Mon. Not. R. Astron. Soc. **392,** 1225–1232 (2009) doi:10.1111/j.1365-2966.2008.14127.x

Galaxy Zoo: chiral correlation function of galaxy spins★

Anže Slosar,[1,2,3]† Kate Land,[2] Steven Bamford,[4,5] Chris Lintott,[2] Dan Andreescu,[6] Phil Murray,[7] Robert Nichol,[4] M. Jordan Raddick,[8] Kevin Schawinski,[2,9,10] Alex Szalay,[8] Daniel Thomas[4] and Jan Vandenberg[8]

[1]*Berkeley Center for Cosmological Physics, Lawrence Berkeley Nat. Lab & Phys. Department, University of California, Berkeley, CA 94720, USA*
[2]*Astrophysics Department, University of Oxford, Oxford OX1 3RH*
[3]*Faculty of Mathematics & Physics, University of Ljubljana, Slovenia*
[4]*Institute of Cosmology and Gravitation, University of Portsmouth, Mercantile House, Hampshire Terrace, Portsmouth PO1 2EG*
[5]*Centre for Astronomy and Particle Theory, University of Nottingham, University Park, Nottingham NG7 2RD*
[6]*LinkLab, 4506 Graystone Ave., Bronx, NY 10471, USA*
[7]*Fingerprint Digital Media, 9 Victoria Close, Newtownards, Co. Down BT23 7GY*
[8]*Department of Physics and Astronomy, The Johns Hopkins University, Homewood Campus, Baltimore, MD 21218, USA*
[9]*Department of Physics, Yale University, New Haven, CT 06511, USA*
[10]*Yale Center for Astronomy and Astrophysics, Yale University, PO Box 208181, New Haven, CT 06520, USA*

Accepted 2008 October 22. Received 2008 October 21; in original form 2008 September 12

ABSTRACT

Galaxy Zoo is the first study of nearby galaxies that contains reliable information about the spiral sense of rotation of galaxy arms for a sizeable number of galaxies. We measure the correlation function of spin chirality (the sense in which galaxies appear to be spinning) of face-on spiral galaxies in angular, real and projected spaces. Our results indicate a hint of positive correlation at separations less than $\sim$0.5 Mpc at a statistical significance of 2σ–3σ. This is the first experimental evidence for chiral correlation of spins. Within the tidal torque theory, it indicates that the inertia tensors of nearby galaxies are correlated. This is complementary to the studies of nearby spin axis correlations that probe the correlations of the tidal field. Theoretical interpretation is made difficult by the small distances at which the correlations are detected, implying that substructure might play a significant role, and our necessary selection of face-on spiral galaxies, rather than a general volume-limited sample.

Key words: galaxies: general – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Understanding the creation and evolution of the angular momentum of dark matter haloes and galaxies is a crucial building block of a comprehensive theory of galaxy evolution. Hoyle (1949) was first to propose that the galaxy spin could be ascribed to the gravitational coupling with the surrounding galaxies. This idea has been formalized and extended in the subsequent work (Peebles 1969; Doroshkevich 1970; White 1984; Heavens & Peacock 1988; Catelan & Theuns 1996) into the modern theory of the evolution of galaxy spin, known as the tidal torque theory [TTT; see Schaefer (2008) for a review]. This theory asserts that protohaloes acquire most of their angular momentum in the early stages of their formation, from the lowest non-vanishing contribution from the linear Lagrangian theory, i.e. a coupling of the quadrupole of the local mass distribution to the external gravitational shear. Compared to *N*-body simulations, theory produces qualitatively correct results, although there are still significant discrepancies at a more quantitative level. Moreover, it seems that at present there are no clear theoretical directions for improving analytical models (Barnes & Efstathiou 1987; Porciani, Dekel & Hoffman 2002a; Bailin & Steinmetz 2005).

On the observational side, most of the work has been done using spiral galaxies. These are characterized by a rotating disc of baryonic matter. The line perpendicular to the plane of the disc determines the axis of rotation, while the spiral arms in most galaxies encode the sense of rotation, i.e. the difference between left- and right-hand screws sense of rotation. For spiral galaxies seen in projection, one can measure the observed galaxy ellipticities, which constrain the *axis* of the galaxy spin (Pen, Lee & Seljak 2000; Lee & Pen 2002; Trujillo, Carretero & Patiri 2006; Lee & Erdogdu 2007). This axis, within a two-fold degeneracy, is known to be associated with the tilt of the galactic plane with respect to the plane of the

★This publication has been made possible by the participation of more than 100 000 volunteers in the Galaxy Zoo project. Their individual contributions are acknowledged at http://www.galaxyzoo.org/Volunteers.aspx.
†E-mail: anze@berkeley.edu

sky. Since the vector can point in two directions on the same axis, the ellipticities constrain the spin vector within a total of four-fold degeneracy. Note that chiral information, *viz.* information about the actual directions of the spin vectors as opposed to spin axis, is completely absent in the study of galactic ellipticities. However, this information contains important clues about the details of the emergence of the spin. As we will explain later in the text, the detection of chiral correlation function implies that the local inertia tensors must be correlated. This lends experimental support to the theoretical expectations that the inertia and gravitational shear tensor are correlated (Porciani, Dekel & Hoffman 2002b).

We now have a unique tool to study the chiral properties of galaxy spins. Through an online project called Galaxy Zoo[1] (Lintott et al. 2008), members of the public have visually classified the morphologies and spin orientations for the entire spectroscopic sample of the Sloan Digital Sky Survey (SDSS; York et al. 2000) Data Release 6 (DR6; Adelman-McCarthy et al. 2008). Data and its reduction are extensively discussed in Lintott et al. (2008).

