Universiteit Leiden
The Netherlands

FLAMINGO: combining kinetic SZ effect and galaxy–galaxy lensing measurements to gauge the impact of feedback on large-scale structure

McCarthy, I.G.; Amon, A.; Schaye, J.; Schaan, E.; Angulo, R.E.; Salcido, J.; ... ; Daalen, M.P. van

Citation
McCarthy, I. G., Amon, A., Schaye, J., Schaan, E., Angulo, R. E., Salcido, J., ... Daalen, M. P. van. (2025). FLAMINGO: combining kinetic SZ effect and galaxy–galaxy lensing measurements to gauge the impact of feedback on large-scale structure. *Monthly Notices Of The Royal Astronomical Society, 540*(1), 143-163. doi:10.1093/mnras/staf731

Version:	Publisher's Version
License:	Creative Commons CC BY 4.0 license
Downloaded from:	https://hdl.handle.net/1887/4290492

Note: To cite this publication please use the final published version (if applicable).

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **540,** 143–163 (2025) https://doi.org/10.1093/mnras/staf731
Advance Access publication 2025 May 2

FLAMINGO: combining kinetic SZ effect and galaxy–galaxy lensing measurements to gauge the impact of feedback on large-scale structure

Ian G. McCarthy [1]★ Alexandra Amon,[2] Joop Schaye [3] Emmanuel Schaan,[4,5] Raul E. Angulo [6,7] Jaime Salcido [1] Matthieu Schaller [3,8] Leah Bigwood,[9,10] Willem Elbers [11] Roi Kugel [3] John C. Helly,[11] Victor J Forouhar Moreno [3] Carlos S. Frenk,[11] Robert J. McGibbon [3] Lurdes Ondaro-Mallea[6,12] and Marcel P. van Daalen [3]

[1]*Astrophysics Research Institute, Liverpool John Moores University, Liverpool, L3 5RF, UK*
[2]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[3]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[4]*Kavli Institute for Particle Astrophysics and Cosmology, 382 Via Pueblo Mall Stanford, CA 94305-4060, USA*
[5]*SLAC National Accelerator Laboratory, 2575 Sand Hill Road Menlo Park, CA 94025, USA*
[6]*Donostia International Physics Center (DIPC), Paseo Manuel de Lardizabal, 4, E-20018 Donostia-San Sebastian, Guipuzkoa, Spain*
[7]*IKERBASQUE, Basque Foundation for Science, E-48013 Bilbao, Spain*
[8]*Lorentz Institute for Theoretical Physics, Leiden University, PO box 9506, NL-2300 RA Leiden, the Netherlands*
[9]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[10]*Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 OHA, UK*
[11]*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham, DH1 3LE, UK*
[12]*Department of Theoretical Physics, University of the Basque Country UPV/EHU, Bilbao, E-48080, Spain*

Accepted 2025 April 30. Received 2025 April 2; in original form 2024 October 25

ABSTRACT

Energetic feedback processes associated with accreting supermassive black holes can expel gas from massive haloes and significantly alter various measures of clustering on ∼Mpc scales, potentially biasing the values of cosmological parameters inferred from analyses of large-scale structure (LSS) if not modelled accurately. Here, we use the state-of-the-art FLAMINGO suite of cosmological hydrodynamical simulations to gauge the impact of feedback on large-scale structure by comparing to *Planck* + ACT stacking measurements of the kinetic Sunyaev–Zel'dovich (kSZ) effect of SDSS BOSS galaxies. We make careful like-with-like comparisons to the observations, aided by high precision KiDS and DES galaxy–galaxy lensing measurements of the BOSS galaxies to inform the selection of the simulated galaxies. In qualitative agreement with several recent studies using dark matter only simulations corrected for baryonic effects, we find that the kSZ effect measurements prefer stronger feedback than predicted by simulations which have been calibrated to reproduce the gas fractions of low redshift X-ray-selected groups and clusters. We find that the increased feedback can help to reduce the so-called S_8 tension between the observed and CMB-predicted clustering on small scales as probed by cosmic shear (although at the expense of agreement with the X-ray group measurements). However, the increased feedback is only marginally effective at reducing the reported offsets between the predicted and observed clustering as probed by the thermal SZ (tSZ) effect power spectrum and tSZ effect – weak lensing cross-spectrum, both of which are sensitive to higher halo masses than cosmic shear.

Key words: methods: numerical – galaxies: clusters: general – galaxies: formation – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

Measurements of how matter is spatially clustered in the Universe can be used to place strong constraints on cosmological models, including allowing one to test theories of gravity as well as the natures of dark matter and dark energy. The standard model of cosmology, the so-called lamba cold dark matter (ΛCDM) model, is based on General Relativity and assumes dark matter is composed of relatively 'cold' and weakly interacting particles and that dark energy takes the form of a cosmological constant. This relatively simple model, which only has six adjustable parameters, describes a wealth of large-scale cosmological data remarkably well. Nevertheless, there are some notable anomalies which may be hinting at deviations from the predictions of ΛCDM (Peebles 2024). One of these is the so-called S_8 tension, where S_8 is defined as $\sigma_8\sqrt{\Omega_{\rm m}/0.3}$, where $\Omega_{\rm m}$ represents the present-day matter density, and σ_8 is the linearly evolved variance of the current matter density field filtered on an

★ E-mail: i.g.mccarthy@ljmu.ac.uk

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

8 Mpc h^{-1} scale. Several low-redshift observations of the large-scale structure (LSS), including measurements of total matter clustering via cosmic shear (e.g. Heymans et al. 2021; Abbott et al. 2022; Amon et al. 2023), yield a best-fitting value of S_8 that is smaller than, and in mild (≈ 1.5–3σ) tension with, the predictions of the standard model based on parameter values specified by the primary CMB and BAO (e.g. Planck Collaboration VI 2020). It is notable that this tension has persisted in some comparisons for nearly a decade, starting with the first *Planck* data release, and encompasses several independent probes that consistently show discrepancies of similar significance and in the same direction (see McCarthy et al. 2018).

Various possible solutions have been put forward to potentially reconcile the low-redshift LSS observations with the primary CMB and BAO data. These include unidentified or mischaracterized systematic uncertainties in the LSS observations, or possibly even in the primary CMB measurements. On the theoretical side, LSS tests of cosmology often probe deep into the non-linear regime, and it has been proposed that new physics on those scales could explain the tension (e.g. Amon & Efstathiou 2022; Preston, Amon & Efstathiou 2023), such as new dark sector models (e.g. Rogers et al. 2023; Elbers et al. 2025), or a mis-understanding of galaxy formation effects (e.g. McCarthy et al. 2018). Modelling non-linear scales necessitates cosmological simulations (or models calibrated on such simulations) to predict the clustering of matter on small scales and at late times. Additionally, as matter collapses to form self-gravitating haloes, densities increase to the point where radiative cooling of the gas becomes efficient, leading to further collapse and galaxy formation. This process is accompanied by various energetic feedback mechanisms related to star formation and the accretion of matter onto supermassive black holes.

Modern cosmological hydrodynamical simulations, particularly of volumes that are large enough for clustering studies, cannot resolve all of the relevant scales in order to include such feedback processes in an ab initio way (see the discussion in Schaye et al. 2015). These processes must be included using subgrid prescriptions and it has been demonstrated that certain predictions of the simulations are sensitive to the details of the feedback implementations. One of these is the fraction of baryons that are retained by (and how they are radially distributed around) massive galaxy groups and clusters (e.g. Le Brun et al. 2014; Planelles et al. 2014; Velliscig et al. 2014; McCarthy et al. 2017; Henden et al. 2018), which dominate the total matter clustering signal (van Daalen & Schaye 2015; Mead et al. 2020). AGN feedback is energetically capable of ejecting large quantities of baryons from haloes and this reduces the amplitude of the clustering signal on small scales (e.g. van Daalen et al. 2011; Mummery et al. 2017; Springel et al. 2018; Chisari et al. 2019; van Daalen, McCarthy & Schaye 2020). If such effects are not accounted for, or are included inaccurately, this can lead to an incorrect prediction for the matter clustering signal and potentially bias the recovered cosmological parameters when fitting to the observed clustering (e.g. Semboloni et al. 2011; Semboloni, Hoekstra & Schaye 2013; Debackere, Schaye & Hoekstra 2020; Schneider et al. 2020; Castro et al. 2021).

In several previous studies, we demonstrated that there is a strong quantitative link between the suppression of the matter power spectrum and the baryon fractions of groups and clusters (Semboloni et al. 2013; van Daalen et al. 2020; Salcido et al. 2023). External observations of the hot gas, which dominates the baryon budget of groups and clusters, can be used to help to evaluate the impact of baryon physics on the matter clustering in cosmological analyses. Traditionally, X-ray observations have provided the highest quality measurements of the state of the hot gas. When combined with measurements of the total mass, either via the X-ray observations themselves (under the assumption of hydrostatic equilibrium) or from other probes such as weak lensing measurements, the hot gas mass fractions can be inferred (e.g. Sun et al. 2009; Lovisari, Reiprich & Schellenberger 2015; Eckert et al. 2016; Akino et al. 2022). Recent cosmological hydrodynamical simulation campaigns, including BAHAMAS (McCarthy et al. 2017, 2018), FABLE (Henden et al. 2018; Henden, Puchwein & Sijacki 2020), and FLAMINGO (Kugel et al. 2023; Schaye et al. 2023), have used X-ray-based measurements of the hot gas fractions to help calibrate the efficiencies of feedback in the simulations, although variations in the feedback efficiencies about the fiducial calibrated models were also considered in BAHAMAS and FLAMINGO. Comparisons of these simulations to LSS observables, including the auto- and cross-power spectra of cosmic shear, the thermal Sunyaev–Zel'dovich effect, and CMB lensing, suggest that baryon feedback is incapable of resolving the S_8 tension (McCarthy et al. 2018, 2023). This is consistent with other recent studies that have used different approaches, such as the baryonification formalism, together with X-ray measurements as constraints on the baryon physics (e.g. Grandis et al. 2024).

While X-ray measurements remain a valuable source of information on the hot gas properties of groups/clusters, particularly with new insights coming in from eROSITA data (e.g. Bulbul et al. 2024; Popesso et al. 2024), observations of the thermal and kinetic Sunyaev–Zel'dovich (tSZ and kSZ) effects around groups and clusters are yielding increasingly high signal-to-noise measurements when stacking analyses are employed. The SZ effects have some advantages over the X-ray measurements. First, the amplitudes of the SZ effects are independent of redshift, in principle allowing one to measure the hot gas over a wide range of cosmic times [by contrast the X-ray surface brightness fades as $(1+z)^4$ and measurements of galaxy groups in particular are generally confined to relatively low redshifts, $z \lesssim 0.3$]. In addition, the tSZ effect yields the total thermal energy density of the gas when integrated over the surface area of the cluster while the kSZ effect yields the gas mass (or gas momentum), both of which are more directly linked to the impact of baryon physics on the matter clustering than the X-ray emission. Note that high angular resolution data are required, so that the groups/clusters under study can be spatially resolved, allowing for measurements of their gaseous properties on the scales where feedback is important (e.g. Le Brun, McCarthy & Melin 2015; Yang et al. 2022).

Planck data have been used to produce full-sky maps of the tSZ effect, enabling the stacking of many sources and detections over a wide range of masses (e.g. Planck Collaboration XI 2013; Greco et al. 2015). However, the typical 10 arcmin angular resolution of *Planck* prevents resolved measurements (radial profiles) for all but most massive and nearby sources. The Atacama Cosmology Telescope (ACT) and the South Pole Telescope (SPT) offer almost an order of magnitude better angular resolution than *Planck* and higher sensitivity measurements, although with a smaller sky coverage and a limited number of frequencies, which can be key to removing foreground and background contaminants (e.g. dust in the Galaxy, clustered radio sources, the cosmic infrared background, or CIB). It is now becoming common place to combine *Planck* data with ACT or SPT data to study groups and clusters (e.g. Aghanim et al. 2019; Melin et al. 2021; Bleem et al. 2022), taking advantage of *Planck*'s multifrequency measurements and sensitivity to large scales together with ACT and SPT's higher angular resolutions and sensitivities.

Schaan et al. (2021) have used a combination of *Planck* and ACT data to measure the stacked kSZ profiles of SDSS BOSS galaxies (Ahn et al. 2014). The kSZ effect is proportional to the line of sight radial (peculiar) velocity of galaxies and, statistically speaking, we

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

are as likely to find a galaxy moving towards us as away from us (if we subtract our motion relative to the CMB), so simply stacking the CMB temperature maps of galaxies without regard for the direction of motion would imply that the kSZ effect will cancel out. To address this, Schaan et al. (2021) used the 3D clustering signal of BOSS galaxies to reconstruct the implied linear velocity field, allowing the galaxies to be weighted according to their predicted velocities in the stack. Strong detections were made for both the CMASS and LOWZ samples from the BOSS survey. The derived radial profiles show that the hot gas around the BOSS galaxies is very extended in comparison to the expectation for the dark matter distribution from dissipationless simulations, implying that non-gravitational processes (particularly feedback) have significantly altered the gas distribution. Using a simple halo model formalism to model the kSZ effect profiles suggests a significant impact on the matter clustering on small scales (Amodeo et al. 2021) with possible implications for the S_8 tension (Amon et al. 2023).

Recently, Schneider et al. (2022) and Bigwood et al. (2024) have used the baryonification formalism (e.g. Schneider & Teyssier 2015; Schneider et al. 2019; Aricò et al. 2021, 2023) to jointly model the Schaan et al. (2021) kSZ effect measurements together with measurements of the cosmic shear correlation functions. The baryonification formalism employs a parametric model to describe the radial distributions of the matter components (gas, stars, dark matter) of haloes which it uses to radially perturb the mass distribution of dissipationless ('dark matter only') simulations. The parameters of the model can either be determined by fitting to external data sets (e.g. X-ray and SZ effect measurements) or they can be left free in the fit to cosmological observables such as the cosmic shear correlation functions. Schneider et al. (2022) and Bigwood et al. (2024) have shown that the inclusion of the kSZ effect measurements in the cosmic shear analysis better constrains the parameters which characterize the impact of baryons. Furthermore, Bigwood et al. (2024) demonstrated that the joint analysis of cosmic shear and kSZ effect data prefer a stronger impact of baryons compared to that implied by models which are fitted to resolved X-ray observations (see also Salcido & McCarthy 2024).

A major source of uncertainty in the modelling of the Schaan et al. (2021) kSZ effect observations is the choice of halo mass of the sample. The amplitude of the kSZ effect scales proportionally with the gas mass, and therefore (approximately) proportionally with the total halo mass. Given that measurements of the mean halo mass of the BOSS CMASS galaxies vary by nearly an order of magnitude between previous studies (see the discussion in Bigwood et al. 2024), marginalizing over this uncertainty significantly weakens the constraints on the feedback model parameters and in turn weakens the cosmological constraints. Note also that the LOWZ sample has no reliable mean halo mass measurement to date, which is the main reason why the stacked kSZ effect measurement from this sample has not yet been utilized as a constraint on the baryon modelling. In addition, previous attempts to model the signal have implicitly assumed that the BOSS CMASS sample is composed entirely of central galaxies with no mis-centring of the galaxies with respect to the total gravitating mass or hot gas distributions. A main reason for adopting these assumptions, whose impact is presently difficult to assess, is that there is currently no straightforward way to select a realistic mock BOSS-like sample in the context of the baryonification formalism.

In this study, we address these issues using self-consistent full cosmological hydrodynamical simulations. Specifically, we use full-sky light-cones constructed using the FLAMINGO hydro simulations, which systematically vary the important parameters (and parametrizations) controlling the impact of baryon feedback. We also explore the cosmology dependence of the signal. We select galaxy populations from the simulations that are constructed to carefully match the observed galaxy–galaxy lensing profiles of the BOSS CMASS and LOWZ samples, as recently measured by Amon et al. (2023) using high-quality Kilo Degree Survey (KiDS 1000) and Dark Energy Survey Year Three (DES Y3) data. We will show that the galaxy–galaxy lensing measurements place very stringent constraints (better than 10 per cent at 2σ) on the mean halo masses of these samples. Using the lensing-selected galaxy populations, we extract the simulation kSZ effect profiles following the methodology of Schaan et al. (2021), stacking the profiles at the locations of the galaxies in the light cone-based maps in a way that is faithful to that done for the real observations. We quantify the impact of satellite contamination on the derived galaxy–galaxy lensing and kSZ effect profiles, showing it to be non-negligible for both. Finally, in agreement with several recent studies, we will show that the kSZ effect measurements imply a stronger impact of feedback relative to that adopted in the fiducial FLAMINGO model (which was calibrated on X-ray observations of low-redshift groups) and we discuss the implications of this finding for the S_8 tension.

The paper is organized as follows. In Section 2, we describe the FLAMINGO simulation suite and the observations employed in this study. In Section 3, we describe our methodology, including the selection of galaxies in the simulations and the derivation of lensing and kSZ effect profiles from light cone-based maps. In Section 4, we present our main results, including an examination of the feedback and cosmological dependencies of the derived lensing and kSZ effect profiles, in comparison with the Amon et al. (2023) and Schaan et al. (2021) measurements. In Section 5, we discuss the implications of our findings for feedback modelling and the S_8 tension. In Section 6, we summarize our main findings and conclude.

