MNRAS 000, 1–8 (2025) Preprint 22 April 2025 Compiled using MNRAS LaTeX style file v3.3

We have an idea for improving measurements of pointing galaxies

~~Optimal intrinsic alignment estimators in the presence of redshift-space distortions~~

Claire Lamman,[1★] Jonathan Blazek,[2] Daniel J. Eisenstein[1]
[1]Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[2]Department of Physics, Northeastern University, Boston, MA, 02115, USA

Accepted XXX. Received YYY; in original form ZZZ

SUMMARY

ABSTRACT

We have ideas for improving measurements you've probably never heard of. This paper was written for the tens of people who have. Although not particularly glamorous, technical papers like these can end up being some of the most useful to other scientists. This annotation gives a glimpse at the types of problems we need to think about for developing new methods.

We present estimators for quantifying intrinsic alignments in large spectroscopic surveys that efficiently capture line-of-sight (LOS) infor[illegible] changing the LOS [illegible] Π_{max} value. This is further improved by replacing the flat Π_{max} cut with a LOS weighting based on shape projection and RSD. Although these estimators incorporate additional LOS information, they are projected correlations that exhibit signal-to-noise ratios comparable [illegible] SUMMIT, we evaluate these estimators and provide recommended Π_{max} values and weights for projected separations of $1 - 100h^{-1}$Mpc. These will improve measurements of intrinsic alignments in large cosmological surveys and the constraints they provide for both weak lensing and direct cosmological applications.

TL;DR: When measuring how much galaxies point at each other, you should care more about galaxies that are close together. Except if they're too close, then you should care about ones a bit further away.

Key words: methods: data analysis – statistical – – cosmology: observations – large-scale structure of Universe

1 INTRODUCTION

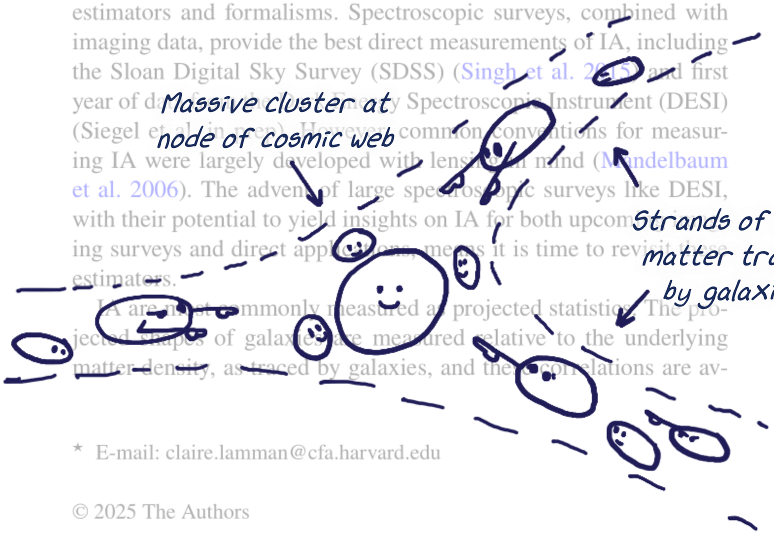
WHY BOTHER?

~~Large-scale structure~~ is subtly imprinted in the orientations of galaxies. The most common form of this intrinsic alignment (IA) is for the long axis of elliptical galaxies to be aligned along the large-scale [illegible] for weak lensing studies; IA can bias the matter power spectrum by 30% as measured by cosmic shear (Hirata et al. 2004), and cosmo[illegible] to IA modeling (Kirk et al. 2015; Dark Energy Survey and Kilo-Degree Survey Collaboration et al. 202[illegible]). In principle, IA can also mea[illegible] scale tidal field, including primordial physics and the nature of dark energy ([illegible] 2015). See Joachimi et al. (2015); Troxel & Ishak (2015) for reviews and Lamman et al. (2024a) for a guide to IA estimators and formalisms. Spectroscopic surveys, combined with imaging data, provide the best direct measurements of IA, including the Sloan Digital Sky Survey (SDSS) (Singh et al. 2015) and first year of d[illegible] Spectroscopic Instrument (DESI) (Siegel et al. in prep). However, common conventions for measuring IA were largely developed with lensing in mind (Mandelbaum et al. 2006). The advent of large spectroscopic surveys like DESI, with their potential to yield insights on IA for both upcoming [illegible]ing surveys and direct applications, means it is time to revisit these estimators.

IA are most commonly measured as projected statistics. The projected shapes of galaxies are measured relative to the underlying matter density, as traced by galaxies, and these correlations are averaged over a total LOS distance $2\Pi_{max}$. Common Π_{max} choices are 60 - 100 h^{-1}Mpc. Some advantages of these projected statistics are that they are less sensitive to LOS uncertainties and can be more directly r[illegible] to transverse modes. As with galaxy clustering, several recent studies explore higher-order statics which take advantage of the full 3D informati[illegible] Kurita et al. 2021; Pyne et al. 2022; Bakx et al. 2025). Unlike galaxy clustering, the nature of tidal alignments and projected shapes leads to most i[illegible]. In the case of tidal alignment, on average the long axis of galaxies are oriented towards [illegible] only the projected shape of galaxies, this results in a weaker signal along the LOS, especially at small transverse separations. Therefore, an IA quadrupole estimator that essentially up-weights the signal in the transverse direction can increase the signal-to-noise ratio (SNR) of measurements (Singh et al. 2024).

As the cosmic web of dark matter grows, the shapes and orientations of galaxies tend to align with it. They point at each other (but really they're pointing at big invisible clumps of matter which tend to form lots of galaxies).

This is a real pain for people who care about weak lensing: a way to measure dark matter by the tiny little ways that its gravity messes with the light of galaxies. Most of the methods we use to measure galaxy alignments come from trying to solve this problem.



A disadvantage of using these 3D estimators statistics is redshift-space-distortions (RSD). In redshift space, galaxy peculiar velocities create a "smearing" along the LOS below separations of about 3 h^{-1}Mpc (Jackson 1972), known as the fingers-of-god effect (FOG). At larger scales, infall into overdense regions produces a LOS "squashing" (Kaiser 1987), known as the Kaiser effect. This means the IA signal does not follow a perfect μ relation in redshift space, and 3D IA correlations must successfully model these effects.

Here we present a set of alternative estimators that preserve the advantages of both classic and 3D measurements: a Π_{max} which varies with tr[illegible] shape projection and RSD. They have the potential to provide similar [illegible]

But now we have big, absolutely GINORMOUS surveys of galaxies and we have a chance to be a little more clever about how we make measurements....

