Testing extensions to ΛCDM on small scales with forthcoming cosmic shear surveys

Sam G. Stafford [1]★ Ian G. McCarthy [1]★ Juliana Kwan,[1] Shaun T. Brown [1] Andreea S. Font [1] and Andrew Robertson [2]

[1]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[2]*Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*

ABSTRACT
We investigate the constraining power of forthcoming Stage-IV weak lensing surveys (Euclid, LSST, and *NGRST*) for extensions to the Lambda cold dark matter model on small scales, via their impact on the cosmic shear power spectrum. We use high-resolution cosmological simulations to calculate how warm dark matter (WDM), self-interacting dark matter (SIDM), and a running of the spectral index affect the non-linear matter power spectrum, $P(k)$, as a function of scale and redshift. We evaluate the cosmological constraining power using synthetic weak lensing observations derived from these power spectra and that take into account the anticipated source densities, shape noise, and cosmic variance errors of upcoming surveys. We show that upcoming Stage-IV surveys will be able to place useful, independent constraints on both WDM models (ruling out models with a particle mass of $\lesssim 0.5$ keV) and SIDM models (ruling out models with a velocity-independent cross-section of $\gtrsim 10$ cm^2 g^{-1}) through their effects on the small-scale cosmic shear power spectrum. Similarly, they will be able to strongly constrain cosmologies with a running spectral index. Finally, we explore the error associated with the cosmic shear cross-spectrum between tomographic bins, finding that it can be significantly affected by Poisson noise (the standard assumption is that the Poisson noise cancels between tomographic bins). We provide a new analytic form for the error on the cross-spectrum that accurately captures this effect.

Key words: gravitational lensing: weak – cosmological parameters – dark matter.

1 INTRODUCTION

In the current concordance cosmological framework, termed the ΛCDM (Lambda cold dark matter) model, structure in the Universe forms hierarchically. The initial density perturbations laid down by inflation eventually grow large enough that they become gravitationally unstable and collapse to form low-mass 'haloes'. These low-mass haloes merge to build up progressively larger systems, eventually culminating in the large-scale structure (LSS) that we observe today (see e.g. Davis et al. 1985). This theoretical picture has been immensely powerful in describing observations of our Universe, accurately reproducing the observed properties of the cosmic microwave background (see e.g. the recent results from Planck Collaboration VI 2020) as well as low-redshift probes such as baryon acoustic oscillations (Eisenstein et al. 2005) and redshift-space distortions (see e.g. Alam et al. 2017). Given its success, this theoretical framework has come to be known as the 'standard model of cosmology'.

It is noteworthy that, while the ΛCDM model accurately describes many observables on large scales, there have been a number of recent mild tensions reported in the best-fitting parameter values for certain cosmological parameters, including the Hubble constant (see Verde, Treu & Riess 2019 for a review) and the LSS parameter

$S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3}$ (see e.g. discussion in McCarthy et al. 2018) as derived from independent measurements. Whether these tensions are signalling the presence of unaccounted for systematic errors in some of the cosmological analyses, that there is new physics (i.e. beyond the standard model), or that they are merely the result of statistical fluctuations, is presently unclear and the subject of much investigation.

Although the ΛCDM model does remarkably well at describing observations of our Universe on large scales (tens to hundreds of Mpc, modulo the recent mild tensions described above), the last decade or so has seen a ramping up of detailed comparisons on smaller scales (typically kpc to Mpc), some of which have reported significant tensions with the predictions of ΛCDM-based cosmological simulations. Three of the most widely discussed tensions are the 'cusp-core' problem (Flores & Primack 1994; Moore 1994), the 'missing satellites' problem (Klypin et al. 1999; Moore et al. 1999) and the 'too-big-to-fail' problem (Boylan-Kolchin, Bullock & Kaplinghat 2011). The common thread between these tensions is that the ΛCDM model appears to predict too much structure (and too high densities) on small scales compared with what is inferred from observations. These tensions are often evaluated with respect to gravity-only calculations in the context of ΛCDM, but recent work has suggested that neglecting important baryonic physics may have significant implications for these tensions. For example, including processes such as reionization, star formation, stellar feedback through supernovae and winds, and feedback associated with active galactic nuclei in the models have been shown to help alleviate some

of these problems with ΛCDM on small scales (see e.g. Efstathiou 1992; Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002; Mashchenko, Wadsley & Couchman 2008; Pontzen & Governato 2012; Sawala et al. 2016; Wetzel et al. 2016).

Nevertheless, it is worthwhile to consider other possibilities, which may act in conjunction with baryon physics, for these small-scale problems, not least because the modelling of baryon physics on small scales is quite complex and requires a degree of fine tuning to resolve some of the aforementioned tensions. Also, it is important that we continue to test the standard model as rigorously as possible on small scales, in order to shed light on the nature of dark matter.

The apparent deficit of small-scale structure has led to the study and development of many extensions to the ΛCDM model that aim to reconcile these differences. Promising extensions to the current ΛCDM paradigm, such as warm dark matter (WDM) and self-interacting dark matter (SIDM), can reduce the formation of structure on small scales. These different cosmological models focus on changing one of the assumed aspects of dark matter. In the case of the former, dark matter decouples in the early universe whilst still relativistic leading to non-negligible thermal velocities and free-streaming, while in the latter case dark matter is allowed to have strong self-interactions (scattering). Both of these extensions have been shown to have success in alleviating the outlined challenges which exist with ΛCDM (see e.g. Colín, Avila-Reese & Valenzuela 2000; Lovell et al. 2012 for the case of WDM and Spergel & Steinhardt 2000; Zavala, Vogelsberger & Walker 2013 for the case of SIDM). One further extension that has not been studied quite so extensively is a running scalar spectral index. A running spectral index is different from the previous two extensions as it does not alter the nature of dark matter but instead changes the initial conditions for structure formation, which is motivated by certain classes of inflation models. Previously this has also been demonstrated to be a promising candidate for reducing the formation of small-scale structure (see e.g. Garrison-Kimmel et al. 2014; Stafford et al. 2020b).

In a recent paper (Stafford et al. 2020b), we compared how effective these three different extensions were in altering various small-scale structure statistics. To do this, we used gravity-only cosmological simulations with the values adopted for the additional parameters associated with each extension being guided by current observational constraints. We found that all of the models can have similar effects on certain statistics, such as the abundance of satellite galaxies inside hosts (which is one of the primary tests of ΛCDM on small scales). As such, it is of interest to explore new observables that could help differentiate these models and the effects they have on structure formation.

In this study, we exploit the fact that these different models alter the non-linear matter power spectrum along with its redshift evolution in different ways (as we will show). Consequently, observational probes which are directly sensitive to the non-linear matter power spectrum may provide a strong test of these extensions and of ΛCDM. One potentially promising observable is the cosmic shear power spectrum. As the light emitted from distant galaxies travels towards us on Earth, its path becomes distorted due to intervening matter, a phenomenon known as gravitational lensing. One can use the correlated effect this has on galaxy shapes to extract information about the non-linear matter power spectrum. This can be done either through the two-point autocorrelation function of galaxy shapes or its Fourier analogue, the cosmic shear power spectrum. It is this latter statistic that we examine in this paper and how it is affected by SIDM, WDM and a running spectral index in comparison with the ΛCDM result.

The impetus for this study stems from the increasing quantity and quality of cosmic shear observations being made by current Stage-III surveys, such as the Dark Energy Survey (DES)[1] (DES Collaboration 2021), the Hyper Suprime-Cam Subaru Strategic Survey (HSC)[2] (Hamana et al. 2020), and the Kilo-Degree Survey (KiDS)[3] (Asgari et al. 2021), and forthcoming Stage-IV cosmic shear surveys, such as Euclid[4] (Euclid Collaboration 2020a), the Rubin Observatory Legacy Survey of Space and Time (LSST)[5] (Zhan & Tyson 2018), and the *Nancy Grace Roman Space Telescope* (*NGRST*)[6] (Spergel et al. 2015). Stage-IV surveys will greatly improve on current observations, by covering a much larger area of the sky and/or being significantly deeper, ultimately resulting in greatly improved measurements of the cosmic shear power spectrum.

To date, most of the forecasting work for Stage-IV surveys has been with regards to anticipated constraints on the standard model of cosmology and extensions that affect LSS (such as evolving dark energy and massive neutrinos). Only a small number of studies have been dedicated to the possible constraints that may be obtained for small-scale extensions. For example, Markovic et al. (2011) explored how future cosmic shear surveys can be used to place constraints on the mass of thermal relic particles. More recently, Hubert et al. (2021) explored constraints on decaying dark matter models. This paper seeks to examine the impact of the three cosmological extensions described above (WDM, SIDM, and a running scalar spectral index) on the cosmic shear power spectrum using the non-linear matter power spectrum extracted from numerical simulations. We then explore the prospect of these upcoming Stage-IV surveys in differentiating these extensions from ΛCDM.

The paper is structured as follows. In Section 2, we describe in more detail the extensions studied in this paper. We also discuss the numerical simulations that we use as well as how we create a non-linear matter power spectrum which covers the full dynamic range of interest for this weak lensing study. In Section 3, we discuss how we compute the cosmic shear power spectrum and the associated uncertainties. In Section 4, we present our results for the cosmic shear auto- and cross-correlation power spectra for each cosmological model and finally in Section 5 we discuss and summarize our results.

2 COSMOLOGICAL EXTENSIONS

In this study, we analyse three separate cosmological extensions, a running scalar spectral index, WDM, and SIDM. We briefly describe these in the following subsections and refer the interested reader to Stafford et al. (2020b) for further details. Note, however, in Stafford et al. (2020b) we only examined two WDM models and two SIDM models, whereas in this study we extend our analysis slightly and look at an additional, more extreme, case for both of these extensions.

2.1 Extensions

2.1.1 *Running of the scalar spectral index*

In the standard model of cosmology, the power spectrum for scalar perturbations generated by inflation is assumed to follow a simple

---

[1]https://www.darkenergysurvey.org/
[2]https://hsc.mtk.nao.ac.jp/ssp/survey/
[3]http://kids.strw.leidenuniv.nl/overview.php
[4]https://www.euclid-ec.org/?page_id = 2581
[5]https://www.lsst.org/scientists/
[6]https://www.jpl.nasa.gov/missions/the-nancy-grace-roman-space-telescope

power law of the form (Guth 1981; Kosowsky & Turner 1995):

$$P(k) = Ak^{n_s}, \qquad (1)$$

where A is the amplitude of the primordial matter power spectrum, k is the wavenumber, and n_s is the spectral index. A first-order extension is that the spectral index has some level of scale dependence, which results in a modification to the functional form of the primordial matter power spectrum (Kosowsky & Turner 1995):

$$P(k) = A_s(k_{pivot}) \left(\frac{k}{k_{pivot}} \right)^{n_s(k_{pivot}) + \frac{\alpha_s}{2} \ln\left(\frac{k}{k_{pivot}} \right)}. \qquad (2)$$

Here, α_s is termed the 'running' of the scalar spectral index and is defined as $dn_s(k)/d\ln(k)$.[7] Here, $n_s(k)$ is still equal to the logarithmic slope of the power spectrum, however, it now has a k-dependence due to the $(\alpha_s/2)\ln(k/k_{pivot})$ term in equation (2). The value adopted for k_{pivot} here corresponds to that for the Planck satellite mission, $k_{pivot} = 0.05$ Mpc^{-1}, and corresponds to the k-scale at which values for A_s and n_s are quoted. In this work, we explore the effects that a positive value for the running (with $\alpha_s = 0.007\,91$) and a negative value for the running (with $\alpha_s = -0.024\,73$) have on the non-linear matter power spectrum and how these propagate through to an observable impact on the cosmic shear power spectrum. Note the cosmological parameter values chosen for the running simulations are in the context of a Planck 2015 cosmology; the maximum-likelihood values are computed from the Markov chains which includes α_s as a free cosmological parameter in the analysis. This results in slight differences to the quoted values of the other cosmological parameters in Section 2.2 (a table of all of the parameter values for each cosmology can be found in Stafford et al. 2020b). The values chosen for α_s are discussed in detail in Stafford et al. (2020a), but in summary represent the $\pm 2\sigma$ values of the posterior distribution extracted from the Planck 2015 Markov chains[8] (Planck Collaboration XVI 2014).

These two cosmological models with a running scalar spectral index were chosen in Stafford et al. (2020a), prior to the release of the Planck 2018 cosmological parameter constraints (Planck Collaboration VI 2020). The updated CMB constraints on α_s are tightened somewhat to $\alpha_s = -0.0045 \pm 0.0067$ (68 per cent CL TT, TE, EE + lowE + lensing) making the posterior distribution for α_s used in Stafford et al. (2020a) slightly out of date. However, it is worth noting that, even with the updated CMB constraints from the Planck team, a mildly negative value for the running of the spectral index is not ruled out (see the discussion in section 7.2.1 in Planck Collaboration VI 2020). Furthermore, updated constraints from the Lyman α forest find a $\approx 3\sigma$ detection for a negative α_s with $\alpha_s = -0.010 \pm 0.003$ (Palanque-Delabrouille et al. 2020).[9] Therefore, the values we simulate are still compatible with current observations, even if they lie on the more extreme end of current constraints. Furthermore, as discussed later on in the paper in the context of WDM models, this work seeks to answer the question whether cosmic shear can be used as a complementary probe to place

independent constraints on these additional cosmological parameters (α_s, M_{WDM}, σ/m).

2.1.2 Warm dark matter

Another assumption of the standard model of cosmology is that dark matter decoupled from the primordial plasma after it became non-relativistic. This results in negligible thermal velocities at early times (hence 'cold' dark matter) and would be expected if dark matter was composed of particles with masses in the GeV range (or larger), such as the currently favoured candidate, the WIMP (weakly interacting massive particle). If, however, dark matter is made up of lighter particles with masses in the keV range, such as thermal relic sterile neutrinos, the dark matter particles decouple while still relativistic. This type of model is referred to as a thermal relic WDM model. The resulting larger thermal velocities that these dark matter particles have (compared to CDM) at early times allows them to free-stream out of density perturbations. The free-streaming of dark matter particles works to suppress the growth of structure on small scales (Bond & Szalay 1983; Bardeen et al. 1986).

