Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **476,** 151–157 (2018)
Advance Access publication 2018 January 12

doi:10.1093/mnras/sty099

Statistical inconsistencies in the KiDS-450 data set

George Efstathiou★ and Pablo Lemos★

Kavli Institute for Cosmology Cambridge and Institute of Astronomy, Madingley Road, Cambridge CB3 OHA, UK

Accepted 2017 November 13. Received 2017 November 13; in original form 2017 July 10

ABSTRACT

The Kilo-Degree Survey (KiDS) has been used in several recent papers to infer constraints on the amplitude of the matter power spectrum and matter density at low redshift. Some of these analyses have claimed tension with the *Planck* Λ cold dark matter cosmology at the $\sim 2\sigma$–3σ level, perhaps indicative of new physics. However, *Planck* is consistent with other low-redshift probes of the matter power spectrum such as redshift-space distortions and the combined galaxy-mass and galaxy–galaxy power spectra. Here, we perform consistency tests of the KiDS data, finding internal tensions for various cuts of the data at $\sim 2.2\sigma$–3.5σ significance. Until these internal tensions are understood, we argue that it is premature to claim evidence for new physics from KiDS. We review the consistency between KiDS and other weak lensing measurements of S_8, highlighting the importance of intrinsic alignments for precision cosmology.

Key words: cosmic background radiation – cosmological parameters – large-scale structure of Universe – cosmology: observations.

1 INTRODUCTION

Precision observations of the cosmic microwave background radiation (CMB) by the *Planck* satellite (Planck Collaboration XVI 2014a; Planck Collaboration XIII 2016a, hereafter P16) and other experiments (Hinshaw et al. 2013; Sievers et al. 2013; Story et al. 2013) have shown that the Λ cold dark matter (ΛCDM) cosmology, with nearly scale invariant, adiabatic, Gaussian initial perturbations, provides an excellent description of our Universe. Measurements of weak lensing of the CMB (Planck Collaboration XV 2016b) show further that the ΛCDM model remains a good description of the Universe down to a redshift of $z \sim 2$, where the CMB lensing kernel peaks.

It is, nevertheless, important to test the model at lower redshifts, particularly at redshifts $z \lesssim 1$ when the Universe becomes dominated by dark energy. Deviations from the ΛCDM model at low redshift could potentially reveal evidence for dynamical dark energy or modifications to General Relativity (see Amendola et al. 2016, for a review).

Weak galaxy lensing is an important probe of the matter power spectrum at low redshifts (Blandford et al. 1991; Miralda-Escude 1991; Kaiser 1992). Several ambitious deep imaging projects have reported results recently. These include the Canada–France–Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al. 2012, 2013; Joudaki et al. 2017a), Deep Lens Survey (DLS; Jee et al. 2016), Dark Energy Survey (DES; Abbott et al. 2016; DES Collaboration et al. 2017; Troxel et al. 2017), and Kilo-Degree Survey (KiDs; Hildebrandt et al. 2017; Köhlinger et al. 2017). Weak lensing analysis of these surveys can be used to constrain the parameter combination[1] $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$, which can be compared to the *Planck* value from P16,[2] $S_8 = 0.825 \pm 0.016$ derived from the *Planck* temperature power spectrum, low multipole polarization, and *Planck* lensing (TT+lowTEB+lensing, in the notation of P16). However, the weak galaxy lensing results span a range of values. The reanalysis of CFHTlenS by Joudaki et al. (2017a) finds $S_8 = 0.732^{+0.029}_{-0.031}$; Jee et al. (2016) find $S_8 = 0.818^{+0.034}_{-0.026}$ from DLS; Abbott et al. (2016) find $S_8 = 0.81 \pm 0.06$ from the DES Science Verification data; (Hildebrandt et al. 2017, hereafter H17) find $S_8 = 0.745 \pm 0.039$ from a tomographic correlation function analysis of KiDs while (Köhlinger et al. 2017, hereafter K17) find $S_8 = 0.651 \pm 0.058$ from a tomographic power spectrum analysis of KiDs. The DES Year 1 weak lensing analysis[3] (Troxel et al. 2017) gives $S_8 = 0.789^{+0.024}_{-0.026}$. Some of these values are in tension with *Planck*. For example, H17 find a 2.3σ discrepancy between KiDs and *Planck*, while K17 find a 3.2σ discrepancy. However, the results from these different surveys do not agree particularly well with each other (even when using the same shear catalogue), showing differences in the value of S_8 at the $\sim 2\sigma$–2.5σ level.

[1] Where σ_8 is the present-day linear theory root-mean-square amplitude of the matter fluctuation spectrum averaged in spheres of radius $8\,h^{-1}$Mpc, Ω_m is the present-day matter density in units of the critical density ρ_c, and h is the Hubble constant in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

[2] Unless stated otherwise, we quote $\pm 1\sigma$ errors on parameters.

[3] DES Year 1 results (DES Collaboration et al. 2017; Troxel et al. 2017) appeared after the submission of this paper and so will not be discussed in detail.

★ E-mail: gpe@ast.cam.ac.uk (GE); pl411@cam.ac.uk (PL)

A statistically significant tension between the *Planck* ΛCDM cosmology and weak galaxy lensing could have important consequences for fundamental physics (e.g. Joudaki et al. 2017b). But how seriously should we take the weak lensing results? A minimal requirement is that a cosmic shear data set should be internally self-consistent. The main purpose of this paper is to show that this does not seem to be the case with KiDS.