Spiral galaxies in the Galaxy Zoo sample are classified as clockwise, anticlockwise or edge-on. The spin direction convention used here is such that clockwise and anticlockwise rotations correspond to the galaxies whose arms are rotating in the sense of the letters Z and S, respectively (Sugai & Iye 1995). For each face-on galaxy, we thus receive one bit of information corresponding to the sign of the galaxy spin vector projected along the line of sight. It is important to note that this information is independent of the tilt of the plane of the galaxy. We will refer to this one-bit information simply as galaxy spin. By the galaxy spin vector, we mean the unit vector that defines the apparent spin of the galaxy: it is perpendicular to the disc plane and points in the direction the right-hand turn screw would move if turned following the spiral arms inwards. This quantity is strongly correlated with the real angular momentum of the gas. The correlation, however, is not perfect and observations show that the angular momentum vector of the gas points in the opposite direction in about 4 per cent of systems (Pasha & Smirnov 1982). In turn, there are theoretical expectations that there is a strong, but not perfect, correlation between the angular momentum vector of gas and that of the dark matter haloes hosting the galaxy (van den Bosch et al. 2002). A detection of the correlation in the galaxy spins would therefore imply a correlation in the dark matter spin vectors. Conversely, a non-detection of the spin correlation can be used to put upper limits on the correlation between angular momentum vectors of dark matter haloes.

This paper is structured as follows. In Section 2, we shortly review the TTT and its main results. Section 3 will connect the correlation function η to an observable correlation function of spins c, while Section 4 will introduce our data and measurement technique. We present our results and discuss systematics in Section 5. Finally, we discuss our results and conclude in Section 6.

2 TIDAL TORQUE THEORY

The TTT derives the following expression for the angular momentum from the first-order linear perturbation theory in Lagrangian space (White 1984; Catelan & Theuns 1996):

$$L_i(t) = a^2(t)\dot{D}\epsilon_{ijk}T_{jl}I_{lk}, \tag{1}$$

where a is the scale factor of the Universe, D is the growth factor and ϵ_{ijk} is the Levi–Civita symbol. The local inertia tensor I_{ij} of the protohalo (the mass that will later form the dark matter halo) in Lagrangian space is given by

$$I_{ij} = \bar{\rho_o}\int_V q_i q_j \mathrm{d}^3 q, \tag{2}$$

where q_i are the Lagrangian coordinates around the centre of mass of the halo and ρ_o is the mean density. The local shear tensor T_{ij} is defined by

$$T_{ij} = \partial_i\partial_j\phi(\boldsymbol{q}), \tag{3}$$

where ϕ is the gravitational potential. In other words, the TTT requires two components: a non-vanishing quadrupole distribution of mass in the halo to be spun up and the cosmological tidal field. In principle, it sounds plausible to assume that while tidal fields between neighbouring protohaloes are correlated, since they are coming from the large-scale modes, the local quadrupole moments of mass distribution are sourced due to random distribution of the local inhomogeneities and should therefore be random. This assumption of a statistical isotropy of the inertia tensor gives the following anatz for the angular moment correlator (Pen et al. 2000):

$$Q_{ij} = \langle L_i L_j|\hat{\boldsymbol{T}}\rangle = \frac{1}{3}\delta_{ij} + c\left(\frac{1}{3} - \hat{T}_{ik}\hat{T}_{kj}\right), \tag{4}$$

where c controls the level of randomization of axial preference due to non-linear and stochastic effects. In fact, this ansatz has been shown to satisfactorily explain the inclinations of axes of spiral galaxies in the vicinity of voids with $c \sim 0.7$ (Trujillo et al. 2006).

Some further algebra gives the probability distribution function for spins $\boldsymbol{s} = \hat{\boldsymbol{L}}$, usually assumed to be Gaussian (Pen et al. 2000):

$$P(\boldsymbol{s}|\boldsymbol{T}) = \frac{|\hat{\boldsymbol{Q}}|^{-1/2}}{4\pi}\exp(-\boldsymbol{s}^T\cdot\hat{\boldsymbol{Q}}^{-1}\cdot\boldsymbol{s}). \tag{5}$$

Using this expression, it is therefore possible to calculate various correlators of $\boldsymbol{s}$ if correlators of $\boldsymbol{T}$ are known.

We will now consider correlation functions. The most general form of possible correlation statistics of the galaxy spins consistent with the homogeneity and isotropy is the spin correlation tensor defined by (see e.g. Groth, Juszkiewicz & Ostriker 1989)

$$\begin{aligned}\Xi_{ij}(r) &= \left\langle s_i(\boldsymbol{x})s_j(\boldsymbol{x}+\boldsymbol{r})\right\rangle \\ &= [\Pi(r) - \Sigma(r)]\hat{r}_i\hat{r}_j + \Sigma(r)\delta_{ij}.\end{aligned} \tag{6}$$

Functions Π and Σ are parallel and perpendicular correlation functions. Following Porciani et al. (2002a), we will be dealing exclusively with the 'dot product' correlation functions given by

$$\eta(r) = \Xi_{ii} = \langle \boldsymbol{s}(\boldsymbol{x})\cdot\boldsymbol{s}(\boldsymbol{x}+\boldsymbol{r})\rangle \tag{7}$$

and

$$\eta_2(r) = \left\langle[\boldsymbol{s}(\boldsymbol{x})\cdot\boldsymbol{s}(\boldsymbol{x}+\boldsymbol{r})]^2\right\rangle - 1/3. \tag{8}$$

It is important to realize what these two quantities measure. The first one measures if the angular momentum vectors are correlated, while the second one measures if the axes of angular momentum vectors are correlated. Note that it is perfectly possible to have $\eta_2(r) > 0$, while $\eta(r) = 0$. For example if all spins were aligned along the z-axis, but with an orientation that is chosen at random from $\hat{z}$ and $-\hat{z}$, $\eta_2 = 1$, while $\eta = 0$. In fact, following the ansatz of equations (4) and (5) results in vanishing $\eta(r)$ and a finite $\eta_2(r)$ (analytical expression for which can be found in Lee & Pen 2001). This is trivially seen from equation (5), since $P(\boldsymbol{s}|\boldsymbol{T}) = P(-\boldsymbol{s}|\boldsymbol{T})$ and is a direct consequence of the assumption of isotropy of local moments of inertia. To put it simply, for a fixed tidal field, averaging over possible realizations of the inertia tensor will, in general, produce a

[1] www.galaxyzoo.org

preferred axis (determined by eigenvectors of $\boldsymbol{T}$), but not a preferred direction. This is due to the fact that for every inertia tensor $\boldsymbol{I}$ that produces a final angular momentum $\boldsymbol{L}$, an equally likely mirror image inertia tensor $-\boldsymbol{I}$ will produce an equal and opposite angular moment $-\boldsymbol{L}$. Therefore, the common assumption that the local moments of inertia are random and uncorrelated will, in general, produce a non-vanishing axis correlation, but a vanishing correlation of the actual spin vectors. The important corollary is that a detection of the chiral correlations in the galaxy spins would directly indicate that the moments of inertia are non-random.

