al. 2010) report that the R band optically selected SDSS survey gives $r_0 = (6.7 \pm 0.1) - (2.0 \pm 0.1)\log_{10}n$. In terms of bias estimates, (Westover 2007) reports $b/b_* = 0.73 + 0.24L/L_*$ while (Norberg et al. 2002) state that $b/b_* = 0.85 + 0.15L/L_*$ for the 2DF survey. In other words, Westover's data show the 2MRS correlations are more dependent on luminosity than are optically selected samples. In view of this luminosity dependent result, it makes the most sense to evaluate the gravity field in a luminosity-weighted manner; it is computed below with a variety of luminosity weightings. Westover (2007) has also made a mock catalog for the missing galaxies at low latitude by interpolating the galaxy density above and be-

low them in three dimensions. We shall use this catalog as an estimate of the local mass density.

2.2 TF sample

Twenty years ago, the mis-calibration of full sky Tully-Fisher data was the problem that led to very discrepant results for the determination of $\beta \equiv \Omega/b$, with $\beta = 0.5 \pm 0.2$ (Davis et al. 1996) and $\beta = 1.0 \pm 0.2$ (e.g. Dekel et al. 1993). The mistaken TF calibration led to a large scale flow that confused both analyses, but in the end, it was a calibration error in the Southern sky which made a false large-scale flow (Willick et al. 1997). In one analysis this led to a higher χ^2 than was acceptable, and in the other it led to a biased result.

For the analysis below, we use the recently completed survey of spiral galaxies with I-band Tully-Fisher distances, SFI++ (Masters et al. 2006; Springob et al. 2007, 2009), which in turn builds on the original Spiral Field I-band Survey (Giovanelli et al. 1994, 1995; Haynes et al. 1999) and Spiral Cluster I-band Survey (Giovanelli et al. 997a,b). We use the published SFI++ magnitudes and velocity widths, and derive our own peculiar velocities, rather than taking the published distances as given. We use the SFI++ catalog as it includes several datasets to give full sky coverage. It is not essential for our analysis that the peculiar velocity sample have uniform sky coverage, but they must have a uniform calibration.

The other major Tully-Fisher catalog was published by (Tully et al. 2008). This survey is restricted to $cz < 3000$ km/s, and includes many of the same galaxies in that redshift range as SFI++. Tully also make use of a different algorithm for measuring spectral line widths, which are not easily comparable to the values derived for SFI++. So while one could in principle combine the two catalogs for this analysis, the small potential gain in sample size is not enough to justify the resulting heterogeneity in observational methods and data analysis.

We shall use the inferred distances, as well as redshifts, to derive an estimate of peculiar velocity for each galaxy. Correlation analysis (Borgani et al. 2000) indicates that peculiar velocities in the SFI++ behave as expected for ΛCDM models. (see esp. (Branchini et al. 2001b; Freudling et al. 1999; da Costa et al. 1998; Feldman & Watkins 2008))

In the analysis below, we shall use the inverse of the Tully-Fisher (ITF) relationship, as given in equation 8. We begin by drawing the published magnitudes, velocity widths, and redshifts from Springob et al. (2007, 2009). We include all field, group, and cluster galaxies, which leaves us with an initial sample of 4859 galaxies. Galaxies in groups and clusters are treated as individual objects, though the redshifts
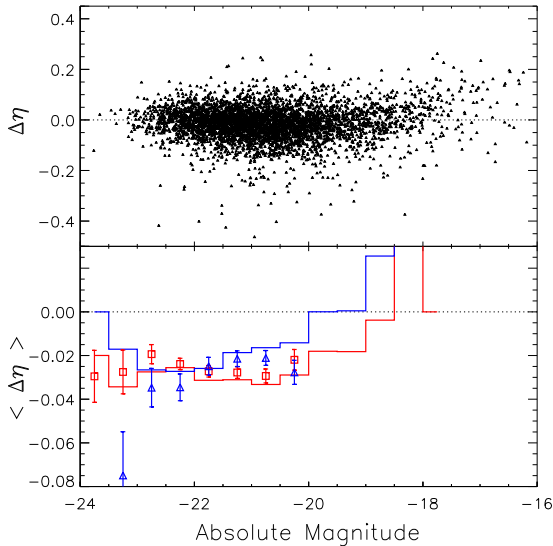
Figure 1. top: The scatter of the $\eta - M_I$ relationship, where $\gamma = 0.12$ Bottom: The mean $\Delta\eta$ in .5 magnitude bins of the raw distribution (recall that $\eta = \log(W)$). The red is for galaxies with $cz > 5000$ km/s, while the blue is for $cz < 5000$ km/s. Note the change where $M_I > -20$, and also how the red and blue curves appear to have different TF relations. The histograms are for the raw data, with no flow model, while the points, with 1σ error bars, are the mean values of $\eta - M_I$ after the fit. The nearby and more distant galaxies now have identical TF relations.

for template cluster galaxies are replaced by the systematic redshift of the cluster. Following Giovanelli et al. (1997a), we brighten the magnitudes of Sb galaxies by 0.10 magnitudes and brighten the magnitudes of spirals earlier than Sb by 0.32 magnitudes, while leaving types later than Sb unchanged. This is done in order to account for subtle differences in the TF relation of different spiral subclasses. We select only objects with inclination $i > 45$ deg to ease problems with inclination corrections. The data must be transformed to the LG frame, and galaxies with $cz < 200$ km/s are deleted. All the analysis is done in the LG frame as the boundary conditions then simply becomes $v_g \to 0$ and $v_{itf} \to 0$.

The few galaxies with large residual $\Delta\eta$ ($\Delta\eta$ is the residual from equation 8; see figure 1) are sufficiently deviant to be a worry for statistics which depend on data with a compact core and no long tails. The typical outlying object is not unusually nearby in redshift, and peculiar motions cannot be the explanation. The vast majority of SFI++ galaxies have well behaved TF relationships; perhaps the outliers are undergoing a merger? There are large negative $\Delta\eta$ outliers, but few corresponding large positive $\Delta\eta$ outliers, and this is resolved by clipping the outliers at $|\Delta\eta| > 0.20$.

The top panel of figure 1 shows the distribution of $\Delta\eta$ *before* galaxies have had their magnitude changed because of peculiar velocities; in the bottom panel is shown the result of averaging the data into .5 mag bins, where the red histogram is for galaxies at $cz > 5000$ km/s, and the blue histogram is for those with $cz < 5000$ km/s, in the raw data. The red and blue points, plus $1 - \sigma$ error bars, are the average $< \Delta\eta >$ versus M *after* the best flow model, described in §3.5, is fit to the data. The blue point with $M < -23$ is deviant, but it only represents 11 galaxies, compared to an average of 170 galaxies in the other bins. Note that the zero point of the $< \Delta\eta >$ behavior makes no difference; only the slope, the constancy of $< \Delta\eta >$ versus M is important. Compared to the different slopes before the flow corrections are applied, the TF relation is now identical in the foreground versus the background of the SFI++ data. This figure is for illustration only, as the data is not binned during the fitting process.

The bend in the TF relationship at $M = -20$ is known to be a result of the reduced mass in the Baryonic Tully-Fisher relation (Stark et al. 2009; Gurovich et al. 2010). We are missing the data to straighten out the curve, and since the ITF method is easiest to apply if there is a linear relationship between η and M, we simply delete all galaxies with $M > -20$ from further consideration. After all these cuts, we are left with 2830 spiral galaxies with $200 < cz_{LG} < 10000$ km/s.

The raw distribution of $\Delta\eta$, after limiting the sample and fitting the best linear curve, is found to approximately fit a gaussian with $\sigma = 0.059$. This is the dispersion with *no flow model* applied. The gaussian width to the distribution is $\sigma = 0.0558$ after the flow model is applied. The small decrease is limited by the intrinsic, dominant noise of the TF relation. This noise has numerous causes, such as the uncertainty in the inclination correction of the SFI++ galaxies, or small variations in the outer limits of the rotation curves of the galaxies.

The following sections, §3 and §5, explain the machinery for effecting this reduction.

3   RECONSTRUCTION OF PECULIAR VELOCITIES

In this section we outline our method described in ND94, ND95 and DNW96 for deriving the smooth peculiar velocities of galaxies from an observed distribution of galaxies in redshift space and, independently, from a sample of spiral galaxies with measured circular velocities η and apparent magnitudes m.

3.1 Peculiar Velocities from the Distribution of Galaxies in Redshift Space

There are several methods for generating peculiar velocities from redshift surveys, using linear (e.g. Fisher et al. 1995) and non-linear relations (e.g. Peebles 1980; Croft & Gaztanaga 1998; Nusser & Branchini 2000; Frisch et al. 2002a; Enßlin et al. 2009) Here we restrict ourselves to large scales where linear-theory is applicable. We will use the method of ND94 for reconstructing velocities from the 2MRS. This method is is particularly convenient, as it is easy to implement, fast, and requires no iterations. Most importantly, this redshift space analysis closely parallels the ITF estimate described below. We next present a very brief summary of the methodology.

We follow the notation of DNW96. The comoving redshift space coordinate and the comoving peculiar velocity relative to the Local Group (LG) are, respectively, denoted by s (i.e. $s = cz/H_0$) and $v(s)$. To first order, the peculiar velocity is irrotational in redshift space (Chodorowski & Nusser 1999) and can be expressed as $v_g(s) = -\nabla\Phi(s)$ where $\Phi(s)$ is a potential function. As an estimate of the fluctuations in the fractional density field $\delta_0(s)$ traced by the discrete distribution of galaxies in redshift space we consider,

$$\delta_0(s) = \frac{1}{(2\pi)^{3/2}\bar{n}\sigma^3} \sum_i \frac{w(L_{0i})}{\phi(s_i)} \exp\left[-\frac{(s - s_i)^2}{2\sigma^2}\right] - 1 \quad . \quad (4)$$

where $\bar{n} = \sum_i w(L_{0i})/\phi(s_i)$ and w weighs each galaxy according to its estimated luminosity, L_{0i}. The 2MRS density field is here smoothed by a gaussian window with a redshift independent width, $\sigma = 350$ km s^{-1}. This is in contrast to DNW96 where the *IRAS* density was smoothed with a width proportional to the mean particle separation. The reason for adopting a constant smoothing for 2MRS is its dense sampling which is nearly four time higher than *IRAS* . We emphasize that the coordinates s are in *observed redshift* space, expanded in a galactic reference frame. The only corrections from pure redshift space coordinates is the collapse of the fingers of god of the known rich clusters prior to the redshift space smoothing (Yahil *et al.* 1991). Weighting the galaxies in equation (4) by the selection function and luminosities evaluated at their redshifts rather than the actual (unknown) distances yields a biased estimate for the density field. This bias gives rise to Kaiser's rocket effect (Kaiser 1987).

To construct the density field, equation 4, we volume limit the 2MRS sample to 3000 km/s, so that $\phi(s < 3000) = 1$, resulting in $\phi(s = 10000) = 0.27$ (Westover 2007). In practice, this means we delete galaxies from the 2MRS sample fainter than $M_* + 2$. Galaxies at 10,000 km/s therefore have

$1/\phi = 3.7$ times the weight of foreground galaxies in the generation of the velocity field, v_g.

If we expand the angular dependence of Φ and $\delta_0(s)$ redshift space in spherical harmonics in the form,

$$\Phi(s) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \Phi_{lm}(s) Y_{lm}(\theta, \varphi) \quad (5)$$

and similarly for δ_0, then, to first order, Φ_{lm} and δ_{0lm} satisfy,

$$\frac{1}{s^2}\frac{d}{ds}\left(s^2\frac{d\Phi_{lm}}{ds}\right) - \frac{1}{1+\beta}\frac{l(l+1)\Phi_{lm}}{s^2} \quad (6)$$
$$= \frac{\beta}{1+\beta}\left(\delta_{0lm} - \kappa(s)\frac{d\Phi_{lm}}{ds}\right) ,$$

where

$$\kappa = \frac{d\ln\phi}{ds} - \frac{2}{s}\frac{d\ln w(L_{0i})}{d\ln L_{0i}} \quad (7)$$

represents the correction for the bias introduced by the generalized Kaiser rocket effect. As emphasized by ND94, the solutions to equation (6) for the monopole ($l = 0$) and the dipole ($l = 1$) components of the radial peculiar velocity in the LG frame are uniquely determined by specifying vanishing velocity at the origin. That is, the radial velocity field at redshift s, when expanded to harmonic $l \leqslant 1$, is not influenced by material at redshifts greater than s.

