Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **465,** 2033–2052 (2017)
Advance Access publication 2016 October 17

doi:10.1093/mnras/stw2665

CFHTLenS revisited: assessing concordance with Planck including astrophysical systematics

Shahab Joudaki,[1]⋆ Chris Blake,[1] Catherine Heymans,[2] Ami Choi,[2]
Joachim Harnois-Deraps,[3] Hendrik Hildebrandt,[4] Benjamin Joachimi,[5]
Andrew Johnson,[1] Alexander Mead,[3] David Parkinson,[6] Massimo Viola[7]
and Ludovic van Waerbeke[3]

[1]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[2]*Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[3]*Department of Physics and Astronomy, The University of British Columbia, 6224 Agricultural Road, Vancouver, BC, V6T 1Z1, Canada*
[4]*Argelander Institute for Astronomy, University of Bonn, Auf dem Hugel 71, D-53121 Bonn, Germany*
[5]*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*
[6]*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[7]*Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*

Accepted 2016 October 12. Received 2016 October 12; in original form 2016 January 21

ABSTRACT

We investigate the impact of astrophysical systematics on cosmic shear cosmological parameter constraints from the Canada–France–Hawaii Telescope Lensing Survey (CFHTLenS) and the concordance with cosmic microwave background measurements by Planck. We present updated CFHTLenS cosmic shear tomography measurements extended to degree scales using a covariance calibrated by a new suite of *N*-body simulations. We analyse these measurements with a new model fitting pipeline, accounting for key systematic uncertainties arising from intrinsic galaxy alignments, baryonic effects in the non-linear matter power spectrum, and photometric redshift uncertainties. We examine the impact of the systematic degrees of freedom on the cosmological parameter constraints, both independently and jointly. When the systematic uncertainties are considered independently, the intrinsic alignment amplitude is the only degree of freedom that is substantially preferred by the data. When the systematic uncertainties are considered jointly, there is no consistently strong preference in favour of the more complex models. We quantify the level of concordance between the CFHTLenS and Planck data sets by employing two distinct data concordance tests, grounded in Bayesian evidence and information theory. We find that the two data concordance tests largely agree with one another and that the level of concordance between the CFHTLenS and Planck data sets is sensitive to the exact details of the systematic uncertainties included in our analysis, ranging from decisive discordance to substantial concordance as the treatment of the systematic uncertainties becomes more conservative. The least conservative scenario is the one most favoured by the cosmic shear data, but it is also the one that shows the greatest degree of discordance with Planck. The data and analysis code are publicly available at https://github.com/sjoudaki/cfhtlens_revisited.

Key words: gravitational lensing: weak – surveys – cosmology: theory.

1 INTRODUCTION

The standard ΛCDM model of cosmology has been successful in describing the expansion history and growth of density perturbations throughout the Universe (e.g. Anderson et al. 2014; Betoule et al. 2014; Ade et al. 2016XIII). At the same time, it is facing challenges through our incomplete understanding of its main ingredients, namely the mechanism that is driving the current accelerated expansion and the dark matter (DM) that constitutes most of the matter in the Universe (e.g. Bertone, Hooper & Silk 2005; Copeland, Sami & Tsujikawa 2006; Feng 2010; Clifton et al. 2012). There are a range of late-time experimental techniques used to improve our understanding of the underlying cosmology of the Universe, such as supernova distances, baryon acoustic oscillations, galaxy cluster counting, and weak gravitational

---

⋆ E-mail: sjoudaki@swin.edu.au

Published by Oxford University Press on behalf of the Royal Astronomical Society

lensing, where lensing is considered to be one of the most promising as a result of its particular sensitivity to both structure formation and universal expansion (e.g. Albrecht et al. 2006; Joudaki & Kaplinghat 2012).

While weak lensing holds significant promise as a cosmological probe, the analyses of lensing data sets are still maturing. In particular, the optimism with weak lensing is predicated on overcoming the vast systematic uncertainties in both observations and theory. On the observational front, there are photometric redshift uncertainties (also denoted as 'photo-z'; e.g. Huterer et al. 2006; Ma, Hu & Huterer 2006; Bernstein & Huterer 2010; Bonnett et al. 2016) and intrinsic alignments (IA) of galaxies (e.g. Hirata & Seljak 2004; Bridle & King 2007; Joachimi et al. 2011, 2015; Troxel & Ishak 2015), along with additive and multiplicative corrections to the lensing observables, for instance due to point spread function (PSF) anisotropies and shear miscalibration (e.g. Hirata & Seljak 2003; Huterer et al. 2006; Heymans et al. 2012; Melchior & Viola 2012; Bernstein et al. 2016).

On the theoretical front, there are higher order correction terms in the lensing integral, for instance due to the Born approximation and lens–lens coupling, but these are negligible even for cosmic variance-limited surveys (e.g. Cooray & Hu 2002; Shapiro & Cooray 2006; Bernardeau, Bonvin & Vernizzi 2010; Krause & Hirata 2010). More crucially, lensing analyses need to account for reduced shear (e.g. Dodelson, Shapiro & White 2006; Shapiro 2009; Krause & Hirata 2010) and uncertainties in the modelling of the non-linear matter power spectrum. The former, if neglected, may induce a bias in the cosmological parameter estimates that exceeds the parameter uncertainties in future surveys such as Large Synoptic Survey Telescope (Shapiro 2009). The latter is true even when assuming all of the matter is collisionless and a cosmological constant drives late-time universal acceleration, both analytically and with simulations (e.g. Bernardeau et al. 2002; Cooray & Sheth 2002; Smith et al. 2003; Heitmann et al. 2014; McQuinn & White 2016; Mead et al. 2015).

There are additional difficulties in modelling the matter power spectrum due to baryonic physics coming from star formation, radiative cooling, and feedback processes (e.g. White 2004; Zhan & Knox 2004; Rudd, Zentner & Kravtsov 2008; van Daalen et al. 2011). The modelling of the non-linear matter power spectrum is also sensitive to extensions of the standard model, for example, to include massive neutrinos (e.g. Saito, Takada & Taruya 2008; Bird, Viel & Haehnelt 2012; Wagner, Verde & Jimenez 2012), dark energy (e.g. McDonald, Trac & Contaldi 2006; Joudaki, Cooray & Holz 2009; Alimi et al. 2010; Heitmann et al. 2014), and modified gravity (e.g. Stabenau & Jain 2006; Zhao, Li & Koyama 2011; Baldi et al. 2014; Hammami et al. 2015).

In this paper, we present a methodical study of three key 'astrophysical' systematic uncertainties affecting the lensing observables from the Canada–France–Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al. 2012; Hildebrandt et al. 2012; Erben et al. 2013; Miller et al. 2013) in the form of intrinsic galaxy alignments, baryonic effects in the non-linear matter power spectrum, and photometric redshift uncertainties. In addition to these astrophysical uncertainties, there are errors on the shear measurement itself which we calibrate through additive and multiplicative shear calibration corrections to the data as a function of galaxy size and signal to noise (Heymans et al. 2012; Miller et al. 2013). In a 2D analysis, Kilbinger et al. (2013) showed that the measured uncertainties in these corrections had a negligible impact on the cosmological constraints for CFHTLenS and so we do not consider 'shear measurement' systematic uncertainties in our analysis (still true with tomography given comparable constraints on $\sigma_8\Omega_{\mathrm{m}}^{0.5}$).

We account for the three key systematic uncertainties more comprehensively than previously, for instance, by incorporating the halo-model based HMCODE (Mead 2015; Mead et al. 2015) to accurately include the baryonic signatures in the non-linear matter power spectrum and by allowing for a possible luminosity and redshift dependence of the IA (in addition to the amplitude dependence). We also account for biases in the measured redshift distribution for each tomographic bin, both by considering random shifts around the fiducial distributions and by considering systematic shifts to the distributions following the analysis of source–lens cross-correlations in Choi et al. (2016). We consider these systematic uncertainties both independently and jointly and ask if the data favour any of the additional degrees of freedom. For the purposes of model selection, we use the deviance information criterion (DIC; Spiegelhalter, Best & Carlin 2002) and complement with calculations of the Bayesian evidence (e.g. Feroz & Hobson 2008; Trotta 2008). These statistical tools are discussed in Section 2.1.7.

We also strive to improve our understanding of the 'discordance' in the cosmological constraints from the cosmic shear and cosmic microwave background (CMB) data sets of CFHTLenS and Planck (e.g. Ade et al. 2014XVI; MacCrann et al. 2015; Grandis et al. 2016; Raveri 2016; Ade et al. 2016XIII). To achieve this, we take a methodical approach. We begin with the minimal scenario where no systematic uncertainties are included in the analysis of CFHTLenS and examine the potential dependence of the results to the choice of cosmological priors. We then consider a whole series of scenarios where the key systematic uncertainties are included independently and jointly, both with informative priors and with non-informative priors. We employ data concordance tests based on the Bayesian evidence and DIC and find that the level of discordance between the CFHTLenS and Planck data sets is sensitive to the assumptions made on the level of systematic uncertainties in the CFHTLenS measurements, such that increasingly conservative scenarios show an increasing degree of concordance between the data sets.

In addition to the comprehensive account of the systematic uncertainties, we update the CFHTLenS measurements first presented in the six-bin tomographic analysis of Heymans et al. (2013; hereafter H13). As described in Section 2.2, we divide the source galaxies into seven tomographic bins with redshift ranges that allow us to more optimally account for the overlap with spectroscopic surveys in forthcoming analyses. We moreover extend the angular coverage of the measurements from [1, 50] arcmin in H13 to [1, 120] arcmin in this work, owing to the increased box size of the new N-body simulations used to determine the data covariance matrix (Harnois-Déraps & van Waerbeke 2015; described in Section 2.3). Thus, instead of the original five angular bins, we now have seven angular bins in the aforementioned range.

In Section 2, we present the theoretical basis of our work, along with our updated CFHTLenS measurements and covariance matrix estimation from N-body simulations. In Section 3, we explore the impact of the systematic uncertainties on the cosmological constraints, independently and jointly. We examine whether the new degrees of freedom are favoured by the data and investigate the level of concordance between the CFHTLenS and Planck data sets. In Section 4, we conclude with a discussion of our results.

2 METHODOLOGY

We give an overview of the theory associated with weak gravitational lensing and intrinsic galaxy alignments. We discuss our new fitting pipeline and the methods by which we account for photometric redshift uncertainties and baryonic uncertainties in the non-linear matter power spectrum. We then proceed to describe our updated CFHTLenS measurements and covariance matrix considering seven tomographic bins. We do not include additional degrees of freedom for the additive and multiplicative shear calibration corrections, but incorporate these directly in our data.

2.1 Theory

2.1.1 *Weak lensing observables*

We follow the standard approach in computing the weak lensing observables (e.g. Bartelmann & Schneider 2001), in the form of the two-point shear correlation functions,

$$\xi_{\pm}^{ij}(\theta)_{\rm GG} = \frac{1}{2\pi}\int {\rm d}\ell\,\ell\,C_{\rm GG}^{ij}(\ell)\,J_{\pm}(\ell\theta), \tag{1}$$

defined at angle θ, where $C_{\rm GG}^{ij}(\ell)$ is the convergence power spectrum for tomographic bin combination $\{i, j\}$ at angular wavenumber ℓ and $J_{\pm}$ are the zeroth- $(+)$ and fourth- $(-)$ order Bessel functions of the first kind. Given seven tomographic bins, i and j both run from 1 to 7, such that there are 28 independent combinations. Using the Limber approximation (Limber 1954, also see Loverde & Afshordi 2008), the convergence power spectrum is then obtained as a weighted integral over the matter power spectrum,

$$C_{\rm GG}^{ij}(\ell) = \int_0^{\chi_{\rm H}} {\rm d}\chi\,\frac{q_i(\chi)q_j(\chi)}{[f_K(\chi)]^2}\,P_{\delta\delta}\left(\frac{\ell+1/2}{f_K(\chi)},\chi\right), \tag{2}$$

where χ is the comoving distance, $\chi_{\rm H}$ is the comoving horizon distance, $f_K(\chi)$ is the comoving angular diameter distance, $P_{\delta\delta}$ is the matter power spectrum, and the geometric weight $q_i(\chi)$ in tomographic bin i is given by

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

Here, $a(\chi)$ is the scalefactor, c is the speed of light, H_0 is the Hubble constant, $\Omega_{\rm m}$ is the present matter density, and $n_i(\chi)$ encodes the source galaxy distribution in a given tomographic bin, normalized to integrate to unity.

2.1.2 *Intrinsic galaxy alignments*

We further extend our theory to account for intrinsic galaxy alignments (Hirata & Seljak 2004; Bridle & King 2007; Joachimi et al. 2011), originating from correlations of intrinsic ellipticities of galaxies with each other and with the shear of background sources. These intrinsic–intrinsic (II) and shear-intrinsic (GI) terms enter the observed correlation function, such that

$$\xi_{\pm}^{ij}(\theta)_{\rm obs} = \xi_{\pm}^{ij}(\theta)_{\rm GG} + \xi_{\pm}^{ij}(\theta)_{\rm II} + \xi_{\pm}^{ij}(\theta)_{\rm GI}. \tag{4}$$

Here, the II and GI terms are defined as in equation (1), except $C_{\rm GG}^{ij}$ is correspondingly replaced by $C_{\rm II}^{ij}$ and $C_{\rm GI}^{ij}$. Following Bridle & King (2007) in using the non-linear matter power spectrum within the linear theory of Hirata & Seljak (2004), we express the II term:

$$C_{\rm II}^{ij}(\ell) = \int_0^{\chi_{\rm H}} {\rm d}\chi\,\frac{n_i(\chi)n_j(\chi)F_i(\chi)F_j(\chi)}{[f_K(\chi)]^2}\,P_{\delta\delta}\left(\frac{\ell+1/2}{f_K(\chi)},\chi\right), \tag{5}$$

and the GI term:

$$C_{\rm GI}^{ij}(\ell) = \int_0^{\chi_{\rm H}} {\rm d}\chi\,\frac{q_i(\chi)n_j(\chi)F_j(\chi)}{[f_K(\chi)]^2}\,P_{\delta\delta}\left(\frac{\ell+1/2}{f_K(\chi)},\chi\right) + \int_0^{\chi_{\rm H}} {\rm d}\chi\,\frac{n_i(\chi)F_i(\chi)q_j(\chi)}{[f_K(\chi)]^2}\,P_{\delta\delta}\left(\frac{\ell+1/2}{f_K(\chi)},\chi\right). \tag{6}$$

We allow for an unknown amplitude A along with a possible redshift (z) and luminosity (L) dependence via η and β, respectively, in defining

$$F_i(\chi) = -AC_1\rho_{\rm cr}\frac{\Omega_{\rm m}}{D(\chi)}\left(\frac{1+z(\chi)}{1+z_0}\right)^{\eta}\left(\frac{L_i}{L_0}\right)^{\beta}, \tag{7}$$

in accordance with Joachimi et al. (2011), where $\rho_{\rm cr}$ is the critical density at present, $D(\chi)$ is the linear growth factor normalized to unity at present, the normalization constant $C_1 = 5\times10^{-14}\,h^{-2}M_{\odot}^{-1}{\rm Mpc}^3$, $z_0 = 0.3$ is an arbitrary pivot redshift, and L_0 is the pivot luminosity corresponding to an absolute r-band magnitude of -22. We determine the luminosities by averaging the individual galaxy luminosities (calculated as $10^{-0.4M}$ for each galaxy, where M is the absolute magnitude), weighted by the *lens*fit weights (defined in Section 2.2), giving us an effective $L_i/L_0 =$ (0.017, 0.069, 0.15, 0.22, 0.36, 0.49, 0.77) for our seven tomographic bins.

2.1.3 *Photometric redshift uncertainties*

We account for uncertainties in the photometric redshift estimation by allowing the redshift distribution in each tomographic bin to shift along the redshift axis by an amount Δz_i, such that

$$n_i^{\rm theory}(z) = n_i^{\rm obs}(z-\Delta z_i), \tag{8}$$

where $n_i^{\rm obs}$ is the observed redshift distribution. This is consistent with the approach used in Abbott et al. (2016). As a minor caveat, as we do not integrate below the minimum redshift of $z_{\rm min} = 0.03$ for the fiducial redshift distributions (the necessity of $z_{\rm min} > 0$ is because $z = 0$ would correspond to wavenumber $k = \infty$), we take a consistent approach and continue to neglect the same lowest end of the redshift distributions when shifted to higher redshifts. A physical interpretation of this would be to consider these outliers as stars at $z = 0$. Since we consider 7 tomographic bins, this introduces an additional seven nuisance parameters that we marginalize over in our analysis, with either uniform or Gaussian priors (e.g. via cross-correlations with an overlapping spectroscopic sample), as further discussed in Section 3.4.

2.1.4 *Baryonic uncertainties in the non-linear matter power spectrum:* HMCODE

We account for baryonic uncertainties in the non-linear matter power spectrum by incorporating HMCODE in Mead et al. (2015)[1] as a separate parallelized module in CosmoMC (Lewis & Bridle 2002[2]). By introducing physically motivated free parameters in the halo-model formalism, and calibrating these to the Coyote N-body DM simulations (Heitmann et al. 2014 and references therein), HMCODE is able to describe the power spectrum to marginally improved accuracy in comparison to the latest incarnation of HALOFIT (Takahashi et al. 2012; Smith et al. 2003).