Before we begin, we make a few remarks concerning cosmic shear analysis. Most analyses involve estimation of correlation functions ξ_+ and ξ_- as a function of relative angular separation θ, or of the cosmic shear E-mode power spectrum $P_\kappa(\ell)$ as a function of multipole ℓ. These are related by

$$\xi_\pm = \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell P_\kappa(\ell) J_{0,4}(\ell\theta). \tag{1}$$

For a cross-power spectrum between redshift bins i and j, the shear power spectrum is related to the non-linear matter power spectrum P_δ by

$$P_\kappa^{ij}(\ell) = \int_0^{\chi_H} \mathrm{d}\chi \frac{q_i(\chi)q_j(\chi)}{[f_K(\chi)]^2} P_\delta\left(\frac{(\ell+1/2)}{f_K(\chi)}, \chi\right), \tag{2}$$

where (following the notation of H17) χ is the comoving radial distance, $f_K(\chi)$ is the comoving angular diameter distance to distance χ, and $q_i(\chi)$ is the lensing efficiency for tomographic redshift bin i:

$$q_i(\chi) = \frac{3H_0^2\Omega_\mathrm{m}}{2c^2}\frac{f_K(\chi)}{a(\chi)}\int_\chi^{\chi_\mathrm{H}} \mathrm{d}\chi' n_i(\chi')\frac{f_K(\chi'-\chi)}{f_K(\chi')}, \tag{3}$$

where χ_H is the comoving Hubble distance and $n_i(\chi)$ is the effective (weighted) number density galaxies in redshift bin i normalized so that $\int n_i(\chi)\mathrm{d}\chi = 1$. Even if the image analysis is assumed to be free of systematic errors and biases, inferences on cosmology require an accurate model of the redshift distribution $n_i(\chi)$, which in turn requires accurate calibration of the photometric redshifts used to define the redshift bin i. A key test of the accuracy of the photometric redshift calibrations would be to demonstrate consistency between distinct cross-correlations i, j. However, this is not straightforward because of intrinsic ellipticity alignments between neighbouring galaxies (II term) and between gravitation shear and intrinsic shear (GI term). The power spectra[4] of these terms are usually modelled as (Hirata & Seljak 2004; Bridle & King 2007)

$$P_{II}^{ij}(\ell) = \int_0^{\chi_H} \mathrm{d}\chi F^2(z)\frac{n_i(\chi)n_j(\chi)}{[f_K(\chi)]^2} P_\delta\left(\frac{(\ell+1/2)}{f_K(\chi)}, \chi\right), \tag{4}$$

$$P_{GI}^{ij}(\ell) = \int_0^{\chi_H} \mathrm{d}\chi F(z)\frac{(q_i(\chi)n_j(\chi)+n_i(\chi)q_j(\chi))}{[f_K(\chi)]^2} \times P_\delta\left(\frac{(\ell+1/2)}{f_K(\chi)}, \chi\right). \tag{5}$$

In these equations,

$$F(z) = -A_{IA}C\rho_c\frac{\Omega_\mathrm{m}}{D(z)}, \tag{6}$$

where $D(z)$ is the linear growth rate of perturbations normalized to unity at the present day, and C is a normalizing constant, usually chosen to be $C = 5\times10^{-14}\,h^{-2}\,\mathrm{M}_\odot^{-1}\,\mathrm{Mpc}^3$. With this choice, the intrinsic alignment amplitude is expected to be of the order of unity (and positive if intrinsic ellipticities are aligned with the stretching axis of the tidal field). This model of intrinsic alignments is heuristic and simplified (see Blazek et al. 2017 for a more complex alignment model). Even in the context of this model, the intrinsic alignment amplitude may vary with redshift, luminosity, and galaxy type. For current weak lensing surveys, intrinsic alignments are not benign. The contributions of equations (4) and (5) are comparable to any claimed tensions between the *Planck* value of S_8 and those inferred from cosmic shear surveys (with positive A_IA tending to raise the value of S_8 and negative values lowering S_8). How can we test the intrinsic alignment model? The conventional solution is to introduce additional nuisance parameters to characterize uncertainties in the intrinsic alignment model (e.g. Kirk et al. 2012), relying on the redshift dependence of the measured signals to disentangle true cosmic shear from intrinsic alignments. This, of course, requires accurate knowledge of the redshift distributions and their errors.

Current cosmic shear data are still relatively sparse, with a small number of measurements in coarse redshift bins. The number of internal consistency checks of the data and the various components of the model (including nuisance parameters) are therefore limited.[5] In Section 2, we perform consistency tests of the KiDS data from H17. In Section 3, we compare the KiDS results with *Planck* and measurements of redshift-space distortions (RSDs) and rich cluster abundances, which provide independent measures of the amplitude of the matter fluctuations at similar redshifts to those of the KiDS galaxies. Section 4 compares the results from various weak lensing analyses. Our main conclusions are presented in Section 5.

2 TESTS OF THE KIDS DATA

We use the KiDS cross-correlation measurements of ξ_+ and ξ_- in four tomographic redshift bins as reported by H17 together with the associated COSMOMC likelihood module and covariance matrix.[6] For reference, the four redshift bins span the following ranges in photometric redshift z_B: $0.1 < z_B \leq 0.3$ (bin 1), $0.3 < z_B \leq 0.5$ (bin 2), $0.5 < z_B \leq 0.7$ (bin 3), and $0.7 < z_B \leq 0.9$ (bin 4). We used the same angular ranges, photometric redshift calibrations and errors, nuisance parameters, and priors as in 'fiducial' analysis in H17 (first entry in their table 4) and verified that we recovered the identical best-fitting χ^2 (162.8) and constraint on S_8 ($S_8 = 0.745\pm0.039$). We then removed all cross-correlations involving one of the photometric redshift bins. The results are summarized in Table 1 and in Fig. 1 .

The first point to note is that the intrinsic alignment amplitude is reasonably stable to the removal of photometric redshift bins. All of the posteriors shown in Fig. 1 are consistent with the intrinsic alignment solution from the full data set ($A_\mathrm{IA} = 1.10^{+0.68}_{-0.54}$). However, it is also clear that redshift bin 4 carries a high weight in fixing A_IA. With redshift bin 4 removed, the posterior distribution develops a long tail to negative values that is cut-off by the lower end of the A_IA prior (uniform between $-6 < A_\mathrm{IA} < 6$). As a consequence of this long tail, the best-fitting value of S_8 with bin 4 removed is driven to lower values and its error increases substantially compared to the full sample (lower panel of Fig. 1 and Table 1). Redshift bin 4 is therefore critical in pinning down the intrinsic alignment solution and reducing the error on S_8.

If redshift bin 3 is removed, S_8 rises and the constraints in the S_8–Ω_m plane become compatible with *Planck* (Fig. 1). This is not

[4] Neglecting B modes.

