

# The dipole anisotropy of AllWISE galaxies

M. Rameez,<sup>1★</sup> R. Mohayaee,<sup>1,2★</sup> S. Sarkar<sup>1,3★</sup> and J. Colin<sup>1,2</sup>

<sup>1</sup>Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

<sup>2</sup>Sorbonne Universités, UPMC Univ Paris 06, CNRS, Institut d'Astrophysique de Paris, 98bis Bld Arago, F-75014 Paris, France

<sup>3</sup>Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford, OX1 3NP, United Kingdom

Accepted 2018 March 2. Received 2018 March 1; in original form 2017 December 11

## ABSTRACT

We determine the dipole in the WISE (Wide Infrared Satellite Explorer) galaxy catalogue. After reducing star contamination to  $<0.1$  per cent by rejecting sources with high apparent motion and those close to the Galactic plane, we eliminate low redshift sources to suppress the non-kinematic, clustering dipole. We remove sources within  $\pm 5^\circ$  of the supergalactic plane, as well as those within  $1''$  of 2MRS sources at redshift  $z < 0.03$ . We enforce cuts on the source angular extent to preferentially select distant ones. As we progress along these steps, the dipole converges in direction to within  $5^\circ$  of the Cosmic Microwave Background (CMB) dipole and its magnitude also progressively reduces but stabilizes at  $\sim 0.012$ , corresponding to a velocity  $> 1000 \text{ km s}^{-1}$  if it is solely of kinematic origin. However, previous studies have shown that only  $\sim 70$  per cent of the velocity of the Local Group as inferred from the CMB dipole is due to sources at  $z < 0.03$ . We examine the Dark Sky simulations to quantify the prevalence of such environments and find that  $< 2.1$  per cent of Milky Way-like observers in a  $\Lambda$ CDM universe should observe the bulk flow ( $> 240 \text{ km s}^{-1}$  extending to  $z > 0.03$ ) that we do. We construct mock catalogues in the neighbourhood of such peculiar observers in order to mimic our final galaxy selection and quantify the residual clustering dipole. After subtracting this, the remaining dipole is  $0.0048 \pm 0.0022$ , corresponding to a velocity of  $420 \pm 213 \text{ km s}^{-1}$ , which is consistent with the CMB. However, the sources (at  $z > 0.03$ ) of such a large clustering dipole remain to be identified.

**Key words:** catalogues – surveys – galaxies: distances and redshifts – large-scale structure of Universe – cosmology: theory.

## 1 INTRODUCTION

The standard cosmological model assumes that the Universe is statistically isotropic and homogeneous on large scales. The isotropy of the Universe has supposedly been observationally confirmed by the Cosmic Microwave Background (CMB) temperature fluctuations on small angular scales which do not show any significant directional dependence (WMAP Collaboration 2011; Planck Collaboration XXVII 2016). However the dipole anisotropy of the CMB is  $\sim 100$  times larger than that of the higher multipoles (see e.g. Kogut et al. 1993). The latter are believed to have originated from an inflationary era in the early universe. Hence the CMB dipole is considered *not* to be of primordial origin and is attributed to the motion of the Solar system in the ‘CMB rest frame’ (in which the universe is exactly isotropic), due to the attraction of a nearby large-density inhomogeneity.

If the dipole anisotropy is indeed due to our motion, then the barycentre of the Solar system is moving at  $369 \text{ km s}^{-1}$  towards  $RA = 168^\circ$ ,  $Dec. = -7^\circ$ , or  $l = 263.85^\circ$ ,  $b = 48.25^\circ$  in Galactic coordinates (Stewart & Sciamia 1967; Peebles & Wilkinson 1968; Hinshaw et al. 2009). Due to the indirect nature of this inference, which relies on there being no primordial dipole, an *independent* direct measurement of this velocity is desirable. This can be done by observing the aberration of the CMB (Challinor & van Leeuwen 2002; Burles & Rappaport 2006), however the effect is too small to have been detected convincingly (with  $> 3\sigma$  significance) even using the latest data (Planck Collaboration XVI 2014).

The standard  $\Lambda$ CDM model extended to first order in perturbation theory does predict some anisotropy in the local Universe on small scales. These are however expected to become progressively smaller as the average is taken over larger volumes, leading to the emergence of homogeneity (and isotropy) on the large scale. That this indeed happens on scales exceeding  $\sim 100 \text{ Mpc}$  has been claimed from observing the scale dependence of counts of galaxies in the SDSS (Hogg et al. 2005) and WiggleZ (Scrimgeour et al. 2012) surveys. As the CMB is an integrated map of the Universe, it cannot trace this transition, however galaxy surveys can indeed do

\* E-mail: mohamed.rameez@nbi.ku.dk (MR); mohayaee@iap.fr (RM); s.sarkar@physics.ox.ac.uk (SS)

so. They show that along the direction of the CMB dipole lie the most massive neighbouring superclusters: Virgo, the Great Attractor Hydra-Centaurus, Coma, Hercules and Shapley, and possibly other yet-to-be-mapped superclusters. However the gravitational attraction of these structures can account at most for 80 per cent of the velocity interpreted from the CMB dipole anisotropy (Lavaux et al. 2010; Colin et al. 2011; Feindt et al. 2013).

A dipole in the distribution of galaxies, which is usually known as the ‘clustering dipole’ ( $\mathcal{D}_{\text{cls}}$ ), can cause us to move in a preferred direction. This motion would lead to an additional anisotropy observable through the aberration and Doppler boost effects by an amount which depends on our velocity, further increasing the total observed dipole. A measurement of this latter effect, often referred to as the ‘kinematic dipole’ ( $\mathcal{D}_{\text{kin}}$ ), can provide an independent confirmation of our velocity. Aberration and Doppler boost effects can be measured only in a sample that is intrinsically isotropic, such as a directionally unbiased catalogue of high redshift sources. More often the total dipole ( $\mathcal{D}$ ) will be a mixture of the clustering and the kinematic dipoles. A further complication is the dipole generated by statistical noise due to the finite size of the galaxy catalogue (see e.g. Itoh, Yahata & Takada (2010)), as well as contamination of the catalogue by foreground stars in our Galaxy.

Most galaxy surveys are shallow and contain significantly large clustering dipoles from sources at distances less than a few hundred Mpc. Estimating our velocity from these catalogues would require the disentangling of the clustering and the kinematic dipoles. Ideally surveys that are very deep in redshift should be able to provide observations at much larger scales, where the clustering dipole is expected to be subdominant. Radio galaxies are expected to provide such an alternative as they are extremely luminous and unaffected by obscuration due to dust, hence probe the Universe at high redshifts  $z > 1$ . Nearly all measurements of the kinematic dipole using radio galaxy catalogues are however *discrepant* with the CMB dipole. While the velocity of the Solar system barycentre inferred from the CMB temperature dipole anisotropy is  $369 \text{ km s}^{-1}$ , the value inferred from radio-galaxy catalogues, e.g. NVSS, ranges from  $700 \text{ km s}^{-1}$  to over  $2000 \text{ km s}^{-1}$  (Blake & Wall 2002; Singal 2011; Gibelyou & Huterer 2012; Rubart & Schwarz 2013; Tiwari et al. 2015; Tiwari & Jain 2015; Colin, Mohayaee & Rameez 2017). The direction of the anomalously high velocity is however found to be quite well-aligned with the CMB dipole.

Several previous studies have tested for homogeneity and isotropy in galaxy surveys with photometric redshifts, e.g. the Sloan Digital Sky Survey (SDSS) Data Release 6 (Itō et al. 2010), the Two Micron All Sky Survey Photometric Redshift (2MRS) survey (Apleby & Shafieloo 2014; Alonso et al. 2015), the Wide Infrared Satellite Explorer (WISE) survey (Yoon et al. 2014), the WISE-2MASS catalogue (Bengaly et al. 2017), and the WISExSUPER-COSMOS catalogue (Bengaly et al. 2018). We go beyond all these works by examining the *largest* unbiased all-sky survey of the infrared sky, namely the Wide-field Infrared Survey Explorer – All Sky (WISE-allsky) and its subsequent extension – the AllWISE catalogues. As the WISE-allsky and AllWISE catalogues also contain stars and other point-like objects from within the Galaxy, these have to be first removed. Star contamination can be suppressed to the level of a few percent using methods described in Kovács & Szapudi (2015). It can be further reduced to  $< 0.05$  per cent by exploiting information from the apparent motion fits made possible by the NEOWISE post-cryogenic phase of the WISE survey, supplied with the AllWISE catalogue. The SDSS (SDSS Collaboration 2009) Data Release 13 is used as a reference catalogue to estimate star contamination through cross-correlation.