---

[1] https://github.com/alexander-mead/hmcode

[2] http://cosmologist.info/cosmomc/

However, by further calibrating to the OverWhelmingly Large (OWL) Simulations (Schaye et al. 2010; van Daalen et al. 2011), the main benefit of HMCODE is its capacity to account for baryonic effects in the matter power spectrum on non-linear scales, for example, due to star formation, radiative cooling, and active galactic nucleus (AGN) feedback. This is achieved by modifying parameters that govern the internal structure of haloes. For example, AGN feedback blows gas out of haloes, which makes them less concentrated. Mead et al. (2015) found that acceptable fits could be made to the OWL simulations using a halo model with less concentrated haloes in the one-halo term. Conversely, the two-halo term is unchanged because feedback only affects small scales. Thus, HMCODE modifies the relationship between halo concentration c and halo mass M, such that

$$c(M, z) = B\frac{1+z_{\rm f}}{1+z}, \tag{9}$$

where $z_{\rm f}$ is the halo formation redshift as a function of halo mass and B is a free parameter that we can marginalize over in our analysis. HMCODE can also change the halo density profile via the halo bloating parameter $\eta_{\rm HMCODE}$ to account for baryonic effects. However, in Mead et al. (2015) it was shown that substantial degeneracy exists between $\eta_{\rm HMCODE}$ and B and that these two parameters can be linearly related to provide a one-parameter baryonic feedback model. We use this prescription and $\eta_{\rm HMCODE}$ therefore does not contribute to an additional degree of freedom in our Markov Chain Monte Carlo (MCMC) analysis.

For scales $k < 10\,h^{-1}$Mpc, HMCODE produces a non-linear matter power spectrum that accounts for baryonic physics (REF, DBLIM, and AGN cases of the OWL simulations, described in van Daalen et al. 2011) accurately at the level of a few per cent. HMCODE's ability to accurately model the non-linear matter power spectrum including baryons with a single parameter can be contrasted with the fitting formula in Harnois-Déraps et al. (2015), which employs 15 free parameters to achieve the same outcome with similar precision (also see Köhlinger et al. 2016, which use the same prescription as in Harnois-Déraps et al. 2015 but only marginalize over a single parameter and MacCrann et al. 2015 for a similar single-parameter marginalization approach based on the AGN case of the OWL simulations).

While HMCODE is calibrated to $k < 10\,h^{-1}$Mpc, it agrees with the matter power spectrum from Takahashi et al. (2012) at the 10 per cent level for $k < 100\,h^{-1}$Mpc (DM-only case). Even if we assume that HMCODE miscalibrates the matter power spectrum (including baryons) for $k > 10\,h^{-1}$Mpc by a factor of 2 for every decade in wavenumber beyond $k = 10\,h^{-1}$Mpc, it would bias the lensing correlations functions by at most 1 per cent for ξ_+ and at the sub-per cent level for ξ_- for the angular scales considered in this work (less for ξ_- than for ξ_+ due to our angular cuts; checked for the case without tomography). Given the statistical power of our data (described in Section 2.2), the accuracy of HMCODE is therefore sufficient for our purposes and it forms an important component of our new pipeline.

2.1.5 New CosmoMC module for WL analyses with systematic uncertainties

In order to account for the systematic uncertainties coming from IA of galaxies, photometric redshift uncertainties, and baryonic effects in the non-linear matter power spectrum, we have developed a new module in CosmoMC (in the language of FORTRAN 90). The module is independent from previous lensing modules and accounts for the systematic uncertainties following the prescriptions in Sections 2.1.2–2.1.4.

The new module allows the user to choose the integration method for the {GG, II, GI} spectra with one of two distinct methods, either with trapezoidal integration or with Romberg integration. We have internally parallelized the code which, with a single eight-core Intel Xeon E5-2660 processor at 2.2 GHz, can calculate the likelihood for a single cosmology, considering six tomographic bins and including IA (i.e. GG, II, and GI), in 0.078 s when using HALOFIT for the non-linear matter power spectrum. Since our module is parallelized the speed would continue to show some improvement with further cores. This can be compared to the existing default lensing module in CosmoMC which, with the same resources, calculates the likelihood for a single cosmology, considering six tomographic bins and without IA (i.e. only GG) in 0.33 s.[3] We note that these are the speeds of only the respective modules, i.e. the numbers do not account for the time it takes the Boltzmann code CAMB (Lewis, Challinor & Lasenby 2000[4]) to compute the matter power spectrum which is fed into both modules.

As we have incorporated HMCODE as a separate parallelized module in CosmoMC, at each new cosmology, the lensing module internally provides HMCODE the linear matter power spectrum obtained from a modified version of CAMB in a (k, z)-array and obtains from it the non-linear matter power spectrum at the same (k, z) values in return. Using the same processor, this transition between linear to non-linear power spectrum takes 0.4 s for a (k, z)-array that is sufficiently dense for our lensing calculation. While the computation of the non-linear matter power spectrum with HMCODE is slower than the computation with HALOFIT, it allows us to account for the baryonic effects on non-linear scales. However, when non-linear baryonic effects are not considered, the agreement between HMCODE and HALOFIT is sufficiently close that either one could be used.

In Fig. 1, we show the impact of the different systematic degrees of freedom on $\{\xi_+, \xi_-\}$ for the tomographic bin combinations {1,7} and {7,7} . As expected, we find that the difference between the HALOFIT and HMCODE prescriptions enters the observables on smaller angular scales. Moreover, varying the HMCODE feedback amplitude leaves an imprint on the observables that increases with smaller angular scales. Analogously, the imprint of the different non-linear prescriptions (both with and without baryons) is larger for ξ_- than for ξ_+, due to the greater sensitivity of ξ_- to non-linear scales in the matter power spectrum for a given angular scale. Meanwhile, the impact of IA and shifts in the photometric redshift distributions seems to be strongest in the cross-bins and fairly independent of angular scale. This illustrates the usefulness of these bins in constraining the IA model and deviations from the fiducial photometric redshift distributions.

We have further extended our module to account for joint analyses of overlapping observations of cosmic shear, galaxy–galaxy lensing,

---

[3] In the development of this module, we verified our results by comparing against a completely independent implementation available to the collaboration (Joachimi) and to the default CosmoMC lensing module. The lensing observables calculated with our new module agree well with those calculated with the collaboration's independent code. There were however discrepancies in the convergence power spectrum with the default CosmoMC lensing module (which seem to be caused by insufficiently accurate integration in the default CosmoMC lensing module; here we only checked GG as the default CosmoMC lensing module does not account for IA). We find the discrepancies to be negligible at the level of the parameter constraints due to the sufficiently weak statistical power of current data.

[4] http://camb.info

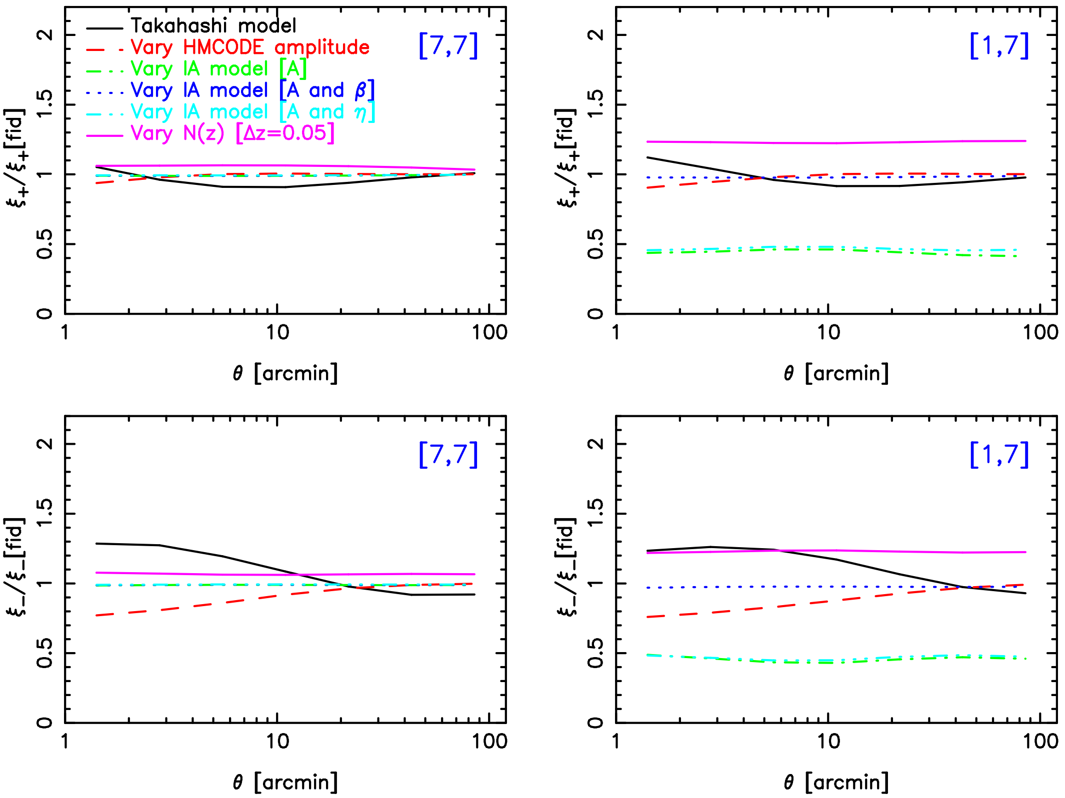

Figure 1. The ratio of shear correlation functions for tomographic bin combinations $\{1{,}7\}$ and $\{7{,}7\}$, taken with respect to HMCODE with feedback amplitude $\log B = 0.496$, defined in equation (9), including no systematic uncertainties (denoted as $\xi_{\pm}[\mathrm{fid}]$). For consistency, we fix the underlying cosmology to that of the best-fitting cosmology of this 'fiducial' case. We allow for the Takahashi et al. (2012) version of HALOFIT (solid black), HMCODE with $\log B = 0.3$ (dashed red), IA with $\{A,\, \eta,\, \beta\} = \{1,\, 0,\, 0\}$ (dot–dashed green), IA with $\{A,\, \eta,\, \beta\} = \{1,\, 0,\, 1\}$ (dotted blue), IA with $\{A,\, \eta,\, \beta\} = \{1,\, 1,\, 0\}$ (dot–dashed cyan), and photometric redshift uncertainties where all bins are positively perturbed by $\Delta z = 0.05$ (solid pink). The $\log B$ value of 0.496 corresponds to the DM-only case, while $\log B = 0.3$ agrees with the AGN case of the OWL simulations. The parameters $\{A,\, \eta,\, \beta\}$ refer to the IA amplitude, redshift dependence, and luminosity dependence, respectively, of the IA model defined in equation (7), while the photo-z shifts are defined in equation (8). The IA model with $\{A,\, \eta,\, \beta\} = \{1,\, 0,\, 1\}$ lies along the unity line because the luminosity $L/L_0 < 1$ in each tomographic bin, such that a positive value of β suppresses the IA signal (analogously, a negative value of η for the redshift dependence would have a similar effect).

and large-scale structure measured through clustering multipoles, including the full covariance, as part of our efforts to constrain modified gravity and neutrino physics. Along with the new CFHTLenS measurements, we are releasing our code (both cosmic shear and HMCODE modules in CosmoMC) pertaining to the calculations presented in this paper at https://github.com/sjoudaki/cfhtlens_revisited. We will be releasing our full code as part of an upcoming paper (Joudaki et al., in preparation).

Lastly, we note that our module is currently independent from the CosmoSIS platform (Zuntz et al. 2015), which combines a range of disparate codes into a single framework for cosmological parameter estimation. There are no technical obstacles to prevent our module from being incorporated into CosmoSIS in the future.

2.1.6 Baseline configurations

In our analysis, we always include the 'vanilla' parameters, given by $\{\Omega_{\mathrm{c}}h^2,\, \Omega_{\mathrm{b}}h^2,\, \theta_{\mathrm{MC}},\, n_{\mathrm{s}},\, \ln{(10^{10}A_{\mathrm{s}})}\}$, which represent the CDM density, baryon density, approximation to the angular size of the sound horizon (in CosmoMC), scalar spectral index, and amplitude of the scalar spectrum, respectively. We note that 'ln' refers to the natural logarithm, while we take 'log' to refer to logarithms with base 10. From these parameters, one can derive the Hubble constant H_0 (also expressed as h in its dimensionless form) and standard deviation of the present linear matter density field on scales of $8\,h^{-1}$Mpc (denoted by σ_8). We impose uniform priors on these cosmological parameters, as discussed in Section 3.

In this baseline ΛCDM model, we include three massless neutrinos, such that the effective number of neutrinos $N_{\mathrm{eff}} = 3.046$ (we have checked that our results are not significantly affected by the approximation of zero mass, as compared to the minimal mass of the normal hierarchy of 0.06 eV). For the primordial fraction of baryonic mass in helium, Y_{p}, we determine the quantity as a function of $\{N_{\mathrm{eff}},\, \Omega_{\mathrm{b}}h^2\}$ in a manner consistent with big bang nucleosynthesis (BBN; see equation 1 in Joudaki 2013). Moreover, we consistently enforce the strong inflation prior on the curvature and running of the spectral index, such that $\{\Omega_k \equiv 0,\, \mathrm{d}n_{\mathrm{s}}/\mathrm{d}\ln k \equiv 0\}$. Thus, with flatness enforced, $\Omega_{\mathrm{m}} > 1$ implies $\Omega_\Lambda < 0$. Lastly, with no running of the spectral index, we define the primordial scalar power spectrum,

$$\ln P_{\mathrm{s}}(k) = \ln A_s + (n_s - 1)\ln(k/k_{\mathrm{pivot}}), \tag{10}$$

where both A_{s} and n_{s} are defined at the pivot wavenumber k_{pivot}.

In order to determine the convergence of our MCMC chains, we use the Gelman & Rubin (1992) R statistic, where R is defined as the variance of chain means divided by the mean of chain variances. Our runs are stopped when the conservative limit

$(R-1) < 2 \times 10^{-2}$ is reached and we have checked that further exploration of the tails does not change our results.

2.1.7 Model selection and data set concordance

We define the best-fitting effective χ^2, via $\chi^2_{\rm eff}(\hat{\theta}) = -2\ln\mathcal{L}_{\rm max}$, where $\mathcal{L}_{\rm max}$ is the maximum likelihood of the data given the model, θ, is the vector of varied parameters and hat denotes the maximum likelihood point. When quoting $\chi^2_{\rm eff}$ without specifying θ, we implicitly assume $\theta = \hat{\theta}$. The reduced χ^2 is then given by $\chi^2_{\rm red} = \chi^2_{\rm eff}/\nu$, where ν is the number of degrees of freedom. Given two separate models, where $\Delta\chi^2_{\rm eff} > 0$, we interpret the model with the higher value of $\chi^2_{\rm eff}$ to be associated with a lower probability of drawing the data at the maximum likelihood point, by a factor given by $\exp(-\Delta\chi^2_{\rm eff}/2)$. For reference, a difference of 10 in $\chi^2_{\rm eff}$ between two models would correspond to a probability ratio of 1 in 148 and therefore constitute strong preference for the more probable model.

When considering the relative performance of two distinct models, it is valuable to compute the DIC (Spiegelhalter et al. 2002, also see Kunz, Trotta & Parkinson 2006; Liddle 2007; Trotta 2008; Spiegelhalter et al. 2014), obtained from the Kullback–Leibler divergence or relative information entropy (Kullback & Leibler 1951). We do not use the Akaike Information Criterion (Akaike 1974), which follows from an approximate minimization of the Kullback–Leibler divergence and does not account for unconstrained directions in parameter space (e.g. Liddle 2007; Trotta 2008). We also do not consider the Bayesian Information Criterion (Schwarz 1978), as it is not grounded in information theory and instead follows from a Gaussian approximation to the Bayesian evidence (e.g. Liddle 2007; Trotta 2008). In practice, we compute

$$\mathrm{DIC} \equiv \chi^2_{\rm eff}(\hat{\theta}) + 2p_D, \tag{11}$$

where $p_D = \overline{\chi^2_{\rm eff}(\theta)} - \chi^2_{\rm eff}(\hat{\theta})$ is the 'Bayesian complexity' and the bar denotes the mean taken over the posterior distribution (Spiegelhalter et al. 2002). The Bayesian complexity is a measure of the effective number of parameters and acts as a penalty against more complex models. Instead of the maximum likelihood point, the Bayesian complexity and DIC are also commonly evaluated at the posterior mean or median of the cosmological parameters (e.g. Spiegelhalter et al. 2002; Trotta 2008). The main limitation of the DIC is the $\chi^2_{\rm eff}$ of the posterior mean can sometimes not be a good fit to the data in multimodal distributions and the alternative (the best-fitting $\chi^2_{\rm eff}$) is a somewhat arbitrary choice. Moreover, the DIC uses the data effectively twice (in that the 'penalty factor' also depends on the data) and its use of a point estimate can be stochastically affected by the data. Finally, beyond brute force, no efficient and accurate method has been developed for computing the errors of the DIC estimates.