[5] The situation is very different to the CMB, where there is a large amount of information to separate a high-amplitude frequency-independent cosmological signal with a distinctive power spectrum from low-amplitude foregrounds with smooth power spectra.

[6] Downloaded from http://kids.strw.leidenuniv.nl.

Table 1. Conditional χ^2 tests removing photometric redshift bins.

$\boldsymbol{y}^D$	S_8	$A_{\rm IA}$	$\chi^2_{\rm cond}$	$N_{\sigma_{\rm cond}}$
minus z-bin 1	0.745 ± 0.040	1.14 ± 0.85	61.0 (52)	0.89
minus z-bin 2	0.754 ± 0.042	1.24 ± 0.80	66.3 (52)	1.40
minus z-bin 3	0.771 ± 0.039	1.25 ± 0.57	78.2 (52)	2.60
minus z-bin 4	0.684 ± 0.071	−0.1 ± 1.7	87.9 (52)	3.52
minus ξ_-	0.778 ± 0.040	1.10 ± 0.73	89.7 (60)	2.71
minus ξ_+	0.705 ± 0.048	0.92 ± 0.97	84.1 (70)	1.20

Notes. The first column defines the portion of the data vector ($\boldsymbol{y}^D$) used to fit the model. The second and third columns give the marginalized mean values of S_8, $A_{\rm IA}$ins, and their 1σ errors. The fourth column gives the conditional $\chi^2_{\rm cond}$, as defined in equation (15), for the rest of data vector, $\boldsymbol{x}^D$. The numbers in parentheses list the length, $N_{\rm x}$, of the vector $\boldsymbol{x}^D$. The fifth column gives the number of standard deviations by which $\chi^2_{\rm cond}$ differs from $N_{\rm x}$, $N_{\sigma_{\rm cond}} = (\chi^2_{\rm cond} - N_{\rm x})/\sqrt{2N_{\rm x}}$.

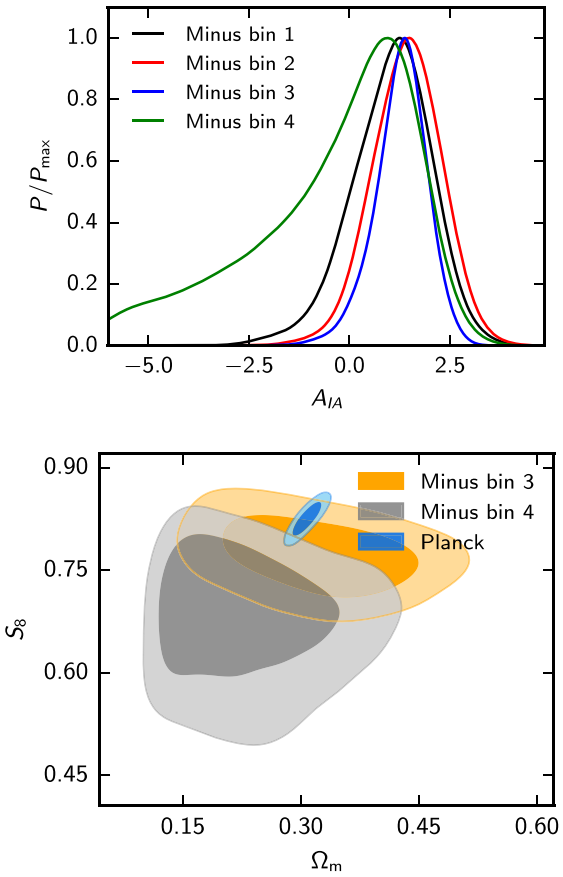

Figure 1. The upper panel shows the posteriors for the intrinsic alignment parameter $A_{\rm IA}$ (equation 6) as we remove all cross-correlations involving a particular redshift bin. The lower panel shows the 68 and 95 per cent constraints on S_8 for the data minus redshift bin 3 (orange) and minus redshift bin 4 (grey). The blue contours show the *Planck* constraints from the TT+lowTEB+lensing data combination as given in P16.

unexpected, because one can see from fig. 5 of H17 that the best-fitting fiducial model tends to sit high for all cross-spectra involving tomographic redshift bin 3 (particularly for ξ_-). With redshift bin 3 removed, there is substantial overlap in the posteriors in the S_8–$\Omega_{\rm m}$ plane with those from the full sample and with the other subsets of the data summarized in Table 1. However, these various estimates of S_8 are highly correlated since they share common data. Are the parameter shifts seen in these subsets statistically reasonable? We turn to this question next.

We can perform a more elaborate statistical consistency test by dividing the data vector into two components:

$$\boldsymbol{z}^D = (\boldsymbol{x}^D, \boldsymbol{y}^D). \tag{7}$$

We can then fit $\boldsymbol{y}^D$ to a model (including nuisance parameters), $\hat{\boldsymbol{y}}$. The model parameters also make a theory prediction for the data partition $\boldsymbol{x}^D$, which we denote $\hat{\boldsymbol{x}}$. We can then write the theory vector for $\boldsymbol{z}^D$ as

$$\hat{\boldsymbol{z}} = (\lambda\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}), \tag{8}$$

introducing a new parameter λ. Evidently, if the data partitions and model are consistent, the new parameter λ should be consistent with unity. The tests described in this section are all based on the ΛCDM model, *but with a free amplitude*. Since cosmic shear measurements have very limited ability to fix shape parameters, and the data cuts that we apply cover similar redshift ranges, it seems reasonable to interpret differences in λ as indicative of systematic errors in the data. To recap, we run Markov chain Monte Carlo (MCMC) chains to determine the model parameters from a data partition $\boldsymbol{y}^D$ and determine a single amplitude parameter λ by fitting to the rest of the data $\boldsymbol{x}_D$. The posterior distributions of λ for the data cuts of Table 1 are shown in Fig. 2.

The upper plot in Fig. 2 compares the amplitudes λ_- (fitting the model parameters to ξ_+) and λ_+ (fitting the model parameters to ξ_-). This agrees with the visual impression given by fig. 5 of H17, namely that ξ_- wants a low amplitude while ξ_+ prefers a high amplitude. Integrating these distributions

$$\int_0^1 P(\lambda_-)\mathrm{d}\lambda_- = 2.9 \times 10^{-3}, \tag{9}$$

$$\int_1^\infty P(\lambda_+)\mathrm{d}\lambda_+ = 4.2 \times 10^{-2}. \tag{10}$$

A value of $\lambda = 1$ therefore lies in the tails of both posterior distributions. These results show that ξ_- sits about 2.8σ low compared to the best-fitting ΛCDM cosmology determined from ξ_+.

