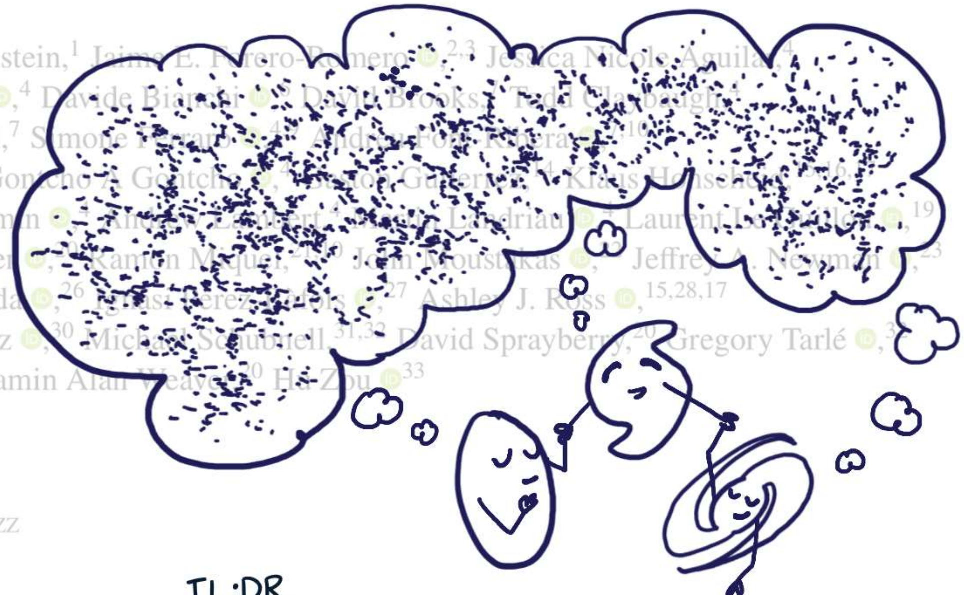
# Tiny groups of galaxies remember their cosmic origins

~~Detection of the large-scale tidal field with galaxy multiplet alignment in the DESI Y1 spectroscopic survey~~

Claire Lamman,[1★] Daniel Eisenstein,[1] Jaime E. Forero-Romero,[2,3] Jessica Nicole Aguilar,[4] Steven Ahlen,[5] Stephen Bailey,[4] Davide Bianchi,[6] David Brooks,[7] Todd Claybaugh,[4] Axel de la Macorra,[8] Peter Doel,[7] Simone Ferraro,[4,9] Andreu Font-Ribera,[10] Enrique Gaztañaga,[11,12,13] Satya Gontcho A Gontcho,[4] Gaston Gutierrez,[14] Klaus Honscheid,[15,16,17] Cullan Howlett,[18] Anthony Kremin,[4] Andrew Lambert,[4] Martin Landriau,[4] Laurent Le Guillou,[19] Michael E. Levi,[4] Aaron Meisner,[20] Ramon Miquel,[21,10] John Moustakas,[22] Jeffrey A. Newman,[23] Gustavo Niz,[24,25] Francisco Prada,[26] Ignasi Pérez-Ràfols,[27] Ashley J. Ross,[15,28,17] Graziano Rossi,[29] Eusebio Sanchez,[30] Michael Schubnell,[31,32] David Sprayberry,[20] Gregory Tarlé,[32] Mariana Vargas-Magaña,[8] Benjamin Alan Weaver,[20] Hu Zou[33]