For two models with the same complexity, the difference in their DIC values is the same as the difference in their respective $\chi^2_{\rm eff}$ values. Analogous to the $\chi^2_{\rm eff}$ scenario, a difference of 10 in DIC between two models constitutes strong preference in favour of the model with the lower DIC estimate, while a difference of 5 in DIC between the two models constitutes moderate preference in favour of the model with the lower DIC estimate. When the difference in DIC between two models is even smaller, the statistic only weakly favours one model over the other. In comparing an extended model with a reference model, we take negative values of ΔDIC to indicate a preference in favour of the extended model as compared to the reference model.

We complement our DIC analysis by using the nested sampling algorithm CosmoChord (Handley, Hobson & Lasenby 2015a,b) to compute the Bayesian evidence (with additional runs using MultiNest to ensure consistency in the results; Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2013). The evidence is given by the average of the likelihood under the prior for a given model,

$$\mathcal{Z} = \int \mathrm{d}^n\boldsymbol{\theta}\, \mathcal{L}(\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \tag{12}$$

where n encapsulates the dimensionality of the parameter space and $\pi(\boldsymbol{\theta})$ is the prior given the vector of parameters $\boldsymbol{\theta}$ (e.g. Feroz & Hobson 2008; Trotta 2008). Equation (12) tells us that the evidence is larger for a simpler theory with a compact parameter space, unless it is significantly worse at explaining the data as compared to a more complicated theory. For model selection purposes, we also compute the Bayes factor (e.g. Feroz & Hobson 2008; Trotta 2008), given by the evidence ratio for two specific models, denoted by $\mathcal{Z}_0$ and $\mathcal{Z}_1$:

$$\mathcal{B}_{01} \equiv \mathcal{Z}_0/\mathcal{Z}_1. \tag{13}$$

For a scenario in which the prior probabilities of the two models are equal, the Bayes factor encapsulates the posterior odds, such that the data favour model 0 as compared to model 1 when the Bayes factor is greater than unity and vice versa. Alternatively, the Bayes factor can be thought of as the change to the prior odds given the data.

From the evidence calculations, we can further construct a measure of the concordance between two data sets D_1 and D_2, given by

$$\mathcal{C}(D_1, D_2) \equiv \frac{\mathcal{Z}(D_1 \cup D_2)}{\mathcal{Z}(D_1)\mathcal{Z}(D_2)}, \tag{14}$$

where $\mathcal{Z}(D_1 \cup D_2)$ is the joint evidence of the two data sets (Marshall, Rajguru & Slosar 2006; Raveri 2016). Thus, $\log\mathcal{C}$ is positive when there is concordance between the two data sets, such that the joint evidence is larger than the product of the individual evidences and similarly $\log\mathcal{C}$ is negative when there is discordance between the two data sets. We will use this concordance test to better assess the potential degree of tension between the updated CFHTLenS and Planck measurements. In this pursuit, we also introduce an analogous but more easily calculable quantity from the DIC estimates:

$$\mathcal{I}(D_1, D_2) \equiv \exp\{-\mathcal{G}(D_1, D_2)/2\}, \tag{15}$$

where

$$\mathcal{G}(D_1, D_2) = \mathrm{DIC}(D_1 \cup D_2) - \mathrm{DIC}(D_1) - \mathrm{DIC}(D_2), \tag{16}$$

and $\mathrm{DIC}(D_1 \cup D_2)$ is the joint DIC of the two data sets. We expect this quantity to diagnose separation or congruence between posterior distributions, through measurement of the relative entropy of one distribution with respect to the other. To describe it in terms of a Gaussian example, if two data sets that agree are added together, we would expect the joint likelihood to have a larger $\chi^2_{\rm eff}(\hat{\theta})$ (since there are more data points), roughly equivalent to the sum of the individual $\chi^2_{\rm eff}$, but the same Bayesian complexity for both, leading to an overall negative value for $\mathcal{G}(D1, D2)$ (since the complexity factor is applied twice) and so to a large $\mathcal{I}$. However, if the two data sets do not agree, there will be a much larger $\chi^2_{\rm eff}$ than the sum of the individual $\chi^2_{\rm eff}$ and this will not be balanced enough by the change in complexity, since the different parameters will not be measured well. In this case, the overall value for $\mathcal{G}(D1, D2)$ will be positive, leading to a small $\mathcal{I}$.

Thus, analogous to the evidence scenario, $\log \mathcal{I}$ is constructed such that there is concordance between the data sets when it is positive and discordance between the data sets when it is negative. For an independently developed concordance test based on the Kullback–Leibler divergence, see Seehars et al. (2016). We further assess the degree of concordance or discordance by employing Jeffreys' scale (Jeffreys 1961, also see Kass & Raftery 1995), such that values for $\log \mathcal{C}$ and $\log \mathcal{I}$ in excess of $\pm 1/2$ are 'substantial', values in excess of ± 1 are 'strong', and values in excess of ± 2 are 'decisive' (where this last case corresponds to a probability ratio in excess of 100).

2.2 Measurements

In this section, we introduce the CFHTLenS data set and new measurements used in our cosmology analysis. The CFHTLenS[5] is a deep multicolour survey optimized for weak lensing analyses, based on data from the Canada–France–Hawaii Telescope (CFHT) Legacy Survey in five optical bands $u^*g'r'i'z'$, using the 1 deg^2 camera MegaCam. The Wide Survey data analysed in this study span four fields W1, W2, W3, and W4, which together cover 154 deg^2.

Galaxy ellipticity components (e_1, e_2) for each source, together with an approximately optimal inverse-variance weight w^s, are determined by the Bayesian model-fitting software *lens*fit (Miller et al. 2013). Photometric redshifts are derived from PSF-matched photometry (Hildebrandt et al. 2012) using the Bayesian photometric redshift code BPZ (Benítez 2000), which also returns a full redshift probability distribution $p_{\rm BPZ}(z)$, with peak $z_{\rm B}$ for each source. The survey pointings have been subjected to a stringent cosmology-independent systematic-error analysis (Heymans et al. 2012), as a result of which a subset of around 25 per cent of the pointings have been flagged as possessing potentially significant systematic errors and are excluded from our analysis. We applied additive shear calibration corrections to the measured ellipticities and multiplicative shear calibration corrections to the cosmic shear measurements (following Heymans et al. 2012 and Miller et al. 2013). We only retain unmasked sources for our analysis.

We perform cosmic shear tomography by dividing the sources according to $z_{\rm B}$ into $N_{\rm t} = 7$ tomographic bins with ranges 0.15–0.29, 0.29–0.43, 0.43–0.57, 0.57–0.70, 0.70–0.90, 0.90–1.10, 1.10–1.30. This choice represents a slight alteration from the original CFHTLenS tomographic analysis (Heymans et al. 2013), which divided the range $0.2 < z_{\rm B} < 1.3$ into six tomographic bins. This modification was motivated by the overlapping spectroscopic data sets now available due to the Baryon Oscillation Spectroscopic Survey (BOSS; Anderson et al. 2014), which are conveniently split into redshift ranges 0.15–0.43 (LOWZ sample) and 0.43–0.70 (CMASS sample). This spectroscopic redshift data may be used to include galaxy–galaxy lensing and redshift-space distortion statistics in the analysis (with appropriate covariance) and further calibrate the photometric redshifts through cross-correlation (e.g. Choi et al. 2016). Fig. 2 displays the stacked BPZ redshift probability distributions, weighted by the *lens*fit weights, for each tomographic source bin. A spline function of these measurements is used as the model source redshift distribution in our cosmology-fitting pipeline.

The effective source density for lensing analyses is defined by

$$n_{\rm eff} = \frac{1}{A_{\rm eff}} \frac{\left(\sum_i w_i^{\rm s}\right)^2}{\sum_i (w_i^{\rm s})^2}, \qquad (17)$$

where $A_{\rm eff}$ is the effective (unmasked) area. In the range $0.15 < z_{\rm B} < 1.3$ used in our study, the values derived for the four survey regions {W1, W2, W3, W4} are $n_{\rm eff} = \{10.9, 9.9, 11.0, 10.4\}$ arcmin^{-2} for unmasked areas {42.9, 12.1, 26.1, 13.3} deg^2; for the whole sample, we find $n_{\rm eff} = 10.7$ arcmin^{-2} over $A_{\rm eff} = 94$ deg^2.

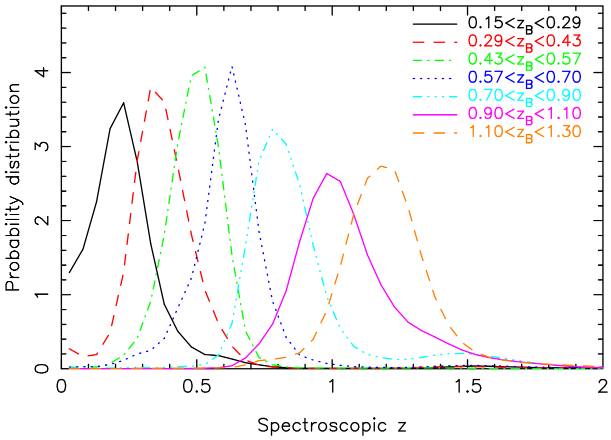

Figure 2. Stacked BPZ redshift probability distributions for CFHTLenS, weighted by the *lens*fit weights, in the seven tomographic photo-z bins used in our analysis.

For each unique pair of tomographic bins, we measured the cosmic shear statistics (ξ_+, ξ_-) in each of the four regions using the ATHENA software (Kilbinger, Bonnett & Coupon 2014). We use $N_\theta = 7$ equally spaced logarithmic bins in the range $1 < \theta < 120$ arcmin. Concretely, for each of the seven tomographic bins, our measurements are evaluated at [1.41, 2.79, 5.53, 11.0, 21.7, 43.0, 85.2] arcmin. This represents a significant increase in the maximum fitted scale of ≈53 arcmin used by H13 (where the earlier reported range of 1.5–35 arcmin corresponded to the central bin values), which is enabled by the increased box size of the N-body simulations now used to determine the data covariance, described below. We also determined jackknife errors in our measurements, splitting the data sample into jackknife regions defined by each individual MegaCam field. We combined the measurements in the different CFHTLenS regions, weighting by the $N_{\rm pairs}$ value for each bin returned by ATHENA. Fig. 3 displays the resulting combined (ξ_+, ξ_-) measurements in panels of pairs of tomographic bins.

We arranged the (ξ_+, ξ_-) measurements into a data vector following the convention of H13, such that for each unique pair of tomographic bins, the ξ_+ values are listed with increasing θ, followed by the ξ_- values. The pairs of tomographic bins ij are then ordered as (11, 12, ..., 17, 22, 23, ..., 77). The length of the full data vector is $p = N_\theta N_{\rm t}(N_{\rm t} + 1) = 392$ elements, although this is further pruned before cosmological fitting. Concretely, we cut our data vector from $p = 392$ elements to $p = 280$ elements by removing angular bins 1–3 for ξ_- and the seventh angular bin for ξ_+. This is motivated by low signal to noise of the removed elements (bins 1–3 for ξ_-), along with roughly 10 per cent covariance underestimation for ξ_+ in the seventh bin due to the finite box size for the simulations (Harnois-Déraps & van Waerbeke 2015).

Lastly, when comparing cosmological constraints from our updated CFHTLenS cosmic shear tomography measurements with CMB measurements from the Planck satelite (Ade et al. 2016XIII; Aghanim et al. 2016a), we include both CMB temperature and polarization information for Planck on large angular scales, limited to multipoles $\ell \leq 29$ (i.e. low-ℓ TEB likelihood) and restrict ourselves to CMB temperature information on smaller angular scales

---

[5] http://www.cfhtlens.org

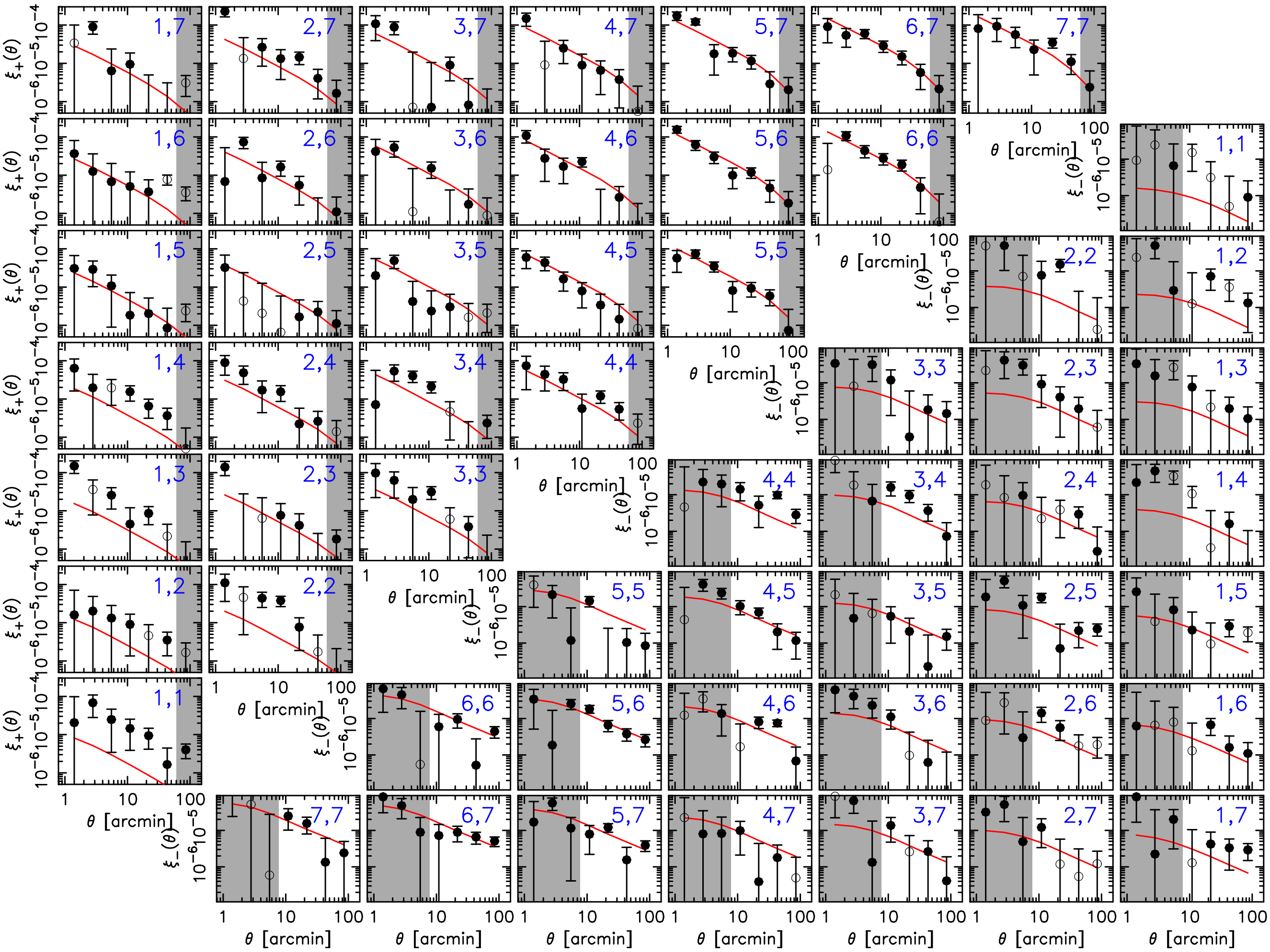

Figure 3. Measurements of the cosmic shear statistics ξ_+ (upper triangle) and ξ_- (lower triangle) against angular scale in arcminutes for all unique pairs of the seven tomographic source bins, defined in Section 2.2. The error bars are determined using the mock catalogues described in Section 2.3. The grey regions correspond to angular scales that were removed from the cosmology analysis, due to low signal-to-noise or covariance underestimation (discussed in Section 2.2). Open circles denote negative points. Fiducial theory lines have been included in red (solid) for comparison.

(via the PLIK TT likelihood). Thus, we conservatively do not include polarization data for the smaller angular scales and we also do not include Planck CMB lensing measurements.

2.3 Covariance

We determined the covariance of our (ξ_+, ξ_-) measurements using a set of mock catalogues created from a large suite of N-body simulations which include a self-consistent computation of gravitational lensing. Our covariance methodology follows the approach of H13, with some enhancements described below.

Our starting point is the Scinet LIght Cone Simulations series (Harnois-Déraps & van Waerbeke 2015), which consists of 500 N-body DM simulations created with the CUBEP^{3}M code (Harnois-Déraps et al. 2013) using a *WMAP9*+BAO+SN cosmological parameter set: matter density $\Omega_{\mathrm{m}} = 0.2905$, baryon density $\Omega_{\mathrm{b}} = 0.0473$, Hubble parameter $h = 0.6898$, spectral index $n_{\mathrm{s}} = 0.969$, and normalization $\sigma_8 = 0.826$. Although the simulations are evaluated at a fixed cosmology, we assume that the cosmology dependence of the resulting covariance matrix has negligible impact on our cosmological parameter constraints following the explicit demonstration of this for a CFHTLenS-like survey in Kilbinger et al. (2013) (also see Eifler, Schneider & Hartlap 2009). The box size of the simulations is $L = 505\, h^{-1}$ Mpc. This is significantly larger than the simulation set used for modelling the earlier CFHTLenS measurements [$L = (147, 231)\, h^{-1}$ Mpc], significantly reducing the suppression of the large-scale signal and variance caused by the finite box size. The simulations follow the non-linear evolution of 1536^3 particles inside a 3072^3 grid cube.