In order to suppress the clustering dipole, we proceed to remove as many local sources as possible in a *directionally unbiased* manner. This is done by first removing the sources that correlate with the 2MRS catalogue, as well as removing sources in symmetric bands around the supagalactic plane along which the most important superclusters in the local Universe lie. The sample can be further reduced by selecting for more distant galaxies by removing extended sources. At each of these steps, we evaluate the dipole. We show that as the clustering dipole is progressively suppressed, the total strength of the dipole is reduced as expected, while the direction converges towards the direction of the CMB dipole. We further test the robustness of this dipole by removing the well-known local superclusters (symmetrically, to avoid producing spurious dipole effects) such as Shapley directly from the catalogue. The Galaxy and Mass Assembly (GAMA) (Liske et al. 2015) is used as a reference spectroscopic survey to estimate the redshift distribution.

At this stage, for a typical observer in a  $\Lambda$ CDM universe in a Milky Way-like halo (i.e. with similar mass and velocity), the remaining dipole should be dominated by the kinematical component. However the dipole we find at this stage is  $\sim 0.0124$ , implying a Solar system barycentre velocity of  $\sim 1039 \text{ km s}^{-1}$  if it is indeed of purely kinematic origin. This is nearly three times as large as the velocity inferred from the CMB dipole.

In order to determine whether this final kinematic dipole is contaminated by any residual clustering dipole, we produce catalogues with the characteristics of our AllWISE selection from the  $z = 0$  halo catalogue of the Dark Sky simulation (Skillman et al. 2014). We construct the catalogue slice by slice in comoving distance to reproduce the redshift distribution of our final selection. The catalogues are constructed around haloes similar in mass and peculiar velocity of the Galaxy. In addition, we also consider constrained observers similar to us, i.e. in environments where the  $z = 0.03$  sphere around them has a bulk motion of  $220\text{--}260 \text{ km s}^{-1}$ . We find that the residual clustering dipole is *larger* around such constrained observers, who constitute however  $< 3$  per cent of Milky Way-like observers. After subtracting this estimate of the average expected clustering dipole from the total observed dipole, we find the remaining dipole to correspond to a velocity of  $402 \pm 183 \text{ km s}^{-1}$ , which is consistent with our inferred motion through the CMB.

This paper is organized as follows. In Section 2 the kinematic dipole is defined while Section 3 describes our methods for estimating the total dipole. In Section 4, we introduce the data set used and Section 5 describes how we minimize contamination by foreground stars. In Section 6 we describe methods to remove sources at low redshift in order to reduce the clustering dipole and in Section 7 we estimate the residual clustering dipole in our final selection, from theory as well as the  $z = 0$  halo catalogue of the Dark Sky simulations. Finally in Section 8, we estimate the velocity of the Solar system barycentre.

## 2 THE KINEMATIC DIPOLE

An observer moving with velocity  $v$  in the rest frame of an intrinsically isotropic distribution of sources observes a dipolar modulation in the number count of sources with amplitude (Ellis & Baldwin 1984; Itō et al. 2010):

$$\mathcal{D}_{\text{kin}} = \frac{v}{c} [2 + x(1 + \alpha)], \quad (1)$$

where  $x$  and  $\alpha$  are flux indices, defined through the integral source counts

$$dN/d\Omega(> S) = kS^{-x}, \quad (2)$$

and the flux density at a fixed observing frequency:

$$S_{\text{obs}} = S_{\text{rest}} \delta^{1+\alpha}. \quad (3)$$

### 3 THE DIPOLE ESTIMATORS

The dipole anisotropy in a catalogue can be estimated by calculating the difference in the number of sources in the upper and lower hemispheres. In order to estimate the dipole direction, the direction of orientation of the hemispheres can be varied until the maximum hemispherical number count is obtained. The strength of the dipole is then given by

$$d_{\text{HC}} = \frac{N^{\text{UH}} - N^{\text{LH}}}{N^{\text{UH}} + N^{\text{LH}}} \times \hat{r}_{\text{max}} \quad (4)$$

where  $N^{\text{UH}}$  and  $N^{\text{LH}}$  are the numbers of sources in the upper and lower hemispheres, respectively. We scan a healpy map of NSIDE = 32, to find the direction with the maximum hemispheric difference in number count (Colin et al. 2017).

While this estimator is robust, its statistical variability is high and this produces a biased estimate of the magnitude of the dipole as shown in Colin et al. (2017). Therefore we also use a 3D linear estimator (Crawford 2009), defined by

$$d_{3\text{D}} = \frac{1}{N} \sum_{i=0}^N \hat{r}_i, \quad (5)$$

where  $N$  is the number of sources in the catalogue and  $\hat{r}_i$  is the unit vector in the direction of the source  $i$ . The statistical properties of this estimator have been studied previously (Rubart & Schwarz 2013; Colin et al. 2017).

The bias of these estimators is a measure of how the statistical noise due to finite sample size affects the measurement of the total dipole. It can be precisely quantified for a sample of a given size as was done in Colin et al. (2017) only when the total expected non-noise dipole is well known. In the presence of an additional component of the dipole such as  $\mathcal{D}_{\text{cls}}$ , the bias factors estimated in Colin et al. (2017) effectively serve as upper limits to the total bias.

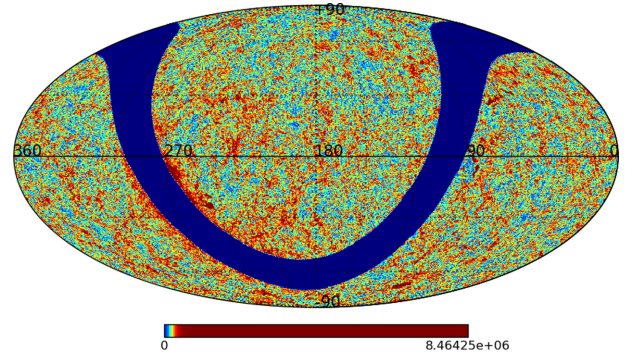
### 4 WISE AND ALLWISE DATA

The Wide-field Infrared Survey Explorer (WISE) mapped the sky in 2010 at 3.4, 4.6, 12, and 22  $\mu\text{m}$  (W1, W2, W3, W4) with an angular resolution of 6''.1, 6''.4, 6''.5, and 12''.0, respectively. The WISE-allsky catalogue contains positions and the four-band photometry for 563 921 584 objects which include both stars and galaxies (Wright et al. 2010).

The AllWISE catalogue supersedes the earlier WISE catalogue by combining data from the WISE cryogenic (Wright et al. 2010) and NEOWISE (Mainzer et al. 2011) post-cryogenic survey phases and presently forms the most comprehensive view of the full mid-infrared sky. The AllWISE source catalogue contains astrometry and photometry for 747 634 026 objects (Cutri et al. 2013). In addition to increased depth, sensitivity, and improved flux variability, the AllWISE Source Catalogue provides an estimate of the apparent motion of each source, exploiting the two independent WISE sky coverage epochs.

### 5 DATA PREPARATION I: STAR-GALAXY SEPARATION

The vast majority of sources in the WISE-allsky and AllWISE catalogues are stars within our own Galaxy. These have to be removed



**Figure 1.** The density of the AllWISE catalogue (number of sources/steradian) with the Galactic plane removed at  $\pm 10^\circ$  and the galaxy selection of Kovács & Szapudi (2015) applied, corresponding to catalogue (I) of table 2. This is a healpix map of nsid 128.

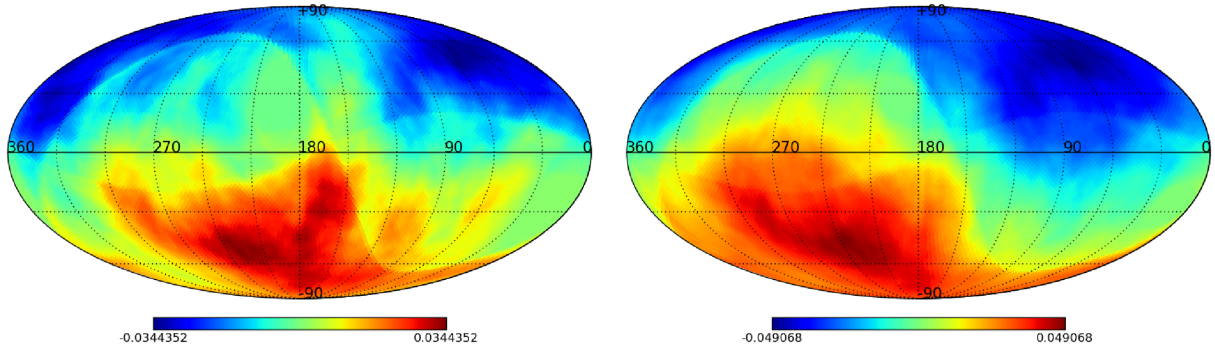
before a meaningful study of the dipole can be carried out. The star contamination can be estimated by cross-matching with SDSS (SDSS Collaboration 2009) with a tolerance of 1 arcsec, as was previously done (Yoon et al. 2014). As SDSS is a spectroscopic survey, it uniquely identifies stars and galaxies and hence serves as a control sample for star-galaxy separation.