3 CORRELATIONS OF PROJECTED SPINS

3.1 Small-scale correlation function of spins

Our life as observers is complicated by the fact that the sense in which spiral arms wind in a projected image of a spiral galaxy measures the sign of the spin vector projected along the line of sight rather than the spin vector itself. We will therefore consider the correlation function of $\tilde{s} = \mathrm{sgn}(\boldsymbol{s} \cdot \hat{z})$:

$$c(r) = \langle \tilde{s}(\boldsymbol{x}) \tilde{s}[\boldsymbol{x} + \boldsymbol{r})] \rangle \,, \tag{9}$$

where we assumed that the radial vectors to positions $\boldsymbol{x}$ and $\boldsymbol{x} + \boldsymbol{r}$ are parallel, i.e. the flat-sky approximation. Note that this only requires *pairs* of galaxies to be close enough so that the flat-sky approximation holds, rather than an entire survey occupying a small portion of the sky. For the time being, we also neglect the difference between the dark-matter angular momentum and the gas angular momentum.

To proceed, we note that $\eta(r)$ is determined entirely by the one-point distribution function $P(\mu|r)$ for cosine angle $\mu = \cos\theta$ between the two spin vectors:

$$\eta(r) = \left\langle s_i(\boldsymbol{x}) s_j(\boldsymbol{x} + \boldsymbol{r}) \right\rangle = \langle \mu \rangle = \int \mu P(\mu|r) \, \mathrm{d}\mu. \tag{10}$$

For a given μ, one will observe two galaxies with the same orientations of spins with the probability (see beginning of Appendix A)

$$P_{+1}(\theta) = 1 - \theta/\pi \tag{11}$$

and with different orientations of spins with the probability

$$P_{-1}(\theta) = 1 - P_{+1}(\theta) = \theta/\pi. \tag{12}$$

The correlation function of the spin signs is then given by

$$c(r) = \frac{1}{N} \int \mathrm{d}\mu P(\mu|r) \left[P_{+1}(\theta) - P_{-1}(\theta) \right], \tag{13}$$

where normalization N is in this case trivially given by

$$N = \int \mathrm{d}\mu P(\mu|r) \left[P_{+1}(\theta) + P_{-1}(\theta) \right] = 1. \tag{14}$$

When there are no correlations, $P(\mu)\,\mathrm{d}\mu = 1/2\mathrm{d}\,\mu$ and both $\eta(r)$ and $c(r) = 0$. When correlations exist, we must specify a one-point probability distribution function for $P(\theta|r)$. We assume the following form:

$$P(\theta|r) = \sin(\theta) \left[1 + e(r) \cos(\theta) \right]. \tag{15}$$

Using this form, one obtains

$$\eta(r) = \frac{1}{3} e(r), \tag{16}$$

$$c(r) = \frac{1}{4} e(r) \tag{17}$$

and so

$$\eta(r) = \frac{4}{3} c(r). \tag{18}$$

Equation (18) hinges on the particular form for $P(\theta|r)$ that we chose. In practice, different forms generically give the results that $\eta(r) = qc(r)$ with q typically between 1 and 3/2.

In reality, however, we measure the correlation function only for galaxies that are sufficiently away from the edge-on orientation not to be classified as a face-on galaxy. For simplicity, let us assume that our sample contains only galaxies, whose spin vector satisfies

$$s(\boldsymbol{x}) \cdot \hat{z} > \cos\alpha. \tag{19}$$

In other words, galaxies that are inclined with an angle greater than α with respect to the line of sight are assumed to have been classified as edge-on. What is the functional form for $P_{+1}(\theta)$ in this case? In Appendix A, we show that for $\alpha > \pi/4$

$$P_{+1}(\theta) = \begin{cases} f(\theta) & \theta < \pi - 2\alpha; \\ f(\theta) & \pi - 2\alpha < \theta < 2\alpha; \\ 0 & 2\alpha < \theta \end{cases} \tag{20}$$

and

$$P_{-1}(\theta) = \begin{cases} 0 & \theta < \pi - 2\alpha; \\ f(\pi - \theta) & \pi - 2\alpha < \theta < 2\alpha; \\ f(\pi - \theta) & 2\alpha < \theta, \end{cases} \tag{21}$$

where

$$\begin{aligned} f(\theta) = {} & 1 - 2\cos(\alpha) \\ & - \frac{1}{\pi} \cos^{-1} \left(\frac{\cos\theta - \cos^2\alpha}{\sin^2\alpha} \right) \\ & + \frac{2\cos\alpha}{\pi} \cos^{-1} \left[\frac{\cos\alpha(\cos\theta - 1)}{\sin\theta \sin\alpha} \right]. \end{aligned} \tag{22}$$

The 'lost' probability, i.e. $1 - P_{+1} - P_{-1}$, corresponds to geometries that are not detected and in general results in $N < 1$. Numerically integrating equation (13), we can obtain a relation between η and the measured $c^{\mathrm{meas}}(r)$.

3.2 Connection to gas angular momentum

As discussed in Pasha & Smirnov (1982), a fraction $f = 0.04$ of galaxies has gas angular momentum that is pointing in the opposite direction to the apparent galaxy spin inferred from orientation of spiral arms. If we momentarily distinguish between the actual and the apparent gas spin correlation functions, we can write

$$\eta_{\mathrm{apparent}}(r) = \left[(1 - f)^2 + f^2 \right] \eta(r) - 2f(1 - f)\eta(r), \tag{23}$$

since correlation function of spins will receive a negative contribution if exactly one spin (but not both) was randomly reversed. This simplifies to

$$\eta_{\mathrm{apparent}}(r) = 4 \left(f - \frac{1}{2} \right)^2 \eta(r). \tag{24}$$

This has the expected properties. The spin correlation function will become zero if exactly half the spin vectors are reversed, effectively randomizing them and exactly following the primary correlation function if all or none spins are reversed. For $f = 0.96$, one gets that $\eta(r) \sim 1.2\eta_{\mathrm{apparent}}(r)$.

