BULK FLOW OF HALOS IN $\Lambda$CDM SIMULATION

MING LI$^{1}$, JUN PAN$^{1,2}$, LIANG GAO$^{2}$, YIPENG JING$^{3}$, XIAOHOU YANG$^{3}$, XUEBIN CHI$^{4}$, LONGLONG FENG$^{1}$, XI KANG$^{1,5}$, WEIPENG LIN$^{7}$, GUIHUA SHAN$^{4}$, LONG WANG$^{4}$, DONGHAI ZHAO$^{3}$, AND PENGJIE ZHANG$^{3}$

1 Purple Mountain Observatory, 2 West Beijing Road, Nanjing 210008, China; jpan@bao.ac.cn
2 National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, China
3 Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, 80 Nan Dun Road, Shanghai 200030, China
4 Supercomputing Center, Computer Network Information Center, Chinese Academy of Sciences, 4 Zhongguancun Nansijie, Haidian District, Beijing 100190, China
5 Partner Group of MPI for Astronomy, PMO, 2 West Beijing Road, Nanjing 210008, China

Received 2012 July 23; accepted 2012 October 29; published 2012 December 4

ABSTRACT

Analysis of the Pangu $N$-body simulation validates that the bulk flow of halos follows a Maxwellian distribution with variance that is consistent with the prediction of the linear theory of structure formation. We propose that the consistency between the observed bulk velocity and theories should be examined at the effective scale of the radius of a spherical top-hat window function yielding the same smoothed velocity variance in linear theory as the sample window function does. We compared some recently estimated bulk flows from observational samples with the prediction of the $\Lambda$CDM model we used; some results deviate from expectation at a level of $\sim 3\sigma$, but the discrepancy is not as severe as previously claimed. We show that bulk flow is only weakly correlated with the dipole of the internal mass distribution, that the alignment angle between the mass dipole and the bulk flow has a broad distribution peaked at $\sim 30^\circ$–$50^\circ$, and also that the bulk flow shows little dependence on the mass of the halos used in the estimation. In a simulation of box size $1 h^{-1}$ Gpc, for a cell of radius $100 h^{-1}$ Mpc the maximal bulk velocity is $>500 \text{ km s}^{-1}$; dipoles of the environmental mass outside the cell are not tightly aligned with the bulk flow, but are rather located randomly around it with separation angles $\sim 20^\circ$–$40^\circ$. In the fastest cell there is a slightly smaller number of low-mass halos; however, halos inside are clustered more strongly at scales $\gtrsim 20 h^{-1}$ Mpc, which might be a significant feature since the correlation between bulk flow and halo clustering actually increases in significance beyond such scales.

Key words: galaxies: halos – large-scale structure of universe – methods: statistical

Online-only material: color figures

1. INTRODUCTION

Bulk flow refers to the apparent coherent peculiar motion of galaxies and galaxy clusters in a considerably large volume around us. In practice, there are several ways to estimate bulk flows from various observation resources, such as galaxy catalogs from peculiar velocity surveys (e.g., Feldman et al. 2010), compiled Type Ia supernovae data (e.g., Dai et al. 2011), and galaxy clusters in combination with cosmic microwave background (CMB) observations (e.g., Kashlinsky et al. 2010). Recently some interesting new methods based on galaxy two-point correlation functions (Song et al. 2011) and galaxy light (Nusser et al. 2011; Abate & Feldman 2012) have also emerged.

Analysis of the spiral galaxy catalog of the SFI++ survey (Springob et al. 2007) shows that within a top-hat spherical window of radius $40 h^{-1}$ Mpc the velocity of the bulk flow is $338 \pm 38 \text{ km s}^{-1}$ toward the Galactic plane $(l, b) = (276^\circ, 14^\circ)$ with a 3$\sigma$ 1$\sigma$ uncertainty, and $257 \pm 44 \text{ km s}^{-1}$ toward $(l, b) = (279^\circ, 10^\circ)$ with a 6$\sigma$ error within a window of radius $100 h^{-1}$ Mpc (Nusser & Davis 2011). These measurements are in agreement with the analysis by Sandage et al. (2010) of data consisting of supernovae, selected nearby galaxies, and galaxy clusters. Feldman et al. (2010) constructed a composite catalog of galaxies with peculiar velocities measured in different surveys, including the SFI++. They estimate that the bulk flow within a Gaussian window of $50 h^{-1}$ Mpc is $416 \pm 78 \text{ km s}^{-1}$ in the direction $(l, b) = (282^\circ \pm 11^\circ, 6^\circ \pm 6^\circ)$ (see also Watkins et al. 2009).

Employing the peculiar velocities of supernovae is another viable route to detect bulk flow, though such samples are usually very sparse and prone to Malmquist bias. Dai et al. (2011) fitted a bulk flow of $188_{-103}^{+119} \text{ km s}^{-1}$ in the direction $(l, b) = (290^\circ_{-31}^{+30}, 20^\circ_{-32}^{+32})$ to the Union2 supernova catalog (Amanullah et al. 2010) for redshifts $z < 0.05$, but no significant bulk flow was detected from data at $z > 0.05$. Colin et al. (2011) used the same data to obtain a similar estimate but with a higher median amplitude of $250_{-80}^{+190} \text{ km s}^{-1}$. However, using a different supernova data set within the redshift shell $z = (0.0043, 0.028)$, Weyant et al. (2011) estimate that the local flow is $538 \pm 86 \text{ km s}^{-1}$ pointing to $(l, b) = (258^\circ \pm 10^\circ, 36^\circ \pm 11^\circ)$, or $446 \pm 101 \text{ km s}^{-1}$ toward $(l, b) = (273^\circ \pm 11^\circ, 46^\circ \pm 8^\circ)$ if a different technique is employed; this is in agreement with the dipole of the CMB $(l, b) = (263^\circ_{-9}^{+9}, 48^\circ_{-26}^{+26})$ (Jarošík et al. 2011). Jha et al. (2007) and Haugbolle et al. (2007) found the same values with similar uncertainties.

The availability of recent galaxy peculiar velocity data is limited to our local universe; the bulk flow at higher redshift, sometimes dubbed dark flow, is mainly explored through the kinetic Sunyaev–Zel’dovich (kSZ) effect of galaxy clusters (Sunyaev & Zeldovich 1980). Kashlinsky et al. (2011) computed kSZ signals of 771 X-ray clusters in the 7 yr Wilkinson Microwave Anisotropy Probe (WMAP) CMB map and conclude that the flow at $z \lesssim 0.16$ is directed to $(l, b) = (278^\circ \pm 18^\circ, 2^\circ \pm 15^\circ)$ and to $(l, b) = (283^\circ \pm 19^\circ, 20^\circ \pm 15^\circ)$ if $z \lesssim 0.25$. They further argue that the flow at these depths should reach a magnitude of $\sim 1000 \text{ km s}^{-1}$ according to an earlier investigation by Kashlinsky et al. (2010). Osborne et al. (2011) derived a conflicting assertion from the same CMB map in conjunction with 736 ROSAT-observed clusters: that there is no significant detection of kSZ effects at low multipoles, basically denying...
the existence of bulk flow. In some cases, however, the thermal Sunyaev–Zel’dovich effect might induce a dipole that could easily be misunderstood as bulk flow of \( \sim 2000–4000\, \text{km}\,\text{s}^{-1} \).