In this paper, we shall consider solutions as a function of β and the parameter α defining a power law form $w_i \propto L_i^\alpha$ for the galaxy weights.

3.2 Peculiar Velocities from the Inverse Tully-Fisher relation

Given a sample of galaxies with measured circular velocity parameters, $\eta_i \equiv \log\omega_i$, linewidth ω_i, apparent magnitudes m_i, and redshifts z_i, the goal is to derive an estimate for the smooth underlying peculiar velocity field. We assume that the circular velocity parameter, η, of a galaxy is, up to a random scatter, related to its absolute magnitude, M, by means of a linear *inverse* Tully-Fisher (ITF) relation, i.e.,

$$\eta = \gamma M + \eta_0. \quad (8)$$

One of the main advantages of inverse TF methods is that samples selected by magnitude, as most are, will be minimally plagued by Malmquist bias effects when analyzed in the inverse direction (Schechter 1980; Aaronson et al. 1982). We write the absolute magnitude of a galaxy,

$$M_i = M_{0i} + P_i \quad (9)$$

where

$$M_{0i} = m_i + 5\log(z_i) - 15 \quad (10)$$

and

$$P_i = 5\log(1 - u_i/z_i) \tag{11}$$

where m_i is the apparent magnitude of the galaxy, z_i is its redshift in units of $\mathrm{km\,s^{-1}}$, and u_i its radial peculiar velocity in the LG frame.

ND95 base a velocity model on spherical harmonics and spherical Bessel functions, for galaxies distributed over the sky to 6000 km/s. With the 2MRS we extend the gravity field to 10,000 km/s. In general, one can write the function P_i in terms of an expansion over j_{m} orthogonal basis functions, F_i^j,

$$P_i = \sum_{j=0}^{j_{\mathrm{m}}} a^j F_i^j \tag{12}$$

with orthonormality conditions,

$$\sum_{i=1}^{N_g} F_i^j F_i^{j'} = \delta_K^{j,j'} \tag{13}$$

and the zeroth mode defined by $F_i^0 = 1/\sqrt{N_g}$, where N_g is the number of galaxies in the sample. The mode F^0 describes a Hubble-like flow in the space of the data set which is degenerate with the zero point of the ITF relation. Here we set $a^0 = 0$, which removes the Hubble-like flow from the gravity field, below. The best fit mode coefficients, a^j, the slope, γ, and the zero point η_0, are found by minimizing the χ^2 statistic

$$\chi^2_{\mathrm{ITF}} = \sum_{i=1}^{N_g} \frac{(\gamma M_{0i} + \gamma P_i + \eta_0 - \eta_i)^2}{\sigma^2_{\eta,\mathrm{int}}} , \tag{14}$$

where $\sigma_{\eta,\mathrm{int}}$ is the rms of the intrinsic scatter in η about the ITF relation, and N_g is the number of galaxies in the sample. Given the orthonormality condition, the solution to the equations $\partial\chi^2_{\mathrm{ITF}}/\partial a^j = 0$, $\partial\chi^2_{\mathrm{ITF}}/\partial\gamma = 0$ and $\partial\chi^2_{\mathrm{ITF}}/\partial\eta_0 = 0$ is straightforward. Thanks to the orthonormality condition, the covariance matrix $<\delta a_j \delta a'_j>$ of the errors in a^j is diagonal with

$$\sigma_a = <(\delta a^j)^2>^{1/2} = \frac{\sigma_{\eta,\mathrm{int}}}{\gamma} . \tag{15}$$

This lack of covariance of the errors in the coefficients is most rewarding as it makes the ITF error analysis exceptionally simple. Therefore, statistical assessment of the match between the data will be done at the level of the modes rather than the peculiar velocities. The interested reader will find details in ND95 and DNW96.

3.3 The orthonormal basis functions

The choice of radial basis functions for the expansion of the modes can be made with considerable latitude. The functions should obviously be linearly independent, and close to

orthogonal when integrated over volume. They should be smooth and close to a complete set of functions up to a given resolution limit. ND95 chose spherical harmonics Y_l^m for the angular wavefunctions and the derivatives of spherical Bessel functions for the radial basis functions, motivated by the desire to use functions which automatically satisfy potential theory boundary conditions at the origin and the outer boundary. That is, they chose

$$P(y,\theta,\phi) = \sum_{n=0}^{n_{max}} \sum_{l=0}^{l_{max}} \sum_{m=-l}^{m=l} \frac{a_{nlm}}{y} \left(j'_l(k_n y) - c_{l1}\right) Y_{lm}(\theta,\phi) . \tag{16}$$

The function $y(z)$ is designed to compress the distance scale, increasing the smoothing scale of the mode to deal with increased noise at large distances. For this analysis we use

$$y(z) = \sqrt{\log(1 + (z/z_s)^2)} \tag{17}$$

where $z_s = 5000$ km/s. The constant c_{l1} is non-zero for the dipole term only and ensures that $P = 0$ at the origin, and is non-zero at the outer boundary. Details of how the orthogonalized functions $\tilde{F}_i^j$ are derived from this expansion are given in ND95.

The spherical harmonics are expanded to a maximum $n = 5$ and $l = 3$, except we delete the $n = 1$ mode for $l = 0$ as this mode can be confused with the false Hubble flow described in the next section. We also include an external quadruple, distinct from the internal quadruple, to describe the gravity induced by material at distances $cz > 10000$ $\mathrm{km\,s^{-1}}$. Summing over the values of m, that makes a total of 72 modes fit toward reducing the χ^2 of equation 14. The use of $y(z)$ is designed to allow the radial resolution to degrade with distance; for example, the $n = 5$ modes have a wavelength in the radial direction of 3000 $\mathrm{km\,s^{-1}}$ at $cz = 8000$ $\mathrm{km\,s^{-1}}$ and a wavelength of 1300 $\mathrm{km\,s^{-1}}$ at $cz = 1000$ $\mathrm{km\,s^{-1}}$.

3.4 Expanding the 2MRS gravity field

In order to assess the match between the velocities by means of the expansion coefficients and to ensure that both fields are smoothed similarly, the 2MRS predicted velocities v_{g} must also be described by an expansion over the basis functions used in the v_{itf} model. Using the machinery for computing a gravity field described in §3.1, one can generate a linear theory predicted peculiar velocity v_{g} for any point in space as a function of its redshift for any value of β. We must ensure that the smoothing scales of the ITF and 2MRS predicted peculiar velocities are matched to the same resolution. Therefore we expand v_{g} in terms of the modes used in the velocity model. Because of the orthonormality, we can write the mode coefficients as

$$a_{\mathrm{g}}^j = 5 \sum_{i=1}^{N_{\mathrm{g}}} \log \left(1 - \frac{(v_{\mathrm{g},i} - H'cz_i)}{cz_i} \right) F_i^j \; , \qquad (18)$$

where the H' term is a correction for Hubble flow and the summation index i is restricted to be over the positions of the same galaxies in the ITF expansion.

This procedure will filter out fluctuations that are not described by the resolution of our basis functions. We do not include any mode such that $P_i = $ constant, which would be a pure Hubble flow. In the fitting for the ITF modes, pure Hubble flow is absorbed into a shift of the zero point η_0 and the orthogonality is ensured. Within a given set of test points occupying a volume smaller than that used to define the gravity field, it is possible for v_g to have a non-zero value of Hubble flow H', which must be removed from v_g before we tabulate the mode coefficients. That is, we tabulate the mean Hubble ratio

$$H' = \frac{\sum_{k=1}^{N_{\mathrm{g}}} v_{g,k} cz_k}{\sum_{k=1}^{N_{\mathrm{g}}} v_{g,k}^2} \qquad (19)$$

and subtract it from the predicted field v_g. This "breathing mode" which mimics a Hubble flow is not trivial in amplitude, and can be a 10% correction on the effective Hubble constant within simulated catalogs. This mode is cosmologically expected to a modest degree, but a bigger portion of the effect is caused by error in the determination of $\overline{n}$, which we estimate by assuming the weighted counts within 12,000 km/s is the mean value.

3.5 Refinements of the ND95 functions

The Nusser & Davis (1995) reconstruction of the base functions can only provide a rough estimate of the spatial distribution of galaxies in the TF data. It does not guarantee that the signal to noise in the filtered fields is uniform all over the sample. The ND95 expansion yields reliable TF velocities of nearby galaxies, but very noisy estimates at larger distances. Moreover, it is difficult to achieve a desired resolution as a function of redshift and to ensure equal resolutions in the radial and angular directions. The ND95 method expands the observed velocities in terms of harmonic functions, but the individual harmonic modes are not regularized and may acquire unrealistically large amplitudes, depending on the spatial coverage of the data.

We aim here at generating base functions which are themselves smoothed with a variable isotropic smoothing window designed to yield a constant signal to noise in the estimated v_{itf}. We construct these new basis functions with the help of the ND95 orthogonal functions denoted here by F^{ND}. Suppose a single radial velocity field, V_{seed}, with the appropriate variable smoothing has been found. We term V_{seed} the *seed* field as the new modes will stem from it. As

will be described below this field will be chosen as the 2MRS predicted velocity field, but any other field representing a viable velocity field could serve as V_{seed}. Given

$$P_{\mathrm{seed},i} = 5 \log(1 - V_{\mathrm{seed},i}/cz_i) \qquad (20)$$

(where i refers to galaxies in the TF sample) we expand $P_{\mathrm{seed},i}$ in base functions constructed according to ND95,

$$a_{\mathrm{seed}}^j = \sum_i P_{\mathrm{seed},i} F_i^{NDj} \; . \qquad (21)$$

Here, the number of the modes F^{ND} is sufficiently large so that the inverse transformation $\sum_j a_{\mathrm{seed}}^j F_i^{NDj}$ reproduces $P_{\mathrm{seed},i}$. In practice we use about 1400 ND95 modes (we go to $l = 17$).

We then form additional fields, P_α, according to

$$P_{\alpha,i} = \sum_j \mathcal{R}_\alpha^j a_{\mathrm{seed}}^j F_i^{NDj} \; , \qquad (22)$$

where $\mathcal{R}$ is a set of normally distributed random numbers with zero mean and standard deviation of unity. This reconstruction of the additional fields preserves the "power" in the modal expansion and randomizes the phases. So far all these fields, $P_{\alpha,i}$, are unfiltered and may contain non-linear small scale fluctuations. Therefore, we smooth all fields $P_{\alpha,i}$ according to

$$P_{\alpha,i}^{\mathrm{smooth}} = \sum_{all\ galaxies} P_{\alpha,i'} W(s_{i,i'}, R_{s,i'}) \; , \qquad (23)$$

where $s_{i,i'}$ is distance (in redshift space) between the galaxies i and i', and W is a gaussian window of width $R_{s,i}$ which depends on the galaxy i. The smoothing width $R_{s,i}$ is tuned such that the expected error in the ITF velocity of galaxy i is $\sim 150 \; \mathrm{km\,s^{-1}}$ and therefore it depends on the redshift and the local density of galaxies near i. The smoothing length at the positions of galaxies in the SFI++ sample is shown in figure (2); it varies roughly linearly with redshift, ranging from $1h^{-1}$ Mpc nearby up-to $30h^{-1}$ Mpc for galaxies at redshifts $\sim 10000 \; \mathrm{km\,s^{-1}}$. The new basis functions are then obtained by orthonormalizing $P_{\alpha,i}^{\mathrm{smooth}}$. We refer to the new functions by the standard notation F_i^j, with F_i^1 being the smoothed $P_{\mathrm{seed},i}$. These new modes will be used in the expansion given in equation (12). In §3.6 we describe how we determine the number of modes, j_{m}, to be used in the expansion. The seed field V_{seed} could be constructed by interpolating any unfiltered cosmological velocity field on the positions of the galaxies in the SFI++ sample. Nevertheless, we could improve on this by constructing V_{seed} from the unfiltered 2MRS velocities given directly by the solution to equation (6). In practice, we use unfiltered v_g obtained with $\beta = 0.2$. The choice of v_g for $\beta = 0.2$ is arbitrary; the predicted field with any other β could be used. If the $V_{\mathrm{seed}} \sim v_g$ and v_{itf} both describe the same underlying velocity field then the
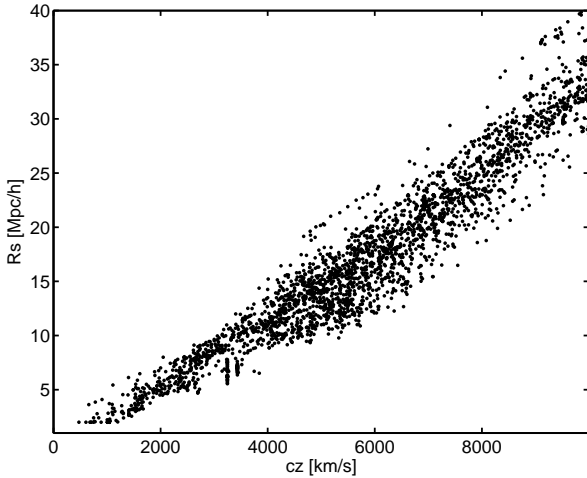
Figure 2. The width of the gaussian smoothing gaussian window versus galaxy redshifts in the SFI++ sample. The scatter reflects the angular variations in the density of galaxies.

additional modes should mainly reflect the covariance of the errors between the two fields.