#### 5.1 Magnitude cuts and Galactic plane removal

In order to remove stars contaminating our sample of galaxies, we follow Kovács & Szapudi (2015), who provide a separation strategy for objects in the WISE-allsky catalogue which have also been observed in 2MASS survey (Skrutskie et al. 2006). A sample of galaxies with 76 per cent galaxy completeness and just  $\sim 2$  per cent star contamination can be obtained by rejecting sources with W1 magnitude greater than 15.2 and  $W1 - J_{2\text{MASS}} > -1.7$ . Making a further  $J_{2\text{MASS}} < 16.5$  cut (Yoon et al. 2014) and masking out contaminated regions in the WMAP Galactic dust mask can reduce the star contamination of the sample to 1.2 per cent (70.1 per cent galaxy completeness). However for the final sample to remain unbiased to dipole direction estimators, a cut removing sources at Galactic latitudes  $|b| < 10^\circ$  is preferred. After these cuts on the WISE-allsky catalogue, we obtain a sample of  $\sim 2.359$  million objects with a star contamination of 1.8 per cent and 74 per cent galaxy completeness, while the AllWISE catalogue yields  $\sim 2.367$  million objects with a star contamination of 1.9 per cent and the same completeness. The density map of this catalogue is shown in Fig. 1.

The hemispherical number count estimator can now be applied to both the WISE-allsky and AllWISE samples to estimate the direction and magnitude of the maximum dipole anisotropy in the sky. Fig. 2 indicates the hemispherical number count differences in different directions for the two catalogues. The direction and magnitude of the maximum anisotropy are summarized in Table 1.

For the sample derived from WISE-allsky, we find that the dipole is in the direction  $\text{RA} = 228.6^\circ$  and  $\text{Dec.} = -52.8^\circ$  (or  $l = 323.67^\circ$  and  $b = 4.2175^\circ$  in Galactic coordinates) and corresponds to a hemispherical number count difference of 3.4 per cent. For the AllWISE catalogue we find a similar direction for the dipole of  $\text{RA} = 237.4^\circ$  and  $\text{Dec.} = -46.6^\circ$  ( $l = 331.9^\circ$  and  $b = 6.02^\circ$ ) and a hemispherical number count difference of 4.9 per cent. These directions are significantly different from that of the CMB dipole, being  $\sim 66^\circ$  and  $\sim 70^\circ$  away, respectively. Our results are in broad agreement with the previous findings, e.g. by Yoon et al. (2014), who also determined the galaxy bias  $b$  at this stage to be  $1.41 \pm 0.07$  (see Section 6.2



**Figure 2.** The left-hand and right-hand panels show the galaxy selection from WISE All-Sky and AllWISE catalogues, in equatorial coordinates. For each direction in the sky, the colour encodes:  $(N^{\text{UH}} - N^{\text{LH}})/(N^{\text{UH}} + N^{\text{LH}})$ , where  $N^{\text{UH}}$  and  $N^{\text{LH}}$  are the numbers of sources in the upper and lower hemisphere, respectively. The directions are scanned using a healpy map of nsid 32. The direction of the dipole is about  $66^\circ$  and  $70^\circ$  away from that given by the CMB for the WISE All Sky and AllWISE selections, while its magnitude is 3.4 per cent and 4.9 per cent, respectively. These are healpix maps of nsid 32.

**Table 1.** Observed dipoles in the galaxy selection described in Section 5.1 with the hemispherical count estimator.

Catalogue	$N$	RA $^\circ$	Dec. $^\circ$	$l^\circ$	$b^\circ$	$\angle$ CMB	$ \mathcal{D} $
WISE All Sky	2,359,212	228.6 $^\circ$	-52.8 $^\circ$	323.7 $^\circ$	4.2 $^\circ$	66.9 $^\circ$	0.034
AllWISE	2,367,162	237.4 $^\circ$	-46.6 $^\circ$	331.9 $^\circ$	6.0 $^\circ$	70.8 $^\circ$	0.049

for more details). While a fully kinematic origin for these large dipoles would require velocities higher than  $\sim 4000 \text{ km s}^{-1}$ , the fact that their direction points to low Galactic latitudes is indicative of residual contamination by stars.

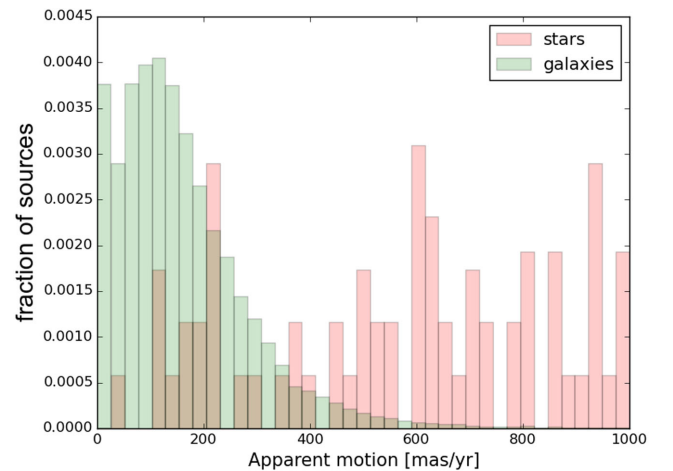
## 5.2 Removing sources with large apparent motion

To further suppress star contamination in the catalogue, we widen the symmetric band around the Galactic plane within which sources are removed to  $|b| < 15^\circ$ , leaving behind 2.09 million sources in AllWISE. Subsequently we use the apparent motion measurements as follows. Most objects in AllWISE have been observed only in two or three epochs, and consequently the proper motion and parallax components of the apparent motion cannot be disentangled. In general, it is the closest objects to the Sun that have substantial proper motions. These objects also have significant parallaxes. The stars that remain in our sample subsequent to the even wider Galactic latitude cuts at  $\pm 15^\circ$ , reside at high Galactic latitudes and are hence nearby. Consequently, the apparent motion provides an excellent discriminator between extragalactic and Galactic objects (Vieira, Garcia-Varela & Sabogal 2017).

The value of the apparent motion,  $am$ , is given by the sum of the best-fitting proper motion in RA and Dec. separately:

$$am = \sqrt{(\cos(\text{Dec.}) \times \text{RA}_{\text{am}})^2 + \text{Dec.}_{\text{am}}^2}, \quad (6)$$

where  $\text{RA}_{\text{am}}$  and  $\text{Dec.}_{\text{am}}$  are the motions in RA and Dec. calculated separately. We combine these into one variable, the distribution of which is shown in Fig. 3. Objects for which the motion fit fails to converge, which constitute  $\sim 2.6$  per cent of the sample, are considered to have zero motion at this stage. Kurcz et al. (2016) examine the quality of the apparent motion fits and conclude that they are reliable only for those sources for which the apparent motion fit has a signal-to-noise ratio larger than 1, albeit with a different definition of the apparent motion. The impact of applying both this more stringent criterion on the apparent motion, as well as assuming the objects with a failed motion fit to have zero motion, is re-examined at the final stage of our analysis (see Section 8).



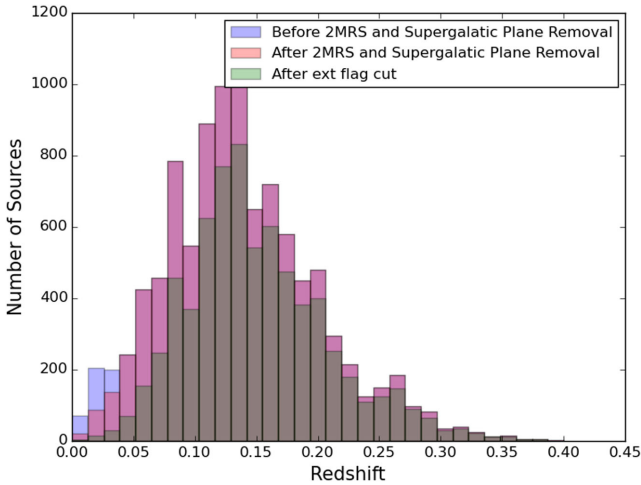
**Figure 3.** The distribution of the AllWISE best-fitting apparent motion for sources identified as galaxies, and those identified as stars through SDSS cross-correlation. The normalization of the stars has been scaled up (by a factor of  $\sim 30$ ) for visibility.