Bulk flow is a topic of long-term interest to observational cosmology (see Strauss & Willick 1995 for a review of early works), and special surveys have been dedicated to it (e.g., Courtois et al. 2011). However, as we see, no consensus on the amplitudes, directions, or convergence depth of bulk flows has yet been achieved to reconcile the different measurements. Nonetheless, some authors have argued that the amplitude of their measured bulk flow is too strong over such large scales, presenting a challenge to the standard \( \Lambda \) CDM model, or at least that of the 5 yr WMAP parameters (Watkins et al. 2009; Kashlinsky et al. 2010; Feldman et al. 2010; Macaulay et al. 2011). Not surprisingly, cosmological models of different flavors have been constructed to explain such anomalies (e.g., Mersini-Houghton & Holman 2009; Afshordi et al. 2009; Wyman & Khoury 2010), but new analysis of similar data sets seems to have nullified support for such a violation (Turnbull et al. 2012; Ma & Scott 2012).

Expectations concerning bulk flow in the \( \Lambda \) CDM universe are generally calculated using linear perturbation theory of sufficiently large-scale structure that its accuracy is believed to be good enough. Often the one-dimensional (1D) velocity variance of dark matter is quoted to compare with measurements; however, we need to note here that it is the rms velocity that should be used rather than the 1D rms velocity. Mak et al. (2011) calculated using linear theory that the rms bulk velocity is typically \( \sim 300\, \text{km}\,\text{s}^{-1} \) and claimed that the uncertainty at 95% confidence due to sample variance is approximately 200 km s\(^{-1}\) for a top-hat window of radius 60 h\(^{-1}\) Mpc. A concern is that nonlinearity might not be negligible even at very large scales (e.g., Scoccimarro 2004) and could act as a systematic bias in any conclusions about the consistency between model and data.

Although linear theory can predict the possibility of observing a bulk flow of a particular amplitude at a certain scale, several key problems cannot yet be easily tackled analytically, e.g., internal properties of the volume demonstrating large bulk flow. In practice, halo catalogs from N-body simulations in a large box with sufficient mass resolution are best suited for such a task. The reason for focusing on halos instead of dark matter is that observational objects used to determine bulk flow are galaxies and galaxy clusters that reside in halos, and in practice the strongly nonlinearity of peculiar velocities of galaxies is largely filtered out so that it is mainly the motion following their host halos that contributes to the bulk flow (e.g., Watkins et al. 2009). In fact, Bahcall et al. (1994) and Moscardini et al. (1996) have analyzed mock halo catalogs and obtained useful results, but their simulations are either at very low mass resolution or based on a compromised simulation method. In this paper we will demonstrate our analysis of the velocity field of halos resolved from a dark-matter-only \( \Lambda \) CDM simulation in a 1 h\(^{-1}\) Mpc box with 3072\(^3\) particles. The large volume and high mass resolution of our simulation allow us to investigate halo behaviors in detail over broad dynamic ranges superseding those of previous works.

In Section 2 we introduce the definition and basic theoretical predictions of bulk flow. Section 3 presents measurements of the bulk flow of simulated randomly placed cells. Section 4 is devoted to the analysis of particular regions showing extraordinarily large bulk flow velocities. The final section contains our summary and discussion.

## 2. Bulk Flow in Linear Theory

Placing a window of characteristic scale \( R \) randomly in the sample space, if \( N \) objects (galaxies, galaxy clusters, or halos—just the latter in this work) are enclosed, the bulk flow of the particular volume indicated by the particular object is

\[
V = \sum_{i=1}^{N} \frac{w_i v_i}{\sum w_i},
\]

in which \( v_i \) is the peculiar velocity of the \( i \)-th object and \( w_i \) is the weight assigned to it. In practice the weights could originate from the radial selection function, the angular selection function (survey mask), luminosity, mass, etc.

### 2.1. Sample Window

The effect of the sample window is purely geometrical and can be easily modeled. In the continuous limit, Equation (1) becomes

\[
\vec{V} = \vec{v} \otimes W(R) = \int \vec{v}(\vec{r}) W(\vec{r}; R) d\vec{r} \times \frac{1}{(2\pi)^3} \int \vec{v}(\vec{k}) \hat{W} d^3k,
\]

where \( W(\vec{r}; R) \) is the window function of characteristic scale \( R \) evaluated at position vector \( \vec{r} \) and \( \hat{W} \) is its Fourier transformation. In principle, \( W \) could be anisotropic, e.g., due to incomplete sky coverage and non-uniform depth. The simplest and most common forms are the spherical top-hat window \( W_{\text{th}} = 3(\sin k R - k R \cos k R)/(k R)^3 \) and the Gaussian window \( W_G = \exp(-k^2 R^2/2) \). In fact, there are few differences between top-hat, Gaussian, and anisotropic windows for bulk flow statistics if the effective scale has been taken care of. Sometimes bulk flow estimates are provided by objects within a spherical shell defined by two radii \( R_1 > R_0 \). It is easy to see that Equation (2) applies with the window function \( W_S = (R_1^3 W_{\text{th}} - R_0^3 W_{\text{th}})/(R_1^3 - R_0^3) \).

The probability distribution function (PDF) of \( V \) can be expressed as

\[
p(V) dV = p(V) dV P(\hat{V}) d\hat{V},
\]

where \( \hat{V} \) denotes the unit vector in the direction of the bulk velocity. The assumption of isotropy leads to \( p(\hat{V}) = 1/4\pi \) and \( \langle \hat{V} \hat{V} \rangle = 0 \), which ensures that \( V = \int V p(V) dV \int \hat{V} p(\hat{V}) d\hat{V} = \langle V \rangle \hat{V} \). Integration over the angular part of Equation (3) yields the PDF of the amplitude of the bulk flow:

\[
p(V) dV = \left( \int p(V) d\hat{V} \right) V^2 dV = 4\pi p(V) V^2 dV.
\]

\( V \) is by definition the velocity field smoothed by the window function. Once the smoothing scale is sufficiently large, the distribution of \( V \) becomes very close to Gaussian so that a Maxwellian distribution could be invoked to model \( p(V) \) (Bahcall et al. 1994):

\[
p(V) dV = \sqrt{\frac{2}{\pi}} \left( \frac{3}{\sigma V^2} \right)^{3/2} V^2 \exp\left( -\frac{3V^2}{2\sigma V^2} \right) dV,
\]

where the variance of \( V \) can be obtained from \( P_{\text{vv}} \), the power spectrum of \( v \), through

\[
\sigma_V^2 = \frac{1}{(2\pi)^3} \int P_{\text{vv}} \hat{W}^2 d^3k.
\]
Given that the radius of the window function \( W(R) \) deployed to measure the bulk flow is fairly large, one would expect that the velocity field smoothed at such a scale will be well described by linear evolution of the initial condition. Then if the initial distribution of velocity is Gaussian, for instance in the case of our simulation, Equation (5) will be a good approximation to the PDF of the amplitudes of bulk flows. With this model the most likely amplitude of bulk flow is simply \( V_p = \sqrt{2/3} \sigma_V ; \) the ranges of variances corresponding to different levels can be computed from the integral

\[
\int p(V) dV = \text{erf} \left( \frac{\sqrt{3V^2}}{2\sigma_V} \right) - \sqrt{\frac{6V^2}{\pi \sigma_0^2}} \exp \left( -\frac{3V^2}{2\sigma_0^2} \right). \tag{7}
\]

If \( \mathbf{v} \) is curl free, a velocity potential field can be defined as \( \theta(\mathbf{r}) = -\nabla \cdot \mathbf{v}/(H a f) \) with the scale factor \( a = 1/(1+z) \), \( f \equiv d \log D(a)/d \log a \approx \Omega_m^{1/7} + (1 + \Omega_m/2) \Omega_L/70 \), and \( D(a) \) is the linear density growth factor at redshift \( z = 1/a-1 \) (Lahav et al. 1991), we have

\[
\sigma^2 = \frac{(H a f)^2}{(2\pi)^3} \int P_{kk} \bar{W}^2 d^3 k. \tag{8}
\]

This expression relies on the assumption of negligible rotational velocity. In the linear regime, if the biasing of halo velocity to dark matter velocity is unity, it can be further simplified with the approximation \( P_{kk} \approx \delta^4 \), where \( \delta^4 \) is the linear matter power spectrum.