The flow patterns of 9 of the modes are shown, respectively, in the 9 panels in figure (3). The color scheme throughout this paper is normal: red (blue) means outflowing (infalling), from the central point.. We have extensively tested a broad variety of choices for the the first mode. None of the results of the analysis reported are sensitive to this choice of the first mode. The figure shows that the higher order modes exhibit smaller scale structures. This is a direct result of the orthogonolization processes. The $j^{th} + 1$ mode has to be orthogonal to the all previous j modes. In order to achieve that, the orthogonalized $j^{th} + 1$ will pick more of the small scale structure.

3.6   The ITF scatter and the number of modes

Once the basis functions for the modal expansion are given, we proceed to solve for the coefficients a^j in $P_i = \sum a^j F_i^j$ by minimizing χ^2_{ITF} in (14). The minimization is also performed, at the same time, with respect to the slope and zero point of the ITF. The estimated slope is $\gamma = -0.1297 \pm 0.0015$ and -0.13 ± 0.0016, respectively, for 20 and 30 modes used in the flow model. The raw slope before fitting the model is $\gamma = -0.1267 \pm 0.0016$. The zero point plays no role at all here and we do not keep track of its estimated values. All estimated parameters, including a^j, are independent of the (assumed constant) intrinsic scatter $\sigma_{\eta,\mathrm{int}}$ in the ITF. The velocity model can be used to estimate the unknown value of $\sigma_{\eta,\mathrm{int}}$. Given the residual

$$\Delta\eta_i = \eta_i - (\gamma M_{0i} + \gamma P_i - \eta_0) \tag{24}$$

we approximate $\sigma_{\eta,\mathrm{int}}$ by

$$\sigma_\eta^2 = \sum_i (\Delta\eta_i)^2 / N_{\mathrm{d.o.f}} \tag{25}$$

where

$$N_{\mathrm{d.o.f}} = N_{\mathrm{g}} - (j_{\mathrm{m}} + 2) \tag{26}$$

is the number of degrees of freedom taking into account that the minimization of χ^2_{ITF} is done with respect to j_{m} coefficients plus the slope and zero point of the ITF. This σ_η will decreases as the number of modes, j_{m}, in the expansion is increased. If j_{m} is too large then the higher order modes will be dominated by noise. If j_{m} is too small then the model may miss significant components of the underlying true galaxy velocities.

The optimal range of j_{m} for our comparison can be seen by inspecting the behavior of σ_η as a function of j_{m}. The (blue) circles in the top panel of figure 5 shows σ_η versus the number of modes for modes generated from the seed field $V_{\mathrm{seed}} = v_{\mathrm{g}}(\beta = 0.2)$ (see §3.5). Most of the reduction in σ_η^2 is already achieved the first mode. This is very encouraging since this means that v_{g} picks up a significant contribution of the velocities as described by the ITF data. It also means that both the 2MRS and the ITF data are likely to provide approximations to the underlying flow field. However, the 2MRS predicted field v_{g} deviates from the underlying field by the presence of correlated errors in the reconstruction scheme. The inclusion of additional expansion modes in the ITF velocity model will dissolve these errors. The average reduction in the variance σ_η^2 per mode becomes insignificant beyond $j_{\mathrm{m}} = 64$; the average reduction per mode for the first 64 modes is 10 times larger than that for the next 75 modes. An F-test also confirms that the reduction in the variance marginal beyond $j_{\mathrm{m}} = 64$. Hence we will approximate $\sigma_{\eta,\mathrm{int}}^2 = 0.0558$, the value acquired by σ_η for 64 modes in the velocity model. Therefore, the F-test argues that 64 is the maximum number modes needed to model the ITF.

The next step is to determine the minimum number of modes needed to describe the ITF flow assuming that $\sigma_{\eta,\mathrm{int}} = 0.0558$. To do so we tabulate χ^2_{ITF} a function of j_{m} and compute the probability $Q = Q(\chi^2_{\mathrm{ITF}}|N_{\mathrm{dof}})$ that the value χ^2_{ITF} is exceeded by chance (c.f. §6.2 in Press et al. 1992). The values of of $\chi^2_{\mathrm{ITF}}/N_{\mathrm{dof}}$ and Q are represented as the (blue) circles, respectively, in the middle and bottom panels of figure 5. For $j_{\mathrm{m}} = 64$, we get $\chi^2_{\mathrm{ITF}}/N_{\mathrm{d.o.f}} = 1$ corresponding to $Q = 0.5$, in accordance with our choice of $\sigma_{\eta,\mathrm{int}}$. Without a velocity model, i.e. $j_{\mathrm{m}} = 0$, we get $(\chi^2_{\mathrm{ITF}}/N_{\mathrm{dof}}, Q) = (1.125, 3 \times 10^{-6})$. This exceedingly low Q rejects a vanishing velocity field with very high confidence level (CL). Including the first mode alone gives a highly significant improvement: $(\chi^2_{\mathrm{ITF}}/N_{\mathrm{d.o.f}}, Q) = (1.031, 0.12)$. The
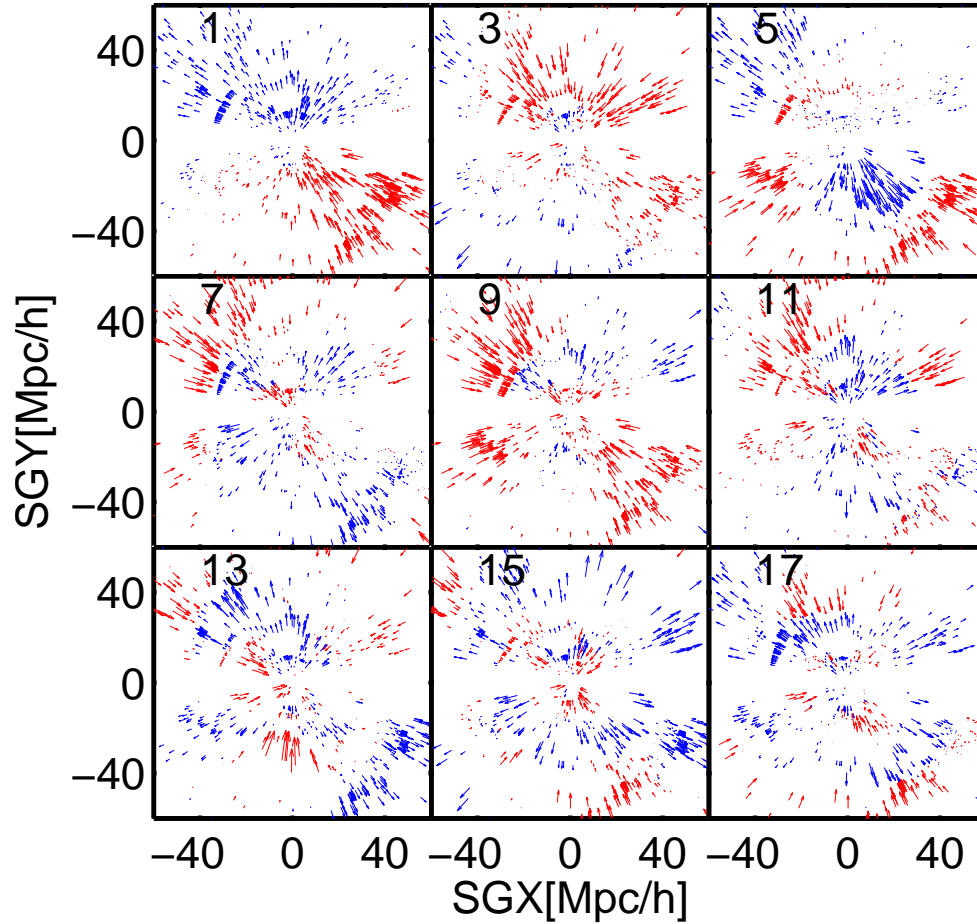
Figure 3. The flow patterns of 9 modes for galaxies within $5h^{-1}$ Mpc of the supergalactic plane. The order of the mode is indicated in the corresponding panel.

hypothesis that χ^2_{ITF} value corresponding to the first mode is obtained by chance is rejected only at the 0.12 CL. This is encouraging since the first mode velocity field is proportional to the 2MRS predicted velocities, v_{g} (for $\beta = 0.2$). For $j_{\mathrm{m}} = 30$ and 20 we get $(\chi^2_{\mathrm{ITF}}/N_{\mathrm{d.o.f}}, Q) = (1.01, 0.29)$ and $(1.014, 0.2)$, respectively.

For comparison with $V_{\mathrm{seed}} = v_{\mathrm{g}}$, the corresponding results for random choice of the seed field, V_{seed}, are shown as the (red) crosses in the top and bottom panels of figure 5. With a random V_{seed}, about 20 modes are needed to reduce σ_η to the level achieved by the single mode v_{g}.

3.7   What is the purpose of the ITF machinery?

The expansion of the gravity field is conceptually very clean when computed in the LG frame (ND95). The Poisson-like equation for the 3-D gravitational field has been solved as a sum over the spherical harmonic functions Y_{lm} times 1-D functions of r that satisfy physically reasonable boundary conditions at the origin. For the purposes of the ITF solution, we furthermore quantize the radial solutions with quantum number n.