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

Variable	Description	Equation
r_p	projected separation [illegible]	
$r_\parallel$	separation along the LOS in real space	
$s_\parallel$	separation along the LOS in redshift-space	
r	[illegible]	
θ	angle with respect to the LOS, $0 < \theta < \pi$	
μ	characterizes θ, $\mu = \cos\theta$	
$\mathcal{E}_+(r)$	projected ellipticity relative to tracer	1
ξ_{g+}	IA correlation function	2
w_{g+}	projected IA correlation function	4
[illegible]	[illegible] IA correlation functions	8

Table 1. A guide to variables used frequently throughout this paper.

Measurements of galaxy alignment are tricky because we can only see the *projected* shape of galaxies. >>

Because of this, people often measure projected shapes relative to projected separation (distance between galaxies on the sky), a 2D measurement. But this throws out useful information, so some people have been working on measuring projected shapes relative to the full 3D separation between galaxies. But this comes with its own challenges...

SNR as the IA quadrupole while being less sensitive to redshift errors and relatively insensitive to RSD.

[illegible] logs from the AbacusSummit simulations (Maksimova et al. 2021), a suite of large, high-accuracy simulations designed to meet the anal- [illegible] ments of DESI. While halos display stronger alignment than galaxies, the alignment strength itself will not significantly impact the relative performance of these estimators. The strong alignment and large volume of these simulations is well-suited to making high-[illegible]

Here we come up with a method that tries to get the ♫ best of both worlds ♪.

Throughout the paper we assume the cosmological parameters of $H_0 = 69.6$, $\Omega_{m,0} = 0.286$, $\Omega_{\Lambda,0} = 0.714$.

2 IA ESTIMATORS

2.1 Existing formalism

EXISTNING METHODS

IA is quantified through a class of correlations involving galaxy positions and intrinsic shapes. Here we will exclusively focus on the intrinsic shape - density correlation, which describes the correlation between the projected shapes of galaxies and underlying density, which is typically traced by galaxy positions. Variables are summarized in Table 1 and marked in Figure 1. We use $r_\perp$ to describe trans[illegible] space, and $s_\parallel$ to describe LOS distance in redshift space. r is the 3D separation and μ the angle relative to the LOS.

This section summarizes the main ways that we make measurements now:

2.1.1 Relative Ellipticity

$\mathcal{E}_+(r)$ is the average projected ellitpicity relative to a neighbor galaxy as a function of separation r (Lamman et al. 2023). It can be thought of as average intrinsic shear, ϵ_+. This is averaged over N shape-tracer pairs. For a shape catalog S and tracer catalog D,

$$\mathcal{E}_+(r) = \frac{1}{N} \sum_{i \in S, j \in D} \epsilon_+(j|i) = \frac{S_+D(r)}{DD(r)} \quad (1)$$

1. The average amount that a galaxies tends to point towards its neighbors.

DD and S_+D are pair and weighted pair counts binned in r. This can also be averaged over bins of r_p and/or $s_\parallel$ to obtain completely projected correlations. The relative ellipticity used to weight the counts in S_+D is quantified as $|\epsilon|\cos(2\phi)$. ϕ is the on-sky angle between the galaxy orientation and separation vector to a tracer galaxy and $|\epsilon|$ is the absolute value of the shape's projected complex ellipticity. We define this in terms of the galaxy's projected axis ratio, q, as $(1-q)/(1+q)$, but some other works use $(1-q^2)/(1+q^2)$. Our results for optimal estimators do not depend on this choice.

$\mathcal{E}_+$ is the estimator we use to determine weights and compare Π_{max}

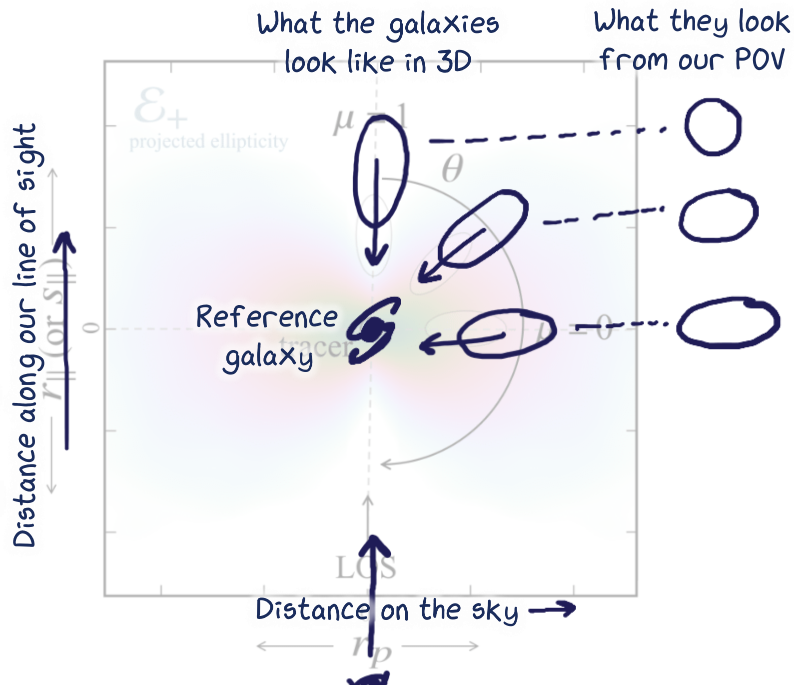


Figure 1. A schematic of coordinates used in measuring IA. r_p is the transverse distance relative to a central tracer, $r_\parallel$ is the distance along the LOS in real space, and $s_\parallel$ is the distance in redshift space. The background shading corresponds to the strength of measured tidal alignment; the extent to which a galaxy [illegible] position will be alignment relative to the tracer. This is fur[illegible] Because we observe projected shapes, the measured alignment will be strongest along r_p and approaches 0 along $r_\parallel$ at small r_p. The ellipticity measurements shown here [illegible] to $100h^{-1}$Mpc in r_p and $r_\parallel$.

This diagram explains the setup: If you have a galaxy in the center, galaxies around it will tend to point at it. But we only observe the shapes of galaxies on the sky! So you get a weird angular dependence if you try to measure this in 3D.

methods as it is related to the most relevant quantity: shape alignment, independent of clustering. Assuming IA noise is dominated by shapes and not clustering (as is almost always the case), the SNR which we use to compare estimators will be the same for both $\mathcal{E}_+$ and the more commonly used projected IA correlation function w_{g+}.