The suppression of small-scale density perturbations due to the thermal velocity associated with the dark matter particles leads to a characteristic cut-off in the WDM power spectrum below a k mode corresponding to the free-streaming scale. In this work, we model the suppression of the initial linear matter power spectrum as a transfer function relative to the corresponding cold dark matter power spectrum:

$$P_{WDM}(k) = T^2_{WDM}(k) P_{CDM}(k). \qquad (3)$$

We compute T_{WDM} using the fitting formula developed in Bode, Ostriker & Turok (2001):

$$T_{WDM}(k) = \left[1 + (\alpha k)^{2\nu} \right]^{-5/\nu}, \qquad (4)$$

where ν is a fitting constant and α dictates the scale of the cut-off in the power spectrum, with this being dependent on the mass of the thermal relic particle. We follow Viel et al. (2005), adopting $\nu = 1.12$ and (assuming the WDM is composed of thermal relics) and computing α as

$$\alpha = 0.049 \left(\frac{M_{WDM}}{1\,keV} \right)^{-1.11} \left(\frac{\Omega_{WDM}}{0.25} \right)^{0.11} \left(\frac{h}{0.7} \right)^{1.22} \text{Mpc } h^{-1}, \quad (5)$$

where M_{WDM} corresponds to the mass of the WDM particle, Ω_{WDM} is the present-day density of WDM[10] in units of the critical density, and h is the reduced Hubble's constant.

It is evident from equation (5) that for lighter WDM particles α increases, pushing the scale of the cut-off in the matter power spectrum to smaller k modes (larger physical scales). In this study, we examine $M_{WDM} = (0.5, 2.5, 5.0)$ keV. Our choice for the parameter values investigated in this study aims to bracket the current observational constraints placed on the mass of a WDM particle from different probes such as the Lyman α forest (Viel et al. 2013; Iršič et al. 2017), the Milky-Way's satellite population (Kennedy et al. 2014; Lovell et al. 2014; Jethwa, Erkal & Belokurov 2018; Nadler et al. 2019; Nadler et al. 2020) and time-delay measurements of strongly gravitationally lensed quasars (Gilman et al. 2019; Hsueh et al. 2019). Although the 0.5 keV model is perhaps currently in tension with constraints from the previous observations, we examine it here to see if cosmic shear, as an independent test with very different

---

[7]Although in the ΛCDM model α_s is assumed to be zero, virtually all models of inflation predict some level of scale dependence for n_s. However, the simplest single field slow-roll inflation models predict that this scale dependence should only be of the order 10^{-3} (Kosowsky & Turner 1995).
[8]http://pla.esac.esa.int/pla/#home
[9]Note, however, that Palanque-Delabrouille et al. (2020) marginalize over neutrino mass at the same time as α_s, whereas the Planck analysis fixed ΣM_ν to 0.06 eV when constraining α_s.

[10]We assume all of the dark matter is in the form of WDM and so $\Omega_{WDM} = \Omega_{CDM}$.

systematics compared to previous methods, can also rule out such a model.

2.1.3 Self-interacting dark matter

In terms of the possible interactions that dark matter particles can experience, the standard model of cosmology adopts the simplest assumption that dark matter interacts only via gravity and is therefore 'collisionless'. The final extension that we investigate is a relaxation of this assumption, allowing the dark matter particles to have strong self-interactions. SIDM was proposed to alleviate the 'cusp–core problem' (Flores & Primack 1994) by Spergel & Steinhardt (2000) and has indeed been shown in the literature to produce strong cores inside dark matter haloes and can have particularly large effects on satellite subhaloes, strongly reducing their masses (see e.g. Peñarrubia et al. 2010; Vogelsberger, Zavala & Loeb 2012; Dooley et al. 2016). We explore how these effects translate through to the non-linear matter power spectrum.

To study the effect of strong self-interactions, we use a version of the GADGET3 *N*-body code (discussed in Section 2.2) that was modified in Robertson et al. (2019) to include dark matter self-interactions, which are assumed to be elastic. The additional parameter this adds to the standard six free parameters of the ΛCDM model is (σ/m), which is the cross-section for interaction of dark matter particles. In this study, the values that we explore are $(\sigma/m) = (0.1, 1.0, 10.0)$ cm^2 g^{-1}. Again, our choices for the values adopted for this cosmological parameter are guided by the current observational constraints. In the case of the self-interaction cross-section, these constraints are placed by probes such as perturbations in strong lensing arcs (Meneghetti et al. 2001; Robertson et al. 2019), dark matter-galaxy off-sets in colliding galaxy clusters (Randall et al. 2008; Harvey et al. 2015; Kahlhoefer et al. 2015; Kim, Peter & Wittman 2017; Robertson, Massey & Eke 2017; Wittman, Golovich & Dawson 2018), cluster shapes (Miralda-Escude 2002; Peter et al. 2013), group shapes (Sagunski et al. 2021), and also from subhalo evaporation arguments (Gnedin & Ostriker 2001). Furthermore, Banerjee et al. (2020) showed that constraints of the order $\sigma/m \lesssim 2$ cm^2 g^{-1} can be placed on the cross-section for self-interaction when combining observations of the distribution of subhaloes and weak lensing measured density profiles.

The largest cross-section we examine is somewhat in tension with some of the observations listed here, but it is worth noting that these are primarily focused on large scales (galaxy clusters). There is evidence to suggest that SIDM models with cross-sections as high as 50 cm^2 g^{-1} are viable when using constraints coming from dwarf galaxies (Elbert et al. 2015; Correa 2021). As our focus is on small scales, the range of cross-sections we explore is therefore plausible given current constraints. Note, although the simulation code developed by Robertson et al. (2019) has the functionality for scattering events to be both angular and velocity dependent, for simplicity we focus on the case where scattering events are velocity-independent and isotropic.

2.2 Simulations

The simulations used in this study are those first introduced in Stafford et al. (2020b). All of the simulations follow dissipationless physics only, meaning they follow either just the gravitational evolution (for CDM, WDM, and the running spectral index models) or the gravitational and self-scattering (SIDM) evolution. The impact of physical processes associated with the baryonic component (i.e.

radiative cooling, star formation, feedback processes) is not included but is discussed further in Section 5.

The simulations are run with the GADGET3 code (last described in Springel 2005). Each simulation is 25 comoving h^{-1} Mpc on a side and contains 1024^3 dark matter particles. We adopt a fixed physical gravitational softening length of 250 h^{-1} pc at $z \leq 3$, with this being a fixed comoving length at higher redshifts. The particle mass for all of the simulations, with the exception of the two with a running spectral index, is $m_{DM} = 1.266 \times 10^6$ M$_\odot$ h^{-1}. For the two 'running' simulations, the particle mass is slightly different, owing to the slightly different values adopted for Ω_m and h (see table 1 in Stafford et al. 2020b). Alongside these main simulations, we also use a secondary suite of simulations of varying box size and resolution. These include a suite of simulations which are 400 comoving h^{-1} Mpc on a side, as well as a suite which are 100 comoving h^{-1} Mpc on a side. Both suites contain 1024^3 collisionless particles. Note, however, that we only run the cosmologies with a running scalar spectral index (as well as the reference ΛCDM cosmology) in these larger volumes, as they are required for the process of creating a spliced power spectrum (see Section 2.3.1).

The initial conditions (ICs) for the simulations are generated using a modified version of the N-GENIC[11] code (Springel et al. 2005) that was modified to include second-order Lagrangian Perturbation Theory corrections. The ICs are generated at a starting redshift of $z = 127$ with each simulation being initialized with the same random phases. The input linear theory matter power spectrum and transfer functions are computed with the Boltzmann code CAMB (Lewis, Challinor & Lasenby 2000, August 2018 version). The initial conditions and the background expansion rate in the simulations are computed assuming the Planck 2015 maximum-likelihood cosmological parameters (Planck Collaboration XVI 2014) ($H_0 = 67.31$ km s^{-1} Mpc^{-1}; $\Omega_{DM} = 0.264$; $\Omega_b = 0.049$; $n_s = 0.966$; $\sigma_8 = 0.830$; $\Sigma M_\nu = 0.06$ eV), with the exception of the two running simulations (as explained above).

Note that we use a version of the GADGET3 code (see McCarthy et al. 2018) that includes the impact of massive neutrinos on the expansion rate and the growth of fluctuations (i.e. accounts for their free streaming), using the semilinear algorithm of Ali-Haïmoud & Bird (2013). As the simulations employed here adopt the minimum allowed neutrino mass ($\Sigma M_\nu = 0.06$ eV), consistent with what is assumed in the Planck analysis, their incorporation will not have important consequences for the present study, but we include them for consistency.

2.3 Matter power spectra

In this study, we probe the theoretical cosmic shear power spectrum over a large range of multipoles/angular scales. This requires that we have a model for the non-linear matter power spectrum which spans a large range of wavenumbers/physical scales. However, the high-resolution 25 h^{-1} Mpc simulations that we use to probe the non-linear effects only have a k-range spanning from ≈ 0.25 to 257.36 h Mpc^{-1}. While the simulations extend to large enough k (i.e. small scales) for upcoming lensing surveys, they are clearly too small to capture all of the relevant structure on large scales. Even though our interest is primarily focused on small scales, we nevertheless want to construct realistic synthetic cosmic shear observations. (Furthermore, as we will show, the running of scalar spectral index models has important contributions from large scales.)

---

[11] https://github.com/sbird/S-GenIC

We thus require a way of modelling the non-linear matter power spectrum over a wide range of scales.

2.3.1 Method for constructing matter power spectra

In order to model the non-linear matter power spectrum over a wide range of wavenumbers, we combine our direct simulation measurements with predictions from `Halofit` (Smith et al. 2003; Takahashi et al. 2012) on larger scales. Briefly, the `Halofit` algorithm provides an empirical correction for scaling the linear power spectrum (e.g. from a Boltzmann code) for a given cosmology to the non-linear power spectrum. The empirical corrections were derived from fits to a large suite of CDM-based cosmological (*N*-body) simulations. We use the Boltzmann code CAMB to compute the `Halofit` prediction at a given redshift. This provides us with the fiducial non-linear ΛCDM power spectrum. To construct a full power spectrum for each of our extensions to ΛCDM, we use the ratios of the non-linear matter power spectra extracted directly from the simulations (with respect to our ΛCDM case) as a multiplicative 'boost factor' for the non-linear `Halofit` ΛCDM prediction via

$$P_{\rm C}(k, z) = P_{\Lambda{\rm CDM}}^{\tt Halofit}(k, z)R_{\rm C}(k, z), \qquad (6)$$

where $P_{\rm C}(k, z)$ is the non-linear matter power spectrum in a given cosmological extension, and $R_{\rm C}(k, z)$ is the constructed ratio (boost factor) for a given cosmology.

This approach allows us to combine multiple simulations of varying box size or resolution (by combining their ratios), following a similar vein to the power spectrum splicing in previous Lyman α forest work (e.g. Palanque-Delabrouille et al. 2015). The combined ratios can then be used to seamlessly scale the power spectra derived on very large scales (e.g. via linear theory, perturbation theory, `Halofit`, etc.) via equation (6) to produce an absolute power spectrum for a cosmological model.

Note that previous studies have shown that cosmological simulations that do not simulate the smaller *k* modes (larger physical scales) do not accurately represent the intermediate *k* modes sampled in the cosmological volume (see e.g. Power & Knebe 2006; Heitmann et al. 2010; Klypin & Prada 2018; Euclid Collaboration 2019), making it difficult to combine the absolute matter power spectra from simulations of different box sizes. This motivates the use of combining ratios of power spectra rather than splicing the absolute power spectra themselves. We have explicitly tested how sensitive the ratio of the different power spectra are to resolution as well as box size effects (see Fig. A1 in Appendix A), concluding that they are more robust to such effects than are the absolute power spectra.

To explain how we construct the non-linear matter power spectrum in a bit more detail, first, we compute the matter power spectra for the different simulations using GENPK[12] (Bird 2017). We then re-bin the matter power spectra to have 10 *k* modes per bin (corresponding roughly to rebinning the power spectra in logarithmic bins of width 0.0143 dex), in order to smooth out some of the associated noise. We compute the mean power in each bin, as well as the mean wavenumber. The ratios of the power spectrum for each cosmological extension with respect to the ΛCDM result are then computed. In the case of the SIDM and WDM models, all six of these simulations have ratios which tend to 1 within the simulated volume (i.e. the physical effects of these modifications are confined to small scales), as such we simply extrapolate these ratios to smaller *k*-values (larger physical scales) assuming they are fixed at unity over the entire range.

[12]https://github.com/sbird/GenPK

In the case of the two simulations that have a running scalar spectral index, the ratios of these power spectra do not tend to unity within the high-resolution 25 h^{-1} Mpc boxes, as such we need to simulate the ratio over a wider *k*-range. To do this, we use two further sets of simulations of size 100 and 400 h^{-1} Mpc, each with 1024^3 dark matter particles, to compute the ratio of these models with respect to a corresponding ΛCDM result. As a result, we can accurately probe the non-linear matter power spectrum in these simulations up to approximately linear scales where we can then make use of the theoretical prediction from `Halofit`. The reason we do not use the `Halofit` prediction over the entire *k*-range for the simulations with running is because it does not tend to reproduce the effects we see in the running simulations on non-linear scales (as shown by the dotted line in Fig. A1), discussed in Stafford et al. (2020a) (see also Smith & Angulo 2019).

