



HAL
open science

The 2M++ galaxy redshift catalogue

Guilhem Lavaux, Michael J. Hudson

► **To cite this version:**

Guilhem Lavaux, Michael J. Hudson. The 2M++ galaxy redshift catalogue. Monthly Notices of the Royal Astronomical Society, 2011, 416, pp.2840-2856. 10.1111/j.1365-2966.2011.19233.x. insu-03645868

HAL Id: insu-03645868

<https://insu.hal.science/insu-03645868>

Submitted on 22 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The 2M++ galaxy redshift catalogue

Guilhem Lavaux^{1,2★} and Michael J. Hudson^{3,4,5}

¹*Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, IL 61801-3080, USA*

²*Department of Physics and Astronomy, The Johns Hopkins University, 3701 San Martin Drive, Baltimore, MD 21218, USA*

³*Institut d'Astrophysique de Paris, UMR7095 CNRS, Univ. Pierre et Marie Curie, 98 bis Boulevard Arago, 75014 Paris, France*

⁴*Department of Physics & Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*

⁵*Perimeter Institute for Theoretical Physics, 31 Caroline St. N., Waterloo, ON N2L 2Y5, Canada*

Accepted 2011 June 10. Received 2011 May 30; in original form 2010 December 12

ABSTRACT

Peculiar velocities arise from gravitational instability, and thus are linked to the surrounding distribution of matter. In order to understand the motion of the Local Group with respect to the cosmic microwave background, a deep all-sky map of the galaxy distribution is required. Here we present a new redshift compilation of 69 160 galaxies, dubbed 2M++, to map large-scale structures of the local Universe over nearly the whole sky, and reaching depths of $K \leq 12.5$, or $200 h^{-1}$ Mpc. The target catalogue is based on the Two-Micron All-Sky Survey Extended Source Catalog (2MASS-XSC). The primary sources of redshifts are the 2MASS Redshift Survey, the 6dF galaxy redshift survey and the Sloan Digital Sky Survey (Data Release 7). We assess redshift completeness in each region and compute the weights required to correct for redshift incompleteness and apparent magnitude limits, and discuss corrections for incompleteness in the zone of avoidance. We present the density field for this survey, and discuss the importance of large-scale structures such as the Shapley Concentration.

Key words: methods: data analysis – methods: numerical – methods: observational – catalogues – galaxies: luminosity function, mass function – large-scale structure of Universe.

1 INTRODUCTION

Peculiar velocities remain the only method to map the distribution of dark matter (DM) on very large scales in the low-redshift Universe. Recently, several intriguing measurements (Kashlinsky et al. 2008, 2010; Watkins, Feldman & Hudson 2009; Feldman, Watkins & Hudson 2010; Lavaux et al. 2010; Colin et al. 2011) of the mean or ‘bulk’ flow on scales larger than $100 h^{-1}$ Mpc suggest a high velocity of our local $\sim 100 h^{-1}$ Mpc volume with respect to the cosmic microwave background (CMB) frame. In the standard cosmological framework, peculiar velocities are proportional to peculiar acceleration and so one expects the bulk flow to arise from fluctuations in the distribution of matter, and hence presumably of galaxies, on very large scales. Another statistic for measuring such large-scale fluctuations is the convergence of the gravity dipole as a function of distance. However, the rate of convergence has been a subject of recent debate (Erdoğdu et al. 2006a, b; Kocevski & Ebeling 2006a; Lavaux et al. 2010; Bilicki et al. 2011, and references therein). A closely related topic is the gravitational influence of the Shapley Concentration (SC), the largest concentration of galaxy clusters in the nearby Universe. It is therefore important to have catalogues that are as full sky and as deep as possible to understand whether

the distribution of matter in the nearby Universe may explain the above-mentioned results.

It is already possible with currently available data to build a redshift catalogue significantly deeper than previous full-sky galaxy redshift catalogues like PSCz (Saunders et al. 2000) or the Two-Micron All-Sky Redshift Survey (2MRS; Huchra et al. 2005; Erdoğdu et al. 2006a; Huchra et al., in preparation). We present here a new catalogue called the 2M++ galaxy redshift compilation. The photometry for this compilation is based on the Two-Micron All-Sky Survey Extended Source Catalog (2MASS-XSC; Skrutskie et al. 2006). We gather the high-quality redshifts from the 2MRS (Huchra et al. 2005; Erdoğdu et al. 2006a; Huchra et al., in preparation) limited to $K = 11.5$, the 6dF galaxy redshift survey Data Release 3 (6dFGRS-DR3; Jones et al. 2009) and the Sloan Digital Sky Survey Data Release 7 (SDSS-DR7; Abazajian et al. 2009).

A summary of this paper is as follows. In Section 2, we describe the steps in constructing the 2M++ redshift galaxy catalogue: source selection, magnitude corrections, redshift incompleteness estimation and correction, the luminosity function (LF) estimation and the final weight computation. In Section 3, we discuss the zone of avoidance (ZoA) in our catalogue, and how its effects can be mitigated. In Section 4, we define groups of galaxies and check some of their overall properties. In Section 5, we compute and analyse the density field, presenting maps of the supergalactic plane and three cluster density and velocity profiles. Section 6 summarizes our key results.

*E-mail: lavaux@iap.fr

2 CATALOGUE CONSTRUCTION

In this section, we describe the construction of the 2M++ galaxy redshift catalogue from the different data sources. First, in Section 2.1, we describe the source data sets that form the basis of our catalogue, as well as the primary steps in the construction of the 2M++ catalogue. We then present the methodology used for merging these different sources. In Section 2.2, we describe the corrections applied to apparent magnitudes to homogenize the target selection. In Section 2.3, we test and apply the redshift cloning procedure to our data to increase the overall redshift completeness. We then estimate redshift completeness (Section 2.4.2) and present the number counts of galaxies as a function of redshift (Section 2.5). Finally, we compute the LF of our sample in Section 2.6 and compute the total weights to apply to each galaxy in Section 2.7.

2.1 Source data sets and construction procedure

Our catalogue is based on the 2MASS (Skrutskie et al. 2006) photometric catalogue for target selection, which has very high completeness up to $K_S = 13.2$ (Cole et al. 2001). Hereafter, for brevity, we use K in place of K_S . As noted above, we will be using redshifts from the SDSS-DR7, the 6dFGRS and the 2MRS. In addition to these main sources, we have gathered additional redshifts from a number of other sources (Schneider et al. 1990; de Vaucouleurs et al. 1991; Binggeli, Popescu & Tammann 1993; Huchra, Geller & Corwin 1995; Falco et al. 1999; Conselice, Gallagher & Wyse 2001; Rines et al. 2003; Koribalski et al. 2004) through NASA/IPAC Extragalactic Database (NED) queries.¹ Due to the inhomogeneity of the target selection between the different redshift surveys, we think that it is more appropriate to define a new target selection rather than using existing target data bases from the above surveys. We used the New York University-Value Added Galaxy Catalog (NYU-VAGC) (Blanton et al. 2005) catalogue for matching the SDSS data to the 2MASS-XSC. The NYU-VAGC provides the SDSS survey mask in MANGLE format (Hamilton & Tegmark 2004).² We sampled the mask on a HEALPIX grid at $N_{\text{side}} = 512$ (~ 10 -arcmin resolution). This angular resolution corresponds to $\sim 1 h^{-1}$ Mpc at $\sim 300 h^{-1}$ Mpc. Because ultimately we will be smoothing the density field on scales of $\sim 4 h^{-1}$ Mpc, the mask has sufficient resolution for our purposes. Additionally, we filter out from our target selection the extended sources that are known not to be galaxies.³

We aim to limit 2M++ at $K \simeq 12.5$ regions of the sky covered by SDSS or by 6dF. The exact cut depends on the adopted definition for the magnitude. As we want to retain as much as possible information from the shallower 2MRS catalogue, we opt to follow closely the magnitude used by 2MRS for target selection. We define as K_{2M++} the magnitude of a galaxy measured in the K_S band, within the circular isophote at $20 \text{ mag arcsec}^{-2}$, after various corrections as described below (Section 2.2). Several of the steps taken to build the catalogue are described in greater detail in the following sections. We now outline these steps.

(i) We import the redshift information for 2MASS-XSC galaxies from the NYU-VAGC for SDSS-DR7, the 6dF-DR3 and the 2MRS.

¹ NED is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

² We use the file named `lss_combmask.dr72.ply`, which gives the geometry of the DR72 sample in terms of target selection with bright stars excised.

³ We require that the `VISUAL_CODE` is not equal to 2.

(ii) We correct for small-scale redshift incompleteness (arising from fibre collisions) by ‘cloning’ the redshifts of nearby galaxies (Section 2.3).

(iii) We correct the apparent magnitudes for Galactic dust extinction (Section 2.2).

(iv) We use the redshift to correct for galaxy evolutionary effects and aperture corrections (Section 2.2). We call this magnitude K_{2M++} . At those magnitudes, we assume that the photometric completeness is one at Galactic latitudes higher than 10° .

(v) We compute two sets of galaxy samples: a target sample with $K_{2M++} \leq 11.5$ in regions not covered by 6dFGRS or SDSS, and a target sample limited to $K_{2M++} \leq 12.5$ in regions covered by SDSS or 6dFGRS.

(vi) We estimate the redshift completeness as a function of position on the sky for each of these regions.

(vii) We place galaxies in groups and clusters using a percolation algorithm.

(viii) We compute the Schechter parameters of the LF of the combined catalogue (Section 2.6). We use this LF to compute the weights to apply to each of the observed galaxies to take into account the unobserved ones (Section 2.7).

In future work we will update the estimated distances for the galaxies using reconstructed velocity field and re-execute step (viii) to update the corrections. The detail of this procedure will be discussed in a later paper (Lavaux & Hudson, in preparation).

2.2 Apparent magnitude corrections

We describe in this section the corrections that must be applied to apparent magnitudes to mitigate the effects of cosmological surface brightness dimming, Galactic extinction and stellar evolution. We choose to use a definition of the magnitudes for target selection that is related to the one used for defining magnitudes in the 2MRS. This ensures that the final completeness is maximized in the parts of the sky where only redshifts from 2MRS are available.

The absolute magnitude M at redshift zero of a galaxy may be written as

$$M = m - A_K(l, b) - k(z) + e(z) - D_L(z), \quad (1)$$

with m the apparent magnitude, $A_K(l, b)$ the absorption by Milky Way’s dust in the direction (l, b) , $k(z)$ the k -correction due to the redshifting of the spectrum, $e(z)$ the correction for evolution of the stellar population and D_L the luminosity distance. We convert the redshifts into luminosity distances assuming a Λ cold dark matter (Λ CDM) cosmology with a mean total matter density parameter $\Omega_M = 0.30$ and a dark energy density parameter $\Omega_\Lambda = 0.70$. All absolute magnitudes are computed assuming $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

The absorption in K_S band is related to the extinction $E(B - V)$ estimated using the maps of Schlegel, Finkbeiner & Davis (1998) by the relation

$$A_K(l, b) = 0.35E(B - V)(l, b), \quad (2)$$

where the constant of proportionality is obtained from the relation between absorption in K band and in V band (Cardelli, Clayton & Mathis 1989).

The adopted k -correction is

$$k(z) = -2.1z \quad (3)$$

from Bell et al. (2003a). Finally, the evolutionary correction is

$$e(z) = 0.8z \quad (4)$$

also from Bell et al. (2003a).

The magnitudes adopted in this work are circular aperture magnitudes defined within a limiting surface brightness of $20 \text{ mag arcsec}^{-2}$. However, various effects will cause not only the observed magnitude to change, but also the observed surface brightness. As the surface brightness of the galaxy profile drops, the isophotal aperture will move inwards and so the aperture magnitude will drop.

The surface brightness will depend on redshift via the usual $(1+z)^4$ cosmological dimming effect as well as the extinction, k -correction and evolutionary effects described above. Therefore, the correction is

$$\begin{aligned} \Delta SB &= SB(z=0) - SB(z) \\ &= -10 \log(1+z) - A_K(l, b) - k(z) + e(z). \end{aligned} \quad (5)$$

Note that the $k+e$ corrections have opposite sign to the cosmological surface brightness dimming, and so there is some cancellation of these effects. However, the cosmological term still dominates, so the net effect is that as galaxies are moved to higher redshift their surface brightness is dimmer.

By simulating simple Sérsic profiles, we have estimated how much the aperture magnitude changes as a result of surface brightness dimming. For a typical 2M++ galaxy with Sérsic index $n = 1.5$ and mean surface brightness within the effective radius of $\langle \mu_e \rangle = 17.5$, we find that the correction to the magnitude due *only* to a change in aperture radius can be approximated by $0.16 \Delta SB$, where ΔSB is the correction in the surface brightness. This term is only the shift in magnitude due to the shift in isophotal radius, and does not include the ‘direct’ effect on the magnitude itself due to extinction and $k+e$ corrections. Thus, the total effect is

$$\begin{aligned} K_{2M++} &= K_{20,c} + 1.16[-A_K(l, b) - k(z) + e(z)] \\ &\quad - 0.16[10 \log(1+z)]. \end{aligned} \quad (6)$$

Note that this is close, but not identical, to the 2MRS-corrected magnitude.

In some cases, only the magnitude $K_{20,e}$, derived from adjusting an ellipsoidal Sérsic profile, is available. In those cases, we have computed the corresponding $K_{20,c}$ using the following relation, obtained by fitting on the galaxies for which the two magnitudes were

available:

$$K_{20,c} = (0.9774 \pm 0.0005)K_{20,e} + (0.288 \pm 0.006). \quad (7)$$

The residual of the fit has a standard deviation equal to 0.11. We also use this relation whenever the predicted $K_{20,c}$ and the actual $K_{20,c}$ from 2MASS-XSC differ by 0.22 and calculate the $K_{20,c}$ from $K_{20,e}$. We have used this relation for 7 per cent of galaxies, both in the target and the final redshift compilation.