2.1.2 IA correlation function

The IA correlation function incorporates the 2-point galaxy clustering correlation function, $\xi(r)$ and uses random catalogs for the shape and tracer samples, R_S and R_D (Mandelbaum et al. 2006). It is given as:

$$\xi_{g+} = \frac{S_+D - S_+R_D}{R_S R_D} \approx \frac{S_+D}{R_S R_D} \quad (2)$$

This is a generalized form of the Landy-Szalay estimator Landy & Szalay (1993). They key difference compared to $\mathcal{E}_+$ is that ξ_{g+} is not averaged on a per-pair basis. They are related as

$$\xi_{g+}(r) = \frac{DD}{[illegible]} \mathcal{E}_+(r) = (1 + \xi(r))\mathcal{E}_+(r) \quad (3)$$

2. Kinda like #1, but taking into account that you tend to see more galaxies close together. The Universe is clumpy.

Especially for [illegible] along the LOS to obtain the IA projected correlation function (e.g. Blazek et al. (2015)):

$$w_{g+}(r_p) = \int_{-\Pi_{max}}^{\Pi_{max}} dr_\parallel \xi_{g+}(r_p, s_\parallel) \approx (2\Pi_{max} + w_p(r_p))\mathcal{E}_+(r_p). \quad (4)$$

Here, w_p is the 2-point projected clustering correlation function integrated along the same Π_{max} bounds. This can generally be related to the $\mathcal{E}_+$ estimator through $\left(2\Pi_{max} + w_p(r_p)\right)$, but differences in binning pairs along the LOS can propagate as $O(0.1)$ amplitude shifts. w_{g+} is the most common convention for direct IA measurements, and Π_{max} is typically 60 - 100 h^{-1}Mpc. While this estimator may be suitable for photometric surveys where there is a large LOS uncertainty, it is known to lose valuable LOS information when measuring IA in spectroscopic surveys.

3. A fancy way to measure alignments in 3D, taking into account the weird angular dependence that the projection of shapes causes.

2.1.3 IA Multipoles

Higher-order correlations, such as the IA quadrupole, seek to reclaim LOS information lost in projected statistics. Unlike galaxy clustering, the tidal alignment of galaxies is much stronger when $r_p > r_\parallel$ due to the geometry of projected shapes. This effect follows a $1 - \mu^2$ dependence, as demonstrated in Figure 3, and is a result of projecting a 3D quadratic form onto a 2D plane. The projected shape orientation is not just a direct projection of the original axes, but a transformation of the entire ellipsoid surface, which involves a $\sin^2(\theta) = 1 - \mu^2$ term. Singh et al. (2024) proposes a "weighted quadrupole", which gives higher weight to shape-tracer pairs when μ is small. The signal is measured as a function of r, made by summing over bins of μ and weighting by $1 - \mu^2 \propto L^{2,2}$:

$$\tilde{\xi}_{g+,2}(r) = \xi(r)\frac{\sum_{i\in\mu}(1-\mu_i^2)\mathcal{E}_+(r,\mu_i)}{\sum_i^n(1-\mu_i^2)} \quad (5)$$

Singh et al. (2024) also removes all signal with $r_p < 5h^{-1}$Mpc to avoid noise from FOG at low separations and the difficulty of modeling IA in the nonlinear regime. From Equation 18 of that work, the "wedged" quadrupole for ξ_{g+} is:

PROPOSED METHODS

$$\tilde{\xi}_{g+}(r) = \frac{[illegible]}{48}\int d\mu_r\Theta(r_p)L^{2,2}(\mu_r)\xi_{g+}(r,\mu_r) \quad (6)$$

4. The new methods that we propose. #3 considers the shape projection, but there's an addition problem that we want to consider...

Redshift-Space Distortions (RSD)!

Here, $L^{\ell,s_{ab}}$ are th[illegible]. This is different from the quadrupole used to measure RSD, which uses L_2 [illegible] $\Theta(r_p)$ is the function that removes the contribution from separations below $5h^{-1}$Mpc. We demonstrate this weighting in Figure 2. μ-based weightings will optimize SNR for triaxial ellipsoids with tidal alignment in real space. Singh et al. (2024) reports around 2 times greater precision using this wedged quadrupole compared to w_{g+} with $\Pi_{max} = 100h^{-1}$Mpc.

2.2 Proposed estimators

In the presence of redshift-space distortions, there are two additional factors we seek to take into account: the Kaiser effect, which smears positions along the LOS on large scales and the small-scale FOG regime below $r_p < 5h^{-1}$Mpc. The Kaiser effect has a minimal impact on the signal compared to FOG, since it distorts volumes in redshift space and not densities (Kaiser 1987; Scoccimarro 2004). Although modeling is challenging in the FOG regime, most information [illegible] Figures 2 and 3 illustrate how the angular dependence of tidal alignment is affected by RSD using Abacus halos. The mathematical relation $1 - \mu^2$ is replicated by tidal alignment in real space, but breaks down, especially at small separations, in redshift space. Here, close pairs of galaxies with strong alignments are smeared along the LOS. At very large r, the expected $1 - \mu$[illegible]

RSD is an effect we see in galaxy surveys due to how we measure their distances. Basically, the positions we measure are squashed and squished from their real-space positions.

Here we explore two new [illegible]

bin of r_p: a LOS weighting and a variable Π_{max}. Both seek to mitigate the impact of RSD, preserve LOS information, minimize modeling challenges, and improve SNR of direct detections. Both also result in a familiar projected correlation function, similar to w_{g+}, and can be modeled similarly (Section 4).

2.2.1 LOS weighting

A weighted Π_{max}, or $\tilde{\Pi}$, functionally replaces the "top hat" of a set Π_{max} cut with a Gaussian. This Gaussian assigns a LOS weight to each shape-tracer pair within a given r_p bin based on the signal dependence on shape projection and RSD. The result is a projected correlation function where each r_p bin contains a unique LOS weighting.

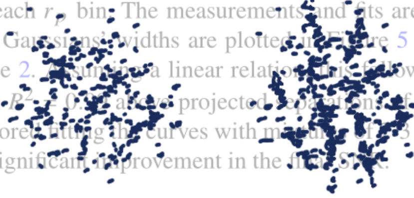
We define optimal weighting as one that maximizes the SNR of the final measurement and assume all noise comes from shape and shot noise. For a given r_p bin, the quantity of interest is the SNR of the signal. This can be characterized by the alignment amplitude, or model re-scaling: $\mathcal{E} = \beta\mathcal{E}_{model}$. The total signal summed over all pairs is proportional to $\mathcal{E}(1+\xi)$ and the noise is proportional to $\sqrt{1+\xi}$. The χ^2 value for pairs in a given r_p bin is summed over their transverse separations $s_\parallel$ as

$$\chi^2_{r_p} = \frac{\left[\sum_{j\in s_\parallel}\mathcal{E}(1+\xi) - \sum_{j\in s_\parallel}\beta\mathcal{E}_{model}(1+\xi)\right]^2}{\sum_{j\in s_\parallel}(1+\xi)}. \quad (7)$$