2 SIMULATION AND OBSERVATIONAL DATA SETS

2.1 FLAMINGO simulations

We provide here a summary of the FLAMINGO simulations, referring the reader to Schaye et al. (2023) and Kugel et al. (2023) for in depth presentations.

The FLAMINGO suite consists of 16 hydrodynamical simulations presented in Schaye et al. (2023), two decaying dark matter variants in Elbers et al. (2025), plus two new hydrodynamical simulations introduced here (see below), and 12 gravity-only simulations. The suite has variations in resolution, box size, subgrid modelling, and cosmology. We mostly use the intermediate resolution ($m_{\rm gas} = 1.09 \times 10^9\,{\rm M}_\odot$) in box sizes of $(1\,{\rm Gpc})^3$. These simulations use 2×1800^3 gas and dark matter particles and 1000^3 neutrino particles, and most adopt cosmological parameters corresponding to the maximum likelihood DES Y3 '3×2pt + All Ext.' ΛCDM cosmology (Abbott et al. 2022), which we refer to as 'D3A'. These values assume a spatially flat universe and are based on the combination of constraints from DES Y3 '3×2-point' correlation functions: cosmic shear, galaxy clustering, and galaxy–galaxy lensing, with constraints from external data from BAO, redshift-space distortions, SN Type Ia, and *Planck* observations of the CMB (including CMB lensing), Big Bang nucleosynthesis, and local measurements of the Hubble constant (see Abbott et al. 2022 for details). We also consider two alternative cosmologies: a run with the *Planck* 2018 maximum likelihood cosmology ('Planck'; Planck Collaboration VI 2020) and the 'lensing cosmology' from Amon et al. (2023) ('LS8').

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

The latter model has a lower amplitude of the power spectrum, $S_8 = 0.766$, compared with 0.815 and 0.833 for D3A and *Planck*, respectively. We assess the resolution dependence of our results by comparing with a higher resolution run, labelled L1_m8. This run adopts the fiducial D3A cosmology in a $(1\,\mathrm{Gpc})^3$ volume but uses 2×3600^3 gas and dark matter particles and 2000^3 neutrino particles. The mass resolution is therefore a factor of 8 higher for this run (i.e. $m_{\rm gas} = 1.34 \times 10^8\,\mathrm{M}_\odot$) compared to the fiducial resolution runs.

The simulations were run with the cosmological smoothed particle hydrodynamics and gravity code *SWIFT* (Schaller et al. 2024) using the SPHENIX SPH scheme (Borrow et al. 2022). The initial conditions are obtained from a modified version of monofonIC (Hahn, Rampf & Uhlemann 2021; Elbers et al. 2022), and neutrinos are implemented with the δf method (Elbers et al. 2021). The modelling of subgrid physics (star formation, stellar evolution, radiative cooling, and sources of feedback) is described in Schaye et al. (2023) and references therein.

The subgrid physics was calibrated by requiring that the simulations should match the $z = 0$ galaxy stellar mass function and the gas fractions in low-z groups and clusters using machine learning-based emulators (Kugel et al. 2023). The emulators are not only used to design simulations that reproduce these observations, but also to create models in which the galaxy stellar mass function and/or cluster gas fractions are shifted to higher/lower values. This allows us to specify model variations in terms of the number of σ by which they deviate from the calibration data. Of particular interest for this work are the variations in the group/cluster gas fractions and AGN model, which are denoted as fgas_$\pm N\sigma$ and Jet_fgas_$\pm N\sigma$. For these models, $N\sigma$ denotes by how many observed standard deviations the gas fractions have been shifted up or down with respect to the fiducial model. The Jet models make use of kinetic jets for the AGN feedback instead of the thermal model used for all other runs. These Jet models are calibrated to match the same data as the corresponding thermal AGN feedback models.

We introduce two new runs, both in 1 Gpc boxes at intermediate resolution. The first is a run denoted 'no cooling' which sets the net radiative cooling + radiative heating rate to zero for gas where the net rate would have been negative (i.e. net cooling). Consequently there is no cooling and also no star formation or feedback present in this simulation. While obviously unrealistic, comparisons to this run are helpful for quantifying the impact of feedback in the other FLAMINGO runs. The second new run, denoted 'LS8_fgas–8σ', is a strong feedback model in the LS8 lensing cosmology. This run provides an opportunity to explore the degeneracy between feedback and cosmology, via comparison to the fgas–8σ run in the fiducial D3A cosmology.

2.1.1 *Light cones:* HEALPIX *maps and galaxy/halo catalogs*

A description of the on-the-fly light-cone implementation in FLAMINGO can be found in the appendix of Schaye et al. (2023). Here, we give a brief summary, focusing on the details most relevant for the present study.

We work primarily with light cone-based maps (as opposed to the particle light-cone output). To produce the maps, the observer's past light cone is split into a set of concentric spherical shells in comoving distance. For each shell, one full sky HEALPIX (Gorski et al. 2005) map for each quantity is created. Whenever a particle is found to have crossed the light cone, we determine which shell it lies in at the time of crossing and accumulate the particle's contributions to the HEALPIX maps for that shell. The shell radii are specified in terms of redshift. From redshifts $z = 0$ to 3, we use shells of thickness $\Delta z = 0.05$, with a larger Δz at higher redshifts.

The HEALPIX map resolution is set to $N_{\rm side} = 16384$, which gives a maximum pixel radius of 13.46 arcsec and $12 * 16384^2$ pixels in each full sky map. We note that the number of pixels exceeds the size of a signed 32-bit integer (2^{31}), which would prevent us from smoothing the kSZ effect maps with the ACT beam (necessary for a like-with-like comparison) using the HEALPY smoothing function (sphtfunc.smoothing), as the function can currently only handle a maximum $N_{\rm side}$ of 8192. We therefore downsample the kSZ effect maps to this resolution using the HEALPY function pixelfunc.ud_grade, preserving the mean of the map in the downsampling operation. We describe the production of galaxy–galaxy lensing and kSZ effect maps in Sections 3.2 and 3.3, respectively.

Computational limitations prevent us from running a structure finder on-the-fly during the simulation. Instead, to produce galaxy/halo catalogues on the light cone, structure finding[1] is done in post-processing on the snapshot particle data using a modified version of the HBT-HERONS algorithm (Han et al. 2012, 2018; Forouhar Moreno et al. 2025; Chandro-Gómez et al. 2025). Note that at $z \leq 3$ snapshots are also written out with a frequency of $\Delta z = 0.05$, which was chosen to minimise any issues arising from the evolution of galaxy/halo properties between the snapshots and the HEALPIX maps. We read in the snapshot subhalo catalog corresponding to a given snapshot and then read in a spherical shell from the black hole (BH) particle light cone which spans the redshift range that is half way to the previous snapshot to halfway to the next snapshot. Every time a BH particle in the light cone shell appears as the most bound BH in a subhalo in the snapshot (identified by matching their unique particle IDs), we place the subhalo at the BH particle's position[2] in the halo light cone. We repeat this process for every snapshot to make the full light-cone halo catalogue. We use the halo light cone to provide the locations of the selected galaxies for stacking the kSZ effect and galaxy–galaxy signals in the maps. We have verified that the profiles extracted from the light cone-based HEALPIX maps using positions from the constructed halo light cone precisely match profiles constructed directly from the snapshots using the snapshot halo catalogues.

2.2 Observational data

2.2.1 *KiDS 1000 + DES Y3 lensing measurements*

The amplitude of the kSZ effect scales approximately with the halo mass of a system, so it is important to match the halo masses of observed and simulated systems to enable a like-with-like comparison of the kSZ profiles. As halo masses are not directly observable, a common approach is to select galaxies from hydrodynamical simulations based on e.g. stellar mass. However,

[1] We have checked that our results and conclusions are not strongly dependent on the choice of substructure finder, by comparing our fiducial results with those derived from employing an independent halo catalog derived using the VELOCIRAPTOR package (Elahi et al. 2019).

[2] Note that subhaloes can potentially contain multiple black holes. To address this issue, we identify which of the BHs in a subhalo exist at both the next and previous snapshots and we pick the most bound BH particle from that subset to use as a tracer. In cases where the most bound BH survives, we then use the same BH as before. If the most bound BH disappears because it merges with a more massive BH before the next snapshot, we pick the next most bound BH in the same subhalo (which is not about to get swallowed) and use that as the tracer.

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

the mapping between such observables and halo mass may not be fully realistic in the simulations, potentially resulting in a halo mismatch. Since halo mass is the key physical quantity that dictates the amplitude of the SZ signal, our approach is to use weak lensing data to ensure that the mean halo mass of the observed and simulated galaxy populations are aligned. Note that the approximate linear scaling of the kSZ signal with halo mass means that a stack of the kSZ effect of many systems will primarily be sensitive to the mean halo mass of the sample. Stacked galaxy–galaxy lensing is directly sensitive to the mean halo mass of the sample. Thus, by adjusting our simulated galaxy selection to match the stacked lensing profiles of the LOWZ and CMASS samples, we can make meaningful predictions for the kSZ effect for the BOSS samples.

We use stacked galaxy–galaxy lensing measurements from Amon et al. (2023) of the BOSS LOWZ and CMASS samples with KiDS 1000 and DES Y3 data. Here, we briefly describe these measurements, referring the reader to Amon et al. (2023) for a more complete description. The LOWZ and CMASS data of the SDSS BOSS Data Release 12 was divided into two distinct lens samples by redshift, with bounds:

$$\mathrm{L1:LOWZ}\ z = 0.15 - 0.31 \quad \& \quad \mathrm{L2:LOWZ}\ z = 0.31 - 0.43$$
$$\mathrm{C1:CMASS}\ z = 0.43 - 0.54 \quad \& \quad \mathrm{C2:CMASS}\ z = 0.54 - 0.70$$

Lensing measurements were made for each sample, using both KiDS and DES. These were shown to be statistically consistent and a combined DES Y3 + KiDS 1000 measurement was produced by taking the inverse-variance weighted average of the derived lensing profiles. Note that Schaan et al. (2021) present a stacked kSZ profile for each of LOWZ and CMASS (i.e. two redshift bins, as opposed to four in Amon et al. 2023). As we will show in Section 4, the lensing analysis yields consistent minimum stellar masses and halo masses for the two bins within LOWZ and CMASS samples, allowing us to jointly fit the two bins to yield a single galaxy selection (each) for the LOWZ and CMASS samples.

The galaxy–galaxy lensing signal is most often expressed in terms of the excess surface density, $\Delta\Sigma$, which is defined as the difference between the mean surface mass density interior to a radius and the surface mass density at that radius. It can be related to the average tangential shear $\langle\gamma_t(\theta)\rangle$ as

$$\Delta\Sigma = \frac{\langle\gamma_t(\theta)\rangle}{\overline{\Sigma_c^{-1}}} \quad , \tag{1}$$

at a projected separation $\theta = R/\chi(z_l)$, where $\chi(z_l)$ is the comoving distance to the lens. For a source redshift distribution $n(z_s)$, the average inverse critical density is given by

$$\overline{\Sigma_c^{-1}}(z_l) = \frac{4\pi G(1+z_l)\chi(z_l)}{c^2}\int_{z_l}^{\infty} \mathrm{d}z_s n(z_s)\frac{\chi(z_l,z_s)}{\chi(z_s)} \quad , \tag{2}$$

where the source redshift distribution is computed for a given lens redshift z_l and normalized such that $\int_0^\infty n(z_s)dz_s = 1$. The division by the inverse critical density in equation (1) removes the dependence of $\Delta\Sigma$ on the background source redshift distribution, $n(z_s)$, as the tangential shear also (implicitly) contains the same geometric factor.

We refer the reader to Amon et al. (2023) for the further details of the estimators of equation (1) used for the DES and KiDS data, including the treatment of additive and multiplicative biases and lens-source pair weightings and the estimation of uncertainties.

2.2.2 Planck + *ACT kSZ effect measurements*

This work uses the kSZ effect measurements presented in Schaan et al. (2021), of the ACT DR5 and *Planck* CMB temperature maps stacked at the locations of galaxies in the BOSS LOWZ and CMASS samples and using their reconstructed velocities as weights. Schaan et al. (2021) presented results at the two ACT frequencies, 90 and 150 GHz. We focus on the 150 GHz measurements here, given their higher sensitivity and angular resolution (1.3 arcmin, compared to 2.1 arcmin for the 90 GHz channel), although we have checked that none of our conclusions are sensitive to this choice.

The kSZ effect induces a fluctuation in the temperature of the CMB, $\Delta T_{\rm kSZ}$, that arises from the Doppler shift of the CMB photons with respect to the bulk motion of the ionized gas (e.g. in groups and clusters) off which CMB photons have scattered. The CMB maintains its black body spectral form but with a fractional temperature change with respect to the mean CMB temperature:

$$\frac{\Delta T_{\rm kSZ}(\boldsymbol{\theta})}{T_{\rm CMB}} = -\sigma_{\rm T}\int n_{\rm e}(\boldsymbol{\theta},z)\frac{v_{\rm e,r}(\boldsymbol{\theta},z)}{c}\frac{\mathrm{d}\chi}{1+z} \quad , \tag{3}$$

where the integration is along the observer's line of sight at given angular coordinates $\boldsymbol{\theta}$, χ is the comoving radial distance, $n_{\rm e}$ is the physical free electron density, $v_{\rm e,r}$ is the free electron peculiar radial velocity, and $\sigma_{\rm T}$ is the Thomson scattering cross-section.

The stacking analysis selectively extracts the kSZ effect associated with galaxies/groups using their reconstructed velocities, so that

$$\frac{\Delta T_{\rm kSZ}(\boldsymbol{\theta})}{T_{\rm CMB}} = -\tau_{\rm gal}(\boldsymbol{\theta})\left(\frac{v_{\rm e,\,r,\,gal}}{c}\right) \quad , \tag{4}$$

where $v_{\rm e,\,r,\,gal}$ is the galaxy's bulk velocity and $\tau(\boldsymbol{\theta})$ is the so-called optical depth to Thomson scattering, which is defined as:

$$\tau_{\rm gal}(\boldsymbol{\theta}) \equiv \sigma_{\rm T}\int n_{\rm e}(\boldsymbol{\theta},z)\frac{\mathrm{d}\chi}{1+z} \quad . \tag{5}$$

For each galaxy, Schaan et al. (2021) apply compensated aperture photometry (CAP) filtering to effectively measure a cumulative kSZ effect profile as a function of an angular disc radius, $\theta_{\rm d}$:

$$\mathcal{T}(\theta_{\rm d}) = \int d^2\theta\, \Delta T_{\rm kSZ}(\theta)\, W_{\theta_{\rm d}}(\theta) \quad , \tag{6}$$

where the CAP filter $W_{\theta_{\rm d}}$ is defined as:

$$W_{\theta_{\rm d}}(\theta) = \begin{cases} 1 & \text{for}\,\theta < \theta_{\rm d}\,, \\ -1 & \text{for}\,\theta_{\rm d} \le \theta \le \sqrt{2}\theta_{\rm d}\,, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

This corresponds to measuring the integrated temperature fluctuation in a c with radius $\theta_{\rm d}$ and subtracting the signal measured in a concentric ring of the same area around the disc. As the disc radius $\theta_{\rm d}$ is increased, the CAP filter output resembles a cumulative profile: for small disc radii, the output vanishes; for large radii, where all the gas profile is included inside the disk, the output is equal to the integrated gas profile. Note that since the filter is compensated (i.e. W integrates over area to zero), it has the desirable property that fluctuations with wavelength longer than the filter size will cancel in the subtraction. This helps to significantly reduce noise from larger scale CMB fluctuations and the correlation of the measurements between different $\theta_{\rm d}$ bins.

Schaan et al. (2021) note that the minimum-variance unbiased linear estimator is the velocity weighted, inverse-variance weighted mean and thus stack the profiles according to:

$$\hat{T}_{\rm kSZ}(\theta_{\rm d}) = -\frac{1}{r_v}\frac{v_{\rm rms}^{\rm rec}}{c}\frac{\sum_i \mathcal{T}_i(\theta_{\rm d})(v_{{\rm rec},i}/c)/\sigma_i^2}{\sum_i (v_{{\rm rec},i}/c)^2/\sigma_i^2} \tag{8}$$

where $v_{\rm rms}^{\rm rec}$ refers to the rms of the radial component of the reconstructed velocity, σ_i^2 is the noise variance for the CAP filter on galaxy i, and the r_v^{-1} factor (discussed below) ensures that the estimator is

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

not biased by the imperfections in the velocity reconstruction. The velocity weighting is key as without it the kSZ signal would cancel in the numerator, since it is linear in the galaxy radial velocities, which are equally likely to be pointing away or towards us if we subtract our motion relative to the CMB. With the velocity weighting, both numerator and denominator scale as the mean squared velocity, avoiding the cancellation and selectively extracting the kSZ signal.