The lower plot in Fig. 2 tests consistency between photometric redshift bins including both ξ_+ and ξ_- in the fits. The parameters λ_i (with i running from 1 to 4) are computed for data partitions in which $\boldsymbol{y}^D$ excludes all cross-correlations involving photometric redshift bin i. In this test, photometric redshift bin 3 is an outlier with

$$\int_0^1 P(\lambda_3)\mathrm{d}\lambda_3 = 1.3 \times 10^{-2}, \tag{11}$$

suggesting that the data involving photometric redshift bin 3 are inconsistent with the rest of the data at about the 2.2σ level. Again, this accords with the visual impression from fig. 5 of H17, which shows that cross-correlations in both ξ_+ and ξ_- involving photometric redshift bin 3 tend to lie below their best-fitting model.

Instead of using an amplitude parameter λ, we can make a prediction for the vector $\boldsymbol{x}^D$ conditional on the fit to $\boldsymbol{y}^D$

$$\boldsymbol{x}^{\rm cond} = \hat{\boldsymbol{x}} + \mathbf{C}_{xy}\mathbf{C}_{yy}^{-1}(\boldsymbol{y}^D - \hat{\boldsymbol{y}}). \tag{12}$$

If the best-fitting model is known exactly, the covariance of $\boldsymbol{x}^{\rm cond}$ is

$$\mathbf{C}_{xx}^{\rm cond} = \mathbf{C}_{xx} - \mathbf{C}_{xy}\mathbf{C}_{yy}^{-1}\mathbf{C}_{yx}. \tag{13}$$

However, in our application the best-fitting model is determined by fitting the data vector $\boldsymbol{y}^D$ and so the uncertainty in the best-fitting

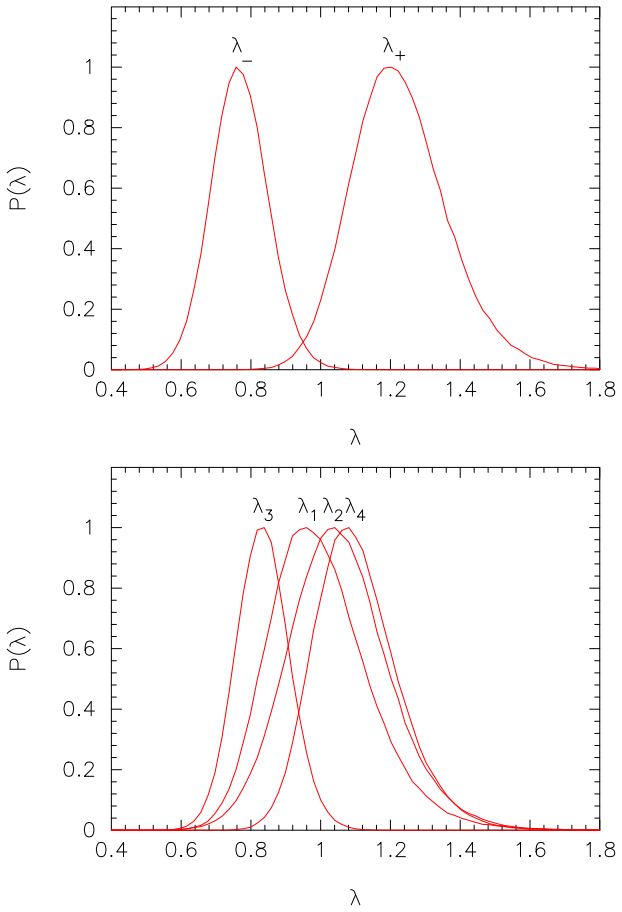

Figure 2. Posterior distributions of the parameter λ defined in equation (8). The upper figure shows the distributions if the model parameters are fitted to ξ_+ (denoted λ_-) and to ξ_- (denoted λ_+). The lower figure shows the posterior distributions of λ for partitions of the data in which all cross-correlations involving a particular tomographic redshift bin are removed from the fit to the theoretical model (e.g. λ_3 corresponds to a theoretical model fitted to all cross-correlations that do not involve tomographic redshift bin 3).

model contributes an additional variance to $\mathbf{C}_{xx}^{\rm cond}$:

$$\mathbf{C}'^{\rm cond}_{xx} = \mathbf{C}^{\rm cond}_{xx} + \Delta\mathbf{C}^{\rm cond}_{xx}, \tag{14}$$

which we determine empirically by sampling over the MCMC chains. In our application, $\Delta\mathbf{C}^{\rm cond}_{xx}$ is a small correction to $\mathbf{C}^{\rm cond}_{xx}$.

As a test of the consistency of the data, we compute a conditional χ^2:

$$\chi^2_{\rm cond} = (\boldsymbol{x}^D - \boldsymbol{x}^{\rm cond})^T (\mathbf{C}'^{\rm cond}_{xx})^{-1} (\boldsymbol{x}^D - \boldsymbol{x}^{\rm cond}). \tag{15}$$

The results of these tests are summarized in Table 1 and are consistent with the λ-tests shown in Fig. 2. Eliminating ξ_- leads to a substantial increase in S_8 that is incompatible with ξ_- at about 2.7σ. The redshift bin 3 component of the data vector is inconsistent with the rest of the data vector at about 2.6σ. However, the $\chi^2_{\rm cond}$ reveals a new inconsistency: the redshift bin 4 component of the data vector is inconsistent with the rest of the data vector at about 3.5σ.

