Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **552,** 1–19 (2026)
Advance Access publication 2026 October 1

https://doi.org/10.1093/mnras/stag1865

# Implications of a contracted dark matter halo for the Milky Way's inferred virial mass

Diego Dado [1,2★], Shaun T. Brown [1,3], Azadeh Fattahi [1,3], Andreea S. Font[4] and Ian G. McCarthy[4]

[1]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[2]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[3]*The Oskar Klein Centre, Department of Physics, Stockholm University, Albanova University Center, SE-106 91 Stockholm, Sweden*
[4]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*

Accepted 2026 September 23. Received 2026 August 28; in original form 2026 March 13

## ABSTRACT

We investigate how reliably the global properties of Milky Way-mass dark matter haloes can be recovered from dynamical data from a limited range of radii, particularly $\lesssim$30 kpc where observations are most sensitive but baryonic processes modify the halo structure. Using the ARTEMIS (Assembly of high-ResoluTion Eagle simulations of MIlkyWay-type galaxieS) simulations, which produce varying degrees of baryon-induced contraction, we fit dark matter profiles over restricted radial ranges using commonly adopted parametric models. Assuming negligible observational uncertainties allows the systematic errors from these choices to be isolated. When fits are confined to inner radii, a Navarro–Frenk–White (NFW) profile underestimates the virial mass by a factor of $\approx$2 on average ($\approx$4 for some systems), and the concentration by a factor of $\approx$2. Einasto and generalized-NFW models provide excellent local fits but retain similar global biases. In contrast, a contracted halo prescription yields stable extrapolations and recovers unbiased halo mass estimates over all radii. The inferred mass improves systematically with increasing radial coverage, and tracers beyond $\gtrsim$50 kpc largely eliminate the mean bias for all models. The local dark matter density at the Solar radius is recovered to within $\lesssim$5% for all profiles other than NFW. These biases are sufficient to reconcile recent low Milky Way-mass estimates derived from inner rotation-curve analyses with the canonical $\approx 10^{12}$ M$_\odot$. We additionally find a halo-to-halo scatter of $\gtrsim$0.1 dex ($\approx$25%) persists even under idealized conditions, setting a likely lower limit for the precision of halo mass estimates.

**Key words:** galaxies: evolution – galaxies: kinematics and dynamics – dark matter.

## 1 INTRODUCTION

In the Lambda cold dark matter ($\Lambda$CDM) cosmogony, galaxies form at the centres of virialized dark matter (DM) haloes, with their subsequent evolution shaped by the interplay between hierarchical accretion, gas cooling, and feedback from stars and active galactic nuclei (e.g. W. H. Press & P. Schechter 1974; J. Binney 1977; C. S. Frenk et al. 1988; A. Cimatti, F. Fraternali & C. Nipoti 2019; A. Marasco et al. 2021). The mass and internal structure of a galaxy's host halo are therefore inextricably linked to its own formation and evolution.

Observationally, we can infer the distribution of DM within galaxies through the kinematics of the stars and gas within different components of a galaxy, with such analysis providing some of the strongest early evidence for the existence of DM (e.g. V. C. Rubin, W. K. Ford & N. Thonnard 1980). Typically, for nearby galaxies similar to the Milky Way (MW), the most robust kinematic data comes from the disc ($\lesssim$30 kpc). With dedicated surveys, we now have robust kinematic data for large samples of galaxies in the nearby Universe (e.g. M. Persic, P. Salucci & F. Stel 1996; F. Walter et al. 2008; D. A. Hunter et al. 2012). Even within this context, observations within the MW offer unparalleled detail to map the DM distribution within our host halo (e.g. reviews by J. S. Bullock & M. Boylan-Kolchin 2017; J. Zavala & C. S. Frenk 2019). In addition to the quality of data within the MW, we also have access to a range of kinematic probes beyond the central disc, including the stellar halo (e.g. X. X. Xue et al. 2008; A. J. Deason et al. 2012; P. R. Kafle et al. 2012), globular clusters (e.g. G. M. Eadie, A. Springford & W. E. Harris 2017; L. Posti & A. Helmi 2019; L. L. Watkins et al. 2019) and satellite galaxies (e.g. L. L. Watkins, N. W. Evans & J. H. An 2010; Z.-Z. Li et al. 2017; E. Patel et al. 2018; T. M. Callingham et al. 2019).

Reliable theoretical models for the distribution of mass with DM are necessary to be able to robustly interpret kinematic data. In the absence of baryonic effects it is well established that haloes follow on average a Navarro–Frenk–White (NFW; J. F. Navarro, C. S. Frenk & S. D. M. White 1997) profile, which exhibits an inner density slope of $\rho \propto r^{-1}$ and an outer slope of $\rho \propto r^{-3}$, with a characteristic scale radius, $r_s$, describing the transitions between the two limits. An NFW profile has two free parameters, typically phrased as the halo mass and concentration, defined

★ E-mail: diego.dado2@durham.ac.uk

© The Author(s) 2026.
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 academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

as the ratio of the virial to scale radius. Cosmological simulations predict these to be strongly correlated, leading to a well-defined concentration-mass relation that varies with redshift and cosmology (e.g. A. A. Dutton & A. V. Macciò 2014; B. Diemer & A. V. Kravtsov 2015; A. D. Ludlow et al. 2016; S. T. Brown et al. 2020). These models predict a MW mass halo at $z = 0$ to have a mean concentration of $c \approx 8$, with a scatter of $\approx$0.1 dex, where the concentration is defined relative to a spherical overdensity radius $R_{200c}$ – i.e. the radius enclosing a mean matter density 200 times the critical density at $z = 0$.

While the internal structure of DM haloes in the absence of baryons is relatively well understood, it has been shown that the inclusion of baryonic processes can significantly impact the central structure of the DM halo. In some cases, strong feedback-driven gaseous outflows can lead to the formation of a DM core due to energy transfer from a time-varying gravitational potential (e.g. A. Pontzen & F. Governato 2012; A. Di Cintio et al. 2014a). The efficiency of this process is strongly mass dependent, with the largest cores forming at intermediate stellar-to-halo mass ratios, $M_\star/M_{\rm halo} \sim 10^{-2}$–$10^{-3}$, and diminishing both at lower and higher masses (e.g. A. Di Cintio et al. 2014a; A. Lazar et al. 2020). Recent work has further shown that this behaviour is closely linked to the burstiness of star formation, helping to reconcile results across different simulation suites and to understand the halo-to-halo scatter in such formalisms (A. M. Garcia et al. 2026; J. Sarrato-Alós et al. 2026). For systems of MW mass or larger, many simulations instead predict that the central concentration of baryons within the galaxy can lead to a contracted DM halo, with an increased inner density relative to a dark-matter-only (DMO) simulation – an effect arising from principles of action conservation and gas cooling when baryons are condensing at the centre of such haloes (e.g. G. R. Blumenthal et al. 1986; O. Y. Gnedin et al. 2004; M. Schaller et al. 2015; R. Poole-McKenzie et al. 2020; T. M. Callingham et al. 2020).

This significantly complicates the task of inferring the host DM halo properties from galactic dynamics, as the host halo not only influences the evolution of the galaxy, but the galaxy also influences the halo itself. The most precise dynamical constraints are often obtained from the disc of a galaxy, where baryons dominate the potential, and therefore from the same region where the impact of baryons on the DM halo is expected to be most significant. Consequently, without a reliable theoretical model to account for baryonic effects, inferences of DM halo properties may be subject to significant systematic errors.

In recent years, a tension has emerged in estimates of the MW halo mass between measurements based on dynamics in the central Galaxy ($5 < r/{\rm kpc} \lesssim 30$) and those using tracers in the outer halo ($r \gtrsim 30$ kpc). Studies using large, precise data sets from *APOGEE* DR17 (S. R. Majewski et al. 2017) and *Gaia* Data Release 3 (Gaia Collaboration 2023), but restricted to the disc, have inferred masses of $\approx 10^{11}$–$5 \times 10^{11}\,{\rm M}_\odot$, depending on the assumed profile (e.g. A.-C. Eilers et al. 2019; X. Ou et al. 2024), while analyses using tracers beyond the disc typically find $\approx 10^{12}\,{\rm M}_\odot$ (e.g. T. M. Callingham et al. 2019; L. L. Watkins et al. 2019).[1] Reconciling this discrepancy is crucial for a broad range of applications, from predictions of Local Group satellite abundances to the local DM density relevant for direct-detection experiments (M. W. Goodman & E. Witten 1985; A. K. Drukier, K. Freese & D. N. Spergel 1986; G. Jungman, M. Kamionkowski & K. Griest 1996).

Addressing this problem requires a statistically meaningful sample of MW-mass haloes simulated at sufficiently high resolution to robustly resolve the inner DM distribution, while simultaneously forming realistic central galaxies. In particular, the simulations must capture the response of the DM halo to baryonic assembly and feedback, which depends sensitively on both numerical resolution and the calibration of subgrid physics. To this end, we use the ARTEMIS (Assembly of high-ResoluTion Eagle simulations of MIlkyWay-type galaxieS) suite of high-resolution cosmological zoom-in hydrodynamical simulations of MW-mass haloes (A. S. Font et al. 2020). ARTEMIS comprises a large sample of haloes with $M_{200c} \sim 10^{12}\,{\rm M}_\odot$, simulated at baryonic mass resolution $m_{\rm b} \simeq 2 \times 10^4\,{\rm M}_\odot$, and employs a stellar feedback model calibrated to reproduce the median stellar mass-halo mass relation inferred from abundance matching applied to large extragalactic surveys. As such, ARTEMIS is not tuned to reproduce the detailed properties of the MW specifically, but instead provides a representative population of MW-mass galaxies suitable for assessing systematic biases in halo inference methods. By fitting the DM profiles of ARTEMIS haloes over a range of radii and assuming a variety of parametric halo models commonly used in the literature, including the contracted halo model of Cautun et al. (2020, hereafter C20), we quantify how accurately halo mass and concentration can be recovered when only kinematic tracers over a limited radial range are available.

The paper is structured as follows. Section 2 summarizes the ARTEMIS simulations, the DM-rescaling procedure, and our fitting methodology. Section 3 quantifies the impact of baryonic contraction on the inner DM distribution and verifies the C20 power-law relation in ARTEMIS. Section 4 presents the performance and characteristic behaviour of each halo model as a function of radial fitting extent, both in local/internal metrics and in global/extrapolated metrics. Section 5 summarizes our proposed interpretation of these findings, especially the implications for MW (and analogues) mass and concentration measurements.

## 2 METHODS

We use the ARTEMIS cosmological hydrodynamical simulations to study how MW-mass galaxies alter their host DM haloes. In this section we outline the details of the simulations, the halo mass definitions used, profile construction, and fitting methodology.

### 2.1 The ARTEMIS simulations

ARTEMIS (A. S. Font et al. 2020) is a suite of 42 cosmological zoom-in hydrodynamical simulations of MW-mass haloes. These simulations are run using the GADGET-3 code (last described in V. Springel 2005) with the same hydrodynamics solver and galaxy formation (subgrid) physics as the EAGLE project (J. Schaye et al. 2015), but at a significantly increased resolution. This increase required a recalibration of key subgrid modules, most notably stellar feedback, to reproduce the median stellar mass-halo mass relation at $z = 0$ for MW-mass systems; a detailed description of the simulation setup and calibration is presented in A. S. Font et al. (2020). The recalibration was not intended to match the detailed properties of the MW specifically, but rather to yield a rep-

[1] In this context, 'halo mass' or 'mass' typically refers to a spherical overdensity definition such as $M_{200c}$ or $M_{\rm vir}$; these definitions differ at the ∼10–20 per cent level and are therefore not sufficient to explain the discrepancy.



Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

resentative population of MW-mass galaxies. As such, ARTEMIS is well suited for studying generic baryonic effects on DM haloes – including baryonic contraction – and for quantifying systematic biases in halo property inference from kinematic data confined to the inner regions of galaxies.

The ARTEMIS simulations adopt a flat ΛCDM cosmology consistent with the nine-year data release of the *Wilkinson Microwave Anisotropy Probe* (*WMAP9*; G. Hinshaw et al. 2013), with parameters: $\Omega_m = 0.2793$, $\Omega_b = 0.0463$, $h = 0.70$, $\sigma_8 = 0.8211$, and $n_s = 0.972$. The simulations are run in a parent volume of $25\,\mathrm{Mpc}/h$ per side, with the gravitational softening length set to $125\mathrm{pc}/h$ for all particles, and with a baryon particle mass of $2.23 \times 10^4 \mathrm{M}_\odot/h$.

In this work, we focus exclusively on the central haloes of the ARTEMIS suite, which have virial masses in the range $0.8 \leq M_{200c}/10^{12}\mathrm{M}_\odot \leq 2$. This yields a sample of 42 central haloes.

### 2.2 Halo finding and virial parameters

Haloes are identified using the friends-of-friends algorithm (M. Davis et al. 1985) with a linking length of 0.2 times the mean particle separation, and they are further split into gravitationally bound structures using the SUBFIND algorithm (V. Springel, N. Yoshida & S. D. M. White 2001). As is common, we use a spherical overdensity definition for the mass and size of a halo. In this approach, the halo is bounded by a radius $R_\Gamma$ within which the mean enclosed density equals Γ times the critical density of the Universe, $\rho_c$. Formally,

$$\frac{3M(<r)}{4\pi\, r^3} = \Gamma \rho_c, \tag{1}$$

where $M(<r)$ is the total mass enclosed, $\rho_c(z) = 3H^2(z)/8\pi G$ is the critical density at redshift $z$, and $H(z)$ is the redshift-dependent Hubble parameter. The mass and radius that satisfy this relation are then defined as halo's mass and radius, $M_\Gamma$ and $R_\Gamma$. Throughout this work, we adopt an overdensity threshold of $\Gamma = 200$ relative to the critical density. The corresponding halo mass and radius are denoted as $M_{200c}$ and $R_{200c}$. It is very common to refer to these definitions simply as the virial mass and radius of the halo, and we will use this terminology throughout to refer to the above definition.

#### 2.2.1 *DM rescaling*

In the halo mass definition described above, it is most common to use the *total* enclosed mass. In Fig. 1, we show the difference in halo mass using the standard definition, comparing the ARTEMIS haloes in a DMO simulation with full hydrodynamics (see blue data points). For the DMO simulations, all matter is in the form of DM, while for the hydrodynamic simulation the standard halo mass definition includes both DM and baryons. When comparing these simulations, this definition of the halo mass leads to the baryonic simulation having systematically less massive haloes by ≈15 per cent, with significant scatter between systems.

