Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **471,** 1619–1633 (2017)
Advance Access publication 2017 July 4

doi:10.1093/mnras/stx1675

KiDS-450: tomographic cross-correlation of galaxy shear with *Planck* lensing

Joachim Harnois-Déraps,[1⋆] Tilman Tröster,[2⋆] Nora Elisa Chisari,[3] Catherine Heymans,[1] Ludovic van Waerbeke,[2] Marika Asgari,[1] Maciej Bilicki,[4,5] Ami Choi,[1,6] Thomas Erben,[7] Hendrik Hildebrandt,[7] Henk Hoekstra,[4] Shahab Joudaki,[8,9] Konrad Kuijken,[4] Julian Merten,[3] Lance Miller,[3] Naomi Robertson,[3] Peter Schneider[7] and Massimo Viola[4]

[1]*Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Blackford Hill, EH9 3HJ Scotland, UK*
[2]*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC, V6T 1Z1, Canada*
[3]*Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[4]*Leiden Observatory, Leiden University, PO Box 9513, 2300RA Leiden, the Netherlands*
[5]*National Centre for Nuclear Research, Astrophysics Division, PO Box 447, PL-90-950 Łódź, Poland*
[6]*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[7]*Argelander-Institut für Astronomie, Auf dem Hügel 71, D-53121 Bonn, Germany*
[8]*Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[9]*ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)*

Accepted 2017 June 30. Received 2017 June 30; in original form 2017 March 15

ABSTRACT

We present the tomographic cross-correlation between galaxy lensing measured in the Kilo Degree Survey (KiDS-450) with overlapping lensing measurements of the cosmic microwave background (CMB), as detected by *Planck* 2015. We compare our joint probe measurement to the theoretical expectation for a flat Λ cold dark matter cosmology, assuming the best-fitting cosmological parameters from the KiDS-450 cosmic shear and *Planck* CMB analyses. We find that our results are consistent within 1σ with the KiDS-450 cosmology, with an amplitude re-scaling parameter $A_{\rm KiDS} = 0.86 \pm 0.19$. Adopting a *Planck* cosmology, we find our results are consistent within 2σ, with $A_{Planck} = 0.68 \pm 0.15$. We show that the agreement is improved in both cases when the contamination to the signal by intrinsic galaxy alignments is accounted for, increasing A by $\sim$0.1. This is the first tomographic analysis of the galaxy lensing – CMB lensing cross-correlation signal, and is based on five photometric redshift bins. We use this measurement as an independent validation of the multiplicative shear calibration and of the calibrated source redshift distribution at high redshifts. We find that constraints on these two quantities are strongly correlated when obtained from this technique, which should therefore not be considered as a stand-alone competitive calibration tool.

Key words: gravitational lensing: weak – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Recent observations of distinct cosmological probes are closing in on the few parameters that enter the standard model of cosmology (see, e.g. Planck Collaboration XIII 2016a, and references therein). Although there is clear evidence that the Universe is well described by the Λ cold dark matter (ΛCDM) model, some tensions are found between probes. For instance, the best-fitting cosmology inferred from the observation of the cosmic microwave background (CMB) in Planck Collaboration XIII (2016a) is in tension with some cosmic shear analyses (MacCrann et al. 2015; Hildebrandt et al. 2017; Joudaki et al. 2016, 2017), while both direct and strong lensing measurements of today's Hubble parameter H_0 are more than 3σ away from the values inferred from the CMB (Bernal, Verde & Riess 2016; Bonvin et al. 2017). At face value, these discrepancies either point towards new physics (for a recent example, see Joudaki et al. 2016) or un-modelled systematics in any of those probes. In this context, cross-correlation of different cosmic probes stands out as a unique tool, as many residual systematics that could contaminate one data set are unlikely to correlate also with the other (e.g. 'additive biases'). This type of measurement can therefore be

⋆ E-mail: jharno@roe.ac.uk (JHD); troester@phas.ubc.ca (TT)

exempt from un-modelled biases that might otherwise source the tension. Another point of interest is that the systematic effects that do not fully cancel, for example 'multiplicative biases' or the uncertainty on the photometric redshifts, will often impact differently the cosmological parameters compared to the stand-alone probe, allowing for degeneracy breaking or improved calibration.

In this paper, we present the first tomographic cross-correlation measurement between CMB lensing and galaxy lensing, based on the lensing map described in Planck Collaboration XV (2016b) and the lensing data from the Kilo Degree Survey[1] presented in Kuijken et al. (2015, KiDS hereafter) and in the KiDS-450 cosmic shear analysis (Hildebrandt et al. 2017). The main advantage in this sort of measurement resides in it being free of uncertainty on galaxy bias, which otherwise dominates the error budget in CMB lensing – galaxy position cross-correlations (Omori & Holder 2015; Baxter et al. 2016; Giannantonio et al. 2016). Over the last 2 yr, the first lensing–lensing cross-correlations were used to measure σ_8 and Ω_m (Hand et al. 2015; Liu & Hill 2015), by combining the CMB lensing data from the Atacama Cosmology Telescope (Das et al. 2014) with the lensing data from the Canada-France-Hawaii Telescope Stripe 82 Survey (Moraes et al. 2014) and from the *Planck* lensing data and the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS hereafter; Erben et al. 2013). Since then, additional effects were found to contribute to the measurement, introducing extra-complications in the interpretation of the signal. For instance, Hall & Taylor (2014) and Troxel & Ishak (2014) showed that the measurement is likely to be contaminated by the intrinsic alignment of galaxies with the tidal field in which they live. At the same time, Liu, Ortiz-Vazquez & Hill (2016) argued that this measurement could point instead to residual systematics in the multiplicative shear bias and proposed that the measurement itself could be used to set constraints on the shear bias (see also Das, Errard & Spergel 2013). Their results showed that large residuals are favoured, despite the calibration accuracy claimed by the analysis of image simulations tailored for the same survey (Miller et al. 2013). A recent analysis from Harnois-Déraps et al. (2016, hereafter HD16) suggested instead that the impact of catastrophic redshift outliers could be causing this apparent discrepancy, since these dominate the uncertainty in the modelling. They also showed that choices concerning the treatment of the masks can lead to biases in the measured signal and that the current estimators should therefore be thoroughly calibrated on full light-cone mocks.

Although these pioneering works were based on Fourier space cross-correlation techniques, more recent analyses presented results from configuration-space measurements, which are cleaner due to their insensitivity to masking. Kirk et al. (2016, hereafter K16) combined the CMB lensing maps from *Planck* and from the South Pole Telescope (SPT; van Engelen et al. 2012) with the Science Verification Data from the Dark Energy Survey.[2] Their measurement employed the POLSPICE numerical tool (Szapudi, Prunet & Colombi 2001; Chon et al. 2004), which starts off with a pseudo-C_ℓ measurement that is converted into configuration space to deal with masks, then turned back into a Fourier space estimator. Soon after, HD16 showed consistency between pseudo-C_ℓ analyses and configuration space analyses of two-point correlation functions, combining the *Planck* lensing maps with both CFHTLenS and the Red-sequence Cluster Lensing Survey (RCSLenS hereafter; Hildebrandt et al. 2016). A similar configuration space estimator was recently used with *Planck* lensing and SDSS shear data (Singh, Mandelbaum & Brownstein 2017), although the signal was subject to higher noise levels.

This paper directly builds on the K16 and HD16 analyses, utilizing tools and methods described therein, but on a new suite of lensing data. The additional novelty here is that we perform the first tomographic CMB lensing – galaxy lensing cross-correlation analysis, where we split the galaxy sample into five redshift bins and examine the redshift evolution. This is made possible by the high quality of the KiDS photometric redshift data, by the extend of the spectroscopic matched sample, and consequently by the precision achieved on the calibrated source redshift distribution (see Hildebrandt et al. 2017, for more details). It provides a new test of cosmology within the ΛCDM model, including the redshift evolution of the growth of structure, and also offers an opportunity to examine the tension between the KiDS and *Planck* cosmologies (reported in Hildebrandt et al. 2017). With the upcoming lensing surveys such as LSST[3] and Euclid,[4] it is forecasted that this type of cross-correlation analysis will be increasingly used to validate the data calibration (Schaan et al. 2017) and extract cosmological information in a manner that complements the cosmic shear and clustering data.

The basic theoretical background upon which we base our work is laid out in Section 2. We then describe the data sets and our measurement strategies in Sections 3 and 4, respectively. Our cosmological results are presented in Section 5. We also describe therein a calibration analysis along the lines of Liu et al. (2016), this time focusing on high-redshift galaxies for which the photometric redshifts and shear calibration are not well measured. Informed on cosmology from lower redshift measurement, this self-calibration technique has the potential to constraint jointly the shear bias and the photo-z distribution, where other methods fail. We conclude in Section 6.

The fiducial cosmology that we adopt in our analysis corresponds to the flat WMAP9+SN+BAO cosmology[5] (Hinshaw et al. 2013), in which the matter density, the dark energy density, the baryonic density, the amplitude of matter fluctuations, the Hubble parameter and the tilt of the matter power spectrum are described by $(\Omega_m, \Omega_\Lambda, \Omega_b, \sigma_8, h, n_s) = (0.2905, 0.7095, 0.0473, 0.831, 0.6898, 0.969)$. Aside from determining the overall amplitude of the theoretical signal from the $[\sigma_8$–$\Omega_m]$ pair, this choice has little impact on our analysis, as we later demonstrate. Future surveys will have the statistical power to constrain the complete cosmological set, but this is currently out of reach for a survey the size of KiDS-450. We note that our fiducial cosmology is a convenient choice that is consistent within 2σ with the *Planck*, KiDS-450, CFHTLenS and WMAP9+ACT+SPT analyses in the $[\sigma_8$–$\Omega_m]$ plane. As such, it minimizes the impact of residual tension across data sets.

2 THEORETICAL BACKGROUND

Photons from the surface of last scattering are gravitationally lensed by large-scale structures in the Universe before reaching the observer. Similarly, photons emitted by observed galaxies are lensed by the low-redshift end of the same large-scale structures. The signal expected from a cross-correlation measurement between the two lenses can be related to the fluctuations in their common foreground

[1] KiDS: http://kids.strw.leidenuniv.nl
[2] DES: www.darkenergysurvey.org
[3] www.lsst.org
[4] sci.esa.int/euclid
[5] Our fiducial cosmology consists of a flat ΛCDM universe in which the dark energy equation of state is set to $w = -1$.

matter field, more precisely by the matter power spectrum $P(k, z)$. The lensing signal is obtained from an extended first-order Limber integration over the past light cone up to the horizon distance $\chi_{\rm H}$, weighted by geometrical factors $W^i(\chi)$, assuming a flat cosmology (Limber 1954; Loverde & Afshordi 2008; Kilbinger et al. 2017):

$$C_\ell^{\kappa^{\rm CMB}\kappa^{\rm gal}} = \int_0^{\chi_{\rm H}} {\rm d}\chi\, W^{\rm CMB}(\chi) W^{\rm gal}(\chi) P\left(\frac{\ell+1/2}{\chi}; z\right). \quad (1)$$

In the above expression, χ is the comoving distance from the observer, ℓ is the angular multipole and z is the redshift. The lensing kernels are given by

$$W^i(\chi) = \frac{3\Omega_{\rm m} H_0^2}{2c^2} \chi\, g^i(\chi)(1+z), \quad (2)$$

with

$$g^{\rm gal}(\chi) = \int_\chi^{\chi_H} {\rm d}\chi' \tilde{n}(\chi') \frac{\chi'-\chi}{\chi'} \quad \text{and}$$

$$g^{\rm CMB}(\chi) = \left[1 - \frac{\chi}{\chi_*}\right] {\rm H}(\chi_* - \chi). \quad (3)$$

The constant c is the speed of light in vacuum, χ_* is the comoving distance to the surface of last scattering. The term $\tilde{n}(\chi)$ is related to the redshift distribution of the observed galaxy sources, $n(z)$, by $\tilde{n}(\chi) = n(z){\rm d}z/{\rm d}\chi$, which depends on the depth of the survey. The Heaviside function ${\rm H}(x)$ guarantees that no contribution comes from behind the surface of last scattering as the integration in equation (1) approaches the horizon.