*Affiliations are listed at the end of the paper*



Accepted XXX. Received YYY; in original form ZZZ

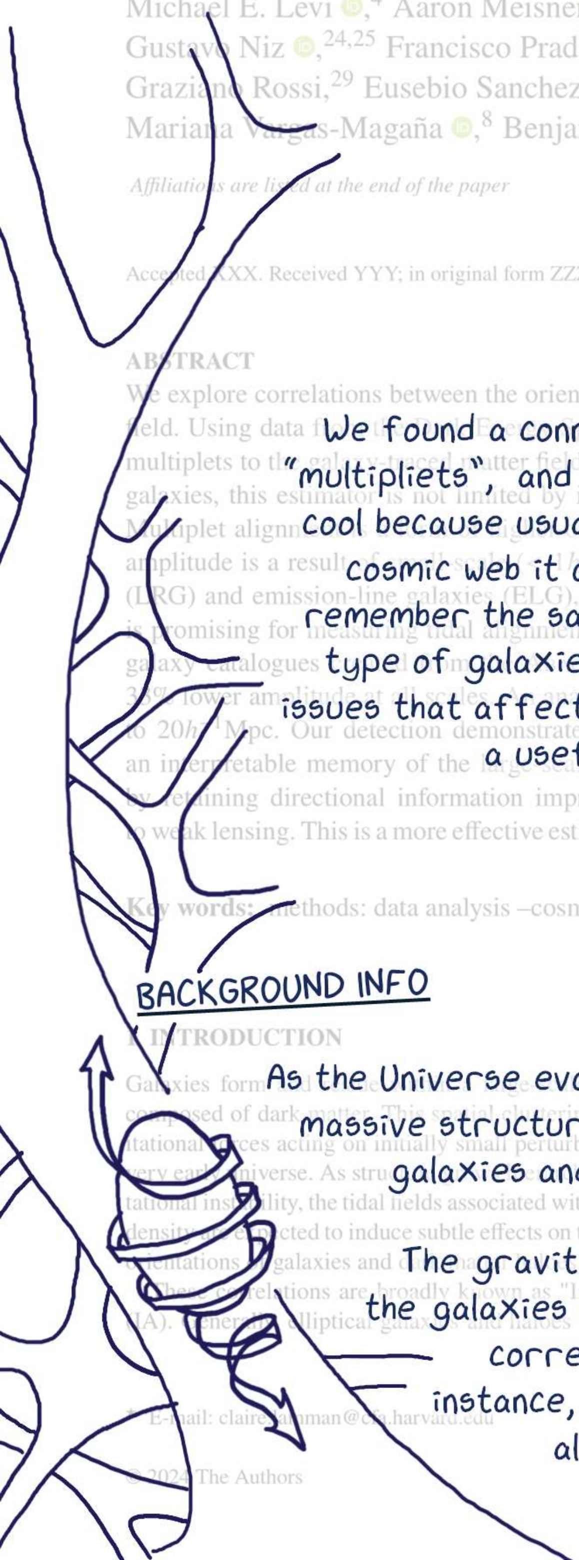

## TL;DR

We found a connection between little groups of galaxies, or "multipliets", and the largest structures in the Universe. This is cool because usually stuff on small scales seems to forget the cosmic web it originated from. We find that all multiplets remember the same large-scale structure, regardless of the type of galaxies in them. This method doesn't have the main issues that affect similar types of measurements, so it could be a useful way to measure the cosmic web.

The more we know about the cosmic web, the more we know about the stuff that shapes it: like dark energy


**Key words:** methods: data analysis –cosmology: observations – large-scale structure of Universe – – cosmology: dark energy

## BACKGROUND INFO



As the Universe evolves, gas and dust fall along massive structures of dark matter, forming galaxies and illuminating the cosmic web.

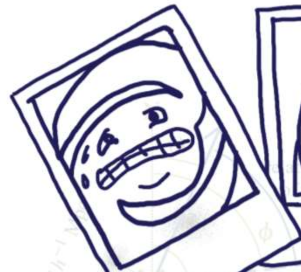
The gravity of the cosmic web affects the galaxies that form along it, creating correlations between the two. For instance, a long galaxy will tend to be aligned along a cosmic strand.

Although we can't see dark matter (it's dark!), we can look at galaxy shapes to figure out more about the invisible structure around it.

However, there are some issue with this approach....