4 DATA AND METHOD

4.1 Data

The basic data reduction is described in great detail in Lintott et al. (2008) and Land et al. (2008). We will briefly summarize the data reduction in the following paragraph, but the reader is invited to read the above papers if interested in the details of the primary data reduction.

In the Galaxy Zoo project, a sample of 893 212 galaxies was visually classified by about 90 000 users. The sample was selected to be sources that were targeted for SDSS spectroscopy, i.e. extended sources with Petrosian magnitude $r < 17.77$. Additionally, we included objects that were not originally targeted as such, but were observed to be galaxies once their spectrum was taken. Where spectroscopic redshifts are available, we find that they have the mean redshift of $z = 0.14$ and the objects with the highest redshift reach $z \sim 0.5$. The galaxies thus probe our local Universe at cosmological scales. Each object has been classified about 40 times from a simplified scheme of six possible classifications: an elliptical, a clockwise spiral galaxy, an anticlockwise spiral galaxy, an edge-on spiral galaxy, a star/unknown object and a merger. Various cuts (hacking attempts, browser misconfigurations, etc.) removed about 5 per cent of our data. Data were reduced into two final catalogues based on whether data were weighted or unweighted. In the unweighted data, each user's classification carries an equal weight, while in the weighted case, user's weights are iteratively adjusted according to how well each user agrees with the classifications of other users. In both cases, the accrued classifications are further distilled into superclean, clean and cleanish catalogues of objects, for which we require 95, 80 and 60 per cent of users to agree on a given classification. In all cases, this is a statistically significant detection with respect to random voting; however, the human 'systematical' error associated with it is difficult to judge. In any case, we are in the limit where taking more data will not change our sample beyond noise fluctuations as the votes are uncorrelated. In Land et al. (2008), a bias of unknown origin towards anticlockwise galaxies was discovered and corrected for by adjusting the cleanliness level for the clockwise galaxies to a slightly lower value. This work uses the same data and bias correction. We note, however, that if unaccounted for, such bias would generate a constant offset in the correlation function that cannot mimic the correlations we are seeing in the data. After bias correction is applied, the numbers of clockwise and anticlockwise galaxies are the same within Poisson noise in each sample.

We decided to use the 80 per cent clean, weighted subsample. We stress that the decision to work with 80 per cent clean sample was made in advance and was not chosen to maximize our signal. We show an example of typical clockwise and anticlockwise spinning galaxies from a clean catalogue in Fig. 1.

Figure 1. This figure shows a pair of typical galaxies from our *clean* catalogue. The left-hand image is an anticlockwise (S-like), while the right-hand image is a typical clockwise (Z-like) galaxy.

4.2 Three-dimensional, angular and projected configuration spaces

In the formalism of Section 3, we have always referred to distance between two galaxies as being r, the physical distance between a pair of galaxies. In practice, it can be any measure of distance between galaxies. In this work, we use the following three different distance measures.

(i) *Angular distances*. These have the advantage of producing the highest number of pairs. We denote the corresponding correlation function with $c(\theta)$.

(ii) *Real-space distance*. We use the distance in the redshift space for pairs of galaxies for which both spectroscopic redshifts are known. These are not the true three-dimensional distances, but are instead distances in the redshift space, and therefore are affected by the fingers-of-god effects (see e.g. Hamilton 1998). Since the axis of subhaloes is correlated with the shape of the parent halo (see e. g. Bailin & Steinmetz 2005), there exist correlations in the ratio of edge-on to face-on spirals as a function of projected distances from the centre of the halo. This considerably complicates any correction for fingers-of-god effects and therefore we do not attempt this correction, since effects are likely to be subdominant. A concordant flat cosmology with $\Omega_{\rm m} = 0.25$ was assumed when calculating distances. We denote the corresponding correlation function with $c(r)$.

(iii) *Projected distances*. There distances are the transversal component of the distance vector connecting two galaxies with known redshift. If only one galaxy in the pair has a known redshift, we assume the other galaxy to have the same redshift. The advantage of this distance measure is that it is not affected by the redshift-space distortions and that the number of pairs is significantly larger than in the case of real-space distances. We denote the corresponding correlation function with $c(p)$.

For each of the above distance measures, we first located all galaxy pairs in our sample that are less than 2000 arcsec or 3 or 1 Mpc h^{-1} projected apart. This gave us three sets, which we describe in Table 1.

We have then removed rogue pairs. In the primary SDSS pipeline analysis, every object is assigned an SDSS ID. Large nearby galaxies are often associated with more than one ID, as various knots and substructure of the galaxy are recognized as sources by the reduction software. All such IDs are therefore classified as the same galaxy, resulting in a spurious positive correlation at the shortest distances. Our automatic mechanism removed all pairs for which their angular separation is less than 1.5 $\max(r_p)$, where $\max(r_p)$ denotes the larger of the two Petrosian radii (Petrosian 1976). This did remove the majority, but not all of the rogue pairs. Therefore, the closest pairs (at angular separations of less than $3r_p$) in each category were examined by hand and 69 additional SDSS objects were removed.

4.3 Determination of α angle

As discussed in Section 3, we need to estimate the value of α, the maximum angle of inclination at which the spirals have a measured spin orientation rather than being classified as 'edge-on' spirals. To do this, we use the adaptive second moments (Bernstein & Jarvis

Table 1. This table shows the basic information about the data sets used in this work. When calculating the mean redshift, only subset of the galaxies with redshift is used, and we average over galaxies and not galaxy redshifts. We also report values of best-fitting $\Delta\chi^2$, Bayesian evidences and parameters of our fits. Note that evidence here is the evidence ratio and not its logarithm.