2.3 Redshift cloning

Within the 6dF and SDSS regions, there is small-scale incompleteness due primarily to fibre collisions. To improve the redshift coverage of the catalogues, we ‘clone’ redshifts of nearby galaxies within each survey region. This procedure, which is related to another one described in Blanton et al. (2005), is as follows. Consider two targets T_a and T_b . If T_a does not have a measured redshift and T_b has one, and furthermore T_b is the nearest target of T_a with an angular distance less than ϵ , we copy the redshift of T_b to T_a . ϵ is determined by the angular distance between two fibres of the measuring instrument, which is $\epsilon = 5.7 \text{ arcmin}$ for 6dF (Jones et al. 2004) and $\epsilon = 55 \text{ arcsec}$ for SDSS (Blanton et al. 2003). We refer to redshifts cloned in this way as ‘fibre-clones’.

To assess the errors on redshifts for the fibre-clones, we randomly split the set of galaxies which has a measured redshifts in two sets $\mathcal{S}_{\text{keep}}$ and $\mathcal{S}_{\text{test}}$. We mark the galaxies belonging to $\mathcal{S}_{\text{test}}$ as having no redshift. We then apply the fibre-cloning procedure to these galaxies.

The result of this test is shown in Fig. 1 for both SDSS galaxies and 6dF galaxies. We note that the Cauchy–Lorentz distribution with width $W = 2.7 h^{-1} \text{ Mpc}$ is a good fit to the central part of the two distributions. We used the formula

$$P(e) = \frac{1}{\pi W} \frac{1}{1 + (e/W)^2} \quad (8)$$

for the modelled probability distribution function in both panels. We checked that a Gaussian distribution manages only to fit the central part of the distribution and is less adequate than a Cauchy–Lorentz

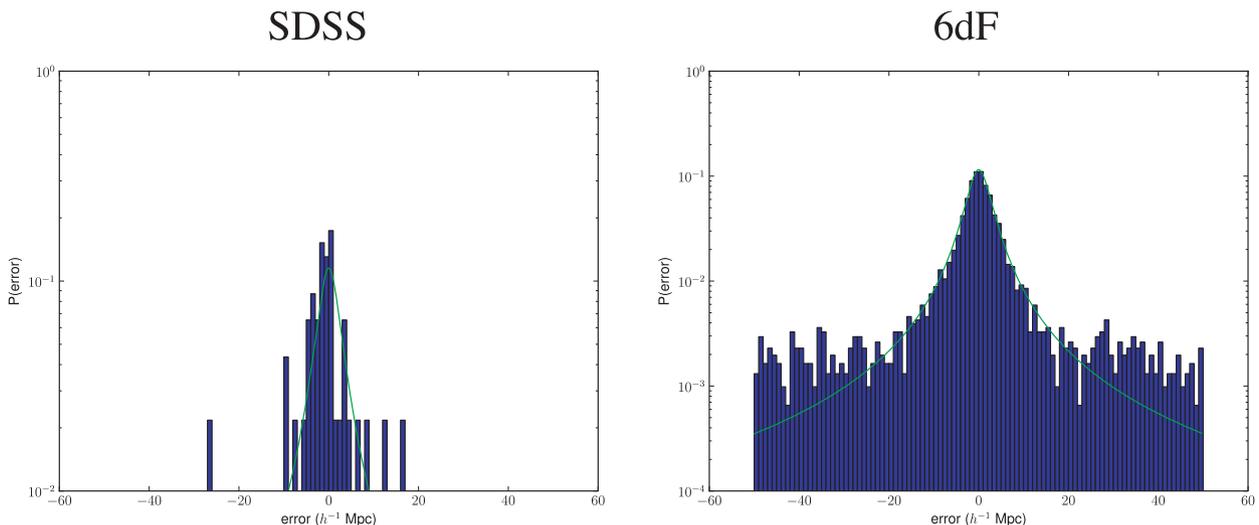


Figure 1. Error distribution due to the redshift cloning procedure – we give here the computed error of the redshifts of either SDSS or 6dF redshift catalogue. We removed the redshift information of half of the objects in these catalogues and tried to recover them using the cloning procedure. The difference is plotted as a histogram in the two plots above. The overlaid continuous curve corresponds to a Cauchy–Lorentz distribution with a width equal to $2.7 h^{-1} \text{ Mpc}$.

distribution. The fibre-clones are given a redshift error of 9×10^{-4} , which corresponds to $\sim 2.7 h^{-1}$ Mpc at redshift $z = 0$.

2.4 Redshift survey masks and completeness

Because of the different redshift catalogues used in $2M_{++}$, we will separate the full sky into the following regions: $K_{2M_{++}} \leq 11.5$ (2MRS); $K_{2M_{++}} \leq 12.5$ (6dF/SDSS) or regions with insufficient redshift data. We begin by assigning preliminary ‘masks’, then measure the redshift completeness and then assign final masks based on a completeness limit of 50 per cent. We now describe these steps in more detail.

2.4.1 Preliminary mask selection

The preliminary mask for the 6dF is as given in Jones et al. (2004): $|b| > 10^\circ$ and in the Southern hemisphere $\delta \leq 0^\circ$.

For SDSS, from the full DR7, we select only the most homogeneous and contiguous portion of the redshift survey. To obtain the geometry corresponding to this portion, we use the mask computed numerically from the MANGLE file, and impose an additional constraint on the positions: we keep only galaxies within the region $90^\circ < \alpha < 250^\circ$, to which we add another region at $250^\circ \leq \alpha < 270^\circ$ and $\delta < 50^\circ$. This selection retains the major contiguous piece of the SDSS in the northern Galactic cap, while removing the southern Galactic strips and the small disjoint piece with $\alpha \sim 260^\circ$. After these cuts, we still retain 90 per cent of the area covered by the complete SDSS-DR7 ‘Legacy’ spectroscopic survey.

The preliminary SDSS mask is shown in Fig. 2. This mask has a relatively simple geometry and is contiguous, except for the presence of very small holes that are due either to stars or to imperfect overlap of the SDSS plates. These imperfections represent about 8 per cent of the SDSS surface area.

For 2MRS, the initial mask is the whole sky with $|b| > 5^\circ$, except in the region $-30^\circ < l < +30^\circ$, where the Galactic latitude is $|b| > 10^\circ$ (Erdođu et al. 2006a), and excluding the regions covered by 6dF or SDSS. We refer to this region as 2Mx6S.

The combination of the three survey masks is shown in Fig. 3, in Galactic coordinates. The grey area near the galactic plane is covered by none of the surveys. There is a small overlap between the SDSS and the 6dFGRS in the north galactic cap (in cyan).

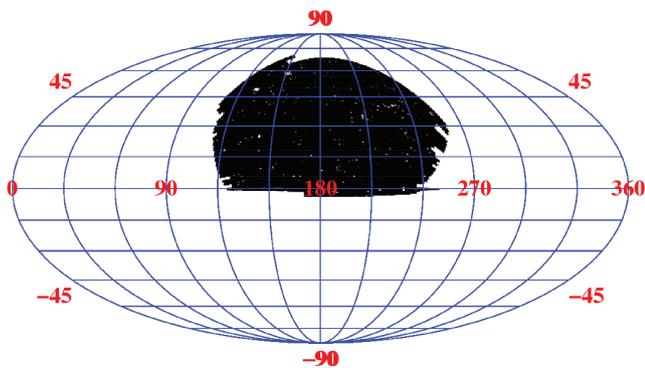


Figure 2. Mollweide projection of the SDSS spectroscopic mask in equatorial coordinates. We removed unconnected parts from the original mask. Note the presence of small holes in the mask due to both the presence of stars and not exact reconnection of the SDSS plates. This corresponds to an intersection of the geometry described by `LSS_GEOMETRY.PLY` and our selection criterion described in Section 2.4.

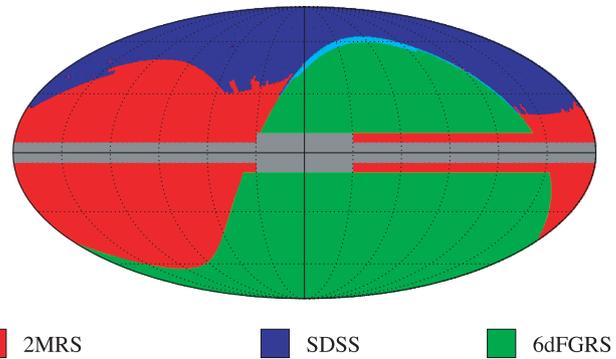


Figure 3. A Mollweide projection in galactic coordinates showing the preliminary masks corresponding to the different redshift catalogues: the 2MRS (red); the 6dFGRS (green) and the SDSS-DR7 (blue). Regions with no redshift data are shown in grey.

2.4.2 Redshift completeness

In order to determine LFs and weights needed for the density field, it is first necessary to assess the completeness of the redshift catalogues on the sky. We will estimate the redshift completeness in some direction of the sky for two different magnitudes cuts: $K_{2M_{++}} \leq 11.5$ (for the 2Mx6S region) and $K_{2M_{++}} \leq 12.5$ (for 6dF/SDSS regions). The maps of these completeness are given in Fig. 4. The completeness of the 2MRS is quite homogeneous and only drops close to the Galactic plane. The 6dF survey is mostly homogeneous except at the location of the Magellanic Clouds. The SDSS completeness is quite homogeneous, and remains at a level of about 80 per cent in the whole contiguous region.

2.4.3 Final masks and the $2M_{++}$ catalogue

Based on the above analysis, we reject from 6dF and SDSS those regions where the completeness at $K_{2M_{++}} \leq 12.5$ is less than 50 per cent, and assign these areas to the 2Mx6S and limit the magnitude there to 11.5. As a result, we note that while the 6dF and SDSS masks are contiguous, the 2Mx6S mask is not. We show in the lower-left panel of Fig. 4 the mask corresponding to the footprint of the $2M_{++}$ compilation. The black corresponds to regions where the survey has a completeness higher than 50 per cent at the limiting magnitude of $K_{2M_{++}} = 12.5$ (the final 6dF and SDSS masks). The grey area represents the same but for a limiting magnitude of $K_{2M_{++}} = 11.5$ (the final 2Mx6S mask). In white, we show the parts of the sky where either there is no redshift information or target galaxies were not present.

The ZoA is clearly visible. There are also important unobserved patches at $K_{2M_{++}} \leq 12.5$ in the southern Galactic hemisphere at the locations of the Magellanic Clouds. The other white patches in the Southern hemisphere are mostly related to local higher absorption by dust in the Milky Way.

Thus, in summary, the final $2M_{++}$ catalogue is defined as all galaxies in 2Mx6S with $K_{2M_{++}} \leq 11.5$, or in 6dF or SDSS with $K_{2M_{++}} \leq 12.5$, and contains 69 160 galaxies with redshifts (including fibre-clones). Table 1 summarizes the statistics and completeness for the different regions of the $2M_{++}$ compilation. We note that the $2M_{++}$ compilation redshifts are nearly 90 per cent complete, and so redshift completeness corrections are small.

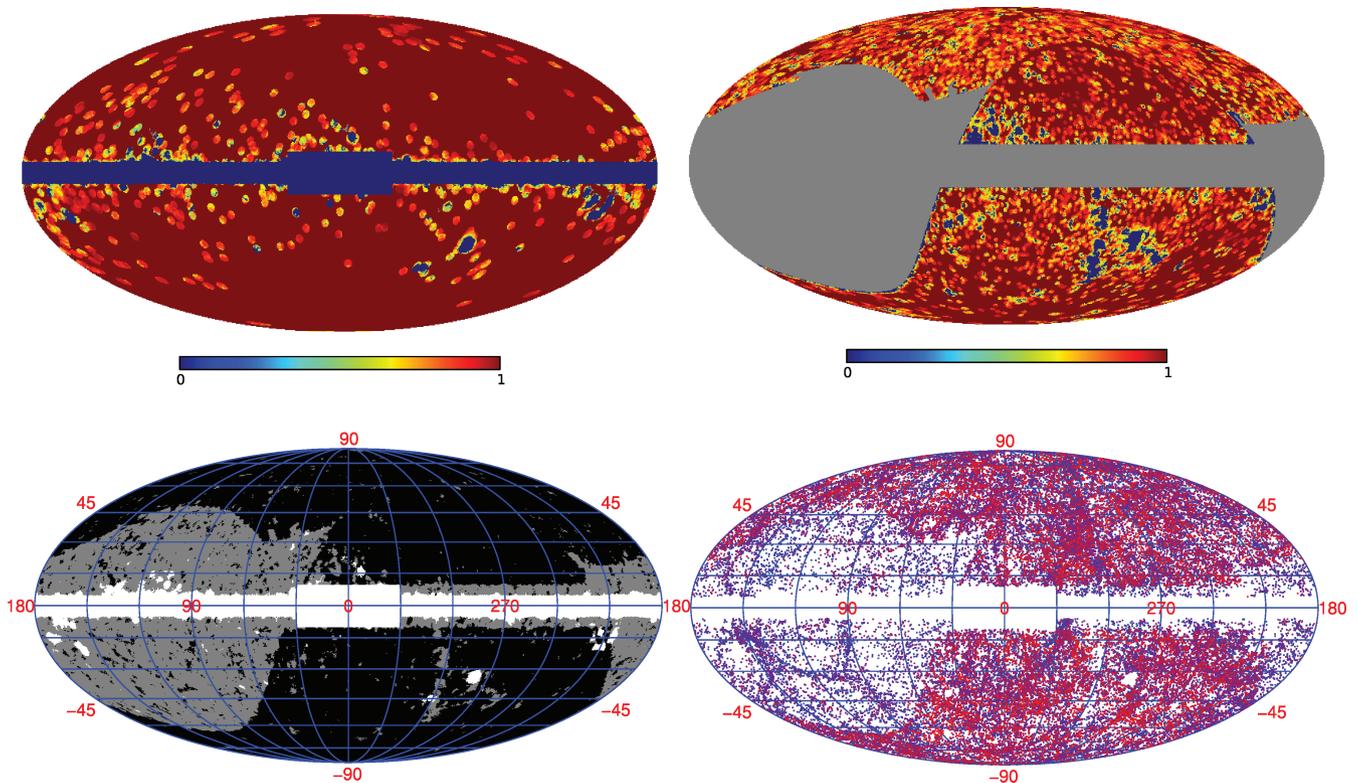


Figure 4. The 2M++ compilation in Galactic coordinates. The top panels give the redshift completeness of 2M++ for the limiting magnitude $K_{2M++} = 11.5$ (left-hand panel) and $K_{2M++} = 12.5$ (right-hand panel). The grey area corresponds to regions with no targets and no redshifts. In the bottom-left panel, we show the final combined mask. The regions in white have been completely excluded because of very low completeness. The regions in grey (black respectively) have a redshift completeness higher than 50 per cent at a limiting magnitude $K_{2M++} \leq 11.5$ ($K_{2M++} \leq 12.5$ respectively). In the bottom-right panel, we show all galaxies with redshift included in the 2M++ compilation. Each galaxy is colour-coded according to its redshift distance, blue for the nearest and red for the farthest.