The origin of the high values of $\chi^2_{\rm cond}$ for these various partitions of the data vector is clear from Fig. 3. The figure shows the data vector (red points) for all cross-correlations involving redshift bin 3 (upper two panels) and those involving redshift bin 4 (lower two panels) compared to the expectations $\boldsymbol{x}^{\rm cond}$ conditional on the rest of the data (equation 12). The grey bands show $\pm1\sigma$ and $\pm2\sigma$ ranges

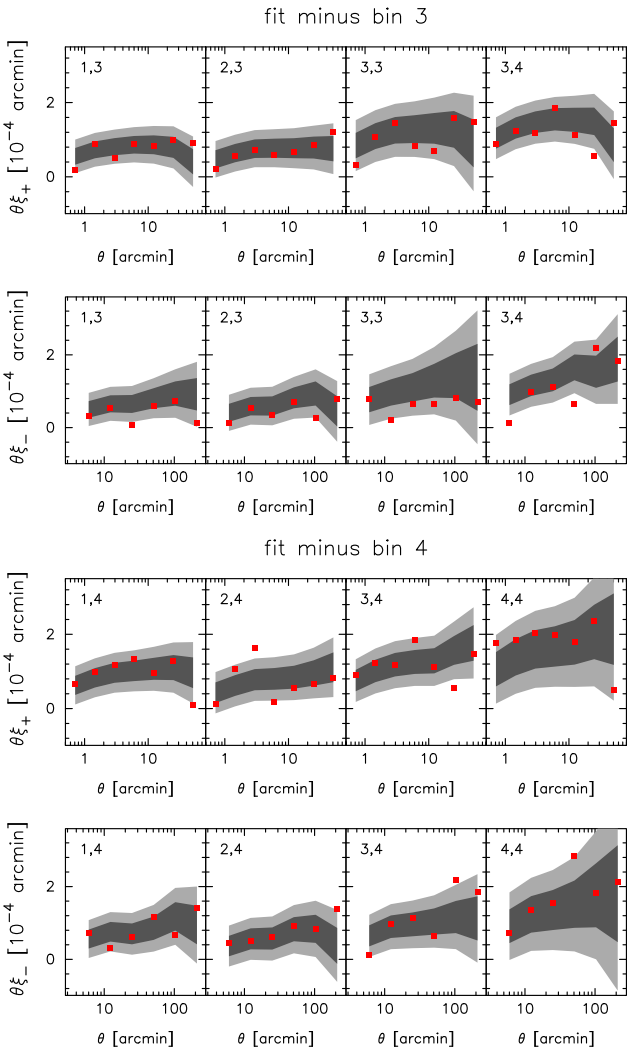

Figure 3. The upper two panels show cross-correlations ξ_+ and ξ_- involving tomographic redshift bin 3 (red points). The numbers in each plot to identify the cross-correlation (e.g. 1, 3 denotes redshift bin 1 crossed with redshift bin 3). The grey bands show the allowed $\pm1\sigma$ (dark grey) and $\pm2\sigma$ (light grey) ranges allowed by the fits to the rest of the data. The lower two panels show the equivalent plots, but for cross-correlations involving tomographic redshift bin 4.

around $\boldsymbol{x}^{\rm cond}$ computed from the diagonal components of equation (14). The top two panels of Fig. 3 show that cross-correlations involving redshift bin 3 want a lower amplitude than the rest of the data. This problem is particularly acute for ξ_- for the (3, 3) and (3, 4) redshift bin cross-correlations. These two cross-correlations carry quite high weight in fits to the full data vector (driving S_8 down), yet they are inconsistent at nearly $\sim2.6\sigma$ with the rest of the data. A possible explanation for this discrepancy is an inaccuracy in the calibration of the photometric redshifts for bin 3. In fact, van Uitert et al. (2017) present evidence for a 2.3σ negative shift of $\Delta z \approx -0.06$ for this redshift bin. They find no evidence for significant shifts in the other redshift bins.

As summarized in Table 1, removing redshift bin 4 lowers the value of S_8 but increases the errors on S_8 substantially because the intrinsic alignment amplitude is less well constrained. From Fig. 3, this low amplitude solution appears to match reasonably well with the general shape of the rest of the data vector, but now we see a high value of $\chi^2_{\rm cond}$ arising from *outliers*. In the lower two panels

of this figure, 8 out of 52 data points sit outside the conditional $\pm 2\sigma$ range.[7] Several of these outliers are at large angular scales and are not obvious in plots using errors computed from the diagonals of the full covariance matrix (e.g. fig. 5 of H17). However, the KiDS covariance matrix tells us that the data vector should be correlated across different tomographic redshift bins. What Fig. 3 shows is that the KiDS correlation functions display significantly higher variance than expected from the KiDS covariance matrix, particularly at large angular scales and for correlations involving redshift bin 4. This excess variance is a serious problem because it means that the KiDS errors on cosmological parameters are systematically underestimated, especially if data at small angular scales are excluded.

Our analysis shows strong evidence for a statistical inconsistency between the KiDS estimates of ξ_+ and ξ_-. H17 and van Uitert et al. (2017) find evidence for non-zero B modes in the KiDS data at small angular scales ($\theta < 4.2$ arcmin), indicative of systematics. If systematic errors contribute equally to the tangential and cross distortions (and this has not been demonstrated for KiDS), then the B modes will affect ξ_+, but not ξ_-. Eliminating ξ_+ entirely from the fits lowers S_8 to 0.705 ± 0.048 (see Table 1) with $\chi^2 = 82.2$ for 50 degrees of freedom (a 3.2σ excess). In other words, if one argues that the difference between ξ_+ and ξ_- is indicative of systematic errors in ξ_+, then the tension between KiDS and *Planck* is exacerbated.

3 COMPARISON WITH OTHER TECHNIQUES FOR MEASURING THE AMPLITUDE OF THE FLUCTUATION SPECTRUM

The results of the previous section show that there are some worrying internal inconsistencies in the KiDS data set as analysed in H17. These inconsistencies suggest that we should be cautious in interpreting the KiDS constraints on cosmology. However, the tests in themselves do not tell us the causes of the inconsistencies, or their impact on the estimates of S_8. Is the amplitude of the matter fluctuations at redshifts $z \lesssim 1$ really lower than expected in the *Planck* ΛCDM cosmology?