Property	Angular	Real space	Projected
Number of pairs	34 031	8005	24 271
Number of gal.	20 827	7979	25 272
Mean z	0.08	0.05	0.07
$\Delta\chi^2$ exponential	9.76	9.12	5.89
Evidence ratio	9	6	1
a	$0.94^{+0.39+0.54+0.56}_{-0.48-0.80-0.92}$	$0.35^{+0.19+0.42+0.86}_{-0.16-0.27-0.34}$	$0.55^{+0.60+0.89+0.95}_{-0.47-0.54-0.55}$
a	$23.41^{+11.37+42.47+73.24}_{-6.35-13.04-22.41}$	$0.37^{+0.16+0.41+0.61}_{-0.11-0.23-0.37}$	$0.02^{+0.08+0.37+0.47}_{-0.01-0.02-0.02}$
$\Delta\chi^2$ Gaussian	9.82	11.52	6.71
Evidence ratio	11	16	1
a	$0.60^{+0.25+0.38+0.40}_{-0.27-0.49-0.58}$	$0.24^{+0.10+0.22+0.41}_{-0.09-0.17-0.23}$	$0.44^{+0.31+0.51+0.55}_{-0.37-0.43-0.44}$
b	$26.76^{+9.36+34.11+68.56}_{-6.35-13.71-25.42}$	$0.44^{+0.13+0.31+0.52}_{-0.10-0.20-0.42}$	$0.02^{+0.05+0.36+0.47}_{-0.01-0.02-0.02}$

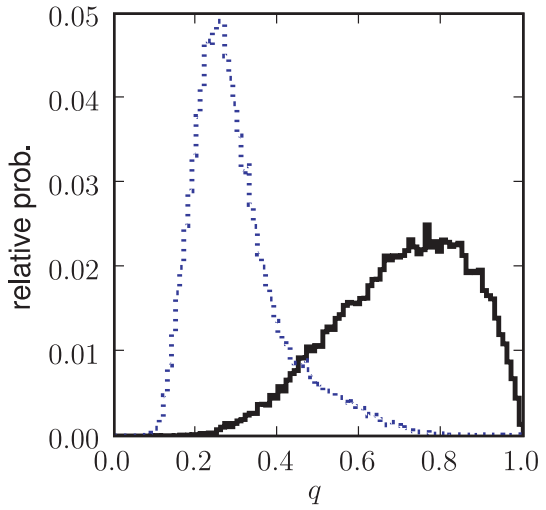

Figure 2. This figure shows the histogram of distributions of q values for galaxies classified as face-on spirals (solid black, Classifications 2 and 3) and edge-on spirals (dashed blue, Classification 4)

2002) from the SDSS pipeline, namely $e_\times$ and e_+, to calculate the axis ratio, following Ryden (2004):

$$q = \left(\frac{1-e}{1+e}\right)^{1/2}, \tag{25}$$

where $e = \sqrt{e_\times^2 + e_+^2}$. In Fig. 2, we show the distribution of q values for spirals galaxies classified as face-on (of either spin orientation) and edge-on spirals. As expected, the two populations occupy the two corners of possible values of q, but there is a significant overlap. Intrinsic ellipticities, non-zero thickness of the disc and potential human-induced selection effects likely complicate things. We have attempted to model intrinsic ellipticities in the spirit of Giovanelli et al. (1997), but difference was negligible.

A plausible range of the cut-off q is 0.2–0.5, giving the values of α between 60° and 80°. If we numerically integrate equation (13) as explained in Section 3.1, we get

$$\eta(r) \sim m c^{\rm meas}(r), \tag{26}$$

with the value of m between ∼0.6 and ∼0.9. We will assume a systematic bias associated with this effect to be $m = 0.75 \pm 0.15$. Adding the effect of the random reversing of galaxy spins to this and allowing a liberal 50 per cent enhancement of the systematic error due to an *ad hoc* assumption in equation (15), we arrive at

$$\eta(r) = (0.9 \pm 0.3) c^{\rm meas}(r). \tag{27}$$

4.4 Correlation function measurement

Our basic method is to measure $c(r)$ and its errors and then to infer constraints on $\eta(r)$.

To measure $c(r)$, we note the following. For a pair of galaxies, whose spins are $\tilde{s}_i$ and $\tilde{s}_j$, the product $\tilde{s}_i\tilde{s}_j$ can be either +1 or −1 with probabilities $p_{\pm1}$. Since $p_{+1} + p_{-1} = 1$ and the expectation value of $\langle\tilde{s}_i\tilde{s}_j\rangle = p_{+1} - p_{-1} = c(r)$, it follows that

$$p_{\pm1} = \frac{1 \pm c(r)}{2}. \tag{28}$$

Therefore, one can write the likelihood function for $c(r)$ as

$$P[c(r)|{\rm data}] \propto P[{\rm data}|c(r)] = \prod_k \left[\frac{1 + d_k c(r_k)}{2}\right], \tag{29}$$

where index k runs over all pairs of galaxies in the sample and $d_k = \tilde{s}_i\tilde{s}_j$ is the spin product for the kth pair whose distance is r_k. In practice, we work with the log likelihoods

$$\log P[c(r)|{\rm data}] = \sum_j \log[1 + d_j c(r_j)] + {\rm const.} \tag{30}$$

We use three possible forms for $c(r)$. First, we assume a stepwise shape for $c(r)$ and measure it in bins. Secondly, we use two two-parameter families of curves that seem to describe our data fairly well: an exponential

$$c(r) = \min\left\{1, ae^{-r/b}\right\} \tag{31}$$

and a Gaussian

$$c(r) = ae^{-r^2/2b^2}. \tag{32}$$

This parameter space is so small that it can be efficiently explored using grid-based methods, and more advanced Markov chain methods are not necessary.

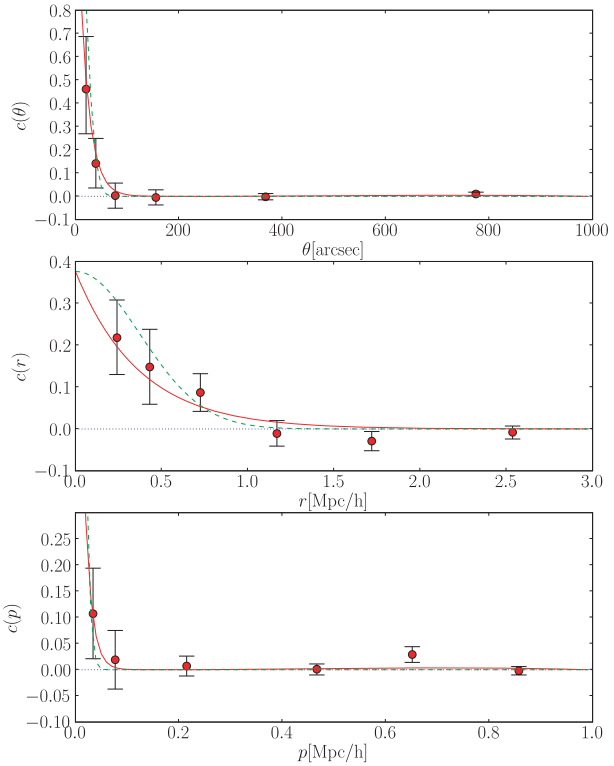

Figure 3. This figure shows the constrains on the binned correlation function c for angular (top panel), redshift (middle panel) and projected (bottom panel) spaces. Two lines correspond to our best-fitting exponential (solid red) and Gaussian (dashed green) fits.