★ E-mail: claire.lamman@cfa.harvard.edu

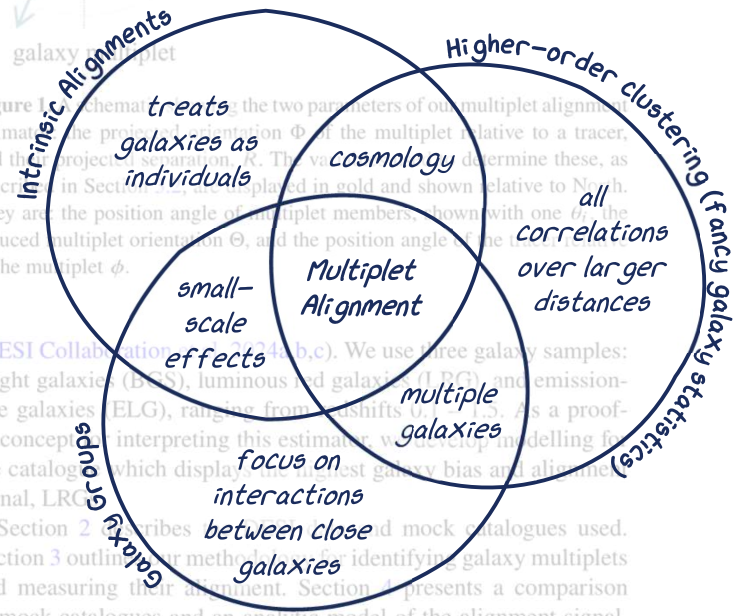
Many people have measured connections between galaxy shapes and the large-scale structure of the Universe. The two main difficulties they face are:



All galaxies are beautiful, but some aren't very photogenic

1. You need really good pictures of galaxies to precisely measure their shapes (this is hard).

2. Many galaxies show no correlation. No one has been able to make this measurement with spiral galaxies.




...so here's our idea:

Instead of galaxy shapes, let's try using the orientation of tiny groups of galaxies

(We'll explain exactly what this means later)

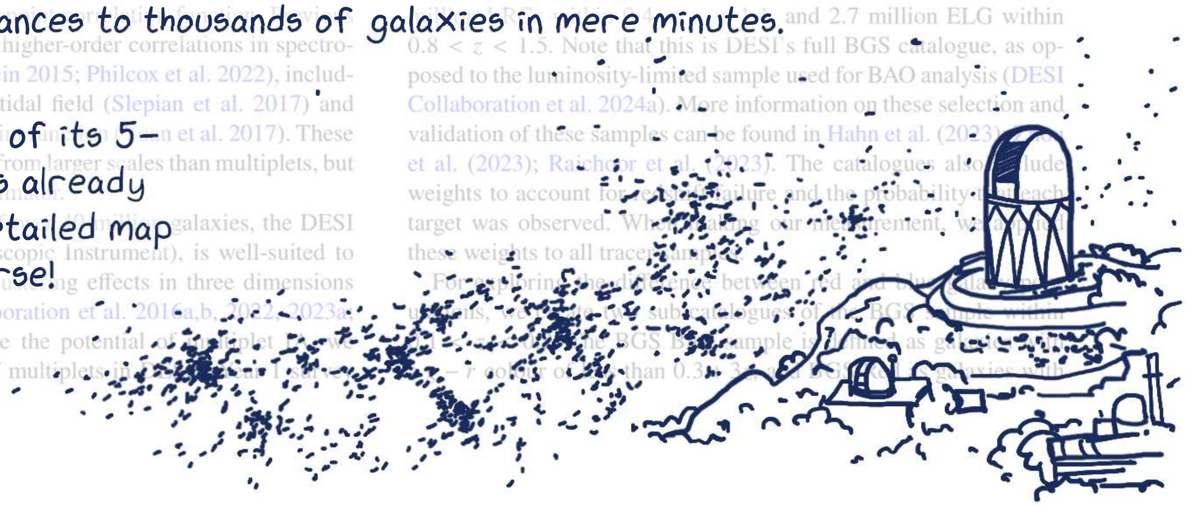
This idea is built upon the work of past scientists and has many connections to it.

## WHERE DOES OUR DATA COME FROM?

Only the most AMAZING TELESCOPE EVER!!!! (although I may be a bit biased)

DESI, or the "Dark Energy Spectroscopic Instrument", sits atop a mountain in Arizona. Inside it are 5000 individually-controlled robots. These allow us to measure the distances to thousands of galaxies in mere minutes.

DESI is in the middle of its 5-year survey, but has already created the most detailed map of the nearby Universe!