$\mathcal{E}$ and ξ are both functions of r_p and the j bin in r_p. To minimize this, we obtain weights W which are proportional to the expected signal in redshift space $\mathcal{E}_+(r_p, s_\parallel)$. For a given r_p, the signal is a weighted sum across each pair, where the $s_\parallel$-dependent weight is determined individually for each r_p bin:

$$\tilde{\mathcal{E}}_+(r_p) = \frac{\sum_{j\in s_\parallel}W(r_p,j)\mathcal{E}_+(r_p,j)}{\sum_{j\in s_\parallel}W(r_p,j)}. \quad (8)$$

[illegible] the expectation of that signal, but the scale-dependence of alignment will have a marginal impact on the determined weights, weights are normalized in each r_p bin and the LOS-dependence in each r_p bin is dominated by shape projection and RSD. Therefore we can also derive weights [illegible] Abacus halo catalogs (Section 3), which can be generalized to any elliptical galaxy population. Note that in real space, the derived weights will recover the $1 - \mu^2$ dependence along the $r_\parallel$ direction. We determine $\mathcal{E}_+$ in 2D bins, shown in Figure 2, then fit a Gaussian to the $s_\parallel$-dependent signal to obtain weights for each r_p bin. The measurements and fits are shown in Figure 4. The Gaussians' widths are plotted in Figure 5 and can be found in Table 2. Assuming a linear relation, this follows $\sigma = 3 + (2/5)r_p$, with $R^2 = 0.9$[illegible] above projected separations of $r_p = 2h^{-1}$Mpc. We explored fitting the curves with mixtures of 2-3 Gaussians but found no significant improvement in the final SNR.



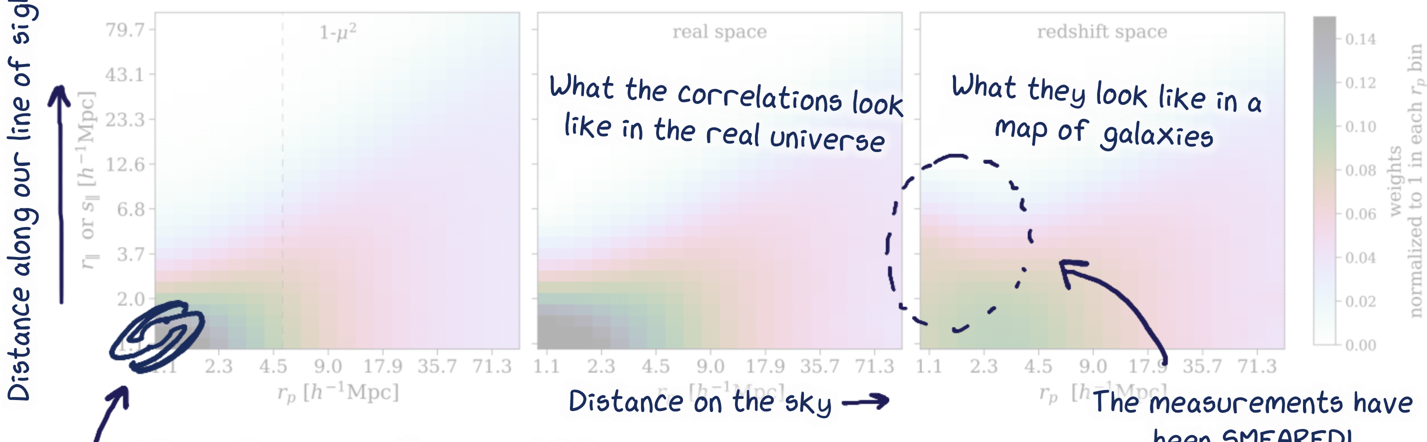
2.2.2 Optimal Π_{max}

In cases where a traditional projected statistic is preferred, we recommend varying the Π_{max} for each r_p bin. A version of this variable Π_{max} cut was used in Lamman et al. (2024b), and can be measured and modeled with most existing IA infrastructure. We use the weights above to produce Π_{max} recommendations for given separation bins [illegible]

[illegible] for a given separation $s_\parallel$ is the

If you've seen my other paper summaries, you might notice that my very first one was about how galaxy alignment can bias RSD. Now I'm looking at how RSD can bias alignments!

These plots attempt to visualize how RSD affects the 3D behavior of galaxy alignment.



If you have a galaxy right here, the color corresponds to how much the (projected) shape of another galaxy would point at it.

Figure 2. Comparisons of tidal alignment and their angular dependence in real and redshift space. The first panel shows the mathematical relation $1 - \mu^2$. This is a visualization of the weighted quadrupole used in Singh et al. (2024), which is cut off below 5 h^{-1}Mpc. This is well-reproduced by the measurement of tidal alignment, $\mathcal{E}_+$ in real space (center panel). The right panel displays alignment binned in redshift space, which smears the signal at low separations. The signal in each pa[illegible] and either LOS distance in real space, $r_\parallel$, or redshift space $s_\parallel$. All signals are normalized in each r_p bin to highlight the LOS difference across all scales.

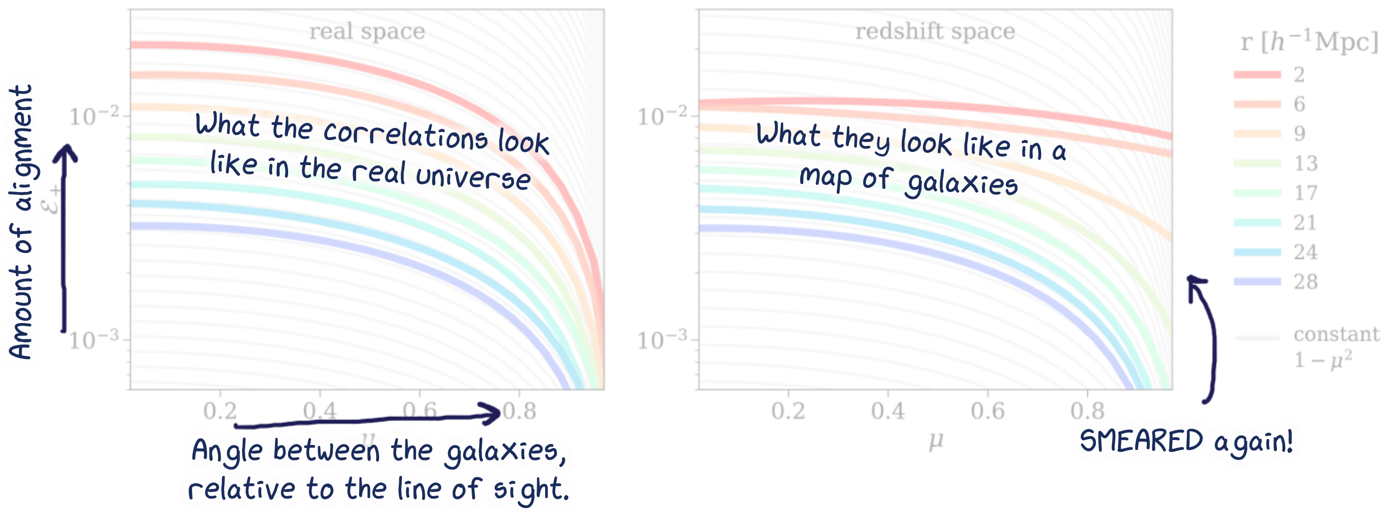


Figure 3. An alternative visualization to Figure 2. Here, each colored line is the mean $\mathcal{E}_+$ measured in a radial annulus with average distance from tracer r as a function of μ. Grey lines show contours of constant $1 - \mu^2$. Without RSD, the angular dependence of these signals is purely a result of shape projection, which follows $1 - \mu^2$. With RSD, the signal is distorted, especially at close separations ($r < 10$ Mpc/h) and close to the LOS ($\mu = 1$). Each line is fit to measurements from AbacusSummit.