The angular cross-spectrum described by equation (1) is related to correlation functions in configuration space, in particular between the CMB lensing map and the tangential shear (Miralda-Escude 1991):

$$\xi^{\kappa^{\rm CMB}\gamma_{\rm t}}(\vartheta) = \frac{1}{2\pi} \int_0^\infty {\rm d}\ell\, \ell\, C_\ell^{\kappa^{\rm CMB}\kappa^{\rm gal}} J_2(\ell\vartheta), \quad (4)$$

where J_2 is the Bessel function of the first kind of order 2, and the quantity ϑ represents the angular separation on the sky. Details about measurements of $C_\ell^{\kappa^{\rm CMB}\kappa^{\rm gal}}$ and the tangential shear $\gamma_{\rm t}$ – relevant to equations (1) and (4), respectively – are provided in Section 4.

Our predictions are obtained from the NICAEA[6] cosmological tool (Kilbinger et al. 2009), assuming a non-linear power spectrum described by the Takahashi et al. (2012) revision of the HALOFIT model (Smith et al. 2003).

3 THE DATA SETS

3.1 KiDS-450 lensing data

The KiDS-450 lensing data that we use for our measurements are based on the third data release of dedicated KiDS observations from the VLT Survey Telescope at Paranal, in Chile, and are described in Kuijken et al. (2015) in Hildebrandt et al. (2017) and de Jong (2017, in preparation). These references describe the reduction and analysis pipelines leading to the shear catalogues and present a rigorous and extensive set of systematic verifications. Referring to these papers for more details, we summarize here the properties of the data that directly affect our measurement.

Although the full area of the KiDS survey will consist of two large patches on the celestial equator and around the South Galactic Pole,

[6] www.cosmostat.org/software/nicaea/

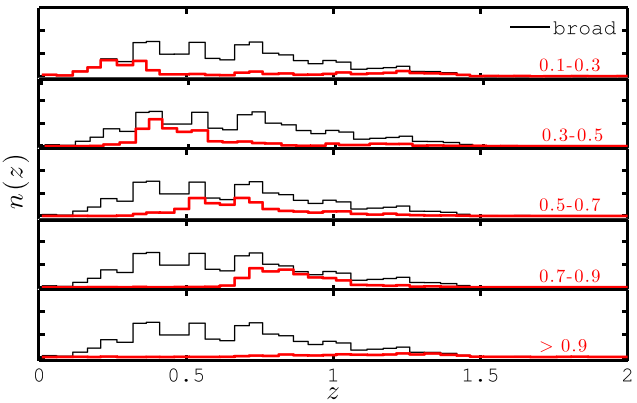

Figure 1. Redshift distribution of the selected KiDS-450 sources in the tomographic bins (unnormalized), calibrated using the DIR method described in Hildebrandt et al. (2017). The $n(z)$ of the broad $Z_B \in [0.1, 0.9]$ bin is shown in black in all panels for reference, while the $n(z)$ for the five tomographic bins are shown in red. The mean redshift and effective number of galaxy in each tomographic bin are summarized in Table 1.

the observing strategy was optimized to prioritize the coverage of the GAMA fields (Liske et al. 2015). The footprint of the KiDS-450 data is consequently organized in five fields, G9, G12, G15, G23 and GS, covering a total of 449.7 deg^2. While the multiband imaging data are processed by Astro-WISE (de Jong et al. 2015), the lensing r-band data are processed by the THELI reduction method described in Erben et al. (2013). Shape measurements are determined using the self-calibrated *lens*fit algorithm (based on Miller et al. 2013) detailed in Fenech Conti et al. (2017).

As described in Hildebrandt et al. (2017), each galaxy is assigned a photometric redshift probability distribution provided by the software BPZ (Benítez 2000). The position of the maximum value of this distribution, labelled Z_B, serves only to divide the data into redshift bins. Inspired by the KiDS-450 cosmic shear measurement, we split the galaxy sample into five redshift bins: $Z_B \in [0.1, 0.3]$, [0.3, 0.5], [0.5, 0.7], [0.7, 0.9] and >0.9. We also define a broad redshift bin by selecting all galaxies falling in the range $Z_B \in [0.1, 0.9]$. The KiDS-450 cosmic shear measurement did not include the $Z_B > 0.9$ bin because the photo-z and the shear calibration were poorly constrained therein. For this reason, we do not use this bin in our cosmological analysis either. Instead, we estimate these calibration quantities directly from our measurement in Section 5.7.

For each tomographic bin, the estimate of the redshift distribution of our galaxy samples, $n(z)$, is not obtained from the stacked BPZ-PDF, but from a magnitude–weighted scheme (in 4-dimensional *ugri* magnitude space) of a spectroscopically matched sub-sample. In Hildebrandt et al. (2017), this 'weighted direct calibration' or 'DIR' method was demonstrated to be the most precise covering our redshift range, among four independent $n(z)$ estimation techniques. Fig. 1 shows these weighted $n(z)$ distributions, which enter the theoretical predictions through equation (1), along with the effective number density per bin. In order to preserve the full description of the data in the high-redshift tail, from where most of the signal originates, we do not fit the distributions with analytical functions, as was done in previous work (Hand et al. 2015; K16; HD16). Fitting functions tend to capture well the region where the $n(z)$ is maximal; however, they attribute almost no weight to the (noisy) high-redshift tail. This is of lesser importance in the galaxy lensing autocorrelation measurements, but becomes highly relevant for the CMB lensing cross-correlation. Instead, we use the actual

Table 1. Summary of the data properties in the different tomographic bins. The effective number of galaxy assumes the estimation method of Heymans et al. (2012).

Z_B cut	$\bar{z}$	$n_{\rm eff}$ (gal/arcmin^2)	σ_ϵ
[0.1, 0.9]	0.72	7.54	0.28
[0.1, 0.3]	0.75	2.23	0.29
[0.3, 0.5]	0.59	2.03	0.28
[0.5, 0.7]	0.72	1.81	0.27
[0.7, 0.9]	0.87	1.49	0.28
>0.9	1.27	0.90	0.33

histograms in the calculation (as in Liu & Hill 2015) recalling that their apparent spikes are smoothed by the lensing kernels in equation (3). What is apparent from Fig. 1, and of importance for this analysis, is that all tomographic bins have a long tail that significantly overlaps with the CMB lensing kernel, especially the first tomographic bin. These tails are caused by inherent properties to the *ugri*-band photo-z of the KiDS-450 data, and given the wavelength range and signal-to-noise ratio (SNR), some high-z tails are expected (Hildebrandt et al. 2017). This feature is well captured by the mean redshift distributions, which are listed in Table 1.

Based on the quality of the ellipticity measurement, each galaxy is assigned a *lens*fit weight w, plus a multiplicative shear calibration factor – often referred to as the m-correction or the shear bias – that is obtained from image simulations (Fenech Conti et al. 2017). This calibration is accurate to better than 1 per cent for objects with $Z_B < 0.9$, but the precision quickly degrades at higher redshifts. As recommended, we do not correct for shear bias in each galaxy, but instead compute the average correction for each tomographic bin (see equation 7). In the fifth tomographic bin, we expect to find residual biases in the m-correction, but apply it nevertheless, describing in Section 5.7 how this correction can be self-calibrated. To be absolutely clear, we reiterate that we do not include this fifth bin in our main cosmological analysis. The effective number density and the shape noise in each tomographic bin are also listed in Table 1.

Following Hildebrandt et al. (2017), we apply a c-correction by subtracting the weighted mean ellipticity in each field and each tomographic bin, but this has no impact on our analysis since this c term does not correlate with the CMB lensing data.

3.2 *Planck* $\kappa_{\rm CMB}$ maps

The CMB lensing data that enter our measurements are the $\kappa_{\rm CMB}$ map obtained from the 2015 public data release,[7] thoroughly detailed in Planck Collaboration XV (2016b). The map making procedure is based on the quadratic estimator described in Okamoto & Hu (2003), which is applicable for a suite of multifrequency temperature and polarization maps. Frequencies are combined such as to remove foreground contamination, while other sources of secondary signal (mainly emissions from the galactic plane, from point sources and hot clusters) are masked in the CMB maps, prior to the reconstruction. If some of these are not fully removed from the lensing maps, they will create systematic effects in the $\kappa_{\rm CMB}$ map that show up differently in the cross-correlation measurement compared to the autospectrum analysis. For example, there could be leakage in the CMB map coming from, e.g., residual thermal Sunyaev-Zel'dovich signal that is most likely located near massive clusters. These same clusters are highly efficient at lensing background galaxies; hence, our cross-correlation measurement would be sensitive to this effect. Indeed, the $\langle {\rm tSZ} \times \gamma_{\rm t} \rangle$, as recently measured in Hojjati et al. (2016), has a very large SNR and could possibly be detected in a targeted analysis. Although it is difficult to assess the exact level of the tSZ signal in our $\kappa_{\rm CMB}$ map, the cleaning made possible from the multifrequency observations from *Planck* is thorough, reducing the residual contaminants to a very small fraction. No quantitative evidence of such leakage has been reported as of yet, and we therefore ignore this in our analysis.

Regions from the full sky lensing map that overlap with the five KiDS footprints are extracted, including a 4 deg extension to optimize the SNR of the measurement (see HD16). The *Planck* release of lensing data also provides the analysis mask, which we apply to the $\kappa_{\rm CMB}$ map prior to carrying out our measurement.[8]

4 THE MEASUREMENTS

This section presents the cross-correlation measurements, which are performed with two independent estimators: $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$ (equation 4) and the POLSPICE measurement of $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ (equation 1). These techniques were used and rigorously validated in previous work, and we refer the interested reader to HD16, K16 and references therein for more details. The reasons for conducting our analysis with these two estimators are twofold. First, they do not probe the same physical scales, which makes them complementary when carried out on surveys covering patchy regions. Secondly, being completely independent codes, residual systematics arising from inaccuracies in the analysis could be identified through their different effect on these two statistics.

4.1 The $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$ estimation

The first estimator presented in this paper, $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$, was recently introduced in HD16, and used later in Singh et al. (2017). It is a full configuration-space measurement that involves minimal manipulation of the data. The calculation simply loops over each pixel of the $\kappa_{\rm CMB}$ maps and defines concentric annuli with different radii ϑ, therein measuring the average tangential component of the shear, $\gamma_{\rm t}$, from the KiDS galaxy shapes. For this reason, it is arguably the cleanest avenue to perform such a cross-correlation measurement, even though there appears to be a limit to its accuracy at large angles in some cases due to the finite support of the observation window (Mandelbaum et al. 2013). That being said, it nevertheless bypasses a number of potential issues that are encountered with other estimators (see HD16 for a discussion). The $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$ estimator is given by

$$\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}(\vartheta) = \frac{\sum_{ij} \kappa^i_{\rm CMB} e^{ij}_{\rm t} w^j \Delta_{ij}(\vartheta)}{\sum_{ij} w^j \Delta_{ij}(\vartheta)} \frac{1}{1 + K(\vartheta)}, \quad (5)$$

where the sum first runs over the $\kappa_{\rm CMB}$ pixels 'i', then over all galaxies 'j' found in an annulus of radius ϑ and width Δ, centred on the pixel i. In this local coordinate system, $e^{ij}_{\rm t}$ is the tangential component of the *lens*fit ellipticity from the jth galaxy relative to

[7] *Planck* lensing package: pla.esac.esa.int/pla/#cosmology

[8] This procedure does not entirely capture the masking analysis since the mask was applied on the temperature field, not on the lensing map. The reconstruction process inevitably leaks some of the masked regions into unmasked area, and vice versa. Applying this mask will therefore only remove the most problematic regions.

pixel i. The exact binning scheme is described by $\Delta_{ij}(\vartheta)$, the binning operator:

$$\Delta_{ij}(\vartheta) = \begin{cases} 1, \text{ if } \left|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j\right| < \vartheta \pm \frac{\Delta}{2} \\ 0, \text{ otherwise} \end{cases} \quad (6)$$

where $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$ are the observed positions of the pixel i and galaxy j. Following HD16, the bin width Δ is set to 30 arcmin, equally spanning the angular range [1, 181] arcmin with six data points. Larger angular scales capture very little signal with the current level of statistical noise. We verified that our analysis results are independent of our choice of binning scheme. In equation (5), w^j is the *lens*fit weight of the galaxy j and $K(\vartheta)$ corrects for the shape multiplicative bias m^j that must be applied to the lensing data (Fenech Conti et al. 2017):

$$\frac{1}{1+K(\vartheta)} = \frac{\sum_{ij} w^j \Delta_{ij}(\vartheta)}{\sum_{ij} w^j (1+m^j) \Delta_{ij}(\vartheta)} \,. \quad (7)$$

The theoretical predictions for $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$ are provided by equation (4). We apply the same binning as with the data, averaging the continuous theory lines inside each angular bin. We show in the upper panel of Fig. 2 the measurements in all tomographic bins, compared to theoretical predictions given by our fiducial WMAP9+BAO+SN cosmology. The estimation of our error bars is described in Section 4.3.