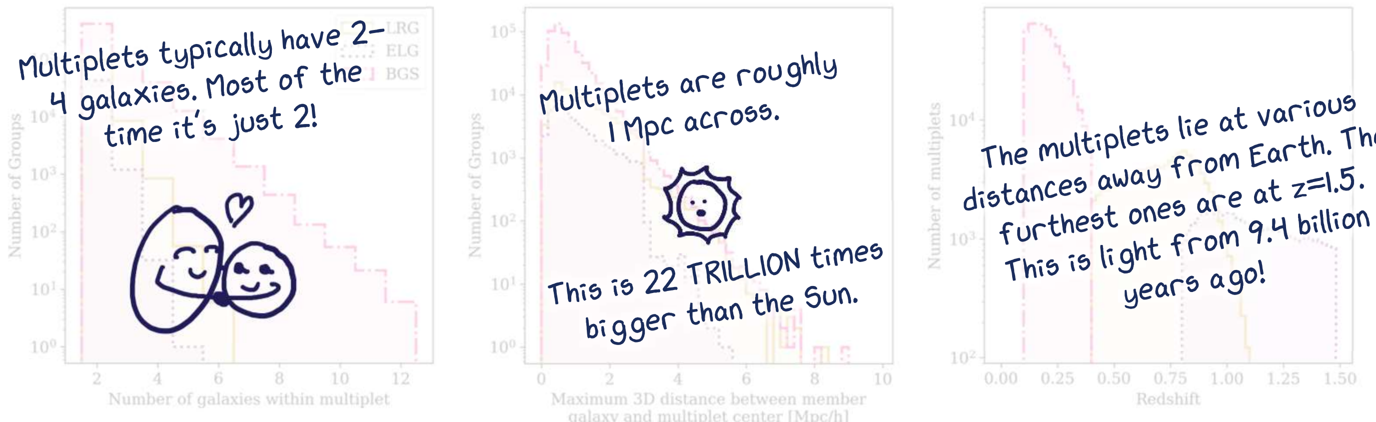



**Figure 2.** Demogr[illegible] ... composed of only two members, even for the densest sample, BGS, where 70% of multiplets are galaxy pairs. The spatial size of multiplets is shown in the middle panel, which is described by the maximum 3D distance betw[illegible] ... is a drop around 3 $h^{-1}$Mpc, corresponding to the maximum distance to the center for a pair of galaxies based on our multiplet definition (Section 3.1). The right panel shows the redshift distribution of multiplets.

These plots show information about the galaxy multiplets we find, which are described in the text below.

| Galaxy type | Redshift Range | N galaxies | N galaxy multiplets | Volume [$Gpc^3 h^{-3}$] |
|---|---|---|---|---|
| ELG | $1.1 < z < 1.5$ | 1.5 M | 21 K | 67.8 |
| ELG | $0.8 < z < 1.1$ | 1.2 M | 22 K | 35.8 |
| [illegible] | [illegible] | [illegible] | [illegible] | 34.6 |
| [illegible] | $0.4 < z < 1.1$ | 2.2 M | 105 K | 34.6 |
| BGS | $0.3 < z < 0.4$ | 0.6 M | 64 K | 3.2 |
| BGS | [illegible] | [illegible] | [illegible] | [illegible] |
| BGS | $0.1 < z < 0.2$ | 1.4 M | 307 K | 0.5 |
| BGS Blue | $0.1 < z < 0.2$ | 0.56 M | 81 K | 0.5 |
| BGS Red | $0.1 < z < 0.2$ | 0.54 M | 100 K | 0.5 |

| | Type of Galaxy | How far away they are | How dense the sample is |
|---|---|---|---|
| BGS | bright, nearby galaxies | "close" | very dense (brownie) |
| LRG | big elliptical galaxies | far away | less dense (cake) |
| ELG | faint spiral galaxies | REALLY far away | sparse (chocolate mouse) |

**Table 1.** Properties of the DESI catalogues used to identify galaxy multiplets. The right column shows the comoving volume of the sample, estimated from the positions of gala[illegible] ... used to make the BGS Blue and Red samples are described in Section 2.