One of the limitations of using all matter to define a halo's mass is that it becomes unclear if this discrepancy is due to changes in the distribution of DM, baryons, or both. In this work, we are primarily focussed on the effect of the baryons on the distribution of DM, and less on the distribution of baryons themselves. Therefore, to isolate the changes to the DM distribution, we choose to

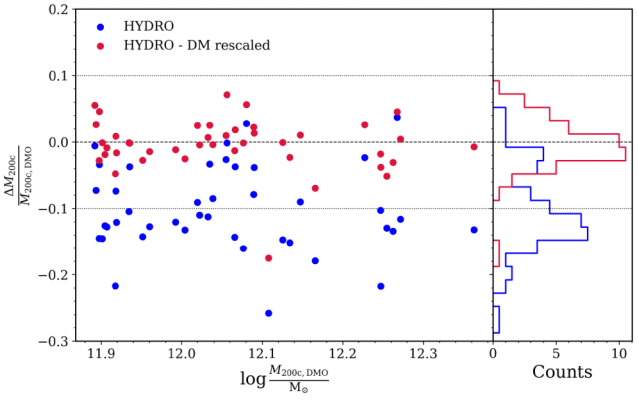



**Figure 1.** Fractional difference in virial mass, $(M_{200c} - M_{200c,DMO})/M_{200c,DMO}$, between the hydrodynamical and DMO counterparts of each ARTEMIS central halo. Colours denote the definition of $M_{200c}$: dark blue for the standard total enclosed mass definition (equation 1), and dark red for the dark matter-rescaled definition, which scales the enclosed DM mass by the cosmic mean baryon fraction (equation 2). *Left panel:* Fractional mass differences as a function of the DMO virial mass; each circle corresponds to a single central halo. *Right panel:* Distribution of these fractional differences. In both panels, the dashed black line marks perfect agreement between hydrodynamical and DMO masses, while the grey dotted lines indicate ±10 per cent deviations. Applying the DM-rescaled definition greatly suppresses the systematic offset, showing that the dominant contribution to the total mass mismatch arises from baryons displaced from the halo rather than from any substantial loss of dark matter.

modify the halo mass definition to the following:

$$\frac{1}{1-f_b}\frac{3M_{DM}(<r)}{4\pi\, r^3} = \Gamma \rho_c, \tag{2}$$

where $f_b$ is the cosmic baryon fraction, which for our assumed cosmology is $f_b = 0.17$. This definition is similar in form to the standard choice (equation 1), and still corresponds to a spherical overdensity mass definition, but differs by using only the distribution of DM, scaled according to the baryon fraction. In a DMO simulation ($f_b = 0$) the two definitions are identical.

In Fig. 1, we show the halo mass difference with this modified definition using only the DM (red points). This halo mass definition largely removes the systematic offset in virial mass seen when using the standard convention, with the distribution having no average systematic difference and a scatter of ≈5 per cent. This highlights that the majority of the virial mass difference between hydrodynamical and DMO runs arises from baryons being expelled from the halo, rather than a substantial change in DM enclosed within the outer halo near the virial radius.[2] Consequently, the DM rescaled definition of virial parameters provides a robust framework for comparing halo properties between hydrodynamical and DMO runs, isolating differences which reflect purely structural changes in the DM distribution.

Throughout the rest of the paper, we only present results using the mass definition based on the rescaled DM distribution (equation 2). For brevity, we will use the terminology 'halo mass',

[2]This result holds for our sample of exclusively MW mass haloes, but may not generalize to all mass scales. For example, it has been shown that for dwarf galaxies there is notable change to the DM halo mass (e.g. T. Sawala et al. 2015).

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

'virial mass' and $M_{200c}$ to refer to this definition. We also note that one outlier halo in Fig. 1, exhibiting a systematic offset of $\sim -17$ per cent under the DM-rescaled mass definition, is removed from the sample for the rest of the paper, and is a particular unrelaxed system with an ongoing major merger.

### 2.3 Circular velocity fitting and theoretical DM profiles

To quantify how well different theoretical halo models reproduce the simulated DM structure, we fit parametric profiles to the radial mass distributions of each halo.

We calculate the spherically averaged density, enclosed mass, and circular velocity radial profiles for all haloes in our sample. Throughout, we show results using $N_{\rm bins} = 128$ evenly spaced logarithmic bins in radius spanning $0.21 \leq r/{\rm kpc} \leq 320$. Roughly corresponding to between $10^{-3}$ and 1.5 times the typical virial radius (the median virial radius of the sample is $R_{200c} = 214\,{\rm kpc}$). We have tested the impact of the number of radial bins by repeating our analysis with $N_{\rm bins} = 32, 64, 256$; we find broadly similar results when fitting enclosed mass and circular velocity profiles, and only mild differences when fitting the density profile. We adopt 128 bins, as this is the regime where the fitting results converge for all profiles, both in local estimates and extrapolated quantities.

We obtain the best-fit parameters and associated uncertainties for each halo model through least-squares optimization. Specifically, we minimize

$$\sigma_{\rm fit} = \frac{1}{N-1}\sum_{i}^{N}(\log M(<r_i) - \log M_{\rm model}(<r_i))^2, \quad (3)$$

where the sum runs over the radial bins $r_i$ used in the fitting procedure. Here $M(<r)$ is the enclosed mass measured from the simulation – $M_{\rm DM,DMO}(<r)$ and $M_{\rm DM,\ rescaled}(<r)$ for the DMO and hydrodynamical runs, respectively – and $M_{\rm model}(<r)$ is the model prediction. Although we perform our fits using the enclosed mass profile, we have verified that fitting the circular velocity profile $V_c(r)$ – which more closely resembles typical observational probes – yields consistent results and does not affect our conclusions.

Furthermore, while the outer radius used in the fitting procedure will be varied, we fix the lower bound to exclude regions where numerical effects can lead to biased estimates of the circular velocity (see e.g. C. Power et al. 2003). Our fixed lower bound is given by $r_{\rm min.\ fit} = 3{\rm max}(r_{\rm conv})$. The convergence radius, $r_{\rm conv}$, is calculated following the prescription of A. D. Ludlow, J. Schaye & R. Bower (2019), and depends on the local particle number and enclosed volume. Although all simulations share the same mass resolution, the convergence radius varies between haloes owing to differences in their internal structure and particle distributions. We therefore adopt a conservative threshold defined as three times the maximum $r_{\rm conv}$ across our halo sample – which corresponds to $\approx$3 kpc for our haloes. Note that some of the most recent works targeting an estimate of the Galactic halo mass adopt a cut at $r \approx 5\,{\rm kpc}$, due to the complexity of modelling the strongly non-axisymmetric potential induced by the Galactic bar (e.g. X. Ou et al. 2024); hence, our cut at $\approx$3 kpc is broadly comparable to those adopted in observational analyses.

Having defined our fitting procedure, we now introduce the analytic DM density profiles used in this work.

#### 2.3.1 NFW

One of the most widely adopted models for DM haloes is the NFW profile (J. F. Navarro et al. 1997). Its functional form is given by

$$\rho_{\rm NFW}(r) = \frac{\rho_0}{\frac{r}{r_s}(1+\frac{r}{r_s})^2}, \quad (4)$$

where $\rho_0$ is the characteristic density and $r_s$ is the scale radius. This profile features an inner slope of $-1$ and an outer slope of $-3$, with a transition at the scale radius. It is common to parametrize NFW haloes in terms of concentration, $c = R_{200c}/r_s$, and virial mass, $M_{200c}$, instead of $\rho_0$ and $r_s$.

To generalize the definition of concentration across different functional density profiles we adopt the convention

$$c = R_{200c}/r_{-2},$$

where $r_{-2}$ is the radius where the density's logarithmic slope is equal to $-2$, i.e.

$$\left.\frac{{\rm d}\log\rho}{{\rm d}\log r}\right|_{r_{-2}} = -2. \quad (5)$$

For an NFW halo $r_{-2} = r_s$.

To allow comparison in extrapolated halo mass across different DM profile models, we report $M_{200c}$ for each model, following the DM-rescaled definition of virial mass in hydrodynamical runs or the standard definition for DMO runs (see Section 2.2.1). For profiles where a closed-form analytical expression for the mass enclosed profile can be derived, $M_{200c}$ is computed analytically; otherwise, it is obtained by numerically integrating the fitted density profile.

#### 2.3.2 Generalized NFW

The generalized NFW (gNFW) profile[3] extends the standard NFW model by introducing a variable inner slope, $\gamma$, which allows for greater flexibility in describing the inner structure of DM haloes; the outer slope is fixed to be $-3$ as in the NFW profile. This additional degree of freedom allows the profile to represent diverse radial density profiles, including cored haloes that depart from the canonical NFW cusp.

The gNFW profile is defined as

$$\rho_{\rm gNFW}(r) = \frac{\rho_0}{r^{\gamma}(r+r_s)^{3-\gamma}}, \quad (6)$$

where $r_s$ is the scale radius (i.e. the slope transition scale), and $\rho_0$ is the characteristic density. For this profile, there exists a simple relationship between the scale radius and the radius we use for the concentration, $r_{-2}$, given by

$$r_{-2} = (2-\gamma)r_s. \quad (7)$$

To ensure physically plausible fits, we constrain the slope parameter to the range $\gamma \in [0, 2]$.

[3]The phrase 'generalized NFW' is not consistently used to refer to the same parametric profile within the literature. While it is common to mean a profile with a free inner and outer slope, we use this to refer to a profile with a fixed outer slope and free inner slope, consistent with recent works inferring the MW's mass from rotation curve analysis (e.g. X. Ou et al. 2024).



Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

### 2.3.3 *Einasto*

The Einasto profile (J. Einasto 1965) is another widely adopted model for describing DM halo density profiles, both in simulations and observational studies. Like gNFW, it allows for a cored or cuspy inner density; however, instead of a broken power law, it features a continuously varying logarithmic slope governed by an exponential function.

Its functional form is

$$\rho_{\rm Ein}(r) = \rho_{-2} \exp(-2/\alpha[(r/r_{-2})^{\alpha} - 1]), \quad (8)$$

where $\alpha$ is the shape parameter, $r_{-2}$ is the radius where the logarithmic slope is $-2$ (following equation 5), and $\rho_{-2}$ is the density at $r_{-2}$. The parameter $\alpha$ controls the steepness of the profile, with larger values of $\alpha$ producing sharper central concentrations, and lower values yielding flatter or core-like distributions. As for gNFW, we ensure physical fits by restricting the shape parameter to $\alpha > 0$. For $\alpha \approx 0.16$ Einasto and NFW profiles look similar over radial ranges typically probed by simulations and observations.

### 2.3.4 *Contracted NFW*

Baryonic processes can significantly modify the inner DM distribution in MW-mass haloes. Among several proposed parametrizations of these effects, we adopt the phenomenological model introduced by C20, which was calibrated using matched hydrodynamical and DMO simulations to relate the degree of halo contraction to the relative enclosed mass profiles of the two runs. Within this section, we introduce the mathematical framework of the model and corresponding definitions, following the original authors' parameter conventions. Later, in Section 3.1, we explicitly test the model predictions against the ARTEMIS simulations, which were not used in the original development and calibration of the mass–ratio relation (equation 11).

C20 found that the ratio of enclosed DM mass,

$$\eta_{\rm DM}(r) = M_{\rm DM}(< r)/M_{\rm DM,DMO}(< r), \quad (9)$$

correlates closely with the ratio of the *total* enclosed mass,

$$\chi_{\rm tot}(r) = M_{\rm DM,DMO}(< r)/(M_{\rm DM}(< r) + M_{\rm bar}(< r)), \quad (10)$$

between a DMO and hydrodynamical simulation of MW-mass haloes. The relation is well described by a power law,

$$\eta_{\rm DM} = A\chi_{\rm tot}^{B}, \quad (11)$$

with $A = 1.023 \pm 0.001$ and $B = -0.540 \pm 0.002$ calibrated from simulations.

The contracted profile can then be approximated as

$$M_{\rm DM,CNFW}(< r) = M_{\rm NFW}(< r) f(\eta_{\rm bar}), \quad (12)$$

with

$$f(\eta_{\rm bar}) = [0.45 + 0.38(\eta_{\rm bar} + 1.16)^{0.53}],$$

where $\eta_{\rm bar}$ is defined as the ratio between the baryonic mass and the NFW mass profile (or any halo model representing the DMO DM mass profile), scaled by the cosmic baryon fraction,

$$\eta_{\rm bar}(r) = \frac{1}{f_{\rm bar}} \frac{M_{\rm bar}(r)}{M_{\rm DM,DMO}(r)}. \quad (13)$$

Equation (12) implicitly assumes that the uncontracted halo follows an NFW distribution, which is why we refer to this as the contracted NFW (CNFW) model. In principle, other 'contracted' models can be formulated equivalently by using another profile for the uncontracted halo model, provided it gives a good description of the halo's matter distribution in DMO simulations.

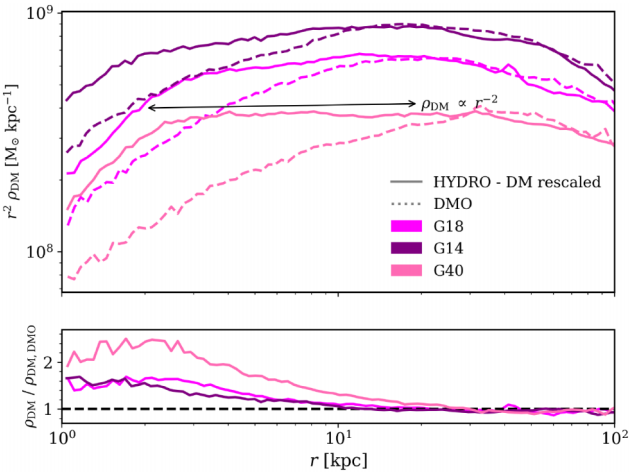



**Figure 2.** *Top panel:* Spherically averaged dark matter density profiles for three representative ARTEMIS central haloes (G40, G18, and G14). Solid lines show the hydrodynamical profiles after rescaling the enclosed dark matter mass by the cosmic mean baryon fraction, while dashed lines of the same colours denote their DMO counterparts. *Bottom panel:* Ratio in dark matter density profiles between the rescaled hydrodynamical and DMO versions of the same haloes. The dashed black line indicates no difference between the two. Colours in both panels distinguish the individual haloes, as indicated in the legend. The impact of baryons is most pronounced at radii $r \lesssim 20$ kpc, where the hydrodynamical profiles deviate significantly from their corresponding DMO counterparts.