For each simulation, the density field is output at 18 redshift snapshots in the range $0 < z < 3$. The gravitational lensing shear and convergence are computed at these multiple lens planes using the Born approximation in the flat-sky approximation and a survey cone spanning 60 deg^2 is constructed by pasting together these snapshots. A DM halo finder is also applied to the particle data at each snapshot, such that self-consistent halo catalogues for each cone are also produced.

Removing some rare cases of failed simulation outputs (e.g. due to cooling failure or internode message passing failure from traffic jam in the network), we use 497 independent simulations in our analysis. We convert these simulation density and shear fields into mock catalogues for our cosmic shear covariance using the following process. We note that although sources in the data set have optimal *lens*fit weights used in cosmic shear analysis, we produce mocks in which all sources have uniform weight $w^s = 1$, by building the mocks using the weighted source densities, weighted redshift distributions and weighted ellipticity variance measured from the data.

(i) In order to gain a sufficient number of realizations that our inverse covariance will not be unduly biased by noise in the sample covariance estimator (Hartlap, Simon & Schneider 2007), we split each 60 deg^2 simulation box into 2 × 2 sub-divisions, producing 1988 'pseudo-independent' sub-realizations (following H13, who used 3 × 3 sub-divisions of a 12.84 deg^2 box; each of our sub-samples is therefore an order of magnitude larger).

(ii) We assigned a source redshift distribution to each survey cone with a weighted effective surface density of 10.7 arcmin^{-2} using the weighted CFHTLenS source redshift probability distribution (as described above), Monte Carlo sampling sources from the density field with bias $b_{\rm source} = 1$. We ensured that sources are produced with a continuous distribution in redshift by linearly interpolating the shear across the finite redshift width of each snapshot.

(iii) We assigned the two-component gravitational shears (γ_1, γ_2) to each source by linearly interpolating the shear fields between the values at adjacent snapshot redshifts at the source positions.

(iv) A photometric redshift $z_{\rm B}$ was assigned for each mock source using a scattering probability function $p(z_{\rm B}|z_{\rm sim})$ as a function of its simulation redshift $z_{\rm sim}$. This scattering function was constructed from the CFHTLenS data set using a Monte Carlo technique sampling from the full `BPZ` probability distribution of each source, $p_{\rm BPZ}(z)$, together with its measured $z_{\rm B}$ value. Specifically, we sampled a redshift $z_{\rm samp}$ from each $p_{\rm BPZ}(z)$ distribution and then binned the values of $(z_{\rm B}, z_{\rm samp})$, weighting each galaxy by the *lens*fit weight and determining the distribution over $z_{\rm B}$ for each $z_{\rm samp}$ bin, normalizing such that $\int p(z_{\rm B})\,{\rm d}z_{\rm B} = 1$.

(v) We applied shape noise to the source catalogues by determining the complex noisy shear $e = (\gamma + n)/(1 + n\,\gamma^*)$ (Seitz & Schneider 1997), where the components of observed shear (e_1, e_2) are found as $e = e_1 + i\,e_2$, the true shear $\gamma = \gamma_1 + i\,\gamma_2$, and the noise $n = n_1 + i\,n_2$. The noise components (n_1, n_2) are drawn from Gaussian distributions with standard deviation σ_e, which we calibrated as a function of $z_{\rm B}$ using the weighted ellipticity variance of the real data:

$$\sigma_e^2 = \sum_i (w_i^s)^2\, e_i^2 \Big/ \sum_i (w_i^s)^2. \tag{18}$$

We find that σ_e as a function of $z_{\rm B}$ ranges between 0.26 and 0.29 with a mean of 0.28.

(vi) We applied small-scale masks to each sub-realization using the 'mosaic masks' provided by the CFHTLenS team. Given that these masks extend beyond the 15 deg^2 area of each sub-realization and that we require each sub-realization to possess identical masking to avoid introducing spurious noise, we consistently applied the same 15 deg^2 cut-out from the mask to every sub-realization. However, given that the fraction of unmasked area varies between the survey regions (owing, for example, to the varying stellar density with Galactic latitude), we repeated this process using mosaic masks for each of the four survey regions {W1, W2, W3, W4} and derived the final covariance as the area-weighted average of the four determinations.

We hence produced $n_\mu = 1988$ pseudo-independent mock CFHTLenS shear catalogues, matching the effective source density, underlying spectroscopic redshift distribution, photo-z scatters, $z_{\rm B}$-dependent shape noise, and small-scale masking to the real data set.

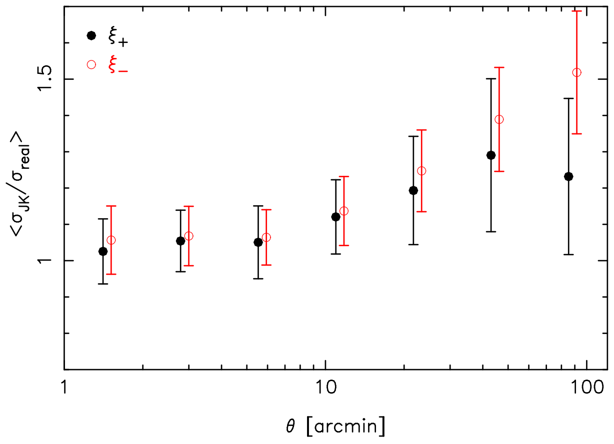

Figure 4. Ratio of the cosmic shear error determined by jackknife sampling using ATHENA to that determined from the suite of mock catalogues, averaged across all tomographic bins as a function of angular scale for ξ_+ (black solid circles) and ξ_- (red open circles).

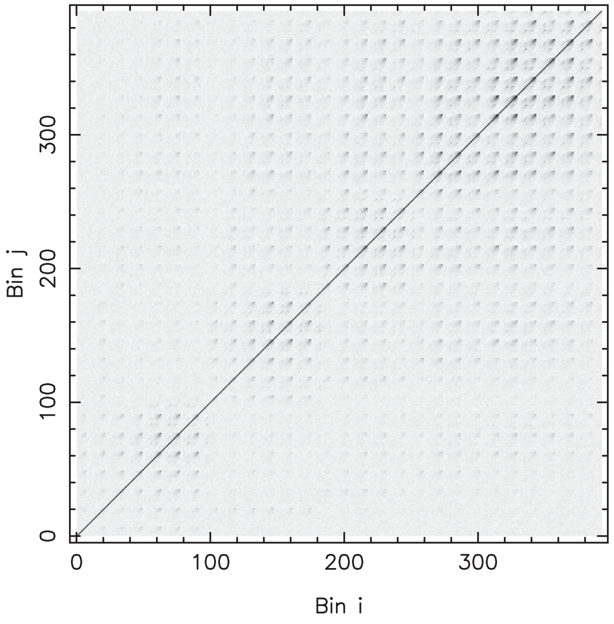

Figure 5. The correlation coefficient of the covariance matrix of the full data vector, plotted using a grey scale where white represents $r = 0$ and black represents $r = 1$.

We divided each mock catalogue by photometric redshift $z_{\rm B}$ into seven tomographic redshift bins and used ATHENA to measure the cosmic shear statistics (ξ_+, ξ_-) for the same angular bins as defined in Section 2.2, which we arranged in a data vector $\boldsymbol{D}$ (writing the measurement of bin i in mock k as D_{ki}). We then derived the data covariance through 'area-scaling' as

$$\mathrm{Cov}(i, j) = \frac{A_{\rm mock}}{A_{\rm eff}\,(n_\mu - 1)} \sum_{k=1}^{n_\mu} \left(D_{ki} - \overline{D_i}\right)\left(D_{kj} - \overline{D_j}\right), \tag{19}$$

where $\overline{D_i} = \sum_{k=1}^{n_\mu} D_{ki}$, $A_{\rm eff}$ is the unmasked area of the data determined above,and $A_{\rm mock}$ is the unmasked area of the sub-realizations. For mocks that include separations up to a few degrees the error induced by area-scaling, as evaluated using equation 36 in Friedrich et al. (2016), is small compared to other factors.

Table 1. Exploring the impact of cosmological priors (applicable to Section 3.1.1 and Fig. 6). The four cases include the same uniform priors on $\{\Omega_c h^2, \Omega_b h^2, \theta_{\rm MC}\}$ and differ in the priors on $\{A_s, n_s, h, k_{\rm pivot}\}$. Concretely, cases I and II have wider priors on $\{A_s, n_s, h\}$ than cases III and IV. The choice of pivot scale further distinguishes case I from case II and case III from case IV (Planck- and *WMAP*-motivated, respectively). The cosmological parameters in this table are defined as 'vanilla' parameters and θ_s denotes the angular size of the sound horizon at the redshift of last scattering.

Parameter	Symbol	Prior case I	Prior case II	Prior case III	Prior case IV
Cold dark matter density	$\Omega_c h^2$	0.001 → 0.99	0.001 → 0.99	0.001 → 0.99	0.001 → 0.99
Baryon density	$\Omega_b h^2$	0.005 → 0.1	0.005 → 0.1	0.005 → 0.1	0.005 → 0.1
100× approximation to θ_s	$100\theta_{\rm MC}$	0.5 → 10	0.5 → 10	0.5 → 10	0.5 → 10
Amplitude of scalar spectrum	$\ln(10^{10}A_s)$	1.7 → 5.0	1.7 → 5.0	2.3 → 5.0	2.3 → 5.0
Scalar spectral index	n_s	0.5 → 1.5	0.5 → 1.5	0.7 → 1.3	0.7 → 1.3
Dimensionless Hubble constant	h	0.2 → 1.4	0.2 → 1.4	0.4 → 1.3	0.4 → 1.3
Pivot scale (Mpc^{-1})	$k_{\rm pivot}$	0.05	0.002	0.05	0.002

The jackknife error estimates are computed by dividing the survey into sub-regions defined by the individual 1 deg^2 MegaCam pointings (also see Friedrich et al. 2016). The ratio of the error in each bin determined from the suite of mock catalogues, to the error determined from jackknife re-sampling, is displayed in Fig. 4, for the full data vector of 392 values ordered as described above. The ratio is close to unity for small angular scales where jackknife errors are reliable, but the jackknife error exceeds the dispersion of the simulations by a factor of more than 1.5 on the largest scales. Fig. 5 displays the full covariance matrix in the form of a correlation coefficient,

$$r(i, j) = \mathrm{Cov}(i, j)/\sqrt{\mathrm{Cov}(i, i)\,\mathrm{Cov}(j, j)}. \quad (20)$$

As in H13, we further obtain an unbiased estimate of the inverse covariance matrix by implementing the multiplicative correction advocated by Kaufman (1967) and Hartlap et al. (2007), such that

$$\mathbf{Cov}^{-1}_{\rm unbiased} = \frac{n_\mu - p - 2}{n_\mu - 1}\mathbf{Cov}^{-1}, \quad (21)$$

where n_μ is the number of pseudo-independent realizations and p is the number of data points. For 280 elements in our data vector and 1988 pseudo-independent realizations, we obtain a correction of 0.86, while for a reduced data vector of 56 elements (considered in the 'max' case defined in Section 3.5), we obtain a correction of 0.97. For both of these cases, our p/n_μ ratios are sufficiently low to avoid overestimating our Bayesian confidence regions by more than ≈5 per cent (Hartlap et al. 2007). While Sellentin & Heavens (2016) extend the analysis of Hartlap et al. (2007) by relaxing the assumption of a Gaussian likelihood, this is mostly visible in the tail of the distribution and does not significantly affect our analysis.

3 RESULTS

We now explore the cosmological constraints using the updated CFHTLenS cosmic shear tomography measurements analysed with the new cosmology fitting pipeline (described in Section 2.1.5). In accordance with Kilbinger et al. (2013) and H13, we mainly illustrate the cosmological constraints in the σ_8–Ω_m plane. We begin with a discussion of the constraints when no systematic uncertainties are included and then methodically include IA, baryonic uncertainties in the non-linear matter power spectrum and photometric redshift uncertainties, independently and jointly. We present the main results of these cases associated with the goodness of fit, DIC, Bayesian evidence, and data set concordance tests in Tables 3 and 5.

3.1 Including no systematic uncertainties

3.1.1 *Impact of cosmological priors*

As a first step, we explore the sensitivity of our weak lensing constraints to the choice of cosmological priors. To this end, we considered four separate cases, listed in Table 1. All of the cases assume the same broad priors for $\{\Omega_c h^2, \Omega_b h^2, \theta_{\rm MC}\}$ and they differ in the priors for $\{A_s, n_s, h, k_{\rm pivot}\}$. Cases I and II have wider priors on $\{A_s, n_s, h\}$ than cases III and IV. We moreover allow for either a *WMAP*-motivated pivot scale $k_{\rm pivot} = 0.002\,{\rm Mpc}^{-1}$ or a Planck-motivated pivot scale $k_{\rm pivot} = 0.05\,{\rm Mpc}^{-1}$, as a different $k_{\rm pivot}$ translates to different values for A_s and n_s and effectively changes the priors on these parameters. This choice of pivot scale distinguishes case I from case II and case III from case IV.

For the four cases considered, in the left-hand panel of Fig. 6, we find significant differences in the marginalized posterior contours along the σ_8–Ω_m plane. However, the four cases show remarkable agreement along the axis perpendicular to the degeneracy direction, such that the 2σ tension with Planck (reported earlier; e.g. MacCrann et al. 2015) effectively remains at the same level of significance regardless of the choice of priors. This is further manifested in the right-hand panel of Fig. 6, where we illustrate the constraints on $\sigma_8\Omega_m^{0.5}$. The marginalized posterior contours in the σ_8–Ω_m plane shrink when using tighter priors and also show a sensitivity to the choice of pivot scale. This implies that current lensing data from CFHTLenS is not sufficiently powerful to constrain the full vanilla parameter space when considering non-informative priors. The only two parameters that are constrained on both ends by the data are $\{\Omega_c h^2, \theta_{\rm MC}\}$, while the other parameters are unconstrained in either one or both directions. Moreover, the four cases differ from each other by at most $\Delta\chi^2_{\rm eff} = 2.3$, such that there is no strong statistical preference between the respective best-fitting points.

3.1.2 *Choice of cosmological priors*

While the contours in Fig. 6 could continue to expand by choosing ever more conservative priors, the cosmological priors for the four cases are all objectively conservative and we do not expect the true values of the parameters to lie outside any of the ranges specified in Table 1. In order to better understand if there exists a real tension between CFHTLenS and Planck, we hereafter adopt our fiducial case, consisting of case III with external priors on the Hubble constant and baryon density from Cepheid data and BBN. Concretely, we uniformly impose the prior $0.61 < h < 0.81$, consistent with Efstathiou (2014) at 99.7 per cent confidence level (CL), and we uniformly impose the extremely conservative prior

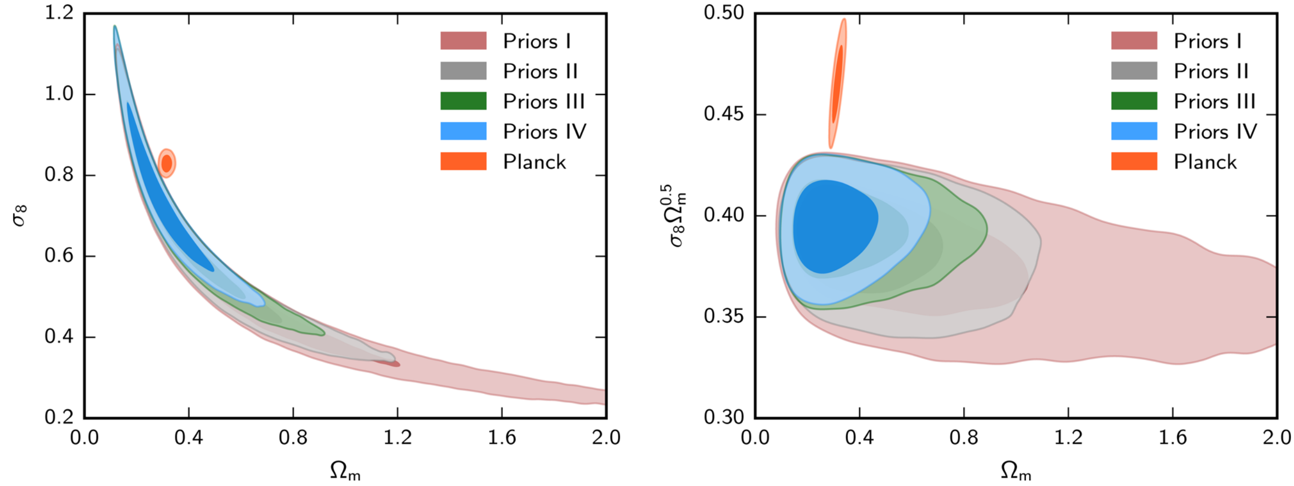

Figure 6. Marginalized posterior contours in the σ_8–Ω_m plane (inner 68 per cent CL, outer 95 per cent CL) from the updated CFHTLenS cosmic shear tomography measurements with different choices of cosmological priors (purple, grey, green, blue, for cases I–IV), defined in Table 1. The Planck contour is included for comparison in red (where our Planck data set is defined in Section 2.2). Right: same as the left-hand panel, except now showing contours in Ω_m against $\sigma_8\Omega_m^{0.5}$, orthogonal to the σ_8–Ω_m degeneracy direction.