Table 1. Summary statistics for the primary regions in the 2M++ compilation.

| Region | m_{lim} | Area | $N_{m < m_{\text{lim}}}$ | N_z | \bar{c} |
|--------|------------------|--------|--------------------------|--------|-----------|
| 2Mx6S | 11.5 | 13 069 | 9419 | 9016 | 0.96 |
| 6dF | 12.5 | 17 041 | 46 734 | 42 442 | 0.91 |
| SDSS | 12.5 | 6970 | 20 333 | 17 702 | 0.87 |
| None | – | 4172 | – | – | – |
| Total | – | 37 080 | 76 451 | 69 160 | 0.90 |

Note. The regions are counted exclusively. We have not enforced the sample to have a local redshift completeness higher than 50 per cent, resulting in a total number of redshifts higher than in the final catalogue.

2.5 Redshift number density of galaxies

Within the three regions outlined above, there are a total of 69 160 galaxy redshifts (including fibre-clones). Fig. 5 shows a histogram of all redshifts, as well as the cumulative counts starting from redshift $z = 0$. Conservatively, the catalogue appears totally complete up to $z = 0.02$ ($\sim 60 h^{-1}$ Mpc). This is due to our use of the 2MRS for one part of the sky.

In Fig. 6, we compare quantitatively the counts in the 2MRS region with $K_{2M++} \leq 11.5$ with the 6dF/SDSS regions with $K_{2M++} \leq 12.5$. Because our magnitude corrections are not precisely equivalent to those used for the 2MRS catalogue, the increase of the magnitude cut to $K_{2M++} = 12.5$ is not strictly equivalent to using only the 6dF and SDSS spectroscopic data but also includes a few

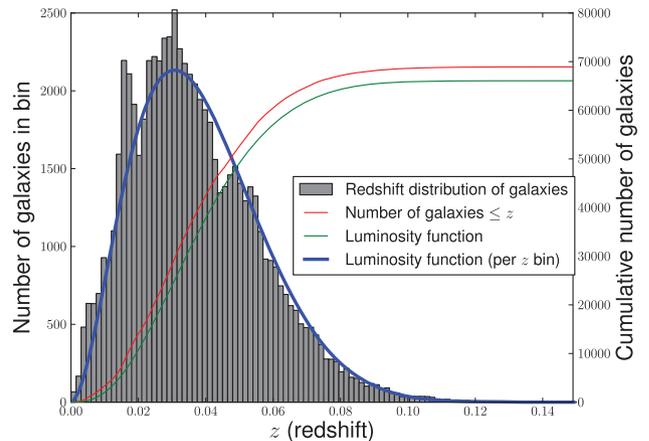


Figure 5. Redshift distribution of 2M++ galaxies. We show, in grey histogram, the number of galaxies within each redshift bin $\delta z = 0.00150$. The cumulative number of galaxies at a redshift less than or equal to z is given by the red curve. The predicted number of galaxies given by our fiducial LF, given at the end of Section 2.6, is shown in solid green for the cumulative number and in solid blue for the number of galaxies in each bin of the grey histogram. The LF has been fitted using a subset of the catalogue for which $5000 \leq cz \leq 20\,000 \text{ km s}^{-1}$ and $-25 \leq M \leq -21$.

2MRS galaxies. None the less, the increase in the magnitude cut corresponds mostly to the sky portions covered by both the SDSS and the 6dF. The deeper redshift data allow us to better probe more distant large-scale structures, particularly in the redshift $0.02 \lesssim z \lesssim$

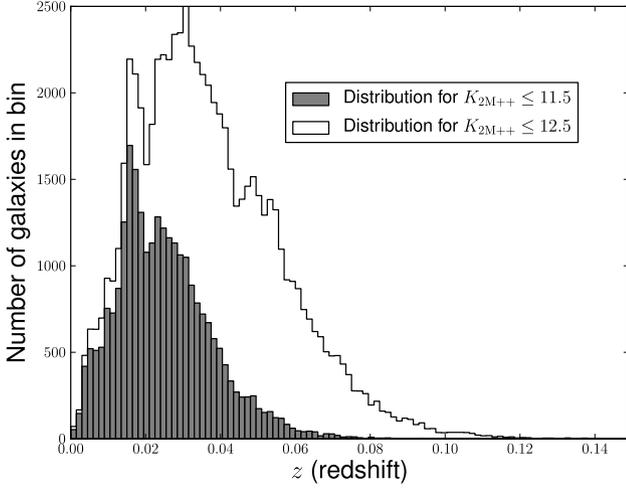


Figure 6. Redshift galaxy distribution in 2M++ for $K_{2M++} \leq 11.5$ (filled histogram) and $K_{2M++} \leq 12.5$ (unfilled histogram).

0.05. For example, the feature in the redshift distribution at $z \sim 0.05$ corresponding to the SC is not present for the subcatalogue $K_{2M++} \leq 11.5$, while it is clearly seen in the 6dF subcatalogue ($K_{2M++} \leq 12.5$).

2.6 Luminosity function

2.6.1 Method

In order to correct for selection effects due to magnitude limits, it is first necessary to measure the LF. We take into account the redshift completeness to measure the LF of galaxies in the combined catalogue. We use a modified version of the likelihood formalism used to find Schechter (1976) function parameters, as described by Sandage, Tammann & Yahil (1979). We assume that evolutionary effects on the luminosity of galaxies have been accounted for by equation (6). The parametrization adopted is the usual Schechter function,

$$\Phi(L) = \frac{n^*}{L^*} \left(\frac{L}{L^*} \right)^\alpha \exp \left(-\frac{L}{L^*} \right), \quad (9)$$

with n^* the density normalization constant, L^* the characteristic luminosity break, or equivalently in terms of absolute magnitudes

$$\begin{aligned} \Phi(M) &= 0.4 \log(10) n^* 10^{0.4(1+\alpha)(M^*-M)} \exp \left[-10^{0.4(M^*-M)} \right] \\ &= n^* \Phi_0(M), \end{aligned} \quad (10)$$

with M^* the characteristic absolute magnitude break in the Schechter function. Above, we have introduced $\Phi_0(M)$, which is the unnormalized Schechter function. We model the probability of observing a galaxy of absolute magnitude M_i given its redshift z_i as

$$P(M_i | z_i, \alpha, M^*, n^*, c) = \frac{c(M_i, \hat{u}_i, d_i) \Phi_0(M_i)}{\int_{M_{\min}}^{M_{\max}} c(M, \hat{u}_i, d_i) \Phi_0(M) dM}, \quad (11)$$

with M_{\min} , M_{\max} the maximum absolute magnitude range from which the galaxies were selected in the catalogue, $c(M, \hat{u}_i, d_i)$ the completeness in the direction \hat{u}_i of the object i , at the absolute magnitude M , and d_i the luminosity distance of the galaxy i at redshift z_i . This expression is simplified using our assumption that redshift incompleteness $c(M, \hat{u}, r)$ may be modelled by two maps

at two apparent magnitude cuts. $c(M, \hat{u}, r)$ is thus

$$c(M, \hat{u}, r) = \begin{cases} c_b(\hat{u}) & \text{if } M + 5 \log_{10} \left(\frac{r}{10 \text{pc}} \right) \leq m_b \\ c_f(\hat{u}) & \text{if } m_b < M + 5 \log_{10} \left(\frac{r}{10 \text{pc}} \right) \leq m_f \\ 0 & \text{otherwise,} \end{cases} \quad (12)$$

with $m_b = 11.5$ and $m_f = 12.5$. The expression of the probability (11) may thus be newly expressed as

$$P(M_i | z_i, \alpha, M^*, n^*, c) = \frac{c(M_i, \hat{u}_i, d_i) \Phi_0(M_i)}{f(d_i, \hat{u}_i, M_{\min}, M_{\max})}, \quad (13)$$

with

$$\begin{aligned} f(r, \hat{r}, M_{\min}, M_{\max}) &= c_b(\hat{r}) \Gamma_{M_{\min}, M_{\max}}^{M^*, \alpha}(m_b, r_{10}) \\ &+ c_f(\hat{r}) \left[\Gamma_{M_{\min}, M_{\max}}^{M^*, \alpha}(m_f, r_{10}) - \Gamma_{M_{\min}, M_{\max}}^{M^*, \alpha}(m_b, r_{10}) \right] \end{aligned} \quad (14)$$

the normalization coefficient for the direction \hat{r} at distance r , and r_{10} defined as the distance in units of 10 pc. In the above, we have also used the function $\Gamma_{M_{\min}, M_{\max}}^{M^*, \alpha}$ defined as

$$\begin{aligned} \Gamma_{M_{\min}, M_{\max}}^{M^*, \alpha}(m, r_{10}) &= \Gamma_{\text{inc}} \left(1 + \alpha, 10^{0.4(M^* - \min(\max(M(m, r_{10}), M_{\min}), M_{\max}))} \right), \end{aligned} \quad (15)$$

with the absolute magnitude

$$M(m, r_{10}) = m - 5 \log_{10}(r_{10}) \quad (16)$$

and $\Gamma_{\text{inc}}(a, y)$ the incomplete Gamma function:

$$\Gamma_{\text{inc}}(a, y) = \int_y^\infty x^{a-1} e^{-x} dx. \quad (17)$$

We write the total probability of observing the galaxies with intrinsic magnitude $\{M_i\}$ and redshift $\{z_i\}$ given the Schechter LF parameters:

$$\begin{aligned} P(\{M_i\} | \{c_i\}, \{z_i\}, \alpha, M^*, n^*) &= \prod_{i=1}^{N_{\text{galaxies}}} P(M_i | z_i, \alpha, M^*, n^*, c_i). \end{aligned} \quad (18)$$

Using Bayes theorem, we now estimate the most likely value taken by α , M^* assuming a flat prior on these parameters.

The normalization constant n^* is determined using the minimum variance estimator of Davis & Huchra (1982), but neglecting the effects of cosmic variance on the weights by setting $J_3 = 0$. While our estimate may be biased relative to the galaxy mean density outside the catalogue, it is less noisy than the optimal case. The estimate also corresponds better to the density in the piece of Universe that we consider than the density corresponding to the optimal weighing. Our choice leads also to a simplification of the mean density as the total number of galaxies divided by the effective volume of 2M++. Consequently, for a survey limited in the absolute magnitude range $[M_{\min}, M_{\max}]$ and with volume V , we compute the mean density of galaxy \bar{n} by the following equation:

$$\bar{n} = \frac{N_{\text{galaxies}}}{\int_V d^3r f(r, \hat{r}, M_{\min}, M_{\max})}, \quad (19)$$

with a standard deviation only from Poisson noise,

$$\frac{\sigma_{\bar{n}}}{\bar{n}} = \frac{\sqrt{N_{\text{galaxies}}}}{N_{\text{galaxies}}}, \quad (20)$$

because we have set $J_3 = 0$. We then convert \bar{n} into n^* using

$$n^* = \frac{\bar{n}}{\int_{M_{\min}}^{M_{\max}} \Phi_0(M) dM}. \quad (21)$$

Similarly, it is possible to define the luminosity density as

$$\bar{L} = n_* 10^{0.4(M_{\odot} - M^*)} \Gamma(2 + \alpha) \times (1 L_{\odot}), \quad (22)$$

with $\Gamma(a) = \Gamma_{\text{inc}}(a, 0)$. The luminosity density is less sensitive than the number density to fluctuations in α .

To determine the LF parameters, we select a subset of the galaxies in our catalogue. We have defined the subset by the following joint conditions.

(i) Galaxies must have a redshift z such that $5000 \leq cz \leq 20000 \text{ km s}^{-1}$. The lower limit reduces the impact of peculiar velocities on absolute magnitude estimation, which is derived using redshifts in the CMB rest frame. By limiting to $cz \leq 20000 \text{ km s}^{-1}$, we avoid more distant volumes with high incompleteness.

(ii) The absolute magnitude estimated from the redshift in CMB rest frame is within the range $[M_{\text{min}} = -25, M_{\text{max}} = -21]$. As mentioned later in this section, this magnitude selection removes the bright objects that do not seem to follow a Schechter LF (as also discussed in Jones et al. 2006).

Absolute magnitudes are determined with $H = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with $h = 1$, and we have assumed a flat Λ CDM cosmology

$\Omega_m = 0.30$ and $\Omega_{\Lambda} = 0.70$. We do not distinguish between early-type and late-type galaxies, and so fit both populations with a single parameter.

