

Dipolar modulation of Large-Scale Structure

by

Mijin Yoon

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in the University of Michigan
2016

Doctoral Committee:

Professor Dragan Huterer, Chair
Professor Fred C. Adams
Professor August E. Evrard
Professor Jeffrey M. McMahon
Professor Christopher J. Miller

©Mijin Yoon

2016

To my God

A C K N O W L E D G M E N T S

First and foremost, I would like to thank my advisor, Dragan Huterer. Due to his support, I could develop my research skills and build a mindset to become a better researcher. I would also like to thank my former advisor, Jeffrey McMahon for encouragement and advice during the early stage of my research. I thank members of my dissertation committee, Fred Adams, August Evrard, Christopher Miller for detailed comments and suggestions during the defense process. I appreciate my collaborators, István Szapudi, Cameron Gibelyou and András Kovács and Maciej Bilicki for exchanging great research ideas and providing helpful comments.

I especially thank Kwanjeong Educational Foundation for financial support during my Ph.D studies.

TABLE OF CONTENTS

Dedication	ii
Acknowledgments	iii
List of Figures	vi
List of Tables	ix
Abstract	x
Chapter	
1 Introduction	1
1.1 Observational Cosmology	1
1.2 Large Scale Structure as a cosmological probe	3
1.3 Testing Statistical Isotropy	7
1.4 Dipolar Modulation of LSS	8
1.5 Outline	10
2 Dipolar modulation in number counts of WISE-2MASS sources	12
2.1 Introduction	12
2.2 Culling of the WISE dataset	13
2.3 Methodology	18
2.3.1 Dipole estimator	18
2.3.2 Foreground Templates and Estimator Validation	20
2.3.3 Theoretical expectation of local-structural dipolar modulation	21
2.4 Results	23
2.5 Conclusions	25
3 Kinematic Dipole Detection with Galaxy Surveys	27
3.1 Introduction	27
3.2 Methodology	29
3.2.1 Theoretical signal	29
3.2.2 Statistical error	30
3.2.3 Systematic bias	33
3.3 Results	35
3.4 Conclusions	40

4 Closing Remarks 42
Appendix 45
Bibliography 51

LIST OF FIGURES

Figure

1.1	Composition of the Universe today according to Planck observations. Adapted from http://www.esa.int/spaceinimages	2
1.2	Distribution of galaxies in Sloan Digital Sky Survey. Adopted from http://www.astronomy.ohio-state.edu/dhw/SDSS08/ofigs.html	4
1.3	Power spectrum $P(k)$ extrapolated to $z = 0$ from various observations. Adapted from [Tegmark et al., 2004].	5
1.4	Cosmological constraints from BAO, SNIa, and CMB observations.	6
1.5	Amplitude of dipolar modulations from different surveys and the rough estimate of local-structure dipole depending on the survey depth. This shows the possibility of detecting kinematic dipole from the deeper surveys. Credit: Dragan Huterer	10
2.1	Stars (blue) and galaxies (red) separation based on WISE mag bands (W1, W2, W3 and W4) and 2MASS mag band (J): It shows obvious usage of 2MASS J band. Using only WISE bands is not enough to separate galaxies and stars. The pink dotted line is the cut used in previous study Goto et al. (2012). The black dotted line is our newly selected cut. Adapted from [Kovács & Szapudi, 2013].	15
2.2	Star contamination (green) and completeness (blue) depending on the color cut (W1-J) in galaxies of WISE combined with 2MASS data. Adapted from [Kovács & Szapudi, 2013].	15
2.3	Map of WISE-2MASS sources that we used with 10 degree Galactic cut (before masking out the contaminated region with the WMAP dust mask). The criteria are described in the text. The map shown is a Mollweide projection in Galactic coordinates with counts binned in pixels of about 0.5° on a side (HEALPix resolution $N_{SIDE} = 128$). The two elliptical sets of contours represent the measured dipole direction when we applied a 10° (left) and 20° (right) Galactic cut, respectively (that is, with $ b < 10^\circ$ and $ b < 20^\circ$). The red, blue, and white colors in those contours represent the 68%, 95%, and 99% confidence regions for the direction.	16
2.4	Number counts of WISE sources as a function of redshift. We obtain redshift information by matching WISE sources to those from the GAMA DR2 catalog. 96.9% of WISE sources are found in GAMA.	18

2.5	Theoretical prediction for the dipole amplitude (horizontal blue line), together with the measured values in WISE (green points). The two sets of error bars on the measurements correspond to 68% and 95% confidence; they have been calculated from the full likelihood in Eq. (2.11) and are rather symmetric around the maximum-likelihood value. The two large horizontal bands around the theory prediction correspond to 1- and 2-sigma cosmic variance error.	23
3.1	Sketch of the problem at hand: we would like to measure the kinematic dipole \vec{v}_{kin} , whose error (represented by a cyan ellipse) can be calculated given the number of extragalactic objects and the sky coverage. The LSS local dipole, \vec{v}_{local} , provides a bias in this measurement. For a survey deep enough (and depending somewhat on the direction of its \vec{v}_{local}), bias in the measurement of \vec{v}_{kin} will be smaller than the statistical error.	33
3.2	The amplitude of local-structure dipole as a function of redshift with a probability distribution $n(z)$	34
3.3	Four footprints of survey coverage: galactic cut $\pm 15^\circ$ (above left), galactic cut $\pm 30^\circ$ (above right), cap cut 60° (below left), and cap cut 90° (below left). Blue color represents the excluded area.	35
3.4	Statistical error in the dipole amplitude, marginalized over direction and other multipoles that are coupled to the dipole, as a function of the number of galaxies in a survey. The top horizontal dashed line shows the amplitude expected based on the CMB dipole measurements ($A = 0.0028$), and is the fiducial value in this work. The two dashed horizontal lines show the 3σ and 5σ detection of dipole with the fiducial amplitude.	36
3.5	Statistical error in the dipole amplitude when $\theta_{\text{kin-dip}} = 0^\circ$ and $\theta_{\text{kin-dip}} = 90^\circ$: the angle $\theta_{\text{kin-dip}}$ is defined as the angle between the galactic north pole and the direction of kinematic dipole. In Fig. 3.4, we assumed the direction of kinematic dipole same as CMB dipole. For galactic cuts, the case when kinematic dipole is pointing toward the galactic north pole is better than the case when kinematic dipole is aligned with the galactic plane because the highest density contrast is captured within the survey area for the former case. For cap cuts, the highest density contrast is captured when it is not pointing toward the galactic north pole so the trend is the opposite to galactic cut scenarios.	38
3.6	$\Delta\chi^2$, defined in Eq. (3.9), corresponding to the bias from the local-structure dipole, as a function of the number of objects N_{gal} . For a fixed amplitude of $\mathbf{d}_{\text{local}}$ the $\Delta\chi^2$ still depends on the direction of this vector; here we show the value averaged over all directions of $\mathbf{d}_{\text{local}}$. Solid lines show cases when the median galaxy redshift is $z_{\text{med}} = 1.0$, while dashed lines are for $z_{\text{med}} = 0.75$. The legend colors are the same in fig.3.4.	39
3.7	Goodness-of-fit ($\Delta\chi^2$ from Eq. (3.9)) dependence of the direction of LSS dipole: θ_{dip} is the angle between the direction of LSS dipole and the galactic north pole. This is the case when $N_{\text{gal}} = 10^7$ and $z_{\text{med}} = 0.5$ and this directional dependence is similar for other conditions.	40
3.8	Statistical error in the dipole angle as a function of the number of galaxies in a survey for four sky cuts.	41

A.1	The comparison between with priors vs. without priors of statistical errors ($\sigma_x, \sigma_y, \sigma_z$) in x, y, z direction as a function of ℓ_{\max}	48
A.2	The statistical errors ($\sigma_x, \sigma_y, \sigma_z$) in x, y, z direction as a function of ℓ_{\max} with priors; priors are based on the expected values of higher multipoles (C_ℓ).	49
A.3	Effects of WMAP mask on the statistical errors in dipole amplitude: this shows the difference between the cases with (black dotted) and without (colored) WMAP mask applied in estimation of statistical errors.	50

LIST OF TABLES

Table

2.1	Measurements of the dipole amplitude in WISE for various Galactic cuts (b_{cut}) corresponding to fractions of the sky covered (f_{sky}). In all cases we marginalized over several foreground templates, as described in the text. The full likelihood for the amplitude A_{WISE} is well approximated by a Gaussian whose mode and standard deviation we quote here. We also show the theoretical expectation A_{theory} due to the local-structure dipole, together with the corresponding cosmic variance given a bias $b = 1.41$	24
3.1	Relative motions of systems contributing to kinematic dipole.	28

ABSTRACT

For the last two decades, we have seen a drastic development of modern cosmology based on various observations such as the cosmic microwave background (CMB), type Ia supernovae, and baryonic acoustic oscillations (BAO). These observational evidences have led us to a great deal of consensus on the cosmological model so-called Λ CDM and tight constraints on cosmological parameters consisting the model. On the other hand, the advancement in cosmology relies on the cosmological principle: the universe is isotropic and homogeneous on large scales. Testing these fundamental assumptions is crucial and will soon become possible given the planned observations ahead.

Dipolar modulation is the largest angular anisotropy of the sky, which is quantified by its direction and amplitude. We measured a huge dipolar modulation in CMB, which mainly originated from our solar systems motion relative to CMB rest frame. However, we have not yet acquired consistent measurements of dipolar modulations in large-scale structure (LSS), as they require large sky coverage and a number of well-identified objects.

In this thesis, we explore measurement of dipolar modulation in number counts of LSS objects as a test of statistical isotropy. This thesis is based on two papers that were published in peer-reviewed journals. In Chapter 2 [Yoon et al., 2014], we measured a dipolar modulation in number counts of WISE matched with 2MASS sources. In Chapter 3 [Yoon & Huterer, 2015], we investigated requirements for detection of kinematic dipole in future surveys.

CHAPTER 1

Introduction

1.1 Observational Cosmology

During the last two decades the field of cosmology has progressed rapidly by relying on abundant observational evidence. We confirmed that our Universe is flat and its expansion is accelerating through observations of the oldest light coming from the early universe called cosmic microwave background (CMB) and a long distance ladder, the supernovae (SNIa). Various observations reached a good agreement with the “standard model” - Λ CDM model (acceleration induced by cosmological constant and containing cold dark matter) and they tightly constrain cosmological parameters of the model.

Friedmann’s equation describes the expansion of the Universe according to its composition:

$$\frac{H^2(a)}{H_0^2} = \Omega_M a^{-3} + \Omega_R a^{-4} + \Omega_{DE} a^{-3(1+w)} - \Omega_K a^{-2} \quad (1.1)$$

where

$$H(a) \equiv \frac{da}{dt} \frac{1}{a}. \quad (1.2)$$

Here scale factor, $a(t)$, quantifies the scale of spatial expansion at a give time relative to the present value, $a(t_0) = 1$, satisfying $d(t) = a(t)d_0$, where $d(t)$ is distance at time t , and d_0 is distance today, and w is the equation of state of dark energy ($w = p/\rho$). The

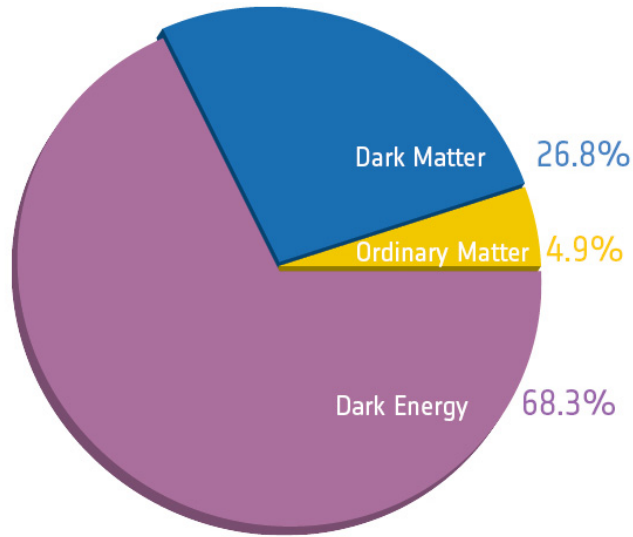


Figure 1.1: Composition of the Universe today according to Planck observations. Adapted from <http://www.esa.int/spaceinimages>.

Hubble parameter, $H(a)$, represents the rate of the expansion at $a(t)$. This equation shows that the rate of expansion at a time t is related to the relative proportion (for species X, $\Omega_X \equiv \rho_X/\rho_c$, $\rho_c \equiv 3H_0^2/8\pi G$: critical density) of components comprising the Universe today. The components are matter (M), radiation (R), dark energy (DE) and curvature (K : $\Omega_K = 0$ for a flat universe). As the scale factor (a) increases, dark energy becomes the main component driving the expansion following periods of radiation-dominated and matter-dominated universe, in that order. The composition is constrained such as dark energy comprises 68.3%, dark matter, 26.8% and ordinary matter, 4.9% according to Planck observations [Ade et al., 2015]. These cosmological parameters are being constrained more tightly by precise observations using different probes, but the search for physical origins of dark energy and elements of dark matter still remain a great challenge for the future.

The Friedmann-Robertson-Walker (FRW) metric is a metric describing an isotropic and homogeneous expanding (or contracting depending on scale factor) universe. Here, isotropy means that the Universe is the same in all directions on large scales while homogeneity means that the Universe looks the same from all locations. Based on these

fundamental assumptions, FRW metric represents the simple standard model:

$$ds^2 = -c^2 dt^2 + a^2(t)[dr^2 + S_k^2(r)d\Omega^2] \quad (1.3)$$

where

$$S_k(r) = R_0 \sin(r/R_0), \quad \kappa = +1, \quad (1.4)$$

$$= r, \quad \kappa = 0, \quad (1.5)$$

$$= R_0 \sinh(r/R_0), \quad \kappa = -1. \quad (1.6)$$

The constant κ is a dimensionless number that signifies the curvature of space. Positive κ implies positive curvature while negative κ implies negative curvature and $\kappa = 0$ implies flat space. In any case, note the fact that the scale factor ($a(t)$) is only a function of time (t), and $S_k(r)$ is a function of radial distance (r), relies on isotropy. Otherwise, if they depended on directions (θ , and ϕ), cosmological modelling would become very complicated.

1.2 Large Scale Structure as a cosmological probe

Large-scale structure (LSS) is, roughly speaking, structure on a bigger scale than a galaxy, which is comprised of galaxies, galaxy groups, galaxy clusters, superclusters, sheets, walls, and filaments, as shown in Fig.1.2 as an example. LSS started to form as matter became trapped into the gravitational potential valleys, which were triggered by primordial perturbations in the early Universe. This structure grew due to gravity while its spatial background stretched. Therefore, investigation of the LSS evolution has the power to constrain cosmological parameters such as density of matter (Ω_m) and dark energy density (Ω_Λ).