We also project the galaxy shape components on to $e_\times$, which is rotated by 45 deg compared to $e_{\rm t}$. This effectively constitutes a nulling operation that can inform us of systematic leakage in analogy to the EB test performed in the context of cosmic shear. For this reason, we loosely refer to EE and EB tests in this paper, when we are in fact comparing $\kappa_{\rm CMB} \times e_{\rm t}$ and $\kappa_{\rm CMB} \times e_\times$, respectively. We note that the past literature referred to such a EB measurement as the 'B-mode test', which can be misleading for the non-expert. Indeed, the proper B-mode test refers to the BB measurement in weak lensing analyses, a non-lensing signal that can be caused by astrophysics and systematics. The EB signal test asserts something more fundamental: since B changes sign under parity, and E does not, a non-zero EB means a violation of the parity of the shear/ellipticity field (Schneider 2003). That is not expected from lensing alone, so could only come from a systematic effect that does not vanish under averaging. For example, signal coming from parts of the survey next to masked cluster regions could be affected by un-masked residuals that correlate with the other data set.

Our EB measurement is shown with the red symbols in Fig. 2. We find by visual inspection that in most tomographic bins, these seem closely centred on zero, but not in all cases. To quantify the significance of this EB measurement, we estimate the confidence at which these red points deviate from zero. We detail in Section 5.2 how we carry out that test and show that they are consistent with noise.

We have carried out an additional null test presented in HD16, which consists in rotating randomly the shapes of the galaxies before the measurement ($\kappa_{\rm CMB} \times$ random). This test is sensitive to the noise levels in the galaxy lensing data and hence affected by the shape noise σ_ϵ listed in Table 1. We find that the resulting signal is fully consistent with zero in all tomographic bins.

4.2 The $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ estimation

The second estimator uses the same data as our $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$ analysis, namely the $\kappa_{\rm CMB}$ map and the KiDS shear catalogues, but requires additional operations on the data, including harmonic space transforms. This is accomplished with the POLSPICE numerical code (Szapudi et al. 2001; Chon et al. 2004) running in polarization mode, where the {T, Q, U} triplets are replaced by $\{\kappa_{\rm CMB}, 0, 0\}$ and $\{0, -e_1, e_2\}$. The code first computes the pseudo-C_ℓ of the maps and of the masks, then transforms the results into configuration space quantities, that are finally combined and transformed back into Fourier space. The output of POLSPICE is therefore an estimate of the cross-spectrum $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$. While POLSPICE is frequently used for CMB analyses, it was applied for the first time in the context of CMB lensing × galaxy lensing by K16 and serves as a good comparison to the configuration estimator described in Section 4.1. One main advantage of this estimator is that in principle different ℓ-bands are largely uncorrelated, which makes the covariance matrix almost diagonal and hence easier to estimate.

The POLSPICE measurement[9] is presented in the lower panel of Fig. 2, plotted against the theoretical predictions given by equation (1). The EB data points are directly obtained from the temperature/B-mode output provided by the polarization version of the code and are further discussed in Section 5.2.

Note that our choice of the $\gamma_{\rm t}$ and POLSPICE estimators was motivated by our desire to avoid producing $\kappa_{\rm gal}$ maps in order to reduce the risks of errors and systematic biases that can arise in the map making stage in the presence of a mask as inhomogeneous as that of the KiDS-450 data. These two estimators produce correlated measurements, but the scales they are probing differ. The $\gamma_{\rm t}$ estimator is accurate at the few per cent level, as verified on full mock data in HD16, and the POLSPICE code has been thoroughly verified and validated on the same mocks as well. We refer the reader to K16 and HD16 for details of these tests.

4.3 Covariance estimation

The $\kappa_{\rm CMB}$ map reconstructed by the *Planck* data is noise dominated for most Fourier modes (Planck Collaboration XV 2016b). It is only by combining the full sky temperature and polarization maps that the *Planck* Collaboration could achieve a lensing detection of 40σ.

Since the noise $N_{\rm CMB}$ is larger than the signal $\kappa_{\rm CMB}$ at every scale included in our analysis (HD16), we can evaluate the covariance matrix from cross-correlation measurements between the 100 *Planck* simulated lensing maps (also provided in their 2015 public data release) and the tomographic KiDS data:

$$\mathrm{Cov}_{\ell\ell'}^{\kappa_{\rm CMB}\kappa_{\rm gal}} \simeq \left\langle \Delta\hat{C}_\ell^{N_{\rm CMB}\kappa_{\rm gal}} \Delta\hat{C}_{\ell'}^{N_{\rm CMB}\kappa_{\rm gal}} \right\rangle \quad (8)$$

and

$$\mathrm{Cov}_{\vartheta\vartheta'}^{\kappa_{\rm CMB}\gamma_{\rm t}} \simeq \left\langle \Delta\hat{\xi}_\vartheta^{N_{\rm CMB}\gamma_{\rm t}} \Delta\hat{\xi}_{\vartheta'}^{N_{\rm CMB}\gamma_{\rm t}} \right\rangle . \quad (9)$$

where the 'hats' refer to measured quantities, $\Delta\hat{x} = \hat{x} - \bar{x}$, and the brackets represent the average over the 100 realizations. This method assumes that the covariance is completely dominated by the CMB lensing and neglects the contribution from the shear covariance. This is justified by the fact that the signal from the former is about an order of magnitude larger, and hence completely drives the statistical uncertainty (HD16). The error bars shown in Fig. 2 are obtained from these matrices (from the square root of the diagonals). For each tomographic bin, the $\mathrm{Cov}_{\ell\ell'}^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ matrix has 25 elements, whereas the $\mathrm{Cov}_{\vartheta\vartheta'}^{\kappa_{\rm CMB}\gamma_{\rm t}}$ matrix has 36. The 100 realizations are enough to invert these matrices one at a time with a controllable

[9] POLSPICE has adjustable internal parameters, and we use THETAMAX = 60 deg, APODIZESIGMA = 60 deg and NLMAX = 3000.

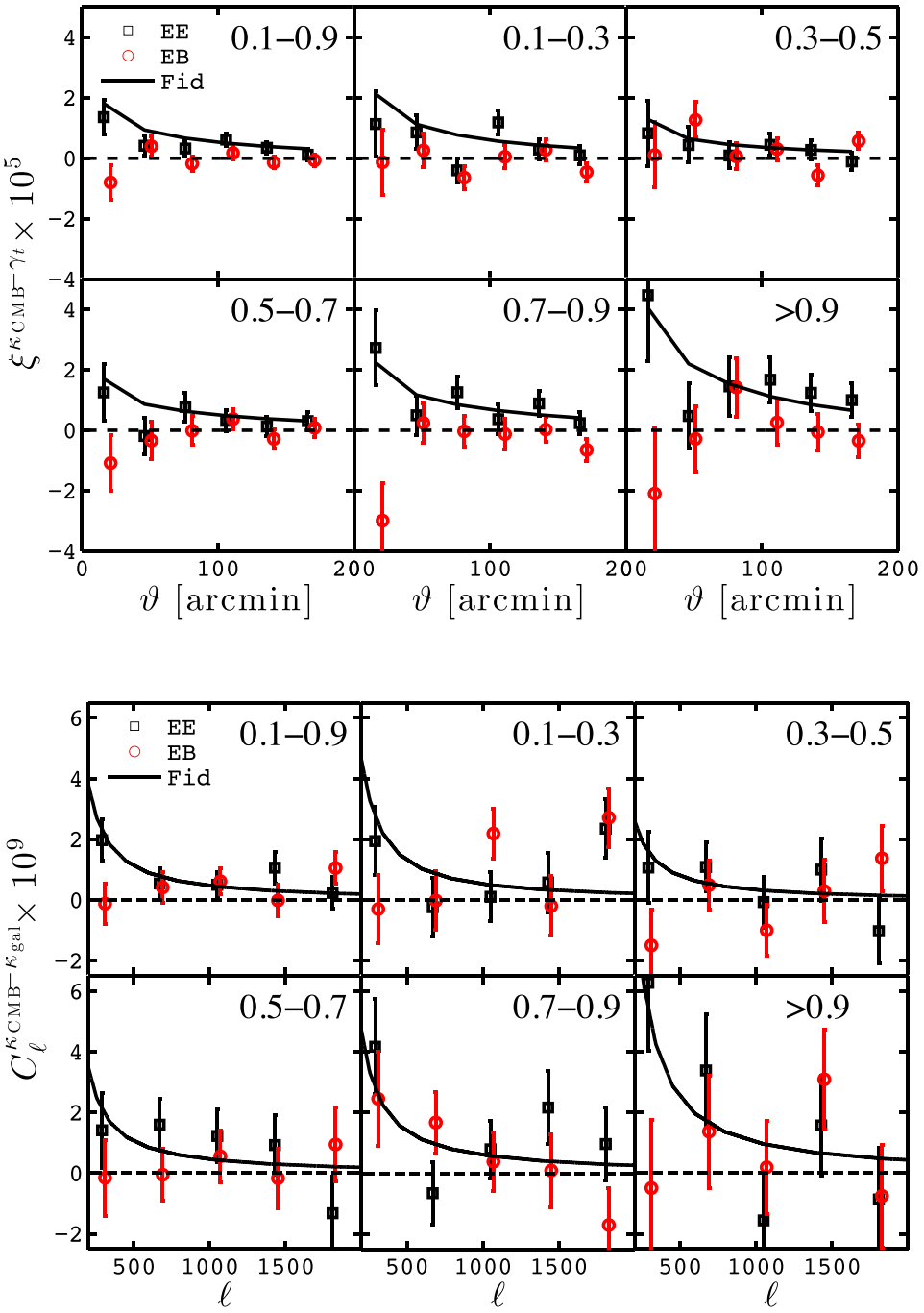

Figure 2. Cross-correlation measurement between *Planck* 2015 $\kappa_{\rm CMB}$ maps and KiDS-450 lensing data. The upper part presents results from the $\xi^{\kappa_{\rm CMB}\gamma_t}$ estimator, while the lower part shows the estimation of $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$. Different panels show the results in different tomographic bins, with predictions (solid curve) given by equations (1) and (4) in our fiducial cosmology. The black squares show the signal, whereas the red circles present the EB null test described in Section 5.2, slightly shifted horizontally to improve the clarity in this figure. The error bars are computed from 100 CMB lensing simulations.

level of noise bias, and the numerical convergence on this inverse is guaranteed (Lu, Pen & Doré 2010).

Note that this strategy fails to capture the correlation between tomographic bins, which are not required by our cosmological analysis presented in Section 5.6. If needed in a future analysis, these could be estimated from full light-cone mock simulations.