We found multiplets using samples of many types of galaxies!

a $g-r$ colour greater than $0.5 + 3z$. The colour $g-r$ is computed [illegible] ... the galaxies' spectroscopic redshifts.

[illegible] ... selection and analysis rely on mock catalogues. These are generated with 25 N-body cosmological simulations from AbacusSummit (Hadzhiyska et al. 2021; Maksimova et al. 2021; Yuan et al. 2024). An important aspect of these is [illegible] ... on thousands of objects at once via individually controlled robotic positioners, which place fiber optic cables on galaxies. Their positions are limited to a set patrol radius and by proximity to other fibers, resulting in an under-sampling of highly clustered targets. Here we use the set of mock catalogues prepared for DESI Y1 clustering measurements, and two implementations of fiber assignment. The first [illegible] ... that targets are observed is estimated by running the fiber assignment algorithm with varying target priorities (Lasker et al. 2024). [illegible] ... the probability that galaxies are observed based on the number of survey [illegible] ... (et al., in [illegible]

## WHAT ARE MULTIPLETS AND HOW DO WE FIND THEM?

Multiples are little sets of galaxies: as in a doublet, triplet, quadruplet, etc.

We find them by using the 3D positions of galaxies measured by DESI. The basic idea is to find the closest neighbor to each galaxy and then use an algorithm to identify sets within those connections.

## 3 ALIGNMENT METHOD

### 3.1 Identifying galaxy multiplets

Our measurement is a projected quantity, relating the orientation of multiplets in the plane of the sky as a function of transverse distance (Figure 1). However, we identify small multi[illegible] ... in 3D comoving space using spectroscopic redshifts. Each galaxy is matched to its nearest neighbour and all pairs are limited to a maximum separation in the plane of the sky $r_p$, and along the line of sight, $r_\parallel$. $r_\parallel$ is necessarily larger than $r_p$ to account for the redshift-space distortions created by peculiar velocities of multiplet members. We then find multiplets within these matches using [illegible] Find [illegible] to i[illegible] ... graph of galaxy pairs (Galler & Fisher 1964). We set no maximum for the number of multiplet members. This is similar to the friends-of-friends algorithm used for identifying haloes in N-body simulations and for constructing group catalogues (Davis et al. 1985; Eke et al. 2004; Robotham et al. 2011). Note that unlike these catalogues, our goal is not to identify complete, gravitationally bound objects. We expect even nonviralized objects to contribute to our final measurement and so set no additional criteria such as completeness or velocity dispersion.

To explore the effectiveness of this algorithm to identify distinct multiplets, we created a catalogue of isolated multiplets, consisting only of multiplets where each member was a minimum of $2r_p$ and $2r_\parallel$ away from the nearest non-multiplet member. This had no significant effect on final results. We tested multiplets constructed from varying criteria, bet[illegible] ... $r_\parallel < 12h^{-1}$Mpc. We found no significant effect on the amplitude of the final [illegible] ... cuts to maxi[illegible] ... model in Section 4, LRGs. For all samples, we use $r_p = 1.0h^{-1}$Mpc and $r_\parallel = 6.0$[illegible] ..., such as scale cuts which depend on density or removing very close pairs (see Section 4.1) may improve the SNR and are worth exploring in future works.

We are careful to vary some of the algorithm parameters and make sure these choices don't impact our final results.

Properties of the DESI samples we identified multiplets in are shown in T[illegible] ... shown in Fig. 2. This displays the number of members within each multiplet, which is most often two. It also shows the spatial size of each mul-

This page shows information about the multiplets we find.

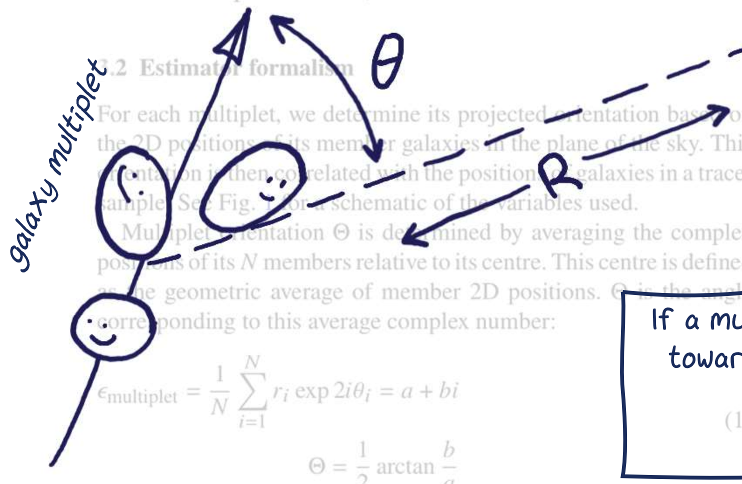
## WHAT EXACTLY ARE WE MEASURING?