With a combined ratio spanning the entire desired *k*-range for each cosmological model, we fit a cubic spline to the data, which is smoothed with a third-order Savitsky–Golay filter (Savitzky & Golay 1964) over the nearest 51 wavenumbers. These two steps are done to ensure a smooth continuous function describing the ratio over the entire *k*-range. We repeat this process at each redshift for which we have a simulation snapshot. Once we have the ratio for each cosmology at each redshift relative to the ΛCDM prediction, we use it as a multiplicative boost factor to a ΛCDM matter power spectrum (at the corresponding redshift) to obtain the absolute power spectrum for each cosmology. Note that we have simulation snapshots at $z = \{0.0, 0.125, 0.25, 0.375, 0.5, 0.75, 1.0\}$ for which we compute the non-linear total matter power spectrum. When we compute our shear power spectra described in Section 3.1, we use a cubic spline to interpolate between both redshift and *k* modes. We set the power to zero for *k* modes outside of the range sampled by our constructed matter power spectrum.

2.3.2 Resulting non-linear matter power spectra

The resultant constructed matter power spectra can be seen in Fig. 1, where we show the absolute power spectra at redshift $z = 0$ in the top panel, and the redshift evolution of the ratios computed with respect to the ΛCDM result in the bottom three panels. We only show here the results up to a maximum redshift of $z = 1.0$ as this is the maximum source redshift we use in our weak lensing calculations (we discuss our choice of a maximum source redshift of $z = 1.0$ in Section 4.1).

A feature common to most of the cosmological extensions examined here is the suppression of small-scale power at $k > 10\, h\, {\rm Mpc}^{-1}$. A cosmological model with a negatively running scalar spectral index has less power on small scales due to a dampening effect on the initial density perturbations generated by inflation. WDM cosmologies have small-scale density perturbations erased due to free streaming. Finally, SIDM cosmologies have small-scale clustering erased at late times due to the self-interactions creating cores (and generally erasing structure) in otherwise high-density dark matter haloes. In detail, however, differences between the three models do exist. For example, whereas the suppression in the matter power spectrum decreases with decreasing redshift for WDM and a negative value for $\alpha_{\rm s}$, the suppression increases with time in a SIDM cosmology. The suppression reduces with time for the other two cosmological models due to mode mixing transferring power from large scales to smaller scales. A similar result to this was seen in Stafford et al. (2020a) for the case of a running spectral index and in Viel et al. (2012) for the case of WDM. However, in the case of SIDM the effect seen in
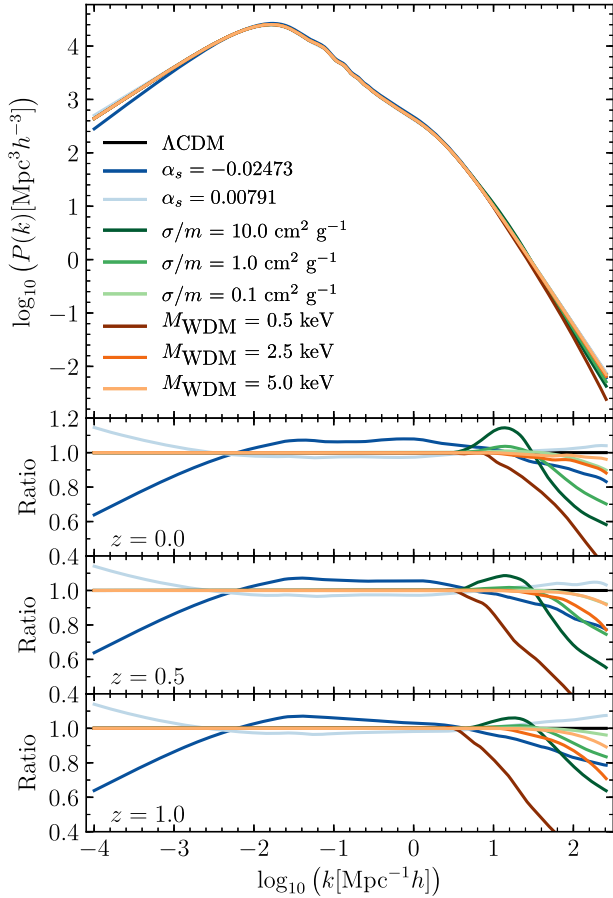
Figure 1. Top: the constructed non-linear matter power spectrum for each cosmological extension indicated with varying colours, computed at $z = 0$. Bottom (three panels): the matter power spectrum of each cosmological extension normalized to the ΛCDM case. Redshifts 0, 0.5, and 1.0, respectively, are shown, demonstrating the redshift evolution of the ratios for each cosmological model. All models, with the exception of a cosmology with a positively running scalar spectral index, result in some level of suppression of the non-linear matter power spectrum compared to the ΛCDM prediction on small physical scales. This suppression ranges from around 5 per cent in the case of the less extreme WDM cosmology to ≈60 per cent for the more extreme WDM cosmology for the smallest scales examined here at the present day.

the non-linear matter power spectrum of the suppression becoming larger with decreasing redshift, which is a feature unique to SIDM, occurs due to the higher virial velocities present at late times. This causes scattering events to be more efficient at redistributing the mass inside the inner regions of haloes. Furthermore, there is a cumulative number of scattering events which occur with decreasing redshift. These combined effects result in the suppression in the matter power spectrum extending to smaller k modes. The differences seen in the redshift evolution of the matter power spectrum are important as measurements at a single fixed redshift could result in different cosmological models having very similar effects. See, for example, the $z = 0.5$ panel in Fig. 1 where the $M_{WDM} = 5.0$ keV and the $\sigma/m = 0.1$ cm^2 g^{-1} are predicted to have an almost indistinguishable effect on the non-linear matter power spectrum. However, at earlier and later redshifts the predictions for their suppression of the matter power spectrum are noticeably different.

The model with a positive running scalar spectral index is the only model to predict an enhancement in power on these same scales,

except for the slight increase in power seen in the more extreme SIDM models on k-scales ≈25 h Mpc^{-1}. The increase in power on small-scales in the SIDM models is due to the re-distribution of dark matter in the central regions of dark matter haloes. In particular, the enhancement mainly occurs due to the outward scattering of dark matter particles from the very central regions to somewhat larger radii (see fig. 5 in Stafford et al. 2020b, for example). This outer radii, where particles tend to gather in stable orbits corresponds approximately to the radius where we expect each particle to have interacted at least once per Hubble time (see section 5.3 of Rocha et al. 2013, for example, for a discussion on this).

One obvious feature of the two cosmologies with a running scalar spectral index is the cross-over regions seen in the ratio panels (in the range $-2 \leq \log_{10}(k\,[h\,\mathrm{Mpc}^{-1}]) \leq 1$). The reason these cross-over regions exist is because of differences in the amplitude of the primordial matter power spectrum, A_s, which is larger in the case of the positive running cosmology and smaller in the negative running case with respect to A_s for the fiducial ΛCDM case. (Note that A_s varies between the models in order to retain a good match to the Planck CMB angular power spectrum, see discussion in Stafford et al. 2020a.) This amplitude difference produces an additional offset with respect to the ΛCDM power spectrum on top of the resulting increase (decrease) in power on large and small scales due to the positive (negative) running of the spectral index.

3 TOMOGRAPHIC WEAK LENSING

In this section, we describe our methodology for computing cosmic shear power spectra from the non-linear matter power spectra described above and how we generate noisy realizations (synthetic observations) of the cosmic shear power spectra.

3.1 Theory

Weak lensing describes the deflection of light rays due to the presence of LSS in the Universe, which results in slight correlated distortions in the observed shapes of galaxies (Blandford et al. 1991; Miralda-Escude 1991; Kaiser 1992). The lensing of galaxies leads to two effects: the dilation or magnification of an image which can be described by the convergence, κ, and the stretching (shearing) of an image, $\gamma_{1,2}$. In this study, we focus on the shears of galaxies that can be used to probe the projected mass distribution via the galaxy shape correlation functions. Additional information on the growth of structure over cosmic time can be obtained if one has redshift measurements of the background galaxies. In this case, the source distribution can be discretized into redshift (or tomographic) bins, allowing one to probe the three-dimensional matter distribution (see Kilbinger 2015 for a recent review).

The two-point correlation function of galaxy shapes, and its Fourier analogue the power spectrum, are directly linked to the underlying matter distribution and its power spectrum integrated along the line of sight. We compute the cosmic shear power spectrum via:

$$P_{ij}^{\gamma}(\ell) = \int_0^{\chi_H} d\chi_1 (1+z_1)^2 W_i(\chi_1) W_j(\chi_1) P_{3D}\left(k = \frac{\ell}{\chi_1}, \chi_1\right) \quad , \quad (7)$$

where χ_1 is the comoving distance to the lens at redshift z_1, P_{3D} is the non-linear matter power spectrum, and W_i and W_j are the lensing efficiencies in the tomographic bins i and j, defined as

$$W_{i,j}(z_1) = \frac{3}{2}\Omega_m \left(\frac{H_0}{c}\right)^2 \int_{z_1}^{z_{max}} \frac{\chi_1 - \chi_s}{\chi_s} n_{i,j}(z_s)\, dz_s. \quad (8)$$

Here, χ_s corresponds to the comoving distance to the source at redshift z_s and n_i is the normalized homogeneous source distribution in tomographic bin i given by

$$n_i(z) = \frac{n(z)}{\int_0^{z_{max}} n(z)\,\mathrm{d}z}, \tag{9}$$

It is worth mentioning that equation (7) is predicated upon some important assumptions. These include the Limber approximation (Limber 1953; Kaiser 1992), which only includes modes in the plane of the sky, neglecting those between structures at different epochs in the line-of-sight integration. The small-angle and flat-sky approximations have also been adopted, which allows one to replace a spherical harmonics transform with a Fourier transform (Hu 1999). Other assumptions that are embedded inside the lensing efficiency $W_{i,j}$ (see equation 8) is that of a homogeneous galaxy distribution, which ignores source–source clustering (Schneider, van Waerbeke & Mellier 2002) as well as source–lens clustering (Bernardeau 1998; Hamana et al. 2002). Spatial flatness has also been assumed in equation (7).

Although the previous assumptions may introduce some level of systematic error in the calculation of the shear power spectrum, we neglect them here owing to the fact that, first, they will be common to each cosmology and should therefore be less important when we focus on the ratios of the cosmic shear power spectra, and, secondly, they mainly affect the largest angular scales ($\ell < 10$) (see for example Schmidt 2008; Giannantonio et al. 2012; Kilbinger et al. 2017) which is not our focus.

The uncertainty in the shear power spectrum can be expressed as (Kaiser 1998; Hu 1999; Euclid Collaboration 2020a)

$$\Delta P_{ij}^\gamma(\ell) = \sqrt{\frac{2}{(2\ell+1)\,\Delta\ell\, f_{sky}}} \left[P_{ij}^\gamma(\ell) + \delta_{ij}\frac{\langle\gamma_{int}^2\rangle}{\overline{n_i}} \right], \tag{10}$$

where $\Delta\ell$ is the multipole bandwidth, f_{sky} is the fraction of sky surveyed, δ_{ij} is the Kronecker delta symbol, $\langle\gamma_{int}^2\rangle$ is the shape noise, representing the variance of observed galaxy ellipticities (which we take to have a value of 0.261, motivated by Gatti et al. 2021), and $\overline{n_i}$ is the surface density of source galaxies in the tomographic bin, expressed in steradians^{-1}. The term under the square root accounts for the limited number of available independent ℓ modes. The first term in the square brackets corresponds to the cosmic variance and the second term is a Poisson noise term. In addition to introducing scatter in the cosmic shear power spectrum (via the Poisson noise term), the shape noise also contributes to an additive shot noise term that biases the auto power spectrum but not the cross-power spectrum, as the shot noise in different tomographic bins is uncorrelated. Thus, the shot noise must be subtracted from the estimated auto power spectra.

As written in equation (10) (see also Euclid Collaboration 2020a, equations 118 and 125), it appears that the Poisson noise only applies to the uncertainty in the autocorrelation power spectrum and not the uncertainty in the cross-correlation power spectrum, in analogy to the way shot noise contributes to the measured power spectrum but not the cross-spectrum. While the Poisson noise is uncorrelated between tomographic bins and therefore does not bias the cross-spectra, we find that there is a significant contribution to the *uncertainty* in the measured cross-spectra due to Poisson noise. Furthermore, equation (10) only provides an estimate of the Gaussian errors; i.e. the diagonal elements of the covariance matrix, but in principle there could be significant non-Gaussian contributions. For these reasons, instead of using the standard analytic error estimate in equation (10),

we instead generate synthetic weak lensing observations using the FLASK software package, as described below.

We note that cosmic variance, shape noise, and Poisson errors are not the only sources of uncertainty for the cosmic shear power spectrum. One of the major astrophysical sources of uncertainty stems from the intrinsic alignment of galaxy shapes. This stems from tidal interactions during the formation period of nearby galaxies that induces an intrinsically correlated orientation of the galaxies' shapes (Joachimi et al. 2015; Kiessling et al. 2015; Kirk et al. 2015) that works to dilute the cosmological signal in the two-point correlation function of these shapes. Additionally, another important source of error stems from photometric redshifts being used for source galaxies. This leads to some galaxies being ascribed the wrong redshift and blurring the edges of tomographic bins. A comprehensive discussion on the systematic errors affecting weak lensing surveys can also be found in Mandelbaum (2018). For simplicity, we ignore these effects in the present study.

3.2 Synthetic cosmic shear observations

For our cosmic shear analysis, there are several choices one needs to make to specify the set-up. For example, as seen in equation (7), the shear power spectrum will directly depend on the redshift distribution of the source galaxies. It will also depend on how many tomographic bins are used in the analysis. Furthermore, the associated uncertainty on the shear power spectrum will depend on the density of source galaxies as well as on the survey sky coverage.