### 2.2. Selection Function

In reality, a large fraction of galaxy samples are magnitude-limited (or flux-limited); galaxies fainter than a certain threshold are missed in the sample, so that the number density of observed galaxies \( n(r) \) as a function of distance to the observer, termed the radial selection function, is not constant. In the presence of the selection function, if no correction is made, the measured bulk flow local to an observer is

\[
\mathbf{V} = \int \mathbf{v}(\mathbf{r}) n(r) W(\mathbf{r}; R) d^3 r / \int n(r) W(\mathbf{r}; R) d^3 r , \tag{9}
\]

in which \( n(r) \) acts as the weighting function. In theoretical modeling, Equation (9) is equivalent to Equation (2) if a new window function is defined through \( W_0 = n(r) W / \int n W d^3 r \); however, there is still the conceptual difference of applying a selection function rather than a purely geometrical window function. A non-constant selection function reflects the fact that the sampling to the velocity field depends on distance; a window function not of top-hat type rather simply denotes that in the estimation the velocity field is weighted by a particular scheme, but the sampling to the field is fair.

The simplest proposal to correct the unfair sampling rate depicted by the selection function is to divide the measured peculiar velocity of an object by the selection function:

\[
\hat{\mathbf{V}} = \frac{\sum_i^N u_i \mathbf{v}_i/n(r_i)}{\sum_i^N u_i/n(r_i)}. \tag{10}
\]

We will check the performance of this equation numerically in Section 3.3.

Note that the discussion here is also applicable to the angular selection function, termed the completeness mask, defined as the ratio of the number of observed objects to the local observable number of that type of object.

---

6 A better approximation to \( f \) can be found in Linder (2005).

## 2.3. Physical Weights

A type of weight different from selection functions arises from the physical properties of astronomical objects such as mass, luminosity, internal velocity dispersion, etc. This kind of weight cannot be assimilated into the sample geometrical window function in theoretical work; rather, we have to develop statistical models to account for the effects of these weighting schemes, which often involve calculation of a series of correlations between the peculiar velocity and the object’s physical quantities. Among the various physical weights, probably the most commonly seen is mass. Mass may not always be the dominant factor determining the properties of galaxies and clusters, but it is always a major factor. For example, the luminosity–mass relation for galaxies in some bands is considerably tight; here weighting by luminosity could be deemed to be roughly equivalent to weighting by a certain power of the mass. We will return to the issue of bulk flow weighted by mass in Section 3.4 with a demonstrative numerical analysis.

### 3. BULK FLOW IN THE PANGU SIMULATION

#### 3.1. The Pangu Simulation and Its Halo Catalog

The Pangu simulation (PS-I) is a large-volume, high-resolution simulation carried out under the scheme of the Computational Cosmology Consortium of China (dubbed C4). PS-I assumes a \( \Lambda \) cold dark matter (ACDM) cosmology model with parameters

\[
\Omega_m = 0.26, \quad \Omega_b = 0.044, \quad \Omega_\Lambda = 0.74, \quad h = 0.71, \quad \sigma_8 = 0.8, \quad n_s = 1.0.
\]

The simulation contains dark matter only and uses \( N = 3072^3 \) particles to follow the distribution and evolution of dark matter within a periodic box with \( L = 1000 h^{-1} \) Mpc on a side. Each particle has a mass of \( 2.48915 \times 10^9 h^{-1} M_\odot \). The Plummer-equivalent force softening length is kept constant at \( 7 h^{-1} \) kpc. The PS-I starts from redshift \( z_{\text{init}} = 127 \) with the initial positions and velocities of particles generated using the Zel’dovich approximation from a glass-like particle set. The input linear power spectrum is computed with CAMB (Lewis et al. 2000). The simulation is then run with L-GADGET, a memory-optimized version of GADGET2 (Springel 2005). L-GADGET2 is designed to meet the requirements of high-performance computations; only the tree-particle mesh algorithm is included to calculate the gravitational forces efficiently. In total, 64 snapshots are saved from \( z_{\text{init}} = 127 \) to redshift \( z = 0 \). The PS-I is performed on the supercomputer Lenovo Deepcomp7000 at the supercomputing Center of Chinese Academy of Sciences. We use 2048 cores and about 6 TB memory at the peak time. The simulation consumes approximately \( 6.5 \times 10^8 \) CPU hr (about 13 days) in total and consists of 6151 time steps.

Dark matter halos are identified on the fly during the simulation for each snapshot, using the standard friends-of-friends (FOF) algorithm with linking length of 0.2 times the mean particle separation. Each FOF group must contain at least 20 particles; at redshift \( z = 0 \) there are \( 4.2 \times 10^4 \) identified particle groups. There is some ambiguity in the definition of halo mass. The definition of halo mass used most in the literature is the mass enclosed by a sphere centered on the halo center with a certain radius, within which the average density is some factor larger than the critical density. A more convenient definition of halo mass is the total mass of all member particles of the FOF group;
Figure 1. Left panel: $p(V)$ measured in spherical top-hat windows with different radius. Solid lines are measurements from N-body simulation, from left to right with decreasing heights corresponding to the radius of window function $R = 100$, 50, and $25\,h^{-1}\text{Mpc}$, respectively. Dashed lines are prediction of the model of Equations (5) and (8) with $P_{\delta\delta} = P_{\delta\delta}^{(L)}$. Middle panel: $p(V)$ measured in one spherical shell window defined by two radii $(R_0, R_1) = (25, 100)\,h^{-1}\text{Mpc}$. Right panel: comparison of $\sigma_V$ predicted by the linear theory (the solid line is for a spherical top-hat window, and the dashed line is for a spherical shell window with $R_0 = R$, $R_1 = 100\,h^{-1}\text{Mpc}$) with the estimation from the simulation (blue crosses) and the $\sigma_V$ that provide the best fit to the PDFs of bulk flow in a simulation using Equation (5) (red squares).

(A color version of this figure is available in the online journal.)

we use this definition. For simplicity, the average of the velocities of the member particles is taken as the velocity of the halo.

