How does the grouping scheme affect the Wiener Filter reconstruction of the local Universe?

Jenny G. Sorce, Elmo Tempel

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **469,** 2859–2868 (2017)
Advance Access publication 2017 May 10

doi:10.1093/mnras/stx1057

How does the grouping scheme affect the Wiener Filter reconstruction of the local Universe?

Jenny G. Sorce[1,2]★ and Elmo Tempel[2,3]

[1] *CNRS, Observatoire astronomique de Strasbourg, UMR 7550, Université de Strasbourg, F-67000 Strasbourg, France*
[2] *Leibniz-Institut für Astrophysik, An der Sternwarte 16, D-14482 Potsdam, Germany*
[3] *Tartu Observatory, Observatooriumi 1, 61602 Tõravere, Estonia*

Accepted 2017 April 28. Received 2017 April 28; in original form 2017 January 10

ABSTRACT

High-quality reconstructions of the three-dimensional velocity and density fields of the local Universe are essential to study the local large scale structure. In this paper, the Wiener Filter reconstruction technique is applied to galaxy radial peculiar velocity catalogues to understand how the Hubble constant (H_0) value and the grouping scheme affect the reconstructions. While H_0 is used to derive radial peculiar velocities from galaxy distance measurements and total velocities, the grouping scheme serves the purpose of removing non-linear motions. Two different grouping schemes (based on the literature and a systematic algorithm) as well as five H_0 values ranging from 72 to 76 km s^{-1} Mpc^{-1} are selected. The Wiener Filter is applied to the resulting catalogues. Whatever grouping scheme is used, the larger H_0 is, the larger the infall on to the local volume is. However, this conclusion has to be strongly mitigated: a bias minimization scheme applied to the catalogues after grouping suppresses this effect. At fixed H_0, reconstructions obtained with catalogues grouped with the different schemes exhibit structures at the proper location in both cases but the latter are more contrasted in the less aggressive scheme case: Having more constraints permits an infall from both sides on to the structures to reinforce their overdensity. Such findings highlight the importance of a balance between grouping to suppress non-linear motions and preserving constraints to produce an infall on to structures expected to be large overdensities. Such an observation is promising to perform constrained simulations of the local Universe, including its massive clusters.

Key words: methods: numerical – techniques: radial velocities – galaxies: groups: general – large-scale structure of Universe.

1 INTRODUCTION

On large scales, where the gravity prevails, the Universe is homogeneous and isotropic enough for the observed velocity field to reflect the evolution of the large scale structure (LSS) and the total underlying mass (i.e. both baryonic and dark) distribution. Therefore, to study the formation and evolution of the LSS, the analysis of observational radial peculiar velocities plays a major role (e.g. Dekel 1994; Strauss & Willick 1995; Dekel & Ostriker 1999; Willick 1999). Consequently, several techniques have been developed to analyse the observed velocity data sets renewing the effort to measure them (e.g. Mathewson, Ford & Buchhorn 1992; Nusser & Davis 1995; Willick et al. 1997; da Costa et al. 1998; Colless et al. 2001; Springob et al. 2007; Tully et al. 2008; Said, Kraan-Korteweg & Jarrett 2014; Tully, Courtois & Sorce 2016) and leading to numerous studies (e.g. Zaroubi et al. 1997; Theureau et al. 1998; Theureau 1999; Zaroubi et al. 2001; Courtois et al. 2012; Hudson & Turnbull 2012; Rathaus, Kovetz & Itzhaki 2013; Hoffman, Courtois & Tully 2015; Watkins & Feldman 2015; Hoffman et al. 2016). In particular, algorithms have been built to reconstruct from the sparse radial observational data sets, the three-dimensional distribution of matter and the three-dimensional velocity field (e.g. POTENT, Wiener Filter, VIRBIUS, respectively; Dekel et al. 1999; Zaroubi, Hoffman & Dekel 1999; Lavaux 2016). Assuming a cosmological model as a prior, these methods are able to produce density and velocity fields of the local Universe on grids using for sole observational information the sparse and noisy radial peculiar velocity data sets.

In this paper, we focus on the Wiener Filter (WF) algorithm (Zaroubi et al. 1995). This technique is very straightforward and Appendix gives detailed equations. Briefly, based on correlation functions, derivation of matrices and their inverse, the WF permits calculating readily the density and velocity fields assuming as a prior the power spectrum of a given cosmological model. While

★ E-mail: jsorce@aip.de

correlation functions are obtained with the power spectrum, the correlation vectors are derived with radial peculiar velocities called 'the constraints'. These latter must be of high quality to allow exquisite reconstructions of the local Universe. Since the WF is a linear minimal variance estimator, removing non-linear motions in the observational catalogues seems primordial. A grouping scheme permits gathering galaxies that belong to a single cluster or group into one point. Subsequently, it produces one linear constraint (one position and radial peculiar velocity) against several non-linear constraints that would damage the reconstruction obtained with the WF. We thus seek to understand the impact of the chosen grouping scheme applied to the observational constraints on the resulting reconstructions. In addition, in view of the recent concerns and discrepancies regarding the Hubble constant value (see Jackson 2015, for a review), we wish to study also the differences between reconstructions obtained with observational catalogues derived using different Hubble constant values. The Hubble constant permits indeed converting distances in h^{-1} Mpc units and most importantly it allows us to derive galaxy peculiar velocities from galaxy total velocities and distance measurements. Constraints are derived from the second sparse and noisy observational distance data set of the Cosmicflows project[1] (Tully et al. 2013).

To summarize, this paper aims at determining the variance of the WF reconstruction with respect to the Hubble constant and grouping scheme choices. The final goal is to select the best choices to build constrained initial conditions of the local Universe within the CLUES[2] collaboration (Gottlöber, Hoffman & Yepes 2010). To study in detail our cosmic environment, the resulting performed simulations should resemble the local Universe down to the clusters. In particular, we expect to optimize the reproduction of the local massive clusters that have been slightly under massive so far, if not for the Virgo cluster (Sorce et al. 2014, 2016a,b; Carlesi et al. 2016a,b). These constrained simulations are the starting point of several projects to study the local Universe in detail, to understand our local environment and to compare it with observations.

This paper starts with a section describing the observational catalogue of radial peculiar velocities or more precisely of galaxy direct distance measurements and the two grouping schemes compared here (Tully, private communication; Tempel et al. 2016). In a subsequent section, the WF algorithm is applied to the observational catalogue grouped with the different schemes and applying different Hubble constant numerical values. First, the effect of the Hubble constant on the reconstructed velocity and overdensity fields are studied then, the impact of the grouping scheme on the reconstructed fields is analysed in detail. An additional analysis made after minimizing the biases in the observational catalogues permits tempering the results. A conclusion presenting the best strategy for the next step (building constrained initial conditions) closes the paper.

2 GROUPING AND RECONSTRUCTION TECHNIQUES

2.1 The catalogue

The second generation catalogue built by the Cosmicflows collaboration is a large publicly released catalogue of radial peculiar velocities or more precisely of direct distance measurements. Published in Tully et al. (2013), it contains more than 8000 galaxy direct distance estimates. These measurements come mostly from the Tully–Fisher (Tully & Fisher 1977) and the Fundamental Plane (Colless et al. 2001) methods. Cepheids (Freedman et al. 2001), tip of the red giant branch (Lee, Freedman & Madore 1993), surface brightness fluctuation (Tonry et al. 2001), supernovae of Type Ia (Jha, Riess & Kirshner 2007) and other miscellaneous methods also contribute to this large data set though to a minor extent (∼12 per cent of the data). Using $H_0 = 75.2$ (=100 h) km s^{-1} Mpc^{-1} (the value given by Tully et al., 2013), it extends up to about 250 h^{-1} Mpc and about 50 per cent of the data are within 70 h^{-1} Mpc and 90 per cent within 160 h^{-1} Mpc. In a companion paper (Sorce, Hoffman & Gottlöber 2017), we have shown that, in absence of a complete catalogue and provided that it is properly grouped, the sampling of this catalogue is optimal for WF reconstructions with respect to uniformly distributed catalogues or catalogues of sole clusters. The goal is then to track the impact of the grouping technique on the resulting reconstructions.