In this study, we are interested in whether or not future Stage-IV weak lensing surveys such as Euclid, LSST, and EUCLI would be able to rule out, or help place constraints on, the different cosmological extensions examined in this study. For our fiducial results, we compute the shear auto- and cross-power spectra for a Euclid-like set-up, however, the survey parameters of an EUCL-like set-up would be very similar, and so the results we present below would not differ much (we have explicitly verified this). We discuss the differences with respect to an IDEUCL set-up in Section 4.1.

When constructing the Euclid-like set-up, we follow as closely as possible that described in Euclid Collaboration (2020a). In particular, the source galaxy distribution is defined via

$$n(z) = z^\alpha \exp\left(-\left[\frac{z}{z_0}\right]^\beta\right), \tag{11}$$

where α, β, and z_0 are survey specific parameters that describe the source distribution. We adopt values of 2, 1.5, and 0.636, respectively, to be similar to the source distribution expected for the Euclid survey.

The planned analysis of the Euclid survey splits the galaxy distribution up into 10 tomographic bins with a redshift range of $0.001 \le z_s \le 2.5$. However, in this study we focus on redshifts ≤ 1, which approximately corresponds to the first six tomographic bins in the IdCLIE set-up. As such, we split our galaxy distribution up into six tomographic bins between $0.001 \le z_s \le 1.0$. The edges of each tomographic bin are $z_i = \{0.001, 0.414, 0.554, 0.669, 0.777, 0.885, 1.0\}$ and are defined such that there is an equal number of galaxies in each bin, essentially fixing the associated shot noise in each bin. We choose a maximum source redshift of 1 as, for a fixed range of angular scales, higher redshift measurements correspond to larger physical scales.[13] As the extensions we explore mostly affect small physical scales, we do not expect them to be easily distinguishable

---

[13]Note the minimum source redshift of 0.001 was chosen to coincide with the minimum redshift used in Euclid Collaboration (2020a).

from ΛCDM at high redshifts (the exception to this are the models which have a running of the scalar spectral index, as we discuss below). Specifically, we compute the cosmic shear power spectrum up to an $\ell_{max} = 4000$, which lies close to the 'optimistic' case for the UclidD survey of $\ell_{max} = 5000$ (with their 'pessimistic' case being $\ell_{max} = 1500$, if one does not include non-Gaussian contributions to the covariance matrix). This ℓ_{max} corresponds to angular scales of around 0.44 arcmin, or $k\,[h\,\mathrm{Mpc}^{-1}] = 2.18$, at $z = 1$.

Note that, as surveys such as Euclid and LSST will make use of photometric redshifts (see e.g. Euclid Collaboration 2020b), one needs to account for the associated uncertainty by convolving the number density distribution in equation (9) with a probability distribution function accounting for uncertainties in the photometric redshifts, along with accounting for catastrophic outliers (see for example equation 115 in Euclid Collaboration 2020a). These are galaxies which have a severely incorrect measurement of their redshift and have therefore been placed in the wrong tomographic bin. We ignore this effect in this paper, however.

To generate the synthetic weak lensing observations using the cosmic shear power spectra computed in Section 3.1 and the details of the Euclid survey choices above, we use the publicly available software package FLASK[14] (Xavier, Abdalla & Joachimi 2016). FLASK can generate lognormal (or Gaussian) realizations of correlated fields on spherical shells. We make use of its ability to generate weak lensing shear fields. Each field can be generated tomographically, with the statistical properties of these fields (including their cross-correlations) being defined via input angular power spectra (where one provides as input the auto- and cross-spectra that they want maps created for). We provide as input the theoretical cosmic shear power spectra as computed in equation (7), which in the flat-sky approximation is equivalent to the convergence power spectrum (Hu 2000; Bartelmann & Schneider 2001; Kilbinger et al. 2017). We use FLASK to produce many realizations of our theoretical shear power spectra, allowing us to evaluate the full covariance matrix in the error analysis, rather than just using the Gaussian errors computed using equation 10, and to evaluate the potential Poisson noise contribution to cross-power spectra between tomographic bins. Note that the measured power spectra computed from the FLASK maps will contain a shot noise contribution. Therefore, when computing the full covariance matrix we subtract this shot noise from the auto-correlation power spectra.

In order to define the log-normality of the shear field, one needs to specify a shift parameter. We compute this shift parameter (λ) following Hilbert, Hartlap & Schneider (2011) who used the Millennium Simulation (Springel 2005) to produce synthetic convergence and shear maps. From these maps, they measure the convergence distribution and find that it is best fitted with a zero-mean shifted lognormal distribution. They do this at multiple redshifts and provide an empirical formula, which we use here, that captures the redshift evolution of the shift parameter λ (κ_0 in Hilbert et al. 2011)

$$\lambda(z) = 0.008z + 0.029z^2 - 0.0079z^3 + 0.00065z^4. \quad (12)$$

We substitute in the mean redshift of each tomographic bin to calculate the associated shift parameter for that field. The values calculated can be found in Table 1.

We want FEucl to output shear maps that will also capture the noise associated with measurements of the cosmic shear power spectrum (see Section 3.1). Therefore, we also supply idEucl with an angular

Table 1. The shift parameters used to define the lognormal realizations of each shear field. The columns from left to right are the tomographic bin number, the mean redshift of the tomographic bin, and the calculated shift parameter for that bin.

i	z_{mean}	λ
1	0.3003	0.0048
2	0.4874	0.0099
3	0.6126	0.0141
4	0.7233	0.0181
5	0.8307	0.0224
6	0.9417	0.0272

selection function, such that the map will be masked to reproduce a survey's specific f_{sky}, as well as providing a redshift selection function so that the number of galaxies in a tomographic bin is reproduced. One also needs to supply the software with a value for the ellipticity dispersion of galaxy shapes, $\langle \gamma_{int}^2 \rangle$, in order to incorporate shape noise into the shear maps created. As already noted, we adopt $\langle \gamma_{int}^2 \rangle = 0.261$.

The maps output by idASK, which are output in HEALPIX[15] format, contain the mean source ellipticity in each pixel, calculated via

$$\epsilon(j) = g(j) + \frac{\epsilon_s}{\sqrt{N_{gal}}}, \quad (13)$$

where $g(j)$ is the reduced shear associated with pixel j, ϵ_s is a value for the intrinsic ellipticity associated with galaxy shapes randomly sampled from a zero-mean Gaussian distribution with width equal to $\langle \gamma_{int}^2 \rangle$, and N_{gal} is the number of galaxies that fall within that pixel. We produce 200 map realizations of each tomographic power spectrum in this way. Following this, we compute the auto- and cross-correlation power spectra of these maps using the PYTHON package HEALPY.[16] We use these power spectra to evaluate the covariance matrix in 14 multipole bins in the range $10 \leq \ell \leq 4000$, which we calculate as

$$\mathrm{cov}\left[P_{ij}^{\gamma}(\ell),\, P_{i'j'}^{\gamma}(\ell') \right]$$
$$= \left\langle \left(P_{ij}^{\gamma}(\ell) - \left\langle P_{ij}^{\gamma}(\ell) \right\rangle \right) \left(P_{i'j'}^{\gamma}(\ell') - \left\langle P_{i'j'}^{\gamma}(\ell') \right\rangle \right) \right\rangle. \quad (14)$$

The resultant cosmic shear auto- and cross-power spectra can be seen in Fig. 2 for all cosmologies, along with the associated uncertainties on the cosmic shear power spectrum, which serves to illustrate the absolute power spectra and how the associated changes due to the different cosmological models are generally quite subtle. The differences in the cosmic shear power spectra are better highlighted in Fig. 3, which shows the ratio of each auto- and cross-correlation power spectrum with respect to the ΛCDM prediction in that same tomographic bin. We discuss here the calculation of the error bars in Figs 2 and 3 and leave the scientific interpretation of these results for Section 4.

The red error bars in Fig. 3 show the theoretical uncertainty associated with the shear power spectrum as computed using equation (10). The black error bars represent the diagonal elements of the covariance matrix for a ΛCDM cosmology calculated using multiple realizations of each shear power spectrum as generated by clidK. Note here that the error bars that are shown are normalized to the ΛCDM power spectrum in each tomographic bin, i.e. $\Delta P_{ij}^{\gamma}(\ell)/P_{ij}^{\gamma}(\ell)$.

[14]https://github.com/hsxavier/flask

[15]https://healpix.sourceforge.io/
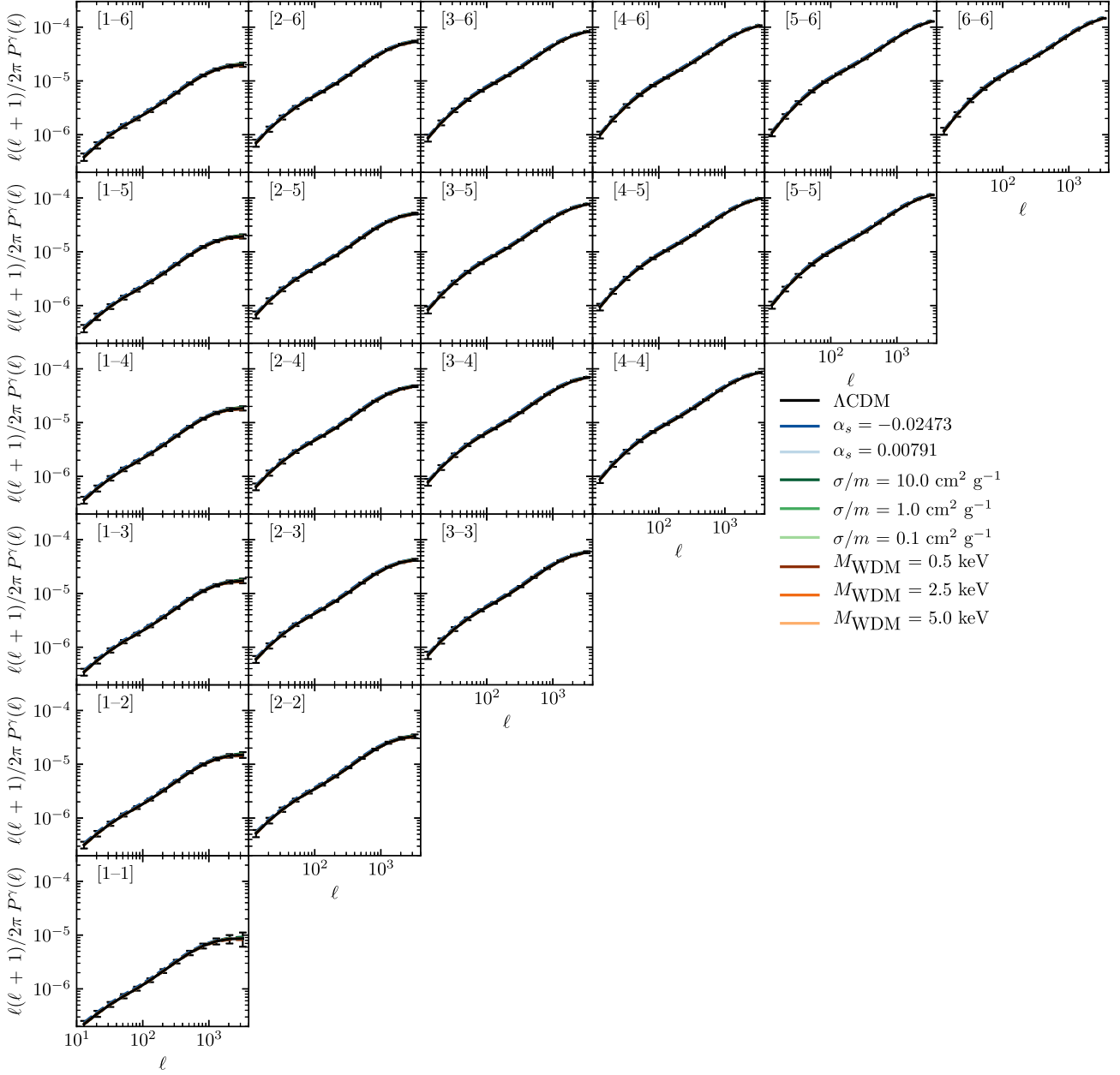[16]https://github.com/healpy/healpy

Figure 2. The shear auto- and cross-correlation power spectra computed for each cosmological model. The source galaxy tomographic bins are distributed between a minimum source redshift of 0.001 and a maximum source redshift of 1.0 and such that each tomographic bin has the same number density of source galaxies. We split the distribution up into six tomographic bins to mimic a Euclid-like survey and plot up to $\ell_{max} = 4000$, which lies between the pessimistic and optimistic case for the EuclEu survey. Here, the error bars show the standard deviation of the realizations of each power spectrum evaluated using the CLIDEU package (see text). In general, the deviations to ΛCDM case are subtle and require a quantitative evaluation of the signal-to-noise ratio.

It can be seen in Fig. 3 when comparing the theoretical error bars shown in red to the black error bars extracted from the covariance matrix computed with the assistance of FEucl, that the errors are in excellent agreement (as one would expect) for the auto-correlation power spectra. There is some disagreement at low multipoles, which stems from cosmic variance issues associated with the maps.[17] However, importantly, this agreement does not hold for the cross-correlation power spectra, with significant disagreements at large multipoles (small angular scales). Specifically, the analytic calculation in equation (10) ignores the effects of Poisson noise on uncertainty in the cross-spectrum, but we find there is always a non-negligible Poisson error.