As discussed in Schaan et al. (2021), peculiar radial velocities for the BOSS galaxies are reconstructed using their 3D clustering densities and solving the linearized continuity equation. The reconstruction is not perfect, however, due to non-linear effects, shot noise, and finite volume effects. Applying their techniques to BOSS mock galaxy catalogues produced using dissipationless cosmological simulations, Schaan et al. (2021) compute the kSZ bias factor, r_v, defined as:

$$r_v = \frac{\langle v_{\rm true} v_{\rm rec} \rangle}{v_{\rm rms}^{\rm true} \, v_{\rm rms}^{\rm rec}} \, , \tag{9}$$

where $v_{\rm rms}^{\rm true}$ and $v_{\rm rms}^{\rm rec}$ are the standard deviations of the true and reconstructed galaxy radial velocities, respectively. They estimate a value $r_v = 0.7$ and use this to correct their stacked kSZ profiles to account for the imperfect velocity reconstruction (see also Ried Guachalla et al. 2024; Hadzhiyska et al. 2024b). Note that r_v is a constant correction factor that scales the amplitude of the derived kSZ temperature profiles. As we will show later, feedback can strongly affect the amplitude of the profile and will thus be degenerate at some level with uncertainties in the velocity reconstruction. It would be interesting to apply the velocity reconstruction technique on BOSS mocks derived from the FLAMINGO hydrodynamical simulations as an independent estimate of r_v, but we leave this exercise for future work, retaining the fiducial estimate from Schaan et al. (2021) for the bias factor. Note that for kSZ temperature profiles derived from the simulations, the value of r_v is 1 (i.e. unbiased), since we use the true radial velocities rather than reconstructed velocities.

3 SYNTHETIC OBSERVABLES FROM FLAMINGO

3.1 Galaxy selection

The BOSS LOWZ sample primarily selects red galaxies while the BOSS CMASS sample targets galaxies at higher redshifts with a surface density of roughly 120 deg^{-2} and a roughly constant minimum stellar mass of a few 10^{11} M$_\odot$. In this study, we do not attempt to implement the precise BOSS selection criteria for the selection of galaxies in the simulations. Such an exercise would be non-trivial, as the mapping from intrinsic stellar properties in the simulations (mass, age, abundances) to observed luminosities requires stellar population synthesis modelling, a treatment of dust and radiative transfer, and would be sensitive to the adopted theoretical nucleosynthetic yields, which have considerable uncertainties. Instead, our approach is to ensure that the simulated sample has a stacked galaxy–galaxy lensing signal that is compatible with the BOSS samples. This ensures that the mean halo masses of the simulated and observed samples are aligned, enabling a fair comparison of the observed and predicted stacked kSZ effect profiles. We discuss below some tests that we have performed to ensure that our selection is realistic.

Our fiducial approach to select the simulated galaxies is to apply a simple minimum stellar mass cut, where we define the stellar mass as the total bound stellar mass within 50 kpc of the most bound particle. For comparison to the four redshift bins in the Amon et al. (2023) lensing analysis, we select simulated 'lens' galaxies from the light-cone shell that is nearest to the mean redshift of the actual (observed) bins. Specifically, we select the shells with $0.225 < z < 0.275$, $0.325 < z < 0.375$, $0.475 < z < 0.525$, and $0.575 < z < 0.625$ for the L1, L2, C1, and C2 bins, respectively, noting that the mean redshifts of the observed bins are: 0.240, 0.365, 0.496, and 0.592. For the kSZ effect comparison, we select galaxies in the shells with $0.275 < z < 0.325$ and $0.525 < z < 0.575$ for LOWZ and CMASS, respectively, noting that the mean redshifts for the two observed samples are $z = 0.31$ and $z = 0.54$, respectively.

As the genuine BOSS LOWZ and CMASS samples do not explicitly exclude satellite galaxies, we should include them in our selection so long as their stellar mass exceeds the minimum stellar mass. We compare the results with a central-only sample to deduce the role that satellites play in derived lensing and kSZ effect profile in Section 4.2. We construct multiple simulated samples by varying the minimum stellar mass cut and compare the derived galaxy–galaxy lensing profiles to the measurements of Amon et al. (2023) to determine the minimum stellar mass cut which best reproduces the lensing measurements. We propagate the uncertainties in the best-fitting minimum stellar mass through to our kSZ effect analysis.

A simple stellar mass cut is unlikely to yield a galaxy sample that matches all aspects of the CMASS and LOWZ samples. However, our selection is constrained to match the mean halo masses of the CMASS and LOWZ samples and, as already noted, the amplitude of the stacked kSZ effect should be mostly sensitive to the mean halo mass of the stack. Nevertheless, as a check, we have explored simultaneous cuts in stellar mass and specific star formation rate, so that we select preferentially 'red' galaxies. In Appendix A, we show that our main results and conclusions are unaltered by adopting this more complex selection function, suggesting that the matching the mean halo mass is sufficient for our purposes. Another test we have performed is to use much narrower bins in the stellar mass selection of 0.1 dex width whose bin centre is adjusted to match the same mean halo mass as our fiducial selection. We find virtually identical kSZ effect predictions for these two cases. In addition, we have compared the predicted and observed projected clustering of the BOSS galaxies for our fiducial selection methodology. For this comparison we used the large-scale clustering measurements presented in Amon et al. (2023) and we adopted the same methodology to derive the clustering of the simulated galaxies selected on stellar mass. In short, we find excellent agreement between the minimum stellar masses and mean halo masses derived from our galaxy–galaxy lensing analysis and our clustering analysis, again suggesting that the simulated galaxy sample is realistic. We leave a detailed presentation of the clustering results for a future paper.

When selecting galaxies, we use the true stellar mass predicted by the simulations. Observationally measured stellar masses have uncertainties. An interesting question is whether our results would be impacted by folding in such uncertainties in the selection of our simulated galaxies. For the same reasons argued above, we do not expect such uncertainties to impact our results since the selection of galaxies (with or without observational uncertainties factored in) is forced to match the galaxy–galaxy lensing measurements. Nevertheless, we explicitly test this hypothesis below.

Note that an alternative possibility for selection would be to select systems based on their halo masses rather than their stellar masses, and to constrain the selected halo mass range based on fits to the galaxy–galaxy lensing measurements as described above. But such a selection would essentially correspond to a central-only population, since virtually all satellites would have halo masses well below the scale of interest here ($\sim 10^{13}$ M$_\odot$) due to tidal stripping by their

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

more massive host and would therefore either not be included in the selection or make no meaningful contribution to the stack. Selection by stellar mass, on the other hand, is closer to the real observational selection function in BOSS and allows us to naturally include both centrals and satellites in the selected simulation population.

3.2 Galaxy–galaxy lensing profiles

A nice property of galaxy–galaxy lensing $\Delta\Sigma$ profiles is that they do not depend on the source redshift distribution of the background galaxies used to measure the profiles. This is in contrast to measurements of cosmic shear. It is thus relatively straightforward to compute predictions from the simulations, using the mass distribution around haloes. In particular, we start from the HEALPIX total mass maps, which provide the sum of all of the mass components (gas, DM, stars, BHs, and neutrinos) in pixels in a given redshift shell. We convert the total mass maps to maps of comoving surface mass density by converting the pixel area (in steradians) into a comoving surface area using the comoving radial distance to the shell centre from the observer and then dividing the total mass maps by this comoving surface area. For a given galaxy included in our selection, we select all the pixels within a certain angular distance which, following Amon et al. (2023), we convert to comoving transverse distances assuming a flat cosmology with $\Omega_{\rm m} = 0.3$ and $h = 0.7$. The $\Delta\Sigma$ profile of a given galaxy is calculated by ordering the pixels by projected comoving distance and then subtracting the surface mass density of a given pixel from the mean surface mass density from all of the pixels interior to it[3].

We compute stacked $\Delta\Sigma$ profiles by defining a set of projected comoving radial bins (logarithmically spaced) and computing the mean $\Delta\Sigma$ in those bins by simply summing the profiles of each galaxy in those bins and dividing by the number of galaxies in the stack. The centre of the bin is not the midpoint between the upper and lower bin bounds but is computed as the $\Delta\Sigma$-weighted radius of all pixels that fall within the radial bin. We have found that such a weighting scheme is more robust to changes in the radial binning strategy.

The lensing profiles of Amon et al. (2023) span a wide radial range ($R \sim 0.1$–$100\,{\rm Mpc}\,h^{-1}$), and extracting the pixels for large numbers of haloes from our fiducial high-resolution HEALPIX ($N_{\rm side} = 16384$) maps is computationally expensive. We therefore adopt a hybrid approach where we use the high-resolution map to extract the signal on small comoving scales ($R < 2\ {\rm Mpc}\,h^{-1}$) and a downsampled version with with $N_{\rm side} = 2048$ to retrieve the signal on large scales. We have verified that there is excellent convergence between the different resolution maps on intermediate scales. Furthermore, we have parallelized the analysis, allowing us to produce lensing profiles for large numbers of haloes at the same time.

3.3 kSZ effect profiles

We construct $\Delta T_{\rm kSZ}$ effect maps and profiles as follows. As described in Schaye et al. (2023), when a gas particle crosses the light cone, we compute its dimensionless Doppler B, b, parameter:

$$b = \frac{n_{\rm e} m_{\rm g} \sigma_{\rm T} v_{\rm r}}{\Omega_{\rm pixel} d_{\rm A}^2 \rho c}, \tag{10}$$

where $m_{\rm g}$, ρ, and $v_{\rm r}$ are the mass, mass density, and radial velocity of the particle, respectively, $\Omega_{\rm pixel}$ is the solid angle of a HEALPIX pixel and $d_{\rm A}$ is the angular diameter distance to the observer. Particles crossing the light cone in a given redshift shell are accumulated to the corresponding HEALPIX map.

Visual inspection of equation (10) reveals a close relation to equation (3). Indeed, the quantity $m_g/[\Omega_{\rm pixel}\ d_{\rm A}^2\ \rho]$ in equation (10) is a discretized (per particle) estimate of the physical path length $\chi/(1+z)$, noting that m_g/ρ is the volume associated with the particle and $\Omega_{\rm pixel} d_{\rm A}^2$ is the physical area of the pixel in which the particle is deposited at the distance of the particle. Aside from this discretization difference, the mapping between the Doppler B is simple: $\Delta T_{\rm kSZ}/T_{\rm CMB} = -b$, which is independent of observing frequency.

Thus, the production of a $\Delta T_{\rm kSZ}$ map from the FLAMINGO Doppler B maps is trivial, requiring only the multiplication of a factor of $-T_{\rm CMB}$ and the summation of the individual maps (shells) along the line of sight. In practice, to achieve convergent results for the stacked kSZ effect profiles, we find that stacking along the full line of sight ($z = 0$ to 3) of the light cones is unnecessary, since most of the line of sight will be uncorrelated with the selected galaxy in the stack. Indeed, we find that using only three shells (each having width $\Delta z = 0.05$, i.e. a padding shell on each side of the shell that contains the galaxy) is sufficient to achieve convergent results. The same holds true for the $\Delta\Sigma$ profiles.

As noted previously, we downsample our full resolution $N_{\rm side} = 16384$ kSZ effect maps to $N_{\rm side} = 8192$ so that we can then smooth the maps with ACT beam. Schaan et al. (2021) used coadded ACT DR5 (day + night) maps in their study. We therefore retrieve the corresponding measured ACT beam for this data set from the NASA Lambda website[4] and convolve it with the simulated maps in multipole space using the HEALPY function sphtfunc.almxfl.

Apart from convolution with a realistic beam, our maps are idealized, in that they do not contain realistic noise. We further assume that the kSZ effect signal can be perfectly recovered in the observational measurements, which would appear to be a strong assumption in the face of significant foreground and background contaminant signals, including the primary CMB, the tSZ effect, radio sources, the CIB, and so forth. However, in general, these sources of contamination are not expected to correlate with the velocity field of the selected galaxies and are therefore suppressed in the velocity stacking process (see discussion in Schaan et al. 2021). Nevertheless, some contamination may still arise from the sources that are truly correlated with the selected haloes (i.e. part of the 2-halo term) and it will be important to evaluate their potential effects by constructing realistic mocks from the hydro simulations that contain all relevant signals. We leave this for future work (T. Yang, in preparation).

We extract and stack the kSZ effect profiles using the same methodology as applied to the genuine *Planck* + ACT data (i.e. equations (6)–(8)) using the true radial velocities of the galaxies in the light cone (i.e. we set $r_v = 1$). Also, as our simulated kSZ effect maps are noiseless, the inverse-variance weighting applied for the observations is not applicable for the simulations. We therefore set $\sigma_i = 1$ in equation (8) when stacking the simulated kSZ effect profiles.

[3] A convenient feature of the HEALPIX maps is that the pixels are of equal area, which implies that the mean surface mass density interior to a given pixel can be simply estimated as the mean value of all interior pixels. For non-equal area pixels, a more cumbersome route of computing a Σ profile, integrating it to the angular radius in question, and then dividing it by the enclosed surface area, is required.

[4] https://lambda.gsfc.nasa.gov/product/act/actpol_dr5_coadd_maps_info.html

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

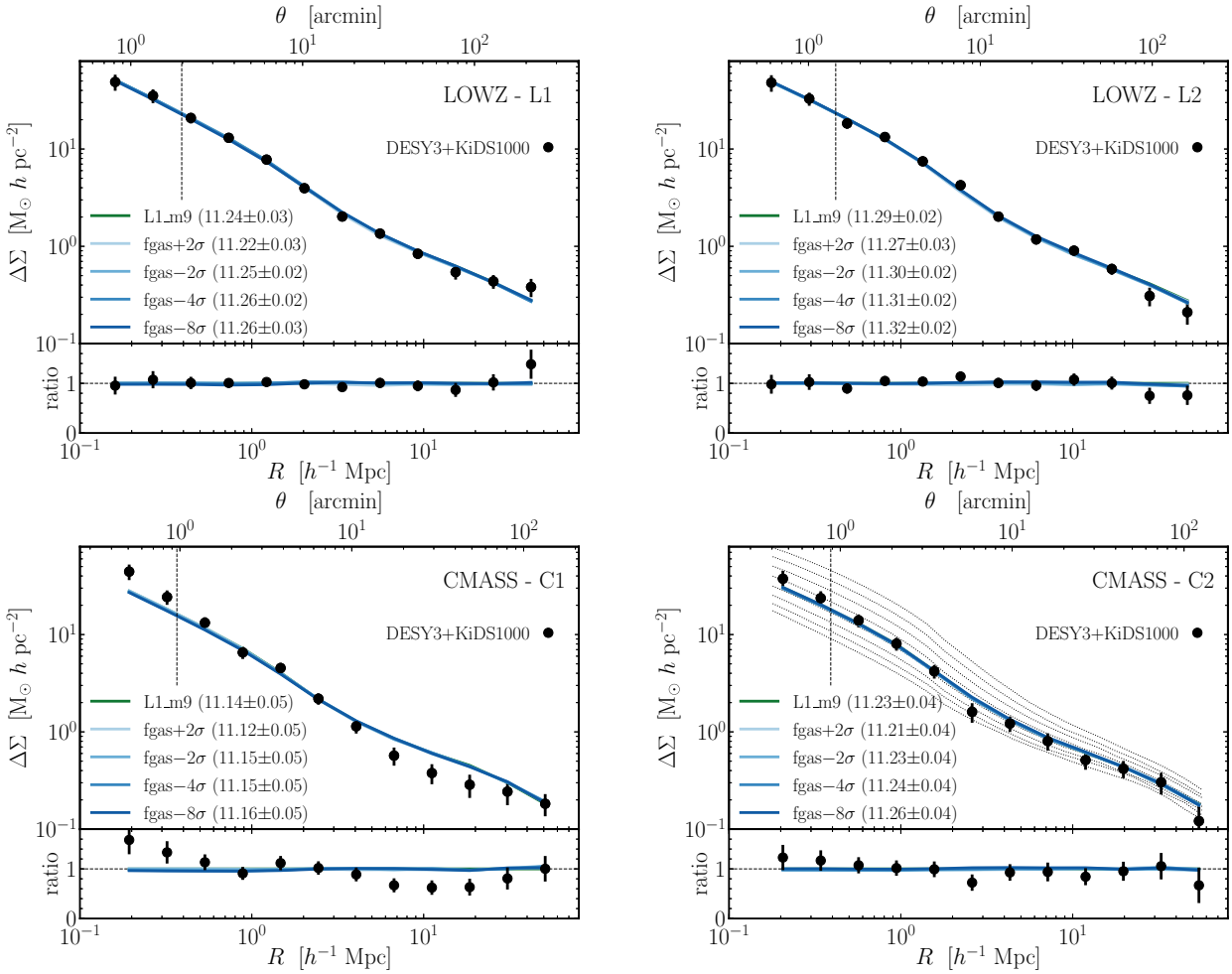

Figure 1. Comparison of the Amon et al. (2023) DES Y3 + KiDS 1000 galaxy–galaxy lensing $\Delta\Sigma$ profiles of BOSS LOWZ and CMASS galaxies (data points with 1σ error bars) with predictions from the fiducial FLAMINGO run (L1_m9) and the fgas feedback variants. There are four main panels corresponding to the four BOSS samples (two redshift bins each for LOWZ and CMASS). The bottom x-axis shows the the comoving projected radius while the top x-axis shows the angular scale (in arcmins) at the mean redshift of each of the redshift bins. The vertical dashed lines indicate the location of R_{500c} derived from the mean halo mass of each redshift bin (see Section 4.3). The solid curves show the best-fitting profiles for each of the FLAMINGO runs, which fall nearly on top of each other. The legend provides the best-fitting ($\log_{10}$) minimum stellar mass (in solar units) of the simulated galaxy selection for each of the runs, with the error bars reflecting the 2σ uncertainties (95 per cent confidence) on the best-fitting minimum mass given the uncertainties on the lensing measurements. The bottom sub-panels in each of the main panels show the ratio of the fgas feedback variants and the observational data with respect to the fiducial feedback model. For reference, the dotted curves in the bottom right panel show the predicted $\Delta\Sigma$ profiles for the fiducial FLAMINGO run with minimum stellar mass $\log_{10}[M_{\rm star}/{\rm M}_\odot]$ values ranging between 10.9 and 11.6 in steps of 0.1. Overall, the simulations reproduce the lensing measurements very well over a wide range of radii, with a simple stellar mass cut (per feedback variant) being sufficient to match the LOWZ ($\log_{10}[M_{\rm star}/{\rm M}_\odot] \approx 11.3$) and CMASS ($\log_{10}[M_{\rm star}/{\rm M}_\odot] \approx 11.2$) measurements.