After removing all sources with  $am > 400 \text{ mas yr}^{-1}$ , we are left with about 1.91 million objects. Among the 36 086 sources that correlate to SDSS sources within 1 arcsec, we find 34 stars, implying a star contamination of less than 0.1 per cent.

## 6 DATA PREPARATION II: SUPPRESSING THE CLUSTERING DIPOLE

The clustering dipole ( $\mathcal{D}_{\text{cls}}$ ) is dominated by the contribution from nearby sources so should decrease at high redshifts. The kinematic dipole ( $\mathcal{D}_{\text{kin}}$ ), due to the motion of the observer caused by the anisotropic distribution of mass seen in the clustering dipole, is however independent of the distance to the sources.

To suppress the contribution of the local clustering dipole to the total dipole and extract the kinematic dipole, it is desirable to remove as many sources as possible at low redshifts, in a directionally



**Figure 4.** Redshift distribution for 5400 sources of AllWISE that are matched to those of GAMA survey. The median redshift is 0.137–0.164, depending on the masks.

unbiased manner. The various steps in the process of suppressing the clustering dipole are described in the following subsections.

WISE being a photometric instrument, the AllWISE catalogue does not provide redshift measurements. We estimate the redshift distribution of these data by cross-matching with the GAMA catalogue (Liske et al. 2015). This is a spectroscopic survey of about 300 000 galaxies down to  $r < 19.8$  magnitude over about  $286 \text{ deg}^2$ . The GAMA survey builds on the previous spectroscopic surveys such as the SDSS which we have used already to estimate the star contamination.

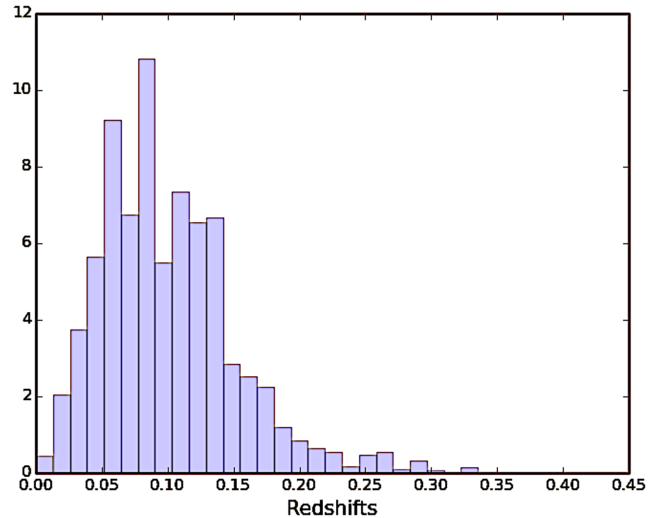
Of the 5620 AllWISE sources at this stage that fall within the solid angle scanned by GAMA, 5491 have cross-matched counterparts. The redshift distribution of these sources is shown in Fig. 4, which also indicates how the distribution evolves in the later stages of this analysis.

### 6.1 Removing the supergalactic plane and sources correlating with 2MRS at $z < 0.03$

A large fraction of the mass in the nearby universe, out to  $z = 0.03$ , is known to be clustered along a planar structure known as the supergalactic plane. In order to exclude this, we add a supergalactic latitude cut of  $\pm 5^\circ$ , which ensures that most of the local superclusters that lie on this plane are removed. Since both the galactic and supergalactic planes form great circles in the celestial sphere, removing an area centred on them leaves the direction of the dipole estimators unbiased.

In order to further suppress any local superstructures that lie outside the supergalactic plane, we cross-correlate our AllWISE galaxy catalogue with the 2MRS catalogue (Huchra et al. 2012) and remove all objects that are common to the two catalogues. This is done by identifying all AllWISE sources that are within 1 arcsec of 2MRS sources out to  $z = 0.03$ , beyond which 2MRS is not complete. Of the 24 648 2MRS sources below redshift  $z = 0.03$ , only 2392 have AllWISE counterparts at this stage – in contrast to Section 5.1, when all 24 648 sources did have counterparts. Consequently, the impact of removing these sources is small.

Subsequent to these cuts, we are left with  $\sim 1.71$  million objects. The median redshift at this stage was found to be  $\sim 0.137$  and the 3D linear estimator of equation 5 finds the direction and the magnitude of the dipole to be  $\text{RA} = 177.4^\circ$ ,  $\text{Dec.} = -49.9^\circ$  ( $l = 292.9^\circ$ ,



**Figure 5.** The redshift distribution of the sources (with  $\text{ext\_flag} \neq 0$ ) which are discarded in Section 6.2. The median redshift is 0.091. As explained in Section 7.2, these sources induce a velocity of at least  $85 \text{ km s}^{-1}$  on the Local Group.

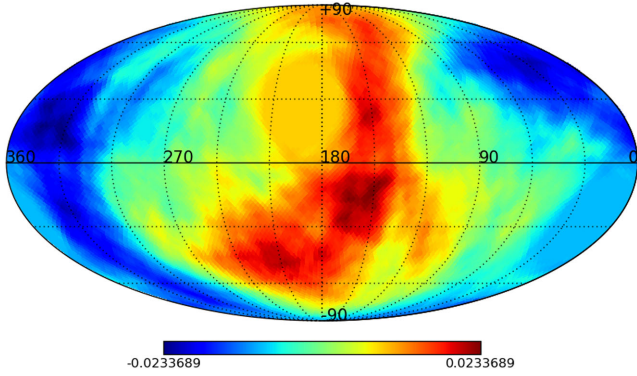
$b = 11.7^\circ$ ) and 0.017, respectively. The dipole direction is now  $43.7^\circ$  away from the CMB dipole. Evidently the removal of local structures slightly reduces the amplitude of the dipole (previous value was 0.018) and brings its direction closer to that of the CMB.

### 6.2 Discarding extended sources

The WISE satellite has an angular resolution of  $\sim 6.1$  arcsec in the  $3.4 \mu\text{m}$  band, which corresponds to  $2.96 \times 10^{-5}$  radians. Galaxies, which are typically a few tens of kpc across, are resolved as extended sources at distances less than a few hundred Mpc. Galaxies of similar size at larger distances are contained within the angular beam size of the detector and appear to be point sources. Hence, discarding extended sources at this stage can significantly suppress the fraction of nearby objects. The AllWISE catalogue provides a variable ‘ext\_flg’, which has a value of 0 if the morphology of the source is consistent with the WISE point spread function, and not associated with a known 2MASS extended source. Higher values of the variable indicate high goodness of fits for extended source profiles.

Therefore we select only sources with ‘ext\_flg=0’, which leaves us with a sample of 1.23 million sources. Fig. 5 illustrates the redshift distribution of the removed sources. The median redshift at this stage is found to have increased to 0.164, testifying to the suppression of low redshift sources. Applying the 3D linear estimator of equation (5) to this sample, we find the dipole to be in the direction  $\text{RA} = 166.2^\circ$ ,  $\text{Dec.} = -15.7^\circ$  ( $l = 269.17^\circ$ ,  $b = 40.17^\circ$ ), i.e. only  $8.8^\circ$  away from the CMB dipole. Its magnitude is now 0.0124, a significant reduction from the previous value of 0.017 (see Section 6.1).

If we further widen the Galactic plane cut to  $\pm 20^\circ$ , then the dipole direction swings to  $\text{RA} = 172.6^\circ$ ,  $\text{Dec.} = -6.6^\circ$  ( $l = 269.7^\circ$ ,  $b = 51.0^\circ$ ), which is just  $4.5^\circ$  away from the CMB dipole, with a magnitude of 0.011 according to the 3D estimator. The hemispheric-count estimator of equation (4) finds the dipole to lie towards  $\text{RA} = 151.9^\circ$ ,  $\text{Dec.} = -15.7^\circ$  ( $l = 255.1^\circ$ ,  $b = 31.5^\circ$ ), which is  $18.0^\circ$  away from the CMB dipole, with a magnitude of 0.023 which is twice that found from the 3D estimator. The hemispherical count map at this stage is shown in Fig. 6. The discrepancy between the



**Figure 6.** The hemispherical count map, in equatorial coordinates, of the AllWISE-galaxy selection as described in Section 6.2. This is a healpix map of nsid 32.