These problems only really show up if you try to do 3D measurements. So do you stick to 2D or try to model RSD and get the full information from 3D?

num[illegible] of pairs times the average alignment $(2s_\parallel + \xi(s_\parallel))\mathcal{E}(s_\parallel)$. The shot noise is given by the **total** number of pairs within $s_\parallel$, $\sqrt{2\Pi_{max} + w_p}$. For a given bin in r_p, the SNR as a function of Π_{max} is

$$\mathrm{SNR}_\Pi = \frac{(2s_\parallel + \xi(s_\parallel))\mathcal{E}}{\sqrt{2\Pi_{max} + w_p}}. \tag{9}$$

The peak of this function, the Π_{max} choice to maximize signal-to-shot noise for IA measurements, is shown in Figure 5 and Table 2. This approximately follows the empirical relation $\Pi_{max,optimal}$ = 8 + (2/3[illegible] R^2 [illegible] 2h^{-1}Mpc. It deviates in the FOG regime below 2-3h^{-1}Mpc. On these smaller scales, the optimal Π_{max} increases in order to capture information from clo[illegible] LOS.

3 COMPARING ESTIMATORS

3.1 Halo catalog

We evaluate the above estimators using halos from the AbacusSummit N-body simulations (Maksimova et al. 2021; Garrison et al. 2021). These large N-body simulations allow us to test many iterations of survey-like samples and precisely compare SNR across methods. We do not expect these simulations to accurately model the alignment of galaxies, which have weaker tidal alignment than halos and depend upon baryon dynamics, even at large scales. However, the relative performance of the estimators is independent of the true IA amplitudes, assuming they have similar scale dependence. We [illegible] r_p bins and the $s_\parallel$ dependence of the signal in each bin is dominated by shape projections and RSD.

We [illegible] AbacusSummit base simulations, generated at $z = 0.8$. Each contains 6912^3 particles in a

Here is our idea: keep making measurements as a function of distance on the sky (2D) but incorporate more 3D information by giving extra weight to pairs of galaxies that are close along the line of sight.

We propose giving more weight to galaxies that are close along the line of sight...

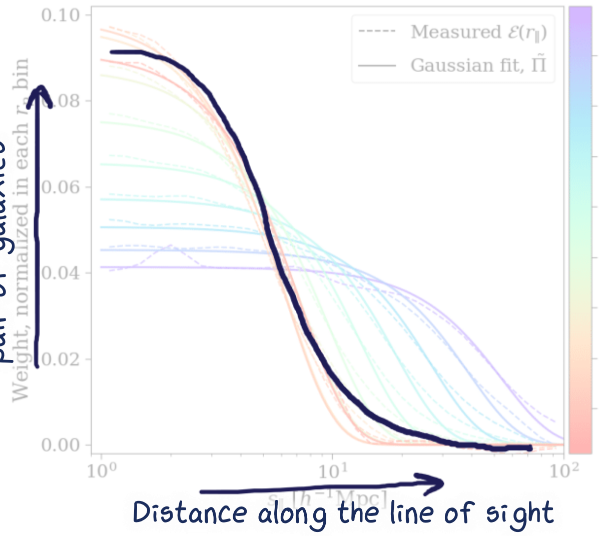


Figure 4. A selection of the LOS weights as a function of $s_\parallel$ in each r_p bin. The dashed lines are measured $\mathcal{E}_+$ and the solid are their Gaussian fits. At smaller projected separations, shown in blue, higher weight is given to close pairs. At very small separations, below 3 h^{-1}Mpc, the functions slightly widen to account for FOG. This general widening of the functions at large separations is a result of the angular dependence of projected shapes, as displayed in Figure 2. The Gaussian widths are plotted in Figure 5.

$r_{p,\mathrm{min}}$	$r_{p,\mathrm{max}}$	σ	Optimal Π_{max}
	$[h^{-1}$Mpc]		
1.0	1.3	5.5	13.6
1.3	1.6	5.1	13.6
1.6	2.0	4.6	11.5
2.0	2.5	4.3	9.7
2.5	3.2	4.2	9.8
3.2	4.0	4.4	10.3
4.0	5.0	4.7	11.0
5.0	6.3	5.2	11.8
6.3	7.9	5.9	12.8
7.9	10.0	6.8	14.0
10.0	12.6	7.9	15.5
12.6	15.8	9.3	17.5
15.8	20.0	11.1	19.9
20.0	25.1	13.3	23.0
25.1	31.6	15.8	26.9
31.6	39.8	18.9	31.9
39.8	50.1	23.0	38.0
50.1	63.1	27.5	45.7
63.1	79.4	32.8	55.5
79.4	100.0	38.4	67.8

Table 2. The points plotted in Figure 5, showing recommendations for summing galaxy alignments along the LOS. σ is the standard deviation of $s_\parallel$-based Gaussian weights for logarithmic bins in r_p. When a traditional Π_{max} is preferred, the last column provides optimal values in these bins.

We're so close to figuring out the question...

Gpc cubic box, which have been used to identify halos with the CompaSO Halo Finder (Hadzhiyska et al. 2021). We select the largest ones, containing an average of about 10,000 particles per halo. Each of the 25 simulations contains 2.5 million halos, with a comoving density of $3.3 \times 10^{-4} h^3$Mpc^{-3}. This is designed to approximately match the density in DESI's Luminous Red Galaxy sample. We

But when galaxies are also close on the sky, ease up! RSD can smear measurements so try to be a bit more generous and even out the weights.