2.2 The grouping schemes

A grouped version designed by Tully, hereafter referred to as Tully Grouping Scheme, and released through the Extragalactic Distance Database[3] was used to build the first generation of constrained initial conditions that result in simulations resembling the local Universe down to 2–3 h^{-1} Mpc (Sorce et al. 2016b). However, clusters reveal themselves to be under massive except for the Virgo cluster (Sorce et al. 2016a), thanks to the prior minimization of biases introduced by Sorce (2015) that reduces the infall on to the local volume, leads the monopole of the velocity field to zero and Gaussianizes the distribution of observed radial peculiar velocities.

The difficulty resides in the definition of 'group' itself. If on the simulation side, groups are well defined thanks to an access to the entire three-dimensional information, on the observational side, calling an ensemble of galaxies a group constitutes a great challenge because of a restricted access to the information. In observations, knowing precisely the fraction of collapsed material becomes quite problematic. Still several schemes have been developed to define groups within galaxy catalogues. They mainly invoke Friends of Friends (FoF) like algorithms based on projected separation, radial velocities and even luminosities to identify what are called 'groups' of galaxies (e.g. Huchra & Geller 1982; Geller & Huchra 1983; Ramella et al. 2002; Eke et al. 2004; Yang et al. 2005; Crook et al. 2007; Lavaux & Hudson 2011; Makarov & Karachentsev 2011; Old et al. 2014; Tempel et al. 2014; Old et al. 2015). This paper does not aim at scrutinizing in detail the methods used to group catalogues. It aims at testing two recently released versions of groups for galaxies in the local Universe to understand the differences in the reconstructions generated by two various grouping schemes as described below:

(i) Tempel et al. (2016) introduced a new grouping method (hereafter Tempel Grouping scheme). This method is based on a widely used FoF percolation method, where different linking lengths in radial (along the line of sight) and in transversal (in the plane of the sky) directions are used but the conventional FoF groups are refined using multimodality analysis. More precisely, Tempel et al. (2016) use a model-based clustering analysis to check the multimodality of groups found by the FoF algorithm and they separate nearby/merging systems. In the current paper, we use published catalogues of groups detected using this new method.

(ii) Tully Grouping scheme is based on literature groups and in that respect is not a systematic scheme. Within 30 Mpc, groups are

[1] http://www.ipnl.in2p3.fr/projet/cosmicflows/
[2] https://www.clues-project.org/
[3] http://edd.ifa.hawaii.edu/

Table 1. Properties of the catalogue of constraints (radial peculiar velocities) according to the grouping scheme: (1) grouping scheme including a short description of the constituents of the catalogue after grouping; (2) Hubble constant; (3) mean velocity; (4) standard deviation of the velocity distribution; (5) skewness of the distribution; and (6) flatness of the distribution.

(1) Grouping scheme + description	(2) H_0 (km s^{-1} Mpc^{-1})	(3) $\langle v \rangle$ (km s^{-1})	(4) σ_v (km s^{-1})	(5) Skewness	(6) Kurtosis
Tully:	72	17.2	1462	−0.16	7.0
4098 isolated, 910 groups	73	− 74.9	1476	−0.29	7.0
Total: 5008 constraints	74	− 166	1493	−0.42	6.9
444 groups with one distance measurement	75	− 258	1511	−0.55	6.9
On average, 4.5 distance measurements per group	76	− 350	1532	−0.66	6.9
Tempel:	72	59.5	1388	−0.055	7.1
3218 isolated, 2344 groups	73	− 27.7	1401	−0.20	7.0
Total: 5562 constraints	74	− 115	1417	−0.34	6.9
1516 groups with one distance measurement	75	− 202	1434	−0.47	6.9
On average, 2.1 distance measurements per group	76	− 289	1454	−0.60	6.8

those identified by Tully (1987), further away groups are those given in the literature like Abell's catalogue (Abell, Corwin & Olowin 1989). Recently, Tully (2015a,b) published a more systematic way of deriving groups based on radii of second turn around and iterations. After comparisons, we find that the catalogue grouped with this last scheme is an intermediate between the catalogues obtained with Tully and Tempel Grouping schemes and as such will result in more mitigated conclusions would we compare it to Tempel Grouping scheme. In addition, Tully Grouping scheme has been used so far with the second catalogue to build constrained initial conditions. We thus stick to Tully Grouping scheme in the rest of the paper.[4]

Tully and Tempel Grouping schemes provide the groups to which the different galaxies that populate the second catalogue of Cosmicflows belong to as well as their total velocity (derived from the observed redshift). We note that the grouping schemes deliver groups built with a complete down to a magnitude limit sample of galaxies. Then, galaxies from the second catalogue of Cosmicflows are distributed into these groups and only the groups to which they belong are retained for further use. The second catalogue of Cosmicflows gives the individual distance modulus (μ) measurements of each galaxy and their uncertainty (σ_μ). To determine the radial peculiar velocity of the groups and their position in real space (by opposition with redshift space), we proceed as follows (Tully, private communication):

$$\mu_{\rm g} = \frac{\sum {\rm w} \times \mu}{\sum {\rm w}};\ \sigma_{\mu {\rm g}} = \sqrt{\frac{1}{\sum {\rm w}}} \text{ where } {\rm w} = \frac{1}{\sigma_\mu^2}, \quad (1)$$

$$d_{\rm g} = 10^{\frac{\mu_{\rm g}-25}{5}};\ \sigma_{\rm dg} = \sigma_{\mu {\rm g}} \times \frac{\log(10)}{5}, \quad (2)$$

$$v_{\rm pec\ g} = v_{\rm tot\ g} - H_0 \times d_{\rm g};\ \sigma_{\rm vpec\ g} = \sigma_{\rm dg} \times d_{\rm g} \times H_0, \quad (3)$$

where the subscript 'g' stands for 'grouped' value and σ for the uncertainty of the given subscript value, d is the distance in real space, $v_{\rm tot}$ is the total velocity of the galaxy/group and $v_{\rm pec}$ is the radial peculiar velocity.

Table 1 reflects the resulting grouped catalogues after application of the two schemes. The first column shows interestingly that while Tully scheme results in more isolated galaxies (i.e. single position and peculiar velocity as constraint for the WF algorithm), Tempel scheme gives less isolated galaxies but more groups (2344 against 910). Overall, Tully scheme is more aggressive than Tempel scheme. While, on average, there is 4.5 distance measurements per group with Tully scheme, there is, on average, only 2.1 distance measurements per group with Tempel scheme. However, this difference could be due to the absence of group identification in Tully scheme when there is only one galaxy measurement. Indeed, summing in both cases the number of isolated galaxies and that of groups with only one measurement, the numbers become similar (4542 for Tully versus 4734 for Tempel). However, excluding groups with a single measurement, there is still, on average, more distance measurements per group with Tully scheme (7.8) than with Tempel scheme (4.1), confirming that Tully scheme groups more (number of groups with more than one measurement about twice smaller). In total, Tully scheme provides 5008 constraints against 5562 for Tempel scheme.