The ITF method is backward from the usual methodology of TF applications; one does not fit curves to the scatter of peculiar velocities. Instead, the χ^2 equation 14
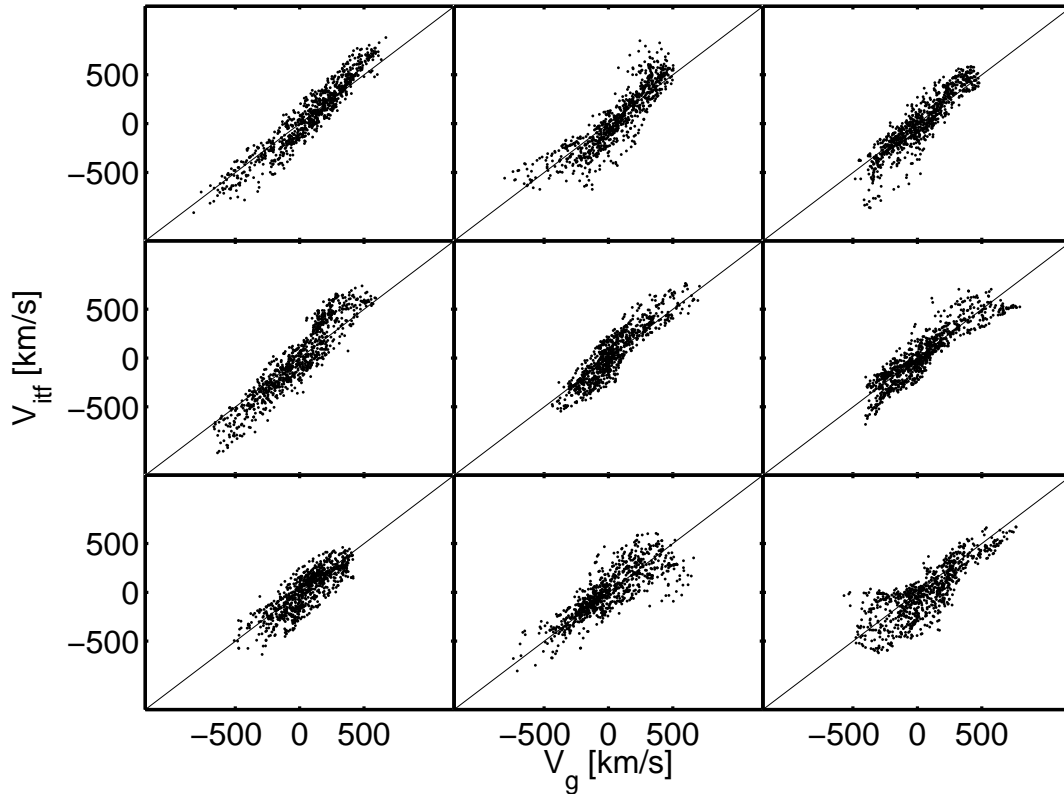
Figure 4. Scatter plots of v_{itf} versus v_{g} (expanded in 20 modes) for galaxies in 9 mock catalogs. About 800 galaxies are plotted for each mock.

is minimized by the addition of linear combinations of the orthonormal functions of n, l, m, where each describes a set of large-scale flow that satisfy the boundary conditions. Furthermore, we have endeavored to form a first, 'seed', mode based on linear growth rate, but in which large scale graininess is filtered out by the use of figure 2.

The individual galaxy's peculiar velocity enters by equation 11, with differential $dP_i \propto du_i/z_i$, and since the uncertainty of peculiar velocity u_i is proportional to redshift, the uncertainty of P_i is redshift independent. This means that each object is given equal weight in a fit, and our window function is therefore equivalent to the display of figure 7, which shows the positions of the SFI++ galaxies.

4   MOCK CATALOGS AND ERROR ANALYSIS

As a measure of the agreement between the TF and predicted velocities by means of the corresponding expansion coefficients, a_{itf} and a_{g}, we will consider the χ^2 function

$$\chi^2 = \sum_{(j,k)=1}^{j_{\mathrm{m}}} \left(a_{\mathrm{g}}^j - a_{\mathrm{itf}}^j\right)\left(\sigma_a^2 + \xi_{\mathrm{g}}\right)_{j,k}^{-1}\left(a_{\mathrm{g}}^k - a_{\mathrm{itf}}^k\right) \ . \tag{27}$$

The parameters α and β will be obtained by minimization of this function. The covariance of the residual, $a_{\mathrm{itf}} - a_{\mathrm{g}}$, is the sum of the covariance matrices of the errors in the estimation of a_{itf} and a_{g}, respectively. Thanks to the orthonormality of the basis functions, the error covariance in the estimation of a_{itf} is diagonal with constant terms $\sigma_a^2 = (\sigma_\eta/\gamma)^2$ (see §3.2 and DNW96). The matrix ξ_{g} represents the covariance of the errors in the determination of a_{g}. The origin of these errors is as follows:

(i) Equation 6 is expected to predict reliable velocity fields only for small amplitude fluctuations. Small scale nonlinear deviations from linear theory inevitably leak to large scales.

(ii) The 2MRS is a finite number sampling of the underlying density field. This leads to "shot-noise" errors in the estimation of the density field.
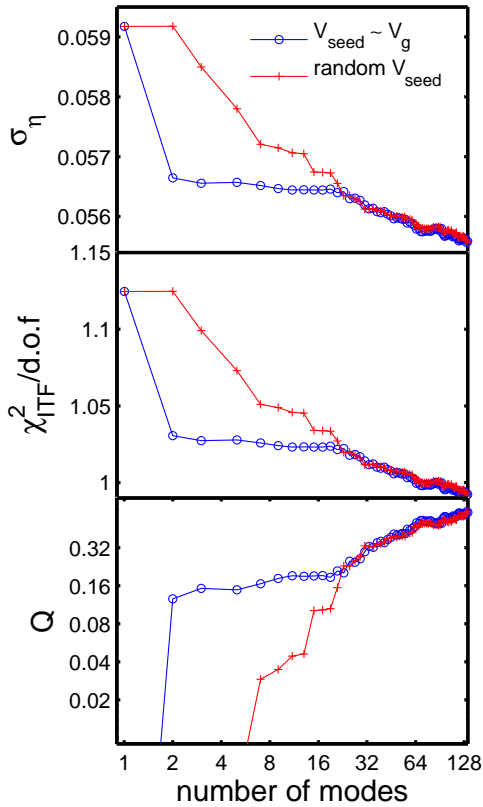
Figure 5. *Top:* The rms value, σ_η, of the ITF as a function of the number of modes used in the velocity model. *middle:* $\chi^2_{\rm ITF}$ per degrees-of-freedom versus the number of modes. It is unity at 64 modes. *Bottom:* The probability that the $\chi^2_{\rm ITF}$ exceeds a certain value by chance, as a function of the number of modes. The chi-square is computed assuming an intrinsic scatter $\sigma_{\eta,{\rm int}} = 0.0558$. The value $Q = 0.5$ is achieved at 64 modes. (see text for details)

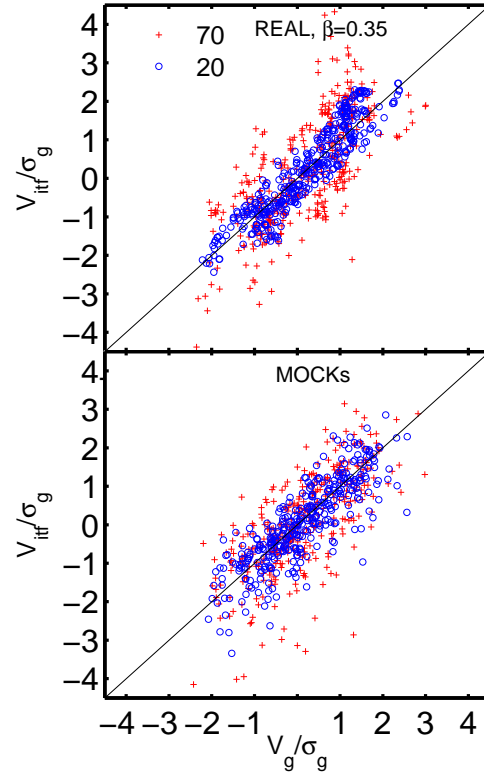**Figure 6.** The peculiar velocities $v_{\rm itf}$ versus $v_{\rm g}$ of galaxies in the real data (top) and the mocks (bottom) for 20 and 70 expansion modes, as indicated in the figure. The real $v_{\rm g}$ has been reconstructed with $\beta = 0.35$ and $\alpha = 0$. Velocities of about 400 randomly selected galaxies are plotted in each panel where each mocks is represented by about 25 galaxies. All velocities in the top panel are normalized by the rms value, $\sigma_g = 233$ km s^{-1} of $v_{\rm g}$, while velocities in the bottom panel are normalized by the rms value of $v_{\rm g}$ of their corresponding catalogs.

(iii) Small scale random motions of galaxies, especially in groups and clusters, give rise to a smearing of the distribution of galaxies along the line of sight in redshift space.

(iv) There is a possible large scale stochastic biasing (Dekel & Lahav 1999; Sigad et al. 2000; Wild et al. 2005) between the galaxy distribution and the mass fluctuations.

The only way to achieve a reliable estimate of $\xi_{\rm g}$ taking into account all of these complicated errors is by means of mock catalogs designed to match the general properties of the 2MRS. A parent simulated catalog of the whole 2MASS catalog has already been prepared (De Lucia & Blaizot 2007) by incorporating semi-analytic galaxy formation models in the Millennium simulation (Springel et al. 2005). From

this parent catalog we have drawn 15 independent mock 2MRS catalogues satisfying the following conditions:

(i) The "observer" in each mock is selected to reside in a galaxy with a quiet velocity field within 500 km/s, similar to the observed universe. That is, the central server sees only one cluster that has high enough peculiar velocities to result in negative redshifts. Recall that in the LG frame, the only galaxies with negative redshift are in the Virgo cluster.

(ii) The motion of the central galaxy is 500 to 700 km/s.

(iii) The density in the environment of the local group, averaged over a sphere of 400 km/s radius , is less than twice the normal.

Corresponding mock ITF catalogs were also prepared. A counts-in-cells statistics shows that the distribution of galaxies in the mocks is unbiased relative to the dark matter, i.e. $b = 1$.

The preparation of the mocks for velocity reconstruction is done in the same way as the real data. Equation (6) is used to generate prediction of mock v_{g} with $\beta = f(\Omega, \Lambda)/b = 0.47$ corresponding to $b = 1$ and $\Omega = 0.25$ and $\Lambda = 0.75$ as in the Millennium simulation. The mean of the rms values of v_g in the mocks is $269 \ \mathrm{km \, s^{-1}}$ and the standard deviation from this mean is $56 \ \mathrm{km \, s^{-1}}$. For contrast, the rms value of v_g derived from the 2MRS with $\beta = 0.35$ is $233 \ \mathrm{km \, s^{-1}}$.

To better illustrate the covariance between the residuals in the mocks, we plot in figure 4 the velocities v_{itf} versus v_{g} for 9 individual mocks. These are velocities expanded with 20 modes. Note the similarity between the structure of the distribution of points in the individual panels and the top panel in figure 6, showing v_{itf} versus v_{g} (with $\beta = 0.35$) for the real data.

4.1 The error covariance matrix

The covariance matrix ξ_{g} is computed from the 15 mocks by projecting the correlation function, ξ_{P}, of the residuals $\Delta P = P_{\mathrm{itf}} - P_{\mathrm{g}}$ onto the basis functions, where $P_{\mathrm{itf}} = 5\log(1 - v_{\mathrm{itf}}/cz)$ and correspondingly for P_{g}. That is

$$\xi_{\mathrm{g}}(j,k) = <(a_{\mathrm{itf}}^j - a_{\mathrm{g}}^j)(a_{\mathrm{itf}}^k - a_{\mathrm{g}}^k)> = \sum_{i,i'} F_i^j \xi_{\mathrm{P}}(i,i') F_{i'}^k \ , \quad (28)$$

where the summation in the last term on the right-hand side is over all data galaxies and F are the basis functions used for the real data. In this calculation, the v_{itf} velocities are reconstructed from an ITF relation without adding the internal scatter of the TF relation. The reason is that the error in a_{itf} resulting from the intrinsic scatter has a simple analytic form given by σ_a. The function ξ_{P} is computed from the 15 mocks as follows. Denote line-of-sight and projected separations in redshift space by, $s_\parallel$ and $s_\perp$, respectively. For each mock we tabulate the average $< \Delta P_1 \Delta P_2 >$ over pairs with separations defined by the grid. We then normalize this quantity by the variance of v_{itf} (for zero ITF instrinsic scatter) in the corresponding mock. This is reasonable since the rms values of the velocity field vary considerably among the mocks and some of them are significantly different than the real data. To minimize this cosmic variance and to derive ξ_{g} given the observed rms value of the velocity, this normalization of $< \Delta P_1 \Delta P_2 >$ for each mock is prudent. The average over all mocks is then computed and interpolated from the grid onto the actual pair separations in the TF catalog to obtain the normalized ξ_{P}. The normalized covariance matrix ξ_{g} is then computed according to (28) and scaled by a factor

matching the velocity variance estimated from the observed v_{itf}.

5 THE RECONSTRUCTED VELOCITIES

This section presents a visual inspection of the fields, and assess the coherence of the residual $v_{\mathrm{itf}} - v_{\mathrm{g}}$ by means of a velocity correlation analysis. The quantification of the agreement between the fields and the extraction constraints on α and β will be deferred to later sections.