## 3 THE EFFECTS OF BARYONIC CONDENSATION ON DARK MATTER HALOES

We begin our analysis by examining the impact of baryonic physics on the inner structure of DM haloes. Specifically, we compare the spherically averaged DM density profiles of central haloes between the hydrodynamical and DMO runs of the ARTEMIS simulations. Throughout this work, we adopt the commonly used assumption of spherical symmetry, which is a reasonable approximation in the inner regions ($r \lesssim 20$ kpc), where baryonic processes tend to make the halo potential more isotropic (e.g. C. S. Frenk et al. 1988; M. G. Abadi et al. 2010; A. Di Cintio et al. 2014b). This region is also where rotation curve measurements are most constraining and where different theoretical halo models deviate more strongly from one another.

Perhaps the most striking consequence of baryonic physics in MW-mass systems is the so-called contraction of the DM distribution, i.e. an enhancement in the central DM density compared to the DMO counterpart of a halo. This effect is illustrated in Fig. 2, which shows the DM density profiles for three representative ARTEMIS haloes (G40, G18 and G14). Colours indicate which halo; hydrodynamical and DMO runs are distinguished by the line styles, solid and dashed, respectively.

Baryonic condensation enhances the central DM density typically for $r \lesssim 20$ kpc. The magnitude and location of this enhancement varies between haloes, reflecting differences in their baryonic content and assembly histories, as shown more directly by the density ratios in the bottom panel of Fig. 2. For example, G40 (pink) exhibits the strongest response, reaching a maximum

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

enhancement of $\rho_{\rm DM}/\rho_{\rm DM,DMO} \approx 2.5$ at $r \approx 2$kpc, while the other haloes show more moderate peak increases of order $\sim$1.7 at their respective peak radii. This baryon-induced contraction steepens the profile slope and brings it closer to an isothermal value. This is particularly evident for G40 again, where the hydrodynamical profile maintains a near-constant slope of -2 over $r \approx 2$–20 kpc. By contrast, its DMO counterpart only briefly approaches the isothermal value near $r \approx 20$ kpc, never sustaining it over a comparable interval. At larger radii, the baryon-induced enhancement in DM density typically diminishes and may even reverse, with the hydrodynamical profiles falling below their DMO counterparts. This outer suppression is substantially weaker in amplitude than the central enhancement (of order a few percent rather than hundreds), but its magnitude also varies systematically between haloes, with more pronounced effects generally associated with differences in stellar mass. This behaviour, sometimes referred to as halo elongation, is visible for example in G18 and G40 beyond the slope transition scale. Around or beyond the virial radius ($r \sim R_{\rm 200c}$), however, the hydrodynamical and DMO profiles converge in both enclosed mass and outer slope, suggesting that baryonic effects become negligible at sufficiently large scales; see also Fig. 1.

These findings are in good agreement with results from other large hydrodynamical simulation projects–including IllustrisTNG, EAGLE, Auriga, and APOSTLE–which similarly show that MW-mass haloes are, on average, contracted (e.g. M. Schaller et al. 2015; R. J. J. Grand et al. 2017; A. Pillepich et al. 2018; M. Cautun et al. 2020). This contraction signature is not universal in amplitude or radial dependence, however: recent cross-suite comparisons (A. Hussein et al. 2025; A. M. Garcia et al. 2026) show that simulations such as Auriga, EAGLE and TNG where the ISM is modelled with an effective equation-of-state and the star formation density is relatively low are well described by simple adiabatic contraction (AC), whereas simulations which model strong, bursty feedback – such as FIRE-2 – show substantial differences, including a DM density enhancement stronger than what predicted by AC and, in some instances, even central core formation at MW masses (A. Lazar et al. 2020). Differences in subgrid physics and modelling choices are therefore known to affect the DM distribution in hydrodynamical realizations of DM haloes, in particular the modelling of bursty and stronger feedback in FIRE-like simulations is likely the source of disagreement on the signature of contraction. Given that ARTEMIS shares a lot of modelling details with the suites M. Cautun et al. (2020) calibrated against–especially EAGLE_recal – we expect close agreement in the signature of contraction we observe and their model; nonetheless, the recalibration of stellar feedback parameters, which ARTEMIS involved, makes the test which we present in the next section necessary to understand whether their contraction model can be applied to our haloes.

### 3.1 Modelling baryon-induced halo contraction

A number of models have sought to describe baryon-induced contraction using adiabatic formalisms, in which the DM distribution in hydrodynamical simulations is assumed to conserve the same action integrals as in its corresponding DMO counterpart (e.g. G. R. Blumenthal et al. 1986; O. Y. Gnedin et al. 2004; J. A. Sellwood & S. S. McGaugh 2005; T. M. Callingham et al. 2020; A. Hussein et al. 2025). These approaches provide physically motivated and interpretable descriptions of halo response from first principles. However, their implementation in analytical profile fitting can require additional assumptions about the baryonic component and halo evolution that are not always directly constrained in observational contexts.

Here, we adopt the semi-empirical model presented in C20 as a complementary approach. This model relates the enclosed DM and total mass profiles in hydrodynamical and DMO simulations through a deterministic power-law relation (see Section 2.3.4). By directly coupling the local DM response to the enclosed baryonic mass profile, $M_{\rm bar}(< r)$, this framework accommodates a wide range of baryonic structures without requiring a prescribed analytical form or detailed knowledge of global quantities, such as the baryon fraction at the virial radius, which are often uncertain observationally. Given the feedback-dependence noted above, we treat this prescription as a practical tool suited to simulations with smooth stellar feedback, rather than as a universally valid model of contraction, and test its applicability to ARTEMIS directly below.

In this work, we apply this contraction prescription to the ARTEMIS haloes to assess how accurately it reproduces the DM structure in the hydrodynamical simulations. Once this is established, our goal is not to adjudicate on the detailed physics of contraction, but rather to quantify how the inclusion or neglect of baryon-induced contraction impacts the inferred DM distribution when fitting halo profiles to rotation curve data. The semi-empirical formalism provides a practical and flexible tool for this purpose. While the formalism can be generalized to other halo models, we adopt an NFW profile for the uncontracted halo due to its simplicity (two parameters) and its established success in describing DMO halo structure. We also emphasize that, unlike C20 who developed this formalism to infer a mass model of the MW from observational data, our aim is to benchmark its performance against commonly used analytic halo profiles that do not include contraction.

The C20 power law – i.e. equation (11) – was calibrated using the Auriga, APOSTLE, and EAGLE_recal suites, all of which model stellar feedback via an effective equation-of-state approach producing smooth, non-bursty outflow histories. ARTEMIS adopts the EAGLE subgrid physics model (J. Schaye et al. 2015), with only the stellar feedback efficiency recalibrated relative to EAGLE_recal (A. S. Font et al. 2020), and so shares a similar feedback implementation. Close agreement with C20 is therefore an expected outcome of this shared physics rather than evidence that the model holds across galaxy formation implementations generally – a genuinely independent test would require comparisons with simulation suites with more significant differences in subgrid models. The comparison is nonetheless a necessary check: It confirms the prescription is internally consistent for the simulations analysed here, which is a prerequisite for using it as a quantitative tool in what follows.

To carry out this check, we compare the relationship between $\eta_{\rm DM}$ and $\chi_{\rm tot}$ for the ARTEMIS haloes against the original C20 power law. Fig. 3 shows this relation, with individual halo profiles represented by dark red solid lines.

Following individual lines that correspond to individual galaxies highlights a subtle but noteworthy feature of the $\eta_{\rm DM}$–$\chi_{\rm tot}$ relation: haloes do not scatter randomly around the best-fit trend, but instead trace coherent trajectories. Most haloes begin in good agreement with the power-law at $\chi_{\rm tot} \sim 1$ before systematically bending downward at smaller $\chi_{\rm tot}$ values, which are associated with the innermost regions. This behaviour suggests that the contraction model performs best when averaged over halo popula-

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

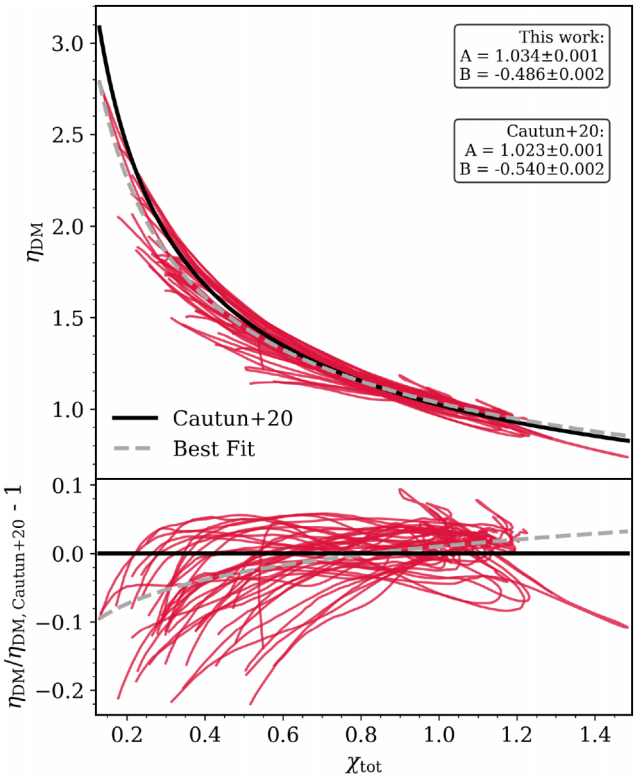



**Figure 3.** Relation between the enclosed dark matter mass ratio, $\eta_{DM}$, and the enclosed total mass ratio, $\chi_{tot}$, for all central ARTEMIS haloes. Ratios are computed from the spherically averaged cumulative mass profiles in the hydrodynamical and DMO versions of each halo. *Top panel:* $\eta_{DM}$–$\chi_{tot}$ relation for individual haloes (dark red curves). The solid black line shows the C20 best-fit power law, and the dashed grey line shows our best fit to the ARTEMIS sample. *Bottom panel:* Fractional deviation of each halo from the C20 prediction. The ARTEMIS haloes follow the same deterministic power-law relation to within a few per cent, with both median offset and sample scatter within the C20 model-quoted bounds, confirming that the contraction formalism remains internally consistent when applied to a simulation suite with a similar (EAGLE-based) feedback implementation.

tions, rather than when applied point-by-point within individual haloes.

Reassuringly, we find that ARTEMIS follows the same general trend as reported in the original study, with a best-fit power law that is remarkably close to the original: our inferred slope and normalization differ by less than 10 per cent from the C20 model. Furthermore, the bottom panel of Fig. 3, which quantifies the fractional deviation from their best fit, reveals that most haloes lie within the ±5 per cent band, and with the largest deviations in our sample remaining bounded by their quoted 20 per cent scatter envelope. The minimal systematic offset, and a halo-to-halo scatter within their quoted bounds, confirm that the power-law relation is internally consistent when applied to ARTEMIS – as expected given a similar feedback prescription – and validate its use as a quantitative tool for the rest of this work. It also strengthens the point that stellar feedback efficiency alone does not seem to have a first-order effect on the response of DM haloes to the assembly of MW-like galaxies in their interiors, given the absence of significant differences between ARTEMIS and EAGLE_recal.

## 4 THE ROLE OF CONTRACTION IN SPATIALLY CONSTRAINED DARK MATTER PROFILE FITS

Having shown that baryons significantly affect the internal structure of DM haloes, particularly for radii $\lesssim$20–30 kpc, a natural question is how well can global halo properties be recovered from observations of the mass profile over only a limited radial range, especially one that largely overlaps with these contracted regions. This is particularly relevant for rotation curve analyses of the MW and similar galaxies, which are most sensitive to the same regions where contraction is strongest. To examine the effects of restricted radial coverage, we fit halo models to the DM mass profiles over various radial ranges. While this approach does not try to reproduce a full mock observation pipeline, it does quantify how well the properties of the host DM halo can be measured assuming that the local DM potential can be accurately measured with minimal errors. In this sense, we are studying what is the fundamental limit to the information to be extracted when reconstructing the properties of the host DM halo using robust data but restricted to relatively small radii, and the potential systematics associated with this process.

We evaluate the performance of four widely used theoretical profiles – Einasto, NFW, gNFW, and CNFW (see Section 2.3) – to spatially restricted regions of our simulated haloes. Each fit region is defined by a fixed inner radius, $r_{\rm min.\,fit}$, and a varying outer radius, $r_{\rm max.\,fit}$, ranging from 10 per cent to 100 per cent of the median $R_{\rm 200c,DMO}$ in our sample (i.e. $\bar{R}_{\rm 200c,DMO} \approx 214$kpc). This strategy enables us to quantify how much radial coverage is required to reliably constrain bulk properties of the halo such as mass, $M_{\rm 200c}$, and concentration, $c$.

### 4.1 Fitting range sensitivity and profile behaviour

We begin by examining how well each profile captures the structure of DM haloes within the fitted radial domain. This step also serves as a sanity check: A model that fails to reproduce the halo structure within the fitted region is unlikely to yield meaningful extrapolation to larger radii. Fig. 4 presents a detailed view of profile behaviour across our model suite. Each panel corresponds to a different model – NFW, CNFW, gNFW, and Einasto – and compares the fits obtained by varying the outer fitting radius while keeping the inner radius fixed.

The top panel in each quadrant shows the DM circular velocity profile, $V_{\rm c}(r)$, of a typical DM-rescaled hydrodynamical halo (solid black) and its DMO counterpart (dashed black). We specifically show halo G2, with other haloes exhibiting qualitatively similar behaviour. Overplotted are best-fit parametric models constrained within three radial intervals: 3–21 kpc (blue), 3–64 kpc (orange), and 3–214 kpc (green). The innermost interval (3–21 kpc) corresponds to a typical range used in stellar rotation curve analysis. The vertical dash–dotted line marks the common inner boundary.