...ed by taking the maximum 3D distance between a multiplet member and the multiplet's centre in redshift space. Note that this can be greater than $r_p$ and $r_{\parallel}$ since we only limit the distance between multiplet members, not its overall size.

### 3.2 Estimator formalism

For each multiplet, we determine its projected orientation based on the 2D positions of its member galaxies in the plane of the sky. This orientation is then correlated with the positions of galaxies in a tracer sample. See Fig. 1 for a schematic of the variables used.

Multiplet orientation Θ is determined by averaging the complex positions of its $N$ members relative to its centre. This centre is defined as the geometric average of member 2D positions. Θ is the angle corresponding to this average complex number:




$$\epsilon_{\mathrm{multiplet}} = \frac{1}{N}\sum_{i=1}^{N} r_i \exp 2i\theta_i = a + bi \quad (1)$$

$$\Theta = \frac{1}{2}\arctan\frac{b}{a}$$

For each member $i$, $r_i$ is the projected distance to the multiplet centre ... We do not consider the full ellipticity of the multiplet, i.e. axis ratio, because this is meaningless for multiplets with two members and it is not expected to increase our signal-to-noise ratio. For single galaxies, measurements with the SDSS-III BOSS LOWZ sample have found that ... 2015). We find a similar result with DESI Year 1 LRGs and Legacy Imaging (Fig. 9).

This orientation angle is then measured relative to the tracer sample. In most cases, the tracer sample is the same as the one used to identify multiplets. For each pair, consisting of a galaxy multiplet and ... the difference between the position angle of the multiplet relative to the tracer, $\phi$, and the multiplet's orientation: (2)

The relevant quantity is $\cos(2\theta)$ ... function of the projected separation between the multiplet and the tracer, $R$. This is then averaged over every multiplet-tracer pair. This is similar to conventions in intrinsic alignments and ensures the relative angle is invariant under rotation. Since we are mostly measuring the orientation of a pair of galaxies ... estimator is similar to a "squeezed" three-point correlation function. For multiplets with more than two members, the pair of points describe an N-weighted orientation ... The code for these measurements, and other analysis throughout the paper, is available in the repository spec-IA.

Our measurement comes down to two numbers:

1. The orientation of a multiplet relative to a tracer, θ
2. The distance between a multiplet and a tracer, R

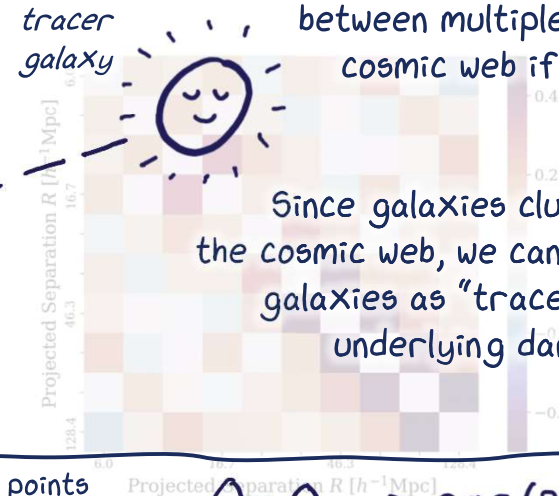
We find millions of multiplet-tracer pairs and find the average orientation as a function of separation.

We expect that multiplets will "point" towards areas of high density, as traced by other galaxies. And that the effect will be strongest when a multiplets is close to a tracer.

Just like long galaxies!

How do we measure correlations between multiplets and the cosmic web if the web is invisible?