In this paper we use a subset of the group catalog as our full halo catalog; this consists of FOF groups of mass larger than $2.5 \times 10^{11}\,M_\odot$ only and provides a total of $1.28 \times 10^7$ entries. As selected FOF groups contain at least 100 member particles, the corresponding Poisson fluctuation is greater than 10, the threshold used for FOF group identification, limiting the discreteness error to the 10% level. Another reason for the mass criterion is that it is not easy to detect a large number of low-mass halos in observations. Our prudence test verifies that including FOF groups with number of particles $<100$ does not affect our final results significantly, even though effects of nonlinearity are stronger. More importantly, we will show that the statistics of bulk flows are not sensitive to the mass of halos in the sample (Section 3.5).

3.2. The Probability Distribution Function of Bulk Flow

To ensure fair sampling to the simulation, the oversampling algorithm of Szapudi (1998) is implemented to generate $\sim 10^9$ cells for $R$ within $25$--$100\,h^{-1}\text{Mpc}$, and bulk flows of halos in random cells are estimated using Equation (1). In our measurements we adopt mainly the spherical top-hat window, i.e., $w = 1$ for all halos inside the window and $w = 0$ otherwise; the shell window function is also deployed as a consistency check. Adopting a top-hat window function simplifies the computation: in principle one could try Gaussian or other more sophisticated window functions; these will not change the main results. Since our simulation box is limited to a cubic box of side length $1\,h^{-1}\text{Gpc}$, probing bulk flows in cells of radius $>100\,h^{-1}\text{Mpc}$ would be statistically unreliable.

Our measurements of bulk flows are displayed in Figure 1. For a top-hat window, $\sigma_V$ decreases with cell radius: the possibility of finding extremely large bulk flow speeds becomes smaller as the volume increases. For a shell window defined by two radii $R_0 < R_1$, if $R_0$ is not very close to $R_1$, then $\sigma_V(R_0, R_1) \sim \sigma_V(R_1)$.

More importantly, the PDFs of the amplitude of simulated bulk flow, no matter whether a spherical top-hat window or a spherical shell window is used, are all well described by the linear model. The agreement between the simulation and the model is better at larger volumes, as expected; the linear theory slightly overpredicts $\sigma_V$ since the nonlinear $P_{\delta\delta}$ is already lower than $P_{\delta\delta}^{(L)}$ at $k \sim 0.1\,h^{-1}\text{Mpc}^{-1}$ by $\sim 20\%$ (Nusser et al. 1991; Ciecielg & Chodorowski 2004; Scoccimarro 2004). The comparison clearly leads to the conclusion that, to a high precision, $p(V)$ obeys a Maxwellian distribution that is completely determined by $\sigma_V$: what really matters is not the exact shape of the window function but rather the corresponding $\sigma_V$. This provides solid ground for us to put different kinds of measurements together for comparison.

3.3. The Simple Correction for Selection Function

In this section we use a numerical approach to assess the effectiveness of the simple correction method of Equation (10) for the selection function. The sample window function is a spherical top-hat of radius $R = 100\,h^{-1}\text{Mpc}$, and the selection function is set to be in the form of the PSCz catalog (Saunders et al. 2000):

$$n(r) = \begin{cases} \frac{1}{n_\ast} \left( \frac{r}{r_\ast} \right)^{1-c} \left[ 1 + \left( \frac{r}{r_\ast} \right)^\gamma \right]^{-\beta/\gamma} & \text{if } r \leq 20\,h^{-1}\text{Mpc}, \\ 0.397 & \text{if } r > 20\,h^{-1}\text{Mpc}, \end{cases}$$

(11)

in which $c = 1.82$, $r_\ast = 86.4$, $\gamma = 1.56$, $\beta = 4.43$, and $n_\ast = 0.397$ to give the normalization $n(r = 20\,h^{-1}\text{Mpc}) = 1$. For the computation Monte Carlo simulation is applied to halos in each sampling cell to generate a mock catalog with radial distribution obeying Equation (11). Next, two kinds of bulk flows for each individual cell are estimated from the mock catalog: the bulk flow estimated directly using Equation (1) and the selection-function-corrected $\vec{V}$ using Equation (10). The two measurements are then compared with the results without a selection function.

For the bulk flow given by Equation (1), the selection function is not corrected at all: the resulting distribution of estimated bulk speed is a Maxwellian distribution function of variance $\sigma_V = 244\,\text{km}\,\text{s}^{-1}$, whereas the variance of the bulk flow without a selection function is $211\,\text{km}\,\text{s}^{-1}$. Apparently a selection function causes the sample to have reduced effective scale, inducing larger $V_p$ and $\sigma_V$. This seriously challenges the claim of Mak et al. (2011) that the selection function has little influence on the estimation of bulk flow.

The comparison of the selection-function-corrected estimation with the results without a selection function is displayed in Figure 2. It appears that the simple correction of Equation (10)
The kSZ effect will be proportional to

\[ \frac{\Delta T_{\text{SZ}}}{T} \propto \hat{V} \cdot \hat{n} \]

within

\[ \delta \]

differs little from

\[ p(V) \]

Figure 3. Differences between \( V \) and \( V_m \) in simulation. \( R \) is the radius of the spherical top-hat window function. Left: the relative difference between \( p(V_m) \) and \( p(V) \); the large fluctuation at large \( V \) is due to the almost zero values of the PDFs in the tail. Middle: PDFs of the amplitude difference between \( V \) and \( V_m \) in the same cell. Right: distribution functions of the angle \( \alpha = \cos^{-1}[V \cdot V_m/(VV m)] \) between \( V \) and \( V_m \) in the same cell.

(A color version of this figure is available in the online journal.)

can recover the bulk flow to a good extent. The PDF \( p(\hat{V}) \) differs little from \( p(V) \), which means that the variance of the smoothed velocity field is actually well recovered. For an individual cell, the deviation of \( \hat{V} \) from \( V \) is small: the amplitude difference shows no systematical bias and is mainly bounded within \( \sim 15 \text{ km s}^{-1} \), which seems not to vary much with the amplitude of \( V \). The shift in direction rarely goes beyond \( \sim 10^\circ \) and has a most likely value of about \( 1^\circ \), but the alignment turns out to be better for larger \( V \).

3.4. Bulk Flow as a Mass-weighted Average of Halo Velocities

It is known that attenuation to CMB temperature results from the kSZ effect \( \Delta T_{\text{SZ}} \propto \hat{V} \cdot \hat{n} \), in which \( \hat{n} \) is the unit vector of the line of sight and \( n_e \) is the density of free electrons in the galaxy cluster. If the aperture used to measure the kSZ effect is sufficiently large and it is taken for granted that the number of hot electrons in the cluster is proportional to the mass of the host halo \( m_h \), the total temperature fluctuation induced by the kSZ effect will be proportional to \( m_h \hat{v} \cdot \hat{l} \). Thus, the bulk flow estimated via the kSZ effect of galaxy clusters in fact is, in principle, the mass-weighted average of the halo velocities:

\[ V_m = \frac{\sum m_i v_i}{\sum m_i}. \tag{12} \]

\( V_m \) is the ratio of two Gaussian random variables, the total mass \( M = \sum m_i \), and the momentum \( P = \sum m_i v_i \). Using the results of Pham-Gia et al. (2006), it is possible to determine a linear theoretical model for \( p(V_m) \). The exact calculation requires knowledge of the power spectrum of matter, the momentum, and the correlation function between matter and momentum. However, a quick inspection provides a rough profile. In the continuous limit \( M = \sum m_i \) becomes \( \langle M \rangle [1 + \delta * W(R)] \), the smoothed density contrast \( \delta * W \ll 1 \) if \( R \) is large enough to enter the linear regime, and therefore \( V_m \sim P/\langle M \rangle \). It has been found that the variance of \( P/\langle M \rangle \) is dominated by \( \alpha^2 = (2\pi)^{-3} \int P_{vv} \tilde{W}^2 d^3 k \) if the smoothing scale is sufficiently large (Park & Park 2006), so we can expect that \( p(V_m) \sim p(V) \).