[17]We tested whether the error bars at small multipole were brought into better agreement if one created full sky realizations of the power spectra using FLASK rather than a masked version and found that this was indeed the case. Therefore, it is likely that the masking is adding some additional spurious noise to the signal on large angular scales.
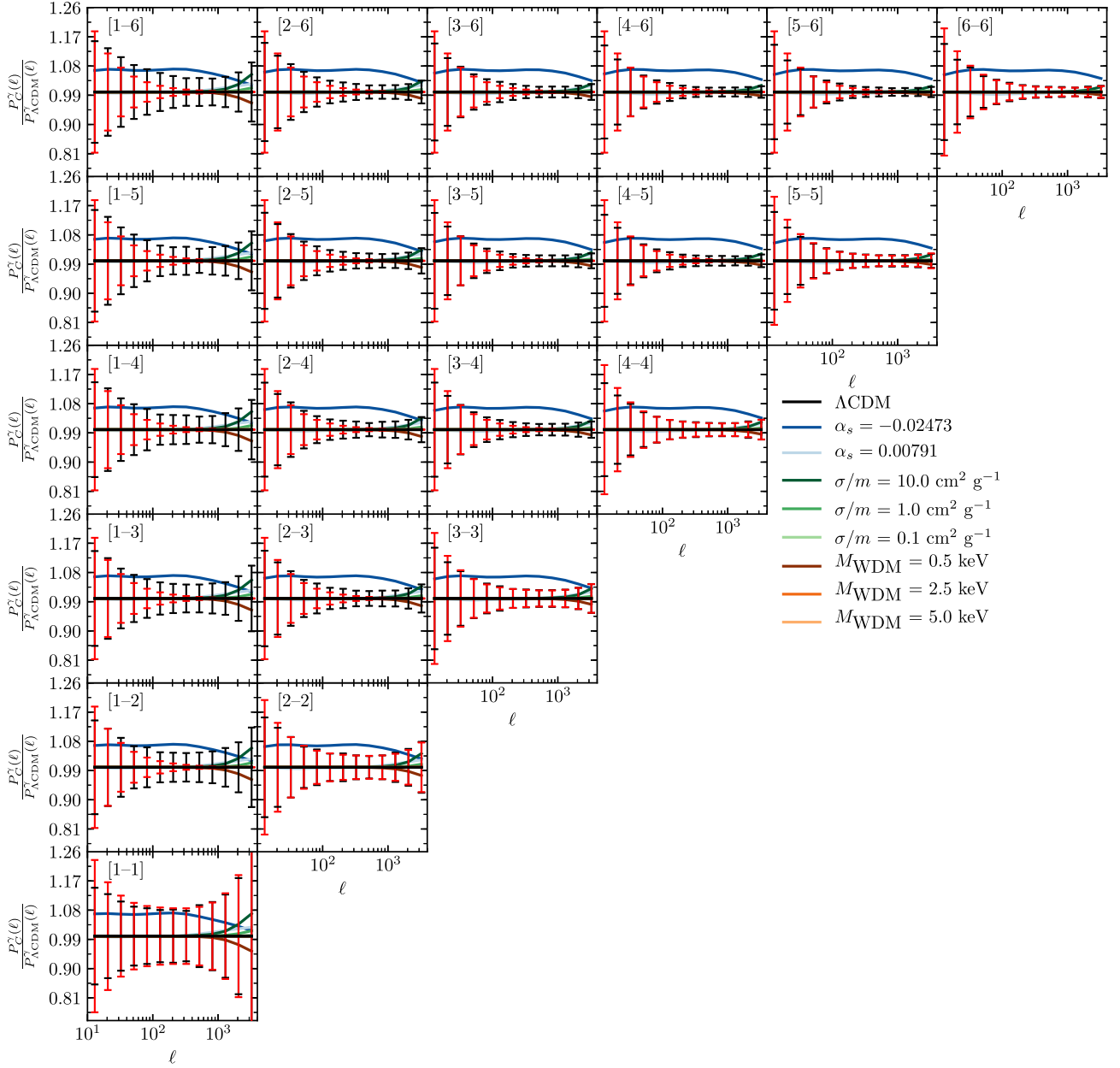
Figure 3. The shear auto- and cross-correlation power spectrum for each cosmological model normalized to the corresponding ΛCDM auto-/cross-correlation power spectrum. The different coloured lines represent the different cosmological modes, indicated in the legend. The red error bars show the standard theoretical (analytic) prediction for the noise on the shear power spectra (which does not include the effects of Poisson noise on the shear cross-spectra), and the black error bars show the noise computed using multiple realizations of each auto and cross power spectrum in a ΛCDM cosmology with the CLIDEU package. This plot helps highlight the resultant changes to the cosmic shear power spectrum, particularly due to the inclusion of a running scalar spectral index as a free parameter in the standard model. The numbers in square brackets in the top left of each panel indicate the tomographic bin number. Increasing tomographic bin number corresponds to higher redshifts (see Table 1 for the mean redshift of each bin).

In Appendix B, we use realizations of pure shape noise fields generated with idEucl to derive a more accurate and general equation for incorporating the impact of shape noise (Gaussian errors only) on the cross-spectrum, namely

$$\Delta P_{ij}^{\gamma}(\ell) = \sqrt{\frac{2}{(2\ell + 1) f_{\text{sky}}}}$$
$$\times \left[P_{ij}^{\gamma}(\ell) + \delta_{ij} \frac{\langle \gamma_{\text{int}}^2 \rangle}{n_i} + (1 - \delta_{ij}) \frac{\langle \gamma_{\text{int}}^2 \rangle}{\sqrt{2 n_i n_j}} \right]. \quad (15)$$

The new term (right most in the square brackets) represents the contribution of Poisson noise to the uncertainty in the cross-power spectrum. Note the similarity of the noise terms for the auto- and cross-spectra, which is expected because they are caused by the same effect: random alignments of galaxy shapes. The only difference is that for the cross-spectrum term we allow for the possibility of different source densities in the tomographic bins being cross-correlated, and the extra factor of square root of 2, which is the result of having more galaxy pairs to evaluate in the cross-spectrum compared to the power spectrum (the error scales as the number of pairs).

Finally, we note that the black error bars plotted in Figs 2 and 3 correspond to just the diagonal elements of the covariance matrix, but we use the full covariance matrix when evaluating signal-to-noise ratios (SNRs) in Section 4.

3.3 Summary of theoretical cosmic shear pipeline

Here, we provide a very brief summary of the steps described above in generating cosmic shear power spectra for the different cosmological models in this study, alongside a realistic estimate of the uncertainties.

(i) Compute $P(k)$ for each cosmological simulation (including box size variations for the two α_s simulations) and re-bin to smooth out some of the initial numerical noise associated with the power spectra.

(ii) Compute the ratio of each power spectrum with respect to a corresponding ΛCDM simulation (of same box size and resolution).

(iii) For cosmologies with a running spectral index, combine the ratios computed from the 400, 100, 25 Mpc h^{-1} box simulations along with the Halofit prediction to a $\log(k_{min}) = -4$. For all other cosmologies, extrapolate $R_C(k, z) = 1$ for large scales not sampled in the 25 Mpc h^{-1} box. Fit a cubic spline to the constructed ratios and smooth.

(iv) Calculate the absolute $P(k)$ for each cosmology using the constructed ratio as a boost factor to modify the ΛCDM prediction computed using Halofit (equation 6).

(v) Generate theoretical tomographic shear power spectra for each cosmological model using equation (7).

(vi) Construct multiple synthetic tomographic weak lensing shear maps for a ΛCDM universe including galaxy shape noise using the idASK package. Compute $C(\ell)$ for each realization to construct a covariance matrix for the ΛCDM prediction using equation (14).

4 RESULTS

In this section, we present the main results of this work. This includes predictions for the tomographic cosmic shear auto- and cross-power spectra for each cosmological model, along with their respective ratios with respect to ΛCDM. We also explore how large the differences in the cosmic shear power spectrum due to a change in cosmology are compared to the expected error associated with the cosmic shear power spectrum.

4.1 Comparisons up to $\ell_{max} = 4000$

We now discuss the main results of our investigation for the fiducial EUCLIE-like set-up, examining the cosmic shear power spectrum up to a maximum multipole of $\ell_{max} = 4000$.

Examining the ratios of the various cosmological extensions with respect to ΛCDM shown in Fig. 3, we can immediately conclude (by eye) that all of the cosmological extensions studied here are capable of producing some level of deviation in the cosmic shear power spectrum compared to the ΛCDM prediction over this range of multipoles. However, in the case of WDM and SIDM, these changes appear most noticeable for the two extremist models examined in this study. The other models do produce differences as well, but they are comparatively smaller and a more quantitative analysis of their 'detectability' is thus required.

We quantitatively characterize the constraining power of the cosmic shear observations via an SNR, which we evaluate as

$$(S/N)^2 = \sum_{\ell, \ell' \leq \ell_{max}} (|R(\ell) - 1|) \left(\frac{\text{Cov}[P_{ij}^{\gamma}(\ell), P_{ij}^{\gamma}(\ell')]}{P_{ij}^{\gamma}(\ell) P_{ij}^{\gamma}(\ell')} \right)^{-1}$$
$$\times (|R(\ell') - 1|), \qquad (16)$$

where $R(\ell)$ is the ratio of each cosmological extension with respect to ΛCDM as plotted in Fig. 3. We make use of the full covariance matrix calculated using equation (14), which we normalize with respect to the absolute power spectra themselves. We do this because we are comparing the error bars to the ratios rather than the absolute power spectra.

The resultant plot of the integrated SNR as a function of ℓ_{max} can be seen in Fig. 4. The solid curves represent the integrated SNR as a function of scale when using the full covariance matrix to estimate the uncertainties. For comparison, the dashed curves show the integrated SNR when using only the Gaussian (diagonal elements) errors. A legend is provided which lists the integrated SNR when summed (in quadrature) over all tomographic bins for the full covariance matrix case and for the diagonal errors only (the latter is in parentheses).

As expected from Fig. 3, the cosmology with a negative value for the running of the spectral index shows strong deviations from the ΛCDM prediction over almost the entire multipole range, with these differences becoming larger at higher redshift. The reason for this stems mainly due to an increase in the absolute power on these multipoles in the later tomographic bins, as seen in Fig. 2. This causes the relative error bars to decrease with increasing redshift resulting in a larger SNR. When summed over all tomographic bins, a Planck-based cosmology with a negative running of the scalar spectral index will likely be easily distinguishable from a Planck-based ΛCDM cosmology with a uclidD-like cosmic shear survey.

One apparent feature in the SNR for the negative running cosmology is the plateau towards higher multipoles. The reason for this behaviour is that the signal we see here is dominated by the region of the power spectrum for which a negative running cosmology produces an enhancement in power (due to the increase in A_s) relative to ΛCDM. However, on the largest multipoles the ratio begins to turn over and decrease, where it would eventually cross over the ΛCDM prediction and predict a suppression in the cosmic shear power spectra, rather than the enhancement that we see on these scales. Therefore, the signal, relative to the error bars, decreases over this range of scales yielding a plateau in the integrated SNR. The same effect is also seen, perhaps somewhat more clearly, in the cosmology with a positively running scalar spectral index where the opposite is true. In this case, there is initially a slight suppression in the cosmic shear power spectra which flips and becomes an enhancement at high multipoles, producing the initial increase in the SNR, followed by a plateau and then a second increase after the ratio has crossed unity. As a result, the positive running cosmology has an integrated SNR in the later multipole bins exceeding 1, showing that upcoming future surveys such as Euclid and LSST are perhaps able to put constraints on a value for the running of the scalar spectral index that are competitive to those from cosmic microwave background and Lyman α forest constraints.

Note that the SNR would likely continue to increase if we increase the maximum source redshift beyond the limit of $z_s = 1.0$ that we adopt here, at least for the running cosmologies. If one was to include tomographic redshift bins beyond $z_s = 1.0$ the relative error bars shown in Fig. 3 would decrease (due to the increase in the cosmic shear signal); hence, this would enhance the SNR seen here. The reason we do not include these higher redshift tomographic bins in
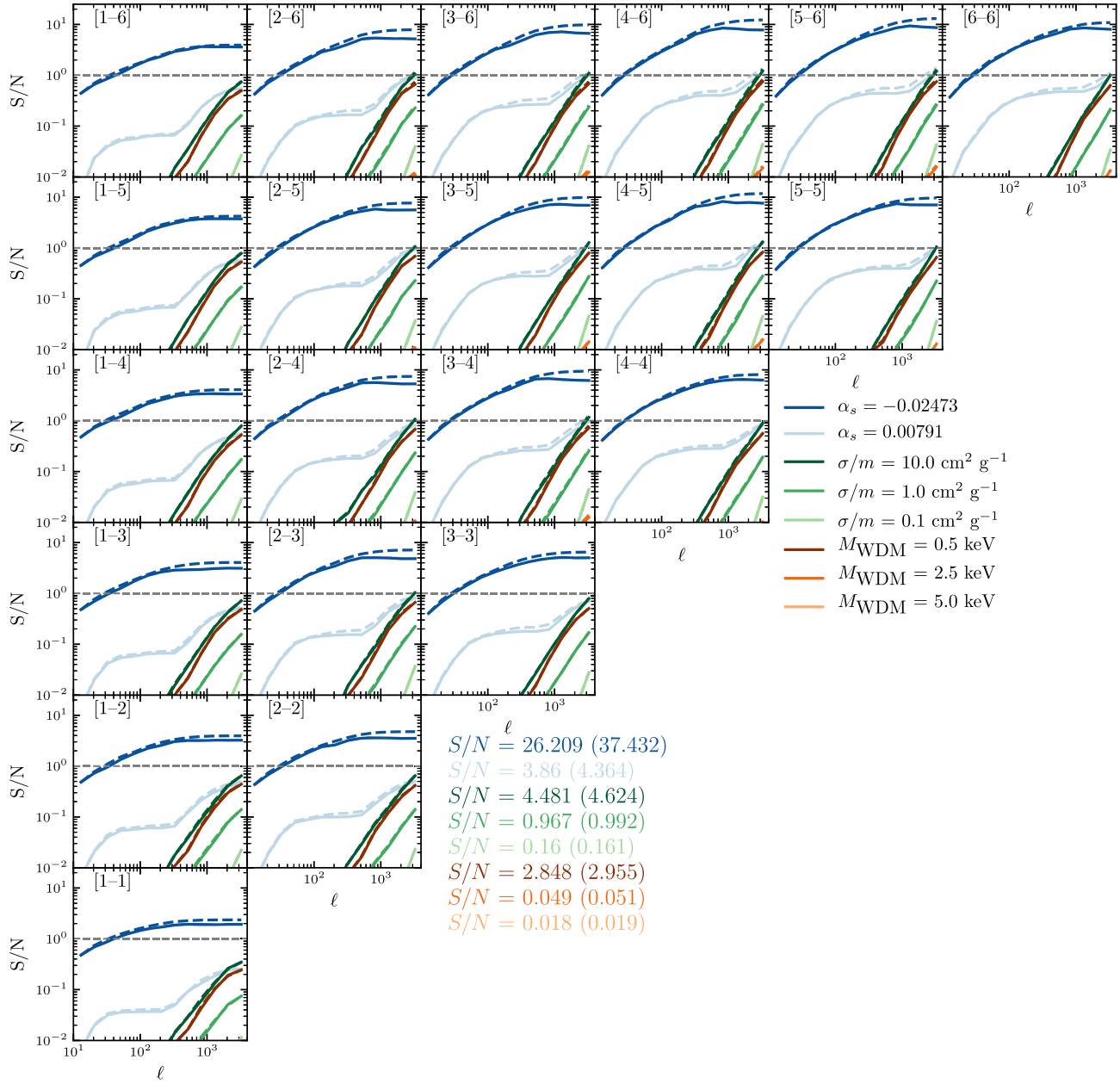
Figure 4. The integrated SNR as a function of ℓ_{max} to demonstrate how significant the deviations in the shear power spectra are for the different cosmological models relative to ΛCDM. We plot here two variants of this statistic shown by solid and dashed lines. In the first case (solid line), the SNR is calculated using the entire covariance matrix, whereas the dashed line only uses the diagonal elements of the covariance matrix. The total integrated SNR, summed (in quadrature) over the tomographic bins, is shown next to the bottom left panel with the values in brackets indicating the SNR if one only uses the diagonal elements of the covariance matrix.