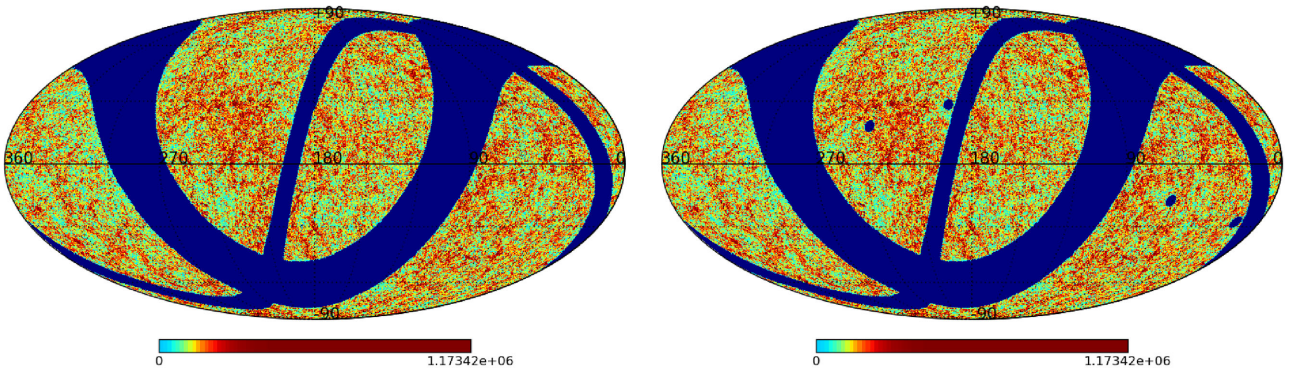
two estimators may be due to the larger bias of the hemispherical count estimator (Colin et al. 2017).

The galaxy bias,  $b$ , of this final sample can be evaluated using the method of Kovács & Szapudi (2015) and the *SPICE* (Szapudi et al. 2000) software package and is found to be  $1.27 \pm 0.12$  (excluding the  $l = 1$  mode from the fit). The lowering of the galaxy bias can be attributed to both the increase in the median redshift of the sample due to suppression of nearby objects, as well as the removal of extended galaxies, which correspond to the largest and most massive galaxies which are known to have a higher bias.

### 6.3 Removal of individual local superclusters

The positions of nearby clusters and superclusters are well-known. To test the robustness of the direction and magnitude of the total dipole estimated in Section 6.2 with respect to contamination by these sources, we further mask out the sky around the Shapley supercluster and the Coma and Hercules Clusters. To keep the dipole estimators directionally unbiased, an equivalent mask has to be applied to the sky diametrically opposite to these directions (see Fig. 7).

The precise details of these ‘punctures’ are as follows:  $4^\circ$  radius around Shapley at RA = 202.5°, Dec. =  $-31.0^\circ$ ,  $3^\circ$  around Hercules at RA = 241.3°, Dec. =  $17.75^\circ$ , and around Coma supercluster at RA = 194.95°, Dec. =  $27.98^\circ$ . After these cuts, there remain 1 193 188 sources. The results are presented in Table 2. The removal of one or more of these sources affects the direction of the dipole by



**Figure 7.** The densities of the final AllWISE-galaxy selections (Number of sources/steradian), corresponding to catalogues IV (left) and V (right) of table 2. The punctures around the most prominent superclusters of the local Universe, Shapley, Hercules, and Coma (and their symmetric images) are performed as a robustness test. These are healpix maps of nsid 128.

only  $1\text{--}4^\circ$  and the change in dipole magnitude is also insignificant. Most of these local structures have already been accounted for in the previous steps, hence these additional cuts have negligible impact as expected.

## 7 THE RESIDUAL CLUSTERING DIPOLE

We can evaluate theoretically the expected clustering dipole in the AllWISE catalogue, from the angular power spectrum  $C_l$  of the galaxy distribution. This is related to the 3D spectrum  $P(k)$  through (see e.g. Huterer et al. 2001; Tegmark et al. 2002)

$$C_l = b^2 \frac{2}{\pi} \int_0^\infty f_l(k)^2 P(k) k^2 dk. \quad (7)$$

Here  $b$  is the galaxy bias and the amplitude of the clustering dipole  $\mathcal{D}_{\text{cls}}$  is related to  $C_1$ , the  $l = 1$  mode of the angular power spectrum, as  $\mathcal{D}_{\text{cls}} = \sqrt{\frac{9}{4\pi} C_1}$ . The filter function  $f_l(k)$  is given by

$$f_l(k) = \int_0^\infty j_l(kr) f(r) dr, \quad (8)$$

where  $j_l$  is the spherical Bessel function which for the  $l = 1$  mode is  $\sin(kr)/(kr)^2 - \cos(kr)/(kr)$ ,  $f(r)$  is the probability distribution for the comoving distance  $r$  to a random galaxy in the survey which is given by the distributions in Fig. 4. For simplicity and lack of better information, we set the correction factor for  $f(r)$  due to evolution to be unity (see e.g. Tegmark et al. 2002) and assume that  $b$  is independent of redshift. The function  $f(r)$  is then proportional to

$$f(r) \sim \frac{H(z)}{H_0 r_0} \frac{dN}{dz}, \quad (9)$$

which is normalized such that  $\int_0^\infty f(r) dr = 1$ .

We set  $r_0 = c/H_0 \simeq 3000 h^{-1}$  Mpc and use the code *astropy* (Robitaille et al. 2013) with cosmological parameters from *Planck* (Planck Collaboration XVI 2014) to obtain both the Hubble parameter and the comoving radius as a function of redshift,  $r(z)$ . The distribution  $dN/dz$  can be approximated by splines fit to the different histograms of Fig. 4.  $P(k)$  can be obtained from CAMB (Lewis et al. 2000), at redshifts corresponding to the median redshift of the selection. Evaluating the above expressions numerically, we find the total average clustering dipole to be  $\sim 0.0095 b$  before the 2MRS-correlated and Supergalactic plane sources are removed, while after their removal it is  $0.0068 b$ . It drops further to  $\sim 0.0052 b$  after sources with `ext_flag`  $\neq 0$  are removed.

**Table 2.** Dipoles obtained with the 3D estimator in the AllWISE-galaxy selection as it evolves through the cuts described in Sections 5 and 6. Where multiple selection options are indicated, the final selection which carries over to the next stage is marked by an asterisk (\*). For reference, the CMB dipole has a magnitude of 0.0012.

Catalogue	$N$	RA $^{\circ}$	Dec. $^{\circ}$	$l^{\circ}$	$b^{\circ}$	$\angle$ CMB	$ \mathcal{D} $	
AllWISE-2MASS								
(I) (Section 5.1); $ b  \geq 10^{\circ}$	2 367 162	281.8 $^{\circ}$	-57.7 $^{\circ}$	338.1 $^{\circ}$	-22.1 $^{\circ}$	96.4 $^{\circ}$	0.032	
$ b  \geq 15^{\circ}$	2 092 276	233.3 $^{\circ}$	-64.7 $^{\circ}$	319.2 $^{\circ}$	-7.01 $^{\circ}$	73.3 $^{\circ}$	0.021	*
(II) (Section 5.2); As in (I) + apparent motion $\leq 200$ mas yr $^{-1}$	1 427 368	156.1 $^{\circ}$	-29.2 $^{\circ}$	268.2 $^{\circ}$	23.6 $^{\circ}$	24.9 $^{\circ}$	0.023	
apparent motion $\leq 400$ mas yr $^{-1}$	1 912 219	182.9 $^{\circ}$	-55.6 $^{\circ}$	297.3 $^{\circ}$	6.8 $^{\circ}$	50.1 $^{\circ}$	0.018	*
(III) (Section 6.1); As in (II) + $ \text{supergalactic-}b  \geq 5^{\circ}$ + sources within $1''$ of 2MRS sources, with $z < 0.03$ removed	1 718 619	177.4 $^{\circ}$	-49.9 $^{\circ}$	292.9 $^{\circ}$	11.7 $^{\circ}$	43.7 $^{\circ}$	0.017	*
(IV) (Section 6.2); As in (III) + sources with ext_flag $\neq 0$ removed	1 233 920	166.2 $^{\circ}$	-15.7 $^{\circ}$	269.17 $^{\circ}$	40.17 $^{\circ}$	8.8 $^{\circ}$	0.0124	*
$+ b  \geq 20^{\circ}$	1 084 178	172.6 $^{\circ}$	-6.6 $^{\circ}$	269.7 $^{\circ}$	51.0 $^{\circ}$	4.5 $^{\circ}$	0.011	
(V) (Section 6.3); As in (IV) + local superclusters removed	1 193 188	174.3 $^{\circ}$	-11.1 $^{\circ}$	275.3 $^{\circ}$	47.7 $^{\circ}$	7.4 $^{\circ}$	0.0119	*

This evolution as we proceed through the various steps of selection towards our final sample is in reasonable agreement with the findings reported in Table 2 (which also include the kinematic dipole contribution).