Another way of studying the amplitude of the matter power spectrum is via RSDs (Kaiser 1987). RSDs provide a measurement of the parameter combination $f\sigma_8$, where f is the logarithmic derivative of the linear growth rate with respect to the scale factor

$$f = \frac{\mathrm{d}\ln D}{\mathrm{d}\ln a}, \tag{16}$$

and $a = (1+z)^{-1}$. In the ΛCDM model, $f \approx \Omega_\mathrm{m}(z)^{0.55}$ and so RSDs measure the parameter combination $\sigma_8\Omega_\mathrm{m}^{0.55}$, i.e. similar to the parameter combination S_8 up to a known constant. Measurements of RSD from the DR12 analysis of the Baryon Oscillation Spectroscopy Survey (BOSS) have been reported by Alam et al. (2017). These measurements are for three redshift slices with effective redshifts $z_\mathrm{eff} = 0.38$, 0.51, and 0.61, substantially overlapping with the redshift range of the KiDS survey. Huterer et al. (2017) have recently used the Supercal Type Ia supernova compilation (Scolnic et al. 2015) together with independent distance measurements of galaxies (Springob et al. 2014) to measure $f\sigma_8$ at $z_\mathrm{eff} = 0.02$. The *Planck* ΛCDM cosmology is in excellent agreement with these measurements of $f\sigma_8$ over the entire redshift range $z = 0.02$–0.61. The consistency between *Planck* and the RSD measurements is illustrated in Fig. 4, where we have combined the BOSS and Supercal RSD measurements to produce constraints in the σ_8–Ω_m plane[8]. The RSD constraints are in mild tension with the KiDS correlation function analysis of H17, and in even greater tension with the tomographic power-spectrum analysis of KiDS described by K17 *using the same shear catalogue.*

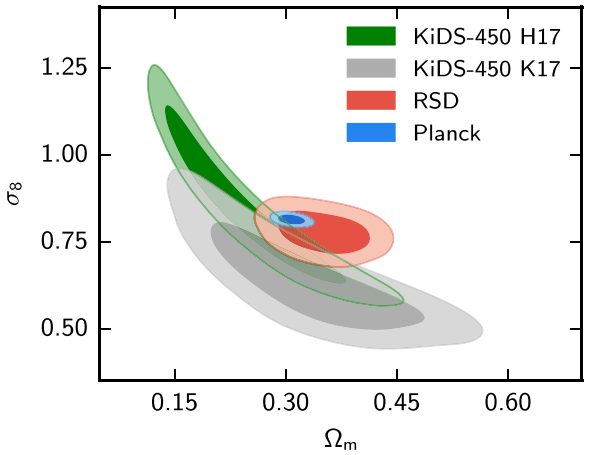

Figure 4. Constraints in the σ_8–Ω_m plane assuming the spatially flat ΛCDM cosmology. The 68 and 95 per cent contours from *Planck* are shown in blue. The constraints from the H17 fiducial KiDS analysis are shown in green. The grey contours show the constraints from the power-spectrum analysis of KiDS reported by K17. The red contours show the constraints from RSDs as discussed in the text.

The abundance of rich clusters of galaxies (selected at various wavelengths) has been used in a number of studies to constrain the amplitude of the fluctuations spectrum at low redshift (e.g. Vikhlinin et al. 2009; Rozo et al. 2010; Hasselfield et al. 2013; Planck Collaboration XX 2014b; Mantz et al. 2015; de Haan et al. 2016; Planck Collaboration XXIV 2016c). As summarized in several of these papers, calibration of cluster masses is a major source of uncertainty in this type of analysis. Two recent studies (Mantz et al. 2015; de Haan et al. 2016) use weak gravitational lensing mass estimates from the 'Weighing the Giants' programme (Applegate et al. 2014; Kelly et al. 2014; von der Linden et al. 2014) to calibrate cluster scaling relations. Mantz et al. (2015) use an X-ray selected sample of clusters from the *ROSAT* All-Sky Survey covering the redshift range $0 < z < 0.5$, finding $\sigma_8(\Omega_\mathrm{m}/0.3)^{0.17} = 0.81 \pm 0.03$. de Haan et al. (2016) use a sample of clusters identified with the South Pole Telescope with median redshift $z_\mathrm{med} = 0.53$ to infer $\sigma_8(\Omega_\mathrm{m}/0.27)^{0.3} = 0.797 \pm 0.031$. Both of these estimates are consistent with the *Planck* P16 ΛCDM cosmology: $\sigma_8(\Omega_\mathrm{m}/0.3)^{0.17} = 0.818 \pm 0.009$, $\sigma_8(\Omega_\mathrm{m}/0.27)^{0.3} = 0.848 \pm 0.012$. Thus, there is no convincing evidence for any discrepancy between rich cluster counts and the expectations from the *Planck*–ΛCDM cosmology. The de Haan et al. (2016) study is particularly interesting because it covers a similar redshift range to those of the BOSS RSD and KiDS measurements, yet is consistent with *Planck* and RSD.

[7] Assuming Gaussian statistics, the p-value for this is about 2.4×10^{-3}.

[8] This is done using the final_consensus_dV_FAP_fsig data files and covariance matrix downloaded from https://sdss3.org/science/boss_publications.php. We then scanned the likelihood, using uniform priors in H_0 and $\Omega_\mathrm{m}h^2$ to rescale the BOSS distance D_V and Alcock–Paczynski (Alcock & Paczynski 1979) parameter F_AP to the fiducial sound horizon used in the BOSS analysis, fixing $\Omega_\mathrm{b}h^2$ to the P16 ΛCDM value.

4 COMPARISON OF WEAK LENSING ESTIMATES OF S_8: THE IMPORTANCE OF INTRINSIC ALIGNMENTS

Fig. 4 shows a discrepancy between the H17 and K17 analyses, which are based on the same shear catalogue. There is little doubt that the H17 and K17 analyses are incompatible, since not one of the 14 469 samples in the K17 MCMC likelihood chain[9] has parameters close to those of the best fit found by H17. In fact, van Uitert et al. (2017, hereafter vU17) have computed cross power spectra from ξ_+ and ξ_- for the KiDS data using the identical redshift bins to those used in K17. Their autospectrum for the highest redshift bin differs substantially from the quadratic estimate of K17. The origin of this difference is not understood.[10] Another pointer that the K17 results are affected by systematic errors comes from the intrinsic alignment solution. K17 find $A_{\rm IA} = -1.72^{+1.49}_{-1.25}$ which has the opposite (and from the theoretical perspective, counterintuitive) sign to that found by H17. This difference drives down the amplitude of S_8 in the K17 analysis. Both the direct comparison of spectra reported by van Uitert et al. (2017) and the shift to a negative intrinsic alignment amplitude suggest that the K17 analysis is suspect.