2.6.2 Results

The derived LF parameters are summarized in Table 2 for our default choice of cuts discussed above as well as for other choices that we discuss below. The error bars are given at 68 per cent confidence limit, estimated using a Monte Carlo Markov chain method.

Fig. 7 shows the LF for our default cuts in the CMB rest frame. We also show, for the entire data set and for each subcatalogue, the non-parametric LFs estimated using the unbiased $1/V_{\text{max}}$ method (Schmidt 1968; Felten 1976). Note that for the $1/V_{\text{max}}$ LFs the volume and magnitude limits are different than for the parametric fit, which explains that the fitted parametric faint-end slope is not a good fit to the $1/V_{\text{max}}$ in the range $[-21, -19]$. In the left-hand panel, we give the LFs and in the right-hand panel the ratio between the estimated LFs and the best fit.

Table 2. Summary of K -band Schechter LF parameters from this paper and selected results from the literature. Magnitude ranges with a ‘ \sim ’ are estimated. n_* is in units of $10^{-2} h^3 \text{ Mpc}^{-3}$ and the luminosity density \bar{L} is in units of $10^8 h L_{\odot} \text{ Mpc}^{-3}$, assuming $M_{K,\odot} = 3.29$.

| Reference | Frame | Magnitude | Redshift range | α | $M^* - 5 \log_{10} h$ | n_* | \bar{L} |
|------------------------|-------|--------------------|------------------|------------------|-----------------------|-----------------|-----------------|
| Kochanek et al. (2001) | | $[-26; -20]$ | $[2000; 14000]$ | -1.09 ± 0.06 | -23.39 ± 0.05 | 1.16 ± 0.10 | |
| Cole et al. (2001) | | $\sim[-26; -20]$ | ? | -0.96 ± 0.05 | -23.44 ± 0.03 | 1.08 ± 0.16 | |
| Bell et al. (2003b) | | $\sim[-25; -18]$ | ? | -0.77 ± 0.04 | -23.29 ± 0.05 | 1.43 ± 0.07 | |
| Eke et al. (2005) | | $\sim[-25; -20]$ | ? | -0.81 ± 0.07 | -23.43 ± 0.04 | 1.43 ± 0.08 | |
| Huchra et al. (2005) | | $\sim[-28.5; -16]$ | ? | -1.02 | -23.4 | 1.08 | |
| Jones et al. (2006) | | $[-28.85; -15.5]$ | $[750; +\infty]$ | -1.16 ± 0.04 | -23.83 ± 0.03 | 0.75 ± 0.08 | |
| This work | CMB | $[-25; -21]$ | $[5000; 20000]$ | -0.73 ± 0.02 | -23.17 ± 0.02 | 1.11 ± 0.02 | 3.94 ± 0.02 |
| This work | CMB | $[-25; -17]$ | $[750; 20000]$ | -0.80 ± 0.01 | -23.22 ± 0.01 | 1.13 ± 0.02 | 4.16 ± 0.02 |
| This work | LG | $[-25; -17]$ | $[750; 20000]$ | -0.86 ± 0.01 | -23.24 ± 0.01 | 1.13 ± 0.02 | 4.25 ± 0.02 |
| This work | LG | $[-25; -21]$ | $[5000; 20000]$ | -0.76 ± 0.02 | -23.18 ± 0.01 | 1.14 ± 0.02 | 4.02 ± 0.02 |
| $ b > 10, K < 11.5$ | CMB | $[-25; -17]$ | $[300; 20000]$ | -0.94 ± 0.02 | -23.28 ± 0.01 | | |
| $1/V_{\text{max}}$ fit | CMB | $[-25; -21]$ | $[750; 20000]$ | -1.03 ± 0.02 | -23.43 ± 0.01 | 0.85 ± 0.06 | 4.22 ± 0.11 |

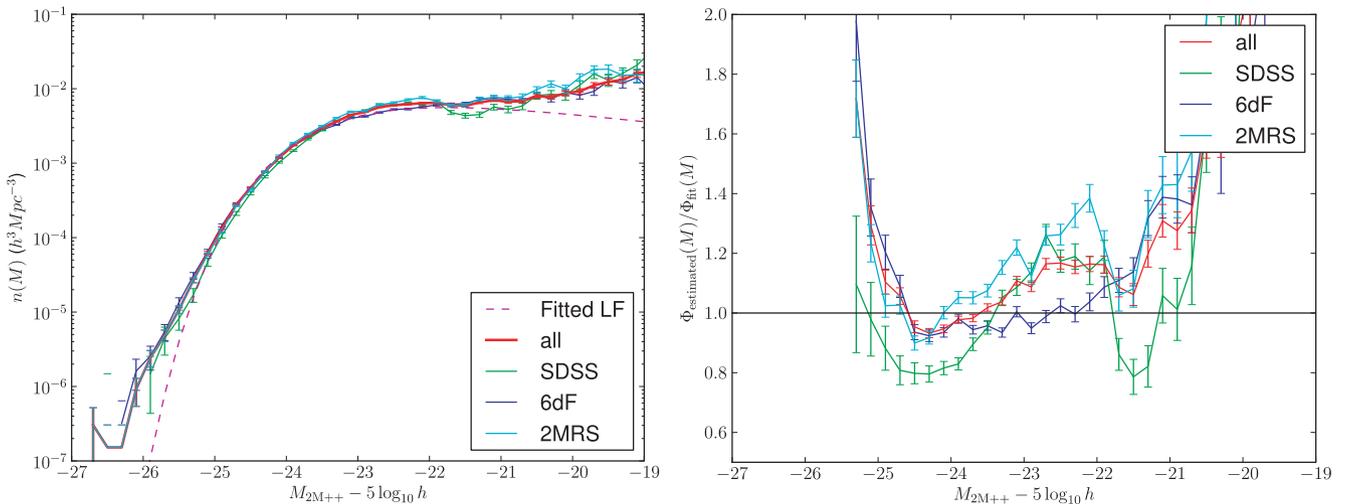


Figure 7. Galaxy LF estimates. The left-hand panel shows the non-parametric LF estimated using the $1/V_{\text{max}}$ method, shown by the solid lines for the regions covered either by 2MRS, 6dF, SDSS or all together. For these plots we use data from 750 to 20000 km s^{-1} . The dashed line shows the parametric LF using the likelihood method of Section 2.6 for galaxies with absolute magnitudes in the range $-25 \leq M \leq -21$, for redshift distances 5000 to 20000 km s^{-1} . The error bars reflect only the uncertainties in galaxy counts and do not include cosmic variance effects. The right-hand panel shows the difference between the $1/V_{\text{max}}$ LFs and the fitted parametric LF.

2.6.3 Discussion and comparison with previous results

Table 2 also lists LF parameters from previous 2MASS studies. Our fitted LF parameters are in agreement with previous studies of the K -band LF (Bell et al. 2003b; Eke et al. 2005) but are somewhat different than those found by Kochanek et al. (2001), Cole et al. (2001), Huchra et al. (2005) and Jones et al. (2006).

The derived LF parameters are sensitive to a number of systematic effects: the magnitude range used, the rest frame used for the redshifts, and the fitting method itself.

The Schechter function itself appears not to be a perfect fit over the whole range of magnitudes. Consequently, the fitted parameters depend on the magnitude (and distance) range of the galaxies used in the fit. Our default minimum distance $r \geq 5000 \text{ km s}^{-1}$ corresponds to $M_K \lesssim -21$ for $K_{2M++} = 12.5$. However, the $1/V_{\text{max}}$ method seems to indicate an inflection in the LF at $M_K \sim -21$. This bend is also seen by Bell et al. (2003b) and Eke et al. (2005). Indeed, Bell et al. (2003b) attempted to fit the part at $M_K > -21$ with a power law instead of a Schechter function. Several studies (Biviano et al. 1995; Yagi et al. 2002) have noted a dip in the LF of cluster galaxies at a similar location (approximately 2 mag below M^*), although other studies suggest that it is a flattening rather than a dip (Trentham 1998). In any case, it seems clear that the choice of magnitude range will affect the Schechter LF parameters. In Jones et al. (2006) and Cole et al. (2001), the magnitude range used in the fit is fainter than our default.

A second issue, which arises when using galaxies with very low redshifts, is the choice of flow model or rest-frame redshifts. Very nearby galaxies are likely to share the peculiar velocity of the Local Group (LG), so the redshift in the LG frame is a better proxy distance than the CMB-frame redshift. For better understanding of the dependence of our results on both local flows and clustering, we have fitted the parameters of the Schechter function in two rest frames (CMB or LG). We find that, for samples extending to $M_K \sim -17$, the faint-end slope α steepens, but only by 0.06.

Finally, the magnitudes, the correction and the fitting method itself are probably the most important systematics.

(i) We note that the studies of Cole et al. (2001), Eke et al. (2005) and Bell et al. (2003b) are based on Kron magnitudes, and that of Jones et al. (2006) is based on total magnitudes, leading to a possible difference in M^* of 0.20 ± 0.04 , as discussed by Kochanek et al. (2001).

(ii) Another notable difference is that Jones et al. (2006) have tried to integrate the effect of uncertainties on the determination of magnitudes, which we do not do here.

(iii) Bell et al. (2003b) match SDSS redshifts to both the 2MASS XSC and the PSC catalogues. Bell et al. (2003b) argue that selection effects bias the raw 2MASS LF compared to the true LF. However, whereas those authors were interested in, for example, the total stellar mass density in the nearby Universe, *our goal* is rather a consistent magnitude system coupled with uniform selection across the sky. Since our primary method will be to weight by luminosity, the small missing contribution from low surface brightness galaxies and the low surface brightness regions of catalogued galaxies is of little concern to us.

(iv) The fitting method itself may also make a difference. Our default parametric fit is pinned to the magnitude range where the formal Poisson errors are smallest, namely $[-25, -22]$. However, we have seen that systematic effects can be important. As an alternative, we have taken the LF given by the $1/V_{\text{max}}$ method, added in quadrature the statistical error bars and the fluctuations from the different subcatalogues, and fitted these data with a Schechter LF.

As indicated in Table 2, we have obtained a steeper faint-end slope and a brighter M_* , which are in better agreement with Kochanek et al. (2001), Cole et al. (2001) and Huchra et al. (2005), but still discrepant with Jones et al. (2006).

We conclude that, given all of these systematics, our LFs are reasonably consistent with those that have been found previously. One aspect which can be improved is peculiar velocity corrections, but we postpone a fully self-consistent treatment of peculiar velocities and the LF determination to a future paper.

We confirm that the bright end part of the LF does not seem to follow a Schechter LF, as already seen by the 6dfGRS (Jones et al. 2006). This effect is clearly seen in the SDSS, 2MRS and 6dfGRS subsamples separately. The deviation becomes significant at $M_K \lesssim -25$, or 2 mag brighter than M_* , and is presumably due to brightest cluster galaxies, which have typical K -band magnitudes of ~ -26 (Lin & Mohr 2004) and have long been known to deviate from the extrapolation of a Schechter function (Tremaine & Richstone 1977).

We may check the consistency of this LF with the number of galaxies in the 2M++ catalogue. We predict that the total number of galaxies of redshifts between the distances $r_{\text{min}}(z_{\text{min}})$ and $r_{\text{max}}(z_{\text{max}})$, assuming the Schechter LF, is

$$N = \int_{M_{\text{min}}}^{M_{\text{max}}} \Phi(M) dM \int_{r_{\text{min}}}^{r_{\text{max}}} d^3r f(r, \hat{r}, M_{\text{min}}, M_{\text{max}}). \quad (23)$$

We plot this function as a solid green line in Fig. 5. We also show the predicted number of galaxies in each bin of the grey histogram by a solid blue line. We see that the prediction in each redshift bin agrees well with the observed number of galaxies, but the total is off by ~ 2 per cent. The difference comes both from the low-luminosity part of the luminosity which is not adjusted because of our cut at $cz \geq 5000 \text{ km s}^{-1}$ and the high-luminosity part for which objects are not following a Schechter LF, as in Fig. 7.

In Table 2, we also give the mean luminosity density \bar{L} as derived from equation (22). \bar{L} is a lot less sensitive than \bar{n} to the faint end of the LF. As before, the errors are dominated by systematics due to the different corrections from peculiar velocities and the adequacy of the Schechter function to fit the observed LF. Taking the average and computing the dispersion in values for \bar{L} for the four tests indicated in Table 2 yields $\bar{L} = (4.09 \pm 0.12) 10^8 h L_{\odot}$.

2.7 Weights

Using the LF, we may now compute the appropriate weights to give to observed galaxies to account for incompleteness of the redshift catalogue. Our long-term goal is to reconstruct the DM density, under the assumption that galaxies trace the DM. There are several ways to link the galaxy density to the DM density: assuming that there is a linear relation between the two fields, one might consider number weighting, in which the DM density is assumed to be related to the number density of galaxies, or luminosity weighting, which can serve as proxy for stellar mass, and so may be a better tracer of DM density. We will consider both of these schemes here. More complicated relationships, for example based on a halo model (Marinoni & Hudson 2002), will be considered in a future paper.

We compute number weighting based on the fraction of observed galaxies:

$$\begin{aligned} f_{\text{observed}}^N(\mathbf{r}, M_{\text{min}}, M_{\text{max}}) &= \frac{N_{\text{observed}}(\mathbf{r})}{N_{\text{average}}} \\ &= \frac{f(r, \hat{r}, M_{\text{min}}, M_{\text{max}})}{\int_{M_{\text{min}}}^{M_{\text{max}}} \Phi_0(M) dM}. \end{aligned} \quad (24)$$

The weight applied to each galaxy is then $1/f_{\text{observed}}^N(\mathbf{r})$. This procedure is common and has been used previously (e.g. Davis & Huchra 1982; Pike & Hudson 2005).