Table 1 also gives the properties of the resulting radial peculiar velocity distributions according to both the grouping scheme and H_0 ranging from 72 to 76 km s^{-1} Mpc^{-1}. The larger H_0 the smaller and more negative the mean velocity, the larger the standard deviation, the more asymmetric and less flat the distribution for both grouping scheme. However, whatever H_0 value considered, Tempel scheme results in smaller mean, standard deviation and skewness values. Note how the mean trend changes for a H_0 value of between 72 and 73 km s^{-1} Mpc^{-1} in both cases.

2.3 The WF technique

We apply the WF technique to the 10 catalogues obtained with the 5 different H_0 values and the two grouping schemes using Planck power spectrum (Planck Collaboration XVI 2014) as a prior. One might argue that using a different H_0 value to build the catalogue of constraints and the cosmological prior could bias the results. Note that tests we made changing the prior (for instance, using WMAP7 instead of Planck power spectrum) show that the prior has only a very small, thus negligible, impact on the reconstruction with respect to the parameters (grouping scheme and H_0 in the observational data) tested in this paper. In other words, the variance between reconstructions obtained with different priors with all the other parameters fixed is much smaller than the variance between reconstructions produced with the same prior but changing H_0 or the grouping scheme.

[4] Note that we reproduced the work with the 2015 Tully Grouping scheme and found that it gives as expected intermediate results between Tully and Tempel Grouping schemes.

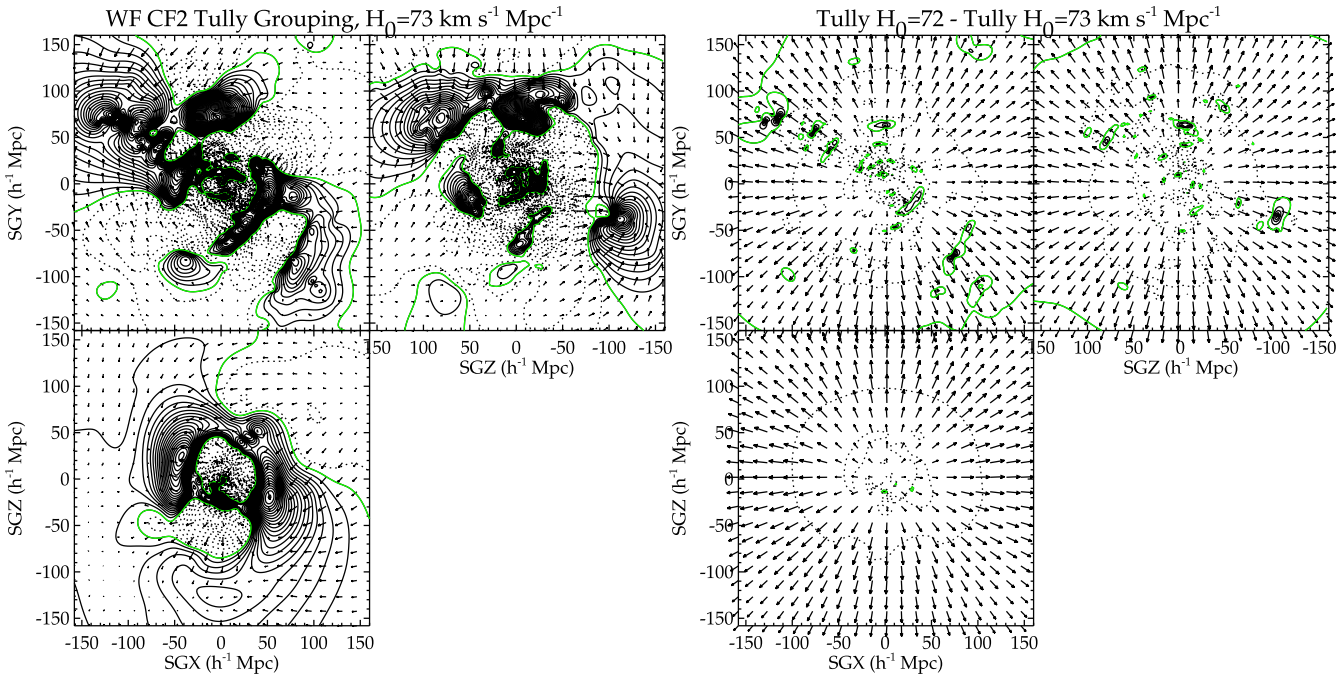

Figure 1. Left-hand panel: supergalactic *XY*, *YZ* and *XZ* slices of the reconstructed velocity (arrow) and overdensity (contour) fields of the local Universe obtained with the catalogue grouped with Tully scheme and $H_0 = 73$ km s^{-1} Mpc^{-1}. The green colour stands for the mean field. Dashed contours are underdense regions while solid contours are overdense areas. The reconstruction shows overall the local structures such as Shapley (top left-hand side in *XY*), Coma (top middle in *XY* and *ZY*) and Perseus Pisces (bottom right-hand side in *XY*). Right-hand panels: same as the left-hand panels but for the residual between reconstructions obtained with Tully Grouping scheme and $H_0 = 72$ and 73 km s^{-1} Mpc^{-1}, respectively. The residual highlights the impact of the Hubble constant value chosen to derive the distances and thus the peculiar velocity constraints. The larger H_0 is, the greater the infall on to the local volume is.

A boxsize of 500 h^{-1} Mpc is retained as the adequate size to contain all the data-constraints. Note that from now on, the discussion will be led in h^{-1} Mpc. Namely, once H_0 has been chosen, every distance is converted in h^{-1} Mpc such that $H_0 = 100$ h km s^{-1} Mpc^{-1}. A grid size of 256^3 cells permits a resolution about 2 h^{-1} Mpc, the linear theory threshold, in agreement with the maximum resolution of the linear WF method. This ensures that differences observed between reconstructions are solely due to the tested parameters and not to non-linear statistical fluctuations.

Additionally, non-linear sigmas, explained in more detail in Appendix, are essential to account for the residual of non-linearities in the data sets. Indeed, even grouped catalogues still contain non-linearities, especially in high-density regions with a poor sampling. The non-linear sigmas correspond to a small additional smoothing applied to the constraints to compensate for their non-linear component that cannot be accounted for directly by the linear WF technique. They are simply added in quadrature to the uncertainties of the constraints. Non-linear sigmas of the same order of magnitude (100–200 km s^{-1}) are found to be required for the different catalogues. Such similar values will prevent any difference due to a significant change in the smoothing. These non-linear sigmas are essential to ensure that only significant differences remained visible between reconstructions obtained with various parameters.

3 WF RECONSTRUCTIONS

3.1 Tully grouping: the results

The left-hand panel of Fig. 1 shows the reconstructed velocity and overdensity fields obtained with the catalogue grouped with Tully scheme and using $H_0 = 73$ km s^{-1} Mpc^{-1}. The right-hand panel of the same figure presents the residual between the reconstructed fields obtained with two different H_0 values but the same grouping scheme. The effect is clear, the larger H_0 is, the greater the infall on to the local volume is. Namely, H_0 value impacts the tidal part of the velocity field.[5] However, the overdensity field is not that much affected: there are only very small and sparse residual contours. It means that H_0 value influences only weakly the divergent part of the velocity field directly linked to the overdensity field. Note that this is the part of the velocity field used to build constrained initial conditions. The infall observed with larger value of H_0 impacts the global density of the local volume. With a smaller value of H_0, not only the infall but also the global local density decreases: The spherical dashed contours on the right-hand side of Fig. 1, indicate, indeed, that globally the reconstructed field obtained with the highest value of H_0 has higher overdensity values than that obtained with the smallest value of H_0. Since an underdensity of the local volume is not excluded (e.g. Keenan, Barger & Cowie 2013) while a large infall on to the local volume is very unlikely, the smallest values of H_0 tested here might be preferred. However, in the last part of this section, we will temper this conclusion by applying the bias minimization scheme introduced in Section 2 and that needs to be applied to the observational catalogue.