When comparing the blue (3–21 kpc fits) and green (3–214 kpc fits) curves, we notice a broadly consistent behaviour in all models that do not explicitly parametrize the effect of baryon-induced contraction (i.e. NFW, gNFW, and Einasto). They all tend to shift their best-fit parameters as more data is included – reflecting the difficulty for these functional forms to model an extended radial range where the density is isothermal, and an outer halo largely untouched by baryonic processes (DMO-like). The CNFW model is less sensitive to this effect, as its formulation explicitly accounts for the role of baryons in shaping the local

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

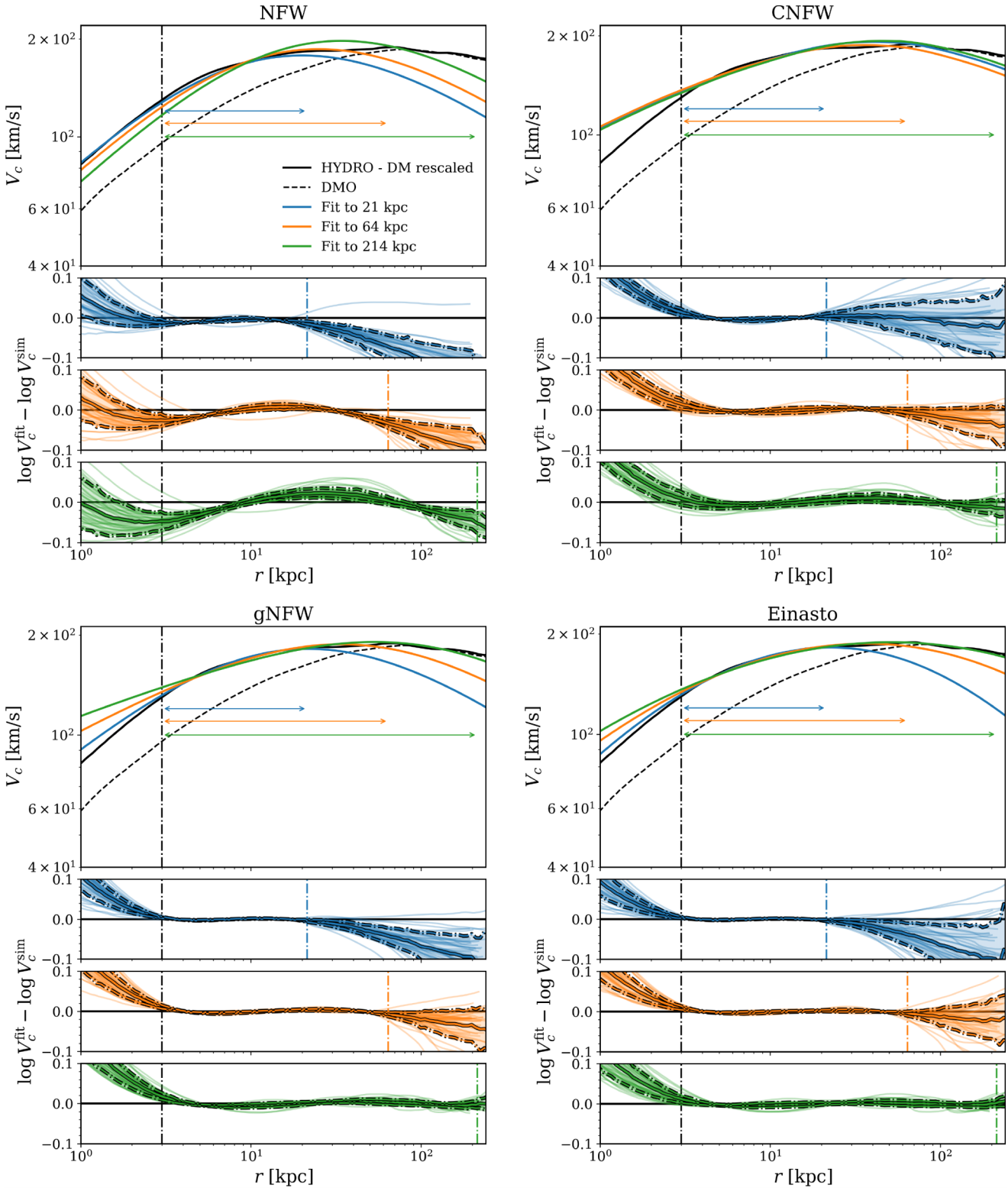



**Figure 4.** Comparison of NFW, CNFW, gNFW, and Einasto fits when constrained to truncated radial ranges of simulated ARTEMIS haloes. Fits are applied to the DM circular velocity profiles of the DM-rescaled hydrodynamical runs. *Top row of each panel:* Circular velocity profiles, $V_c(r)$, of a representative halo (solid black) and its DMO analogue (dashed black), together with best fits obtained when the outer fitting radius is limited to 21 kpc (blue), 64 kpc (orange), and 214 kpc (green). Coloured arrows indicate the radial interval used for each fit, following the same colour scheme. *Bottom rows of each panel:* Logarithmic deviations (in dex) between the fitted, $V_c^{\rm fit}$, and simulated hydrodynamical, $V_c^{\rm sim}$, profiles. Thin curves show individual haloes, the thick curve the sample mean, and shaded regions the $1\sigma$ halo-to-halo scatter. The solid black line denotes zero deviation. In each sub-panel, the coloured vertical dash–dotted line marks the corresponding outer fitting radius. In all panels, the vertical dash–dotted line marks the fixed inner fitting radius (i.e. 3 kpc). Across all profile families, restricted radial coverage increases systematic deviations in the extrapolated profile, with the magnitude and character of the bias strongly dependent on the chosen functional form.

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

DM distribution. This results in less variance in halo parameters across different radial coverages, or, equivalently, stable fits. The behaviour just described is particularly noticeable in the example NFW profile (top left panel), where the 3–21 kpc fit (blue) gives a circular velocity peaking at $r \approx 10$ kpc, while in the 3–214 kpc fit (green) it flattens at $r \approx 30$ kpc – i.e. roughly a factor of 3 larger $r_{\rm peak}$.

The bottom three panels in each quadrant of Fig. 4 show the relative deviation between the best-fit and simulated circular velocity profiles for all haloes in our sample. We quantify this deviation as the difference in their logarithms, measured in dex – i.e. $\log V_c^{\rm fit} - \log V_c^{\rm sim}$. The colours indicate the fitted radial range (as in the top panels), and each fitting range is displayed in a separate subplot. Thin lines represent individual haloes, while thick lines show the sample mean. Shaded regions denote the $1\sigma$ scatter.

The NFW fits reveal several clear trends. When only the inner galactic region is constrained, extrapolated circular velocities are systematically biased low – by ≈0.10 dex (26 per cent) at 100 kpc and ≈0.17dex (48 per cent) near the virial radius – despite minimal residual amplitudes and structure over the fitted range. This demonstrates that high formal fit quality within a limited radial interval (small compared to the virial radius) should not be mistaken for accurate inference of global halo properties. As seen in the bottom two panels of the first quadrant, widening the fitted radial range reduces the median offset at the virial radius, though the halo-to-halo scatter decreases more modestly. Notably, improving long-range accuracy comes at the cost of increased residual structure and larger scatter within the constrained domain, which is reflected in the enhanced wave-like patterns of the orange and green curves.

Interestingly, the more flexible three-parameter models – gNFW and Einasto – show similar qualitative behaviour. When fits are restricted to the inner galaxy (3–21 kpc, blue curves), they also yield substantially biased estimates of the circular velocity near the virial radius. Although the fits within the probed region are formally excellent, with absolute residuals never exceeding 0.005 dex (1 per cent) and virtually no scatter, the extrapolated circular velocity at the virial radius is still biased by $\approx -0.11$ dex ($-22$ per cent) on average. As before, the scatter increases within the fitted domain and decreases in the extrapolation region with an increase in radial coverage, though sample variance is comparatively suppressed for the three-parameter models. Furthermore, while NFW fits never recover the circular velocity at the virial radius perfectly [median offset $\approx -0.05$ dex ($-10$ per cent) even under the most favourable conditions], the gNFW and Einasto profiles can achieve negligible median bias – but this requires unrealistically large radial ranges, spanning essentially the entire virialized region, to constrain their parameters.

As briefly mentioned when discussing the fit results from the representative halo shown in the top panel of each quadrant, but shown clearly for the entire ARTEMIS sample in the bottom three panels, the CNFW model yields broadly consistent fits almost independently of the radial coverage. When extrapolating fits with a range of order tens of kiloparsecs, the median offset at the virial radius is $\approx -0.03$ dex ($-7$ per cent), and $\approx -0.02$ dex ($-5$ per cent) with a range of order hundred kiloparsecs. As for all models discussed, increasing the radial coverage leads to increased sample variance in the fitted domain and decreased variance in the extrapolated domain – in this sense CNFW sits somewhere between the NFW and the three-parameter models (gNFW and Einasto).

Interpreting these tendencies requires both mathematical and physical considerations. Mathematically, it is natural that parametric profiles adjust their degrees of freedom to improve the likelihood, resulting in formally higher quality fits within the probed radial domain. Physically, the preferred direction in parameter space is dictated by baryonic effects. As discussed in Section 3, a prominent consequence of baryon-induced contraction is the emergence of an extended radial interval in which the density slope is very close to $-2$, forming a broad isothermal shell. As shown in Fig. 2, this region can span tens of kiloparsecs, typically bounded within 2–20 kpc, and features an enhanced density compared to a DMO realization. Models that do not explicitly encode contraction physics can mimic this feature, but only by distorting the overall halo structure. We note that the specific radial structure underlying the arguments that we present – a broad, near-isothermal shell produced by smooth, centrally concentrated contraction – is itself a property of the ARTEMIS/EAGLE feedback model at MW-mass regimes, rather than a generic feature of baryonic physics in DM haloes. Simulations such as FIRE-2 with burstier star formation can show a qualitatively different radius-dependent response, including local core formation at the very innermost regions rather than monotonic contraction (A. Lazar et al. 2020; A. Hussein et al. 2025). The specific degeneracies we identify below, particularly the interplay between the scale radius and the width of the isothermal shell, are therefore not guaranteed to generalize to simulations with different feedback implementations, and may be weaker or absent where no comparably broad isothermal region is produced. We would nonetheless expect the qualitative conclusion – that fits restricted to the inner, baryon-affected region distort global parameters – to persist in some form: even in bursty-feedback simulations, baryonic effects on the DM profile remain of similar magnitude and strongest near the centre and dilute outward, so profiles lacking an explicit contraction term would still have to compensate for this central enhancement at the expense of the fitted large-scale shape.

The NFW profile, being the least flexible among the models we consider, is most strongly affected. With no parameter governing the curvature of the profile, NFW must compromise both the scale radius ($r_s = r_{-2}$) and normalization ($\rho_0$) to absorb the influence of the extended isothermal region. When the fitted radial range is small (3–21 kpc), sampling falls entirely within the degenerate isothermal region. To reproduce the observed density profile, NFW compresses $r_s$ inward towards $r \approx 2$ kpc (relative to the typical DMO value of ∼20 kpc), producing artificially high concentrations. The extrapolation to large radii with these distorted parameters then systematically underestimates circular velocities at the virial radius. As the radial coverage increases, the inferred $r_s$ shifts outward and partially approaches the true (DMO) value, reflecting the higher statistical weight of outer regions, which are weakly affected by baryonic effects. However, because NFW profiles cannot accommodate the broadness of the isothermal shell, the fitted haloes remain significantly overconcentrated and under-massive even with extensive radial information.

The more flexible gNFW and Einasto models achieve even better local fits by adjusting their curvature parameter ($\gamma$ and $\alpha$, respectively), but this extra flexibility simply allows them to overfit the isothermal shell's appearance in a narrow radial slice (compared to the virial radius). In both cases, parameters optimized for local residual minimization carry little information about structure beyond the fitted domain, as the structure of the

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

halo past 20–30 kpc is largely uncorrelated with the isothermal shell. This results in behaviours similar to the NFW model, i.e. systematic biases in extrapolated circular velocities, as they all arise from the same tendency to optimize for a region that carries little information about the outer halo. In fact, only when the fitting range extends across nearly the full virialized region are the models' degrees of freedom sufficiently constrained to recover the true global structure; before then, strong degeneracies between the scale radius and inner slope parameters appear when examining the covariance matrices for gNFW fits (see Section 4.2 for a more detailed discussion).

The CNFW model, on the other hand, incorporates the expected change to the DM structure explicitly due to baryon-induced contraction. Essentially, the model knows where to locate the isothermal shell given the observed baryonic mass profile $M_{\rm bar}(< r)$, thereby preventing the fitted parameters (e.g. $\rho_0$ and $r_s$) from compensating for baryon-induced contraction by distorting the global halo structure.

Breaking the degeneracy between fitted parameters and the physics of contraction results in stable fits across a wide range of radial scales for this model. This is evident in the representative halo shown, where the blue and green curves are nearly indistinguishable – note that the former is constrained by an outer radius 1 dex smaller than the latter.

These trends are not driven by numerical resolution or sampling choices; varying the number of radial bins in the analysis has a negligible effect on our results. Instead, the biases arise from the limited radial leverage of the fit, with profile parameters compensating for the locally enhanced density when baryonic effects are not modelled explicitly.

Degeneracies between $r_s$ and $\gamma$ for the gNFW model have already been highlighted when fitting the MW's rotation curve (e.g. E. V. Karukes et al. 2019; X. Ou et al. 2024), indicating that the estimates of the Galactic DM halo structure might be affected by these systemic biases. Overall, our results demonstrate that high-fidelity fits – i.e. formally small residuals – within a limited radial range, particularly one that overlaps with the contracted region of a halo, do not necessarily guarantee an accurate recovery of global halo properties. While all models perform well where they are expected to – in the regime where data are available – their ability to correctly extrapolate beyond this region varies significantly, as we explore in greater detail in the next section.

### 4.2 Model-dependent bias in extrapolated halo properties

Having established that DM profile fits constrained to limited radial domains can, in some cases, exhibit significant residual structure in the extrapolated regions and signs of overfitting, we now focus explicitly on the impact of these effects on key structural properties. Specifically, we examine how the inferred values of halo mass, $M_{200c}$, and concentration, $c$, vary with the choice of profile model and the radial extent used for its fit.