For both estimators, the covariance matrix is dominated by its diagonal, with most off-diagonal elements of the cross-correlation coefficient matrix being under ±10 per cent. Some elements reach larger values, ±40 per cent correlation at the most, but these are isolated, not common to all tomographic bins, and are consistent with being noise fluctuations, given that we are measuring many elements from 'only' 100 simulations. This partly explains why our cosmological results are not based on a joint tomographic analysis. We keep the full matrices in the analysis, even though we could, in principle, include only the diagonal part in the POLSPICE

measurement. Nevertheless, we have checked that our final results are only negligibly modified if we use this approximation in the χ^2 calculation, suggesting that one could reliably use a Gaussian approximation to the error estimation in this type of measurement (see equation 23 in HD16).

5 COSMOLOGICAL INFERENCE

Given the relatively low SNR of our measurement (Fig. 2), we do not fit our signal for the six parameters ΛCDM cosmological model. Instead, we follow the strategy adopted by earlier measurements: we compare the measured signal to our fiducial cosmological predictions, treating the normalization as a free parameter 'A'. If the assumed fiducial cosmology is correct and in absence of other systematic effects, A is expected to be consistent with unity. As discussed in previous studies, A is affected by a number of effects that can similarly modulate the overall amplitude of the signal. Aside from its sensitivity to cosmology – our primary science target – this re-scaling term will absorb contributions from residual systematic errors in the estimation of $n(z)$, from mis-modelling of the galaxy intrinsic alignments, from residual systematic bias in the shear multiplicative term m (equation 7), from astrophysical phenomena such as massive neutrinos and/or baryonic feedback, and from residual systematics in the cross-correlation estimators themselves (K16 and HD16).

In this section, we first present our constraints on A; we then quantify how the different effects listed above can impact our measurements, and finally present our cosmological interpretation. Our primary results assume the fiducial WMAP9+BAO+SN cosmology, i.e. we first place constraints on $A_{\rm fid}$; however, we also report constraints on $A_{\rm KiDS}$ and A_{Planck}, obtained by assuming different baseline cosmologies.

5.1 Significance

To measure A, we first compute the χ^2 statistic:

$$\chi^2 = \Delta x^T \, {\rm Cov}^{-1} \, \Delta x \tag{10}$$

with

$$\Delta x = \hat{\xi}^{\kappa_{\rm CMB}\gamma_{\rm t}} - A\xi^{\kappa_{\rm CMB}\gamma_{\rm t}} \text{ or } \Delta {\rm x} = \hat{C}^{\kappa_{\rm CMB}\kappa_{\rm gal}} - AC^{\kappa_{\rm CMB}\kappa_{\rm gal}} \tag{11}$$

for the configuration space and POLSPICE estimators, respectively. As before, quantities with 'hats' are measured, and the predictions assume the fiducial cosmology, unless stated otherwise. The SNR is given by the likelihood ratio test, which measures the confidence at which we can reject the null hypothesis (i.e. that there is no signal, simply noise) in favour of an alternative hypothesis described by our theoretical model with a single parameter A (see Hojjati et al. 2016, for a recent derivation in a similar context). We can write SNR $= \sqrt{\chi^2_{\rm null} - \chi^2_{\rm min}}$, where $\chi^2_{\rm null}$ is computed by setting $A = 0$, and $\chi^2_{\rm min}$ corresponds to the best-fitting value for A. The error on A is obtained by varying the value of A until $\chi^2_A - \chi^2_{\rm min} = 1$ (see, e.g. Wall & Jenkins 2003).

We include two additional statistical corrections to this calculation. The first is a correction factor that multiplies the inverse covariance matrix, $\alpha = (N_{\rm sim} - N_{\rm bin} - 2)/(N_{\rm sim} - 1) = 0.94$, to account for biases inherent to matrix inversion in the presence of noise (Hartlap, Simon & Schneider 2007). Here, $N_{\rm bin}$ is the number of data bins (5 for $C^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ and 6 for $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$) and $N_{\rm sim}$ is the number of simulations (100) used in the covariance estimation. There exists an improved version of this calculation based on assuming a t-distribution in the likelihood; however, with our values of $N_{\rm bin}$ and $N_{\rm sim}$, the differences in the inverted matrix would be of order 10–20 per cent (Sellentin & Heavens 2016), a correction on the error that we ignore given the relatively high level of noise in our measurement.

Table 2. Summary of χ^2, SNR and p-values obtained with the two different pipelines. The $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ measurements have 4 degrees of freedom (5 ℓ-bins – 1 free parameter), whereas the configuration space counterpart $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}(\vartheta)$ has one more, with 6 ϑ-bins. $A_{\rm fid}$ is the best-fitting amplitude that scales the theoretical signals in the fiducial cosmology, according to equation (11), also shown in Fig. 3. The numbers listed here include the covariance debiasing factor α and the extra error ϵ due to the noise in the covariance (see the main text of Section 5.1 for more details).

Z_B	Estimator	$\chi^2_{\rm min}$	$\chi^2_{\rm null}$	SNR	p-values	$A_{\rm fid}$
[0.1, 0.9]	$C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$	2.80	18.21	3.93	0.53	0.77 ± 0.19
	$\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$	2.88	22.94	4.48	0.64	0.69 ± 0.15
[0.1, 0.3]	$C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$	5.48	8.89	1.85	0.20	0.55 ± 0.30
	$\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$	7.93	13.38	2.34	0.12	0.53 ± 0.24
[0.3, 0.5]	$C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$	2.95	4.95	1.42	0.50	0.71 ± 0.51
	$\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$	1.44	4.19	1.66	0.84	0.60 ± 0.37
[0.5, 0.7]	$C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$	4.00	10.13	2.47	0.35	0.87 ± 0.35
	$\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$	2.00	6.45	2.11	0.77	0.55 ± 0.26
[0.7, 0.9]	$C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$	5.12	10.04	2.22	0.23	0.79 ± 0.36
	$\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$	2.78	15.41	3.55	0.65	1.02 ± 0.29
>0.9	$C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$	4.70	12.92	2.87	0.26	0.83 ± 0.29
	$\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$	4.68	22.64	4.24	0.38	0.95 ± 0.22

The second correction was first used in HD16 and consists of an additional error on A due to the propagated uncertainty coming from the noise in the covariance matrix (Taylor & Joachimi 2014). This effectively maps $\sigma_A \rightarrow \sigma_A(1 + \epsilon/2)$, where $\epsilon = \sqrt{2/N_{\rm sim} + 2(N_{\rm bin}/N^2_{\rm sim})} = 0.145$. These two correction factors are included in the analysis. The results from our statistical investigation are reported in Table 2, where we list $\chi^2_{\rm min}$, $\chi^2_{\rm null}$, SNR and A for every tomographic bin. The theoretical predictions provide a good fit to the data, given that for our degrees of freedom $\nu = N_{\rm bin} - 1$, $\nu - \sqrt{2\nu} < \chi^2_{\rm min} < \nu + \sqrt{2\nu}$. In other words, all our measured χ^2 fall within the expected 1σ error. We also compute the p-value for all these χ^2 measurements at the best-fitting A in order to estimate the confidence at which we can accept or reject the assumed model. Assuming Gaussian statistics, p-values smaller than 0.01 correspond to a 99 per cent confidence in the rejection of the model (the null hypothesis) by the data, and are considered 'problematic'. Our measured p-values, also listed in Table 2, are always larger than 0.12, meaning that the model provides a good fit to the data in all cases.

These tomographic measurements are re-grouped in Fig. 3, where we compare the redshift evolution of A for both estimators. We mark the 1σ region of the broad bin $n(z)$ with the solid horizontal lines, and see that all points overlap with this region within 1σ. This is an indication that the relative growth of structure between the tomographic bins is consistent with the assumed ΛCDM model. For the broad $n(z)$, the signal prefers an amplitude that is ∼23–31 per cent lower than the fiducial cosmology, i.e. the 1σ region shown by the horizontal solid lines in Fig. 3 is offset from unity by that amount. The main cosmological result that we quote from the $Z_B \in [0.1, 0.9]$ measurement is that of the $\gamma_{\rm t}$ estimator due to its higher SNR, as seen from comparing the top two rows of Table 2. For our fiducial cosmology, we find

$$A_{\rm fid} = 0.69 \pm 0.15. \tag{12}$$

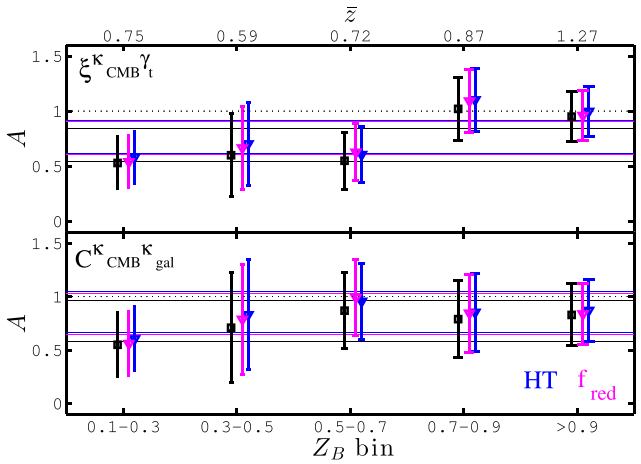

Figure 3. Tomographic measurement of $A_{\rm fid}$, defined in equation (11), assuming our fiducial cosmology. The two panels present results from the two cross-correlation estimators (labelled in the top left corner). Black symbols assume no IA, while colour symbols include correction factors from two IA models ($f_{\rm red}$ in magenta and HT in blue, see Section 5.3). The horizontal solid lines of a given colour enclose the 1σ region measured in the broad $Z_B \in [0.1, 0.9]$ bin, while the dotted horizontal lines indicate the fiducial values ($A_{\rm fid} = 1$). The mean source redshift in each bin is indicated at the top and summarized in Table 1. The mean in the first bin is high because of the long tail, visible in Fig. 1. The best-fitting values in different cosmologies are $A_{\rm fid} = 0.69 \pm 0.15$, $A_{\rm KiDS} = 0.86 \pm 0.19$ and $A_{Planck} = 0.68 \pm 0.15$.

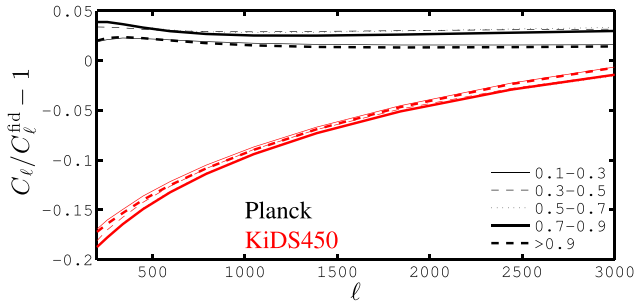

Figure 4. Fractional effect on the signal when changing the fiducial cosmology to *Planck* or KiDS-450. Different symbols show the impact in different tomographic bins, relative to the fiducial predictions. Current measurements are limited to $\ell < 2000$.

Varying the cosmology to the best-fitting KiDS-450 and *Planck* cosmologies,[10,11,12] we obtain

$$A_{\rm KiDS} = 0.86 \pm 0.19 \quad \text{and} \quad A_{Planck} = 0.68 \pm 0.15. \qquad (13)$$

The relative impact of these different cosmologies on our signal is presented in Fig. 4, where we see that the KiDS-450 cosmology mostly differ from the other two at large scales. The signal from our fiducial cosmology agrees with that assuming the best-fitting *Planck* cosmology to better than 5 per cent in all tomographic bins.

Fig. 3 demonstrates there are small but noticeable differences between the two estimators at fixed cosmology, especially at high redshift. As mentioned in Section 4.2, the scales being probed are not identical, and therefore some differences in the recovered values of A are expected. Nevertheless, within the current statistical accuracy, the two estimators are fully consistent with one another.