The first half of Table 2 summarizes the properties of the reconstructions obtained with Tully Grouping scheme and different H_0 values to support our findings based on Fig. 1. On the one hand, it clearly shows that for large H_0 values, the infall is large: The monopole term of the velocity field is highly negative at large radii.

[5] The velocity field can be decomposed into two components, the tidal part due to the objects outside of the volume considered and the divergent part generated by the objects within the volume considered.

Table 2. Properties of the reconstructed velocity and overdensity fields for different H_0 values and grouping schemes: (1) grouping scheme; (2) Hubble constant; (3) standard deviation of the velocity field; (4) standard deviation of the overdensity field; (5) dipole value of the velocity field at 10 h^{-1} Mpc; (6) dipole value of the velocity field at 240 h^{-1} Mpc, the edge of the box/data; and (7) monopole value of the velocity field at 240 h^{-1} Mpc.

(1) Grouping scheme	(2) H_0 (km s^{-1}Mpc^{-1})	(3) σ_v (km s^{-1})	(4) σ_ρ	(5) Dipole at $r = 10\ h^{-1}$ Mpc (km s^{-1})	(6) Dipole at $r = 240\ h^{-1}$ Mpc (km s^{-1})	(7) Monopole at $r = 240\ h^{-1}$ Mpc (km s^{-1})
	72	320	0.20	477	144	− 83
	73	332	0.21	476	144	− 359
Tully	74	369	0.21	478	144	− 634
	75	424	0.22	477	144	− 914
	76	491	0.23	479	144	− 1192
	72	324	0.22	476	138	− 62
	73	339	0.22	445	138	− 264
Tempel	74	369	0.22	478	140	− 611
	75	615	0.23	624	140	− 1201
	76	495	0.24	420	183	− 797

The infall for larger H_0 values, deduced from the observed outflow in the subtraction of reconstructions obtained with increasing values of H_0 in the right-hand panel of Fig. 1, is confirmed. At both large and small radii, the dipole of the velocity field is on the other hand quasi-unchanged, in agreement with the fact that the overdensity (or divergent part of the velocity) field is quite unaffected by a change in H_0 value. These two points are visible in another form on Fig. 2 where both monopole and dipole of the velocity field are shown at all radii. While the dipole is quite independent of H_0 value at all radii, the monopole tends to get smaller and smaller at all radii with H_0 getting larger and larger. On another aspect, the standard deviation of the overdensity and velocity fields increases slightly with the value of H_0.

While Table 2 shows properties of reconstructions obtained with different H_0 values independently of each other, the first third of Table 3 summarizes the comparisons between reconstructed fields obtained with different H_0 values but the same (Tully) grouping scheme. Standard deviation of the residual between two different H_0 reconstructed overdensity and velocity fields obviously increase with the difference between the two H_0 values but are quite stable for a given difference between the two H_0 values. In any case, the standard deviation of the residual is smaller than the standard deviation of the compared velocity and overdensity fields taken independently except when the reconstructed velocity field obtained with 76 km s^{-1} Mpc^{-1} is compared to that obtained with the smallest H_0 value (i.e. 72 km s^{-1} Mpc^{-1}), namely when the separation between H_0 values, chosen for this paper, is maximal. Regardless, 76 km s^{-1} Mpc^{-1} seems to be a very unlikely value in light of the above observations.

3.2 Tempel Grouping: the results

Fig. 3 shows the reconstructed velocity and overdensity field obtained with $H_0 = 73$ km s^{-1} Mpc^{-1}, like in Fig. 1, but with Tempel Grouping scheme. The observations made in the previous section still stand. Namely, H_0 value impacts clearly the tidal part of the velocity field while it barely affects the overdensity or divergent part of the velocity field. As H_0 gets larger, the infall on to the local volume increases.

The second part of Table 2 summarizes the different values obtained for the reconstructions obtained with Tempel Grouping scheme and different H_0 values. Again, the same findings as with Tully Grouping are valid except that the standard deviations of both

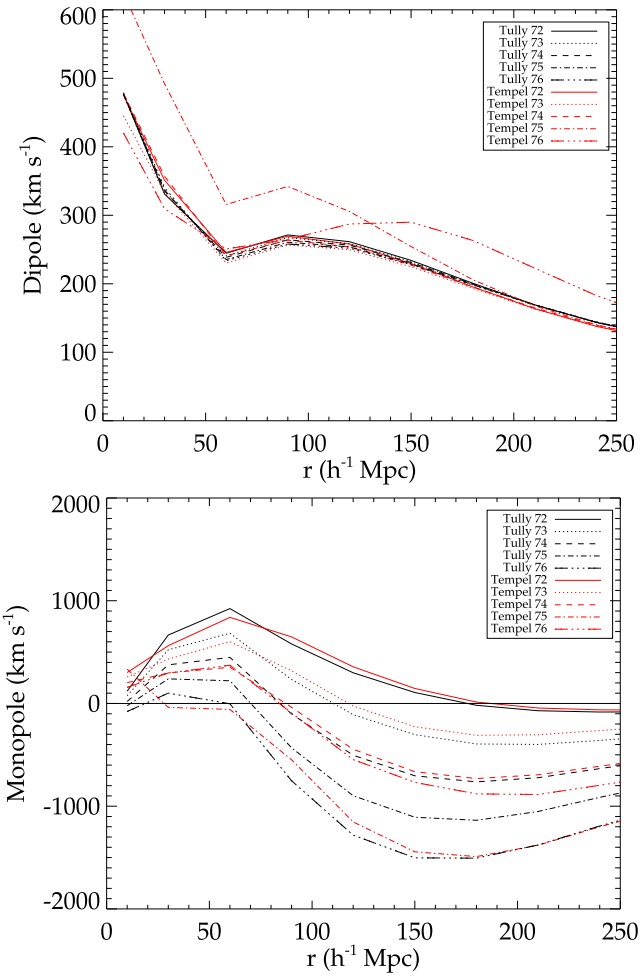

Figure 2. Dipole (top panel) and monopole (bottom panel) of the WF reconstructed fields for different H_0 values (linestyle) and different grouping schemes (colour), as a function of the distance from us. If H_0 impacts only weekly the dipole, the monopole term is profoundly affected. The larger H_0, the larger is the infall on the local volume. The grouping scheme has only a weak influence on the dipole and monopole of the velocity field except very locally.

Table 3. Properties of the residual between reconstructed velocity and overdensity fields obtained with different H_0 values and different grouping schemes: (1) grouping scheme of the reconstruction number 1 – grouping scheme of the reconstruction number 2; (2) H_0 value used for the first reconstructed field – H_0 value used for the second reconstructed field; (3) standard deviation of the residual velocity field; (4) standard deviation of the residual overdensity field; and (5) maximum of the residual overdensity field.