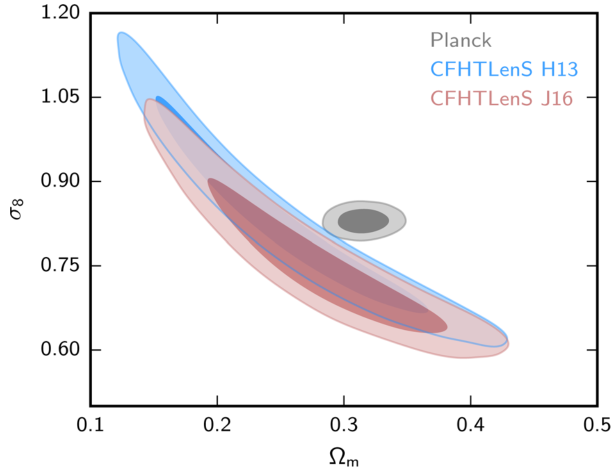

Figure 7. Marginalized posterior contour in the σ_8–Ω_m plane (inner 68 per cent CL, outer 95 per cent CL) from the updated CFHTLenS cosmic shear tomography measurements (CFHTLenS-J16; in purple), with fiducial cosmological priors listed in Table 2. For comparison, including the corresponding contour using the H13 measurements with our fiducial cosmological priors (CFHTLenS-H13; in blue) and the CMB measurements from Planck (in grey).

$0.013 < \Omega_b h^2 < 0.033$ (allowing for potential systematics and exotic physics), consistent with Burles, Nollett & Turner (2001), Olive & Particle Data Group (2014), and Cyburt et al. (2016). As these external priors are completely consistent with Planck, any further tension with Planck would therefore derive from CFHTLenS.

Given these choices for the priors, we show the resulting contours in the σ_8–Ω_m plane in Fig. 7. In addition to using the measurements and covariance described in this paper (denoted ‘CFHTLenS-J16’ with a data vector consisting of 280 elements given seven tomographic bins), we also show the resulting contour using the original CFHTLenS measurements and covariance from H13 (denoted ‘CFHTLenS-H13’ with a data vector consisting of 210 elements given six tomographic bins). We find that the two analyses agree well and that there seems to be a marginal increase in the tension with Planck for the new measurements. As for the statistical goodness of the lensing fits, we find $\chi^2_{\rm red} = 1.51$ for the new measurements, as compared to $\chi^2_{\rm red} = 1.19$ for the old measurements.

The reduction in the ‘goodness of fit’ between the two analyses derives from two changes in the analysis. The first change is the use of a new suite of *N*-body simulations to determine the covariance matrix. In H13, the field of view of the 184 simulations used was only 12.84 deg^2. In order to gain enough mock realizations to accurately invert the covariance matrix, they split the simulations into 3 × 3 sub-realizations such that each sub-realization was close in size to the ≈53 arcmin maximum scale measured for the lensing statistics. Pairs on those scales were therefore ‘missing’ due to edge effects and as a result, the error on large scales was overestimated. In our analysis, the field of view of the 497 simulations used is 60 deg^2 and we can therefore measure the large-scale simulated covariance accurately. As the CFHTLenS data are a poorer fit to the model on large scales, the reduction in errors on large scales results in an increased $\chi^2_{\rm red}$.

While our new covariance analysis is certainly an improvement on H13, it also does not include supersample variance terms (Takada & Hu 2013). These supersampling variance errors contribute to all angular scales and are missing from our calculation as very large-scale modes in the density field are not simulated in the finite box of the *N*-body simulations. However, from the good agreement between the jackknife and simulated errors in Fig. 4, we can conclude that these supersample terms are not significant on small scales where the majority of the cosmological information is accessed. On large scales, including supersample terms is likely to improve the goodness of fit of the data, an analysis that we will pursue in future work.

The second change in our analysis is the use of angular scales larger than the 50 arcmin limit of H13, introduced owing to the limitation of their simulations. Asgari et al. (2017) have recently presented an optimal E/B mode decomposition analysis of CFHTLenS using the COSEBIs statistic (Schneider, Eifler & Krause 2010; Asgari, Schneider & Simon 2012). This analysis reveals significant B modes on large angular scales ($\theta > 40$ arcmin) that do not derive from gravitational lensing, which exhibits a pure E-mode signal. These B modes are further enhanced when the data are analysed in tomographic bins.

Table 2. Priors on the systematic degrees of freedom when considered independently (applicable to Sections 3.1.2–3.4 and Figs 7–11). We also list the fiducial cosmological priors (applicable everywhere from Section 3.1.2). The '→' sign indicates uniform priors and the '±' sign indicates Gaussian priors. When the IA parameters and photo-z bins are fixed, they are set to zero. Since we do not have external information on the sixth and seventh photo-z bins, when considering an informative photo-z scenario, we keep non-informative priors on these bins. Please note that we have imposed informative priors on the baryon density and Hubble constant, as described in Section 3.1.2. The cosmological parameters in this table are defined as 'vanilla' parameters and $\theta_{\rm s}$ denotes the angular size of the sound horizon at the redshift of last scattering.

Parameter	Symbol	Prior
Cold dark matter density	$\Omega_{\rm c}h^2$	$0.001 \rightarrow 0.99$
Baryon density	$\Omega_{\rm b}h^2$	$0.013 \rightarrow 0.033$
100 × approximation to $\theta_{\rm s}$	$100\theta_{\rm MC}$	$0.5 \rightarrow 10$
Amplitude of scalar spectrum	$\ln(10^{10}A_{\rm s})$	$2.3 \rightarrow 5.0$
Scalar spectral index	$n_{\rm s}$	$0.7 \rightarrow 1.3$
Dimensionless Hubble constant	h	$0.61 \rightarrow 0.81$
Pivot scale (Mpc^{-1})	$k_{\rm pivot}$	0.05
IA amplitude	A	$-50 \rightarrow 50$
– informative case		$-6 \rightarrow 6$
IA redshift dependence	η	$-50 \rightarrow 50$
– informative case		0
IA luminosity dependence	β	$-50 \rightarrow 50$
– informative case		1.13 ± 0.25
HMCODE feedback amplitude	$\log B$	$0 \rightarrow 2$
– informative case		$0.3 \rightarrow 0.6$
– when fixed		0.496
Photo-z bin 1	Δz_1	$-0.1 \rightarrow 0.1$
– informative case		-0.045 ± 0.013
Photo-z bin 2	Δz_2	$-0.1 \rightarrow 0.1$
– informative case		-0.014 ± 0.010
Photo-z bin 3	Δz_3	$-0.1 \rightarrow 0.1$
– informative case		0.008 ± 0.008
Photo-z bin 4	Δz_4	$-0.1 \rightarrow 0.1$
– informative case		0.042 ± 0.017
Photo-z bin 5	Δz_5	$-0.1 \rightarrow 0.1$
– informative case		0.042 ± 0.034
Photo-z bin 6	Δz_6	$-0.1 \rightarrow 0.1$
Photo-z bin 7	Δz_7	$-0.1 \rightarrow 0.1$

Asgari et al. (2017) also present a compressed-COSEBIs analysis of CFHTLenS (formalism introduced in Asgari & Schneider 2015), where the COSEBIs are optimally combined to extract cosmological information. In this compressed analysis, the recovered B modes are consistent with zero (except for the blue galaxy case in the 40–100 arcmin range with six tomographic bins). If we assume that the systematics that introduce B modes into the data contribute equally to the E and B modes, we can conclude that these systematics will impact on the goodness of fit of the E mode, particularly on large scales where the B modes are found to be at their strongest. However, as the compressed cosmological parameter analysis results in a zero B mode, in particular when the full galaxy sample is considered, these B modes are not degenerate with cosmological parameters and are therefore fairly benign in the cosmological analysis that follows, particularly when we allow for uncertainty in the three astrophysical sources of systematics that we focus on in this paper. We will investigate the origin of these B modes further in future work.

Table 3. Exploring changes in $\chi^2_{\rm eff}$ and DIC for different choices of systematic uncertainties, given fiducial cosmological priors. The reference vanilla model without systematic uncertainties gives $\chi^2_{\rm eff} = 414.6$ and DIC $= 421.7$ when using HALOFIT and $\chi^2_{\rm eff} = 416.4$ and DIC $= 423.3$ when using HMCODE (with fixed $\log B = 0.496$). Since the size of the data vector for the 'max' case is significantly smaller than the size of the fiducial data vector (where the max case keeps only 'large' angular scales and is defined in Section 3.5), we calculate the difference in $\chi^2_{\rm eff}$ and DIC with respect to the measurements used for the 'max' case but without systematic uncertainties. For this reduced data vector, $\chi^2_{\rm eff} = 86.8$ and DIC $= 92.0$, considering HMCODE with $\log B = 0.496$.

Model	$\Delta\chi^2_{\rm eff}$	ΔDIC
vanilla + A	−5.8	−4.6
– informative case	−5.7	−4.9
vanilla + $\{A, \eta, \beta\}$	−21	12
– informative case	−0.72	2.4
vanilla + B	−1.9	−0.64
– informative case	−0.22	1.4
vanilla + 7 photo-z	−6.0	0.97
– informative case	1.5	11
vanilla + min case	−10	1.6
vanilla + mid-case	−1.1	12
vanilla + max case	−25	19

We now proceed to exploring the impact of three distinct systematic uncertainties on our results: intrinsic galaxy alignments, baryonic uncertainties in the non-linear matter power spectrum, and photometric redshift uncertainties.

3.2 Including intrinsic galaxy alignments

We begin by including the three systematic uncertainties independently, before accounting for them jointly. The first of these systematic uncertainties comes from the IA of galaxies. We consider two separate scenarios, one in which we only allow for a variation of the amplitude A (defined in equation 7) and a second scenario in which we also allow for a possible redshift and luminosity dependence of the IA signal via the two parameters η and β, respectively. For each of the two scenarios, we consider both informative and non-informative priors. We specify these priors in Table 2.

More specifically, when varying the vanilla parameters along with A, we let the amplitude vary uniformly in the range $\{-50, 50\}$ for the non-informative case and uniformly in the range $\{-6, 6\}$ for the informative case. When varying the vanilla parameters with $\{A, \eta, \beta\}$, we let each of the IA parameters vary uniformly in the range $\{-50, 50\}$ for the non-informative case. For the informative case, we let A vary uniformly in the range $\{-6, 6\}$, we fix $\eta = 0$, and we impose $\beta = 1.13 \pm 0.25$ as a Gaussian prior (motivated by Joachimi et al. 2011).

While our non-informative priors are reasonably wide, our informative priors are driven by the fact that our sample is dominated by blue galaxies, which are known to be less sensitive to IA effects. We have therefore taken the tightest available luminosity and redshift dependent constraints determined from red galaxies as the 'worst-case scenario' for the luminosity and redshift dependence of the dominant blue galaxies in the sample (from Joachimi et al. 2011, consistent with Singh, Mandelbaum & More 2015), while also encompassing the luminosity and redshift dependence in the red sample. For the informative case, we set $\eta = 0$ given the lack of evidence for redshift evolution in Joachimi et al. (2011) and Singh et al. (2015). We further allow for negative values of the IA

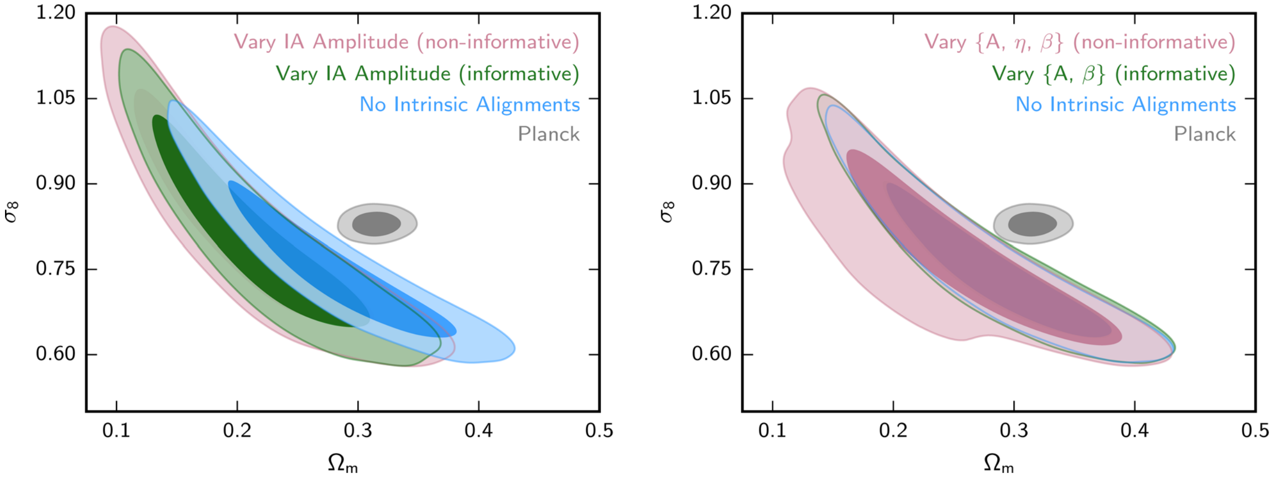

Figure 8. Left: marginalized posterior contours in the σ_8–$\Omega_{\rm m}$ plane (inner 68 per cent CL, outer 95 per cent CL) from the updated CFHTLenS cosmic shear tomography measurements, considering fiducial cosmological priors, where the IA amplitude A is allowed to vary with both informative and non-informative priors on the amplitude (in green and purple, respectively). We moreover include the fiducial (no-systematics) cosmic shear and Planck CMB contours for comparison (in blue and grey, respectively). The fiducial cosmological and IA priors are listed in Table 2. Right: marginalized posterior contours where all three IA parameters $\{A, \eta, \beta\}$, encapsulating the amplitude, redshift dependence, and luminosity dependence of the IA, respectively, are allowed to vary jointly (with both informative and non-informative priors on the three parameters, in green and purple, respectively). Fiducial cosmic shear and Planck CMB contours included for comparison (in blue and grey, respectively).

amplitude as the best-fitting model for a mixed population could result in such values, as discussed in H13.

In Fig. 8, we show the marginalized posterior contours in the σ_8–$\Omega_{\rm m}$ plane for different IA models. When varying the vanilla parameters with the IA amplitude (i.e. without luminosity or redshift dependence), the contours slightly degrade due to the extra degree of freedom despite the additional cosmological information contained in the II and GI terms (described in Section 2.1.2). The contours moreover shift towards larger values of σ_8 and smaller values of the matter density (evident from equation 7), increasing the tension with Planck. The degradation and shift in the contours are in agreement with H13, who pointed out that σ_8 is driven towards larger values by negative best-fitting estimates of the IA amplitude.

The observed behaviour applies to both the non-informative and informative cases, as the non-informative constraint on the IA amplitude is $A = -3.6 \pm 1.6$ (corresponding to the mean of the posterior distribution along with the symmetric 68 per cent confidence interval about the mean), which implies that the informative case will give similar constraints. As seen in Table 3, the informative and non-informative cases only differ from each other by $\Delta\chi^2_{\rm eff} = 0.1$. They further differ from the $A = 0$ scenario by $\Delta\chi^2_{\rm eff} = -5.8$ for the non-informative case and $\Delta\chi^2_{\rm eff} = -5.7$ for the informative case. The penalty due to the increased Bayesian complexity gives $\Delta{\rm DIC} = -4.6$ for the non-informative case and $\Delta{\rm DIC} = -4.9$ for the informative case.

Thus, there seems to be substantial preference in favour of a non-zero and negative IA amplitude (as also found in Fig. 9). While the IA amplitude for a single sample must be positive, for a mixed sample, the best-fitting model can be negative, as described in H13. Thus, this could be a sign that the model is a good fit to red galaxies, while the majority of our sample is blue. Alternatively, the negative IA amplitude could imply that the IA model we use is too simplistic or the result of unaccounted systematics. For instance, photometric redshift errors could also mimic IA-like behaviour, so this could be a sign that the low-redshift source distributions are inaccurate.