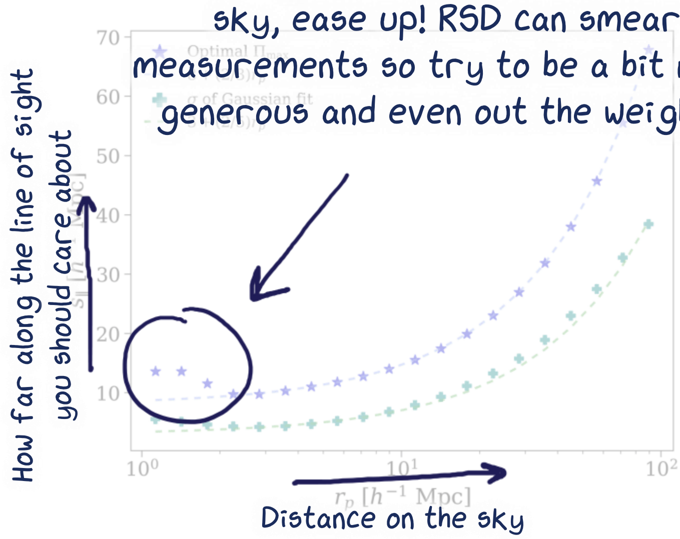


Figure 5. A guide for how to treat the LOS direction in projected IA measurements. Generally, a larger Π_{max} choice is recommended for tidal alignment at larger projected separations, following $(2/3)r_p$. Below 3-4 h^{-1}Mpc, a slightly larger Π_{max} is recommended as RSD smears the signal along the LOS. A Gaussian LOS weight can further improve the signal. Widths of these Gaussians are shown as teal plus signs and based on the expected tidal alignment in a given r_p bin (Figure 2). These follow $3 + (2/5)r_p$.

The region where RSD has a big effect

create a light cone by placing an observer 2590 h^{-1}Mpc away from the center of each box along one axis (corresponding to comoving distance at z=0.8), and determine a redshift without RSD and one with RSD.

TESTING THE NEW METHODS

3.2 Alignment Measurements

To produce the weights displayed in Figures 2 - 4, we find pairs of galaxies up to projected separations of $r_p < 100h^{-1}$ Mpc and transverse separations of $|r_\parallel| < 100h^{-1}$ Mpc in real comoving space. For each pair, we calculate its separation in the plane transverse to the LOS, r_p, and along the LOS in both real and redshift space, $r_\parallel$ and $s_\parallel$. The LOS is defined for each pair as the vector between the observer and pair center. We also obtain the projected ellipticity of each halo along its unique LOS, following (Lamman et al. 2023).

Now time to put our galaxies where our mouths are and see just how well these new methods compare to traditional ones...

For comparing alignment estimators, we trim the Abacus catalogs to a uniform survey geometry using RA and DEC within ±10 deg and redshifts $0.58 < z < 0.95$. Within these mock catalogs, we measure $\mathcal{E}_+(r_p)$ in each simulation using the $\tilde{\Pi}$ weights described in Section 2.2.1, the variable Π_{max} described in Section 2.2.2, and a selection of flat Π_{max} cuts. The error is estimated in each simulation by measuring the signal independently in 100 regions and taking the standard error. This results in 25 determinations of SNR for each estimator. We average and take the standard error of these 25 SNRs to evaluate the performance of each estimator in each r_p bin. The results are unchanged when using bootstrap error estimation from these regions.

We do this by using 25 different simulated universes (!) so we can get a very good idea of the average performance of each method.

Code for these measurements and the associated modeling is publicly available in the repository github.com/cmlamman/spec-IA.

3.3 Results

The measurements of $\mathcal{E}_+$ made with each Π_{max} choice, the corresponding plot for w_{g+}, and their average SNR are shown in Figure 6. The LOS weighting $\tilde{\Pi}$ has a 2.3 times improvement over a flat $\Pi_{\mathrm{max}} = 100h^{-1}$ Mpc, followed closely by using the optimal

I am shamelessly stealing a cartoon about simulated universes from another annotation because I am defending in two weeks and why did I commit to doing this for every paper in my phd ....