(1) Grouping scheme 1–2	(2) H_0 1–2 (km s^{-1} Mpc^{-1})	(3) σ_v (km s^{-1})	(4) σ_ρ	(5) ρ_{max}
	72–73	95	0.03	1.0
	72–74	187	0.06	1.4
	72–75	282	0.1	1.5
	72–76	375	0.13	2.1
Tully–Tully	73–74	93	0.03	0.9
	73–75	187	0.06	1.2
	73–76	280	0.1	2.1
	74–75	94	0.03	0.9
	74–76	187	0.06	1.3
	75–76	93	0.03	1.3
	72–73	128	0.03	1.0
	72–74	187	0.06	1.4
	72–75	448	0.1	2.0
	72–76	427	0.1	2.4
Tempel–Tempel	73–74	147	0.03	0.8
	73–75	401	0.06	1.3
	73–76	383	0.1	1.8
	74–75	298	0.03	0.9
	74–76	366	0.06	1.1
	75–76	519	0.03	0.8
	72–72	43	0.07	4.0
	73–73	111	0.07	4.1
Tully–Tempel	74–74	43	0.07	4.3
	75–75	254	0.07	4.2
	76–76	400	0.07	4.7

the velocity and overdensity fields are slightly higher, a first hint that structures are more contrasted in the WF reconstructions obtained with Tempel Grouping. Tempel Grouping reconstructions are also less affected by the infall or in other words for a given H_0 value, the monopole term is less negative in the reconstructions obtained with Tempel Grouping than with Tully Grouping. The dipole varies slightly more in Tempel Grouping scheme's case than in Tully Grouping scheme's case probably because of the higher number of constraints: At small radii, the larger number of constraints generates more non-linearities; at large radii, the larger number of constraint slows the fields in their pace to reach the mean value. The two largest values of H_0 (75 and 76 km s^{-1} Mpc^{-1}) present exceptions that deserve attention. A value of 75 km s^{-1} Mpc^{-1} results in a larger dipole value than the average at small radii while a value of 76 km s^{-1} Mpc^{-1} gives a field with a larger dipole value than the average at large radii. In addition, the monopole value at large radii for $H_0 = 75$ km s^{-1} Mpc^{-1} is extremely high in absolute value. It clearly looks like there is a transition between values of 74 and 75 km s^{-1} Mpc^{-1} linked to the grouping scheme since none of these observations are valid for Tully Grouping scheme. This seems to imply that a more aggressive grouping has to be preferred for a better stability of the dipole and monopole of the velocity field whatever H_0 value is used.

Tests we made varying the default linking length (0.25 h^{-1} Mpc at redshift zero changed to 0.20 or 0.30 h^{-1} Mpc) in Tempel Grouping scheme and applying the WF technique to the resulting grouped catalogues show that indeed a large linking length permits increasing the stability but an excessive grouping (no more field galaxies) leads to wrong dipole values. This is in agreement with Sorce et al. (2017) that show that galaxies in the fields are an absolute necessity. Additionally, H_0 has to be chosen with more care: Minimizing in absolute value, the mean of the velocity distribution seems a reasonable approach to choose the value of H_0. Again, we will temper this conclusion within the last part of this section.

The second and third columns of Table 3 show the properties of the residual between reconstructions obtained with Tempel

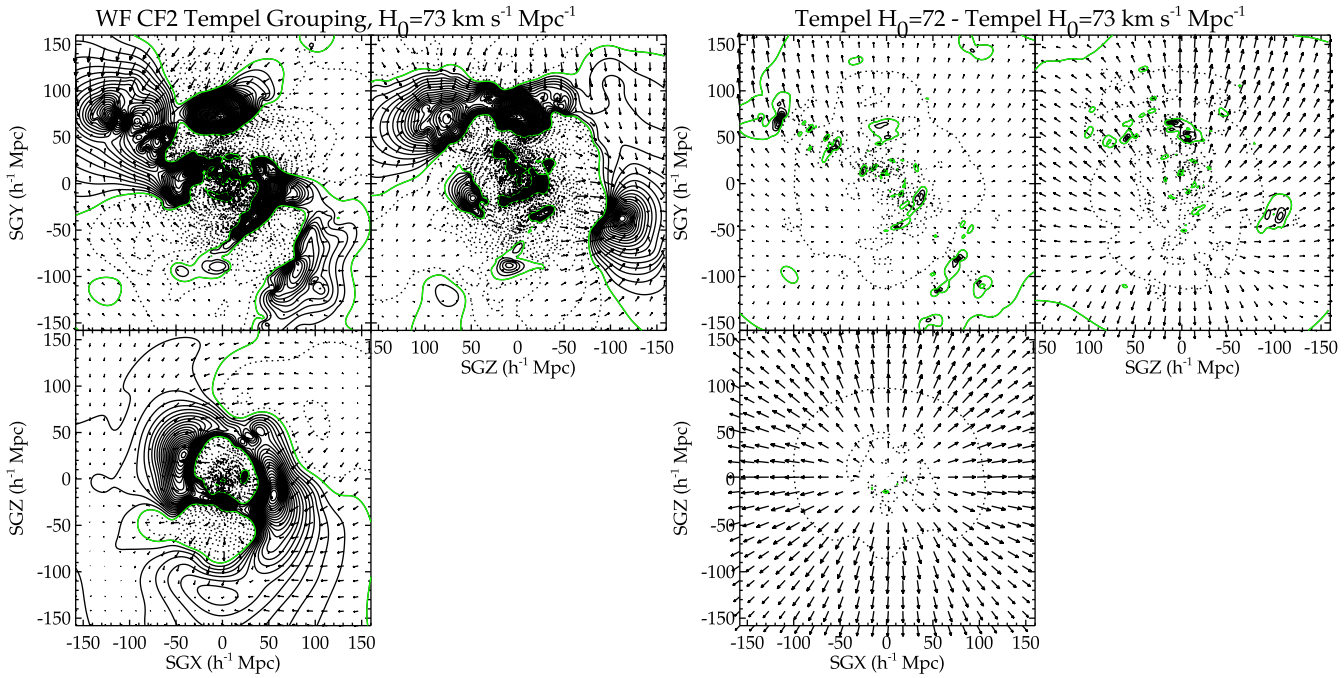

Figure 3. As Fig. 1 but obtained with Tempel Grouping scheme.

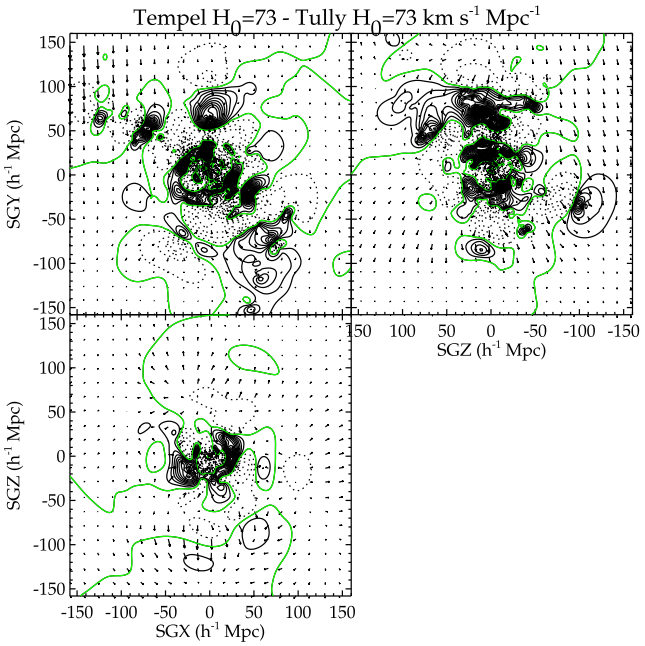

Figure 4. Supergalactic *XY*, *YZ* and *XZ* slices of the residual between reconstructed velocity (arrow) and overdensity (contour) fields of the local Universe obtained with the catalogue grouped with Tempel and Tully schemes. The green colour stand for the null value. The residual shows that overall the local structures such as Coma (top middle in *XY* and *ZY*) are more pronounced in the reconstruction obtained with Tempel Grouping scheme than in that obtained with Tully Grouping scheme.

Grouping scheme but different H_0 values. Overall, the same observations as with Tully Grouping scheme apply. One might notice that the residual values are larger than those obtained with Tully Grouping scheme. This is again due to the aggressiveness of the grouping. Indeed, in the tests made varying the default linking length, we observe that the variance between the two reconstructions obtained with different H_0 values is larger for the smallest linking length than for the default linking length used in the tests. Namely, grouping more eases slightly the dependence on H_0.