Visually, the $\xi^{\kappa_{\rm CMB}\gamma_t}(\vartheta)$ estimator seems to show a mild trend for decreasing values of A in lower redshift bins. Although such an effect could point towards a number of interesting phenomena suppressing power for source galaxies at $z \lesssim 0.7$ (e.g. modification to the growth history compared to the fiducial cosmology, additional feedback processes from baryons or massive neutrinos, or redshift-dependent contamination from IA), the significance of this redshift dependence is too low to draw any robust conclusions.

What is significantly seen from Fig. 3 is that the signal is generally low compared to the fiducial and *Planck* cosmologies. Our measurements of the amplitude A prefer instead the KiDS-450 cosmology, which also aligns with the CFHTLenS cosmic shear results (Kilbinger et al. 2013). We further quantify this comparison in Section 5.6, first presenting results from our set of null tests, and then examining three sources of contamination and systematic biases that potentially affect our signal.

In this work, we neglect the effect of source-lens coupling (Bernardeau 1998), which could possibly act as another secondary signal, biasing the signal low. As it is the case for cosmic shear, this effect should be too small (<10 per cent) to affect our results significantly, and further investigation will be required to interpret correctly the measurements from future surveys.

5.2 Null tests

We have shown in Section 4 and in Fig. 2 (red circles) that the parity violation EB test seemed consistent with noise in most tomographic bins, but occasionally that was not obvious. In this section, we investigate the significance of these measurements. Statistically, this is accomplished by measuring the confidence at which we can reject the null hypothesis 'parity is not violated'. We therefore re-run the full χ^2 statistical analysis[13] and measure the p-value about the model with $A = 0$. Low p-values correspond to high confidence of rejection, i.e. that some residual systematic effect might be causing and apparent parity violation. This type of measurement strongly probes the tail of the χ^2 distribution, hence assuming a Gaussian likelihood would provide inaccurate estimations of the p-values, even when including the Hartlap et al. (2007) debiasing α factor. Instead, we follow Sellentin & Heavens (2016) and assume a t-distribution for the likelihood, which better models the tail of the likelihood. Table 3 lists all these p-values, highlighting in bold one that seems slightly problematic (p-value ≤ 0.01). Since this single low p-value is only seen in one of the two estimators, we conclude that it must originate from expected noise fluctuations, and not from the data itself. This conclusion is additionally supported by the fact that the level of B-modes in the KiDS data (i.e. the BB measurement) is consistent with zero on the scales we are probing (Hildebrandt et al. 2017).

For the sake of testing the robustness of the EB POLSPICE measurement, we have additionally investigated the effect of changing the number of bins from 5 to 10. In the EE case, the recovered values of A and the SNR are similar to those presented in Table 2, from which we conclude that this comes with no gain. However, when applied to the EB null test, something interesting happens: the 'problematic' measurement (p-value = 0.01 in the $Z_B \in [0.1, 0.3]$ bin, Table 3) relaxes to 0.09, as seen in the column labelled

[10] Fiducial $(\Omega_{\rm m}, \Omega_\Lambda, \Omega_{\rm b}, \sigma_8, h, n_s) = (0.29, 0.71, 0.047, 0.83, 0.69, 0.97)$.
[11] KiDS-450 $(\Omega_{\rm m}, \Omega_\Lambda, \Omega_{\rm b}, \sigma_8, h, n_s) = (0.25, 0.75, 0.047, 0.85, 0.75, 1.09)$.
[12] *Planck* $(\Omega_{\rm m}, \Omega_\Lambda, \Omega_{\rm b}, \sigma_8, h, n_s) = (0.32, 0.68, 0.049, 0.80, 0.67, 0.97)$.

[13] Due to the absence of parameters in the null hypothesis, the EB case has one additional degree of freedom compared to the EE case.

Table 3. p-values for the EB test obtained for the six tomographic bins. Highlighted in bold is the p-value $\leq$0.01. The column labelled C_{10} refers to POLSPICE measurements in which the data are organized in 10 bins instead of 5. These calculations assume t-distributed likelihoods (following Sellentin & Heavens 2016).

Z_B	$C^{t\text{-dist}}$	$C_{10}^{t-\text{dist}}$	$\xi^{t\text{-dist}}$
[0.1, 0.9]	0.20	0.09	0.58
[0.1, 0.3]	**0.01**	0.09	0.52
[0.3, 0.5]	0.39	0.63	0.08
[0.5, 0.7]	0.94	0.57	0.78
[0.7, 0.9]	0.21	0.20	0.16
>0.9	0.53	0.54	0.68

C_{10}. This is another indication that the cause of the low p-value originates from fluctuations in the noise – which is affected in the binning process – without pointing to residual systematic effects in the data.

We have verified that our measurement of A is robust against the removal of some scales. When we exclude the largest or the smallest angular bin in the $\xi^{\kappa_{\rm CMB}\gamma_t}$ measurement, results change by at most 0.7σ, generally by less than 0.2σ. This gives us confidence in the robustness of our measurement. The same holds when removing the highest ℓ bin from the POLSPICE measurement, but not for the lowest ℓ bin, which captures the peak of the signal, and therefore contributes significantly to the SNR. At the same time, this test illustrates that we are currently not sensitive to the effect of massive neutrinos nor to baryonic feedback, which affect mainly these non-linear scales.

5.3 Effect of intrinsic alignments

Intrinsic alignments (IA) are a known secondary effect to the cross-correlation of galaxy lensing and CMB lensing that lowers the amplitude of the measured signal (Hall & Taylor 2014; Troxel & Ishak 2014; Chisari et al. 2015). It is therefore important to investigate how much IA could contribute to the observed low values of A reported in Table 2. To estimate the contamination level, we compare two different models, which we then apply equally to both estimators, $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ and $\xi^{\kappa_{\rm CMB}\gamma_t}(\vartheta)$.

First, we follow Hall & Taylor (2014, 'HT-IA' model hereafter) in using the 'linear non-linear alignment' model of Bridle & King (2007) with the SuperCOSMOS normalization found in Brown et al. (2002). We recall that this prescription comes from constraints at $z = 0.1$ that are independent of galaxy type or colour, and that the effect of IA in this model is to reduce the amplitude of the observed signal, as the galaxies tend to align radially towards each other. The scale-dependence of the alignment contribution is similar to the lensing signal, as seen in Hall & Taylor (2014) and in Fig. 5, hence we only quote the percentage of contamination at $\ell = 1000$ for reference. This also allows us to use with confidence the same IA contamination levels for the configuration space estimator, since re-scaling $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ by a constant re-scales $\xi^{\kappa_{\rm CMB}\gamma_t}$ by the same constant (as per equation 1). For each of the five redshift bins considered in this paper, starting from the lower redshift, we estimate a {10, 17, 10, 8, 5} per cent contamination to the signal, respectively. For the broader tomographic bin $Z_B \in [0.1, 0.9]$, we estimate a 11 per cent contamination. In other words, within the HT-IA model, the measured value of $A_{\rm fid}$ in the broad bin (equation 12) should be

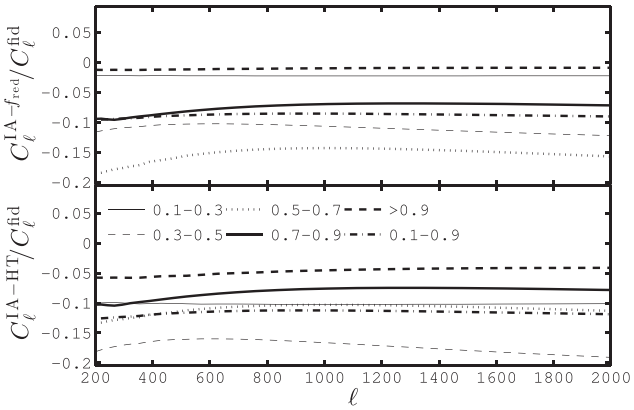

Figure 5. Strength of the contamination by intrinsic galaxy alignments for different tomographic bins, assuming our fiducial cosmology and the linear non-linear alignment model. The difference between lines is caused by changes in $n(z)$ (and in the red fraction in the $f_{\rm red}$-IA model).

corrected to

$$A_{\rm fid}^{\rm HT} = 0.77 \pm 0.15. \quad (14)$$

The error bars are not modified compared to the no-IA case since this contamination signal is additive. This model is the simplest as it assumes no luminosity or redshift dependence of the alignment normalization, and adopts the same alignment prescription for all galaxies regardless of morphological type/colour.

Secondly, we estimate the contamination from the alignment model of Chisari et al. (2015, '$f_{\rm red}$-IA' model hereafter) that allows for differential contributions based on galaxy colour/morphology. We assume that blue galaxies do not contribute at all, consistent with observations (Mandelbaum et al. 2011; Heymans et al. 2013), even though this null measurement remains poorly constrained. We estimate the red fraction directly from the data in each redshift bin using the best-fitting spectral template returned by BPZ for each source, referred to as T_B. Motivated by Simon et al. (2015), we identify red galaxies as objects with T_B <1.5. For the five tomographic bins, we obtain fractions of red galaxies $f_{\rm red} = \{0.04, 0.12, 0.27, 0.18, 0.04\}$; we estimate $f_{\rm red} = 0.15$ for the broad bin. We then use the alignment amplitude for the red galaxies from Heymans et al. (2013) to obtain an estimate of alignment contamination given our red fractions. These results are presented in the upper panel of Fig. 5. With this method, we estimate a {2, 11, 14, 7, 1} per cent contamination from intrinsic alignments in the tomographic bins, and a 9 per cent contamination in the $Z_B \in [0.1, 0.9]$ bin.[14] Then we can estimate

$$A_{\rm fid}^{f_{\rm red}} = 0.75 \pm 0.15. \quad (15)$$

One caveat with this model is that the K-correction and evolutionary corrections are uncertain at high redshift, which could result in biased estimates of the red fraction (see discussion in Chisari et al. 2015). This has an impact on the exact level of intrinsic alignment contamination by red galaxies, but we neglect this effect in this work.

Both methods are broadly consistent even though they differ in details, especially in the lowest redshift bin. For instance, the

[14] We measured the field-to-field variance in $f_{\rm red}$ and observed that it hardly varies except in the highest redshift bin, where the scatter could turn the 1 per cent IA contamination into a 0.0–2.5 per cent contamination. This remains small and should have a negligible impact.

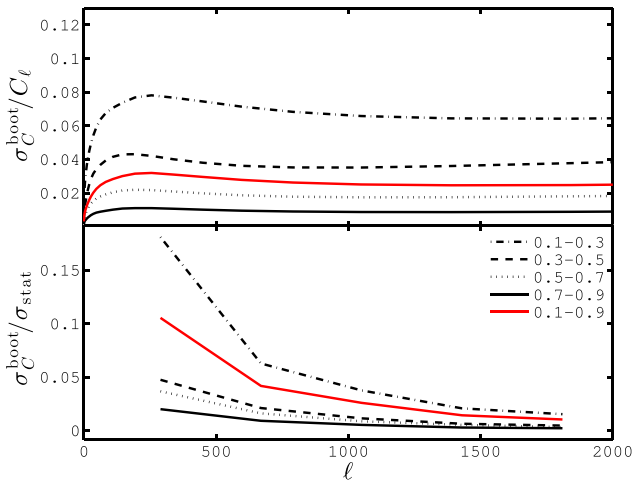

Figure 6. Top: fractional effect on the $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ signal when varying the $n(z)$ between 100 bootstrap re-samplings, for the four tomographic bins with $Z_B < 0.9$. Shown is the 1σ scatter divided by the signal. Bottom: ratio between the error from the $n(z)$ and the statistical error in our measurements.

second method captures the redshift differences observed in the data and takes into account the split in contributions arising from different galaxy types, which introduces a slightly different redshift dependence of the IA signal. The overall trends between the HT and the $f_{\rm red}$-IA models are similar though, but that is not the case for all IA models (see, for example, the tidal torque theory from Codis, Pichon & Pogosyan 2015, in which the sign of the IA effect on the signal is the opposite). There remains a large uncertainty in the modelling of the IA contamination, and we do not know which model, if any, should enter in our cosmological interpretation.