The top two plots show measurements
made with several different methods

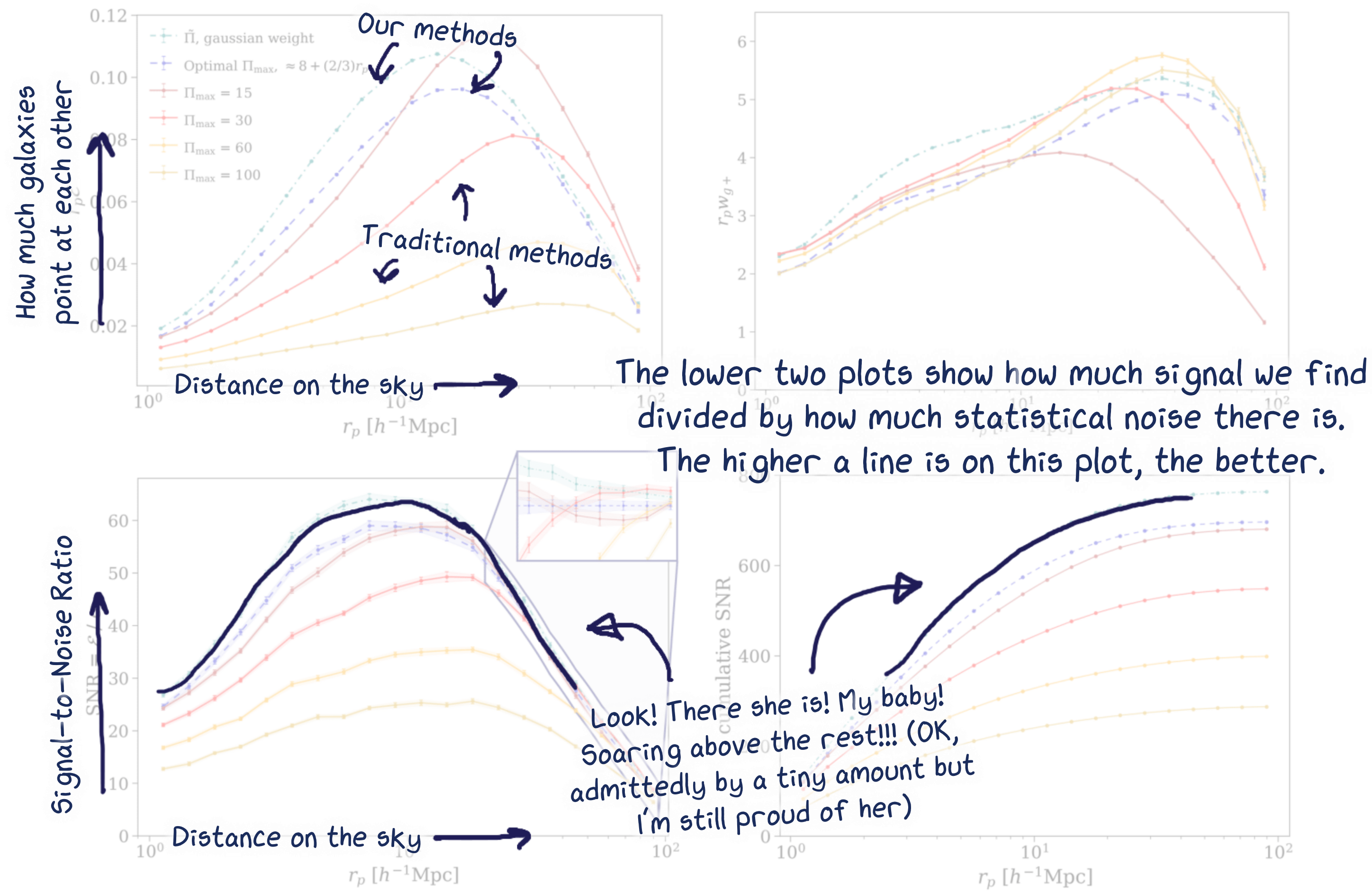


Figure 6. Comparison between different LOS treatments for $\mathcal{E}_+$ and w_{g+}, estimated as $(2\Pi_{max} + w_p(r_p))\mathcal{E}_+(r_p)$. The measurements made with various Π_{max} choices and averaged over 25 Abacus Summit simulations are shown in the top two panels. Since the errors are dominated by shape noise, the SNR is the same for the two estimators. The bottom left panel shows the SNR of each measurement, and the bottom right shows the cumulative SNR as r_p increases. The SNR was estimated independently in each simulation and then averaged for this plot. As expected, the SNR is highest for the gaussian weights across all scales, followed by the optimal Π_{max}. Among the single Π_{max} choices, generally smaller Π_{max} perform better, at least at scales bellow 30 h^{-1} Mpc. The insert shows a closer look at where the estimators' SNR cross, although many become statistically indistinguishable at the largest scales.

The fancy 3D measurement we mentioned before is about 2x better than the "normal" method. Our 2D-3D hybrid is about the same improvement, but you don't need to worry about annoying RSD effects!

Π_{max} in each r_p bin. As expected, smaller LOS distances tend to be more advantageous at small $r_\parallel$. However, it is worth noting that $\Pi_{max} = 15h^{-1}$ Mpc performs well too, significantly better than the common choice of $\Pi_{max} = 100h^{-1}$ Mpc. Since tidal alignment concentrates most of the signal along the [illegible] smaller Π_{max} [illegible] preferred compared to when measuring standard 2-point clustering. Above $r_p = 12h^{-1}$ Mpc, several estimators perform similarly as the signal has less variation in $r_\parallel$ at very large r_p. All SNR curves on this plot peak around $7 - 11h^{-1}$Mpc, a scale which depends on our choice of logarithmic binning.

MODELING THE NEW METHODS

4 MODELING VARIABLE Π_{MAX}

Now the mathy part -> explaining how people can build this measurement into their theoretical models. This is the vegetables that even an observer like I must take for a healthy science paper...

As seen in the top two panels of Figure 6, the modified Π_{max} estimators have a different scale dependence than the classic flat Π_{max} cuts. Th[illegible] models to account for these modifications. For the variable Π_{max} approach, modeling follows the same procedure as traditional projected IA statistics, but applied separately to individual r_p bins with their respective Π_{max} values. For the $\tilde{\Pi}$ approach with Gaussian weights, we present a modified modeling framework below.

The strength of intrinsic alignments is quantified by the alignment amplitude A_{IA}, which describes the response of galaxy shapes to the large-scale gravitational tidal field. The tidal tensor T_{ij} is defined as:

$$T_{ij} = \partial_i \partial_j \phi - \frac{1}{3}\delta^K_{ij}\nabla^2\phi \tag{10}$$

where ϕ is the gravitational potential and δ^K_{ij} is the Kronecker delta. In Fourier space, this becomes

$$T_{ij}(\vec{r}) = \int \frac{d^3k}{(2\pi)^3}\left(\frac{k_i k_j - \frac{1}{3}\delta^K_{ij}k^2}{k^2}\right)\tilde{\delta}_m(\vec{k})e^{i\vec{k}\cdot\vec{r}}, \tag{11}$$

following Lamman et al. (2023). Here, $\tilde{\delta}_m(\vec{k})$ is the Fourier transform of the matter density field. In the linear and nonlinear alignment [illegible] field according to:

$$\epsilon = \tau[T_{xx} - T_{yy} + 2iT_{xy}] \equiv \tau\mathbf{T}_{\mathcal{E}}, \tag{12}$$

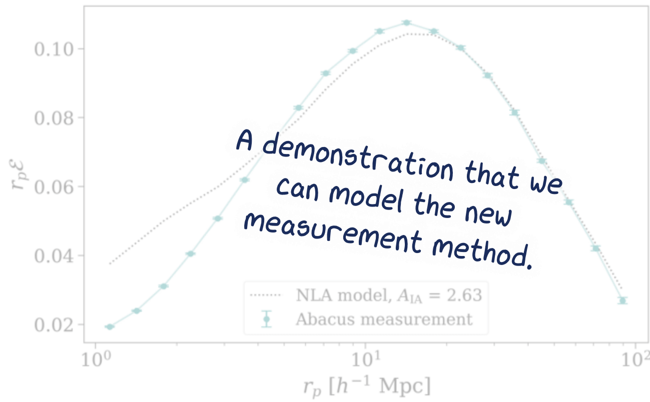


Figure 7. A comparison of the Gaussian-weighted IA signal shown in Figure 6 to the nonlinear alignment model, predicted using the nonlinear Abacus matter power spectrum. The alignment amplitude was estimated from the signal strength above 12 h^{-1}Mpc. As expected, this model breaks down at scales below 6 h^{-1}Mpc.

Fortunately, it's not that different from the usual way that people model these correlations.

where τ is a proportionality constant. The intrinsic shear is related to the alignment amplitude by:

$$\gamma^I_{ij} = -A_{\rm IA}(z) C_1 \frac{\rho_{\rm m,0}}{D(z)} T_{ij}. \tag{13}$$

Here, $\rho_{\rm m,0}$ is the present-day matter density, $D(z)$ is the linear growth factor (normalized so $\bar{D}(z) = (1+z)D(z)$ is unity at present day), and $C_1 = 5\times10^{-14} M_\odot^{-1} h^{-2}{\rm Mpc}^3$ is a historical normalization constant introduced by Brown et al. (2002). The relationship between our alignment parameter τ and the conventional $A_{\rm IA}$ is:

$$A_{\rm IA}(z) = -\frac{\tau}{C_1}\frac{D(z)}{\rho_{\rm m,0}} \tag{14}$$

Following Lamman et al. (2024b), we can express the alignment signal in terms of the relevant matter power spectrum:

$$\tilde{\mathcal{E}} = \frac{-\tau}{(2\Pi_{\rm max} + \bar{w}_p)} \int KdK \mathcal{J}_2(K,R) \mathcal{P}_\Pi(K) \tag{15}$$

...but that didn't mean it was easy to figure out!