5 RESULTS

In Fig. 3, we plot the results of our binned estimation of $c(r)$. From the two figures, it is immediately clear that there is a hint of an excess at low values of r. The statistical significance of this excess is marginal, at about $\Delta\chi^2$ of 7.5, 14.2 and 5.6 for angular, real and projected distances with six extra degrees of freedom associated with six bins. This corresponds to 2σ–3σ detection in the redshift space but a non-detection in other spaces.

To understand this excess better, we calculate the probability contours on the $a - b$ plane using exponential and Gaussian likelihoods. These are plotted in Fig. 4 and the relevant numbers are given in Table 1. How significant are these detections? The improvement in χ^2 is between nine and 12 with respect to zero correlation in angular and redshift cases with two free parameters. Within a frequentist approach, this is significant at 2σ–3σ level. The excess at low redshift is not significant in the case of projected distances, although visually, the low-distance points are not incompatible with an excess.

A more appropriate statistical procedure is the Bayesian evidence (Slosar et al. 2003; Beltrán et al. 2005; Trotta 2007) which we calculate for all our two model parametrizations and are also shown in Table 1. These can be calculated exactly for a simple problem like ours. Evidence depends weakly on the prior size, and in this we chose the prior on a between 0 and 1/1.5 for Gaussian/exponential case and b between 0 and 1000 arcsec or 1 or 0.5 Mpc h^{-1} projected. Regardless of the exact number employed, the evidence ratio is between a few and a few tens units implying a weak evidence or a hint for angular and redshift spaces, but not for the projected space. This is consistent with results from the frequentist approach above.

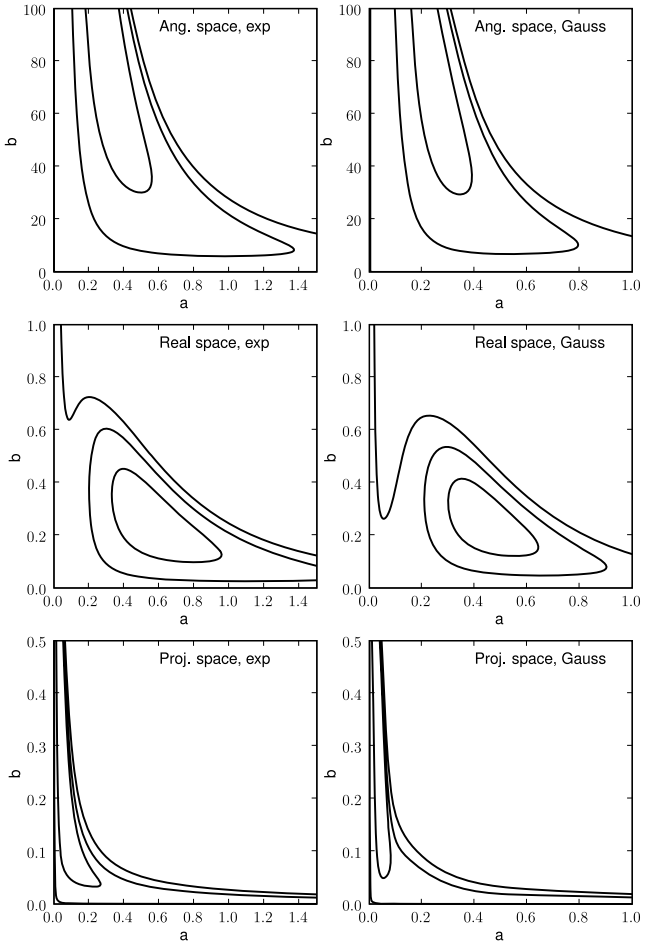

Figure 4. This figure shows the constrains on the a–b plane for all data sets and models under consideration. Thick lines enclose 68.3, 99.4 and 99.8 per cent likelihood volume for the weighted sample. Thin lines are the same for unweighted sample. The top and bottom rows show results in real and angular spaces, respectively. The left- and right-hand columns are the exponential and the Gaussian fittings exponentially.

Finally, we acknowledge the fact that the exponential and Gaussian forms were chosen a posteriori after seeing the data, and hence the improvements in fits contain a subjective a posteriori factor.

5.1 Systematics

We can now briefly discuss some of the main systematic effect that might affect our measurements.

Rogue pairs. As discussed in Section 4.1, we manually looked at all pairs in the clean sample and discarded rogue pairs. It is an important systematic check, because we have at the same time convinced ourselves that manually classifying a small subset (80 galaxies) of the total sample gave consistent results.

Weighting. Repeating our measurements with unweighted data, changes result by less than 5 per cent.