5.1 Visual inspection of the flows

Blue dots and red crosses of figure 6 show velocities expanded in 20 and 70 modes, respectively. In the bottom panel, v_{itf} versus v_{g} from the mocks are shown. In mock v_{itf} velocities are obtained from a fake ITF relation with an intrinsic slope $\gamma = -0.1$ and a scatter with $\sigma_\eta = 0.05$. For the sake of clarity only 400 randomly selected galaxies are shown in either panel. Further, each mock is represented by about 25 galaxies (randomly selected). The velocities are scaled by the corresponding rms value of v_{g} in the corresponding catalog.

There is a an excellent overall agreement between v_{itf} and v_{g} for 20 modes, both in the real data and the mocks. A good agreement prevails even for 70 modes despite the clear enhanced noise contamination. In the real data, the rms of $v_{\mathrm{itf}} - v_{\mathrm{g}}$ for 20 expansion modes is $99 \ \mathrm{km \, s^{-1}}$, significantly smaller than σ_g. For 70 modes the rms values of v_{g} and $v_{\mathrm{itf}} - v_{\mathrm{g}}$ are $238 \ \mathrm{km \, s^{-1}}$ and $231 \ \mathrm{km \, s^{-1}}$, respectively. Both panels show clear structures in the distribution of points, implying strong covariance between the residuals, $v_{\mathrm{itf}} - v_{\mathrm{g}}$, in the real, or mock, data. Because the bottom panel represents random selections of galaxies from all the mocks, the covariance pattern between the velocities is diluted in the distribution of points. The covariance pattern is, however, clear in figure 4 where scatter velocity plots for a few mocks are shown individually.

In the aitoff projections in Figure 7 we plot the TF peculiar velocities, v_{itf} and the derived gravity modes, v_{g}, for galaxies in redshift shells, $cz < 2000$, $2000 < cz < 4000$, $4000 < cz < 6000$, and $6000 < cz < 10000 \ \mathrm{km \, s^{-1}}$. The projections are in galactic coordinates centered on $l, b = 0$ and with $b = 90$ at the top. Figures 7 and 8 show v_{g} with $\beta = 0.35$. The rightmost plots are the residuals $v_{\mathrm{itf}} - v_{\mathrm{g}}$. The key point is to note that the residuals are small for the entire sky and have amplitude that is constant with redshift. The amplitude and coherence of the residuals $v_{\mathrm{itf}} - v_{\mathrm{g}}$ is the same as for the mock catalogs in figure 8, where for example the lower right picture shows $v_{\mathrm{itf}} - v_{\mathrm{g}}$ for a mock catalog. It is not very dissimilar from the real plot of $v_{\mathrm{itf}} - v_{\mathrm{g}}$ in
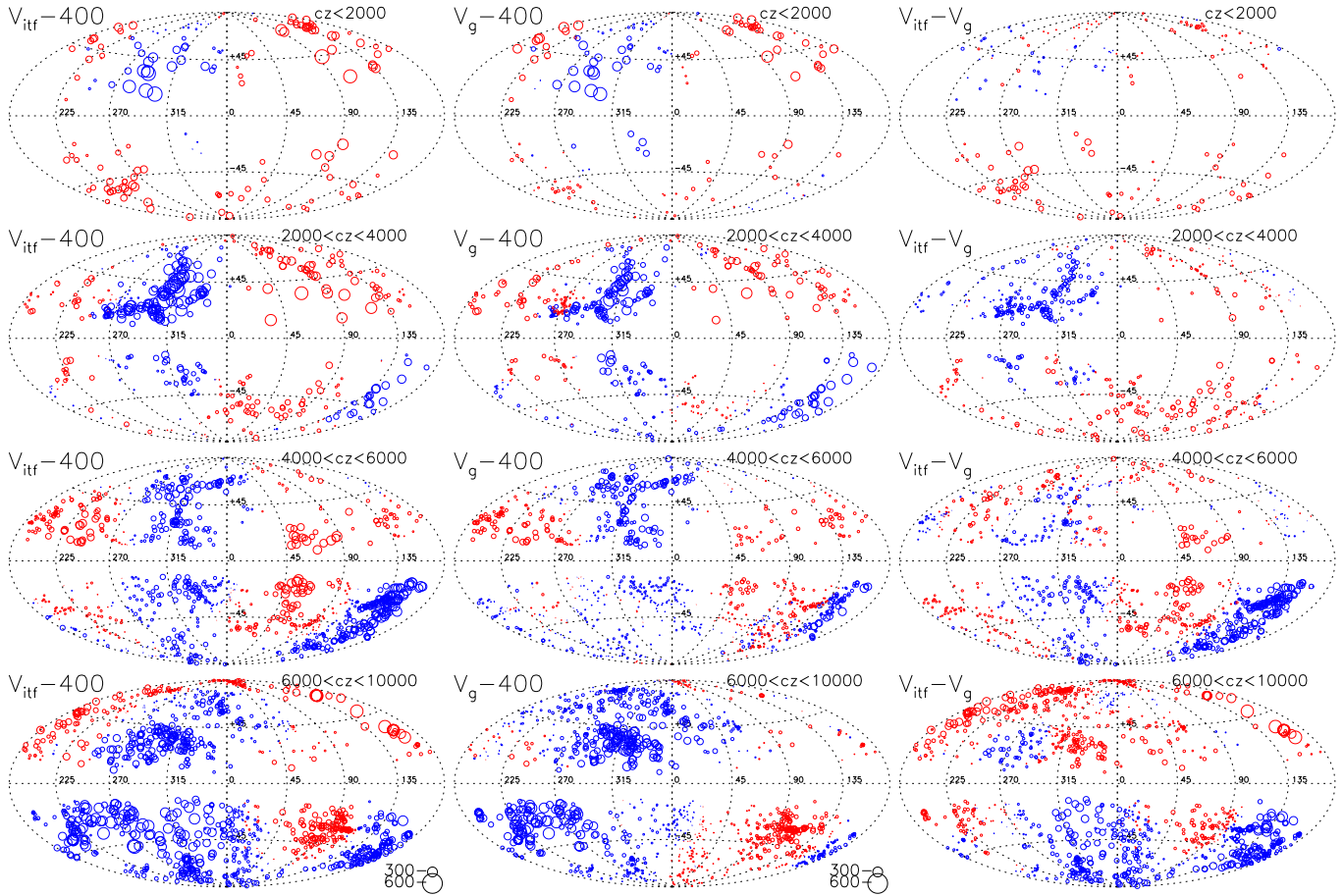
Figure 7. The derived peculiar velocities v_{itf}, v_{g}, and $v_{\mathrm{itf}} - v_{\mathrm{g}}$ of galaxies on aitoff projections on the sky in galactic coordinates. The rows correspond to galaxies with $cz < 2000$, $2000 < cz < 4000$ $4000 < cz < 6000$ km s^{-1} and $6000 < cz < 10000$ km/s, respectively. The size of the symbols is linearly proportional to the velocity amplitude (see key to the size of the symbols given at the bottom of the figure). In order to better see the differences, a 400 km/s dipole, in the direction of the CMB dipole, has been subtracted from the v_{itf} and v_{g} velocities. Note that $v_{\mathrm{itf}} - v_{\mathrm{g}}$ is considerably smaller than v_{itf} or v_{g}, even for the most distant galaxies.

the upper right, demonstrating the feasibility of the entire method.

Note the quadrupole pattern for $cz < 4000$ km s^{-1} in figures 7 and 8, visible after 400 km s^{-1} has been subtracted from the flow. This has been previously noted by Haugbølle et al. (2007) on the basis of the flows detected in 133 SNe. The quadrupole is the typical pattern observed in nbody simulations and is the principle mode of collapse to a 1-D structure.

There is amazingly overall good agreement between the large scale motions as described by v_{itf} and v_{g}. The residual velocities are coherent over large scales but they are clearly of smaller amplitude than v_{itf} and v_{g}. Note that residuals

shown in $v_{\mathrm{itf}} - v_{\mathrm{g}}$, particularly visible in the shell $6000 < cz < 10,000$ km s^{-1}, are dominated by $l = 4$, because the fit for the reduction of the TF χ^2 is limited at $l = 3$ modes.

5.2 Correlations

The residuals, both in the real and mock data, have error fields, $v_{\mathrm{itf}} - v_{\mathrm{g}}$, that show large regions of coherence. To address the significance of these errors, we show in figure 9 the velocity correlation function, (Gorski et al. 1989), defined as

$$\Psi(s; u) = \frac{\sum_{pairs} u_1 u_2 \cos\theta_{12}}{\sum_{pairs} \cos^2\theta_{12}} \qquad (29)$$
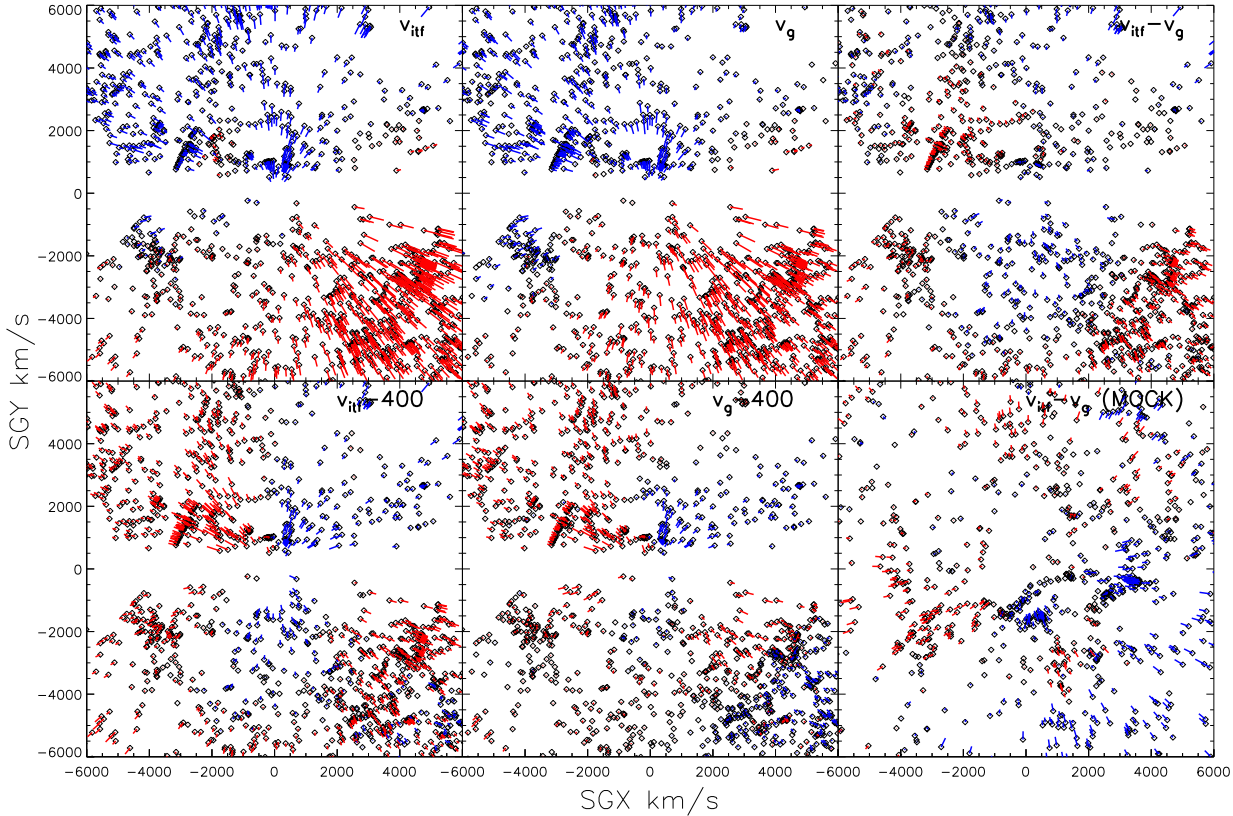
Figure 8. Supergalactic plane projection, $|SGB| < 30$, of the derived flows. To better see the differences in the plots, a dipole of 400 km/s towards the CMB pole has been subtracted from the fields, and is shown in the bottom left and bottom center. A quadruple velocity is now visible in the plots. The points are drawn at the estimated distance of an SFI++ galaxy, and the line, blue or red, is drawn to the galaxy's redshift. In other words, the length of the arrow is the peculiar velocity. The lower right plot shows $v_{\rm itf} - v_{\rm g}$ for a mock catalog, and the upper right shows $v_{\rm itf} - v_{\rm g}$ for the data. They have very similar degrees of coherence.

where the sum is over all pairs, 1 and 2, separated by vector distance $\mathbf{s_{12}}$ (in redshift space), θ_{12} is the angle between points 1 and 2, and u is either $v_{\rm itf}$ (dashed red) or $v_{\rm itf} - v_{\rm g}$ (red for data, blue for 15 mock catalogs), At small lags for the real data, the function $\Psi(r; v_{\rm itf} - v_{\rm g})$ is a factor of 3 less than $\Psi(s; v_{\rm itf})$, about the same as for the mock catalogs. Note how the large coherence of $v_{\rm itf}$ is enormously diminished in $\Psi(s > 2000 {\rm km/s}; v_{\rm itf} - v_{\rm g})$. This shows that the coherence seen in the residual field, figure 9, is expected and is not a problem. The large scale drift of a sample is demonstrated by the persistent amplitude of Ψ beyond $\approx 60 - 80$ Mpc.