this study is because, as mentioned, at fixed ℓ_{max}, but increasing z_s the region of the matter power spectrum one becomes sensitive to tends towards smaller k modes. As such, although this would still result in a measurable signal for the two models with a running spectral index, it would not result in a measurable signal for the WDM and SIDM cosmologies (their effects are confined to small physical scales). However, these higher redshift tomographic bins will be included in future weak lensing surveys, resulting in potentially increased constraining power on α_s compared to what is shown here.

For the SIDM and WDM cosmologies, it appears that one needs to probe to much smaller scales (higher multipoles) to be able to distinguish most of the models we explored from ΛCDM. The only

exceptions are the most extreme WDM (0.5 keV) and SIDM ($\sigma/m = 10 \text{ cm}^2 \text{ g}^{-1}$) models. Our calculations suggest that a Euclid-like survey with realistic source densities and shape noise may be able to (marginally) distinguish these models from ΛCDM. Pushing to higher multipoles would help in these cases as well.

An important aspect of a Euclid-like survey is the fact that it has a large sky coverage, meaning the cosmic variance error associated with the measurements of the cosmic shear power spectrum is strongly reduced on all but the largest angular scales. Therefore, the limiting factor in the ability of detecting a difference between some of these cosmological models and ΛCDM, particularly at the upper end of the multipole range we analyse in this study, is a result

of the uncertainty in galaxy shapes. Thus, we have also investigated a similar tomographic set-up as that expected for the EUCLI survey which instead of a large f_{sky} (DGRST is envisioned to have f_{sky} = 0.0485), is expected to have a larger number density of source galaxies equal to n_s = 51 galaxies arcmin^{-2} compared to n_s = 30 galaxies arcmin^{-2} for a EUCEuc-like (or LIDEU-like) set-up.

For the CLIDT set-up, we use five tomographic bins up to redshift 1.0, defining the source galaxy distribution using equation (11) (here, we adopt the parameter values: $\alpha = 2$, $\beta = 0.9$, $z_0 = 0.28$ in order to closely mimic the galaxy distribution described in Eifler et al. 2020). We find that, although the shot noise error is decreased in such a survey set-up, the increase in the cosmic variance error due to the reduction in sky coverage compensates, resulting in no improvement on the SNR.[18] Therefore, it seems that one would need both a large f_{sky} (although potentially not as large as EEucli or EUCL), and a large number density of source galaxies, n_s, to use cosmic shear to better constrain these cosmological models without having to push to higher multipoles.

It is apparent from the above analysis that, if it was possible to push to higher multipoles, this could considerably increase the constraining power for the SIDM and WDM scenarios. However, if we were to evaluate the uncertainties using a full covariance matrix approach this would be much more computationally expensive, as each realization has around 2×10^8 pixels in each tomographic map and $[N_T(N_T + 1)/2]$ auto- and cross-correlation power spectra (where N_T is the number of tomographic bins).[19] However, we can considerably simplify the process if the Gaussian errors (in equation 15) are sufficient to calculate an SNR. Comparing the solid and dashed curves in Fig. 4, we conclude that the difference in the integrated SNRs are fairly significant for the running cases (particularly the negative running scenario), but not for the SIDM or WDM cases, at least over the range of multipoles examined here. As such, below we examine if pushing to $\ell_{max} = 20\,000$ improves the SNR for the models with WDM, SIDM and positive running scalar spectral indices, using the estimated Gaussian errors rather than a full covariance matrix approach.

4.2 Comparisons up to $\ell_{max} = 20\,000$

We show in Fig. 5 the new integrated SNR, extended up to an l_{max} = 20 000, with the vertical dashed grey line in each panel showing the previous l_{max}. Note that we have replaced the reduced covariance matrix term in equation (16), with simply $\left(\Delta P_{ij}^\gamma / P_{ij}^\gamma\right)^{-2}$, where ΔP_{ij}^γ is calculated using equation (15), which includes the additional Poisson error on the cross-correlation power spectra.

Pushing to higher multipoles does indeed reveal regions of the cosmic shear power spectrum where the differences due to changes in cosmology are not yet completely drowned out by the associated noise. This is particularly true for the positive running cosmology and the two SIDM cosmologies with the smallest cross-sections, which all show a marked improvement in the total integrated SNR when summed over tomographic bins, as shown at the bottom of Fig. 5. However, although the total integrated SNR is now >1 for the less extreme SIDM cosmology of the two, the integrated SNR never exceeds 1 in any single tomographic bin, meaning the full

tomographic information will be needed to place constraints on the cross-section for interaction.

Note that the signal present in each tomographic bin will be somewhat sensitive to the tomographic binning strategy. For example, at fixed source density for a given survey, the fewer tomographic bins one has the larger the effective number density of source galaxies in each tomographic bin. This will work to decrease the shot noise associated with each tomographic power spectrum, thus increasing the integrated SNR in an individual bin. The integrated SNR when summed over all bins, though, should be a more robust quantity.

In the case of the WDM cosmologies, pushing to these higher multipoles does not result in a significant gain in SNR, in so far as detection is concerned. As explained previously, this is in part because the suppression in the matter power spectrum in a WDM cosmology is maximal at larger redshifts. However, the k-scales affected in a WDM cosmology move out of range of the ℓ modes covered in this study at higher redshifts.

While it is clear from the above analysis that pushing to higher multipoles in general results in increased SNRs, a potentially important caveat is that our analysis does not take into account additional sources of uncertainty that may be prevalent on such small angular scales, including source deblending difficulties. This refers to the phenomenon where sources (be them galaxies or stars) overlap on the sky, disrupting the shear estimation of a source galaxy. This can be accounted for by rejecting objects which are flagged as blended. However, as was shown in, for example Hartlap et al. (2011) and MacCrann et al. (2017), this can lead to a selection bias on the source galaxies used in the cosmic shear analysis. In particular, this selection bias works to exclude galaxies in high-density environments, which will have a higher convergence than average, resulting in a biased (low) two-point correlation function of galaxy shapes, particularly on small scales (see e.g. fig. 8 in MacCrann et al. 2017). However, recently Hoekstra, Kannawadi & Kitching (2021) showed that this bias can be effectively mitigated using a process of METADETECTION (Sheldon et al. 2020). In addition to this, an effective modelling of baryonic physics will be needed on these scales to fully extract the cosmological information on these small scales (see below).

5 DISCUSSION AND CONCLUSIONS

In this study, we have explored the effects that different extensions to the standard model of cosmology have on the non-linear matter power spectrum, particularly on small scales. This was achieved using a suite of numerical simulations which contain three cosmological variations (in addition to the fiducial ΛCDM realization): (i) a running scalar spectral index (α_s), WDM (M_{WDM}), and SIDM (σ/m). We focus on these extensions in particular as they have previously been shown to suppress small-scale structure and therefore offer a potential means to mitigate small-scale challenges that have been highlighted with the ΛCDM model. We combined the small-scale power spectra extracted from the simulations with the predictions of Halofit on large scales to construct non-linear matter power spectra for the different extensions spanning over six orders of magnitude in wavenumber ($-4 \leq \log_{10}(k\,[h\,\mathrm{Mpc}^{-1}]) \leq 2.4$). These power spectra were used to compute the cosmic shear power spectrum. Finally, We evaluated via synthetic lensing observations (generated with idASK) whether forthcoming Stage-IV lensing surveys (Euclid, LSST, and *NGRST*) will potentially be able to differentiate these extensions from the standard ΛCDM model.

The main findings of our study are as follows:

---

[18]The total integrated SNRs calculated for an *NGRST*-like set-up for each cosmology, using the full covariance matrix are {9.657, 1.733, 2.203, 0.481, 0.082, 1.39, 0.023, 0.009}, in the same order as that displayed in Fig. 4.

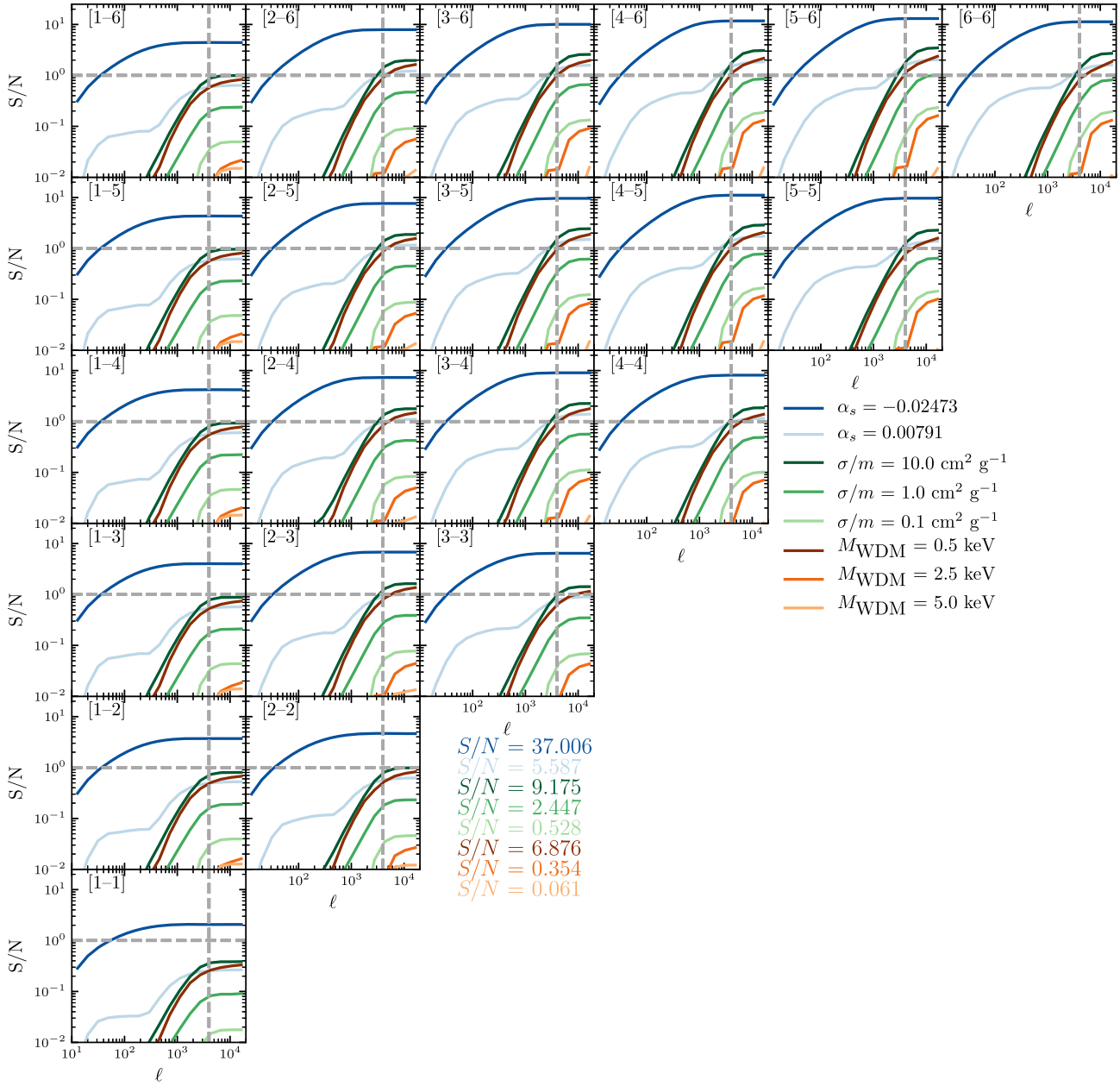[19]Note that we produce 200 unique realizations in order to evaluate the covariance matrix.

Figure 5. Same as Fig. 4, however, here we plot the integrated SNR up to an increased ℓ_{max} of 20 000. To do this we only make use of the Gaussian terms of the covariance matrix, computed using the updated error equation shown by equation (15). This shows that if one was able to push to these smaller angular scales, these observations could potentially be fruitful in putting constraints on the cross-section for interaction of dark matter particles, as well as increased constraining power on the running of the spectral index. The vertical grey dashed line corresponds to the previous $\ell_{max} = 4000$.