While it is impossible to predict the locations of individual structures, the two-point correlation function and power spectrum of the LSS are predictable based on the cosmological models. The two-point correlation function, $\xi(r)$, measures the excess probability

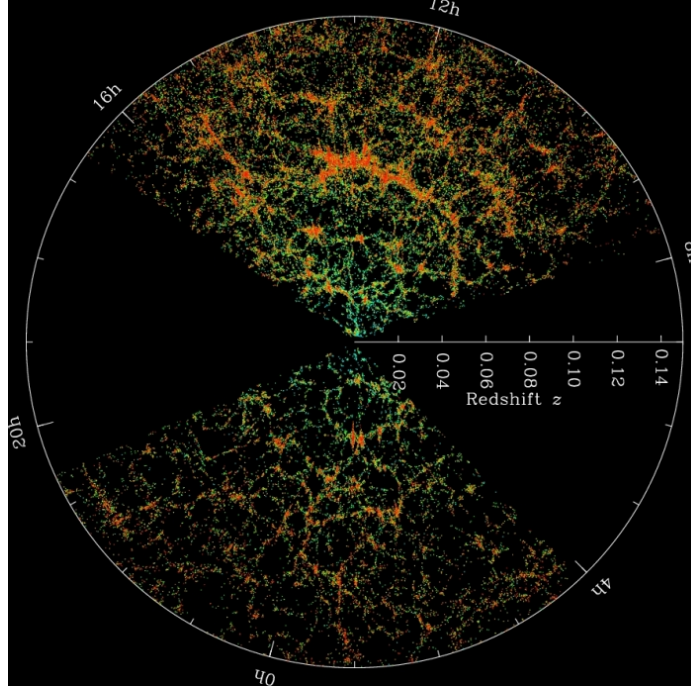


Figure 1.2: Distribution of galaxies in Sloan Digital Sky Survey. Adopted from <http://www.astronomy.ohio-state.edu/dhw/SDSS08/ofigs.html>

to find another galaxy at a distance (r) around a galaxy averaged over space, \vec{x} . (Bracket in the equation, $\langle \rangle$, means averaging over all samples.)

$$\xi(r) = \langle \delta(\vec{x} + \vec{r})\delta(\vec{x}) \rangle, \quad (1.7)$$

where the overdensity $\delta(\vec{x})$ is defined in terms of a matter density $\rho(\vec{x})$ and its mean density $\bar{\rho}$:

$$\delta(\vec{x}) = \frac{\rho(\vec{x}) - \bar{\rho}}{\bar{\rho}}. \quad (1.8)$$

The power spectrum $P(k)$ is the Fourier transform of the two-point correlation.

$$\delta_{\vec{k}} = \int \delta(\vec{r}) e^{i\vec{k}\cdot\vec{r}} d^3r \quad (1.9)$$

$$\langle \delta_{\vec{k}} \delta_{\vec{k}'}^* \rangle = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}') P(k) \quad (1.10)$$

where \vec{k} is the wave number. In deriving these functions, we assumed the statistical isotropy

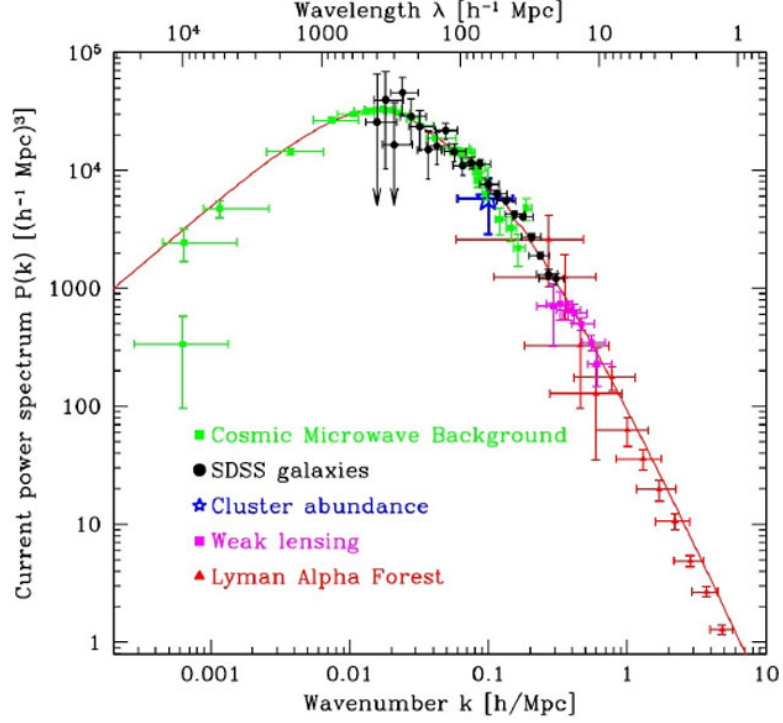


Figure 1.3: Power spectrum $P(k)$ extrapolated to $z = 0$ from various observations. Adapted from [Tegmark et al., 2004].

so we average over all directions on the sky. This simplifies the two-point correlation function $\xi(\vec{r})$ to $\xi(r)$ and the power spectrum $P(\vec{k})$ to $P(k)$.

Fig.1.3 [Tegmark et al., 2004] shows various measurements of power spectrum extrapolated to $z = 0$. Combined measurements include CMB and galaxy LSS, weak lensing of galaxy shapes, and Lyman alpha forest. They agree well with the prediction of the Λ CDM model shown as a solid line.

In this thesis, we will often deal with angular power spectrum C_ℓ , which is the spherical harmonic-space representation of the two point correlation function:

$$C_\ell = \frac{1}{(2\ell + 1)} \sum_m \langle a_{\ell m} a_{\ell m}^* \rangle, \quad (1.11)$$

$$a_{\ell m} = \int \delta(\theta, \phi) Y_{\ell m}^*(\theta, \phi) d\Omega. \quad (1.12)$$

Observable objects such as galaxies and galaxy clusters are biased tracers of the under-

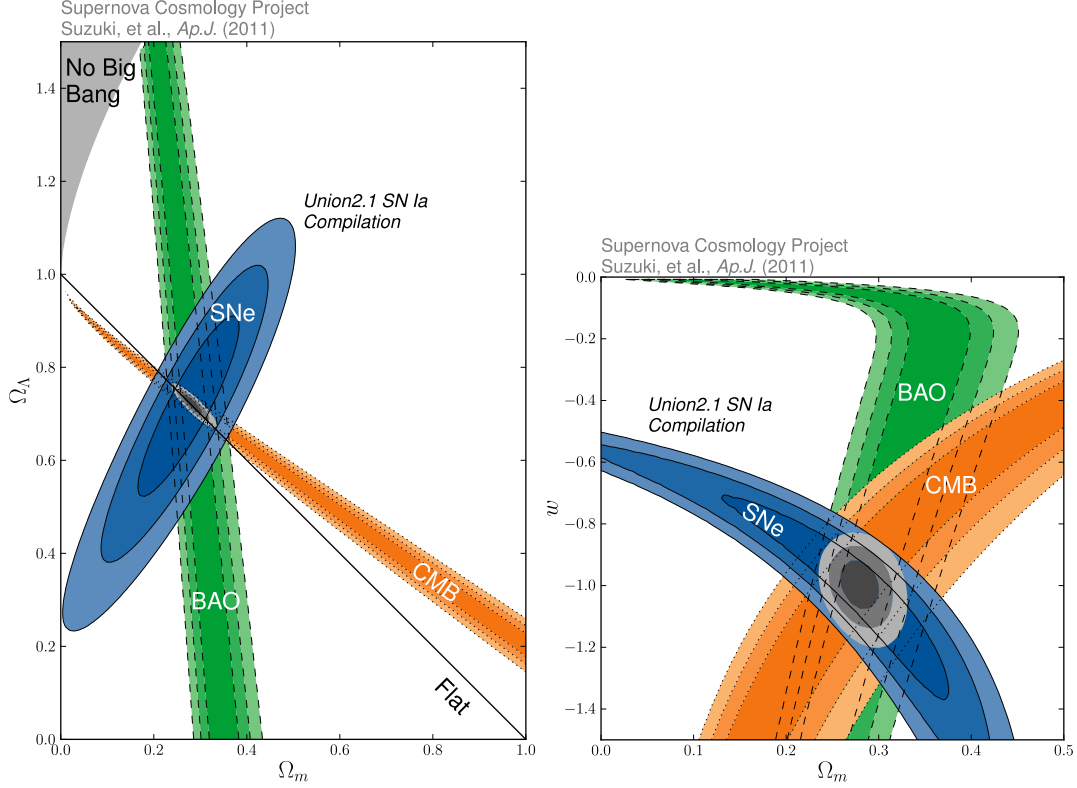


Figure 1.4: Cosmological constraints from BAO, SNIa, and CMB observations.

lying matter density with bias, b , set by the ratio between the overdensity $\delta_g(\vec{x}, t)$ of visible objects and that of the total underlying matter including dark matter density $\delta(\vec{x}, t)$,

$$b = \delta_g / \delta. \quad (1.13)$$

This unknown bias factor depends on the properties of observed sources and their redshifts. Thus, precise calibration of bias is crucial to match the amplitude of observed power spectrum of LSS with the one of theoretical power spectrum.

One way of constraining cosmological parameters based on LSS comes from the measurement of baryonic acoustic oscillations (BAO). The BAO signal, shown as a small bump in the power spectrum of LSS on scales $\sim 150\text{Mpc}$ (in today's universe), plays a role of a standard ruler. Not the amplitude but the scale of the peak reveals the sound horizon scale (at the given redshift) which is the imprinted oscillation due to the pressure waves gener-

ated in the photon-baryonic fluid in the early Universe. The BAO signal corresponds to the peaks of the Cosmic Microwave Background (CMB) power spectrum. The Fig. 1.4 shows that degeneracies among cosmological parameters - Ω_Λ , Ω_m , and w - break by combining observations of BAO, CMB and SNIa.

1.3 Testing Statistical Isotropy

Modern cosmological models rely on two basic assumptions: isotropy and homogeneity. The statistical isotropy specifically means isotropy on average over all realizations. These fundamental assumptions of the statistical isotropy and homogeneity have helped the advancement of cosmology because they simplify the analysis and permit the generalization of local observations to the whole Universe.

For example, we can constrain cosmological parameters using the CMB angular power spectrum C_ℓ which is defined in Eq. 1.11. In practice, C_ℓ is measured by averaging over the $2\ell + 1$ values of m for each ℓ . If the statistical isotropy did not hold, instead of C_ℓ we would have to measure

$$C_{\ell m \ell' m'} \equiv \langle a_{\ell m} a_{\ell' m'}^* \rangle \quad (1.14)$$

for each (ℓ, m, ℓ', m') . It is obvious that without these assumptions, we lose statistical power to constrain the cosmological parameters. Therefore, testing these two assumptions is essential to verifying the foundation of cosmological model. This is particularly important given that we do not fully understand the physics behind dark energy.

Since cosmologists already assume that statistical isotropy holds most of the time, testing it requires developing special statistical tools. For example, the Bipolar Power Spectrum Biposh (BiPS) is one of the tools that has been used for testing violations of the statistical isotropy from WMAP temperature and polarization maps. BiPS extracts non-statistical isotropic information from the off-diagonal elements in the $\langle a_{\ell m} a_{\ell' m'}^* \rangle$ correlation. While there was no evidence of violation of statistical isotropy found in the WMAP

temperature map, in the polarization maps, broken statistical isotropy was detected without a complete understanding of its origin [Hajian & Souradeep, 2006]. Other modified estimators were developed to achieve direction dependency in spherical harmonics. Maximum angular momentum direction (MAMD) is defined as a direction that gives maximum value of L^2 which sums over $|a_{\ell m}|^2$ weighted by m^2 for each m . Using MAMD, anomalous alignment of dipole, quadrupole and octopole in WMAP was detected [Copi et al., 2006]. People have investigated to check the difference among cosmological parameters constrained on different patches of simulated CMB maps and found that A_s is most susceptible [Mukherjee et al., 2015]. These findings are nontrivial and need to be further investigated to confirm their origins.

While there is no solid theory which strongly supports the statistical isotropy of our Universe, inflationary model suggests that if our observable Universe is only a small patch of the exponentially expanded Universe after inflation, it is natural to assume a smooth isotropic Universe. Therefore, if any violation of the statistical isotropy is observed, then it would have far-reaching implications for our understanding of the early Universe. Until now, CMB observations have been utilized extensively to test the statistical isotropy, but obtaining sufficiently strong signal only from CMB is challenging because of the cosmic variance which arises from the fact that we have only one sample of universe. LSS has potential to contribute to test the statistical isotropy and to confirm the origins of anomalies found in CMB observations. In this thesis, we will focus on observations of LSS as a probe to test the statistical isotropy.

1.4 Dipolar Modulation of LSS

The statistical isotropy in LSS implies that the number counts of objects should be the same on average in all directions. For testing the statistical isotropy, the measurement of dipolar modulation in number counts of the LSS tracers could be used as a basic probe. The

dipolar modulation represents a large angle feature which quantifies the trend of observing more objects in a certain direction than the opposite direction. The dipolar modulation is the largest anisotropy quantifiable in angular power spectrum (C_ℓ), and its amplitude is proportional to $\sqrt{C_1}$.

In CMB measurements, the dipolar modulation induced by the relative motion of our solar system to the CMB rest frame was already measured accurately. However, it has been more difficult to measure dipolar modulation in LSS because it requires a survey covering a large area and a number of objects to mitigate the systematic effects. The possible origins of the dipolar modulation in LSS are categorized into three different categories: local-structure, kinematic, and intrinsic dipoles.

- Local-structure dipole: the matter distribution predicted by LCDM could generate the dipolar modulation within the limited depth of the survey. The amplitude of this kind of dipole decreases as we observe deeper region of the sky because the number density fluctuation decreases as we include more objects on large scales.
- Kinematic dipole: the motion of our solar system relative to the LSS rest frame causes the dipolar modulation due to the relativistic aberration and the Doppler effect and . The relativistic aberration causes the observed objects to look bunched up in the direction of motion. The Doppler effect blue-shifts the observed frequencies of the objects on the direction of motion and red-shifts those in the other direction. Therefore, the number of objects in the limited band width of a survey changes depending on the location relative to the direction of motion. The dipole measured in CMB has the same kinematic origin.
- Intrinsic dipole: this is the dipole generated by perturbation in early time of the Universe. There are several models aimed at explaining possible origins of intrinsic dipole. [Hirata \[2009\]](#) analyzed SDSS quasars to constrain isocurvature perturbations and ruled out the simplest curvaton-gradient model. However, there still remains the-

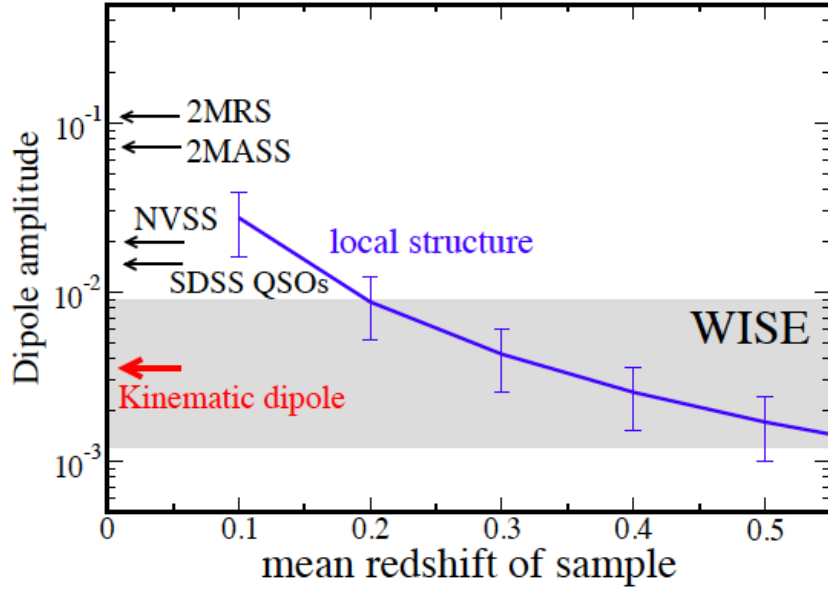


Figure 1.5: Amplitude of dipolar modulations from different surveys and the rough estimate of local-structure dipole depending on the survey depth. This shows the possibility of detecting kinematic dipole from the deeper surveys. Credit: Dragan Huterer

oretical support for an intrinsic dipole, since superhorizon-scale density fluctuation in the early Universe could appear as a dipolar modulation in our observable Universe.

1.5 Outline

The studies of the local-structure dipole and the kinematic dipole were developed in two peer-reviewed papers, respectively [Yoon & Huterer, 2015; Yoon et al., 2014].