4 RESULTS

In this section, we explore the feedback and cosmology dependencies of the kSZ effect predictions (Section 4.1), as well as the role that satellite galaxies play (Section 4.2). As already described, we optimize the selection of galaxies from the simulations by fitting to the galaxy–galaxy lensing profiles of BOSS galaxies, thus ensuring that our selection has the correct underlying mean halo mass. In Section 4.3, we discuss the implied halo masses and radii of the BOSS samples.

4.1 Dependence on feedback and cosmology

Fig. 1 shows the best-fitting galaxy–galaxy lensing profiles for the comoving excess surface mass density, $\Delta\Sigma$, for the fiducial FLAMINGO run (L1_m9) and the fgas feedback variants. There are four panels corresponding to the four BOSS redshift bins from Amon et al. (2023) (two bins each for LOWZ and CMASS). We show measurement as a function of the comoving projected radius, R, on the bottom x-axis, but we also show the angular scale at the median redshift of the sample on the top x-axis, in order to facilitate comparisons with the kSZ effect measurements.

The solid curves show the best-fitting profiles for each of the FLAMINGO runs, which fall nearly on top of each other. The legend provides the best-fitting ($\log_{10}$) minimum stellar mass (in solar masses) of the simulated galaxy selection for each of the runs. For example, in the LOWZ-L1 bin, the FLAMINGO fiducial feedback run with a minimum stellar mass of $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.24 \pm 0.03$ provides the best fit to the DES Y3 + KiDS 1000 $\Delta\Sigma$ measurements. The error bars reflect the 2σ uncertainties (95 per cent confidence) on the best-fitting minimum mass given the uncertainties on the lensing measurements. The best-fitting minimum stellar mass and

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

its uncertainties are estimated by first calculating the $\Delta\Sigma$ profiles at different values of the minimum stellar mass and then interpolating to obtain the best-fitting result. The bottom right panel of Fig. 1 shows the process, where the dotted curves correspond to the predicted $\Delta\Sigma$ profiles for the fiducial FLAMINGO run with $\log_{10}[M_{\rm star}/{\rm M}_\odot]$ values ranging between 10.9 and 11.6 in steps of 0.1 dex. We then interpolate the $\Delta\Sigma$ values to a much finer grid of $\log_{10} M_{\rm star}$ Specifically, we linearly interpolate $\log_{10}\Delta\Sigma$ in stellar mass bins of width 0.01 dex at projected radius, R. The best-fitting minimum stellar mass is determined through χ^2 minimisation with respect to the DES Y3 + KiDS1000 measurements and their uncertainties from Amon et al. (2023). Note that Amon et al. (2023) quote diagonal uncertainties for their combined DES Y3 + KiDS 1000 measurements only; i.e. the radial bins are assumed to be uncorrelated. Calculation of the off-diagonal elements of covariance matrix for the combined data set is non-trivial and beyond the scope of this work.

It is interesting that a simple stellar mass cut applied to the various FLAMINGO runs is capable of yielding excellent fits to the lensing measurements over approximately 2.5 decades in radius, spanning both the 1-halo and 2-halo regimes (the transition between the two regimes is clearly visible at $R \approx 3$–$4\,{\rm Mpc}\,h^{-1}$). Varying the stellar mass has the impact of varying the mean halo mass of the simulated galaxy sample which affects the amplitude of the predictions. Thus, the fact that the simulations reproduce the amplitude of the observed galaxy–galaxy lensing signal is not surprising. But the *shape* of the profile is a genuine prediction of the simulations and ΛCDM generally, and the fact that the profiles accurately match these precise measurements over a very large range of radii is remarkable. Note that the detailed shape is expected to be cosmology dependent in ΛCDM, since the 1-halo and 2-halo terms themselves have different cosmology dependencies. The main cosmological dependence of the 1-halo regime is through the halo concentration (e.g. Bullock et al. 2001; Eke, Navarro & Steinmetz 2001; Correa et al. 2015; Diemer & Kravtsov 2015; Brown et al. 2022), whereas at large radii (2-halo) it is via the halo bias (e.g. Sheth & Tormen 1999; Tinker et al. 2010).

In a given redshift bin (e.g. LOWZ-L1), all of the runs yield similarly good fits to the data and the different runs prefer only slightly different stellar masses. This is not unexpected, since the fiducial FLAMINGO run and the fgas feedback variants in Fig. 1 have each been independently calibrated to reproduce the local galaxy stellar mass function. Thus, the mapping between stellar mass and halo mass is expected to be nearly the same in each of the runs. The fact that the stronger feedback variants prefer a slightly higher stellar mass cut is likely because the halo masses themselves have been reduced slightly more through baryon ejection in the stronger feedback variants. Thus, to get back to the same mean halo mass required to match the lensing data, a slightly higher stellar mass cut is required in the models with stronger feedback. If the measurements could be extended to smaller projected radii ($R < 0.1\,{\rm Mpc}\,h^{-1}$), it is possible that the lensing measurements themselves could be used to place constraints on the feedback models, through deviations in the profile shapes on small scales (e.g. fig. 6 of Velliscig et al. 2014). In this study, the lensing measurements are used to constrain the galaxy selection, so that more sensitive kSZ effect measurements can be used to discriminate between the different feedback models.

Comparing the fits across the two LOWZ bins (top two panels), for the majority of the cases the preferred minimum stellar masses are consistent within a few sigma for a given run. For example, the fiducial FLAMINGO model prefers a minimum stellar mass of $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.24 \pm 0.03$ for the LOWZ-L1 bin and $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.29 \pm 0.02$ for the LOWZ-L2 bin. As the two LOWZ bins are consistent with a single LOWZ selection for a given model, we jointly fit the LOWZ L1 and L2 bins to determine the minimum stellar mass cut for the kSZ effect predictions. The same is true for the higher redshift CMASS bins (C1 and C2, with $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.14 \pm 0.05$ and $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.23 \pm 0.04$, respectively), although we note that for the CMASS-C1 bin the shape of the best-fitting simulated profiles do not match the measurements perfectly. In particular, we note that the 1-halo regime dominates the fit (in terms of signal to noise), and the simulations provide a good match to the data there, but they predict a signal that is too large compared to the measurements at larger radii, in the 2-halo regime. The CMASS-C1 bin was also identified as an outlier in Amon et al. (2023), who found that even their flexible HOD framework (with 5 free parameters) was unable to reproduce the lensing measurements in detail. Nevertheless, the best-fitting stellar masses are consistent within a few sigma between the C1 and C2 bins for all the runs and we therefore jointly fit them to derive a single stellar mass cut (for each feedback variant) for the kSZ effect CMASS analysis. Furthermore, we highlight that the kSZ effect measurements mostly probe relatively small radii, in the 1-halo regime, where the simulation predictions match the CMASS lensing measurements well even for the C1 bin.

An alternative to approach to handling the two redshift bins in the LOWZ and CMASS samples of Amon et al. (2023) would be to predict the kSZ effect profiles for each bin given their respective best-fitting minimum stellar masses and then to average the profiles (e.g. by inverse-variance weighting). In practice, we find that this gives nearly identical results to our default method of combining the bins. Indeed, even if we uniformly applied the lower or higher of the two stellar mass estimates to the whole sample, our general conclusions with regards to the strength of feedback required to match the kSZ effect measurements would be unchanged.

While we have elected to constrain our simulated galaxy selection based on the galaxy–galaxy lensing signal rather than a more observable quantity such as stellar mass, it is interesting to note that our best-fitting minimum stellar masses, which range from $\log_{10}[M_{\rm star}/{\rm M}_\odot] \approx 11.15$ to 11.3, align remarkably well with the actual observed peak and mean stellar masses of the BOSS LOWZ and CMASS samples. In particular, Maraston et al. (2013) find mean $\log_{10}[M_{\rm star}/{\rm M}_\odot]$ values of 11.33 at $0.2 \lesssim z \lesssim 0.4$ (LOWZ), 11.27 at $0.4 \lesssim z \lesssim 0.5$ (CMASS), and 11.26 at $0.5 \lesssim z \lesssim 0.6$ (CMASS). Thus, the calibrated FLAMINGO runs have realistic stellar mass to halo mass ratios at the typical stellar mass scale probed by the BOSS survey.

Armed with strong constraints on the galaxy selection (minimum stellar mass) from the galaxy–galaxy lensing comparisons, we compare the kSZ effect profiles predicted by the fiducial FLAMINGO run and the fgas feedback variants with the 150 GHz stacking measurements of Schaan et al. (2021) in Fig. 2. The solid curves represent the predictions for the best-fitting minimum stellar masses. The legend provides the number of standard deviations that the model deviates from the observational measurements with the error bars corresponding to the propagated uncertainties on the stellar mass cut given the uncertainties on the galaxy–galaxy lensing measurements. Note that the number of standard deviations is computed using the full covariance matrices of the data, which is important since the outermost bins are highly correlated. Taking into account this correlation, the strongest constraint on the goodness of fit comes from the innermost three or four radial bins.

We see visually from Fig. 2 and from the computed number of standard deviations that the fiducial calibrated FLAMINGO run is statistically ruled out by the kSZ effect measurements, at ≈5 sigma, for both the LOWZ and CMASS samples, which are independent.

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

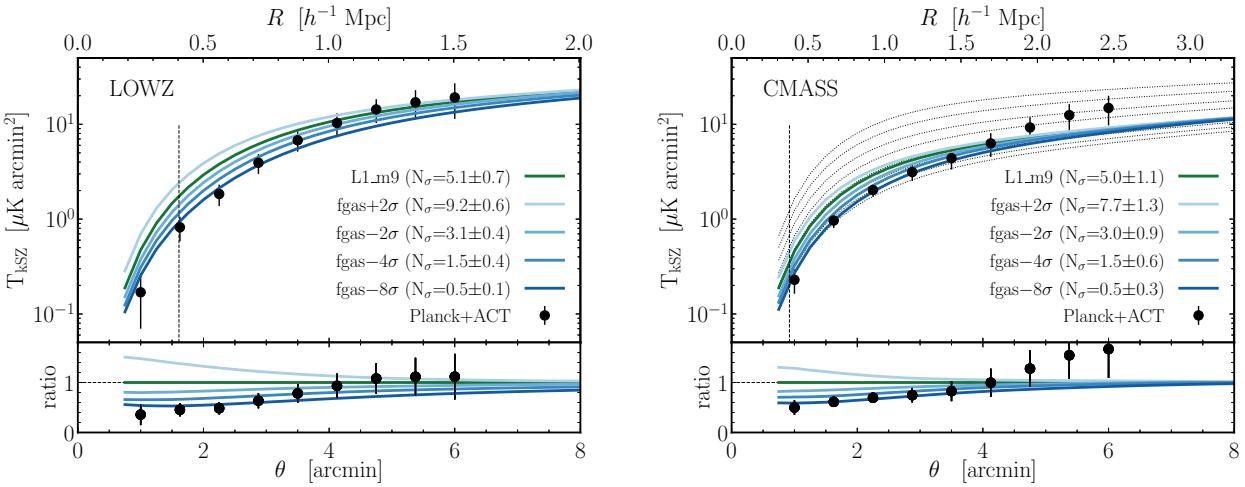

Figure 2. Comparison of the 150 GHz *Planck* + ACT kSZ effect temperature profiles of BOSS LOWZ and CMASS galaxies from Schaan et al. (2021) (data points with 1σ error bars) with predictions from the fiducial FLAMINGO run (L1_m9) and the fgas feedback variants. The bottom x-axis shows the angular scale (in arcmins) while the top x-axis shows the comoving projected radius at $z = 0.31$ and $z = 0.54$ for the LOWZ and CMASS panels, respectively. The vertical dashed lines indicate the location of R_{500c} derived from the mean halo mass of each redshift bin (see Section 4.3). The solid curves represent the predictions for the best-fitting minimum stellar mass (fitted to the lensing, see Fig. 1). The legend provides the number of standard deviations that the model deviates from the observational measurements with the error bars corresponding to the propagated uncertainties on the stellar mass cut given the uncertainties on the galaxy–galaxy lensing measurements. Note that the number of standard deviations is computed using the full covariance matrices of the data, which is important since the outermost bins are highly correlated. The error bars on the data points reflect only the diagonal elements of the covariance matrices. The bottom sub-panels in each of the main panels show the ratio of the fgas feedback variants and the observational data with respect to the fiducial model. For reference, the dotted curves in the right panel show the predicted $\Delta T_{\rm kSZ}$ profiles for the fiducial FLAMINGO run with minimum stellar mass $\log_{10}[M_{\rm star}/{\rm M}_\odot]$ values ranging between 10.9 and 11.6 in steps of 0.1. The fiducial calibrated FLAMINGO run is statistically ruled out by the kSZ effect measurements at about the 5σ level, for both the LOWZ and CMASS samples which are independent. Only the two strongest feedback models in this comparison (fgas-4σ, fgas-8σ) are formally consistent with the measurements.

Only the two strongest feedback models in this comparison, fgas-4σ and fgas-8σ, are formally consistent with the measurements, with $\approx 1.5\sigma$ and $\approx 0.5\sigma$ deviations from the measurements, respectively. Note that in computing the number of standard deviations, we have not marginalized over uncertainties in cosmology (which we expect to be small, as discussed below) or uncertainties in the velocity reconstruction used in the observational measurements. Thus, the quoted level of tension may be slightly overestimated.

Note that because the kSZ effect is proportional to the gas mass, it will be affected by any physical process that alters the gas mass fractions of groups and clusters. Ejection of gas due to AGN feedback is believed to be the main mechanism for altering the gas fractions, but gas is also removed via radiative cooling leading to neutral gas and star formation. An important aspect of the FLAMINGO simulations shown in Figs 1 and 2 is that they have all been calibrated to reproduce the observed galaxy stellar mass function. Thus, the differences between the models in Fig. 2 are due entirely to differences in the level of gas ejection, and the relatively low observed kSZ signal indicates that a relatively high level of ejection is required. A caveat, of course, is if there are significant biases in the observed stellar masses (e.g. due to uncertainties in stellar population synthesis modelling or the extrapolation of surface brightness profiles) this would in turn bias the estimates of the required level of feedback.

Our results are qualitatively consistent with Bigwood et al. (2024), who used the baryonification formalism to model the kSZ effect jointly with cosmic shear, in the sense that the kSZ effect data appears to prefer stronger feedback relative to that inferred by modelling X-ray-based baryon fractions of galaxy groups. However, given the use of self-consistent full cosmological hydrodynamical simulations, galaxy–galaxy lensing to strongly pin down the mass scale of the BOSS samples, inclusion of satellites and mis-centring effects, and the use of both the LOWZ and CMASS samples, our quantitative results are more robust.

The dotted curves in the right panel of Fig. 2 illustrate the importance of the galaxy–galaxy lensing measurements in pinning down the mass scale of the BOSS galaxies and therefore the required level of feedback. These correspond to the same variations in the minimum stellar mass as shown in the bottom right panel of Fig. 1 for the fiducial feedback model. With the kSZ effect data alone to go by, we would not be able to easily distinguish the fiducial feedback scenario with a lower minimum stellar mass, of $\log_{10}[M_{\rm star}/{\rm M}_\odot] \approx 10.9$, from a stronger feedback scenario with higher stellar mass (noting that it is the innermost 3 or 4 bins which dominate the fit). The lensing data are therefore crucial to break the degeneracy between halo mass and feedback-driven gas ejection.

We note that the measurements of Schaan et al. (2021) extend out to $\approx 2\,{\rm Mpc}\,h^{-1}$ which is the scale where the 2-halo term becomes visible in the galaxy–galaxy lensing measurements in Fig. 1. The 2-halo term is easily visible in the lensing measurements given their precision, the large dynamic range of the measurements, and that the signal is a differential one. In the case of the kSZ profiles, the measurement is effectively a cumulative signal which will partially mask the transition to the 2-halo regime. Nevertheless, the correlated clustering of nearby haloes is expected to modestly contribute to the outermost bins in these measurements (see also Amodeo et al. 2021) and the profiles should therefore not be regarded as being due solely to the selected galaxies in the stack.