Since galaxies cluster along the cosmic web, we can use other galaxies as "tracers" of the underlying dark matter.




If a multiplet points towards a tracer $\rightarrow \theta = 0 \rightarrow \cos(2\theta) = 1$

If multiplets have random orientations, on average: $\rightarrow \cos(2\theta) = 0$

**Figure 3.** The reduced covariance matrix corresponding to the LRG signal ... has been subtracted ... There is covariance in the multiplet alignment measurement between bins of projected separation, particularly in the largest bins.

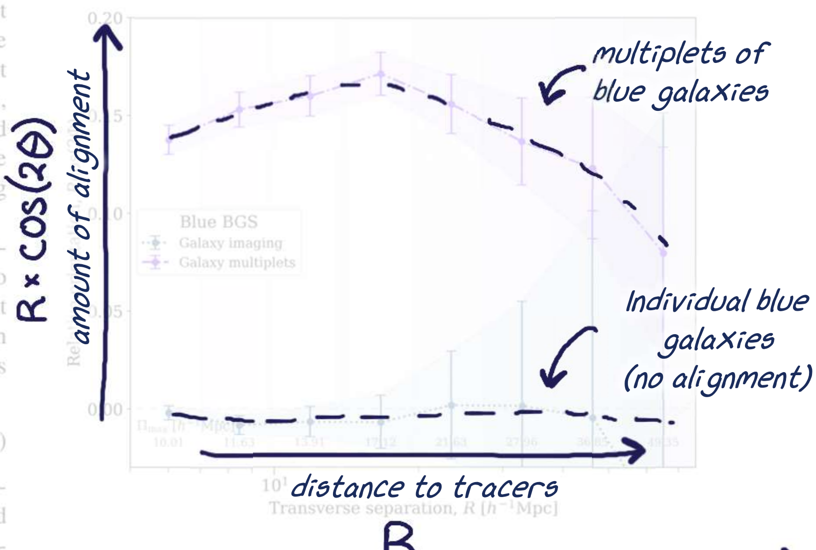



**Figure 4.** A demonstration of the advantages of using multiplet alignment. Here we show the tidal alignment of galaxy and multiplet orientations within a dense, blue sample. The alignment of ... galaxies is highly sensitive to ... and, as expected, consistent with zero. However, the alignment of multiplets displays a clear signal.

Why so blue??

Here we show how we can detect a correlation to the cosmic web using multiplets of blue galaxies. This is much better than the alignment of individual blue galaxies, which is 0.

### 3.3 ...

When measuring the projected orientation of multiplets relative to a tracer catalogue, we limit the multiplet-tracer pairs to a line-of-sight separation that is unique to each bin of projected separation, $\Pi_{\max}(R_{\mathrm{bin}})$. This is to maximise the signal-to-noise of our measurement. In the case of positive tidal alignment, shapes are elongated along the stretching direction of the tidal field. In this situation, the tidal field ... measurable orientation in the plane of the sky. Therefore, multiplet-tracer pairs that are close in the plane of the sky but distant along the line-of-sight direction will have a relatively low contribution to the total alignment

signal. At large ... contribution from radially distant galaxies and it becomes advantageous to increase $\Pi_{\max}$. We chose $\Pi_{\max}(R_{\mathrm{bin}}) = 6h^{-1}\mathrm{Mpc} + \frac{2}{3}R_{\mathrm{bin}}$ based on the signal-to-noise ratio of our final LRG signal. Our model estimate is computed in these same $R$ bins. Throughout plots in this paper, the varying values of $\Pi_{\max}$ are shown through shaded regions and marked explicitly. We use this projected statistic, as opposed to keeping the measurement as a function of $r_p$ and $r_{\parallel}$, because most of the signal is along the LOS for tidal alignments due to the projection of shapes. Additionally, a projected statistic allows for more direct modelling as it is less sensitive to redshift-space distortions (Figure 7).

For each measurement, we separate the multiplet catalogue into 100 sky regions by right ascention and declination, with equal numbers of multiplets in each. The orientation of multiplets ... measured

## WHAT ARE THE MEASUREMENTS?

The measurement are.... On the next page! But there's a sneak-peak here.