3.3 Comparisons between the grouping schemes

Fig. 4 shows the residual between two WF reconstructions obtained with a different grouping scheme but with the same H_0 value. The figure is clear and irrevocable: While the velocity field is weakly affected by a different grouping scheme, the density field is largely impacted. Note that the position of structures is not impacted, structures are reconstructed at the proper location in both grouping scheme cases but their density value varies. In other words, the infall on to the large structures is slightly more important with Tempel Grouping scheme than with Tully Grouping scheme.

Actually, the last third column of Table 3 gives the standard deviation of the residual velocity and overdensity fields of two reconstructions obtained with the exact same H_0 value but different grouping schemes. These values confirm that the velocity fields are quasi-identical except for the largest value of H_0 (76 km s^{-1} Mpc^{-1}) but this value has been shown to be slightly unrealistic. This is in agreement with the fact that at larger H_0 values, Tully Grouping scheme produces overall a larger infall than Tempel Grouping scheme with the exception of $H_0 = 75$ km s^{-1} Mpc^{-1}. Interestingly, the standard deviation of the residual overdensity fields is not exceptionally high, although Fig. 4 clearly shows that the structures are affected by the grouping schemes. The answer is in the maximum value of the residual overdensity fields. The standard deviation might be quite low but the maximum value is higher than when comparing, for example, two reconstructions obtained with different H_0 values but with the same grouping scheme. Consequently, on average, the values of the residual are higher for reconstructions obtained with different grouping schemes than with different H_0 values confirming the observations made looking at the right-hand panels of Figs 1 and 3 and at Fig. 4.

To understand the difference emanating from the two grouping schemes in more detail, we look at the distribution of constraints in the *XY* supergalactic slice of the local Universe. Fig. 5 shows the constraints as dots at galaxies' position: A blue dot means a radial peculiar velocity pointing towards us while a red dot stands for a radial peculiar velocity going away from us. The dot sizes are proportional to the radial peculiar velocity value in absolute value. In a first approximation, i.e. on large scales, the distributions of constraints and their values look overall very similar. Next, we focus on particular regions of interest such as the Coma cluster area that has been shown to present a structure with a greater contrast using Tempel Grouping rather than Tully Grouping. This particular region is plotted in the small top right-hand inset in both panels of Fig. 5. The differences are striking: An infall from both sides of the Coma cluster region (red and blue) is visible in Tempel Grouped catalogue, while the more aggressive grouping used by Tully removes most of these constraints. The same goes for the Centaurus cluster region visible in the small bottom left-hand inset in both panels of the same Figure. Tests made using different linking lengths in Tempel Grouping scheme confirm that a lesser grouping increases the infall/outflow and thus the contrasts of structures. This highlights the importance of a balance between grouping and removing non-linear motions.

3.4 H_0: not a real dependence

In this last part, we investigate whether the WF reconstruction has a real strong dependence on the Hubble constant value. Indeed, in the above tests, the bias minimization scheme developed originally to suppress the infall observed in the reconstructions has not been applied to the observational catalogue. However, to build adequate constrained initial conditions, the observational catalogue must undergo a bias minimization. We apply the method described in Sorce (2015) to the different H_0 value catalogues grouped with Tempel scheme and run the WF technique on each one of them. Results are visible on Fig. 6 in form of the monopole of the velocity fields. The bias minimization scheme strongly reduces the effect of the H_0 value selected to derive the peculiar velocities. There is clearly no strong infall anymore on to the local volume (no large negative values for the monopole at large radii) whatever H_0 value is used. This observation drastically minimizes the previous conclusions about the dependence of the reconstruction on H_0 and removes concerns about choosing adequately H_0 providing that the catalogue is bias minimized.

4 CONCLUSION

Reconstructions of the three-dimensional velocity and density fields of the local Universe are essential to study the local large scale structure. Numerous methodologies have been developed to perform such reconstructions using observational data. In this paper, we use the WF technique applied to galaxy radial peculiar velocity catalogues to obtain reconstructed velocity and overdensity fields

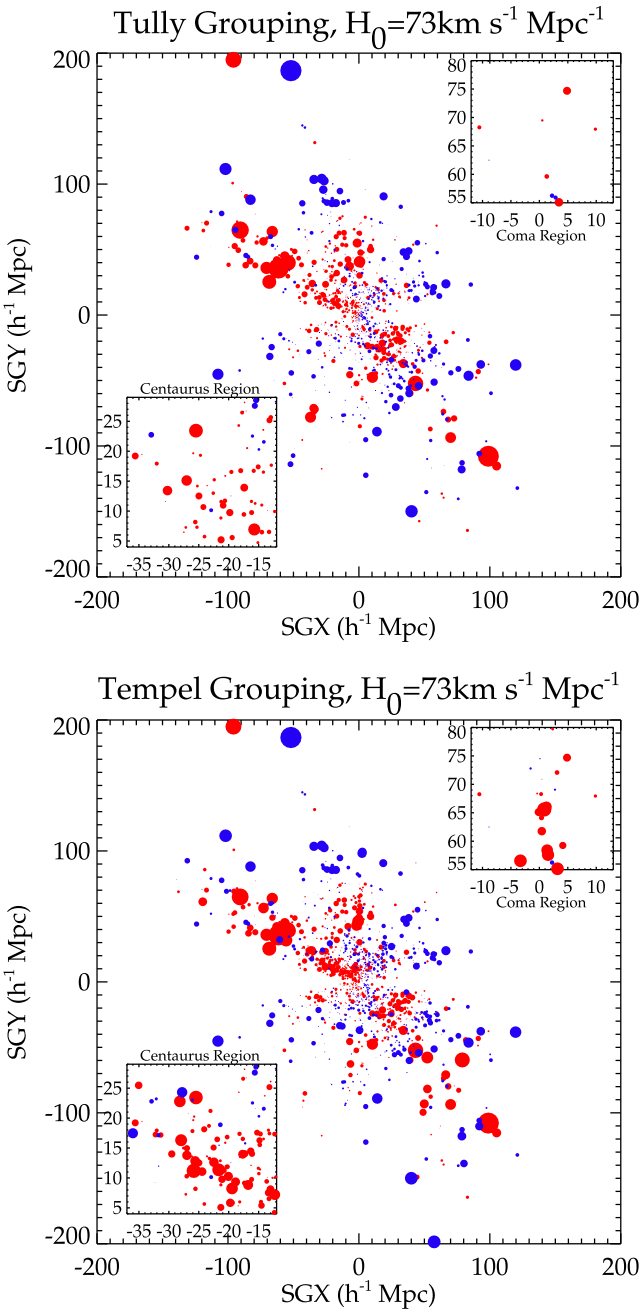

Figure 5. *XY* supergalactic slice (10 h^{-1} Mpc) of the local Universe showing the constraints (radial peculiar velocity at galaxies' position) obtained with Tully (top panel) and Tempel (bottom panel) Grouping schemes. A red dot means that the radial peculiar velocity is positive while a blue dot means that it is negative. The dot size is proportional to the absolute value of the radial peculiar velocity. Overall, the two grouping scheme exhibits catalogues in agreement with each other, the constraints are quite similar. However, zooming on a particularly dense region, like the Coma cluster area or the Centaurus cluster region, differences are more pronounced. Tempel Grouping scheme presents more constraints with large values reinforcing the infall on to Coma/Centaurus (from both sides) that explains the contrast between Coma/Centaurus areas reconstructed using the second catalogue of Cosmicflows obtained with the two different grouping schemes.