The theoretical approach employed so far refers to a typical observer in a  $\Lambda$ CDM universe. A more precise analysis should consider only galaxies similar to the Milky Way and its environment in a  $N$ -body simulation as we do below.

### 7.1 Milky Way-like environments in the Dark Sky simulations

We quantify the size and variance of the expected clustering dipole by looking at the  $z = 0$  snapshot halo catalogue of ‘Dark Sky’ – a publicly available Hubble volume trillion-particle  $N$ -body simulation (Skillman et al. 2014). Objects are sampled from the  $z = 0$  snapshot according to a redshift distribution which mimics that of the AllWISE galaxy selection. Subsequently, we use the 3D estimator (5) to evaluate the dipole. Since we have not included the effect of aberration and Doppler boosting and calculate the angles from the coordinates of the objects in rest frame of the simulation, the dipole that we find is entirely *non*-kinematical, i.e. purely due to the clustering effects.

The intrinsic dipole in the distribution of mass around an observer determines the magnitude and the direction of the velocity at that position. Therefore to compare with our findings in the AllWISE galaxy selection, we examine catalogues constructed from the  $z = 0$  halo catalogue around observer haloes similar to that of the Milky Way in mass and velocity, viz. with virial mass  $M_{200}$  in the range  $2.2 \times 10^{11} - 1.4 \times 10^{12} M_{\odot}$  (Cautun et al. 2014) and velocity in the range 600–650 km s $^{-1}$ .

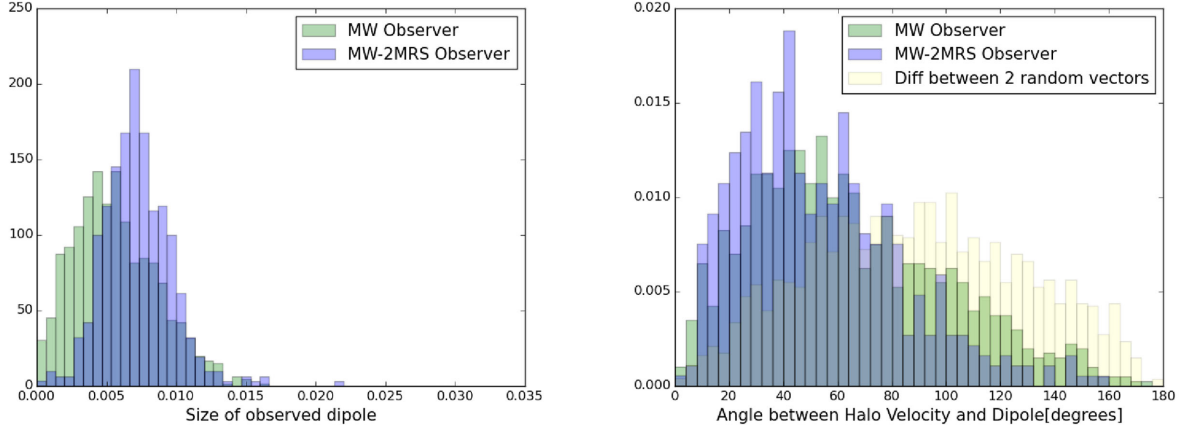
Each halo is assumed to correspond to a galaxy, and the most massive haloes are selected in radial shells centred around the selected Milky Way-type observer in a directionally unbiased manner so as to produce a comoving distance distribution corresponding to the redshift distributions shown in Fig. 4. Such a catalogue is expected

to have the same bias as the final AllWISE selection, although the galaxy-dark matter bias cannot be quantified from a simulation that includes only dark matter particles. Subsequently, the 3D estimator (5) is applied to such a catalogue to extract a purely intrinsic clustering dipole.

We assume that linear theory holds on these large scales, and that perturbations grow as  $\delta(k, t) = \delta(k, t)D(t)/D(t)$ . In order to accurately compare a catalogue constructed from a  $z = 0$  snapshot with an observed catalogue which is a light cone centred around the observer, the dipole contribution in each shell is scaled by the corresponding cosmological structure growth factor  $D(z)/D(z=0)$ . An examination of the catalogues constructed around 500 such observers yields dipoles with magnitudes and directions as shown in Fig. 8.

To mimic our environment as closely as possible, we now restrict ourselves to observers satisfying a more stringent criterion. Previous work has shown that the velocity of the Local Group in a rest frame of radius of  $\sim 120$  Mpc (corresponding to  $z \sim 0.03$ ) is  $\sim 350$  km s $^{-1}$  (see e.g. Lavaux et al. 2010), implying that a bulk flow of velocity  $\sim 240$  km s $^{-1}$  persists beyond  $z = 0.03$  (Colin et al. 2011; Feindt et al. 2013). Of the  $\sim 23\,800$  haloes satisfying the previous criterion, only 500 are found to satisfy the criterion that the average velocity of all haloes within a 120 Mpc sphere around it should be greater than 240 km s $^{-1}$ , making the probability of such a system  $\sim 0.021$ . 15 973 Milky Way-like observer haloes had to be examined before 500 were found with a similar bulk flow greater than 220 km s $^{-1}$ , corresponding to probability of  $\sim 0.031$ , while for a threshold of 250 km s $^{-1}$ , the probability is only  $\sim 0.009$ .

To estimate the residual clustering dipole in our sample, we constrain ourselves to such observers, excluding also neighbourhoods with velocities higher than the allowed ranges to avoid biases from regions with unusually large bulk flows. We find that the residual clustering dipole has a value of  $0.0076 \pm 0.0022$  (see Fig. 8) for the range 240–280 km s $^{-1}$ . The value is  $0.0079 \pm 0.0017$  and  $0.0071 \pm 0.0021$  for the ranges 250–290 km s $^{-1}$  and



**Figure 8.** Left: Intrinsic clustering dipoles observed in 500 realizations of AllWISE-like galaxy catalogues from Milky Way-like haloes (green) and Milky Way-like haloes in an environment as found in 2MRS by Lavaux et al. (2010), corresponding to a bulk flow velocity of the  $z = 0.03$  sphere in the range  $240\text{--}280\text{ km s}^{-1}$  (blue). Right: The observed angle between the observer velocity and the observed clustering dipole direction. The distribution of angles expected between two isotropic random dipoles is shown (in yellow) for comparison.

$220\text{--}260\text{ km s}^{-1}$ , respectively. The quoted uncertainties on these values correspond to  $1\sigma$ .

## 7.2 The clustering dipole and the velocity

While 2D catalogues such as the AllWISE-galaxy selection do not have redshift information, the velocity imposed by the dipole anisotropy on the Local Group of galaxies can still be estimated by employing the fluxes as proxies for the distance, assuming a narrow range of intrinsic luminosities for the sources under consideration. This idea was initially proposed by Gott and subsequently used for numerous surveys (see e.g. Yahil et al. 1986; Maller et al. 2003; Erdoğan et al. 2006; Bilicki et al. 2011).