(i) A negative running spectral index, WDM, and SIDM are all capable of producing a significant suppression in the non-linear matter power spectrum at late times (Fig. 1). At $z = 0$, this suppression can range from ≈ 5 per cent to ≈ 40 per cent in the case of WDM, at $k \approx 100\ h\,\mathrm{Mpc}^{-1}$. Furthermore, in the case of WDM, the suppression in the matter power spectrum increases with increasing redshift, rising to $\gtrsim 60$ per cent at $z = 1$. A similar trend is seen for the negative running cosmology, although not to the same extent as for WDM. Conversely, the suppression in the matter power spectrum increases with decreasing redshift in an SIDM cosmology. This is due to the strong cores developing in the density profiles of haloes as structures collapse and the relative velocities of particles increase.

(ii) From the different cosmological extensions we have examined, a running scalar spectral index looks the most promising in having a measurable effect on the cosmic shear power spectrum for upcoming surveys such as Euclid. This can be seen, somewhat counter-intuitively, in the enhancement that is produced in the cosmic shear signal on intermediate scales, due to the change in the amplitude of the non-linear power spectrum with respect to ΛCDM on the important scales (Fig. 3). We find that there is a significant signal for both a negatively running spectral index and a positively running one, shown through the SNR (Fig. 4). This illustrates that cosmic shear could be an additional probe which, if used in combination with other probes such as the CMB, could help place strong constraints on the running α_s.

(iii) The constraining power for the other two cosmological extensions of interest here, the mass of the WDM particle and the self-interaction cross-section, is slightly weaker, owing to the fact that these extensions affect only the smallest scales. We have shown that upcoming cosmic shear measurements should be able to rule out SIDM models with $\sigma/m > 10 \, \mathrm{cm}^2 \, \mathrm{g}^{-1}$ or WDM models with thermal relic masses $M_{\mathrm{WDM}} < 0.5 \, \mathrm{keV}$. While these are perhaps weaker constraints than what can be obtained from other methods (e.g. satellite abundances, strong lensing time delay, Lyman α forest), we note that cosmic shear is independent test that has very different systematics than previous small-scale probes, making it still a very worthwhile test of these small-scale extensions. In addition, we have demonstrated that, if it is possible to push to higher multipoles with these experiments, there is the potential that cosmic shear could help place some of the strongest constraints on the SIDM cross-section (σ/m) (Fig. 5).

(iv) Finally, we have illustrated that the standard analytic prediction for the error associated with the cosmic shear cross-correlation power spectrum (between tomographic bins) significantly underestimates the true error that one recovers when computing the same power spectrum from a map which includes a prescription for galaxy shape noise by a factor of around 20 at $\ell \approx 1000$ (Fig. B1). This is because the standard prescription ignores the associated Poisson error. We have introduced a modification to the analytic form of the error which modifies the error associated with cross-correlation power spectrum (equation 15) and is found to bring the analytic errors into much better agreement with our empirical findings based on cross-correlating shape noise maps.

One of the main focuses of upcoming weak lensing surveys, such as Euclid, LSST, and *NGRST*, is to help place constraints on the dark energy equation-of-state parameter w, or the time-varying dark energy equation-of-state parameters w_0, w_a. It is well established that cosmic shear tomography provides a sensitive probe of the growth of structure which, in turn, depends on the evolution of dark energy. However, here we illustrate that the constraining power of cosmic shear measurements also extends beyond dark energy and the other Friedmann parameters. Specifically, we have shown that cosmic shear observations can potentially place constraints on the running of the spectral index, as well as the cross-section for interaction of dark matter particles (SIDM) and the thermal relic mass (WDM). As a final test, we have calculated the source density required to improve constraints on the less extreme cosmological models such as the $M_{\mathrm{WDM}} = 2.5, 5.0$ and $\sigma/m = 0.1 \, \mathrm{cm}^2 \, \mathrm{g}^{-1}$ models. We find that for a sky coverage equal to that of Euclid, the required source density of background galaxies should be $\approx \{85, 525, 60\}$ galaxies arcmin^{-2}, respectively, in order to obtain an SNR of >1, when pushing the analysis to $\ell_{\mathrm{max}} = 20\,000$.

The tests performed in this study consisted of whether forthcoming cosmic shear surveys could distinguish (on the basis of SNR) a number of extensions, with specific parameter values, from the baseline ΛCDM model. The tests were particularly simple in that, in most cases, all of the cosmological parameters were held fixed, apart from the new parameters describing the extension. A more realistic test would be to allow the various parameters to be free and to marginalize over them when estimating the uncertainty (and potentially bias) in the recovered extension parameters (e.g. WDM particle mass). However, to achieve this requires many more simulations than produced here and potentially sophisticated methods (e.g. emulators) for interpolating the results for arbitrary cosmological

parameter values. While our promising results demonstrate that this is clearly worthwhile, it is a large undertaking and we leave this for future work.

In this study, we focused on the constraining power of future cosmic shear measurements. However, there are additional complementary two-point statistics that may be helpful in placing constraints on these cosmological extensions. In particular there is galaxy clustering, which describes the clustering between lens galaxies, and galaxy–galaxy lensing, which describes the overdensity of mass around lens galaxies. Together with cosmic shear, these 3×2pt statistics have been shown to help place tighter constraints on the cosmological parameters Ω_{m} and σ_8 (see e.g. fig. 6 of DES Collaboration 2021). However, combining all three of these probes is beyond the scope of this work, which focused on the constraining power of cosmic shear alone. In future work, we will explore how the constraints in this paper change when combining these 3×2 pt statistics.

An important caveat to the work we have presented is that the simulations we used neglected the effects of baryonic physics. Previous work has demonstrated that when numerical simulations include complex galaxy formation physics, such as feedback from supernovae and active galactic nuclei, they can produce relatively large effects (typically 5–20 per cent) on the non-linear matter power spectrum (e.g. van Daalen et al. 2011; Chisari et al. 2018; Springel et al. 2018; van Daalen, McCarthy & Schaye 2020). Furthermore, these effects due to galaxy formation physics have been shown to be visible in the cosmic shear two-point correlation functions at a similar level (Semboloni et al. 2011). Our work has focused mostly on small scales where, as opposed to baryons suppressing the power spectrum via the expulsive effects of feedback on relatively large scales (the focus of most previous studies), it is more likely that cooling and star formation will lead to an *enhancement* in the power spectrum. Regardless of whether baryons produce a suppression or an enhancement, the effects may be degenerate with the cosmological extensions we have examined here. As discussed in Stafford et al. (2020b), ultimately what is required is a systematic and simultaneous exploration of the (uncertain) impact of baryons and cosmological variations on small scales and an understanding of how these effects propagate through to observables on small scales such as the cosmic shear power spectrum.

In closing, to help mitigate potential degeneracies, recent studies have shown the power of combining complementary probes when placing constraints on additional cosmological (and baryon) parameters. For example, Enzi et al. (2020) and Nadler et al. (2021) have illustrated the power of combining multiple probes (Lyman α forest, strong lensing and the abundance of Milky Way satellites) in placing constraints on the WDM particle mass. Measurements of cosmic shear on small scales provide an important new tool in this regard and one that has very different systematic uncertainties from currently used small-scale probes. Forthcoming Stage-IV lensing surveys therefore offer a promising new window to study cosmological and galaxy formation physics on small scales.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Alam S. et al., 2017, MNRAS, 470, 2617

Ali-Haïmoud Y., Bird S., 2013, MNRAS, 428, 3375

Asgari M. et al., 2021, A&A, 645, A104

Banerjee A., Adhikari S., Dalal N., More S., Kravtsov A., 2020, J. Cosmol. Astropart. Phys., 2020, 024

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002, MNRAS, 333, 156

Bernardeau F., 1998, A&A, 338, 375

Bird S., 2017, Astrophysics Source Code Library, record ascl:1706.006

Blandford R. D., Saust A. B., Brainerd T. G., Villumsen J. V., 1991, MNRAS, 251, 600

Bode P., Ostriker J. P., Turok N., 2001, ApJ, 556, 93

Bond J. R., Szalay A. S., 1983, ApJ, 274, 443

Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, 40

Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517

Chisari N. E. et al., 2018, MNRAS, 480, 3962

Colín P., Avila-Reese V., Valenzuela O., 2000, ApJ, 542, 622

Correa C. A., 2021, MNRAS, 503, 920

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

DES Collaboration, 2021, preprint (arXiv:2105.13549)

Dooley G. A., Peter A. H., Vogelsberger M., Zavala J., Frebel A., 2016, MNRAS, 461, 710

Efstathiou G., 1992, MNRAS, 256, 43P

Eifler T. et al., 2021, MNRAS, 507, 1514

Eisenstein D. J. et al., 2005, ApJ, 633, 560

Elbert O. D., Bullock J. S., Garrison-Kimmel S., Rocha M., Oñorbe J., Peter A. H. G., 2015, MNRAS, 453, 29

Enzi W. et al., 2020, MNRAS, 506, 5848

Euclid Collaboration, 2019, MNRAS, 484, 5509

Euclid Collaboration, 2020a, A&A, 642, A191

Euclid Collaboration, 2020b, A&A, 644, A31

Flores R. A., Primack J. R., 1994, ApJ, 427, L1

Garrison-Kimmel S., Horiuchi S., Abazajian K. N., Bullock J. S., Kaplinghat M., 2014, MNRAS, 444, 961

Gatti M. et al., 2021, MNRAS, 504, 4312

Giannantonio T., Porciani C., Carron J., Amara A., Pillepich A., 2012, MNRAS, 422, 2854

Gilman D., Birrer S., Treu T., Nierenberg A., Benson A., 2019, MNRAS, 487, 5721

Gnedin O. Y., Ostriker J. P., 2001, ApJ, 561, 61

Guth A. H., 1981, Phys. Rev. D, 23, 347

Hamana T., Colombi S. T., Thion A., Devriendt J. E. G. T., Mellier Y., Bernardeau F., 2002, MNRAS, 330, 365

Hamana T. et al., 2020, PASJ, 72, 16

Hartlap J., Hilbert S., Schneider P., Hildebrandt H., 2011, A&A, 528, A51

Harvey D., Massey R., Kitching T., Taylor A., Tittley E., 2015, Science, 347, 1462

Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, ApJ, 715, 104

Hilbert S., Hartlap J., Schneider P., 2011, A&A, 536, A85

Hoekstra H., Kannawadi A., Kitching T. D., 2021, A&A, 646, A124

Hsueh J.-W., Enzi W., Vegetti S., Auger M. W., Fassnacht C. D., Despali G., Koopmans L. V. E., McKean J. P., 2019, MNRAS, 3059, 3047

Hu W., 1999, ApJ, 522, L21

Hu W., 2000, Phys. Rev. D, 62, 043007

Hubert J., Schneider A., Potter D., Stadel J., 2021, preprint (arXiv:2104.07675)

Iršič V. et al., 2017, Phys. Rev. D, 96, 023522

Jethwa P., Erkal D., Belokurov V., 2018, MNRAS, 473, 2060

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

Kahlhoefer F., Schmidt-Hoberg K., Kummer J., Sarkar S., 2015, MNRAS, 452, L54

Kaiser N., 1992, ApJ, 388, 272

Kaiser N., 1998, ApJ, 498, 26

Kennedy R., Frenk C., Cole S., Benson A., 2014, MNRAS, 442, 2487

Kiessling A. et al., 2015, Space Sci. Rev., 193, 67

Kilbinger M., 2015, Rep. Prog. Phys., 78, 086901

Kilbinger M. et al., 2017, MNRAS, 472, 2126

Kim S. Y., Peter A. H., Wittman D., 2017, MNRAS, 469, 1414

Kirk D. et al., 2015, Space Sci. Rev., 193, 139

Klypin A., Prada F., 2018, MNRAS, 478, 4602

Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82

Kosowsky A., Turner M. S., 1995, Phys. Rev. D, 52, R1739

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

Limber D. N., 1953, ApJ, 117, 134

Lovell M. R. et al., 2012, MNRAS, 420, 2318

Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, MNRAS, 439, 300

MacCrann N. et al., 2017, MNRAS, 465, 2567

Mandelbaum R., 2018, ARA&A, 56, 393

Markovic K., Bridle S., Slosar A., Weller J., 2011, J. Cosmol. Astropart. Phys., 2011, 022

Mashchenko S., Wadsley J., Couchman H. M., 2008, Science, 319, 174

McCarthy I. G., Bird S., Schaye J., Harnois-Deraps J., Font A. S., van Waerbeke L., 2018, MNRAS, 476, 2999

Meneghetti M., Yoshida N., Bartelmann M., Moscardini L., Springel V., Tormen G., White S. D., 2001, MNRAS, 325, 435

Miralda-Escude J., 1991, ApJ, 380, 1

Miralda-Escude J., 2002, ApJ, 564, 60

Moore B., 1994, Nature, 370, 629

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Nadler E. O., Gluscevic V., Boddy K. K., Wechsler R. H., 2019, ApJ, 878, L32

Nadler E. O. et al., 2020, ApJ, 893, 48

Nadler E. O., Birrer S., Gilman D., Wechsler R. H., Du X., Benson A., Nierenberg A. M., Treu T., 2021, ApJ, 917, 7

Palanque-Delabrouille N. et al., 2015, J. Cosmol. Astropart. Phys., 2015, 011

Palanque-Delabrouille N., Yèche C., Schöneberg N., Lesgourgues J., Walther M., Chabanier S., Armengaud E., 2020, J. Cosmol. Astropart. Phys., 2020, 038

Peñarrubia J., Benson A. J., Walker M. G., Gilmore G., McConnachie A. W., Mayer L., 2010, MNRAS, 406, 1290