Chapter 2 describes the measurement of local-structure dipole in number counts of WISE (Wide-field Infrared Survey Explorer) matched with 2MASS (Two Micron All Sky-Survey) galaxies. By matching sources in two surveys and using combined cuts, we generated a suitable map of galaxies to measure dipolar modulation. This measurement utilized many more objects (~ 2 millions) than the previous dipole measurements. The results appear robust since the subsets of the map have consistent results of amplitudes and directions considering the estimated error range. We compared the measured dipole amplitude with

the theoretical prediction and discovered a result somewhat off from the expectation.

Chapter 3 discusses the requirements for the detection of the kinematic dipole in the future surveys. Interestingly, the local-structure dipole appears as a systematic contamination in the detection of kinematic dipole. As we observe deeper, the amplitude of local-structure dipole decreases because it arises due to the limited depth of the observations. Therefore, the dipolar modulation for deeper survey mainly originates from the kinematic effect caused by our Solar system's motion relative to large-scale structure rest frame. We investigated the requirements for the number of objects, sky coverage, coverage shapes and the redshift range. As shown in Fig.1.5, compared to other previous surveys, WISE itself has potential to detect the kinematic dipole if all of the observed galaxies are selected well from the raw data. We expect to have a first detection of the kinematic dipole in the future surveys.

Chapter 4 summarizes the results from Chapter 2 & 3 and discusses the future surveys of large-scale structure with their implications for testing the statistical isotropy.

CHAPTER 2

Dipolar modulation in number counts of WISE-2MASS sources

2.1 Introduction

Modern surveys of large-scale structure allow tests of some of the most fundamental properties of the universe – in particular, its statistical isotropy. One of the most fundamental such tests is measuring the dipole in the distribution of extragalactic sources. One expects a nonzero amplitude consistent with the fluctuations in structure due to the finite depth of the survey; this “local-structure dipole” in the nomenclature of [Gibelyou & Huterer \[2012\]](#) is of order 0.1 for shallow surveys extending to $z_{\max} \sim 0.1$, but significantly smaller ($A \lesssim 0.01$) for deeper surveys. The motion of our Galaxy through the cosmic microwave background (CMB) rest frame also contributes to the dipole, but only at the level of $v/c \simeq 0.001$; while this kinematic dipole was detected in the CMB a long time ago, and more recently even solely via its effects on the higher multipoles in the CMB fluctuations [[Aghanim et al., 2014](#)], it has not yet been seen in large-scale-structure (LSS) surveys.

Measurements of the dipole in LSS therefore represent consistency tests of the fundamental cosmological model, and have in the past been applied to the distribution of sources in NVSS [[Blake & Wall, 2002](#); [Fernández-Cobos et al., 2014](#); [Hirata, 2009](#); [Rubart & Schwarz, 2013](#)]. Detection of an anomalously large (or small) dipole in LSS could indicate new physics: for example, motion between the CMB and LSS rest frames, or the

presence of superhorizon fluctuations [Itoh, Yahata & Takada, 2010; Zibin & Scott, 2008]. Moreover, in recent years, measurements of the bulk motion of nearby structures have been conducted, out to several hundred megaparsecs, using CMB-LSS correlations [Kashlinsky et al., 2008], or out to somewhat smaller distances, using peculiar velocities [Feldman, Watkins & Hudson, 2010; Watkins, Feldman & Hudson, 2009].

In this study, for the first time we test statistical isotropy using WISE (Wide-field Infrared Survey Explorer) [Wright et al., 2010]. WISE is, at least at first glance, perfectly suited to tests of statistical isotropy since it is deep and covers nearly the full sky. Moreover, its selection functions have been increasingly well understood over the past few years based on its observations in four bands sensitive to 3.4, 4.6, 12, and 22 μm wavelengths (called $W1$, $W2$, $W3$, and $W4$ respectively) with resolution in the 6''-12'' range [Ménard et al., 2013; Yan et al., 2013].

2.2 Culling of the WISE dataset

Our measurement of the dipole relies on a suitable selection of a representative sample of sources. The most important goal is to exclude Galactic sources – mainly stars. Galactic sources are expected to be concentrated around the Galactic plane, with density falling off to the north and south. While they are therefore expected to look like a Y_{20} quadrupole in Galactic coordinates, the residual contamination of the dipole may still be significant. Hence, in what follows we pay particular attention to magnitude and color cuts applied to WISE in order to leave a trustworthy set of extragalactic sources.

WISE is a space-based mission which was launched in Dec. 2009 and decommissioned in Feb. 2011. The Nov. 2013 release of WISE data includes 747 million objects in total. The redshift depth estimated by matching with SDSS sources almost reaches as deep as $z \sim 1$ [Yan et al., 2013]. Due to its sensitivity, the sources observed in W1 band contains most of the entire sources observed by WISE. In the WISE raw data, individual objects

were not identified so data selection is the key part of the analysis. We therefore apply carefully chosen criteria to define a map as uncontaminated by Galactic objects as possible. As argued in Kovács & Szapudi [2013], color cuts using only the WISE bands are not sufficient as shown in three panels of Fig. 2.1, except the one (above, left) with J magnitude from 2MASS. We therefore have applied 2MASS-PSC¹ magnitude ($J_{2\text{mass}}$) to distinguish between stars and galaxies.

2MASS is a survey observed in three infrared bands - J ($1.25 \mu\text{m}$), H ($1.65 \mu\text{m}$), and K_s ($2.17 \mu\text{m}$) - with a pixel size of $2'' \times 2''$. The observations were taken from 1997 to 2001 by twin telescopes located at Mt. Hopkins, Arizona for the northern hemisphere and at Cerro Tololo Inter-American Observatory, Chile for the southern hemisphere. 2MASS has two main catalogs: a Point Source Catalog (PSC) containing ~ 500 million stars and galaxies and an Extended Source Catalog (XSC) containing ~ 1.6 million resolved galaxies.

Both surveys cover almost full sky and 93% of WISE sources with $W1 < 15.2$ have the matched sources in 2MASS. The finally selected sources has 1.2% stellar contamination and 70.1% completeness as shown in Fig. 2.2 [Kovács & Szapudi, 2013].

Every source we use is observed in both WISE and 2MASS, though we refer to our sample as “WISE” because using that survey is crucial to give our sample greater depth. To cull a uniform, extragalactic sample of sources, we adopt the following color cuts:

- $W1 < 15.2$,
- $J_{2\text{mass}} < 16.5$,
- $W1 - J_{2\text{mass}} < -1.7$

These cuts, and in particular the cross-survey $W1 - J_{2\text{mass}}$ cut, ensure two highly desirable properties of the selected sample: sufficient depth and spatial uniformity. Note also

¹Two Micron All Sky Survey [Skrutskie et al., 2006] Point Source Catalog

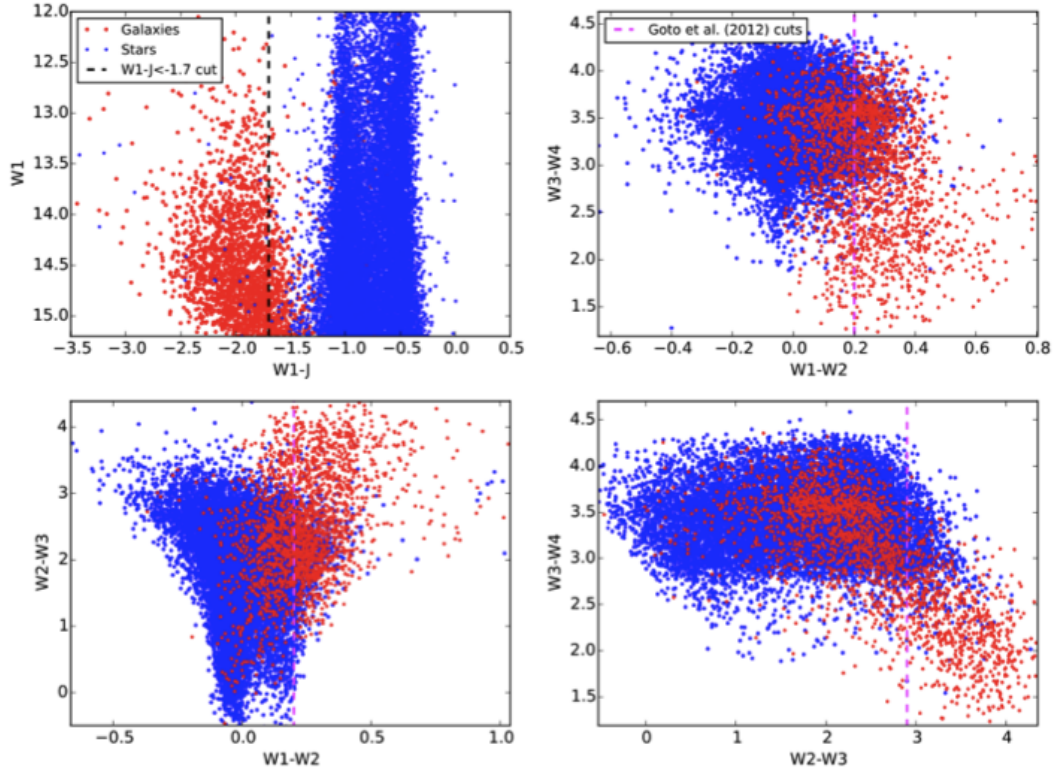


Figure 2.1: Stars (blue) and galaxies (red) separation based on WISE mag bands (W1, W2, W3 and W4) and 2MASS mag band (J): It shows obvious usage of 2MASS J band. Using only WISE bands is not enough to separate galaxies and stars. The pink dotted line is the cut used in previous study Goto et al. (2012). The black dotted line is our newly selected cut. Adapted from [Kovács & Szapudi, 2013].

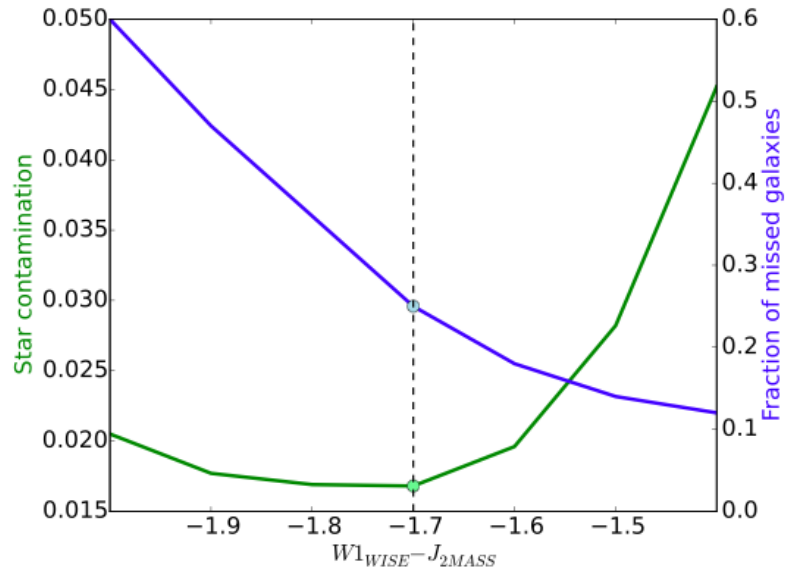


Figure 2.2: Star contamination (green) and completeness (blue) depending on the color cut ($W1-J$) in galaxies of WISE combined with 2MASS data. Adapted from [Kovács & Szapudi, 2013].

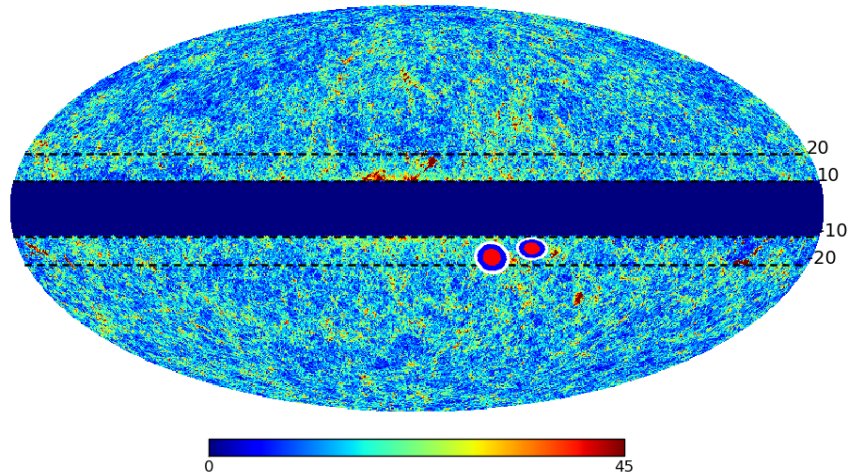


Figure 2.3: Map of WISE-2MASS sources that we used with 10 degree Galactic cut (before masking out the contaminated region with the WMAP dust mask). The criteria are described in the text. The map shown is a Mollweide projection in Galactic coordinates with counts binned in pixels of about 0.5° on a side (HEALPix resolution $N_{\text{SIDE}} = 128$). The two elliptical sets of contours represent the measured dipole direction when we applied a 10° (left) and 20° (right) Galactic cut, respectively (that is, with $|b| < 10^\circ$ and $|b| < 20^\circ$). The red, blue, and white colors in those contours represent the 68%, 95%, and 99% confidence regions for the direction.

that the 2MASS Point Source Catalog contains many more objects than the previously-used but shallower 2MASS Extended Source Catalog. With the benefit of WISE colors the cuts listed above ensure a robust selection of galaxies (~ 2 millions) in 2MASS-PSC that are not in the XSC. This is a huge improvement compared to the previous 2MASS map which was generated from XSC only with ~ 0.4 million galaxies [Gibelyou & Huterer, 2012]. We have measured dipolar modulations with different sets of galaxies by changing this color-cut ($W1 - J_{2\text{mass}}$), but the results did not vary meaningfully.

Note that the first two criteria simply remove the faintest objects in the respective band. To account for the effects of extinction by dust, we correct the magnitudes for these two cuts using the SFD [Schlegel, Finkbeiner & Davis, 1998] map². The third criterion above represents the color cut that serves to separate galaxies from stars and we optimized for low contamination and high completeness at the same time, as shown in Fig.2.2. The detailed

²http://lambda.gsfc.nasa.gov/product/foreground/fg_sfd_get.cfm

analysis on the data selection was described in this paper [Kovács & Szapudi \[2013\]](#); the resulting WISE map is shown in Fig. 2.3.

Unlike the previous studies that used WISE for cosmological tests [[Ferraro, Sherwin & Spergel, 2014](#); [Kovács et al., 2013](#)], our map does not show obvious contamination in regions affected by the appearance of the Moon. Therefore, we do not need to make further (and typically severe) cuts that remove these regions. We do use the WMAP dust map [[Bennett et al., 2013](#)] to mask out the pixels with remaining contamination; these mostly fall within $\pm 15^\circ$ Galactic latitude. In addition, we cut out all pixels with $E(B - V) > 0.5$ from the SFD map (most of these have already been excluded by the WMAP dust map). We also checked for any unusual gradients with Galactic latitude, especially around the Galactic plane, due to contamination from stars. These tests were consistent with zero gradient.

In our analysis, there are of order 2 million galaxies. Because WISE is a photometric dataset, we do not have redshift information for individual sources. We can determine the redshift distribution of our objects by matching the WISE objects to spectroscopic sources. We used the Galaxy and Mass Assembly (GAMA) spectroscopic dataset Data Release 2 [[Driver & Gama Team, 2008](#)] to find sources in the WISE dataset that are within the radius of $3''$ around GAMA sources. For multiply matched sources – 0.15% out of the total 8,493 matched sources – we took the average of their redshifts to determine the redshift of WISE source with multiple matches. The GAMA survey has three observational regions 48 sq. deg. each and down to r-band magnitude limit 19.8. In the 144 sq. deg. overlapping region on the sky, the matching rate is 96.9% which means that 96.9% of WISE sources are matched with GAMA sources. This matching rate was similar for all three distant regions and it suggests the reliability of our matching and redshift estimation. The redshift distribution of matched objects, $N(z)$, is shown in Figure 2.4; the mean is $\bar{z} = 0.139$. We use a smooth fit to the full distribution to obtain our theoretical expectation for the local-structure dipole below.