Fig. 5 presents a systematic evaluation of these estimates across the ARTEMIS halo sample, comparing best-fit values from profiles truncated at varying outer radii. Here different columns correspond to different models (NFW, CNFW, gNFW, and Einasto; from left to right), and rows provide summary diagnostics for halo mass and concentration. Coloured thin lines indicate tracks from individual haloes, while thick lines and shaded regions indicate the sample median and 68th percentiles (or sample scatter); the grey shaded band represents the realistic radial range currently probed by kinematic tracers in the Galaxy. Specifically, we quantify the bias in halo mass via the deviation between the logarithms of the inferred and values computed via the DM-rescaled $M_{200c}$ convention (see Section 2.2.1) – i.e. $\log M^{\rm fit}_{200c} - \log M^{\rm sim}_{200c}$. The bias in the inferred concentration is measured relative to the best-fit value when the full virialized region is used to constrain the parameters of the DMO counterpart of a halo – i.e. $\log c^{\rm fit} - \log c^{\rm fullfit}_{\rm DMO}$.

The NFW fits (leftmost column, light blue curves) lead to extrapolated values of $M_{200c}$ systematically underestimating the true halo mass, with the 68th percentile of the population never intersecting the zero-deviation line across all choices of $r_{\rm maxfit}$. Even when fitting out to $r_{\rm maxfit} \approx \bar{R}_{200c,\rm DMO}$, the upper bound still falls short by $\approx$0.05 dex (11 per cent); conversely, the inferred concentrations are biased high – with an offset of $\approx$0.12 dex (32 per cent) for the lower 16th percentile, as a direct consequence of the fitted scale radius, $r_s$, shifting inward to accommodate for the contraction of the halo. The halo-to-halo scatter in extrapolated mass decreases from a value of $\sigma_{\rm M200} \approx 0.15$ dex (41 per cent) for fits limited to 21 kpc to $\sigma_{\rm M200} \approx 0.07$ dex (17 per cent) when limited to 214kpc – roughly a factor of two over one decade in radius. The scatter in concentration estimates similarly declines by a factor of two at the typical virial radius from $\sigma_c \approx 0.13$ dex (35 per cent) at $r_{\rm maxfit} = 21$kpc. From an observational perspective, these biases are particularly concerning. When only data within the inner 30kpc are available (denoted by the vertical shaded grey band), a regime common in Galactic and extragalactic rotation curve studies, the median extrapolated mass is underestimated and concentration is overestimated by a factor of two (i.e. −0.3dex and +0.3 dex, respectively). Moreover, in our most extreme case the halo mass is underestimated and concentration overestimated by nearly a factor of five. These results imply that NFW fits in the MW-mass regime are unreliable for precision estimates of $M_{200c}$ and $c$, especially when extrapolating from parameters constrained using the inner, contracted galactic regions.

Looking at the inferred halo mass when assuming a gNFW or Einasto profile (top row, last two columns; dark green and orange curves, respectively), they start with large yet comparatively suppressed underestimates of the halo mass, with median offsets of −0.21 dex (−38 per cent) when the inner 21 kpc are probed, and reach negligible deviations when the maximum fitting radius is pushed to $\log r_{\rm maxfit}/{\rm kpc} \gtrsim 2$. Their halo-to-halo scatter differs slightly especially for $\log r_{\rm maxfit}/{\rm kpc} \lesssim 2$, as the sample spread in gNFW fits is of $\sigma_{\rm M200} \approx 0.13$ dex (35 per cent), compared to $\sigma_{\rm M200} \approx$ 0.17dex (48 per cent) in Einasto fits. This causes the 86th percentile in Einasto fits to cross the zero-deviation line at $\log r/{\rm kpc} \approx 1.6$, roughly 0.2dex earlier than gNFW fits. These models are also qualitatively similar in the recovered concentrations (bottom row, last two columns). Einasto fits restricted to 21 kpc exhibit a median offset and halo-to-halo scatter in recovered halo concentration of +0.26 dex (82 per cent) and 0.15 dex (41 per cent), respectively, while gNFW fits exhibit a median offset of +0.29 dex (95 per cent) and halo-to-halo scatter of 0.12 dex (32 per cent). Even when the fitting range is increased to the full halo (i.e. 3–214 kpc), there persist significant biases in median offsets and sample spread in estimated concentrations; quantitatively, gNFW fits show a median offset of +0.12 dex (32 per cent) and halo-to-halo scatter of 0.07 dex (17 per cent), and Einasto fits median offset of +0.20 dex (58 per cent) and halo-to-halo scatter of 0.09 dex (23 per cent).

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

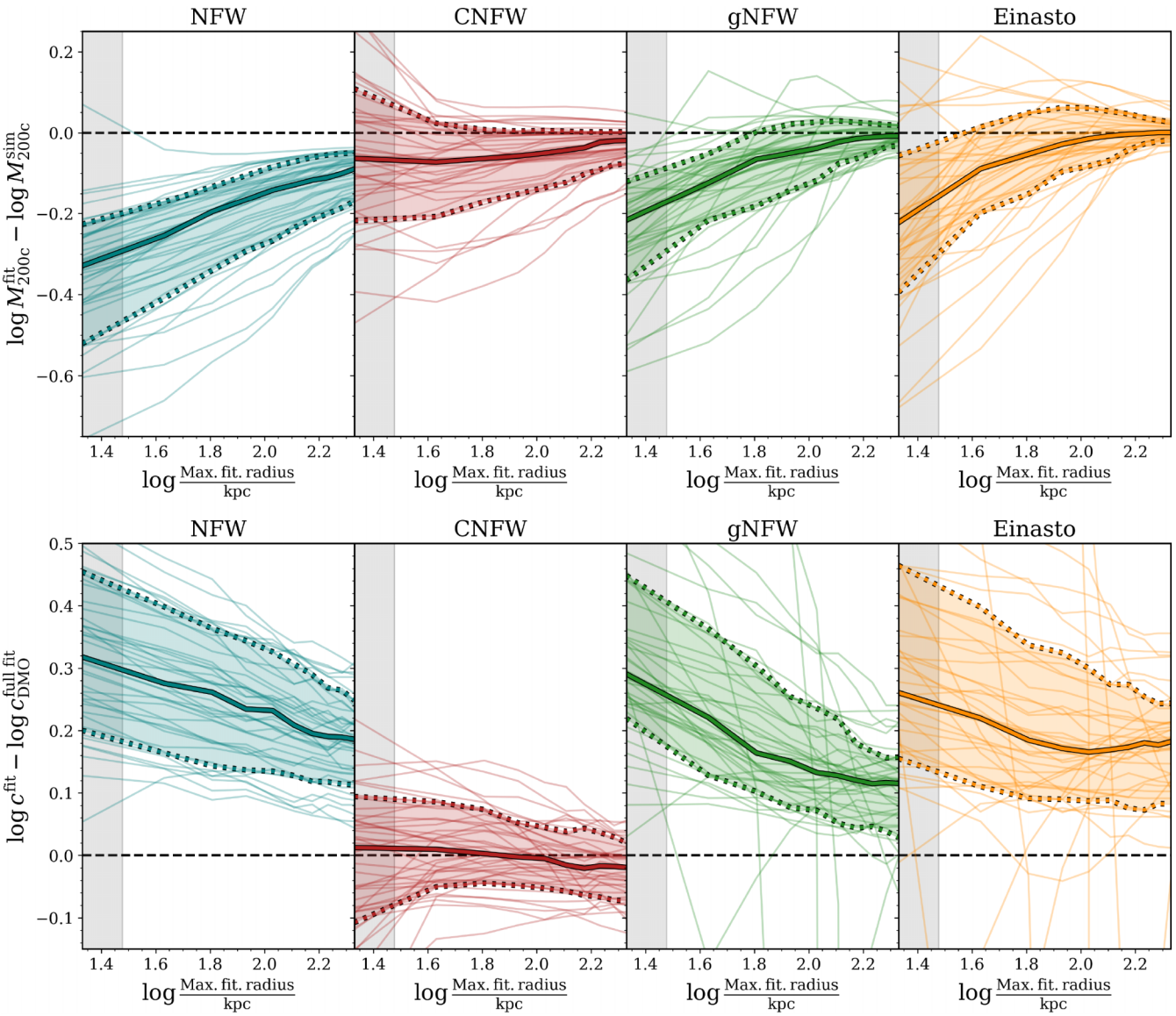



**Figure 5.** Model-dependent biases in extrapolated halo properties as a function of the maximum fitted radius, $r_{\rm max,fit}$. For each ARTEMIS halo, we plot the logarithmic offset between the fitted parameter and its true/reference value. *Top panel:* Logarithmic offset between virial masses – $\log M_{200c}^{\rm fit} - \log M_{200c}^{\rm sim}$ – against the logarithm of the maximum fitting radius. *Bottom panel:* Same as the top panel but focusing on the concentration parameter. Hence showing deviations in the recovered concentration relative to that of the corresponding DMO halo – $\log c^{\rm fit} - \log c_{\rm DMO}^{\rm fullfit}$. Thin coloured curves show individual haloes, thick curves the sample median, and shaded regions the 16th–84th percentile range. Columns, as well as colours, correspond to the four fitted models (NFW, CNFW, gNFW, and Einasto, from left to right). In both panels, the grey shaded band indicates the typical observational range of the outermost radius probed by the Milky Way rotation curve. The figure demonstrates that restricting the fitted domain to small radii induces systematic, model-dependent biases in both halo mass and concentration, despite excellent local fits within the constrained region.

Despite an intuitive explanation for the trends exhibited by the vast majority of the sample, a small number of haloes show erratic evolution of their extrapolated parameters, particularly for the gNFW and Einasto models (last two columns of Fig. 5). For this reason, we quote the 16th–84th percentile range rather than the standard deviation throughout, as the latter is disproportionately sensitive to these non-Gaussian tails and would misrepresent typical model performance. We find these outliers to be associated with significant substructure near the virial radius (e.g. a massive satellite or close companion), which distorts the spherically averaged circular velocity profile in a way that cannot be captured by smooth parametric forms; the additional flexibility of the gNFW and Einasto models allows them to chase these localized features via extreme excursions in their shape parameters, at the cost of poorly constrained concentration estimates. We examine a representative case in detail in Appendix A.

Strikingly, the CNFW model largely mitigates the systematic biases seen in the rest of our model suite, even when the fit is restricted to the inner region ($\lesssim$20 kpc). Here, the median offset in estimated $M_{200c}$ is of $-0.06$ dex ($-13$ per cent) and the halo-to-halo scatter is of 0.15 dex (41 per cent), meaning a non-negligible portion of the sample returns an estimate within $\pm$25 per cent of the true value. However, estimates do not converge to the zero-deviation line even when the entire virialized

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8559319 by guest on 04 November 2026

halo is used for their fits, retaining a median offset and scatter of 0.02 dex (4.7 per cent) and 0.03 dex (7 per cent), respectively. Similar to the Einasto fits, the median concentration estimate does not evolve as the radial range increases, but the halo-to-halo scatter does modestly decrease with larger radial coverage. Offsets are always of order $\pm 0.02$ dex ($< 5$ per cent), the sample scatter evolves from $\sigma_c \approx 0.10$ dex (26 per cent) for $r_{\rm maxfit} = 21$kpc to 0.05 dex (12 per cent) at the maximal range. We also note that this estimate should not be interpreted as the concentration of the contracted halo. By explicitly accounting for baryon-induced contraction, the CNFW model is able to robustly infer concentration values that trace the intrinsic structural parameters of the halo, rather than those imprinted by baryonic condensation.

### 4.3 Implications for DM inference in MW-mass systems

The results presented above highlight several points of practical relevance for both modelling strategies and the interpretation of MW mass estimates.

First, our analysis shows that the CNFW model provides the least biased inferences of global halo properties, $M_{200c}$ and $c$, when the available kinematic information is restricted to the radial ranges where baryon-induced contraction dominates. In the MW mass regime, this is precisely where most observational tracers reside (e.g. *APOGEE*, *Gaia*; see the discussions in A.-C. Eilers et al. 2019; X. Ou et al. 2024). Although more flexible three-parameter models (Einasto, gNFW) yield excellent local fits with smaller formal residuals, their ability to capture the correct global structure is compromised unless the fitted domain extends to large radii.

Expanding on this point, we find that fits restricted to $50 \lesssim r/\mathrm{kpc} \lesssim 214$ remove the significant systematic bias in halo mass: for all profiles the median offset is less than 1 per cent, with sample scatter of ≈10 per cent. This is not simply an effect of including the virial radius, as fits over the narrower 50–70 kpc range exhibit nearly identical behaviour, with only slightly increased halo-to-halo scatter. Inferred concentrations, on the other hand, become largely unconstrained in these outer-halo fits, with any bias being small compared to the sample scatter (never smaller than ∼100 per cent for any profile); this reflects the fact that both radial ranges exclude the typical scale radius ($r_s \sim 10$–$20$ kpc for MW-mass systems in our sample). These results reinforce our interpretation that the biases in halo mass arise primarily from the inner halo, where contraction alters the DM density profile: Models adjust parameters to accommodate the near-isothermal shell (i.e. $\rho \propto r^{-2}$), whereas the largely DMO-like outer halo provides unbiased constraints when sampled. They also suggest that, ignoring other potential observation systematics, kinematic probes residing in the outer halo should be considered of higher fidelity when attempting a precision estimate of the MW's halo mass.

Secondly, we find that internal diagnostics of fit quality are generally poor indicators of whether a given halo will yield unbiased extrapolations. Monte Carlo sampling of each best-fit parameter and covariance matrix (100 realizations per halo, per profile, per $r_{\rm maxfit}$) confirms that the intra-halo scatter is consistently one to two orders of magnitude smaller than the halo-to-halo scatter. Consequently, well-behaved fits – i.e. small residuals, smooth profiles, stable parameter posteriors – do not distinguish 'good' from 'bad' when extrapolating to larger radii.

Thirdly, while the mean bias can be negligible, there remains significant scatter between systems for all fitted profiles when using limited radii (typically ≈40–50 per cent). In particular, internal goodness-of-fit metrics and the enclosed baryonic mass profile alone carry little information about the reliability of the extrapolation beyond the fitted radial range. The population scatter therefore represents a hard lower bound on the uncertainty in extrapolated halo properties – even under the idealized assumptions adopted here, i.e. full knowledge of the baryonic distribution, negligible observational errors, and a perfectly reconstructed DM mass profile. In real applications, the error budget can only be larger; hence, for fitting pipelines bounded by $r_{\rm max\,fit} \approx 20$–$30$ kpc, we regard halo mass estimates with quoted uncertainties below 0.20 dex (58 per cent) as likely overoptimistic in the MW-mass regime.