According to these estimations, both the HT-IA and $f_{\rm red}$-IA models help to bring A closer to unity. From the contamination levels listed above, at most 17 per cent of the observed cross-correlation signal can be cancelled by IA contamination in our tomographic bins. After correcting for this effect in each tomographic bin, most points agree with $A_{\rm fid} = 1$ within 1σ. This is shown with the colour symbols in Fig. 3.

Finally, we note that the uncertainty on the level of IA contamination quoted in the section is high, especially because of the unknown signal from the blue galaxies. For instance, at the 1σ level and assuming the linear non-linear alignment model, the IA contamination from blue galaxies could range from $-\{10, 15, 8, 6, 4\}$ per cent to $+\{6, 9, 5, 4, 3\}$ per cent in each tomographic bin, and from -10 per cent to $+6$ per cent in the $Z_B \in [0.1, 0.9]$ bin.

5.4 Effect of $n(z)$ errors

We investigate here the impact on our measurement of A from the uncertainty on the source redshift distribution. This is estimated from 100 bootstrap re-samplings of the source catalogue, as detailed in Hildebrandt et al. (2017, the DIR method described therein). These samples consist of internal fluctuations in the $n(z)$, which we turn into fluctuations in the signal with equations (1)–(3). We present in the top panel of Fig. 6 the fractional error on the signal, i.e. $\sigma_{C_\ell}^{\rm boot}/C_\ell$. According to this error estimate, the uncertainty on $n(z)$ is up to 8 per cent of the signal in the first redshift bin, then 4, 2 and 1 per cent for the others, and about 3 per cent for the $Z_B \in [0.1, 0.9]$ tomographic bin.

Note that this quantity is a measure of how the DIR $n(z)$ varies – and how it impacts the signal – across sub-samples of the re-weighted spectroscopically matched catalogue. This catalogue is by itself sub-sampling the full KiDS sources, and hence subject to sampling variance. It is therefore likely that the error quoted above slightly under-estimates the true error on the signal due to the $n(z)$, as discussed in Section C3.1 of Hildebrandt et al. (2017). The bottom panel of Fig. 6 show the importance of this error on $n(z)$ relative to the statistical uncertainty in our measurement. We see that this is sub-dominant in all bins and at all scales, and is therefore not expected to affect our results. Future surveys with higher precision will need to take this into account and fully propagate the uncertainty on $n(z)$ in the final error budget.

For comparison purposes, we also investigated estimates of the redshift distribution determined using the cross-correlation between spectroscopic and photometric samples (known as the CC method in Hildebrandt et al. 2017; Morrison et al. 2017). This scheme has a high level of noise compared to the fiducial DIR method and we find that the error on the recovered C_ℓ in our analysis increases from $\sim$5 per cent in the DIR case to $\sim$30 per cent in the CC case. From this we can draw the same conclusion, as the KiDS-450 cosmic shear analysis, that determining the redshift distribution using the cross-correlation CC method will remove any discrepancy with a *Planck* cosmology through the inflation of the error bars. We believe, however, that the error on the CC estimate is not representative of our actual knowledge of the $n(z)$ in the KiDS data and refer instead to the redshift distribution defined using DIR method in the rest of this paper.

Precision on the KiDS source redshift distribution will soon increase thanks to the ongoing processing of near-IR VIKING data (Hildebrandt et al., in preparation), which primarily impact the high-redshift tail so crucial to our measurement. Finally, note that in the DIR method we are using a calibrated $n(z)$, estimated from weighted spectroscopic data, hence we do not have to worry nearly as much about catastrophic photo-z outliers. This was not the case for the analysis presented in HD16, which showed that for $n(z)$ estimated directly from photometric data (for e.g. CFHTLenS and RCSLenS), these can easily dominate the error budget, with systematic effects on the signal of the order of 15 per cent. If our measurement contains more high-redshift objects than our $n(z)$ suggests, our predictions are too low; correcting for this would lower A.

5.5 Baryon feedback, massive neutrinos and non-linear modelling

As shown in HD16, baryonic feedback and massive neutrinos can cause an important decrease of the cross-correlation signal, which would translate into lower values of A when compared to a fiducial dark matter only cosmology.

To investigate how this could affect our cosmological results, we modify the $P(k, z)$ term in equation (1) to include 'massive neutrino bias' and 'baryon feedback bias' as detailed in Harnois-Déraps et al. (2015). The baryon bias was extracted from the OWL simulations, assuming the AGN model (van Daalen et al. 2011), while the neutrino bias was extracted from the recalibrated HALOFIT code (Takahashi et al. 2012) with total neutrino masses $M_\nu = 0.05$, 0.2, 0.4 and 0.6eV. Our results are presented in Fig. 7 for two simplified cases, in which the source galaxy populations are placed on single planes at $z_s = 0.5$ (in red) and at $z_s = 1.5$ (in blue). The figure focuses on the 0.05 eV scenario, showing the suppression of power caused by massive neutrinos (3.5 per cent effect on A for both z_s planes, averaged over the ℓ-modes that we measured),

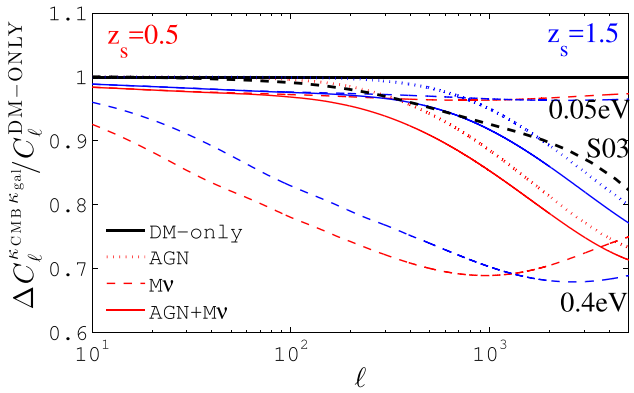

Figure 7. Fractional effect of the AGN baryon feedback and massive neutrinos on the cross-spectra for different combinations of source planes. The red solid line shows the combined effect on the cross-spectrum for sources placed on a single plane at $z_s = 0.5$. The effect of 0.05eV massive neutrinos and AGN feedback are shown separately by the upper dashed and the dotted line (also in red). The lower dashed redline shows the impact of 0.4 eV neutrinos. Blue lines show the same quantities, but for sources placed at $z_s = 1.5$. The dashed black line shows the ratio between the predictions from Smith et al. (2003) and that of Takahashi et al. (2012).

by baryonic feedback (5.0 per cent for $z_s = 1.5$ and 10.6 per cent for $z_s = 0.5$) and by the combination of both (8.2 per cent and 13.7 per cent for $z_s = 1.5$ and 0.5, respectively). The reason why the effect of baryons is larger on the lower redshift source plane is simply a projection effect: the same physical scales subtend different angles on the sky, which contribute differently to our measurement restricted to the $\ell \in$ [20–2000] range. We also show the effect of 0.4 eV neutrinos (28 per cent and 30 per cent in the two z_s slices), which demonstrates a scaling of 7 per cent per 0.1eV. We note that Mead et al. (2015) proposes an alternative method to account simultaneously for baryons and neutrinos based on the halo model, which might prove useful in future work.

These two effects contribute at some level to the measurement of A, but it is too early to put constraints on them based on our measurement. First, the cosmology is not guaranteed to be that of *Planck*; secondly, the exact feedback mechanism that is at play in the Universe remains largely unknown; and thirdly, other effects (e.g. IA contamination or error in the $n(z)$) could explain why our measured $A_{\rm fid}$ is low. However, if the fiducial cosmology is correct and if the intrinsic alignments are well described by the HT model described in Section 5.3, then $A_{\rm fid}$ would be brought to unity with $M_\nu = 0.33 \pm 0.22$ eV in absence of baryonic feedback, and $M_\nu = 0.19 \pm 0.22$ eV within the AGN model.

We have verified that the uncertainty in the non-linear modelling does not affect our measurement of A significantly. This is mainly because the angles and redshifts probed by our measurement correspond to scales that are mostly in the linear and mildly non-linear regime. Replacing the non-linear power spectrum from the Takahashi et al. (2012) model with that of Smith et al. (2003), a radical change in the non-linear predictions at small scales shown in Fig. 7 (black dashed line), affects our measurement of A by 1–2 per cent only. This is well within the statistical uncertainty and can be safely neglected. Fig. 7 shows that there is a clear degeneracy between differences in the two models, and the effect of baryonic feedback. However, the Smith et al. (2003) predictions are known to suffer from a significant loss of power at small scales, visible in Fig. 7, and the state-of-the-art precision on the non-linear power spectrum, from e.g. the Cosmic Emulator (Heitmann et al. 2014) deviates from the Takahashi et al. (2012) model by less than 5 per cent (Mead et al. 2015). This alleviates the degeneracy between modelling and baryonic feedback effects and further supports our (model-dependent) neutrino mass contraints presented above.

5.6 Cosmology from broad $n(z)$

In this section, we investigate how our cross-correlation measurement can constrain cosmology, and specifically compute confidence regions in the $[\sigma_8$–$\Omega_{\rm m}]$ plane. For this calculation, we assume massless neutrinos, no baryonic feedback, we ignore the error on $n(z)$, but examine our results for the two IA models (as well as the 'no-IA' case) described in Section 5.3.

It was shown in Liu & Hill (2015) that the amplitude of the cross-correlation signal scales approximately with $[\sigma_8^2\Omega_{\rm m}^{-0.5}]$ at large, linear scales ($\ell <$ few hundred), and as $[\sigma_8^3\Omega_{\rm m}^{1.3}]$ at small scales ($\ell > 1000$). Most of our constraints come from small scales, but our measurement includes some large modes down to $\ell \sim 200$. For this reason, we strike a compromise: we keep the $\Omega_{\rm m}^{1.3}$ dependence, as suggested by Liu & Hill (2015), but use a $\sigma_8^{2.5}$ dependence, to capture the gradual transition between both. Future measurements will require MCMC algorithms to be run to better capture these dependencies, but this does not seem to be necessary in this case given the relatively large uncertainty on A.

As discussed before, we use the $\xi^{\kappa_{\rm CMB}\gamma_t}$ results in the broad $Z_B \in$ [0.1, 0.9] tomographic bin because it has the highest signal to noise; however, our results would not change significantly if we used the $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ measurement instead. We could also have used the tomographic results, i.e. the $A(z)$ in the four bins. However, these measurements are all correlated, probing common low-redshift lenses. This would require us to calculate and include cross-correlation coefficients between the different tomographic bins when solving for the best-fitting cosmology. These could be evaluated from mock data, but this is not required when working with a single data point for A. Combining this scaling relation with equations (12)–(14), we get

$$A = A_{\rm fid}\left(\frac{\sigma_8}{0.831}\right)^{2.5}\left(\frac{\Omega_{\rm m}}{0.2905}\right)^{1.3} = 0.69 \pm 0.15 \qquad (16)$$

and

$$A = A_{\rm fid}^{\rm HT}\left(\frac{\sigma_8}{0.831}\right)^{2.5}\left(\frac{\Omega_{\rm m}}{0.2905}\right)^{1.3} = 0.77 \pm 0.15, \qquad (17)$$

which we use to propagate the error on A into confidence regions in the $[\sigma_8$–$\Omega_{\rm m}]$ plane. We show in Figs 8 and 9 how these constraints compare to the results from KiDS-450 cosmic shear (with IA), *Planck* and pre-*Planck* CMB experiments.[15] Our cross-correlation measurement has a larger overlap with the KiDS-450 constraints but is still consistent with the *Planck* cosmology in the sense that their $\lesssim$95 per cent confidence regions overlap. Including IA reduces the offset from *Planck*.

Given that our signal has different dependences on cosmological (e.g. $\Omega_{\rm m}$, σ_8) and nuisance (e.g. m, $n(z)$) parameters, we can see

[15] The MCMC chains entering these contour plots can be found on the KiDS-450 website: kids.strw.leidenuniv.nl/cosmicshear2016.php. Note also that the WMAP9+SPT+ACT cosmology presented in Figs 8 and 9 differs from the fiducial WMAP9+BAO+SN cosmology and is described in Calabrese et al. (2013) and Calabrese et al. (2017).