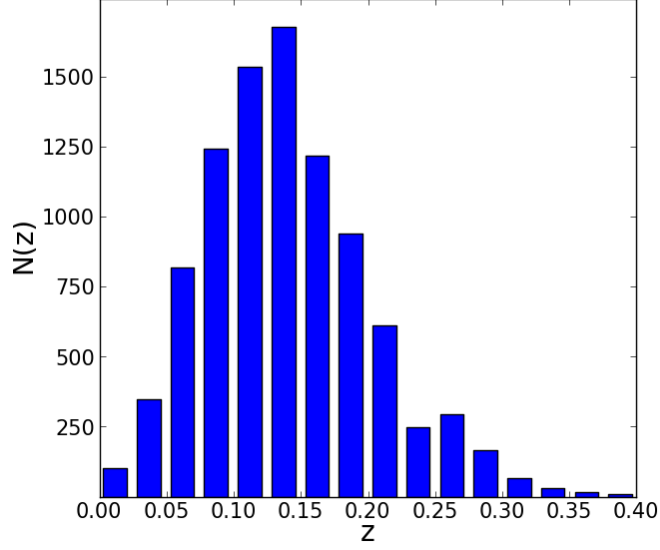


Figure 2.4: Number counts of WISE sources as a function of redshift. We obtain redshift information by matching WISE sources to those from the GAMA DR2 catalog. 96.9% of WISE sources are found in GAMA.

2.3 Methodology

2.3.1 Dipole estimator

A robust and easy-to-implement dipole estimator was first suggested by [Hirata \[2009\]](#), who measured hemispherical anomalies of quasars, and later adopted by [Gibelyou & Huterer \[2012\]](#) to measure the dipole in a variety of LSS surveys. The number of sources in direction $\hat{\mathbf{n}}$ can be written as

$$N(\hat{\mathbf{n}}) = [1 + A \hat{\mathbf{d}} \cdot \hat{\mathbf{n}}] \bar{N} + \epsilon(\hat{\mathbf{n}}) \quad (2.1)$$

where A and $\hat{\mathbf{d}}$ are the amplitude and direction of the dipole, and ϵ is noise. One can further write the contribution from fluctuations as a mean offset [[Hirata, 2009](#)].

$$\delta N / \bar{N} = A \hat{\mathbf{d}} \cdot \hat{\mathbf{n}} + \sum_i k_i t_i(\hat{\mathbf{n}}) + C, \quad (2.2)$$

where the last two terms correspond to $\epsilon(\hat{\mathbf{n}})$ from Eq. (2.1) divided by \bar{N} . Here $t_i(\hat{\mathbf{n}})$ represent the systematics maps, while the coefficients k_i give the amplitudes of the contributions of these systematics to the observed density field. The presence of the monopole term, C , allows us to account for covariance between the monopole and other estimated parameters, especially covariance between the monopole and any systematic templates. The best linear unbiased estimator of the combination (\mathbf{d}, k_i, C) , with corresponding errors, is obtained as follows. First, we rewrite the above equation as

$$\delta N/N = \mathbf{x} \cdot \mathbf{T}(\hat{\mathbf{n}}), \quad (2.3)$$

where

$$\mathbf{x} = (d_x, d_y, d_z, k_1, \dots, k_N, C), \quad (2.4)$$

$$\mathbf{T}(\hat{\mathbf{n}}) = (n_x, n_y, n_z, t_1(\hat{n}), \dots, t_N(\hat{n}), 1), \quad (2.5)$$

$$n_x^2 + n_y^2 + n_z^2 = 1. \quad (2.6)$$

The best linear unbiased estimator of \mathbf{x} is

$$\hat{\mathbf{x}} = F^{-1}g, \quad (2.7)$$

where the components of the vector g are

$$g_i = \int T_i(\hat{n}) \delta N^\Omega(\hat{n}) d^2 \hat{n}, \quad (2.8)$$

and the Fisher matrix F is given by

$$F_{ij} = \bar{N}^\Omega \int T_i(\hat{n}) T_j(\hat{n}) d^2 \hat{n}, \quad (2.9)$$

where $N^\Omega \equiv dN/d\Omega$ is the number of galaxies per steradian (Ω is a solid angle). The integrals from which the vector g and the Fisher matrix F are calculated are discretized in our survey. We adopt a HEALPix [Górski et al. \[2005\]](#) pixelization with `NSIDE=128`, so that each pixel corresponds to about half a degree on a side and contains roughly 14 sources.

The formalism above returns the best-fit dipole components (first three elements of the vector \mathbf{x}), together with their covariance (inverse of the corresponding Fisher matrix). We are however most interested in the likelihood of the amplitude of the dipole:

$$A = (d_x^2 + d_y^2 + d_z^2)^{1/2}. \quad (2.10)$$

We can construct a marginalized likelihood function for the amplitude A [[Hirata, 2009](#)]:

$$\mathcal{L}(A) \propto \int \exp \left[-\frac{1}{2} (A\hat{\mathbf{n}} - \mathbf{d}_{\text{best}}) \text{Cov}^{-1} (A\hat{\mathbf{n}} - \mathbf{d}_{\text{best}}) \right] d^2\hat{\mathbf{n}}, \quad (2.11)$$

where $d^2\hat{\mathbf{n}}$ indicates integration over all possible directions on the sphere and \mathbf{d}_{best} is the best-fit dipole vector calculated using Eq. (2.7). Thus we readily obtain a full likelihood for the amplitude. In our results, we quote the 68% region around the best-fit amplitude.

2.3.2 Foreground Templates and Estimator Validation

Despite our carefully chosen magnitude and color cuts, it is likely that there is some star contamination to our extragalactic source map. Moreover, on a cut sky, the dipole is not completely decoupled from the monopole, quadrupole, and other multipoles, and hence we need to marginalize over some of them in order to get correct results. We therefore include several templates – maps $t_i(\hat{\mathbf{n}})$ in the parlance of Eq. (2.2) – with amplitudes k_i over which we marginalize:

- To deal with the remaining star contamination, we add a star map as a template.

The star map was generated based on the Tycho 2 catalog [Høg et al., 2000], as suggested in Kovács et al. [2013]. The inclusion of this template affects the measured dipole negligibly, reinforcing our confidence that star contamination does not affect the result.

- To account for the other multipoles, we add the monopole (corresponding to the constant C in Eq. (2.2) with no spatial dependence), as well as the quadrupole and octopole that include 5 and 7 extra parameters. We therefore marginalize over these 13 parameters in addition to the amplitude of the star map. We experimented with marginalization over a few more ($\ell \geq 4$) multipoles, but for small Galactic cuts ($b_{\text{cut}} \lesssim 15^\circ$), the shift in the dipole direction and magnitude were small.

We validated our estimator by running simulations with an input dipole of a given amplitude assuming various sky cuts and marginalizing over templates. We verified that the input dipole is recovered within the error bars.

2.3.3 Theoretical expectation of local-structural dipolar modulation

We calculate the theoretical expectation for the local-structure dipole using standard methods (see e.g. Sec. 2.2 of Gibelyou & Huterer [2012]). We calculate the angular power spectrum of large-scale structure for the given source distribution $N(z)$, and evaluate it at the dipole (C_ℓ at $\ell = 1$); this calculation does not assume the Limber approximation since the latter is inaccurate at these very large scales. The exact equation of angular power spectrum is

$$C_\ell = 4\pi \int d \ln k \Delta^2(k, z = 0) I^2(k), \quad (2.12)$$

where the dimensionless power spectrum is defined as $\Delta^2(k, z = 0) \equiv k^3 P(k)/(2\pi^2)$ and the intensity function $I(k)$ is given as

$$I(k) \equiv \int dz W(z) D(z) j_\ell(k\chi(z)). \quad (2.13)$$

Here $\chi(z)$ is the radial distance and $D(z)$ is a linear growth function of density fluctuation satisfying $\delta(z) = D(z)\delta(0)$. The window function $W(z) = b(z)n(z)$ is a function of the bias $b(z)$ and the probability distribution of galaxies $n(z)$.

The amplitude of dipolar modulation is then given as

$$A_{\text{theory}} = (9C_1/(4\pi))^{1/2}. \quad (2.14)$$

The theory error arising from cosmic variance for C_ℓ is

$$\delta C_\ell = \sqrt{\frac{2}{(2\ell + 1)f_{\text{sky}}}} C_\ell \quad (2.15)$$

where f_{sky} is the fraction of sky covered by used data. Therefore, the error of amplitude of dipole A which is related to C_1 is

$$\frac{\delta A_{\text{theory}}}{A_{\text{theory}}} = \frac{1}{2} \sqrt{\frac{2}{(2\ell + 1)f_{\text{sky}}}} = (6f_{\text{sky}})^{-1/2}, \quad (2.16)$$

Evaluating the theoretically expected dipole for the source distribution shown in Fig. 2.4, we get

$$A_{\text{theory}} = (0.0233 \pm 0.0094 f_{\text{sky}}^{-1/2}) \times \left(\frac{b}{1.41} \right) \quad (2.17)$$

Here we make explicit the dependence of the cosmic variance error on the fraction of the sky covered (f_{sky}), and also on the bias of WISE sources (bias parameter: b). To obtain the latter, we followed Kovács & Szapudi [2013], and estimated the bias of the galaxy catalog using SPICE [Szapudi, Prunet & et al., 2001] and the Python CosmoPy³ package. We note that the estimation of the bias is particularly sensitive to σ_8 because they both act to renormalize the angular power spectrum, and in linear theory $C_\ell^{gg} \propto (b\sigma_8)^2$. We fix $\sigma_8 = 0.8$ in our measurements, finding $b = 1.41 \pm 0.07$. This value is comparable to earlier findings [Rassat, Land & et al., 2007] that measured a value of $b = 1.40 \pm 0.03$ for

³<http://www.ifa.hawaii.edu/cosmopy/>

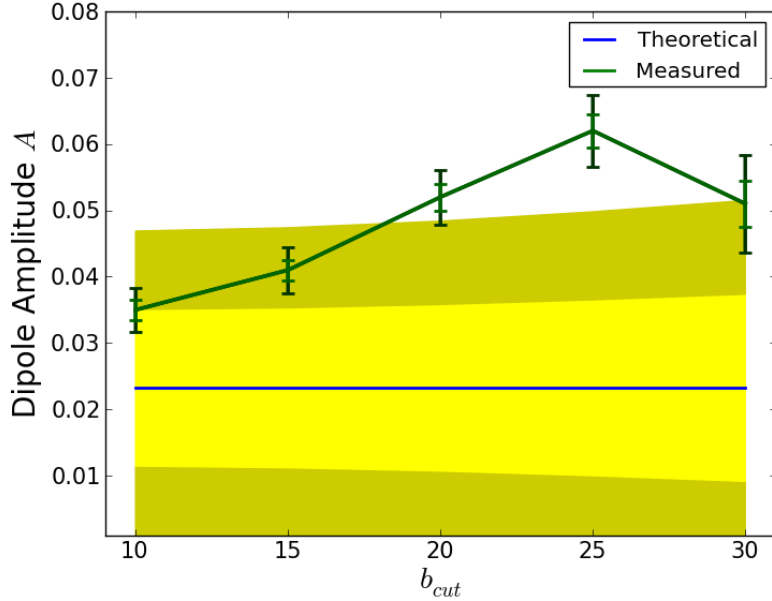


Figure 2.5: Theoretical prediction for the dipole amplitude (horizontal blue line), together with the measured values in WISE (green points). The two sets of error bars on the measurements correspond to 68% and 95% confidence; they have been calculated from the full likelihood in Eq. (2.11) and are rather symmetric around the maximum-likelihood value. The two large horizontal bands around the theory prediction correspond to 1- and 2-sigma cosmic variance error.

a 2MASS selected galaxy sample.

2.4 Results

Our measurements of the dipole’s amplitude and direction, as a function of the (isolatitude) Galactic cut, are presented in Table 2.1 and shown in Figure 2.5. The best-fit direction of the dipole is also shown in Fig. 2.3 for the 10° and 20° Galactic cut, the two cases roughly illustrating the dependence of the direction on the Galactic cut.

We first note a reasonably good consistency between the recovered directions, despite the fact that the number of sources decreases by a factor of ~ 1.4 as we increase the Galactic cut in the range shown. We also note that the overall amplitude is roughly 1.5 - 2.7 times larger than the theoretically expected one, and is roughly 1-2 σ high, where σ corresponds

b_{cut}	f_{sky}	A_{WISE}	A_{theory}	$\hat{\mathbf{d}}(l^\circ, b^\circ)$
10°	0.65	0.035 ± 0.002	0.023 ± 0.012	$(326 \pm 3, -17 \pm 2)$
15°	0.62	0.042 ± 0.002	0.023 ± 0.012	$(316 \pm 3, -15 \pm 2)$
20°	0.57	0.052 ± 0.002	0.023 ± 0.012	$(308 \pm 4, -14 \pm 2)$
25°	0.51	0.062 ± 0.003	0.023 ± 0.013	$(315 \pm 6, -12 \pm 2)$
30°	0.45	0.051 ± 0.004	0.023 ± 0.014	$(335 \pm 6, -18 \pm 3)$

Table 2.1: Measurements of the dipole amplitude in WISE for various Galactic cuts (b_{cut}) corresponding to fractions of the sky covered (f_{sky}). In all cases we marginalized over several foreground templates, as described in the text. The full likelihood for the amplitude A_{WISE} is well approximated by a Gaussian whose mode and standard deviation we quote here. We also show the theoretical expectation A_{theory} due to the local-structure dipole, together with the corresponding cosmic variance given a bias $b = 1.41$.

to cosmic variance since the measurement error is much smaller (see Table 2.1). Finally, we note that while the dipole amplitude does vary with b_{cut} more than its typical measurement errors, it is overall consistent at $A_{\text{WISE}} \simeq 0.04\text{-}0.05$, which is rather robustly stable given the large decrease of the number of sources with increasing Galactic cut.

It is interesting to note that 2MASS Extended Source Catalog data, as analyzed in [Gibelyou & Huterer \[2012\]](#) (redshift $0 < z < 0.2$, $N = 3.8 \times 10^5$), give $A_{2\text{MASS}} = 0.104 \pm 0.004$, $(l, b) = (268.4^\circ, 0.0^\circ)$ – the direction is not far from ours, but the amplitude is larger because 2MASS data used in this previous work is shallower than our WISE-2MASS sample. Relative to [Gibelyou & Huterer \[2012\]](#), we have therefore made progress by pushing down a factor of 2.5 in the dipole amplitude. This is a welcome development toward being able to probe the kinematic dipole due to our motion relative to the overall LSS rest frame, which will require reaching the level $A \sim 10^{-3}$, and therefore a deeper survey (or a deeper sample of WISE sources).

2.5 Conclusions

We measured the clustering dipole in the WISE survey, using a carefully culled sample that contains 2 million extragalactic sources with a known redshift distribution. The direction of the dipole is $\simeq (310^\circ \pm 5, -15^\circ \pm 2)$. The amplitude of the measured dipole is $A \simeq 0.05 \pm 0.01$, where we quote the central value corresponding to the 20° cut case and error that shows the dispersion of central values for $15^\circ \leq b_{\text{cut}} \leq 25^\circ$. While the amplitude is therefore roughly twice as large as the theoretical expectation given in Eq. (3.2), the large cosmic variance on the theoretical prediction calculated in Sec. 2.3.3 makes the measured amplitude $\sim 2.5\text{-}\sigma$ high — in tension with theory but not sufficiently statistically significant to claim departures from the standard Λ CDM prediction.