Our simulation data results indeed reveal that differences between \( p(V_m) \) and \( p(V) \) are small (Figure 3). But in an individual cell \( V_m \) does differ from \( V \), both in direction and in amplitude. As we can see in Figure 3, \( p(V_m - V) \) has a width of several tens of \( \text{km s}^{-1} \), which decreases as the cell volume increases. If the sample volume is small, it could appear that \( V_m \) deviates from \( V \) by \( \sim 100 \text{ km s}^{-1} \), though this possibility is tiny. We also note that \( V_m - V \) does not show any apparent trend with \( V \) or \( V_m \). The pointing of the mass-weighted bulk flow does not coincide with \( V \). The most likely angle between them is around \( 1^\circ - 3^\circ \) for a top-hat window of \( R \in (25, 100) \text{ h}^{-1} \text{ Mpc} \) and becomes smaller for larger volumes. Note that the distribution of the difference angle has a rather long tail: for instance, if \( R = 50 \text{ h}^{-1} \text{ Mpc} \), the probability of misalignment greater than \( 10^\circ \) is \( \sim 7.3\% \), not yet trivial.

Hitherto only the kSZ measurements can provide an estimation of the mass-weighted bulk flow. However, mass weighing
does not introduce significant statistical differences, as is actually supported by real observations (Lavaux et al. 2012); thus, hereafter we will just concentrate on the unweighted bulk flow.

### 3.5. Halo Mass Dependence

There is a possibility that the bulk flow may depend on the typical mass of the halo sample. The full halo catalog is divided into six subsamples by halo mass, and the measured $\sigma_V$ is plotted in Figure 4 as a function of the mean halo mass of the subsample. If the smoothing scale is large, it is obvious that there is little dependence of bulk flow on the mass of the sampled halos. For small-sized windows, e.g., $R = 25 h^{-1} \text{Mpc}$, the measured $\sigma_V$ of the low-mass subsample is slightly lower than that of the high-mass subsample, but this could just be statistical fluctuation.

For individual cells, the diversity in bulk flows measured from different halo mass bins might be non-trivial. If both $R$ and $V$ are not very large (Figure 4), Considering the fact that intrinsic properties of galaxies and galaxy clusters are more or less correlated with their host halo mass, in the case that the sample depth is shallow and the estimated bulk flow is of low amplitude, it would not be surprising to find it difficult to achieve tight convergence among different samples.

### 3.6. Consistency between Observation and Model

Systematical biases in dark flow, the bulk flow measured at high redshift, are not fully understood and precisely controlled, so we refrain from discussing the high-redshift case. Most of the local (or nearby) bulk flow measurements have redshift less than $z = 0$, the resulting redshift evolution of $\sigma_V$ with respect to $z = 0$ is of magnitude a few percent at most and can be comfortably ignored. Window functions used in different works are not all the same, but the excellent performance of the linear model provides a unified scheme. Since the PDFs of bulk flow are determined solely by $\sigma_V$, independent of the type of window, the radius of a top-hat window that gives the same linear amplitude, it would not be surprising to find it difficult to achieve the effective scale corresponding to a particular sample.

To check if an observed bulk flow is consistent with the $\Lambda$CDM model, we need to determine the variance range of $V$. The most likely amplitude of $V$ is $V_p = \sqrt{2/3} \sigma_V$, derived via $dP(V)/dV = 0$, and the variance range at different levels is computed through Equation (7). Given significance levels of 1σ, 2σ, and 3σ, with corresponding confidence probabilities $\epsilon_{1,2,3} = 68.3\%$, 95.5\%, and 99.7\%, respectively, we choose to define the variance range $\Delta V$ of $V$ around $V_p$ at a specified level through

$$P(|V - V_p| \geq \Delta V_{1,2,3}) \leq \epsilon_{1,2,3}. \quad (13)$$

Since by definition $V \equiv |V| \geq 0$ and the probability $\int_0^{2V_p} p(V)dV = 95.4\% < \epsilon_{2,3}$, to continue with Equation (13) the variance ranges at the 2σ and 3σ levels must be translated to $(0, V_p + \Delta V_{2,3})$ with $\int_{0}^{V_p + \Delta V_{2,3}} p(V)dV = \epsilon_{2,3}$, while the 1σ variance range is the usual one, $(V_p - \Delta V_1, V_p + \Delta V_1)$ with $\int_{V_p - \Delta V_1}^{V_p + \Delta V_1} p(V)dV = \epsilon_1$. Numerical computation with Equation (7) then shows that $\Delta V_1 = 0.38875\sigma_V$, $\Delta V_2 = 0.81904\sigma_V$, and $\Delta V_3 = 1.35577\sigma_V$.

Several recent measurements from observational data are overplotted on the model prediction in Figure 5. One must keep in mind that our calculation of effective scales has assumed that all other factors affecting the estimation, such as nonlinearity, the selection function, and sky coverage of incompleteness, have been perfectly corrected. The details of how these samples have been constructed are often too sketchy to render appropriate weights for our calculation. However, the imperfection actually reduces the effective volumes so that the true effective scales will be smaller, i.e., the data points will shift leftward along the horizontal axis in Figure 5. We shall therefore deem Figure 5 to be the most conservative judgment of the consistency between observation and theory. Nonetheless, from Figure 5, it appears that observation of our local universe does not rule out the $\Lambda$CDM model. The results of Haugboelle et al. (2007), Feldman et al. (2010), and Weyant et al. (2011), often quoted as supporting evidence disfavoring the standard $\Lambda$CDM model, are around the 3σ level, but the significance will be smaller if the error bars are taken into account. In addition, considering that many other measurements (including the report of Wang (2007) are in fact consistent with the $\Lambda$CDM model, we prefer to take a conservative standpoint on the issue.

### 3.7. Bulk Flow and Mass Distribution in the Cell

It is interesting to investigate the relation between bulk flow and the mass distribution in the sample volume. One might wonder if one could infer the bulk flow from the mass distribution if peculiar velocity data are absent, since in the linear regime Fourier modes of the velocity field can be derived from modes of the density field. However, from Equation (2) it is clear that the bulk flow is determined by those wavelength models
larger than the characteristic scale of the window function. If the Fourier transformation of the density field is restricted to the same volume in which the bulk flow is measured, those long-wavelength modes accounting for the bulk flow are missing. In fact, it has been clearly shown by Nusser & Davis (1994) that the bulk flow is completely immune to the internal mass distribution.

Our measurements confirm the expectations. The first quantity we checked is the mass monopole, the total mass \( \sum m_i \), or the total number \( N \) of halos in the volume, which is equivalent to the density fluctuation smoothed by the window function. Correlation coefficients are computed to denote the correlation strength between the amplitudes of the bulk flow and mass monopole (Table 1). Bulk flow is apparently not correlated with the total mass and the total number of halos in the volume at all.