The bottom panel of figure 9 shows velocity correlations for 15 mock catalogs where the actual velocity, v_{true}, generated in the nbody code and then smoothed with the 20 mode expansion, can be compared to either $v_{\rm itf}$ or $v_{\rm g}$. Note that the raw velocities, v_{itf} (red), have enormous correlation

that reaches large lag, while the correlations, $(v_{true} - v_{\rm itf})$, (blue) are extremely small. This is because the only difference with v_{true} is the gaussian error in $\Delta\eta = .05$ that affects $v_{\rm itf}$. The blue curves show this error is not a problem, because the mode expansions are insensitive to gaussian noise in the 2500 galaxies, i.e. they are essentially perfect. This demonstrates that even though the TF noise is as large as for the actual data, the ability to find the correct flow, when characterized by only 20 numbers, is intact.

Note also that the auto-covariance of $(v_{true} - v_{\rm g})$ (dot-dashed curves) is also is greatly reduced from that of $v_{\rm itf}$. Recall that $v_{\rm g}$ assumes linear theory estimated from the distribution of 20000 galaxies. Occasionally the correlations are badly mistaken, when a large cluster (much larger than Virgo) is in the foreground and complicates the difference between physical and redshift space separations, but $v_{\rm g}$ is always an excellent approximation to the TF velocity.
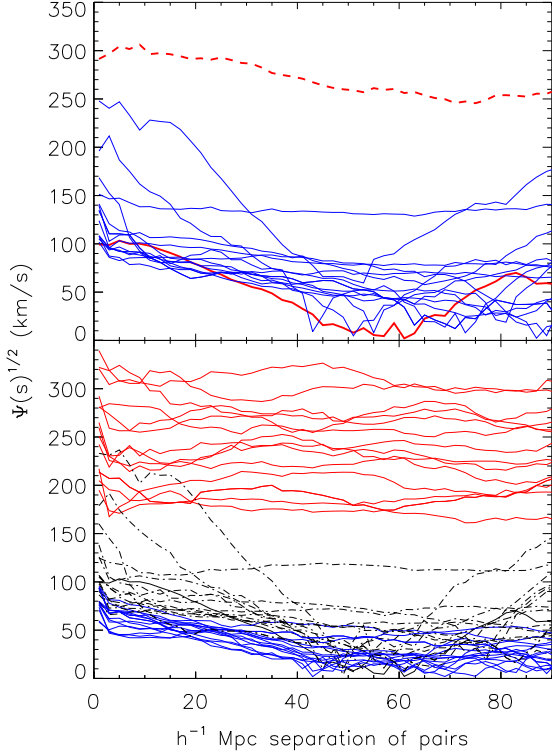
Figure 9. *Top:* The velocity correlation of the real data and 15 mock catalogs. The dashed red and solid red curves curve are, respectively, the correlations of v_{itf} and $v_{\mathrm{itf}} - v_{\mathrm{g}}$ in the real data. The blue lines are each correlations of $v_{\mathrm{itf}} - v_{\mathrm{g}}$ for the mock catalogs. *Bottom:* Velocity correlations for 15 mock catalogs. The red curves are the velocity of v_{itf}, the dot-dashed curves show the correlation of $(v_{true} - v_g)$, and the blue curves correspond to $v_{true} - v_{itf}$). Both v_{true} and v_g are first smoothed with the 20 mode expansion before the autocovariance is computed. Note that the correlation of $v_{\mathrm{itf}} - v_{\mathrm{g}}$ is only slightly worse than the correlation of $v_{true} - v_{gs}$, showing that the velocity reconstruction dominates the errors. Note also that we are plotting the square root of the velocity correlation Ψ.

6   THE CONSTRAINTS ON α, β

Equipped with the error covariance matrices, we proceed to minimize χ^2 in equation (27) with respect to α and β. We shall present detailed results for α and β for fields expanded in 20 modes and 30 modes.

The minimization is done by computing χ^2 on a grid of values in the plane α and β. At the minimum point $\chi^2 = \chi^2_{\min} = 21.5$ which is very reasonable given that the standard deviation from the expected value of 22 (20 mode coefficients plus TF slope and zero point) is $\sim \sqrt{44} \approx 7$ (Press et al. 1992).
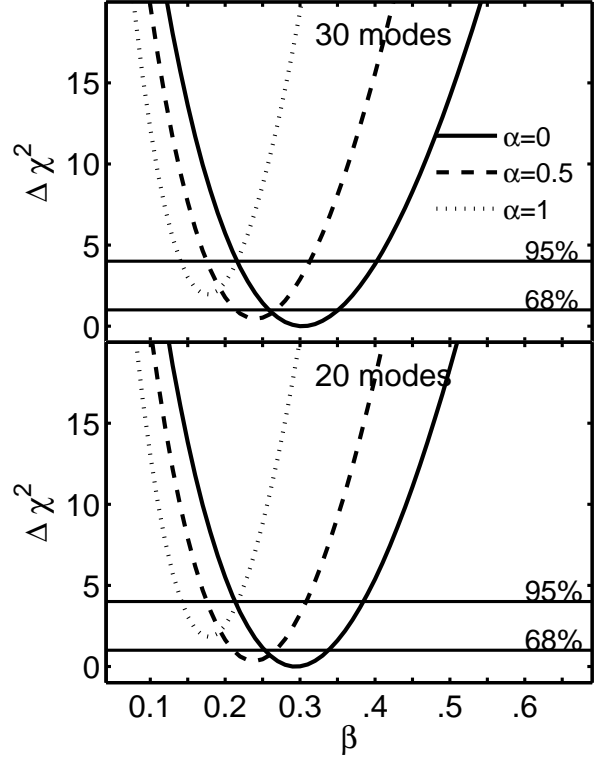

Figure 10. The difference $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$ versus β computed for three values of α, as indicated in the figure. Horizontal lines mark the 68% and 95.4 % CLs.

Figure 10 shows the difference $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$ as a function of β for three fixed values of α, as indicated in the figure. The horizontal lines indicate 68% and 95% confidence levels (CLs). Figure 12 is a contour plot of $\Delta\chi^2$ in the plane of α and β.

The higher frequency basis functions should probe smaller scales. Hence, if our assumption of linear bias which is independent of scale is valid then varying the number of modes, j_{m}, should yield consistent constraints on β. Figure 11 shows the best fit β (thick red curve) and the corresponding 95% (2σ) CLs (thin blue solid lines) versus the number modes in the expansion, for $\alpha = 0$. The circles show the "differential" best fit β obtained from a single mode as a function of the order of the mode. The 2σ errorbars on this differential β are significantly enhanced beyond the second order mode. A few points lie at the ends of the errorbars corresponding to best fit β obtained at either 0.05 or 0.7 which are the bounds of the range of β values used in the 2MRS reconstruction. There is a hint that β declines with increasing j_{m} but this is completely dominated by the noise.
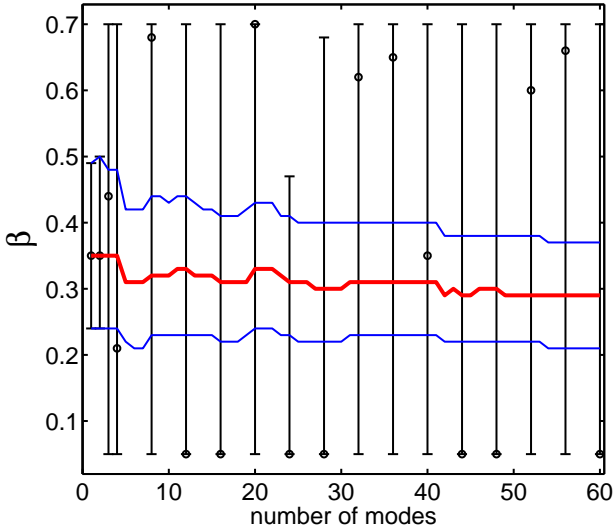
Figure 11. The thick solid red line is best fit β as a function of the number of modes in the velocity model, for $\alpha = 0$. The thin blue lines mark the 95% CL on the best fit. The circles denote the "differential" best fit β obtained with the single j^{th} mode. After the first 4 modes, only 1 in 4 modes are represented. The errorbars attached to the circles are correspond to 95% CLs on the differential best fit.
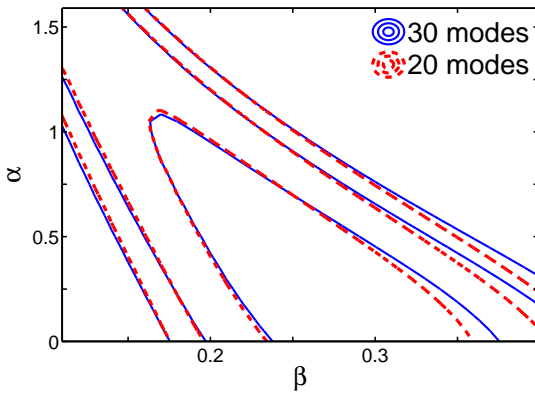
Figure 12. Contour plot of $\Delta\chi^2$ in the plane of α and β. The contours are 2.3, 6.17 and 9.2 corresponding to CLs of 68% 95.4% and 99%, respectively.

This figure shows clearly that we get consistent constraints on β (the red curve) when varying the number of modes in the expansion. Further, it shows that most of the signal is contained the very few first modes.

© 2010 RAS, MNRAS **000**,

7 DISCUSSION

The analysis reported demonstrates a good match between the 2MRS predicted and TF observed velocities. The analysis is unique in several respects. Firstly, it completely avoids dealing with covariance matrices of errors in the velocities estimated from the TF sample, Secondly, it uses elaborate mock galaxy catalogues to compute the error covariance in the predicted velocities from the redshift survey. Thirdly, the TF and predicted velocities are filtered in a very similar fashion, taking special care to minimize the effects of noise in the comparison. In this analysis faint galaxies with $M > -20$ are excised from the TF catalog since they systematically show strong deviations from the linear TF relation.

The comparison yields $\beta = f/b = 0.33\pm0.04$ (1σ error). The quoted error is not actually far from the limit of what the current data can constrain in the absence of any errors on the 2MRS predicted velocities [2]. A moderate reduction of the errors by a factor of two requires a significant enlargement in the number of peculiar velocity measurements by a factor of 4, which could be done if dedicated time is available, but the TF samples already use the best local galaxies. Going to larger distance is not the answer, as the error of a peculiar velocity increases linearly with the distance, and the 2MRS density field becomes very dilutely sampled. Two surveys, WALLABY, to be undertaken by the ASKAP telescope in Western Australia, and ALFALFA, an ongoing project at Arecibo, will hopefully produce good TF data for $cz < 12000$ km/s. Another strategy would involve peculiar velocities inferred from more precise distance indicators than the TF relation. SNe and surface brightness fluctuations techniques are likely candidates, but such measurements are available for a much smaller number of galaxies. Larger samples of local SNe are turning out to have increased errors (Ganeshalingam, private communications), significantly larger than previously measured (e.g. Riess et al. 1997).