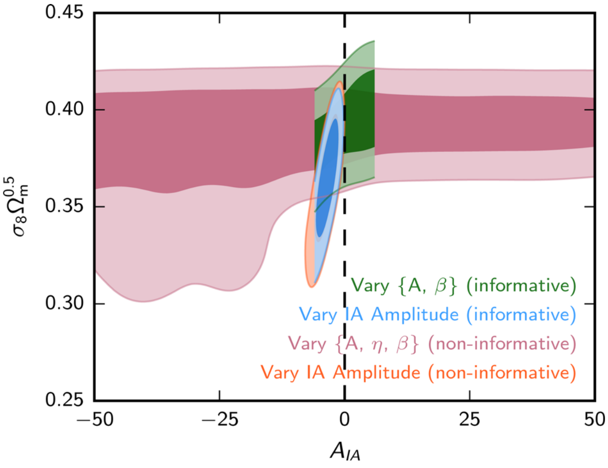

Figure 9. Marginalized posterior contours in the plane given by $\sigma_8\Omega_{\rm m}^{0.5}$ and IA amplitude A (inner 68 per cent CL, outer 95 per cent CL) from the updated CFHTLenS cosmic shear tomography measurements, considering both informative and non-informative priors on the IA parameters $\{A, \eta, \beta\}$, encapsulating the amplitude, redshift dependence, and luminosity dependence of the IA, respectively. The fiducial cosmological and IA priors are listed in Table 2.

In Fig. 8, we further show the cosmological constraints as we let all three IA parameters vary simultaneously. We note that this three-parameter model has not been fit to large-scale cosmic shear data before; for example, H13 only considered varying the IA amplitude and Abbott et al. (2016) considered a two-parameter model with a varying amplitude and redshift dependence. When imposing informative priors on $\{A, \eta, \beta\}$, the additional degree of freedom from the luminosity dependence causes the contour in the σ_8–$\Omega_{\rm m}$ plane to effectively transform back to the original contour given by the $A = 0$ scenario. This is because our prior on β decreases the strength of the IA II and GI terms in the lensing calculation, even though A is unbounded within the region given by the prior. Hence, the cosmological constraints with a large value of β mimic the constraints for the scenario with no IA. There is therefore less tension with Planck for this model than the one-parameter IA model

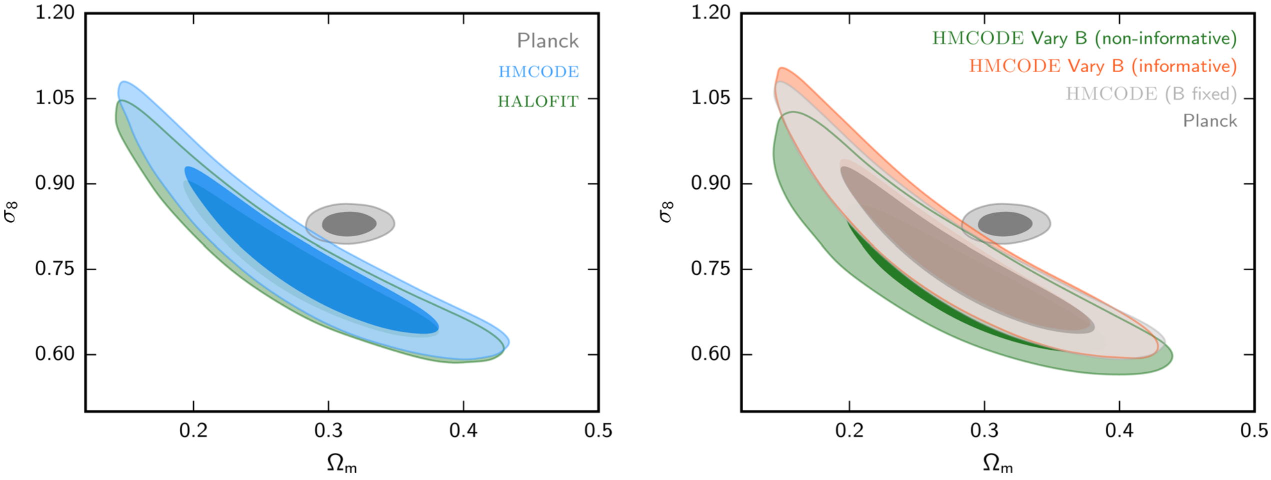

Figure 10. Left: marginalized posterior contours in the σ_8–Ω_m plane (inner 68 per cent CL, outer 95 per cent CL) from the updated CFHTLenS cosmic shear tomography measurements when using either HMCODE or HALOFIT for the non-linear matter power spectrum (in blue and green, respectively), where the HMCODE feedback amplitude log *B* is fixed to its fiducial DM-only value of 0.496. The Planck contour is included for comparison (in grey) and the cosmological priors are listed in Table 2. Right: marginalized posterior contours where the HMCODE feedback amplitude is allowed to vary (with both informative and non-informative priors on the amplitude, listed in Table 2, in red and green, respectively). The fiducial cosmic shear contour (where log $B = 0.496$) and the Planck CMB contour are included for comparison (in brown and grey, respectively).

that we first considered. Here, we find $\Delta\chi^2_{\rm eff} = -0.72$, while ΔDIC $= 2.4$ shows a weak preference against the extended IA model.

When considering non-informative priors on $\{A, \eta, \beta\}$, we find an enlarged contour towards smaller values of the matter density (where the enlarged region is consistent with extremely negative values of the IA amplitude as seen in Fig. 9). While the IA amplitude is completely unconstrained, we find one-sided bounds on $-50 < \eta < 8.0$ and $0.54 < \beta < 50$ at 95 per cent CL (where $\beta > 0$ even at 99.7 per cent CL). In other words, η and β are taking on values that exclude potentially enormous IA signals and are, in fact, consistent with a negligible signal. The range for η is moreover consistent with no redshift evolution, while the range for β shows a weak preference for a luminosity dependence of the IA. Interestingly, we find $\Delta\chi^2_{\rm eff} = -21$ for this three-parameter IA model, while the increased Bayesian complexity of the model is severely penalized in ΔDIC $= 12$. This demonstrates the extreme usefulness of the information criterion, as it determines the three-parameter IA model to be less preferred than the reference model without IA.

3.3 Including baryonic uncertainties in the non-linear matter power spectrum

We now proceed to another important systematic coming from baryonic uncertainties in the non-linear matter power spectrum. We account for the baryonic effects by varying the HMCODE feedback amplitude *B*, described in Section 2.1.4. For the scenario with only DM, we fix log $B = 0.496$ as advocated in Mead et al. (2015).

In Fig. 10, we first show the marginalized posterior contours in the σ_8–Ω_m plane corresponding to the use of either HALOFIT or HMCODE for the non-linear extension to the matter power spectrum, considering no baryonic effects on non-linear scales (such that log *B* is fixed to its fiducial value of 0.496 for HMCODE). As one would expect from Mead et al. (2015), which is in excellent agreement with Takahashi et al. (2012) for cosmologies where no baryonic effects are included, the contours agree remarkably well for this DM-only scenario. This implies the two matter power spectrum prescriptions can be used interchangeably when baryonic effects are not included on non-linear scales. Since HALOFIT is faster than HMCODE (as discussed in Section 2.1.5), this has allowed us to comfortably use HALOFIT for our runs where HMCODE is not directly needed (i.e. when non-linear baryonic effects are not included). As for the relative fits for the DM-only scenario, using either HMCODE or HALOFIT, we find a difference of $\Delta\chi^2_{\rm eff} = 1.8$ and ΔDIC $= 1.5$ between the models, such that they are close to equally preferred (with the weak preference in favour of HALOFIT).

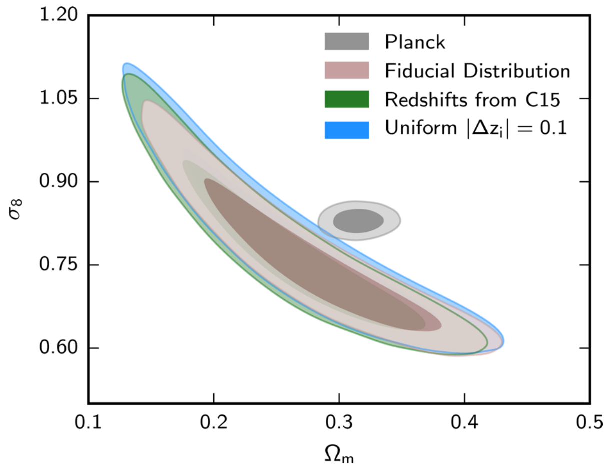

Figure 11. Marginalized posterior contours in the σ_8–Ω_m plane (inner 68 per cent CL, outer 95 per cent CL) from the updated CFHTLenS cosmic shear tomography measurements for different treatments of the photometric redshift uncertainties. The contour where the fiducial redshift distribution is used is given in brown, the contour where the redshift distribution is perturbed according to the results from source–lens cross-correlations in Choi et al. (2016) is given in green, and the contour where the redshift distribution is perturbed with uniform priors of $|\Delta z_i| = 0.1$ in each tomographic bin is given in blue. The Planck contour is included for comparison in grey, while the fiducial cosmological and photometric redshift priors are listed in Table 2.

In Fig. 10, we further show the constraints when allowing the amplitude *B* in HMCODE to vary freely, considering both informative and non-informative priors (listed in Table 2). We take the

Table 4. Exploring changes in systematic priors for three joint scenarios (applicable to Section 3.5 and Fig. 12). The 'min' case corresponds to the most optimistic scenario for the priors and the 'max' case corresponds to the most conservative scenario, while the 'mid-' case lies between these two scenarios. In our language, keeping 'all' angular scales implies the data vector consists of 280 elements, while keeping 'large' angular scales implies the data vector consists of 56 elements, as discussed in Section 3.5. The '→' sign indicates uniform priors and the '±' sign indicates Gaussian priors. Moreover, the priors on the underlying cosmology are the same as in Table 2.

Parameter	Symbol	Min-case	Mid-case	Max case
IA amplitude	A	$-6 \rightarrow 6$	$-6 \rightarrow 6$	$-50 \rightarrow 50$
IA luminosity dependence	β	0	1.13 ± 0.25	$-50 \rightarrow 50$
IA redshift dependence	η	0	0	$-50 \rightarrow 50$
HMCODE feedback amplitude	$\log B$	$0.3 \rightarrow 0.6$	$0.3 \rightarrow 0.6$	$0 \rightarrow 2$
Photo-z bin 1	Δz_1	-0.045 ± 0.013	-0.045 ± 0.050	$-0.1 \rightarrow 0.1$
Photo-z bin 2	Δz_2	-0.014 ± 0.010	-0.014 ± 0.050	$-0.1 \rightarrow 0.1$
Photo-z bin 3	Δz_3	0.008 ± 0.008	0.008 ± 0.050	$-0.1 \rightarrow 0.1$
Photo-z bin 4	Δz_4	0.042 ± 0.017	0.042 ± 0.050	$-0.1 \rightarrow 0.1$
Photo-z bin 5	Δz_5	0.042 ± 0.034	0.042 ± 0.050	$-0.1 \rightarrow 0.1$
Photo-z bin 6	Δz_6	$-0.1 \rightarrow 0.1$	$-0.1 \rightarrow 0.1$	$-0.1 \rightarrow 0.1$
Photo-z bin 7	Δz_7	$-0.1 \rightarrow 0.1$	$-0.1 \rightarrow 0.1$	$-0.1 \rightarrow 0.1$
Angular scales	θ	All	All	Large

informative prior to uniformly cover the range $0.3 < \log B < 0.6$, which effectively corresponds to the range given by the best-fitting values of the {DMONLY, REF, DBLIM, AGN} cases in the OWL simulations (Schaye et al. 2010; van Daalen et al. 2011), as demonstrated by Mead et al. (2015). This informative case gives a contour that marginally prefers larger values of σ_8, while $\Delta\chi^2_{\rm eff} = -0.22$ and $\Delta{\rm DIC} = 1.4$ (as compared to the case where $\log B$ is fixed to 0.496), suggesting that there is no strong preference for or against the additional degree of freedom. In other words, the constraining power of the weak lensing data set from CFHTLenS seems to be inadequate to distinguish the DM-only model for the non-linear matter power spectrum from one of the models of the OWL simulations.

For the non-informative case, the HMCODE feedback amplitude is constrained by the data such that $0.37 < \log B < 1.1$ at 95 per cent CL. The resulting contour is both expanded and shifted to smaller values of σ_8 (in the plane with $\Omega_{\rm m}$). This downward shift in σ_8 is caused by unnaturally large values of B that are allowed by the data, despite the baryonic OWL simulation models preferring values of B below the fiducial value (as B and σ_8 are anticorrelated). Thus, as compared to the fiducial model, allowing for an extra degree of freedom in the non-linear matter power spectrum does not seem to alleviate the tension between CFHTLenS and Planck. Moreover, $\Delta\chi^2_{\rm eff} = -1.9$ and $\Delta{\rm DIC} = -0.64$, which again implies the data does not strongly prefer B to stray away from its fiducial DM-only value.

3.4 Including photometric redshift uncertainties

We now turn to the third key systematic: photometric redshift uncertainties. We consider seven tomographic bins in our analysis and therefore introduce seven new parameters to allow the source distribution of each tomographic bin to shift along the redshift axis, preserving the shape of each distribution.

As shown in Table 2, we consider two distinct cases for our priors, one where $-0.1 < \Delta z_i < 0.1$ for each tomographic bin (varied uniformly) and one where Gaussian priors are obtained from Choi et al. 2016 (with minor variations; hereafter also denoted C15) for the first five tomographic bins, while the last two bins are varied uniformly between -0.1 and 0.1. The informative priors are derived by fitting angular cross-correlation function measurements between sources in each tomographic bin and an overlapping spectroscopic sample from BOSS, as detailed in Choi et al. (2016). The priors are only available for the first five tomographic bins, given the redshift coverage of BOSS and the numbers differ slightly from those presented in Choi et al. (2016) because of more conservative error estimation and a different normalization scheme.[6] The two sets of numbers agree well within the 1σ error bars and the choice of normalization has a negligible impact on our analysis.

In Fig. 11, we show the marginalized posterior contours in the σ_8–$\Omega_{\rm m}$ plane for the different treatments of the redshift distribution. For both informative and non-informative prior cases, we find only small changes in the contours along the σ_8–$\Omega_{\rm m}$ plane as compared to the fiducial scenario where the redshift distribution is fixed, in agreement with a similar analysis in Abbott et al. (2016). Here, the informative contour is marginally expanded, but not in the region that would increase the agreement with Planck, while the non-informative contour is marginally expanded and shifted towards the Planck contour.

While the tomographic shifts are given by C15 for the informative case (aside from the last two bins, which are found to be entirely unconstrained within the prior), at 95 per cent CL, we find $-0.06 < \Delta z_1 < 0.1$, $-0.03 < \Delta z_2 < 0.1$, $-0.1 < \Delta z_3 < 0.1$, $-0.05 < \Delta z_4 < 0.1$, $-0.1 < \Delta z_5 < 0.1$, $-0.1 < \Delta z_6 < 0.06$, $-0.1 < \Delta z_7 < 0.1$ for the case with uniform priors ($|\Delta z_i| = 0.1$). These bounds demonstrate that the constraints are weak and the uncertainties would increase with wider priors. For the C15 case, we find $\Delta\chi^2_{\rm eff} = 1.5$ and $\Delta{\rm DIC} = 11$, such that the more complex model seems to be strongly disfavoured by the data as compared to the fiducial model. This finding is in agreement with the conclusions of Choi et al. (2016), who showed that the best-fitting model for the one-parameter shift of the $n(z)$ used in our 'min' case (and advocated in Abbott et al. 2016) is actually insufficient to encompass the errors in the CFHTLenS redshift distributions. Significantly better fits to the spectroscopic–photometric cross-correlation clustering measurements can be obtained when the width of the redshift distribution is allowed to vary along with the peak. We will explore this

---

[6] See section 2.5 of Choi et al. (2016) for details of the redshift probability distribution shifting procedure. Whenever negative shifts were applied, the BPZ probability distributions were renormalized by the integrated probability including negative redshifts. The renormalization performed in the process of obtaining our priors neglects the probability distribution shifted to negative redshifts.