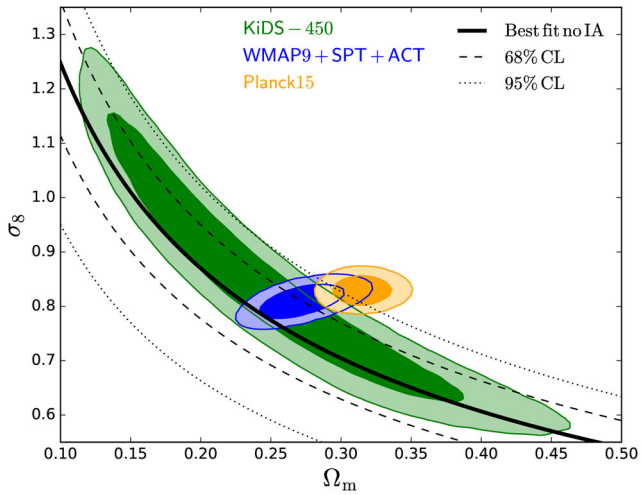

Figure 8. Constraints on σ_8 and Ω_m as estimated from the cross-correlation measurement, ignoring potential contamination by intrinsic galaxy alignments (shown in black). The solid line shows the best fit, while the dashed and dotted lines indicate the 68 per cent and 95 per cent confidence level (CL) regions, respectively. The cross-correlation results can be compared to KiDS-450 (green, where IA are accounted for), *Planck* (orange) and WMAP9+SPT+ACT (blue).

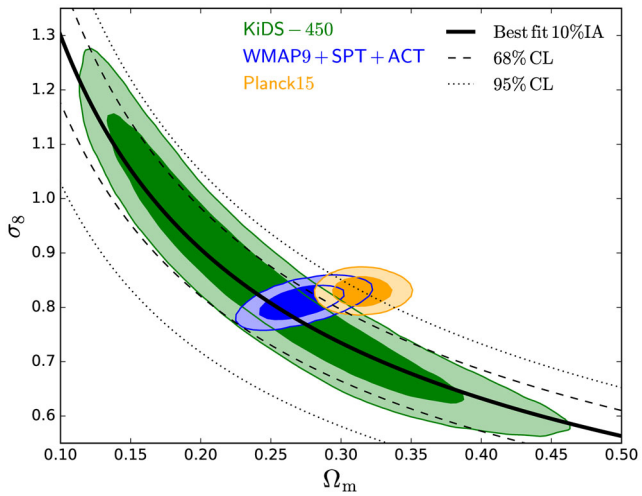

Figure 9. Same as Fig. 8, but here assuming 10 per cent contamination from IA in the cross-correlation measurement (equation 17), consistent with both the 'HT-IA' and the '$f_{\rm red}$-IA' models.

how this can provide new insights in resolving tensions between the cosmic shear and CMB measurements. For example, whereas the KiDS-450 and CFHTLenS cosmic shear results scale as $[m^2n^2(z)]$, our KiDS-450 × *Planck* lensing measurement scales as $[m\,n(z)]$. This difference could therefore allow us to break the degeneracy in a joint probe analysis. Also note that in general, we should not exclude possibility that there could be residual systematics left over in a CMB temperature and polarization analysis – driving the cosmology to higher $[\sigma_8, \Omega_m]$ values – that do not make their way to the CMB lensing map or into the joint probes measurement, in analogy with the cosmic shear c-term. This is certainly the case for the additive shear bias (the c-correction) described in Kuijken et al. (2015). Having this new kind of handle can help to identify the cause of disagreements between different probes, and will be central to the cosmological analyses of future surveys. We explore further how cross-correlation analyses can be turned into a calibration tool in the next section.

5.7 Application: photo-z and m-calibration

The CMB lensing – galaxy lensing cross-correlation signal has been identified as a promising alternative to calibrate the cosmic shear data without relying completely on image simulations (Das et al. 2013; Liu et al. 2016; Schaan et al. 2017). This statement relies on the fact that A absorbs all phenomena that affect the amplitude of the measurement, i.e. cosmology, intrinsic alignment, $n(z)$, shear calibration and that we can marginalize over some of these in order to solve for others.

Most of the attention so far has been directed towards the multiplicative term in the cosmic shear calibration – the m^j factor in equation (7) – which has an important impact on the cosmological interpretation. In the case of the KiDS-450 data, the shear calibration is known at the per cent level from image simulations for objects with $Z_B < 0.9$ (see Fig. 11), but the precision on m^j quickly degrades at higher redshift (Fenech Conti et al. 2017). Similar conclusions can be drawn from the photometric redshift estimation, which becomes unreliable at high redshift when only using optical bands (Hildebrandt et al. 2017). We see in our cross-correlation measurement a unique opportunity to place a joint-constraint on these two quantities in the highest redshift bin, informed by our measurement at lower redshift. We ignore the contribution from IA due to the high level of statistical noise in our measurement. However, this will need to be included in similar analyses of upcoming surveys with higher statistical precision.[16]

Our approach is to fix the scaling term A to the value preferred by the $Z_B \in [0.1, 0.9]$ data, which we label $A_{\rm low_z}$ here for clarity, and to jointly fit for the mean shear bias and mean redshift distribution in the $Z_B > 0.9$ bin. Forcing A to this value in the high redshift bin provides constraints on $\langle m_{\rm high_z}\rangle$ and $\langle n_{\rm high_z}(z)\rangle$, which we extract by varying these quantities in the predictions.

The correction to the shear bias is trivial to implement as it scales linearly with A, so we simply write $A_{\rm high_z} = A_{\rm low_z}(1 + \delta_m) = 0.95 \pm 0.22$ (from Table 2) and solve for δ_m. If this was the only correction, we could write $\delta_m = A_{\rm high_z}/A_{\rm low_z} - 1 = 0.38 \pm 0.44$, which is consistent with zero but not well constrained.

Corrections to the photometric distribution can be slightly more complicated as the full redshift distribution that enters our calculation is not simple, as seen in Fig. 1. There are a number ways with which we could alter the $n(z)$ and propagate the effect on to the signal, e.g. by modifying the overall shape, the mean or the tail of the distribution. We opted for arguably the simplest prescription, which consists in shifting the $n(z)$ along the z direction by applying the mapping $z \rightarrow z + \delta_z$ (thereby shifting $\langle n_{\rm high_z}(z)\rangle$ by the same amount). We propagate this new $n(z)$ through equation (1) and solve for values of δ_z that satisfy constraints on A. In this process, we allow δ_z to vary by up to 0.5, which is rather extreme.

[16] One might well object that the uncertainty in IA modelling and its evolution is already larger than the uncertainty in shear calibration, and hence that our strategy is flawed to start with. Instead, we should be placing simultaneous constraints on the photo-z, m-calibration and IA. A full MCMC will certainly be required in the future to disentangle these effects, exploiting their different shape dependence to break the degeneracy between these parameters. As an illustration of this strategy, however, we present a simple case and assume no IA contamination in the rest of this section.

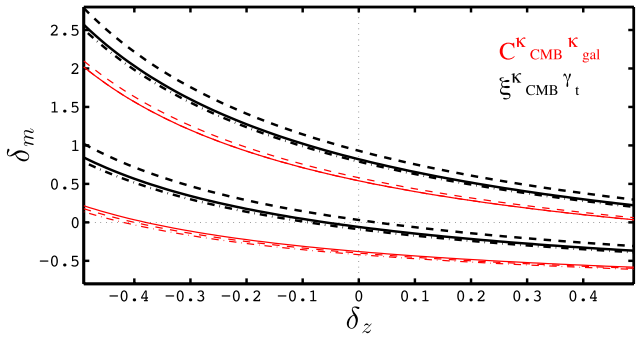

Figure 10. 1σ contour regions on the shear calibration correction δ_m and the redshift distribution correction δ_z in the bin $Z_B > 0.9$, from the cross-correlation measurements. Black and red correspond to constraints from $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$ and $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$, respectively. The multiple lines present the results in three different cosmologies (fiducial is solid, KiDS-450 is dot–dashed, *Planck* is dashed), which are shown here to have a small impact on the constraints. Other independent measurements and improved image simulations could tighten the region of allowed values. The upper red solid and dot–dashed lines perfectly overlap.

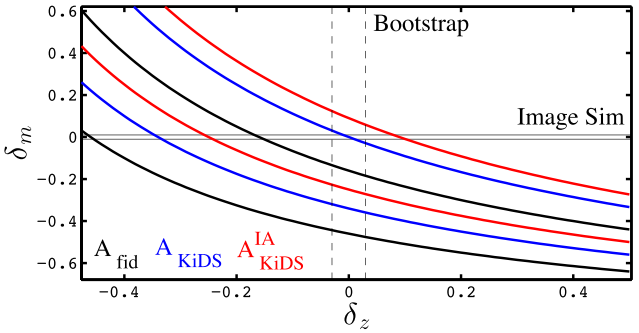

Figure 11. 1σ contour regions on the shear calibration correction δ_m and the redshift distribution correction δ_z derived from the $Z_B \in [0.1, 0.9]$ measurement of A in three different cases. Results from $A_{\rm fid}$ (no-IA) are shown in solid black, results from $A_{\rm KiDS}$ (no IA) are shown in solid blue and results from $A_{\rm KiDS}$ with 10 per cent IA are shown in solid red. The pair of solid horizontal lines shows the region of δ_m values allowed from image simulations, while the pair of dashed vertical lines shows the region of δ_z values allowed from bootstrap re-sampling the $n(z)$.

Following the simple reasoning described for the shear calibration, we can see that if m was trusted at the per cent level in this high-redshift bin, constraints on the redshift distribution could be simply derived by computing $A_{\rm highz}/A_{\rm lowz} = C_{\delta_z}/C_{\rm fid} = (1 + \delta_z)$. We therefore obtain the exact same constraints as for δ_m, namely $\delta_z = 0.38 \pm 0.44$. We place constraints on the $[\delta_m$–$\delta_z]$ plane by requiring $(1 + \delta_m)(1 + \delta_z) = A_{\rm highz}/A_{\rm lowz}$ and present the 1σ constraints in Fig. 10. The data are still consistent with $\delta_m = \delta_z = 0$, but these two biases are not currently well constrained.

We also show in Fig. 10 the results from the $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ estimator (in red dashed), but these have a lower SNR hence are not included in the analysis. At first sight, the difference observed between the results from the two estimators could seem worrisome. Given that these constraints on $[\delta_m$–$\delta_z]$ are obtained from the same data and that the only difference is the analysis method, it is justified to question whether we could use this measurement for precise self-calibration if two methods on the same data give such different values for δ_m and δ_z. We recall that differences are expected since both techniques are probing different scales; however, the calibration technique presented here is sensitive to these differences. The calibration is only weakly sensitive to the fiducial cosmology adopted, as shown in Fig. 10 with the different line styles.

A significant improvement will come from the future data sets (advanced-ACT, SPT-3G, LSST, Euclid), in which the noise will be much lower, allowing for more accurate measurements of $\xi^{\kappa_{\rm CMB}\gamma_{\rm t}}$ and $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$ to start with. In addition, including other measurements in this self-calibration approach will greatly enhance the achievable precision. For example, one could measure the galaxy–galaxy lensing signal from the same KiDS-450 source galaxies, using i.e. the GAMA galaxies as lenses (van Uitert et al. 2016), selecting the sources in the same tomographic bin (i.e. $Z_B \in [0.1, 0.9]$ and $Z_B > 0.9$). Fixing the cosmology from the low redshift bin, one could then similarly constrain $[\delta_m$–$\delta_z]$ in the high redshift bin. The idea here is that the trend can be made *opposite* to that seen in Fig. 10: an increase in δ_z pushes the sources away from the lenses, which, depending on the geometry, could reduce the signal. To compensate for this, the m-calibration would need to increase as well. In such a set-up, the preferred region in parameter space would inevitably intersect with ours, and exploiting this complementarity might lead to competitive constraints. Further investigation on this combined measurement will be explored in upcoming work. We are aware that our bi-linear modelling of the m and $n(z)$ calibration is an oversimplification of our knowledge (and uncertainty) about these quantities in the highest redshift bin, and one could envision improving this strategy in the future. For instance, the high-redshift objects are often the hardest to measure spectroscopically; hence, there are higher chances that the DIR method fails at higher redshifts. To capture this effect, instead of shifting the $n(z)$, one could modify only the high-redshift tail, moving 1 per cent, 5 per cent or 10 per cent of our source galaxies from (very) low redshifts to $z > 1$, propagating the effect on the signal, and use our measurement of A to constrain the fraction of such 'missing' high-redshift galaxies. However, given the size of our error bars, it is not clear that we would learn more from this approach at the moment.