The key point that we want to emphasize here is that the intrinsic alignment parameter $A_{\rm IA}$ is not a benign 'nuisance' parameter (for reviews, see e.g. Joachimi et al. 2015; Troxel & Ishak 2015). The modelling of intrinsic alignments is degenerate with the cosmological parameters of interest, σ_8, $\Omega_{\rm m}$, and S_8, and so the model and associated parameters matter. Systematic errors in the data can be absorbed by the intrinsic alignment model and this will have an impact on cosmology. For example, van Uitert et al. (2017) have noted that the parameter $A_{\rm IA}$ can absorb systematic errors in the calibrations of photometric redshift distributions (this can also be inferred from Fig. 1 which shows the sensitivity of the intrinsic alignment solution for the KiDS data to the highest photometric redshift bin). Implausible (e.g. strongly negative) values of $A_{\rm IA}$ suggest systematic errors and should therefore be followed up.

As an example, one of the lowest weak lensing determinations of S_8 comes from the reanalysis of the revised CFHTLenS data (Joudaki et al. 2017b). However, these authors find a strongly negative value of $A_{\rm IA} = -3.6 \pm 1.6$, a value which seems unlikely for any reasonable mix of galaxy types. The recent DES analysis of Troxel et al. (2017) uses a redshift-dependent amplitude: $A_{\rm IA}[(1+z)/(1.62)]^{\eta}$, finding $A_{\rm IA} = 1.3^{+0.5}_{-0.6}$, $\eta = 3.7^{+1.0}_{-2.3}$.[11] Troxel et al. (2017) also test a more elaborate 'mixed' alignment model based on the work of Blazek et al. (2017). This model leads to a downward shift of S_8 by about 1σ, demonstrating that uncertainties in the modelling of intrinsic alignments make a non-negligible contribution to the errors in cosmological parameters.

Returning to the KiDS survey, one way of achieving better control of intrinsic alignments and photometric redshift calibration errors is to add additional types of data. vU17 have analysed the shear power spectra from KiDS, P^E (constructed by integrating over ξ_+ and ξ_-). In addition, they use the Galaxies Mass Assembly (GAMA) redshift survey (Driver et al. 2011; Liske et al. 2015) to compute the galaxy-mass power spectra, $P^{\rm gm}$ by cross-correlating the KiDS shear measurements with GAMA galaxies, and the galaxy–galaxy power spectra $P^{\rm gg}$. From $P^{\rm gm} + P^{\rm gg}$, they find $S_8 = 0.853 \pm 0.042$. Combining with P^E, they find $A_{\rm IA} = 1.30 \pm 0.40$ and $S_8 = 0.801 \pm 0.032$ (consistent with the *Planck* and RSD results shown in Fig. 4).

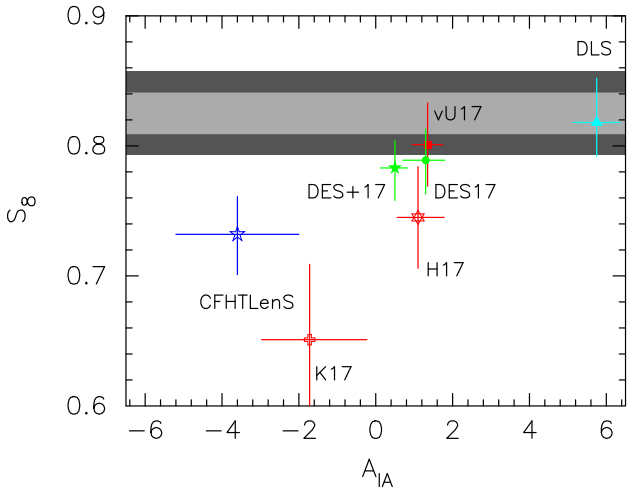

Figure 5. S_8 plotted against the intrinsic alignment amplitude for various surveys together with 1σ errors on S_8 and $A_{\rm IA}$. The grey bands show the 1σ and 2σ constraints from *Planck*. The data points are as follows: CFHTLens (Joudaki et al. 2017a); DLS (Jee et al. 2016)[12]; K17 shows the power spectrum analysis of KiDS (K17); H17 shows the correlation function analysis of KiDS (H17); vU17 shows the constraints from combining $P^{\rm gg}$, $P^{\rm gm}$, and P^E measurements from KiDS and GAMA data (vU17); DES17 shows the cosmic shear constraints from DES year 1 data (Troxel et al. 2017, note that the DES analyses use a redshift-dependent model of intrinsic alignments, as described in the text); DES+17 shows the combination of DES year 1 cosmic shear results with galaxy–galaxy and galaxy-shear measurements (DES Collaboration et al. 2017).

Fig. 5 gives a summary of the results discussed in this section. The two analyses that are most discrepant with the S_8 value from *Planck* (CFHTLenS and K17) both have strongly negative intrinsic alignment solutions. The H17 results are in tension with *Planck* but become consistent with *Planck* with the addition of galaxy–galaxy and galaxy-mass data (vU17). The DES year 1 analyses plotted in Fig. 5 are both consistent with *Planck*. The intrinsic alignment solutions of vU17 and DES Collaboration et al. (2017, i.e. $A_{\rm IA} \sim 1$) seem physically plausible given the mix of galaxy types expected in these surveys.

5 CONCLUSIONS

The main purpose of this paper has been to highlight and quantify internal inconsistencies in the KiDS cosmic shear analysis. Our main conclusion is that more effort is needed to resolve inconsistencies in the KiDS data. This includes understanding the origin of the B modes, systematic differences between ξ_+ and ξ_-, the parameter

[9] KiDS450_QE_EB_4bins_3zbins_basez_ia_bary_nu.txt, downloaded from http://kids.strw.leidenuniv.nl.