Until recently, there was a broad consensus for the mass of the MW (≈$10^{12}\mathrm{M}_\odot$), with many different tracers and methods agreeing within their quoted uncertainties (W. Wang et al. 2020). However, with an increase in the quality and quantity of data measurements using the inner Galaxy, researchers have begun to systematically infer lower halo masses, with significantly narrower quoted errors. A.-C. Eilers et al. (2019) inferred a mass of $M_{200c} = (7.25 \pm 0.25) \times 10^{11}\,\mathrm{M}_\odot$ assuming and NFW DM profile. While the recent work of X. Ou et al. (2024) inferred a mass of $M_{200c} = (6.94 \pm 0.25) \times 10^{11}\,\mathrm{M}_\odot$ or $M_{200c} = (1.81 \pm 0.25) \times 10^{11}\,\mathrm{M}_\odot$ when assuming a gNFW or Einasto DM profile, respectively. Our results show that when using profiles that do not account for halo contraction (i.e. NFW, gNFW, and Einasto), it is common to underestimate the halo mass by a factor of ≈2, and in the most extreme haloes 4. This alone is enough to significantly reduce, and in most cases remove completely, the tension between these recent constraints on the MW's virial mass and the traditional ≈$10^{12}\mathrm{M}_\odot$, without even considering other systematic errors.

A more recent revision by X. Ou et al. (2025) is an interesting external check in this context. Their updated NFW and gNFW fits yield MW masses of ≈7–8 × $10^{11}\mathrm{M}_\odot$, closer to the canonical ≈$10^{12}\mathrm{M}_\odot$ value than their Einasto fit (≈$10^{11}\mathrm{M}_\odot$), and both are also in better agreement with independent outer-halo probes (see their fig. 10). As noted above, fits restricted to the outer halo show no systematic tendency to bias the recovered mass, since this region lies beyond the reach of contraction, making agreement with outer probes a stronger indicator of a reliable extrapolation than internal goodness-of-fit metrics, which carry little information about extrapolation accuracy (Section 4.1). Within our sample, Einasto fits remain less biased than NFW on average, and recover systematically higher $M_{200c}$, for essentially any choice of maximum fitting radius. However, the halo-to-halo scatter in this bias is substantial: at a fitting radius comparable to that adopted by X. Ou et al. (2025) (∼30 kpc), haloes for which an NFW fit is less biased and extrapolates more reliably than an Einasto fit lie well within the $1\sigma$–$2\sigma$ range of the population. The fact that NFW and gNFW outperform Einasto for the MW in X. Ou et al. (2025) is therefore consistent with our results: It reflects a plausible individual realization of the halo-to-halo scatter we find, rather than a departure from the population-level trend.

We additionally examined the impact of the radial fitting range on the inferred DM density at the Solar radius, $\rho_{\rm DM}(r_\odot)$ where $r_\odot = 8$ kpc. We find that all models except the standard NFW yield highly precise estimates: median deviations are $\lesssim$0.01dex(2 per cent), and the halo-to-halo scatter never exceeds 0.02 dex (5 per cent). Only the NFW profile exhibits a noticeable bias, typically inferring densities too high by 5 per cent (for 3–

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

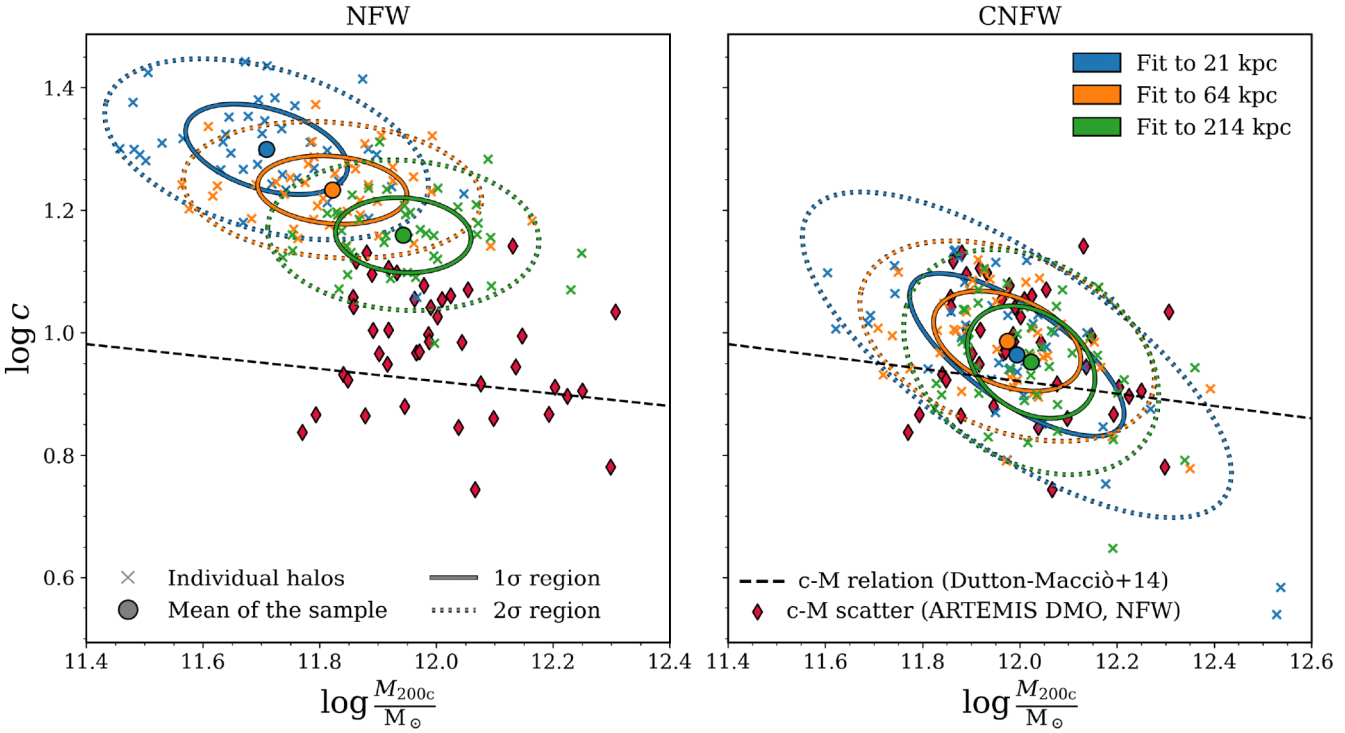



**Figure 6.** Recovered concentration–mass relations extrapolating from NFW (left panel) and contracted NFW (right panel) fits. Coloured crosses show individual haloes, with colours indicating the maximum fitted radius. Large filled circles give sample means for each fitting range, and solid/dashed ellipses indicate the corresponding $1\sigma/2\sigma$ covariance contours. The black dashed line marks the $c$–$M_{200c}$ relation from A. A. Dutton & A. V. Macciò (2014), shown as a reference. Dark red diamonds show the intrinsic $c$–$M_{200c}$ distribution of the ARTEMIS DMO haloes, measured directly from the full profiles, and therefore provide the appropriate simulation-specific benchmark for comparison. The figure illustrates how, even with limited radial coverage, modelling baryon-induced contraction helps in recovering the correct underlying concentration–mass relation.

21 kpc fits) and up to 20 per cent for larger fitting ranges. These results indicate that while global properties can be strongly biased when using restricted radial ranges, the local DM density at the Solar radius is remarkably robust to these modelling choices.

Taken together, these findings reinforce the view that halo contraction is a critical ingredient in dynamical modelling at MW masses. They also provide a quantitative pathway for re-interpreting rotation-curve-based MW mass estimates in a manner consistent with the behaviour of realistic hydrodynamical galaxy formation simulations. Crucially, the biases we identify are not random; they follow systematic trends that depend on how contraction steepens the inner density profile and on how much of the halo is radially sampled. This implies that any inference relying solely on inner-halo kinematics necessarily encodes information about contraction, even if the model used does not. This perspective naturally motivates a broader question: to what extent do the biases introduced by different modelling choices distort the recovered concentration–mass relation itself? The next section examines this issue directly by comparing how NFW and CNFW fits populate the $c$–$M_{200c}$ plane as a function of radial coverage.

### 4.4 The impact of modelling contraction on the inferred $c$–$M_{200c}$ relation

The concentration–mass relation is a cornerstone of ΛCDM structure formation, encoding the connection between halo mass, assembly history, and internal structure (e.g. J. F. Navarro et al. 1997; J. S. Bullock et al. 2001; A. A. Dutton & A. V. Macciò 2014; A. D. Ludlow et al. 2016). In DMO simulations, this relation follows a tight, monotonic trend, with more massive haloes being less concentrated. In hydrodynamical simulations, however, baryonic processes – including halo contraction, star formation, and energetic feedback – can alter the inner density profiles and shift haloes away from the DMO prediction (e.g. O. Y. Gnedin et al. 2004; A. Di Cintio et al. 2014b; M. Schaller et al. 2015). Whether the canonical DMO relation remains a reliable benchmark for real galaxies, particularly at MW masses, is therefore an important open question.

Earlier we showed that virial masses and concentrations inferred from NFW fits can be significantly biased when baryon-induced contraction is not modelled, with the magnitude and direction of the bias depending strongly on radial coverage. It is therefore natural to ask whether such mis-modelling affects the population-level $c$–$M_{200c}$ relation itself, and whether a model that incorporates contraction (CNFW) can recover the expected scaling even when only the inner ∼20 kpc are available. Fig. 6 examines this by comparing how NFW and CNFW fits populate the $c$–$M_{200c}$ plane for different choices of maximum fitted radius (left and right panels, respectively). Coloured crosses show individual haloes, with colours indicating the maximum fitted radius – blue for 3–21 kpc fits, orange for 3–64 kpc fits, and green for 3–214 kpc fits (as in Fig. 4). Large filled circles give sample means for each fitting range, solid/dashed ellipses indicate the corresponding $1\sigma/2\sigma$ contours, and the dashed black line represents the theoretical $c$–$M_{200c}$ relation from A. A. Dutton & A. V. Macciò (2014), also referred to in the figure as Dutton-Macciò+14. We emphasize that the ARTEMIS DMO haloes themselves provide

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

the most direct benchmark for assessing the accuracy of the recovered $c$–$M_{200c}$ relation. These are shown in Fig. 6 as dark red diamonds, constructed from fits to the full radial extent of the corresponding DMO simulations. The A. A. Dutton & A. V. Macciò (2014) relation is included for reference only, to provide a commonly used external comparison. In practice and as shown in the figure, ARTEMIS haloes scatter around this relation with a small systematic offset, consistent with differences in cosmology and halo selection, and alternative calibrations (e.g. a concentration–mass relation from A. D. Ludlow et al. 2016) would shift the line vertically by only $\sim$0.02 dex in this mass range, well below the effects discussed here.

As shown in the left panel of Fig. 6, standard NFW fits systematically overestimate concentrations while underestimating halo masses for all choices of radial coverage. The discrepancy is largest when the fit is restricted to $r \lesssim 20$ kpc, where haloes lie well away from the DMO relation by more than $2\sigma$, even though the individual mass and concentration offsets fall within this range (see Fig. 5). Crucially, the joint bias does not move haloes parallel to the standard $c$–$M_{200c}$ relation. Instead, the correlated increase in concentration and decrease in mass shifts haloes roughly perpendicular to the expected trend, displacing them systematically off the canonical scaling. In other words, the biases reinforce one another in a direction that increases the apparent tension with the DMO prediction. This behaviour persists even for full-halo NFW fits, indicating that (i) the NFW functional form cannot reproduce the contracted profile, and (ii) the resulting parameter covariance drives coherent departures from the expected relation even when information well beyond the contracted region is included.

In contrast, CNFW fits (right panel of Fig. 6) remain closely aligned with the DMO $c$–$M_{200c}$ relation for all radial ranges, avoiding the systematic displacement seen for NFW fits. Even when restricted to 21 kpc, the sample mean lies on the expected scaling and the scatter remains small – i.e. the median offset and $1\sigma$ scatter mostly land on the dashed black line. Extending the fits to larger radii reduces the halo-to-halo scatter primarily along the mass axis, reflecting the increased radial leverage, but does not alter the central trend or the covariance orientation. An equivalent analysis for the DMO counterparts shows that mild systematic offsets persist for NFW fits even with full coverage and are qualitatively consistent with what is shown in the right panel of Fig. 6. These residual offsets are consistent with the intrinsic scatter and mild systematic offset of the ARTEMIS DMO haloes relative to the A. A. Dutton & A. V. Macciò (2014) relation (see dark red diamonds in Fig. 6), rather than indicating a shortcoming of the contraction model itself. Crucially, independent of the radial range used for the profile fitting, the posterior distribution of CNFW fits closely follows the intrinsic ARTEMIS DMO scatter in the $c$–$M_{200c}$ plane, which constitutes the appropriate simulation-based benchmark. In contrast, NFW posteriors remain largely off the concentration–mass relation, whether compared to the A. A. Dutton & A. V. Macciò (2014) relation or to the DMO haloes themselves.

It is worth clarifying the conceptual status of this comparison, since DM haloes in hydrodynamical simulations are not, in general, expected to reproduce the $c$–$M_{200c}$ relation of DMO simulations. Baryonic condensation, cooling and feedback introduce non-gravitational physics with no counterpart in a DMO run, so there is no first-principles reason for a hydrodynamical halo's DM distribution to sit on, or near, the same locus in the $c$–$M_{200c}$ plane as its DMO analogue. The relevance of the contraction formalism is not that it enforces such an expectation, but that it provides a physical route between the two: Contraction arguments, whether adiabatic (e.g. O. Y. Gnedin et al. 2004) or semi-empirical (e.g. M. Cautun et al. 2020), relate the contracted profile to an assumed uncontracted, DMO-like progenitor through the (observable) enclosed baryon mass, via the local conservation of orbit-averaged action-like quantities or deterministic relations between baryon mass and total mass. The CNFW model implements this explicitly, parametrising the departure from an uncontracted NFW profile as a function of the enclosed baryon mass fraction, rather than assuming the hydrodynamical halo itself is DMO-like. Fig. 6 should therefore be read not as a test of whether hydrodynamical haloes obey the DMO $c$–$M_{200c}$ relation but as a test of whether the uncontracted component recovered by the fit, once the baryon-induced distortion has been modelled and effectively subtracted, is consistent with the relation obeyed by the true DMO realization of the same haloes. That CNFW recovers this relation, while NFW, gNFW and Einasto – which do not separate a contraction term and instead absorb the baryonic distortion into a single set of DMO-motivated shape parameters – systematically fail to, constitutes a non-trivial validation of the contraction ansatz itself: it shows that the formalism correctly isolates the baryon-induced response, leaving an inferred uncontracted profile whose statistical properties match those of the halo's actual DMO counterpart. This is a property of the fitting model, not a prior assumption imposed on the hydrodynamical haloes.