The velocity–acceleration relation is (Peebles 1980)

$$v = \frac{2}{3} \frac{f(\Omega_m)}{H_0 \Omega_m} g, \quad (10)$$

where the acceleration  $g$  is given by

$$g = \frac{G}{b} \sum_i \frac{M_i}{r_i^2} \hat{r}_i. \quad (11)$$

Here  $b$  is the bias between the mass and galaxy distribution,  $M_i$  is the mass of galaxy  $i$  and  $r_i$  its distance from the observer,  $H_0$  is the Hubble parameter at  $z = 0$  and  $f(\Omega_m) \sim \Omega_m^{0.55}$ , is the derivative of the growth factor with respect to the natural logarithm of the redshift. The above expression can be rewritten as

$$g = \frac{G}{b} \sum_i L_i \frac{M_i}{L_i} \frac{\hat{r}_i}{r_i^2}, \quad (12)$$

where  $L_i$  is the apparent luminosity of galaxy  $i$ . Under the assumption of a universal mass-to-light ratio this writes

$$g = \frac{4\pi G}{b} \frac{M}{L} \sum_i \frac{L_i}{4\pi r_i^2} \hat{r}_i = \frac{4\pi G}{b} \frac{M}{L} \sum_i S_i \hat{r}_i, \quad (13)$$

where we have used the flux–luminosity relation  $S_i = L_i/(4\pi r_i^2)$ . The universal mass-to-light ratio can be evaluated for a given survey as (Peebles 1993)

$$\frac{M}{L} = \frac{3\Omega_m}{8\pi G} \frac{H_0^2}{j}, \quad (14)$$

where the luminosity density  $j$  is evaluated from the luminosity function  $\Phi(L)$  of galaxies in a particular wavelength band using

$$j = \int_0^\infty L \Phi(L) dL \quad (15)$$

However, most flux-limited catalogues have a lower flux cut-off which too needs to be taken into account. For the WISE catalogue the  $2.4\mu$  luminosity density has been evaluated to be  $j_{\text{wise}, 2.4\mu} = 3.8 \times 10^8 L_{2.4\mu\odot} \text{Mpc}^{-3}$ , where the Solar luminosity  $L_{2.4\mu\odot} = 3.34 \times 10^{-8} \text{ JyMpc}^2$  (Lake et al. 2017). However, this was obtained by selective spectroscopy of a small subset of galaxies with median redshift  $z \sim 0.35$ . We correct this by a factor of  $(1.35/1.14)^{0.8}$  in order to adapt the measurement to the median WISE redshift.

Putting all this together, we have

$$v = \frac{f(\Omega_m)}{b} \frac{H_0}{j} \sum_i S_i \hat{r}_i. \quad (16)$$

We evaluate  $\sum_i S_i \hat{r}_i$  for different subsamples of the AllWISE galaxy selection and the corresponding induced velocity of the Local Group, given in Table 3. While the final galaxy selection induces a velocity of just  $\sim 50\text{ km s}^{-1}$  on the Local Group, only lower limits can be inferred for the contributions of nearby subsamples to this velocity as these are dominated by extended sources for which WISE photometry is significantly underestimated (Wright et al. 2010). These velocities serve as consistency checks for the analysis. However, they do not affect the conclusions drawn in the previous or following sections.

## 8 RESULTS AND DISCUSSION

The total observed dipole of the final sample of 1 233 920 galaxies, with star contamination less than 0.1 percent and sources at low redshift suppressed as much as possible with WISE photometry, is found to be  $\mathcal{D} = 0.0124$  in the direction  $\text{RA} = 166.2^\circ$ ,  $\text{Dec.} = -15.7^\circ$  ( $l = 269.17^\circ$ ,  $b = 40.17^\circ$ ), which is just  $8.8^\circ$  away from the CMB dipole. Of the total 1 323 200 objects that would have remained had those with an apparent motion above  $400\text{ mas yr}^{-1}$  not been removed, 88 121 are found to have a signal-to-noise ratio  $> 1$ . Discarding only these objects, we obtain a catalogue of 1 235 079 objects, with  $\mathcal{D} = 0.0123$  in the direction  $\text{RA} = 167.5^\circ$ ,  $\text{Dec.} = -16.3^\circ$  ( $l = 270.4^\circ$ ,  $b = 40.1^\circ$ ), i.e.  $9.3^\circ$  away from the

**Table 3.** Induced velocity from the clustering dipole in different subsamples. When dominated by extended sources, the WISE photometry significantly underestimates the flux and consequently only a lower limit is provided.

Catalogue	N	RA <sup>°</sup>	Dec. <sup>°</sup>	∠ CMB	$v$ (km s <sup>-1</sup> )
AllWISE - 2MRS ( $z < 0.03$ ) $\Delta b = \pm 5^\circ$	24 648	159.7°	6.9°	16.2°	>116
AllWISE (III) + sources with ext_flag > 0 2MRS sources removed; $z < 0.03$	541 840	176.9°	-79.0°	72.1°	>85
AllWISE (IV) $\Delta b = \pm 15^\circ$ $\Delta \text{SGb} = \pm 5^\circ$ 2MRS sources removed; $z < 0.03$	1 358 233	167.9°	-40.9°	33.9°	~39

CMB dipole. If the 32 576 sources for which the motion fit failed to converge are also discarded in addition to those with an apparent motion above 400 mas yr<sup>-1</sup>, then the direction moves to RA = 156.4°, Dec. = -5.4° ( $l = 250.1^\circ$ ,  $b = 42.0^\circ$ ), i.e. 11.6° away from the CMB dipole, with  $\mathcal{D} = 0.0132$ . Thus both the magnitude and the direction of the final observed dipole are reasonably robust with respect to the details of the apparent motion measurement, which is an essential ingredient of the process of suppressing the star contamination. They are also robust with respect to the removal of individual local superclusters as described in Section 6.3.

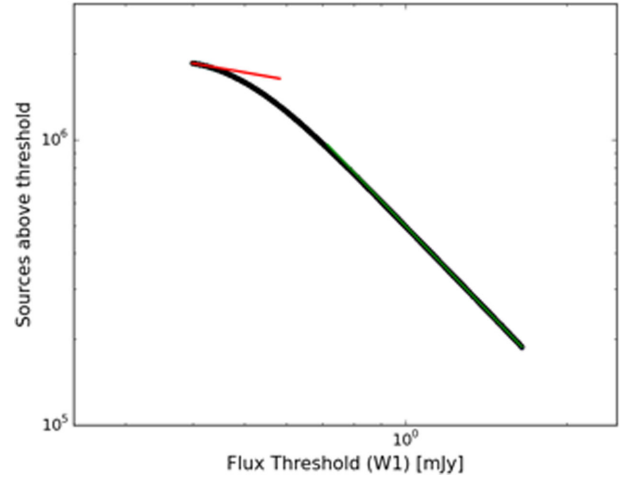
Subtracting the best estimate of the residual clustering dipole,  $|\mathcal{D}_{\text{cls}}| = 0.0076 \pm 0.0022$  (Section 7.1) from the total observed dipole  $|\mathcal{D}| = 0.0124$  (Section 6.2), we obtain  $|\mathcal{D} - \mathcal{D}_{\text{cls}}| = 0.0048 \pm 0.0022$ . A catalogue of this size (1.2 million sources) is also expected to have a *random* dipole of size  $\sim 0.001$ , implying  $|\mathcal{D}_{\text{kin}}| = 0.0048 \pm 0.0024$  if we do a scalar subtraction. While this subtraction ought to be done vectorially, the precise direction of the structure dipole in the local Universe is unknown. However the close alignment of the total dipole observed in data with the CMB dipole, despite  $|\mathcal{D}_{\text{cls}}|$  being significantly larger than  $|\mathcal{D}_{\text{kin}}|$ , suggests that the two are closely parallel. Hence the vector subtraction can be approximated with a scalar subtraction of the magnitudes.

It is straightforward to evaluate the flux power-law index  $x$  in equation (2) for a given catalogue in a single-frequency band. However for WISE and AllWISE, the initial cuts and the cuts applied for star–galaxy separation depend on magnitudes in different bands, hence the index changes between the different bands (Griffith et al. 2015). Since our galaxy selection is driven primarily by a W1 magnitude cut, we confine ourselves to the W1 band.

The index of the flux function can be fitted from the data (Colin et al. 2017). The Doppler shift is more important for faint galaxies, hence the value of  $x$  near the threshold is most relevant and is found to be 0.75 as shown in Fig. 9. The spectral index  $\alpha$  (3) for galaxies in infrared depends on the classification of the galaxy. However, in the W1 band range, for most galaxy types, the spectral index varies between 0.8 and 1.0 (Griffith et al. 2015). Using a median spectral index of 0.9 and  $x = 0.75$  in equation (1) yields a velocity of  $420 \pm 213$  km s<sup>-1</sup> for the Solar system barycentre. The uncertainties include the statistical error in the clustering dipole, the error in the spectral index, and the shot noise due to the random dipole from a finite sample, all added in quadrature.

## 9 SUMMARY

The total observed dipole in the final AllWISE galaxy selection after suppressing star contamination and local source contribution

**Figure 9.** Variation of the AllWISE galaxy selection source count with the lower cut in flux. At the lower flux threshold, the best-fitting value (red line) of the power-law exponent  $x$  in equation (2) is 0.75.

is 0.0124, corresponding to a velocity of 1110 km s<sup>-1</sup> if interpreted as solely kinematic in origin.