[10] Note that the quadratic estimator used by K17 is sensitive to noise estimation, particularly if there are B-mode systematics (which are known to be present in the KiDS data). Inaccurate noise estimation would primarily affect the autospectra, where the noise levels are high compared to the cosmological signal (see fig. 4 of H17).

[11] These constraints become $A_{\rm IA} = 0.5^{+0.32}_{-0.38}$, $\eta = 0^{+2.7}_{-2.8}$ with the addition of galaxy–galaxy and galaxy-shear data, DES Collaboration et al. (2017). These authors argue that an amplitude of $A_{\rm IA} \sim 0.5$ is consistent with their selection criteria if only red galaxies contribute to the intrinsic alignments.

[12] Note that the Jee et al. (2016) 'baseline' analysis of DLS uses a luminosity-dependent model of intrinsic alignments and imposes a flat prior of $5.14 < A_{\rm IA} < 6.36$, motivated by the results of Joachimi et al. (2011). However, they find that their results on S_8 are insensitive to $A_{\rm IA}$ (see their fig. 12), presumably because of the huge depth of DLS.

shifts seen by excluding photometric redshift bin 3, the large excess χ^2 and scatter at large angular scales. Until this is done, it seems premature to draw inferences on new physics from KiDS.

A comparison of *Planck* with other measures of the amplitude of the mass fluctuations, principally RSDs from BOSS, reveals no evidence for any inconsistencies with the Planck-based ΛCDM cosmology. We have also reviewed cosmic shear constraints on S_8, emphasizing the degeneracy between intrinsic alignments and cosmology. As summarized in Fig. 5, the two analyses which yield the lowest values of S_8 both have strongly negative values of $A_{\rm IA}$. The DES 1-yr analyses are consistent with the *Planck* ΛCDM value for S_8 (DES Collaboration et al. 2017; Troxel et al. 2017) and give physically plausive values for $A_{\rm IA}$. The H17 value of S_8 from KiDS sits about 2.3σ low compared to *Planck*, but is pulled upwards with the addition of galaxy–galaxy, galaxy-mass data (vU17). Overall, we conclude there is no strong evidence for any inconsistency between the *Planck* ΛCDM cosmology and measures of the amplitude of the fluctuation spectrum at low redshift.

ACKNOWLEDGEMENTS

We thank Hiranya Peiris, Benjamin Joachimi, Fergus Simpson, and the referee for helpful comments on the pre-print version of this paper. We thank Frankie Nobis-Efstathiou for help with the early stages of this project. We also thank Anthony Challinor, Steven Gratton, and members of the KiDS team for comments on aspects of this analysis. We also thank members of the *Planck* Parameters team. PL acknowledges support from an Isaac Newton Studentship at the University of Cambridge and from the Science and Technologies Facilities Council.

REFERENCES

Abbott T. et al., 2016, Phys. Rev. D, 94, 022001
Alam S. et al., 2017, MNRAS, 470, 2617
Alcock C., Paczynski B., 1979, Nature, 281, 358
Amendola L. et al., 2016, preprint (arXiv:1606.00180)
Applegate D. E. et al., 2014, MNRAS, 439, 48
Blandford R. D., Saust A. B., Brainerd T. G., Villumsen J. V., 1991, MNRAS, 251, 600
Blazek J., MacCrann N., Troxel M. A., Fang X., 2017, preprint (arXiv:1708.09247)
Bridle S., King L., 2007, New J. Phys., 9, 444
de Haan T. et al., 2016, ApJ, 832, 95
DES Collaboration et al., 2017, preprint (arXiv:1708.01530)
Driver S. P. et al., 2011, MNRAS, 413, 971
Hasselfield M. et al., 2013, J. Cosmology Astropart. Phys., 7, 008
Heymans C. et al., 2012, MNRAS, 427, 146
Heymans C. et al., 2013, MNRAS, 432, 2433
Hildebrandt H. et al., 2017, MNRAS, 465, 1454(H17)
Hinshaw G. et al., 2013, ApJS, 208, 19
Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526
Huterer D., Shafer D. L., Scolnic D. M., Schmidt F., 2017, J. Cosmology Astropart. Phys., 5, 015
Jee M. J., Tyson J. A., Hilbert S., Schneider M. D., Schmidt S., Wittman D., 2016, ApJ, 824, 77
Joachimi B., Mandelbaum R., Abdalla F. B., Bridle S. L., 2011, A&A, 527, A26
Joachimi B. et al., 2015, Space Sci. Rev., 193, 1
Joudaki S. et al., 2017a, MNRAS, 465, 2033
Joudaki S. et al., 2017b, MNRAS, 471, 1259
Kaiser N., 1987, MNRAS, 227, 1
Kaiser N., 1992, ApJ, 388, 272
Kelly P. L. et al., 2014, MNRAS, 439, 28
Kirk D., Rassat A., Host O., Bridle S., 2012, MNRAS, 424, 1647
Köhlinger F. et al., 2017, MNRAS, 471, 4412(K17)
Liske J. et al., 2015, MNRAS, 452, 2087
Mantz A. B. et al., 2015, MNRAS, 446, 2205
Miralda-Escude J., 1991, ApJ, 380, 1
Planck Collaboration XVI, 2014a, A&A, 571, A16
Planck Collaboration XX, 2014b, A&A, 571, A20
Planck Collaboration XIII, 2016a, A&A, 594, A13(P16)
Planck Collaboration XV, 2016b, A&A, 594, A15
Planck Collaboration XXIV, 2016c, A&A, 594, A24
Rozo E. et al., 2010, ApJ, 708, 645
Scolnic D. et al., 2015, ApJ, 815, 117
Sievers J. L. et al., 2013, J. Cosmology Astropart. Phys., 10, 060
Springob C. M. et al., 2014, MNRAS, 445, 2677
Story K. T. et al., 2013, ApJ, 779, 86
Troxel M. A., Ishak M., 2015, Phys. Rep., 558, 1
Troxel M. A. et al., 2017, preprint (arXiv:1708.01538)
van Uitert E. et al., 2017, preprint (arXiv:1706.05004)
Vikhlinin A. et al., 2009, ApJ, 692, 1060
von der Linden A. et al., 2014, MNRAS, 439, 2(vU17)

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