We do not show gNFW and Einasto fits in Fig. 6; however, we have verified their behaviour with analogous plots, which are presented in Appendix B. Qualitatively, both gNFW and Einasto follow trends similar to standard NFW. Fits restricted to small radii systematically overestimate concentrations while underestimating masses, displacing haloes away from the expected $c$–$M_{200c}$ relation by significantly more than the $2\sigma$ population scatter. The main difference with NFW is that when the outer fitting radius exceeds $\sim$60 kpc, the $2\sigma$ covariance contours begin to overlap the expected relation, and when the full halo is sampled (up to 214 kpc) the $1\sigma$ ellipses also cross the canonical scaling. These results confirm that the coherent biases seen for NFW are broadly representative of profiles not explicitly modelling contraction, especially when restricted to the inner regions.

These results have direct implications for analyses that impose the DMO $c$–$M_{200c}$ relation as a prior when fitting halo models to observational data. Because NFW fits produce correlated biases – overestimated concentrations associated with underestimated masses – these shifts occur in a direction that does not follow the slope of $c$–$M_{200c}$ relation itself. A tight prior based on the DMO relation therefore does not correct the joint bias; instead, it can result in apparently precise but systematically biased estimates, particularly when only the inner regions of the halo are probed, as is typical for the MW. In contrast, CNFW fits recover the correct scaling without requiring a $c$–$M_{200c}$ prior; the use of any prior becomes statistically innocuous rather than compensatory. The effectiveness of $c$–$M_{200c}$ -prior-based inference therefore depends primarily on the physical adequacy of the underlying halo model: When the profile family cannot reproduce and/or account for the contracted shape, no reasonable prior can eliminate the resulting joint bias.

In summary, explicitly modelling baryon-induced contraction is essential for recovering an unbiased concentration–mass relation from hydrodynamical haloes, especially when observational

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

constraints reach only the inner ∼10 per cent of the virial radius. CNFW fits preserve the population-level scaling across all radial ranges tested, whereas NFW fits introduce large, coherent distortions that propagate directly into analyses that employ $c$–$M_{200c}$ priors (and similarly for gNFW and Einasto fits restricted to the inner regions). Importantly, these conclusions hold when comparing directly to the intrinsic $c$–$M_{200c}$ distribution of the ARTEMIS DMO haloes, and are therefore not contingent on the choice of any external calibration of the relation. This underscores the need for contraction-aware models in dynamical studies of the MW halo, even more so in efforts to interpret the concentration–mass relation in real galaxies.

## 5 CONCLUSION

In this work, we have quantified how baryonic processes modulate the structure of DM haloes in the ARTEMIS simulation suite, and how these effects propagate into the inference of global halo properties when observations are limited to the inner, baryon-dominated regions. Our results demonstrate that the inclusion of baryons produces an extended isothermal-like ($\rho \propto r^{-2}$) shell in the DM halo spanning ≈2–20 kpc, accompanied by an enhancement of the inner density (Fig. 2). This behaviour is consistent with previous findings from other hydrodynamical simulations and confirms that contraction is a generic outcome of galaxy formation in haloes of MW mass.

We verified that the semi-empirical contraction model of M. Cautun et al. (2020) accurately captures this response within ARTEMIS. The ARTEMIS haloes reproduce the same deterministic power-law relation between enclosed DM and total mass ratios in hydrodynamical against DMO runs, with deviations from this trend remaining within the originally quoted scatter (Fig. 3). This agreement reinforces the physical interpretation of contraction as a predictable coupling between the baryonic and DM components.

A central goal of this study was to assess how the limited radial coverage available to kinematic tracers and the choice of halo model impact the recovery of structural halo parameters. By fitting several widely used parametric profiles – NFW, gNFW, Einasto, and CNFW – to restricted radial ranges, we found:

(i) Excellent local fits do not guarantee accurate extrapolation. When constrained only by radii $\lesssim$20–30 kpc, NFW, gNFW, and Einasto models infer systematically biased halo mass and concentrations, even when their residuals within the fitted domain are minimal. Moreover, internal metrics – such as parameter degeneracies or posteriors–often carry little to no information on the quality of the extrapolation.

(ii) The biases arise due to baryon-induced contraction creating a broad isothermal shell that strongly influences local curvature but carries little information about the outer halo. Models without explicit contraction physics compensate by distorting their fitted parameters, typically producing overconcentrated and under-massive haloes.

(iii) The CNFW model – which incorporates baryonic effects through the M. Cautun et al. (2020) formalism – exhibits substantially more stable behaviour. Even when fitted only to ∼20 kpc, CNFW returns extrapolated halo masses within ∼0.06 dex (15 per cent) of the true value on average, with a significant fraction of haloes accurate to within 25 per cent. While some scatter remains, CNFW is the only model that avoids the strong systematic biases observed for the NFW, gNFW, and Einasto profiles. Importantly, it yields concentration values that trace the intrinsic structural parameters of the primordial halo, rather than those imprinted by baryonic condensation. Furthermore, it is the only model in our suite able to recover the underlying DMO concentration–mass relation when only the inner regions are probed.

(iv) Increasing the maximum fitting radius towards the typical virial radius in our sample ($\tilde{R}_{200,\mathrm{DMO}} \approx 214$ kpc) substantially improves the recovery of halo mass and, to a lesser extent, concentration for all profile families – NFW, gNFW, Einasto, and CNFW. The improvement is smallest for CNFW, which already achieves very stable fits through explicit contraction modelling. It is worth noting that NFW estimates never fully reach zero median deviation in halo mass, and that all non-contraction-aware models continue to exhibit significant bias in concentration even when the full halo is probed. In practice, we find that accurate (1 per cent bias using CNFW, gNFW and Einasto) and precise (10 per cent scatter using CNFW, gNFW, and Einasto) recovery of halo mass is best achieved using kinematic tracers restricted to the outer halo, largely unaffected by contraction (i.e. $r \gtrsim 50$ kpc), whereas inner-halo coverage is more informative for constraining concentration provided a contraction-aware model.

Taken together, our findings have direct implications for dynamical studies of the MW and galaxies of similar mass. Kinematic data for the inner Galaxy primarily probe regions shaped by contraction; hence, widely used DM profiles can yield biased halo mass and concentration estimates even in the absence of observational uncertainties. This suggests that discrepancies among recent MW mass estimates may originate not only from data choices or modelling assumptions, but also from the intrinsic limitations of profile forms that do not encode baryonic physics. Our results therefore advocate for the routine incorporation of contraction-aware models when interpreting mass tracers confined to the inner ∼30 kpc.

We emphasize that the 'best' profile to use depends on the scientific question and application. For inferring the host halo properties, such as mass and concentration, the CNFW is the most reliable. While the flexible gNFW and Einasto profiles do not accurately extrapolate to larger radii, they do provide a better fit to the inner DM profile where we have the best kinematic data. By providing a more accurate description of the inner DM halo, these profiles are likely an equivalent or better choice for direct DM detection studies or searches for DM annihilation signals in the Galactic centre.

Future work will benefit from extending this analysis to a large observational rotation curve compilation, applying contraction-calibrated models directly to MW-mass systems, especially given the (expected) reliable extraction of the underlying concentration-mass relation plane without the need of imposing the theoretical relation as a prior. Nonetheless, the conclusions from the present study are clear: Baryon-induced contraction is both universal and dynamically significant, and accounting for it is essential for achieving unbiased inferences of the global properties of MW-mass DM haloes.

## AUTHOR CONTRIBUTIONS

DD led the analysis, interpretation, and presentation of all results under the supervision of SB and AF. Other authors contributed to creating the ARTEMIS simulations, and provided input on the interpretation of results and draft versions of the manuscript.

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

## ACKNOWLEDGEMENTS

DD acknowledges support by the Royal Society through a Dorothy Hodgkin Fellowship (DHF/R1/231105). STB and AF acknowledge support by a UK Research and Innovation (UKRI) Future Leaders Fellowship (grant number MR/T042362/1) and Sweden's Wallenberg Academy Fellowship. ASF acknowledges support from UKRI grant ST/W006766/1. IGM acknowledges support from STFC grant ST/Y002733/1.

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. This work has made use of NASA's Astrophysics Data System Bibliographic Services.

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

*Software:* The following software packages were used in this work: NUMPY (C. R. Harris et al. 2020), SCIPY,[4] PANDAS (The pandas development team 2020), ASTROPY (Astropy Collaboration 2022), and H5PY (A. Collette et al. 2022).

## DATA AVAILABILITY

The data underlying the plots in this article are available from the corresponding author upon request. The raw data from the ARTEMIS simulation suite are not publicly available; researchers wishing to access the simulations may contact the ARTEMIS Collaboration.

## REFERENCES

Abadi M. G., Navarro J. F., Fardal M., Babul A., Steinmetz M., 2010, MNRAS, 407, 435
Astropy Collaboration, 2022, ApJ, 935, 167
Binney J., 1977, ApJ, 215, 483
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Brown S. T., McCarthy I. G., Diemer B., Font A. S., Stafford S. G., Pfeifer S., 2020, MNRAS, 495, 4994
Bullock J. S., Boylan-Kolchin M., 2017, ARA&A, 55, 343
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
Callingham T. M. et al., 2019, MNRAS, 484, 5453
Callingham T. M., Cautun M., Deason A. J., Frenk C. S., Grand R. J. J., Marinacci F., Pakmor R., 2020, MNRAS, 495, 12
Cautun M. et al., 2020, MNRAS, 494, 4291 (C20)
Cimatti A., Fraternali F., Nipoti C., 2019, preprint (arXiv:1912.06216)
Collette A. et al., 2022, h5py/h5py: 3.7.0, Zenodo. Available at:
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Deason A. J. et al., 2012, MNRAS, 425, 2840
Di Cintio A., Brook C. B., Macciò A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014a, MNRAS, 437, 415
Di Cintio A., Brook C. B., Dutton A. A., Macciò A. V., Stinson G. S., Knebe A., 2014b, MNRAS, 441, 2986
Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108
Drukier A. K., Freese K., Spergel D. N., 1986, Phys. Rev. D, 33, 3495
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Eadie G. M., Springford A., Harris W. E., 2017, ApJ, 835, 167
Eilers A.-C., Hogg D. W., Rix H.-W., Ness M. K., 2019, ApJ, 871, 120
Einasto J., 1965, Trudy Astrofiz. Inst. Alma-Ata, 5, 87
Font A. S. et al., 2020, MNRAS, 498, 1765
Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507
Gaia Collaboration, 2023, A&A, 674, A1
Garcia A. M. et al., 2026, ApJ, 1002, 8
Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16
Goodman M. W., Witten E., 1985, Phys. Rev. D, 31, 3059
Grand R. J. J. et al., 2017, MNRAS, 467, 179
Harris C. R. et al., 2020, Nature, 585, 357
Hinshaw G. et al., 2013, ApJS, 208, 19
Hunter D. A. et al., 2012, AJ, 144, 134
Hussein A., Necib L., Kaplinghat M., Kim S. Y., Wetzel A., Read J. I., Rey M. P., Agertz O., 2025, preprint (arXiv:2501.14868)
Jungman G., Kamionkowski M., Griest K., 1996, Phys. Rep., 267, 195
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98
Karukes E. V., Benito M., Iocco F., Trotta R., Geringer-Sameth A., 2019, J. Cosmol. Astropart. Phys., 2019, 046
Lazar A. et al., 2020, MNRAS, 497, 2393
Li Z.-Z., Jing Y. P., Qian Y.-Z., Yuan Z., Zhao D.-H., 2017, ApJ, 850, 116
Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214
Ludlow A. D., Schaye J., Bower R., 2019, MNRAS, 488, 3663
Majewski S. R. et al., 2017, AJ, 154, 94
Marasco A., Cresci G., Posti L., Fraternali F., Mannucci F., Marconi A., Belfiore F., Fall S. M., 2021, MNRAS, 507, 4274
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Ou X., Eilers A.-C., Necib L., Frebel A., 2024, MNRAS, 528, 693
Ou X., Necib L., Wetzel A., Frebel A., Bailin J., Oeur M., 2025, ApJ, 994, 128
Patel E., Besla G., Mandel K., Sohn S. T., 2018, ApJ, 857, 78
Persic M., Salucci P., Stel F., 1996, MNRAS, 281, 27
Pillepich A. et al., 2018, MNRAS, 473, 4077
Pontzen A., Governato F., 2012, MNRAS, 421, 3464
Poole-McKenzie R., Font A. S., Boxer B., McCarthy I. G., Burdin S., Stafford S. G., Brown S. T., 2020, J. Cosmol. Astropart. Phys., 2020, 016
Posti L., Helmi A., 2019, A&A, 621, A56
Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14
Press W. H., Schechter P., 1974, ApJ, 187, 425
Rubin V. C., Ford W. K., Jr, Thonnard N., 1980, ApJ, 238, 471
Sarrato-Alós J., Bullock J., Di Cintio A., Brook C., Valenciano F., Macciò A. V., 2026, A&A, 710, A95
Sawala T. et al., 2015, MNRAS, 448, 2941
Schaller M. et al., 2015, MNRAS, 451, 1247
Schaye J. et al., 2015, MNRAS, 446, 521
Sellwood J. A., McGaugh S. S., 2005, ApJ, 634, 70
Springel V., 2005, MNRAS, 364, 1105
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
The Pandas Development Team, 2020, pandas-dev/pandas: Pandas, Zenodo. Available at: 10.5281/zenodo.3509134
Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt R. C., Jr, Thornley M. D., Leroy A., 2008, AJ, 136, 2563
Wang W., Han J., Cautun M., Li Z., Ishigaki M. N., 2020, Sci. China: Phys. Mech. Astron., 63, 109801
Watkins L. L., Evans N. W., An J. H., 2010, MNRAS, 406, 264
Watkins L. L., van der Marel R. P., Sohn S. T., Evans N. W., 2019, ApJ, 873, 118
Xue X. X. et al., 2008, ApJ, 684, 1143
Zavala J., Frenk C. S., 2019, Galaxies, 7, 81