Taking for the moment the excess dipole measured relative to theoretical expectation at face value, we can ask: what could explain it? The systematics, while an obvious first suspect, are not necessarily at fault given the rather extensive care we took to account for them: we carefully culled the dataset by imposing cuts based on WISE and 2MASS magnitudes; we included cuts based on Galactic latitude and on the WMAP dust map, and we further marginalized over a carefully derived star-map template as well as templates corresponding to the quadrupole and octopole.

Another possibility is that the excess signal is cosmological. For example, a large void might generate the excess observed here [Rubart, Bacon & Schwarz, 2014]. Such a void was incidentally just detected in the analysis of the WISE data itself [Finelli et al., 2014; Szapudi et al., 2015]. At this time it is too early to tell whether the WISE void is contributing significantly to the excess dipole that we measured, though a rough comparison with numbers in Rubart, Bacon & Schwarz [2014] appears to indicate that it is not.

It is also interesting to note that Planck found a best-fit modulation with both amplitude and direction roughly (within $\sim 3\sigma$ of their errors) in agreement with ours [Ade et al., 2013]: $A_{\text{Planck}} = 0.078 \pm 0.021$, $(l, b) = (227^\circ, -15^\circ) \pm 19^\circ$. It is not clear at this time what, if any, significance to assign to the comparable-looking modulations in WISE and

Planck since their sources are at vastly different redshifts ($z \sim 0.15$ and ~ 1000), and the agreement in amplitude and direction is only approximate. Finally, the direction we find is *also* close to the peculiar-velocity bulk-flow directions found using type Ia supernovae [Dai, Kinney & Stojkovic, 2011; Kalus et al., 2013; Rathaus, Kovetz & Itzhaki, 2013], galaxies [Feldman, Watkins & Hudson, 2010; Ma, Gordon & Feldman, 2011; Ma & Pan, 2014; Turnbull et al., 2012], and the kinetic Sunyaev-Zeldovich effect [Lavaux, Afshordi & Hudson, 2013]. While the agreement between the directions is suggestive, it is not immediately clear how our WISE dipole is related to these. For example, interpreting the excess dipole amplitude $\delta A \sim 0.03$ as a bulk motion is clearly out of the question, since it would correspond to a huge velocity of $v \simeq 0.015c = 4500$ km/s, an order of magnitude larger than what typical bulk-motion measurements indicate.

With recent measurements of the cross-correlation of its sources with the CMB and the detection of a large underdense void, WISE is finally making major contributions to cosmology. Its nearly all-sky coverage is a huge asset and gives the survey a big advantage on that front over most other LSS surveys. In this paper we have taken another step in testing fundamental cosmology with WISE by measuring the clustering dipole in the distribution of its extragalactic sources. We look forward to further investigations of this result, especially in conjunction with other related findings in the CMB and LSS.

CHAPTER 3

Kinematic Dipole Detection with Galaxy Surveys

3.1 Introduction

Measurements of the motion of our Solar System through the cosmic microwave background (CMB) rest frame represent one of the early successes of precision cosmology. This so-called kinematic dipole corresponds to a velocity of $(369 \pm 0.9) \text{ km s}^{-1}$ in the direction $(l, b) = (263.99^\circ \pm 0.14^\circ, 48.26^\circ \pm 0.03^\circ)$ [Hinshaw et al., 2009]. The kinematic dipole has even been detected (though not *as* precisely measured) by observing the relativistic aberration in the CMB anisotropy that it causes, which is detected via the coupling of high CMB multipoles in Planck [Aghanim et al., 2014]. Table 3.1 shows relative motions of different scale systems where the Earth resides in. The vector sum of these relative motions contributes to the measured CMB kinematic dipole. The CMB dipole was measured by already subtracting out the contribution from the Earth’s motion around the Sun so the CMB dipole only corresponds to our Solar System’s motion relative to CMB rest frame.

Independently, the past few decades have seen significant progress in measuring the dipole in the distribution of extragalactic sources. The contribution of our motion through the large-scale structure (LSS) rest frame – the kinematic dipole – also leads to relativistic aberration, this time of galaxies or other observed LSS sources. We define the dipole

Motion	Approximate speed (kms ⁻¹)	Direction (l°, b°)
Earth around Sun	~ 30	Annually varying
Solar system w.r.t. Local Group [Courteau & Van Den Bergh, 1999]	~ 306	(99, -4)
Local Group w.r.t. CMB [Maller et al., 2003]	~ 622	(272, 28)
Overall CMB kinematic dipole [Hinshaw et al., 2009]	~ 370	(264, 48)

Table 3.1: Relative motions of systems contributing to kinematic dipole.

amplitude via the amount of its “bunching up” of galaxies in the direction of the dipole

$$\frac{\delta N(\hat{\mathbf{n}})}{\bar{N}} = A \hat{\mathbf{d}} \cdot \hat{\mathbf{n}} + \epsilon(\hat{\mathbf{n}}), \quad (3.1)$$

where N is the galaxy number in an arbitrary direction $\hat{\mathbf{n}}$, $\hat{\mathbf{d}}$ is the dipole direction, and ϵ is random noise. The dipole amplitude A is approximately (but not exactly) equal to our velocity through the LSS rest frame in units of the speed of light; the precise relation is given in the following section.

However, the dominant contribution to the LSS dipole is typically not our motion through the LSS rest frame, but rather the fluctuations in structure due to the finite depth of the survey. The dipole component of the latter – the so-called “local-structure dipole” in the nomenclature of Gibelyou & Huterer [2012] – has amplitude $A \sim 0.1$ for shallow surveys extending to $z_{\max} \sim 0.1$, but is significantly smaller for deeper surveys. The local-structure dipole is the dominant signal at multipole $\ell = 1$ in all extant LSS surveys. It has been measured and reported either explicitly [Alonso et al., 2015; Appleby & Shafieloo, 2014; Baleisis et al., 1998; Blake & Wall, 2002; Fernández-Cobos et al., 2014; Gibelyou & Huterer, 2012; Hirata, 2009; Rubart & Schwarz, 2013; Yoon et al., 2014], or as part of the angular power spectrum measurements. No LSS survey completed to date therefore had a chance to separate the small kinematic signal from the larger local-structure dipole contamination due to insufficient depth and sky coverage. This will change drastically with

the new generation of wide, deep surveys.

Standard theory based on the adiabatic initial perturbations predicts that the kinematic dipole measured by the LSS should agree with the one measured by the CMB. Detection of an anomalously large (or small) dipole or the disagreement of its direction from that of the CMB dipole could indicate new physics: for example, the presence of superhorizon fluctuations in the presence of isocurvature fluctuations [Erickcek, Carroll & Kamionkowski, 2008; Itoh, Yahata & Takada, 2010; Turner, 1991; Zibin & Scott, 2008]. Clearly, a kinematic dipole detection and measurement represent an important and fundamental consistency test of the standard cosmological model.

3.2 Methodology

3.2.1 Theoretical signal

The expected LSS kinematic dipole signal amplitude is given by [Burles & Rappaport, 2006; Itoh, Yahata & Takada, 2010]

$$A = 2\tilde{\beta} = 2[1 + 1.25x(1 - p)]\beta \quad (3.2)$$

where $\beta = v/c = 0.00123$ (assuming the CMB dipole). The contribution 2β comes from relativistic aberration, while the correction $[1 + 1.25x(1 - p)]$ corresponds to the Doppler effect; here x is the faint-end slope of the source counts, $x \equiv d \log_{10}[n(m < m_{\text{lim}})]/dm_{\text{lim}}$, and p is the logarithmic slope of the intrinsic flux density power-law, $S_{\text{rest}}(\nu) \propto \nu^p$.

Clearly, the parameters x and p depend on the population of sources selected by the survey, and on any population drifts as a function of magnitude. We now estimate these parameters – note also that we only need the quantity A to set our fiducial model, so very precise values of the population parameters are not crucial for this paper. Marchesini et al. [2012] find that the faint end of the V-band galaxy luminosity function does not vary much

over the redshift range $0.4 \leq z \leq 4$ and is equal to, in our notation, $x = 0.11 \pm 0.02$. Moreover, for optical sources the flux density slope p varies significantly with the age of the source, but in the infrared it is more consistent, with measurements indicating $p \sim 0$ [Mo, van den Bosch & White, 2010; White & Majumdar, 2004]. Here we adopt $p = 0$. Applying all these values to Eq. (3.2), we get

$$A \simeq 0.0028 \quad (\text{expectation from CMB}). \quad (3.3)$$

While the actual value of the kinematic dipole is of course unknown prior to the measurement, standard cosmology theory predicts it takes this value, plus or minus $O(20\%)$ changes depending on the source population selected. We adopt Eq. (3.3) as the fiducial amplitude.

The fiducial direction we adopt is the one of the best-fit CMB dipole, $(l, b) = (263.99^\circ, 48.26^\circ)$. Note, however, that the results may vary depending on the relative orientation between the actual dipole direction and the coverage of the observed sky. Finally, note that bias (of the galaxy clustering relative to the dark matter field) enters into the contamination of the kinematic dipole measurements, but not the signal. The former quantity – the local-structure dipole – is linearly proportional to the bias b . Therefore, the bigger the bias, the more contamination the local-structure dipole provides for measurements of the kinematic effect. In this work we assume bias of $b = 1$. Note that the kinematic signal itself, being due to our velocity through the LSS rest frame, is independent of bias.

3.2.2 Statistical error

Rewriting Eq. (3.1) somewhat, the modulation in the number of sources is given at each direction $\hat{\mathbf{n}}$ can be written as

$$\frac{\delta N(\hat{\mathbf{n}})}{\bar{N}} = \mathbf{x} \cdot \mathbf{T}(\hat{\mathbf{n}}) + \epsilon(\hat{\mathbf{n}}), \quad (3.4)$$

where $\mathbf{x} = (d_x, d_y, d_z, k_1, \dots, k_M)$ is the vector of the three dipole component coefficients in the three spatial coordinates, plus coefficients corresponding to other multipoles (one for the monopole, five for that many components of the quadrupole, etc), as well as any desired systematic templates. The vector $\mathbf{T}(\hat{\mathbf{n}}) = (n_x, n_y, n_z, t_1(\hat{\mathbf{n}}), \dots, t_M(\hat{\mathbf{n}}))$, contains the three dipole unit vectors (with $n_x^2 + n_y^2 + n_z^2 = 1$), plus M additional spatial patterns for all templates included. Note that the choice of the fiducial values of the non-dipole template coefficients k_i is arbitrary, since we will fully marginalize over each of these, effectively allowing k_i to vary from zero to plus infinity. The optimal estimate of \mathbf{x} is given by $\hat{\mathbf{x}} = F^{-1}g$ [Hirata, 2009], where the components of the vector g are

$$g_i = \int T_i(\hat{\mathbf{n}}) \delta N^\Omega(\hat{\mathbf{n}}) d^2\hat{\mathbf{n}} \quad (3.5)$$

and the best-fit dipole \mathbf{d}_{best} is given by the first three elements of \mathbf{x} . Here the Fisher matrix F is given by

$$F_{ij} = \bar{N}^\Omega \int T_i(\hat{\mathbf{n}}) T_j(\hat{\mathbf{n}}) d^2\hat{\mathbf{n}}, \quad (3.6)$$

where $N^\Omega \equiv dN/d\Omega$ is the number of galaxies per steradian and Ω is a solid angle. Note that the Fisher information is proportional to the number of sources, and unrelated to the depth of the survey. It is therefore the number of sources, together with the sky cut (not just the fraction of the sky observed f_{sky} but also the shape of the observed region relative to the multipoles that need to be extracted) that fully determines the statistical error in the various templates including the dipole.

The Fisher matrix contains information about how well the three Cartesian dipole components, as well as the multipole moments of all other components, can be measured in a given survey. Our parameter space has a total of $M_{\text{par}} = (\ell_{\text{max}} + 1)^2$ parameters, where ℓ_{max} is the maximum multipole included to generate the templates (see below for more on the choice of ℓ_{max}). With this Fisher matrix in hand, we then marginalize over the $M \equiv M_{\text{par}} - 3$ non-dipole parameters, using standard Fisher techniques, to get the 3×3

Fisher matrix describing the final inverse covariance matrix for the dipole components. Finally, we perform a basis change, converting from Cartesian coordinates $\{d_x, d_y, d_z\}$ to spherical coordinates $\{A, \theta, \phi\}$ (where A is the amplitude of dipole), by using a Jacobian transformation to obtain the desired 3×3 Fisher matrix in the latter space, $F_{(3 \times 3)}^{\text{marg}}$. (See detailed derivation in Appendix A.1.)

The forecasted error on A is then given in terms of this matrix as

$$\sigma(A) = \sqrt{[(F_{(3 \times 3)}^{\text{marg}})^{-1}]_{AA}}. \quad (3.7)$$

In a realistic survey with partial-sky coverage, the presence of other multipoles (monopole, quadrupole, etc) will be degenerate with the dipole, degrading the accuracy in determining the latter. We have extensively tested for this degradation, in particular with respect to how many multipoles need to be kept – that is, what value of ℓ_{max} (and therefore M) to adopt. We explicitly found that the prior information on the “nuisance” C_ℓ , corresponding to how well they can be (and are being) independently measured, is of key value: once the prior information on the C_ℓ – corresponding to cosmic variance plus measurement error – is added, very high multipoles are not degenerate with the dipole. Our tests show that keeping all multipoles out to $\ell_{\text{max}} = 10$ is sufficient for the dipole error to fully converge. (See detailed test results in Appendix A.2.) We also experimented with adding additional individual templates $t_i(\hat{\mathbf{n}})$ corresponding to actual sky systematics and with modified coverage (corresponding to e.g. dust mask around the Galactic plane), but found that these lead to negligible changes in the results. (See Appendix A.3.) Moreover, we envisage a situation where the maps have already been largely cleaned of stars by the judicious choice of color cuts prior to the dipole search analysis. For these two reasons, we choose not to include any additional systematic templates in the analysis.

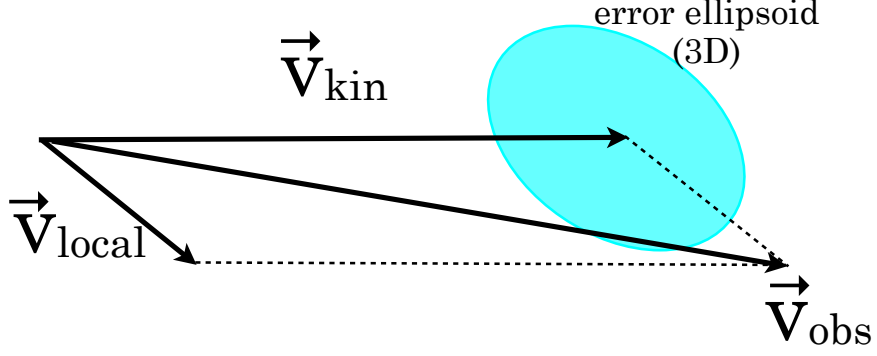


Figure 3.1: Sketch of the problem at hand: we would like to measure the kinematic dipole \vec{v}_{kin} , whose error (represented by a cyan ellipse) can be calculated given the number of extragalactic objects and the sky coverage. The LSS local dipole, \vec{v}_{local} , provides a bias in this measurement. For a survey deep enough (and depending somewhat on the direction of its \vec{v}_{local}), bias in the measurement of \vec{v}_{kin} will be smaller than the statistical error.

3.2.3 Systematic bias

The local-structure dipole $\mathbf{d}_{\text{local}}$ will also provide a contribution to the kinematic signal \mathbf{d}_{kin} that we seek to measure. The observed dipole in any survey will be the sum of the two contributions:

$$\mathbf{d}_{\text{obs}} = \mathbf{d}_{\text{kin}} + \mathbf{d}_{\text{local}}. \quad (3.8)$$

Without any loss of generality, we consider the kinematic dipole as the fiducial signal in the map, whose errors are therefore given by the Fisher matrix worked out above. We consider the local-structure dipole to represent the contaminant whose magnitude, ideally, should be such that the resulting observed dipole \mathbf{d}_{obs} is still within the error ellipsoid around the kinematic dipole direction and amplitude. This is illustrated in Fig. 3.1. It is possible to measure the kinematic dipole with relatively small contamination from the local-structure dipole because the amplitude of local-structure dipole drops drastically as we observe deeper. (See Fig. 3.2.)