We follow a similar procedure for correcting the local luminosity density of galaxies by estimating how much light we are missing at the distance of each galaxy located at position \mathbf{r} . The fraction of luminosity is

$$f_{\text{observed}}^L(\mathbf{r}) = \frac{L_{\text{observed}}(\mathbf{r})}{L_{\text{average}}} = \frac{1}{L_{\text{average}}} \left[(c_b - c_f)(\hat{r}) \Gamma_{M_{\min}, M_{\max}}^{M^*, 1+\alpha}(m_b, r_{10}) + c_f(\hat{r}) \Gamma_{M_{\min}, M_{\max}}^{M^*, 1+\alpha}(m_f, r_{10}) \right], \quad (25)$$

with L_{observed} the mean luminosity expected to be observed in a small volume at position \mathbf{r} , $\hat{r} = \mathbf{r}/|\mathbf{r}|$, and L_{average} the mean luminosity emitted by galaxies in the Universe. The value of L_{average} is

$$L_{\text{average}} = \int_{M_{\min}}^{M_{\max}} L(M) \Phi_0(M) dM. \quad (26)$$

The weight to apply to each intrinsic luminosity of a galaxy is then $1/f_{\text{observed}}^L(\mathbf{r})$. This procedure has already also been used with success with observation and mock catalogues (e.g. Lavaux et al. 2008, 2010; Davis et al. 2011).

For our choice of absolute magnitudes, $M_{\min} = -25$ and $M_{\max} = -20$, the $2M++$ is volume limited up to $r_{\min} \sim 20 h^{-1}$ Mpc, and extends up to $r_{\max} = 300 h^{-1}$ Mpc. We find that at a distance of $\sim 150 h^{-1}$ Mpc, the galaxy number weights are typically between 10 and 400, depending on whether the region is limited to $K_{2M++} \leq 12.5$ or $K_{2M++} \leq 11.5$, respectively. Similarly, the luminosity weights range between ~ 2 and ~ 40 . So weighing by luminosity has the advantage that it is less noisy at large distances.

3 TREATING THE ZONE OF AVOIDANCE

The ‘ZoA’ is the region of the Galactic plane where observations of galaxies are difficult due to the extinction by Galactic dust and stellar confusion. We show in Fig. 8 the number of galaxies in $2M++$ with $K_{2M++} \leq 11.5$, in bins of $\sin(b)$, and corrected for incompleteness effects. We see that the distribution is close to flat as a function of Galactic latitude, except for a hole contained in the range $-10^\circ \leq b \leq 10^\circ$ of the latitudes. We define the ZoA in $2M++$

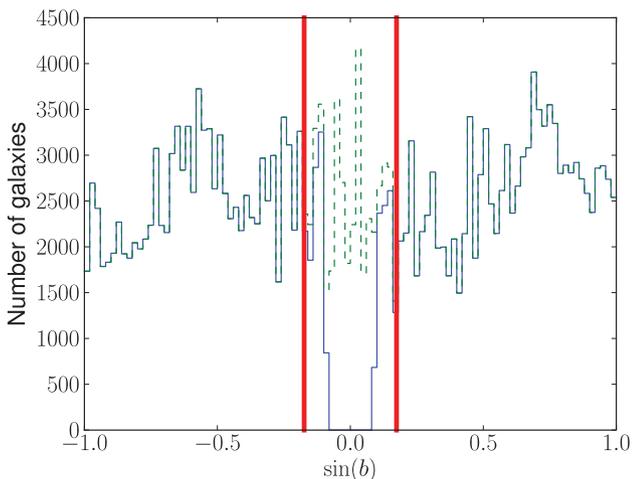


Figure 8. The effect of the ZoA on $2M++$. The weighed number density of galaxies in each bin of $\sin(b)$ is shown by the thin solid histogram. The dashed line shows the number density of galaxies once ZoA is filled with cloned galaxies. The two thick vertical lines correspond to $b = \pm 10^\circ$. Here we used only galaxies for which $c_z \leq 15000 \text{ km s}^{-1}$.

as this band for galactic longitudes $-30 \leq l \leq 30$, but reduce it to 5° outside this range. In addition, we impose the constraint for the absorption not to exceed $A_K = 0.25$ in regions devoid of galaxies.

In order to reconstruct the density field over the full sky, it is clearly necessary to fill the ZoA. One option is to fill it with mock galaxies so that their density of these objects matches the mean density outside the ZoA. This option would, however, fail to interpolate large-scale structure observed above and below the ZoA. The option adopted here, following Lynden-Bell, Lahav & Burstein (1989), is to ‘clone’ galaxies immediately above and below the ZoA. The procedure of creating a galaxy clone at a latitude b_c of a galaxy at latitude b is simply to shift the latitude:

$$\sin(b_c) = \sin(b_{\text{zoa}}) - \sin(b), \quad (27)$$

where

$$b_{\text{zoa}} = \begin{cases} \text{sign}(b) \times 5^\circ & \text{if } |b| > 5^\circ \text{ and } |l| > 30^\circ \\ \text{sign}(b) \times 10^\circ & \text{if } |b| > 10^\circ \text{ and } |l| < 30^\circ. \end{cases} \quad (28)$$

We refer to these as ‘ZoA-clones’. Fig. 8 shows the distribution of galaxies as a function of Galactic latitude before and after cloning. After cloning, the distribution shows no dependence on latitude.

4 GROUPING GALAXIES

We use redshifts to estimate galaxy distances, but in the presence of peculiar velocities this relationship is not perfect. In addition to the so-called ‘Kaiser (1987) effect’ which affects very large scales, there is also a contamination by the ‘finger-of-god effect’ due to the velocity dispersion of galaxies in clusters of galaxies. This causes a significant amount of noise on the redshift-estimated distance. One way to deal with the problem is to group galaxies, which by simple averaging improves the distances estimated from the redshifts. The grouping information is also interesting to study the statistics and properties of galaxy groups.

In this section, we describe the algorithm used to assign galaxies to groups and clusters. We use this information in the next sections for deriving a better density field (Section 5.1) and, in a future work, peculiar velocities. Grouping also allows a better determination of the centre of mass of superclusters (Section 5.2) and their infall pattern (Section 5.3). As a byproduct of $2M++$, we provide a catalogue of groups and their properties in Appendix B.

4.1 Grouping algorithm

To assign galaxies to groups we use the standard percolation or ‘friends-of-friends’ (FoF) algorithm developed by Huchra & Geller (1982). The algorithm is designed to identify cone-like structures in redshift space. Two galaxies are considered to be part of the same group if

- (i) their estimated angular distance separation is less than D_{sep} and
- (ii) their apparent total velocity separation is less than V_0 .

V_0 is kept fixed for the whole volume of the catalogue. D_{sep} is adapted such that the detected structures are always significant compared to the apparent local number density of galaxies, by explicitly accounting for selection effects. The constraint of a constant local overdensity at a redshift distance z leads to

$$D_{\text{sep}} = D_0 \left(\frac{\int_{L_{\min}(z)}^{+\infty} \Phi(L) dL}{\int_{L_{\min}(c_f)}^{+\infty} \Phi(L) dL} \right)^{-1/3}, \quad (29)$$

Table 3. Parameters for manual grouping of galaxies. All galaxies which are in the direction (l, b) and within the redshifts $[z_{\min}; z_{\max}]$ with a maximum angular separation to (l, b) equal to θ_{sep} are considered part of the group indicated in the first column.

| Group name | cz_{\min} | cz_{\max} (km s^{-1}) | θ_{sep} ($^{\circ}$) | l | b |
|------------|-------------|---------------------------------------|---|-----|-----|
| Virgo | $-\infty$ | 2500 | 10 | 279 | 74 |
| Fornax | $-\infty$ | 1600 | 8 | 240 | -50 |

with $L_{\min}(z)$ the minimum absolute luminosity observable at redshift z , $\Phi(L)$ the galaxy LF, z_{F} the fiducial redshift and D_0 the selection angular distance at fiducial redshift. The parameter D_0 is linked to the sought overdensity $\delta_{\text{overdensity}}$ for group detection by the relation

$$\delta_{\text{overdensity}} = \left(\frac{4\pi}{3} D_0^3 \int_{L_{\min}(z_{\text{F}})}^{+\infty} \Phi(L) dL \right)^{-1} - 1. \quad (30)$$

This density is computed at the fiducial redshift distance z_{F} .

We have chosen the following parameters for defining our groups: $V_{\text{F}} = cz_{\text{F}} = 1000 \text{ km s}^{-1}$ and $\delta_{\text{overdensity}} = 80$. These parameters have been used in previous studies (Ramella, Geller & Huchra 1989). With the LF, for our choice of fiducial parameters, we compute that the transverse linking length is $D_0 = 0.45 h^{-1} \text{ Mpc}$. We count 4002 groups with three or more members within 2M++, for redshift distance less than $20\,000 \text{ km s}^{-1}$. We do not group galaxies farther than $20\,000 \text{ km s}^{-1}$ where the catalogue becomes sparse. For the very nearby Virgo and Fornax clusters, the FoF algorithm fails and so we manually assign galaxies to nearby clusters according to the parameters given in Table 3.

4.2 Results

In Fig. 9, we plot the richness of detected groups as a function of redshift. In Fig. 10, we have plotted the total luminosities of the same groups. Finally, in Fig. 11, we give the velocity dispersion of the galaxies within these groups. As expected, the mean velocity dispersion does not vary significantly with distance, and has a mean value of $\sim 95 \text{ km s}^{-1}$. The richness is approximately constant up to $\sim 150 h^{-1} \text{ Mpc}$, which is a design feature of the group finder. The minimal luminosity of the groups increases with distance, as

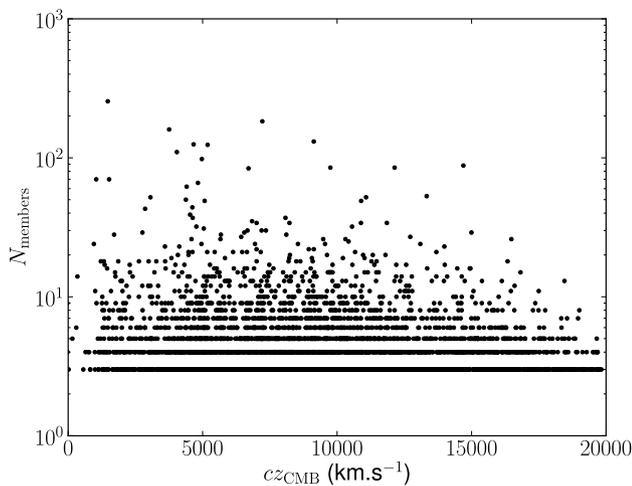


Figure 9. Number of group members (richness) as a function of redshift. The richness is not corrected for incompleteness.

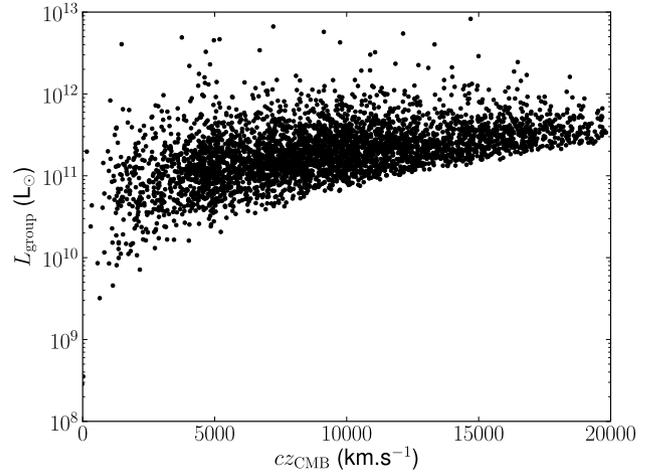


Figure 10. Group luminosity as a function of redshift. The luminosities are not corrected for incompleteness.

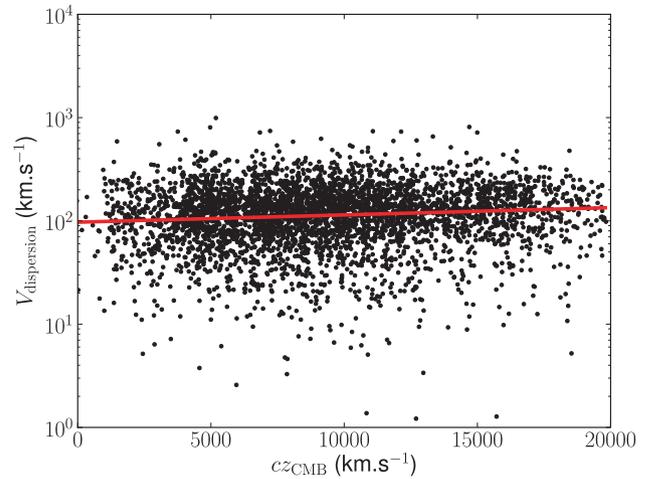


Figure 11. Group velocity dispersions as a function of redshift. The thick solid line indicates the trend of the evolution of the average velocity dispersion with redshift. The scale of the variation is $\sim 1.4 h^{-1} \text{ Gpc}$, far larger than the depth of the 2M++ catalogue.

we are losing the fainter objects at larger distances because the 2M++ catalogue is limited in apparent magnitude. The catalogue of group properties is given in Appendix B. We have checked that the parameters of the fitted Schechter LF do not change significantly after the grouping of the galaxies.