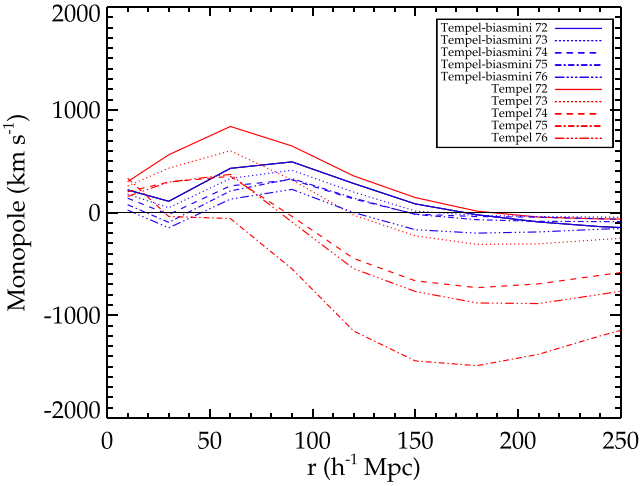

Figure 6. Monopole of the WF reconstructed fields as a function of the distance from us. The WF has been applied to catalogues with different H_0 values (linestyle) using a unique grouping scheme but applying (blue) or not (red) a method to minimize the observational biases. H_0 impacts strongly the monopole term only for the catalogues without minimization of biases. The larger H_0, the larger is the infall on the local volume. The minimization bias scheme has a strong influence on the monopole of the velocity field: it clearly suppresses the infall for all the values of H_0 considered.

of the local volume. These reconstructions are useful as such for direct study of the linear local Universe today but also to build constrained initial conditions that permit performing constrained simulations of the local Universe, i.e. simulations that resemble the local Universe down to the cluster scales. We seek to understand how the Hubble constant value chosen to derive the radial peculiar velocities from galaxy distance measurements and total velocities, and the grouping scheme used to remove non-linear motions affect the reconstructions and by extension impact the quality of the constrained simulations.

To this end, two different grouping schemes (Tully based on the literature and Tempel based on a systematic algorithm) are selected as well as five reasonable locally derived H_0 values (from 72 to 76 km s^{-1} Mpc^{-1}, for the most recent values see e.g. Beaton et al. 2016; Riess et al. 2016; Singh et al. 2016; Tully et al. 2016). A total of 10 grouped versions of the second radial peculiar velocity catalogue of Cosmicflows are produced accordingly: 5 per grouping scheme with each one of the H_0 values. These catalogues differ by the number of isolated galaxies and groups as well as by their radial peculiar velocity distribution. Tully Grouping scheme results in more isolated galaxies but less groups and as a result less peculiar velocity-constraints when compared to Tempel Grouping scheme. Namely, the latter is found to be more aggressive than the former. In addition, the larger H_0, the more asymmetric the distribution, the larger the standard deviation and the more negative the mean.

The WF algorithm is applied to these 10 catalogues and the resulting velocity and overdensity fields are compared. Whatever grouping scheme is used, the larger H_0 is, the larger the infall on to the local volume is. If the tidal part of the velocity field due to objects outside of the local volume is greatly affected by H_0, the divergent part due to the objects inside the Volume and tightly tied to the overdensity field is weakly impacted by a change in H_0. Note that it is the latter that is used to build constrained initial conditions. Actually, the latter is greatly affected by the grouping scheme. Comparing at fixed H_0, reconstructions obtained with catalogues grouped with

different grouping schemes, we observe that structures, although they are present at the proper location in both cases, are more contrasted in Tempel Grouping scheme's case then in Tully Grouping scheme's case. This is in particular true for the Coma cluster area and the Centaurus cluster region. Looking for the reasons of such observations, we compare the distribution of radial peculiar velocity in the *XY* supergalactic slice and notice that overall the agreement between the catalogues grouped with the two different schemes is very good: positions of constraints (peculiar velocities) and their values match quite closely. However, when focusing on smaller areas to study the details, like the Coma cluster region or the Centaurus cluster region, we note quite a lot of differences mostly due to the difference in terms of aggressiveness of the grouping schemes. Tempel Grouping scheme allows more constraints in these regions than Tully Grouping scheme. Consequently, the infall from both sides on to these areas are reinforced providing an explanation for the greater overdensity value in the reconstruction obtained with the Tempel grouped catalogue than in that obtained with the Tully grouped catalogue. Such findings highlight the importance of a balance between grouping to remove non-linear motions and preserving some constraints to produce an infall on to structures that are expected to be large overdensities.

The main conclusions of the paper are as follows. The choice of H_0 impacts overall the velocity field in a given direction, i.e. it creates a general infall/outflow patterns but it does not really affect the overdensity field. Namely, the tidal part of the velocity field changes quite a lot with H_0 but not the divergent part. However, this conclusion has to be strongly mitigated. Indeed, the bias minimization scheme described in Sorce (2015) applied to the grouped observational catalogue strongly suppresses the dependence of the reconstructions on H_0. There is no more drastic infall on to the local volume. In contrast, the grouping scheme affects greatly the overdensity field accentuating or diminishing the contrast between the structures. Still overall structures are reconstructed at the proper location with both grouping schemes studied here. Then in terms of H_0, we simply recommend either to choose the value giving the more neutral result (i.e. monopole term close to zero at large radii) or to apply the bias minimization scheme described in Sorce (2015) after grouping. Note that this bias minimization scheme also erases the bump entirely due to biases in the dipole term and makes the radial peculiar velocity distribution Gaussian. It is worth noticing that again, the dipole at large radii is proven to be very stable whatever choices is made to built the WF reconstruction providing that the grouping is properly done and that the catalogue contains both clusters/groups and galaxies in the field (e.g. Sorce et al. 2017). Regarding the grouping scheme, there is a clear need for a balance between grouping to remove non-linear motions to preserve the quality of the WF reconstruction (Sorce et al. 2017) and its stability with respect to H_0 choice, and keeping some constraints to contrast the high overdensity regions with respect to other regions. If a more aggressive grouping like Tully Grouping scheme permits stabilizing the WF reconstruction across a large range of H_0 values, preserving more constraints like with Tempel Grouping scheme provides Coma and Centaurus regions with a greater contrast with respect to other regions. Such an observation is promising to perform constrained simulations of greater quality than those of the first generation in terms of cluster (Virgo excluded since its mass is already in good agreement with observations) masses. Therefore, the next step consists in using the catalogue grouped with Tempel Grouping scheme as constraints in order to build constrained initial conditions of the local Universe with the local massive clusters.

ACKNOWLEDGEMENTS

We thank Brent Tully and Noam Libeskind for useful discussions as well as Yehuda Hoffman and Romain Graziani for interesting discussions and Stefan Gottlöber for useful comments. We acknowledge the use of the Extragalactic Distance Database to extract the Cosmicflows catalogues and thank the Cosmicflows team led by Brent Tully and Helene Courtois for the release of the catalogues. JS acknowledges support from the Astronomy ESFRI and Research Infrastructure Cluster ASTERICS project, funded by the European Commission under the Horizon 2020 Programme (GA 653477). ET acknowledges support from the ETAg grant IUT40-2 and from the European Regional Development Fund. We thank the anonymous referee for his/her comments.