We now quantify the systematic bias, due to the local structure, relative to statistical error in the measurements of the kinematic dipole. First note that we are in possession of the (statistical) inverse covariance matrix for measurements of the kinematic dipole, $F_{(3 \times 3)}^{\text{marg}}$,

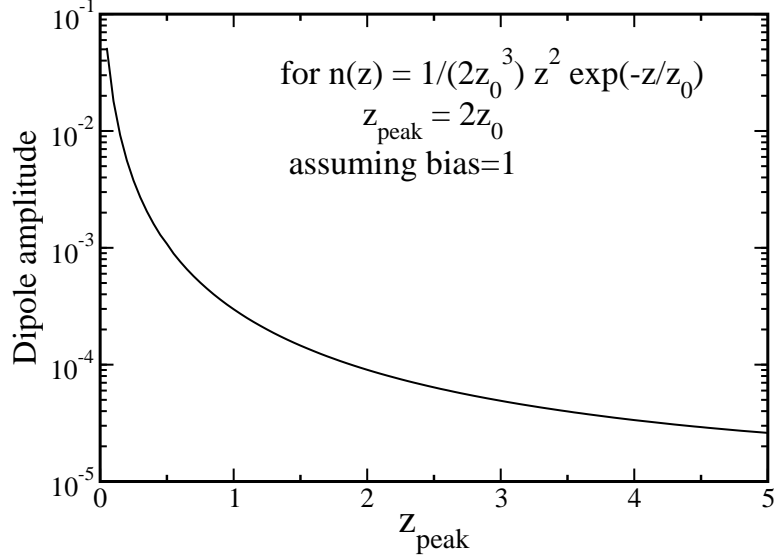


Figure 3.2: The amplitude of local-structure dipole as a function of redshift with a probability distribution $n(z)$.

which is already fully marginalized over other templates. The quantity

$$\begin{aligned}
 \Delta\chi^2(\mathbf{d}_{\text{local}}) &= (\mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{kin}})^T F_{(3 \times 3)}^{\text{marg}} (\mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{kin}}) \\
 &= \mathbf{d}_{\text{local}}^T F_{(3 \times 3)}^{\text{marg}} \mathbf{d}_{\text{local}}
 \end{aligned}
 \tag{3.9}$$

then represents “(bias/error)²” in the kinematic dipole measurement due to the presence of the local-structure contamination. This chi squared depends quadratically on the expected local-structure dipole, and is therefore expected to sharply drop with deeper surveys which have a lower $|\mathbf{d}_{\text{local}}|$, as we find in the next section. With three parameters, requiring 68% confidence level departure implies $\Delta\chi^2 = 3.5$. We therefore require that, for a given survey, the local-structure dipole magnitude and direction are such that the value in Eq. (3.9) is smaller than this value.

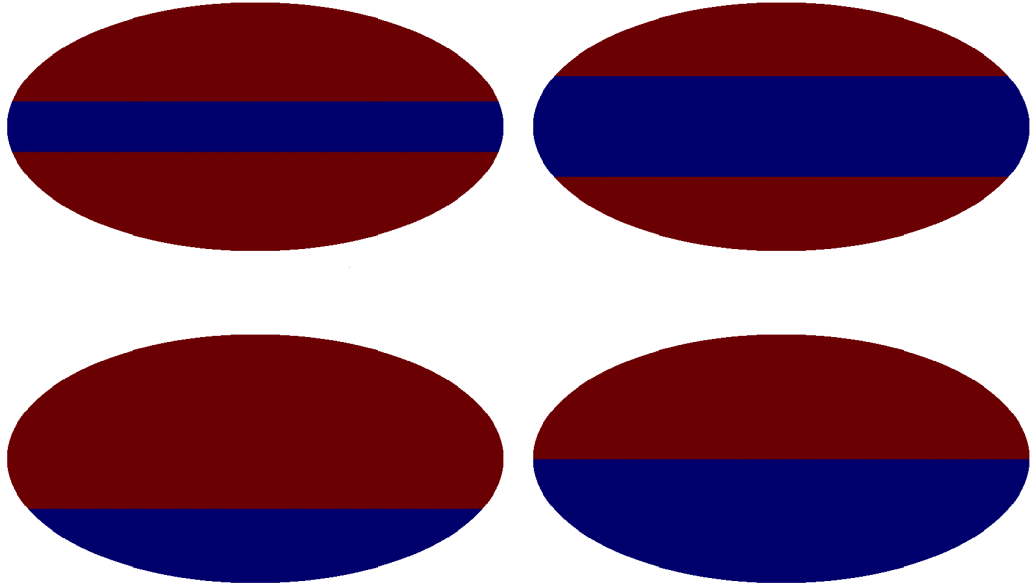


Figure 3.3: Four footprints of survey coverage: galactic cut $\pm 15^\circ$ (above left), galactic cut $\pm 30^\circ$ (above right), cap cut 60° (below left), and cap cut 90° (below left). Blue color represents the excluded area.

3.3 Results

For a fixed sky cut and number of sources in the survey, we first calculate the error on the amplitude of the dipole $\sigma(A)$. In the Fig. 3.4 we show errors as a function of the number of galaxies in a survey. As previously noted, this statistical error does not depend on the depth of the survey, but does depend on both f_{sky} and the shape of the sky coverage. Here we show results for an isolatitude cut around the equator of ± 15 deg, and ± 30 deg and isolatitude cap-shaped cuts of 90 deg (i.e. half the sky removed) and 60 deg (i.e. leaving out a circular region around a pole) as shown in Fig. 3.3. Note that the Galactic ± 15 deg cut and the cap cut of 60 deg both have $f_{\text{sky}} = 0.75$, while the Galactic ± 30 deg and the cap 90 deg cuts both have $f_{\text{sky}} = 0.5$. The results will also depend on the fiducial amplitude of the dipole, and here and throughout we assume the CMB-predicted value of $A = 0.0028$. Even for a fixed f_{sky} of the survey, the cut geometry clearly matters, and the Galactic-cut cases have a smaller error in the dipole amplitude due to symmetrical covering of the two

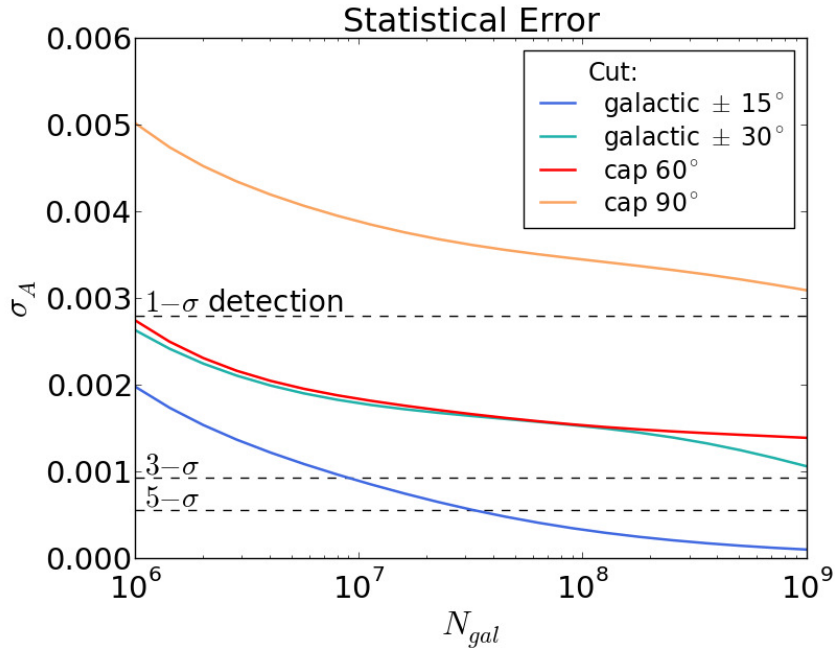


Figure 3.4: Statistical error in the dipole amplitude, marginalized over direction and other multipoles that are coupled to the dipole, as a function of the number of galaxies in a survey. The top horizontal dashed line shows the amplitude expected based on the CMB dipole measurements ($A = 0.0028$), and is the fiducial value in this work. The two dashed horizontal lines show the 3σ and 5σ detection of dipole with the fiducial amplitude.

hemispheres. For Galactic ± 15 deg case, $3\text{-}\sigma$ and $5\text{-}\sigma$ detections are easily achievable, requiring only $N_{gal} = 9 \times 10^6$ and 3×10^7 objects, respectively. Note that if the actual direction of the LSS kinematic dipole deviates from the assumed dipole direction (the CMB direction), the result changes. For a 5-sigma detection and the same ± 15 deg isolatitude cut, a dipole pointing along toward a Galactic pole, which is the best-case scenario, only requires 8 million objects; if instead the dipole points toward the Galactic plane, then 70 million objects are required. (See Fig. 3.5 for checking the dependence on the direction of kinematic dipole.) For the Galactic ± 30 deg cut, the $3\text{-}\sigma$ detection is more challenging since it requires having over $N_{gal} = 10^9$ sources. The cap 60 deg cut mostly follows the trend of the Galactic ± 30 case. Lastly, the 90 deg cap cut cannot detect the signal even at the 1-sigma level and with $N_{gal} = 10^9$. We conclude that dual-hemisphere sky coverage is crucial in the ability of the survey – or a combined collection of surveys – to detect the

kinematic dipole.

Fig. 3.6 shows the systematic bias in the dipole measurement due to the presence of the local-structure contamination, showing the quantity defined in Eq. (3.9). Because $\mathbf{d}_{\text{local}}$ has an a-priori unknown direction and its amplitude changes according to the depth of the survey, the systematic error is a function of direction of $\mathbf{d}_{\text{local}}$ and the depth of the survey. Therefore, we choose to plot $\Delta\chi^2$ averaged over all directions of $\mathbf{d}_{\text{local}}$. (See Fig. 3.7 to check the variation of systematic bias depending on the direction of the local-structure dipole.)

To calculate the amplitude of $\mathbf{d}_{\text{local}}$, we model the radial distribution of objects as $n(z) = z^2/(2z_0^3) \exp(-z/z_0)$ [Huterer, 2002], where the parameter z_0 is related to the median redshift as $z_0 = z_{\text{med}}/2.674$. A deeper survey (larger z_0) has a smaller local-structure dipole. Note that one could additionally cut out low-redshift objects in order to further reduce the contamination from the nearby structures, as well as the star-galaxy confusion. We have tested the case when all $z < 0.5$ objects are removed from the analysis; while helpful, this step is not crucial since the resulting additional benefits in decreased bias are moderate, decreasing $\Delta\chi^2$ for example by a factor of 1.8 for $z_{\text{med}} = 0.75$ relative to the case where no cut has been applied and for $z_{\text{med}} = 1.0$, a factor of 1.4. The number of objects decreases to $\sim 71\%$ and 83% respectively for the two cases. This way of cutting out the low- z data is helpful but the cut-out should be applied carefully by considering the loss in statistical error with fewer objects.

The dashed lines in Fig. 3.6 represent the cases when $z_{\text{med}} = 0.75$ and the solid ones are when $z_{\text{med}} = 1.0$. Since $\Delta\chi^2$ is inversely proportional to the statistical error squared, the best cases in Fig. 3.4 have a larger bias in Fig. 3.6. In particular, the more galaxies the survey has, the more it is susceptible to systematic bias (for a fixed depth and thus $|\mathbf{d}_{\text{local}}|$). For example, a survey with ± 15 deg Galactic cut with 30 million sources can detect the kinematic dipole at $5\text{-}\sigma$, but needs to have a median redshift of at least $z_{\text{med}} = 0.75$ in order for this not to be excessively biased due to local structures.

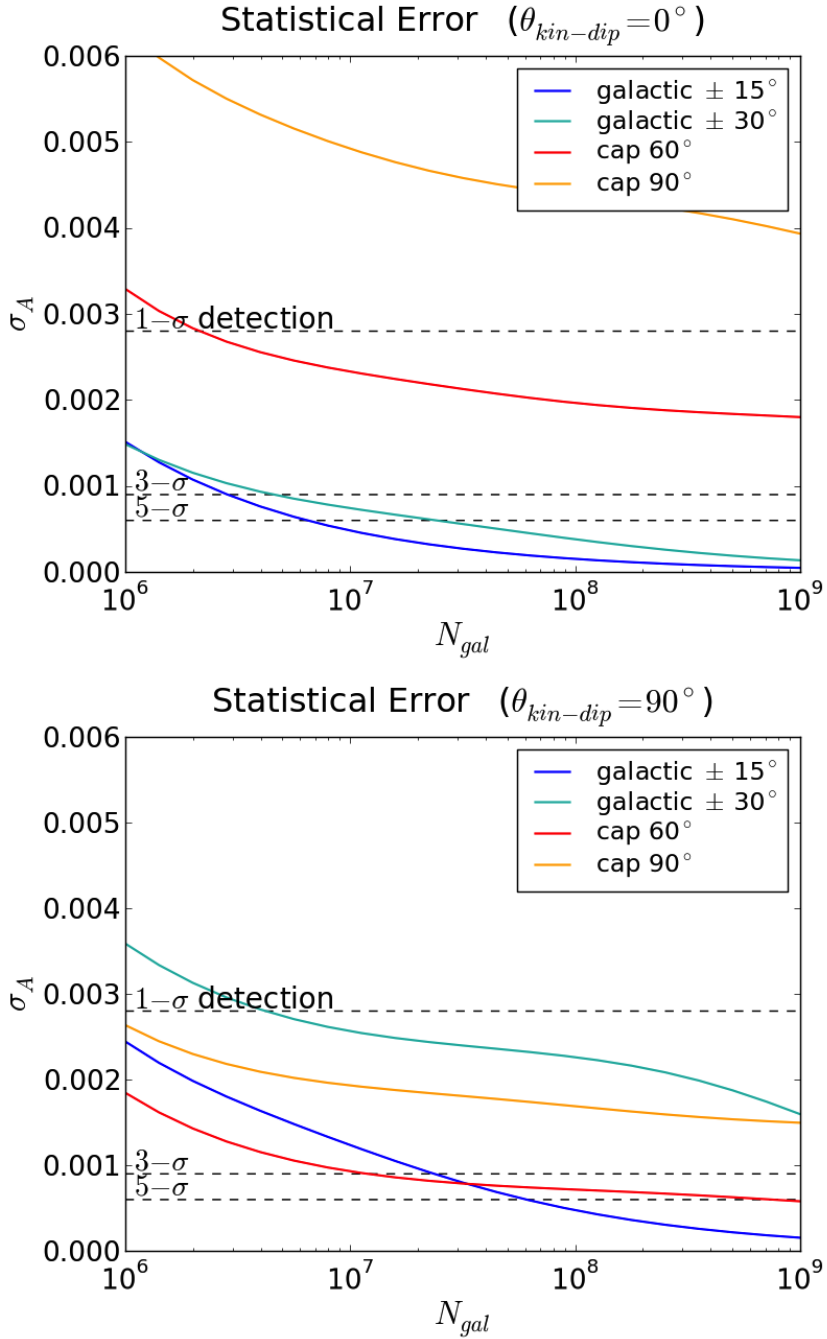


Figure 3.5: Statistical error in the dipole amplitude when $\theta_{kin-dip} = 0^\circ$ and $\theta_{kin-dip} = 90^\circ$: the angle $\theta_{kin-dip}$ is defined as the angle between the galactic north pole and the direction of kinematic dipole. In Fig. 3.4, we assumed the direction of kinematic dipole same as CMB dipole. For galactic cuts, the case when kinematic dipole is pointing toward the galactic north pole is better than the case when kinematic dipole is aligned with the galactic plane because the highest density contrast is captured within the survey area for the former case. For cap cuts, the highest density contrast is captured when it is not pointing toward the galactic north pole so the trend is the opposite to galactic cut scenarios.