While this seems anomalously high, theoretical expectations for a  $\Lambda$ CDM universe suggests that a clustering dipole of  $\sim 0.006$  is expected in a sample with the same redshift distribution as our final selection. This does not however account for our special local environment. To do so, we examine mock AllWISE galaxy selection-like catalogues generated from a  $\Lambda$ CDM Hubble volume simulation. We search for haloes with velocities similar to that of the Milky Way embedded in an environment as observed in 2MRS with a bulk velocity of  $\sim 240$  km s<sup>-1</sup> extending beyond  $z = 0.03$ . We find that an intrinsic clustering dipole of size  $0.0071 \pm 0.0022$  can arise for these observers. This lowers the inferred velocity of the Solar system barycentre to  $430 \pm 213$  km s<sup>-1</sup>, which is compatible with the value inferred from the CMB dipole. However, the estimate of the residual clustering dipole from theory is model dependent (here, a  $\Lambda$ CDM model with parameters fitted to *Planck* data) consequently the final value of the velocity *cannot* be considered an independent measurement and serves only as a consistency test.

The structures in the redshift range 0.03–0.3 which give rise to such a large clustering dipole remain to be identified. In any case it is evident that we are *not* typical observers – the observed local velocity field is expected for only 2 per cent of Milky Way-like observers in the standard  $\Lambda$ CDM cosmology and this fraction drops rapidly as the local bulk velocity and its extent increases. The Six Degree Field Galaxy Redshift Survey (6dFGRS) extending out to  $z \sim 0.05$  provides the most uniform galaxy peculiar velocity

sample to date. If we adopt their estimate of the bulk velocity  $397 \pm 68 \text{ km s}^{-1}$  (Springob et al. 2014; Magoulas et al. 2016), then following the procedure in Section 7.1 this fraction is found to be only 0.14 per cent for the median velocity, and 0.8 per cent for the  $3\sigma$  lower bound of  $193 \text{ km s}^{-1}$ . This provides a fresh perspective on interpretations drawn from analysis of cosmological data which assume that the observations are being made by a *typical* observer.

Note added: While this manuscript was being revised, an eprint by Hellwing et al. (2018) appeared – they too find that Local Group-like observers are rather untypical and explore the consequences for the interpretation of future cosmological data.

## ACKNOWLEDGEMENTS

MR acknowledges hospitality and support from the Institut d’Astrophysique Paris and SS acknowledges a DNRf Niels Bohr Professorship. We thank the anonymous referee for critical comments which improved the manuscript.

## REFERENCES

- Alonso D., Salvador A. I., Sánchez F. J., Bilicki M., Garcia-Bellido J., Sánchez E., 2015, *MNRAS*, 449, 670
- Appleby S., Shafieloo A., 2014, *JCAP*, 10, 070
- Bengaly C. A. P., Bernui A., Alcaniz J. S., Xavier H. S., Novaes C. P., 2017, *MNRAS*, 464, 768
- Bengaly C. A. P., Novaes C. P., Xavier H. S., Bilicki M., Bernui A., Alcaniz J. S., 2018, *MNRAS*, 475, L106
- Bilicki M., Chodorowski M., Jarrett T., Mamon G. A., 2011, *ApJ*, 741, 31
- Blake C., Wall J., 2002, *Nature*, 416, 150
- Burles S., Rappaport S., 2006, *ApJ*, 641, L1
- Cautun K., Frenk C. S., van de Weygaert R., Hellwing W. A., Jones B. J. T., 2014, *MNRAS*, 445
- Challinor A., van Leeuwen F., 2002, *Phys. Rev. D*, 65, 103001
- Colin J., Mohayaee R., Sarkar S., Shafieloo A., 2011, *MNRAS*, 414, 264
- Colin J., Mohayaee R., Rameez M., Sarkar S., 2017, *MNRAS*, 471, 1045
- Crawford F., 2009, *ApJ*, 692, 887
- Cutri R. M. et al., 2012, Explanatory Supplement to the WISE All-Sky Data Release Products, Tech. rep.
- Ellis G. F. R., Baldwin J. E., 1984, *MNRAS*, 206, 377
- Erdoğdu P. et al., 2006, *MNRAS*, 368, 1515
- Feindt U. et al., 2013, *A&A*, 560, A90
- Gibelyou C., Huterer D., 2012, *MNRAS*, 427, 1994
- Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759
- Griffith R. L., Wright J. T., Maldonado J., Povich M. S., Sigurdsson S., Mullan B., 2015, *ApJS*, 217, 25
- Hellwing W. A., Bilicki M., Libeskind N. I., 2018, [arXiv:1802.03391](https://arxiv.org/abs/1802.03391) [astro-ph.CO].
- Hinshaw G. et al., 2009, *ApJS*, 180, 225
- Hogg D. W., Eisenstein D. J., Blanton M. R., Bahcall N. A., Brinkmann J., Gunn J. E., Schneider D. P., 2005, *ApJ*, 624, 54
- Huchra J. P. et al., 2012, *ApJS*, 199, 26
- Huterer D., Knox L., Nichol R. C., 2001, *ApJ*, 555, 547
- Itoh Y., Yahata K., Takada M., 2010, *Phys. Rev. D*, 82, 043530
- Kogut A. et al., 1993, *ApJ*, 419, 1
- Kovács A., Szapudi I., 2015, *MNRAS*, 448, 1305
- Kurcz A., Bilicki M., Solarz A., Krupa M., Pollo A., Malek K., 2016, *A&A*, 592, A25
- Lake S. E., Wright E. L., Assef R. J., Jarrett T. H., Petty S., Stanford S. A., Stern D., Tsai C.-W., 2017, *ApJ*, preprint ([arXiv:1702.07829](https://arxiv.org/abs/1702.07829))
- Lavaux G., Tully R. B., Mohayaee R., Colombi S., 2010, *ApJ*, 709, 483
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- Liske J. et al., 2015, *MNRAS*, 452, 2087
- Magoulas C., Springob C., Colless M., Mould J., Lucey J., Erdoğdu P., Jones D. H., 2016, *Proc. IAU Symp.* 308, 336
- Mainzer A. et al., 2011, *ApJ*, 731, 53
- Maller A. H., McIntosh D. H., Katz N., Weinberg M. D., 2003, *ApJ*, 598, L1
- Peebles P. J. E., Wilkinson D. T., 1968, *Phys. Rev. D*, 174, 2168
- Peebles P. J. E. 1980, *The Large-Scale Structure of the Universe*. Princeton Univ. Press, Princeton, NJ, 1980
- Peebles P. J. E. 1993, *Principles of Physical Cosmology*. Princeton Univ. Press, Princeton, NJ
- Planck Collaboration XXVII, 2016, *A&A*, 594, A16
- Planck Collaboration XVI, 2014, *A&A*, 571, A27
- Robitaille T. P. et al., 2013, *A&A*, 558, A33
- Rubart M., Schwarz D. J., 2013, *A&A*, 555, A117
- Scrimgeour M. I. et al., 2012, *MNRAS*, 425, 116
- SDSS Collaboration: Abazajian K. N., Adelman-McCarthy J. K., Agueros M. A., 2009, *ApJ*, 182, 543
- Singal A. K., 2011, *ApJ*, 742, L23
- Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., preprint ([arXiv:1407.2600](https://arxiv.org/abs/1407.2600))
- Skrutskie M. F. et al., 2006, *AJ*, 131, 1163
- Springob C., Magoulas C., Colless M., Mould J., Erdoğdu P., Jones D. H., Lucey J. R., Campbell L., Fluke C. J., 2014, *MNRAS*, 445, 2677
- Stewart J. M., Sciamia D. W., 1967, *Nature*, 216, 748
- Szapudi I., Prunet S., Pogonyan D., Szalay A. S., Bond J. R., 2001, *ApJ*, 548, L115
- Tegmark M. et al., 2002, *ApJ*, 571, 191
- Tiwari P., Jain P., 2015, *MNRAS*, 447, 2658
- Tiwari P., Kothari R., Naskar A., Nadkarni-Ghosh S., Jain P., 2015, *As-tropart. Phys.*, 61, 1
- Vieira K., Garcia-Varela A., Sabogal B., 2017, *MNRAS*, 469, 4175
- WMAP Collaboration et al., 2011, *ApJS*, 192, 17
- Wright E. L. et al., 2010, *AJ*, 140, 1868
- Yahil A., Walker D., Rowan-Robinson M., 1986, *ApJ*, 301, L1
- Yoon M., Huterer D., Gibelyou C., Kovács A., Szapudi I., 2014, *MNRAS*, 445, L60

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