Two kinds of dipoles of the halo distribution are measured, \( \langle r_i/\langle r \rangle \rangle \) and the mass-weighted one \( \langle \sum m_i r_i/\langle r \rangle \rangle / \sum m_i \). Our results show that using halo mass as weight to estimate mass dipole makes little difference. As can be seen in Table 1 and Figure 6, mass dipole correlates with bulk flow very weakly both in amplitude and in direction. The correlation becomes slightly tighter as sample volume increases. The peak of the distribution function of the misalignment angle between mass dipole and bulk velocity is at \( \sim 45^\circ \) for \( R = 25 \ h^{-1} \text{ Mpc} \) and shifts to \( \sim 32^\circ \) for \( R = 100 \ h^{-1} \text{ Mpc} \) (Figure 6).

**4. THE FASTEST BULK FLOW**

There are some works that claim detection of unusually large bulk flows (e.g., Feldman et al. 2010; Weyant et al. 2011). It is therefore interesting to check the properties of these areas in the CDM universe. The largest amplitudes of bulk motion measured in our simulations for \( R = 25, 50, \) and \( 100 \ h^{-1} \text{ Mpc} \) are 1070, 778, and \( 514 \text{ km s}^{-1} \), respectively. The value for \( R = 100 \ h^{-1} \text{ Mpc} \) is already very close to those unusual observational results. The possibility of residing in the cell with the fastest bulk motion is defined by the ratio of the cell volume to the total volume of the simulation, \( \sim 0.42\% \) for \( R = 100 \ h^{-1} \text{ Mpc} \). However, given that the diameter of the cell is as large as \( 200 \ h^{-1} \text{ Mpc} \), a huge volume moving at a speed more than \( \sim 500 \text{ km s}^{-1} \) will yield observable features too prominent to be missed.

Here we choose the \( R = 100 \ h^{-1} \text{ Mpc} \) case as the example to study the peculiarities of the cell showing the fastest bulk motion. The first physical quantity checked is the halo mass function in the cell; the halo mass functions of the fastest 10 cells (centers separated by at least \( 100 \ h^{-1} \text{ Mpc} \)) are also measured for comparison. Figure 7 shows the measured cell halo mass functions divided by the halo abundance in the full catalog. Of the 10 cell mass functions, most (more than 7 of 10) are smaller than the halo mass function of the full catalog for \( m_h \lesssim 10^{11} \ h^{-1} \text{ M}_\odot \). In the high-mass regime we cannot draw any reliable conclusions due to the very small number of high-mass halos, though the mass function of the fastest cells demonstrates a high tail. So far we can only cautiously conclude that there is a tendency to find a smaller number of low-mass halos in high bulk flow regions.

The distribution of all halo velocities in the cell is illustrated in Figure 8, which also shows the velocity distribution functions for these halos in the mass bins (0.25–0.3) \( \times 10^{12} \ h^{-1} \text{ M}_\odot \), (0.62–1.12) \( \times 10^{12} \ h^{-1} \text{ M}_\odot \), and \( > 1.62 \times 10^{12} \ h^{-1} \text{ M}_\odot \). Mass-binned halos do not exhibit any significant differences with

---

**Table 1**

| Cell Radius \( R (h^{-1} \text{ Mpc}) \) | Mass Monopole \( \sum m_i \) | Mass Monopole \( N \) | Mass Dipole \( \langle r_i/\langle r \rangle \rangle \) |
|----------------------------------------|----------------|-----------------|----------------------|
| 25                                     | 0.024          | -0.0004         | 0.139                |
| 50                                     | 0.004          | -0.0188         | 0.188                |
| 100                                    | 0.014          | -0.0031         | 0.245                |

**Figure 5.** Measured local bulk flows against the prediction of the linear model. Symbols are some recently estimated amplitudes of local bulk flows from observational data of galaxies (Feldman et al. 2010; Nusser & Davis 2011) and supernovae (Haugbølle et al. 2007; Colin et al. 2011; Dai et al. 2011; Weyant et al. 2011). The solid line is the most likely bulk speed predicted by the model, the dashed lines indicate variance at the 1σ level (68.3%), and the dot-dashed line and the dotted line are at levels of 2σ (95.5%) and 3σ (99.7%), respectively.
Figure 6. Probability distribution function of the angle between the bulk flow and the mass dipole. The dash-dotted line is the expectation of null correlation. (A color version of this figure is available in the online journal.)

Figure 7. Difference in halo abundance between the fastest cell with $R = 100\, h^{-1}$ Mpc and the full halo catalog, $d\nu_{\text{Cell}}/d\nu_{\text{All}}$. The red solid circles connected by a solid line are for the fast cell, which is enclosed by dashed lines marking its Poisson variance. The blue dotted line is the average of the fastest 10 cells. The horizontal solid line is for $d\nu_{\text{Cell}}/d\nu_{\text{All}} = 1$. (A color version of this figure is available in the online journal.)

respect to their velocity distributions, which eases the worry of possible bias in mass-selected halo samples. The distribution function of the angle between the halo velocity and the bulk flow is very skewed toward small misalignment: the peak is around $20^\circ$–$30^\circ$ but not $0^\circ$, and about 90% of halos are moving in a direction within $60^\circ$ to the bulk flow. It appears that halos in the cell with largest bulk velocity are more likely to have higher speed: the peak of the velocity amplitude distribution of halos in the cell is at around $600\, \text{km s}^{-1}$, while that of the full halo catalog is at $\sim 380\, \text{km s}^{-1}$ (Figure 8).

We have shown that bulk flow is basically weakly correlated with the internal mass dipole on average. But for the cell with the largest bulk flow, intuitively one would conjecture that there should be certain very massive clumps neighboring the cell whose gravitational action might play a dominant role in causing such extreme bulk flow of nearby halos. As an attempt to justify the paradigm, mass dipoles in shells within $R = 100$–$300\, h^{-1}$ Mpc of the center of the $100\, h^{-1}$ Mpc cell with the largest bulk flow are calculated in four layers with the help of the Healpix package (Górski et al. 2005). If there is an
Figure 8. Velocity distribution of halos in the cell with the largest bulk flow ($R = 100 h^{-1}$ Mpc). The left panel presents the distribution of amplitudes of halo velocities; the thick solid line is for halos in the cell, the thin solid line is for the full halo catalog, and the other colored lines are for halos in three different mass bins as labeled in the figure legend. The right panel presents the distributions of the angle $\alpha = \cos^{-1}\left(\mathbf{V} \cdot \mathbf{v}_h / (|\mathbf{V}_h|)\right)$ between the bulk flow $\mathbf{V}$ and the halo velocity $\mathbf{v}_h$. (A color version of this figure is available in the online journal.)

Figure 9. Mollweide projection of directions of the largest bulk flow and mass dipoles. Mass dipoles are computed from the mass field represented by halos, $\sum m_i \mathbf{r}_i / r_i$. The plus marks the direction of the largest bulk flow ($514 \text{ km s}^{-1}$, cell radius $R = 100 h^{-1}$ Mpc), and the cross is the direction of the dipoles of halos in the cell; the angle between them is $37^\circ 4$. The triangle square, diamond, and circle are the directions of mass dipoles within shells $(100, 150), (150, 200), (200, 250)$, and $(250, 300) h^{-1}$ Mpc, which deviate from the bulk flow by about $30^\circ, 33^\circ, 21^\circ, 17^\circ$ respectively. The color map is the surface mass density contrast of halos in the cell, plotted in the Healpix scheme (dark color for high density contrast). (A color version of this figure is available in the online journal.)