Cleanliness level. We have repeated the analysis with the superclean sample. There are many fewer galaxies in the superclean sample (Lintott et al. 2008) and so the statistical significance decreases considerably. We have no significant detection in any of the spaces considered. The error bars increase by a factor of 2 to 2.5, but the central values in individual bins remain consistent. While

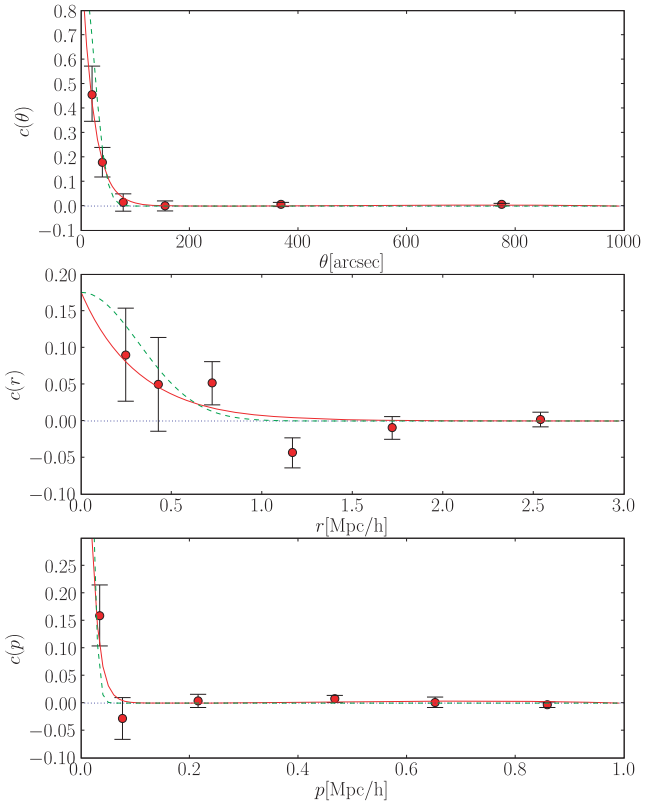

Figure 5. Same as Fig. 3, but for *cleanish* sample. These results are likely to be affected by the rogue pairs (see text for discussion).

the statistical power is decreased, the final signal is consistent with the results presented above.

We have also repeated our measurements with the *cleanish* sample that requires 60 per cent of votes to agree. The results imply strong detections in both angular and projected samples, but with a lower significance in redshift space. We show their results in Fig. 5. The high confidence with which the results are detected in angular and projected spaces is likely to be deceiving, as the rogue pairs have not been manually cleaned for these samples. The decrease of signal in the redshift space indicates that the signal is indeed getting lower due to the noise introduced by low significance.

Selection effects. Another important question is whether there is any physical difference between our redshift sample and the angular sample, and how do these samples compare to the general SDSS sample. To do this, we have divided the galaxies that formed our pairs at closest distances into those for which we have redshift information and those for which redshift information is not available. When comparing colours and magnitudes, we find that there is no evidence that objects with and without redshifts are drawn from different magnitude, $u - r$ colour or petro radius size distributions. The reason for some objects lacking redshift is therefore probably incompleteness due to fibre collisions. We therefore find no evidence that the correlation in the angular sample is of a different physical origin than the correlation in the real-space sample.

6 DISCUSSION AND CONCLUSIONS

We are now in a position to make a synthesis of our results. The redshift-space results show that there is a significant correlation of $c(r) \sim 0.15$ in the projected galaxy spins up to the $\sim 0.5h$ Mpc^{-1}. The angular correlation shows larger correlations of $c(\theta) \sim 0.4$ that are significant correlations up to 30 arcsec, which roughly corresponds to projected distances of about 0.03 Mpc h^{-1}, since mean redshift of 0.08 corresponds to a distance of $\sim$230 Mpc h^{-1}. This is consistent with the redshift-sample results – both exponential and Gaussian fits do predict $c(r)$ to rise to $\sim$0.3–0.4 as r goes to zero. A consistent picture is therefore the following. The angular sample detects correlations at the shortest distances, where majority of pairs are physical associations, but these get diluted at larger distances due to interlopers. The redshift-space correlations track these correlations to larger physical distances. Projected space pairs do not have enough signal-to-noise ratio to detect these correlations at high significance.

How do these results compare to theoretical predictions? Simple models as those suggested in Pen et al. (2000) (equations 4 and 5) predict a vanishing η, and hence we have directly detected a non-random distribution of inertia tensors. Within the standard model, the reason for correlations of moments of inertia is the correlations of these with the (slowly varying) tidal field. On the other hand, if moments of inertia are *perfectly* aligned with the tidal field, the tidal torque cannot produce any angular momentum, and therefore the resulting angular momentum is due to the residual 10 per cent of misalignment (Porciani et al. 2002b). The stunning outcome of our result, if confirmed, is that even these 10 per cent misalignments are correlated from (sub)halo to (sub)halo.

What is also interesting is, however, that in Porciani et al. (2002a), η correlations have not been detected in simulations at $z = 0$ at all separations. In particular, $\eta < 0.02$ at $r = 0.5$ Mpc h^{-1}. A virtually identical result has been found by Bailin & Steinmetz (2005), who also find $\eta < 0.02$ at $r = 0.5\,\mathrm{Mpc}\,h^{-1}$ (our η is their ξ_{LL}). This is in tension with our results even after conversion factors in equation (27) are taken into account. There are many reasons that explain why our results are not directly comparable to the above work. First, they are comparing individual dark matter haloes. In our case, we see the signal at pair separations of less than 1 Mpc. At such distances, one-halo pairs (pairs of galaxies that reside in the same dark matter halo) dominate over two-halo pairs (pairs in which two galaxies occupy two different haloes). By selecting spiral galaxies, we are essentially selecting pairs that are composed of satellites residing in the same halo, rather than pairs compromising central halo galaxies. The latter are bright ellipticals and hence inaccessible in using our method. Unfortunately, not very much theoretical work has been done for spin correlations of substructure. The most relevant paper in the literature is Lee, Kang & Jing (2005), which, however, still uses the chirality agnostic model of Pen et al. (2000) and does not calculate the chiral correlation function. More work on the theoretical side and N-body side is required to understand the implications of our results. Hopefully, the results could be turned around and help us understand what kind of substructure spiral galaxies occupy in a typical dark matter halo.

It is therefore imperative that our observational protocol is simulated on a large enough N-body simulation, for example the *Millennium Simulation* (Springel et al. 2005) or *MareNostrum Universe* (Gottlöber & Yepes 2007) simulations. There haloes and subhaloes hosting spiral galaxies can be identified, and those whose inclination with respect to a given observer is small enough to be considered face-on should be correlated. This would result in a quantity $c(r)$ that is directly comparable to the observables that we constrain with the Galaxy Zoo data.

Another interesting aspect of our results is that, for spiral galaxies, we essentially exclude large and random misalignments between gas and dark matter angular momenta. Since the dark matter is

dynamically dominant, gas angular momenta can only be correlated if they are so due to the correlations between dark matters.