Peter A. H. G., Rocha M., Bullock J. S., Kaplinghat M., 2013, MNRAS, 430, 105

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

Planck Collaboratio VI, 2020, A&A, 641, A6

Pontzen A., Governato F., 2012, MNRAS, 421, 3464

Power C., Knebe A., 2006, MNRAS, 370, 691

Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., 2008, ApJ, 679, 1173

Robertson A., Massey R., Eke V., 2017, MNRAS, 465, 569

Robertson A., Harvey D., Massey R., Eke V., McCarthy I. G., Jauzac M., Li B., Schaye J., 2019, MNRAS, 488, 3646

Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Oñorbe J., Moustakas L. A., 2013, MNRAS, 430, 81

Sagunski L., Gad-Nasr S., Colquhoun B., Robertson A., Tulin S., 2021, J. Cosmol. Astropart. Phys., 2021, 024

Savitzky A., Golay M. J. E., 1964, Anal. Chem., 36, 1627

Sawala T. et al., 2016, MNRAS, 457, 1931

Schmidt F., 2008, Phys. Rev. D, 78, 043002

Schneider P., van Waerbeke L., Mellier Y., 2002, A&A, 389, 729

Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020

Sheldon E. S., Becker M. R., MacCrann N., Jarvis M., 2020, ApJ, 902, 138

Smith R. E., Angulo R. E., 2019, MNRAS, 486, 1448

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

Spergel D. N., Steinhardt P. J., 2000, Phys. Rev. Lett., 84, 3760

Spergel D. et al., 2015, preprint (arXiv:1503.03757)

Springel V., 2005, MNRAS, 364, 1105

Springel V. et al., 2005, Nature, 435, 629

Springel V. et al., 2018, MNRAS, 475, 676

Stafford S. G., McCarthy I. G., Crain R. A., Salcido J., Schaye J., Font A. S., Kwan J., Pfeifer S., 2020a, MNRAS, 493, 676

Stafford S. G., Brown S. T., McCarthy I. G., Font A. S., Robertson A., Poole-McKenzie R., 2020b, MNRAS, 497, 3809

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

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

van Daalen M. P., McCarthy I. G., Schaye J., 2020, MNRAS, 491, 2424

Verde L., Treu T., Riess A. G., 2019, Nat. Astron., 3, 891

Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534

Viel M., Markovič K., Baldi M., Weller J., 2012, MNRAS, 421, 50

Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, PRD, 88, 1

Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740

Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, ApJ, 827, L23

Wittman D., Golovich N., Dawson W. A., 2018, ApJ, 869, 104

Xavier H. S., Abdalla F. B., Joachimi B., 2016, MNRAS, 459, 3693

Zavala J., Vogelsberger M., Walker M. G., 2013, MNRAS, 431, 20

Zhan H., Tyson J. A., 2018, Rep. Prog. Phys., 81, 066901

APPENDIX A: RESOLUTION AND BOX SIZE STUDY

In Section 2.3.1, we detail our method for constructing the non-linear matter power spectrum spanning a large range of k modes, which is done using the ratio of the matter power spectrum extracted from the numerical simulations with respect to the result from the ΛCDM simulations. As we show here, splicing ratios of power spectra, rather than absolute power spectra, is more robust to changes in resolution and box size. We do this using an additional sub-suite of simulations, examining the running spectral index cosmological extensions, which consists of various box sizes and resolutions. First, there are two set-ups that are used in the analysis detailed in the paper, which have box sizes of 25 and 400 h^{-1} Mpc, each with 1024^3 particles. We then compare the non-linear matter power spectrum extracted from these simulations to that extracted from simulations which are a step down in resolution. These simulations are 25 and 100 h^{-1} Mpc on a side and have 512^3 and 256^3 particles respectively. Thus, this allows us to explore the effects of both box size and resolution on both the absolute matter power spectrum and the ratio with respect to ΛCDM.

Fig. A1 shows the non-linear matter power spectrum for two running of the spectral index models. In the left-hand panel, we show how the simulation box size affects the absolute power spectra (top) and the ratio with respect to a complementary ΛCDM simulation of the same box size (bottom), both at fixed resolution. In the right-hand panel, we show how resolution affects these two statistics at fixed box size. Therefore, it is important to note that when comparing the solid to dashed lines (or dotted) in the bottom panels, both the numerator and denominator have changed. Comparing the different linestyles allows us to assess the convergence of the ratio of power spectra from the running cosmologies with respect to a ΛCDM power spectrum as a function of varying box size at fixed resolution, or vice versa.

One can see that while the absolute power spectra tend to disagree with each other on the largest scales sampled in the box, the ratio is a much better converged quantity. This is particularly relevant for the smaller volume boxes, where the size of the simulations cause an almost order-of-magnitude suppression in the matter power spectrum on the largest scales sampled by the simulated volume (comparing the simulation curves to the Halofit curve). Conversely, if we compare the ratios at a k-scale of around 10 h Mpc^{-1}, one can see that these agree to within a few per cent across the varying box sizes and resolutions.
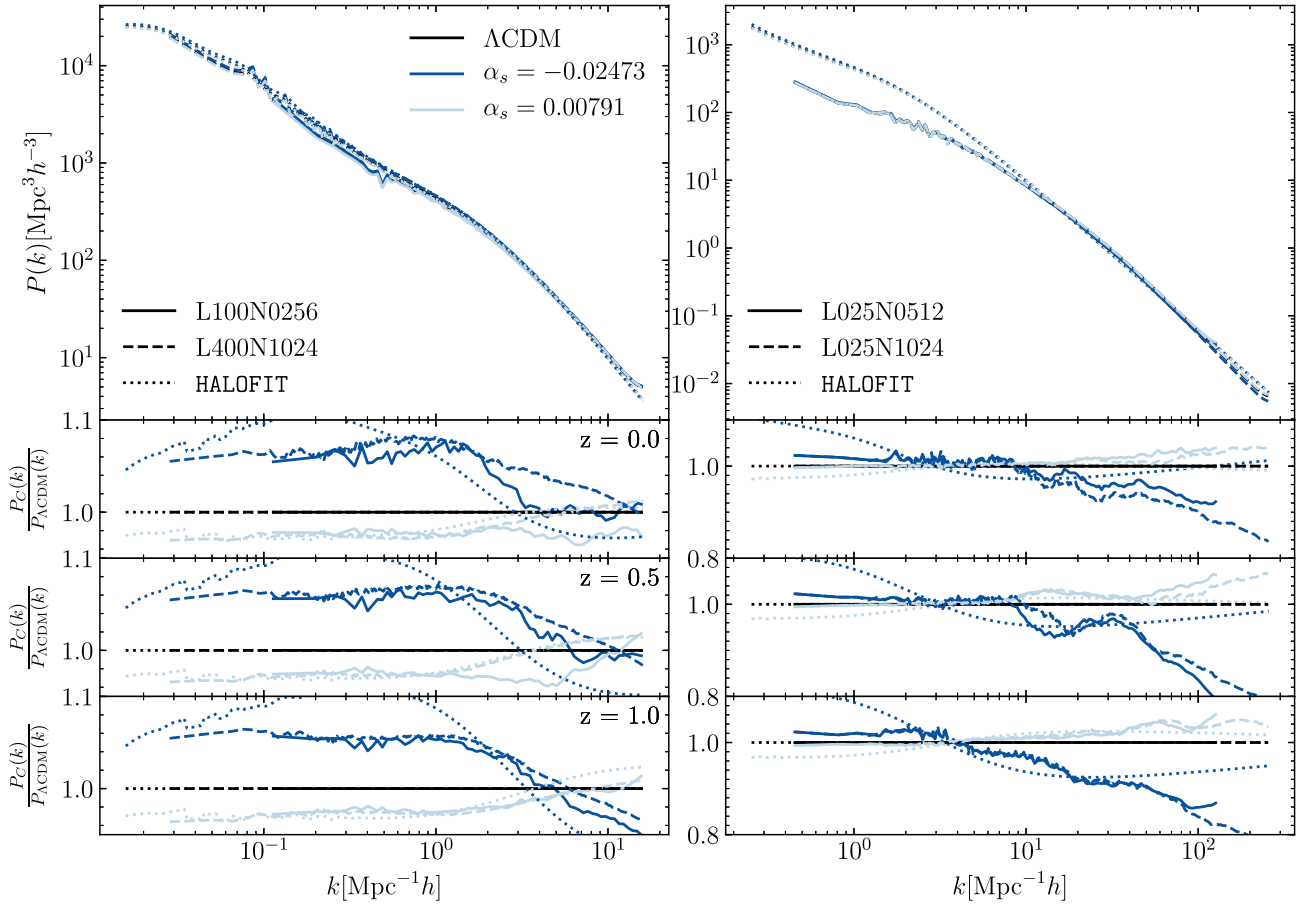
Figure A1. Testing how both box size and resolution affect the ratios of the power spectrum extracted from a cosmology with a running spectral index with respect to the ΛCDM result. Left-hand panels show the results for fixed resolution and varying box size, while the right-hand panels show the opposite. The top panel in each case shows the absolute power spectra at $z = 0$. The three sub-panels below show the ratio as a function of redshift. The ratio is computed with respect to the power spectrum from the ΛCDM simulation with the same box size and resolution as in the running spectral index case.

The ratio of the power spectra appears to be well converged for the various box sizes and resolutions examined here which motivates us to combine the ratios over the extended k-scale, as described in Section 2.3.1, rather than the absolute power spectra.

APPENDIX B: ANALYTIC ERROR ON CROSS POWER SPECTRA

As discussed in Section 3.1, there are two sources of error associated with measurements of the cosmic shear power spectrum (ignoring other systematic errors which exist such as intrinsic alignment errors), these being cosmic variance and Poisson noise. It is commonplace to assume that Poisson noise only affects the uncertainty in the auto-correlation power spectrum. However, as we show here, cross-power spectra can still have a significant Poisson noise term.

As shown in Fig. 3, we find there to be a significant difference in the error bars on small angular scales (large multipoles) when comparing equation (10) to the error bars we derive from synthetic weak lensing maps using FLASK. The level of disagreement between the two sets of error bars is more clearly illustrated in Fig. B1, which shows the ratio of the errors extracted directly from the FLASK power spectra to the errors computed using equation 10. Note that for clarity we only show here the result for the bottom three panels of Fig. 3, although the results presented here is true

for the other auto- and cross-correlation spectra. This shows how well the two sets of errors agree with one another in the case of the auto-correlation power spectra, whereas, examining the red curve in the cross-correlation power spectrum panel, the error bars diverge significantly at high multipoles, due to equation 10 having no treatment for Poisson noise on the cross-correlation power spectrum.

The fiducial error associated with the cosmic shear power spectrum (shown in equation 10) is unable to capture the residual error which exists on the cross-correlation power spectrum. This motivates us to produce a more general formula which is able to capture this residual error without having to rerun FLASK for each possible tomographic set-up. For this reason, we have run a set of FLASK noise-only realizations (with the weak lensing signal due to gravity removed) to derive a functional form for the residual Poisson noise error on the cross-correlation power spectrum. The maps were produced using the same tomographic set-up described in Section 4.1, for varying levels of the source density of galaxies. In total, we produced four sets of realizations sampling: $n_s = 15, 30, 45, 60$ galaxies arcmin^{-2}, with each set having 100 realizations of each tomographic bin.
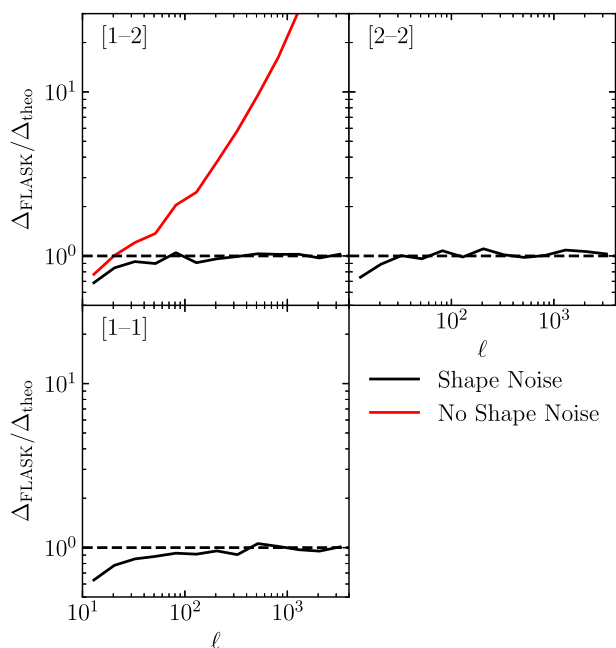
We then computed the auto- and cross-correlation power spectra of these maps, along with cross-correlation power spectra of maps with varying source densities (to test the case where two tomographic bins may not have the same effective number density of galaxies).

We find that an additional term, which only contributes to the cross-correlation power spectrum, is required. See equation (15). This term includes a multiplicative combination of the two effective source densities in each tomographic bin and is able to reproduce the additional error on the cross exceptionally well. This is illustrated again in Fig. B1 in the cross-correlation power spectrum panel (top left), where the black line now shows the ratio of the theoretical error bars, now computed using equation (15), with the error bars extracted from FLASK. Note that here we show the results from the full analysis (i.e. including an intrinsic lensing signal, as well as the noise). Note that the slight disagreement that exists between the errors on large angular scales is strongly dependent on the sky coverage of the survey. We computed this same test where we had a Euclid-like galaxy sample, but a full sky survey and this brought the errors into excellent agreement even at low multipoles.

Figure B1. The ratio of the errors extracted directly from the multiple realizations of a given ΛCDM auto/cross-correlation power spectrum as computed by FLASK with respect to the theoretical error bars computed using equations (10) and (15). Note the tomographic bins shown here are the same as those in the previous plots, however, we do not show the full tomographic set-up for brevity as we only need to focus on one cross-correlation power spectrum (with the result being the same for the rest). The black line shows the result when one includes a prescription for the error associated with Poisson noise in the theoretical calculation of the error on the auto/cross-correlation power spectrum, i.e. computed using equation (15), with the red line being the result if one does not, i.e. computed using equation (10).

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