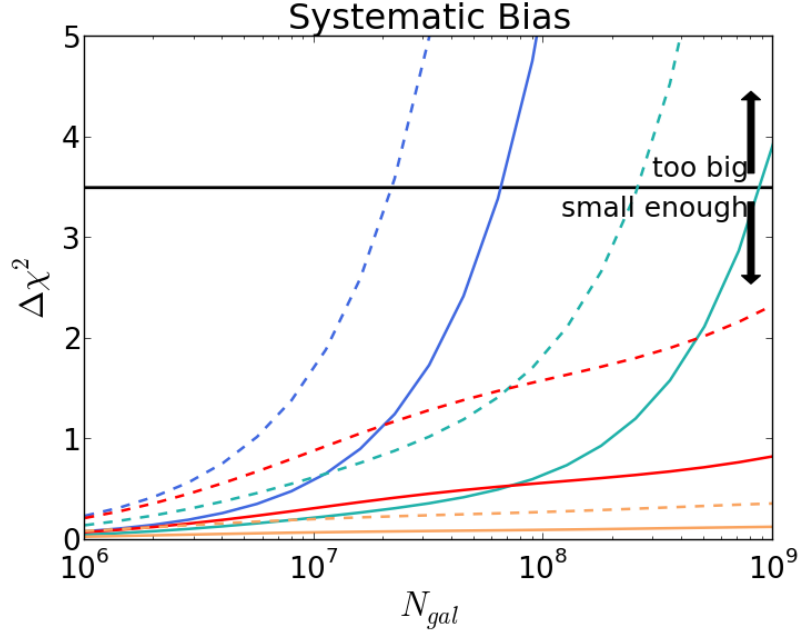


Figure 3.6: $\Delta\chi^2$, defined in Eq. (3.9), corresponding to the bias from the local-structure dipole, as a function of the number of objects N_{gal} . For a fixed amplitude of \mathbf{d}_{local} the $\Delta\chi^2$ still depends on the direction of this vector; here we show the value averaged over all directions of \mathbf{d}_{local} . Solid lines show cases when the median galaxy redshift is $z_{med} = 1.0$, while dashed lines are for $z_{med} = 0.75$. The legend colors are the same in fig.3.4.

On the whole, Fig. 3.4 indicates that the convincing detection of the kinematic dipole expected given the CMB measurements is entirely within reach of future surveys, as long as those surveys have good coverage over both hemispheres and, given the source density, are deep enough not to be biased by the local-structure dipole. All requirements can be straightforwardly satisfied by surveys like some combination of LSST [Ivezic et al., 2008], Euclid [Laureijs et al., 2011] and DESI [Levi et al., 2013] and, especially, by deep, all-sky surveys with good redshift information such as SPHEREX [Doré et al., 2014].

Finally, we have also calculated the statistical error in the direction of the kinematic dipole, based on the fiducial amplitude we had adopted as shown in Fig. 3.8. The direction's error is generally rather large, e.g. an area of about $\simeq 10$ deg in radius for $N_{gal} = 10^8$ and the Galactic ± 15 deg cut. Nevertheless, a combination of the kinematic dipole's amplitude *and* direction that roughly match the CMB dipole would present a convincing confirmation

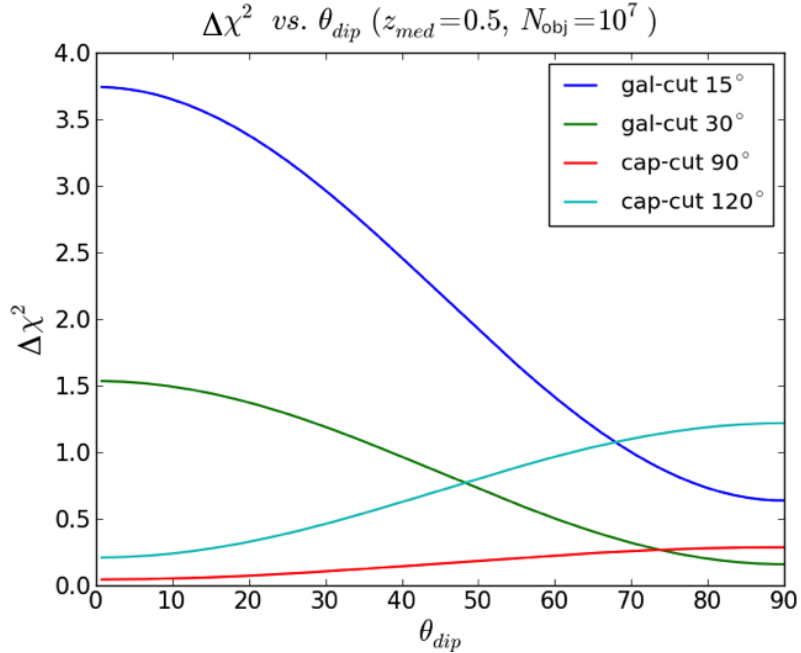


Figure 3.7: Goodness-of-fit ($\Delta\chi^2$ from Eq. (3.9)) dependence of the direction of LSS dipole: θ_{dip} is the angle between the direction of LSS dipole and the galactic north pole. This is the case when $N_{gal} = 10^7$ and $z_{med} = 0.5$ and this directional dependence is similar for other conditions.

of the standard assumption. One could further carry out detailed forecasts of what various findings could rule out the null hypothesis; we leave that for future work.

3.4 Conclusions

We have studied the prospects for measuring the kinematic dipole – our motion through the LSS rest frame – as revealed by the relativistic aberration of tracers of the large-scale structure. The standard theory predicts that the kinematic dipole should agree with the CMB dipole, but this expectation could be violated due to a number of reasons. Therefore, verifying the standard expectation is an important null test in cosmology. The challenge comes from the fact that the dipole amplitude is small ($A \sim 0.003$), and easily contaminated by the intrinsic clustering of galaxies (the “local-structure dipole”).

A successful measurement of the kinematic dipole therefore has two qualitatively dif-

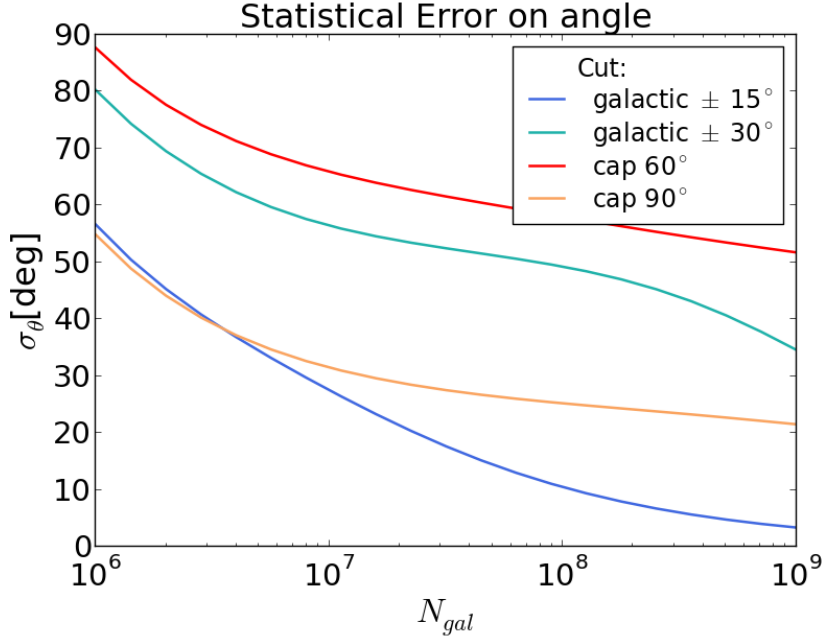


Figure 3.8: Statistical error in the dipole angle as a function of the number of galaxies in a survey for four sky cuts.

ferent requirements: the survey should cover most of the sky and have enough objects to have sufficient signal-to-noise to detect the aberration signature of the dipole, but it should also be deep enough, so that the local-structure dipole contamination is sufficiently small. The two requirements are displayed in Fig. 3.4 and Fig. 3.6 respectively. For a $5\text{-}\sigma$ detection, a survey covering $\gtrsim 75\%$ of the sky in both hemispheres (our “Galactic ± 15 deg cut” case), with $N_{gal} \gtrsim 30$ million galaxies, is required. For a negligible bias, this same survey should have median redshift greater than about 0.75 or higher, with increasing depth requirements as N_{gal} increases.

Fortunately these requirements can be satisfied by upcoming surveys, including DESI, Euclid, and LSST if they are properly combined, and potentially with SPHEREX alone. Even current all-sky surveys such as WISE (Wide-field Infrared Survey Explorer, [Wright et al., 2010]) are not out of the question, provided a sufficiently deep sample can be selected photometrically; current WISE samples have typical galaxy redshifts $z_{med} \simeq 0.2$ [Bilicki et al., 2014] and are not yet deep enough to measure the kinematic dipole.

CHAPTER 4

Closing Remarks

The cosmological principle asserts that our Universe is statistically isotropic and homogeneous at large scales. This principle has guided us to establish the standard cosmological model and to understand overall picture of our Universe. Meanwhile, the cosmological principle is becoming more testable and necessary to be tested as we enter the era of precision cosmology. In this thesis, we suggested the measurement of dipolar modulation in LSS as a test of the statistical isotropy. We have discussed two topics related to the measurement of local-structure dipole and observational prediction of kinematic dipole in LSS.

In Chapter 2 [Yoon et al., 2014], we measured dipolar modulation in the number counts of WISE - 2MASS sources. The matched sources relied on the selection criteria carefully chosen to ensure the high purity (free from star contamination), completeness and uniformity, all of which are crucial for dipole measurement. We matched WISE with 2MASS to select galaxies and obtained a map containing a larger number of galaxy sources, more complete sky coverage, and greater depth than the previous dipole measurements. The measured amplitude of the dipole is somewhat larger than the theoretical estimate based on Λ CDM model, and could therefore suggest the need for new physics at large scales.

In Chapter 3 [Yoon & Huterer, 2015], we investigated the conditions for the detection of kinematic dipole in future LSS surveys. The kinematic dipole originates from our solar systems motion relative to the LSS rest frame. This kinematic dipole that has been well measured in CMB observations should be the same as the one in LSS, assuming the stan-

standard cosmological model holds true. However, if the kinematic dipole in LSS is different from the one measured by the CMB, this could reveal the misalignment between CMB and LSS rest frames. The kinematic dipole in the LSS has been difficult to measure in previous surveys mainly because of their shallow depths and systematic bias from local effect. The future surveys could improve this, because the effect from local structures reduce as we observe deeper. Using the Fisher matrix analysis technique, we concluded that a survey with sky coverage around 75%, covering both hemispheres, and having 3 million objects would be able to detect the kinematic dipole as long as its median redshift is 0.75. This result suggests that a future survey such as SPHEREx [Doré et al., 2014] would be ideal for kinematic dipole detection, because this all-sky satellite mission not only covers both hemispheres but is also optimized for large cosmological volume at low redshifts, aiming for 300 million galaxies with 10% redshift accuracy and 9 million with 0.3% redshift accuracy.

The next decade will witness a drastic improvement in the quantity and the quality of LSS observations. DESI (Dark Energy Spectroscopic Instrument) [Levi et al., 2013] is a spectroscopic survey to obtain data of 18 million emission-line galaxies, 4 million luminous red galaxies and 3 million quasi-stellar objects distributed at redshift $z < 2$. One of the largest planned photometric surveys is LSST (Large Synoptic Survey Telescope), targeting millions of Type Ia supernovae and 4 billion galaxies (with 1-2% photometric redshift accuracy and the third quartile at $z \sim 2$) uniformly covering almost a half of the sky. Finally, Euclid [Laureijs et al., 2011] is a space-based mission reaching to galaxies at high redshifts covering over 15,000 deg². Euclid aims for both photometric observations of a billion galaxies and for spectroscopic observations of millions of galaxies. These upcoming surveys will equip us with very deep and wide three-dimensional LSS maps. At the same time, improving utilization of currently available surveys by combining with other surveys, as we did in this thesis with WISE and 2MASS, will be another good way of achieving better LSS maps. Based on these improved maps we will be able to distinguish

the different origins of imprint in dipolar modulation of LSS beyond our current limitation.

To understand the alignment of our measured dipoles in LSS with other measurements as well as their contributions to our measurements, we need to better quantify how local structures such as supervoids or superclusters and bulk motions of observed objects are attributable to the local-structure dipole measurement. While it is more important to have more objects in the map for dipole detection, to make sure measurement more precise, it may also be worth investigating how much signal uncertainty comes from the redshift uncertainty of observed objects.

Future surveys may find a different kinematic dipole in the LSS from the one observed in the CMB when there is misalignment of CMB and LSS rest frames. This means some recalibration would be necessary when one cross-correlates CMB with LSS data. For example, estimators on Integrated Sachs-Wolfe effect (temperature change of light coming through LSS potentials) and CMB lensing (light deflection due to matter distribution) map may require modification. The CMB lensing map reconstruction is a great tool that provides additional useful information on LSS. Therefore, to fully benefit from novel cross-correlations and cross-calibrations between CMB and LSS, we would need modification of estimators related to both of CMB and LSS observations.

More fundamentally, if we could confirm the statistical anisotropy, it would have a tremendous impact on the whole of precision cosmology. As a result, it would provide clues to construct more precise models of the early Universe; constrained cosmological parameters would be recalibrated accordingly. The modified models may be able to embrace observed anomalies and to relieve some observed tensions among different measurements.

APPENDIX A

Kinematic Dipole Statistical Error Estimation

A.1 Coordinate transformation of Fisher matrix

Let F_{old} be the original Fisher matrices in the old parameters, and F_{new} be the desired new Fisher matrix. The relation between these two Fisher matrix is given as (see e.g. Eq. (12) in the Figure of Merit Science Working Group Report, <http://arxiv.org/pdf/0901.0721v1.pdf>):

$$F_{\text{new}} = M^T F_{\text{old}} M \quad (\text{A.1})$$

where M is the matrix of derivatives of old with respect to new:

$$M_{ij} \equiv \frac{\partial p^i}{\partial q^j}. \quad (\text{A.2})$$

In our case the bases are

$$p^i \in \{d_x, d_y, d_z\} \quad (\text{A.3})$$

$$q^i \in \{A, \theta, \phi\}. \quad (\text{A.4})$$

The are related via

$$d_x = A \sin \theta \cos \phi \quad (\text{A.5})$$

$$d_y = A \sin \theta \sin \phi \quad (\text{A.6})$$

$$d_z = A \cos \theta \quad (\text{A.7})$$

So that the matrix M is given by

$$M = \begin{pmatrix} \sin \theta \cos \phi & A \cos \theta \cos \phi & -A \sin \theta \sin \phi \\ \sin \theta \sin \phi & A \cos \theta \sin \phi & A \sin \theta \cos \phi \\ \cos \theta & -A \sin \theta & 0 \end{pmatrix} \quad (\text{A.8})$$

Values to plug in correspond to the fiducial (A, θ, ϕ) dipole. The matrix F_{new} corresponds to the full 3×3 info on the three new parameters, and simply marginalizing over the other brings the error on A :

$$\sigma(A) = \sqrt{(F_{\text{new}}^{-1})_{11}}. \quad (\text{A.9})$$

F_{new} can be simplified based on the fact F_{old} is diagonal and $F_{\text{old}(11)} = F_{\text{old}(22)}$ due to the symmetry of the cuts for x and y axis. Non-zero elements of F_{new} are:

$$F_{AA} = F_{11} \sin^2 \theta + F_{33} \cos^2 \theta \quad (\text{A.10})$$

$$F_{\theta\theta} = A^2 (F_{11} \cos^2 \theta + F_{33} \sin^2 \theta) \quad (\text{A.11})$$

$$F_{\phi\phi} = A^2 F_{11} \sin^2 \theta \quad (\text{A.12})$$

$$F_{A\theta} = F_{\theta A} = A \sin \theta \cos \theta (F_{11} - F_{33}) \quad (\text{A.13})$$

so σ s calculated based on F_{new}^{-1} are:

$$\sigma(A) = \sqrt{\sin^2 \theta / F_{11} + \cos^2 \theta / F_{33}} \quad (\text{A.14})$$

$$\sigma(\theta) = 1/A \sqrt{(\cos^2 \theta / F_{11} + \sin^2 \theta / F_{33})} \quad (\text{A.15})$$

$$\sigma(\phi) = 1/A \sqrt{1/(\sin^2 \theta F_{11})} \quad (\text{A.16})$$

A.2 Adding higher multipole templates

The multipoles ($Y_{\ell m}$) in spherical harmonic space are orthogonal bases which are uncorrelated in full sky. For cut-sky, however, multipoles become correlated so the dipole estimator should include higher multipoles as templates T_i (in Eq. 2.3) to properly measure dipolar modulation. Adding higher multipoles increases the statistical error and we need to determine the maximum multipoles to estimate errors. Setting the maximum multipoles (ℓ_{\max}) is only possible when the errors converge, but we found they do not converge without priors based on the expected C_ℓ values. Fig. A.1 shows the statistical errors in x, y, and z directions ($\sigma_x, \sigma_y, \sigma_z$) for different sky-cuts with/without priors. Fig. A.2 enlarged the cases with priors and shows that setting $\ell_{\max} = 10$ is an enough approximation.