This situation will improve significantly with future CMB and galaxy surveys. According to Schaan et al. (2017), the lensing data provided by a Stage-4 CMB experiment, combined with 10 tomographic bins for LSST, will enable a m-calibration that is accurate to better than 0.5 per cent. This is marginalizing over a number of nuisance parameters that unfortunately does not include catastrophic photometric redshift outliers, so the actual accuracy will likely degrade compared to this impressive benchmark. Nevertheless, this is an avenue that is certainly worth exploiting with the upcoming data.

The $Z_B < 0.9$ redshift data in the KiDS survey has been calibrated on image simulations whose precision on δ_m largely surpasses that of the cross-correlation technique presented in this section. Fig. 11 shows the 1σ constraints in the $[\delta_m$–$\delta_z]$ plane in the $Z_B \in [0.1, 0.9]$ bin, assuming the fiducial cosmology without IA (black), the KiDS-450 cosmology without IA (blue) and the KiDS-450 with 10 per cent IA (red), consistent with both the HT-IA and the $f_{\rm red}$-IA models. For comparison, the 1 per cent precision on δ_m obtained from image simulations and the 3 per cent precision on δ_z obtained from bootstrap re-sampling the $n(z)$, described in Section 5.4, are shown as the pairs of horizontal and vertical lines, respectively. For these redshifts at least, the measurement provides interesting constraints on the cosmology, IA and δ_z, but not on the m-calibration.

6 CONCLUSIONS

We perform the first tomographic lensing-lensing cross-correlation by combining the *Planck* 2015 lensing map with the KiDS-450 shear data. Our measurement is based on two independent estimators, the POLSPICE measurement of $C_\ell^{\kappa_{\rm CMB}\kappa_{\rm gal}}$, and the configuration-space measurement of $\xi^{\kappa_{\rm CMB}\gamma_t}(\vartheta)$. The two techniques agree within 1σ in all tomographic bins, although the former exhibits a lower SNR.

We compare our tomographic results against a two-dimensional lensing analysis of a single broad redshift bin ($Z_B \in [0.1, 0.9]$), and fit the measured amplitude of the signal with a single multiplicative parameter A that scales the predictions. We obtain $A_{\rm fid} = 0.69 \pm 0.15$ in our fiducial cosmology and show that the constraints on the $[\sigma_8 - \Omega_{\rm m}]$ plane are consistent with the flat ΛCDM *Planck* cosmology at the 95 per cent level, with $A_{Planck} = 0.68 \pm 0.15$ and with all previous results (Hand et al. 2015; Liu & Hill 2015; Singh et al. 2017; K16 & HD16). The KiDS-450 cosmology is preferred, however, in which we obtain $A_{\rm KiDS} = 0.86 \pm 0.19$.

Photometric redshifts have been examined carefully and are unlikely to be affecting these results significantly (<8 per cent effect on the signal), unless the spectroscopic sample that is used to estimate the $n(z)$ suffers from significant sampling variance. Multiplicative shear calibration is also highly unlikely to be affecting A, since it is known to be accurate at the percent level over the redshift range that enters our cosmological measurement. However, including different models of intrinsic alignment, massive neutrinos and baryon feedback in the predictions all affect the signal by tens of percent, pushing the recovered A to higher values.

Fixing the cosmology to that favoured by our low-redshift measurements ($Z_B < 0.9$), we calibrate the high-redshift ($Z_B > 0.9$) photometric $n(z)$ and the multiplicative shear calibration, which are not robustly constrained. We find that the high-redshift data are consistent with no residual systematics, but that these are still allowed and only weakly constrained. Improved results on this high-redshift calibration will come in the future from larger data sets, from improved image simulations and from the combination with other independent measurements.

Tomographic measurements such as that presented in this paper are insensitive to galaxy bias, and hence opening the possibility to obtain cosmological constraints from measurements of the growth factor. Upcoming and future lensing surveys will have excellent opportunities for combining probes and improving their cosmological analyses.

ACKNOWLEDGEMENTS

Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme IDs 177.A-3016, 177.A-3017 and 177.A-3018, and on data products produced by Target/OmegaCEN, INAF-OACN, INAF-OAPD and the KiDS production team, on behalf of the KiDS consortium. OmegaCEN and the KiDS production team acknowledge support by NOVA and NWO-M grants. Members of INAF-OAPD and INAF-OACN also acknowledge the support from the Department of Physics & Astronomy of the University of Padova and of the Department of Physics of University Federico II (Naples).

JHD acknowledges support from the European Commission under a Marie-Skłodowska-Curie European Fellowship (EU project 656869). TT, LvW and AH are supported by the NSERC of Canada, TT is additionally funded by CIFAR. HHi is supported by an Emmy Noether grant (No. Hi 1495/2-1) of the Deutsche Forschungsgemeinschaft. MV, CH and MA acknowledge support from the European Research Council under European Union's Seventh Framework Programme (FP7) grant number 279396 (MV) and 647112 (CH, MA). MV is also supported by the Netherlands Organization for Scientific Research (NWO) through grants 614.001.103. Parts of this research were conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020. JM has received funding from the People Programme (Marie Curie Actions) under the FP7 REA grant agreement number 627288. ENC is supported by a Beecroft fellowship. PS is supported by the Deutsche Forschungsgemeinschaft by the TR33 'The dark Universe'. LM is supported by STFC grant ST/N000919/1, and KK acknowledges support by the Alexander von Humboldt Foundation.

Author Contributions: All authors contributed to the development and writing of this paper. The authorship list reflects the lead authors of this paper (JHD, TT and ENC), followed by two alphabetical groups. The first alphabetical group (CH and LvW) consists of authors who contributed both to the infrastructure of the KiDS data that were used in this work and to the analysis itself. Members of the second group are infrastructure contributors whose products are directly used in this work.

REFERENCES

Baxter E. et al., 2016, MNRAS, 461, 4099
Benítez N., 2000, ApJ, 536, 571
Bernal J. L., Verde L., Riess A. G., 2016, JCAP, 10, 019
Bernardeau F., 1998, A&A, 338, 375
Bonvin V. et al., 2017, MNRAS, 465, 4914
Bridle S., King L., 2007, New J. Phys., 9, 444
Brown M. L., Taylor A. N., Hambly N. C., Dye S., 2002, MNRAS, 333, 501
Calabrese E. et al., 2013, Phys. Rev. D, 87, 103012
Calabrese E. et al., 2017, Phys. Rev. D, 95, 063525
Chisari N. E., Dunkley J., Miller L., Allison R., 2015, MNRAS, 453, 682
Chon G., Challinor A., Prunet S., Hivon E., Szapudi I., 2004, MNRAS, 350, 914
Codis S., Pichon C., Pogosyan D., 2015, MNRAS, 452, 3369
Das S., Errard J., Spergel D., 2013, preprint (arXiv:1311.2338)
Das S. et al., 2014, JCAP, 4, 14
de Jong J. T. A. et al., 2015, A&A, 582, A62
Erben T. et al., 2013, MNRAS, 433, 2545
Fenech Conti I., Herbonnet R., Hoekstra H., Merten J., Miller L., Viola M., 2017, MNRAS, 467, 1627
Giannantonio T. et al., 2016, MNRAS, 456, 3213
Hall A., Taylor A., 2014, MNRAS, 443, L119
Hand N. et al., 2015, Phys. Rev. D, 91, 062001
Harnois-Déraps J., van Waerbeke L., Viola M., Heymans C., 2015, MNRAS, 450, 1212
Harnois-Déraps J. et al., 2016, MNRAS, 460, 434 (HD16)
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399
Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, ApJ, 780, 18
Heymans C. et al., 2012, MNRAS, 427, 146
Heymans C. et al., 2013, MNRAS, 432, 2433
Hildebrandt H. et al., 2016, MNRAS, 463, 635
Hildebrandt H. et al., 2017, MNRAS, 465, 1454
Hinshaw G. et al., 2013, ApJS, 208, 19
Hojjati A. et al., 2016, MNRAS, preprint (arXiv:1608.07581)
Joudaki S. et al., 2016, MNRAS, preprint (arXiv:1610.04606)
Joudaki S. et al., 2017, MNRAS, 465, 2033
Kilbinger M. et al., 2009, A&A, 497, 677, 0810.5129
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Kilbinger M. et al., 2017, preprint (arXiv:1702.05301)
Kirk D. et al., 2016, MNRAS, 459, 21 (K16)
Kuijken K. et al., 2015, MNRAS, 454, 3500

Limber D. N., 1954, ApJ, 119, 655
Liske J. et al., 2015, MNRAS, 452, 2087
Liu J., Hill J. C., 2015, Phys. Rev. D, 92, 063517
Liu J., Ortiz-Vazquez A., Hill J. C., 2016, Phys. Rev. D, 93, 103508
Loverde M., Afshordi N., 2008, Phys. Rev. D, 78, 123506
Lu T., Pen U.-L., Doré O., 2010, Phys. Rev. D, 81, 123015
MacCrann N., Zuntz J., Bridle S., Jain B., Becker M. R., 2015, MNRAS, 451, 2877
Mandelbaum R. et al., 2011, MNRAS, 410, 844
Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544
Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958
Miller L. et al., 2013, MNRAS, 429, 2858
Miralda-Escude J., 1991, ApJ, 380, 1
Moraes B. et al., 2014, in Mateus A., Gregorio-Hetem J., Cid Fernandes R., eds, Rev. Mex. Astron. Astrofis. Conf. Ser. Vol. 44, XIV Latin American Regional IAU Meeting. Instituto de Astronomía. RMxAC, UNAM, Mexico, p. 202
Morrison C. B., Hildebrandt H., Schmidt S. J., Baldry I. K., Bilicki M., Choi A., Erben T., Schneider P., 2017, MNRAS, 467, 3576
Okamoto T., Hu W., 2003, Phys. Rev. D, 67, 083002
Omori Y., Holder G., 2015, preprint (arXiv:1502.03405)
Planck Collaboration XIII, 2016a, A&A, 594, A13
Planck Collaboration XV, 2016b, A&A, 594, A15
Schaan E., Krause E., Eifler T., Doré O., Miyatake H., Rhodes J., Spergel D. N., 2017, Phys. Rev. D, 95, 123512
Schneider P., 2003, A&A, 408, 829
Sellentin E., Heavens A. F., 2016, MNRAS, 456, L132
Simon P. et al., 2015, MNRAS, 449, 1505
Singh S., Mandelbaum R., Brownstein J. R., 2017, MNRAS, 464, 2120
Smith R. E. et al., 2003, MNRAS, 341, 1311
Szapudi I., Prunet S., Colombi S., 2001, ApJ, 561, L11
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Taylor A., Joachimi B., 2014, MNRAS, 442, 2728
Troxel M., Ishak M., 2014, Phys. Rev. D, 89, 063528
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
van Engelen A. et al., 2012, ApJ, 756, 142
van Uitert E. et al., 2016, MNRAS, 459, 3251
Wall J. V., Jenkins C. R., 2003, in Ellis R., Huchra J., Kahn S., Rieke G., Stetson P. B., eds, Practical Statistics for Astronomers. Vol. 3, Cambridge Univ. Press, Cambridge

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