As discussed in Section 3.1, observational stellar masses have measurement uncertainties and it is interesting to ask what the effect of such uncertainties might be if we applied them to the simulated galaxies and repeated the above analysis. Behroozi et al. (2019) find that a lognormal scatter with a standard deviation

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

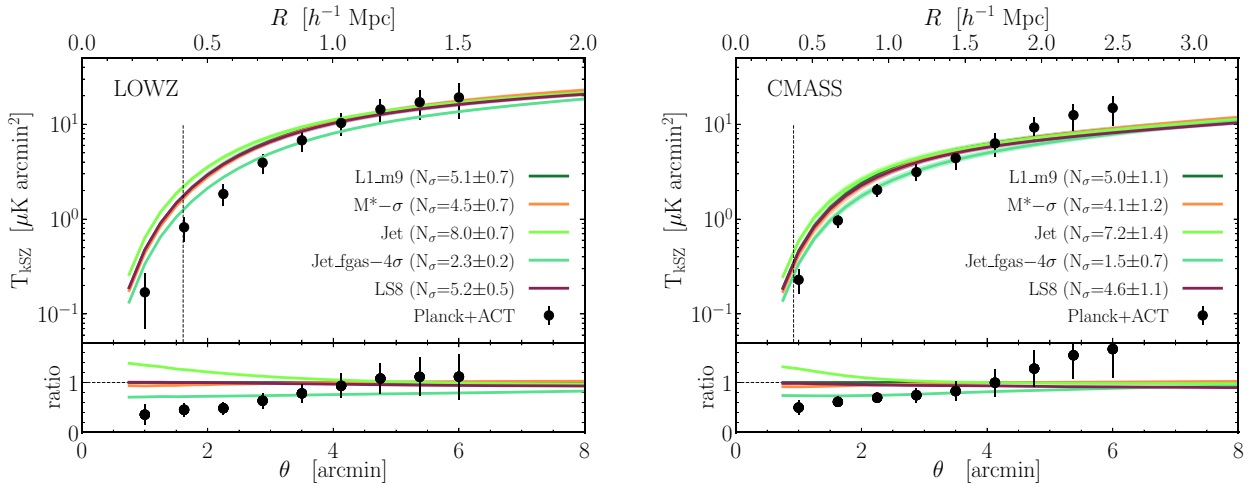

Figure 3. As Fig. 2, but showing the dependence of the predicted kSZ effect profiles on other model variations, namely variations in the galaxy stellar mass function and the fiducial and strong jet models of AGN feedback, as well as the lensing (LS8) cosmology run with fiducial feedback. The solid curves correspond to the different FLAMINGO simulations, with the number of standard deviations from the data provided in the legend. The best performing model shown is the Jet model with reduced gas fractions (Jet_fgas−4σ), which is consistent with the findings from Fig. 2 that enhanced gas ejection is required by the measurements.

$\sigma(\log_{10} M_{\rm star}) = \min(0.070 + 0.071z, 0.3)$ dex describes typical random measurement errors in the observed stellar masses. As a test, we have applied this scatter to the true simulated stellar masses to mimic an observed stellar mass (note that we also applied this scatter during the calibration of FLAMINGO; see Schaye et al. 2023). We then analysed the simulations as described above, by determining the best-fit minimum stellar mass cut required to match the galaxy–galaxy lensing measurements and then predicting the kSZ profiles for this selection.

We find that the best-fitting minimum stellar mass required to match the galaxy–galaxy lensing measurements typically increases by 0.04–0.05 dex with respect to our fiducial analysis with no measurement scatter. This reflects the fact there are more lower mass objects than high mass objects, thus leading to a slight net up-scattering. To reproduce the galaxy–galaxy lensing signal in the presence of lower stellar mass galaxies entering the selection, a slightly higher minimum stellar mass cut is therefore required to match the stacked lensing profiles, compared to the case with no scatter. Nevertheless, the best-fitting mean halo mass is virtually identical to the no-scatter case, as is the predicted stacked kSZ profile. We therefore conclude that measurement uncertainties in the observed stellar masses do not significantly impact our results or conclusions, by virtue of the fact that the selection is constrained to match the lensing signal.

We have also tested the sensitivity of our results and conclusions to numerical resolution, by analysing the high-resolution calibrated FLAMINGO model (L1_m8) and comparing the results with the fiducial resolution, calibrated model (L1_m9). This comparison is presented in Appendix B. In short, we find that the high-resolution run displays a similar level of tension with the observed kSZ effect measurements of Schaan et al. (2021) compared to the fiducial resolution, suggesting that our conclusions are robust to changes in resolution.

In Fig. 3, we explore kSZ effect predictions using several other feedback variations in the FLAMINGO suite at the fiducial resolution, namely variations in the galaxy stellar mass function at the fiducial gas fraction (M*−σ) and the fiducial and strong jet models of AGN feedback (Jet and Jet_fgas−4σ). The corresponding galaxy–galaxy lensing profiles are shown in Appendix C (see Fig. C1). We find that none of the models provide an acceptable fit to the data. The best performing model is the jet model with reduced gas fractions (Jet_fgas−4σ), which is consistent with the findings in Fig. 2 that enhanced gas ejection is required by the measurements. The lensing LS8 cosmology run, which gives a similarly poor fit to the measurements as the fiducial L1_m9 run. This is likely because both models adopt the same feedback model and that, intrinsically, the kSZ effect profile is not expected to depend significantly on cosmology. The underlying matter profile depends weakly on cosmology through the dependence of the concentration on cosmology, but this dependence will likely be even weaker when dealing with the hot gas due to the effects of non-gravitational processes such as cooling and feedback. The kSZ effect would be expected to scale with the universal baryon fraction, $f_{\rm b} \equiv \Omega_{\rm b}/\Omega_{\rm m}$, as this quantity dictates the fraction of baryons that haloes can accrete (at least in the absence of feedback). However, $f_{\rm b}$ is precisely determined from several cosmological probes and all of our runs have very similar values of $f_{\rm b}$. Lastly, we have also analysed the FLAMINGO runs that vary the summed neutrino mass (which also adopt the fiducial feedback model) and find that they, too, yield similarly poor fits to the kSZ effect measurements, but we do not show them here for brevity.

4.2 The role of satellite galaxies

Some fraction of the galaxies comprising the BOSS LOWZ and CMASS samples will be satellites. Previous attempts to model the kSZ effect measurements have not accounted for the impact of satellites, but whether this is a significant omission is unclear. On the one hand, satellites will be mis-centred with respect to their host haloes and this might be expected to lead to a reduced kSZ signal compared to a galaxy that is centred on the hot gas distribution. On the other hand, a stellar mass-based selection implies that the satellites will typically be in hosts that are more massive than a host which has a central of similar stellar mass. This will tend to boost the kSZ signal.

Here, we use the FLAMINGO simulations and associated subhalo catalogues to explore the impact of the inclusion of satellite galaxies on the derived stacked lensing and kSZ effect profiles. In Fig. 4, we compare the ratios of the stacked lensing and kSZ signals for a central-only sample with that for the fiducial sample for different

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

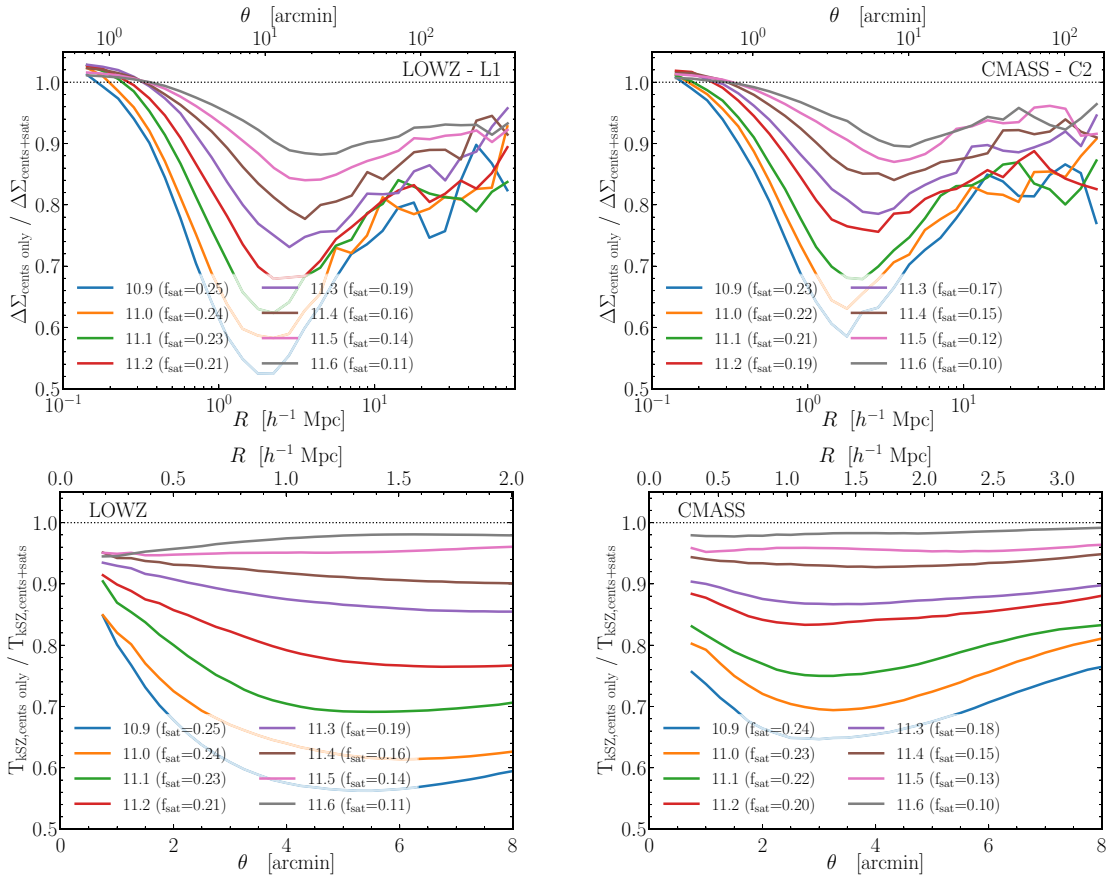

Figure 4. Ratios of the stacked lensing and kSZ signals for a central-only sample with respect to the fiducial sample for different choices of the minimum stellar mass cut. For the lensing comparison (top panels) we show only the L1 and C2 bins, for brevity, noting that the L2 and C1 bins give similar results. The solid curves show the ratios for the fiducial FLAMINGO feedback model and correspond to different minimum stellar masses $\log_{10}[M_{\rm star}/{\rm M}_\odot]$ values ranging between 10.9 and 11.6 in steps of 0.1. The effects of including satellites tend to be stronger for selections including lower mass galaxies. For minimum stellar masses of ≈ 11.2, appropriate for the LOWZ and CMASS selections, the inclusion of satellites can boost the lensing and kSZ effect signals by up to ≈ 30 per cent and ≈ 20 per cent, respectively.

choices of the minimum stellar mass cut. For the lensing comparison (top panels), we show only the L1 and C2 bins for brevity, noting that the L2 and C1 bins give similar results. The solid curves show the ratios for the fiducial FLAMINGO feedback model and correspond to different minimum stellar masses: $\log_{10}[M_{\rm star}/{\rm M}_\odot]$ values ranging between 10.9 and 11.6 in steps of 0.1 and in the legend we quote the satellite fractions corresponding to these minimum stellar mass selections.

It is first worth noting that all of the curves in Fig. 4 are below 1, meaning that the fiducial selection including satellites leads to boosted mean lensing and kSZ signals relative to a central-only selection. This implies that of the two effects discussed above; i.e. mis-centring versus satellites living in a higher-mass host, it is the latter that dominates. It also implies that analyses that do not account for satellites will likely tend to overestimate the halo mass required to match the lensing and kSZ measurements and therefore potentially underestimate the impact of feedback.

From our previous analysis, we found that minimum stellar masses in the range of ≈ 11.15–11.25 best describe the selection for the BOSS LOWZ and CMASS samples. For these selections, Fig. 4 implies that the inclusion of satellites can boost the lensing and kSZ effect signals by up to ≈ 30 per cent and ≈ 20 per cent, respectively. The effect is scale dependent, which is likely due to the mis-centring of satellites with respect to their hosts. One can therefore potentially constrain the satellite fraction through detailed measurements of the shape of the lensing and kSZ effect profiles.

While the effects of satellites are not large enough to alter our conclusion that stronger feedback is required to match the kSZ effect measurements relative to an X-ray-based calibration strategy, they nevertheless should be factored in for quantitative analyses. Furthermore, according to Fig. 4, as observations push to lower stellar masses, the role of satellites will become more significant (given the large satellite fractions) and will need a careful accounting.

4.3 Implied halo masses

Using the simulated galaxy selection that best fits the lensing measurements it is straightforward to compute a mean halo mass. We quote the mean halo masses in terms of $M_{\rm 500c}$ in order to place the them in the context of X-ray samples. Note that for satellite galaxies, we use the $M_{\rm 500c}$ value associated with the FOF group in which the satellite resides. Weighting each selected galaxy equally yields mean halo masses of $\log_{10}[M_{\rm 500c}/{\rm M}_\odot] = 13.53 \pm 0.02$ and 13.34 ± 0.04 for the LOWZ and CMASS samples, respectively. Thus, the lensing data yield a $\lesssim 10$ per cent constraint on the halo mass at 2σ uncertainty. In the above, we have combined the L1 and L2 (C1 and C2) constraints into a single mean halo mass using inverse-variance weighting for LOWZ (CMASS), given that

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

the individual bins are consistent (within 2σ). We find that these constraints on the mean halo mass are virtually independent of which FLAMINGO model we use to compute the mean halo masses, which we attribute to the fact that we adjust the minimum stellar mass for each model to refit the lensing data. This includes runs that vary the galaxy stellar mass function, which require a significantly different minimum stellar mass to match the lensing data.

We note that if we apply the same minimum stellar mass cuts as in the fiducial selection but limit our analysis to central galaxies only, we find lower mean halo masses[5] of $\log_{10}[M_{500c}/\mathrm{M}_\odot] = 13.32 \pm 0.03$ and 13.13 ± 0.06 for the LOWZ and CMASS samples, respectively. This confirms that the inclusion of satellite galaxies in the selection boosts the mean halo mass of the sample and by consequence also the stacked lensing and kSZ effect signals. If we were instead to re-fit the minimum stellar mass for the central-only selection, the derived mean halo mass would increase, as expected. But note that a central-only selection does not provide a statistically good fit to the lensing data, so the derived mean halo mass would have questionable value.

5 DISCUSSION

5.1 Implications for feedback models

Here, we discuss the interpretation of our findings in the context of feedback modelling. This work makes crucial steps in confirming the preference for stronger feedback from WL + kSZ without the reliance on a baryonification model. The origin of the difference in the implied required strength of feedback from the kSZ effect measurements and the X-ray-based baryon fraction measurements is unclear. Taking the observations at face value and as discussed previously by Bigwood et al. (2024), the discrepancy between X-ray and kSZ effect constraints on the strength of feedback can be explained by: (1) differences between the real and simulated mass- and/or redshfit-dependencies of feedback; (2) differences between the real and simulated scale-dependence of feedback effects; and/or (3) the possibility of unaccounted for systematics or selection effects in the measurements.

5.1.1 Mass and redshift dependence of feedback models

One way to potentially reconcile the kSZ effect profiles with X-ray-based gas fraction constraints is to appeal to the different halo masses and redshifts that they probe. For example, the BOSS CMASS sample has a mean redshift of $z \approx 0.54$ and a halo mass of $M_{500c} \approx 2.2 \times 10^{13}\,\mathrm{M}_\odot$ (see Section 4.3), whereas the X-ray measurements of gas fractions are generally confined to groups at $z \lesssim 0.3$ with halo masses of $10^{14}\,\mathrm{M}_\odot$, although with considerable variation about these typical values. Therefore, if the effective halo mass and/or redshift dependencies of feedback in the simulations differ from those in nature (such that the simulated feedback has a weaker dependence on mass and/or redshift than in reality), one could potentially understand the difference between the kSZ effect and group gas fractions calibrations.

However, such an explanation is made more difficult by the inclusion of the LOWZ sample here, since its mean redshift is lower at $z \approx 0.31$ and its mean halo mass is slightly higher at $3.4 \times 10^{13}\,\mathrm{M}_\odot$. These have some overlap with the X-ray samples used to calibrate the simulations, including FLAMINGO.

5.1.2 Scale dependence of feedback

The mean halo masses of the LOWZ and CMASS samples imply that the angular scales corresponding to R_{500c} at $z = 0.31$ (LOWZ) and $z = 0.54$ (CMASS) are $\theta_{500c} \approx 1.61$ and 0.92 arcmins, respectively. These scales are indicated by vertical dashed lines in Fig. 2, for example. For LOWZ the first data corresponds roughly to $0.65\,R_{500c}$, whereas for CMASS the first data point is $\approx 1.1\,R_{500c}$. Most of the constraining power on the feedback models is therefore coming from $\approx$ 1.5-3 R_{500c} for LOWZ and from $\approx$ 2.5-5 R_{500c} for CMASS. The X-ray-constrained baryon fractions, by contrast, are measured within R_{500c}.

Therefore, a way to potentially reconcile the X-ray and kSZ measurements is to invoke more steeply declining gas density profiles at large radii (beyond R_{500c}) than are seen in the simulations. What physical mechanism could result in the required steepening is unclear. In addition, it would be interesting to calculate if the required steepening is consistent with the constraint that, according to Planck Collaboration XI (2013), the tSZ effect within $5R_{500c}$ scales self-similarly with halo mass over a very wide range of masses, implying that haloes are fully baryon loaded within that aperture, at least at low redshift. But we leave this question for future work.