$\bar{w}_p$ represents the projected 2-point cross-correlation function between the shape and tracer catalogs within an annulus defined by $R_{\rm min}$ and $R_{\rm max}$. The variable K corresponds to the Fourier-space wavenumber in the transverse plane (encompassing k_x and k_y), while k_z represents the wavenumber along the LOS direction, with the full 3D wavenumber being $k^2 = K^2 + k_z^2$. $\mathcal{J}_2$ is a modified Bessel function integrated over a given radial bin (see Eq. 22), and $\mathcal{P}_\Pi(K)$ encapsulates the relevant galaxy-matter power spectrum contribution for a given $\Pi_{\rm max}$. The expectation value of the cross-correlation between projected shapes and the underlying matter field, denoted as Q, is:

$$\tilde{\mathcal{E}}_{\rm model} = \frac{1}{2}\langle \epsilon^* Q + \epsilon Q^* \rangle \tag{16}$$

$$Q(R_{\rm bin}, \pm\tilde\Pi) = \frac{\int d^3 r W(\bar r)\delta_g e^{2i\theta_r}}{\int d^3 r W(\bar r)(1+\xi_{\epsilon g})} \tag{17}$$

Q describes the 3D matter field in a specific bin of transverse separation $R_{\rm bin}$ and weighted LOS separation $\tilde\Pi(s_\parallel)$. The function $W(\bar r)$ serves as a window function that selects this region in configuration space. The LOS weights for each r_p bin are modeled as a normalized Gaussian with standard deviation σ:

$$\tilde\Pi(s_\parallel) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(s_\parallel^2/2\sigma^2)} \tag{18}$$

This modified window function now needs to be carried through the expansion of Equation 16:

$$\mathcal{E}_{\rm model} = \frac{-\tau}{\int d^3 r W(\bar r)(1+\xi_{\epsilon g})} \int d^3 r W(\bar r) \int \frac{dk_z}{2\pi} \int KdK J_2(KR) \frac{K^2}{k^2} P_{gm}(k) e^{ik_z z}. \tag{19}$$

The volume element in our weighted space becomes:

$$\int d^3 r W(\bar r) = \pi(R^2_{\rm max} - R^2_{\rm min}) \int r_\parallel d_\parallel \tilde\Pi(r_\parallel) \tag{20}$$

Similarly, the normalization factor accounting for the clustering signal is:

$$\int d^3 r W(\bar r)(1+\xi_{\epsilon g}) = \pi(R^2_{\rm max} - R^2_{\rm min}) \int r_\parallel dr_\parallel \tilde\Pi(r_\parallel)(1+\xi(r_\parallel)). \tag{21}$$

The radial kernel $\mathcal{J}_2$ remains unchanged from the standard case with a fixed $\Pi_{\rm max}$:

$$\mathcal{J}_2(K) = \frac{2}{(R^2_{\rm max} - R^2_{\rm min})} \int_{R_{\rm min}}^{R_{\rm max}} RdRJ_2(KR), \tag{22}$$

[illegible] that $\Pi_{\rm max}$ cut produces a sinc function in Fourier space, our Gaussian weighting transforms to:

$$\int_{-\infty}^{\infty} dz \tilde\Pi(z) \exp^{ik_z z} = \exp^{-k_z^2\sigma_i^2/2}$$

.

This leads to a modified power spectrum contribution:

$$\mathcal{P}_{\tilde\Pi}(K) = \frac{1}{2\pi}\int dk_z \frac{K^2}{K^2+k_z^2} P_{gm}\left(\sqrt{K^2+k_z^2}\right) \sigma \exp^{-k_z^2\sigma^2/2}. \tag{23}$$

P_{gm} is the galaxy-matter power spectrum, which we estimate using [illegible] of $b = 2.7$. For our weighted estimator, the "effective $\Pi_{\rm max}$", $(2\Pi_{\rm max} + \bar{w}_p)$, becomes:

$$\tilde\Pi_{\rm eff} = \frac{\int s_\parallel dr_\parallel \tilde\Pi_{r_p}(s_\parallel)(1+\xi(r_p, s_\parallel))}{\int dr_\parallel \tilde\Pi_{r_p}(s_\parallel)} \tag{24}$$

Our final expression for the alignment signal is therefore:

$$\mathcal{E}(r_p) = \frac{-\tau}{\tilde\Pi_{\rm eff}} \int KdK \mathcal{J}_2(K, r_p) \mathcal{P}_{\tilde\Pi}(K). \tag{25}$$

Figure 7 compares the LOS-weighted $\mathcal{E}_+$ to the normalized model, which reproduces the signal above $r_p = 10h^{-1}$Mpc.

IA models often use the Limber approximation, assuming that [illegible] transverse to the line of sight, $k_z = 0$ modes, will contribute (Blazek et al. 2015). In this case, Equation 25 simply becomes the galaxy power spectrum, [illegible] $\Pi_{\rm max}$, and can be sufficient for $\Pi_{\rm max} = 30-100h^{-1}$Mpc alignment signals at small r_p. The $\Pi_{\rm max}$ recommendations we provide here are much smaller and [illegible] **approximation**. While modeling the full power spectrum is computationally more expensive, our LOS weighting comes with its own computational advantage. The Gaussian weights eliminate the oscillatory behavior of the sinc function that arises from a sharp $\Pi_{\rm max}$ cut, leading to more numerically stable and efficient integration.

Fun fact: There is a "shortcut" that people often use for modeling these measurements, but it doesn't really work for ours. HOWEVER our math results in a DIFFERENT shortcut which sort of cancels out the added complexity.

CONCLUSION

I am impressed that you made it this far in a rather specific, technical paper. As a reward, I will now attempt to summarize it with as much PIZZAZ and DRAMA as possible:

In a vast kingdom of people who explore galaxy alignments (a handful of researchers),

there is a raging battle (mild discussion) about if we should use 2D or 3D measurements.

2.5D?

We come up with a completely new method (based on lots of other people's ideas) that ANNIHILATES (works slightly better then) the other measurement methods.

Maximizing the precision of IA measurements directly improves constraints on the alignment amplitude. By more efficiently handling LOS information, these estimators will enhance the cosmological information extracted from large spectroscopic surveys, benefiting both weak lensing analyses and the exploration of direct cosmological applications of IA.

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

ACKNOWLEDGEMENTS

If anyone has a problem with this you can come say it to my face (please feel free to take a look at the data yourself and email me if you have questions).

DATA AVAILABILITY

AbacusSummit data products are publicly available at abacusnbody.org (Maksimova et al. 2021).

Thank you for reading! Check out ClaireLamman.com for other accessible annotations made by myself and others!
– Claire

I made this in power point using the XKCD font:
https://github.com/ipython/xkcd-font

Claire Lamman