An unusual distribution of matter in a layer, the mass dipole of the layer will be an efficient indicator of this. Projected directions of bulk flow and mass dipoles are displayed in Figure 9. When the shell moves outward, the mass dipole pointing walks fairly randomly around the bulk flow; the misalignment angle varies between $\sim 20^\circ$ and $40^\circ$, which is analogous to the typical value in Figure 6. It seems that the dipoles of local environmental mass are only aligned crudely with the bulk flow. The correlation is not negligible, however, and since these misalignment angles are not small, we have no strong support from the simulation to attribute the extremely large bulk flow in such a huge volume mainly to the inhomogeneous environment.

An anomaly is detected in the clustering of halos in the special cell—see Figure 10, which also shows the two-point correlation function of the halos in 1000 randomly located cells of radius $100 h^{-1}$ Mpc and the full halo catalog. $\xi_h(r)$ averaged over the 1000 measurements agrees with $\xi_h$ of the full halo catalog at scales $r \lesssim 30 h^{-1}$ Mpc and then drops down more quickly to zero at larger scales due to the integral constraint resulting from the finite volumes of the cells (Landy & Szalay 1993). As we are interested in $\xi_h$ in a finite volume, we did not bother to apply the relevant correction. $\xi_h$ of the cell with the largest bulk velocity is higher than the average of random cells by around $2\sigma$ at scales $r \lesssim 20 h^{-1}$ Mpc; at larger scales the excess of clustering power rises to levels of $\sim 2-3\sigma$. In order to assess the statistical significance of the event, correlation coefficients of the bulk flow amplitude and $\xi_h(r)$ are computed from the 1000 random cells (bottom panel in Figure 10). Little correlation is detected...
from measurements of 1000 randomly selected cells. We further check the halo clustering at large scales, which appears to be in line with the velocity it is indeed more possible to find a power excess in halo constraint in this regime is negative and will bend the integral constraint, because the leading term of the integral find that 7 of the 10 demonstrate a power excess at similar scales. 

\( R \) (radius \( r \) at scales \( R = 100 h^{-1} \) Mpc). Top panel: the solid line is the two-point correlation function \( \xi_h \) of halos in the cell with the largest bulk velocity, the dotted red line is the average over 1000 \( \xi_h \) values measured in randomly selected cells while the yellow shadow marks the corresponding 1σ variance, and the dashed blue line is the result for all halos in our halo catalog. Bottom: correlation coefficients between amplitude of bulk velocity and \( \xi_h(r) \) as a function of \( r \), summarized from measurements of 1000 randomly selected cells.

(A color version of this figure is available in the online journal.)

at scales \( r \lesssim 20 h^{-1} \) Mpc, but at larger scales the correlation becomes much stronger, indicating that in cells with high bulk velocity it is indeed more possible to find a power excess in halo clustering at large scales, which appears to be in line with the findings of Macaulay et al. (2011). We further check the halo two-point correlation functions in the fastest 10 cells, and we find that 7 of the 10 demonstrate a power excess at similar scales. Note that the power excess at large scales cannot be ascribed to the integral constraint, because the leading term of the integral constraint in this regime is negative and will bend \( \xi_h \) downward (Landy & Szalay 1993).

5. SUMMARY AND DISCUSSION

Through analysis of the Pangu simulation, it is confirmed that the bulk flow of halos follows a Maxwellian distribution that is completely determined by a single parameter, the bulk velocity dispersion. We find that the dispersion measured in simulations agrees very well with the predictions of linear perturbation theory of structure formation; nonlinearity only becomes important when the sampling volume is very small. In most cases the mass-weighted bulk flow has some minor statistical differences from the unweighted one, but this will not affect the overall statistics significantly. It is also revealed that, statistically, the bulk flow has little systematic dependence on the mass of halos used for the estimation. Based on the results, we propose a unified scheme to compare results from observational samples with theories. In the proposal, the scale at which bulk flow in a particular region of the universe is measured is chosen to be the effective scale \( R \), which is the radius of a spherical top-hat window function \( W_{th} \) that yields the same bulk velocity dispersion \( \sigma_V \) as the practical window function \( W_O \) for the observational sample in linear theory. \( W_O \) not only is determined by the sample geometry but also contains weights emerging from the selection function, incompleteness, etc., and is analogous to the window function used in estimating power spectra. Numerical experiments indicate that the effects of selection functions on the bulk flow estimation can be corrected to a good accuracy by the simple treatment of Equation (10).

Variance ranges of bulk flow are also clarified on the basis of a Maxwellian distribution. We make a rough comparison of some recent measurements with the ΛCDM model adopted in Pangu simulation and find that some results do deviate from the model by about 3σ, but the tension between observation and model is not as strong as the original works claimed. The estimated effective scales for observational results in Figure 5 are in fact the upper limits; the true effective scales could be even smaller since we have assumed that these samples have full sky coverage and that their selection functions have been corrected for during the estimation. Furthermore, the observed bulk velocity contains residuals from thermal motion of galaxies in their host halo, and this has not been included in calculations of the velocity dispersion so far. More accurate modeling could be developed by assuming that the velocities of galaxies relative to their halos obey a certain simple distribution, but this requires explicit knowledge of the occupation details of galaxies in their host halos, which in itself is already a challenging problem. It is simpler to deduct the random motion component in the estimation procedure, as in the treatment of Wang (2007).

The correlation between the bulk flow and the dipole of internal mass is very weak but is stronger for larger volumes. If the bulk flow resides in a volume of radius greater than \( 100 h^{-1} \) Mpc with large bulk velocity, the observed mass dipole in the volume will have a considerable chance of being unusually strong. The typical misalignment angle between the bulk flow and the mass dipole is mostly likely around (~30°–50°). This might introduce non-negligible systematic bias to cosmological probes involving local mass distribution, such as the late-time integrated Sachs–Wolfe effect (Rees & Sciama 1968).

In our simulation volumes of scale extending to 200 \( h^{-1} \) Mpc in diameter moving with extreme large bulk velocity (>500 km s\(^{-1}\)) do exist. Most halos inside the volume are moving in alignment with the bulk flow within 60°, and the flow shows no dependence on halo mass. Such group motion of numerous halos will generate prominent kSZ signals; this is a rare event, but given its high speed and huge scale (200 \( h^{-1} \) Mpc versus 1 \( h^{-1} \) Gpc), the probability of detection is actually not very small, and of course it also depends on the inclination between the bulk flow and the line of sight. Another possible observation effect is that galaxies in the volume could be dimmed or brightened on average by the extreme bulk flow from galaxies in other places: this was the starting point of the efforts of Nusser et al. (2011) and Abate & Feldman (2012).