5 DENSITY FIELD

In this section, we consider some properties of the peaks in the three-dimensional density field obtained from the distribution of galaxies in the 2M++ galaxy redshift catalogue. We assume that the number density and luminosity density of galaxies follow a Poisson distribution. As such, the mean smoothed density contrast $\rho(\mathbf{x})$ given the galaxy weights w_i is

$$\rho(\mathbf{s}) = \frac{1}{\bar{\rho}} \sum_{i=1}^{N_{\text{galaxies}}} W(\mathbf{s} - \mathbf{s}_i) w_i, \quad (31)$$

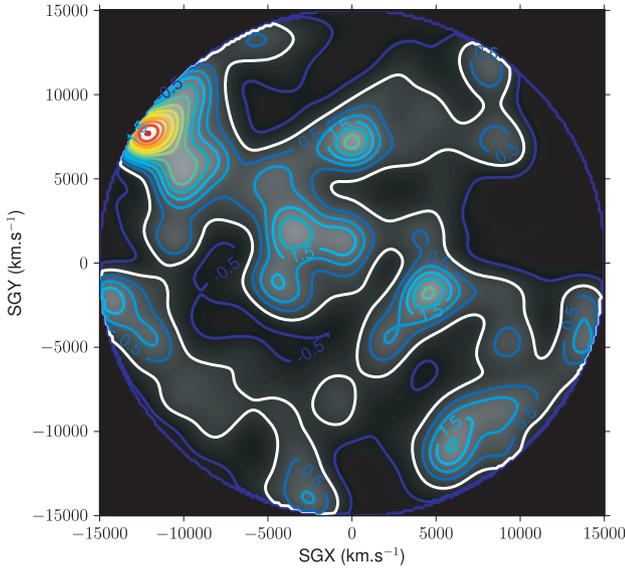


Figure 12. The 2M++ number density field in supergalactic plane. The density field is smoothed with a Gaussian kernel of 1000 km s^{-1} radius. Colour contours show the overdensity in units of the mean density and are separated by 0.5. The mean density is highlighted by a white contour.

and the standard deviation is

$$\sigma_\rho^2(\mathbf{s}) = \frac{1}{\rho^2} \sum_{i=1}^{N_{\text{galaxies}}} W(\mathbf{s} - \mathbf{s}_i)^2 w_i^2, \quad (32)$$

with \mathbf{s} the coordinate in redshift space and $W(\mathbf{x})$ the smoothing kernel considered. To compute the position of the peaks in this density field, we use an iterative spherical overdensity algorithm:

- (i) we initialize the algorithm with an approximation \mathbf{x}_c^0 of the expected position of the cluster;
- (ii) we compute the barycentre \mathbf{x}_c^{N+1} of the set of galaxies contained in a sphere centred on \mathbf{x}_c^N and with radius R_N ;
- (iii) we iterate (ii) until convergence, setting $R_{N+1} = R_N$;
- (iv) we reduce $R_{N+1} = 0.80R_N$. If $R_{N+1} > 1 h^{-1} \text{ Mpc}$, then we go back to step (ii), in the other case we terminate the algorithm.

We define the position of the structure as the one given by the last step in the above algorithm. This position is used in the following sections to compute mean densities and infall velocities on clusters.

5.1 Cosmography

Fig. 12 shows the galaxy number density field of our catalogue in the supergalactic plane, smoothed to $10 h^{-1} \text{ Mpc}$ with a Gaussian kernel. The SC in the upper-left corner, near $(\text{SGX}, \text{SGY}) \simeq (-10000, 7000) \text{ km s}^{-1}$, is particularly prominent and is the largest density fluctuation in the 2M++ catalogue. The Shapley region is covered by the 6dF portion of the survey which extends to a depth $K_{2M++} \leq 12.5$. Shapley is thus correctly sampled and is not a result of overcorrection of data limited to $K_{2M++} \leq 11.5$. When smoothed with a Gaussian kernel of $10 h^{-1} \text{ Mpc}$ radius, the SC peaks at $(l, b) = (312, 30)$ and $d = 152 h^{-1} \text{ Mpc}$ with a density $1 + \delta_g = 8.83 \pm 0.46$, in galaxy number density contrast, and $1 + \delta_L = 9.51 \pm 0.54$ in terms of luminosity density contrast.

The second most important structure in the supergalactic plane of the 2M++ catalogue is the Perseus-Pisces (PP) supercluster. It is clearly seen in the supergalactic plane in Fig. 12 at $(\text{SGX}, \text{SGY}) \simeq (5000, -1000) \text{ km s}^{-1}$. Its highest redshift space density, smoothed

with Gaussian kernel of $10 h^{-1} \text{ Mpc}$ radius, is about $1 + \delta_g = 4.46 \pm 0.18$ in terms of number density contrast, $1 + \delta_L = 4.47 \pm 0.20$ in terms of luminosity density contrast. The position of the peak corresponds to the Perseus cluster at $(l, b) = (150, -13)$, which is quite near the ZoA, and a distance of $52 h^{-1} \text{ Mpc}$. It is quite possible that the filling of the ZoA by galaxies cloned from the Perseus itself amplifies the overdensity of this supercluster.

The extended overdense structure in the central part of the supergalactic plane, at about $(\text{SGX}, \text{SGY}) \simeq (-5000, 0) \text{ km s}^{-1}$, is the Hydra–Centaurus–Virgo (HC) supercluster. At $10 h^{-1} \text{ Mpc}$ smoothing scale, the highest peak, located at $(l, b) = (302, 21)$, $d = 38 h^{-1} \text{ Mpc}$, coincides with the Centaurus cluster and has a height of $1 + \delta_g = 3.02 \pm 0.08$ in number density contrast and $1 + \delta_L = 3.40 \pm 0.14$ in luminosity density contrast.

Finally, Fig. 13 is a three-dimensional representation of the catalogue in Galactic coordinates, which means the Galactic plane goes through the middle of the vertical sides of the box, near the Norma cluster. We plot the 2M++ galaxies as points. Strong overdensities are highlighted by a transparent dark-red isosurface of density fluctuation of luminosity $\delta_L = 2$. This density has been smoothed at $10 h^{-1} \text{ Mpc}$ with a Gaussian kernel from the corrected number distribution. The Shapley supercluster is located at the top-left corner of the cube. A number of overdensities in the right part of the cube arise from the high weights, as this region has a depth of only $K_{2M++} \leq 11.5$.

5.2 Supercluster masses

We show in Fig. 14 the mean overdensity and excess mass within a sphere of $50 h^{-1} \text{ Mpc}$ for four important superclusters in the 2M++ catalogue: the SC, the PP supercluster, the Horologium–Reticulum (HR) supercluster centred at $(l, b) = (265, -51)$ and a distance of $193 h^{-1} \text{ Mpc}$ and the Hydra–Centaurus supercluster (see also Table 4). The profiles are centred on the position where the density peaks for each supercluster.

For the four superclusters, we note that the profiles obtained through number weighing and luminosity weighing are nearly equivalent. The bumps in the mean density, shown in the left-hand panels, are reproduced in both weighing schemes. This is particularly striking for the PP supercluster, even for scales as small as $10 h^{-1} \text{ Mpc}$. In all cases, the luminosity-weighted contrast is slightly lower than the number-weighted density contrast. In the following discussion, we adopt the luminosity-weighted number contrast.

The excess masses of all superclusters converge at radii of $\sim 50 h^{-1} \text{ Mpc}$. While Shapley is the most massive supercluster, we find that HR is very similar when measured on scales of $50 h^{-1} \text{ Mpc}$. Both have masses close to $10^{17} h^{-1} M_\odot$. The PP and HC superclusters are less massive, but, being considerably closer, these have more impact in the motion of the LG and nearby galaxies, as we discuss below.

Our estimate of Shapley’s mass and density contrast is similar to that of Proust et al. (2006) who measured a density contrast $\delta_n = 5.4 \pm 0.2$ in a truncated cone of 225 deg^2 between 90 and $180 h^{-1} \text{ Mpc}$ with a volume equivalent to a sphere of effective radius $30.3 h^{-1} \text{ Mpc}$. In a sphere of this radius centred on Shapley, we find a luminosity density contrast of $\delta_{K,L} = 4.1 \pm 0.15$.

Muñoz & Loeb (2008) calculated the mass of SC based on the overdensity of rich clusters and obtained a mass $3.3 \pm 0.3 \times 10^{16} h^{-1} M_\odot$ within a sphere of $35 h^{-1} \text{ Mpc}$. On the same scale, we obtain a mass of $4.87 \pm 0.18 \times 10^{16} h^{-1} M_\odot$, assuming $\Omega_m = 0.3$ and $b_{K,L} = 1$ for K -band luminosity. Using similar arguments, Sheth & Diaferio (2011) quote a mass of $1.8 \times 10^{16} h^{-1} M_\odot$

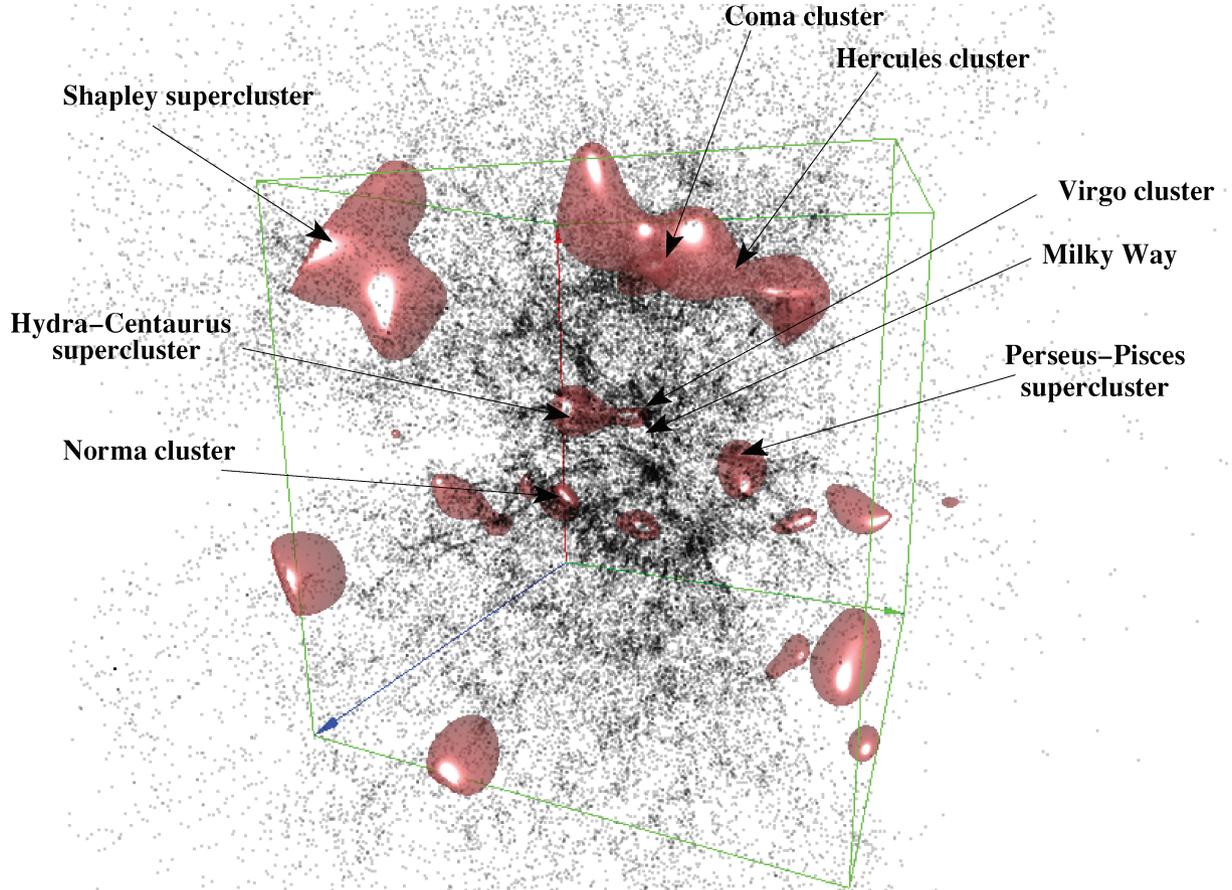


Figure 13. The 2M++ galaxy distribution and density field in three dimensions. The cube frame is in Galactic coordinates. The Galactic plane cuts orthogonally through the middle of the back vertical red arrow. The length of a side of the cube is $200 h^{-1}$ Mpc and is centred on Milky Way. We highlight the isosurface of number fluctuation, smoothed with a Gaussian kernel of radius 1000 km s^{-1} , $\delta_L = 2$ with a shiny dark red surface. The positions of some major structures in the local Universe are indicated by labelled arrows. We do not show isosurfaces beyond a distance of $150 h^{-1}$ Mpc, so HR is, for example, not present.