---

[2] In the absence of gravity errors and for β close to the best fit value β_0, the χ^2 function is approximated as

$$\chi^2 = \sum_j [a_{\text{itf}}^j - B(\beta)a_{\text{g0}}^j]^2/\sigma_a^2$$

where a_{g0}^j correspond to reconstruction with $\beta = \beta_0$ and $B = (\beta/\beta_0)(1 + 1.5\beta_0)/(1 + 1.5\beta)$ approximates the dependence of v_{g} on β (in contrast to the dependence β/β_0 in reconstruction from galaxy distribution in real space). The 1σ error in B is $\sigma_{\text{a}}/\sqrt{\sum_j (a_{\text{g0}}^j)^2} \approx 0.06$, where we took $\sigma_{\text{a}} = \sigma_\eta/\gamma = 0.43$ and $\sum_j (a_{\text{g0}}^j)^2 = 52$ as given from the solution with $\beta = 0.35$. This error in B translates into an error of 0.03 in β which is close to the error obtained with the full analysis.

The good match between the gravity and velocity fields implies that they probe the same underlying potential field within the framework of the gravitational instability paradigm for structure formation. The agreement is achieved assuming a linear biasing relation between mass and galaxies on large scales. Linear biasing is consistent with theoretical predictions (e.g. Kauffmann et al. 1997) for the large scale clustering of galaxies. Further, no scale-dependent biasing seems to be required by the velocity comparison. However, the theoretically expected scale dependence of the bias factor (e.g. Peacock & Smith 2000; Desjacques et al. 2010) is well below the level which can be probed by the velocity comparison carried out here. There is also no clear indication for a scale dependence bias from the observed galaxy clustering on the relevant scales (e.g. Verde et al. 2002).

DNW96, which compared the predicted velocities from the *IRAS* 1.2 Jy redshift surveys and the MARK III Tully-Fisher data, revealed systematic discrepancies that could not be attributed to errors in the data and the reconstruction methods. Inspection of the flow fields obtained in the current work with the those presented in DNW96 (see their figures 9–13) clearly show that the problem lies in the MARK III data set. The velocity fields predicted from the *IRAS* and 2MRS surveys have similar patterns which grossly deviate from MARK III but are in accordance with SFI++. The 2MRS has all the attributes that one would want for estimating the gravity field including a very weak bias. The survey was done by an instrument that was photometrically stable, which is important to avoid large scale drifts in the derived gravity field. The survey is far superior to the *IRAS* survey, the first full sky galaxy survey, which detected galaxies at $60\mu m$, a sign of star formation and not a good indicator of mass. However, that survey led to sensible results, and was not at fault for the disagreement 15 years ago.

Radburn-Smith et al. (2004) compared the predicted velocities from the *IRAS PSCZ* survey have also been compared with measured peculiar velocities of SNe. They found a best-fitting $\beta_{IRAS} = f/b_{IRAS} = 0.55 \pm 0.06$. The lower value of β derived in our work could be due to a difference in the biasing factor between 2MRS and the *IRAS* galaxies, but we also emphasize that our estimation of the error in the predicted velocities should be more reliable as it is based on realistic mock catalogs. Pike & Hudson (2005) performed a comparison of the 2MRS predicted velocities with direct velocity measurements from three different samples, including 836 SFI++ galaxies within $cz = 5000$ km s^{-1}. Their analysis yields $\beta = 0.55 \pm .05$ for the comparison of gravity with the SFI++. They derive $\Omega = 0.55 \pm 0.05$, inconsistent with our result at more than the 2.5σ level. However, they did not calibrate their methods with advanced mock catalogues nor included the expected covariance of the predicted velocities.

Lavaux et al. (2010) employed a sophisticated version

of the nonlinear MAK reconstruction method (Frisch et al. 2002b) to compare the 2MRS predicted velocities with the 3K velocity catalog (Tully et al. 2008) of 1791 galaxies with redshifts < 3000 km s^{-1}. They derive $\Omega = 0.31 \pm 0.05$, corresponding to $\beta \sim 0.52$. However, as they point out their error analysis is incomplete. Their method is promising as it takes intro account of nonlinear effects. Nevertheless, they do not account for the covariance of the errors in their smoothed observed velocities and predicted velocities.

Both Pike & Hudson (2005) and Lavaux et al. (2010) use iterative schemes based on Yahil et al. (1991) for deriving the peculiar velocities from redshift surveys. These schemes rely on a relation between the peculiar velocity and density in real space. At any iteration, this relation is solved for new peculiar velocity given real space coordinates obtained from the observed redshifts by subtracting the old peculiar velocities derived in the previous iteration. We caution here that the these schemes are intrinsically biased: error in velocities used to estimate the distances will yield a biased density field in real space (see the §A for details). Hence the estimation of the velocity field is actually done from a biased distribution in real space. The bias produces an undesired smoothing of density field along the radial direction. The smoothing width (in km s^{-1}) is equal to the rms random error in the velocities $\sim 200 - 300$ km s^{-1} (Branchini et al. 2002; Nusser & Branchini 2000). Therefore, the bias is more pronounced in nonlinear methods which aim at probing small scales.

Checks to find the best way of estimating the gravity field did not lead to improvements. Weighting the galaxy maps by the 2MRS luminosity led to a worse agreement, and recall that 2MRS is selected in K band, which is closest to a measure of the stellar mass. Giving the elliptical galaxies double weight, as indicated by lensing analysis (Mandelbaum et al. 2006), did not improve the agreement. This in itself is not too surprising, because on large scale the 2MRS survey is dominated by spiral galaxies. It seems that the galaxies brighter than $M_* + 2$ are each surrounded by a dark matter halo that has the same mass on average. There is no hint that the dark matter mass is larger if the luminosity is increased.

Using our estimate for β we can constrain the amplitude of mass fluctuations. As a measure of the amplitude we consider the rms of density fluctuations in spheres of $8h^{-1}$ Mpc in radius, denoted by σ_8 and σ_{8g} for the mass and galaxy distributions, respectively. Adopting $\Omega = 0.266$ (Larson et al. 2010) gives $f(\Omega, \Lambda = 1 - \Omega) = 0.483$ (Linder 2005). Comparing this to our result $\beta = f/b = 0.33 \pm 0.04$ (1σ error) we get a bias factor $b = 1.46 \pm .20$ between the dark matter and the 2MRS galaxy distribution. Taking $\sigma_{8g} = 0.97 \pm 0.05$ (Westover 2007; Reid et al. 2010), yields $\sigma_8 = \sigma_{8g}/b = 0.65 \pm 0.11$ for the underlying mass density

field, marginally consistent with the latest WMAP results (Larson et al. 2010) of $\sigma_8 = 0.8 \pm 0.03$ (see also Jarosik et al. (2010)).

8 CONCLUSIONS

We summarize the major conclusions of our work:

• After a detailed examination of the 2MRS and SFI++ catalogs, we find the local gravity field to be a fine predictor of the local velocity field. Such a conclusion is a comfort for linear perturbation theory in an expanding universe and was certainly expected. It is interesting that the counts of galaxies give the best possible gravity field, reinforcing the old idea that the mass of the halo around a galaxy is not very well correlated with the luminosity of that galaxy.

• We see no evidence that the dark matter does not follow the galaxy distribution, and is consistent with constant bias on large scales. There is no evidence for a non-linear bias in the local flows. A smooth component to the universe is not something testable with these methods.

• Linear perturbation theory appears to be adequate for the large scales tested by our method.

• The solution favors $\alpha = 0.$, no correlation between luminosity and mass, and $\beta = .33 \pm .04$, which is consistent, but more than twice as tight as Erdoğdu et al. (006a). Using the derived Ω from WMAP (Jarosik et al. 2010), leads to an estimate $\sigma_8 = 0.65 \pm 0.10$, deviant from WMAP's reports at the 1.5σ level.

• Our estimate of σ_8 gives the most precise value at $z \sim 0$ and is useful for tests of the growth rate and Dark Energy.

• The velocity-gravity comparison measures the acceleration on scales up to $30 - 50$ Mpc. and since we derived a similar value of β as for clusters of galaxies, we conclude that dark matter appears to fully participate in the clustering on scales of a few Megaparsecs and larger.

• We find no evidence for large-scale flows such as reported by, for example (Hudson et al. 1999; Hudson & Ebeling 1997; Feldman et al. 2010). Note that our analysis has not used the CMBR dipole, but we see a velocity field that is fully consistent with those previously reported (Erdoğdu et al. 2006, 006a; Erdoğdu & Lahav 2009), which are consistent with the CMBR dipole radiation. We see no evidence that the dipole in the CMBR is produced by anything other than our motion in the universe.

9 ACKNOWLEDGMENTS

REFERENCES

Aaronson M., Huchra J., Mould J., Schechter P. L., Tully R. B., 1982, ApJ, 258, 64
Abate A., Erdoğdu P., 2009, MNRAS, 400, 1541
Abazajian K. N., Adelman-McCarthy J. K., Agüeros M. A., Allam S. S., Allende Prieto C., An D., Anderson K. S. J., Anderson S. F., Annis J., Bahcall N. A., et al. 2009, ApJ. S, 182, 543
Bernardeau F., Juszkiewicz R., Dekel A., Bouchet F. R., 1995, MNRAS, 274, 20
Borgani S., da Costa L. N., Zehavi I., Giovanelli R., Haynes M. P., Freudling W., Wegner G., Salzer J. J., 2000, Astrophysical. J, 119, 102
Branchini E., Eldar A., Nusser A., 2002, MNRAS, 335, 53
Branchini E., Freudling W., Da Costa L. N., Frenk C. S., Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., Zehavi I., 2001a, MNRAS, 326, 1191
Branchini E., Freudling W., Da Costa L. N., Frenk C. S., Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., Zehavi I., 2001b, MNRAS, 326, 1191
Bridle S. L., Zehavi I., Dekel A., Lahav O., Hobson M. P., Lasenby A. N., 2001, MNRAS, 321, 333
Chodorowski M. J., Nusser A., 1999, MNRAS, 309, L30
Colless M., Peterson B. A., Jackson C., Peacock J. A., Cole S., Norberg P., Baldry I. K., Baugh C. M., Bland-Hawthorn J., et al. 2003, ArXiv Astrophysics e-prints
Croft R. A. C., Gaztanaga E., 1998, ApJ, 495, 554

da Costa L. N., Nusser A., Freudling W., Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., 1998, MNRAS, 299, 425

Davis M., Huchra J., Latham D. W., Tonry J., 1982, ApJ, 253, 423

Davis M., Nusser A., Willick J. A., 1996, ApJ, 473, 22

de Lapparent V., Geller M. J., Huchra J. P., 1986, ApJL, 302, L1

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

Dekel A., Bertschinger E., Yahil A., Strauss M. A., Davis M., Huchra J. P., 1993, ApJ, 412, 1

Dekel A., Lahav O., 1999, ApJ, 520, 24

Desjacques V., Crocce M., Scoccimarro R., Sheth R. K., 2010, ArXiv e-prints

Enßlin T. A., Frommert M., Kitaura F. S., 2009, Physical Review D., 80, 105005

Erdoğdu P., Huchra J. P., Lahav O., Colless M., Cutri R. M., Falco E., George T., Jarrett T., Jones D. H., Kochanek C. S., Macri L., Mader J., Martimbeau N., Pahre M., Parker Q., Rassat A., Saunders W., 2006a, MNRAS, 368, 1515

Erdoğdu P., Lahav O., 2009, Physical Review D., 80, 043005

Erdoğdu P., Lahav O., Huchra J. P., Colless M., Cutri R. M., Falco E., George T., Jarrett T., Jones D. H., Macri L. M., Mader J., Martimbeau N., Pahre M. A., Parker Q. A., Rassat A., Saunders W., 2006, MNRAS, 373, 45

Feldman H. A., Hudson M. J., Watkins R., 2008, ArXiv e-prints

Feldman H. A., Watkins R., 2008, MNRAS, 387, 825

Feldman H. A., Watkins R., Hudson M. J., 2010, MNRAS, 407, 2328

Fisher K. B., Lahav O., Hoffman Y., Lynden-Bell D., Zaroubi S., 1995, MNRAS, 272, 885