5.1.3 Unaccounted for systematics in the measurements

It is possible that, instead of there being a fundamental issue with the feedback in the simulations, there could be unaccounted for (or mischaracterized) systematic errors in the gas fraction, kSZ effect, or galaxy–galaxy lensing measurements. For example, there are considerable uncertainties in the X-ray selection function of galaxy groups (e.g. Pearson et al. 2017; Giles et al. 2022; Marini et al. 2025) and some studies make no attempt to account for selection effects. For the kSZ effect measurements, a significant underestimate of the bias in the velocity reconstruction, or from tSZ effect leakage in the kSZ effect stacking analysis, could potentially reconcile the measurements with the simulation predictions using X-ray-calibrated feedback. Ideally, we would use the same theory (hydrodynamical simulations) to select our systems, analyse them in precisely the same way as done for the real systems, and compare the processed observables in a like-with-like fashion. While we have taken important steps in this direction in the present study, the issues of the X-ray selection function and kSZ effect velocity reconstruction (and/or tSZ effect leakage) remain open questions. Finally, we have used galaxy–galaxy lensing measurements to constrain the selection of simulated galaxies and therefore any biases present in the lensing measurements will impact our feedback conclusions. However, as previously noted, we find excellent consistency in the derived mean halo masses from the lensing with that inferred from fitting to the large-scale projected clustering, suggesting any biases in the lensing masses are likely to be small.

5.2 Implications for the S_8 tension

Our analyses of the independent LOWZ and CMASS samples suggests that feedback stronger than adopted in the fiducial FLAMINGO model is required to match the stacked kSZ effect measurements of Schaan et al. (2021). The strong FLAMINGO fgas-8σ variant yields a reasonably good fit to the kSZ effect profiles. We examine the impact of the stronger feedback in this model on the 3D matter power spectrum as well as on other observable measures of clustering.

[5]We find consistent constraints with the central galaxy only mean halo masses if, instead of computing the mean of the selected central galaxies, we simply fit a composite NFW profile + 2-halo term (using the halo bias model of Tinker et al. 2010) to the lensing data, using the COLOSSUS package (Diemer 2018).

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

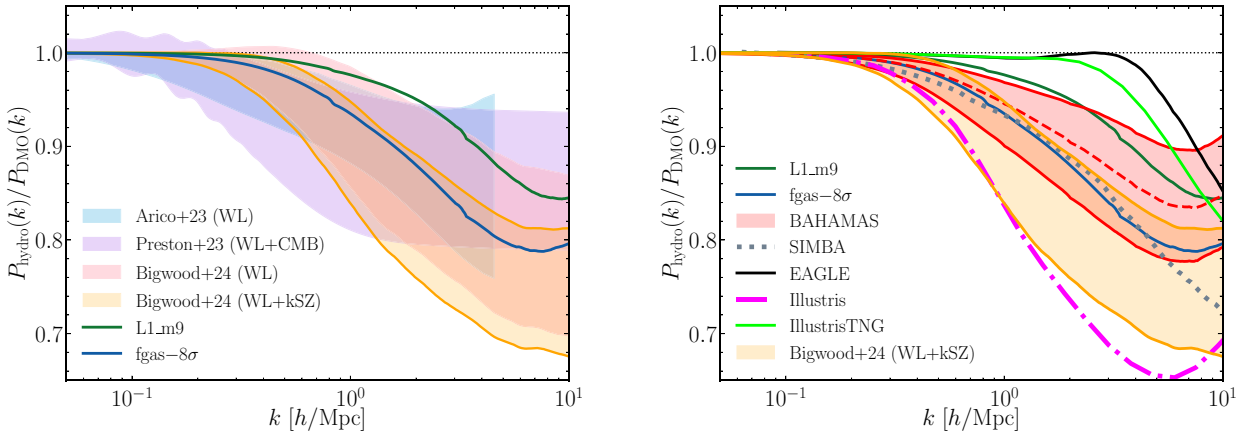

Figure 5. Suppression of the 3D matter power spectrum at $z = 0$ due to baryon feedback. *Left:* Comparison of the fiducial FLAMINGO feedback model and the strong fgas-8σ variant with recent observational constraints on the suppression. The solid curves shows the suppression predicted by the FLAMINGO models. The shaded regions correspond to constraints from: Preston et al. (2023), who jointly model the DES Y3 cosmic shear and *Planck* 2018 primary CMB data; Aricò et al. (2023) and Bigwood et al. (2024) (WL), who re-analyse DES Y3 data including small-scale shear measurements (see the text); and Bigwood et al. (2024) (WL + kSZ), who perform a joint analysis of cosmic shear and the kSZ effect CMASS measurements of Schaan et al. (2021). *Right:* Comparison with other recent hydrodynamical simulations, including: BAHAMAS (McCarthy et al. 2017, 2018) (the red shaded region encloses the low-AGN and hi-AGN variants, with the dashed red curve corresponding to the fiducial BAHAMAS model); SIMBA (Davé et al. 2019), EAGLE (Schaye et al. 2015), the original Illustris simulations (Vogelsberger et al. 2014), IllustrisTNG (Pillepich et al. 2018). Recent WL-based constraints favour stronger suppressions and are compatible with the predictions of the FLAMINGO fgas-8σ variant, as well as SIMBA and BAHAMAS hi-AGN.

5.2.1 *Cosmic shear/3D matter power spectrum*

In Fig. 5 (left panel), we compare the predictions of the suppression of the matter power spectrum from fiducial FLAMINGO model and the strong fgas-8σ variant with the recent constraints from Bigwood et al. (2024), Aricò et al. (2023), and Preston et al. (2023) (see also Schaller et al. 2025). The Preston et al. (2023) constraint comes from a joint analysis of the DES Y3 data and the *Planck* 2018 primary CMB. The constraints shown in Fig. 5 correspond to their best-fitting empirical 'Amod' model, which is intended to parameterize possible modifications of the non-linear part of the matter power spectrum due to baryonic physics and/or non-standard dark matter. The two constraints from Bigwood et al. (2024) come from an analysis of DES Y3 cosmic shear data (labelled WL) or from a joint analysis of the DES Y3 cosmic shear and the kSZ effect CMASS profile of Schaan et al. (2021) (labelled WL + kSZ), modelled using the BCEmu baryonification formalism (Schneider et al. 2019; Giri & Schneider 2021). Note that Bigwood et al. (2024) do not adopt the small scale cuts employed in the fiducial DES Y3 cosmic shear analysis (e.g. Krause et al. 2021; Amon et al. 2022; Secco, Samuroff et al. 2022) that were designed to mitigate the impact of baryonic physics, but instead include the small scale measurements and use the baryonification formalism to model the baryonic effects, aided by the inclusion of kSZ effect data. Aricò et al. (2023) use the BACCO baryonification emulator[6] (Aricò et al. 2021) to model the DES Y3 cosmic shear data (including small scale measurements).

Examining Fig. 5, we see that the fgas-8σ variant is in good agreement with the suppression constraints from Bigwood et al. (2024), for both their WL only and WL + kSZ analysis, as well as those from Aricò et al. (2023). There is also reasonable agreement with the constraints from Preston et al. (2023), with a slight mismatch on intermediate scales ($0.1 \lesssim k[h/\mathrm{Mpc}] \lesssim 1$), such that the fgas-8$\sigma$ variant predicts a bit too much power. As noted above, a difference between the constraints of the other studies and those of Preston et al. (2023) is that the latter also fit to the primary CMB measurements. Nevertheless, the agreement with the other constraints is generally good. It is worth highlighting that the observational constraints shown in Fig. 5 are likely to be model dependent, particularly on small scales (k greater than a few h/Mpc) which current cosmic shear measurements are less sensitive to.

Bigwood et al. (2024) derive a value of $S_8 = 0.818^{+0.017}_{-0.024}$ from their cosmic shear-only analysis and $S_8 = 0.823^{+0.019}_{-0.020}$ from their cosmic shear + kSZ effect analysis, which are consistent with the *Planck* primary CMB estimate of $S_8 = 0.832 \pm 0.013$ (Planck Collaboration et al. 2020) at the 0.7σ and 0.4σ levels, respectively (i.e. no significant tension). Using the same DES Y3 cosmic shear data, Aricò et al. (2023) find a slightly lower value of $S_8 = 0.795^{+0.015}_{-0.017}$. Note that even prior to the inclusion of small scales and associated baryonic modelling, the DES Y3 cosmic shear results were not in strong tension with the *Planck* primary CMB (typically 2σ level). The inclusion of small scales, updated modelling of intrinsic alignments, and inclusion of baryonic effects has weakened this tension. The KiDS 1000 constraints, by contrast, typically show stronger levels of tension ($\gtrsim 3\sigma$), which would require feedback in excess of that implied by the kSZ effect measurements studied here (e.g. Heymans et al. 2021; Amon & Efstathiou 2022; Schneider et al. 2022; McCarthy et al. 2023).

In the right panel of Fig. 5 we compare the fiducial FLAMINGO model and the fgas-8σ variant with the predictions of other recent hydrodynamical simulations, including: BAHAMAS (McCarthy et al. 2017, 2018, $L_{\rm box} = 571$ Mpc) (the red shaded region encloses the low-AGN and hi-AGN variants, with the dashed red curve corresponding to the fiducial BAHAMAS model); SIMBA (Davé et al. 2019, $L_{\rm box} = 147$ Mpc), EAGLE (Schaye et al. 2015, $L_{\rm box} = 100$ Mpc), the original Illustris simulations (Vogelsberger et al. 2014, $L_{\rm box} = 107$ Mpc), andIllustrisTNG (Pillepich et al. 2018, $L_{\rm box} = 300$

[6]We note that the BCEmu and BACCO model differences have been studied and shown to impact the shape of the predicted power spectrum (Grandis et al. 2024). A primary difference is that the former displaces formally all the particles in the simulation, whereas the latter displaces only particles in haloes.

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

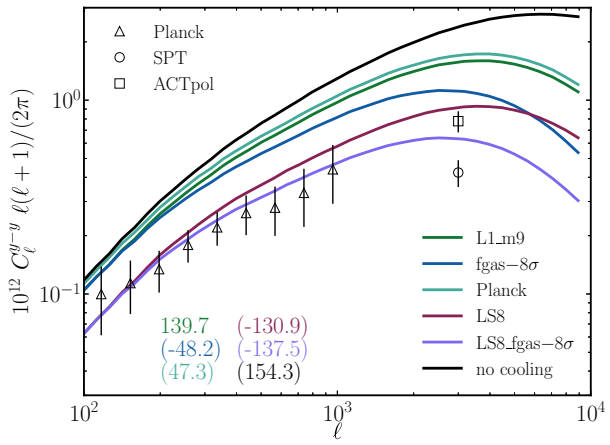

Figure 6. The tSZ effect angular power spectrum. The open triangles correspond to the *Planck* tSZ measurements of Bolliet et al. (2018) while the open circle and open square correspond to the SPT and ACTpol measurements of Reichardt et al. (2021) and Choi et al. (2020), respectively. Following the same colour coding, the numbers in the bottom left indicate χ^2 for the L1_m9 run and the $\Delta\chi^2$ (in parentheses) of the other runs with respect to L1_m9. Note that the χ^2 is computed here with respect to the Planck tSZ measurements only, using the diagonal elements of the covariance matrix. A negative value for the $\Delta\chi^2$ indicates a better match to the data. A lensing LS8 cosmology yields a significantly better fit to the data relative to the fiducial D3A cosmology, whereas a *Planck* CMB cosmology yields a worse fit. Increasing the efficiency of feedback (fgas-8σ), as preferred by kSZ effect results in the present study, improves the fit on small angular scales. But the offset on large scales cannot be reconciled through feedback alone.

Mpc). Recent WL-based constraints favour stronger suppressions and are compatible with the predictions of the FLAMINGO fgas-8σ variant, as well as SIMBA and BAHAMAS hi-AGN.

5.2.2 *Thermal SZ power spectrum*

Moving on from cosmic shear, previous studies have shown that various measures of the tSZ effect are in tension with the standard model fit to the primary CMB (e.g. McCarthy et al. 2014; Planck Collaboration XXI 2014, XXII 2016a, XXIV b; Bolliet et al. 2018; McCarthy et al. 2018). Note that the tSZ effect (e.g. its PDF, power spectrum, and number counts) is typically sensitive to the most massive clusters and is therefore expected to be more sensitive to parameters such as σ_8 and $\Omega_{\rm m}$ than cosmic shear. The 'catch' is that the tSZ effect needs to be cleanly separated from various other signals present in the CMB temperature maps, including the primary CMB, the kSZ effect, radio sources, the CIB, etc., which is non-trivial.

We revisit the tSZ effect power spectrum and the tSZ effect-shear cross-spectrum examined for FLAMINGO in McCarthy et al. (2023). We refer the reader to that study for a full description of how the cosmological observables are calculated.

In Fig. 6, we compare selected FLAMINGO runs to the latest tSZ effect power spectrum measurements, namely the *Planck*-based measurements reported in Bolliet et al. (2018), the SPT measurement from Reichardt et al. (2021), and the ACTpol measurement from Choi et al. (2020). Note that Bolliet et al. (2018) present an improved re-analysis of the *Planck* 2015 tSZ data set from Planck Collaboration XXII (2016a), by taking into account the tri-spectrum in the covariance matrix and placing physical constraints on the amplitudes of foreground contaminants (particularly radio and infrared point sources and the clustered infrared background, or CIB). The ACTpol measurement was not considered in McCarthy et al. (2023) as we were unaware of it at the time. The measurement from ACTpol is significantly higher in amplitude compared to previous ACT measurements (e.g. Sievers et al. 2013) and compared to the SPT measurement of Reichardt et al. (2021). The origin of the differences between the ACTpol measurements and the previous ACT measurements was not discussed in Choi et al. (2020).

Comparing the simulations to the observational measurements, we see that both the fiducial feedback model (L1_m9) and fgas-8σ variant in the fiducial D3A cosmology are strongly in tension with the *Planck* tSZ power spectrum measurements on large scales, which is sensitive mainly to very massive clusters that are generally unaffected by feedback (e.g, McCarthy et al. 2014). At the smaller scales probed by SPT and ACT, increased feedback yields a qualitatively better match to the observational measurements though the discrepancy between the SPT and ACTpol measurements, prevents a quantitative assessment of the goodness of fit. A *Planck* primary CMB cosmology with fiducial feedback yields a slightly worse match to the measurements than the D3A cosmology. The LS8 cosmology with fiducial feedback, by contrast, yields a significantly improved fit to the *Planck* tSZ power spectrum measurements on large scales and with the ACTpol measurements on small scales. Increasing the feedback in an LS8 cosmology (LS8_fgas-8σ) further improves the fit, though the improvement is small compared to the improvement that resulted from lowering S_8.

5.2.3 *Thermal SZ–cosmic shear cross-spectrum*

In Fig. 7, we compare selected FLAMINGO runs to the tSZ effect-cosmic shear cross-spectrum measurements of Tröster et al. (2022), who cross-correlated the 5 tomographic KiDS1000 bins with tSZ effect maps constructed from the *Planck* 2015 data set (Planck Collaboration XXII 2016a). In agreement with the tSZ effect power spectrum analysis, we see that both the fiducial feedback model and fgas-8σ variant in the fiducial D3A cosmology are strongly in tension with the measurements on large scales. At smaller scales ($\ell \gtrsim 700$), increased feedback yields a better match to the observational measurements. A *Planck* primary CMB cosmology with fiducial feedback yields a slightly worse match to the measurements than the D3A cosmology. The LS8 cosmology with fiducial feedback, by contrast, provides a significantly better match on all scales compared to the fiducial model in either the D3A or *Planck* cosmologies. Increasing the feedback in the LS8 cosmology slightly improves the fit relative to the fiducial feedback model in the LS8 cosmology.

5.2.4 *Summary*

The above comparisons indicate that when adopting a *Planck*-like cosmology (such as D3A or the *Planck* maximum likelihood cosmology), increasing the strength of baryonic feedback generally serves to improve the match to various observed clustering statistics, primarily through reducing the power on small scales. Importantly, though, on large scales the tSZ effect power spectrum and tSZ effect-shear cross-spectrum are insensitive to the effects of feedback even when the feedback is increased significantly to match the kSZ effect measurements examined here. Increasing the feedback improves the match on small scales in a *Planck*-like cosmology but only marginally improves it in an LS8 cosmology.