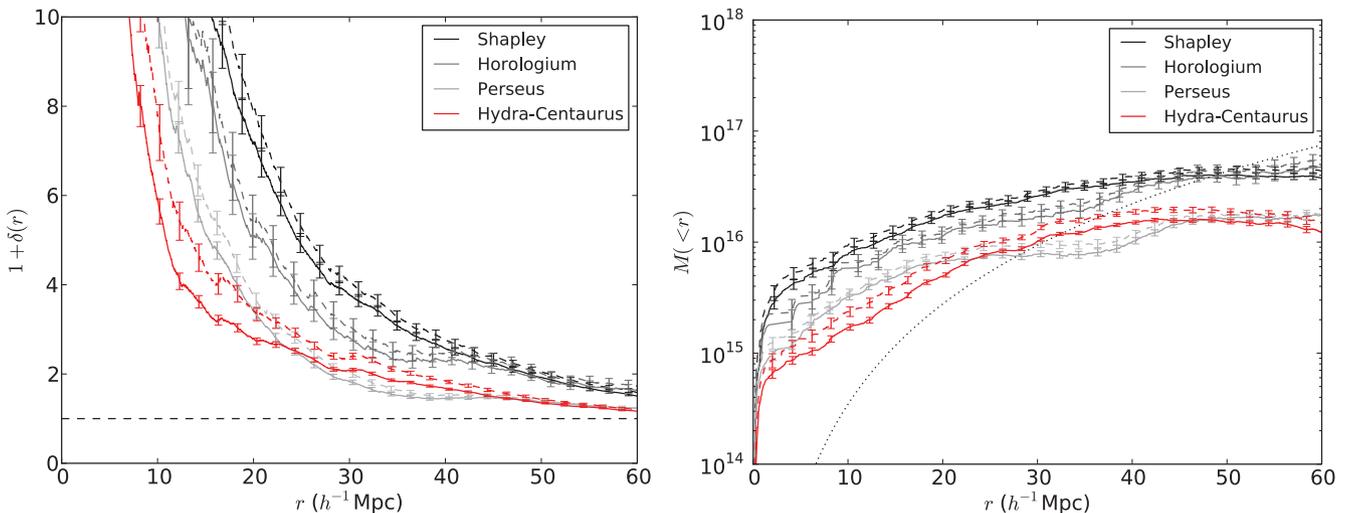


Figure 14. The cumulative average density profile and the excess mass as a function of radius from four major superclusters in the 2M++ redshift catalogue. In the two panels, we both show the profiles computed using the number weighed (solid lines) and the luminosity weighed (dashed lines) scheme. In the left-hand panel, the horizontal black dashed line corresponds to the mean density. In the right-hand panel, the dotted lines indicate the mass of a sphere of the given radius at the mean density. Note that since we are plotting excess mass, to obtain the total mass one must add this value. The black, dark grey, light grey and red lines correspond, respectively, to the SC, the HR supercluster, the PP supercluster and the Hydra-Centaurus supercluster. The error bars are estimated assuming that galaxies follow Poisson distribution for sampling the matter density field, as given by equation (32).

Table 4. Luminosity density contrast and estimated masses (assuming $\Omega_m = 0.3$ and $b_{K,L} = 1$) of the four superclusters from the distribution of galaxy light within a sphere of radius $50 h^{-1}$ Mpc.

| Supercluster | Sphere centre | | | $1 + \delta_L$ | Mass ($10^{16} h^{-1} M_\odot$) |
|----------------------|------------------------|---------------------|---------------------|-----------------|--------------------------------------|
| | r (h^{-1} Mpc) | l ($^\circ$) | b ($^\circ$) | | |
| Shapley | 152 | 312 | 30 | 2.05 ± 0.05 | 8.9 ± 0.2 |
| Horologium-Reticulum | 193 | 265 | -51 | 2.01 ± 0.10 | 8.7 ± 0.5 |
| Perseus-Pisces | 52 | 150 | -13 | 1.41 ± 0.03 | 6.1 ± 0.1 |
| Hydra-Centaurus | 38 | 302 | 21 | 1.43 ± 0.03 | 6.2 ± 0.1 |

within a slightly smaller radius of $31 h^{-1}$ Mpc. On the same scale, we find $4.00 \pm 0.17 \times 10^{16} h^{-1} M_\odot$. These values could be brought into rough agreement if luminosity-weighted 2MASS galaxies are strongly biased, with $b = 2-3$. Such a strong biasing would, however, conflict with the measurement $b_{K,n} = 1.05 \pm 0.10 (\Omega_m/0.3)^{0.55}$ by Pike & Hudson (2005) but may be marginally consistent with the lower value $b_{K,n} = 1.56 \pm 0.16 (\Omega_m/0.3)^{0.55}$ found recently by Davis et al. (2011).

A further caveat is that our density estimates are in redshift space, and so are enhanced by a factor up to $b_s = 1.2$ (Kaiser 1987)⁴ compared to the estimates of Muñoz & Loeb (2008) and Sheth & Diaferio (2011). In a future paper, we will reconstruct the density field in real space and calibrate the biasing factor directly using peculiar velocity data; so a detailed comparison of overdensities awaits future work.

Hudson et al. (2004) studied the overdensity of the SC as traced by *IRAS*-selected galaxies. Within a $50 h^{-1}$ Mpc-radius sphere they found that the overdensity of *IRAS*-selected galaxies is only 0.2. Here we find that the overdensity of 2MASS-selected galaxies on the same scale is ~ 1 . Clearly, the relationship between *IRAS*- and 2MASS-selected galaxies is not well described by a relative linear bias, since a value of ~ 5 would be required in the Shapley supercluster, whereas the field requires a relative bias between 2MASS- and *IRAS*-selected galaxies of ~ 1 (Pike & Hudson 2005).

5.3 Supercluster infall

We now discuss the impact these structures have on large-scale flows in the nearby Universe. We have estimated the infall velocity on to each of these structures using linear theory:

$$v_{\text{infall}} = \frac{1}{3} \beta H \bar{\delta}(R) R, \quad (33)$$

with H the Hubble constant, $\beta \equiv f/b$, where f is the linear density perturbation growth rate and b is a biasing parameter, and $\bar{\delta}(R)$ the mean density inside a sphere of radius R and centred on the supercluster. For a Λ CDM cosmology, $f \simeq \Omega_m^{5/9}$ (Bouchet et al. 1995). We use $\beta = 0.5$ (as determined for 2MASS galaxies by Pike & Hudson 2005) whenever we need to estimate a velocity. This value corresponds to $\Omega_m \simeq 0.30$ and $\Omega_\Lambda = 1 - \Omega_m$ with $b = 1$.

Fig. 15 shows the infall velocity profiles of the four superclusters. Although we plot the linear theory infall down to small radii ($R \lesssim 10 h^{-1}$ Mpc), we note that linear theory does not apply in these regions and focus the discussion on distances $R \gtrsim 10 h^{-1}$ Mpc. The infall velocities at $10 h^{-1}$ Mpc are all at least 2000 km s^{-1} , with Shapley having the highest infall at nearly 4000 km s^{-1} . At $50 h^{-1}$ Mpc, the Shapley and HR superclusters have an infall of

⁴ $b_s = \sqrt{1 + 2f/3 + f^2/5}$, with $f = 0.5$.

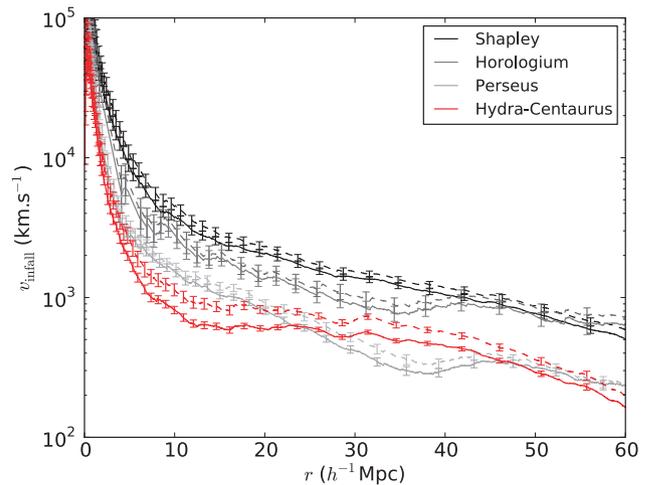


Figure 15. The infall velocities as a function of distance for four major superclusters in the 2M++ redshift catalogue. Curves and error bars are as in Fig. 14.

$\sim 800 \text{ km s}^{-1}$. The average overdensity of the SC within a sphere of $50 h^{-1}$ Mpc is $1 + \delta_L = 2.05 \pm 0.05$. Neglecting structures beyond $50 h^{-1}$ Mpc, linear theory implies that the supercluster is responsible for attracting the LG with a peculiar velocity of $90 \pm 10 \text{ km s}^{-1}$. This motion represents ~ 15 per cent of the total velocity of the LG with respect to the CMB rest frame. Although the excess mass of HR is similar to Shapley, its effect on the LG's motion is less than that of Shapley due to its greater distance: we estimate 60 km s^{-1} . Added vectorially, the net peculiar velocity from these two superclusters is approximately 110 km s^{-1} towards $(l, b) = (297, -1)$. This direction is within the errors of the direction of the 407 km s^{-1} bulk flow found by Watkins et al. (2009), but is lower in amplitude.

Because they are closer to the LG, the HC and PP superclusters have a greater impact. Approximating HC as a sphere, the infall at the LG's distance of $38 h^{-1}$ Mpc is $588 \pm 26 \text{ km s}^{-1}$. Whereas PP is denser, its greater distance of $52 h^{-1}$ Mpc puts it on the losing side of the gravitational tug of war with HC: the infall of the LG towards PP is only $313 \pm 24 \text{ km s}^{-1}$. Thus, the net motion is towards HC.

Note that it is likely that underdense regions also contribute a push. Kocevski & Ebeling (2006b) have noted the deficit of rich clusters in the northern sky, particularly in the distance range $130-180 h^{-1}$ Mpc. Thus, a full analysis of peculiar velocities requires integration over the entire density field, a topic we defer to a later paper.

6 SUMMARY

We have compiled a new, nearly full-sky galaxy redshift catalogue, dubbed 2M++, based on the data from three redshift surveys: the 2MRS ($K \leq 11.5$), the SDSS and the 6dFGRS. After having calculated corrected magnitudes and having calculated redshift completeness, we have determined LFs and weights that allow us to determine the redshift density field to a depth of $200 h^{-1}$ Mpc. The most prominent structure within $200 h^{-1}$ Mpc is the SC: its luminosity density within a sphere of radius $50 h^{-1}$ Mpc is 2.05 times the mean, and is thus responsible for approximately 90 km s^{-1} of the LG's motion with respect to the CMB rest frame. We have compared the density profile of four massive superclusters that are present in the 2M++ catalogue: the SC, the PP supercluster, the HR supercluster and Hydra–Centaurus. The SC is clearly the most massive of the four, but HR is only slightly less massive.

This new, deep full-sky catalogue will be used in future work to study the peculiar velocity of the LG and other nearby galaxies. Our hope is that the distribution of density in the 2M++ volume will account for the high-amplitude bulk motions on scales of $100 h^{-1}$ Mpc (Watkins et al. 2009; Lavaux et al. 2010).

ACKNOWLEDGMENTS

We thank Roya Mohayaee for stimulating discussions, and for helping to initiate this project at the Institut d'Astrophysique de Paris (IAP). We also thank the IAP for its hospitality.

This publication makes use of data products from the Two-Micron All-Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

We thank Lucas Macri and the 2MRS team for making their redshifts available in advance of publication.

We thank the 6dFGS team for the survey and for making their redshifts available at <http://www.aao.gov.au/local/www/6df/>.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society and the Higher Education Funding Council for England. The SDSS website is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

GL acknowledges financial support from NSF Grants AST 07-08849 and AST 09-08693. MJH acknowledges support from NSERC. GL and MJH acknowledge the financial support of the 'Programme visiteur de l'IAP' and the French ANR via the OTARIE project.

Finally, we acknowledge the lifetime work of John Huchra, a pioneer of galaxy redshift surveys, as well as the driving force behind the 2MASS Redshift Survey that is such a critical part of the data set compiled here. He will be missed.

REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543
 Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003a, *ApJS*, 149, 289
 Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003b, *ApJS*, 149, 289
 Bilicki M., Chodorowski M., Mamon G. A., Jarrett T., 2011, preprint (arXiv:1102.4356)
 Binggeli B., Popescu C. C., Tammann G. A., 1993, *A&AS*, 98, 275
 Biviano A., Durret F., Gerbal D., Le Fevre O., Lobo C., Mazure A., Slezak E., 1995, *A&A*, 297, 610
 Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003, *AJ*, 125, 2276
 Blanton M. R. et al., 2005, *AJ*, 129, 2562
 Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, *A&A*, 296, 575
 Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245
 Cole S. et al., 2001, *MNRAS*, 326, 255
 Colin J., Mohayaee R., Sarkar S., Shafieloo A., 2011, *MNRAS*, 414, 264
 Conselice C. J., Gallagher J. S., III, Wyse R. F. G., 2001, *ApJ*, 559, 791
 Davis M., Huchra J., 1982, *ApJ*, 254, 437
 Davis M., Nusser A., Masters K. L., Springob C., Huchra J. P., Lemson G., 2011, *MNRAS*, 413, 2906
 de Vaucouleurs G., de Vaucouleurs A., Corwin H. G. Jr., Buta R. J., Paturel G., Fouqué P., 1991, *Third Reference Catalogue of Bright Galaxies: Vol. I, Explanations and References; Vol. II, Data for Galaxies between 0^h and 12^h ; Vol. III: Data for Galaxies between 12^h and 24^h* . Springer-Verlag, New York
 Eke V. R., Baugh C. M., Cole S., Frenk C. S., King H. M., Peacock J. A., 2005, *MNRAS*, 362, 1233
 Erdođu P. et al., 2006a, *MNRAS*, 368, 1515
 Erdođu P. et al., 2006b, *MNRAS*, 373, 45
 Falco E. E. et al., 1999, *PASP*, 111, 438
 Feldman H. A., Watkins R., Hudson M. J., 2010, *MNRAS*, 392, 756
 Felten J. E., 1976, *ApJ*, 207, 700
 Hamilton A. J. S., Tegmark M., 2004, *MNRAS*, 349, 115
 Huchra J. P., Geller M. J., 1982, *ApJ*, 257, 423
 Huchra J. P., Geller M. J., Corwin H. G., Jr., 1995, *ApJS*, 99, 391
 Huchra J. et al., 2005, in Colless M., Staveley-Smith L., Stathakis R. A., eds, *Proc. IAU Symp. 216, 2MASS and the Nearby Universe*. Kluwer, Dordrecht, p. 170
 Hudson M. J., Smith R. J., Lucey J. R., Branchini E., 2004, *MNRAS*, 352, 61
 Jones D. H. et al., 2004, *MNRAS*, 355, 747
 Jones D. H., Peterson B. A., Colless M., Saunders W., 2006, *MNRAS*, 369, 25
 Jones D. H. et al., 2009, *MNRAS*, 399, 683
 Kaiser N., 1987, *MNRAS*, 227, 1
 Kashlinsky A., Atrio-Barandela F., Kocevski D., Ebeling H., 2008, *ApJ*, 686, L49
 Kashlinsky A., Atrio-Barandela F., Ebeling H., Edge A., Kocevski D., 2010, *ApJ*, 712, L81
 Kocevski D. D., Ebeling H., 2006a, *ApJ*, 645, 1043
 Kocevski D. D., Ebeling H., 2006b, *ApJ*, 645, 1043
 Kochanek C. S. et al., 2001, *ApJ*, 560, 566
 Kogut A. et al., 1993, *ApJ*, 419, 1
 Koribalski B. S. et al., 2004, *AJ*, 128, 16