Dipoles of mass outside the environment of the largest bulk flow region are not tightly aligned with the bulk velocity but deviate from it by around ~20°–40°. Interestingly, we identified that halo clustering of the particular volume is apparently strengthened at scales \( r \gtrsim 20 h^{-1} \) Mpc; the simulation results indicate that such an enhancement is not completely accidental, as at large scales the two-point correlation function of halos in a finite volume is indeed mildly correlated with bulk velocity. The bulk velocity is dominated by Fourier modes of velocity at scales larger than the characteristic scale of the sample, while in linear theory \( v(k) \propto k\delta(k)/k^2 \). Unusually large bulk velocity seems to imply that there should be an extraordinarily super-large mode of density fluctuation topping up the region. However, bulk flow is barely correlated with the total number or mass of enclosed halos inside the sample volume, and, as we checked, the total number or mass of halos in the cell with the largest bulk velocity...
is less than the mean value but still within the 2σ variance range. Moreover, in the linear regime Fourier modes of density fluctuation are independent—the $\xi$ of halos inside the volume is actually controlled by modes of scale less than the characteristic scale of the sample.

Aside from the theoretical puzzle, there is the question of whether the power excess of halo clustering in a region at scales $\gtrsim 20 \, h^{-1} \, \text{Mpc}$ can be used as an indicator of candidate space for extremely large bulk flow. The advantage of using the two-point correlation function is that clustering does not rely on the direction of the line of sight and the complication resulting from redshift distortion can in principle be overcome by the ratio of $\xi_r(\pi \sim 20–50 \, h^{-1} \, \text{Mpc})$ to $\xi_0$ at small scales, e.g., $\sim 10 \, h^{-1} \, \text{Mpc}$ where $\xi_0$ is not correlated with bulk flow. A more serious concern is that if we are unfortunately (or luckily) in a special region as large as our current largest galaxy survey with extremely large bulk flow, the measured clustering strength at and beyond scale of baryonic acoustic oscillation would be significantly leveled out. Could this be the cause of the power excess at very large scales in the Baryon Oscillation Spectroscopic Survey elaborated by Ross et al. (2012)? To answer these questions, one surely needs multiple realizations of simulations of volumes much larger than those of our Pangu simulation. For the time being we must leave these questions open.

This work is partly supported by the NSFC through grant nos. 10873027, 10873035, 11073055, and 11133003. Y.P.J., W.P.L., X.H.Y., and P.J.Z. are members of the Innovation Group funded by NSFC (no. 11121062). J.P. and X.K. acknowledge the One-Hundred Talent fellowships of CAS. We appreciate stimulating discussion with Jiasheng Huang, Cheng Li, Guoliang Li, and Lifan Wang, as well as the very helpful comments and suggestions of the anonymous referee.

The Pangu simulation was carried out in the Supercomputing Center of CNIC, CAS, under the collaboration scheme of the Computational Cosmology Consortium of China (C4); participating institutions are NAOC, PMO, SHAO, and CNIC.

REFERENCES

Abate, A., & Feldman, H. A. 2012, *MNRAS*, 419, 3482
Afshordi, N., Gershnizjani, G., & Khoury, J. 2009, *J. Cosmol. Astropart. Phys.*, **JCAP08(2009)030**
Amanullah, R., Lidman, C., Rubin, D., et al. 2010, *ApJ*, 716, 712
Bahcall, N. A., Cen, R., & Gramann, M. 1994, *ApJ*, 430, L13
Ciccielg, P., & Chodorowski, M. J. 2004, *MNRAS*, 349, 945
Colin, J., Mohayae, R., Sarkar, S., & Shafieloo, A. 2011, *MNRAS*, 414, 264
Courtois, H. M., Tully, R. B., Makarov, D. I., et al. 2011, *MNRAS*, 414, 2005
Dai, D.-C., Kinney, W. H., & Stoikovic, D. 2011, *J. Cosmol. Astropart. Phys.*, **JCAP04(2011)015**
Feldman, H. A., Watkins, R., & Hudson, M. J. 2010, *MNRAS*, 407, 2328
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
Haughbølle, T., Hamnestad, S., Thomsen, B., et al. 2007, *ApJ*, 661, 650
Jarosik, N., Bennett, C. L., Dunkley, J., et al. 2011, *ApJS*, 192, 14
Jha, S., Riess, A. G., & Kirshner, R. P. 2007, *ApJ*, 659, 122
Kashlinsky, A., Atrio-Barandela, F., & Ebeling, H. 2011, *ApJ*, 732, 1
Kashlinsky, A., Atrio-Barandela, F., Ebeling, H., Edge, A., & Kocevski, D. 2010, *ApJ*, 712, L81
Laahv, O., Lilje, P. B., Primack, J. R., & Rees, M. J. 1991, *MNRAS*, 251, 128
Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
Lavaux, G., Afshordi, N., & Hudson, M. J. 2012, arXiv:1207.1721
Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, 538, 473
Linder, E. V. 2005, *Phys. Rev. D*, 72, 043529
Ma, Y.-Z., & Scott, D. 2012, arXiv:1208.2028
Macaulay, E., Feldman, H., Ferreira, P. G., Hudson, M. J., & Watkins, R. 2011, *MNRAS*, 414, 621
Mak, D. S. Y., Pierpaoli, E., & Osborne, S. J. 2011, *ApJ*, 736, 116
Mersini-Houghton, L., & Holman, R. 2009, *J. Cosmol. Astropart. Phys.*, **JCAP02(2009)006**
Moscardini, L., Branchini, E., Brunozzi, P. T., et al. 1996, *MNRAS*, 282, 384
Nusser, A., Branchini, E., & Davis, M. 2011, *ApJ*, 735, 77
Nusser, A., & Davis, M. 1994, *ApJ*, 421, L1
Nusser, A., & Davis, M. 2011, *ApJ*, 736, 93
Nusser, A., Dekel, A., Bertschinger, E., & Blumenthal, G. R. 1991, *ApJ*, 379, 6 Osbome, S. J., Mak, D. S. Y., Church, S. E., & Pierpaoli, E. 2011, *ApJ*, 737, 98
Park, C.-G., & Park, C. 2006, *ApJ*, 637, 1
Pham-Gia, T., Turkkan, N., & Marchand, E. 2006, *Commun. Stat.—Theory Methods*, 35, 1569
Rees, M. J., & Sciama, D. W. 1968, *Nature*, 217, 511
Ross, A. J., Percival, W. J., Sánchez, A. G., et al. 2012, *MNRAS*, 424, 564
Sandage, A., Reindl, B., & Tammann, G. A. 2010, *ApJ*, 714, 1441
Saunders, W., Sutherland, W. J., Maddox, S. J., et al. 2000, *MNRAS*, 317, 55
Scoccimarro, R. 2004, *Phys. Rev. D*, 70, 083007
Song, Y.-S., Sabiu, C. G., Kayo, I., & Nichol, R. C. 2011, *J. Cosmol. Astropart. Phys.*, **JCAP05(2011)020**
Springel, V. 2005, *MNRAS*, 364, 1105
Springob, C. M., Masters, K. L., Haynes, M. P., Giovannelli, R., & Marinoni, C. 2007, *ApJS*, 172, 599
Strauss, M. A., & Willick, J. A. 1995, *Phys. Rep.*, 261, 271
Sunyaev, R. A., & Zeldovich, I. B. 1980, *MNRAS*, 190, 413
Szapudi, I. 1998, *ApJ*, 497, 16
Turnbull, S. I., Hudson, M. J., Feldman, H. A., et al. 2012, *MNRAS*, 420, 447
Wang, L. 2007, arXiv:0705.0363
Weyant, A., Wood-Vasey, M., Wasserman, L., & Freeman, P. 2011, *ApJ*, 732, 65
Wyman, M., & Khoury, J. 2010, *Phys. Rev. D*, 82, 044032