As the tSZ effect measurements appear strongly in tension with the standard model fit to the primary CMB and on the largest scales are generally insensitive to baryonic effects, it will be important to further scrutinize the robustness of these measurements, including the robustness of the component separation techniques employed to

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

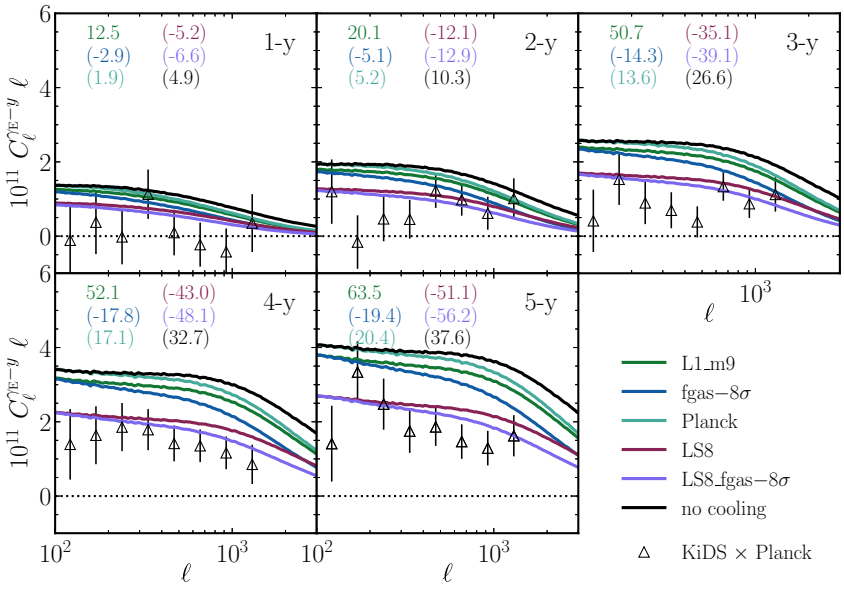

Figure 7. The cosmic shear–tSZ effect angular cross-power spectrum. The open triangles correspond to the KiDS1000 × *Planck* measurements of Tröster et al. (2022). The different panels correspond to the cross-spectrum between different KiDS tomographic bins (1–5) and the tSZ data. Similar to case of the tSZ power spectrum (see Fig. 6), a lensing LS8 cosmology yields a significantly better fit to the data relative to the fiducial D3A cosmology, whereas a Planck CMB cosmology yields a worse fit. Increasing the efficiency of feedback (fgas-8σ and LS8-fgas-8σ), improves the fit on small angular scales. But the offset on large scales cannot be reconciled through feedback alone.

separate the tSZ effect from the other signals present in the data such as the CIB and radio sources. In this regard, cross-correlations with other signals, such as cosmic shear, are invaluable since they will be differently affected by any residual contamination than is the case for the tSZ effect power spectrum. Relativistic effects should also be included in future analyses (e.g. Remazeilles et al. 2019; Remazeilles & Chluba 2025). Furthermore, we note that the baryonification formalism has recently been extended to tSZ effect modelling by Aricò & Angulo (2024) and it would be interesting to compare constraints from this methodology with those derived from the hydrodynamical simulations.

6 CONCLUSIONS

In this work, we have used the state-of-the-art FLAMINGO suite of cosmological hydrodynamical simulations to compare to kSZ effect stacking measurements of Schaan et al. (2021), based on a sample selection derived from fitting galaxy–galaxy lensing measurements (Amon et al. 2023). This is the first time the powerful combination of kSZ effect and galaxy–galaxy lensing measurements has been used as a benchmark for feedback modelling in cosmological hydrodynamical simulations.

We used associated full-sky light cone-based HEALPIX maps and catalogues to perform a like-with-like comparison with these observations for the first time in the context of self-consistent hydro simulations. We explored the dependence of the predictions on feedback efficiency and implementation as well as on cosmology. The use of cosmological simulations also allowed us to quantitatively assess the impact that satellite galaxies included in the stack have on the resulting lensing and kSZ effect profiles.

We have highlighted that a quantitative interpretation of the kSZ effect measurements requires precise knowledge of the halo mass scale of the BOSS samples, since the amplitude of the kSZ effect scales approximately with the halo mass. A meaningful comparison to the predictions of simulations (whatever the feedback implementation) requires that we select simulated systems that have an implicit mean halo mass matching that of the observed systems in the kSZ effect stacking. To this end, a unique aspect of this work is that we have employed high-precision measurements of the stacked galaxy–galaxy lensing profiles of the BOSS LOWZ and CMASS from Amon et al. (2023) using DES Y3 + KiDS1000 data to better than 10 per cent mass constraints (2σ). We have used the lensing measurements to select the simulated galaxies required for a fair comparison of the kSZ effect predictions to the measurements of Schaan et al. (2021).

Our main finding is that the kSZ effect measurements imply that more aggressive feedback is required in the simulations compared to that inferred from X-ray cluster observations and that this goes some way to alleviating the difference in the observed and predicted clustering on small scales. However, we have shown that the offsets on large scales between measurements and predictions of the tSZ power spectrum and its cross-spectrum with cosmic shear cannot be resolved through increased feedback. In more detail, our findings may be summarized as follows:

(i) A simple minimum stellar mass-based selection employed on the simulations yields galaxy–galaxy lensing profiles which match the measurements of Amon et al. (2023) over the wide range of observed radii, with the best-fitting minimum stellar masses, but not the best-fitting haloes masses, depending slightly on the feedback model (Figs 1 and C1) and the adopted cosmology. Higher minimum stellar masses (and halo masses) are required to match the LOWZ sample compared to the CMASS sample. Invoking more complicated selection functions that depend on both stellar mass and specific star formation rate yields results consistent with our fiducial analysis (see Appendix A). The derived mean halo masses for the CMASS and LOWZ samples are $\log_{10}[M_{500c}/\mathrm{M}_\odot] = 13.53 \pm 0.02$ and 13.34 ± 0.04 (at 2σ uncertainty), respectively.

(ii) Using the lensing measurements as a strong constraint on the selection for the kSZ effect predictions, we find that the fiducial

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

FLAMINGO feedback model, which was calibrated on the X-ray gas fractions of low-redshift galaxy groups and clusters, is ruled out at the 5σ level by both the BOSS LOWZ and CMASS samples (Fig. 2). The strongest feedback variant considered in the FLAMINGO suite, the so-called fgas-8σ model, provides a good match to the measurements, though is inconsistent with the local X-ray measurements.

(iii) Our conclusions are robust to the details of the feedback implementation, in that we find similar results for the jet-based AGN feedback implementation as we do for the fiducial thermal (isotropic) implementation when both models are calibrated to the same gas fraction data (Fig. 3). The main determining factor of the predicted kSZ signal is the fraction of baryons that are retained by haloes. The kSZ effect results are also generally insensitive to reasonable variations in the cosmological parameters.

(iv) Both the galaxy–galaxy lensing and kSZ effect profiles are boosted by the inclusion of satellite galaxies included in the stacking analyses. For the BOSS selection, as implied by the lensing measurements, the boost is up to ≈ 30 per cent and ≈ 20 per cent, respectively (Fig. 4). This effect increases with decreasing galaxy stellar mass and will therefore need to be carefully modelled as observations push to lower masses.

(v) We have discussed various ways to reconcile the differing constraints on feedback modelling from X-ray and kSZ effect measurements (see Section 5.1), including (1) that there could be differences between the real and simulated halo mass- or redshfit-dependencies of feedback; (2) differences between the real and simulated scale-dependence of feedback effects; and/or (3) unaccounted for systematics or selection effects in either the X-ray or kSZ effect measurements. We have used lensing-based halo mass measurements to highlight that the kSZ effect measurements of the BOSS galaxies typically probe gas at several virial radii, whereas the X-ray measurements are confined to smaller scales.

(vi) In light of the finding that stronger feedback is required to reproduce the kSZ effect measurements, we have reassessed the impact of feedback on several clustering statistics, including the matter power spectrum, relevant for cosmic shear (Fig. 5), the tSZ effect power spectrum (Fig. 6) and the tSZ effect cross-spectrum (Fig. 7). When adopting a *Planck*-like cosmology (such as D3A or the *Planck* maximum likelihood cosmology), increasing the strength of baryonic feedback generally serves to improve the match to various observed clustering statistics, primarily through reducing the power on small scales, in general agreement with the findings of McCarthy et al. (2023) and previous studies employing dark matter only simulations with corrections for baryonic effects (e.g. Amon & Efstathiou 2022; Preston et al. 2023). The level of tension between a *Planck*-like cosmology and the DES Y3 cosmic shear measurements was already relatively low in the case of standard feedback modelling and the inclusion of stronger feedback reduces the tension further. However, the tSZ effect power spectrum and tSZ effect-shear cross-spectrum are dominated by more massive haloes and are insensitive to the effects of feedback on large scales even when the feedback is increased significantly to match the kSZ effect measurements examined here. Increasing the feedback improves the match on small scales in a *Planck*-like cosmology but the improvement is only marginal in the 'lensing' cosmology (LS8).

Whilst in the final stages of preparation of this study, Hadzhiyska et al. (2024a) posted to the arxiv a new *Planck* + ACT kSZ effect stacking analysis of galaxies in the DESI Legacy Imaging Survey, using photometric redshifts in the velocity reconstruction. The use of photometric redshifts leads to a larger bias in the reconstructed velocities, which has been estimated using DESI mocks (Ried Guachalla et al. 2024; Hadzhiyska et al. 2024b). Hadzhiyska et al. (2024a) define four tomographic bins spanning $0.4 \lesssim z \lesssim 1.0$ (i.e. similar to the range probed by BOSS LOWZ and CMASS) and compare their stacking results with the Illustris-TNG (Pillepich et al. 2018, $L_{\rm box} = 300$ Mpc) and the original Illustris (Vogelsberger et al. 2014, $L_{\rm box} = 107$ Mpc) simulations, although without using light cones and adopting a simplified abundance matching approach for the selection of simulated galaxies. Note that the original Illustris simulation has considerably stronger feedback than Illustris-TNG and that the gas fractions of the former are considerably below what is observed for X-ray-selected galaxy groups (whereas the latter lie considerably above the X-ray measurements). Consistent with this study, Hadzhiyska et al. (2024a) find that strong feedback is required to match their stacked profiles and that the original Illustris simulation provides a good fit to the data, whereas Illustris-TNG is ruled out at many sigma.

Upcoming LSS surveys demand an accurate and precise feedback model that is consistent with a wide range of observations. Toward this goal, X-ray measurements remain a valuable source of information on the hot gas properties of groups/clusters with new insights coming from eROSITA data. Understanding the impact of X-ray selection effects will be crucial (e.g. Kugel et al. 2024; Popesso et al. 2024). kSZ effect measurements binned by mass and redshift (e.g. via stacking on DESI galaxies) will be a rich source of information that is complementary to that of the X-ray measurements in that it can probe larger scales and higher redshifts. For more nearby and massive galaxies, there will also be the opportunity to directly compare the X-ray and kSZ effect measurements to test for consistency. Comparisons with complementary probes of the hot gas, such as through radio-based dispersion measure data (e.g. Macquart et al. 2020) and so-called 'patchy screening' of the CMB (e.g. Coulton et al. 2024) should also provide important insights.

ACKNOWLEDGEMENTS

We thank the referee for a helpful report that improved the paper. This work was supported by the Science and Technology Facilities Council (grant number ST/Y002733/1). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 769130). REA acknowledges support from the Spanish Ministry of Science under grant number PID2021-128338NB-I00. CSF acknowledges support from the European Research Council (ERC) through Advanced Investigator grant DMI-DAS (GA 786910). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

DATA AVAILABILITY

The data supporting the plots within this article are available on reasonable request to the corresponding author. The FLAMINGO simulation data will eventually be made publicly available, though we note that the data volume (several petabytes) may prohibit us from simply placing the raw data on a server. In the meantime, people interested in using the simulations are encouraged to contact the corresponding author.

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

REFERENCES

Abbott T. M. C. et al., 2022, Phys. Rev. D, 105, 023520
Aghanim N. et al., 2019, A&A, 632, A47
Ahn C. P. et al., 2014, ApJS, 211, 17
Akino D. et al., 2022, PASJ, 74, 175
Amodeo S. et al., 2021, Phys. Rev. D, 103, 063514
Amon A. et al., 2022, Phys. Rev. D, 105, 023514
Amon A. et al., 2023, MNRAS, 518, 477
Amon A., Efstathiou G., 2022, MNRAS, 516, 5355
Aricò G., Angulo R. E., 2024, A&A, 690, A188
Aricò G., Angulo R. E., Contreras S., Ondaro-Mallea L., Pellejero-Ibañez M., Zennaro M., 2021, MNRAS, 506, 4070
Aricò G., Angulo R. E., Zennaro M., Contreras S., Chen A., Hernández-Monteagudo C., 2023, A&A, 678, A109
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Bigwood L. et al., 2024, MNRAS, 534, 655
Bleem L. E. et al., 2022, ApJS, 258, 36
Bolliet B., Comis B., Komatsu E., Macías-Pérez J. F., 2018, MNRAS, 477, 4957
Borrow J., Schaller M., Bower R. G., Schaye J., 2022, MNRAS, 511, 2367
Brown S. T., McCarthy I. G., Stafford S. G., Font A. S., 2022, MNRAS, 509, 5685
Bulbul E. et al., 2024, A&A, 685, A106
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Castro T., Borgani S., Dolag K., Marra V., Quartin M., Saro A., Sefusatti E., 2021, MNRAS, 500, 2316
Chandro-Gómez Á., Lagos C. D. P., Power C., Moreno Forouhar V. J., Helly J. C., Lacey C. G., McGibbon R. J., Schaller M., Schaye J., 2025. MNRAS, 539, 776
Chisari N. E. et al., 2019, Open J. Astrophys., 2, 4
Choi S. K. et al., 2020, J. Cosmol. Astropart. Phys., 2020, 045
Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, MNRAS, 452, 1217
Coulton W. R. et al., 2024, preprint (arXiv:2401.13033)
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Debackere S. N. B., Schaye J., Hoekstra H., 2020, MNRAS, 492, 2285
Diemer B., 2018, ApJS, 239, 35
Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108
Eckert D. et al., 2016, A&A, 592, A12
Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114
Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019, PASA, 36, e021
Elbers W. et al., 2025, MNRAS, 537, 2160
Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2021, MNRAS, 507, 2614
Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2022, MNRAS, 516, 3821
Forouhar Moreno V. J., Helly J., McGibbon R., Schaye J., Schaller M., Han J., Kugel R., 2025, preprint (arXiv:2502.06932)
Giles P. A. et al., 2022, MNRAS, 511, 1227
Giri S. K., Schneider A., 2021, J. Cosmol. Astropart. Phys., 2021, 046
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Grandis S., Aricò G., Schneider A., Linke L., 2024, MNRAS, 528, 4379
Greco J. P., Hill J. C., Spergel D. N., Battaglia N., 2015, ApJ, 808, 151
Hadzhiyska B. et al., 2024a, preprint (arXiv:2407.07152)
Hadzhiyska B., Ferraro S., Ried Guachalla B., Schaan E., 2024b, Phys. Rev. D, 109, 103534
Hahn O., Rampf C., Uhlemann C., 2021, MNRAS, 503, 426
Han J., Cole S., Frenk C. S., Benitez-Llambay A., Helly J., 2018, MNRAS, 474, 604
Han J., Jing Y. P., Wang H., Wang W., 2012, MNRAS, 427, 2437
Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, MNRAS, 479, 5385
Henden N. A., Puchwein E., Sijacki D., 2020, MNRAS, 498, 2114
Heymans C. et al., 2021, A&A, 646, A140
Krause E. et al., 2021, preprint (arXiv:2105.13548)
Kugel R. et al., 2023, MNRAS, 526, 6103
Kugel R., Schaye J., Schaller M., McCarthy I. G., Braspenning J., Helly J. C., Moreno V. J. F., McGibbon R. J., 2024, MNRAS, 534, 2396
Kukstas E. et al., 2023, MNRAS, 518, 4782
Le Brun A. M. C., McCarthy I. G., Melin J.-B., 2015, MNRAS, 451, 3868
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118
Macquart J. P. et al., 2020, Nature, 581, 391
Maraston C. et al., 2013, MNRAS, 435, 2764
Marini I. et al., 2025, A&A, 694, A207
McCarthy I. G. et al., 2023, MNRAS, 526, 5494
McCarthy I. G., Bird S., Schaye J., Harnois-Deraps J., Font A. S., van Waerbeke L., 2018, MNRAS, 476, 2999
McCarthy I. G., Le Brun A. M. C., Schaye J., Holder G. P., 2014, MNRAS, 440, 3645
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
Mead A. J., Tröster T., Heymans C., Van Waerbeke L., McCarthy I. G., 2020, A&A, 641, A130
Melin J. B., Bartlett J. G., Tarrío P., Pratt G. W., 2021, A&A, 647, A106
Mummery B. O., McCarthy I. G., Bird S., Schaye J., 2017, MNRAS, 471, 227
Pearson R. J. et al., 2017, MNRAS, 469, 3489
Peebles P. J. E., 2024, Philo. Trans. A, 383, 20240021
Pillepich A. et al., 2018, MNRAS, 473, 4077
Planck Collaboration VI, 2020, A&A, 641, A6
Planck Collaboration XI, 2013, A&A, 557, A52
Planck Collaboration XXI, 2014, A&A, 571, A21
Planck Collaboration XXII, 2016a, A&A, 594, A22
Planck Collaboration XXIV, 2016b, A&A, 594, A24
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Popesso P. et al., 2024, MNRAS, 527, 895
Preston C., Amon A., Efstathiou G., 2023, MNRAS, 525, 5554
Reichardt C. L. et al., 2021, ApJ, 908, 199
Remazeilles M., Bolliet B., Rotti A., Chluba J., 2019, MNRAS, 483, 3459
Remazeilles M., Chluba J., 2025, MNRAS, 538, 1576
Ried Guachalla B., Schaan E., Hadzhiyska B., Ferraro S., 2024, Phys. Rev. D, 109, 103533
Rogers K. K., Hložek R., Laguë A., Ivanov M. M., Philcox O. H. E., Cabass G., Akitsu K., Marsh D. J. E., 2023, J. Cosmol. Astropart. Phys., 2023, 023
Salcido J., McCarthy I. G., 2024, preprint (arXiv:2409.05716)
Salcido J., McCarthy I. G., Kwan J., Upadhye A., Font A. S., 2023, MNRAS, 523, 2247
Schaan E. et al., 2021, Phys. Rev. D, 103, 063513
Schaller M. et al., 2024, MNRAS, 530, 2378
Schaller M., Schaye J., Kugel R., Broxterman J. C., van Daalen M. P., 2025, MNRAS, 539, 1337
Schaye J. et al., 2015, MNRAS, 446, 521
Schaye J. et al., 2023, MNRAS, 526, 4978
Schneider A., Giri S. K., Amodeo S., Refregier A., 2022, MNRAS, 514, 3802
Schneider A., Stoira N., Refregier A., Weiss A. J., Knabenhans M., Stadel J., Teyssier R., 2020, J. Cosmol. Astropart. Phys., 2020, 019
Schneider A., Teyssier R., 2015, J. Cosmol. Astropart. Phys., 2015, 049
Schneider A., Teyssier R., Stadel J., Chisari N. E.,Le Brun A. M. C., Amara A., Refregier A. , 2019, J. Cosmol. Astropart. Phys., 2019, 020
Secco L. F., Samuroff S., et al., 2022, Phys. Rev. D, 105, 023515
Semboloni E., Hoekstra H., Schaye J., 2013, MNRAS, 434, 148
Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sievers J. L. et al., 2013, J. Cosmol. Astropart. Phys., 2013, 060
Springel V. et al., 2018, MNRAS, 475, 676
Sun M., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., 2009, ApJ, 693, 1142