[4] docs.scipy.org/doc/



Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

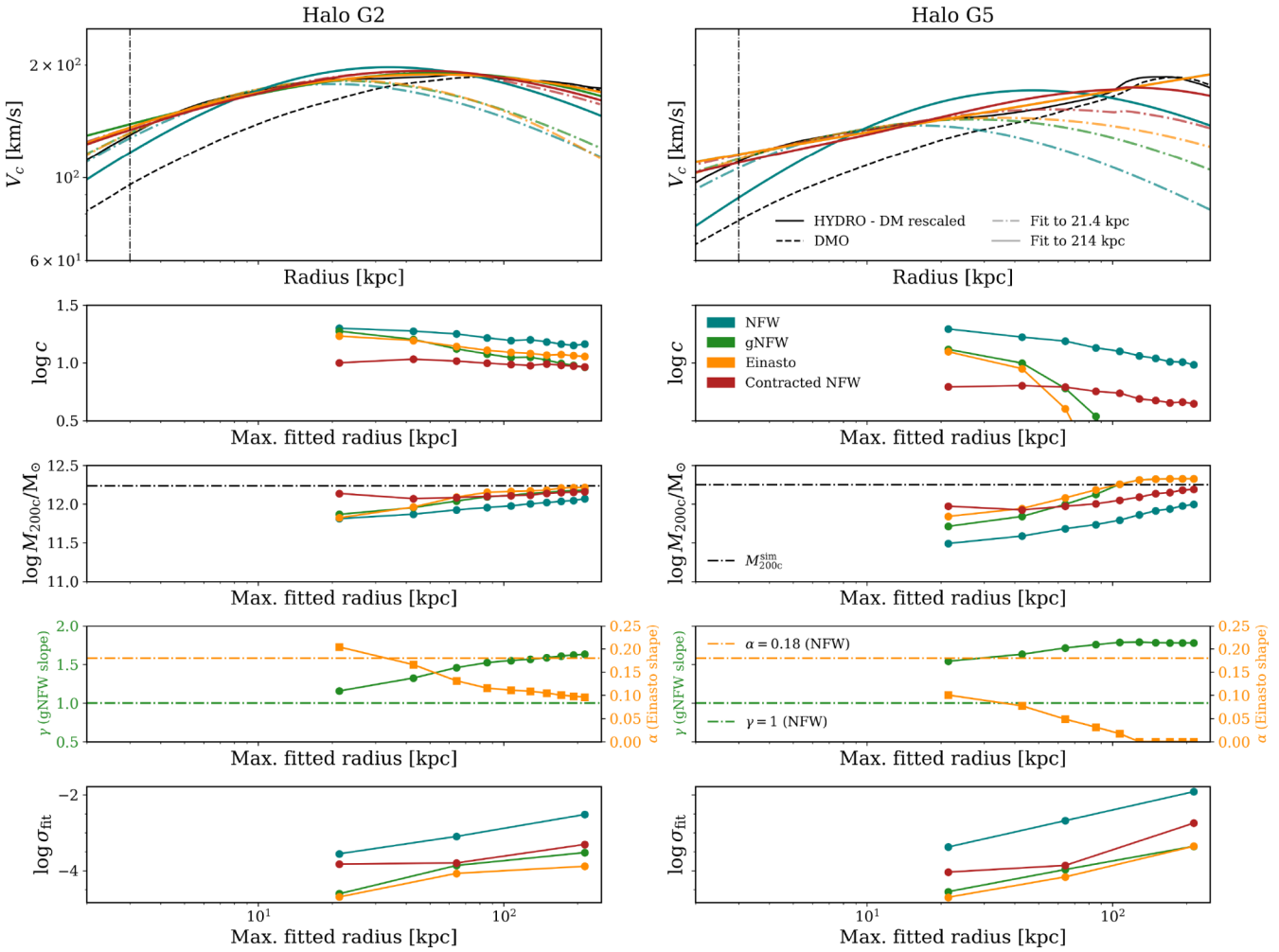



**Figure A1.** Comparison between a representative halo (halo G2, left) and one of the extreme outliers identified in Fig. 5 (halo G5, right). *First row:* Circular velocity profile of the simulated hydrodynamical halo (solid black) and its DMO counterpart (dashed black), with best-fit NFW (teal), gNFW (green), Einasto (orange), and contracted NFW (red) profiles constrained to $r_{\rm max,fit} = 21.4$ kpc (dash–dotted) and 214 kpc (solid). *Second and third rows:* Recovered concentration and $M_{200c}$ as a function of maximum fitted radius; the dash–dotted line in the third row marks the true simulated $M_{200c}$. *Fourth row:* Best-fit gNFW slope $\gamma$ (green, left axis) and Einasto shape parameter $\alpha$ (orange, right axis) as a function of maximum fitted radius; horizontal dash–dotted lines mark the equivalent NFW values ($\gamma = 1, \alpha \approx 0.18$). *Fifth row:* Logarithmic fit residual scatter, $\log \sigma_{\rm fit}$, as a function of maximum fitted radius. Halo G5 exhibits a localized excess in circular velocity for $r \gtrsim 50$ kpc, associated with a massive subhalo near the virial radius, which drives the gNFW and Einasto shape parameters to extreme values once this radial range enters the fit, producing erratic swings in extrapolated mass and concentration despite low, well-behaved fit residuals.

## APPENDIX A: THE ORIGIN OF EXTREME OUTLIERS IN EXTRAPOLATED HALO PROPERTIES

As noted in Section 4.2, a small number of haloes in our sample show erratic evolution of their extrapolated halo parameters as the fitted radial range is increased, particularly for the gNFW and Einasto models. Here we examine one such case in detail and contrast it with a representative halo, to clarify the origin of this behaviour.

Fig. A1 compares halo G2, representative of the bulk of our sample, with halo G5, one of the extreme outliers visible in the last two columns of Fig. 5. The top row shows the simulated circular velocity profile (HYDRO-DM rescaled, solid black; DMO, dashed black) together with best-fit models (NFW, teal; gNFW, green; Einasto, orange; CNFW, red) constrained to an inner radius of 21.4 kpc (dash–dotted) and to the full typical virial radius, 214 kpc (solid). For halo G2, all four models track the simulated profile closely at all radii probed in the fit. For halo G5, in contrast, the simulated profile itself departs from a smooth, monotonically declining shape at $r \gtrsim 50$ kpc, developing a localized excess in $V_c$ that is absent from the bulk of the sample. Cross-matching against the halo catalogue confirms that this feature coincides with the presence of a massive subhalo (or close companion) projected near the virial radius – both in the hydrodynamical and DMO runs, which breaks the spherical symmetry assumed by all four fitting models.

The consequences of this localized departure are most clearly seen in the shape parameters of the flexible models (fourth row). For halo G2, the gNFW slope $\gamma$ and Einasto shape parameter $\alpha$ evolve smoothly and monotonically from their NFW-equivalent values ($\gamma = 1$, $\alpha \approx 0.18$) as the fitted range widens, consistent

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

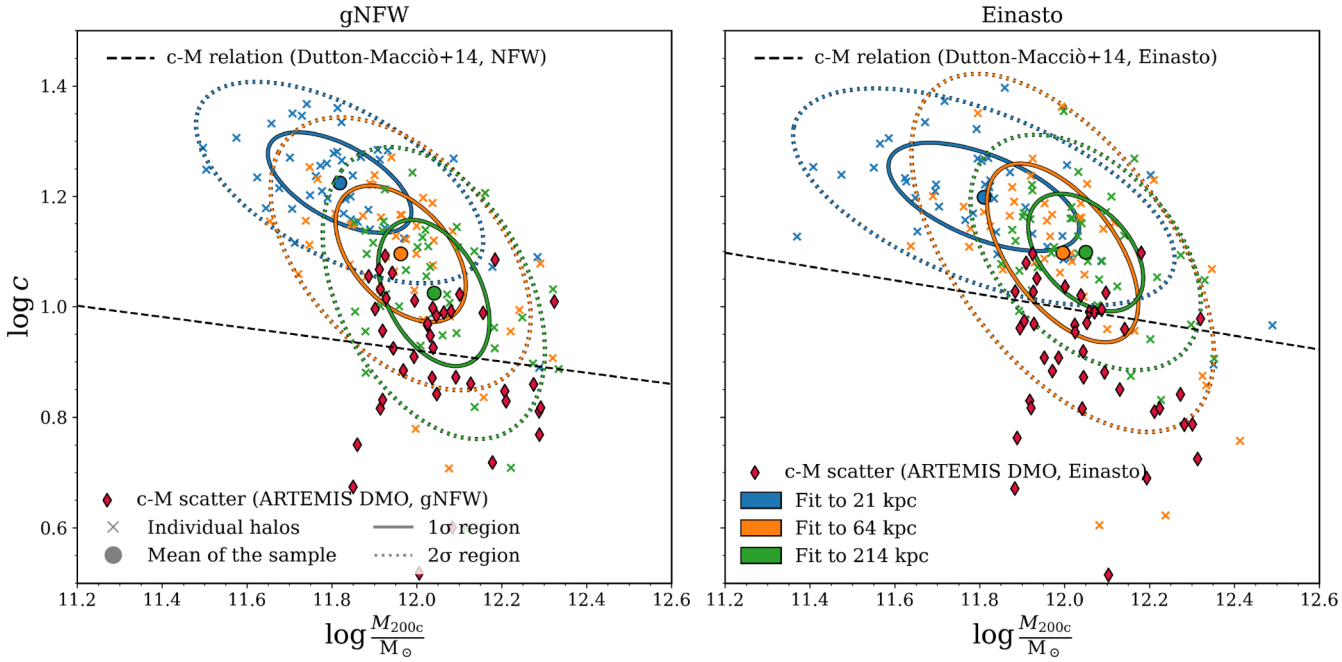



**Figure B1.** Recovered concentration–mass relations for gNFW (left) and Einasto (right) profile fits. Coloured crosses show individual haloes for different maximum fitted radii, while large filled circles indicate the corresponding sample means. Solid and dashed ellipses denote the $1\sigma$ and $2\sigma$ covariance regions, respectively. Dark red diamonds show the intrinsic $c$–$M_{200c}$ distribution of the ARTEMIS DMO haloes, obtained from fits to their full radial extent, and therefore provide the appropriate simulation benchmark. The dashed black line indicates the A. A. Dutton & A. V. Macciò (2014) relation (for NFW parametrizations on the left, Einasto on the right) shown for reference.

with the behaviour discussed in Section 4.1 and approaching realistic values. For halo G5, this evolution proceeds smoothly only while the fitted domain remains inside $\sim$50 kpc; once the substructure-induced feature enters the fit, $\alpha$ is driven towards zero and $\gamma$ saturates near its imposed upper bound, as both models attempt to reproduce the localized excess by adopting increasingly extreme, quasi-power-law density profiles. This directly produces the abrupt swings in concentration and mass seen in the second and third rows: $\log c$ drops sharply and $\log M_{200c}$ rises sharply over the same narrow range of $r_{\rm max,fit}$ where the shape parameters saturate. The NFW and CNFW models, lacking this additional shape freedom, cannot reproduce the feature as closely, but for the same reason their extrapolated parameters remain comparatively stable and, in the case of concentration, realistic.

Notably, the bottom row shows that this instability is not signalled by the internal fit-quality metric $\log \sigma_{\rm fit}$: For halo G5, gNFW and Einasto retain residual scatter as low as, or lower than, the well-behaved models, even in the regime where their extrapolated mass and concentration are least reliable. This reinforces the point made in Section 4.3 that internal diagnostics of fit quality are, on their own, poor indicators of the reliability of an extrapolation. We find this behaviour – a localized departure from spherical symmetry driving extreme shape-parameter excursions in the most flexible models – to be characteristic of all outliers identified in Fig. 5, and we have verified that it is insensitive to our choice of radial binning.

## APPENDIX B: GENERALIZED NFW AND EINASTO MODELS IN THE MASS-CONCENTRATION PLANE

In this Appendix, we present the behaviour of gNFW and Einasto profile fits in the $c$–$M_{200c}$ plane, complementing the analysis of NFW and CNFW models discussed in Section 4.4. The results are shown in Fig. B1, constructed in the same manner as Fig. 6.

Both gNFW and Einasto fits exhibit systematic offsets in the $c$–$M_{200c}$ plane when constrained to the inner regions, consistent with the behaviour discussed in Section 4.4. In particular, fits restricted to $r \lesssim 20$ kpc overestimate concentrations and underestimate halo masses, displacing haloes away from the intrinsic DMO relation by amounts comparable to or exceeding the $2\sigma$ population scatter.

However, clear differences emerge between the two models. Einasto fits (right panel) lie systematically closer to both the intrinsic ARTEMIS DMO distribution and the reference $c$–$M_{200c}$ relation than either gNFW or NFW fits, even when restricted to small fitting radii. In contrast, gNFW fits (left panel) display a much stronger dependence on the fitted radial range: Inner-halo fits are significantly offset, while extending the fits to larger radii progressively shifts the recovered parameters towards the DMO locus.

Another notable distinction is that, unlike NFW, both gNFW and Einasto fits recover distributions that overlap with the intrinsic DMO scatter when the full radial extent of the halo is included. This indicates that the additional degree of freedom in these models can partially absorb the structural differences

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

between hydrodynamical and DMO haloes, thereby reducing the apparent tension in the $c$–$M_{200c}$ plane. However, this flexibility does not eliminate the underlying bias when only inner-halo data are available; instead, it redistributes the mismatch across parameters. This effect is particularly pronounced for gNFW fits, where strong degeneracies between shape parameters lead to increased scatter and sensitivity to the fitted radial range.

Overall, these results highlight that increased model flexibility can mitigate discrepancies with the DMO $c$–$M_{200c}$ relation when wide radial coverage is available, but does not provide a physically robust correction for baryon-induced contraction. In contrast, the CNFW model achieves consistency with the expected scaling by explicitly accounting for these effects, rather than effectively marginalising over them through additional degrees of freedom.vs

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

Downloaded from academic.oup.com/mnras/article/552/4/stag1865/8859319 by guest on 04 November 2026

© The Author(s) 2026.
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.