REFERENCES

Abell G. O., Corwin H. G. Jr, Olowin R. P., 1989, ApJS, 70, 1
Beaton R. L. et al., 2016, ApJ, 832, 210
Carlesi E., Hoffman Y., Sorce J. G., Gottlöber S., Yepes G., Courtois H., Tully R. B., 2016a, MNRAS, 460, L5
Carlesi E. et al., 2016b, MNRAS, 458, 900
Colless M., Saglia R. P., Burstein D., Davies R. L., McMahan R. K., Wegner G., 2001, MNRAS, 321, 277
Courtois H. M., Hoffman Y., Tully R. B., Gottlöber S., 2012, ApJ, 744, 43
Crook A. C., Huchra J. P., Martimbeau N., Masters K. L., Jarrett T., Macri L. M., 2007, ApJ, 655, 790
da Costa L. N., Nusser A., Freudling W., Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., 1998, MNRAS, 299, 425
Dekel A., 1994, ARA&A, 32, 371
Dekel A., Ostriker J. P., 1999, The Observatory, 119, 292
Dekel A., Eldar A., Kolatt T., Yahil A., Willick J. A., Faber S. M., Courteau S., Burstein D., 1999, ApJ, 522, 1
Eke V. R. et al., 2004, MNRAS, 348, 866
Freedman W. L. et al., 2001, ApJ, 553, 47
Geller M. J., Huchra J. P., 1983, ApJS, 52, 61
Gottlöber S., Hoffman Y., Yepes G., 2010, preprint (arXiv:1005.2687)
Hoffman Y., Courtois H. M., Tully R. B., 2015, MNRAS, 449, 4494
Hoffman Y., Nusser A., Courtois H. M., Tully R. B., 2016, MNRAS, 461, 4176
Huchra J. P., Geller M. J., 1982, ApJ, 257, 423
Hudson M. J., Turnbull S. J., 2012, ApJ, 751, L30
Jackson N., 2015, Living Rev. Relativ., 18, 2
Jha S., Riess A. G., Kirshner R. P., 2007, ApJ, 659, 122
Keenan R. C., Barger A. J., Cowie L. L., 2013, ApJ, 775, 62
Lavaux G., 2016, MNRAS, 457, 172
Lavaux G., Hudson M. J., 2011, MNRAS, 416, 2840
Lee M. G., Freedman W. L., Madore B. F., 1993, ApJ, 417, 553
Makarov D., Karachentsev I., 2011, MNRAS, 412, 2498
Mathewson D. S., Ford V. L., Buchhorn M., 1992, ApJS, 81, 413
Nusser A., Davis M., 1995, MNRAS, 276, 1391
Old L. et al., 2014, MNRAS, 441, 1513
Old L. et al., 2015, MNRAS, 449, 1897
Planck Collaboration XVI 2014, A&A, 571, A16
Ramella M., Geller M. J., Pisani A., da Costa L. N., 2002, AJ, 123, 2976
Rathaus B., Kovetz E. D., Itzhaki N., 2013, MNRAS, 431, 3678
Riess A. G. et al., 2016, ApJ, 826, 56
Said K., Kraan-Korteweg R. C., Jarrett T. H., 2014, preprint (arXiv:1410.2992)
Singh M., Gupta S., Pandey A., Sharma S., 2016, JCAP, 8, 026
Sorce J. G., 2015, MNRAS, 450, 2644
Sorce J. G., Courtois H. M., Gottlöber S., Hoffman Y., Tully R. B., 2014, MNRAS, 437, 3586
Sorce J. G., Gottlöber S., Hoffman Y., Yepes G., 2016a, MNRAS, 460, 2015
Sorce J. G. et al., 2016b, MNRAS, 455, 2078
Sorce J. G., Hoffman Y., Gottlöber S., 2017, MNRAS, 468, 1812

Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2007, ApJS, 172, 599
Strauss M. A., Willick J. A., 1995, Phys. Rep., 261, 271
Tempel E. et al., 2014, A&A, 566, A1
Tempel E., Kipper R., Tamm A., Gramann M., Einasto M., Sepp T., Tuvikene T., 2016, A&A, 588, A14
Theureau G., 1999, 167, 328
Theureau G., Rauzy S., Bottinelli L., Gouguenheim L., 1998, A&A, 340, 21
Tonry J. L., Dressler A., Blakeslee J. P., Ajhar E. A., Fletcher A. B., Luppino G. A., Metzger M. R., Moore C. B., 2001, ApJ, 546, 681
Tully R. B., 1987, ApJ, 321, 280
Tully R. B., 2015a, AJ, 149, 54
Tully R. B., 2015b, AJ, 149, 171
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
Watkins R., Feldman H. A., 2015, MNRAS, 447, 132
Willick J. A., 1999, preprint (arXiv:astro-ph/9909003)
Willick J. A., Courteau S., Faber S. M., Burstein D., Dekel A., Strauss M. A., 1997, ApJS, 109, 333
Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., 2005, MNRAS, 357, 608
Zaroubi S., Hoffman Y., Fisher K. B., Lahav O., 1995, ApJ, 449, 446
Zaroubi S., Zehavi I., Dekel A., Hoffman Y., Kolatt T., 1997, ApJ, 486, 21
Zaroubi S., Hoffman Y., Dekel A., 1999, ApJ, 520, 413
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

APPENDIX

The WF technique is the optimal minimal variance estimator given a data set and an assumed prior power spectrum. Data dominate the reconstruction in region where they are dense and accurate. On the opposite when they are noisy and sparse, the reconstruction is a prediction based on the assumed prior model. Briefly, the overdensity $\delta^{\rm WF}$ and velocity $\boldsymbol{v}^{\rm WF}$ fields of the WF are expressed in terms of the following correlation matrixes. For a list of M constraints c_i:

$$\delta^{\rm WF}(\boldsymbol{r}) = \sum_{i=1}^{M} \langle \delta(\boldsymbol{r}) c_i \rangle \eta_i, \tag{A1}$$

$$v_\alpha^{\rm WF} = \sum_{i=1}^{M} \langle v_\alpha(\boldsymbol{r}) c_i \rangle \eta_i \quad \text{with} \quad \alpha = x, y, z, \tag{A2}$$

where $\eta_i = \sum_{j=1}^{M} \langle C_i C_j \rangle^{-1} C_j$ are the components of the correlation vector η. $C_i = c_i + \epsilon_i$ are observational constraints plus their uncertainties. Hence, $\langle C_i C_j \rangle$ is equal to $\langle c_i c_j \rangle + \epsilon_i^2 \delta_{ij}$ assuming statistically independent errors. The constraints can be either densities or velocities. $\langle \mathrm{AB} \rangle$ notations stand for the correlation functions involving the assumed prior power spectrum.

The associated correlation functions are given by

$$\begin{aligned} \langle \delta(\boldsymbol{r}\,') v_\alpha(\boldsymbol{r}\,' + \boldsymbol{r}) \rangle &= \frac{\dot{a} f}{(2\pi)^3} \int_0^\infty \frac{\mathrm{i} k_\alpha}{k^2} P(\boldsymbol{k}) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{k} \\ &= -\dot{a} f r_\alpha \zeta(r) \end{aligned} \tag{A3}$$

$$\begin{aligned} \langle v_\alpha(\boldsymbol{r}\,') v_\beta(\boldsymbol{r}\,' + \boldsymbol{r}) \rangle &= \frac{(\dot{a} f)^2}{(2\pi)^3} \int_0^\infty \frac{k_\alpha k_\beta}{k^4} P(\boldsymbol{k}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{k}, \\ &= (\dot{a} f)^2 \Psi_{\alpha\beta} \end{aligned} \tag{A4}$$

where P is the assumed prior power spectrum, a the scalefactor and f the growth rate.

Because data sample a typical realization of the prior model, i.e. the power spectrum, $\frac{\chi^2}{\rm dof}$ should be close to 1 where $\chi^2 = \sum_{i=1}^{M} \sum_{j=1}^{M} C_i \langle C_i C_j \rangle^{-1} C_j$ and dof is the degree of freedom. However, data include non-linearities which are not taken into account in the model. Consequently, a non-linear sigma ($\sigma_{\rm NL}$) such that $\langle C_i C_j \rangle = \langle c_i c_j \rangle + \delta_{ij}^k \epsilon_j^2 + \delta_{ij}^k \sigma_{\rm NL}^2$ is required to compensate for the non-linearities to drive $\frac{\chi^2}{\rm dof}$ closer to 1.

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