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G.,Gottlöber S. , 2010, ApJ, 724, 878
Tröster T. et al., 2022, A&A, 660, A27
van Daalen M. P., McCarthy I. G., Schaye J., 2020, MNRAS, 491, 2424
van Daalen M. P., Schaye J., 2015, MNRAS, 452, 2247
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Yang T., Cai Y.-C., Cui W., Davé R., Peacock J. A., Sorini D., 2022, MNRAS, 516, 4084

APPENDIX A: GALAXY SELECTION INCLUDING SSFR CUTS

As described in Section 3.1, our fiducial analysis employs a simple minimum stellar mass cut for the simulated galaxies. We vary the minimum stellar mass to determine the cut which results in the best match to the galaxy–galaxy lensing measurements. The real BOSS selection functions employ colour cuts, such that the CMASS and (particularly) the LOWZ samples are composed primarily of red galaxies that lack significant ongoing star formation. As a check of whether our fiducial stellar mass cut-only selection method is sufficient, we have also explored selections which apply cuts in both stellar mass and sSFR. In this appendix, we show the impact of this more complicated 'quenched' selection function for the fiducial FLAMINGO feedback model.

We first impose a stellar mass cut as per the default analysis. From the selected galaxies we compute a histogram of sSFRs. The peak corresponds to the star-forming main sequence. Following Kukstas et al. (2023), we use the galaxies with sSFRs above the star-forming main sequence to determine the standard deviation (width) of this sequence. We designate quenched galaxies as those with sSFRs that are at least 3 sigma below the sSFR of the star-forming main sequence. We select these galaxies and compute their galaxy–galaxy lensing signal. We repeat this process for different choices of the minimum stellar mass cut and determine the value which best matches the galaxy–galaxy lensing measurements.

We show the results for the fiducial feedback model in Fig. A1. The quality of the fit to the measurements is similar to that for our default stellar mass-limited analysis (Fig. 1) and the best-fitting minimum stellar mass for the quenched sample is consistent with that for the default selection.

We show the resulting kSZ effect predictions for this quenched selection in Fig. A2. In agreement with our default stellar mass cut-only analysis (in Fig. 2), we find the fiducial FLAMINGO feedback model is ruled out at high significance for the quenched selection. Thus, our main conclusion, that the kSZ effect measurements of Schaan et al. (2021) prefer stronger feedback than in the fiducial FLAMINGO model, for which the feedback was calibrated using X-ray cluster observations, is insensitive to the details the simulated galaxy selection so long as the selection matches the galaxy–galaxy lensing measurements.

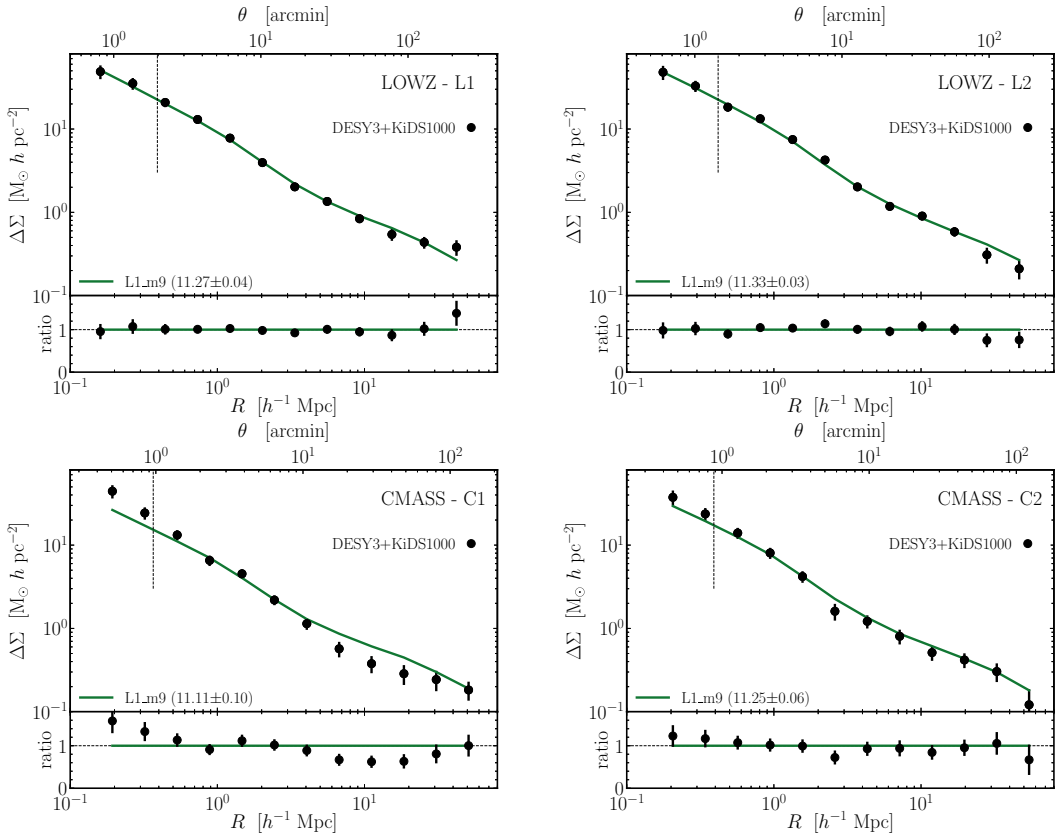

Figure A1. As Fig. 1, but showing the impact of selecting only simulated galaxies without significant ongoing star formation for the fiducial FLAMINGO model. The best-fitting minimum stellar mass consistent with that for our default stellar mass-limited selection in Fig. 1.

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

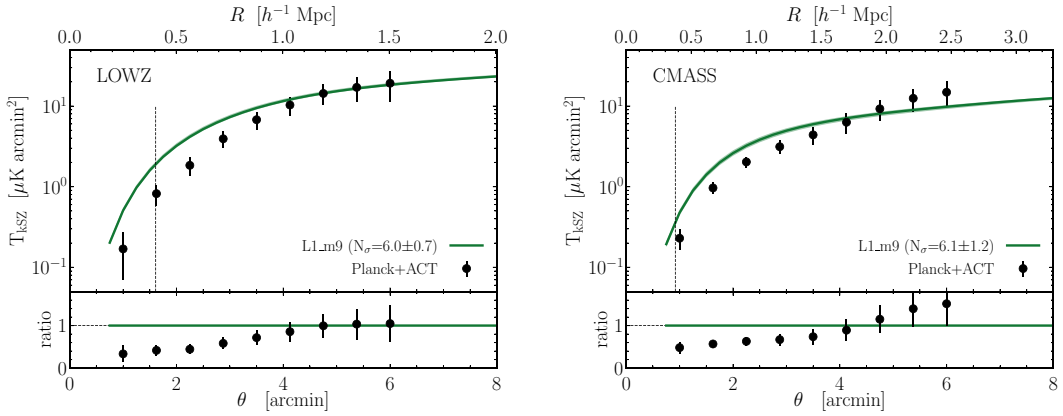

Figure A2. As Fig. 2, but showing the impact of selecting only simulated galaxies without significant ongoing star formation for the fiducial FLAMINGO model. The solid curves represent the predictions for the best-fitting minimum stellar mass (fitted to the lensing, see Fig. A1). Selection of quenched galaxies only slightly improves the fit to the observational measurements, but the fiducial model is still strongly ruled out by the data.

APPENDIX B: RESOLUTION DEPENDENCE

In Fig. B1, we compare the predictions of the high-resolution calibrated FLAMINGO model (L1_m8) with the fidicual resolution calibrated model (L1_m9). As per our default analysis, to predict the kSZ effect we first determine the minimum galaxy stellar mass limit that best matches the galaxy–galaxy lensing measurements of Amon et al. (2023) independently for the two models. For the high-resolution model, the best-fitting masses are $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.35 \pm 0.04$ and 11.26 ± 0.06 for the LOWZ and CMASS samples, respectively. These are the best-fitting masses from fitting the two redshift bins in each sample (e.g. LOWZ-L1 and LOWZ-L2) jointly. These values are somewhat higher than the best-fitting joint masses for the fiducial resolution run, which are $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.26 \pm 0.02$ and 11.18 ± 0.04 for LOWZ and CMASS, respectively, (though only the LOWZ difference is statistically significant). This difference is expected since, although both models were independently calibrated to the observed galaxy stellar mass function, the calibration to the data is not perfect. Furthermore, the simulations were only calibrated up to $\log_{10}[M_{\rm star}/{\rm M}_\odot] = 11.5$ and higher mass do affect the sample. We note that in the galaxy stellar mass function the high-resolution run is offset from the fiducial resolution run (and the observational measurements) by ≈ 0.1–0.2 dex at a stellar mass scale of $\sim 10^{11}{\rm M}_\odot$, resulting in larger stellar masses compared to the fiducial resolution run at a halo mass scale of $\sim 10^{13}{\rm M}_\odot$ (see fig. 9 of Schaye et al. 2023). Thus, to match the galaxy–galaxy lensing signal, a higher stellar mass cut is required for the high-resolution run.

When the galaxy selection is fixed by the lensing, the predicted kSZ effect profiles are similar (see Fig. B1), typically deviating from each by less than the observational measurement errors, and both are in strong tension with the measurements of Schaan et al. (2021).

APPENDIX C: GALAXY–GALAXY LENSING OF ADDITIONAL FLAMINGO VARIATIONS

In Fig. C1, we show the best-fitting galaxy–galaxy lensing profiles for the FLAMINGO variations shown in Fig. 3. Consistent with the analyses of the other runs, we find that all models yield an acceptable fit to the DES Y3 + KiDS 1000 measurements given an appropriate

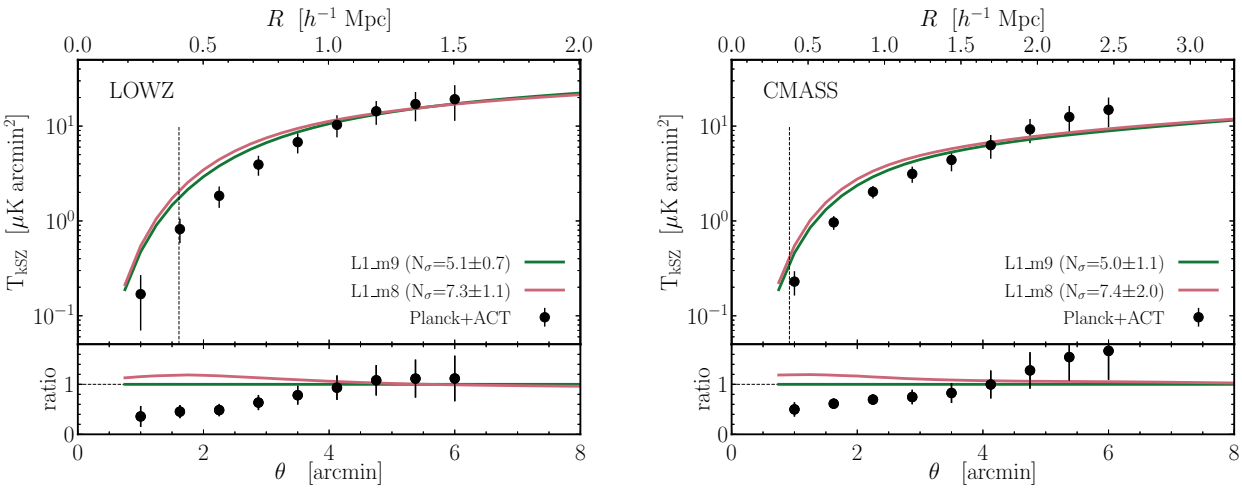

Figure B1. As Fig. 2 but comparing the fiducial resolution calibrated model (L1_m9) with the high-resolution calibrated model (L1_m8). Overall the profiles are similar (typically deviating from each by less than the observational measurement errors) and both are in strong tension with the measurements of Schaan et al. (2021).

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

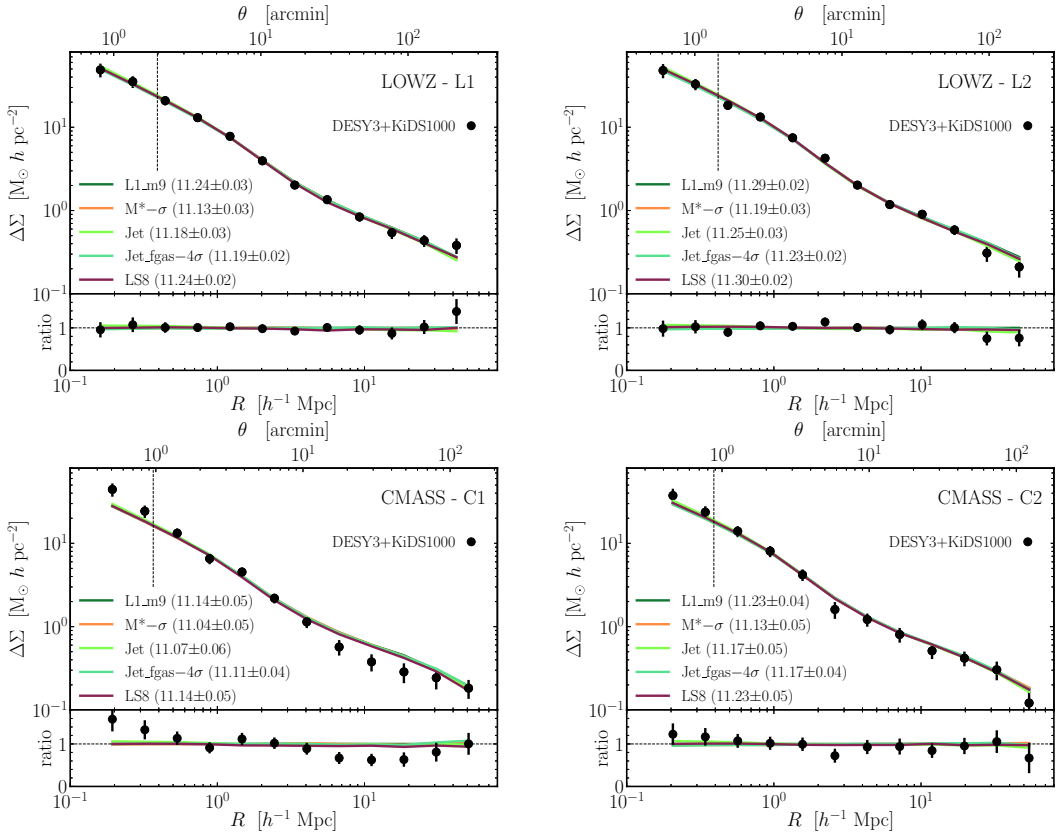

Figure C1. As Fig. 1, but showing the dependence of the galaxy–galaxy lensing profiles on other feedback variations, namely variations in the stellar mass function (both the fiducial and reduced cluster gas fractions) and the fiducial and strong jet models of AGN feedback. The solid coloured curves correspond to the FLAMINGO simulations as baryon models are varied, with the best-fitting stellar masses provided in the legend.

choice of the minimum stellar mass. Note that, as expected, the minimum stellar mass required to match the lensing measurements differs significantly for the run that varied the stellar mass function (M∗−σ) by approximately 0.1 dex. While the minimum stellar mass is lower, we highlight that the derived mean halo mass of the selected sample agrees remarkably well with that of the other runs.

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

Downloaded from https://academic.oup.com/mnras/article/540/1/143/8124379 by guest on 29 January 2026

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.