Finally, it is tempting to combine our measurements with the ellipticity measurements to improve signal-to-noise ratio and remove some systematic. Note, however, that our one-bit signal divides a four-fold degeneracy into a two-fold one, and thus this is a non-trivial task, which will be left for the future.

To conclude, we have tentatively detected a chiral correlation function in the spins of spiral galaxies. This correlation function vanishes in the simplest models based on the TTT. Our results indicate that moments of inertia of protohaloes that end up hosting the spiral galaxies are correlated at distances less than $\sim 0.5\ \mathrm{Mpc}\, h^{-1}$. These short distances imply that these protohaloes are often likely to be substructures of massive haloes. More work is required to understand these results at a quantitative level.

ACKNOWLEDGMENTS

We would be nowhere without the amazing contributions from all the Galaxy Zoo (GZ) members, forum contributors and other essential volunteers!

AS thanks Cristiano Porciani for bringing up the important distinction between one- and two-halo pairs and Uroš Seljak for useful literature tips.

AS is supported by the inaugural BCCP Fellowship. CJL acknowledges support from a STFC Science in Society Fellowship.

REFERENCES

Adelman-McCarthy J. K. et al., 2008, ApJS, 175, 297
Bailin J., Steinmetz M., 2005, ApJ, 627, 647
Barnes J., Efstathiou G., 1987, ApJ, 319, 575
Beltrán M., García-Bellido J., Lesgourgues J., Liddle A. R., Slosar A., 2005, Phys. Rev. D, 71, 063532
Bernstein G. M., Jarvis M., 2002, AJ, 123, 583
Catelan P., Theuns T., 1996, MNRAS, 282, 436
Doroshkevich A. G., 1970, Astrofizika, 6, 581
Giovanelli R., Haynes M. P., Herter T., Vogt N. P., Wegner G., Salzer J. J., da Costa L. N., Freudling W., 1997, AJ, 113, 22
Gottlöber S., Yepes G., 2007, ApJ, 664, 117
Groth E. J., Juszkiewicz R., Ostriker J. P., 1989, ApJ, 346, 558
Hamilton A. J. S., 1998, in Hamilton D., ed., The Evolving Universe. Linear Redshift Distortions: a Review. Astrophysics and Space Science Library, Vol. 231, p. 185
Heavens A., Peacock J., 1988, MNRAS, 232, 339
Hoyle F., 1949, in Burgers J. M., van de Hulst H. C., eds, Problems of Cosmical Aerodynamics. Central Air Documents Office, Dayton, Ohio, p. 195
Land K. et al., 2008, MNRAS, 388, 1686
Lee J., Erdogdu P., 2007, ApJ, 671, 1248
Lee J., Pen U.-L., 2001, ApJ, 555, 106
Lee J., Pen U.-L., 2002, ApJ, 567, L111
Lee J., Kang X., Jing Y. P., 2005, ApJ, 629, L5
Lintott C. J. et al., 2008, MNRAS, 389, 1179
Oat C., Sander P. V., 2007, in I3D '07: Proc. ACM Symp. Interactive 3D Graphics and Games Ambient Aperture Lighting. ACM, New York, p. 61
Pasha I. I., Smirnov M. A., 1982, Ap&SS, 86, 215
Peebles P. J. E., 1969, ApJ, 155, 393
Pen U., Lee J., Seljak U., 2000, ApJ, 543, L107
Petrosian V., 1976, ApJ, 209, L1
Porciani C., Dekel A., Hoffman Y., 2002a, MNRAS, 332, 325
Porciani C., Dekel A., Hoffman Y., 2002b, MNRAS, 332, 339
Ryden B. S., 2004, Astrophys. J., 601, 214
Schaefer B. M., 2008, preprint (arXiv:0808.0203)
Slosar A. et al., 2003, MNRAS, 341, L29
Springel V. et al., 2005, Nat, 435, 629
Sugai H., Iye M., 1995, MNRAS, 276, 327
Trotta R., 2007, MNRAS, 378, 72
Trujillo I., Carretero C., Patiri S. G., 2006, ApJ, 640, L111
van den Bosch F. C., Abel T., Croft R. A. C., Hernquist L., White S. D. M., 2002, ApJ, 576, 21
White S. D. M., 1984, ApJ, 286, 38
York D. G. et al., 2000, AJ, 120, 1579

APPENDIX A: CALCULATIONS OF 'SEEN PAIR' PROBABILITIES

The question that we want answer is: for a given pair of galaxies, whose spin vectors are at angle $\cos\theta = \mu$, what is the probability of an observer seeing the pair with the same sense of galaxy rotation or not seeing them at all due to selection criteria?

If $\alpha = \pi/2$, the result can be obtained by considering each spin in turn. The first spin divides the unit sphere of possible observer directions into two half spheres, depending on the sign of its projected spin. When two spins are considered, the intersection of the two half spheres is two lunes. The thickness of the lune of opposite spins is θ/π, leading to the result in equations (11) and (12).

When $\alpha < \pi/2$, the dividing line between the two half spheres becomes a band of angular thickness $2(\pi/2 - \alpha)$ and the two half spheres shrink to two spherical caps of radius α. The overlapping area of the spherical caps separated by θ is $4\pi f(\theta)$, where $f(\theta)$ is given by equation (22) (Oat & Sander 2007).

If $\alpha > \pi/4$, the intersection of the four cups gives four 'trimmed' lunes. There are three possibilities.

(i) $\theta < 2(\pi/2 - \alpha)$. Both spins are in roughly same direction, and the opposite spin 'trimmed' lunes are squeezed to zero area. Hence, $P_{+1} = f(\theta)$ and $P_{-1} = 0$.

(ii) $2(\pi/2 - \alpha) < \theta < 2\alpha$. General situation in which all four 'trimmed' lunes have finite area. We have $P_{+1} = f(\theta)$ and $P_{-1} = f(\pi - \theta)$.

(iii) $\theta > 2\alpha$. Both spins are in roughly opposite directions and opposite spin 'trimmed' lunes are squeezed to zero area. In this case, $P_{+1} = 0$ and $P_{-1} = f(\pi - \theta)$.

These results imply equations (20) and (21). We have tested these analytical predictions using a Monte Carlo code.

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