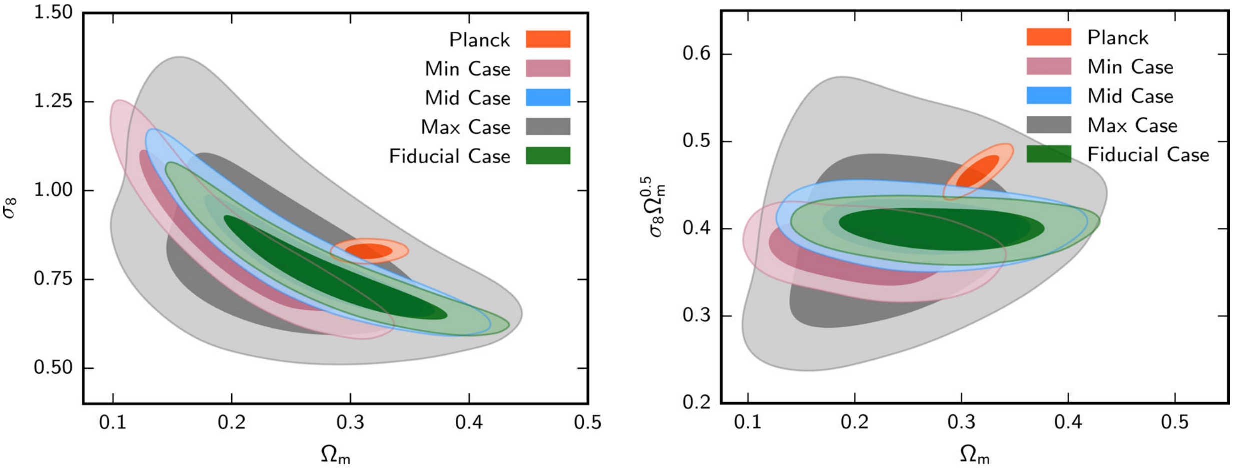

Figure 12. Left: marginalized posterior contours in the σ_8–Ω_m plane (inner 68 per cent CL, outer 95 per cent CL) from the updated CFHTLenS cosmic shear tomography measurements for the joint analysis of the systematic uncertainties, where the priors on the cosmological and systematic degrees of freedom are listed in Tables 2 and 4, respectively. We include the 'min' case in purple, 'mid-' case in blue, 'max' case in grey, and fiducial case in green. The fiducial case keeps the HMCODE feedback amplitude $\log B$ fixed at the DM-only value of 0.496 and does not include any systematic uncertainties. We further include the Planck CMB contour for comparison in red. Right: same as the left-hand panel, except now showing contours in Ω_m against $\sigma_8\Omega_m^{0.5}$, orthogonal to the σ_8–Ω_m degeneracy direction.

further freedom in the photometric redshift distribution uncertainty in future work.

For the case with uniform priors on all tomographic bins, $\Delta\chi^2_{\rm eff} = -6.0$ and $\Delta{\rm DIC} = 0.97$. Thus, the goodness of fit improves when allowing for uniform deviations around the fiducial distribution, as opposed to the shifts advocated by C15, but keeping the new degrees of freedom is not preferred by the data.

3.5 Joint account of systematic uncertainties

Within the ΛCDM cosmology, we now consider the joint analysis of systematic uncertainties coming from IA of galaxies, non-linear baryonic physics in the matter power spectrum, and photometric redshift errors. To this end, we set up three distinct cases for the priors on the new degrees of freedom, a minimum ('min') case with informative priors, a maximum ('max') case with conservative priors, and a middle ('mid-') case with a combination of informative and conservative priors. We list the priors for these three cases in Table 4. Meanwhile, for the underlying cosmology, we continue to impose the fiducial priors given in Table 2.

For the 'min' case, we assume an informative uniform prior on the IA amplitude of $-6 < A < 6$ and exclude a luminosity or redshift dependence of the IA signal. We assume an informative uniform prior on the HMCODE amplitude $0.3 < \log B < 0.6$, while the fiducial tomographic redshift distributions are perturbed by the shifts given in Choi et al. 2016 (aside from the last two bins, as discussed in Section 3.4). For the 'mid-' case, we keep the same settings as for the 'min' case, except we now allow for a luminosity dependence via $\beta = 1.13 \pm 0.25$ and we increase the error bars on the C15 redshift shifts to $|\Delta z_i| = 0.05$. Lastly, for the 'max' case, we impose wide priors on all systematic degrees of freedom, such that $-50 < \{A, \eta, \beta\} < 50$, $0 < \log B < 2$, and $-0.1 < \Delta z_i < 0.1$.

As described in Section 2.2, our measurements are evaluated at seven angular bins (for each of the seven tomographic bins). However, for the 'max' case, we consider removing the dependence on non-linear scales in the matter power spectrum altogether. To this end, we follow Ade et al. (2016XIV) and cut our data vector by removing ξ_- entirely and keeping ξ_+ for $\theta > 17$ arcmin. In practice, this implies we only keep our measurements of ξ_+ at 21.7 and 43.0 arcmin (since we already remove the ξ_+ measurements at 85.2 arcmin in the fiducial data vector), such that the downsized data vector consists of 56 elements (from the fiducial vector of 280 elements, itself originally downsized from 392 elements).

In Fig. 12, we show the marginalized posterior contours for the three systematic cases along the σ_8–Ω_m plane. For the 'min' case, which includes informative priors on the systematic uncertainties, the main change to the contour comes from the freedom in the IA amplitude. This is because the impact of the baryonic and photometric redshift uncertainties is marginal when imposing informative priors (as seen in Sections 3.3 and 3.4). As expected from the left-hand panel of Fig. 8, where only the IA amplitude is varied freely (in addition to the vanilla cosmological parameters), the expanded 'min' contour primarily shifts towards smaller values of the matter density and away from the Planck contour, as compared to the fiducial case with no systematic uncertainties included. Thus, for the min case, $\sigma_8\Omega_m^{0.5} = 0.372^{+0.023}_{-0.022}$ (68 per cent CL), as compared to $\sigma_8\Omega_m^{0.5} = 0.401^{+0.016}_{-0.017}$ (68 per cent CL) for the fiducial case.

Proceeding from the 'min' case to the 'mid-' case, the expanded contour is shifted back to values of $\{\sigma_8, \Omega_m\}$ that overlap with those of the fiducial scenario due to the additional degree of freedom from the luminosity dependence. More specifically, the parameter β is sufficiently large to decrease the IA signal, consistent with the behaviour seen for the informative case in the right-hand panel of Fig. 8. The contour for the 'mid-' case is moreover expanded in the σ_8 direction due to the photometric redshift uncertainties, consistent with the behaviour seen in Fig. 11. The combination of these two effects brings the 'mid' contour in greater agreement with Planck (as compared to the fiducial case), as also quantified by $\sigma_8\Omega_m^{0.5} = 0.404^{+0.021}_{-0.021}$ (68 per cent CL).

In the 'max' case, a combination of conservative priors and downsized data vector increases the size of the contour to such an extent that it encloses all of the aforementioned contours, including the Planck contour. It is clear that in this pessimistic scenario, the CFHTLenS data set is only able to place weak constraints in the σ_8–Ω_m plane, with $\sigma_8\Omega_m^{0.5} = 0.395^{+0.074}_{-0.064}$ (68 per cent CL), although it does retain the anticorrelated shape between σ_8 and the matter

Table 5. Exploring changes in $\log \mathcal{C}$, $\log \mathcal{I}$, and $\ln \mathcal{B}_{01}$ for different choices of systematic uncertainties given fiducial cosmological priors. For $\log \mathcal{C}$ and $\log \mathcal{I}$, positive values indicate concordance between CFHTLenS and Planck, while negative values indicate discordance between the data sets. For $\ln \mathcal{B}_{01}$, which only considers CFHTLenS data, positive values indicate the fiducial model is preferred, while negative values indicate that the extended model is preferred. We note that 'ln' refers to the natural logarithm, while 'log' refers to the common logarithm (base 10). The Planck CMB evidence is given by $\ln \mathcal{Z} = -5680.5$, while the vanilla CFHTLenS evidence is given by $\ln \mathcal{Z} = -214.9$ when using HMCODE with feedback amplitude $\log B = 0.496$. The vanilla CFHTLenS evidence for the downsized data vector of the 'max' case is $\ln \mathcal{Z} = -47.9$ (considering $\log B = 0.496$). From these numbers and those in the table, the individual evidences can be reconstructed. Moreover, Planck's DIC = 11297.1, which can be used to reconstruct the joint DIC estimates from the $\log \mathcal{I}$ estimates in this table and the DIC estimates from Table 3. The errors in our calculations of $\log \mathcal{C}$ and $2 \ln \mathcal{B}_{01}$ are approximately 0.2 and 0.3, respectively. For the joint systematics calculations, we consistently use HMCODE for the non-linear matter power spectrum.

Model	$\log \mathcal{C}$	$\log \mathcal{I}$	$2 \ln \mathcal{B}_{01}$
vanilla (HMCODE, B fixed)	−0.47	−0.98	0
vanilla (HALOFIT)	−0.96	−1.6	−1.9
vanilla + min case	−1.8	−2.6	−6.6
vanilla + mid-case	0.51	−0.28	1.3
vanilla + max case	0.90	0.62	−0.52

density. Thus, from the marginalized posterior contours, it seems the introduction of key systematic uncertainties from IA, baryons, and photometric redshifts is able to alleviate the tension with Planck for the more conservative 'mid-' and 'max' cases (as compared to the fiducial case). However, as pointed out in e.g. Raveri (2016), there is a risk of biasing one's conclusions when assessing data set concordance from marginalized posterior contours. We therefore proceed to evaluate the three cases of the joint analysis more quantitatively.

In the joint analysis, we find $\Delta\chi^2_{\rm eff} = -10$, $\Delta{\rm DIC} = 1.6$, and $2 \ln \mathcal{B}_{01} = -6.6$ for the 'min' case (with respect to the fiducial case and where $\mathcal{B}_{01}$ refers to the Bayes factor defined in Section 2.1.7). This illustrates the usefulness of model selection based on multiple statistics, as the DIC and evidence estimates point in somewhat different directions. This may be a reflection of the parameter priors, analogous to the 'Jeffreys–Lindley paradox' (Lindley 1957; Jeffreys 1961; also see Cousins 2014) We follow the prescription in Kunz et al. (2006) and conclude that despite the increased complexity, the improvement in the evidence is sufficiently large to warrant the 'min' case as favoured by the data.

For the 'mid-' case, $\Delta\chi^2_{\rm eff} = -1.1$, $\Delta{\rm DIC} = 12$, and $2 \ln \mathcal{B}_{01} = 1.3$, which implies a preference against the more complex model at a greater significance when employing the DIC as compared to the evidence. For the 'max' case, we find $\Delta\chi^2_{\rm eff} = -25$, $\Delta{\rm DIC} = 19$, and $2 \ln \mathcal{B}_{01} = -0.52$, such that this model is roughly equally favoured to the 'vanilla' model when employing the evidence, but highly disfavoured when accounting for its increased complexity. We note that the changes in $\chi^2_{\rm eff}$, evidence, and DIC for the 'max' case are when compared to a fiducial case using the downsized data vector keeping 'large' angular scales but without systematics. For this fiducial case, $\chi^2_{\rm red} = 1.74$, up from $\chi^2_{\rm red} = 1.51$ when including 'all' scales, suggesting that the high $\chi^2_{\rm red}$ is exacerbated on large scales (in agreement with the discussion in Section 3.1.2).

Turning to the question of data set concordance between CFHTLenS and Planck, we computed both the $\log \mathcal{C}$ and $\log \mathcal{I}$ statistics (defined in equations 14 and 15, respectively), with results shown in Table 5. These two data set concordance tests are, respectively, based on the Bayesian evidence and information theory, detailed in Section 2.1.7. For the scenario without systematic uncertainties included, we find $\log \mathcal{C} = -0.47$ and $\log \mathcal{I} = -0.98$ when employing HMCODE for the non-linear matter power spectrum (with DM-only feedback amplitude of $\log B = 0.496$), showing substantial degree of discordance between CFHTLenS and Planck for the two statistics. When instead employing HALOFIT for the non-linear matter power spectrum, $\log \mathcal{C} = -0.96$ and $\log \mathcal{I} = -1.6$, pointing towards strong discordance between the data sets. The increase in the discordance between the data sets when employing HALOFIT (as compared to HMCODE) is in agreement with the increased separation in the marginalized posterior contours in the σ_8–$\Omega_{\rm m}$ plane in Fig. 7. As shown in Table 5, the Bayes factor for the vanilla ΛCDM model with HMCODE relative to HALOFIT is $2 \ln \mathcal{B}_{01} = -1.9$, such that the evidence is marginally improved when employing HALOFIT as compared to HMCODE (with $\log B = 0.496$).

Including systematic uncertainties, we find $\log \mathcal{C} = -1.8$ and $\log \mathcal{I} = -2.6$ for the 'min' case, which demonstrates strong degree of discordance between the data sets for the $\log \mathcal{C}$ statistic and decisive degree of discordance for the $\log \mathcal{I}$ statistic. This increase in the discordance between the data sets is consistent with the increased separation between the Planck and CFHTLenS marginalized posterior contours in the σ_8–$\Omega_{\rm m}$ plane in Fig. 12. For the 'mid-' case, we find $\log \mathcal{C} = 0.51$ and $\log \mathcal{I} = -0.28$, such that $\log \mathcal{I}$ is consistent with weak discordance, while $\log \mathcal{C}$ lies on the border between weak concordance and substantial concordance. This again seems to be consistent with the partial overlap in the CFHTLenS and Planck marginalized posterior contours in Fig. 12. For the 'max' case, we find $\log \mathcal{C} = 0.90$ and $\log \mathcal{I} = 0.62$, such that $\log \mathcal{I}$ is consistent with substantial concordance between the data sets, while $\log \mathcal{C}$ lies on the border between substantial concordance and strong concordance. Thus, even though the marginalized posterior contour for this case completely envelopes the Planck contour in Fig. 12, the degree of concordance between the data sets is not as impressive as naively expected prior to the execution of the $\log \mathcal{C}$ and $\log \mathcal{I}$ tests.

We can conclude that the question of data set concordance between CFHTLenS and Planck is sensitive to the exact details of the systematic uncertainties coming from intrinsic alignments, photometric redshift uncertainties, and baryonic uncertainties in the non-linear matter power spectrum. For our 'min' scenario there seems to strong-to-decisive discordance between the data sets, for the 'mid-' scenario there seems to be weak discordance to substantial concordance, and for the 'max' scenario there seems to be substantial concordance. These results are largely in agreement with the a priori expectation from the marginalized posterior contours in the σ_8–$\Omega_{\rm m}$ plane. For the three joint systematics cases, the 'min' case is the one most favoured by the cosmic shear data, but it is also the one that shows the greatest degree of discordance with Planck. The general agreement between the results from the $\log \mathcal{C}$ and $\log \mathcal{I}$ tests indicates that one may be able to compute the degree of concordance between data sets more easily from existing MCMC chains for parameter estimation, instead of embarking on new evidence calculations.

While it is more than plausible that CFHTLenS contains unaccounted systematics beyond those considered here (see Asgari et al. 2017), we note that Planck itself may suffer from internal discordance in its measurements, as pointed out in Addison et al. (2016, but disputed in Aghanim et al. 2016b). Moreover, the discordance between CFHTLenS and Planck CMB temperature bears resemblances to that between the Planck CMB temperature and Planck Sunyaev–Zel'dovich cluster counts (Ade et al. 2014XX, 2016c), where the latter is also a probe of the low-redshift universe (as compared to the CMB temperature) and exhibits a similar tension with the Planck CMB temperature in the σ_8–$\Omega_{\rm m}$ plane. Although the tension between the Planck observables can be reduced by allowing for a larger uncertainty in the mass bias estimates, the cluster count systematics seem to mainly cause shifts along the degeneracy direction and the fiducial cluster count constraint on $\sigma_8\Omega_{\rm m}^{0.3}$ is consistent with CFHTLenS.

4 CONCLUSIONS

We have revisited the analysis of the CFHTLenS data set with new cosmic shear measurements and covariance from extensive *N*-body simulations, along with a new CosmoMC fitting pipeline that accounts for key systematic uncertainties from intrinsic galaxy alignments, baryonic uncertainties in the non-linear matter power spectrum, and photometric redshift errors. Our data vector comprises seven tomographic bins covering redshifts up to $z = 3.5$ and seven angular bins extending to 120 arcmin. The covariance is constructed from a large suite of 497 *N*-body simulations, which were 2×2 sub-divided to gain a sufficient number of realizations, preventing our inverse covariance to be unduly biased by noise in the sample covariance estimator.

We used the new measurements and covariance to explore the consistency of CMB data from Planck with cosmic shear data from CFHTLenS, given increasing degrees of freedom from the systematic uncertainties. To this end, our CosmoMC pipeline calculates the cosmic shear likelihood and allows for 3 degrees of freedom for the IA, in the form of an amplitude along with a redshift and luminosity dependence. The pipeline further allows for one degree of freedom for the baryonic effects on non-linear scales in the matter power spectrum in HMCODE, which we have incorporated as a distinct CosmoMC module that internally communicates with our likelihood module. Lastly, the pipeline allows for 7 degrees of freedom for the photometric redshift uncertainties, which are manifested by shifts in each of the tomographic source distributions along the redshift axis, with either uniform or Gaussian priors. Thus, the pipeline allows for a total of 11 nuisance parameters, in addition to the cosmological parameters.

We first applied the pipeline to the measurements considering four different sets of cosmological priors, finding that the data are not sufficiently powerful to constrain the marginalized posterior contours in the σ_8–$\Omega_{\rm m}$ plane without prior-dependence of the results. However, the four cases show remarkable agreement along the axis perpendicular to the degeneracy direction, such that the 2σ tension with Planck effectively remains at the same level of significance regardless of the choice of priors. We proceeded with the lensing analysis by imposing external priors from local Hubble constant and BBN measurements. As these external priors are completely in agreement with Planck, any further discrepancy between Planck and CFHTLenS with external priors must be coming from CFHTLenS. The new marginalized posterior contours continue to show discrepancy with Planck at the 2σ level.