- Lavaux G., Mohayaee R., Colombi S., Tully R. B., Bernardeau F., Silk J., 2008, *MNRAS*, 383, 1292
- Lavaux G., Tully R. B., Mohayaee R., Colombi S., 2010, *ApJ*, 709, 483
- Lin Y., Mohr J. J., 2004, *ApJ*, 617, 879
- Lynden Bell D., Lahav O., Burstein D., 1989, *MNRAS*, 241, 325
- Marinoni C., Hudson M. J., 2002, *ApJ*, 569, 101
- Muñoz J. A., Loeb A., 2008, *MNRAS*, 391, 1341
- Pike R. W., Hudson M. J., 2005, *ApJ*, 635, 11
- Proust D. et al., 2006, *A&A*, 447, 133
- Ramella M., Geller M. J., Huchra J. P., 1989, *ApJ*, 344, 57
- Rines K., Geller M. J., Kurtz M. J., Diaferio A., 2003, *AJ*, 126, 2152
- Sandage A., Tammann G. A., Yahil A., 1979, *ApJ*, 232, 352
- Saunders W. et al., 2000, *MNRAS*, 317, 55
- Schechter P., 1976, *ApJ*, 203, 297
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schmidt M., 1968, *ApJ*, 151, 393
- Schneider S. E., Thuan T. X., Magri C., Wadiak J. E., 1990, *ApJS*, 72, 245
- Sheth R. K., Diaferio A., 2011, preprint (arXiv:1105.3378)
- Skrutskie M. F. et al., 2006, *AJ*, 131, 1163
- Tremaine S. D., Richstone D. O., 1977, *ApJ*, 212, 311
- Trentham N., 1998, *MNRAS*, 293, 71
- Tully R. B., Shaya E. J., Karachentsev I. D., Courtois H. M., Kocevski D. D., Rizzi L., Peel A., 2008, *ApJ*, 676, 184
- Watkins R., Feldman H. A., Hudson M. J., 2009, *MNRAS*, 392, 743
- Yagi M., Kashikawa N., Sekiguchi M., Doi M., Yasuda N., Shimasaku K., Okamura S., 2002, *AJ*, 123, 87

APPENDIX A: THE 2M++ DATA

Table A1. The 2M++ catalogue. This is a sample of the full table, which is available as Supporting Information with the online version of the article. Column (1): the name of the galaxy as given in the 2MASS-XSC data base. Column (2): Galactic longitude in degrees. Column (3): Galactic latitude in degrees. Column (4): apparent magnitude in band K_s as defined in Section 2.2. Column (5): heliocentric total apparent velocity. Column (6): total apparent velocity in CMB rest frame, using relation from Kogut et al. (1993) and Tully et al. (2008). Column (7): total apparent velocity error (equal to zero if not measured). Column (8): unique group identifier obtained from the algorithm of Section 4. Column (9): redshift incompleteness at magnitude $K_{2M++} \leq 11.5$. Column (10): redshift incompleteness at magnitude $K_{2M++} \leq 12.5$. It may be empty; in that case the catalogue is limited to $K_{2M++} \leq 11.5$ in the portion of the sky holding the galaxy. Column (11): flag to indicate whether this is a fake galaxy to fill the ZoA following the algorithm of Section 3. Column (12): flag to indicate if the redshift has been obtained by the cloning procedure of Section 2.3. Column (13): flag to indicate whether this galaxy lies in the exclusive region covered by the 2MRS target mask (2Mx6S region). Column (14): flag to indicate whether this galaxy lies in the non-exclusion region covered by the SDSS. Column (15): same as (14) but for the 6dFGRS. Column (16): bibliographic code for the origin of the redshift information. The code is truncated in this table but available in full in the electronic version of the catalogue.

| Name | l ($^{\circ}$) | b ($^{\circ}$) | K_{2M++} | V_{helio} (km s^{-1}) | V_{CMB} (km s^{-1}) | V_{err} (km s^{-1}) | Group | $c_{11.5}$ | $c_{12.5}$ | ZoA | Cloned | M_0 | M_1 | M_2 | Bibcode |
|------------------|-----------------------|-----------------------|------------|--|--|--|-------|------------|------------|------|--------|-------|-------|-------|--------------|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
| 07345116-6917029 | 281.00 | -21.54 | 7.90 | 1367 | 1486 | 69 | 4996 | 1.0 | 1.0 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 21100305-5448123 | 342.30 | -41.63 | 12.31 | 18717 | 18571 | 0 | | 1.0 | 0.9 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 20353522-4422308 | 356.19 | -36.74 | 11.44 | 7066 | 6893 | 198 | 4388 | 1.0 | 1.0 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 13271270-2451409 | 313.15 | 37.30 | 12.04 | 12132 | 12429 | 0 | | 1.0 | 1.0 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 21112498-0849375 | 41.49 | -35.05 | 12.37 | 8296 | 7988 | 0 | 4177 | 1.0 | 0.9 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 02581778-0449064 | 182.17 | -52.44 | 10.50 | 9235 | 9036 | 0 | 3733 | 1.0 | 0.9 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 14303940+0716300 | 357.34 | 59.22 | 9.97 | 1370 | 1601 | 10 | | 1.0 | 0.8 | 0 | 0 | 0 | 1 | 0 | 1998AJ..... |
| 00362801+1226414 | 117.21 | -50.26 | 11.44 | 10339 | 9993 | 11 | | 0.8 | 0.0 | 0 | 0 | 1 | 0 | 0 | 20112MRS.... |
| 18490084+4739293 | 77.11 | 20.08 | 10.40 | 4671 | 4536 | 10 | | 0.8 | 0.0 | 0 | 0 | 1 | 0 | 0 | 1991RC3.9... |
| 23054906-8545110 | 305.13 | -30.91 | 12.45 | 19770 | 19788 | 270 | | 1.0 | 0.7 | 0 | 1 | 0 | 0 | 1 | none |
| 07511205-8540159 | 298.28 | -25.94 | 12.44 | 25990 | 26053 | 270 | | 1.0 | 0.6 | 0 | 1 | 0 | 0 | 1 | none |
| 03355460-8537067 | 299.59 | -30.39 | 11.90 | 12702 | 12736 | 270 | 3775 | 1.0 | 0.7 | 0 | 1 | 0 | 0 | 1 | none |
| 08423963-8430223 | 297.59 | -24.47 | 11.70 | 12284 | 12357 | 270 | 4039 | 1.0 | 1.0 | 0 | 1 | 0 | 0 | 1 | none |
| 08431792-8429053 | 297.58 | -24.44 | 12.35 | 12284 | 12357 | 270 | 4039 | 1.0 | 1.0 | 0 | 1 | 0 | 0 | 1 | none |
| 08424060-8427453 | 297.55 | -24.44 | 11.96 | 12284 | 12357 | 270 | 4039 | 1.0 | 1.0 | 0 | 1 | 0 | 0 | 1 | none |
| ZOA0000000 | 330.07 | -0.41 | 10.21 | 5238 | 5387 | 0 | | 0.9 | 0.7 | 1 | 0 | 0 | 0 | 1 | zoa |
| ZOA0000001 | 330.49 | -0.33 | 11.31 | 8841 | 8989 | 0 | | 0.9 | 0.6 | 1 | 0 | 0 | 0 | 1 | zoa |
| ZOA0000002 | 330.01 | -1.18 | 11.96 | 11830 | 11983 | 0 | | 1.0 | 0.9 | 1 | 0 | 0 | 0 | 1 | zoa |
| ZOA0000003 | 331.29 | -1.10 | 11.09 | 3158 | 3306 | 70 | | 0.9 | 0.8 | 1 | 0 | 0 | 0 | 1 | zoa |
| 07243410-8543223 | 298.20 | -26.43 | 11.51 | 5301 | 5361 | 69 | 4638 | 1.0 | 0.7 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 03403012-8540119 | 299.56 | -30.29 | 12.41 | 12714 | 12749 | 66 | 3775 | 1.0 | 0.6 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 07400785-8539307 | 298.20 | -26.13 | 11.09 | 1845 | 2465 | 11 | 4638 | 1.0 | 0.7 | 0 | 0 | 0 | 0 | 1 | 2008ApJ..... |
| 03355460-8537067 | 299.59 | -30.39 | 11.90 | 12702 | 12736 | 270 | 3775 | 1.0 | 0.7 | 0 | 1 | 0 | 0 | 1 | none |
| 07420104-8525161 | 297.96 | -26.04 | 10.31 | 5150 | 5213 | 14 | 4638 | 1.0 | 0.7 | 0 | 0 | 0 | 0 | 1 | 20096dF.... |
| 02090195-8520255 | 301.12 | -31.51 | 11.52 | 12675 | 12699 | 11 | 3776 | 1.0 | 0.9 | 0 | 0 | 0 | 0 | 1 | 20112MRS.... |

APPENDIX B: THE 2M++ GROUP CATALOG

Table B1. The 2M++ group catalogue. This is a sample of the full table, which is available as Supporting Information with the online version of the article. Column (1): group identifier in the catalogue. It corresponds to column (8) of Table A1. Column (2): Galactic longitude. Column (3): Galactic latitude. Column (4): apparent magnitude in band K_S as defined in Section 2.2. The magnitude is derived from the 2M++ galaxies. This is a magnitude uncorrected for incompleteness effect. Column (5): richness, uncorrected for incompleteness effect. Column (6): heliocentric total apparent velocity. Column (7): total apparent velocity in CMB rest frame, using relation from Kogut et al. (1993) and Tully et al. (2008). Column (8): velocity dispersion in the group.

| Id | l ($^\circ$) | b ($^\circ$) | K_{2M++} | N_{galaxies} | V_{helio} (km s^{-1}) | V_{CMB} (km s^{-1}) | σ_V (km s^{-1}) |
|------|---------------------|---------------------|------------|-----------------------|--|--|--------------------------------------|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1 | 281.26 | 73.47 | 6.60 | 65 | 1223 | 1556 | 599 |
| 1000 | 182.41 | -13.06 | 10.65 | 4 | 3 | -25 | 97 |
| 1001 | 91.75 | 51.01 | 6.15 | 4 | 691 | 752 | 51 |
| 1002 | 137.70 | 12.33 | 5.39 | 9 | 1139 | 1055 | 118 |
| 1003 | 316.23 | -10.88 | 10.42 | 3 | -80 | 17 | 21 |
| 1004 | 123.60 | 74.51 | 3.80 | 9 | -249 | -33 | 156 |
| 1005 | 184.67 | 83.05 | 4.96 | 24 | 675 | 956 | 312 |
| 1006 | 144.05 | 66.22 | 3.59 | 69 | 845 | 1049 | 227 |
| 1007 | 171.51 | 32.81 | 8.58 | 4 | 506 | 646 | 46 |
| 1008 | 108.51 | 58.06 | 5.83 | 6 | 186 | 302 | 109 |
| 1009 | 319.07 | -12.21 | 4.75 | 4 | -146 | -65 | 89 |
| 1010 | 144.70 | 36.20 | 3.55 | 5 | 92 | 157 | 82 |
| 1011 | 33.58 | 14.01 | 8.21 | 4 | 1865 | 1780 | 101 |
| 1012 | 41.38 | 14.94 | 8.00 | 3 | 2291 | 2188 | 12 |
| 1013 | 134.90 | 32.70 | 6.94 | 4 | 1330 | 1348 | 51 |
| 1014 | 103.67 | 33.03 | 8.10 | 4 | 1174 | 1129 | 63 |
| 1015 | 41.06 | 12.68 | 9.01 | 3 | 2741 | 2626 | 93 |
| 1016 | 45.04 | 17.71 | 8.24 | 3 | 2261 | 2162 | 60 |
| 1017 | 150.98 | 5.93 | 8.69 | 4 | 5096 | 5028 | 112 |
| 1018 | 147.84 | 7.90 | 8.10 | 6 | 4807 | 4736 | 237 |
| 1019 | 129.32 | 8.92 | 7.37 | 5 | 3276 | 3146 | 203 |
| 1020 | 103.21 | 12.39 | 7.94 | 5 | 2692 | 2523 | 131 |
| 1021 | 50.91 | 6.89 | 8.34 | 4 | 4831 | 4659 | 193 |
| 1022 | 69.13 | 8.13 | 8.96 | 4 | 4556 | 4359 | 81 |
| 1023 | 75.62 | 6.03 | 8.22 | 4 | 4711 | 4497 | 94 |
| 1024 | 269.12 | 5.57 | 9.91 | 3 | 5052 | 5324 | 75 |
| 1025 | 264.14 | 7.22 | 8.89 | 4 | 4685 | 4965 | 234 |
| 1026 | 275.43 | 8.94 | 9.48 | 3 | 4008 | 4291 | 88 |
| 1027 | 264.36 | 8.40 | 8.29 | 7 | 4796 | 5081 | 121 |
| 1028 | 295.46 | 8.81 | 8.36 | 4 | 4449 | 4701 | 113 |

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table A1. The 2M++ catalogue.

Table B1. The 2M++ group catalogue.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.