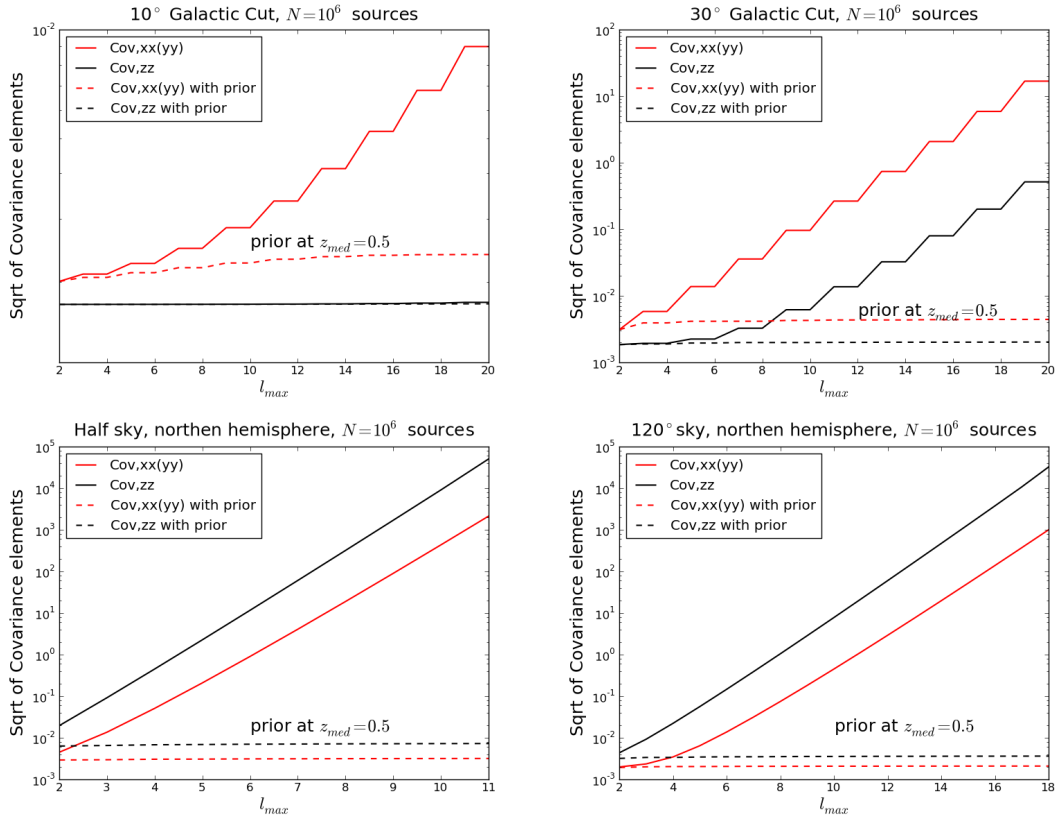


Figure A.1: The comparison between with priors vs. without priors of statistical errors ($\sigma_x, \sigma_y, \sigma_z$) in x, y, z direction as a function of ℓ_{\max}

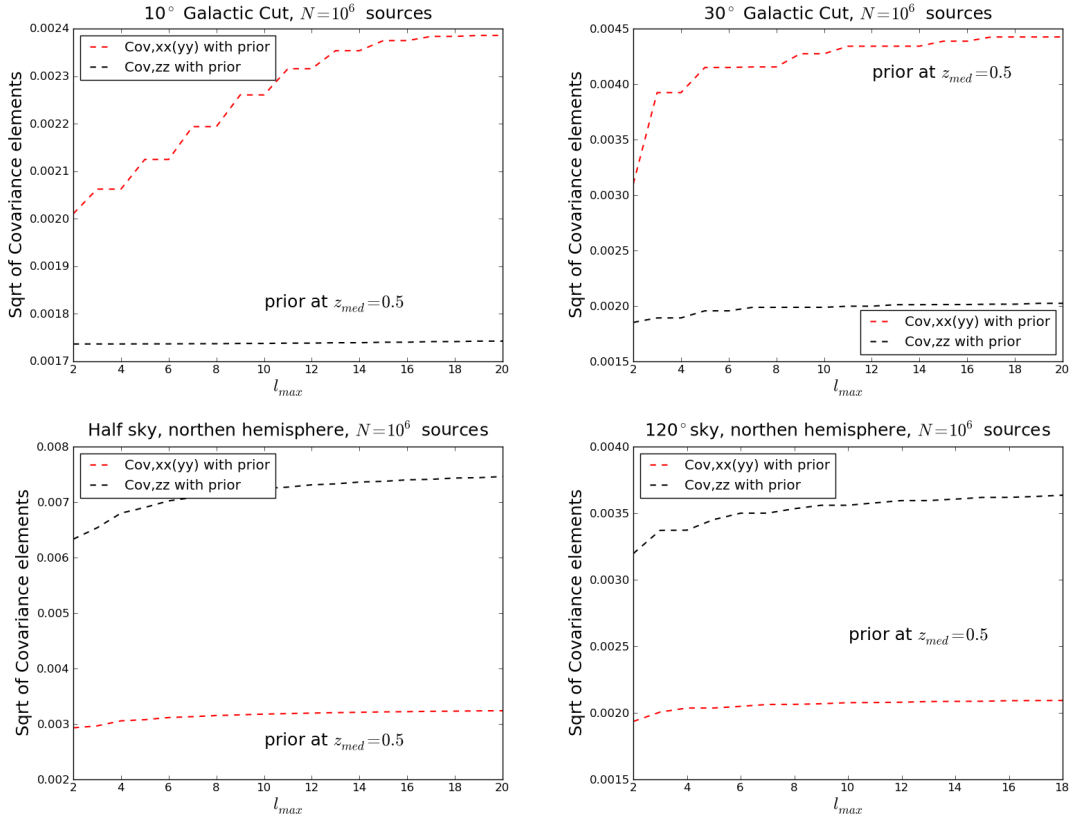


Figure A.2: The statistical errors ($\sigma_x, \sigma_y, \sigma_z$) in x, y, z direction as a function of l_{max} with priors; priors are based on the expected values of higher multipoles (C_ℓ).

A.3 Effects of the sky mask

Since we tested only the simplified footprints of coverage (Fig. 3.3), one may expect the results are unrealistic. We thus applied an additional mask (WMAP mask) to each footprint for checking possible deviation from small features of mask. The results, however, are not very different from the simplified footprints. This is illustrated in Fig. A.3 which shows the statistical error of dipole amplitude for four sky-cuts. The cases the additional mask (black dotted) are not far from the cases without the mask (colored). This figure corresponds to Fig. 3.4 and the four lines are cap 90° cut, 60° cut, galactic $\pm 30^\circ$ cut and galactic $\pm 15^\circ$ cut from the top in order. The last three pairs of lines are almost overlapped and the last pair, cap 90° case, shows slight deviation up to $N_{obj} \sim 10^7$.

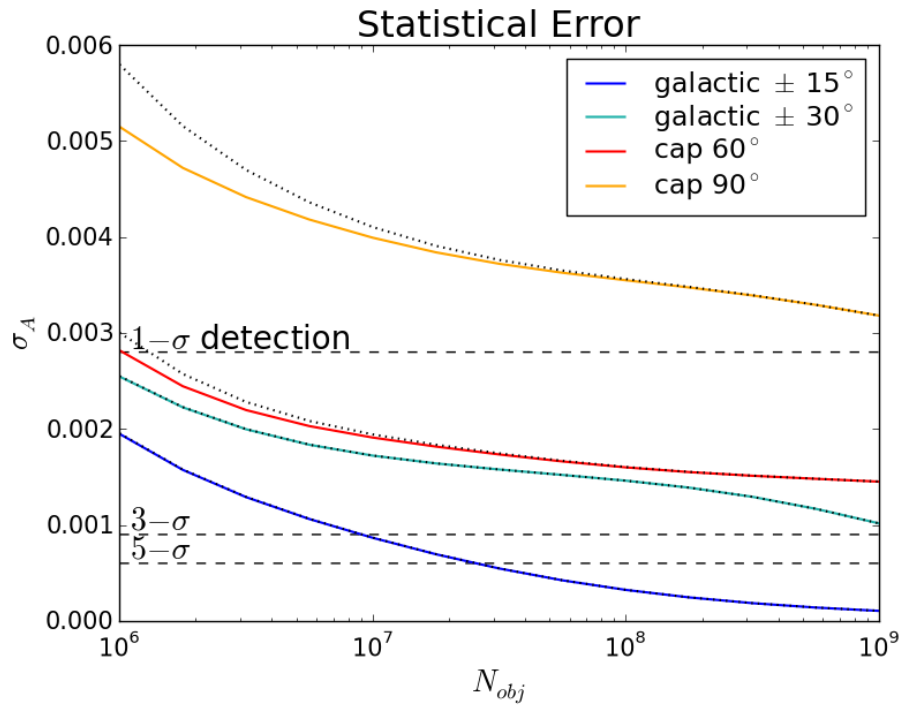


Figure A.3: Effects of WMAP mask on the statistical errors in dipole amplitude: this shows the difference between the cases with (black dotted) and without (colored) WMAP mask applied in estimation of statistical errors.

BIBLIOGRAPHY

- Ade P., et al., 2013, arXiv:1303.5083
- Ade P. A. R., et al., 2015
- Aghanim N., et al., 2014, A&A, 571, A27
- Alonso D., Salvador A. I., Sánchez F. J., et al., 2015, MNRAS, 449, 670
- Appleby S., Shafieloo A., 2014, JCAP, 10, 70
- Baleisis A., Lahav O., Loan A. J., Wall J. V., 1998, MNRAS, 297, 545
- Bennett C., et al., 2013, ApJS, 208, 20
- Bilicki M., Jarrett T. H., Peacock J. A., Cluver M. E., Steward L., 2014, ApJS, 210, 9
- Blake C., Wall J., 2002, Natur, 416, 150
- Burles S., Rappaport S., 2006, ApJL, 641, L1
- Copi C. J., Huterer D., Schwarz D. J., Starkman G. D., 2006, Mon. Not. Roy. Astron. Soc., 367, 79
- Courteau S., Van Den Bergh S., 1999, The Astronomical Journal, 118, 337
- Dai D.-C., Kinney W. H., Stojkovic D., 2011, JCAP, 4, 15
- Doré O., et al., 2014
- Driver S. P., Gama Team, 2008, Anglo-Australian Observatory Epping Newsletter, 114, 3
- Erickcek A., Carroll S., Kamionkowski M., 2008, PhRvD, 78, 083012
- Feldman H. A., Watkins R., Hudson M. J., 2010, MNRAS, 407, 2328
- Fernández-Cobos R., Vielva P., Pietrobon D., et al., 2014, MNRAS, 441, 2392
- Ferraro S., Sherwin B. D., Spergel D. N., 2014, arXiv:1401.1193
- Finelli F., Garcia-Bellido J., Kovacs A., Paci F., Szapudi I., 2014, ArXiv e-prints
- Gibelyou C., Huterer D., 2012, MNRAS, 427, 1994

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759

Hajian A., Souradeep T., 2006, *Physical Review D*, 74, 123521

Hinshaw G., et al., 2009, *ApJS*, 180, 225

Hirata C. M., 2009, *JCAP*, 0909, 011

Høg E., Fabricius C., Makarov V. V., Urban, et al., 2000, *A&A*, 355, L27

Huterer D., 2002, *PhRvD*, 65, 063001

Itoh Y., Yahata K., Takada M., 2010, *PhRvD*, 82, 043530

Ivezic Z., Tyson J. A., Allsman R., Andrew J., Angel R., 2008

Kalus B., Schwarz D. J., Seikel M., Wiegand A., 2013, *A&A*, 553, A56

Kashlinsky A., Atrio-Barandela F., Kocevski D., Ebeling H., 2008, *The Astrophysical Journal Letters*, 686, L49

Kovács A., Szapudi I., 2013, *arxiv:1401.0156*

Kovács A., Szapudi I., Granett B. R., Frei Z., 2013, *MNRAS*, 431, L28

Laureijs R., et al., 2011

Lavaux G., Afshordi N., Hudson M. J., 2013, *MNRAS*, 430, 1617

Levi M., et al., 2013

Ma Y.-Z., Gordon C., Feldman H. A., 2011, *PhRvD*, 83, 103002

Ma Y.-Z., Pan J., 2014, *MNRAS*, 437, 1996

Maller A. H., McIntosh D. H., Katz N., Weinberg M. D., 2003, *The Astrophysical Journal Letters*, 598, L1

Marchesini D., Stefanon M., Brammer G. B., Whitaker K. E., 2012, *ApJ*, 748, 126

Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., et al., 2013, *arXiv:1303.4722*

Mo H., van den Bosch F., White S., 2010, *Galaxy Formation and Evolution*. Cambridge University Press

Mukherjee S., Aluri P. K., Das S., Shaikh S., Souradeep T., 2015

Rassat A., Land K., et al., 2007, *MNRAS*, 377, 1085

Rathaus B., Kovetz E. D., Itzhaki N., 2013, *MNRAS*, 431, 3678

Rubart M., Bacon D., Schwarz D. J., 2014, *arXiv:1402.0376*

Rubart M., Schwarz D. J., 2013, *A&A*, 555, A117

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *Astrophys.J.*, 500, 525

Skrutskie M., et al., 2006, *AJ*, 131, 1163

Szapudi I. et al., 2015, *MNRAS*, 450, 288

Szapudi I., Prunet S., et al., 2001, *ApJ*, 548, L115

Tegmark M., et al., 2004, *Phys. Rev.*, D69, 103501

Turnbull S. J., Hudson M. J., Feldman H. A., Hicken M., Kirshner R. P., Watkins R., 2012, *MNRAS*, 420, 447

Turner M., 1991, *PhRvD*, 44, 3737

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

White M., Majumdar S., 2004, *ApJ*, 602, 565

Wright E., et al., 2010, *AJ*, 140, 1868

Yan L., et al., 2013, *AJ*, 145, 55

Yoon M., Huterer D., 2015, *Astrophys. J.*, 813, L18

Yoon M., Huterer D., Gibelyou C., Kovács A., Szapudi I., 2014, *MNRAS*, 445, L60

Zibin J., Scott D., 2008, *PhRvD*, 78, 123529