We then examined if the introduction of systematic degrees of freedom could alleviate the tension between the data sets, and whether any of the extensions are statistically preferred. To this end, we employed the DIC, which accounts for the Bayesian complexity of models. We find that a negative IA amplitude, such that $A = -3.6 \pm 1.6$, is preferred by the data at the level of $\Delta{\rm DIC} \simeq -5$ as compared to the fiducial model with no systematics included. However, this model seems to be at even greater tension with Planck, at the 3σ level. We find that an extension of the IA model to allow for redshift and luminosity dependence brings the relative tension between CFHTLenS and Planck back to its fiducial 2σ level, but this is because the redshift and luminosity dependence terms allow for values that diminish the IA signal. The three-parameter IA model is marginally disfavoured by the data at the level of $\Delta{\rm DIC} = 2.4$ with informative priors on the IA parameters and more strongly disfavoured at $\Delta{\rm DIC} = 12$ with non-informative priors on the parameters.

Next, we did not find a preference for non-linear baryonic physics in the CFHTLenS data, as the extension to allow for a varying amplitude in HMCODE is only favoured at $\Delta{\rm DIC} = -0.64$ when considering non-informative priors and disfavoured at $\Delta{\rm DIC} = 1.4$ when considering informative priors. Allowing for the HMCODE feedback amplitude to account for non-linear baryonic physics has a marginal impact on the tension between CFHTLenS and Planck. We moreover allowed for photometric redshift uncertainties by imposing uniform priors of $|\Delta z_i| = 0.1$ for each tomographic bin. Allowing for deviations around the fiducial redshift distributions produces an improvement in the goodness of fit, at the level of $\Delta\chi^2_{\rm eff} = -6.0$, but we find that introducing the new degrees of freedom is not preferred by the data at $\Delta{\rm DIC} = 0.97$. We considered a case where the redshift perturbations are obtained from the cross-correlation analysis of Choi et al. (2016), but found that these redshifts are even more disfavoured, at the level of $\Delta{\rm DIC} = 11$. As for the tension between CFHTLenS and Planck, the photometric redshift uncertainties only have a marginal impact.

Thus, when introducing the systematic uncertainties independently, only the IA amplitude is substantially preferred by the CFHTLenS data. However, the negative amplitude may be considered unphysical, likely caused by overly simplistic IA modelling and/or unaccounted systematics and it only increases the tension between the Planck and CFHTLenS data sets. Aside from the question of whether the systematic degrees of freedom are preferred by the data, we also find no strong relief in the tension between the two data sets when independently allowing for baryonic physics on non-linear scales, photometric redshift uncertainties, and non-minimal extensions to the IA model.

We moreover considered three distinct cases for the joint account of the systematic uncertainties (detailed in Table 4). The first of these cases is the 'min' case where we impose an informative prior on the IA amplitude (excluding a possible luminosity or redshift dependence), an informative prior on the HMCODE feedback amplitude, and informative priors on the source redshift distributions given by Choi et al. (2016). The second case is the 'mid-' case, where we impose informative priors on the IA amplitude, redshift, and luminosity dependence. For this case, we continue to impose an informative prior on the HMCODE feedback amplitude. We also continue to use the Choi et al. (2016) shifts in the fiducial redshift distributions, but with error bars given by $|\Delta z_i| = 0.05$. Lastly, the third case is the 'max' case, where we impose non-informative priors on all three IA parameters, a non-informative prior on the HMCODE feedback amplitude and uniform priors on the redshift distributions given by $|\Delta z_i| = 0.1$. For the 'max' case, we also only keep 'large' scales, by which we remove all ξ_- measurements and only keep ξ_+ measurements for $\theta > 17$ arcmin.

For the 'min' case, the main imprint comes from varying the IA amplitude, which increases the tension with Planck, analogously to when the parameter is considered independently. The goodness of fit significantly improves with $\Delta\chi^2_{\rm eff} = -10$ and the Bayes factor favours the 'min' case by $2\ln\mathcal{B}_{01} = -6.6$ (as compared to the fiducial cosmological model with no systematic uncertainties). Despite the increased complexity, manifested in $\Delta{\rm DIC} = 1.6$, the evidence is sufficiently improved to warrant the 'min' case as more favoured as compared to the vanilla ΛCDM model. For the 'mid-' case, the marginalized posterior contour shows strong overlap with that of the fiducial case, as the luminosity dependence diminishes the IA signal, while the baryonic and photometric redshift errors each contribute to a marginal increase in the area of the posterior contour, bringing CFHTLenS into greater concordance with Planck. We find a marginal improvement in the goodness of fit, given by $\Delta\chi^2_{\rm eff} = -1.1$. However, the Bayes factor disfavours the 'mid-' case by $2\ln\mathcal{B}_{01} = 1.3$ and the information criterion disfavours it more strongly by $\Delta{\rm DIC} = 12$.

The largest impact on the cosmological constraints comes from the 'max' case, where the cutting of angular scales and non-informative priors on the 11 systematic parameters results in cosmological constraints so weak that the marginalized posterior contour for CFHTLenS completely envelopes the Planck contour. For this case, we find a significant improvement in the best fit, at $\Delta\chi^2_{\rm eff} = -25$, while the Bayes factor shows a marginal improvement of $2\ln\mathcal{B}_{01} = -0.52$ (as compared to a fiducial cosmological model using the downsized data vector). However, the increased complexity of the model renders it highly disfavoured at $\Delta{\rm DIC} = 19$, which highlights the usefulness of considering multiple statistical tools for purposes of model selection.

In more carefully assessing the degree of concordance or discordance between CFHTLenS and Planck, we further employed 'data concordance tests' as quantified by $\log\mathcal{C}$ and $\log\mathcal{I}$ (defined in Section 2.1.7), grounded in the Bayesian evidence and DIC, respectively. With these statistical tools, we find strong-to-decisive discordance between the two data sets for the 'min' case, as evidenced by $\log\mathcal{C} = -1.8$ and $\log\mathcal{I} = -2.6$, respectively. We further find weak discordance to substantial concordance for the 'mid-' case, as evidenced by $\log\mathcal{I} = -0.28$ and $\log\mathcal{C} = 0.51$, while there is substantial concordance for the 'max' case, as evidenced by $\log\mathcal{C} = 0.90$ and $\log\mathcal{I} = 0.62$. The outcome of these concordance tests generally agree with the a priori expectation from the marginalized posterior contours, although the degree of concordance is weaker for the 'max' case than naively expected. For the three joint systematics cases, it is interesting to note that the case most discordant with Planck is also the one most favoured by the data. We can also conclude that the results from the $\log\mathcal{C}$ and $\log\mathcal{I}$ data concordance tests generally agree with one another, indicating that either one may be used to assess the degree of concordance between data sets in future analyses.

Our new measurements and fitting pipeline are publicly available at the address https://github.com/sjoudaki/cfhtlens_revisited. We have extended the pipeline to account for joint analyses of cosmic shear tomography, galaxy-galaxy lensing, and redshift space distortion measurements, including the full covariance for overlapping surveys. We plan to release this extended pipeline as part of an upcoming paper to constrain modified gravity and neutrino physics (Joudaki et al., in preparation).

ACKNOWLEDGEMENTS

We much appreciate useful discussions with Jason Dossett, Marian Douspis, Farhan Feroz, Will Handley, Manoj Kaplinghat, Antony Lewis, Niall MacCrann, Gregory Martinez, and Marco Raveri. We greatly thank Luke Hodkinson for help with parallelizing our code. We also greatly thank Robin Humble and Jarrod Hurley for HPC support. We acknowledge the use of ASTAC time on Swinburne's swinSTAR and NCI's Raijin machines, with special thanks to Amr Hassan and Jarrod Hurley for these resources. We acknowledge the use of CAMB and CosmoMC packages (Lewis & Bridle 2002; Lewis et al. 2000). CB acknowledges the support of the Australian Research Council through the award of a Future Fellowship. CH and AC acknowledge support from the European Research Council under grant number 240185 and 647112. BJ acknowledges support by an STFC Ernest Rutherford Fellowship, grant reference ST/J004421/1. HH is supported by an Emmy Noether grant (No. Hi 1495/2-1) of the Deutsche Forschungsgemeinschaft. MV acknowledges support from the European Research Council under FP7 grant number 279396 and the Netherlands Organisation for Scientific Research (NWO) through grants 614.001.103. Computations for the *N*-body simulations were performed on the GPC supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund – Research Excellence; and the University of Toronto.

REFERENCES

Abbott T. et al., 2016, Phys. Rev., D94, 022001
Addison G. E., Huang Y., Watts D. J., Bennett C. L., Halpern M., Hinshaw G., Weiland J. L., 2016, ApJ, 818, 132
Aghanim N. et al., 2016a, A&A, 594, A11
Aghanim N. et al., 2016b, preprint (arXiv:1608.02487)
Akaike H., 1974, IEEE Trans. Autom. Control, AC-19, 716
Albrecht A. et al., 2006, preprint (arXiv:0609591)
Alimi J.-M., Füzfa A., Boucher V., Rasera Y., Courtin J., Corasaniti P.-S., 2010, MNRAS, 401, 775
Anderson L. et al., 2014, MNRAS, 441, 24
Asgari M., Schneider P., 2015, A&A, 578, A50
Asgari M., Schneider P., Simon P., 2012, A&A, 542, A122
Asgari M., Heymans C., Blake C., Harnois-Deraps J., Schneider P., Van Waerbeke L., 2017, MNRAS, 464, 1676
Baldi M., Villaescusa-Navarro F., Viel M., Puchwein E., Springel V., Moscardini L., 2014, MNRAS, 440, 75
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Benítez N., 2000, ApJ, 536, 571
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Bernardeau F., Bonvin C., Vernizzi F., 2010, Phys. Rev. D, 81, 083002
Bernstein G., Huterer D., 2010, MNRAS, 401, 1399
Bernstein G. M., Armstrong R., Krawiec C., March M. C., 2016, MNRAS, 459, 4467
Bertone G., Hooper D., Silk J., 2005, Phys. Rep., 405, 279
Betoule M. et al., 2014, A&A, 568, A22
Bird S., Viel M., Haehnelt M. G., 2012, MNRAS, 420, 2551
Bonnett C. et al., 2016, Phys. Rev. D, 94, 042005
Bridle S., King L., 2007, New J. Phys., 9, 444
Burles S., Nollett K. M., Turner M. S., 2001, Phys. Rev. D, 63, 063512
Choi A. et al., 2016, MNRAS, 463, 3737
Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Phys. Rep., 513, 1
Cooray A., Hu W., 2002, ApJ, 574, 19
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Copeland E. J., Sami M., Tsujikawa S., 2006, Int. J. Mod. Phys. D, 15, 1753
Cousins R. D., 2014, Synthese, preprint (arXiv:1310.3791)
Cyburt R. H., Fields B. D., Olive K. A., Yeh T.-H., 2016, Rev. Mod. Phys, 88, 015004
Dodelson S., Shapiro C., White M., 2006, Phys. Rev. D, 73, 023009
Efstathiou G., 2014, MNRAS, 440, 1138
Eifler T., Schneider P., Hartlap J., 2009, A&A, 502, 721
Erben T. et al., 2013, MNRAS, 433, 2545
Feng J. L., 2010, ARA&A, 48, 495
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint (arXiv:1306.2144)
Friedrich O., Seitz S., Eifler T. F., Gruen D., 2016, MNRAS, 456, 2662

Gelman A., Rubin D., 1992, Stat. Sci., 7, 457

Grandis S., Seehars S., Refregier A., Amara A., Nicola A., 2016, J. Cosmol. Astropart. Phys., 5, 034

Hammami A., Llinares C., Mota D. F., Winther H. A., 2015, MNRAS, 449, 3635

Handley W. J., Hobson M. P., Lasenby A. N., 2015a, MNRAS, 450, L61

Handley W. J., Hobson M. P., Lasenby A. N., 2015b, MNRAS, 453, 4384

Harnois-Déraps J., van Waerbeke L., 2015, MNRAS, 450, 2857

Harnois-Déraps J., Pen U.-L., Iliev I. T., Merz H., Emberson J. D., Desjacques V., 2013, MNRAS, 436, 540

Harnois-Déraps J., van Waerbeke L., Viola M., Heymans C., 2015, MNRAS, 450, 1212

Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399

Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, ApJ, 780, 111

Heymans C. et al., 2012, MNRAS, 427, 146

Heymans C. et al., 2013, MNRAS, 432, 2433 (H13)

Hildebrandt H. et al., 2012, MNRAS, 421, 2355

Hirata C., Seljak U., 2003, MNRAS, 343, 459

Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526

Huterer D., Takada M., Bernstein G., Jain B., 2006, MNRAS, 366, 101

Jeffreys H., 1961, Theory of Probability, 3rd edn. Oxford Univ. Press, Oxford

Joachimi B., Mandelbaum R., Abdalla F. B., Bridle S. L., 2011, A&A, 527, A26

Joachimi B. et al., 2015, Space Sci. Rev., 193, 1

Joudaki S., 2013, Phys. Rev. D, 87, 083523

Joudaki S., Kaplinghat M., 2012, Phys. Rev. D, 86, 023526

Joudaki S., Cooray A., Holz D. E., 2009, Phys. Rev. D, 80, 023003

Kass R. E., Raftery A. E., 1995, J. Am. Stat. Assoc., 90, 773

Kaufman G. M., 1967, Some Bayesian Moment Formulae, Report No. 6710, Centre for Operations Research and Econometrics, Catholic University of Louvain, Heverlee

Kilbinger M. et al., 2013, MNRAS, 430, 2200

Kilbinger M., Bonnett C., Coupon J., 2014, Astrophysics Source Code Library, record ascl:1402.026

Köhlinger F., Viola M., Valkenburg W., Joachimi B., Hoekstra H., Kuijken K., 2016, MNRAS, 456, 1508

Krause E., Hirata C. M., 2010, A&A, 523, A28

Kullback S., Leibler R. A., 1951, Ann. Math. Stat., 22, 79

Kunz M., Trotta R., Parkinson D. R., 2006, Phys. Rev. D, 74, 023503

Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Liddle A. R., 2007, MNRAS, 377, L74

Limber D. N., 1954, ApJ, 119, 655

Lindley D. V., 1957, Biometrika, 44, 187

Loverde M., Afshordi N., 2008, Phys. Rev. D, 78, 123506

McDonald P., Trac H., Contaldi C., 2006, MNRAS, 366, 547

McQuinn M., White M., 2016, J. Cosmol. Astropart. Phys., 1, 043

MacCrann N., Zuntz J., Bridle S., Jain B., Becker M. R., 2015, MNRAS, 451, 2877

Ma Z., Hu W., Huterer D., 2006, ApJ, 636, 21

Marshall P., Rajguru N., Slosar A., 2006, Phys. Rev. D, 73, 067302

Mead A., 2015, Astrophysics Source Code Library, record ascl:1508.001

Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958

Melchior P., Viola M., 2012, MNRAS, 424, 2757

Miller L. et al., 2013, MNRAS, 429, 2858

Olive K. A., Particle Data Group, 2014, Chin. Phys. C, 38, 090001

Planck Collaboration XVI, 2014a, A&A, 571, A16

Planck Collaboration XX, 2014b, A&A, 571, A20

Planck Collaboration XIII, 2016a, A&A, 594, A13

Planck Collaboration XIV, 2016b, A&A, 594, A14

Planck Collaboration XXIV, 2016c, A&A, 594, A24

Raveri M., 2016, Phys. Rev. D, 93, 043522

Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19

Saito S., Takada M., Taruya A., 2008, Phys. Rev. Lett., 100, 191301

Schaye J. et al., 2010, MNRAS, 402, 1536

Schneider P., Eifler T., Krause E., 2010, A&A, 520, A116

Schwarz G., 1978, Ann. Stat., 6, 461

Seehars S., Grandis S., Amara A., Refregier A., 2016, Phys. Rev. D, 93, 103507

Seitz C., Schneider P., 1997, A&A, 318, 687

Sellentin E., Heavens A. F., 2016, MNRAS, 456, L132

Shapiro C., 2009, ApJ, 696, 775

Shapiro C., Cooray A., 2006, J. Cosmol. Astropart. Phys., 3, 7

Singh S., Mandelbaum R., More S., 2015, MNRAS, 450, 2195

Smith R. E. et al., 2003, MNRAS, 341, 1311

Spiegelhalter D., Best N. G., Carlin B. P., 2002, J. R. Stat. Soc. B, 64, 583

Spiegelhalter D., Best N. G., Carlin B. P., van der Linde A., 2014, J. R. Stat. Soc. B, 76, 485

Stabenau H. F., Jain B., 2006, Phys. Rev. D, 74, 084007

Takada M., Hu W., 2013, Phys. Rev. D, 87, 123504

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152

Trotta R., 2008, Contemp. Phys., 49, 71

Troxel M. A., Ishak M., 2015, Phys. Rep., 558, 1

van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649

Wagner C., Verde L., Jimenez R., 2012, ApJ, 752, L31

White M., 2004, Astropart. Phys., 22, 211

Zhan H., Knox L., 2004, ApJ, 616, L75

Zhao G.-B., Li B., Koyama K., 2011, Phys. Rev. D, 83, 044007

Zuntz J. et al., 2015, Astron. Comput., 12, 45

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