Freudling W., Zehavi I., da Costa L. N., Dekel A., Eldar A., Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., Zaroubi S., 1999, ApJ, 523, 1

Frisch U., Matarrese S., Mohayaee R., Sobolevski A., 2002a, Nature, 417, 260

Frisch U., Matarrese S., Mohayaee R., Sobolevski A., 2002b, Nature, 417, 260

Giovanelli R., Haynes M. P., Herter T., Vogt N. P., da Costa L. N., Freudling W., Salzer J. J., Wegner G., 1997b, AJ, 113, 53

Giovanelli R., Haynes M. P., Herter T., Vogt N. P., Wegner G., Salzer J. J., da Costa L. N., Freudling W., 1997a, AJ, 113, 22

Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., da Costa L. N., Freudling W., 1994, AJ, 107, 2036

Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., da Costa L. N., Freudling W., 1995, AJ, 110, 1059

Gorski K. M., Davis M., Strauss M. A., White S. D. M.,

Yahil A., 1989, ApJ, 344, 1

Gurovich S., Freeman K., Jerjen H., Staveley-Smith L., Puerari I., 2010, Astrophysical. J, 140, 663

Haugbølle T., Hannestad S., Thomsen B., Fynbo J., Sollerman J., Jha S., 2007, ApJ, 661, 650

Haynes M. P., Giovanelli R., Chamaraux P., da Costa L. N., Freudling W., Salzer J. J., Wegner G., 1999, Astrophysical. J, 117, 2039

Hudson M. J., Dekel A., Courteau S., Faber S. M., Willick J. A., 1995, MNRAS, 274, 305

Hudson M. J., Ebeling H., 1997, ApJ, 479, 621

Hudson M. J., Smith R. J., Lucey J. R., Schlegel D. J., Davies R. L., 1999, ApJL, 512, L79

Jaffe A. H., Kaiser N., 1995, ApJ, 455, 26

Jarosik N., Bennett C. L., Dunkley J., Gold B., Greason M. R., Halpern M., Hill R. S., Hinshaw G., Kogut A., et. al. 2010, ArXiv e-prints

Jones D. H., Read M. A., Saunders W., Colless M., Jarrett T., Parker Q. A., Fairall A. P., Mauch T., Sadler E. M., Watson F. G., Burton D., Campbell L. A., Cass P., et. al. 2009, MNRAS, 399, 683

Jones H., Saunders W., Colless M., Read M., Parker Q., Watson F., Campbell L., 2005, in A. P. Fairall & P. A. Woudt ed., Nearby Large-Scale Structures and the Zone of Avoidance Vol. 329 of Astronomical Society of the Pacific Conference Series, The 6dF Galaxy Survey: First Data Release. pp 11−+

Juszkiewicz R., Ferreira P. G., Feldman H. A., Jaffe A. H., Davis M., 2000, Science, 287, 109

Kaiser N., 1987, MNRAS, 227, 1

Kaiser N., Efstathiou G., Saunders W., Ellis R., Frenk C., Lawrence A., Rowan-Robinson M., 1991, MNRAS, 252, 1

Kauffmann G., Nusser A., Steinmetz M., 1997, MNRAS, 286, 795

Larson D., Dunkley J., Hinshaw G., Komatsu E., Nolta M. R., Bennett C. L., Gold B., Halpern M., Hill e. a., 2010, ArXiv e-prints

Lavaux G., Tully R. B., Mohayaee R., Colombi S., 2010, ApJ, 709, 483

Linder E. V., 2005, Physical Review D., 72, 043529

Lynden-Bell D., Faber S. M., Burstein D., Davies R. L., Dressler A., Terlevich R. J., Wegner G., 1988, ApJ, 326, 19

Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715

Masters K. L., Springob C. M., Haynes M. P., Giovanelli R., 2006, ApJ, 653, 861

Norberg P., Baugh C. M., Hawkins E., Maddox S., Madgwick D., Lahav O., Cole S., Frenk C. S., Baldry I., et. al. 2002, MNRAS, 332, 827

Nusser A., Branchini E., 2000, MNRAS, 313, 587

Nusser A., da Costa L. N., Branchini E., Bernardi M.,

Alonso M. V., Wegner G., Willmer C. N. A., Pellegrini P. S., 2001, MNRAS, 320, L21

Nusser A., Davis M., 1994, ApJL, 421, L1

Nusser A., Davis M., 1995, MNRAS, 276, 1391

Nusser A., Davis M., 2011, arXiv:1101.1650 [astro-ph.CO]

Nusser A., Dekel A., 1993, ApJ, 405, 437

Nusser A., Dekel A., Bertschinger E., Blumenthal G. R., 1991, ApJ, 379, 6

Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144

Peebles P. J. E., 1980, The large-scale structure of the universe. Princeton University Press

Pike R. W., Hudson M. J., 2005, ApJ, 635, 11

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing. Cambridge: University Press, —c1992, 2nd ed.

Radburn-Smith D. J., Lucey J. R., Hudson M. J., 2004, MNRAS, 355, 1378

Reid B. A., Percival W. J., Eisenstein D. J., Verde L., Spergel D. N., Skibba R. A., Bahcall N. A., 23+ SDSS collaborators 2010, MNRAS, 404, 60

Riess A. G., Davis M., Baker J., Kirshner R. P., 1997, ApJL, 488, L1+

Schechter P. L., 1980, Astronomical Journal, 85, 801

Sigad Y., Branchini E., Dekel A., 2000, ApJ, 540, 62

Skrutskie M. F., Cutri R. M., Stiening R., Weinberg M. D., Schneider S., Carpenter J. M., Beichman C., Capps R., Chester T., et al. 2006, AJ, 131, 1163

Springel V., White S. D. M., Jenkins A., Frenk C. S., Yoshida N., Gao L., Navarro J., Thacker R., Croton D., Helly J., Peacock J. A., Cole S., Thomas P., Couchman H., Evrard A., Colberg J., Pearce F., 2005, Nature, 435, 629

Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2007, ApJ. S, 172, 599

Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2009, VizieR Online Data Catalog, 217, 20599

Stark D. V., McGaugh S. S., Swaters R. A., 2009, Astrophysical. J, 138, 392

Strauss M. A., Davis M., 1988, Large-Scale Motions in the Universe: A Vatican study Week. Rubin, V. C. & Coyne, G. V. Eds. . Princeton University Press, NJ, pp 255–274

Strauss M. A., Willick J. A., 1995, Physics Reports, 261, 271

Tully R. B., Shaya E. J., Karachentsev I. D., Courtois H. M., Kocevski D. D., Rizzi L., Peel A., 2008, ApJ, 676, 184

Verde L., Heavens A. F., Percival W. J., Matarrese S., Baugh C. M., Bland-Hawthorn J., Bridges T., Cannon R., Cole S., et. al. 2002, MNRAS, 335, 432

Watkins R., Feldman H. A., Hudson M. J., 2009, MNRAS, 392, 743

Westover M., 2007, PhD thesis, Harvard University

Wild V., Peacock J. A., Lahav O., Conway E., Maddox S., Baldry I. K., Baugh C. M., Bland-Hawthorn J., Bridges e. a., 2005, MNRAS, 356, 247

Willick J. A., Courteau S., Faber S. M., Burstein D., Dekel A., Strauss M. A., 1997, ApJ. S, 109, 333

Yahil A., 1988, Large-Scale Motions in the Universe: A Vatican study Week. Rubin, V. C. & Coyne, G. V. Eds.. Princeton University Press, NJ, pp 219–253

Yahil A., Strauss M. A., Davis M., Huchra J. P., 1991, ApJ, 372, 380

Zaroubi S., 2002, ArXiv Astrophysics e-prints

Zaroubi S., Branchini E., Hoffman Y., da Costa L. N., 2002, MNRAS, 336, 1234

Zehavi I., Zheng Z., Weinberg D. H., Blanton M. R., Bahcall N. A., Berlind A. A., Brinkmann J., Frieman J. A., Gunn J. E., et. al. 2010, ArXiv e-prints

APPENDIX A: A NOTE ON ITERATIVE SCHEME

We caution of a possible systematic bias which may be important in iterative schemes for reconstructing velocities from the distribution of galaxies in redshift space. These schemes rely on the availability of a relation between the peculiar velocity and density in real space, e.g., the linear relation $f \mathrm{div} \mathbf{v} = -\delta$. At the end of any iteration intermediate peculiar velocities are provided, which are used as to derive the distances from the redshifts in the next iteration. Given those distances, the adopted real space relation between δ and $\mathbf{v}$ is then solved to obtain a new guess for the velocities. The loop is continued until the change in the peculiar velocity between successive iterations becomes smaller than a certain threshold. Vanishing peculiar velocities could be taken as input for the first iteration.

We demonstrate here that the real space distribution of galaxies as obtained from the output from any iteration scheme is biased. Hence the corresponding peculiar velocity is also biased. We will first show that a biased distribution in real space is obtained even if unbiased but noisy peculiar velocities are used to get the distances. We write the density of galaxies in real space at distance r in a given direction on the sky as

$$n_{\mathrm{e}}(r)r^2 = (2\pi\sigma^2)^{-1/2} \int \mathrm{d}s\, s^2 n_{\mathrm{s}} \mathrm{e}^{-\frac{[s-v(s)-r]^2}{2\sigma^2}} , \qquad (A1)$$

where $s = r+v$ is the radial redshift space coordinate, $v(s)$ is the peculiar velocity of a galaxy present at s, and σ is the rms of the error in the determination of the $v(s)$. In the above we assume normal error distribution and a one-to-one mapping

between s and v, i.e. we neglect fingers-of-god effects and triple value zones. Working with the variable $r_1 = s - v(s)$ we get

$$n_\mathrm{e}(r)r^2 = (2\pi\sigma^2)^{-1/2} \int \mathrm{d}r_1 r_1^2 n_\mathrm{t}(r_1) \mathrm{e}^{-\frac{(r_1-r)^2}{2\sigma^2}} , \qquad (A2)$$

where

$$n_t(r_1,) = \frac{\mathrm{d}s}{\mathrm{d}r_1} \left[1 + \frac{v(s(r_1))}{r_1}\right]^2 n_\mathrm{s}(s(r_1)) , \qquad (A3)$$

is the actual real space density at r_1. Therefore, errors in the peculiar velocities (even if unbiased relative to the true ones) cause a smearing of structure in the radial direction. This anisotropic smearing is important for scales $\sim \sigma$ (in $\mathrm{km\,s}^{-1}$). The bias is similar to the traditional inhomogeneous Malmquist bias which is usually encountered in studies of distance indicators (Lynden-Bell et al. 1988).

A self-consistent treatment of the bias should take into account the fact that $v(s)$ used in A2 and A3 is the biased peculiar velocity obtained from n_e. This could be done in the far observer limit, $|v/r| << 1$, and for small perturbations where equation A3 reduces to

$$\delta_\mathrm{t} = \delta_\mathrm{s} + \frac{\mathrm{d}v}{\mathrm{d}s} . \qquad (A4)$$

Substituting this into A2 and Fourier transforming the result we get

$$\tilde{\delta}_\mathrm{e}(\mathbf{k}) = \left[\tilde{\delta}_\mathrm{s}(\mathbf{k}) + ik_r\tilde{v}(\mathbf{k})\right]\mathrm{e}^{-k_r^2\sigma^2/2} , \qquad (A5)$$

where k_r is the component of of $\mathbf{k}$ parallel to the line of sight and the tilde denotes quantities in k-space. Using this last equation in the linear $\delta - v$ relation, $f\tilde{v} = -i(k_r/k^2)\tilde{\delta}_\mathrm{e}$, we find

$$\tilde{\delta}_\mathrm{e}(\mathbf{k}) = \frac{\tilde{\delta}_\mathrm{s}(\mathbf{k})\mathrm{e}^{-k_r^2\sigma^2/2}}{1 + f(k_r/k)^2\mathrm{e}^{-k_r^2\sigma^2/2}} , \qquad (A6)$$

instead of the usual unbiased expression obtained with $\sigma = 0$.

A more complete analysis of the bias must incorporate the covariance of the errors in the derived peculiar velocity field.